UNIVERSITÀ DEGLI STUDI
DI TRENTO

DEPARTMENT OF INFORMATION ENGINEERING AND COMPUTER SCIENCE
ICT International Doctoral School

Methods for the analysis of time series of multispectral remote sensing images and application to climate change variable estimations

Iwona Katarzyna Podsiadlo

Advisor

Prof. Lorenzo Bruzzone

Università degli Studi di Trento

Co-Advisor

Dr. Claudia Paris

Università degli Studi di Trento

May 2021

Dedicated to my grandmother and each person without possibility of education.

Abstract

In the last decades, the increasing number of new generation satellite images characterized by a better spectral, spatial and temporal resolution with respect to the past has provided unprecedented source of information for monitoring climate changes. To exploit this wealth of data, powerful and automatic methods to analyze remote sensing images need to be implemented. Accordingly, the objective of this thesis is to develop advanced methods for the analysis of multitemporal multispectral remote sensing images to support climate change applications. The thesis is divided into two main parts and provides four novel contributions to the state-of-the-art.

In the first part of the thesis, we exploit multitemporal and multispectral remote sensing data for accurately monitoring two essential climate variables. The first contribution presents a method to improve the estimation of the glacier mass balance provided by physically-based models. Unlike most of the literature approaches, this method integrates together physically-based models, remote sensing data and in-situ measurements to achieve an accurate and comprehensive glacier mass balance estimation. The second contribution addresses the land cover mapping for monitoring climate change at high spatial resolution. Within this work, we developed two processing chains: one for the production of a recent (2019) static high resolution (10 m) land cover map at subcontinental scale, and the other for the production of a long-term record of regional high resolution (30 m) land cover maps.

The second part of this thesis addresses the common challenges faced while performing the analysis of multitemporal multispectral remote sensing data. In this context, the third contribution deals with the multispectral images cloud occlusions problem. Differently from the literature, instead of performing computationally expensive cloud restoration techniques, we study the robustness of deep learning architectures such as Long Short Term Memory classifier to cloud cover. Finally, we address the problem of the large scale training set definition for multispectral data classification. To this aim, we propose an approach that leverages on available low resolution land cover maps and domain adaptation techniques to provide representative training sets at large scale.

The proposed methods have been tested on Sentinel-2 and Landsat 5, 7, 8 multispectral images. Qualitative and quantitative experimental results confirm the effectiveness of the methods proposed in this thesis.

Keywords

Remote sensing, multispectral images, multitemporal images, time series, machine learning, deep learning, land cover mapping, regression models, glacier mass balance estimation.

Acknowledgements

Now, when I am towards the end of this PhD journey, I would like to look back and express my gratitude to all people without whom I would not have been able to complete this path. First of all, I would like to express my sincere gratitude to my supervisor prof. Bruzzone for giving me the opportunity to do this PhD under his supervision, for his guidance and support. Although your very tight schedule, you have always found time to give advises of excellent quality, regardless the time of day or night. I would like to thank my co-advisor Claudia. Your support have been of tremendous importance in my PhD journey. Since I have started to work with you, I have learned so much from research, communication and organization perspective. I feel extremely grateful that you have been my co-superior. You have been not only a great mentor to me but also a good friend, who was always there to give me insightful advises, feedback, words of encouragement and was there to listen. I would also like to thank people, who I met in the ESRIN dr. Henri Laur and Nuno Miranda for reinforcing me to take this PhD path and recommending to do it under the supervision of prof. Bruzzone.

I would like to thank to my colleagues from RS lab, especially Elisa and Sanchari, for your support and making this PhD much more funny and full of 'giretti'. Thank to Massimo for his technical and scientific support. Thank to Federico for our cooperation on CCI project, and sending me the images and classification maps late at night. Next thank to people that I met in Trento, who will always remain my dear friends Marina, Athanasia, Valentina, Paulina, Carletto, Piotr, Stefano x2. Thank you for supporting me and making sure that I was doing well. I would like to thank my good friend Stasiu for being such a wonderful friend your support is so meaningful to me. Special thanks go to my friends, who have been always in contact with me especially during this strange covid times Victoria, Laure, Wenting, Hayret, Monika, Ewelina, Kamil, Agnieszka and Marcin.

I would like to thank to my family, especially to my bother Damian for his passion in science (which happened to by contagious), good advice and support with all problems related to the English language. Dziekuje mojej rodzinie, a w szczegolnosci moim rodzicom, za ich troske i pomoc. Dziekuje mojemu rodzenstwu Mariuszowi i Marzenie za wiare we mnie.

From the bottom of my heart I would like to say big thank to Bruno, who appeared in my life just in the beginning of this PhD journey. You know exactly how difficult this path has been for me. Thank you for your unconditional love, encouragement and support in every possible way even if at the same time you are combating covid-19 in the first line. Vorrei anche ringraziare Adriana, Robi e Pluto per trattarmi come parte della famiglia.

Abreviations

AAR Accumulation area ratio

ALS Airborne laser scanner

AMORGOS Accurate MERIS Ortho-Rectified Geo-location Operational Software

AMUNDSEN Alpine Multiscale Numerical Distributed Simulation Engine

AVHRR Advanced Very High Resolution Radiometer

CCA Canonical Correlational Analysis

CCI Climate Change Initiative

CGLC Copernicus Global Land Cover

CGLS Copernicus Global Land Service

CNES Centre National d'Études Spatiales

CNN Convolutional neural network

CR Continuum Removal

CSI Cloud shadow index

DANN Domain-adversarial neural networks

DDI Data density indicator

DTM Digital Terrain Model

ELA Equilibrium-line altitude

ENVISAT Environmental Satellite

ESA European Space Agency

ETM+ Enhanced Thematic Mapper Plus

FIR Far-infrared

Fmask Function of mask

FROM-GLC Finer Resolution Observation and Monitoring of Global Land Cover

GMB Glacier mass balance

GCOS Global Climate Observing System

GFK Geodesic Flow Kernel

GHOST Global Hierarchical Observing Strategy

HBV Hydrologiska Byrans Vattenbalansavdelning

HJ-1 Huan Jing-1

HR High resolution

IPCC Intergovernmental Panel on Climate Change

KCCA Kernel canonical correlation analysis

KEMA Kernel based manifold alignment method

k-NN k-nearest neighbors

k-NNG k-nearest neighbor graph

LC Land cover

LCCS United Nation Land Cover Classification System

LDA Linear discriminant analysis

LiDAR Light Detection and Ranging

LSTM Long short Term Memory

MAJA MACCS-ATCOR Joint Algorithm

MERIS Medium Resolution Imaging Spectrometer

MIR Middle-infrared

MME Minimax entropy

MLC Maximum likelihood classifier

MODIS Moderate Resolution Imaging Spectroradiometer

MSE Mean Absolute Error

NASA National Aeronautics and Space Administration

NDSI Normalized Difference Snow Index

NDVI Normalized Difference Vegetation Index

NDWI Normalized Difference Water Index

NIR Near-infrared

NSPI Neighborhood similar pixel interpolation

OA Overall accuracy

OLI Landsat Operational Land Imagery

PAN Panchromatic

PCA Principal component analysis

R	Correlation coefficient
RBF	Radial Basis Function
RF	Random forest
RMSD	Root-mean-square deviation
RMSE	The Root Mean Square Error
RNN	Recurrent neural networks
RS	Remote sensing
SAR	Synthetic Aperture Radar
SLA	Snow line altitude
SLC	Scan Line Corrector
SPOT	Système Pour l'Observation de la Terre
SRTM	Shuttle Radar Topography Mission
SSMA	Semi-supervised manifold alignment
SSTs	Sea surface temperatures
SVM	Support vector machine
SVR	Support vector regression
SWIR	Short wave infrared bands
TCA	Transfer Component Analysis
TIR	Thermal InfraRed
TM	Landsat Thematic Mapper
TS	Time Series
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
WGMS	Word Glacier Monitoring Service
WMO	World Meteorological Organization

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this chapter we provide the background of the thesis. Moreover, we present the motivations and the objectives of this dissertation, along with the main novel contributions of the research developed.

1.1 Background

Throughout centuries Earth's climate has changed significantly. The most notable of these changes is global temperature rise. According to the World Meteorological Organization (WMO), the world temperature is about one Celsius degree warmer than before the era of industrialization. Moreover, the Intergovernmental Panel on Climate Change (IPCC) forecasts that the temperature will continue to rise [2]. The temperature rise is determined mainly by the increase of emissions of carbon dioxide and other gases (known for their heat-trapping nature [3]) in the atmosphere. Such climate change has serious implications across the globe. For instance, the water from melting glaciers is causing the rise of the average sea level, with an estimated increase of 3.6 mm per year (in between 2005 and 2015). Such sea level rises increased the possibility of inundation and coastal erosion, which result in the loss of coastal wetlands and beaches [4]. Indeed, even a small variation of the global average temperature, triggers a chain reaction of climatic change around the world. Apart from the changes of the Earth' surface, the climate change is impacting many other domains, such as agricultural productivity, energy use, human health and water quality. Therefore, climate monitoring is of paramount importance to better understand its complexity and to adequately invest in climate mitigation and adaptation strategies, which are necessary to limit the harmful impacts of climate change.

Remote sensing (RS) is one of the most common technologies used for Earth observation and climate change monitoring. Indeed, a lot of effort have been devoted to understand how RS can support climate change monitoring [5]. The Global Climate

Observing System (GCOS) operated by the Word Glacier Monitoring Service (WGMS) defined a list of the essential climate variable, which are representing the key components of the climate system. Indeed, out of 54 essential climate variables more than half can be supported by the information derived from satellites. This is because RS satellites provide data covering extensive areas, with good temporal resolution, high spatial resolution and in a cost-efficient manner. Therefore, RS techniques play a crucial role to support and monitor the climate dynamics [5]. In particular, RS techniques allow for a worldwide monitoring of the sea surface temperatures (SSTs) [6], land cover (LC)[7], deforestation [8], glacier mass balance (GMB)[9], sea level change [10], soil moisture [11] and many others climate variables.

Due to the open data policy of high resolution (HR) and middle resolution (MR) multispectral missions (e.g., Sentinel 2 and Landsat, respectively), these data are provided free of charge to all data users. In particular, historical data available in the archives (e.g., Landsat data are available since 1972) represent an extremely important source of information to better capture long term climate change dynamics. The possibility of exploiting this large archives of multitemporal MR and HR RS multispectral data plus high-performance computing capabilities currently available, enables the accurate mapping of LC changes occurred on the Earth' surface and predict their impact on the climate [12]. Thus, more effort has to be devoted to the definition of processing chains able to ingest long and dense time series (TS) of MR and HR optical images. Other essential climate variable, which monitoring benefits from the availability of the RS data is GMB. In this context, there is the need of developing system architectures capable of accurately monitoring complex environmental phenomena of glacier melting. In particular, to properly capture their dynamic, it is necessary to define approaches able to exploit and properly integrate complementary information provided by different data sources such as RS data, *in-situ* measurements and physical-based models.

Analyzing long-term large scale RS data in an automatic and efficient way is not trivial. Processing large archives of RS data is challenging, and requires the definition of robust approaches, which are computationally efficient. In particular, one of the most important challenges to be addressed when classifying multispectral data is related to the cloud and shadow problem. In accordance with the American Meteorological Society estimates, the mean annual global cloud cover over land is ranging from 35% up to 66%, depending on latitude [13]. When performing large-scale mapping, the cloud cover issue should be properly handled since it can severely affect the results obtained. Another important challenge that strongly affects the possibility of performing LC classification at large scale is the lack of large volumes of training data. The accuracy of the produced LC maps largely depends on the quantity and the quality of labeled samples, which are

equally important to the learning paradigm. On one hand, the photo-interpretation is a very time consuming and laborious task. On the other hand, the publicly available sample databases are limited and usually dedicated to peculiar applications, which makes it difficult to reuse them. In this context, there is the need of investigate possible solutions able to deal with the cloud coverage problem and the collection of reference data in a fast and effective way.

1.2 Motivations of the Thesis

To support the monitoring of essential climate variables, and thus improve our understanding of climate change, many approaches leveraging on RS data have been proposed in the literature. In this thesis, we focus the attention on two important essential climate variables: GMB and LC.

The most common methods for the estimation of the distributed GMB are based on exploitation of the *in-situ* measurements, physically based simulations and information provided through RS sensors. Usually, this data sources are exploited separately or the integration of two of them is considered. Indeed, on one hand the integration of physically based simulations and RS data proved to enhance the distributed GMB estimation [14]. On the other hand the assimilation of RS data into the physically based simulations is an effective approach for the simulations calibration [15], [16]. Although these approaches demonstrated to be effective, they are computationally expensive and do not fully exploit the complementary information provided by the different data sources. In particular, they do not take advantage from high accuracy of *in-situ* measurements and the generalization capabilities of physically based simulations. Moreover, the estimates provided by the physically based simulations suffer from low resolution (i.e S3M (Snow Multidata Mapping and Modeling)[17] at 500 meters, AMUNDSEN at 50 meters and PROSNOW at 25/50 meters [18].

By focusing the attention on the LC, GCOS has defined the list of requirements, which are critical to properly understand the role of LC for climate modeling and climate mitigation studies. The requirements consider the compromise that need to be taken between spatial, temporal resolutions and the product accuracy. It is expected that climate modeling would benefit from LC maps provided at sub-continental or global scale, with high spatial resolution between 10 and 30 meters. To answer the climate mitigation questions, a more detailed studies should be carried out at regional or local scale (especially within regions known as hot spot of climate change), with at least 5 years temporal sampling starting from 1990. Even though there are several LC maps publicly available, these products partially answer the needs of climatologists. This is because only few LC prod-

ucts have been generated at large scale and they are typically provided at coarse spatial resolution [19], [20]. Moreover, as they do not fully exploit the long term LC information provided by RS data archives, to enable detailed long-term LC monitoring, these products should be harmonized. However, it is a challenging task as this maps considers different legend definitions and are not consistent between each other. In recent years, a Climate Change Initiative (CCI) medium resolution map have been produced [7], representing the first attempt of producing long-time records of LC maps. Although the results are promising to support climate change analysis, there are still several requirements not fulfilled. These LC maps do not provide sufficient thematic detail and their spatial resolution should be improved. Moreover, the temporal consistency of such data need to be better investigated.

As mentioned above, multitemporal multispectral images represent an extremely useful source of information for continuously supporting studies of glaciers and LC. However, many challenges occur when these data are used at large scale. First, their use is strongly affected by the cloud cover and shadow, which if not properly handled strongly impacts the quality of image analysis. Although, different cloud detection and restoration methods have been proposed in the literature, typically the methods that obtain a high accuracy are associated with high computational burden, and thus they cannot be used at the operational level [21]. In this context, recent advancements of deep learning based architectures have opened new possibilities to the RS data analyses. The qualitative studies have shown that the deep architectures are not only very promising for the long TS analyses but are also robust to the cloudy images in the TS [22]. However, so far no detailed study have been conducted to better understand its robustness and possible limitations.

Second, another important challenge related to large scale classification of RS data is the scarcity of ground reference data to properly train the classifier. It is very well known that the production of a proper training set is often challenging and is a time-consuming and complex task. In the literature, there are very few annotated large training sets available [23] [24]. However, usually they can be used in very specific applications. Indeed, it is very challenging to produce training set that could properly represent diverse environmental conditions given the size of our Planet. To address this problem, domain adaptation methods have been proposed in the literature to reduce the need of collecting training samples. This is done by adapting the classifier trained on samples coming from source domain to classify the samples from different but related target domain. From the operational view point, the most interesting domain adaptation methods are those unsupervised [25], which are typically based on the exploitation of the data distribution structure. However, the possibility of achieving accurate results depends on both the similarity between the source and target distributions, and the peculiarities of the considered

set of LC classes [26]. In [27], the authors demonstrated that assuming the availability of some target labeled samples, the probability to succeed in solving complex domain adaptation problem sharply increases. Even though a few samples from the the target domain are required, this constrain can be too demanding in some operational scenarios. In this context, the availability of LC products represents valid source of information, which can be used for the extraction of labeled units from the target domain. However, publicly available LC maps are typically obsolete, characterized by low spatial resolution and their legend is pre-defined. For these reasons, their used is not straightforward. In this context, there is need of developing strategy able to ensure the selection of reliable samples (correctly associated to their labels) and representative of the considered LC classes.

1.3 Objectives and Novel Contributions

In this thesis, we aim to develop automatic methods for the analysis of multitemporal and multispectral RS data at large scale to support climate change studies. The contributions of this dissertation considers two main direction. On one hand, we propose methods for monitoring the following essential climate variables: i) GMB by integration of complementary data sources, and ii) long-term HR multitemporal LC mapping. On the other hand, we propose two methods to support the large scale classification of the multitemporal and multispectral RS data. In particular, we propose the methods to address the problem of: i) cloudy images in the TS, and ii) large scale training set generation. In the following, these contributions are briefly described.

1.3.1 Integrating Models and Remote Sensing Data for Distributed Glacier Mass Balance Estimation

The main goal of this contribution is developing an approach being able to improve the GMB estimates provided by a physically-based model by using RS data and few GMB *in-situ* measurements. Differently from state-of-the-art methods, we do not assimilate the observations into the model [15], but combine the information provided by the different sources to correct the model estimates. This condition allows us to preserve good generalization capability of the theoretical model while improving the accuracy of its estimate. The considered approach exploits multispectral and multitemporal images and a digital terrain model (DTM) to define a feature space being able to characterize the glacier properties. While the DTM accurately models the topography of the scene (strictly correlated to the amount of GMB), the multispectral data characterize the spectral properties of ice and snow. In such an informative feature space, the approach exploits the *in-situ* measurements to correct GMB estimates of the physically-based simulations. The correc-

tion is computed according to feature-space division strategy based on a local k-Nearest Neighbors (k-NN) estimation technique. In particular, the applied correction is the average difference between the reference samples and the model estimates. In the last step, the correction is applied to the entire glacier area, thus producing a better GMB estimates at pixel-level. The proposed approach is able to provide GMB estimates at the higher spatial resolution with respect to the one provided by the physically-based model (due to the better geometrical information provided by the multispectral sensors) and can be applied to different test sites for which *in-situ* measurements and physically-based simulations are available.

1.3.2 Multitemporal Land Cover Mapping at High Spatial Resolution for Sub-Continental Monitoring of Climate Change

Within this contribution, we define a system architecture to perform LC mapping for supporting climate change studies at high spatial resolution. In particular, two processing chains have been defined and developed to answer the requirements of the climate modeling community. The former, i.e., static processing chain, is tailored to the characteristics of Sentinel-2 images and aims at generating a HR LC map at 10 meters spatial resolution on sub-continental scale to support large scale climate modeling studies. The latter, historic processing chain is tailored to the properties of Landsat 5, 7 and 8 images. The LC maps are produced at 30 meters spatial resolution, and the historic processing chain is focused on the production of the long-term record of regional HR LC maps to support the climate mitigation and adaptation studies. The outputs from both processing chains will be then used as an input to the climate models, which aim is to understand the role of spatial and temporal resolution on the consistency of LC classification. Taking into account the large scale at which such analysis need to be carried out, the developed processing chains should be able to handle a very large amount of optical images. Moreover, going back to the past leads to different temporal sampling. The archives of MR and HR data present high variability in data availability from the temporal view point. The proposed static and historic processing chains are fully automatic and can be deployed at regional and sub-continental scale. Both processing chains are based on three mains phases: (i) images preprocessing, (ii) feature extraction, and (iii) supervised classification of the multitemporal RS data. The static map is generated using largely available Sentinel-2 data, as it is expected that the higher spatial and temporal resolution of the sensor will have a positive impact of the final LC map accuracy. The historic maps are produced every 5 years between 1990 and 2015. Due to availability of the Landsat data in the whole considered period the maps were produced at 30 meters spatial resolution. We test both the processing chains over the three peculiar study areas: African Sahel,

Amazon and Siberia.

1.3.3 A Long Short Term Memory Classifier Robust to Images with Cloud Cover

Cloud cover is a very well known problem in the RS literature. For almost all analysis of multispectral and multitemporal RS data it is mandatory to perform a proper cloud detection and restoration. In this contribution, we aim to investigate the robustness of the deep Long Short Term Memory (LSTM) classifier to the cloud presence. Although LSTM classifier has proven to be able to handle the presence of clouds [28], [29], no work which extensively analyse the robustness of LSTM to clouds can be found in the literature. In this study, we aim to quantitatively asses the capability of the network of handling different amounts of cloud coverage under different lengths of the TS. In greater detail, we analyze the effect of the cloud coverage on the classification maps produced by the LSTM by considering: (i) simulated cloud values, (ii) detected clouds represented by zeros values, and (iii) restored images by simple linear temporal gap filling and the effects of the cloud positions within the TS. The experimental analysis has been carried out on a TS of Sentinel-2 images acquired in 2019 over the Amazon. To perform a quantitative evaluation, we compared the classification map obtained on the cloud-free TS of images (i.e., cloud-free map) with those generated with the simulated cloudy TS (i.e., cloudy maps).

1.3.4 An Approach Based on Low Resolution Land-Cover-Maps and Domain Adaptation to Define Training Sets at Large Scale

The accurate classification of RS data at large scale is typically hampered by the availability of training data representative of the whole study area. The main goal of this contribution is developing a method that aims to enlarge existing training sets leveraging publicly available thematic products available at large scale. Indeed, such maps are typically obsolete and provided at low spatial resolution, thus cannot be directly use to enlarge existing training databases. For this reason, the proposed approach aims to exploit these thematic products to detect labeled samples having the highest probability to be correctly labeled by associating them with the reliable HR spectral information provided by the RS data. Such samples, extracted in an area geographically distant from the training set (target domain), are jointly used with the annotated samples of the source domain (where training set is available) to find a mapping space (called also latent space) where the data are aligned. The alignment of target and source domains is done by using the kernel based manifold alignment method. By projecting samples into the common latent

space, we are able to enlarge the training set in an unsupervised (no annotated samples from the target domain are required) but reliable way.

1.4 Structure of the Thesis

This Chapter provided a brief background and described both the objectives and the main contributions of this thesis. The rest of the thesis is organized in two main parts. The first part is dedicated to the application of the multitemporal multispectral RS data to support climate change monitoring. In greater detail, in Chapter 2 we present the state of the art related to the use of multitemporal multispectral RS data to monitor GMB and LC at large scale. Chapter 3, describes the GMB estimation method based on integration of RS data, physically based simulations and *in-situ* measurements. Chapter 4 elaborates the proposed processing chain for the generation of the LC maps at large scale.

The second part of this thesis, focuses on the challenges in the analysis of the multispectral and multitemporal RS data at the operational level. Chapter 5 describes the state of the art regarding the cloud cover analysis and the representative training set generation. Chapter 6 reports the study on the robustness of the LSTM classifier, while Chapter 7 presents the approach for large scale training set generation using domain adaptation methods and low resolution LC maps. Finally, in the last chapter the conclusions of the thesis are drawn and further developments of the research activities are also discussed.

Part I

Multispectral and Multitemporal Remote Sensing Data for Supporting Climate Change Monitoring

Chapter 2

Background and basic concepts

This chapter aims at providing background and basic concepts of multitemporal multispectral RS data and methods based on RS data to estimate the two climate essential variables investigated in this thesis: GMB and LC. In particular, we introduce the importance of GMB and LC monitoring for climate sciences. For the distributed GMB estimation, we describe the methods based on RS data, while for LC monitoring we discuss the global LC maps available. [1]

2.1 Background on Multitemporal Multispectral Remote Sensing Data

Generally, RS sensors can be divided into active (such as Synthetic-aperture radar (SAR) and light detection and ranging (LiDAR)) and passive systems (spectrometer and radiometer) and can be characterized by different spatial and temporal resolutions. Passive systems detect energy which is naturally available (e.g., solar radiation), while active systems provide their own energy source. Multispectral sensors are passive RS systems, which measure the sunlight radiation reflected by the objects on the Earth's surface in a specific spectral interval. Passing through the atmosphere solar radiation is strongly affected by the particles and gases. The incoming light and radiation can be scattered or absorbed by gas molecules in a very specific part of the spectrum. The parts of spectrum not severely influenced by these mechanism are used by existing passive remote sensors, the so called atmospheric windows. Most of RS instruments operate in one or more of these windows with detectors tuned to wavelengths that pass through the atmosphere. Indeed, according to the chemical and physical properties of the target, different surface

[1] Part of this chapter appears in I. Podsiadlo et al., "Integrating Models and Remote Sensing Data for Distributed Glacier Mass Balance Estimation," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 6177-6194, 2020, doi: 10.1109/JSTARS.2020.3028653.

types (such as vegetation, water or bare areas) reflect, absorb or transmit solar radiation differently in different portions of the electromagnetic spectrum. This relation is called the spectral signature and describes radiation reflected as a function of the wavelength. The spectral signature is extremely useful in identifying and discriminating between different objects present in the RS data.

Different optical RS sensors provide different trade-off of spatial, temporal and spectral resolution. The broader the wavelength range for a particular channel/band the higher spatial resolution. Therefore, optical RS can be divided into the following categories: panchromatic (PAN), multispectral and hyperspectral. Panchromatic sensors does not distinguish the individual wavelengths of the spectrum. This is why PAN images report only the total radiation for each pixel and are of very high spatial resolution (tens of centimeters). Multispectral sensors register energy over several separate wavelength ranges and they are available at very high (VHR) (2 to 5 meters) and moderate spatial resolution (MS)(from 10 to hounded of meters). The VHR multispectral sensors usually records the radiation in the following ranges of spectrum: visible (390 - 700 nm), the near-infrared (NIR) (from 700 to 900 nm) and in the shortwave infrared (SWIR) region (from 1100 to 3000 nm). While, the MS are able to measure shorter radiation intervals in the visible and NIR ranges as well as register the middle-infrared (MIR) and far-infrared (FIR or thermal) parts of spectrum. Moreover, due to their usually large swaths they have short revisit time, and thus are very attractive for the climate change monitoring. Hyperspectral sensors detect hundreds of very narrow spectral bands and are of coarser resolution. Therefore the are particularly capable to provide detailed identification of different targets (such as distinction of different types of rocks/forests).

In 1972 NASA (National Aeronautics and Space Administration) have launched first satellite of a highly successful Landsat program designed to monitor Earth's surface. During Landsat program eight satellites have been launched (out of which two are still active and one is under development) and till today provide large amount of MR multispectral data. The images spatial resolution is 30 meters and 16 day temporal resolution. Landsat 8 provides nine spectral bands, visible, NIR and SWIR bands have spatial resolution of 30 meters, PAN is of 15 meters while thermal bands are of 100 meters. The Landsat program main objectives, includes: i) acquisition and archivization of multispectral images, and thus assurance of seasonal coverage of the global landmasses, ii) data consistency between Landsat missions in terms of calibration, spectral and temporal characteristics, acquisition geometry, coverage characteristics and product quality, and iii) open data policies, enabling users allover the world access to the multispectral imagery. Indeed, Landsat program has been very helpful to accelerate various research and application developments. Another source of HR multispectral data are SPOT (Système Pour l'Observation de la

Terre) satellites, from which five were launched by the CNES (Centre National d'Études Spatiales) and recent two by the Airbus Defence & Space. SPOT is a commercial provider of Earth observation data and each of the satellites can provide data in a high spatial resolution panchromatic mode, or a lower spatial resolution multispectral mode. In 2015, the European Space Agency (ESA) have launched Sentinel-2 program with an improved (with respect to Landsat) spatial resolution of 10 meters and similarly to Landsat program an open data policy. The Sentinel-2 mission includes the constellation of two polar-orbiting satellites, phased at 180° to each other. Moreover, the sensors can be characterized by the wide swath (290 km), and thus enable the mission to provide data with high temporal resolution of 5 days (depending on cloud cover and latitude). The Sentinel-2 provides 13 spectral bands out of which visible and NIR bands have a spatial resolution of 10 meters, SWIR of 20 meters while coastal/aerosol, water vapor, and cirrus bands of 60 meters. The program is designed to ensure the continuity of ongoing multispectral observations, such as Landsat program. The program has been already very successful resulting in the high number of the scientific publications. Spectral resolution of Sentinel-2 and Landsat 7 and 8 are compared on the Fig. 2.1. As it can be seen on a graphic the spectral resolutions of the satellites are very similar with exception for thermal bands for Landsat 7 and 8.

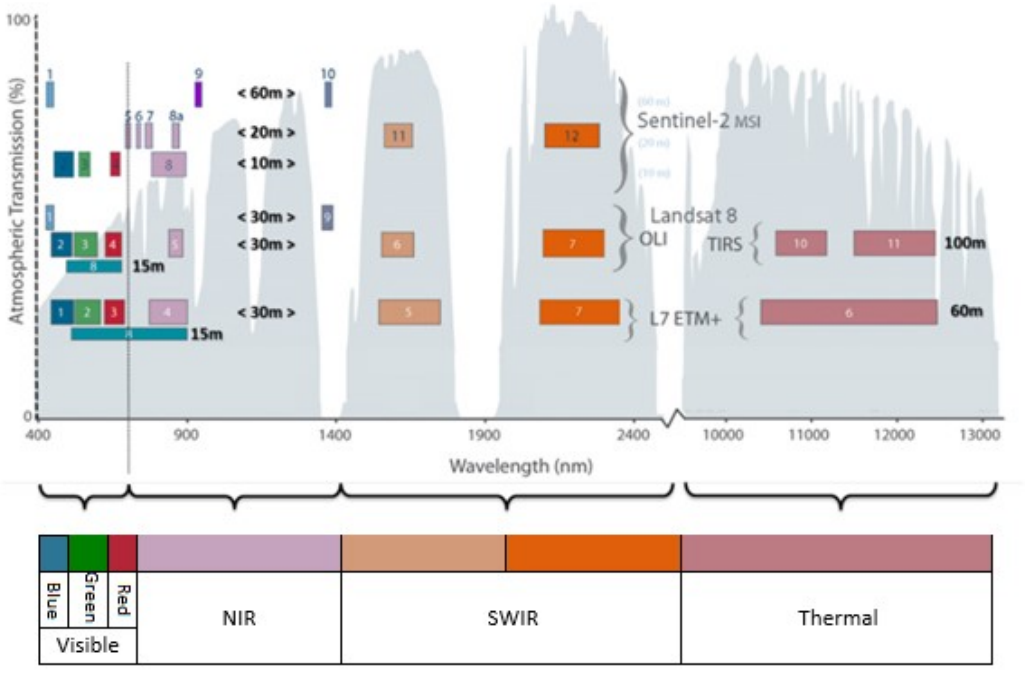


Figure 2.1: Comparison of the spectral resolution of the Sentinel-2 , Landsat 7 and 8 [1]. The colors of the graphic correspond to the electromagnetic spectrum for each sensor being considered.

2.2 Overall Importance of Glacier Mass Balance Estimation and Land Cover Mapping

The GCOS recognizes glaciers and LC as key climate change indicators [30]. The GMB represents the difference between accumulation and ablation of a glacier within a given time period. The glacier mass changes are the response to changing meteorological conditions at the snow/ice surface of the glacier in the longer term. The high levels of melting (or diminishing) of glaciers contributes to the sea level rise. Glaciers act as naturally storing and releasing reservoirs for water in a cyclic manner. Moreover, meltwater from glaciers plays an important role in hydro-power generation, irrigation and drinking water supply. Monitoring of glaciers is therefore crucial for climate sciences. Thus, glacier monitoring follows the Global Hierarchical Observing Strategy (GHOST), which includes better understanding of glacier and snow processes and analysis of GMB TS [31]. Although, some glaciers are being permanently monitored using *in-situ* campaigns, this solution is unfeasible at large scale due to its high financial and logistic cost. This is why RS sensors mounted on satellites provide a complementary source of information (available periodically even for inaccessible areas) at large scale. Hence, several methods in the literature exploit these sources of information for monitoring of distributed GMB.

The LC is the second essential climate variable analyzed in this thesis. LC mapping is essential to provide information about the physical coverage of Earth's surface. It distinguishes different coverage types (classes), such as, water, grasslands, croplands, forest, wetlands, ice. The analyses of LC helps to quantify surface energy and the sources of greenhouse gasses. The variations in LC are used to assess and monitor the impact on occurrence of extreme events e.g., droughts, floods or hurricanes. Indeed, LC change can be seen as both a cause and a consequence of climate change either if it is caused by the human actions or by the natural events itself. Moreover, the LC changes lead to the degradation of the Earth ecosystems, such as reduction of vegetation cover, loss of biodiversity or pollution. Therefore, the LC is an essential input to climate change models to understand the global changes [32], as the evolution of LC has dramatically increased over past decades and is expected to continue [33]. LC monitoring is made possible due to the availability of RS data, which periodically provide information over large areas. Although the LC classification was one of the first application of RS and several global LC maps have been produced, new missions with better temporal and spatial resolution require methodologies, which have to deal with big Earth observation repositories (of data coming from different sensors).

2.3 Methods for Glacier Mass Balance Estimation

In this chapter, we describe methods based on the analyses of the optical RS data to estimate GMB. In greater detail, we discuss two main group of methods: i) based on the optical RS data, and ii) based on the integration of the optical RS data with the physically-based simulations. The first group includes methods based on the multitemporal analysis of the multispectral RS data to identify the glacier surface characteristics as well as its spatio-temporal changes. In particular, we focus the attention on methods based on the analysis of the: equilibrium-line altitude (ELA), glacier-wide surface albedo, glacier snow cover mapping and regression models that exploit *in-situ* measurements and RS data. The ELA methods are based on the detection of the equilibrium-line, which is then used to estimated the GMB. This method, which does not require *in-situ* measurements, is particularly interesting for those glaciers where no ground truth is available. The equilibrium-line is the boundary between accumulation and ablation on a glacier, where the GMB is equal to zero. Therefore, by analyzing the ELA variations, the GMB can be estimated. It is commonly accepted that the end-of-summer glacier snow line altitude (SLA) (which can be detected from the RS data) is a good approximation of the ELA. Thus, the ELA method relies only on RS data and can be applied to glaciers where no ground truth exist. In order to obtain GMB from the ELA variations, first the difference between the theoretically ELA, which is expected when the glacier is in steady state (i.e., distributed GMB = 0) and the actual SLA is computed for each year. Then, the computed ELA variation is multiplied by the mass-balance gradient across the ELA. Mean values of the mass-balance gradient are publicly available for some glaciers or can be computed if the ground-truth measurements exists. The ELA approach has been applied and validated on mid-latitude glaciers showing promising results [9], [34]. Most of the methods proposed in the literature for the detection of ELA rely on MR imagery, such as Landsat 5, 7, 8 and HR imagery, such as Sentinel-2 or SPOT. This is because the higher is the spatial resolution of the sensor the more accurate is the detection of the ELA. However, it should be mention that these methods can work with medium resolution data when the analyzed glaciers are wide. In 2020 the method have been applied at a regional scale to estimate the annual GMB of 239 glaciers in the European Alps. The authors exploited the publicly available Landsat 5, 7, 8 and Sentinel-2 data to drive the estimates between 2000 and 2016 [35]. However, differently from the standard ELA approach (when SLA is manually detected), in this study the authors semi-automatically identified the SLA. This was done by detecting the steepest gradient of the ratio between the near infrared and the shortwave infrared channels. Then, the highest transient SLA was selected as the representation of the glacier's annual SLA. Although the SLA is calculated on the central part of the

glacier, due to ambiguities in the melting process and the cloud-coverage the equilibrium-line may not always be clearly visible. This strongly limits the use of such method from the operational view point.

The glacier-wide surface albedo and glacier snow cover mapping methods can be applied for the estimation of the annual and seasonal GMB. In particular, the albedo method provides accurate results for the estimation of the summer GMB, since the correlation between in-situ measurements and the albedo signal is lower during winter season [36]. In contrast, the methods based on snow mapping provide good results in winter when high variations of seasonal altitude of snow can be observed. Albedo describes the ratio between the solar radiation that is reflected and absorbed by a surface. The albedo metric is well correlated to the GMB since it increases during accumulation period and decreases with ablation [37]. This kind of approach converts the spectral reflectances to albedo, and averages them over the glacier surface. Then the absorption sum (1 - averaged albedo) is weighted by the solar radiation and integrated over the summer period. Nevertheless, it should be pointed out that to properly convert the spectral reflectances to albedo, the method relies on using *in-situ* data. Such method achieves more robust results compared to the ELA approach since it does not rely only on the boundary between accumulation and ablation (which for some cases can be ambigious), but considers on the mean surface albedo all over the glacier. The method has been widely used [38], [39], achieving particularly good results in estimating summer GMB. Nevertheless, it requires to use several cloud-free images for the considered period, which may be a critical requirement over glacier's areas especially during the winter season. In contrast to the ELA method, the spatial resolution of RS data is less important for efficacy of the albedo and snow mapping methods. However, more attention should be paid to the temporal resolution of the sensor as these methods require a few cloud-free images. This is why, in the literature these methods have been often used with Advanced Very High Resolution Radiometer (AVHRR), Moderate Resolution Imaging Spectroradiometer (MODIS) and PROBA-V imagery[38], [39].

Alternative approach performs snow mapping by combining a DTM and the Normalized Difference Snow Index (NDSI) retrieved a from set of the multispectral images. The NDSI is proportional to the amount of snow in the pixel and is efficient for snow monitoring [40]. First, the NDSI is calculated for all available images and is averaged for winter and summer seasons. Then, the seasonal attitudinal distribution of the NDSI is retrieved by combining the seasonal NDSI maps with the glacier topography provided by the DTM. The mean altitude of snow is estimated by intersecting the seasonal attitudinal distribution of the NDSI and the average NDSI. Finally, a linear regression is estimated between the mean regional altitude of snow and the *in-situ* measurements of GMB. The method

have been applied to 55 glaciers in the European Alps, achieving particularly good results during the winter season due to high winter variation of snow [41]. However, it requires availability of constantly updated DTM and similarly to the albedo based method several cloud-free optical RS images for each season. Finally, in 2020, a novel approach that combines RS data, *in-situ* measurements and statistical techniques to estimate GMB was proposed [42]. In particular the method defined the regression model between *in-situ* GMB measurements and the accumulation area ratio (AAR), which is computed as the accumulation area divided by the total glacier area. Both the accumulation area and the total glacier area are derived from the analysis of the NDSI of the multitemporal multispectral RS data. The method has been successfully used for the analysis of Parvati glacier in Himalaya from 1998 to 2016. However, this approach requires *in-situ* measurements, which are unavailable or limited for many glaciers around the world.

The second group of methods is based on the integration and assimilation of optical RS and hydro-climatological simulations. Callegari et al. proposed to estimate the GMB by integrating the AMUNDSEN (Alpine Multiscale Numerical Distributed Simulation Engine) hydro-climatological simulations with binary snow cover maps derived from Landsat data [14]. The authors proposed two correction strategies. The former, correction on disagreement, is limited to pixels where the binary snow cover map disagrees with the GMB computed with AMUNDSEN. The latter, spatially distributed correction, defines a generic parametric function by maximizing the agreement between a binary snow cover map and the snow cover map generated by the model. The method improves the GMB estimates but does not provide spatially distributed GMB maps. The optical RS data have been also assimilated into the physically-based models to perform model calibration. In an attempt to improve discharge simulations, Thirel et al. (2013) assimilated the snow cover areas of the MODIS products into the LISFLOOD physical distributed hydrological model using a particle filter [15]. Although the method provides good results in small basins, the results obtained in spatially heterogeneous areas are not satisfactory. In [16] Finger at al. have investigated the value of different data sources, such as: discharge and GMB measurements, and MODIS imagery to improve the performance of the physically-based simulations. The analysis was carried out in between 1994 and 2007 and showed that the combination of all three data sources leads to the best model performance. While in [43] authors have calibrated HBV (Hydrologiska Byrans Vattenbalansavdelning) using MODIS images over 148 glacier catchments in Austria. The obtained results demonstrate that the assimilation helps to improve snow cover and runoff simulations. Nevertheless, the physically-based simulations are computationally expensive and the estimates further away from the meteorological stations become more uncertain. Finally, the obtained results are highly dependent on the quality of the observations and the input model.

This brief analysis of the literature points out that several methods regarding the analysis of optical RS data for the estimation of the distributed GMB have been proposed. This can be attributed to the comprehensive spatial representation of the glacier area that can be achieved by the RS techniques. Nevertheless, it is noteworthy that SAR sensor constitutes valuable source of information as it can detect the start and stop of the melting period, due to the presence of liquid water [44]. Thus, methods based on SAR data have been widely applied for the GMB estimation [45], [46]. It has been shown that the analysis of the SAR interferometric data are able to demonstrate that ice–ocean interaction have a significant impact on the acceleration of increase in ice-sheets mass loss in the Antarctic and Greenland [47]. Moreover, the aforementioned ELA method have been also successfully adopted with SAR imagery providing reliable results[48]. However, main drawback of using SAR sensors for snow and glacier monitoring is the possibility to detect just the wet snow as in the case of dry snow the reflectance comes mainly from the underlying layers [44]. Finally, it it noteworthy that technique based on satellite altimetry have been also successfully exploited for the estimation of mass losses [49], [45]. Among the others, the recent study has estimated the mass losses in the Greenland from 1992 to 2018 [50]. However, looking at the literature one can notice that methods being able to take advantage from different information sources are still not enough investigated. It is well known that *in-situ* campaigns are the most accurate techniques for GMB estimation, but as they are very laborious and time consuming, they can not be used over large areas. On the other hand the physically-based hydrological simulations are also valid source of information. However, they are resource intensive and the estimates further away from the meteorological stations are associated with higher uncertainty. In this context, the complementary information provided by the RS data, *in-situ* measurements and physically based-models represent an interesting solution for the accurate estimation of the distributed GMB.

2.4 Land Cover Mapping

To respond to the need of monitoring the Earth's surface at a large scale, in the last decades many global LC products have been produced. Tab. 2.1 presents an overview of the main global LC products publicly available by comparing their temporal coverage, spatial resolution and the number of classes present in the map legend.

IGBP-DISCover global land was created by the United States Geological Survey (USGS), the University of Nebraska-Lincoln, and the European Commission's Joint Research Center. The 1 km spatial resolution AVHRR data acquired from April 1992 to March 1993 were used. The majority of the LC classes were mapped by using first un-

supervised clustering, which was then followed by a post-classification refinement applied separately to each continent. The snow and ice and the barren were classified by thresholding the maximum Normalized Difference Vegetation Index (NDVI) (of 12-months composites) while the water bodies were classified by using hydrography layer of the Digital Chart of the World. The accuracy of the IGBP-DISCover global LC was validated by three independent photo-interpreters in accordance to two strategies: i) when decision of all photo-interpreters agree and ii) when the majority of photo-interpreters agree. The classification accuracy was equal to 59.4% and 73.5% for the first and the second strategy, respectively. Similarly to IGBP-DISCover the University of Maryland generated the global LC map by using the 1 km spatial resolution AVHRR data in between 1992-1993. However, the number of LC classes has been reduced to 14 in comparison to the 17 classes considered in IGBP LC classes. In particular, the crop-natural vegetation, wetland, and ice / snow classes were ignored. The decision tree classifier was used for discrimination of the majority of classes (build-up class and water class were taken from the ancillary data). In [19] authors reported the map classification accuracy equal to 69%. The other global LC map produced at 1 km spatial resolution (by using SPOT 4 satellite) in 2000 is the Global Land Cover 2000. This LC map is the result of international cooperation of 30 research groups organized by the European Commission's Joint Research Center. The idea was to divide global mapping into 19 regions and exploiting knowledge from the local experts. The proposed legend was based on the United Nation Land Cover Classification System (LCCS) and had two levels: first level, which includes regionally optimized 44 LC classes, and second level which joins first level regional LC classes into 22 classes at the global scale. For each of 19 regions the corresponding scientific group performed classification independently using monthly and seasonal composites. Ancillary data such as: information from Defense Meteorological Satellite Program were exploited to improve classification accuracy of build up and swamp forest. The Overall Accuracy (OA) of the Global Land Cover 2000 is equal to 68.6%.

