UNIVERSITÀ DEGLI STUDI
DI TRENTO

DEPARTMENT OF INFORMATION ENGINEERING AND COMPUTER SCIENCE
ICT International Doctoral School

ADVANCED METHODS FOR THE ANALYSIS OF MULTITEMPORAL MULTISPECTRAL SATELLITE IMAGES

Yady Tatiana Solano Correa

Advisor

Dr. Francesca Bovolo

Fondazione Bruno Kessler

Co-Advisor

Prof. Lorenzo Bruzzone

Università degli Studi di Trento

April 2018

To Leo,
for being love and light in space and time

"If we knew what we were doing,
It wouldn't be called "Research",
Would it?"

attributed to Albert Einstein

"All our dreams can become true;
If we have the courage to pursue them"

Walt Disney

Abstract

Thanks to the revisit property of the Earth observation satellites, a huge amount of multitemporal (MT) images are now available in archives. Such kind of images allows us to monitor land surface changes in wide geographical areas according to both long term (e.g., yearly) and short term (e.g., daily) observations. Evolution on the acquisition sensor technology has resulted in the availability of MT and multispectral satellite images with: i) Very High spatial Resolution (VHR) (e.g., QuickBird, WorlView-2) and, ii) very high temporal and high spatial resolutions (e.g., Sentinel-2 (S2)). Images acquired by such sensors allow for a detailed geometrical and temporal analysis when compared to medium or high spatial and temporal resolution data. Nevertheless, factors like satellite revisit period, the possible competing orders of different users on the satellite pointing (for VHR images only), the limited life of a satellite mission, and weather conditions can lead to: i) lack of enough images acquired by a single sensor to perform MT analysis (VHR case), and ii) lack of regular and continuous Time Series (TS) to perform short term MT analysis (very high temporal and high spatial resolution sensors case) at the level of small objects. Both problems arise from the application requirements on temporal resolution and are being of particular interest in the last years. Two main solutions to the above mentioned problems can be considered: i) use of multisensor VHR optical images to replace the missing acquisitions when a single VHR sensor is considered and; ii) development of regression techniques to reconstruct regular and continuous TS from both high temporal and very high temporal resolution sensors. In the literature, most of the MT analysis techniques have been designed to work with: i) VHR images acquired by single sensors, and ii) multispectral images acquired by high spatial resolution sensors, but with low temporal resolution or very high temporal resolution, but with low spatial resolution. Therefore, the effectiveness of existing techniques when applied to the complex MT problems in both VHR multisensor and very high temporal resolution images is reduced. Accordingly, the goal of this thesis is to develop novel techniques for the automatic analysis of MT multispectral satellite images such that images acquired by multisensor VHR and very high temporal resolution sensors can be analyzed.

The thesis provides four main novel contributions to the state-of-the-art. The first three contributions address the problems arising from the analysis of multisensor VHR multispectral images, whereas the fourth one deals with the problems faced while working with long TS acquired by sensors with high spatial and very high temporal resolutions. The first contribution presents an approach for unsupervised CD in multisensor MT VHR images, where the possible sources of noise/changes are studied in detail and a strategy to mitigate them at the levels of pre-processing and feature extraction is presented. In the second contribution, further attention is paid to the homogenization step and a method to generate homogeneous VHR TS focused on the homogenization of intrinsic spectral induced differences is presented. The third contribution further focuses on the detection of multiple changes, while relaxing the knowledge on the statistical distribution of the classes. To this aim, a method based on iterative clustering and adaptive thresholding is implemented. Comprehensive qualitative and quantitative experimental results, with real VHR multisensor datasets, confirm the effectiveness of the proposed approaches and led to the development of 3 contributions that allow to perform unsupervised bi-temporal CD by means of MT multisensor VHR images.

In the fourth contribution, an approach to handle images acquired by both high spatial and very high temporal resolution sensors is presented. To this aim, spectral, spatial and temporal information of S2 satellite images TS is exploited in different phases and in a fully automatic way, allowing for the derivation of different relevant products in the precision agriculture field. Comprehensive qualitative and, to some extent, quantitative experimental results confirm the capacity of the method to automatically exploit TS containing both high spatial and very high temporal resolution information. The proposed method can be easily extrapolated for other applications and other sensors with similar characteristics to those of S2.

Keywords. Remote Sensing, Multispectral Images, Multitemporal Images, Very High Spatial Resolution, Change Detection, Time Series, Multisensor, Homogenization, Very high temporal resolution.

Acknowledgements

When I first decided to search abroad for a PhD, I never thought I would have had the real opportunity to come and learn from first level researchers on my field. All I was sure about was that it was going to be a long trip full of adventures and challenges. For sure nobody could have ever told me how challenging it would be, neither how many adventures I would really be able to pursue and live.

Along this path, I must first thank to Francesca, my supervisor and guidance towards this trip. I know it might have not been an easy job during all these years from her side. But I also know that at the end we both manage to find our ways and grow together from all this experience. Though it was tough and hard to understand how things were supposed to be and to work, I was able to find the encouragement and support to continue searching and continue learning from her. So, thank you Francesca for having taken the risk to accept me as your first PhD student, and for having walked with me through all the process. Thank you for helping me to grow and for letting me see you growing at the same time. And thank you for the opportunity to continue learning from you and from this amazing international environment.

My second thanks go to Lorenzo, my co-supervisor. Having someone with lots of experience is definitely something that helps and enlightens a path. With time I was able to grow and learn from mistakes and assets. We were able to enter in that nice state in which we were all contributing with ideas and were all as equals. I surely still have a lot to learn, but I thank Lorenzo for being an example of experience and for giving a bit of it to my research life.

On the third place I must, with no doubts, thank my good friends in RSLab and RSDE. Some of them I have met from the very beginning, some others later on. But for sure all of them have been of equal relevance for me. Ana-Maria "Ferro"-Ilisei, you know me well, not many words, but for sure no words could ever express my gratitude to the very first person who ever provided me encouragement and understanding words in research and life. Thank you for being there, for helping me and for introducing me to another great person, Adamo. To both of you, thanks for the most amazing life times: hiking, ciaspole, eating, playing, skating, talking in 4 language at the same time and, the best birthday (26). Thank you!. Mahdi Khodadadadadadadazadeh, how to forget that it was you who helped me to remember my Colombian roots. Being able to talk in Spanish with you was just the beginning of what today is a great friendship. Thank you for the spontaneous, happy and nice way of life that you have. Thank you for reminding me that happiness is in the smallest things in the world, like coming to help you pick new clothes (missing it). And thank you for taking me out from home even when I was not willing to do it. Aravind, you came just one year after me and even though I was not the most experienced person, it was really nice to be helpful for you. I think we were mainly feeling lost together, but yet, we were also growing up and learning from each other. Thank you for giving me a place in research, for your constant encouragement, even when I myself was in doubts about everything. Thanks for your innocence that helped me to see things without thinking in the bad ones. Thanks for becoming a friend. Christopher, you are the kid of the story. I wouldn't have imagined to have the time to meet yet another special person and friend. I don't even know how did all started, but thanks for trusting me and letting me know you. Thanks for making me laugh with your geniality and "handsomeness".

I would also like to thank to those friends that I met outside the research laboratories, yet in the research environment. Ionuț, I will never ever again forgot that there is a Country called Romania. Thanks for those interesting times spent while listening to your unique life stories. Thanks for letting me be part of your life and to see how you grew up from being a student to being a doctor and a husband. Thanks for inviting me to make part of yours and Alesia's life. To both of you thanks for being there. To the Latin-American "community": Denisse, Aura and Ramiro. Guys, you brought me back to being at home. Happiness, complicity, spontaneity, talking, eating, being in family. Thanks for receiving me as part of your lives and for showing me what is there in the future as researchers and people. Don Jairo Tálaga!!, thanks for your

happiness and craziness. Thanks for feeding me and teaching me to cook and survive on my first year. Femi, thanks for all the conversations and time spent talking about random things.

To all the RSLab and RSDE members, thanks for being part of this research path. Special thanks to Milad and Elena, I have barely met you for few months now, yet you have made a big contribution to the happiness in my everyday research life. To Claudia, Massimo Z. and Manuel. Guys, it has been the most interesting experience to share and learn together in research. Claudia, thanks for being a closer example of what a woman can do in research, looking forward to reach close to you. Massimo Z., too smart to be true, yet so nice and funny to talk with you. Manuel, stop looking at every single detail!!!. Yet, thanks for showing me that sometimes, only sometimes, it is really important to do it.

Finally, but none the less, the most special thanks go for my family and friends in Colombia. I don't know what my life would have been without all the support that you have always given to me and that I am sure you will continue to give me. To my Mom, the strongest woman I have ever met in the world, thanks for your constant support and comforting words in the distance. Thanks for coming to visit me, even though it was for sure the toughest thing for you. Thanks for not giving up on me and making me feel proud of being your daughter. To my Dad, thanks for showing me always your support and proudness on my achievements. Thanks for being my "social representative" in all. Caro, no words could explain how happy I feel for you and for all the support that you have always given to me, for showing me that from your eyes I look great and that I can succeed on whatever I start. Thanks for the best gift ever: Mariana. Freddy and Sandra, thanks for your support and time in the distance. And also thanks for Isabella, the purest love on Earth. Angelica and Camila, no matter how I look at it, you are way too young, yet, you have shown me from your own life experiences how strong a woman can be in life when pursuing her dreams. Thanks to both for the crazy and nice conversations. Luza, you are like my family. Always there, no matter what, where or when. Thanks for your time, your passion, your trust and your love towards me. To Dario, David, Wilmer, Jaz, Ospi, Mawel, Mario, Guillermo, Leidy, Katerin, Félix, Willy, Fabio, FBI: thanks for keeping in contact throughout the years and distance. Leo, my friend, my family, my love. You have been, are and will be always my pole to ground. You and only you know better than nobody else how hard this path has been. You are the last name in this acknowledgements, but you also know you are not the last one on my life. Thanks for your kindness, love, patience, time, words. Thanks for trusting me, even before I would do it. Thanks for being there ALWAYS, from the beginning to the end. In good and bad, in happiness and sadness. You deserve this degree as much as I do.

Work acknowledgements

Part of this research was developed under the project "MS-TS – Analysis of Multisensor VHR image Time Series" with Digital Globe Foundation. Authors would like to thank DigitalGlobe Foundation for data crosscheck and images provided for the research development.

Part of this research is being developed under the project "SEOM - Scientific Exploitation of Operational Missions - S2-4SCI - Land and Water – Multi-temporal Analysis" funded by ESA.

Contents

List of Tables

List of Figures

Chapter 1

1. Introduction

In this chapter the basic concepts of remote sensing systems and multitemporal analysis in multispectral images are briefly overviewed. Two challenging problems in the context of multitemporal analysis are introduced: i) performing change detection in multisensor multispectral Very High Resolution (VHR) images and ii) analysing long-dense-discontinuous high temporal resolution time series. Then, the main motivations, objectives and the novel contributions of this thesis are presented. Finally, the whole structure and organization of the thesis are described.

1.1. Analysis of Multitemporal Multispectral Satellite Images

In the last years, a strong interest has been devoted to the development of novel methodologies for multitemporal information extraction and analysis. This is demonstrated by the sharp increase in the number of papers published in the major remote sensing journals, the increased number of sessions in international conferences and the increased number of projects related to multitemporal images and data.

The main reasons for this are: i) the increased number of satellites with higher revisit period that allow the acquisition of either long time series or frequent bi-temporal images, ii) the increased number of satellites acquiring images with Very High spatial Resolution (VHR) that allow for the study of changes at the level of small and single objects (e.g., IKONOS, QuickBird, GeoEye), iii) the new policy for data distribution of archive data that makes possible a retrospective analysis on large scale (e.g., the Landsat Thematic Mapper archive), and iv) the new policies for the distribution of new satellites data (e.g., ESA Sentinel).

Multitemporal information extraction methodologies differ because of both the specific investigated application and the kind of data available. Accordingly, different kinds of multitemporal products are more suitable to be considered in certain applications rather than others. The most widely addressed applications are related to products obtained through multitemporal classification (for change-detection analysis and detection of land-cover transitions) and trend analysis of temporal series of data (for change identification or forecasting/prediction).

According to an information theory perspective, the information in multitemporal data is associated with the dynamic of the variables that are measured, which is linked with the changes occurred between successive acquisitions. Thus, the most interesting applications are related to the classification/integration of multitemporal data/images for the detection of changes. We can distinguish among abrupt changes that occur in a short time (e.g., the ones caused by forest fires, floods and earthquakes) which can be appreciated by comparing bi-temporal images; or medium/long term changes, which can be appreciated only by comparing long time series of images (e.g., desertification, urban growth, vegetation monitoring). The above-mentioned applications can be addressed by using images acquired at different times by: i) the same sensor; ii) different sensors with similar properties (multi-sensor optical images); iii) different sensors with different properties (multi-sensor optical and radar images).

Assuming a change-detection perspective, the main methodological approaches proposed in the literature for the analysis of changes in multi-temporal remote-sensing images can be categorized in relation to the use of bi-temporal data or time series. Accordingly, three groups of methods for multitemporal images analysis can be considered: i) Unsupervised bi-temporal data analysis methods; ii) Supervised/semi-supervised bi-temporal data analysis methods; and iii) Time series analysis methods.

1. Unsupervised Bi-Temporal Data analysis: includes algorithms where multitemporal information is extracted by means of integration of multitemporal features/images. Multitemporal information is associated with differences in the spectral signatures of the land covers. After multitemporal

comparison the distinctions between changed and unchanged areas (i.e., each pixel is associated with one of two possible classes: the class of changed patterns or the class of unchanged patterns) is performed mainly by unsupervised decision approaches. Sometimes land-cover transitions can be distinguished, but without explicit labelling.

2. Supervised/Semi-Supervised Bi-Temporal Data analysis: includes algorithms that elaborate the multitemporal signature performing classification at the level of decision. Approaches in this category are mainly supervised or semi-supervised. They explicitly identify land-cover classes in each considered time instant and land-cover transitions are labelled accordingly (these methods can also be used when there are no changes between images for generating land-cover maps).
3. Time series analysis: the algorithms for time series analysis depend on the application perspective and the temporal scale. At very fine temporal resolution, changes such as the ones occurred on agricultural or urban areas become interesting and can be studied, whereas at lower temporal resolution, changes such as desertification/revegetation can be investigated. In both cases, a proper sampling of time series is required. Nevertheless, factors like atmospheric conditions or revisit period of the satellites can lead to the availability of irregular and non-continuous time series. Therefore, techniques that aim at building/studying adequate time series by means of gap filling and curve fitting methods, warping methods and by using multi-sensor and multi-source data can be found in the literature. Once proper time series are available, the analysis over different trends, at pixel or region levels, can lead to the classification/detection of intra-annual or seasonal changes and inter-annual changes.

Within each category, sub-categories can be identified according to the strategy for extracting multitemporal information. Figure 1.1 shows the tree of multitemporal analysis approaches, for bi-temporal data and time series. It is worth noting that this is a possible categorization of the methods in the literature. However, alternative categorizations could be considered.

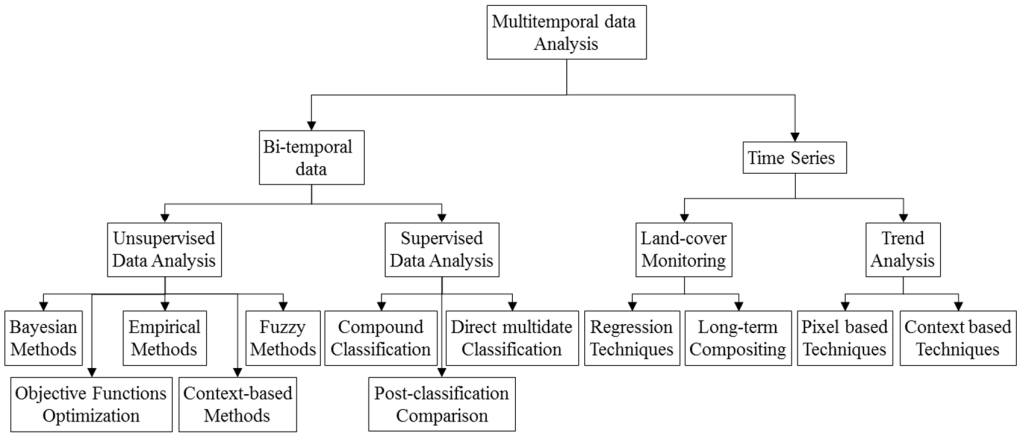


Figure 1.1. Tree of multitemporal data analysis approaches.

Methods in the tree should be implemented considering the characteristics of the kind of data. However, in general the overall block scheme of multitemporal analysis is the one in Figure 1.2. Multitemporal analysis for change-detection commonly assumes that images X_1, …, X_N acquired over the same geographical area at different times are accurately pre-processed in order to mitigate differences that do not depend on real changes occurred on the ground. In other words pre-processing aims at making multitemporal images as similar as possible to each other. Pre-processing usually includes radiometric corrections (absolute or relative normalization [1]–[6]), geometric corrections (co-registration, ortho-rectification, geo-referencing [7]–[10]), etc.

The first pre-processing step aims at making the multitemporal images radiometrically comparable. Ideally, a ground object should show the same brightness values if no change has occurred. In reality, measured

intensity values are sensitive to differences in acquisition geometry and environmental conditions (e.g., atmospheric conditions for optical passive systems). Radiometric conditions can be influenced by many factors such as imaging seasons, electromagnetic source incidence angles, meteorological conditions, etc. Acquisition geometry, such as sensor viewing angle, local incident angle, and solar orientation have strong effects on the acquired images. Atmospheric conditions have a serious impact on the measured intensity when using optical remotely sensed images. Absolute or relative normalization is often used to reduce this impact and make the multitemporal optical images comparable. Absolute normalization converts digital numbers to scaled surface reflectance and requires information about the atmospheric condition during image acquisition, which is not always easy to obtain. Relative normalization consists of the linear transformation of the spectral characteristics of the image to be corrected to match those of a reference image.

The second pre-processing step aims at making the multitemporal images geometrically comparable. This process ensures that corresponding pixels in the multitemporal images refer to the same geographic location on the ground and is referred to as geometric correction step. Geometric corrections are accomplished by image-to-image registration (co-registration) and/or image ortho-rectification in mountainous areas [11], and in urban areas for VHR images. Co-registration is often carried out by manually selecting ground control points. Automatic techniques, with different levels of success, also exist [12]–[17]. Co-registration is particularly difficult when the analysis involves high spatial resolution images, or when the images contain high frequency components (e.g., edges and linear features). Inaccurate co-registration is one of the main sources of errors in multitemporal analysis, and can lead to a significant degradation in accuracy [18]–[21].

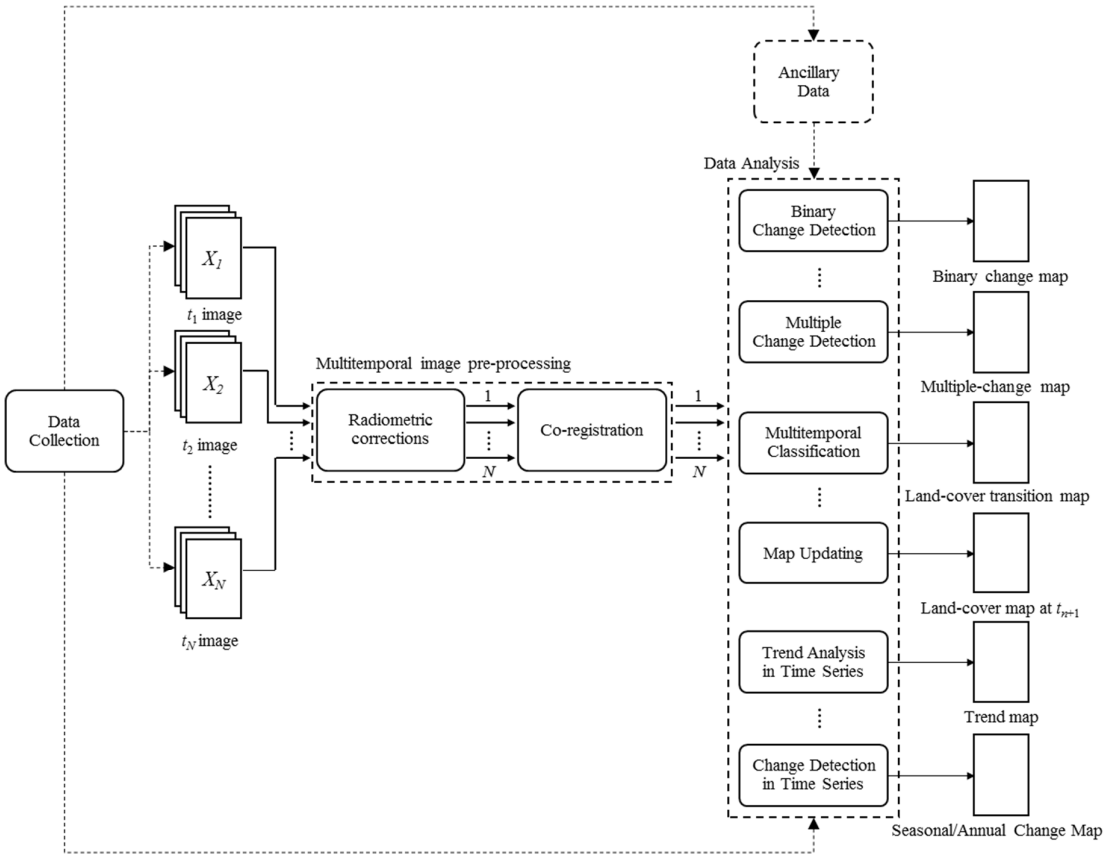


Figure 1.2. Overall block scheme of multitemporal multispectral data analysis.

Once pre-processing is completed, data analysis of multitemporal images can be conducted in different ways, resulting in different products (see Figure 1.2). Of particular interest is the use of unsupervised methods where, in general, the same approaches used for bi-temporal data can be used for multitemporal and time series data. Here, the aim is to identify changes and/or trends without the need of any ground reference

data. This is usually achieved by integrating multitemporal images and generating new features that highlight multitemporal information. These features are often referred to as change or trend indices since they are employed to highlight changes occurred in image pairs (X_1 and X_2 acquired over the same area at different times t_1 and t_2) or time series ($X_1, \cdots, X_N$ images acquired over the same area at different times $t_1, \cdots, t_N$). Change/trend indices are the main input to unsupervised change-detection or time series analysis procedures that generate maps without the use of ground reference information. Figure 1.3 (a) and (b) summarize the basic processing chain for multitemporal information integration and extraction (at both bi-temporal and time series levels), where the high level process is basically the same with the exception of the amount of images used as input. Due to their unsupervised nature, this kind of approaches are widely employed since, at an operational level, multitemporal ground reference information is often not available (e.g., the user is interested to investigate a phenomenon occurred in the past for which no information was collected), costly (i.e., it requires in situ surveys by experts with proper equipment) or impossible (i.e., ground truth is required over a very large or difficult to access area) to be collected.

Unsupervised change detection approaches mainly distinguish: i) between changed and unchanged pixels, for bi-temporal and time series data and; ii) between long term directional trend, seasonal and systematic movements; and additional irregular, unsystematic short term fluctuations, for time series data only. In the former case, some techniques allow identifying different kinds of changes as well. However, they do not give any explicit label to land-cover transitions. In the final map, each pixel is associated with one among the following classes: no-change (ω_n) or change (Ω_c). In the case that land-cover transitions can be distinguished, the change class can be further detailed in K kinds of changes as $\Omega_c = \{\omega_{c_1}, \omega_{c_2}, \ldots, \omega_{c_K}\}$ [22]. In the latter case, and depending on the research focus, rather the long term trend, seasonal, or short term fluctuations can be of special interest (or all three simultaneously) [23]. Climate scientists are especially interested in long term trends (e.g., land surface temperature, snow cover duration, sea level rise), but short term fluctuations are often relevant for an immediate response in managing natural resources (e.g., plant disease, fires). Such short term fluctuations (sometimes also called residuals) are the remaining component if the trend and seasonal components are removed from a time series.