Moving to higher accuracy, the GlobCover map was produced by using 300 m spatial resolution imagery from Medium Resolution Imaging Spectrometer (MERIS) on board of Environmental Satellite (ENVISAT) in 2009. The LC legend was based on LCCS considering 22 global and 51 regional LC classes. First, the entire globe was divided into 22 equal areas characterized by similar ecological conditions. This is done to enhance the spectral differentiation between classes. The processing chain included standard optical image preprocessing, where the geometric correction is done using AMORGOS (Accurate MERIS Ortho-Rectified Geo-location Operational Software) tool [51], followed by the atmospheric corrections and cloud/snow pixels screening. The input to the classifier are seasonal composites generated as the average of acquisitions over considered period. The

unsupervised classification based on ISODATA algorithm was carried out at pixel level, followed by the extraction of the phenological parameters for each of the identified classes. Then the clustering is performed to identify classes with similar spectro-temporal characteristics. Finally, the identified classes are labeled using referenced-based function defined by regional LC experts [52]. The OA of the map is equal to 67.5%. The other maps available at 300 m spatial resolution were produced in 2000, 2005 and 2010 in the framework of the CCI produced by ESA. These maps were generated by using MERIS-ENVISAT acquired from 2003 to 2012 to produce a 10 year global LC map. Then, based on back up-dating techniques and using SPOT Vegetation TS (acquired between 1998- 2013) the 2000, 2005 and 2010 maps were generated. The legend is hierarchically structured and consist of 22 and 16 LC classes in the first and second level, accordingly. The images were first radiometrically, geometrically and atmospherically corrected and pixel identification was carried out (i.e. attribution of a status to each pixel, being "land", "water", "cloud/ cloud shadow", "snow", and invalid pixels). MERIS-ENVISAT and SPOT images are then processed to generate averaged seasonal based composites, which are input to the classification module. The world is first stratified into equal-reasoning areas, which are characterize by similar ecological conditions (i.e., seasonality, cloud coverage). The classification is performed independently for each strata, and thus the processing chain can be properly tuned to the specific conditions of considered regions. The training set leverages the publicly available global, regional and local LC maps, which are translated in accordance with CCI legend, then resampled to 300 meters spatial resolution and merged. The classification is divided into 4 steps. In the first two steps the supervised and unsupervised classification is performed, resulting in two classification maps. Next, in the third step those maps are merged, while in the fourth step the auxiliary information (i.e., CCI water bodies product, CCI urban product) are integrated to provide the final LC map. The validation of the CCI LC map in 2015 was carried out using the GlobCover 2009 validation set. The OA was reported by using only certain samples labeled by experts and it was equal to 71.1% when considering "homogeneous" (i.e. made of a single LC class) or "heterogeneous" (i.e. made of several or mosaic LC classes) and 71.7% when considering only "homogeneous" samples.

The Copernicus Global Land Service (CGLS) delivered Copernicus Global Land Cover (CGLC) map for 2015, 2016, 2017, 2018 and 2019. This global LC product is a medium spatial resolution map having pixel grid of 100 m. The map is produced by considering the TS of multispectral PROBA-V 100 m images, a dataset of more than 140 K crowdsourced (Geo-WIKI) LC training points and other ancillary data. Since groundtruth points were collected through manual visual interpretation of Google Maps and Bing images at 10 m spatial resolution, the CGLC product at 100 m derives also information

Table 2.1: Comparison of the main publicly available global LC maps in terms of: temporal coverage, spatial resolution and number of classes present in the map legend.

LC map name	Temporal resolution	Spatial resolution	Spatial coverage	# Classes
IGBP-DISCover [20]	1992	1 km	Global	17
University of Maryland [19] Land Cover	1992	1 km	Global	14
Global LC 2000 [53]	2000	1 km	Global	44
GlobCover 2009 [54]	2009	300 m	Global	22
MODIS GLC [55] (Collection 5)	2001 to 2013 (annually)	500 m	Global	17
MODIS GLC [56] (Collection 6)	2001 to 2019 (annually)	500 m	Global	23
GLC-SHARE 2014 [57]	2014	1 km	Global	11
Climate Change Initiative [7]	2005, 2010, 2015	300 m	Global	38
Copernicus Global [58] Land Cover	2015 to 2019 (annually)	100 m	Global	23
FROM-GLC [59]	2017, 2010	30 m	Global	10
GlobeLand30 [60]	2000, 2010	30 m	Global	10

about the fractional cover layers for the ten base LC classes: “forest”, “shrub”, “grass”, “moss and lichen”, “bare and sparse vegetation”, “cropland”, “built-up / urban”, “snow and ice”, “seasonal and permanent inland water bodies”. Moreover, the CGLC product provides a forest type layer offering twelve types of forest, quality indicators for input data (data density indicator), for the discrete map (probability) and for six of the fractional cover layers. The CGLC map provides discrete classification at three levels according to LCCS. While in the first level the map presents the main twelve LC classes (i.e., “forest”, “shrub”, “herbaceous vegetation”, “herbaceous wetland”, “moss and lichen”, “bare / sparse vegetation”, “cultivated and managed vegetation - cropland”,“ built-up”, “snow and ice” and “permanent inland water bodies”), in the third level more detailed information are available by distinguishing 23 classes. For instance, the forest classes are divided into six types of closed and six types of open forest. The PROBA-V images were first geometrically projected into the UTM (Universal Transverse Mercator) coordinate system (to reduce distortions in the high north latitudes) and atmospherically corrected. Next the median 5-days-based composites were generated and the data density indicator (DDI) was calculated. Instead of using spectral bands as an input to the classifier the following indices have been computed: NDVI, Enhanced Vegetation Index, Structure Intensive Pigment Index, Normalized Difference Moister Index, NIR reflectance of vegetation, Angle at NIR, HUE, and the Hue Saturation Value color system transformation, Area Under the Curve, and Normalized Area Under the Curve. For each of the aforementioned indices the statistical metrics (i.e., minimum, maximum, median, quantile) and textural metrics (i.e., median variation) were computed. Moreover, external metrics retrieved from DTM (i.e., height, slope, aspect) were considered. Due to very long TS the data were condensed trough extraction of temporal features (done by using harmonic model). The classification was run independently within clusters defined using ecological information. In order to reduce the number of the metrics, the analyses of their separability were carried out and the metrics identified as redundant were removed. The training set plus the DDI quality metric are used as an input to the random forest (RF) classifier to provide the pixel based LC map and class probability. Moreover, the RF regression trained on the photo-interpreted cover fractions was used to provide the vegetation coverage per pixel. To generate the final CGLC map, the obtained classification, probability and regression maps are integrated together with the auxiliary data, such as World Settlement Footprint [61] and Global Surface Water [62]. The validation of CGLC product was performed by using independent validation data of around 21 600 samples generated in collaboration with experts. The OA of the CGLS 100m map is equal to 80.2 % +/-0.7 % depending on the continent. The highest classification accuracy (> 85%) is achieved for the forest, snow/ice, bare /sparse vegetation, and permanent water. For the herbaceous vegetation,

croplands, urban CGLC achieve moderate accuracy, the lowest accuracy of (< 65%) is performed for herbaceous wetland, lichen and moss and shrubs.

Only few gobal products are available at high spatial resolution, i.e., the Finer Resolution Observation and Monitoring of Global Land Cover (FROM-GLC) [59] and GlobeLand30 [60]. The FROM-GLC is the first 30 m spatial resolution global LC product [59] available for 2010 and 2017. The map was produced by using Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) data. The legend includes the following 10 classes: "cropland", "forest", "grassland", "shrubland", "wetland", "water", "tundra", "impervious surface", "bareland" and "snow/ice". The authors compared the classification accuracy by using four standard classifiers i.e., conventional maximum likelihood classifier (MLC), J4.8 decision tree classifier, RF classifier and support vector machine (SVM). On the global validation data set the best OA equal to 64.9% was produced by SVM, compared with 59.8% by RF, 57.9% by J4.8 and 53.9% by MLC. However, due to the lack of temporal information used in producing the FROM-GLC, in [63] authors have used a segmentation-based approach to introduce temporal information from the MODIS and other auxiliary data, such as bioclimatic, DTM and world maps on soil-water conditions. It is important to note, that temporal information are particularly necessary for proper discrimination between some vegetation cover types (e.g. grasslands vs croplands). After introducing temporal information the OA have increased to 67.08%. In 2017 by using all-season training and validation sets together with spatial-temporal partition strategy the accuracy of FROM-GLC was improved to 70.17% [64]. In 2020 LC mapping at 30-meters resolution was further improved to 81.5% [65]. This was done by integrating Landsat Operational Land Imager (OLI) and ETM+ imagery with free and public high-resolution Google Earth images. The approach is a hybrid of three components: convolutional neural network (CNN) to extract features from GE imagery, SVM and RF based on Landsat spectral features and the final decision fusion module. The other global LC product available at 30 meters resolution is the GlobeLand30, which overall classification accuracy is 80.33% [60]. The product is available for the two base-line years 2000 and 2010. The legend similarly to FROM-GLC includes 10 classes: "cultivated land", "forest", "grassland", "shrubland", "wetland", "water bodies", "tundra, "artificial surfaces", "bareland" and "permanent snow and ice". The primary images used for the GlobeLand30 generation are the Landsat TM/ETM+ and Chinese Environmental and Disaster satellite HJ-1 (Huan Jing-1), due to its similar spectral and spatial characteristic to Landsat data. Moreover, to enhance the temporal resolution the MODIS-NDVI were also included. Due to the wide swath of the HJ-1 the images are affected by the nonlinear distortion, and thus they were first geometrically corrected. Next, Landsat and HJ-1 images were atmospherically corrected. The topographic correction was done by

using Continuum Removal (CR) method, which leverages spectral information of shadow and neighboring no shadow pixels without the use of DEM [66]. The cloud cover and missing data due to the failure of the Landsat ETM+ Scan Line Corrector (SLC) have been restored using the neighborhood similar pixel interpolation (NSPI) approach [67]. The classification is based on pixel-object-knowledge approach, which first exploits pixel based classification and than to minimize the "salt-and-pepper effect" the object based classification was performed (using eCognition (v8) software). Finally, the pixel based results, the segmented objects and auxiliary data (i.e., existing LC maps, DEM, Geo-Wiki) are integrated using decision based fusion to provide the final LC map.

Looking at the available thematic products, it can be noticed that no recent global HR maps have been released. Moreover, differently from the medium resolution products such maps have been produced considering only the main LC categories and only for the specific years. However, in order to perform accurate climate change monitoring more detailed classification scheme and more frequent LC maps are required. Furthermore, high spatial resolution classification maps help to better resolve local climate features (which are connected to the heterogeneity of the surface), and thus support climate modeling research. Therefore, within this thesis we would like to address these issues, by proposing a processing chain that can be applied at sub-continental scale to produce multitemporal HR LC maps having a detailed classification scheme.

Chapter 3

Integrating Models and Remote Sensing Data for Distributed Glacier Mass Balance Estimation

This chapter presents an approach to improve the estimation of the GMB. based on combination of three complementary data sources, hydro-climatological model, RS data and ground measurements. The hydro-climatological model provides spatially distributed mass balances at low resolution. RS supplies spatially distributed surface characteristics at HR. The ground point measurements provide the mass balance at the local scale. The combination of these data sources allows us to improve the spatial resolution of the model output and its GMB estimates. We used the AMUNDSEN model, which considers the processes of accumulation and ablation of snow and ice for the area of an entire glacier. In the proposed integration approach, we first compute the deviations between the GMB simulations (afforded by the hydro-climatological model) and the ground measurements. Then, the RS data are used to define a feature space in which objectively characterizing the glacier surface properties). The method estimates the adjustment required to the model at a pixel level. This is done by leveraging each considered sample neighborhood of labeled samples (for which the deviation is known) in the defined feature space. Thus, we are able to apply similar adjustment to samples sharing similar glacier surface conditions. Experimental results confirm the effectiveness of the proposed approach. [1]

[1]Part of this chapter appears in I. Podsiadlo et al., "Integrating Models and Remote Sensing Data for Distributed Glacier Mass Balance Estimation," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 6177-6194, 2020, doi: 10.1109/JSTARS.2020.3028653.

and I. Podsiadlo et al. "Integration of hydro-climatological model and remote sensing for glacier mass balance estimation." Image and Signal Processing for Remote Sensing XXV. Vol. 11155. International Society for Optics and Photonics, 2019.

3.1 Introduction

Glaciers are very sensitive to climate change and this is why they have been identified as one of the terrestrial essential climate variables. There are many glacier characteristic (e.g., surface area, length), which can help to monitor glacier cover. The GMB is particularly interesting for climatologist as it is driven by the atmospheric conditions.

Apart from the approaches based on optical RS data to estimate the spatial distribution of GMB (see chapter 2.3) the glaciological, the geodetical and the hydrological methods [68] have been widely used. The glaciological method is based on local observations of snow accumulation (measured with snow pits) and local ablation of glacier ice (measured with stakes). This method is therefore time-intensive and limited to a few glaciers. The geodetic method requires accurate topographies of the glacier surface for the moment of minimum snow coverage. In the early days, these were provided by handmade terrestrial geodetic observations, whereas today mostly RS techniques, like laser scanning, are applied. The hydrological method is usually applied by means of a (more or less) physically-based model that describes (or estimates) the physical processes of snow and ice mass accumulation, redistribution and ablation. For that purpose, a variety of modelling approaches exist [69, 70]. In this chapter, we apply AMUNDSEN [71]. It is a hydro-climatological model that has proven its reliability in many applications [72]. AMUNDSEN computes hourly balances of energy and water mass (i.e., snow and ice) at the pixel scale, usually in resolutions of 10 to 50 m. To consider glacier motion, the ice body geometry is re-adapted once per year following the Δh approach [69]. The model has been intensively validated in all dimensions of time and space, with a collection of different data sets (including HR TS of local observations, spatially distributed satellite data and glacier-wide mass balances) [73].

The AMUNDSEN model facilitates the simulation of all relevant fluxes of energy which strongly depend on the topographical conditions (affecting all processes in high mountain areas). As a result, the model computes hourly melt rates of snow, ice and firn, and integrates the hourly computations of accumulation and ablation to annual mass balances for the entire area of a glacier. The accuracy of physically-based models largely depends on the forcing data [74]. Other than empirical models, physically-based models do not require excessive tuning or parameter calibration from a theoretical point of view.

Forcing data, parameter choice, model initialization, model structure and the related uncertainties propagate through the model chain and can affect simulation results in non-linear ways [75]. This is especially the case when meteorological forcing data are subject to high uncertainties (like for precipitation and its phase which are difficult to measure in mountain regions). In these circumstances the model results can diverge from ground-

truth observations. This can be problematic in remote regions, with only few stations, where even nowadays simple models have to be applied. In regions like the European Alps, meteorological forcing data are interpolated from several stations around a glacier or derived from output of climate models. Their bias may also be corrected with proper methods. Apart from the difficulties in obtaining accurate precipitation estimates, the complex topography of high alpine regions complicates the prediction of wind fields and solar radiance. The applied hydro-climatological model also needs to adjust the simulated spatial discretization to the required resolution in order to represent a physical process (depending on topography) in accordance with the scale of the observations. For fully spatially distributed (gridded) simulations these considerations often lead to a grid-size of 10s to 100s.

From the analysis of the literature one can conclude that the combination of physically-based mass balance modeling, RS data acquired over the glacier surface, and ground measurements could be a promising way of improving the quantification of a GMB, and to overcome some of the weaknesses inherent in each of these methods. One of the weakness of physically-based models is that the further away we are from a meteorological station, the higher the uncertainty in the interpolation method of any model. In contrast, RS data provide a comprehensive spatial representation of the glaciers, but are affected by clouds occlusions and require a large number of *in-situ* measurements to train supervised methods to determine the glacier surface characteristics. Finally, the collection of field data is time-consuming and hampered by lack of access in such complex environmental areas, so typically only few measurements are collected [76]. The aim of this contribution is to propose a novel integration approach, which makes use of the complementary advantages afforded by model simulations, RS data and *in-situ* measurements. This could help to estimate a spatially distributed annual GMB that is better than any of the classical methods can provide. Even though RS data with bias deviation correction strategies have been proposed in the literature to improve model parameter estimation [11], this approach has never been used to improve existing physical-models and GMB model calculations. The main contributions of this work are: (i) using an existing (state-of-the-art) distributed hydro-climatological model to compute the processes of mass accumulation and ablation, (ii) using RS data to define an informative feature space (where samples characterized by similar glacier surface characteristics can be easily detected) and (iii) efficiently integrating the hydro-climatological model, the RS data and *in-situ* observations to accurately estimate the spatially distributed annual GMB. The method was applied to two clusters of glaciers: one in the Austrian Alps, and another in the Italian Alps. We used HR Sentinel-2 multispectral images for the experiments performed after 2015, while MR Landsat-5 and Landsat-7 data were considered for the ones carried out before 2015.

A 10 m spatial resolution DTM was used to extract the topographic parameters of all four glaciers in HR.

The remainder of this chapter is structured as follows. Section 2 describes in details the proposed integration method. Section 3 presents the study areas, the RS data, and the local observations, and section 4 presents the results obtained with our integration approach. Finally, in section 5 we draw the conclusion and propose ideas for future developments.

3.2 Proposed GMB Estimation Method

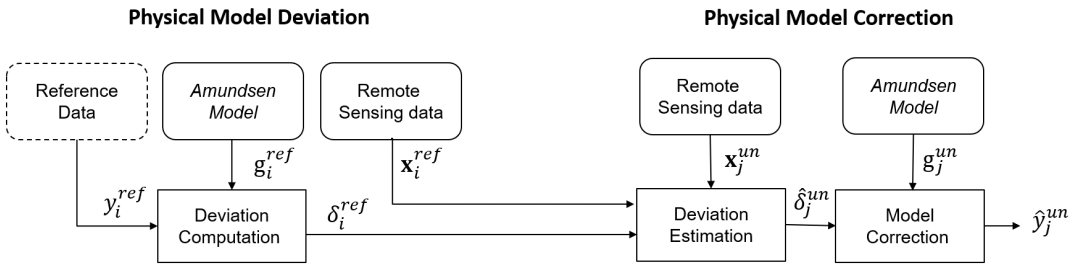


Figure 3.1: Architecture of the proposed integration approach. The method takes advantage of the complementary information provided by *in-situ* measurements, RS data and hydro-climatological model to accurately estimate GMB.

The proposed integration method aims to correct the downscaled GMB estimates, provided by the hydro-climatological model AMUNDSEN, by efficiently integrating: *in-situ* measurements, the MR and HR multispectral optical images and the DTM. In contrast to previous studies, the proposed integration approach does not assimilate *in-situ* measurements into a model but performs the fusion of RS data, physical model and *in-situ* measurements. This fusion preserves the generalization of AMUNDSEN (to properly describe the physical processes of the whole glacier) while improving the accuracy of its estimates. The AMUNDSEN model is run for the whole study area to generate a spatially continuous GMB map. The proposed integration method assumes that the *in-situ* measurements and the physical model data are available on the considered test sites over the same time period. The joint use of these complementary data sources allows us to: (i) improve the accuracy of the GMB estimates, and (ii) increase the spatial resolution of the GMB maps provided by the AMUNDSEN model.

Fig. 3.1 represents the block diagram of the proposed integration method that can be divided into two steps: (i) physical model deviation computation - where the hydro-climatological model outputs and the *in-situ* measures (reference data) are used to com-

pute the model residuals on a small set of available reference data; and (ii) physical model correction - where the topographic and spectral information provided by the RS data defines a feature-space to accurately model the environmental properties of the glaciers. By representing the problem in a feature-space, for each unlabeled sample (i.e., glacier pixel without associated *in-situ* measurement), we are able to detect the reference samples sharing similar glacier conditions (i.e., located in the same feature sub-space). Finally, by leveraging the deviations of the reference data (computed in the previous step), we correct the model estimates. In the following subsections, each step of the proposed integration approach is described in detail. Moreover, the feature space used to perform the deviation estimation is presented. In particular, we explain the physical meaning of the multispectral and topographic variables employed to model the glacier environmental conditions.

3.2.1 Feature Space Definition

Selecting a proper feature space is a crucial step in order to effectively represent the variables affecting the environmental conditions of a glacier, and thus the behaviour of GMB. To define a representative feature space, we select the most informative spectral bands and topographic parameters for snow and ice monitoring, in accordance with the Earth observing system analysis [77] and experimental analysis performed. From the DTM, we extract the altitude, slope and aspect parameters that characterize the topography of a glacier, which is correlated to the GMB.

From the multispectral optical data we consider the blue spectral channel (450 nm – 520 nm), which provides information about potential snow pixels, and the SWIR bands (1050 nm - 2500 nm), which are typically used for snow and ice detection. Moreover, due to the sensitivity of the SWIR bands to moisture content, they can distinguish old snow from fresh snow. This is due to the fact that old snow often develops a solid crust which corresponds to a rise in moisture content, and thus it is less reflective in the SWIR than fresh snow. In addition, we also include the Normalized Difference Water Index (NDWI), which is correlated to the amount of water per unit area. This metric has been proven to be correlated to the beginning of the snow melt period [78], [79].

It is important to mention that the annual GMB is reported on the September 30^{th}. However, in the proposed integration method we utilize multispectral images acquired at the end of August or at the beginning of September (see Tab. 3.1a and 3.1b). This is because the maximum ablation of the considered glaciers can be observed in this period. On the 30^{th} of September glaciers are usually fully covered with fresh snow. Thus, we are no longer able to differentiate the types of a glacier cover (such as firn, old snow, fresh snow or ice) as the spectral signatures correspond to fresh snow only. The experimental results

and wider literature review suggest [14], these images are a good proxy for estimating the maximum glacier ablation. Moreover, these images are typically characterized by a low level of cloud coverage leading to a good approximation of annual GMB.

3.2.2 Physical Model Deviation Computation

In the first step, the proposed integration method takes advantage from the availability of a small set of *in-situ* GMB measurements to compute the deviation between the reference samples and the corresponding physical model estimates. Let $\{y_1^{ref}, y_2^{ref}, \cdots, y_N^{ref}\}$ be the set of measured GMB and let $\{g_1^{ref}, g_2^{ref}, \cdots, g_N^{ref}\}$ be the corresponding hydro-climatological model estimates. Due to uncertainties associated with numerical simulations, for each sample we can compute the deviation between the physical model and the true GMB value. This can be expressed as follows:

$$\begin{aligned} \delta_i^{ref} &= y_i^{ref} - \mathrm{g}_i^{ref} \\ &\quad where \quad i = 1, \cdots, N \end{aligned} \tag{3.1}$$

The computed training sample deviations are then used to perform the physical model correction. This helps to estimate the final GMB values of the unlabeled samples (i.e., samples for which no ground reference data are available). In order to infer the deviation correction, the method exploits the information provided by the RS data. This enables us to model the physical properties of different parts of the glaciers. In particular, for each unlabeled sample we aim to detect the most similar reference ones, for which the model deviation has been computed. Then, based on the reasonable assumption that samples having similar physical properties are affected by similar model deviations, we estimate the model deviation. For each unlabeled sample (δ^{un}), we compute a correction as an average deviation from the reference samples located in the δ^{un}'s neighborhood.

3.2.3 Physical Model Correction

The aim of this step is to estimate for each unlabeled sample its deviation $\widehat{\delta}_j^{un}$, and thus leading to an accurate GMB estimate $\widehat{y}_j$. To this end, the proposed integration method takes advantage from the spectral and topographic information provided by the MR and HR optical data and the DTM, to define a feature space where the reference and the unlabeled samples are re-projected. Let $\{\mathbf{x}_1^{ref}, \mathbf{x}_2^{ref}, \cdots, \mathbf{x}_N^{ref}\}$ be the RS features associated to the reference samples, with $\mathbf{x}_i^{ref} \in \mathcal{R}^d$ made up of d components derived from the multispectral RS information and the DTM. Let us define with $\{\mathbf{x}_1^{un}, \mathbf{x}_2^{un}, \ldots, \mathbf{x}_M^{un}\}$ the RS features associated to the unlabeled samples, with $M >> N$, and let $\{\mathrm{g}_1^{un}, \mathrm{g}_2^{un}, \cdots, \mathrm{g}_M^{un}\}$ be the corresponding hydro-climatological model estimates. By analyzing this feature

space we can identify for each $\mathbf{x}_j^{un}$ the reference samples characterized by similar environmental and snow/ice conditions, i.e., those located in the same sub-space. Based on the reasonable assumption that the physical model estimation accuracy varies in the different portions of the feature-space, we exploited a local deviation correction strategy. Therefore, in our proposed approach the correction is computed and applied to local feature sub-spaces.

To partition the feature space we considered a simple and efficient non-parametric k-NN algorithm as proposed in [11]. First, the distances (similarities) between each unlabeled sample $\mathbf{x}_j^{un}$ and the training reference samples $\{\mathbf{x}_1^{ref}, \mathbf{x}_2^{ref}, \cdots, \mathbf{x}_N^{ref}\}$ are calculated. The estimated correction value is computed as an average of the k nearest (most similar) samples deviations. Let us that after applying k-NN partition, the feature-space is divided into R regions and let r_z be a generic region. For each region r_z we compute the correction value as follows:

$$\begin{aligned} \widehat{\delta}_j^{un} &= \frac{1}{k}\sum_{i=1}^{k}\delta_i^{ref} = \frac{1}{k}\sum_{i=1}^{k} y_i^{ref} - \mathrm{g}_i^{ref} \\ where \quad \delta_i^{ref} &\in r_z, \quad and \quad j = 1, \cdots, M \end{aligned} \tag{3.2}$$

Where $\widehat{\delta}_j^{un}$ is the correction for the samples within r_z. The tuning of the k-parameter is performed with cross-validation. Here, the distance is computed as Euclidean distance. However, other metrics can be applied. It is worth mentioning that this strategy partitions the feature-space according to the training reference data distribution. The final estimation of the GMB is calculated as the sum of the estimates provided by the physical model and of the correction computed according to the local strategy, i.e.,:

$$\begin{aligned} \widehat{y}_j^{un} &= \mathrm{g}_j^{un} + \widehat{\delta}_j^{un} \\ & where \quad j = 1, \cdots, M \end{aligned} \tag{3.3}$$

This strategy allows the correction of the GMB estimates by using the average deviation of the reference samples (located in a neighborhood of the unlabeled samples to be corrected).

3.3 Study Areas and Dataset Description

3.3.1 Study Areas

To carry out the experimental analysis, we considered two study areas located in the Austrian and Italian Alps (see Fig. 3.2). The first one is the Rofental Valley which

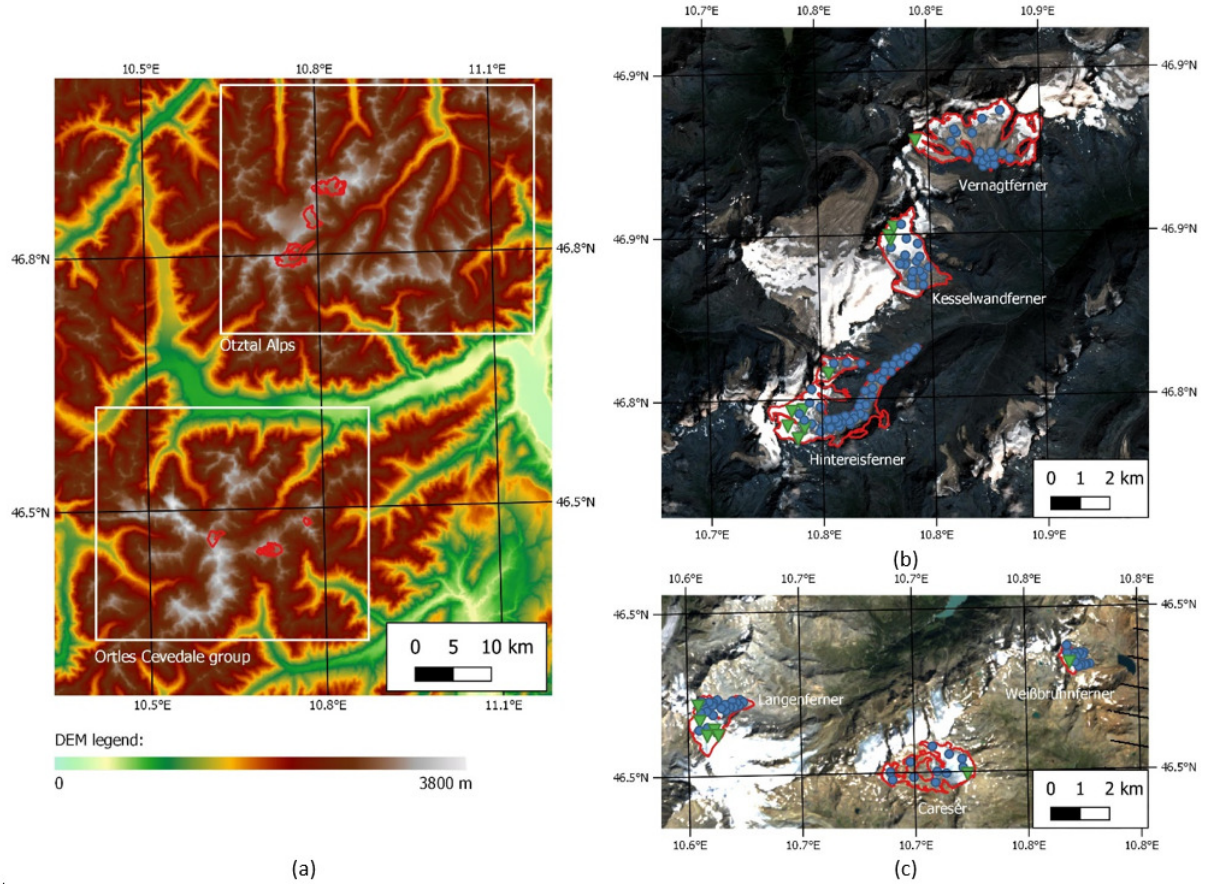
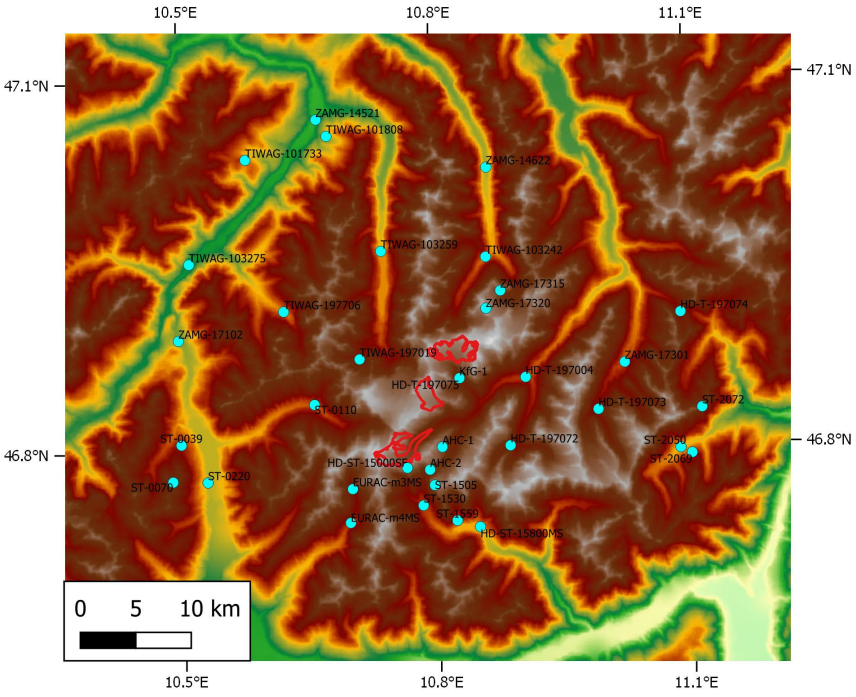
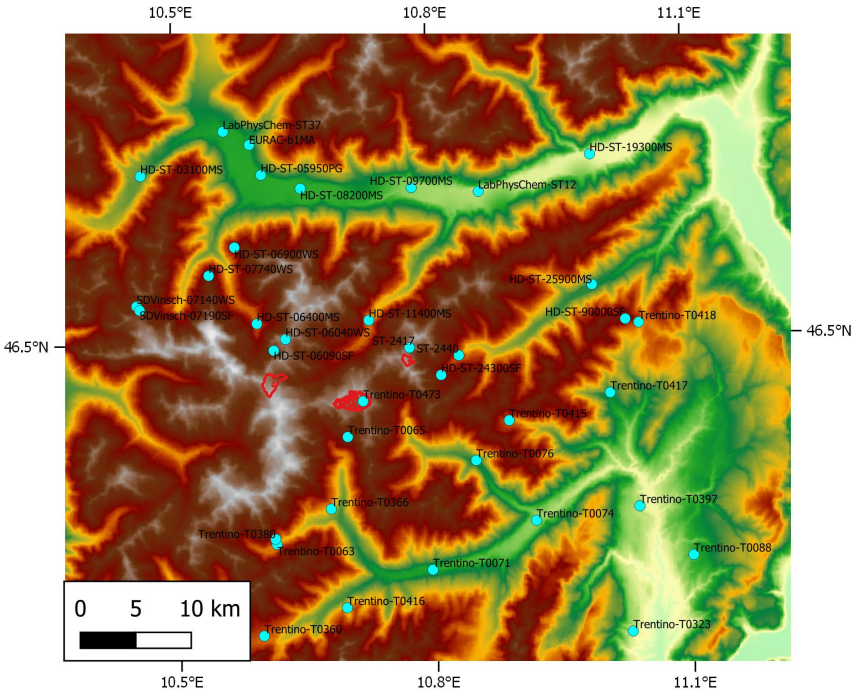


Figure 3.2: (a) Location of the study area with the outlines of the considered glaciers in red. (b) Test site in the Ötztal Alps three studied glaciers: Kesselwandferner, Hintereisferner and Vernagtferner. (c) Test sites on the Ortles-Cevedale group with three considered glaciers: Langenferner – Lunga Vedretta, Careser and Weissbrunnferner – Fontana Bianca. The images (b) and (c) also represent the location of the *in-situ* measurements marked as blue dots for ablation stakes and green triangle for snow pits. The red line represents the glacier outlines.



(a)



(b)

Figure 3.3: Overview of the meteorological stations (cyan circles) used as input to the AMUNDSEN model for the test glaciers in the Oetztal Alps (a) and Orltes-Cevedale group (b) (highlighted in red).

belongs to the Otzal Alps, in the Tyrol region, Austria. This valley, which is one of the most studied in the Austrian Alps, is characterized by well-instrumented and equipped test-sites with weather stations from which measurements are used to drive AMUNDSEN simulations [80]. In particular, we considered the: Kesselwandferner, Hintereisferner and Vernagtferner glaciers. The second study area is located in Ortles-Cevedale group in South Tyrol (Italy). The Ortles-Cevedale is the largest glacierized group in the Italian Alps. In this study we focused on the following glaciers within the group: Langenferner (Lunga Vedretta), Careser and Weissbrunnferner (Fontana Bianca). We would like to remark that in our study area the considered glaciers represent mainly ablation, and thus glacier zero balance line is not present.

3.3.2 Field Reference Data

The GMB reference data were downloaded from the WGMS [81], the service of the International Association of the Cryospheric Sciences. Since 1984, the WGMS has collected glacier data from over 30 countries. Hintereisferner, Kesselwandferner and Langenferner GMB measurements available at the WGMS website were collected by the Institute of Atmospheric and Cryospheric Sciences of the University of Innsbruck. For the Vernagtferner, the data were provided by the Commission for Geodesy and Glaciology of the Bavarian Academy of Sciences. The GMB information for Careser were provided by the Comitato Glaciologico Trentino of the Società degli Alpinisti Tridentini, and for Weissbrunnferner by the Ufficio Idrografico of the Provincia Autonoma di Bolzano - Alto Adige.

The spatial distribution of the GMB *in-situ* point measurements are shown in Fig. 3.2 (c) for the Rofental valley and in Fig. 3.2 (b) for the Ortles-Cevedale group. Tab. 3.1a and 3.1b present the number of GMB point measurements available for each glacier per year.

3.3.3 RS Data

To define an informative feature space for modeling the glacier environmental conditions, we exploited the multispectral information provided by optical satellite sensors and the DTM acquired by the airborne laser scanner (ALS) systems. For the experimental analysis carried out between 2006 and 2013, the multispectral data acquired by the Landsat were used, whereas between 2015 and 2017 the most recent HR multispectral Sentinel-2 data were considered (see Tab. 3.1a and 3.1b).

Table 3.1: Dataset description of the glaciers located in: (a) the Tyrol (Austria), (b) south Tyrol (Italy). For each glacier, the number of *in-situ* measurements (reference samples) available per year and the considered multispectral satellite optical image employed are reported.

Tyrol (Austria)				
Glacier name	**Year**	**# ref. samples**	**Sensor**	**Acquisition date**
Hintereisferner	2015	48	Sentinel-2	2015-08-26
	2016	54	Sentinel-2	2016-08-27
	2017	55	Sentinel-2	2017-08-30
Kesselwandferner	2015	20	Sentinel-2	2015-08-26
	2016	21	Sentinel-2	2016-08-27
	2017	21	Sentinel-2	2017-08-30
Vernagtferner	2012	29	Landsat-7	2012-09-16
	2013	35	Landsat-7	2013-08-02

South Tyrol (Italy)				
Glacier name	**Year**	**# ref. samples**	**Sensor**	**Acquisition date**
Langenferner	2010	26	Landsat-7	2010-08-26
	2011	25	Landsat-7	2011-08-29
Careser	2007	19	Landsat-5	2007-07-25
	2010	12	Landsat-7	2010-08-26
	2011	14	Landsat-7	2011-08-29
Weissbrunnferner	2006	16	Landsat-5	2006-08-23
	2007	16	Landsat-5	2007-07-25
	2010	16	Landsat-7	2010-08-26
	2011	17	Landsat-7	2011-08-29

3.3.3.1 DTM Topographic Maps

The DTM is used to represent topographic parameters of the considered glaciers, such as altitude, slope and aspect. In Tyrol (Austria) we used the DTM coming from the Land Tirol, available at Open Data Austria [82]. In south Tyrol (Italy) the DTM was provided by the Istituto Nazionale di Geofisica e Vulcanologia [83]. Both DTMs were acquired by the airborne laser campaign and are characterized by 10 m spatial resolution.

3.3.3.2 Sentinel-2 Multispectral Images

Sentinel-2 is a mission of the ESA, which comprises two twin polar-orbiting satellites. It provides multispectral images of both high spatial (10 m - 20 m) and high temporal resolutions (5 days on the equator under cloud-free condition). In our study, the 20 m spectral bands are resampled to 10 m with the nearest neighbor approach. The Sentinel-2 images are used over the Rofental valley for the years 2015 - 2017.

3.3.3.3 Landsat-5 and Landsat-7 Multispectral Images

Landsat-5 and Landsat-7 provide MR images at 30 m spatial resolution with temporal resolution of 16 days. The Landsat-5 is used over the Ortles Cevedale group for the years 2006 - 2007, while the Landsat-7 is used over Rofental valley for the years 2012-2013 and over the Ortles Cevedale between 2010 - 2011.

3.3.4 Hydro-climatological Model

The GMB estimates are derived from the hydro-climatological model AMUNDSEN [71]. This physically-based model, designed to simulate snow and ice accumulation and ablation in high mountains areas, has been widely used and evaluated in different climatological conditions [71], [84], [85], [86]. AMUNDSEN distinguishes four types of snow and ice, namely: fresh snow, old snow, ice and firn. New snow is considered as the snow coming from the recent snowfall. It is converted into old snow when settling and/or melting processes increase the density of the new snow above 200 kg/m^3. On September 30^{th} old snow is converted into firn by definition and firn turns into ice when the density reaches 900 kg/m^3 [87]. For each grid cell AMUNDSEN computes the energy and mass balance of these four snow and ice types, and hence, provides separate predictions of the water equivalent of fresh snow, old snow, firn and ice at each timestep. The sum of water equivalents of firn, old snow, fresh snow and ice represents the annual GMB, which is by definition the difference of the GMB calculated on the September 30^{th} of the considered year and the GMB on October 1^{st} of the previous year.

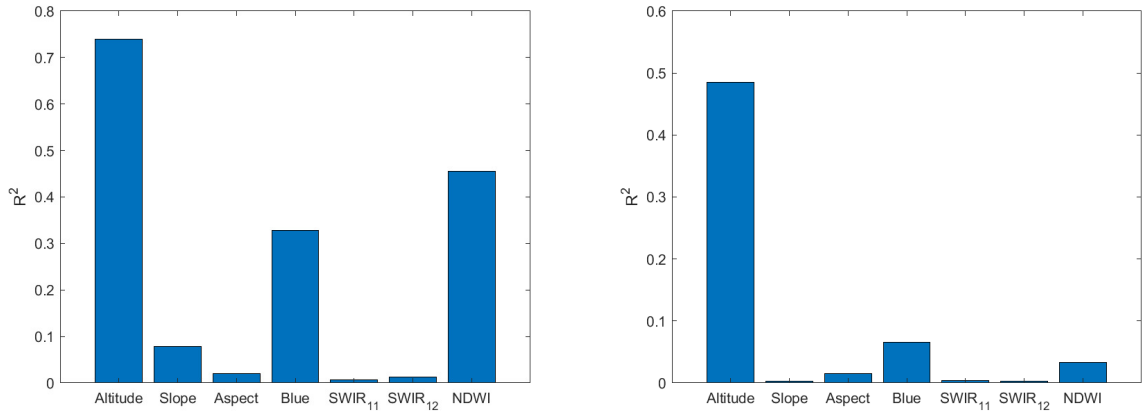
As a forcing data model, AMUNDSEN uses hourly meteorological variables observed at climate stations, such as: precipitation, wind speed, relative humidity, air temperature and global radiation. Other variables are parameterized (e.g., longwave fluxes). The location of the meteorological stations is presented in Fig. 3.3. The model and its particular features are described in [88, 89, 84, 73]. All meteorological forcing variables are distributed across the simulation domain by using a topography-considering inverse distance weighting scheme [84]. From the DTM of the considered area the following topographic features were computed: slope, aspect, openness and sky-view factor, all in the original resolution of 50 m, i.e., the one from DTM grid. In a post-processing step, the model results are re-sampled with linear interpolation to 10 m resolution. The model solves the coupled energy and mass balance of the snow-pack, and does not rely on parameter calibration for meaningful predictions. The initialization of the glacier areas is described in [73] which established the model in the Oetztaler Alps [72]. In this work, we extended the simulations to include the glaciers of the Ortler group, by utilizing climate station data provided by the Meteorological Office of Trentino, which include stations up to 2600 m a.s.l.. Initial glacier ice thickness distributions for the year 2003 were provided by Matthias Huss and are obtained by a method based on glacier mass turnover and ice flow mechanics. The method requires glacier outlines and a DTM [90].

3.4 Experimental Results

In this section, we present the GMB estimates obtained by applying the proposed integration method to both RS datasets (i.e., Sentinel-2 and Landsat data). First, we present the experimental setup by analyzing the correlation between the RS features and the GMB *in-situ* measurements. Then, the quantitative and qualitative results are reported. The quantitative results are presented in terms of: i) GMB estimates evaluated over the *in-situ* measurements, and ii) GMB of the whole glacier. Moreover, a sensitivity analysis is reported to evaluate the impact of the number of *in-situ* measurements on the proposed integration approach. Finally, to allow a visual interpretation of the results, we also present: i) scatter plots (between *in-situ* measured and estimated GMB aggregated per year), ii) altitudinal GMB profiles of *in-situ* measured and estimated GMB, and iii) GMB maps.