Several mathematical operators can be applied to extract a change/trend index. The choice of the specific mathematical operator gives rise to different kinds of techniques [22]–[29]. Change/trend indices by bi/multi-temporal image comparison highlight information associated with changes in the spectral signature. In many occasions the limited availability of data forces the generation of a change/trend index by integrating/comparing an image pair or time series acquired with different spatial and spectral resolutions (multi-sensor data), or with different technologies, SAR and optical data (multi-source data). In order to extract the change/trend information after comparison, a proper unsupervised image analysis technique should be adopted.

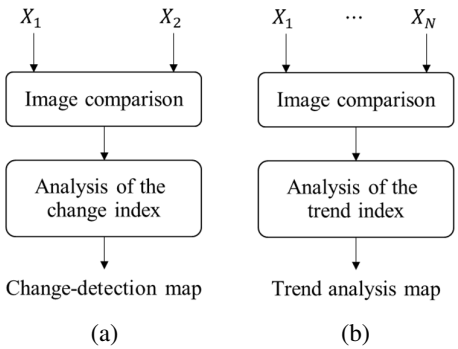


Figure 1.3. Block scheme of a standard change detection approach based on integration/extraction at feature level for (a) bi-temporal images and (b) multitemporal and time series images.

1.2. Motivation and Objectives of the Thesis

For decades, Earth Observation (EO) satellites have provided a unique way to observe our living planet from space. Thanks to the revisit property of the EO satellites, a huge amount of multitemporal images are now available in archives and continue to be acquired. This allows us to monitor land surface changes in wide geographical areas according to both long term (e.g., yearly) and short term (e.g., daily) observations. A comprehensive understanding of the global change is necessary for sustainable development of human society. As one of the interesting subtopics in global change study, detection of anthropogenic and natural impacts on land surface is essential for environmental monitoring [30]. Change Detection (CD) and temporal evolution investigation are among the hottest remote sensing application topics, which are continuously attracting attention in remote sensing community. Due to the improvement of both acquisition sensor technology and data processing algorithms, it is possible to get an accurate and automatic identification and extraction of features for understanding the environmental changes occurring on the ground due to natural and anthropogenic interactions.

The technological evolution has resulted in the availability of multitemporal and multispectral satellite images with: i) Very High spatial Resolution (VHR) at relatively low temporal resolution (e.g., QuickBird, WorlView-2, GeoEye, Pleiades), and ii) very high temporal resolution at high spatial resolution (e.g., Sentinel-2). These images allow a detailed geometrical and temporal analysis when compared to medium or high spatial and temporal resolution data [31]–[34]. Such detailed geometrical and temporal information is of greatest relevance, especially in cases where: i) rapid and precise impact assessment after a catastrophic event is crucial for initiating effective emergency response actions, ii) detailed information, at single object level, is required to study the behavior of vegetation and, ii) highly frequent and detailed information is required to follow long time events with high speed of change. Nevertheless, factors like satellite revisit period, the possible competing orders of different users on the satellite pointing (for VHR images only), the limited life of a satellite mission, the weather conditions and the intrinsic characteristics of the satellite (for very high temporal resolution images only) can lead to: i) lack of enough images acquired by a single sensor to perform multitemporal analysis (in the case of VHR images) [35], [36], and ii) lack of regular and continuous time series to perform short term multitemporal analysis (in the case of very high temporal resolution images) [29], [37], [38]. The former issue has an impact from the application perspective, where no continuous analysis can be performed. The latter one has an impact from both the methodological and application perspectives. From the methodological part, state-of-the-art methods have focused on managing data with regular and continuous acquisitions (only) coming from very high temporal resolution images, with low spatial resolution. As a consequence, from the application part, data are usually analyzed as the sum of many small objects, but seldom at small/single object level.

Regarding the lack of enough images acquired by a single VHR sensor to perform continuous multitemporal analysis, since a considerable number of satellites with VHR sensor on board, have been launched in the last decades, one possible solution could be to build time series (TS) by considering images acquired by different multispectral VHR sensors. The definition of such multisensor TS increases the probability to have sequences of multitemporal images with a proper time sampling (as per application requirements) but, at the same time, poses some challenges. In addition to real changes occurred on the ground, multisensor multitemporal images are affected by differences induced by the acquisition conditions (e.g., atmospheric conditions and acquisition system). Some of the differences in atmospheric conditions (e.g., cloud cover), and the differences in acquisition system (e.g., view angle and seasonal effects) affect single-sensor multitemporal image processing as well [30]. Whereas the system induced differences due to the type of sensor and the sensor acquisition modes are specific issues posed by multisensor TS that result in spectral and geometric differences (additional to the ones expected from real changes occurring on the ground). Therefore, such differences reduce the effectiveness of CD state-of-the-art methods where the assumption is that multitemporal images are acquired by the same sensor and under similar acquisition conditions. Thus, spectral and geometric differences are only expected when real changes are present in the study area. Such

observations hold for both physical (reflectance) and non-physical (Digital Numbers (DN)) quantities, making information non-homogeneous and thus non-comparable when using images acquired by different sensors [39], thus also increasing CD error. An appropriate multitemporal image homogenization is therefore required to reduce both spectral and geometrical differences, and to ensure that differences in the multitemporal images can be associated to real changes occurred on the ground. Furthermore, definition of CD steps and methods that take into account the intrinsic characteristics of VHR images are also required.

Regarding the lack of regular and continuous time series to perform short term multitemporal analysis, solutions to get algorithms suitable to handle irregular and non-continuous TS depend on the application itself. To give some examples, the perspective changes when the topic of study goes from: i) forest (big/large land cover with slightly small variability) to agriculture (small objects (crop field) with high variability) and; ii) snow (extensive land cover with constant and known variability) to fire (small or big event with unknown appearance or variability). Applications related to big/extensive land covers, can be easily addressed with TS acquired by sensors such MODIS or sometimes even Landsat series. Whereas the ones involving small objects are limited by the low spatial resolution offered by those sensors. Therefore, of particular interest, to contribute to state-of-the-art, are the applications involving small objects with very high variability, such as the agricultural monitoring [29], [37], [38], [40]. TS of vegetation indices derived from satellite spectral measurements are usually used to gain information on seasonal vegetation development. Independently of the spatial resolution of data, TS of vegetation indices, such as the normalized difference vegetation index (NDVI), are known to be hindered by noise arising from varying atmospheric conditions and sun-sensor-surface viewing geometries [41]–[43]. Off-nadir viewing and low sun zenith angles can also cause a similar effect [42]. Such problems contribute to render the TS acquired by very high temporal resolution sensors even more irregular and discontinuous. The literature contains reference to a broad variety of strategies designed to reduce the impacts of both missing and noisy acquisitions, based on smoothing and gap filling algorithms [29], [37], [38], but none of them account for the study at the level of small objects nor for irregular acquisitions. Therefore, there is a clear need to develop and/or adapt state-of-the-art methods: i) to render the very high temporal resolution acquisitions regular and continuous and; ii) to properly analyze long and dense TS with high spatial resolution.

In summary, in this thesis, the main objective is to define advanced methods to solve the problems of CD in multitemporal multispectral images and TS analysis by focusing in two important questions:

1. How to perform unsupervised CD by means of multitemporal multisensor VHR images?;
2. How to automatically exploit the TS containing both high spatial and very high temporal resolutions information?.

1.3. Novel Contributions of the Thesis

Based on the main motivations and objectives of the thesis, attention is focused on the development of advanced techniques for unsupervised and automatic analysis of multitemporal multispectral images. Research activities are mainly carried out to develop novel and robust techniques for addressing the considered challenging problems while working with: i) multisensor VHR images and; ii) high spatial and very high temporal resolution images. The main contributions and novelties of the thesis are briefly reported in the next:

i. An approach for unsupervised change detection in multitemporal VHR images acquired by different multispectral sensors (multisensor optical images).

Here, inspiration is taken from the framework for the design of CD systems for VHR images presented in [30]. From there, an approach for unsupervised CD in multisensor multitemporal VHR images is developed. The proposed approach exploits some of the concepts in [30] and use them to first analyze the sources of the non-homogeneous properties of multitemporal images acquired by different VHR multispectral sat-

ellite systems in the context of CD. Then it extends and integrates these concepts with a strategy for mitigating the non-homogeneous properties effects at both pre-processing and feature extraction/change detection level. The main steps of the proposed approach are: i) mitigation of differences induced by the use of VHR multitemporal images acquired by different sensors; and ii) detection of multiple changes occurred on the ground by means of high level physical features. The first step is conducted by defining homogenization procedures that address radiometric, spectral and geometrical differences. Thus, multitemporal multisensor images become more comparable (i.e., more homogeneous) across time. Homogenization is further improved by extracting proper features from multisensor images that allow for an effective multitemporal comparison across sensors at a given level of abstraction. Features are designed to detect multiple changes relevant to the user. Although the approach is general, here we concentrated on the selection of high level features suitable to detect changes in vegetation and urban areas. Linear/Orthogonal transformation features were selected with no loss of general validity for the detection step. They are shown to be effective for representing the multisensor information in a coherent physical way versus the considered sensor, but other features can be considered as well. The second step is conducted by means of Change Vector Analysis (CVA) in spherical coordinates, where the changes are represented by magnitude and direction variables. Separation among the changes is carried out in an automatic way. In the case of the magnitude variable, changed and non-changed pixels are separated by means of a Bayesian decision rule [44]. Whereas along direction variables, an adaptation of the Two-Stage Multithreshold Otsu (TSMO) method [45] is used.

ii. Generation of homogeneous VHR time series by non-parametric regression of multisensor multitemporal images.

Coming from the contributions presented in (i), it was seen that while working with multisensor multitemporal VHR images, further attention should be paid to the homogenization step. In this contribution, a method for generating homogeneous VHR TS focused on the homogenization of intrinsic spectral differences was faced. It is based on non-parametric regression and aims at generating consistent multisensor TS showing a homogeneous spectral representation. To achieve this goal, a model that represents the spectral relationship between two sensors (S_1 and S_2) is derived. For the derivation of that model, selection of invariant Radiometric Control Set Samples (RCSS) is required [36], [46], [47]. The proposed method jointly exploits the capabilities of Absolute Radiometric Normalization (ARN) [48] and Relative Radiometric Normalization (RRN) [49] approaches, by adapting them to the complexity of multisensor and VHR images. In the first step, a multisensor homogenization is carried on by transforming DN to physical values (ARN) and by transforming the multisensor images into a common spatial resolution, by means of state-of-the-art methods [14], [15], [17], [35], [39]. The second homogenization step is based on RRN approaches and performs a non-parametric regression (prediction) by means of a model derived from invariant RCSS. The model itself is derived by means of a machine learning algorithm i.e., Artificial Neural Networks (ANN) or Support Vector Regression (SVR). Because of the lack of one-to-one correspondence among VHR multisensor bands, spectral information appears to be non-comparable. Thus, a fusion of S_1 and S_2 bands must take place during RRN step. Such fusion helps to guarantee the selection of radiometrically stable features (because of a one-to-one correspondence), which in turn results in an easier selection of RCSS. The selection of the reliable RCSS is based on a CD-driven approach that takes, as input, features derived from the fused bands and is fully automatic. To demonstrate the effectiveness of the homogenization procedure, multitemporal information is extracted, specifically by applying binary CD by means of CVA.

iii. An approach to multiple change detection in VHR optical images based on iterative clustering and adaptive thresholding.

In order to separate changed from unchanged samples, a magnitude-direction space (derived from CVA) is widely used in literature. In such space, sub-optimal solutions are often employed that identify decision thresholds along the magnitude and direction independently [50]–[52]. Such methods rely on a priori knowledge of both class statistical models and expected number of changes [50]–[52]. Furthermore, no-

change class is usually assumed to be a single class [50], [53]–[56], whereas change class is assumed and modeled as multiple change classes [51], [57]. Two main limitations arise from the above-mentioned assumptions. The first one is given by the definition of the threshold value along the magnitude, which is defined by considering all kinds of change as a single large metaclass [57]. This implies that the threshold along the magnitude is defined according to the average properties of the different kinds of change and does not gather the peculiarities of each kind. The second one is given by the definition of different thresholds along the direction variable, which is usually defined by assuming a known statistical distribution and number of classes (usually as a single class) [50]–[52]. Both hypothesis are usually satisfied when low and medium spatial resolution images are considered, but are hard to be when working with VHR optical images. Thus, we propose an approach to multiple CD, which jointly exploits the histogram distribution of the magnitude and direction variables in (hyper-) spherical coordinates. The approach is distribution free (for the change class) and uses adaptive thresholds in accordance with each kind of change in order to tune the no-change class. Thus, it is particularly suitable for VHR images. The proposed approach first performs an initial discrimination between unchanged and changed samples. Then, automatically identifies the sectors associated to different kinds of change by iteratively performing density-based clustering and region growing ([58], [59]), while adding samples from around the peaks of the joint magnitude-direction histogram. Finally, it tunes the magnitude threshold by adapting it according to the characteristics of different kinds of change. The method is unsupervised, application-independent and identifies the number of changes automatically.

iv. Spatio-temporal evolution of crop fields in Sentinel-2 Satellite Image Time Series.