3.4.1 Experimental Set-up

To generate the feature space used to estimate the bias deviation, we extracted the blue and the SWIR spectral bands from the multispectral images. In greater detail, for the Sentinel-2 we considered band 2 (490 nm), band 11 (1610 nm) and band 12 (2190 nm),



(a) DTM and Sentinel-2 Features

(b) DTM and Landsat Features

Figure 3.4: The correlation between the RS features and *in-situ* measured GMB for the two datasets: (a) DTM and Sentinel-2, and (b) DTM and Landsat.

while for the Landsat Thematic Mapper we used band 1 (485 nm), band 5 (1650 nm) and band 7 (2220 nm). We also computed the NDWI as bellow:

$$NDWI = \frac{NIR - SWIR}{NIR + SWIR} \tag{3.4}$$

Where for the Sentinel-2, we used band 8 (842 nm) and band 11 (1610 nm), while for the Landsat we used band 4 (840 nm) and band 5 (1650 nm). Fig. 3.4 presents the correlation between the RS features and the *in-situ* measured GMB for both datasets. As expected, the altitude is strongly correlated to the GMB since it is strictly related to the snow distribution over glacier. Moreover, for Sentinel-2, the blue band and the NDWI present good correlation with the GMB values. Note that even though it is interesting to evaluate the correlation between each feature and the GMB, the whole feature space enables accurate modeling of the complex environmental properties of the glacier thanks to the complementary information provided by each feature.

To improve AMUNDSEN GMB estimates in the considered feature space, we exploited the k-NN correction strategy by using Euclidean distance. A parameter selection was first carried out to tune the k parameter which was set equal to 4 for all the experiments. Due to the small number of reference points, we adopted a five-folds cross validation procedure to evaluate the proposed approach. First, the reference samples were randomly divided into five subsets. At each iteration, one subset was used for evaluating the correction (validation subset), and the other subsets were used for computing the deviation correction. After five iterations, all the samples were used four times for the deviation correction

and once for the validation. The evaluation metrics were averaged on the five validation subsets. This condition allows us to ensure that the obtained results are reliable from the statistical view point. To evaluate the performance of the proposed integration method, the resulting GMB estimates were compared with the true GMB values (i.e., reference samples), considering the following metrics: the Mean Absolute Error (MSE), the Root Mean Square Error (RMSE), Root-mean-square deviation (RMSD) and the Correlation coefficient (R). The RMSD is calculated according to the formula:

$$RMSD = \frac{RMSE}{max - min} * 100\% \tag{3.5}$$

For comparison purposes, we also included the error metrics achieved by AMUNDSEN model with no correction technique applied. Moreover, to assess the effectiveness of the proposed integration method, we compared it with the standard supervised Support Vector Regression (SVR) technique. SVR is often used for geo/bio-physical variable estimation as this technique achieves good estimation accuracy and demonstrates strong generalization capabilities [91], [92]. We tested both linear and Radial Basis Function (RBF) kernels. The SVR parameters were tuned with a standard cross-validation. The SVR was trained on the GMB *in-situ* measurements by using as input the same variables as the one considered for the feature space definition.

3.4.2 Quantitative Results of the Physical Model Correction

To evaluate the proposed integration approach to GMB estimation, first we present the results obtained using HR Sentinel-2 data, AMUNDSEN simulations and *in-situ* measurements available over Hintereisferner and Kesselwandferner glaciers in the Rofental Valley (Austria), for the 2015, 2016 and 2017. Tab. 3.2 and 3.3 report the results obtained by the proposed integration approach, the original AMUNDSEN model, the linear SVR and the RBF SVR. The results were derived on the reference points collected per year separately. One can notice that the proposed integration method sharply reduces the estimation error with respect to AMUNDSEN model for all the cases. For instance, for the Hintereisferner glacier in 2016 the proposed method compared to the original AMUNDSEN achieves a MAE of ∼382 mm vs ∼940 mm, a RMSE of ∼560 mm vs ∼1207 mm a RMSD of ∼10% vs ∼33% and a R of 0.94 vs 0.91. Moreover, the proposed integration approach achieves better results compared to the baseline methods for almost all the cases. By focusing the attention on the Hintereisferner glacier in 2016, the proposed method compared to SVR RBF achieves a MAE ∼382 mm vs ∼430 mm, a RMSE of ∼560 mm vs ∼598 mm a RMSD of ∼10% vs ∼11% and a R of 0.94 vs 0.93. The error reduction is even more significant with respect to the linear SVR, where for Hintereisferner glacier in 2016 we obtained a MAE of ∼382 mm vs ∼556 mm, a RMSE of ∼560 mm vs ∼756 mm a RMSD of ∼10%

Table 3.2: Mean Absolute Error (MAE [mm]), Root Mean Square Error (RMSE [mm]), Root-mean-square deviation (RMSD) and correlation coefficient (R) obtained for GMB estimation of Hintereisferner glacier in 2015, 2016 and 2017. The error metrics are computed on reference samples using fivefold cross validation for AMUNDSEN, the proposed integration approach, SVR linear and SVR RBF.

Glacier Name		**HINTEREIS F.**			
Year	**Method**	**MAE** (mm)	**RMSE** (mm)	**RMSD (%)**	R
2015	**AMUNDSEN**	617.8	755.94	16	0.90
	Proposed	377.52	568.72	9	0.94
	SVR Linear	632.73	795.83	11	0.89
	SVR RBF	429.84	532.38	8	0.95
2016	**AMUNDSEN**	939.72	1206.64	33	0.91
	Proposed	382.25	559.90	10	0.94
	SVR Linear	555.83	755.89	16	0.90
	SVR RBF	429.7	598.21	11	0.93
2017	**AMUNDSEN**	790.87	908.57	17	0.89
	Proposed	548.37	747.93	13	0.89
	SVR Linear	614.64	789.31	11	0.89
	SVR RBF	622.09	925.7	16	0.83

Table 3.3: Mean Absolute Error (MAE [mm]), Root Mean Square Error (RMSE [mm]), Root-mean-square deviation (RMSD) and correlation coefficient (R) obtained for GMB estimation of Kesselwandferner glacier in 2015, 2016 and 2017. The error metrics are computed on reference samples using fivefold cross validation for AMUNDSEN, the proposed integration approach, SVR linear and SVR RBF.

Glacier Name		**KESSELWAND F.**			
Year	**Method**	**MAE** (mm)	**RMSE** (mm)	**RMSD (%)**	R
2015	**AMUNDSEN**	824.48	981.68	85	0.73
	Proposed	336.93	394.86	18	0.93
	SVR Linear	338.08	417.21	15	0.96
	SVR RBF	347.98	551.75	16	0.85
2016	**AMUNDSEN**	864.34	946.51	83	0.82
	Proposed	306.88	391.55	18	0.86
	SVR Linear	394.47	494.25	19	0.91
	SVR RBF	335.46	429.6	22	0.86
2017	**AMUNDSEN**	431.62	556.12	25	0.88
	Proposed	413.30	527.44	21	0.87
	SVR Linear	483.28	623.02	20	0.82
	SVR RBF	413.15	603.9	20	0.82

Table 3.4: Mean Absolute Error (MAE [mm]), Root Mean Square Error (RMSE [mm]), Root-mean-square deviation (RMSD) and correlation coefficient (R) obtained for GMB estimation of Vernagtferner glacier in 2012 and 2013. The error metrics are computed on reference samples using fivefold cross validation for AMUNDSEN, the proposed integration approach, SVR linear and SVR RBF.

Glacier Name		**VERNAGT F.**			
Year	**Method**	**MAE** (mm)	**RMSE** (mm)	**RMSD (%)**	R
2012	**AMUNDSEN**	467.48	567.97	18	0.91
	Proposed	420.57	514.01	15	0.90
	SVR Linear	551.66	678.47	14	0.84
	SVR RBF	461.98	591.57	31	0.86
2013	**AMUNDSEN**	577.99	755.83	38	0.97
	Proposed	244.94	321.14	10	0.96
	SVR Linear	889.78	1126.93	24	0.64
	SVR RBF	862.08	1112.70	31	0.64

vs ∼16% and R of 0.94 vs 0.90. However, when many *in-situ* samples are available, the RMSDs obtained by the baseline methods are slightly lower than the one obtained by the proposed method e.g. Kesselwandferner in 2015 RMSD of ∼18% for the proposed method vs ∼15% for SVR Linear vs ∼16% for SVR RBF.

In order to further analyze the capability of the proposed integration approach of improving the GMB estimation, we extended our analysis to historic data by considering the Landsat-7 optical images. According to the availability of AMUNDSEN simulations and the RS images representing maximum ablation, we analyzed Vernagtferner in the Rofental valley (Austria) in 2012 and 2013. The results are reported in Tab. 3.4. From the Tab., one can notice that the proposed integration approach significantly reduces the MAE, RMSE and RMSD values compared to the original AMUNDSEN, i.e., ∼245 mm vs ∼578 mm, ∼321 mm vs ∼ 756 mm and a RMSD of ∼10% vs ∼38%, respectively. However, in some cases AMUNDSEN shows slightly higher correlation coefficient than the proposed method e.g., on the Vernagtferner 2012 R of 0.91 vs 0.90. The proposed

Table 3.5: Mean Absolute Error (MAE [mm]), Root Mean Square Error (RMSE [mm]), Root-mean-square deviation (RMSD) and correlation coefficient (R) obtained for GMB estimation for Careser glacier in 2007, 2010 and 2011. The error metrics are computed on reference samples using fivefold cross validation for AMUNDSEN, the proposed integration approach, SVR linear and SVR RBF.

Glacier Name		**CARESER**			
Year	**Method**	**MAE** (mm)	**RMSE** (mm)	**RMSD (%)**	R
2007	**AMUNDSEN**	497.15	609.97	28	0.81
	Proposed	258.18	302.93	15	0.89
	SVR Linear	679.29	849.85	56	0.18
	SVR RBF	457.46	647.14	39	0.45
2010	**AMUNDSEN**	681.22	771.68	72	0.85
	Proposed	402.32	477.93	38	0.77
	SVR Linear	469.30	583.16	39	0.55
	SVR RBF	374.88	491.58	28	0.74
2011	**AMUNDSEN**	728.93	801.59	42	0.96
	Proposed	267.05	347.93	19	0.94
	SVR Linear	538.93	626.85	28	0.63
	SVR RBF	454.96	558.92	22	0.72

Table 3.6: Mean Absolute Error (MAE [mm]), Root Mean Square Error (RMSE [mm]), Root-mean-square deviation (RMSD) and correlation coefficient (R) obtained for GMB estimation for Weissbrunnferner glacier in 2006, 2007, 2010 and 2011. The error metrics are computed on reference samples using fivefold cross validation for AMUNDSEN, the proposed integration approach, SVR linear and SVR RBF.

Glacier Name		**WEISSBRUNN F.**			
Year	**Method**	**MAE** (mm)	**RMSE** (mm)	**RMSD (%)**	R
2006	**AMUNDSEN**	407.73	544.91	76	0.72
	Proposed	319.53	375.83	30	0.70
	SVR Linear	247.69	362.40	27	0.72
	SVR RBF	321.58	421.75	35	0.59
2007	**AMUNDSEN**	671.44	768.26	43	0.78
	Proposed	338.71	410.97	22	0.75
	SVR Linear	534.69	624.12	31	0.52
	SVR RBF	314.39	386.96	26	0.75
2010	**AMUNDSEN**	324.81	410.89	36	0.73
	Proposed	279.48	351.99	28	0.75
	SVR Linear	449.76	514.14	19	0.66
	SVR RBF	341.70	413.77	21	0.68
2011	**AMUNDSEN**	512.47	639.36	45	0.47
	Proposed	405.65	493.04	31	0.55
	SVR Linear	406.21	503.59	27	36
	SVR RBF	427.07	517.60	30	0.65

Table 3.7: Mean Absolute Error (MAE [mm]), Root Mean Square Error (RMSE [mm]), Root-mean-square deviation (RMSD) and correlation coefficient (R) obtained for GMB estimation for Langenferner glacier in 2010 and 2011. The error metrics are computed on reference samples using fivefold cross validation for AMUNDSEN, the proposed integration approach, SVR linear and SVR RBF.

Glacier Name		**LANGEN F.**			
Year	**Method**	**MAE** (mm)	**RMSE** (mm)	**RMSD (%)**	R
2010	**AMUNDSEN**	509.63	615.93	44	0.91
	Proposed	436.69	528.01	46	0.90
	SVR Linear	524.67	664.43	54	0.85
	SVR RBF	478.02	591.39	50	0.87
2011	**AMUNDSEN**	478.01	599.00	38	0.91
	Proposed	408.22	517.75	31	0.91
	SVR Linear	562.29	825.01	44	0.80
	SVR RBF	490.84	606.05	34	0.87

integration approach also obtained reduction of the MAE and RMSE errors and increase of the correlation coefficient compared to the SVR. For instance, by focusing on the Vernagtferner in 2013 the proposed method compared to the linear SVR achieves a MAE of ∼245 mm vs ∼890 mm, a RMSE - ∼321 mm vs ∼1127 mm, a RMSD of ∼10% vs ∼24% and a R of 0.96 vs 0.64.

Similar results were obtained in the Ortles-Cevedale group in Italy, where we considered three glaciers: Careser, Weissbrunnferner and Langenferner. Tab. 3.5, 3.6 and 3.7 report the results obtained for the 2006, 2007, 2010 and 2011. Similarly to the previous analysis, the proposed method strongly improves AMUNDSEN GMB estimates (see Tab. 3.5, 3.6 and 3.7). For instance, in Careser in 2007, the RMSE is ∼303 mm compared to AMUNDSEN one of ∼610 mm. Nevertheless, in some cases it shows slightly smaller linear correlation with *in-situ* measurements than the original AMUNDSEN model i.e., Weissbrunnferner 2006, R of 0.70 vs 0.72. Moreover, the proposed approach confirms its effectiveness compared to the SVR baselines, which in few cases provides results similar to the proposed method (e.g., Weissbrunnferner in 2006, where the RMSE of the proposed method is ∼376 mm and the one of the linear SVR is ∼362 mm). However, for most of the cases the baseline methods achieve poor accuracy due to the small amount of labeled samples available (e.g., Langenferner in 2011, where the RMSE of the proposed method is ∼518 mm and the one of the linear SVR is ∼825 mm).

To allow a visual interpretation of the results obtained, we present the scatter plots aggregated per year of the *in-situ* measured vs AMUNDSEN GMB (Fig. 3.5(a), 3.5(c), 3.5(e), 3.6(a), 3.6(c), 3.7(a), 3.7(c)), 3.7(e)) and scatter plots aggregated per year of the *in-situ* measured GMB vs the predicted one with the proposed method (Fig. 3.5(b), 3.5(d), 3.5(f), 3.6(b), 3.6(d), 3.7(b), 3.7(d), 3.7(f)). The red line corresponds to the perfect linear correlation, and the blue one represents the considered method correlation. These scatter-plots demonstrate the capability of the proposed approach to sharply improve the GMB estimates compared to the initial AMUNDSEN estimates, regardless of the number of samples and the RS dataset. Consequently, good results are obtained in 2006 (see Fig.3.5(b)) even though the *in-situ* measurements available are much less compared to those of 2016 (see Fig. 3.7(c)). However, one of the limitations of the proposed approach is that, for some samples, there might not be *in-situ* samples representing similar glacier condition, which will lead to poor correction accuracy. One example of this can be seen in Fig. 3.5(e) and 3.5(f), where the measured accumulation is ∼1000 mm and the proposed method estimates ∼100 mm (similarly to AMUNDSEN). This is due to the few accumulation measurement points available in 2010. The closest accessible samples are mainly ablation measurements, which leads to low correction values.

Fig. 3.8 reports the glacier-wide mass balance obtained over different glaciers for the

different years, considering the *in-situ* measurements (in black), AMUNDSEN (in green), the proposed integration approach (in blue), the SVR linear (in orange) and the SVR RBF (in yellow). These results confirm the importance of the correction performed by the proposed method at the glacier scale. The proposed approach obtains results very similar to those provided by the WGMS for all the glaciers by strongly improving AMUNDSEN estimates (see Kesselwandferner 2016 or Careser 2010). Moreover, it achieves better glacier-wide mass balance with respect to both the baselines methods.

Finally, to investigate the impact of the number of available *in-situ* measurements on the method accuracy, we performed a sensitivity analysis. As an example we considered, the Hintereisferner in 2016, where 54 samples were available. In our experiment, we subsampled the data into 5 sets made up of 10, 20, 30, 40, and 54 samples. Similarly to the other experiments, we adopted a five-folds cross validation procedure. The results are shown in Fig. 3.9. For instance, the proposed method with only 10 samples, compared to the original AMUNDSEN, achieves a MAE of ∼440 mm vs ∼890 mm, a RMSE of ∼495 mm vs ∼981 mm and a R2 of 0.93 vs 0.93. The result obtained by considering all 54 samples is nearly the same with MAE of ∼340 mm, a RMSE of ∼428 mm and R2 of 0.94. Thus, we can conclude, that the method performs well, regardless of the number of considered samples, as the MAE and RMSE are always improved.

3.4.3 Qualitative Results of the Physical Model Correction

Fig.3.10 presents the altitudinal GMB profiles per glacier per year obtained through *in-situ* measures (in black), AMUNDSEN model (in green) and the proposed approach (in blue). As expected for most of the profiles, the *in-situ* GMB increases when the altitude increases. A similar trend is visible in both AMUNDSEN and the proposed method GMB estimates. However, from the analysis of the obtained profiles, one can notice that the proposed method is able to match the altitudinal GMB provided by the *in-situ* measurements by sharply improving AMUNDSEN GMB values.

Finally, Fig. 3.11 presents the corrected GMB maps obtained by the proposed integration approach, the original AMUNDSEN GMB simulations and the true color compositions of the Sentinel-2 data acquired in August 2015, 2016 and 2017, respectively. The corrected maps (see Fig. 3.11 c, g, i) are characterized by a better geometrical detail than the original AMUNDSEN simulations (see Fig. 3.11 b, e, h), due to the higher spatial resolution of Sentinel-2 data (10 m) compared to AMUNDSEN simulation one (50 m). The high correlation between the Sentinel-2 images and the obtained GMB maps demonstrates the capability of the proposed method to accurately detect pixels characterized by similar environmental conditions when applying the deviation correction. This allows us to generate GMB maps which are more representative of the environmental properties of

the glacier than the original AMUNDSEN ones. We would like to remark that consistent results are provided also for Vernagtferner glacier, for which no reference samples are available for the 2015, 2016 and 2017.

3.5 Discussion and Conclusion

In this chapter, we presented a novel integration approach to GMB estimation. The proposed approach takes advantage of the complementary information provided by the: hydro-climatological AMUNDSEN simulations, multispectral RS data, DTM and *in-situ* measurements to generate accurate estimation results. The novelties of the proposed method are: (i) the definition and use of a feature space that can represent the environmental properties of the glaciers, (ii) the use of a deviation correction strategy to improve the accuracy of the GMB provided by the AMUNDSEN model and its spatial resolution, while preserving the reliability of the hydro-climatological model, and (iii) an efficient integration of the complementary information coming from physically-based simulations, RS data, DTM and *in-situ* measurements, which overcomes their individual limitations.

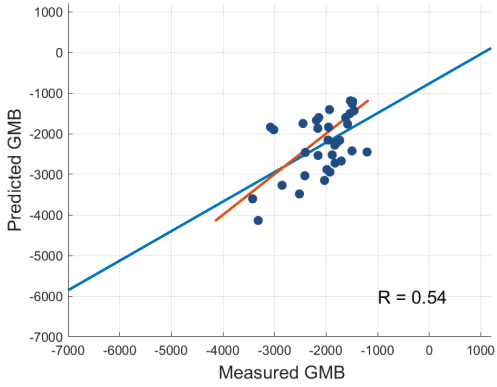
To assess the performance of the proposed integration method, an extensive analysis of the GMB correction was carried out in Tyrol (Austria) and South Tyrol (Italy), in 2006, 2007, 2010, 2011, 2012, 2013, 2015, 2016 and 2017. While for the analysis before 2015 we considered the Landsat-5 and Landsat-7 data, after 2015 we employed multispectral data from Sentinel-2. Compared to the AMUNDSEN model, the proposed approach demonstrates a significant reduction of the RMSE and MAE values for all the considered cases and an increase of the correlation coefficient factor in most of the cases. Overall, it can be noticed that the proposed approach performs better with RS data coming from Sentinel-2 despite the relevant spectral bands of both Sentinel-2 and Landsat being very similar (i.e., band blue 490 nm for Sentinel-2 vs 485 nm for Landsat). This can be explained by the better spatial resolution of Sentinel-2 (10 m vs 30 m for Landsat), which leads to better exploitation of the spectral information (i.e., pure spectral pixels). At a spatial resolution of 30 m the probability of having a mixed spectral signature is higher than at 10 m (e.g., both accumulation and ablation may be present in an area of 30x30 m^2). Moreover, since the proposed method was applied at a pixel level, the obtained GMB maps have a better spatial resolution compared to the original GMB simulations (50 m). This is due to the better geometrical information provided by the multispectral sensors.

The results obtained were compared with a standard supervised SVR regression technique, which is widely used for geo/bio-physical parameter estimation. In the analysis, we considered both linear SVR and RBF SVR. As expected, RBF SVR achieves better

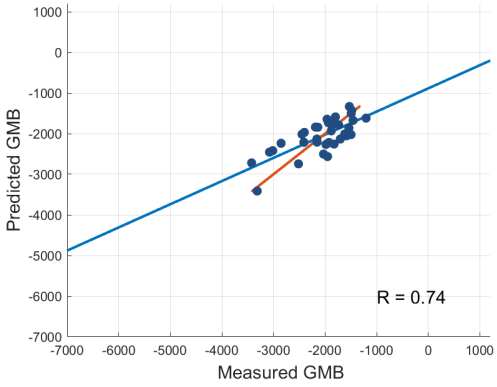
accuracy than the linear SVR. However, the proposed approach outperforms the baselines in most of the cases. This is because the SVR obtains accurate estimates when enough training samples are available. However, due to the complex application scenario, the collection of *in-situ* data is expensive and hampered by lack of access. Hence, the capability of the proposed method to exploit the complementary information provided by the different data sources is fundamental to obtain accurate results when few labeled samples are available. Moreover, since the method is data driven it can be easily applied to different test-sites by tuning only the value of the k parameter, depending on the available number of training samples, without the need of tuning any other parameters.

We would like to remark that the definition of a representative RS feature space is crucial to accurately model the various environmental conditions on glaciers. The feature space was defined by exploiting the spectral and the topographic information of the RS data. The effectiveness of the considered feature space is confirmed by the obtained GMB maps (both from the qualitative and quantitative view point).

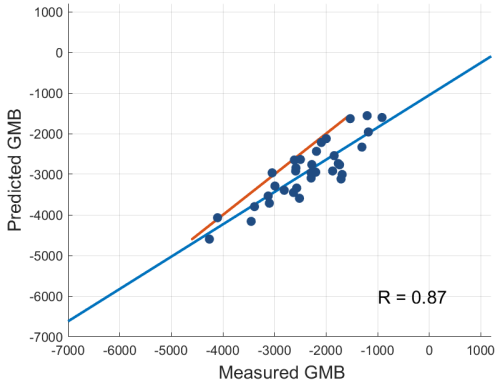
As future developments of this work, we would like to improve the feature space definition by integrating also SAR data (e.g., Sentinel-1). This would be particularly useful as SAR data are known for providing information about snow wetness, which is crucial for the GMB monitoring. Moreover, we would like to investigate the generalization capability of the proposed method. Thus, we plan to analyze the possibility of extending the correction to the other neighboring glaciers (in respect to the glaciers where the analyses was carried out), where no *in-situ* measurements are available. A final future development could consist of different validation scenarios. A possible option here is the validation of the glacier-wide mass balance by using the geodetic method or the continuous TS of ablation (from an ultrasonic depth gauge). Both methods could be informative and should be further analyzed.



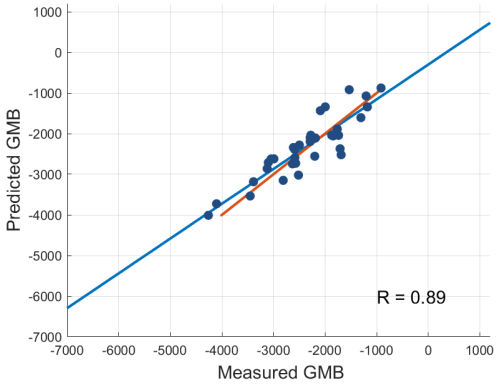
(a) AMUNDSEN 2006



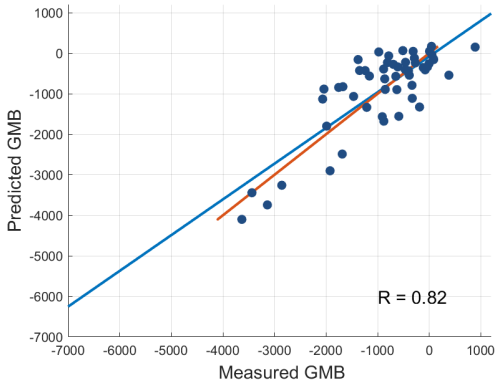
(b) Proposed Method 2006



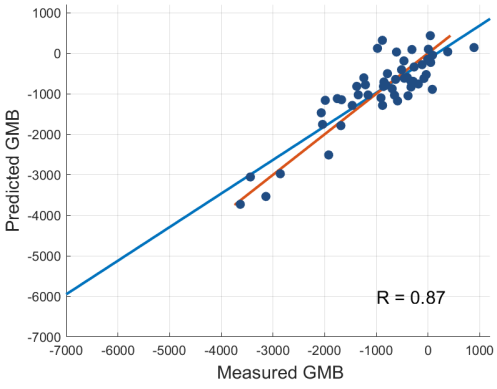
(c) AMUNDSEN 2007



(d) Proposed Method 2007

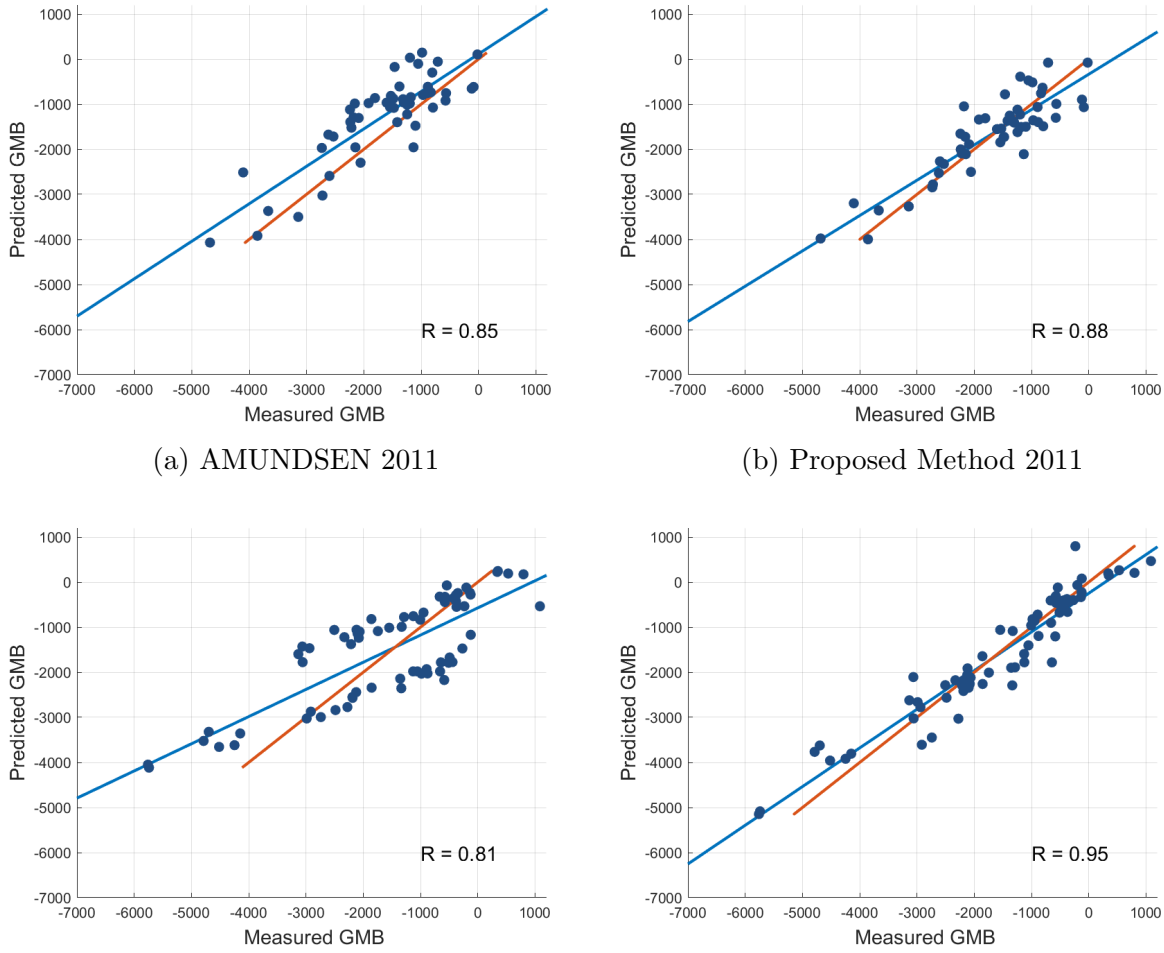


(e) AMUNDSEN 2010



(f) Proposed Method 2010

Figure 3.5: Scatter plots of correlation between GMB *in-situ* measured and AMUNDSEN modelled (a), (c), (e) and scatter plots of correlation between *in-situ* measured and GMB estimated by the proposed integration approach (b), (d), (f) in the Ortles-Cevedale group. Reference correlation line (i.e., maximum correlation) is highlighted in blue, while the model correlation line is reported in red.



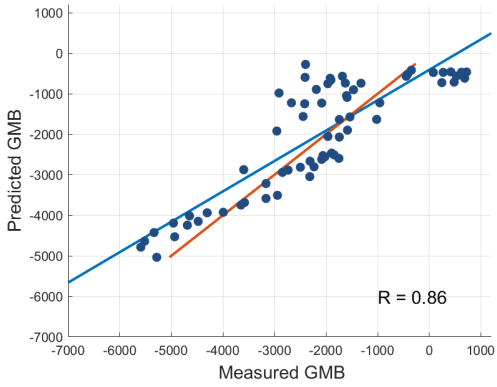
(a) AMUNDSEN 2011

(b) Proposed Method 2011

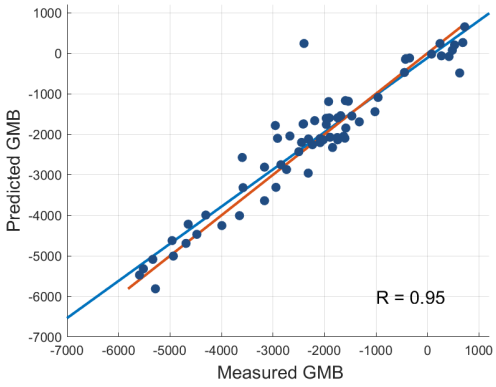
(c) AMUNDSEN 2013

(d) Proposed Method 2013

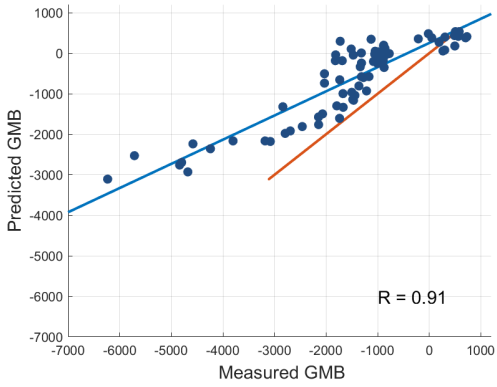
Figure 3.6: Scatter plots of correlation between GMB *in-situ* measured and AMUNDSEN modelled (a), (c) and scatter plots of correlation between *in-situ* measured and GMB estimated by the proposed integration approach (b), (d) in the Ortles-Cevedale group (a),(b) and Ötztal Alps (c) and (d). Reference correlation line (i.e., maximum correlation) is highlighted in blue, while the model correlation line is reported in red.



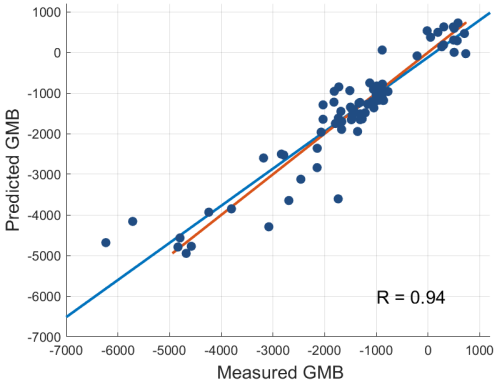
(a) AMUNDSEN 2015



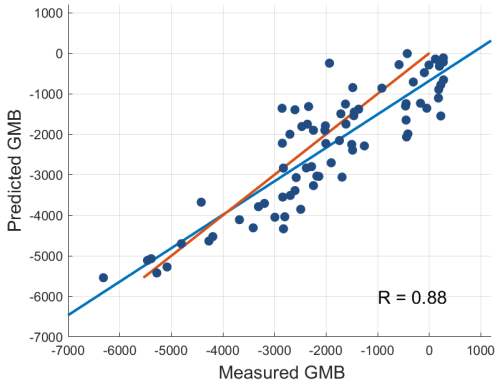
(b) Proposed Method 2015



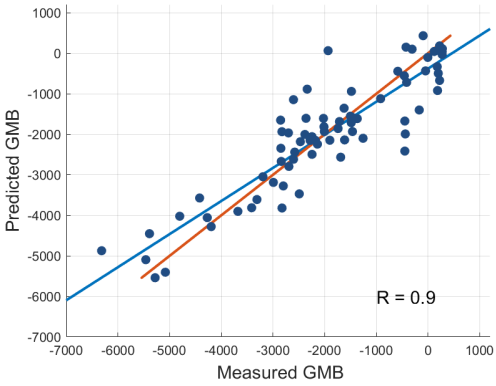
(c) AMUNDSEN 2016



(d) Proposed Method 2016

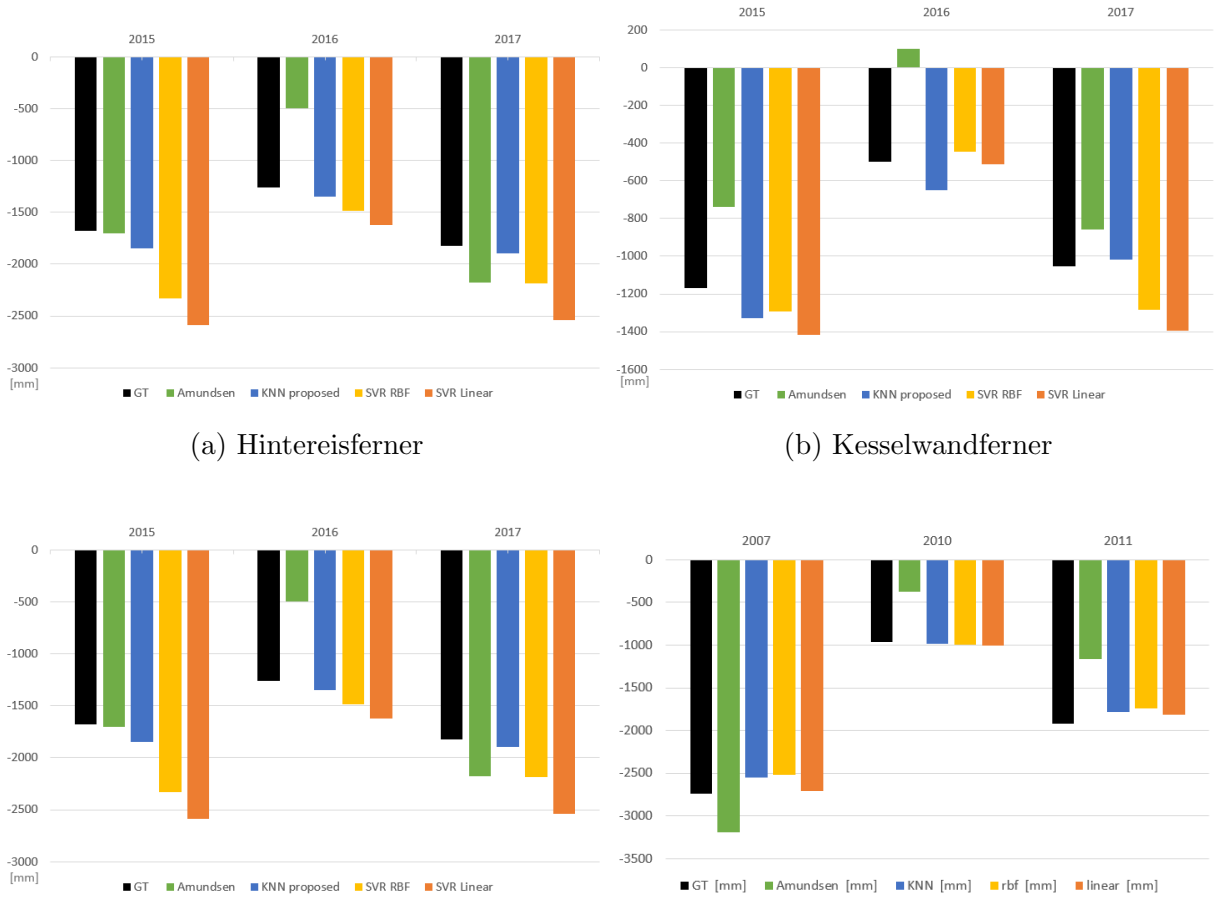


(e) AMUNDSEN 2017



(f) Proposed Method 2017

Figure 3.7: Scatter plots of correlation between GMB *in-situ* measured and AMUNDSEN modelled (a), (c), (e) and scatter plots of correlation between *in-situ* measured and GMB estimated by the proposed integration approach (b), (d), (f) in the Ötztal Alps. Reference correlation line (i.e., maximum correlation) is highlighted in blue, while the model correlation line is reported in red.



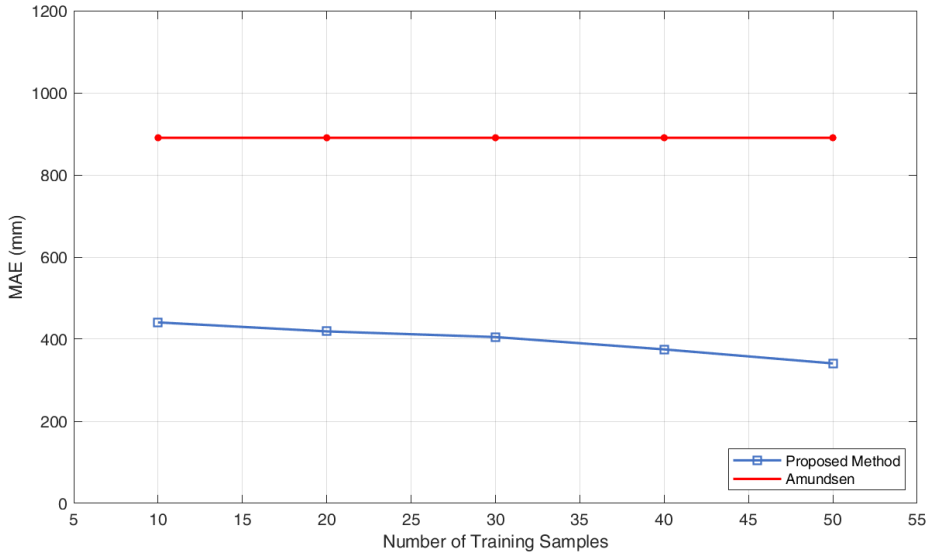
(a) Hintereisferner

(b) Kesselwandferner

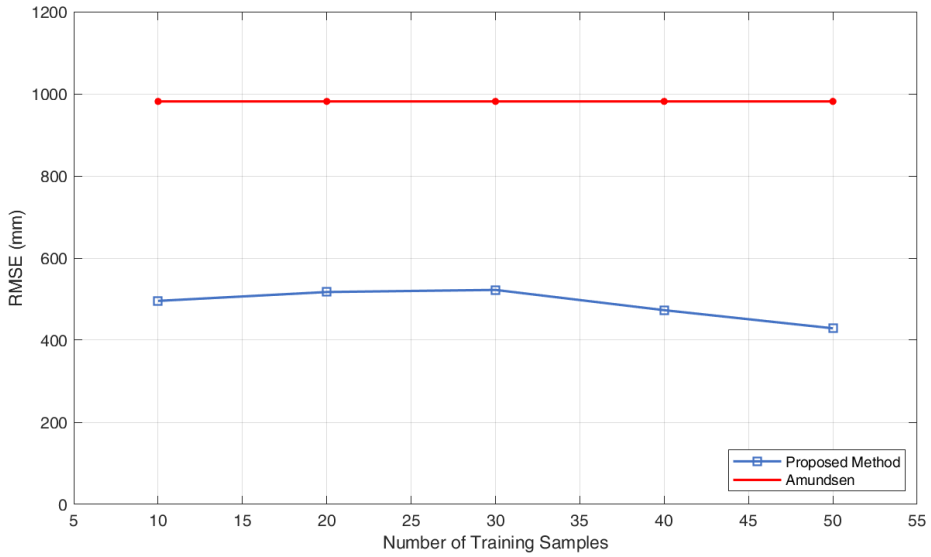
(c) Weissbrunnferner

(d) Careser

Figure 3.8: Glacier-wide mass balance for (a) Hintereisferner, (b) Kesselwandferner, (c) Weissbrunnferner and (d) Careser. The glacier-wide mass balance was estimated by: *in-situ* measurements (in black), AMUNDSEN (in green), the proposed integration approach (in blue), SVR RBF (in yellow) and SVR linear (in orange).

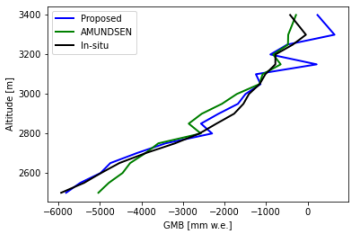


(a)

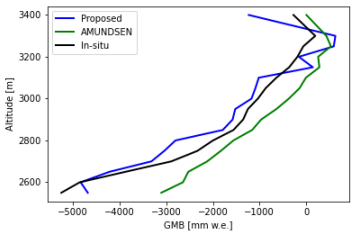


(b)

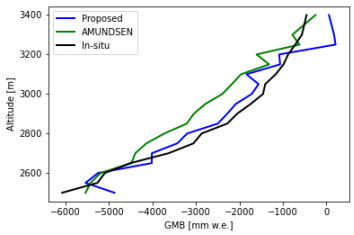
Figure 3.9: The sensitivity analysis of the Mean Absolute Error (a) and Root Mean Square Error (b) obtained for GMB estimation of Hintereisferner in 2016 for different number of considered samples (10, 20, 30, 40 and 54).