The main goal of this contribution is to exploit the information contained in Satellite Image Time Series (SITS) acquired with both high spatial and very high temporal resolution. A clear example of such SITS, are the images acquired by the Sentinel-2 (S2) constellation, where opposite to up-to-date sensors, both characteristics are present. Such characteristics end up producing SITS with irregular, high frequency and high spatial resolution acquisitions, which are of particular interest in many applications (i.e., forest, snow, fire, sea level, and vegetation dynamics), but cannot be easily handled by state-of-the-art methods. Other sensors, such as MODIS or Landsat, have been traditionally used in literature to perform the study of such applications, but none of them is suitable when considering small objects like the fields in precision agriculture that show a high change frequency. Thus, here we focus on the agricultural analysis application. In the proposed approach, the S2-SITS images are processed in a number of automatic steps that lead to the derivation of different relevant products. In the first step, a crop field map is constructed that accounts for: i) fields/areas that have had vegetation at least once in a given period; and ii) the temporal evolution and rotation practices of the fields [60]. Here, spectral, spatial and temporal information are exploited. In a second step, regular and continuous S2-SITS are derived based on an adaptive non-parametric regression, opposite to state-of-the-art techniques, this process is performed at single crop field level [29], [37], [38], [61]. In a third step, and as a consequence of having a regular and continuous SITS, adaptation of state-of-the-art methods is done in order to derive phenological parameters and to integrate rules allowing to differentiate among objects corresponding to crop fields and those that not [29]. Additional products are derived (e.g., phenological maps, cumulative indices maps) that allow for the understanding of agricultural areas, and therefore are useful in precision agriculture. The proposed approach is fully automatic, can easily account for the new upcoming images, is able to deal with irregularly sampled SITS and can be extrapolated for other applications and other sensors with similar characteristics to S2.

1.4. Structure of the Thesis

This thesis is organized in seven chapters. The present chapter gives a brief introduction on the analysis of multitemporal multispectral satellite images and the main steps to be followed while extracting multitemporal information. It presents the motivation and objectives of the thesis, it provides a brief summary of each novel contribution and describes the structure of the whole thesis.

Chapter 2 presents a review of the state-of-the-art on multitemporal analysis of remotely sensed images acquired by multispectral sensors. Attention is paid to: i) unsupervised bi-temporal image analysis, ii) (semi-)supervised bi-temporal image analysis; and iii) time series analysis. The first set of methods mainly exploits multitemporal image comparison techniques for the detection of presence/absence of changes. The second one relies on classification methods for detecting land-cover transitions. And the third one makes use of both unsupervised and supervised/semi-supervised methods for land-cover monitoring and trend analysis in long time series.

Chapter 3 introduces a novel approach for unsupervised change detection in multitemporal VHR images acquired by different multispectral sensors. First, the challenges arising from the intrinsic characteristics of VHR images, plus the use of images acquired by different multispectral sensors are discussed, as well as the state-of-the-art techniques used to mitigate them. Second, traditional features extracted to separate changes of interest among the ones of no-interest along with their advantages and disadvantages are also analysed. Finally, a novel approach is presented that extracts higher-level physical features for both the mitigation of multisensor differences and the detection of multiple changes in the context of VHR images.

Chapter 4 presents an approach for the generation of homogeneous VHR TS by means of a non-parametric regression of multisensor multitemporal images. To this aim, suitability of ARN and RRN methods, available in the literature, to perform homogenization of VHR multisensor images is studied. The proposed approach takes then advantage of both ARN and RRN methods to perform the homogenization, but introduces a novel concept based on a non-parametric regression and CD-driven RRN.

Chapter 5 describes an approach for multiple change detection in VHR optical images (single- and multi-sensor) based on iterative clustering and adaptive thresholding. Problems arising from the most common assumptions on the number of classes (changed and unchanged) and their statistical distribution are first discussed. Then, state-of-the-art solutions are presented along with their critical characteristics. Finally, a solution is proposed that combines the best from the state-of-the-art methods and integrates it with a novel iterative clustering and adaptive thresholding. The proposed approach is automatic and based on 3 steps, and does not require a statistical model for the classes of change.

Chapter 6 presents an approach for the analysis of images acquired with both high spatial and very high temporal resolutions. Spatio-temporal evolution of crop fields in S2-SITS is studied as an example of such images. Where first, a state-of-the-art review is made to show the relevance of developing new algorithms that account for the characteristics of sensors such as S2, as well as their importance for precision agriculture applications. Then, S2-SITS spatial, spectral and temporal characteristics are jointly exploited to derive products suitable for the agricultural analysis. To this aim, the proposed approach is divided in different stages that work in an automatic way and without need of in-situ data.

Chapter 7 draws the conclusions for this thesis. The remaining open issues and further developments of the research activities are also discussed.

Chapter 2

2. State-of-the-art: Multitemporal Analysis of Remotely Sensed Image Data[1]

In this chapter, literature is reviewed about: i) unsupervised bi-temporal images analysis, ii) (semi-)supervised bi-temporal images analysis; and iii) time series analysis. The first set of methods mainly exploits multitemporal image comparison techniques for the detection of presence/absence of changes. The second one relies on classification methods for detecting land-cover transitions. The third one makes use of both unsupervised and supervised/semi-supervised methods for land-cover monitoring and trend analysis in long time series. Images acquired by multispectral optical systems at medium, high and very high spatial resolution are considered.

2.1. Unsupervised bi-temporal data analysis

Following the block scheme, for standard CD approach, presented in Figure 1.3.a., and in order to extract the change information, we follow two steps: i) comparison of an image pair, and ii) multitemporal information extraction by a change index. The first step relies on the use of a comparison operator, e.g., univariate image differencing, vegetation index differencing, regression, change vector analysis, etc [62]. Whereas the second step relies on CD methods that identify the changes according to the feature space created by the comparison operator. In this subsection, we focus the attention in unsupervised image analysis techniques traditionally adopted for bi-temporal images. In the literature approaches have been developed including pixel-based [22], [53], context-based [53], [63], single-scale [22], [24], [26], [28], [64]–[66] and multi-scale [63], [67]–[72] approaches. Among pixel-based techniques, the most widely used are based on the selection of a decision threshold that aims at separating changed from unchanged pixels. The decision threshold can be selected either with a manual trial-and-error procedure (according to the desired trade-off between false and missed alarms) or with automatic techniques (e.g., by analysing the statistical distribution of the image after comparison, by fixing the desired false alarm probability [73], [74], following a Bayesian minimum-error/cost decision rule [53], [75], using methods based on fuzzy theory [66], [76], etc.). Among context-based techniques there are the ones based on fixed size sliding windows [53], [77], and the ones based on adaptive segmentation [64], [78]. Among multi-scale techniques, three main strategies can be identified: adaptive multiscale techniques for SAR images [68], multilevel parcel-based technique suitable for very high resolution images [67], [78], and approaches based on the use of similarity measures [77], [79]. In the next, a further literature review is presented for the specific cases of single-sensor and multi-sensor/source images.

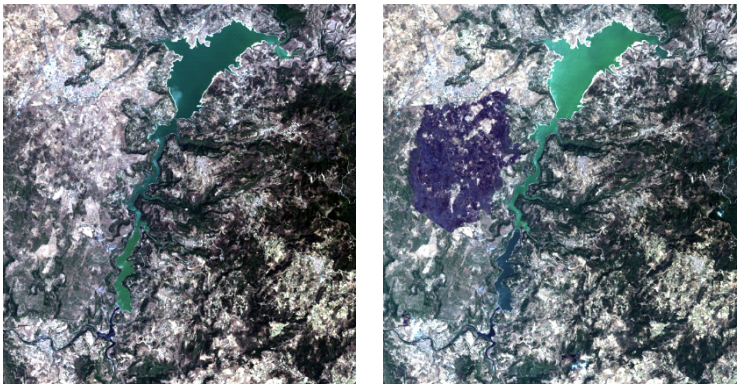
2.1.1 Comparison approaches for single-sensor bi-temporal analysis

When dealing with single-sensor optical data acquired by passive sensors, image comparison mainly relies on the difference operator. This is because the noise model in optical images is additive and the natural classes have a Gaussian distribution. Thus, the difference operator results to be the most effective one.

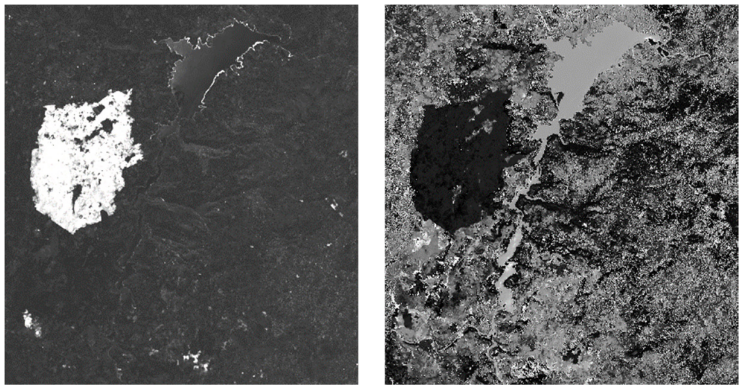
The simplest way to apply the difference operator is to consider the same spectral band for X_1 and X_2 and perform subtraction pixel-by-pixel. This technique is referred to as Univariate Image Differencing (UID) [22], [24], [25], [27]. The follow up of this approach leads to the use of multiple spectral bands [22]. This

[1] Part of the contents in this chapter is taken from:
Book Chapter: F. Bovolo, L. Bruzzone, and Y. T. Solano-Correa, "Multitemporal Analysis of Remotely Sensed Image Data," in Comprehensive Remote Sensing, 1st Edition, Elsevier, vol. 2, pp. 156-185, Nov. 2017.

technique takes the name of Change Vector Analysis (CVA) because the result of differencing is a multi-dimensional Spectral Change Vector (SCV) [30], [48], [50], [51], [55], [80]–[85]. Under the assumption of Gaussian distributed natural classes and being the difference a linear operator, classes of change and no-change in the SCV feature space result to be Gaussian distributed as well (if low-medium spatial resolution images are considered) [30], [56], [86]. However when non-linear features are extracted from SCVs, the analysis becomes more complex. In fact, in order to better characterize the properties of changes it is common to compute the magnitude and the directions of SCVs by applying Cartesian to Spherical coordinates transformation [30], [48], [50], [51], [55], [81], [83]–[85]. The magnitude image is such that pixels associated with land-cover changes present values significantly higher than those of pixels associated to unchanged areas [22], [50]. Both change and no-change classes are often assumed to follow a Gaussian [55] or nearly Gaussian [80] statistical distribution. However, in [50] and later on in [56] it has been demonstrated that, under some reasonable assumptions, they are better approximated by one or more Rayleigh and Rice distributions, respectively. Direction variables carry less information about unchanged samples since they result to be uniformly distributed [50]. They become highly relevant when analysing the classes of change instead, since they characterize different kinds of change. Changes assume preferred directions depending on their kind. Examples can be found in the literature where the direction information is used in the change detection process [22], [50], [51], [55], [83], [84]. Figure 2.1 gives an example of a change detection problem in multispectral optical images and of the change index after comparison by CVA. Nevertheless, the detection is performed without following a specific statistical distribution of the data. A detailed analytical derivation of class statistical distribution in the magnitude-direction domain can be found in [50], [56], but only the unchanged class is further analyze. The above mentioned methods were developed under the assumption that multitemporal images are well co-registered and radiometrically corrected. If the assumptions on image pre-processing are not satisfied, the statistical distribution of spectral change vectors becomes more complex and the change-detection process rather difficult and less effective. This points out the importance of a proper pre-processing [50]. Cartesian to Spherical coordinates transformation preserves the dimensionality problem. Sometimes this can be a drawback since it hampers the visualization of the feature space when the dimensionality becomes higher than three. A possible alternative is to use Compressed Change Vector Analysis (C^2VA) [51]. C^2VA compresses the information present in SCVs by computing the direction as the angular distance between the multispectral difference vector and a reference vector. If combined with the magnitude we obtain a 2-dimensional feature space that can be easily visualized and where no information is neglected. Both characteristics become highly interesting and are successfully applied when multitemporal hyperspectral images are considered [83]–[85]. A limitation of this approach is that the lossy compression of the direction information may result in recognizing some classes of change as a single class. Figure 2.2 gives an example of a change detection problem in hyperspectral optical images and C^2VA features.