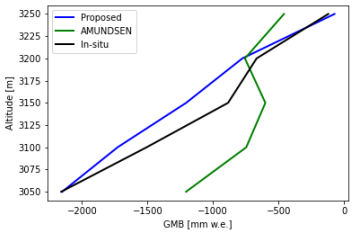
(a) Hintereis F. 2015



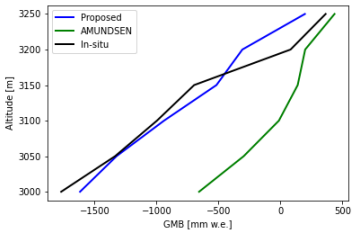
(b) Hintereis F. 2016



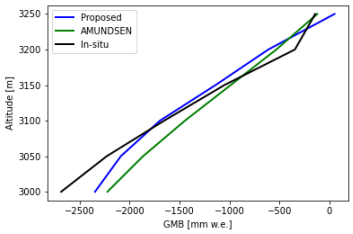
(c) Hintereis F. 2017



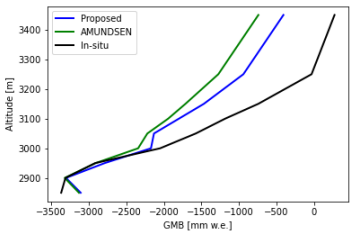
(d) Kesselwand F. 2015



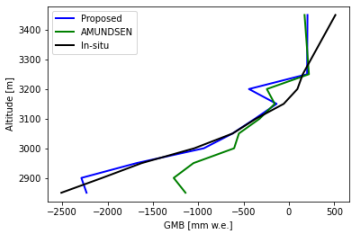
(e) Kesselwand F. 2016



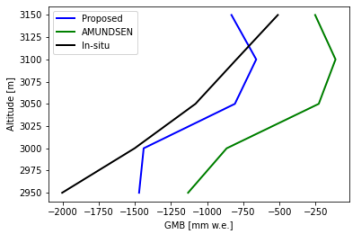
(f) Kesselwand F. 2017



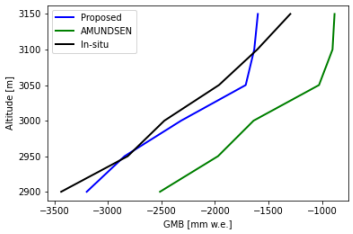
(g) Vernagt F. 2012



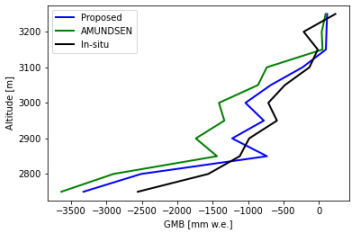
(h) Vernagt F. 2013



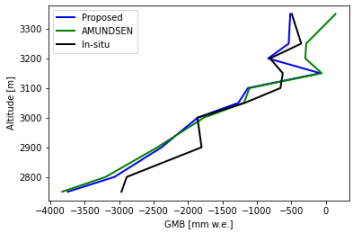
(i) Caresar 2010



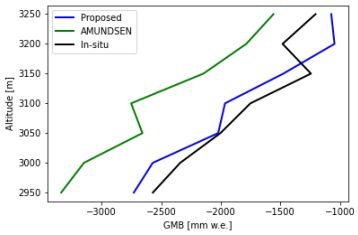
(j) Caresar 2011



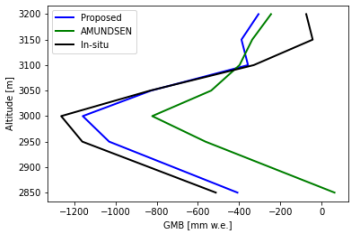
(k) Langen F. 2010



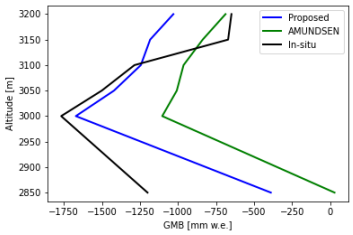
(l) Langen F. 2011



(m) Weissbrunn F. 2007



(n) Weissbrunn F. 2010



(o) Weissbrunn F. 2011

Figure 3.10: The altitudinal GMB profiles of *in-situ* measured (in black), modelled by AMUNDSEN (in green) and corrected by the proposed integration approach (in blue).

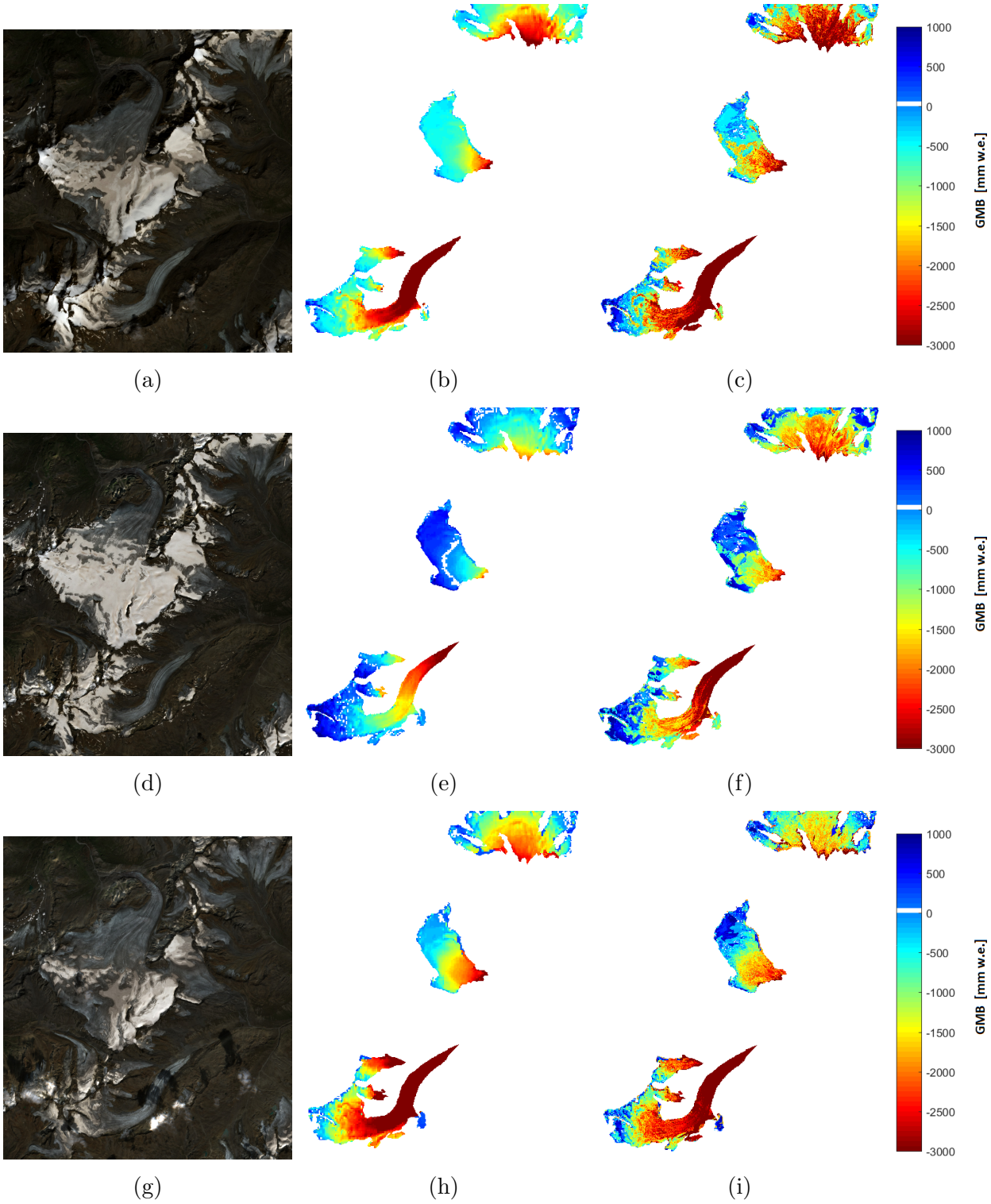


Figure 3.11: True color compositions of the Sentinel-2 bands acquired respectively on (a) 2015-08-26, (d) 2016-08-27 and (g) 2017-08-30, and related AMUNDSEN GMB simulations (b)(e)(h), and corrected AMUNDSEN simulations using the proposed integration approach (c)(f)(i).

Chapter 4

Multitemporal Land Cover Mapping at High Spatial Resolution for Sub-Continental Monitoring of Climate Change

This chapter presents a processing chain for the production of multitemporal high resolution land cover (HRLC) maps on a sub-continental scale for climate change monitoring. The considered processing chain generates LC products every five years from 1990 to 2019 leveraging multitemporal multispectral optical images acquired by different sensors. In particular, a recent static map (2019) produced at 10 m spatial resolution is produced considering TS of Sentinel-2 images, while the historic maps (from 1990 to 2015) are produced at 30 m spatial resolution using Landsat data. Due to the need of exploiting these maps for refined climate change analysis, a very detailed classification scheme is considered. The processing chain is based on three mains phases: (i) the optical images preprocessing, (ii) the feature extraction, and (iii) the supervised classification of the multitemporal RS data. The preprocessing phase aims to handle variability in data availability as well as large amount of optical images acquired by different sensors at different spatial resolutions. To this end, several processing steps are computed, namely, the spectral filtering, cloud/shadow detection, cloud/shadow restoration, topographic shadow restoration and monthly/seasonal or annual composite depending on the considered environmental conditions. To enhance the spatial information, the textural features are extracted from the multispectral images. Finally, the production of the multitemporal HRLC maps is performed in the supervised classification phase.

4.1 Introduction

To perform accurate LC monitoring, it is necessary to have LC maps produced at high spatial and temporal resolution. However, as presented in Chapter 2.4, a lot of effort has been devoted to produce medium resolution LC maps, while the production of HR large scale LC maps is still very limited. This is due to the fact that medium spatial resolution RS data are typically characterized by high temporal resolution (e.g., SPOT-Vegetation have daily based acquisitions), which allows the accurate characterization of the temporal profiles of different LC classes. For the HRLC mapping, the RS data available are characterized by a quite limited number of acquisitions, especially going back in the past (e.g., Landsat 5 acquires images every 16 days). Therefore, to compensate a low temporal resolution we need to deal with images acquired by different sensors at different spatial and temporal resolution, which requires a peculiar preprocessing phase to properly harmonize the data. Moreover, the few HRLC products available at global scale are focused on the most important LC categories, and cannot be used to perform detailed environmental monitoring such as the one required for climate change analysis. Furthermore, the climate change monitoring task requires the availability of LC maps frequently updated. This is because from the climatological prospective the consistency between maps is crucial to understand LC changes, and thus ongoing processes, such as, deforestation, urbanization and arctic greening.

In this context, we proposed two automatic processing chains. First one tailored to the characteristics of Sentinel-2 imagery, which provides a large static HRLC map at subcontinental level (at 10 meters resolution). The second one tailored to the properties of Landsat 5, 7 and 8 imagery, generates a long-term record of regional HRLC maps (at 30 meters) has been developed. This work has been defined in the framework of the CCI High Resolution Land Cover project, which is the continuation of CCI medium resolution project (with LC maps available at 300 meters spatial resolution) [7] and aims to better understand the role of the spatial resolution of the detected LC and LC changes for the climate modeling research. In particular, the produced HR maps aim to support both: the sub-continental climate modeling simulations and regional climate mitigation and adaptation studies. We consider three areas located in Amazon, Africa and Siberia. This areas are known for the exposure to the extreme climate conditions and climate change. The HRLC maps are generated by fusing the classification results obtained by the optical-based and SAR-based pipelines. The high-level work-flow of the whole processing chain is presented in Fig. 4.1, where the red and dark blue colors correspond to the modules of the optical and SAR processing chain, respectively. The light blue color represents the modules designed for the optical and SAR data integration and geolocation. The optical

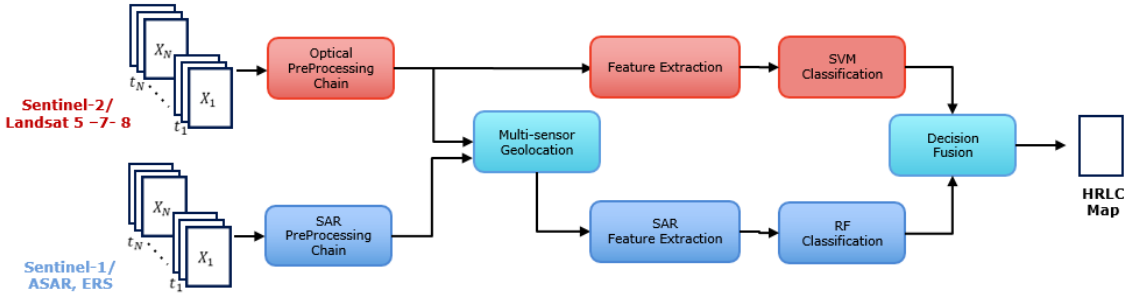


Figure 4.1: Block-based representation of the processing chain for the production of HRLC maps.

processing chain exploits Sentinel-2 data while the SAR chain works with Sentinel-1 data from 2019. To generate the 5 years LC maps in the past (from 1990 to 2015), ERS and ASAR data are considered from the SAR processing chain and Landsat 5, Landsat 7 and Landsat 8 from the optical one. The combination of these two data sources allows for accurate LC mapping due to the complementary information provided by the SAR and optical sensors. This integration is done in the final data fusion module. Therefore, we are able to develop advanced and ad-hoc processing approaches for multisensor data, while keeping the system scalable and modular.

In this thesis, we focused the attention on the optical-based pipeline made up of three main parts: (i) images preprocessing, (ii) feature extraction, and (iii) the supervised classification of the multitemporal RS data. The main objective of the optical preprocessing part is the harmonization of the multisource multitemporal optical RS data to ensure data consistency. In particular, the optical data preprocessing, includes the following steps: spectral filtering, cloud/shadow detection, cloud/shadow restoration, composite generation (the composite can be monthly/seasonal or annual based depending on the considered environmental conditions and data availability) and topographic shadow restoration. Next, in order to introduce spatial information to the classifier (and thus to improve classes discrimination) we computed textural features. Finally, the processing chains perform the supervised automatic classification which is crucial to produce accurate LC maps.

The remainder of this chapter is structured as follows. Section 2 presents the study areas for which the proposed processing chain was developed. Section 3 describes in detail the proposed optical multitemporal HRLC processing chain, while Section 4 presents the obtained results. Finally, in Section 5 we draw the conclusion and propose further developments.

4.2 Study Areas and Data Set Description

Fig. 4.2 shows the location of the 3 regions selected for the study the impact of LC and LC change on the Earth climate. In accordance with GCOS requirements (see Chapter 1.2) the static area was selected to support climate modeling simulations, while historic area is dedicated to more detailed long-term study of climate mitigation and adaptation. The study areas were selected in locations, where the interactions between the LC and climate are known to climatologists. Moreover, due to the scarce availability of RS data in the past, the location of historic areas maximizes the Landsat data availability between 1990 and 2015. The green contours delimit the area for which the static map is produced, while the yellow contours correspond to the historic products. For each study areas, we delimited in gray the experimental site for which the qualitative results are reported in this thesis. Please note that the quantitative evaluation are reported within the whole historic area (in each of the considered study areas) for both static and historic processing chain. The historic areas includes 283 Sentinel-2 tiles in the African Sahel, 265 in the Amazon, and 204 in the Siberia. Finally the proposed static processing chain will be also applied to the whole static area. However, at the time of writing this thesis the maps are not yet available.

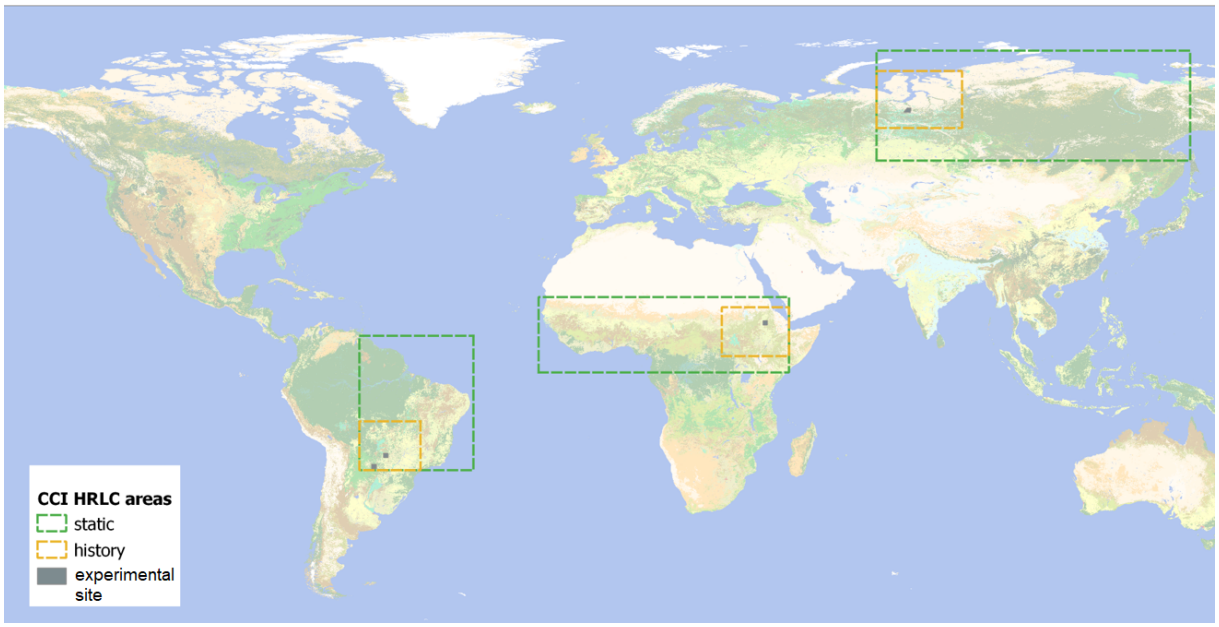


Figure 4.2: Location of the 3 regions to study climate/LC interactions (the green contours delimit the large regions where the static HRLC map will be produced, the yellow ones, the restricted areas where LC will be mapped over the last 30 years).

4.2.1 Study Areas: Climate Regions

The considered climate regions are located in three different continents and represent different climate conditions (tropical, semi-arid and boreal). The selected areas corresponds to the ones, which were already identified as a hot-spots of climate change by the CCI medium resolution project. Indeed, due to the peculiar location, these areas present complex surface atmosphere interactions, which affect not only regional climate but also global climate structures. Moreover, these regions covers major biomes that are particularly vulnerable to climate changes. At the same time, they are of paramount importance for the global carbon cycle, as they remove human-made carbon emissions from the atmosphere (tropical forest and permafrost).

4.2.1.1 Amazon

For several decades the Amazon basin has been deeply analyzed by climatologists. This is because of the large deforestation that leads to losses of carbon storage and strongly affects climate processes at the large-scale. The selected region in Amazon includes the hot spots of deforestation, cropping, grazing and urbanization. In this context, the HRLC maps are of particular importance for climate models to better identify the forest landscape structure (e.g., spatial organization and gap size distribution) and landscape interplay (e.g, croplands present in the flooded areas). The geographical coordinates of the selected regions are as follows: static map - (24°S -9°N; 34°W - 62°W) and historic LC map - (24°S - 12°S; 47°W - 62°W).

4.2.1.2 African Sahel

In Africa, we focused on the Sahel region including the West Africa. Within this region, we are able to catch the whole Senegal and Awash river basins as well as the extensive swamp area in the South Soudan, which is particularly sensitive to floods and droughts. From the climatic point of view, it is a complex region due to the ongoing processes, such as: land degradation, desertification, cropland extension, reforestation and urbanization increase. These processes are often attributed to climate warming. Previous simulations performed within this region were attributed with systematic errors related to the poor description of surface properties. Therefore, it is expected that the higher resolution of LC maps will help to better characterize surface heterogeneity and its properties (i.e., soil thermal inertia, surface albedo). Moreover, the higher accuracy of LC map will help to better identify classes that were classified as mixed in the middle resolution LC map. The geographical coordinates of selected regions are as follows: static map - (0°N - 18.5°N; 18°W - 43.5°E) and historic LC map -(4°N - 16°N; 27°E - 43.5°E).

4.2.1.3 Siberia

The last region is located in the Siberia and includes the whole catchment of the Lena river. For this region, permafrost degradation is already visible and future climate changes (such as polar amplification) are expected to be particularly evident. The selected region covers a large diversity of frozen soil types, ranging to continuous, discontinuous and sporadic permafrosts. The Siberia represents 10% of land surface and 30% of forested surfaces globally. The high quality LC maps can help to better understand the greening of the arctic region and the associated displacement of the forest-shrubs-grasslands transition zone and snow. The geographical coordinates of selected regions are as follows: static map - (52°N - 79°N; 65°E - 142°E) and historic LC map - (60°N - 74°N; 65°E - 86°E).

4.2.2 LC Classes: Reference Data Collection

The legend of the HRLC was defined in accordance with The Food and Agriculture Organization and LCCS [93]. Tab.4.1 reports the considered detailed classification scheme made up of the following LC types: "Tree Cover Evergreen Broadleaf," "Tree Cover Evergreen Needleaf," "Tree Cover Deciduous Broadleaf," "Tree Cover Deciduous Needleaf," "Shrub Cover Evergreen," "Shrub Cover Deciduous," "Grasslands," "Croplands," "Woody Vegetation Aquatic or regularly flooded," "Grassland Vegetation Aquatic or regularly flooded," "Lichens and Mosses," "Bare areas," "Built-up," "Open water seasonal," "Open water permanent," and "Permanent snows and/or ice". While in Africa and Amazon this peculiar classification scheme has been considered, in Siberia for most of the year the region is covered by snow and seasonal LC classes can not be observed. Therefore, for Siberia we excluded the seasonal LC classes for both the static and historic processing chains. In particular, the shrub evergreen and deciduous have been merged into unique shrub class, grassland vegetation aquatic or regularly flooded and woody vegetation aquatic or regularly flooded have been merged into unique vegetation aquatic while seasonal and permanent water have been merged into water class.

The collection of the labeled samples has been carried out via photo-intepretation. Indeed, even though publicly available thematic products (such as Copernicus Global Land Cover [58] and Permafrost [94]) represent a valid source of information, given the difference in temporal resolution, legend definition and spatial resolution between existing thematic products and the proposed HRLC map, they could not be used to extract labeled samples as done in CCI medium resolution LC mapping [7]. However, the 2015 CGLC map was used to study the peculiar landscape of the three study areas and to estimate the prior probabilities of the different LC classes present in the scene. In particular, a stratified random sampling strategy has been considered [95]. The label of each sample

was defined by the photo-interpretation of both 2019 Sentinel-2 data and HR 2019 SPOT images for the static maps as well as the multitemporal analysis of trend of NDVI and NDWI indices. For those areas where SPOT images were not available, we exploited the public HR Google Satellite and ESRI images (i.e., 50 cm). The data were pixel-wise labeled. Thus, we avoided the strong positive correlation between samples units, which is the case for polygon-wise labeling. Due to the lack of the HR images for all the considered years in the past, the training set extracted in 2019 was back-propagated up to 1990. In particular, to generate a temporally consistent training set for the long-term record maps, for each labeled sample collected in 2019 the photo-interpreter was asked to confirm, remove or change its label. This is because, as the images available in the past are of lower spatial and temporal resolution and the LC changes in time, it maybe impossible to accurately understand the correct sample label for each considered year. Moreover, it may happen that due to heavy cloud cover and corruption of Landsat-7 data, no information is available. As a consequence, the training set produced in the past have smaller number of samples compared to the ones used to classify the static map. Fig. 4.3 shows an example view of the photo-interpretation task, where highlighted dots represents the samples to be labeled. The presented images comes from the HR Google Satellite and corresponds to the Sentinel-2 tile 21KUQ in the Amazon.

Table 4.1: The legend of the HRLC classification map.

Code	Class Name	Definition
10	Tree Cover Evergreen Broadleaf	A tree is a woody, perennial plant with a simple and well-defined stem, more or less defined crown and a minimum height of 5 m. Tree canopy cover (>50 %) composed of trees that are broadleaf and never entirely without green foliage.
20	Tree Cover Evergreen Needleaf	Tree canopy cover (>50 %) composed of trees that are needleaf and never entirely without green foliage.
30	Tree Cover Deciduous Broadleaf	Tree canopy cover (>50 %) composed of trees that are broadleaf and leafless for a certain period during the year.
40	Tree Cover Deciduous Needleaf	Tree canopy cover (>50 %) composed of trees that are needleaf and leafless for a certain period during the year.

50	Shrub Cover Evergreen	A shrub is a woody perennial plant with persistent woody stems and without any defined main stem, being less than 5 m tall. Shrub canopy cover (>50 %) composed of shrubs that are never entirely without green foliage.
60	Shrub Cover Deciduous	Shrub canopy cover (>50 %) composed of shrubs that are leafless for a certain period during the year.
70	Grasslands	Primarily vegetated areas with an herbaceous cover of more than 50% at the time of fullest development. Herbaceous plants are defined as plants without persistent stem or shoots above ground and lacking definite firm structure.
80	Croplands	Primarily vegetated areas with a herbaceous cover of more than 50 % at the time of fullest development. Croplands are mainly herbaceous plants are sowed/planted and harvestable at least once within the 12 months after the sowing/planting date. Cropland includes rain fed crops, irrigated crops, aquatic crops and annual pastures. Croplands exclude permanent crops like woody plantations that are part of the tree or shrub classes.
90	Woody Vegetation Aquatic or regularly flooded	Primarily vegetated areas with trees, shrubs covering more than 50 % of the area flooded by water for more than 4 months throughout the year. The water can be saline, fresh or brackish.
100	Grassland Vegetation Aquatic or regularly flooded	Primarily vegetated areas with grasslands or lichens and mosses covering more than 50 % of the area flooded by water for more than 4 months throughout the year. The water can be saline, fresh or brackish.
110	Lichens and Mosses	Primarily vegetated areas with a cover of more than 50% at the time of fullest development. Mosses are a group of photo-autotrophic land plants without true leaves, stems or roots. Lichens are composite organisms formed from the symbiotic association of fungi and algae.

120	Bare area	Areas where the sum of vegetation cover is less than 50% at the time of fullest development. Bare rock areas, sands and deserts are classified as bare areas. Extraction sites (open mines and quarries) and salt flats covered by water for less than 5 months are classified as bare areas.
130	Built-up	Areas where any predominant type of linear and non-linear artificial surface covers at least 50%. Built-up areas include buildings, roads, airports, greenhouses, etc. but exclude temporary settlements.
140	Open water seasonal	Areas where open water covers at least 50% of the surface and remains between 5 and 9 months a year, except in special circumstances (particularly dry year, construction of dams, etc.). Water bodies can be natural or artificial. Water can be saline, fresh or brackish.
150	Open water permanent	Areas where open water covers at least 50% of the surface and remains for more than 9 months a year, except in special circumstances (particularly dry year, construction of dams, etc.). Water bodies can be natural or artificial. Water can be saline, fresh or brackish.
160	Permanent snows and/or ice	Areas where snow and/or ice cover at least 50% of the surface for more than 9 months a year. Built-up areas and open water cover less than 50% of the surface.

4.2.3 RS Data

4.2.3.1 Sentinel-2 Images

The input data of the processing chain that generates the 2019 static map at 10 m spatial resolution are the atmospherically corrected Sentinel-2 data (i.e., L2A products). In particular, we considered the following Sentinel-2 spectral bands: blue (B2 - 490 nm), green (B3 - 560 nm), red (B4 - 665 nm), the four vegetation red edge (B5 - 705 nm, B6 - 740 nm, B7 - 0.783 nm and B8A - 865 nm), NIR (B8 - 842 nm) and the two SWIR

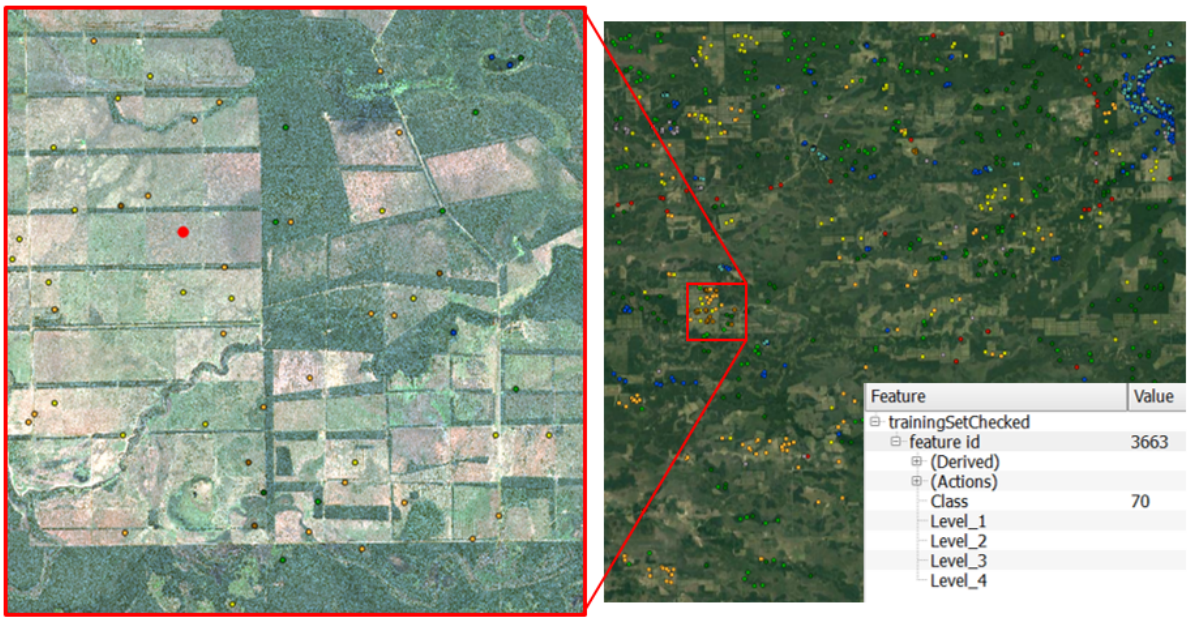


Figure 4.3: Example view of the photo-interpretation task in the Amazon over the Sentinel-2, 21KUQ tile. The presented HR images come from the Google Satellite.

bands (B11 - 1610 nm and B12 - 2190 nm). A nearest neighbor interpolation technique was used to match the 20 m bands to the 10 m bands. All the images considered in the experiments have cloud coverage smaller than 40%. The number of images per tile varies between 13-84 in the Amazon, 5-100 in the African Sahel and 60-103 in the Siberia.

4.2.3.2 Landsat Images

The input data of the processing chain that generates long-term record of regional HRLC maps in 1990, 1995, 2000, 2005, 2010 and 2015 are the atmospherically corrected Landsat 5, 7 and 8 images (i.e., L2 products). Tab. 4.2 specifies the Landsat sensor used for the different years and areas. The following spectral bands are considered for Landsat 8: band blue (B2 - 480 nm), band green (B3 - 560 nm), band red (B4 - 650 nm), band NIR (B5 - 865 nm), band SWIR 1 (B6 - 1610 nm) and band SWIR 2 (B7 - 2200 nm), while for Landsat 5/7 we used: band blue (B1 - 485 nm), band green (B2 - 560 nm), band red (B3 - 660 nm), band NIR (B4 - 835 nm), band SWIR 1 (B6 - 1145 nm) and band SWIR 2 (B7 - 2200 nm). Also in this case the images considered in the experiments have cloud coverage smaller than 40%. The number of images per tile varies between 17-72 in the Amazon, 18-63 in the African Sahel and 7-14 in the Siberia.

Table 4.2: The specification of the imagery considered for the historic maps generation.

Area	1990	1995	2000	2005	2010	2015
Siberia	Landsat 5	Landsat 5	Landsat 5 -7	Landsat 5 -7	Landsat 5 -7	Landsat 7 -8
Amazon	Landsat 5	Landsat 5	Landsat 5 -7	Landsat 5 -7	Landsat 5 -7	Landsat 7 -8
African Sahel	Landsat 5	Landsat 5	Landsat 5 -7	Landsat 5 -7	Landsat 5 -7	Landsat 7 -8

4.2.3.3 DTM Topographic Maps

The DTM has been used to represent topographic characteristics of the considered areas, such as altitude, slope and aspect. As DTM we have considered the Shuttle Radar Topography Mission (SRTM) digital elevation data provided by NASA JPL. The SRTM is the result of international research effort to produce the near-global set of elevation models. The SRTM is available at 30 meters spatial resolution, and it was produced by using interferometric synthetic-aperture radar technique [96].

4.3 Multitemporal HRLC Mapping

To answer the requirements regarding LC monitoring defined by GCOS, we defined two processing chains static and historical one. At the best of the author's knowledge, this work is the first that address both: LC monitoring at high spatial resolution (10 meters) on a subcontinental scale and to provide a long-term record of regional HRLC maps at 30 meters every 5 years from 1990 in the context of climate studies. Both processing chains includes a set of standard modules for image preprocessing and classification presented in Fig. 4.4. The input to the static processing chain includes 10 meters spatial resolution Sentinel-2 data, while to the historic processing chain 30 meters spatial resolution Landsat 5, 7, 8 data. The specific differences for Sentinel-2 and Landsat image processing are reported in the following.

4.3.1 Preprocessing Phase

The purpose of the preprocessing is to provide TS of spatially, temporally and spectrally harmonized RS data. This is necessary to achieve accurate classification results, in particular when dealing with the detailed considered classification scheme, which includes many seasonal classes. The details about each preprocessing steps are given in the following.

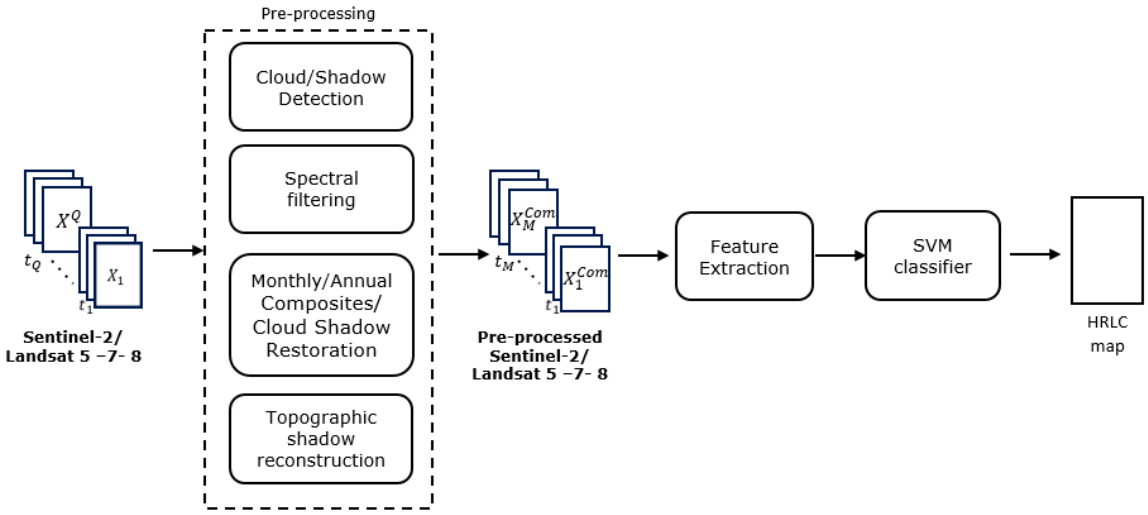


Figure 4.4: Block-based representation of the optical processing chain for the production of HRLC maps.

4.3.1.1 Cloud and Cloud Shadow Detection

The detection of both clouds and the related shadows is one of the most important challenges when processing optical RS data. This is because there are many clouds types and they can be characterize with different transparency [97]. Moreover, clouds can be easily confused with bright landscapes, while clouds shadows are often confused with topographic shadows, water and burnt areas. Since cloud and cloud shadow detection algorithms are computationally expensive, we decided to rely on the cloud and cloud shadow information provided with the L2A product (for Sentinel-2) and L2 product (for Landsat). The ESA distribution of the Sentinel-2 cloud mask is based on Sen2Cor tool while the USGS distribution of Landsat on Fmask (Function of mask). The details about Sen2Cor tool and Fmask are provided in Section 5.1. However, the OA of cloud and shadow mask provided by the Sen2Cor (84%) is on average lower than the one provided by Fmask (90%) [97]. Therefore, the Sen2cor masks should be further enhanced to achieve the required accuracy.

Fig. 4.5 presents the overview of the block-based representation of the cloud detection method used to improve the Sentinel-2 mask. The main goal of the proposed method is to identify the clouds present in each optical image (of the considered TS) by comparing it with a cloudless background image. This is a standard approach widely used in the literature [98]. In the considered implementation of the method, we work at seasonal level to reduce the computational burden. For each season, the cloudless background is computed as the quarter quantile of the blue band. Let $\{X^1, X^2, \ldots, X^S\}$ be the

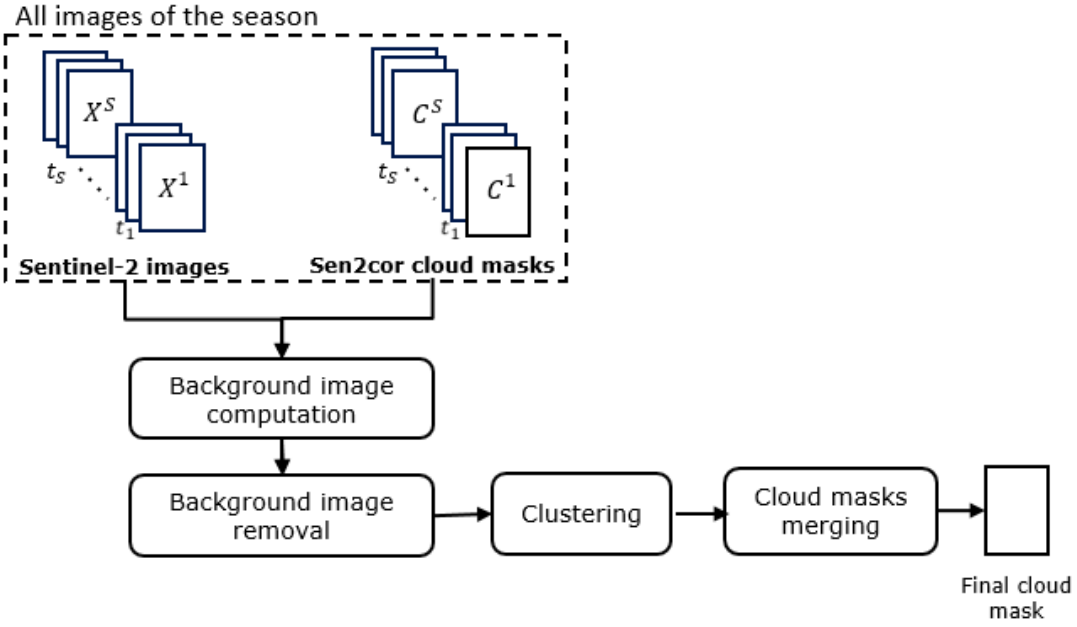


Figure 4.5: Block-based representation of the implemented cloud detection for Sentinel-2.

considered TS, which includes optical images acquired over a season. The multitemporal pattern of band blue associated to the jth pixel of $\{X^1, X^2, \ldots, X^S\}$ can be defined as $[x^1_{j,Blue}, x^2_{j,Blue}, \ldots, x^S_{j,Blue}]$. The jth pixel of the background image x^B_j is generated by computing quarter quantile of the blue band of the cloud-free multispectral pixels of the images present in $\{X^1, X^2, \ldots, X^S\}$ as follows:

$$x^B_j = quantile_{0.25}\{x^1_{j,B_{Blue}}, x^2_{j,B_{Blue}}, \ldots, x^S_{j,B_{Blue}}\} \tag{4.1}$$

After the background image is computed, we use it as a cloudless reference image and the problem of cloud detection is treated as a change detection problem. Thus, we calculate the difference between the blue bands of each image from the TS and the background image (a standard procedure in the change detection domain). However, instead of performing thresholding on the diffrence image, we cluster it in order to find homogeneous areas of surface and clouds. To understand which from the obtained clusters belong to cloud cover, the mean of each cluster is compared with the blue band mean of the cloudy pixels overall image. Finally, we perform logical disjunction of the obtained cloud mask with the original Sen2cor mask.

To enhance the cloud shadow mask, we exploited the cloud shadow index (CSI) proposed in [99]. The CSI index is based on the physical reflective characteristic of cloud

shadow. The CSI index is computed by combining information provided by the NIR and SWIR bands. In this way CSI index exploits the reflective characteristics of cloud shadows in the longer wavelengths. The CSI index is computed as:

$$CSI = \frac{1}{2} * (B_{NIR} + B_{SWIR}) \tag{4.2}$$

Nevertheless, due to the spectral similarity between shadows and water in the considered bands, an additional condition including shorter wavelengths, i.e., the blue band reflectance, should also be analyzed. Thus, the cloud shadow is identified in areas where the following conditions are fulfilled:

$$CSI < min(CSI) + t_1 * (mean(CSI) - min(CSI)) \tag{4.3}$$

$$B_{Blue} < min(B_{Blue}) + t_2 * (mean(B_{Blue}) - min(CSI)) \tag{4.4}$$

. Where the adjusting coefficients t_1 and t_2 are fine-tuned according to [99]. Finally, the obtained shadow mask is merged together with the original Sen2Cor shadow mask.

4.3.1.2 Multitemporal Composite Generation

When working at large scale, it is necessary to harmonize the TS of images acquired over different tiles. This is because TS of each tile are characterized by different lengths and can be acquired at different times. To solve this problem, in the preprocessing step we generate monthly, seasonal and annual composites in accordance with the sensor, the considered study area and the amount of RS data available. Details and examples of different choices are given in the experimental result session. This condition allows us to mitigate cloud occlusions problem and minimize the processing resources. To this end, we consider a statistic-based approach that computes the median value for each pixel. This approach is able to generate consistent results at large scale in an automatic way by sharply reducing the spatial noise. Let $\{X^1, X^2, \ldots, X^Q\}$ be the considered TS, which includes the optical images acquired over a month, a season or the whole year. The multitemporal pattern associated to the jth pixel of $\{X^1, X^2, \ldots, X^Q\}$ can be defined as $[x_j^1, x_j^2, \ldots, x_j^Q]$, where $x_j^1 = [x_{j,B_1}^1, x_{j,B_2}^1, \ldots, x_{j,B_B}^1]$ represents the B spectral values of the jth pixel in the first image of the TS. The jth pixel of the composite x_j^{Com} is generated by computing the band-wise median of the cloud-free multispectral pixels of the images present in $\{X^1, X^2, \ldots, X^Q\}$, as follows:

$$x_{j,B_1}^{Com} = Med\{x_{j,B_1}^1, x_{j,B_1}^2, \ldots, x_{j,B_1}^Q\} \tag{4.5}$$

$$x_{j,B_2}^{Com} = Med\{x_{j,B_2}^1, x_{j,B_2}^2, \ldots, x_{j,B_2}^Q\} \tag{4.6}$$

$$x_{j,B_B}^{Com} = Med\{x_{j,B_B}^1, x_{j,B_B}^2, \ldots, x_{j,B_B}^Q\} \tag{4.7}$$

In accordance with the availability of RS data, in the Amazon and African Sahel , we considered 12 monthly composites for the static products and 4 seasonal composites for the historic products. In the Siberia, due to the heavy cloud and snow cover, we generated only one annual composite for both static and historic products.

4.3.1.3 Multitemporal Cloud and Shadow Restoration

Cloud restoration is an important step in the optical images preprocessing part. Although we are considering the composites instead of original TS of images, missing information due to poor atmospheric conditions (e.g., thick clouds and related shadows) or defective sensors may be present in the composites. In the literature, a large effort has been devoted to solve this problem. However, to properly recover missing information, sophisticated and usually computationally intensive techniques should be used, and thus increasing significantly the computational complexity of the preprocessing part. Instead of considering computationally demanding approaches we used a simple and effective linear temporal gap filling. In this method the missing information for the jth pixel of the composite $x_{j,B_1}^{Com_T}$ at time T is restored as the band-wise average of the spectral values acquired in the previous $x_{j,B_1}^{Com_{T-1}}$ and the following $x_{j,B_1}^{Com_{T+1}}$ composites in the TS as follows:

$$x_{j,B_1}^{Com_T} = Avg(x_{j,B_1}^{Com_{T-1}}, x_{j,B_1}^{Com_{T+1}}) \tag{4.8}$$

$$x_{j,B_2}^{Com_T} = Avg(x_{j,B_2}^{Com_{T-1}}, x_{j,B_2}^{Com_{T+1}}) \tag{4.9}$$

$$x_{j,B_B}^{Com_T} = Avg(x_{j,B_B}^{Com_{T-1}}, x_{j,B_B}^{Com_{T+1}}) \tag{4.10}$$

If clouds are present in the first or last composite in TS, the second or the one before last composite are considered, respectively.

4.3.1.4 Topographical Shadow Reconstruction

The final step of optical images preprocessing is the topographical shadow reconstruction. Here, similarly to the shadow detection, we identified topographic shadow by using thresholding of the CSI index. Moreover, in the detected topographic shadow areas we further check the topographical slope. This is done in order to avoid confusion between shadow

and water bodies as they both have very similar spectral signatures associated with their low reflectances. To enhance the detected topographic shadow areas, we perform simple image dilatation.