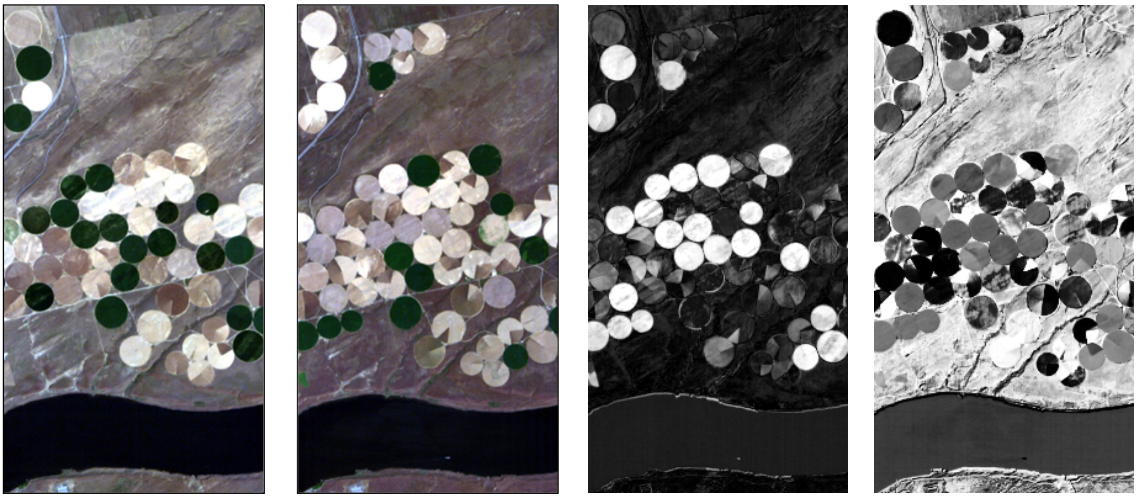


(a) (b)



(c) (d)

Figure 2.1. Example of image comparison in bi-temporal optical images. RGB true colour composition of Landsat-8 images acquired in: (a) July 2013, and (b) August 2013. (c) Magnitude and (d) direction images computed according to CVA. The area of interest is located close to the Lake Omodeo in Sardinia Island (Italy). Changes occurred between acquisition dates are associate to a forest fire (left-top) and the increase of the lake surface (center-top).



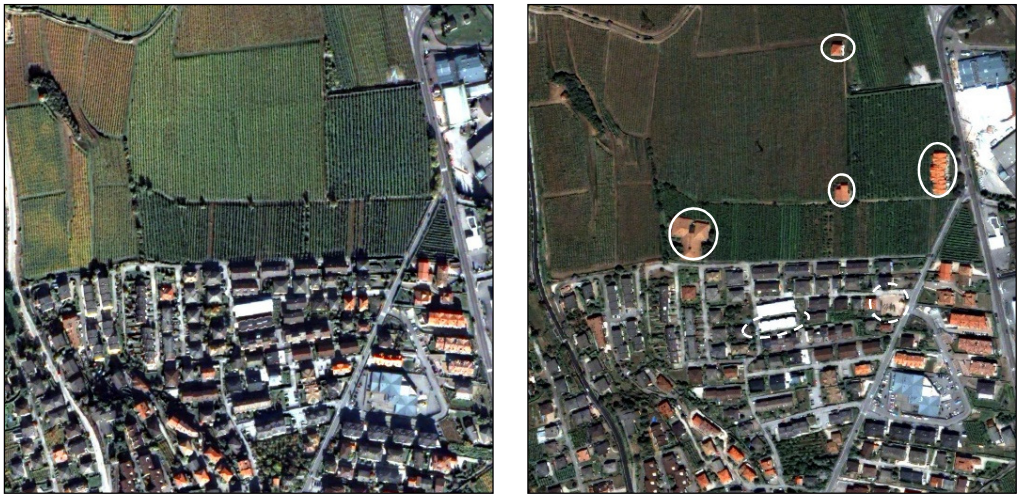
(a) (b) (c) (d)

Figure 2.2. Example of image comparison in bi-temporal optical images. RGB true colour composition of Hyperion EO-1 images acquired in: (a) 1st May 2004, and (b) 1st May 2007 (images downloaded from Geological Survey (USGS) website http://earthexplorer.usgs.gov/). (c) Magnitude and (d) direction images computed according to C^2VA. The area of interest is located close to Hermiston city in Umatilla County, U.S. The study area is an agricultural land and changes are mainly associated to crops.

CVA and difference operator have been mainly applied to the original image feature space. However examples can be found where they are applied to the posterior probability space [87] as well as to vegetation indices (Vegetation Index Differencing) [22], [28] or other linear (e.g., Tasselled Cap Transformation [28], Multivariate Alteration Detection [88]–[91]) or non-linear combinations of spectral bands. Transformation-based techniques like Multivariate Alteration Detection (MAD) [91] have been widely investigated resulting in several subsequent ameliorations like the Iteratively Reweighted (IR)-MAD [88], [90] combined with Maximum Autocorrelation Factor (MAF) transformation to find maximum change areas and its kernel version [89]. In the transformed feature space and after differencing, similar to CVA, unchanged and changed areas will show significantly different values. MNF (Maximum Noise Fraction)/MAD [89] has been employed for change detection in multisensor multitemporal images as well (for low and medium spatial resolution images only). The main characteristics that make the MNF/MAD suitable for multisensor change detection are: i) random variables associated to the acquisitions before and after the event of change should not necessarily have same dimension (i.e., multitemporal images are not required to have the same

number of spectral channels); ii) the MNF/MAD method is invariant to linear transformations, which implies that the impact of missing radiometric normalization and rectification is lower than for other change-detection approaches. Nevertheless, MNF/MAD method lacks from the detailed statistical and formal analysis on how to separate among changed and unchanged samples that CVA like methods have. An alternative approach based on transformations is to use Principal Component Analysis (PCA). PCA can be applied separately to the feature space at single time image [22], [26], [27] or jointly to the stacked image features [92], [93]. In the first case, comparison should be performed in the transformed feature space before performing change detection; in the second case, the minor components of the transformed feature space contain change information. Other linear transformations have been used such as Tasselled Cap Transformation (TCT), and Gram-Schmidt orthogonalization [35], [94], [95], all of them when considering low and medium spatial resolution images.

More recently, the multiscale/resolution concept has been introduced in the multitemporal image analysis literature. The first works were devoted to SAR data because of their complexity. However, their use resulted to be effective in optical images as well. As an example the Wavelet decomposition was used in [63], [70], [96], and the Contourlet transform was used in [97]. Such transformations have been applied either before or after image comparison. Multiresolution profiles for multitemporal images have been elaborated by using features extracted from multiresolution segmentation [67], [98], morphological profiles and their improvements [80], [82], and methods based on scale-invariant feature transform (SIFT) [99]. More sophisticated approaches to the representation of multiresolution information have been developed when VHR images are considered (Figure 2.3 shows an example of a change detection problem in VHR optical images). Such approaches aim at modelling the high level semantic information available in VHR images effectively [30]. Though no single method exists in literature that can deal with the whole complexity added by working with VHR images from-to end.



(a) (b)

Figure 2.3. Example of change detection problem in VHR optical images. True colour composition of QuickBird pansharpened images acquired in: (a) October 2005, and (b) July 2006. The area of interest is a sub-urban area located close to Trento (Italy). Changes are mainly associated to buildings (see white circles in the right image).

From the whole literature presented in this subsection, it is clear that a lot of attention has been devoted to the development of methods for multitemporal analysis in the context of low and medium multispectral images. But less or few attention has been devoted to the same type of analysis when considering VHR images. This is in part a consequence of the sensors evolution, and thus image availability, but also a consequence of the complexity added while working with VHR images.

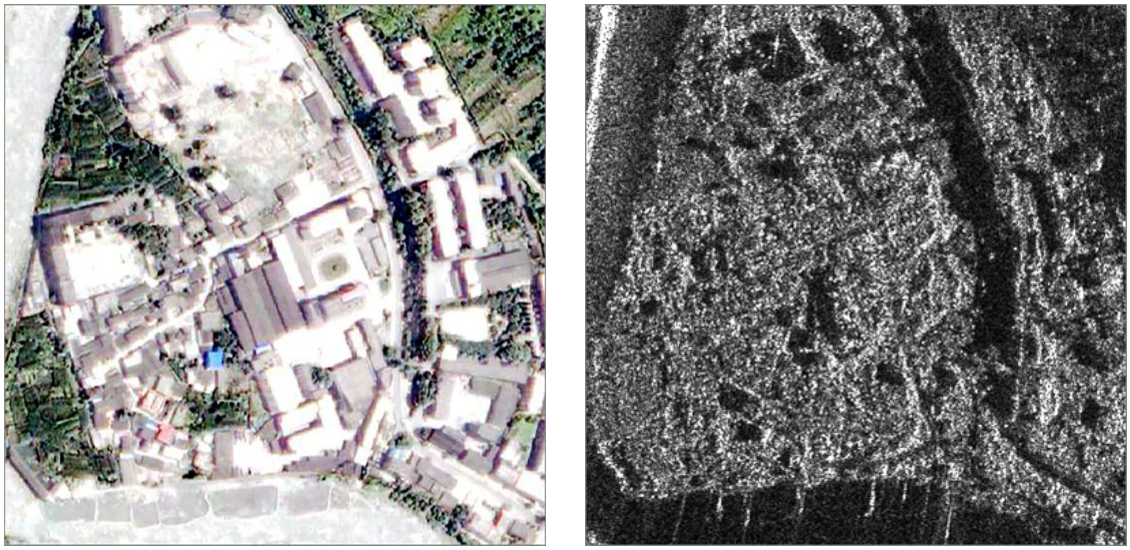
2.1.2 Comparison approaches for multi-sensor and multi-source bi-temporal analysis

As a complement to the use of single-sensor optical data for multitemporal analysis, one may consider the use of multi-sensor (multitemporal images acquired by different sensors, either passive or active) and multi-source data (multitemporal images acquired by different sensors or modes and ancillary data) to enhance the final product or to integrate missing temporal data due to atmospheric conditions (i.e., presence of clouds). Methods able to process/compare this kind of data fall in the category of remote sensing image fusion. Where image fusion can be seen as: i) fusion methods for images from the same sensor to increase the spatial and spectral information (also known as pansharpening), or ii) comparison of images from different sensors. Here we concentrate our attention on the latter one, where different kind of images, i.e., images with different spatial or spectral resolutions (e.g., Landsat vs MODIS), or, in many cases, images with different acquisition technologies (e.g., SAR and optical images), are used. Of particular interest is the case where images with different spatial or spectral resolutions are considered. Since up to few years ago, this was referred only to the integration of sensors having rather big (10-20 times bigger) differences in spectral, spatial and temporal resolutions. Nowadays, similar approaches can be found while talking about VHR optical images, where the spatial and spectral resolutions differences occur at a different scale. Yet, fusion of multisensor VHR images result in a problem given the complexity added by their intrinsic characteristics. Few approaches can be found in the literature that use multisensor VHR images as a solution to mitigate the missing temporal data [6], [35], [95], [100]–[104].

In change detection, the fusion of SAR and optical data is important from two perspectives. First, on many occasions the limited availability of data forces the generation of a change indicator through the comparison of images acquired over the same area, but with different technologies. Although the images were acquired with sensors that have different technologies, they are two different representations of the same physical reality and consequently can be compared [105]. Recently, similarity measures have played an essential role in performing such image comparison. Mercier et al. [79], successfully used Kullback-Leibler divergence to compare an ERS SAR image with a SPOT image. Liu et al. [106] proposed a multidimensional evidential reasoning approach to extract change information from heterogeneous multitemporal images. Second, single-source multitemporal images (i.e., optical or SAR) do not provide exhaustive documentation of changes since they do not capture all the available information. Unlike optical, SAR images are immune to atmospheric problems, such as cloud presence, which increases the probability of acquiring images suitable for change detection. The existence of speckle, however, impedes the accurate identification of shapes and edges. Optical images show more details and allow the detection of sharp edges and region boundaries [107]. Change detection can benefit from the complementary nature of the change information represented by each type of data.

In the case of catastrophic events, the availability of pre and post event images acquired as closer as possible, in time, to the event becomes necessary. Nevertheless, the probability of having such images, acquired by the same technology sensor is reduced because of atmospheric conditions (in the case of optical data) or because of the revisit period of the sensors. Thus, the possibility of having an optical and SAR images as pre-event and post-event acquisition becomes an interesting option for satisfying operational time constraints. Information on the impact of an event can be then derived from suitable imagery by comparing data from a chosen reference before the event to imagery acquired shortly after the event. Optical VHR sensors availability has increased a lot in the last decade as well as their frequently updated image archives, making this kind of data well suited as the pre-event reference data source. If post-event VHR optical data are also available, it can be used to investigate the impact of the event. If no VHR optical data are available, then SAR one become the suitable option as post-event data. In this case, approaches must be considered to analyze, detect and extract the change information. Brunner et al. [108] presented a novel method that detects buildings destroyed in an earthquake using pre-event VHR optical imagery (QuickBird and WorldView-1) and post-event VHR SAR imagery (TerraSAR-X and COSMO-SkyMed). Figure 2.4 shows an example of the pre-event and post-event VHR optical and SAR images, where the complexity

and challenges of using this kind of data for change detection can be seen. The method operates at the level of individual buildings and assumes that they have a rectangular footprint and are isolated. Building parameters are estimated from the pre-event optical imagery, and used to predict the expected signature of the building in a simulated post-event SAR scene. The similarity between the predicted image and the actual SAR image is analyzed. If the similarity is high, the building is likely to be still intact, whereas a low similarity indicates that the building is destroyed.



(a) (b)

Figure 2.4. Image subsets of optical and SAR images over the area in Yingxiu with Earthquake occurred on May 12 of 2008. (a) Pre-QB image acquired on June 26 of 2005 and (b) Post-TSX image acquired on June 07 of 2008, with viewing direction from left to right [108].

Fusion in remote sensing analysis is not restricted to images acquired with different resolution or technologies; it is also extended to include the fusion of different kinds of information extracted from the same source. In an attempt to improve the quality of the binary change detection, Bruzzone and Fernández-Prieto [109] proposed an unsupervised change detection approach that uses consensus theory to integrate several change variables. Ban et al. [110] demonstrated that the fusion of both optical and SAR data could improve change detection. In this research, a multidimensional change image was constructed by combining SAR and optical change variables. An iterative classification strategy is then adopted to separate changed and unchanged classes. Mishra and Susaki [111] demonstrated the improvement of change detection while using SAR and optical data when compared to the result obtained by using a single pair of SAR or optical images. The change detection was carried out by using the Normalized Difference Ratio (NDR) for SAR image and the Normalized Difference Vegetation Index (NDVI) for the optical one, and applying the difference operator. Authors proved a correlation between the increase/decrease of backscattering and the increase/decrease of NDVI values that was used to separate between change and non-change. Comparison with the results obtained by using NDR and NDVI in single source change detection proved the effectiveness of the approach. Gong et al. [71] improved the quality of the final change indicator through the fusion of different change variables extracted from the same multitemporal dataset by using different comparison operators: i) the mean-ratio image that emphasizes changed areas in the scene; and ii) the log-ratio image that reflects more the background information, i.e., no-change. The fusion of the change variables is performed in the wavelet domain, where different rules were developed for the low- and high-frequency components. The same wavelet-based fusion approach was adopted by Ma et al. [112], but with different fusion rules. In a similar way, Hou et al. [72] fused two change images using the Gauss log-ratio and log-ratio

comparison operators. In contrast, in Du et al. [113] each band of a multispectral image was used to generate a change map and fusion performed at decision level.

2.1.3 Multitemporal information extraction

In order to effectively extract information after image comparison, the change index should be analyzed. Depending on the comparison operator, changes can be identified in different positions of the feature space. Methods for detection available in the literature can be classified into: i) empirical methods; ii) methods based on the Bayesian decision theory; iii) methods based on the optimization of an objective function; iv) methods based on fuzzy theory; and v) methods based on the use of spatial-context information.

In empirical methods, changes are identified/classified by thresholding a change index. The threshold identification can be performed according to empirical strategies [114] that often employ manual trial-and-error procedures, which significantly affect the reliability and accuracy of the final change-detection map. The basic assumption when applying empirical strategies is that changed pixels are few and show values significantly different from the unchanged ones. Thus, changed pixels are those far from the mode of the density function associated to the change index. A simple strategy consists in fixing the decision threshold as $n\mu + \sigma$, being μ and σ the mean and the standard deviation of the considered change index, respectively, and n is a real number derived by a trial-and-error procedure. In this context, the selection of the parameter n strongly depends on the end-user's subjective criteria, which may lead to unreliable change-detection results. In addition, such a selection usually requires several trials and hence a non-negligible computation time. It is worth noting that these approaches are not fully automatic and objective from an application point of view, as they depend on the user sensitivity in constraint definition with respect to the considered kind of change. These properties may represent a critical limitation.

An interesting alternative consists in formulating the change-detection problem in the framework of the Bayesian decision theory in order to optimize the separation between changed and unchanged pixels in an unsupervised way. The main problem to be solved for the application of the Bayes decision theory consists in the estimation of the statistical terms associated to the classes of change and no-change (i.e., their prior probabilities and probability density functions) [53], [75] without any ground-truth information (i.e., without any training set). The starting point of methodologies based on the Bayesian decision theory is the hypothesis that the statistical distribution of pixels in the change index can be modelled as a mixture of densities. Mixture components are associated to changed and unchanged pixels. In the literature, explicit estimation of class statistical parameters has been addressed with the Expectation-Maximization (EM) algorithm which is an iterative approach to maximum-likelihood (ML) estimation for incomplete data problems [115]. The iterative equations that characterize the EM algorithm are different according to the statistical model adopted for the distributions of the classes. The most suitable statistical model varies according to the kind of data. If optical passive sensor data are considered, the most common statistical models are: i) Gaussian [27], [116]–[118]; ii) mixture of Gaussians [55]; iii) Rayleigh or mixture of Rayleigh (for the magnitude of unchanged samples computed according to CVA) [50], [56]; iv) Rice (for the magnitude of changed samples computed according to CVA) [50], [86]; v) Uniform (for the direction of unchanged samples computed according to CVA) [50]; and vi) Non-uniform (for the direction of changed samples computed according to CVA) [50]. The iterative equations needed for performing EM parameter optimization under the Gaussian, mixture of Gaussian and Generalized Gaussian class models can be found in [75], [78], [80], respectively, whereas more details on the validity of the Rayleigh and Rice models can be found in [50].

Once the statistical parameters are computed, pixel-based or context-based decision rules from the pattern recognition literature can be applied. Here we recall the most widely used approaches in the context of change detection. Concerning pixel-based methods, we can mention: i) Bayes rule for minimum error [50], [75], [80], [116]; ii) Bayes rule for minimum cost [116]; and iii) Neyman-Pearson criterion [116]. The Bayes rule for minimum cost and the Neyman-Pearson criterion allow considering the costs of false and/or missed alarms in the decision process. Bayesian decision theory can be used also in multisensor change

detection [109]. Here integration is carried out according to the consensus theory by integrating the estimates of statistical terms over different sensors. In the integration step, a weight is associated to each source according to its expected reliability. Within the Bayesian decision theory framework, different techniques for reducing the effects of the residual registration noise between multitemporal images have been integrated [18], [20], [119].

Another set of methods is based on the optimization of objective (cost) functions. The fact that, generally, the change index is one-dimensional makes this process easy. The choice of the cost function plays a fundamental role in the accuracy of the results. In the change detection literature, several objective functions have been employed based on: i) discriminant analysis and inter- and intra-class measures [120]; ii) Bayes decision rule for minimum error [121]; and iii) distribution free fuzzy entropy measure [122]. The optimization of objective functions leads to an implicit estimation of the class statistical parameters [121], [123]–[130]. According to the kind of data, different assumptions on the statistical distribution of classes can be made. As an example, the Kittler and Illingworth criterion has been used under both the Gaussian [127] and the Generalized Gaussian [75], [131] assumptions for the statistical distribution of classes. In addition, methods based on Machine Learning and clustering that minimize a cost function can be listed in this category. In the literature examples can be found based on Support Vector Machine [123], [128], clustering and kernel-based clustering [124], [130], [132], and neural networks [125], [133].

The use of fuzzy theory is another possibility. These kinds of techniques rely on the assumption that some ambiguity exists that arises from the overlapping nature of classes or image properties [66], [76]. The ambiguity of an image can be expressed in terms of radiometry (e.g., fuzzy entropy, hybrid entropy, correlation, etc.) or geometry (e.g., compactness, high and width, length and breadth, index of area coverage, degree of adjacency, etc.). The decision threshold is selected as the value where the membership function shows a global minimum or maximum depending on the selected ambiguity measure [66], [76]. Fuzzy clustering approaches belongs to this group as well [71].

Some approaches involve spatial-context information in the decision process. This is justified by the reasonable assumption that changes are large if compared with the spatial resolution of the sensor. Thus, a pixel is likely to be surrounded by pixels of the same class. The use of inter-pixel dependence may yield to more reliable and accurate change-detection results. A fully automatic approach to the unsupervised analysis of the change index, which exploits the spatial contextual information to reduce the effect of noise in the detection procedure, has been proposed in [75], [80]. The solution is developed in the context of Bayesian decision theory, where the spatial context of each pixel is modelled by Markov Random Fields. More advanced methods perform a context-sensitive analysis by considering adaptive neighborhoods modeled by multitemporal parcels, i.e., small homogeneous regions shared by both original images [67], [78]. The adaptive nature of parcels allows spatial-contextual information to be exploited so that noise may be reduced without damaging the boundaries of the changed areas. On the other hand, they better capture the spatial correlation information present in the scene and become particularly promising for VHR images showing complex objects (e.g., buildings and other manmade structures). In order to effectively model objects in the images at different scales, some of the concepts employed in the previously mentioned papers such as morphological filters, multitemporal parcels, and even Markov random fields can be adapted and used in multiscale/multilevel analysis [63], [67], [80], [134], [135], together with specific multiscale/multilevel representation tools such as wavelet transform [70], [96], [136].

2.2. (Semi-)Supervised bi-temporal data analysis

As opposed to unsupervised change detection, a set of multitemporal image techniques can be found that aims at multitemporal analysis by using some ground reference information. Methods belonging to this category mainly rely on supervised or semi/partially-supervised/unsupervised classification. The terms partially supervised classification and partially unsupervised classification have been used in the literature for defining change detection solutions for bi-temporal images where reference data (and thus a training set) are available only for one of the two images. The two terms refer to the same concept considered from the

initial perspective either of supervised or unsupervised classification. It is worth noting that most recently similar problems in the context of multitemporal classification have been defined as domain adaptation in the framework of transfer learning. The term semi-supervised refers in general to the use of labelled and unlabeled data in the learning phase of a classifier. Partially supervised and unsupervised methods (as well as domain adaptation methods) exploit semi-supervised techniques implemented across two domains (associated with the two images). After this clarification, for avoiding confusion, here we refer to all these approaches with the terms which they have been presented in the literature. Such kind of approaches do not only highlight the changes, they explicitly identify the pair of classes (i.e., land-cover transition) associated with each detected change. Note that they can be successfully applied to bi-temporal images both when changes exist or not. However, in the following we concentrate our attention to their use in the context of change detection applications. The (semi)supervised nature of these kind of approaches reduces the sensitivity to radiometric differences. The use of fully or partially supervised methods depends on the availability of ground truth information. On the one hand, if multitemporal ground truth information is available supervised techniques can be applied. This information is used in the learning phase of supervised data classification for modelling the kind of land-cover transitions. On the other hand, if ground truth is available for one or some of the images in the multitemporal sequence, partially-supervised techniques should be considered. If no ground truth is available, unsupervised clustering techniques should be used.