In order to restore the detected topographic shadow, we considered the method presented in [100]. This approach is based on the assumption that the signal registered in the shadow areas is weak but can be exploited for shadow reconstruction. The shadow region can be retrieved according to the following equation:

$$y_j = S_f/S_s(x_j - M_s) + M_f \tag{4.11}$$

. Where $x_j = [x_{j,B_1}, x_{j,B_2}, \ldots, x_{j,B_B}]$ represents the B spectral values of the jth pixel in the shadow area, $y_j = [y_{j,B_1}, y_{j,B_2}, \ldots, y_{j,B_B}]$ represents the B restored spectral values of the jth pixel in the shadow area, and S_s and S_f are the standard deviation of the shadow and shadow free areas, respectively. M_s and M_f are the statistical mean value of the shadow and shadow free areas, respectively. To harmonize the obtained shadows restoration results and the surrounding spectral signatures of the original data, we used the inpainting technique based on the fast marching method presented in [101]. This condition allows us to avoid noticeable distortions on the shadow's boundaries,

4.3.2 Feature Extraction

Texture allows the accurate characterization of the contextual information of a pixel in the image. It is well known in the literature that the use of textural features can significantly improve the classification results, as they can be more distinctive than spectral features for some LC classes. Instead of considering complex textural features, we considered local standard deviation and morphological gradient. Both features are effective for the considered LC classification [102]. Local standard deviation is a widely used difference statistic measure. The morphological gradient allows the estimation of the partial derivatives in each surrounding direction of the considered pixel. The morphological gradient is defined as the difference between dilation and erosion. The dilatation and erosion for the jth pixel of the composite $x_j^{Com_T}$ by structuring element B with its origin at $x_j^{Com_T}$ are computed as follows:

$$dilatation(x_j^{Com_T}) = \max_{\beta \in B}(x_j^{Com_T}(j + \beta)) \tag{4.12}$$

$$erosion(x_j^{Com_T}) = \min_{\beta \in B}(x_j^{Com_T}(j + \beta)) \tag{4.13}$$

Both the textural features were computed in a local neighborhood of 3x3 pixels.

4.3.3 SVM Classification

In the last phase of the considered processing chain, we performed the supervised automatic classification to produce accurate LC maps. Due to the need of processing a huge amount of data, the classification algorithm must achieve a good trade-off between classification accuracy and computational burden. The classical approaches to LC mapping are based on machine learning and statistical methods. In the category of the most recent methods, deep architectures are taking off, achieving state of the art performances [103], [104]. This is because deep architectures are able to automatically extract both, low-level and high-level features and extensively exploit temporal dependencies of the RS data. Nevertheless, shallow approaches such as SVM remain competitive to deep based methods in many applications since they do not require very large training-set, are usually faster and have high generalization capability, which lead to accurate classification results [104]. Indeed, the SVM is one of the most effective methods in pattern and texture classification for the LC mapping [105]. Thus, taking into account the limited number of samples available, we selected the SVM as a classifier.

In particular, to generate the high-resolution LC maps, the SVM classifier is trained on the TS of optical composites generated in the preprocessing step. To accurately represents the spatial information together with the temporal one, an optical feature extraction step is performed to extract textural features from the first composite. In order to perform the supervised training of the classifier, a lot of effort has been devoted to generating a training set by photo-interpretation. The training is performed once for each year in each considered area independently (Amazon, African Sahel and Siberia). The SVM parameter tuning was performed with the standard cross-validation approach.

4.4 Experimental Results

In this section, we present the results of the proposed optical processing chain obtained for both the static map (Sentinel-2) and the historic maps (Landsat 5, 7, 8). The quantitative and qualitative results are reported. To perform the quantitative analysis the labeled samples were divided into independent training and validation sets. The results are presented in terms of: i) OA, ii) user accuracy and iii) producer accuracy. Moreover, to allow a visual interpretation of the results, we also present: i) composites, ii) classification maps, iii) comparison of the obtained classification maps with the medium resolution product of the CCI LC maps produced at 300 m spatial resolutions, and (iv) comparison of the obtained classification maps with the CGLC 2019 maps produced at 100 m spatial resolutions.

Table 4.3: Characteristics of composites generated for the different study areas according to the availability of cloud-free optical images.

Area	Sentinel-2	Landsat 5 - 7 - 8
Siberia	1 yearly composite (July - August)	1 yearly composite (April - September)
Amazon	12 monthly composites	4 seasonal composites
African Sahel	12 monthly composites	4 seasonal composites

4.4.1 Generation of Image Composites

Tab. 4.3 summarizes the kinds of composites generated per study area according to different optical sensors. Due to the increased revisit time of Sentinel-2 (5 days) with respect to Landsat (16 days), denser TS are available in 2019, and thus we are able to generate monthly based composites. In the case of Sentinel-2 data over the Amazon and the African Sahel, we computed 12 monthly composites. Considering that the cloud coverage over some regions can be quite intense, each monthly composite is systematically computed using the considered and the following month images (e.g., January and February are used for January composite generation, the only exception is December, where only December images are considered). This conservative choice allows us to sharply reduce the probability of having cloudy pixels in the TS. For Sentinel-2 data over the Siberia, we generated yearly composites due to the heavy cloud and snow coverage problem that hampered the use of images acquired for most of the year. The Siberia yearly composite is computed as the median of data acquired between July and August. In the case of Landsat data, we similarly considered yearly composite for the Siberia, which is computed as the median of data acquired between April and September. Finally, for Landsat data over the Amazon and the African Sahel, we computed four seasonal composites considering the optical data acquired in the following months: (i) January – March, (ii) April – June, (iii) July- September, and (iv) October – December.

Fig. 4.6 (a)(e)(i)(m) shows the results obtained when only original Sen2Cor cloud and shadow detection was applied to the composite; after the proposed cloud mask detection enhancement was applied (b)(f)(j)(n); after the proposed cloud shadow detection enhancement (c)(g)(k)(o) and after the proposed topographical shadow reconstruction (d)(h)(l)(p). It can be noticed that the original Sen2Cor mask did not detect all the clouds and shadows present on the image. However, the improvement given by the pro-

Table 4.4: Producer Accuracy (PA), User Accuracy (UA), Fscore (F1) and Overall Accuracy (OA) of the proposed processing chain for the Sentinel-2 images acquired over the Amazon, 2019.

Quality metrics	**Ever. Brod.**	**Decid. Brod.**	**Shrub Ever.**	**Shrub Deci.**	**Grass**	**Crop**	**Grass Aqua.**	**Bare**	**Build up**	**Water Ses.**	**Water Per.**
UA [%]	80.00	76.92	75.00	56.25	57.89	84.61	84.61	90.90	77.78	80.00	66.67
PA [%]	85.71	71.42	42.85	64.28	78.57	78.57	78.57	71.42	100	57.14	85.71
F1 [%]	82.75	74.07	54.54	60.00	66.67	81.48	81.48	80.00	87.50	66.67	75.00
OA [%]	74.03										

posed enhanced cloud and shadow detection can be clearly seen. Although for some cases the shadow is not perfectly removed (see Fig. 4.6 (f) and (g)), the results present a significant improvement with respect to the original Sen2cor results. Moreover, one can also note that the topographic shadow reconstruction correctly identified and reconstructed the shadow regions.

4.4.2 Static HRLC Map

The SVM models have been trained independently for each area (i.e., Amazon, African Sahel and Siberia) and for each considered year. To validate the obtained HRLC maps, the available labeled datasets have been divided into training and validation sets. The training sets are composed of 10150 samples for the Amazon, 5694 for the African Sahel and 13139 for the Siberia. The validation sets are composed of 154 samples for the Amazon, 140 for the African Sahel and 126 for the Siberia.

Tables 4.4, 4.5 and 4.6 report the user, producer and OA obtained in 2019 for the Amazon, the African Sahel and the Siberia, respectively. The validation set consists of 14 photo-interpreted samples for each considered class. It can be noticed that in the Amazon OA is equal to 74.03%, with a particularly good F1 score for tree cover evergreen broadleaf 82.75%, cropland 81.48%, grassland vegetation aquatic or regularly flooded 81.48%, bare areas 80.00% and build up 87.50%. Nevertheless, one can note that the F1 score of water permanent is quite low 75.00%, which can be attributed to the confusion with water seasonal and the grassland vegetation aquatic or regularly flooded (due to the similarity in the class definition). The OA obtained in the African Sahel is higher 78.57%. On the one hand, half of the classes (i.e., grassland vegetation aquatic or regularly flooded, bare areas, build up, water seasonal and water permanent) have F1 score higher

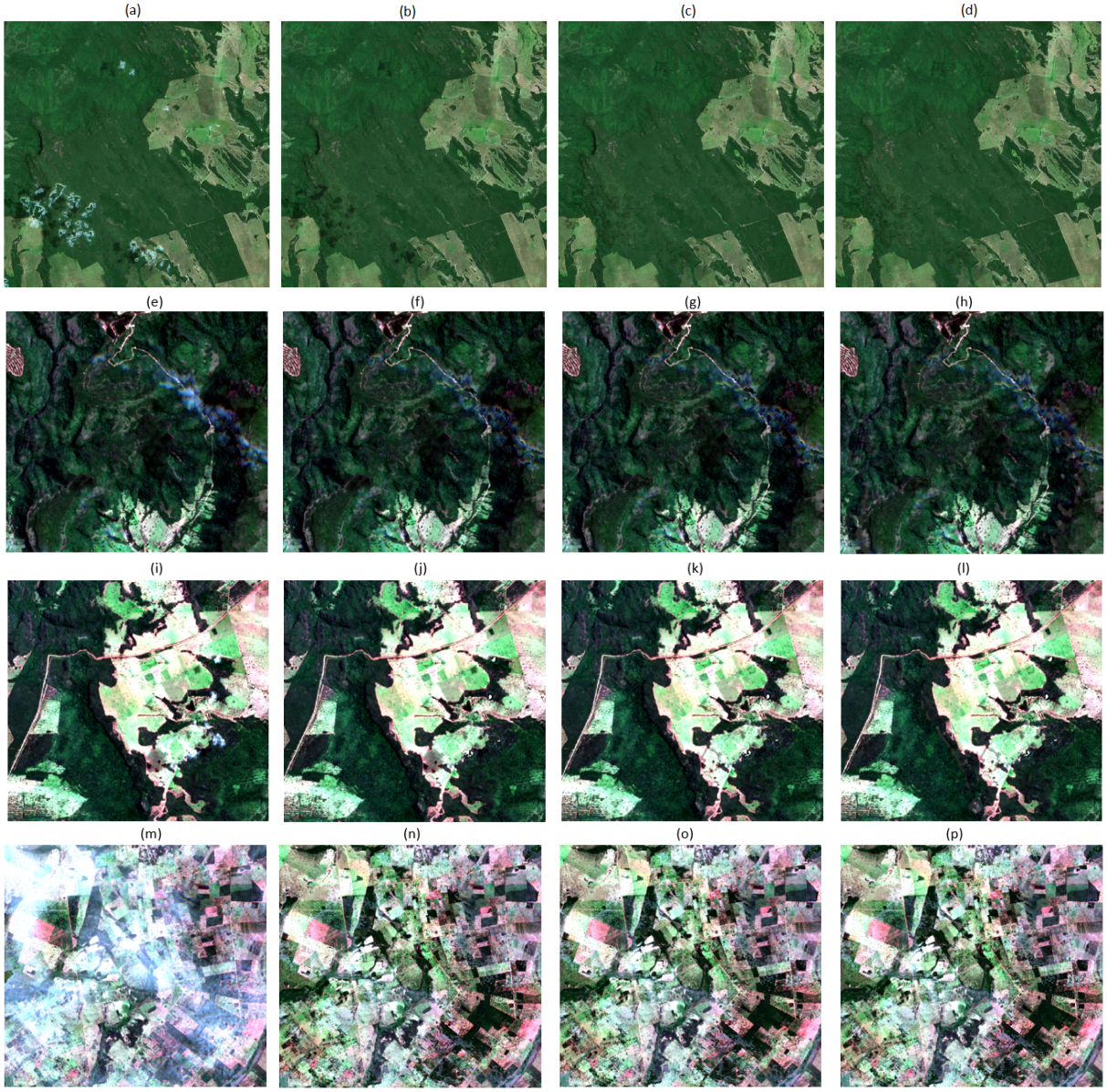


Figure 4.6: The January 2019 composite of tile 21KXT, over the Amazon obtained by (a)(e)(i)(m) the original composite; (b)(f)(j)(n) after cloud mask detection; (c)(g)(k)(o) after cloud shadow detection and (d)(h)(l)(p) after topographical shadow reconstruction.

Table 4.5: Producer Accuracy (PA), User Accuracy (UA), Fscore (F1) and Overall Accuracy (OA) of the proposed processing chain for the Sentinel-2 images acquired over the African Sahel, 2019.

Quality metrics	**Ever. Brod.**	**Decid. Brod.**	**Shrub Ever.**	**Grass**	**Crop**	**Grass Aqua.**	**Bare**	**Build up**	**Water Ses.**	**Water Per.**
UA [%]	83.33	60.00	47.36	71.42	100	100	73.68	100	92.31	87.50
PA [%]	71.43	85.71	64.29	35.71	64.29	78.57	100	100	85.71	100
F1 [%]	76.92	70.59	54.54	47.62	78.26	88.00	84.84	100	88.89	93.33
OA [%]	78.57									

Table 4.6: Producer Accuracy (PA), User Accuracy (UA), Fscore (F1) and Overall Accuracy (OA) of the proposed processing chain for the Sentinel-2 images acquired over Siberia, 2019.

Quality metrics	**Ever. Brod.**	**Ever. Need.**	**Shrub**	**Grass**	**Veget. Aqua.**	**Lichens Mosses**	**Bare**	**Build up**	**Water**
UA [%]	92.30	70.00	55.00	57.90	64.71	62.50	45.45	100	93.34
PA [%]	85.71	50.00	78.57	78.57	78.57	35.71	35.71	92.86	100
F1 [%]	88.89	58.33	64.70	66.67	70.97	45.45	40.00	96.29	96.55
OA [%]	70.64								

than 80.00%. On the other hand, some classes (i.e., shrub evergreen and grassland) have particularly low accuracy. The lowest OA accuracy 70.64% is obtained in the Siberia. This can be attributed to the poorer temporal information as only one annual composite is available. The obtained F1 score is very high for the build up 96.29% and water 96.55% classes. The lichens and mosses 45.45% as well as bare class 40.00% suffer from the lowest accuracy.

Fig. 4.7 shows classification maps of Sentinel-2 21KXT tile in the Amazon, obtained by: (a)(d)(g) the CCI middle resolution 2015, (b)(e)(h) the proposed static processing chain, while (c)(f)(i) represents the corresponding HR Google images. The CCI middle resolution 2015 product have been translated into HRLC legend, with exception for three MR classes which does not have corresponding class definition in HRLC legend. Although medium resolution product is able to correctly identify build up areas, it can be noticed that it has many problems. On one hand these classification errors can be attributed to

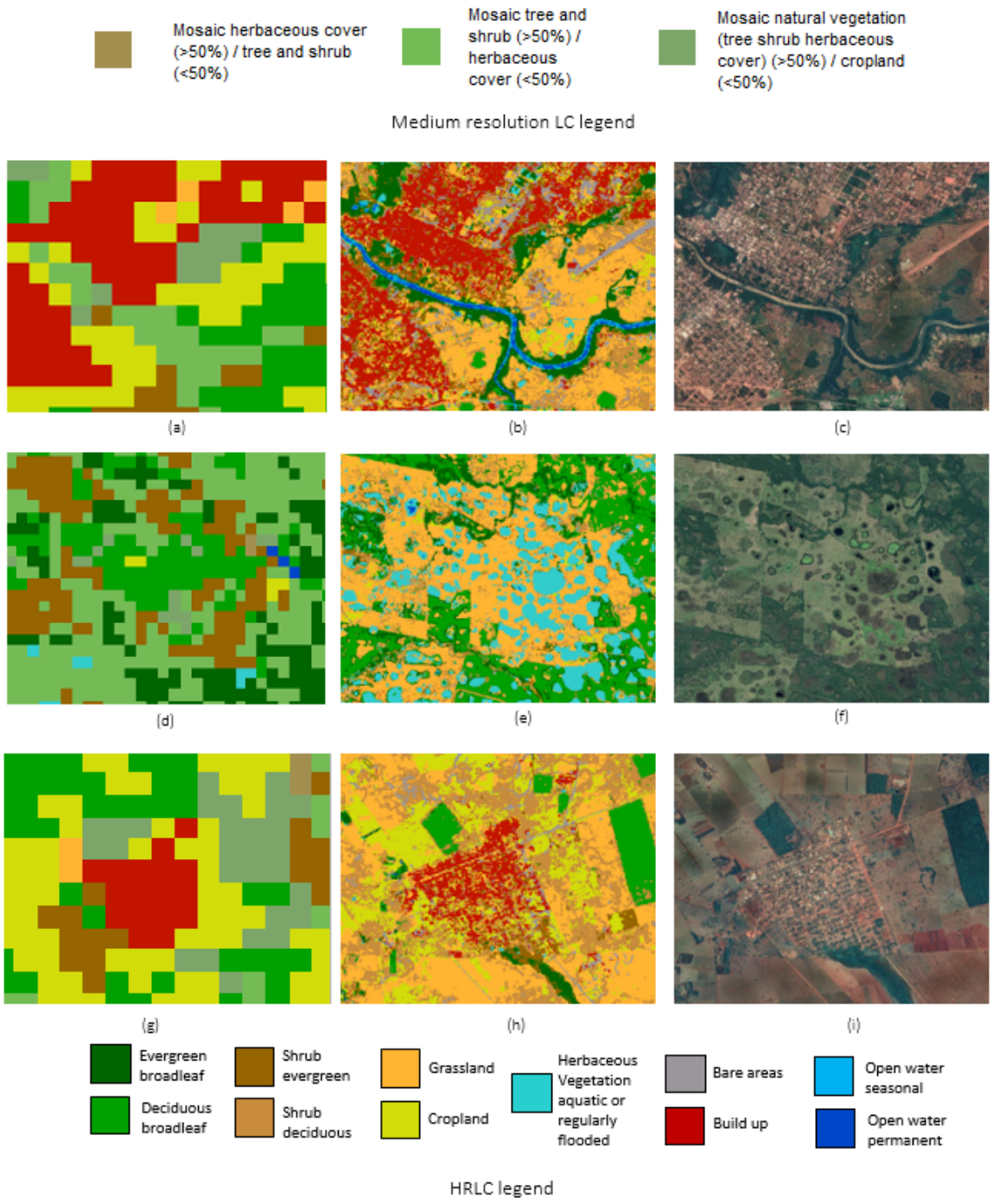


Figure 4.7: Classification maps of Sentinel-2 21KXT tile in the Amazon, obtained by: (a)(d)(g) the CCI middle resolution 2015, (b)(e)(h) the proposed static processing chain, and (c)(f)(i) the corresponding HR Google images. The MR LC legend considers only the classes which could not be translated into the HRLC legend.

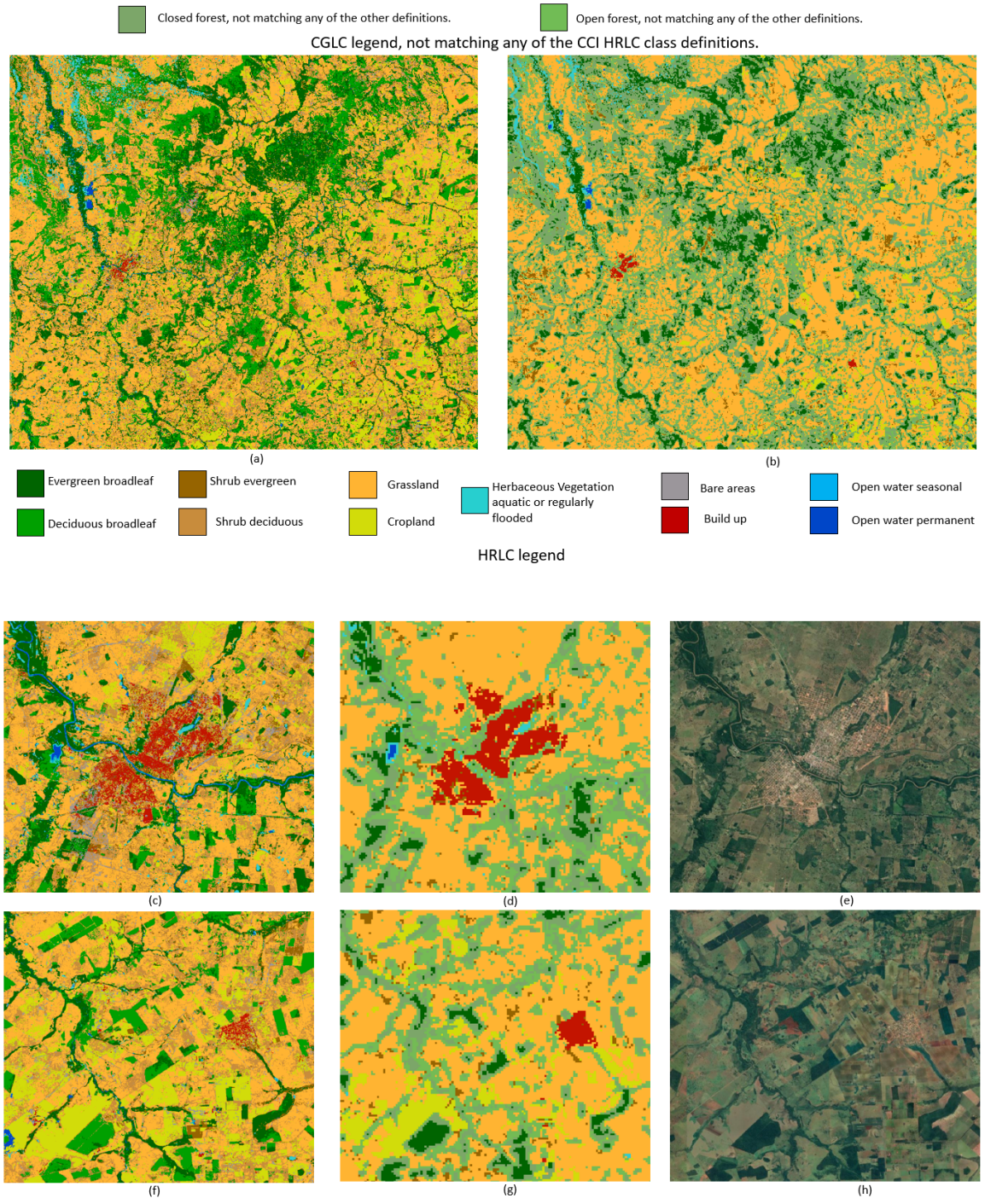


Figure 4.8: Classification maps obtained by: (a)(c)(f) the proposed processing chain 2019; (b)(d)(g) the CGLC 2019; and (e) (h) the corresponding HR Google images.

Table 4.7: The number of considered training samples for the historic maps generation.

Area	1990	1995	2000	2005	2010	2015
Siberia	6837	7167	9355	10390	12618	13010
Amazon	9507	9513	9536	9607	9615	9632
African Sahel	3535	3937	4794	4504	4644	4712

the lower spatial resolution. For instance, the medium resolution map is not able to detect river and bare areas (see Fig. 4.7 (a) and (b)). On the other hand, the map is wrong even for objects of adequate resolution 300 meters, i.e., the aquatic vegetation class is misclassified with mosaic natural vegetation (see Fig. 4.7 (d) and (e)) and the deciduous broadleaf tree are misclassified with shrubland (see Fig. 4.7 (g) and (h)). Overall, it can be concluded that the HRLC map provides not only greater spatial details but also a better quality LC maps.

Finally, in the Fig. 4.8 we have also compared the obtained classifications maps with the CGLC 2019 product, which OA accuracy is equal to 80.3% [58]. The CGLS 2019 product have been translated into HRLC legend, with exception for the two classes: closed forest, not matching any of the other definitions and open forest, not matching any of the other definitions. It can be noticed that the overall agreement between maps is high. In particular, the two maps agree in the classification of classes: build up, permanent and seasonal water, and herbaceous vegetation aquatic or regularly flooded. While quite good correspondence between tree classes is observed, the CGLS 2019 assigns most of them as mixed forest, and HRLC map provides more detailed class identification. As expected, the confusion between grassland and cropland can be observed which can be explained by the very similar class definition.

4.4.3 Historic HRLC Maps Production

Tab. 4.7 shows the number of considered training samples for the historic maps generation, while the validation set for each of the considered cases includes 126 photo-interpreted samples.

Tables 4.8 and 4.9 and 4.10 show the producer accuracy, user accuracy, F1 score and OA of the proposed processing chain for the Landsat images acquired in 1990, 1995, 2000, 2005, 2010 and 2015 over the Amazon, the African Sahel and the Siberia, respectively. For the Amazon the OA varies between 61.11% in 1995, 2000 and 78.57% in 2005.

For all the considered years the classifier achieved the lowest accuracy for the cropland, whereas good and stable results were obtained for shrub, build up and water. On the other hand, in the African Sahel OA is lower than in the Amazon, ranging from 46.03% up to 62.98%. For most of the years, the classifier performs well for the water and tree evergreen broadleaf. Similarly to the Amazon, the lowest accuracy is achieved for cropland and grassland, which are very similar, and thus very challenging to distinguish. One can also notice, that the better OA accuracy achieved in 2000 (62.98%) can be attributed to higher number of training samples available for this year. Thus, suggesting that the number of training samples in the African Sahel is overall insufficient. The OA in the Siberia varies between 57.93% in 1990 and 65.07% in 2005 and 2015. The best F1 score is obtained for water and build up, whereas for the lichens and mosses the accuracy is the lowest. This can be attributed to the lichens and mosses similarity to the aquatic vegetation, which at 30 m spatial resolution can be very challenging to differentiate.

Figures 4.9 and 4.10 show the classification maps of the 21 KXT tile in the Amazon and the 42 WXS tile in the Siberia, respectively. The obtained maps were produced by the proposed processing chain in: (a) 2015; (b) 2010; (c) 2005; (d) 2000; (e) 1995 and (f) 1990. Although the classification maps in diverse years present quite good correspondence, it can be noticed that the accuracy of the grassland aquatic vegetation class varies quite significantly in the Amazon between 1990 and 2015. Similarly in the Siberia some confusion in the grassland aquatic vegetation with grassland class can be observed. Moreover, the spatial distribution of the lichens and mosses varies significantly in diverse years, which depends on the low accuracy in this class identification.

4.5 Discussion and Conclusion

In this chapter, we have presented a processing chain for the classification of the high spatial resolution multitemporal and multispectral images at the large scale with an objective to support climate change studies. Both processing chains includes preprocessing, which aims to harmonize the images, textural features extraction and classification based on spectral, temporal and textural features. Due to the variability in the data availability for static and historic products, the main difference between the two processing chains lies in the composites generation. For the static product, we consider 12 monthly composites, while for the historic 4 seasonal composites.

Please note, that the proposed static processing chain is designed to work at subcontinental scale. However, at the time of writing this thesis, the proposed static and the historic processing chains have been applied over the considered historic area, in particular in the Amazon, the African Sahel and the Siberia. The Sentinel-2 images have

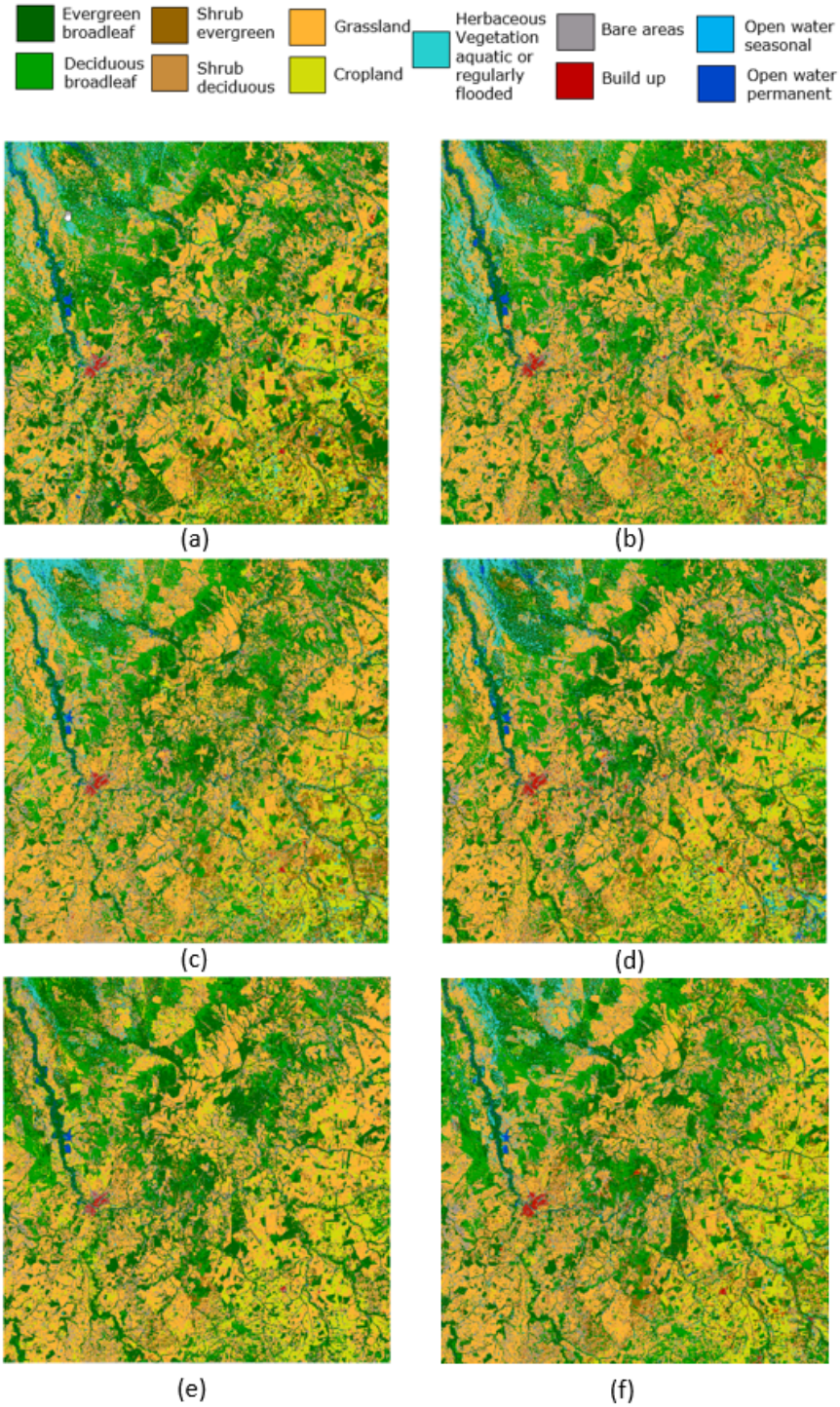


Figure 4.9: Classification maps of the 21 KXT tile in the Amazon obtained by the proposed processing chain in: (a) 1990; (b) 1995; (c) 2000; (d) 2005; (e) 2010 and (f) 2015.

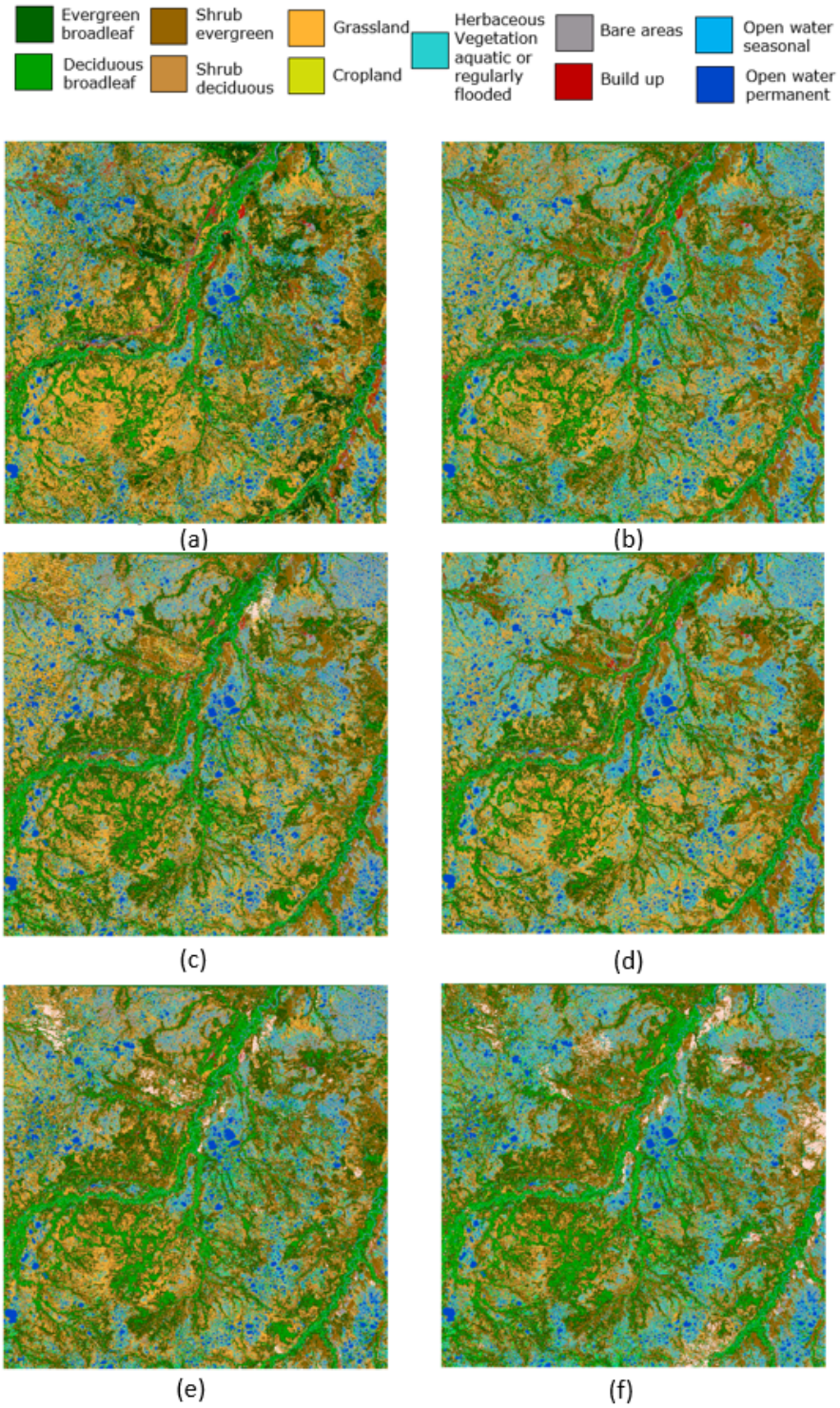


Figure 4.10: Classification maps of the 42WXS tile in the Siberia obtained by the proposed processing chain in: (a) 1990; (b) 1995;b (c) 2000; (d) 2005; (e) 2010 and (f) 2015.

been used for the static processing chain in 2019, while Landsat 5, 7 and 8 images have been used in the 1990, 1995, 2000, 2005, 2010 and 2015. The generated composites have been qualitatively analyzed and demonstrated the effectiveness of the cloud, cloud shadow and topographical shadow detection and reconstruction. Moreover, the obtained thematic products demonstrate the effectiveness of the proposed processing chain. The OA of the static thematic products is equal to 74.03%, 78.57% and 70.64% for the Amazon, the African Sahel and the Siberia, respectively. The obtained OA is comparable to the other thematic products available at that scale (i.e., FROM-GLC). For the historic thematic products the obtained accuracy is as expected lower than for the static products (ranging from 61.11% - 78.57% for the Amazon, 46.03% - 62.98% in the African Sahel and 57.93% - 65.07% in the Siberia). The better OA for the historic analysis in the Amazon can be associated to the higher number of training samples and images available in this area in respect to the African Sahel and Siberia. It should be noted, that the particularly low accuracy for the historic products in the African Sahel and in the Siberia, require further analysis and improvement. It is expected that this kind of low accuracy maybe attributed to the cloudy acquisitions, low number of images available in the past, insufficient number of training samples and the quality of training samples. Therefore, we plan to further investigate this accuracy drops and to enlarge the number of samples in the training set. Finally, the obtained HRLC maps have been qualitatively compared with the CCI medium resolution LC maps. The analysis carried out over one Sentinel-2 tile in the Amazon, showed that apart from significant increase of spatial details the produced maps are also more reliable.

The requirement of the HR photo-interpreted training set can be considered as a possible limitation of the proposed method. This, however, can be possibly solved by applying domain adaptation techniques. Moreover, we plan to investigate more sophisticated textural features (Gabor filter) as well as the topographical features (DTM, slope and aspect).

Table 4.8: Producer Accuracy (PA), User Accuracy (UA), Fscore (F1) and Overall Accuracy (OA) of the proposed processing chain for the Sentinel-2 images acquired over the Amazon in 1990, 1995, 2000, 2005, 2010 and 2015.

Year	Quality metrics	Ever. Brod.	Decid. Brod.	Shrub	Grass	Crop	Grass. Aqua.	Bare	Build up	Water
1990	**UA** [%]	50.00	50.00	88.89	47.05	75.00	85.71	64.70	100	87.50
	PA [%]	64.28	71.42	57.14	57.14	42.85	85.71	78.57	50.00	100
	F1 [%]	56.25	58.82	69.56	51.61	54.54	85.71	70.96	66.67	93.33
	OA [%]	67.46								
1995	**UA** [%]	42.85	47.36	100	39.28	66.67	83.33	87.5	83.33	80.00
	PA [%]	71.42	64.28	78.57	78.57	14.28	35.71	50.00	71.42	85.71
	F1 [%]	52.38	54.54	88.00	52.38	23.52	50.00	63.63	76.92	82.75
	OA [%]	61.11								
2000	**UA** [%]	64.58	42.85	100	52.63	60.00	72.72	42.85	82.75	85.71
	PA [%]	77.78	71.42	42.85	77.78	21.42	77.78	57.14	35.71	85.71
	F1 [%]	71.42	52.63	60.00	63.63	32.25	76.47	50.00	50.00	85.71
	OA [%]	61.11								
2005	**UA** [%]	80.00	54.16	100	75.00	100	88.89	77.78	100	76.47
	PA [%]	85.71	92.85	100	85.71	35.71	57.14	100	57.14	92.85
	F1 [%]	82.75	68.442	100	80.00	52.63	69.56	87.50	72.72	83.87
	OA [%]	78.57								
2010	**UA** [%]	53.84	60.00	100	46.15	77.78	77.78	72.73	84.62	85.71
	PA [%]	50.00	85.71	78.57	85.71	50.00	50.00	57.14	78.57	85.71
	F1 [%]	51.85	70.58	88.00	60.00	60.87	60.87	64.00	81.48	85.71
	OA [%]	69.04								
2015	**UA** [%]	70.58	76.47	83.33	50.00	55.56	100	75.00	92.30	66.67
	PA [%]	85.71	92.85	71.43	78.57	35.71	50.00	42.85	85.71	100
	F1 [%]	77.41	83.87	76.92	61.11	43.47	66.67	54.54	88.89	80.00
	OA [%]	71.43								

Table 4.9: Producer Accuracy (PA), User Accuracy (UA), Fscore (F1) and Overall Accuracy (OA) of the proposed processing chain for the Sentinel-2 images acquired over the African Sahel in 1990, 1995, 2000, 2005, 2010 and 2015.

Year	Quality metrics	Ever. Brod.	Decid. Brod.	Shrub	Grass	Crop	Grass. Aqua.	Bare	Build up	Water
1990	**UA** [%]	58.82	35.00	41.17	27.27	53.84	83.33	58.33	75.00	100
	PA [%]	71.42	50.00	50.00	21.42	50.00	35.71	100	21.43	100
	F1 [%]	64.51	41.17	45.16	24.00	51.85	50.00	73.68	33.33	100
	OA [%]	55.56								
1995	**UA** [%]	73.33	52.38	36.00	29.41	50.00	100	41.67	0.00	100
	PA [%]	78.57	78.57	64.28	35.71	14.28	42.85	71.42	0.00	100
	F1 [%]	75.86	62.85	46.15	32.25	22.22	60.00	52.63	0.00	100
	OA [%]	53.97								
2000	**UA** [%]	85.71	58.82	40.00	35.92	75.00	100	50.00	100	100
	PA [%]	85.74	71.42	71.42	42.85	42.85	64.28	42.85	60.00	100
	F1 [%]	85.74	64.51	51.28	38.70	54.54	78.26	46.15	75.00	100
	OA [%]	62.98								
2005	**UA** [%]	64.28	26.67	18.51	15.00	60.00	100	52.17	100	100
	PA [%]	64.28	28.57	35.71	21.42	21.42	35.71	85.71	21.42	100
	F1 [%]	64.28	27.58	24.39	17.64	31.57	52.63	64.86	35.29	100
	OA [%]	46.03								
2010	**UA** [%]	75.00	41.67	26.31	100	100	69.23	61.53	88.89	93.33
	PA [%]	64.28	71.42	71.42	7.14	7.14	62.28	57.14	57.14	100
	F1 [%]	69.23	52.63	38.46	13.33	13.33	66.67	59.25	69.56	96.55
	OA [%]	55.56								
2015	**UA** [%]	90.90	35.71	35.00	46.15	100	100	54.54	57.14	100
	PA [%]	71.42	71.42	64.51	35.29	35.29	28.57	78.57	28.57	100
	F1 [%]	80.00	47.62	46.15	40.00	52.63	42.85	64.51	38.46	100
	OA [%]	57.43								

Table 4.10: Producer Accuracy (PA), User Accuracy (UA), Fscore (F1) and Overall Accuracy (OA) of the proposed processing chain for the Sentinel-2 images acquired over the Siberia in 1990, 1995, 2000, 2005, 2010 and 2015.

Year	Quality metrics	Ever. Brod.	Decid. Brod.	Shrub	Grass	Veget. Aqua.	Lichens moses	Bare	Build up	Water
1990	**UA** [%]	47.36	58.33	72.72	29.03	64.28	0.00	57.14	87.5	86.67
	PA [%]	64.28	50.00	57.14	64.28	64.28	0.00	28.57	100	92.85
	F1 [%]	54.54	53.84	64.00	40.00	64.28	0.00	38.09	93.33	89.65
	OA [%]	57.93								
1995	**UA** [%]	57.14	100	57.89	34.61	50.00	0.00	61.53	100	100
	PA [%]	85.71	57.14	78.57	64.28	50.00	0.00	57.14	78.57	100
	F1 [%]	68.57	72.72	66.67	45.00	50.00	0.00	59.25	88.00	100
	OA [%]	63.49								
2000	**UA** [%]	55.00	100	60.00	45.45	62.50	0.00	62.50	86.67	80.00
	PA [%]	78.57	64.28	85.71	71.42	71.42	0.00	35.71	92.85	85.71
	F1 [%]	64.70	78.26	70.58	55.56	66.67	0	45.45	89.65	82.75
	OA [%]	65.07								
2005	**UA** [%]	60.00	80.00	43.75	41.37	66.67	0.00	53.84	100	81.25
	PA [%]	64.28	57.14	50.00	85.71	71.42	0.00	50.00	85.71	92.85
	F1 [%]	62.06	66.67	46.67	55.81	68.96	0.00	51.85	92.30	86.67
	OA [%]	61.90								
2010	**UA** [%]	55.00	87.50	45.45	35.71	55.56	100	50.00	100	87.50
	PA [%]	78.57	50.00	71.42	71.42	35.71	14.28	28.57	92.85	100
	F1 [%]	64.70	63.63	55.56	47.61	43.47	25.00	36.36	96.29	93.33
	OA [%]	60.31								
2015	**UA** [%]	75.00	80.00	44.44	30.00	84.61	100	62.50	86.67	87.50
	PA [%]	85.71	57.14	85.71	42.85	78.57	7.14	35.71	92.85	100
	F1 [%]	80.00	66.67	58.53	35.29	81.48	13.33	45.45	89.65	93.33
	OA [%]	65.07								

Part II

Challenges in the Analysis of the Multispectral and Multitemporal RS data at the Operational Level

Chapter 5

Background

This chapter aims at providing an overview of the state-of-the-arts methods that cope with some of the main challenges related to the classification of multitemporal multispectral RS data at operational level. In particular, we focus on: i) mitigating the effects of the cloud coverage, and ii) addressing the definition of large-scale training sets. For the first challenge, we describe the cloud detection methods widely used to generate cloud and shadow masks for the high spatial resolution systems considered in this thesis, i.e., Sentinel 2 and Landsat. Moreover, a brief overview of the main cloud restoration approaches used to recover the missing spectral information is presented. With regard to the problem of the large-scale training set generation, we focus the attention on the supervised, unsupervised and semi-supervised domain adaptation techniques typically employed to address this issue.

5.1 Multitemporal Multispectral Data: Cloud Cover Problem

ESA's Sentinel-2 and NASA's Landsat satellites provide huge amount of optical images with high spatial and temporal resolution every year. These dense TS of multispectral data are used for a wide range of applications enabling multitemporal monitoring of physical phenomena. Nevertheless, one of the main challenges in their use is related to missing information caused by cloud occlusions. Cloud cover impedes optical RS sensors from acquiring a clear image of the Earth's surface. Therefore, to properly use multitemporal optical data, one needs to first automatically detect the clouds and then reconstruct the missing spectral information. This is particularly true for operational applications, when large-scale datasets are considered and the cloud coverage issue is unavoidable. However, as already mentioned in Section 4.3.1.1 cloud detection is a complex task. This is because there are many cloud types that can be characterized by different transparency. Thus, depending on the underneath landscape, the resulting signal can be a diverse mixture of land and cloud spectral signature.

In the literature, many cloud detection algorithms have been proposed for Landsat data [106, 107]. In [106], the authors present the Automated Cloud Cover Assessment method, which aims to automatically detect the clouds present in Landsat images by using the information provided by the thermal infrared band. Another way to address the cloud detection problem is by considering a multispectral method, which takes advantage from the information in different bands. The Haze Optimized Transformation approach exploits the correlation between the red and blue bands that occurs when the pixel is cloud-free [108]. In [107], the cloud masks for Landsat 8 data are derived by using spectral and statistical properties. In [109], authors proposed a linear model, which uses the near-infrared to detect the haze in the visible bands. In [110], an empirical radiative transfer model was proposed for Landsat 8. In particular, to model cloud effects, the method assumes the linear relation between the Landsat 8 cirrus band and the rest of the bands. However, this group of methods works properly only with thin-cloud cover, as the thicker clouds are impacting all the spectral bands. The USGS distribution of the Landsat L2 products is based on the FMask method for cloud and shadow detection [111]. The method uses the information from the surface reflectances and the brightness temperatures of the Thermal InfraRed (TIR) band. The TIR band is particularly useful since the temperature of cloud-free surfaces is usually much higher than the cloud top temperature. In the first step, by using a pre-defined threshold in the TIR band, the Fmask method detects potential cloud and shadow pixels. In the second step, the method computes the cloud probability for all the pixels not identified as potentially cloudy. The ones having the highest probability are included into the group of potential cloud pixels. Finally, the potential cloud pixels are segmented, and by iterating on the cloud altitude the method tries to match the clouds with shadows. If the match with a shadow is found the cloudy pixel if confirmed.

Few cloud detection methods have been defined for Sentinel-2, due to the recent acquisition of these images that started in 2015. The ESA distribution of the Sentinel-2 cloud mask is based on Sen2Cor [112], which uses the four thresholds (in the visible range, near infra-red range, NDSI and threshold on the reflectance in the SWIR-Cirrus band) and each of them produces a probability map of cloud presence. Next, the global probability mask is generated by multiplying all the probability masks. Finally, by applying various thresholds to the global probability mask the following masks are obtained: a low probability cloud mask (recently renamed "unclassified"), a medium probability cloud mask and a high probability cloud mask. The cloud shadows detection is based on the fact that shadows are darker and their positions have to be correlated to the clouds. Recently, a new version of the Fmask algorithm has been released, i.e., Fmask 4 [113]. This version does not require the thermal band to be applied, and thus can be applied

to both Landsat and Sentinel-2 data. Indeed, the USGS distribution of the Sentinel-2 products is based on Fmask 4. In particular, the method is based on the view angle parallax of the NIR bands. To minimize the confusion with the snow and build-up class, Fmask 4 uses contextual and spectral information. Moreover, it relies on a global surface water map to avoid confusion with water pixels. The cloud shadows are detected on the base of cloud and clouds shadows similarity. The drawback of the Fmask 4 applied to Sentinel-2 data is its 20 m spatial resolution. As the use of deep learning techniques is taking off in the RS literature, advanced deep methods to cloud and shadow detection have been also introduced. Based on residual learning and semantic segmentation, CloudNet proves to provide robust results for cloud and haze classification for Sentinel-2 [114]. The authors have compared its results to other state-of-the-art architectures, such as Deeplab v3+. CloudNet proves to overcome other architectures especially for the cirrus cloud recognition. In [21], authors compare the classification accuracy of a fully CNN architecture trained on Landsat 8 to detect cloudy pixels on Landsat-8 and Sentinel-2 images. The obtained results show that their architecture outperforms standard operational threshold-based algorithms when tested on the same Landsat-8 dataset. However, the results are similar to standard operational algorithms when validated on datasets different from Sentinel-2 or Landsat-8. Although these methods achieve higher detection accuracy they are computationally expensive and require large collection of statistically significant training data. Therefore, the computational burden hampers their use at the operational level.

Both Sen2Cor and FMask are mono-temporal methods, which means that they use only the information provided by the image to analyze for detecting clouds and shadows. On the contrary, the recently released MACCS-ATCOR Joint Algorithm (MAJA) is based on the analysis of the multitemporal information to detect clouds and shadows [115], [97]. MAJA is a recurrent method, and thus it needs to be initialized when analyzing the first image in a TS. This is done by applying a set of thresholds in the blue, red, NIR and SWIR bands in order to define a mono-temporal cloud mask. For a given image, the corresponding cloud mask is generated by using a reference cloud-free image previously acquired. If no recent cloud-free images are available, a composite reference image is created by using the last cloud-free observation per pixel. Please note that every time a new image is acquired, its reference image has to be defined. The reflectance values of the blue band belonging to the cloud-free reference image are compared with the ones present in the image to be analyzed. If a rapid increase of reflectance values is detected, the pixel is labeled as cloudy. However, as ground surface reflectances may change in time, the method also takes into account the difference in time between the acquisition of the analyzed and the reference images. The longer the time difference is, the higher is

the threshold used to evaluate the reflectance increase. The pixel is confirmed as cloudy if none of the following condition is true: i) the reflectance variation in the red band is higher than the reflectance variation in the blue band, and ii) a similar correlation of the pixel with its neighboring pixels can be observed for the last 10 acquisitions (this is because it is very improbable that the cloud will be observed in 10 last acquisitions in the same place and shape). Also MAJA shadow detection is based on a multitemporal approach. The approach provides information about two kinds of masks, the geometric and the radiometric shadow masks. The geometric cloud detection is an intersection of regions where there is a high probability of cloud shadow to occur (due to presence of a cloud) and regions where the darkening in the red band can be observed. Then, by considering cloud altitude, MAJA detects the cloud shadow. This is done by computing a geometric projection, which uses solar and viewing directions. Then, within detected possible shadow region, MAJA thresholds the variation of the reflection in the red band, and thus confirms or discards the cloud shadow pixels. The goal of the radiometric mask is to identify the shadow of clouds, which may lye outside the analyzed image. This is done by applying the detected threshold in the red band (done in the geometric step) for the whole image. Finally, similarly to the cloud detection, in the last step the correlation between different acquisitions is also checked to discard false shadows. Few studies have been carried out to compare the accuracy of the Fmask, Sen2Cor and MAJA on Sentinel 2 images. In [97], the OA of cloud and shadow mask provided by the Sen2Cor is equal to 84%, much lower than the one provided by Fmask 4 (OA equal to 90%), and MAJA (OA equal to 91%). In [116], authors compared the cloud cover detection algorithms using Sentinel-2 images over the Amazon rain forest, which is known for its dense cloud coverage. The reported accuracy indicated the best OA for FMask (90%), followed by Sen2cor's (79%) and MAJA (69%). However, in [117], the authors clarified that such low accuracy achived by the MAJA was due to the small number of images in the TS since only Sentinel-2A were considered [117], thus neglecting the main assumptions of this approach.

In addition to the accurate cloud and shadow detection task, a large effort has been devoted in the RS literature to perform cloud restoration. Several multitemporal approaches have been defined to reconstruct the spectral signature covered by the clouds leveraging the temporal cloud-free information available in previously acquired optical images. In particular, such a goal can be achieved cloning information from multitemporal satellite images [118], by using dictionary learning [119] or multitemporal dictionary learning [120] as well as by considering approaches based on deep architectures [121]. Although multitemporal methods perform well, they need a cloud-free reference image, which in some cases can be challenging. Moreover, they are usually computationally

intensive, increasing significantly the preprocessing complexity. To mitigate this issue, several mono-temporal spatial-based methods have been proposed. Such methods rely on the spatial autocorrelations between local and non-local regions of the image. The spatial based methods assume that the missing regions have similar statistical or geometrical properties to the rest of the image. If cloudy pixels constitute small portion of the image, the interpolation methods (such as nearest-neighbor [122] and kriging [123]) proved to be effective. For the larger gap filling of textured region the exemplar-based method [124] or the bandelet-based inpainting [125] have been proposed. Nevertheless, these techniques proves to perform poorer then mutlitemporal or spectral ones when dealing with large cloud cover or complex texture.

This analysis of the literature points out that a large effort has been devoted to cloud detection and restoration techniques in the past. This is because it is a necessary preprocessing step especially for shallow classifiers. Although many successful cloud detection and cloud restoration techniques have been proposed in the literature, the ones that achieve the highest performance are usually sophisticated and computationally intensive. By including them in the image pre-processing phase of a standard processing chain for the classification of multitemporal data, the computational complexity significantly increases. Therefore, their use becomes unfeasible from the operational view point. On the other hand, the recent advances of sophisticated multitemporal deep learning models changed the perspective. Indeed, nowadays instead of running accurate cloud detection and restoring methods, new deep architectures (e.g. Recurrent neural networks (RNN), LSTM) have been proposed. These methods are able to extensively exploit dense TS and in accordance with recent qualitative studies they prove to be robust to cloud cover [28], [29]. This is why in this thesis, we aim to quantitatively evaluate what is the impact of the cloud coverage on the classification performances of a deep LSTM when the clouds are restored, neglected or kept.

5.2 Generation of the Representative Training Set at Large Scale

The supervised classification of RS data at the large scale represents a challenging task from the operational view point. To achieve accurate classification results, the quality and the quantity of training samples, as well as sampling strategy are equally important to the learning paradigm. Let us consider a domain D, where $P(x, y)$ is the distribution of the analyzed data, $x \in X$ represents a vector of features of the considered RS image and $y \in \Omega$ corresponds to set of class labels. In the classical supervised classification problem, the algorithm will try to estimate the distribution $\widehat{P}(x, y)$ given the training set

$T = (x_i, y_i)_i$, where $x_i \in X, y_i \in \Omega$. Thus, the more representative is the training set of the underlying distribution of the considered study area, the better is the classification result. However, the statistical distribution of the LC classes may change across the study area because of the spatial variability of the spectral signature. Indeed, the same LC class may present different spectral behavior in satellite images acquired in different regions due to many variables, including atmospheric conditions, acquisition dates or topographic condition [126]. Moreover, at the large scale, a proper sampling strategy is needed to capture labeled samples representative of the spectral behavior of the LC classes for the whole study area. However, gathering large amount of labeled samples is costly, time consuming, and thus often unfeasible. Therefore, in many operational applications, the quality of available training set is too poor to appropriately estimate the data distribution $\widehat{P}(x, y)$ [127].

A way to reduce the need and effort of collecting reference samples is adapting the classifier trained on the source domain (i.e., images for which the training set is available) to classify the target domain (i.e., images for which no labeled samples are available). To properly handle this issue, transfer learning methods have been introduced in the RS literature [128]. These methods aim to reduce the discrepancy across domains to re-use the information available in the source to classify the target domain, assuming that the two domains have different, but related, distributions [127]. In this literature review, we will focus on a special kind of transfer learning called domain adaptation. Let us consider two domains: target and source, which distribution can be described as $P_T(x, y)$ and $P_S(x, y)$, respectively. The goal of domain adaptation methods is to approximate well $P_T(x, y)$ by using information provided in $P_S(x, y)$. In general, domain adaptation methods can be divided into four main categories: (i) invariant feature selection, (ii) representation matching, (iii) adaptation of classifiers, and (iv) selective sampling. In this thesis, we focus attention on the case of the representation matching. The aim of the representation matching is to adapt the source and target domain distribution to the common latent space, where a single classier can be applied. Most of the domain adaptation methods available in the literature, assumes that target and source domain consider the same set of classes [128]. Depending on the availability of the labeled units in the domains, representation matching can be further subdivided into three subcategories: supervised, semi-supervised and unsupervised.

The unsupervised methods are usually based on the exploitation of the data geometrical structure. Looking back in time, a common approach to the unsupervised domain adaptation was the Canonical Correlational Analysis (CCA) [129], proposed by Hotelling in 1936 [130]. This approach has been widely used for finding a common latent space across different applications, e.g., computer vision [131]. The method aims to find linear

projections of two multidimensional variables that are maximally correlated. The extension of this method is the Kernel canonical correlation analysis (KCCA), which increases the flexibility of the feature selection kernelized CCA [132]. Nevertheless, both methods require that the points are pair-wise between the two domains, which can be challenging in many applications. Another important group of unsupervised domain adaptation methods, such as Transfer Component Analysis (TCA), is based on the idea of finding a projection that minimizes the discrepancy between two domains data distributions [133]. However, in this case the same features have to be considered in the source and target domains. Finally, the geodesic distances along domains have been studied in [134] and then extended to Geodesic Flow Kernel (GFK) [135], but both methods requires the same data dimensionality. In the category of the most recent approaches, the domain adaptation based on adversarial learning is a popular direction. The idea is to train a feature generator to produce domain invariant features for target and source domains to fool a domain discriminator, which is trained to identify domain labels for features produced by generator [136], [137]. [138]. However, it might happen that the discriminative distribution of the target domain samples may be lost, and thus result in the low classification accuracy [139]. To solve this problem in [140], authors proposed dual-adversarial network, which performs a domain specific feature adaptation and classification.

Semi-supervised methods try to cope with a more relaxed problem, where the labeled samples are available only in the source domain and for the target domain only the unlabeled sampled are considered. It is important to note that this group of methods are based on the assumption that the distributions of source and target domains are different, but still close enough. Otherwise, source domain samples would be of no help for solving problem in the target domain. An example of the semi-supervised method is the abovementioned GFK approach, which can be also used with labeled data in the source domain when the source domain eigenvectors are found with a discriminative feature extractor (e.g. partial least squares [135]). TCA can be extended to a semi-supervised setting by maximizing the dependence (measured by Hilbert-Schmidt Independence Criterion) between kernel of labels and features [133]. Alternative methods, called Optimal Transport, are looking for the transportation plan of the probability distribution function between the source and the target domain. In this way the methods estimate the invertible and non-linear transformation. In [141], authors proposed a regularized unsupervised optimal transportation model, which put the constrain on the source domain samples belonging to the same class to remain close each other during transportation. Therefore, the method is able to exploit the domain distributions and the source domain labeled samples. Domain-adversarial neural networks (DANN)[137] can be applied to semi-supervised and unsupervised application. DANN learns domain independent labels by training CNN with

loss function, which minimizes class specific component while maximizing domain specific component. This way, the model improves the classification accuracy, while losing the ability to identify from which domain the sample is coming from. In [142], authors have proposed Sourcerer, which first trains the CNN classifier on source domain samples, then continues the training on the target domain samples using novel Bayesian-inspired regularizer. The approach has been tested on LC mapping providing state-of-the-art accuracy. Minimax entropy (MME) is another recent method proposed in the literature [143]. For each class from the considered labeled samples MME aims to learn a representative datapoint (prototype) and than minimizes its distance with unlabeled samples. As the labeled samples are coming mainly from the source domain, MMA uses adversarial method to align prototype samples to the target domain distribution. However, both DANN and MME requires very large number of labeled samples.

Supervised domain adaptation assumes to have labeled sampled in both the target and source domains. However, the number of labeled sampled in the target domain is usually much smaller with respect to the one available in the source domain. The above-mentioned GFK and TCA methods can also be used when both domains include labeled samples. In [144], authors proposed Max-Margin Domain Transforms method. By using the classification loss, the method optimizes both the projection of target features into the source domain and the classifier parameters. The method is able to cope with different feature dimensionalities and is scalable to large training sets. Manifold alignment is a special group of supervised methods, which aims to find unified representation of multiple domains. In other words, this group of methods reduces to finding a projection to the common latent space, where the classes separability is enhanced. In 2014 Tuia et al. [145] proposed semi-supervised manifold alignment (SSMA). SSMA projects data from different domains into a common latent space, where the geometry of each manifold is preserved, while the same class samples are placed close to each other and different class samples are placed far apart. In [27], the authors proposed kernelized version of SSMA, called kernel based manifold alignment method (KEMA), which performs better with a high-dimensional data and with strong nonlinear deformations. Nevertheless, the method requires labeled samples for each domain. In [146], this requirement has been relaxed, by exploiting semantic ties, i.e. samples coming from the same object, which are in common between domains. However, in this case KEMA requires partial spatial overlap between the domains.

Even though some supervised domain adaptation techniques, such as KEMA requires only a few training set samples for source and target domains, this condition can be challenging in some operational applications. On the other hand, the unsupervised domain adaptation techniques do not requires any labeled data but if the statistical distributions of

the two domains are too different they fail. In this thesis we propose a novel unsupervised domain adaptation technique, which leverages on the publicly available low resolution thematic products for the target domain and on a small set of labeled samples for the source domain. The use of the low resolution thematic product of the target domain allows us to sharply increasing the reliability of the results obtained without the need of collecting ground reference data. The proposed solution is thus effective from the operational view point, especially when there is the need of defining training sets at large scale.

Chapter 6

A Long Short Term Memory Classifier Robust to Images with Cloud Cover

LSTM has not only proven to be one of the most promising approaches to the processing of the TS but qualitative analysis has also shown its robustness to the cloudy images within the TS. In this chapter we aim to quantitatively assess the capability of the LSTM to handle different amounts of cloud coverage present in the images of TS having different lengths. In particular, we study the impact of diverse clouds representation on the classifier accuracy, by assigning cloudy pixels values of: (i) simulated cloud values, (ii) zeros values, and (iii) restored values by simple linear temporal gap filling (i.e., average of the spectral values acquired in the previous and following cloud-free images in the TS). Moreover, we study the effects of the cloud positions within the TS. [1]

6.1 Introduction

Each year the increasing number of multispectral multitemporal images is available. The possibility to define TS of multispectral images enables the development of many applications ranging from hazard assessment to agricultural monitoring. However, as mentioned in the previous chapter, their usage is hampered by the the cloud coverage that occludes the ground cover that has to be monitored. To address this issue in some applications only cloud-free images are considered, thus reducing significantly the number of items

[1]Part of this chapter appears in I. Podsiadlo, C. Paris, L. Bruzzone, "A study of the robustness of the long short-term memory classifier to cloudy time series of multispectral images." Image and Signal Processing for Remote Sensing XXVI. Vol. 11533. International Society for Optics and Photonics, 2020.

available in the TS. For this reason (as mentioned in Chapter 5.1) many cloud restoration techniques have been proposed in the literature. By properly recovering missing information, almost all the images in the TS can be used. However, to obtain reliable cloud restoration results, sophisticated and usually computationally intensive techniques should be used [147] [148]. Significantly increases the computational complexity of the preprocessing part. Frequently, the final goal of the TS processing chain is the production of LC maps, and thus by performing restoration approach the classifier is able to correctly identify the LC class of each pixel. Therefore, a cloud restoration step is typically implemented to support the analysis of the multitemporal data.

Recent studies pointed out that multitemporal deep learning architectures, such as the LSTM classifier, are able to handle the presence of clouds. In the [29] authors proposed a convolutional recurrent layers for multitemporal LC classification of TSs of Sentinel-2 images [22]. By performing the visual interpretation of the internal gate activations (over a sequence of cloudy and cloud-free data), the authors verified that LSTM is able to discriminate clouds as a noise, without the need of any preprocessing step. However, the authors provided only a qualitative evaluation, and no quantitative analysis of LSTM robustness to cloud coverage was given. Sharma *et al* proposed a patch based LSTM for LC classification [28]. The authors first detected cloud/shadow pixels by using the Fmask algorithm. Then, they mitigated the cloud/shadow effects on the classification results by masking cloudy pixels with zero values. This choice allowed them to regulate the input gates of the LSTM, in order to dismiss the cloud information [28]. However, similarly to [22], no quantitative analysis has been carried out to assess the capability of the network to handle different amounts of cloud coverage. Moreover, none of the above mentioned work, investigate the role of the clouds position within the TS as the final results.

In this chapter, we focus the attention on the robustness of the LSTM classifier to cloud coverage since it is one of the most promising deep learning architectures for classification of dense TS of images [149]. This is investigated by quantitatively assessing its effectiveness in handling different amounts of cloud coverage under different TS lengths (in the framework of the LC mapping). In greater detail, we analyze the effect of the cloud coverage on the classification maps produced by LSTM considering: (i) different amounts of clouds in the TS, (ii) the use of zeros values to represent the detected clouds, and (iii) the use of restored images obtained by simple linear temporal gap filling (i.e., average of the spectral values acquired in the previous and following cloud-free images in the TS). It is worth noting that by representing cloud values as zeros, we investigate the possibility to regulate the input gates of the LSTM in order to dismiss the cloud information. Moreover, as the temporal profile plays a key role in TS classification, we study the role of cloud positions within the TS.

The rest of this chapter is organized as follows. Section 2 describes the theoretical bases of the LSTM classifier followed by the description of the proposed methods to investigate the robustness of LSTM to cloud coverage. Section 3 describes the dataset used, while in Section 4 the obtained results are presented. Finally, Section 5 draws the conclusion of the chapter.

6.2 Study on the Robustness of LSTM to Cloud Coverage

6.2.1 Long Short-Term Memory Network

RNN have been used with a great success for many applications dealing with sequential data, such as speech recognition, robot control and video tagging [150]. LSTM is a special type of RNN, which has been proposed to overcome the problem of the vanishing gradient on long data sequences [151]. In particular, by using a forget gate's vector, LSTM has a better control of the gradients values at each time step, and thus avoiding the gradient vanishing in a long TS. LSTM is able to catch the temporal dependency of sequential data, since the network is using feedback loops connected to their past decisions. Thus, the network has a state and makes decisions on the basis of the network output from the previous input (x_{t-1}), the current input (x_t) and the previous hidden state (h_{t-1}).

The scheme of the LSTM cell is presented in Fig. 6.1. The cell state is marked as a straight line on top with letter C. The data which flow into the LSTM are regulated by the input, forget and output gates. Let W_i, W_f, and W_o be the learnable weights and b_i, b_f, and b_o the biases for the input, forget and output gate, respectively. The three gates of the LSTM can be described as follows:

- **forget gate** (f) - It controls the information coming from the previous cell state C_{t-1} and the one provided by the current input (x_t). The output of this gate ranges from 0 (discard information) to 1 (keep the information). This gate is described by the following equation:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{6.1}$$

  where σ is the sigmoid function.

- **input gate** (i) - It updates the cell state. First, the sigmoid function decides which values from previous cell state C_{t-1} and current input (x_t) have to be updated. Then, the *tanh* function determines the rescaling factor. The computed value is added to the LSTM cell state. The input gate can be described by the following equations:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{6.2}$$

$$\widehat{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{6.3}$$

- **output gate** (o) - It decides the new hidden state. The previous hidden state and the current input are passed through the sigmoid function. Then, the output is multiplied by the current cell state passed through the *tanh* function to decide the information kept by the hidden state. The gate can be described by the following two equations:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{6.4}$$

$$h_t = o_t * tanh(C_t) \tag{6.5}$$

The weights associated with each gate are trained by a recurrent learning process.

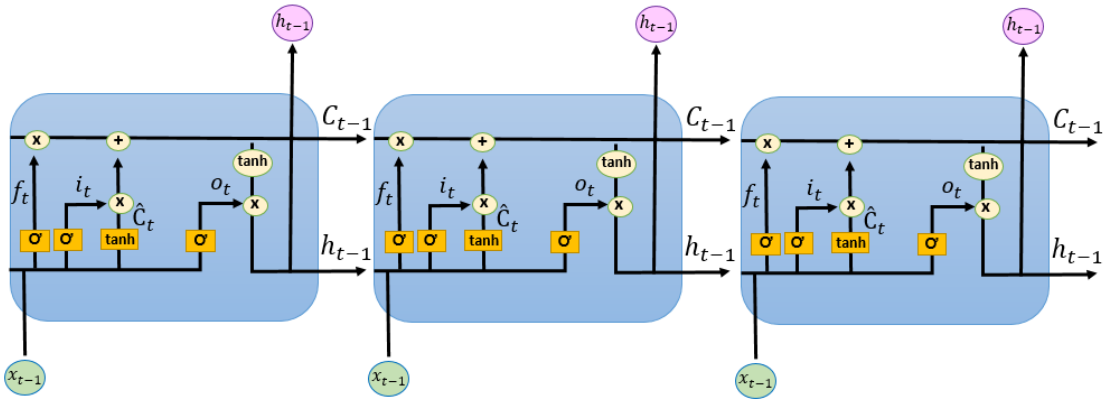


Figure 6.1: Illustration of the LSTM cell. The internal structure of the LSTM cell is represented, which contains the input gate, the forget gate, and the output gate.

In this work, we are considering a sequence-to-label classification task. The LSTM is trained on a TS of optical multispectral images to generate LC maps. In particular, the considered multi-layer LSTM network provides the classification result at pixel level. After testing different LSTM architectures, we decided to use LSTM network made up of two layers, each of 400 hidden units, a fully connected layer and a softmax function, which generates the probability distribution. Please note that the aim of the work is to evaluate the impact of the clouds in the classification results by comparing cloud-free and cloudy TS of optical images. Hence, more sophisticated models can be used to generate the LC map.

6.2.2 Assessment to Cloud Robustness

In this work, we aim to analyze the effects of the cloud coverage on the LSTM classifier accuracy. In greater details, we first investigate the role of the TS length on the LSTM classification accuracy. This is done by considering TSs made up of 10 and 30 Sentinel-2

images (see Fig. 2b). Then, we analyze the robustness of the classifier to cloud coverage by comparing the classification maps obtained using the cloud-free TS and the TS where cloudy images are introduced. We considered different rates of fully cloudy images within the TSs ranging from 10% to 40% of the TS' length. For each cloudy image, we analyzed different representations of the cloudy pixels: i) simulated clouds, ii) zeros values and iii) restored values (linear temporal gap filling). The details about each representation are given in the following. Please note, that when clouds are represented by zeros or restored values a cloud detection step is required as the cloudy pixels have to be detected. This increases the computational complexity of the preprocessing part and can be affected by the quality of the cloud detection accuracy. However, in our experiments we used simulated data, thus neglecting the effects of possible cloud detection errors.

6.2.2.1 Simulated Clouds

In the first scenario, we introduced an artificial cloud with the spectral values corresponding to the real cloudy pixels retrieved from the Sentinel-2 images. In this way, we are able to investigate the robustness of LSTM classifier to cloud cover when no clouds filtering is applied. Let $\{X^1, X^2, \ldots, X^S\}$ be the considered TS. The multitemporal pattern associated to the jth pixel of the considered TS can be defined as $[x_j^1, x_j^2, \ldots, x_j^S]$, where $x_j^1 = [x_{j,B_1}^1, x_{j,B_2}^1, \ldots, x_{j,B_B}^1]$ represents the B spectral values of the jth pixel in the first image of the TS. The cloudy image is randomly introduced in the TS. Thus, the missing data for the jth pixel for each band B of the randomly selected image of the TS $x_{j,B_{1,\ldots,B}}^R$ are as follows:

$$x_{j,B_{1,\ldots,B}}^R = [3621, 3654, 3541, 4354, 4996, 5118, 5201, 5348, 4635, 3745] \tag{6.6}$$

, where $[3621, 3654, 3541, 4354, 4996, 5118, 5201, 5348, 4635, 3745]$ represent an qualitative example of a cloudy pixel in corresponding 10 bands of the Seninel-2 image.

6.2.2.2 Zeros Values

In the second case, we investigated the LSTM classifier robustness to cloud cover when the artificially introduced cloudy pixels are band-wise equal to zero values. Thus, we are able to investigate the possibility to regulate the LSTM input gates in order to ignore the cloud information. The missing data for the jth pixel for each band B of the randomly selected image of the TS $x_{j,B_{1,\ldots,B}}^R$ are equal to zeros:

$$x_{j,B_{1,\ldots,B}}^R = [0, ..., 0] \tag{6.7}$$

6.2.2.3 Restored Values - Linear Temporal Gap Filling

In the last considered scenario, we considered a very simple and efficient cloud restoration method based on linear temporal gap filling. The missing data for the jth pixel of the image x^T_{j,B_1} at time T are restored as the band-wise average of the spectral values acquired in the previous x^{T-1}_{j,B_1} and the following x^{T+1}_{j,B_1} images in the TS as follows:

$$x^T_{j,B_1} = Avg(x^{T-1}_{j,B_1}, x^{T+1}_{j,B_1}) \tag{6.8}$$

$$x^T_{j,B_2} = Avg(x^{T-1}_{j,B_2}, x^{T+1}_{j,B_2}) \tag{6.9}$$

$$x^T_{j,B_B} = Avg(x^{T-1}_{j,B_B}, x^{T+1}_{j,B_B}) \tag{6.10}$$

If clouds are present in the first or last composite in TS, the second or the one before last composite are considered, respectively.

6.2.3 Assessment of Cloudy Image Position in Time-Series

In this step, we analyze the impact of cloudy image position in the TS on the obtained LC map since the temporal profile plays an important role in the LC classification task. Also in this case we considered the three clouds representations values. It is worth noting that in order to mitigate any possible effect of clouds and shadows already present in the TSs, we detected them and excluded from our analysis. We classify only the TS with simulated clouds, and thus we are able to control the experimental set-up. In greater detail, we inserted in the cloud-free TS the different representations of fully cloudy images.

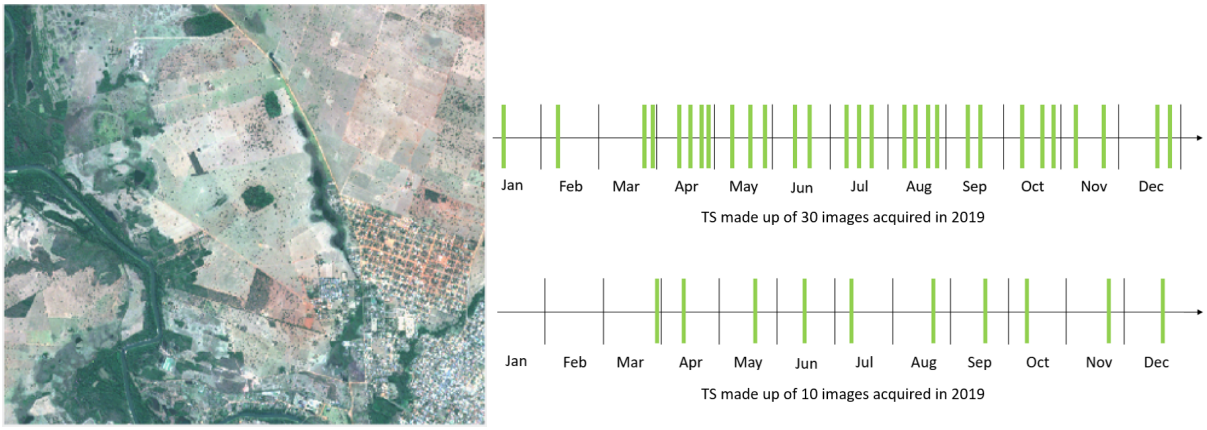
6.3 Data Set Description

In order to test the capability of the LSTM to handle cloud coverage, we classified a TS of Sentinel-2 images acquired over Aquidauana, Brasil in 2019. The area is located 140 km from the state capital, Campo Grande and it can be characterize by the hot, muggy and cloudy summers, while winters are warm and mostly clear. The Aquidauana belongs to Pantanal natural region, which is the world's largest tropical wetland and flooded grassland area. In particular, we considered the Sentinel-2 21KXT tile. Fig. 6.2a shows the Sentinel-2 image acquired on the 30^{th} March 2019 over the considered study area, while Fig. 6.2b presents the considered TS made up of 10 and 30 images. Both TSs approximately represent whole year. Although, due to quite heavy cloud cover in January and February the number of available images is smaller, and thus we have excluded this months from TS made up of 10 images. The input to the LSTM is a TS of Sentinel-2 images, which were first atmospherically corrected by using ESA's Sen2Cor tool [112].

The following Sentinel-2 spectral bands were considered as features in input of the LSTM classifier: blue (B2 - 490 nm), green (B3 - 560 nm), red (B4 - 665 nm), the four vegetation red edge (B5 - 705nm, B6 - 740 nm, B7 - 0.783 nm and B8A - 865 nm), NIR (B8 - 842 nm) and the two SWIR bands (B11 - 1610 nm and B12 - 2190 nm). A nearest neighbor interpolation technique was used to match the 20 m bands to the 10 m bands. The spectral outlier detection and removal was performed by discarding the reflectance values higher than the 0.999 quantile and lower than the 0.001 quantile of each spectral band. All the images considered in the experiments have cloud coverage smaller than 15%.

The considered classification task aims to detect 8 LC types, namely: "Evergreen broadleaves," "Deciduous broadleaves," "Shrub cover," "Cropland," "Herbaceous Vegetation Aquatic," "Bare areas," "Built-up areas" and "Open water permanent". In our analysis, we are considering a training set composed of 8105 samples selected by photo-interpretation. The results are evaluated as the agreement between the classification maps obtained by using the cloud-free TS of images (i.e., cloud-free map) with the maps obtained when simulated clouds are introduced in the TSs.

The LSTM model was trained with the following hyperparameters: (a) 5 epochs; (b) Adam optimizer with an initial learning rate set equal to 0.001 and the weight decay of 0.0001; and (c) mini-batch size of 128. The cross-entropy loss function calculated by comparing the predicted and the ground truth class probabilities [152] is used at each training step to adjust the model weights. The considered training set is composed of 8105 samples, while the maps agreement is evaluated on 312409 samples.



(a) (b)

Figure 6.2: The considered study area: (a) Sentinel-2 RGB image on 30.03.2019; (b) the considered TSs made up of 10 and 30 images.

6.4 Experimental Results

In this section, we present first the results obtained when assessing the robustness of the LSTM to cloud coverage by considering different rates of cloud images (10% - 40%) in TSs characterized by different lengths (10 and 30 images). The results are evaluated both from the quantitative and qualitatively view point. Then, we present the analysis performed to assess the impact of the cloudy image position in the TS on the classification results obtained.

6.4.1 Analysis of the Robustness of the LSTM to the Amount of Clouds in the TS

Tab. 6.1 and Tab. 6.2 present the % of the LC map agreement and standard deviation (σ) obtained by comparing the result achieved on the cloud-free TS and those by considering cloudy TS characterized by different cloud percentages (10 - 40%) on the TS made up of 10 images and 30 images, respectively. For each experiment different clouds representations have been considered: (i) simulated cloud, (ii) represented with zero values, and (iii) restored cloud by linear interpolation. Since the cloud position in the TS has an impact on the LSTM accuracy, the results reported in the table are the average of five independent experiments where the posterior of the clouds has been randomly determined.

From the obtained LC map agreement, one can conclude that the LSTM is able to generate a LC map very similar to the one obtained with a cloud-free TS when simple linear interpolation cloud restoration step is applied. In this case, the LC map agreement is very high for all the considered cases, e.g., 94% even when the 40% of images in the TS of 30 images are covered with clouds (i.e., 12 cloudy images out of 30) and standard deviation is relatively low (2.21). As expected, when restored clouds have been considered it can be noticed that the longer TS is, the LSTM obtains better and more stable results. Indeed, for the TS composed of 10 images with 30% of cloudy images (i.e., 3 cloudy images) the obtained agreement is 89.24% with $\sigma = 8.22$, whereas for the TS composed of 30 images considering the same amount of cloudy images (i.e., 9 cloudy images) the agreement is equal to 94.34% with $\sigma = 1.31$. One can also conclude that when clouds are represented with zeros values, on average the LSTM performs better than when the simulated clouds values are kept. For example, when cloudy images are the 10% of the TS, the agreement between the cloud-free and the cloudy maps is 77% and 44% when clouds are detected and represented with zeros and with their original values (no detection), respectively. This indicates that by assigning zero value to clouds, LSTM is able to mitigate the effect of the input cloud information.

This quantitative evaluation is confirmed by the qualitative analysis which can be

Table 6.1: LC map agreement ([%]) and standard deviation (σ) between the cloud-free classification map and the one obtained with the cloudy TS by considering 10%, 20%, 30% and 40% of cloudy images in the TS of images made up of 10 images. Reported results are the average of five independent experiments.

	LC Map Agreement							
% cloudy images	**10%**		**20%**		**30%**		**40%**	
Quality metrics	[%]	σ	[%]	σ	[%]	σ	[%]	σ
Simulated Clouds	47.20	15.62	20.22	19.62	16.76	9.53	8.77	7.63
Zero Value Clouds	70.66	3.69	48.50	10.20	21.79	0.38	12.19	0.48
Restored Clouds	95.31	7.66	89.47	6.24	89.24	8.22	86.41	2.82

Table 6.2: LC map agreement ([%]) and standard deviation (σ) between the cloud-free classification map and the one obtained with the cloudy TS by considering 10%, 20%, 30% and 40% of cloudy images in the TS of images made up of 30 images. Reported results are the average of five independent experiments.

	LC Map Agreement							
% cloudy images	**10%**		**20%**		**30%**		**40%**	
Quality metrics	[%]	σ	[%]	σ	[%]	σ	[%]	σ
Simulated Clouds	43.64	26.10	10.77	1.84	5.49	0.98	4.62	0.06
Zero Value Clouds	77.09	3.00	51.51	14.17	29.09	8.94	16.87	5.75
Restored Clouds	96.82	1.52	94.95	0.81	94.34	1.31	93.75	2.21

performed by checking the LC maps reported in Fig. 6.3. The reported analysis represents two regions: vegetated area (images from a to f) and build up area (images from g to l). The results presented are the ones obtained by inserting the 20% of cloud coverage in the TS of 30 images. Fig. 6.3a, g, Fig. 6.3b, h, Fig. 6.3d, j and Fig. 6.3e, k present: (i) the cloud-free map, (ii) the map obtained with restored clouds, (iii) the map obtained with simulated clouds and (iv) the map obtained with zero clouds, respectively. Moreover, for better comparison images (c)(i) corresponds to RGB Sentinel-2 image from and 30.03.2019, while (f)(l) represents corresponding Google Satellite HR image. Both the zeros and the simulated clouds LC maps are heavily affected by the presence of clouds. In particular, even if only 6 images out of 30 present cloud coverage, the obtained classification map for simulated clouds is strongly corrupted. Hence the high values of spectral reflectance for clouds lead to the classification of most of the image pixels as the "Built-up areas" class (see Fig. 6.3d and j). This is because "Built-up areas" class has a spectral signature very similar to the clouds. By masking the clouds as zeros, the results obtained are improved (see Fig. 6.3e and k). However, the phenological trend of the "Cropland" class is confused with the multitemporal trend of the "Deciduous broadleaves" vegetation. In contrast, the LC map obtained with the considered simple cloud restoration achieves very strong agreement with the original cloud-free map.

6.4.2 Analysis of the Robustness of LSTM Classifier to the Position of the Cloudy Images in the TS

Fig. 6.4 and Fig. 6.5 show the LC map agreement accuracy obtained between the cloud-free map and the one obtained when one fully cloudy image is introduced at different positions in TS of 10 and 30 images, respectively. Also in this experiment, we considered the three cloud representations. Similarly to the previous analysis, the results show that the network is very stable (for both TSs with 10 and 30 images) when the simple clouds restoration technique is applied. Moreover, also in this case, the use of zero values with respect to the high reflectance values of the simulated clouds strongly mitigates the effect of the cloud coverage in the TS. As expected, LSTM is more robust to the presence of cloud noise when longer TSs of images are considered. However, from the results obtained, one can observe that the network is more sensitive to the clouds occlusions when they are present at the beginning (see Fig. 6.4 for all the clouds representation) or at the end of the TS (see Fig. 6.4 and 6.5 for both the zeros and the simulated clouds). This can be attributed to the temporal profiles of classes (e.g., grassland vs bare soil or vegetation), which play a crucial role in the TS classification.

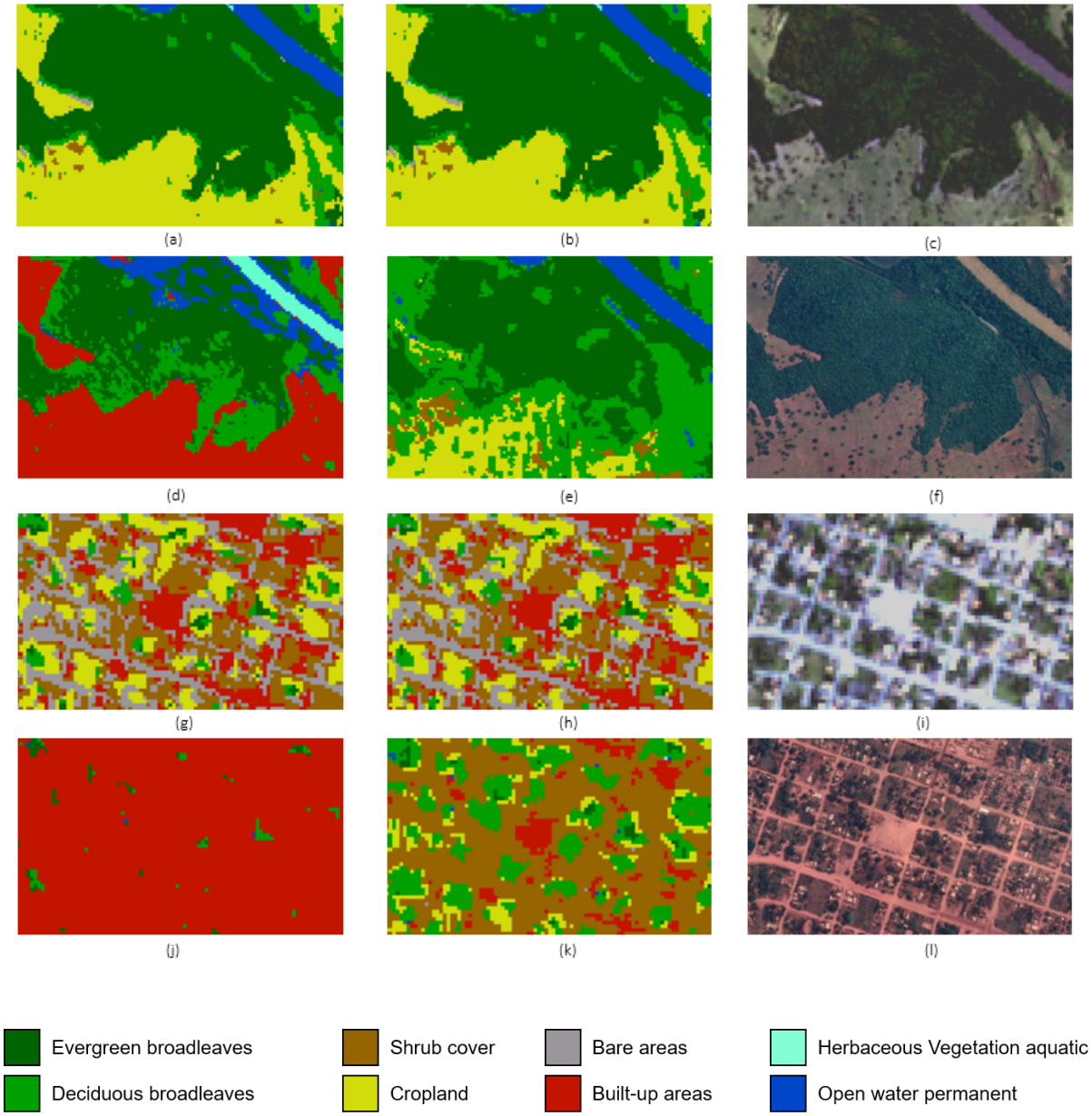


Figure 6.3: Classification maps obtained for two study areas (a-f corresponds to vegetated area and g-l build up area) considering the TSs made up of 30 images where the 20% of cloud coverage has been inserted. The classification maps have been generated by using the: (a)(g) cloud-free TS; (b)(h) restored clouds TS; (d)(j) TS with simulated clouds values; (e)(k) TS with clouds set as zeros. While, (c)(i) corresponds to RGB Sentinel-2 image from and 30.03.2019 and (f)(l) Google Satellite HR image.