Three main general approaches to integration at decision level can be found in the literature: Post-Classification Comparison [22], Supervised Direct Multidate Classification [22], [137] and Compound Classification [138]–[141]. In the literature, many different classifiers have been used in the context of the analysis of temporal series of remote sensing images. Among the others, we recall the Maximum Likelihood classifier [139], Neural Networks [142], Fuzzy Classifiers [142], Support Vector Machines [143], [144], and Deep Learning [145], [146], which are either the most widely used or the most effective ones. The reader is referred to the literature for more details on the behavior and mathematical details of each single classifier.

The use of supervised classification is in general more accurate than unsupervised approaches. Nevertheless, it is less appealing in operational applications. This is due to the difficulties in collecting proper ground-truth information (necessary for supervised techniques), which is a complex, time consuming and expensive process (in many cases this is not consistent with the application constraints). It becomes even more complex since there is a need of multitemporal information. Semi-supervised approaches represent a trade-off between the two above-mentioned conditions. It is worth noting that all the techniques based on classifiers cited in this section are intrinsically suitable to be used with different kind of data and also with multisensor, multisource and multiresolution information.

The post-classification comparison (PCC) (also referred to as delta classification [22]) is the simplest technique among supervised approaches. It performs change detection by comparing the classification maps obtained by classifying independently the bi-temporal images. For each change, the land-cover transition is obtained in an explicit way. The main advantage of delta classification lies in the fact that multitemporal images are classified independently, thereby minimizing the problem of radiometric calibration. Although PCC has been extensively used in several applications, its performance strongly depends on the classification accuracies of single date classification. After comparison the accuracy is close to the product of the accuracies yielded by the independent classifiers [22], [140], making the approach often unsatisfactory [147]. This is a direct result of the fact that PCC does not take into account the temporal correlation. However, it has been widely employed in the literature both at pixel and region level [148]–[151]. Attempts to increase PCC accuracy have been done by using more than two images in the integration step [152].

Supervised direct multidate classification (DMC) [22], [140], [153], unlike PCC, takes into account the dependence existing between multitemporal images. The main idea is to characterize pixels by stacking the feature vectors related to the images acquired at the two different times. Then the identification of the land-cover transitions is carried out by considering each transition as a single class and by training a classifier to recognize such transitions. It is worth noting that a complex constraint to satisfy for using this technique is

to have a training set composed of training pixels related to the same points on the ground at the two times. In addition, training pixels should represent accurately the proportions of all the transitions in the whole area of interest. This represents a serious drawback as, in real applications, it is difficult to obtain training sets with such characteristics. In [154] it is remarked that, since spectral and temporal features have equal status in the combined data set, they cannot be easily separated in the pattern recognition process. Therefore, class labelling may be difficult if relatively simple classification algorithms are considered. In [153], DMC has been adapted to the use in VHR images.

A more realistic approach is compound classification (CC) [140]. Similarly, to the DMC, also in this case the objective is to perform the classification of pixels of the two images according to the maximization of the posterior joint probability of classes. Conditional probabilities of classes can be estimated according to different techniques and under different assumption on their statistical distribution. On the one hand, with respect to the PCC, the CC technique allows the temporal correlation between images to be considered in the change-detection process. On the other hand, with respect to the DMC method, the CC technique allows the constraints related to the training sets to be relaxed [140]. In particular, training pixels should not necessarily be related to the same area on the ground [138], [140].

In real problems, it may happen that given a series of multi-temporal images, a ground truth is not available for all the items of the series. In such realistic cases, supervised approaches cannot be employed. In [155], an ensemble of non-parametric multitemporal partially-supervised classifiers was defined and integrated in the context of a multiple classifier system. Each multitemporal classifier was developed in the framework of the compound classification decision rule. In [138], a partially supervised methodology was proposed able to update the parameters of an already trained parametric maximum-likelihood (ML) classifier whenever a new image lacking the corresponding ground truth has to be analyzed. The updating is performed by means of the EM algorithm [115] that allows tuning the parameters of the trained ML classifier on the basis of the distribution of the new image. In this way, it is possible to classify multi-temporal data of a given area (and hence to derive land-cover transition maps) without relying on a multi-temporal ground truth. These methods have been recently referred to as Domain Adaptation (DA) methods. In [156], a partially unsupervised technique based on Markov Random Fields is proposed for the identification of the land cover transitions of interest for the end-user only, by exploiting training samples belonging exclusively to the land covers involved in the specific kind of changes to be mapped. In [157], an advanced context-sensitive classification technique that exploits a temporal series of remote sensing images for a regular updating of land-cover maps is proposed. The authors introduced a classifier which is based on an iterative partially supervised algorithm that jointly estimates the class-conditional densities and the prior model for the class labels on the image to be classified by accounting for spatial-context information.

All the aforementioned supervised and partially (semi) supervised methods based on classification are intrinsically suitable to process also multisensor/multisource data. In fact, if proper distribution-free non-parametric classifiers are used for the analysis of the images, data acquired from different sensors (e.g., multispectral images and SAR images) can be processed to produce the map of land-cover transitions. Under the assumption that the considered images are re-sampled at the same geometrical resolution, it is also possible to employ (semi)supervised approaches with different sensors at the two dates: in fact, the comparison process is carried out at a classification-map level. This property is very important as it allows: i) to produce land-cover transition maps related to large temporal differences (large temporal differences involve the availability of data acquired by different sensors of different generation); and ii) to obtain land-cover transition maps also when data acquired from a specific sensor at the first date are not available at the second date (e.g., multispectral images at the second date might be not available depending on atmospheric conditions. In these cases, SAR images could be compared with multispectral images). In [158] the authors describe both neuro-fuzzy and statistical approaches to the exploitation of the contextual information and the classification, and different schemes for the multisensor fusion. The presented technique fuses the information of active and passive sensors and results in a good change-detection precision and in the best possible classification accuracy.

Due to the complexity in constructing ground truth information for the training of classifiers, integration at decision level methods have been less used. Recently a novel interest from the scientific community was devoted to these methods because of the methodological developments in the context of domain adaptation (DA) and active learning (AL) context. As mentioned above, DA approaches allow to take advantage of the ground truth information available for one acquisition (i.e., the source domain) and to adapt it to images for which ground truth is not available (i.e., the target domain). The adaptation mechanism can be significantly improved by AL approaches that guarantee a minimum amount of new labelled data for the target domain. Thus, they handle in a better way the possible significant differences between statistical distributions of the source and target domains. Examples of such approaches to multitemporal image classification can be found in [159]–[161]. However, in this thesis we focus our attention on unsupervised methods only. This is because of two main reasons: i) the complexity given by managing multisensor VHR images and; ii) the lack of any ground truth for applying a (semi)-supervised method.

2.3. Time Series analysis

In this section, we move from the extraction of change information from bi-temporal images to the analysis of long time series trends. Where the same block scheme presented in Figure 1.3.b., is followed, and derivation of a trend-index is required. Time series represent a huge information source for Earth Observation (EO) purposes. The extraction of information on the temporal behaviors of imaged ground areas requires the definition of ad hoc features being able to capture land-cover trends over time. The starting point of such kind of analysis is the temporal signature of pixels. Each element of this signature corresponds to a spectral reflectance value. From the temporal signature, several features can be extracted to highlight temporal behaviors. Features can be computed on each single element in the temporal signature (i.e., for each image) in order to generate new temporal signatures associated to the time evolution of the specific features. Other features can be computed by involving all (or some) elements of the temporal signature in the calculation. This allows one to capture the time periodicity of events.

When dealing with time series of optical images, the temporal signature can be built based on indices such as the original reflectance, or more complex features such as NDVI [162]–[166]. More complex indices can be designed to model other physical phenomena. For example temporal statistics, i.e., statistics computed along time rather than in the spatial domain (e.g., mean, standard deviation, normalized standard deviation, saturation, maximum-minimum ratio) [167] or by arithmetic operations (such as the difference) applied to the spectral signature or other features of subsequent pairs of images [166]. They can be obtained by involving both spatial and temporal information, as it happens for multitemporal filters or isochronous information [40].

More complex features can be obtained by applying data transformation to spectral reflectance or to one among the above mentioned features. These transforms are applied along the time dimension and include Principal Component Analysis (PCA) [168], Fourier Transform [169] and Wavelet Transform [170], empirical mode decomposition and Hilbert spectrum [171]. Except PCA, the mentioned methods locate specific behaviors in the frequency spectrum. The first two methods mainly locate such components that show a time periodicity, whereas the last two characterize non-stationary behavior of time-series in the frequency domain. Such features capture the temporal behaviors of specific phenomena, and have been employed for analyzing trends and seasonal changes, and separate them from other kind of changes [23], [38]. Available works mainly concentrate on specific applications such as vegetation [172], floods [173], forest fires [174], etc. Methods rely on state-of-the-art classifiers and/or information extraction tools sometimes only readapted and not specifically developed for handling the time dimension. Methods are often only semi-automatic. Moreover, they deal with dense time series of low spatial resolution images (e.g., MODIS time series) or with less-dense time series of high resolution images (e.g., Landsat, SPOT). New generation sensors, such as the S2, acquire dense time series with high spatial resolution and in several spectral bands, allowing for new applications and introducing new challenges that require development or adaptation of new approaches to account for these kind of characteristics [60], [175]–[178].

In addition to the exploitation of both the temporal inter-correlation among images in the same series and the temporal correlation between two series, in practical applications sub-optimal solutions can be considered. Semantic products can be obtained over long series by combining the results obtained by techniques based on the processing of pairs of temporal images. At application level this strategy allows a theoretical (and computational) simplification of the problem of time series analysis at the cost of sub-optimal (but often acceptable at an operative perspective) solutions. In this context, it is possible to combine the behaviours of:

- Binary change detection maps obtained by applying unsupervised change detection techniques to neighbouring couple of images for identifying the temporal trend of abrupt (or relatively "fast") changes (e.g., deforestation) [179];
- Binary change detection maps obtained with different kinds of sensors (i.e., active or passive) for a joint analysis of different effects related to the same phenomenon (e.g., tsunami and earthquakes damage assessment);
- Maps of land-cover transitions obtained by supervised, semi-supervised or unsupervised techniques applied to neighbouring couple of images for identifying temporal patterns related to land-use and/or land cover and analyse their temporal dynamic (e.g., crop rotation, weather changes that can indirectly cause avalanches, desertification, etc) [180], [181];
- Phenomena inherent results achieved by applying techniques for trend analysis to pairs of temporal series extracted from a long set of temporal series for identifying the annual trend of changes (e.g., vegetation health monitoring, desertification monitoring, glaciers monitoring);
- Specific phenomena by integrating the products of multitemporal analysis on pairs of neighbouring images with other ancillary data and prior knowledge (e.g., risk assessment and analysis);
- Short term trend analyses together with long term trend analysis (and study of past events) to perform more robust and reliable short-term forecasts of catastrophic events and hazard maps (e.g., volcanoes eruption, tsunamis and avalanches monitoring and prediction, activity planning in risky areas);

In the scientific literature, it is possible to find many examples of these derivate products that can be obtained by integrating the dynamic change information extracted from remote sensing images with prior information and other data coming from different information sources [168], [182]–[185]. Many other approaches that aim at analysing the available time series, from the temporal perspective, are also available in the literature. However, some constraints arising from the technological and thus methodological evolution, result in the lack of regular data in time and in periodicity. Therefore, attention has first been oriented to the construction of regular and dense time series by means of methods based on gap filling or curve fitting. And then to the extraction of temporal information and derivation of relevant products (according to the application itself).Once regular and dense time series are available, analysis can be carried on both at pixel and context based levels.