Figure 6.4: LSTM accuracy for a three considered cases, when a fully cloudy image is included in the TSs. The x-axis indicates the position (date) of the cloudy image (from 1^{st} up to the 10^{th} position of the TS).

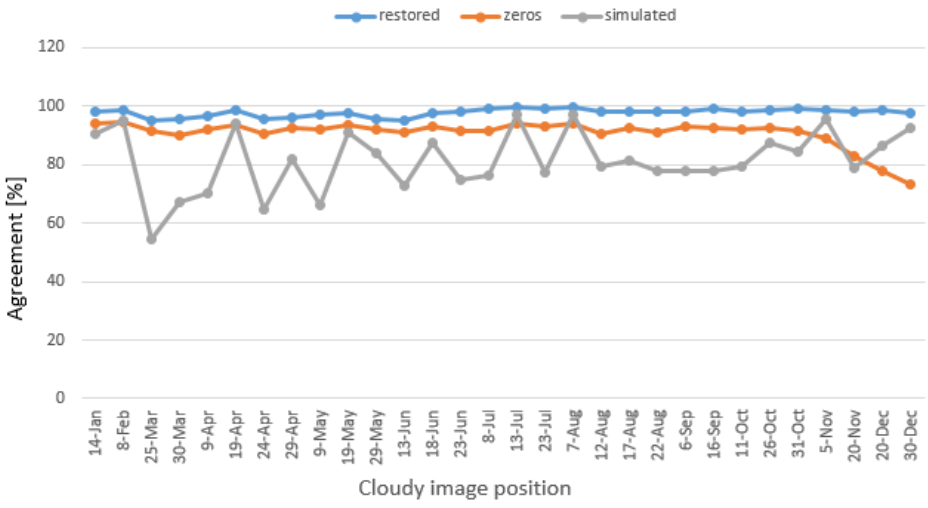


Figure 6.5: LSTM accuracy for a three considered cases, when a fully cloudy image is included in the TSs. The x-axis indicates the position (date) of the cloudy image (from 1^{st} up to the 30^{th} position of the TS).

6.5 Discussion and Conclusion

In this chapter a quantitative and qualitative evaluation of the impact of the cloud coverage on the classification accuracy obtained by the multitemporal deep learning LSTM method has been presented. In particular, the experimental analysis has been carried out to evaluate the capability of the LSTM to handle the cloud coverage considering different: (i) the length of the TS, (ii) the position of the cloudy images in the TS, and (iii) the cloud representation values. In the experiments, we considered TSs of 10, 20 and 30 images, by inserting a percentage of cloudy images ranging from 10% to 40% of the TS length.

From the results obtained one can conclude that when clouds are restored with very simple and fast linear temporal gap filling, the map agreement between the cloud-free and the cloudy map is very high and stable regardless of the percentage of cloud coverage in the considered range. Moreover, when clouds are represented as zero the classifier performs better than when the simulated clouds values are used. This indicates that by assigning zero value to clouds, LSTM is able to mitigate the effect of the input cloud information. However, one should also notice that when clouds are represented by zeros or restored values, first cloud detection algorithm should be applied. Hence, increasing the computational complexity of the preprocessing part. In contrast, by representing clouds with their original values no preprocessing step is needed, and thus computational cost of this solution is comparably low. Finally, the impact of the position of the cloudy image in the classification result depends on the TS length. As expected the longer the TS is, the smaller is the impact of the position in the TS.

The obtained results area extremely important to demonstrate that the computational effort of more complex deep learning models may be compensated by a simpler preprocessing phase compared to the one required by shallow machine learning models. Moreover, such multitemporal deep learning model does not require any hand-crafted feature selection and is able to accurately capture the multitemporal trend of the land cover classes. Nevertheless, it should be mentioned that one known limitation of the deep based architectures is the need of very large training sets, which can be challenging for many applications. As a future development, we plan to investigate the effectiveness of the LSTM with respect to standard machine learning models typically employed for LC mapping such as RF and CNN. Moreover, we would like to better analyze the impact of the cloud coverage on the different LC classes.

Chapter 7

An Approach Based on Low Resolution Maps and Domain Adaptation to Define Training Sets at Large Scale

In this chapter, we present an approach that aims to enlarge existing training sets leveraging global thematic products and domain adaptation techniques. The method assumes available a small amount of reference data collected in a portion of the study area (source domain), which is not sufficient to be representative of the whole region. Global scale thematic images can be used to mitigate this problem. However, these maps are in many areas at low resolution and typically unreliable. To address this issue, the proposed approach relies on a domain adaptation strategy to detect the labeled units in the map having the highest probability to be correctly associated to their labels according to the information provided by the initial training set. In particular, the method performs: i) a preliminary sample selection from the low resolution thematic product (in the target domain), ii) the identification of a mapping features space where source and a target samples are aligned, and iii) the definition of the final enlarged training set. Experimental results obtained using the CGLS-LC map demonstrate the effectiveness of the method. [1]

[1]Part of this chapter appears in I. Podsiadlo, C. Paris, L. Bruzzone, "An approach based on low resolution land-cover-maps and domain adaptation to define representative training sets at large scale." IGARSS 2021 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2021.

7.1 Introduction

As discussed in previous chapters LC mapping is of high importance for many applications in land management and climate change monitoring. Large scale LC classification is possible due to the large amount of RS data currently available. Nevertheless, frequently producing accurate and up-to-date LC maps is a challenging task from the operational view point. This is because the landscape is constantly changing, and thus it requires up to date labeled samples, which is very laborious and time-expensive. Moreover, the quantity and the quality of training samples play a crucial role on the classification accuracy. On the one hand, when dealing with large scale LC mapping, collecting labeled samples in the whole study area is unfeasible. On the other hand, when training samples used to train the classifier are related to regions spatially distant to the region to be classified, then the classifier may produce poor results. This is due to possible different statistical distributions of classes in spatially disjoint regions, a well-known problem from the operational view point, the so-called sample selection bias. Sample selection bias occurs when test and training data are coming from the same domain D. However, the estimated distribution $\widehat{P}(x,y) = \widehat{P}(x)\widehat{P}(y|x)$ does not correctly approximates the real distribution of the data $P(x,y)$. This problem usually occurs when the training set is not representative of the whole study area or the number of training samples is not sufficient. This results in the bias estimation of class prior distribution $\widehat{P}(y) \neq P(y)$ and the poor estimation of prior distribution $\widehat{P}(x) \neq P(x)$. A special case of sample selection bias is the covariate shift problem, which is referred to cases when estimated and the real distributions differ only in the prior distribution $\widehat{P}(x) \neq P(x)$. To mitigate this problem, several domain adaptation methods able to handle to the shifts in data distributions, have been proposed in the RS literature (see Chapter 5.2). These methods are able to use information from one domain (source domain) and apply it to the other domain (target domain) under the assumption that the two domains have different but still correlated distributions.

In this chapter, we propose a novel strategy to define a representative training set, which can be used to handle large scale classification problem. The method assumes the availability of a training set for at least one RS image of the study area (i.e., source domain). The main goal of the method is adding labeled units of RS data geographically distant from the available reference data (i.e., target domain). To solve the problem in an unsupervised way, we rely on publicly available thematic products, used to extract few map-labeled samples of the target domain. The small set of map-labeled samples is used together with the annotated samples to align the statistical distributions of the source and target domains in a common latent space. This mapping space allows us to assign reliable labels to a set of target samples. Thus, those target samples can be added to

the initial training set to better represent the whole study area. The proposed approach can be applied to any RS data and adopted without geographical constrains due to the current availability of many global LC maps.

The remaining of the chapter is organized as follows. Section 2 presents the proposed method to define representative training sets at large scale. Section 3 describes the dataset and Section 4 presents the obtained experimental results. Finally, in section 5 the conclusions of this chapter are drawn.

7.2 Proposed Method

The proposed method assumes the availability of a training set for at least one RS data of the study area (i.e., only for the source domain) and an existing LC product for the target domain (RS data of the study area where no labeled samples are available). Please note that the latter assumption is reasonable from the operational view point, since many LC maps are available at global scale (see Sec. 2.4). Fig. 7.1 shows the architecture of the proposed approach made up of three main parts: (1) target map samples selection, (2) obsolete low resolution LC-Map-based domain adaptation strategy, and (3) representative training set generation.

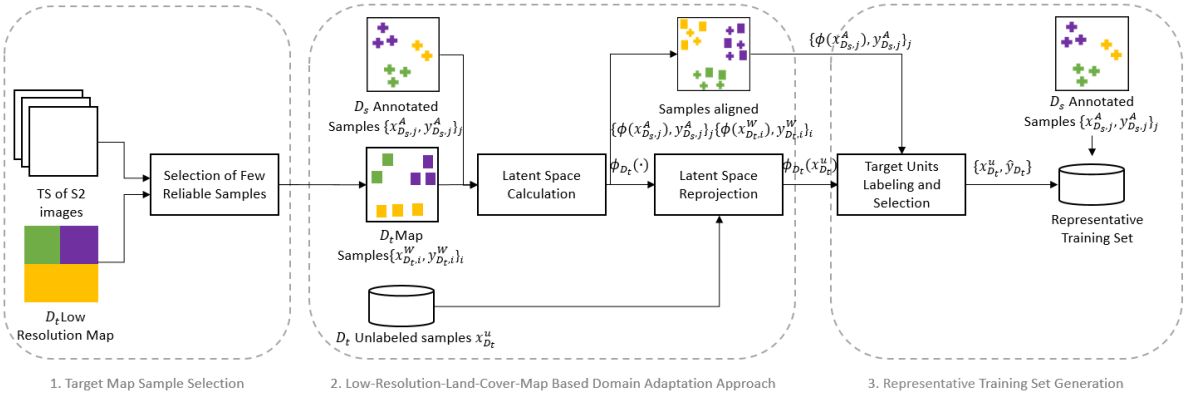


Figure 7.1: Architecture of the proposed LC-Map-Based domain adaptation strategy.

7.2.1 Target Map Samples Selection

Continental or global thematic products contain a large amount of information about the LC classes present in the scene. However, they can be obsolete and sometimes have low spatial resolution (e.g., 100 m). For these reasons, their usage is not straightforward. By directly including labeled samples extracted from the map in the initial training set, there

is a high possibility to add poor quality labeled units. To address this issue, the goal of this step is to select few map target units (having the highest probability to be correctly labeled) that can be used to learn the shift of the LC distributions between source and target domains [27].

First, the thematic product is re-scaled to the spatial resolution of the RS images to classify. Then, its legend is revised to select only the classes present in the source domain training set. This is done in accordance with [153]. Let us define the set of considered LC classes as $\Omega = \{\omega_c\}_c$. The thematic product is used to generate N statistically independent noisy training sets, which are then employed to train an ensemble of N classifiers. In order to ensure proper LC classes representation in each training set, the samples were selected according to the stratified random sampling strategy [95]. In this chapter we use ensemble of three neural networks but any other classifier could be used. To detect the most reliable target units we select the samples: (i) where the original thematic product agrees with the N noisy classification maps, (ii) having the highest confidence score of the classifier. At the end of this step, we collect few labeled target units of the target domain $\{\mathbf{x}^W_{D_t,i}, y^W_{D_t,i}\}_i$, where $\mathbf{x}^W_{D_t,i}$ represents the RS features (spectral bands) of the i-th sample of the target domain and $y^W_{D_t,i}$ the associated map label. This set of target domain selected samples will be used in the next step in order to align target and source domains.

7.2.2 LC-Map based Domain Adaptation strategy

The second step is the core of the proposed approach. Let $\{\mathbf{x}^A_{D_s,j}, y^A_{D_s,j}\}_j$ be the set of annotated samples of the source domain. Let $\mathbf{x}^u_{D_t}$ be the pool of unlabeled samples of the target domain. The distribution of the source domain can be written as:

$$p_{D_s}(\mathbf{x}) = \sum_{\omega_c \in \Omega} P_{D_s}(\omega_c) P_{D_s}(\mathbf{x}|\omega_c) \tag{7.1}$$

and the distribution of the target domain:

$$p_{D_t}(\mathbf{x}) = \sum_{\omega_c \in \Omega} P_{D_t}(\omega_c) P_{D_t}(\mathbf{x}|\omega_c) \tag{7.2}$$

where $P_{D_s}(\omega_c)$ is the prior probability of class ω_c and $P_{D_s}(\mathbf{x}|\omega_c)$ is its conditional probability. We assume that $P_{D_s}(\omega_c) \neq P_{D_t}(\omega_c)$. To learn the shift of the probability distributions of the source and target domains (i.e., $p_{D_s}(\mathbf{x}) \neq p_{D_t}(\mathbf{x})$), we use the few map-labeled target samples $\{\mathbf{x}^W_{D_t,i}, y^W_{D_t}\}_i$ and the source annotated samples $\{\mathbf{x}^A_{D_s,j}, y^A_{D_s,j}\}_j$. To this end, we considered KEMA proposed in [27] (see Sec. 7.2.4), which projects the source and target samples into a common latent space, where units belonging to the same LC classes are close to each other. The common projection matrix ϕ, which keeps the original geometry

of each manifold preserved, contains the domain specific projection function, i.e.,

$$\sum_{\omega_c \in \Omega} P_{D_s}(\phi_{D_s}(\mathbf{x})|\omega_c) \approx \sum_{\omega_c \in \Omega} P_{D_t}(\phi_{D_t}(\mathbf{x})|\omega_c) \tag{7.3}$$

In the latent space, we train a classifier by using only the projected source domain samples $\{\phi_{D_s}(\mathbf{x}^A_{D_s,j}), y^A_{D_s,j}\}_i$ to classify a set of unlabeled samples of the target domain projected onto the latent space $\mathbf{x}^u_{D_t}$. Therefore, the manifold alignment allows us to provide source samples that approximate the target distribution. Please note that by using manually annotated labeled samples sharply increases the probability to obtain reliable classification results.

The classification is performed using a simple linear discriminant analysis (LDA) since the data were subject to a non-linear projection to detect the latent space, where the classes can be linearly discriminated. The more reliable units, classified with higher confidence by LDA, are added to enlarge the initial training set, i.e., $\{\phi_{D_t}(\mathbf{x})_{D_t,z}, \hat{y}_{D_t,z}\}_z$.

7.2.3 Representative Training Set Generation

At the end of the proposed method, we merge the two training sets: the original photo-interpreted sampled from the source domain and the target domain samples added in the previous step. First, the added target samples $\{\phi_{D_t}(\mathbf{x})_{D_t,z}, \hat{y}_{D_t,z}\}_z$ are re-projected into the original feature space $\{\mathbf{x}_{D_t,z}, \hat{y}_{D_t,z}\}_z$. To validate the quality of the training set, we perform the pixel-based classification of the target domain image by using united training set (in the original feature space). In the proposed implementation, we used a SVM classifier, which has proven good generalization capability even if a small training set is available [154]. Moreover, the Gaussian kernel demonstrates low computational performance. The SVM parameters were tuned using the standard cross-validation procedure.

7.2.4 Kernel Manifold Alignment

Let us consider two domains D_s and D_t, where $D_s \in R^{\{\mathbf{x}_{D_s,j}\}_j \times dim_{D_s}}$ and $D_t \in R^{\{\mathbf{x}_{D_t,i}\}_i \times dim_{D_t}}$. The D_s contains set of annotated examples $\{\mathbf{x}^A_{D_s,j}, y^A_{D_s,j}\}_j$ and D_t contains set of samples with label extracted from the thematic product $\{\mathbf{x}^W_{D_t,i}, y^W_{D_t,i}\}_i$. The SSMA method aligns data from both domains by projecting them to a common latent space so that samples belonging to the same class are placed close to each other and samples of different classes are pushed away, while the geometry of the manifolds is preserved. This is done by using a domain specific projection function v_{D_s}. The projection functions are grouped in a projection matrix $V := [v_{D_s}, v_{D_t}]^T$. The SSMA finds a data projection matrix V that minimizes the cost function:

$$\Theta = \frac{\mu GEO + SIM}{DIS} \tag{7.4}$$

where GEO is a similarity matrix, which represents the topology of the considered domain W, e.g., it can be constructed using a k-nearest neighbor graph (k-NNG); SIM is a similarity matrix between $\{\mathbf{x}^A_{D_s,j}, y^A_{D_s,j}\}_j$ samples belonging to D_s and $\{\mathbf{x}^W_{D_t,i}, y^W_{D_t,i}\}_i$ samples belonging to D_t, if samples belong to the same class is equal to 1, or 0 otherwise; and DIS is a dissimilarity matrix of $\{\mathbf{x}^A_{D_s,j}, y^A_{D_s,j}\}_j$ samples belonging to D_s and $\{\mathbf{x}^W_{D_t,i}, y^W_{D_t,i}\}_i$ samples belonging to D_t, if samples belong to the same class is equal to 0 or 1 otherwise. This three matrices lead to three graph Laplacians: L, L_s and L_d, respectively. After solving equation 7.4 the problem comes down to finding eigenvalues of the generalized eigenproblem:

$$Z(L + \mu L_s)Z^T V = \lambda Z L_d Z^T V \tag{7.5}$$

,where μ is the weight parameter, Z is the block diagonal matrix containing the data to be aligned, V is the researched common projection matrix, organized in rows for the particular domain. After the projection V is found, each sample $\{\mathbf{x}^A_{D_s,j}, y^A_{D_s,j}\}$ from domain D_s can be projected into the common latent space by using $dim_{D_s} \times d$, where $d = dim_{D_s} + dim_{D_t}$ blocks of eigenvectors v_{D_s}:

$$P(x_{D_s,j}) = v^T_{D_s}\{\mathbf{x}^A_{D_s,j}, y^A_{D_s,j}\} \tag{7.6}$$

In order to better cope with the high dimensional data, the kernelized version of SSMA was proposed. Thus, first the data are projected to the higher dimensional Hilbert Space, where the linearly based classifier can be performed. Kernalization of SSMA is done by mapping data from D_s and D_t to a Hilbert feature space $\varphi : D \to H$. We replace all the samples with their mapped feature vectors $\varphi = (\{\varphi_{D_s}(\mathbf{x}^A_{D_s,j}), y^A_{D_s,j}\}_j, \{\varphi_{D_t}(\mathbf{x}^W_{D_t,i}), y^W_{D_t,i}\}_i)^T$. However, as this data live in high dimensional spaces, the computational complexity increases dramatically. Therefore, we use the Riesz representation theorem, according to which eigenvectors defined in Hilbert space $U = [u_{D_s}, u_{D_t}]^T$ can be expressed as a linear combination of mapped samples $U = \varphi\Lambda$. Than after replacing dot products with corresponding kernel matrices $K = \varphi^T\varphi$, we get:

$$K(L + \mu L_s)K\Lambda = \lambda K L_d K\Lambda \tag{7.7}$$

where K is the block diagonal matrix containing the Kernel matrices. Thus, the projection of sample $\{\mathbf{x}^A_{D_s,j}, y^A_{D_s,j}\}$ from domain D_s to the latent space, requires first mapping to the corresponding Hilbert space $\{\varphi_{D_s}(\mathbf{x}^A_{D_s,j}), y^A_{D_s,j}\}$ and then applying the corresponding projection vector u_{D_s}.

$$\{\phi_{D_s}(\mathbf{x}^A_{D_s,j}), y^A_{D_s,j}\} = u^T_{D_s}\{\varphi_{D_s}(\mathbf{x}^A_{D_s,j}), y^A_{D_s,j}\} \tag{7.8}$$

7.3 Dataset Description

7.3.1 Study Areas

The proposed approach was tested over a study area located in the Amazon. In particular, target domain was identified in the State of Mato Grosso do Sul, Brasil, while the source domain in the Presidente Hayes Dept, Paraguay. The distance between two domains is around 400 km. The area of the target and the source domain is equal around 10000 km^2. The corresponding geographical coordinates of the selected regions are following: 21KXT (19°53'4.26"S - 20°53'6.85"S; 54°59'0.02"W - 56° 2'41.27"W) and 21KUQ (22°36'14.88"S - 23°34'45.38"S; 57°53'19.61"W - 58°56'55.88"W). Both areas have similar environmental characteristics, they are mostly covered by grassland, aquatic vegetation and forest. However, it can be noticed that the target domain is more urbanized (within this area there are few towns, such as Aquidauana) than the source domain.

7.3.2 RS Data

The RS data employed to transfer the knowledge from the source to the target domain are optical atmospherically corrected Sentinel-2 data (i.e., L2A products). In particular, we considered the following Sentinel-2 spectral bands: blue (B2 - 490 nm), green (B3 - 560 nm), red (B4 - 665 nm), the four vegetation red edge (B5 - 705 nm, B6 - 740 nm, B7 - 0.783 nm and B8A - 865 nm), NIR (B8 - 842 nm) and the two SWIR bands (B11 - 1610 nm and B12 - 2190 nm). We considered two Sentinel-2 tiles, namely the 21KXT (target domain) and the 21KUQ (source domain) in 2019. For each domain, we generate a TS of 12 monthly composites retrieved from the Sentinel-2 images acquired in 2019. The Sentinel-2 20 m bands were re-sampled at 10 m spatial resolution using a nearest neighbor interpolation technique. The spectral outlier detection and removal was performed by discarding the reflectance values higher than the 0.999 quantile and lower than the 0.001 quantile of each spectral band. To generate the monthly composites all the images having cloud coverage smaller than 40% were considered. For each month, we computed the pixel-based median value of all the images acquired within the month. Clouds and shadow have been detected by using Sen2cor mask and discarded from the composite generation. For the generation of the target domain composites we used 53 Sentinel-2 images, while for the the source domain 41 Sentinel-2 images.

7.3.3 Thematic Product

In order to assess the effectiveness of the proposed method, we considered the CGLS-LC map, which aims to monitor the LC surface status and evolution at the global scale. The map was generated by considering the TS of multispectral PROBA-V 100 m images. The training set includes a dataset of more than 140 K crowd-sourced (Geo-WIKI) LC points and other ancillary data. The CGLS-LC map is an yearly based (2015 - 2019) moderate-resolution map having a classification accuracy of almost 80%. The map spatial resolution is equal to 100 m. In this chapter we have considered the CGLS-LC map from 2019.

The CGLC map provides discrete classification at three levels according to LCCS. In this chapter in accordance to photo-interpreted samples available the following LC classes were considered: "Evergreen broadleaf","Deciduous broadleaf", "Grassland", "Herbaceous Vegetation Aquatic", "Built-up areas" and "Open water permanent".

7.4 Experimental Results

To properly validate the proposed domain adaptation approach to define representative training sets at large scale, we performed the quantitative and the qualitative analysis reported in the following.

7.4.1 Experimental Set-up

To quantitatively evaluate the proposed approach, we collected manually annotated samples in both the source and the target domains for the considered study area. The manually annotated samples of the target domain were divided into training and test sets. In this way, we were able to compare the results obtained by the proposed method with the SVM RBF classifier trained on the reference target samples (SVM D_t), which provides an upperbound of the achievable accuracy. While the lowerbound of the achievable accuracy is provided by training the SVM model on the manually annotated samples from the source domain (SVM D_s). Moreover, in order to provide a good understanding of the added value of this work, we report the results obtained at each step of the proposed method. In particular, we present the classification results obtained by training the: SVM model on the target domain samples retrieved from the LC map (output of the step 1 of the proposed method), LDA classifier using the KEMA aligned (manually annotated) samples from the source domain (output of the step 2 of the proposed method), and finally the SVM using the enlarged training set made up of the manually annotated samples from the source domain and the unsupervised target samples labeled in the latent space (output of the proposed method).

The obtained results have been also compared to the GFK [135] method mentioned in the section 5.2. GFK is the domain adaptation method widely applied in the RS. The source domain samples were used for training the GFK to perform the classification of the target domain. Finally, to prove that the target domain samples extracted from the low-resolution thematic product add value, we have labeled them by using SVM trained on the D_s samples. Then, we used them together with the original D_s samples to align in the latent space (Step 2 CGLS).

All the classification results refer to pixel-based classification and were evaluated on the manually annotated samples of the target domain, in terms of OA and Fscore (F1). Moreover, for each class, we have reported the Bhattacharyya distance [155] between samples coming from the target and the source domain. Bhattacharyya distance measures the similarity of the probability distributions, and thus is a helpful metrics to understand how much shifted are the two considered domains probability distributions. We would like to remark that the proposed method is totally unsupervised and no annotated samples from the target domain are required. For the source domain, we considered a training set composed of 678 photo-interpreted samples. For the target domain, the proposed method added 614 samples. In the target map samples selection step, the N parameter (number of classifiers in the ensemble) have been set to 3.

7.4.2 Results

Fig. 7.2 and 7.3 present the classification maps obtained in the two areas of the target domain by considering different steps of the proposed method. It can be noticed that the map produced by the output of the first step of the proposed processing chain (see Fig. 7.2a, 7.3a) provides reasonable classification of grassland and evergreen broadleaf classes. However, it seems to misclassify the open water permanent class and overestimate the built-up and the herbaceous vegetation aquatic or regularly flooded classes. The classification map provided in the second step of the proposed approach (see Fig. 7.2b, 7.3b) properly classifies the open water permanent class and the herbaceous vegetation aquatic or regularly flooded, while it seems to overestimate the deciduous broadleaf class. Even though the proposed approach seems to slightly overestimate the grassland class the obtained classification map provides the most reliable results (see Fig. 7.2c, 7.3c). Although the map produced by the SVM trained on the manually annotated sampled from the source domain provides already reasonable result (see Fig. 7.2d, 7.3d) the map provided by the proposed approach provides better classification of build-up class as well as the deciduous broadleaf class (see Fig. 7.2c, 7.3c). Whereas, the classification map provided by the GFK overestimates the deciduous broadleaf class and significantly underestimates build-up and grassland classes (see Fig. 7.2e, 7.3e).

Table 7.1: Bhattacharyya distance (Bhat. Dist.) between samples coming from the target D_t and the source domain D_s and the classification results obtained on the manually annotated samples of the D_t by training the: (i) SVM D_t - SVM model on the manually annotated samples from the D_t (upperbound of the achievable accuracy), (ii) SVM D_s - SVM model on the manually annotated samples from the D_s, (iii) Step 1 - SVM model on the D_t samples retrieved from the LC map, (iv) Step 2 - LDA classifier using the KEMA aligned (manually annotated) samples from the D_s, (v) Step 2 (CGLS) - LDA classifier using the KEMA aligned (manually annotated) samples from the D_s (KEMA alignment done by using the D_s manually annotated samples and the D_t samples labeled by using SVM D_s), (vi) PM - SVM using the enlarged training set made up of the manually annotated samples from the D_s and the unsupervised D_t labeled in the latent space, and (vii) the GFK trained on the D_s samples. The F scores and the OA metrics per method are reported.

Class	**Bhat. Dist.**	**SVM D_t**	**SVM D_s**	**Step 1**	**Step 2**	**Step 2 (CGLS)**	**PM**	**GFK**
		F1 [%]	F1 [%]	F1 [%]	F1 [%]	F1 [%]	F1 [%]	F1 [%]
Everg. brod.	2.21	0.93	0.85	0.66	0.86	0.82	0.88	0.88
Decid. brod.	4.64	0.90	0.62	0.05	0.78	0.69	0.83	0.76
Grassland	2.41	0.94	0.77	0.84	0.76	0.71	0.83	0.80
Herb. Veg. A.	1.32	0.93	0.71	0.36	0.70	0.74	0.85	0.84
Build-up	6.98	0.95	0.80	0.78	0.84	0.81	0.83	0.63
Water pern.	2.38	0.98	0.92	0.64	0.89	0.93	0.94	0.93
OA		**0.94**	**0.80**	**0.58**	**0.82**	**0.80**	**0.87**	**0.83**

Tab. 7.1 shows the results obtained by training the: (i) SVM model on the manually annotated samples from the target domain (SVM D_t), (ii) SVM model on the manually annotated samples from the source domain (SVM D_s), (iii) SVM model on the samples selected from the LC map in the target domain (Step 1), (iv) LDA classifier using the KEMA aligned (manually annotated) samples from the source domain (Step 2), (v) LDA classifier using the KEMA aligned (manually annotated) samples from the source domain (KEMA alignment done by using the source domain manually annotated samples and the target domain samples labeled by using SVM trained on source domain) (Step 2 (CGLS)), (vi) SVM using the enlarged training set made up of the manually annotated samples from the source domain and the unsupervised target samples labeled in the latent space (PM), and (vii) by the GFK trained on source domain samples. All the results are provided on the test set of manually annotated samples of the target domain.

It can be observed that the OA obtained by the proposed method is relatively high 0.87 in comparison to the upperbound 0.94 and the lowerbound 0.80 produced by the manually labeled samples from the target and source domain, respectively. Although the OA provided by the CGLS samples is the lowest 0.58, it can be noticed that it represents the reliable source of information. This can be proved by looking at the Step 2-CGLS results, where the label provided by the CGLS samples was changed to the result of the SVM classifier trained on only source domain samples. Indeed, the OA accuracy obtained in the latent space in this scenario is lower 0.80 than the one obtained when the CGLS labels were considered 0.82. Moreover, one can notice that by projecting the data onto the latent space (Step 2) the classification accuracy improve with respect to the SVM model trained in the original domain (SVM D_s). For instance, for the deciduous broadleaf class, the LDA classifier (in the latent space) compared to the baseline SVM achieves an F score of 0.78 vs 0.62. However, the possibility of exploiting the enlarged training set in the original features space (after the re-projection) further increases the accuracy obtained for some classes (PM). By focusing the attention on the grassland, the SVM classifier achieves 0.83 accuracy vs 0.76 by LDA in the latent space. Overall, it can be noticed that the significant improvement of the classification accuracy is obtained for most of the classes. When comparing with other domain adaptation method, the proposed method achieves higher or equal F1 score for each considered class and the higher OA. Finally, one can notice that the proposed method improves the F1 score, even if the Bhattacharyya distance is relatively high. For instance, the build up class the Bhattacharyya distance is equal to 6.98, the F1 score is equal to 0.80 when classifier is trained with the source domain samples and with 0.83 when the proposed method is applied.

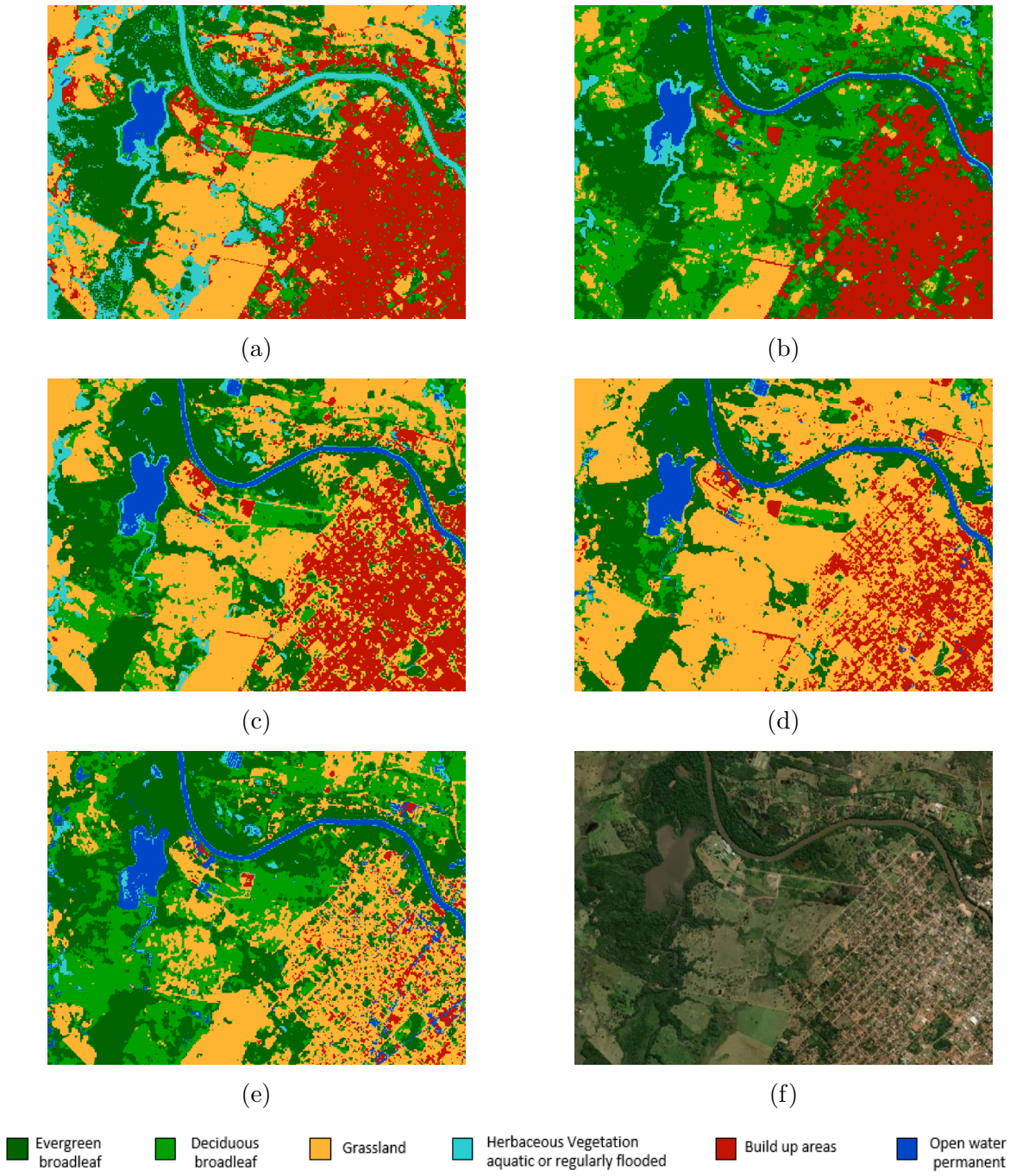


Figure 7.2: Classification results obtained in the target domain by training: (a) SVM on the output of the step 1, (b) LDA classifier on the output of the step 2, (c) proposed method, (d) SVM on reference samples from source domain, (e) the GFK, and (f) presents ESRI image of the corresponding area.

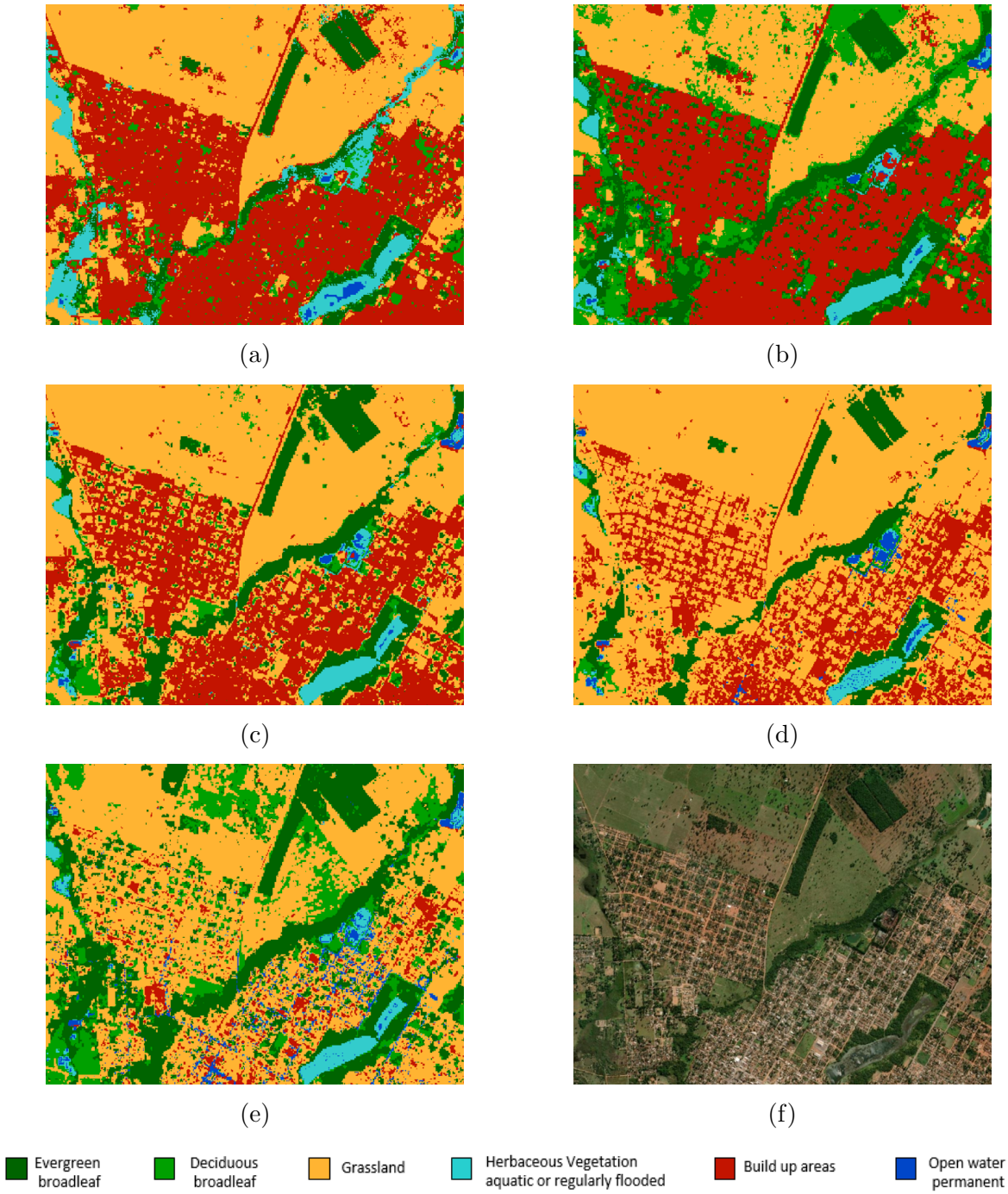


Figure 7.3: Classification results obtained in the target domain by training: (a) SVM on the output of the step 1, (b) LDA classifier on the output of the step 2, (c) proposed method, (d) SVM on reference samples from source domain, (e) the GFK, and (f) presents ESRI image of the corresponding area.

7.5 Conclusion

In this chapter, we presented an approach based on obsolete LC-maps and domain adaptation to define representative training sets at large scale. The method assumes availability of a training set for at least one RS image of the study area (the source domain), and the thematic product for the RS data where no labeled data are available (target domains). The thematic product is used to extract few weak labeled target samples that are used together with the annotated source samples to detect a common latent space, where the LC distributions are aligned. This allows us to to assign the labels to several target units in an unsupervised but reliable way. Then, this target domain units are re-projected into original feature space and are united with the source domain samples. Finally, to validate the quality of the training samples, we run the pixel-based classification using the united training set.

The proposed method have been applied to two Sentinel-2 tiles in the Amazon study area, namely: 21KXT and 21KUQ in 2019. The results obtained by the proposed method have been validated by qualitatively and quantitatively comparing its accuracy to each step of the proposed architecture. Moreover, we have analyzed the value of the information provided by the low resolution LC map. This was done by classifying the samples extracted from the low resolution LC map with the SVM trained on the source domain samples. The analysis carried out proves that the proposed method effectively aligns the two domains and take advantage from the small amount of manually labeled source domain samples as well as the low resolution LC map. Finally, the method outperforms the widely used GFK domain adaptation method (0.87 vs 0.83).

The proposed method is able to: i) perform well even if the small set of manually labeled sampled is available (for the source domain), and ii) add labeled units for RS data geographically distant from the available reference data. Moreover, we plan to test this strategy on different RS data, and with more complex classification schemes. Finally, we would like to test the proposed method on more challenging assumptions regarding the shift of domain distributions.

Chapter 8

Conclusion

This chapter draws the conclusions of the research activities presented in the thesis. It summaries and discusses the results presented in this dissertation and proposes future research directions of the presented work.

8.1 Novel Contributions

In this thesis, we have presented four novel contributions to the field of multitemporal multispectral automatic RS data analyses. In the following, the novel contribution of this dissertation are discussed considering advantages and disadvantages of the proposed methods.

In Chapter 3, we have addressed the estimation of the GMB at detailed spatial scale. The proposed contribution accurately improves the GMB estimation provided by a physically-based model. Differently from the methods in the literature, which are based on assimilation of RS data into physically-based simulations, the proposed methods integrates the complementary information afforded by the physically-based models, RS data and a few *in-situ* measurements. This allows us to overcome some of the weaknesses inherent in each of these data sources. In particular, physically-based simulations provide good generalization capability but simulations further away from a meteorological station are more uncertain. Moreover, the spatial resolution of such products is typically relatively low. On the other hand, *in-situ* measurements are very precise but also time-consuming and laborious. Whereas, the RS data provide a comprehensive spatial representation of glacier by their usage is hampered by cloud occlusions. The proposed approach was tested on glaciers in the Rofental Valley in Otzal Alps, Austria and Ortles-Cevedale group in South Tyrol, Italy. Sentinel-2 data were employed for analysis in 2015, 2016 and 2017, while Landsat 5 and 7 were used in 2006, 2007, 2010, 2011, 2012 and 2013. Both qualitative and quantitative results prove the effectiveness of integrating complementary data

sources, obtaining an average root-mean-square error of 460 mm compared to 732 and 661 mm obtained by the physically-based model and the standard regression models, typically used for parameters estimation. Moreover, due to spatial details provided by the RS imagery (10 and 30 m) the estimated GMB maps have a high spatial resolution with respect to the maps produced by the considered physically-based model (50 m). The proposed method is data-driven and can be applied to different glaciers by tuning the parameter k in accordance with the number of *in-situ* measurements available. A limitation of the proposed method is the requirement of *in-situ* measurements.

In Chapter 4, we have addressed the problem of multitemporal high spatial resolution LC monitoring to support climate change analysis. In contrast to the LC maps currently available, the proposed processing chains aim at performing a detailed LC mapping considering a classification scheme tailored to the specific needs of large scale climate change analysis. The proposed system architecture is general and is used for generating specific products: the 2019 HR land cover maps at 10 m spatial resolution at subcontinental scale to address the needs of climate modeling community and a long-term record (1990 - 2015) of regional HRLC (30 m) maps to support climate mitigation and adaptation analysis. The produced LC maps enable climate community to understand the role of spatial and temporal resolution on the LC maps consistency. The proposed system processes TS of high spatial resolution Sentinel-2 in 2019 and Landsat 5, 7 and 8 images in 1990, 1995, 2000, 2005, 2010 and 2015. The work presented in this chapter is part of the CCI HRLC project, which is the continuation of the CCI middle resolution project. We have qualitatively compared both products. The analysis carried out over few areas, showed that the the HRLC maps provide not only a higher spatial detail but are also more accurate than the CCI middle resolution. Although, the proposed static processing chain is designed to work at large scale, at the time of writing this thesis the quantitative analysis of both processing chains have been performed at regional scale (over 283, 265 and 204 Sentinel-2 tiles in the African Sahel, the Amazon, and the Siberia, respectively.). The obtained OA in accordance with the considered study areas is equal to 74.03%, 78.57%, 70.64% for the static map in the Amazon, the African Sahel and the Siberia, respectively. While, for the historic LC maps, the OA varies for all the considered years between 61.11% - 78.57% for the Amazon, 46.03% - 62.98% in the African Sahel and 57.93% - 65.07% in the Siberia. As expected the OA of the historic products is lower than that of the static ones, especially for the Siberia and the African Sahel areas. This can be associated to the small number of images available in these areas (especially in the Siberia) and the insufficient number of training samples (in the African Sahel). This is one of the main motivations for the development of the work of Chapter 7. Both processing chains have modular structure, and can be easily modified and adapted. Moreover, they are able to deal with the large

volumes of RS images coming from different sensors and characterized by different spatial, spectral and temporal resolution in a computationally efficient manner.

In Chapter 6, we have addressed the cloud coverage issue in TS of optical images, which is an extremely important problem from the operational point of view. Our work highlights the possibility of exploiting the capability of the deep learning based algorithms to provide accurate multitemporal classification results having lighter pre-processing steps compared to standard classifiers. We quantitatively studied the robustness of the multitemporal LSTM deep learning architecture under different cloud cover conditions. In particular, we analyzed the agreement between the classification maps obtained by using an LSTM classifier trained on a cloud-free TS and a TS with artificially introduced clouds. The analysis was carried out using Sentinel-2 imagery in the Amazon, which is known for its heavy cloud cover. The obtained results show that the LSTM is able to preserve very high agreement of maps when the clouds are detected using Sen2cor tool and a fast linear temporal gap filling is applied. The obtained results show that the classifier preserves an agreement of 93.75% even when 40% of a TSs made up of 30 images is cloudy. Moreover, when cloud are represented as zeros the maps agreement is better than when the clouds value are not removed. This because assigning zero values to the cloudy pixels remove this information from the input gates of the LSTM. As expected the longer is the considered TS the better is the classification result obtained and LSTM is more robust to the presence of cloud cover. For instance, when 40% of images in the TS are restored using linear temporal gap filling, the agreement of maps is equal to 86.41% and 93.75% for TS of 10 and 30 images, respectively. Finally, the obtained results show that the cloud position within TS plays a significant role for the classifier accuracy, especially for short TSs. This can be explained by the temporal variability of spectral signatures, which is crucial for the proper class identification.

In Chapter 7, we have address the issue of the definition of a reliable training set when working on large study areas. The proposed method highlights the importance of exploiting the available thematic products, which represent a valid source of information at global scale. Qualitative and quantitative analyses have been carried out over two Sentinel-2 tiles acquired in the Amazon in 2019. For the source domain a small set of photo-interpreted samples have been considered (20 samples per class), while for the target domain samples have been extracted from the CGLC 2019 map, available at 100 meters spatial resolution. The obtained results show the effectiveness of the proposed approach, which obtained OA of 87% compared to 80% obtained with the source domain photo-interpreted training set. Whereas the accuracy of the original target domain training set extracted from the CGLS map was equal to 58%. We would like to mention that given the current availability of the LC products, the proposed method can be applied to other

RS data without geographical constraint. Although the proposed method is promising further analysis should be carried out in order to understand when the classes dissimilarity makes the method fail.

8.2 Future Developments

In the following, we present some future developments to enhance the presented works and expand its use to more applications.

Regarding the estimation of the GMB spatial maps, three main areas should be further investigated. First, we would like to test the transferability of corrections across years, which from an operational viewpoint is an extremely interesting open research question. To answer it, it might prove helpful to analyze the correlation between the: (i) physically based simulations across years, and (ii) environmental conditions of the study areas (e.g., possible changes on the ground). If the hydro-climatological processes strongly vary from year to year, the deviations estimated by the proposed method cannot be simply inherited, but require an automatic adaptation step that should be properly defined. Second, we plan to investigate the possibility of extending the correction to other glaciers where no *in-situ* data are available. This can be assessed for glaciers that are close each other and have similar properties. Third, we would like to analyze a possible integration of SAR data (e.g., Sentinel-1) for feature-space definition. This is expected to be particularly helpful to better characterize different snow and ice types as SAR data provides information about snow and ice roughness, water content, grain size, and impurities [44].

Regarding the multitemporal large scale LC mapping, there are several future developments to be analyzed. An important research direction is the use of a multitemporal strategy to enhance the temporal consistency of the produced historic LC maps. This is because from the climate change perspective the maps consistency plays a crucial role to monitor and understand the impact of the LC changes on the climate. In this context, the use of cascade classifiers can obtain more temporally consistent classification maps. On the other hand, taking into consideration still not satisfying the OA obtained, one possible option would be to further enlarge the training set. This can be addressed or by reinforcement of the photo-interpretation activities or by applying domain adaptation strategies. To this end, we would like to test the proposed approach that leverages low resolution LC maps and domain adaptation methods to define a representative training set at large scale. So far, the proposed domain adaptation technique has been only tested under the covariate shift assumption. However, looking at the low OA obtained in the African Sahel, one may expect more challenging classification assumptions regarding training set distribution shifts, such as sample selection bias or other problems related to

the training set definition such as missing representation of some subclasses e.g., different kind of bare areas (deserts, volcanoes). Moreover, within this thesis, the domain adaptation method was tested with quite simple classification scheme (only six LC classes have been considered). Therefore, we would like to test its accuracy when applied to more complex classification scenarios, such as the one considered for the large scale HR LC mapping. In the context of the image classification, we plan to further analyze the LSTM classifier robustness to cloudy images. In particular, we aim to study the temporal variability of the spectral signature of different classes to better understand the land covers, which recognition is more impacted by cloud occlusions.

Chapter 9

List of Publications

INTERNATIONAL JOURNALS

[J1] **I. Podsiadlo**, C. Paris, M. Callegari, C. Marin, D. Günther, U. Strasser, C. Notarnicola, L. Bruzzone, "Integrating Models and Remote Sensing Data for Distributed Glacier Mass Balance Estimation," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 6177-6194, 2020, doi: 10.1109/JSTARS.2020.3028653.

[J2] **I. Podsiadlo**, C. Paris, L. Bruzzone, "On the robustness of the long short-term memory classifier to cloudy time series of multispectral images," in IEEE Geoscience and Remote Sensing Letters, in Preparation.

[J2] **I. Podsiadlo**, C. Paris, L. Bruzzone, "A Novel Approach Based on Low Resolution Maps and Domain Adaptation to Define Representative Training Sets at Large Scale" IEEE Transactions on Geoscience and Remote Sensing,, in Preparation.

[J2] **I. Podsiadlo**, C. Paris, L. Bruzzone, "Sub-Continental High Resolution Multi-Temporal Land Cover Maps Production (1990-2019) for Climate Change Analysis" in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, in Preparation.

INTERNATIONAL CONFERENCES

[C1] **I. Podsiadlo**, C. Paris, F. Bovolo, M. Callegari, L. Gregorio, D. Günther, C. Marin, T. Marke, M. Niroumand-Jadidi, C. Notarnicola, U. Strasser, M. Zebisch, L. Bruzzone, "Integration of hydro-climatological model and remote sensing for glacier mass

balance estimation." Image and Signal Processing for Remote Sensing XXV. Vol. 11155. International Society for Optics and Photonics, 2019.

[C2] **I. Podsiadlo**, C. Paris, L. Bruzzone, "A study of the robustness of the long short-term memory classifier to cloudy time series of multispectral images." Image and Signal Processing for Remote Sensing XXVI. Vol. 11533. International Society for Optics and Photonics, 2020.

[C3] **I. Podsiadlo**, C. Paris, L. Bruzzone, "An approach based on low resolution land-cover-maps and domain adaptation to define representative training sets at large scale", 2021 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Brussels, Belgium, 2021, in press.

DOCTORAL PROGRAM IN
INFORMATION AND COMMUNICATION TECHNOLOGY

Doctoral candidate

Iwona Katarzyna Podsiadlo

Cycle	**33**
Thesis	Methods for the analysis of time series of multispectral remote sensing images and application to climate change variable estimations
Advisor	**Prof. Lorenzo Bruzzone (** University of Trento **)**
Co-advisor	**Dr. Claudia Paris (** University of Trento **)**

1. List of publications

Published:

Podsiadlo, I., Paris, C., Callegari, M., Marin, C., Günther, D., Strasser, U., Notarnicola, C. and Bruzzone, L., 2020. Integrating Models and Remote Sensing Data for Distributed Glacier Mass Balance Estimation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, pp.6177-6194.

Podsiadlo, I., Paris, C. and Bruzzone, L., 2020, September. A study of the robustness of the long short-term memory classifier to cloudy time series of multispectral images. In *Image and Signal Processing for Remote Sensing XXVI* (Vol. 11533, p. 115331O). International Society for Optics and Photonics.

Podsiadlo, I., Paris, C., Bovolo, F., Callegari, M., De Gregorio, L., Günther, D., Marin, C., Marke, T., Niroumand-Jadidi, M., Notarnicola, C. and Strasser, U., 2019, October. Integration of hydro-climatological model and remote sensing for glacier mass balance estimation. In *Image and Signal Processing for Remote Sensing XXV* (Vol. 11155, p. 1115513). International Society for Optics and Photonics.

In press:

Podsiadlo, I., Paris, C. and Bruzzone, L., 2021. An approach based on low resolution land-cover-maps and domain adaptation to define representative training sets at large scale", 2021 *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Brussels, Belgium, in press.

In preparation:

Podsiadlo, I., Paris, C. and Bruzzone, L., On the robustness of the long short-term memory classifier to cloudy time series of multispectral images," in *IEEE Geoscience and Remote Sensing Letters*, in Preparation.

Podsiadlo, I., Paris, C. and Bruzzone, L. A Novel Approach Based on Low Resolution Maps and Domain Adaptation to Define Representative Training Sets at Large Scale" *IEEE Transactions on Geoscience and Remote Sensing*, in Preparation.

Podsiadlo, I., Paris, C. and Bruzzone, L. Sub-Continental High Resolution Multi-Temporal Land Cover Maps Production (1990-2019) for Climate Change Analysis" in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, in Preparation.

2. Research/study activities

Research projects:

European Space Agency, Climate Change Initiative (CCI), The High Resolution Land Cover

Euregio Scientific Research Found, CRYOMON- SciPro

Teaching activities:

Bachelor thesis co-supervision = 1 student

Bibliography

[1] USGS EROS Archive - Sentinel-2 - Comparison of Sentinel-2 and Landsat. [Online]. Available:https://web.archive.org/web/20170515022022/https://landsat.gsfc.nasa.gov/wp-content/uploads/2015/06/Landsat.v.Sentinel-2.png . [Accessed: 25- April- 2021].

[2] D. Smith, A. Scaife, E. Hawkins, R. Bilbao, G. Boer, M. Caian, L.-P. Caron, G. Danabasoglu, T. Delworth, F. J. Doblas-Reyes, *et al.*, "Predicted chance that global warming will temporarily exceed 1.5 c," *Geophysical Research Letters*, vol. 45, no. 21, pp. 11–895, 2018.

[3] T. R. Anderson, E. Hawkins, and P. D. Jones, "Co2, the greenhouse effect and global warming: from the pioneering work of arrhenius and callendar to today's earth system models," *Endeavour*, vol. 40, no. 3, pp. 178–187, 2016.

[4] K. Von Schuckmann, L. Cheng, M. D. Palmer, J. Hansen, C. Tassone, V. Aich, S. Adusumilli, H. Beltrami, T. Boyer, F. J. Cuesta-Valero, *et al.*, "Heat stored in the earth system: where does the energy go?," *Earth System Science Data*, vol. 12, no. 3, pp. 2013–2041, 2020.

[5] J. Y. et al., "The role of satellite remote sensing in climate change studies," *Nature Climate Change*, vol. 3, p. 875–883, 2013.

[6] E. P. McClain, W. G. Pichel, and C. C. Walton, "Comparative performance of avhrr-based multichannel sea surface temperatures," *Journal of Geophysical Research: Oceans*, vol. 90, no. C6, pp. 11587–11601, 1985.

[7] M. N. H. A. K. I. A. O. B. R. B. S. B. M. B. C. D. P. H. S. H. M. K. G. L. C. L. D. O. C. P. M. Poulter, B. and P. Peylin, "Plant functional type classification for earth system models: results from the european space agency's land cover climate change initiative," *Geosci. Model Dev.*, vol. 8, p. 2315–2328, 2015.

[8] S. D. McCracken, E. S. Brondizio, D. Nelson, E. F. Moran, A. D. Siqueira, and C. Rodriguez-Pedraza, "Remote sensing and gis at farm property level: Demogra-

phy and deforestation in the brazilian amazon," *Photogrammetric engineering and remote sensing*, vol. 65, pp. 1311–1320, 1999.

[9] A. Rabatel, J.-P. Dedieu, and C. Vincent, "Using remote-sensing data to determine equilibrium-line altitude and mass-balance time series: validation on three french glaciers, 1994–2002," *Journal of Glaciology*, vol. 51, no. 175, pp. 539–546, 2005.

[10] A. F. Rahman, D. Dragoni, and B. El-Masri, "Response of the sundarbans coastline to sea level rise and decreased sediment flow: A remote sensing assessment," *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3121–3128, 2011.

[11] 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," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4764–4774, 2016.

[12] K. L. Findell, A. Berg, P. Gentine, J. P. Krasting, B. R. Lintner, S. Malyshev, J. A. Santanello, and E. Shevliakova, "The impact of anthropogenic land use and land cover change on regional climate extremes," *Nature communications*, vol. 8, no. 1, pp. 1–10, 2017.

[13] W. B. Rossow and R. A. Schiffer, "Advances in understanding clouds from isccp," *Bulletin of the American Meteorological Society*, vol. 80, no. 11, pp. 2261–2287, 1999.

[14] M. Callegari, C. Marin, D. Günther, P. Rastner, L. Bruzzone, B. Demir, T. Marke, U. Strasser, M. Zebisch, and C. Notarnicola, "Integration of remote sensing with a hydroclimatological model for an improved monitoring of alpine glaciers," in *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 6239–6242, IEEE, 2018.

[15] G. Thirel, P. Salamon, P. Burek, and M. Kalas, "Assimilation of modis snow cover area data in a distributed hydrological model using the particle filter," *Remote Sensing*, vol. 5, no. 11, pp. 5825–5850, 2013.

[16] D. Finger, F. Pellicciotti, M. Konz, S. Rimkus, and P. Burlando, "The value of glacier mass balance, satellite snow cover images, and hourly discharge for improving the performance of a physically based distributed hydrological model," *Water Resources Research*, vol. 47, no. 7, 2011.

[17] e. a. Avanzi, Francesco, "S3m 5.1: a distributed cryospheric model with dry and wet snow, data assimilation, glacier mass balance, and debris-driven melt," *Geoscientific Model Development Discussions*, pp. 1–50, 2021.

[18] S. Morin and D. Ghislain, "Prosnow-provision of a prediction system allowing for management and optimization of snow in alpine ski resorts," *EGU General Assembly Conference Abstracts*, p. 12933, 2017.

[19] M. Hansen, R. DeFries, J. R. Townshend, and R. Sohlberg, "Global land cover classification at 1 km spatial resolution using a classification tree approach," *International journal of remote sensing*, vol. 21, no. 6-7, pp. 1331–1364, 2000.

[20] T. R. Loveland, B. C. Reed, J. F. Brown, D. O. Ohlen, Z. Zhu, L. Yang, and J. W. Merchant, "Development of a global land cover characteristics database and igbp discover from 1 km avhrr data," *International Journal of Remote Sensing*, vol. 21, no. 6-7, pp. 1303–1330, 2000.

[21] D. López-Puigdollers, G. Mateo-García, and L. Gómez-Chova, "Benchmarking deep learning models for cloud detection in landsat-8 and sentinel-2 images," *Remote Sensing*, vol. 13, no. 5, p. 992, 2021.

[22] M. Rußwurm and M. Körner, "Multi-temporal land cover classification with sequential recurrent encoders," *ISPRS International Journal of Geo-Information*, vol. 7, no. 4, p. 129, 2018.

[23] I. D. et al., "A challenge to parse the earth through satellite images," p. 17172–17209, Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern RecognitionWorkshops (CVPRW), 2018.

[24] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," Conference on Advances in Geographic Information Systems, 2010.

[25] W. Liu and R. Qin, "A multikernel domain adaptation method for unsupervised transfer learning on cross-source and cross-region remote sensing data classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 6, pp. 4279–4289, 2020.

[26] S. Hanneke and S. Kpotufe, "On the value of target data in transfer learning," *Advances in Neural Information Processing Systems*, vol. 32, pp. 9871–9881, 2019.

[27] D. Tuia and G. Camps-Valls, "Kernel manifold alignment for domain adaptation," *PloS one*, vol. 11, no. 2, p. e0148655, 2016.

[28] X. Y. A. Sharma, X. Liu, "Land cover classification from multi-temporal, multi-spectral remotely sensed imagery using patch-based recurrent neural networks," *Neural Networks*, vol. 105, pp. 346–355, 2018.

[29] K. M. Rußwurm M., "Multi-temporal land cover classification with sequential recurrent encoders," *ISPRS International Journal of Geo-Information*, vol. 7, 2018.

[30] R. Young and L. Norby, *Geological monitoring*. Geological Society of America, 2009.

[31] P. G. Knight, *Glacier science and environmental change*. John Wiley & Sons, 2008.

[32] W. M. A. Gómez C., White J. C., "Optical remotely sensed time series data for land cover classification: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 116, pp. 55–72, 2016.

[33] J. N. et al., "A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring," *Remote Sensing*, vol. 8, no. 1, pp. 2072–4292, 2016.

[34] A. Rabatel, J.-P. Dedieu, E. Thibert, A. Letréguilly, and C. Vincent, "25 years (1981–2005) of equilibrium-line altitude and mass-balance reconstruction on glacier blanc, french alps, using remote-sensing methods and meteorological data," *Journal of glaciology*, vol. 54, no. 185, pp. 307–314, 2008.

[35] L. Davaze, A. Rabatel, A. Dufour, R. Hugonnet, and Y. Arnaud, "Region-wide annual glacier surface mass balance for the european alps from 2000 to 2016," *Observational Assessments of Glacier Mass Changes at Regional and Global Level*, 2020.

[36] L. Davaze, A. Rabatel, Y. Arnaud, P. Sirguey, D. Six, A. Letreguilly, and M. Dumont, "Monitoring glacier albedo as a proxy to derive summer and annual surface mass balances from optical remote-sensing data," *The Cryosphere*, vol. 12, no. 1, pp. 271–286, 2018.

[37] A. Rabatel, P. Sirguey, V. Drolon, P. Maisongrande, Y. Arnaud, E. Berthier, L. Davaze, J.-P. Dedieu, and M. Dumont, "Annual and seasonal glacier-wide surface mass balance quantified from changes in glacier surface state: A review on existing methods using optical satellite imagery," *Remote Sensing*, vol. 9, no. 5, p. 507, 2017.

[38] M. S. D. R. De Wildt, J. Oerlemans, and H. Björnsson, "A method for monitoring glacier mass balance using satellite albedo measurements: application to vatnajökull, iceland," *Journal of Glaciology*, vol. 48, no. 161, pp. 267–278, 2002.

[39] W. Greuell and J. Oerlemans, "Assessment of the surface mass balance along the k-transect (greenland ice sheet) from satellite-derived albedos," *Annals of Glaciology*, vol. 42, pp. 107–117, 2005.

[40] A. Chaponniere, P. Maisongrande, B. Duchemin, L. Hanich, G. Boulet, R. Escadafal, and S. Elouaddat, "A combined high and low spatial resolution approach for mapping snow covered areas in the atlas mountains," *International Journal of Remote Sensing*, vol. 26, no. 13, pp. 2755–2777, 2005.

[41] V. Drolon, P. Maisongrande, E. Berthier, E. Swinnen, and M. Huss, "Monitoring of seasonal glacier mass balance over the european alps using low-resolution optical satellite images," *Journal of Glaciology*, vol. 62, no. 235, pp. 912–927, 2016.

[42] A. K. Tak Swati, Keshari, "Investigating mass balance of parvati glacier in himalaya using satellite imagery based model," *Scientific Reports*, vol. 10, 2020.

[43] J. Parajka and G. Blöschl, "The value of modis snow cover data in validating and calibrating conceptual hydrologic models," *Journal of hydrology*, vol. 358, no. 3-4, pp. 240–258, 2008.

[44] J. Shi and J. Dozier, "Inferring snow wetness using c-band data from sir-c's polarimetric synthetic aperture radar," *IEEE transactions on geoscience and remote sensing*, vol. 33, no. 4, pp. 905–914, 1995.

[45] J. L. Bamber and A. Rivera, "A review of remote sensing methods for glacier mass balance determination," *Global and Planetary Change*, vol. 59, no. 1-4, pp. 138–148, 2007.

[46] F. Rau, M. Braun, M. Friedrich, F. Weber, and H. Goßmann, "Radar glacier zones and their boundaries as indicators of glacier mass balance and climatic variability," in *Proceedings of the 2nd EARSeL Workshop-Special Interest Group Land Ice and Snow*, pp. 317–327, 2000.

[47] E. R. et al., "Accelerated ice discharge from the antarctic peninsula following the collapse of larsen b ice shelf," *GEOPHYSICAL RESEARCH LETTERS*, vol. 31, no. 18, 2004.

[48] M. Demuth and A. Pietroniro, "Inferring glacier mass balance using radarsat: results from peyto glacier, canada," *Geografiska Annaler: Series A, Physical Geography*, vol. 81, no. 4, pp. 521–540, 1999.

[49] H. D. Pritchard, S. Luthcke, and A. H. Fleming, "Understanding ice-sheet mass balance: progress in satellite altimetry and gravimetry," *Journal of Glaciology*, vol. 56, no. 200, pp. 1151–1161, 2010.

[50] A. Shepherd, E. Ivins, E. Rignot, B. Smith, M. van Den Broeke, I. Velicogna, P. Whitehouse, K. Briggs, I. Joughin, G. Krinner, *et al.*, "Mass balance of the greenland ice sheet from 1992 to 2018," *Nature*, vol. 579, no. 7798, pp. 233–239, 2020.

[51] F. E. ACRI-ST, "The amorgos meris cfi (accurate meris ortho-rectified geo-location operational software) software user manual & interface control document," 2007.

[52] O. Arino, D. Gross, F. Ranera, M. Leroy, P. Bicheron, C. Brockman, P. Defourny, C. Vancutsem, F. Achard, L. Durieux, *et al.*, "Globcover: Esa service for global land cover from meris," in *2007 IEEE international geoscience and remote sensing symposium*, pp. 2412–2415, IEEE, 2007.

[53] E. B. C. author and A. S. Belward, "Glc2000: a new approach to global land cover mapping from earth observation data," *International Journal of Remote Sensing*, vol. 26, no. 9, pp. 1959–1977, 2005.

[54] S. Bontemps, P. Defourny, E. V. Bogaert, O. Arino, V. Kalogirou, and J. R. Perez, "Globcover 2009-products description and validation report," 2011.

[55] M. A. Friedl, D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X. Huang, "Modis collection 5 global land cover: Algorithm refinements and characterization of new datasets," *Remote sensing of Environment*, vol. 114, no. 1, pp. 168–182, 2010.

[56] S. P. A. M. A. F. Damien Sulla-Menashe, Josh M. Gray, "Hierarchical mapping of annual global land cover 2001 to present: The modis collection 6 land cover product," *Remote sensing of Environment*, vol. 222, no. 1, pp. 183–194, 2019.

[57] I. R. John Latham, Renato Cumani and M. Bloise, "Fao global land cover (glc-share) beta-release 1.0 database," Land and Water Division, 2014.

[58] M. B. et al., "Copernicus global land service: Land cover 100m: epoch 2015: Globe. dataset of the global component of the copernicus land monitoring service.," 2019.

[59] P. G. et al., "Finer resolution observation and monitoring of global land cover: first mapping results with landsat tm and etm+ data," *International Journal of Remote Sensing*, vol. 34, p. 2607–2654, 2013.

[60] L. S. Jun Chen, Ban Yifang, "Open access to earth land-cover map," *Nature*, vol. 514, no. 434, pp. 1476–4687, 2014.

[61] M. Marconcini, A. Metz-Marconcini, S. Üreyen, D. Palacios-Lopez, W. Hanke, F. Bachofer, J. Zeidler, T. Esch, N. Gorelick, A. Kakarla, *et al.*, "Outlining where humans live, the world settlement footprint 2015," *Scientific Data*, vol. 7, no. 1, pp. 1–14, 2020.

[62] J.-F. Pekel, A. Cottam, N. Gorelick, and A. S. Belward, "High-resolution mapping of global surface water and its long-term changes," *Nature*, vol. 540, no. 7633, pp. 418–422, 2016.

[63] L. Yu, J. Wang, and P. Gong, "Improving 30 m global land-cover map from-glc with time series modis and auxiliary data sets: a segmentation-based approach," *International Journal of Remote Sensing*, vol. 34, no. 16, pp. 5851–5867, 2013.

[64] C. Li, P. Gong, J. Wang, Z. Zhu, G. S. Biging, C. Yuan, T. Hu, H. Zhang, Q. Wang, X. Li, X. Liu, Y. Xu, J. Guo, C. Liu, K. O. Hackman, M. Zhang, Y. Cheng, L. Yu, J. Yang, H. Huang, and N. Clinton, "The first all-season sample set for mapping global land cover with landsat-8 data," *Science Bulletin*, vol. 62, no. 7, pp. 508–515, 2017.

[65] W. Li, R. Dong, H. Fu, J. Wang, L. Yu, and P. Gong, "Integrating google earth imagery with landsat data to improve 30-m resolution land cover mapping," *Remote Sensing of Environment*, vol. 237, p. 111563, 2020.

[66] Y. Zhou, J. Chen, Q. Guo, R. Cao, and X. Zhu, "Restoration of information obscured by mountainous shadows through landsat tm/etm+ images without the use of dem data: A new method," *IEEE transactions on geoscience and remote sensing*, vol. 52, no. 1, pp. 313–328, 2013.

[67] X. Zhu, F. Gao, D. Liu, and J. Chen, "A modified neighborhood similar pixel interpolator approach for removing thick clouds in landsat images," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 521–525, 2011.

[68] J. G. Cogley, R. Hock, L. Rasmussen, A. Arendt, A. Bauder, R. Braithwaite, P. Jansson, G. Kaser, M. Möller, L. Nicholson, *et al.*, "Glossary of glacier mass balance and related terms," *IHP-VII technical documents in hydrology*, vol. 86, 2011.

[69] M. Huss, G. Jouvet, D. Farinotti, and A. Bauder, "Future high-mountain hydrology: a new parameterization of glacier retreat," *Hydrology and Earth System Sciences*, vol. 14, no. 5, pp. 815–829, 2010.

[70] F. Maussion, A. Butenko, N. Champollion, M. Dusch, J. Eis, K. Fourteau, P. Gregor, A. H. Jarosch, J. M. Landmann, F. Oesterle, *et al.*, "The open global glacier model (oggm) v1. 1," *Geoscientific Model Development*, vol. 12, no. 3, pp. 909–931, 2019.

[71] U. Strasser, "Modelling of the mountain snow cover in the berchtesgaden national park," *FORSCHUNGSBERICHTE-NATIONALPARK BERCHTESGADEN*, vol. 1, no. 55, pp. ALL–ALL, 2008.

[72] F. Hanzer, K. Förster, J. Nemec, and U. Strasser, "Projected cryospheric and hydrological impacts of 21st century climate change in the ötztal alps (austria) simulated using a physically based approach," *Hydrology and Earth System Sciences 22 (2018), Nr. 2*, vol. 22, no. 2, pp. 1593–1614, 2018.

[73] F. Hanzer, K. Helfricht, T. Marke, and U. Strasser, "Multilevel spatiotemporal validation of snow/ice mass balance and runoff modeling in glacierized catchments," *The Cryosphere*, vol. 10, no. 4, pp. 1859–1881, 2016.

[74] D. Günther, T. Marke, R. Essery, and U. Strasser, "Uncertainties in snowpack simulations—assessing the impact of model structure, parameter choice, and forcing data error on point-scale energy balance snow model performance," *Water Resources Research*, vol. 55, no. 4, pp. 2779–2800, 2019.

[75] T. Sauter and F. Obleitner, "Assessing the uncertainty of glacier mass-balance simulations in the european arctic based on variance decomposition," *Geoscientific Model Development*, vol. 8, no. 12, p. 3911, 2015.

[76] G. Kaser, A. Fountain, P. Jansson, *et al.*, *A manual for monitoring the mass balance of mountain glaciers*. Unesco Paris, 2003.

[77] "Earth observing system: snow and ice". [Online]. Available: https://eos.com/snow-and-cloud/ . [Accessed: 21- Feb- 2020].

[78] N. S. Sekhon, Q. K. Hassan, and R. W. Sleep, "Evaluating potential of modis-based indices in determining "snow gone" stage over forest-dominant regions," *Remote Sensing*, vol. 2, no. 5, pp. 1348–1363, 2010.

[79] N. Delbart, L. Kergoat, T. Le Toan, J. Lhermitte, and G. Picard, "Determination of phenological dates in boreal regions using normalized difference water index," *Remote Sensing of Environment*, vol. 97, no. 1, pp. 26–38, 2005.

[80] U. Strasser, T. Marke, L. Braun, H. Escher-Vetter, I. Juen, M. Kuhn, F. Maussion, C. Mayer, L. Nicholson, K. Niedertscheider, *et al.*, "The rofental: a high alpine

research basin (1890–3770 m asl) in the ötztal alps (austria) with over 150 years of hydrometeorological and glaciological observations," *Earth System Science Data*, vol. 10, no. 1, p. 151, 2018.

[81] WGMS (2019): Fluctuations of Glaciers Database. World Glacier Monitoring Service, Zurich, Switzerland. DOI:10.5904/wgms-fog-2019-12. Online access: http://dx.doi.org/10.5904/wgms-fog-2019-12.

[82] Land Tirol, Innsbruck, Austria: Digital terrain model Tirol. 2018. [Online]. Available: https://www.data.gv.at/katalog/dataset/land-tirol_tirolgelnde . [Accessed: 21- Feb- 2020].

[83] Istituto NazionaIe di Geofisica e Vulcanologia, Rome, Italy: TINITALY/01. 2007. [Online]. Available: http://tinitaly.pi.ingv.it/ . [Accessed: 21- Feb- 2020].

[84] F. Hanzer, K. Förster, J. Nemec, and U. Strasser, "Projected hydrological and cryospheric impacts of 21st century climate change in the ötztal alps (austria) simulated using a physically based approach," *Hydrol. Earth Syst. Sci.(submitted)*, 2017.

[85] F. Hanzer, K. Helfricht, T. Marke, and U. Strasser, "Multilevel spatiotemporal validation of snow/ice mass balance and runoff modeling in glacierized catchments," *The Cryosphere*, vol. 10, no. 4, pp. 1859–1881, 2016.

[86] T. Marke, U. Strasser, F. Hanzer, J. Stötter, R. A. I. Wilcke, and A. Gobiet, "Scenarios of future snow conditions in styria (austrian alps)," *Journal of Hydrometeorology*, vol. 16, no. 1, pp. 261–277, 2015.

[87] F. Hanzer, U. Strasser, T. Marke, M. Warscher, B. Hynek, M. Olefs, and W. Schöner, "Latest developments of the alpine snowcover model amundsen: new modules, projects and perspectives," in *Geophysical Research Abstracts*, 2011.

[88] U. Strasser, *Die Modellierung der Gebirgsschneedecke im Nationalpark Berchtesgaden*. Nationalparkverwalt. Berchtesgaden, 2008.

[89] U. Strasser, M. Warscher, and G. E. Liston, "Modeling snow–canopy processes on an idealized mountain," *Journal of Hydrometeorology*, vol. 12, no. 4, pp. 663–677, 2011.

[90] M. Huss and D. Farinotti, "Distributed ice thickness and volume of all glaciers around the globe," *Journal of Geophysical Research: Earth Surface*, vol. 117, no. F4, 2012.

[91] L. Bruzzone and F. Melgani, "Robust multiple estimator systems for the analysis of biophysical parameters from remotely sensed data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 1, pp. 159–174, 2005.

[92] L. Pasolli, C. Notarnicola, G. Bertoldi, L. Bruzzone, R. Remelgado, F. Greifeneder, G. Niedrist, S. Della Chiesa, U. Tappeiner, and M. Zebisch, "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*, vol. 8, no. 1, pp. 262–283, 2015.

[93] U. d. S. di Trento, U. d. S. di Genova, and P. Italia, "Climate change initiative extension (cci+) phase 1 new essential climate variables (new ecvs) high resolution land cover ecv (hr_landcover_cci),"

[94] A. B. et al., "Feasibility of tundra vegetation height retrieval from sentinel-1 and sentinel-2 data," *Remote Sensing of Environment*, vol. 237, no. 111515, 2020.

[95] S. V. Stehman and R. L. Czaplewski, "Design and analysis for thematic map accuracy assessment: fundamental principles," *Remote Sensing of Environment*, vol. 64, no. 3, pp. 331–344, 1998.

[96] T. G. Farr, P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, *et al.*, "The shuttle radar topography mission," *Reviews of geophysics*, vol. 45, no. 2, 2007.

[97] L. Baetens, C. Desjardins, and O. Hagolle, "Validation of copernicus sentinel-2 cloud masks obtained from maja, sen2cor, and fmask processors using reference cloud masks generated with a supervised active learning procedure," *Remote Sensing*, vol. 11, no. 4, 2019.

[98] G. Mateo-García, L. Gómez-Chova, J. Amorós-López, J. Muñoz-Marí, and G. Camps-Valls, "Multitemporal cloud masking in the google earth engine," *Remote Sensing*, vol. 10, no. 7, 2018.

[99] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Cloud/shadow detection based on spectral indices for multi/hyperspectral optical remote sensing imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 144, pp. 235–253, 2018.

[100] Q. Zhan, W. Shi, and Y. Xiao, "Quantitative analysis of shadow effects in high-resolution images of urban areas," 2005.

[101] A. Telea, "An image inpainting technique based on the fast marching method," *Journal of Graphics Tools*, vol. 9, no. 1, pp. 23–34, 2004.

[102] X. Huang, L. Zhang, and L. Wang, "Evaluation of morphological texture features for mangrove forest mapping and species discrimination using multispectral ikonos imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 3, pp. 393–397, 2009.

[103] G. C. Z. Li and T. Zhang, "A cnn-transformer hybrid approach for crop classification using multitemporal multisensor images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 847–858, 2020.

[104] D. C. S. Jozdani, B.Johnson, "Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification," *Remote Sensing*, vol. 11, no. 14, 2019.

[105] D. Shi and X. Yang, "Support vector machines for land cover mapping from remote sensor imagery," in *Monitoring and Modeling of Global Changes: A Geomatics Perspective*, pp. 265–279, Springer, 2015.

[106] R. R. Irish, "Landsat 7 automatic cloud cover assessment," in *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI*, vol. 4049, pp. 348–355, International Society for Optics and Photonics, 2000.

[107] Y. Yang, H. Zheng, and H. Chen, "Automated cloud detection algorithm for multispectral high spatial resolution images using landsat-8 oli," in *Chinese Conference on Image and Graphics Technologies*, pp. 396–407, Springer, 2015.

[108] Y. Zhang, B. Guindon, and J. Cihlar, "An image transform to characterize and compensate for spatial variations in thin cloud contamination of landsat images," *Remote Sensing of Environment*, vol. 82, no. 2-3, pp. 173–187, 2002.

[109] C. Ji, "Haze reduction from the visible bands of landsat tm and etm+ images over a shallow water reef environment," *Remote Sensing of Environment*, vol. 112, no. 4, pp. 1773–1783, 2008.

[110] H. Lv, Y. Wang, and Y. Shen, "An empirical and radiative transfer model based algorithm to remove thin clouds in visible bands," *Remote Sensing of Environment*, vol. 179, pp. 183–195, 2016.

[111] S. Foga, P. L. Scaramuzza, S. Guo, Z. Zhu, R. D. Dilley Jr, T. Beckmann, G. L. Schmidt, J. L. Dwyer, M. J. Hughes, and B. Laue, "Cloud detection algorithm comparison and validation for operational landsat data products," *Remote sensing of environment*, vol. 194, pp. 379–390, 2017.

[112] Y. L. et al., "Evaluation of sentinel-2a surface reflectance derived using sen2cor in north america," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 6, p. 1997–2021, 2018.

[113] S. Qiu, Z. Zhu, and B. He, "Fmask 4.0: Improved cloud and cloud shadow detection in landsats 4–8 and sentinel-2 imagery," *Remote sensing of environment*, vol. 231, p. 111205, 2019.

[114] C.-C. Liu, Y.-C. Zhang, P.-Y. Chen, C.-C. Lai, Y.-H. Chen, J.-H. Cheng, and M.-H. Ko, "Clouds classification from sentinel-2 imagery with deep residual learning and semantic image segmentation," *Remote Sensing*, vol. 11, no. 2, p. 119, 2019.

[115] V. P. D. G. D. Olivier Hagolle, Mireille Huc, "A multi-temporal method for cloud detection, applied to formosat-2, venμs, landsat and sentinel-2 images.," *Remote Sensing of Environment*, vol. 114, no. 8, 2010.

[116] A. H. Sanchez, M. C. A. Picoli, G. Camara, P. R. Andrade, M. E. D. Chaves, S. Lechler, A. R. Soares, R. F. B. Marujo, R. E. O. Simões, K. R. Ferreira, and G. R. Queiroz, "Comparison of cloud cover detection algorithms on sentinel–2 images of the amazon tropical forest," *Remote Sensing*, vol. 12, no. 8, 2020.

[117] O. Hagolle and J. Colin, "Comment on "comparison of cloud cover detection algorithms on sentinel-2 images of the amazon tropical forest"," *Remote Sensing*, vol. 13, no. 5, 2021.

[118] C.-H. Lin, P.-H. Tsai, K.-H. Lai, and J.-Y. Chen, "Cloud removal from multitemporal satellite images using information cloning," *IEEE transactions on geoscience and remote sensing*, vol. 51, no. 1, pp. 232–241, 2012.

[119] M. Xu, X. Jia, M. Pickering, and A. J. Plaza, "Cloud removal based on sparse representation via multitemporal dictionary learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2998–3006, 2016.

[120] X. Li, H. Shen, L. Zhang, H. Zhang, Q. Yuan, and G. Yang, "Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 11, pp. 7086–7098, 2014.

[121] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial–temporal–spectral deep convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4274–4288, 2018.

[122] A. C. Siravenha, D. Sousa, A. Bispo, and E. Pelaes, "Evaluating inpainting methods to the satellite images clouds and shadows removing," in *International Conference on Signal Processing, Image Processing, and Pattern Recognition*, pp. 56–65, Springer, 2011.

[123] C. Yu, L. Chen, L. Su, M. Fan, and S. Li, "Kriging interpolation method and its application in retrieval of modis aerosol optical depth," in *2011 19th International Conference on Geoinformatics*, pp. 1–6, IEEE, 2011.

[124] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on image processing*, vol. 13, no. 9, pp. 1200–1212, 2004.

[125] A. Maalouf, P. Carré, B. Augereau, and C. Fernandez-Maloigne, "A bandelet-based inpainting technique for clouds removal from remotely sensed images," *IEEE transactions on geoscience and remote sensing*, vol. 47, no. 7, pp. 2363–2371, 2009.

[126] C. Paris and L. Bruzzone, "A sensor-driven hierarchical method for domain adaptation in classification of remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 3, pp. 1308–1324, 2017.

[127] L. Bruzzone and M. Marconcini, "Domain adaptation problems: A dasvm classification technique and a circular validation strategy," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 5, pp. 770–787, 2010.

[128] D. Tuia, C. Persello, and L. Bruzzone, "Domain adaptation for the classification of remote sensing data: An overview of recent advances," *IEEE geoscience and remote sensing magazine*, vol. 4, no. 2, pp. 41–57, 2016.

[129] S. S. David Hardoon and J. Shawe-Taylor, "Canonical correlation analysis: an overview with application to learning methods," *Neural Comput*, vol. 16, p. 2639–2664, 2004.

[130] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, p. 321–377, 1936.

[131] A. Sharma and D. W. Jacobs, "Bypassing synthesis: Pls for face recognition with pose, low-resolution and sketch," in *CVPR*, p. 593–600, Ieee, 2011.

[132] P. L. Lai and C. Fyfe, "Kernel and nonlinear canonical correlation analysis," *J. Neural Sys*, vol. 10, no. 5, p. 365–377, 2000.

[133] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE Transactions on Neural Networks*, vol. 22, no. 2, pp. 199–210, 2011.

[134] R. Gopalan, R. Li, and R. Chellappa, "Domain adaptation for object recognition: An unsupervised approach," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 999–1006, IEEE, 2011.

[135] B. Gong, Y. Shi, F. Sha, and K. Grauman, "Geodesic flow kernel for unsupervised domain adaptation," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pp. 2066–2073, IEEE, 2012.

[136] X. Ma, T. Zhang, and C. Xu, "Gcan: Graph convolutional adversarial network for unsupervised domain adaptation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8266–8276, 2019.

[137] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *Journal of Machine Learning Research*, vol. 17, no. 59, pp. 1–35, 2016.

[138] M. B. Bejiga, F. Melgani, and P. Beraldini, "Domain adversarial neural networks for large-scale land cover classification," *Remote Sensing*, vol. 11, no. 10, p. 1153, 2019.

[139] Y. Luo, L. Zheng, T. Guan, J. Yu, and Y. Yang, "Taking a closer look at domain shift: Category-level adversaries for semantics consistent domain adaptation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2507–2516, 2019.

[140] J. Lin, T. Yu, L. Mou, X. Zhu, R. K. Ward, and Z. J. Wang, "Unifying top-down views by task-specific domain adaptation," *IEEE Transactions on Geoscience and Remote Sensing*, 2020.

[141] N. Courty, R. Flamary, D. Tuia, and A. Rakotomamonjy, "Optimal transport for domain adaptation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016.

[142] B. Lucas, C. Pelletier, D. Schmidt, G. I. Webb, and F. Petitjean, "A bayesian-inspired, deep learning-based, semi-supervised domain adaptation technique for land cover mapping," *Machine Learning*, pp. 1–33, 2021.

[143] K. Saito, D. Kim, S. Sclaroff, T. Darrell, and K. Saenko, "Semi-supervised domain adaptation via minimax entropy," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8050–8058, 2019.

[144] J. Hoffman, E. Rodner, J. Donahue, T. Darrell, and K. Saenko, "Efficient learning of domain-invariant image representations," *arXiv preprint arXiv:1301.3224*, 2013.

[145] D. Tuia, M. Volpi, M. Trolliet, and G. Camps-Valls, "Semisupervised manifold alignment of multimodal remote sensing images," *Geoscience and Remote Sensing, IEEE Transactions on*, 2014.

[146] G. C.-V. Diego Marcos Gonzalez and D. Tuia, "Weakly supervised alignment of multisensor images," in *IEEE Int. Geosci. and Remote Sens. Symp (IGARSS)*, p. 2588–2591, Ieee, 2015.

[147] M. M. Xu, X. Jia, "Cloud removal based on sparse representation via multitemporal dictionary learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, p. 2998–3006, 2016.

[148] X. Z. et al., "A modified neighborhood similar pixel interpolator approach for removing thick clouds in landsat images," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, p. 521–525, 2012.

[149] M. Rußwurm and M. Korner, "Temporal vegetation modelling using long short-term memory networks for crop identification from medium-resolution multi-spectral satellite images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 11–19, 2017.

[150] A. Graves and N. Jaitly, "Towards end-to-end speech recognition with recurrent neural networks," in *International conference on machine learning*, pp. 1764–1772, PMLR, 2014.

[151] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[152] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Deep learning for time series classification: a review," *Data Mining and Knowledge Discovery*, vol. 33, no. 4, pp. 917–963, 2019.

[153] M. Herold, C. E. Woodcock, A. Di Gregorio, P. Mayaux, A. S. Belward, J. Latham, and C. C. Schmullius, "A joint initiative for harmonization and validation of land cover datasets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 7, pp. 1719–1727, 2006.

[154] U. Maulik and D. Chakraborty, "Remote sensing image classification: A survey of support-vector-machine-based advanced techniques," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 1, pp. 33–52, 2017.

[155] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc.*, vol. 35, pp. 99–109, 1943.