2.3.1 Building regular/dense time series

With almost 40 years of data from the National Oceanic and Atmospheric Advanced Very High Resolution Radiometer (NOAA AVHRR), more than 15 years of data from the Moderate Resolution Imaging Spectroradiometer (MODIS), more than 40 years of data from Landsat program, nearly 3 years of data from Sentinel-1 and Sentinel-2, and many new time-series products being developed, there is a need for efficient and practical methods for handling these data. However, given the large influence of noise on the satellite-derived measurements, the direct use of the data is often difficult. Cloud conditions and hazy atmosphere are the primary causes of noise in optical satellite sensor data, although other factors (e.g., directional effects, geometric inaccuracies, and sensor disturbances) also contribute to the noise in the final products [186]. While several highly processed data sets exist, they are usually far from being noise free. This noise

severely affects the possibility to estimate accurate land surface parameters, particularly during the rainy parts of the season (which is year-round in many humid tropical areas), leading to the availability of incomplete or irregular time series. Before the launch of S2 satellite, there were two main situations where the time series could be considered as irregular:

- When no image exists for a given time: This situation arises from the satellite acquisition schema, which can lead to periods of time where no image was taken, because of satellite programming constraints or because cloud percentage is that high to make images useless. In this case, the problem is to build an intermediate image that should be a good approximation of what could have been acquired. The interest of this approximation resides in the ability to provide a regularly sampled time series for trend analysis.
- When pixels are unknown in an image at a given time sample: This situation arises when clouds partially occlude the scene, or when there have been acquisition problems on some lines or columns of the image, or because the footprint of the image does not cover pixels covered by other images in the time series. There is an interest of finding a good approximate of these missing pixels to fill the holes in the image.

Nevertheless, S2 sensor possess a particular orbiting characteristic that, combined with its high spatial resolution, adds an extra situation in which irregular time series could be acquired. In order to cope for possible missing pixels in the proximity of observations acquired by different orbits, but over continuous physical/geographical areas, an overlap of those orbits was introduced in S2 [175]. This similar concept exists for other sensors such as Landsat, but the overlap is much smaller than for S2. This plus the S2 high spatial resolution results in the possibility of mapping some areas: i) with double frequency than initially expected and, in turn, ii) at an irregular frequency. Such situation does not always occur, but introduces a new definition to the concept of irregularity that did not exist in literature before. The closest irregularity situation to that of S2 is the one where no image exists for a given time.

In this subsection we concentrate our attention on state-of-the-art-methods for mitigating the situations where no images is usable due to conditions such as the atmospheric ones. Early methods are relatively simple, e.g., the use of maximum-value compositing or Moving Average Window (MAW) [187], [188]. These method relies on the ability of the NDVI to minimize noise when it assumes the highest value during a specific time-period. The principle of selecting the highest value is also the basis for the Best Index Slope Extraction (BISE) [189]. Though simple and generally effective in removing many disturbances, these methods ignore the fact that also positively biased noise may be present in the NDVI series, and they are limited to NDVI or similar vegetation indices in which noise is negatively biased. For more general noise reduction, several mathematical filters, which use either temporal or spatial information, have been proposed for remotely sensed data and produce temporally smoothed and spatially continuous products [43], [186], [190], [191].

Spatial filters using pixel-level or regional ecosystem statistical data include geostatistical and regression methods [192]–[194]. Nevertheless, spatial filters may fail in the case of physical products (e.g., Leaf Area Index –LAI) derived from coarse resolution satellites to represent the complexity of real landscapes mainly over mixed pixels where the pixel value may vary within a short distance widely. To overcome this limitation, some studies tried to combine both temporal and spatial information by using historical high-quality data and temporal curves from neighboring pixels. Fang et al. [195] proposed a temporal spatial filtering algorithm for MODIS LAI data applying an improved ecosystem curve fitting method based on the MODIS vegetation continuous fields product. This product imposes regional dependent phenological behavior onto each target pixel temporal data in order to maintain pixel-level spatial and temporal integrity. Gao et al. [196] proposed using preferentially a seasonal-variation curve within a small window around the pixel of the MODIS LAI product and an ancillary seasonal curve within the tile when high-quality data are unavailable within the defined maximum distance.

Temporal filters include a broad variety of strategies such as the well-known Maximum Value Compositing (MVC) [187], BISE [189], Fourier-based fitting methods [185], wavelet decomposition [197], asymmetric Gaussian filtering [198], Savitzky-Golay filtering [199], logistic function fitting [200] or curve-fitting procedure [38]. The choice of the smoothing gap filling or compositing method may have a large impact on the accuracy of the phenology extracted from the reconstructed time series [37], [201]. In the case of Fourier series, the periodic annual fluctuations in data is utilized [185], [191], [202], [203]. Though the parameters of these harmonic functions contain information about the timing of the seasons, the method is generally too inflexible for remotely sensed time series data, in which the timing of seasons can vary considerably between different years. An example is shown in Figure 2.5, where the Harmonic ANalysis of Times Series (HANTS) method proposed by Roerink et al. [185] is applied over a S2 NDVI time series in order to reconstruct a continuous time series. HANTS uses the predefined most significant frequencies expected in a time series profile and harmonic components to apply a least-squares function-fitting procedure to the data. Unlike standard Fourier analyses, this technique does not require observations to be equidistant in time, and is therefore more flexible in its application, but it involves the setting of several parameters and more time for processing, resulting in a non-fully suitable method.

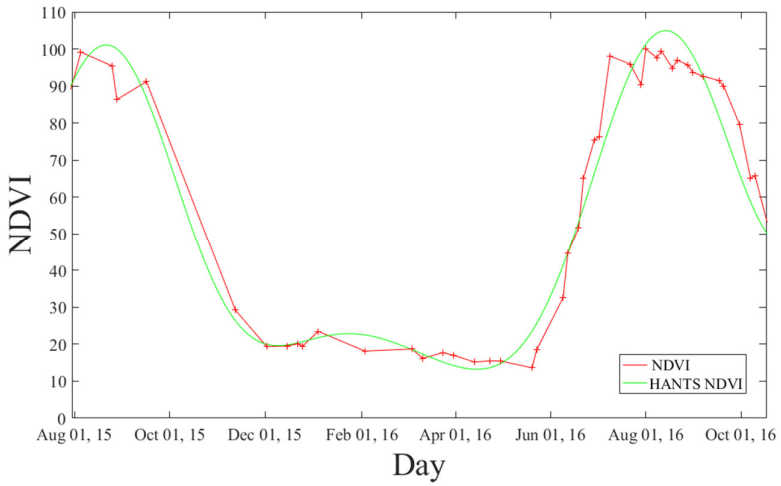


Figure 2.5. NDVI and NDVI HANTS reconstruction for a single crop field in a Sentinel-2 time series.

More recent methods are generally based on fitting mathematical functions to the time-series data. Quantitative comparisons of alternative temporal filters are relatively rare. Chen et al. [190] showed the effectiveness of a modified Savitzky–Golay filter in comparison to the BISE algorithm and fast Fourier transform technique for reconstructing SPOT VEGETATION high-quality NDVI time-series. Later, Hird and McDermid [201] revealed the general superiority of the Beck et al. [204] double logistic and Jönsson and Eklundh [198] asymmetric Gaussian function fitting methods over more simple local filtering methods. It is however still difficult to identify the potentials and limitations associated with different methods since most of these studies focus on a small sample of global conditions and they have been applied to NDVI rather than on a true biophysical variable such as LAI, with few exceptions [195], [196], [205], [206]. Additional methods exist that combine high temporal/low spatial resolution data (e.g., MODIS) with low temporal/high spatial resolution (e.g., Landsat) in order to integrate/obtain time series that account for high temporal/high spatial resolution sequences [207]–[211]. Both approaches to filling of missing data allow to build higher temporal frequency time series with temporal limitations given by the highest temporal resolution of the employed images. Still the high temporal resolution time series suffers from atmospheric conditions issues.

2.3.2 Land Cover Monitoring

In land-cover monitoring, trend detection and forecasting are of special interest. For this kind of monitoring, a long-time span of data is needed and the number of observations must be large and dense enough in accordance with the kind of phenomena to be monitored and the targeted area. Land-cover monitoring requires the comparison of temporal development curves, also called time trajectories or time profiles, of relevant indicators. This is used as a reference for successive growing seasons, for agricultural areas monitoring, or for successive years for forest or vegetation monitoring in general. The inherent high temporal frequency in data acquisition not only expedites the detection of ecosystem modifications, but also facilitates the characterization of phenological variations in ecosystem status. If we consider a given pixel in a temporal series, we can observe the temporal evolution of its radiometry information along the time. These evolutions are specific of the phenomena that affects the underlying object and can be used to characterize it. The example given in Figure 2.6, shows the mean NDVI value of a single crop field in an agricultural area over 16 months in a time series built by using S2 images (time series acquired from July 5th, 2015 to November 30th, 2016). NDVI values have been scaled by 100 for visualization clarity. Starting on the left from early-October 2015, vegetation starts to grow until it reaches its maximum value by mid-March. During the beginning of October 2016, a new increase of the vegetation can be seen, but with a faster speed. Such a difference in the growing speed can be an indicator of many things, among them: the environmental conditions, the plantation of a new crop type, the last grow of remaining vegetation, etc.

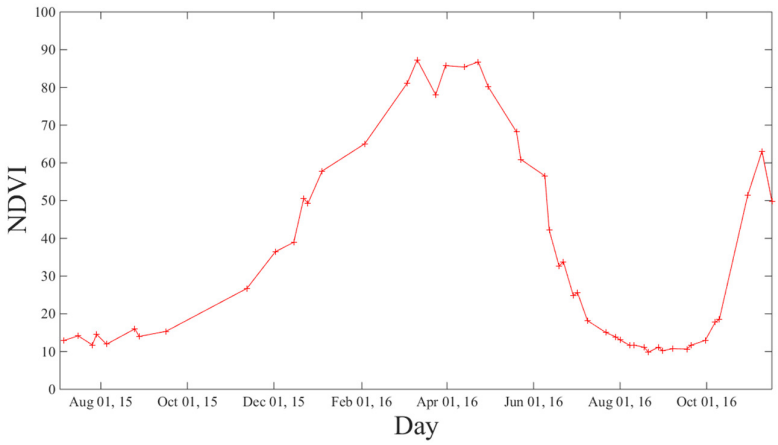


Figure 2.6. Mean NDVI temporal behavior for a single crop field in a S2 time series. Vertical axis corresponds to 100 times the NDVI value.

Such kind of curves can be used to extract several information. However, this interpretation can be done when the nature of the underlying object in the scene is known. Temporal signature analysis is affected by the possible presence of temporal outliers, like a cloud or haze. These phenomena behave as randomly spread noise on the temporal behavior of a pixel/crop and could lead to the impossibility to interpret its behavior correctly over time. These effects affect time series even after temporal frequency improvement (section 2.3.1). Therefore, effective models should be designed able to model the behavior of time series in a small number of parameters, thus filtering the remaining outliers. Such parameters should be rich enough to preserve the essence of the temporal behavior. Such a model should be able to:

- Adapt itself to the different object behaviors in the scene. Objects obey different evolution laws. e.g., i) vegetated pixels will be subject to weather conditions like sun exposure, rain, dryness and also to seasonal variations; and be different depending on the species/land cover; ii) water pixels will be affected by weather conditions too, but in a very different way than the vegetated pixels: after a heavy rainfall, the rivers turn brighter, and if there is wind, the surface wavelets may impact

on the reflectance or backscattering; and iii) urban pixels will be less subject to weather conditions, but will depend on societal conditions like construction, destruction, growing, among others.

- Avoid the possible temporal outliers like clouds which may be present in some parts of the scene at some moment. This condition can be fulfilled by detecting such outliers and by discarding them, or even by interpolating the corrupted samples with the methods shown in Section 2.3.1.
- Represent the pixel evolution in a small number of coefficients, in order to make a compact description of the behaviors.

A first attempt to extract all the characteristics of interest, while considering most of the problems mentioned above, was made by Jönsson and Eklundh in [198]. The whole method was integrated in a software called TIMESAT and is based on fitting asymmetric Gaussian functions to time series. The fitted functions were used to portray the seasonal growth and decline of curves of the land vegetation, and to estimate phenological parameters. In the newest version of TIMESAT [29], the function fitting can be achieved by means of Savitsky-Golay filter, asymmetric Gaussian or double logistic smooth functions. Three steps are followed to extract the parameters of interest: i) the number of seasons and their approximate timing is defined; ii) data are filtered or fitted by smooth functions; and iii) the seasonality parameters are computed and written to output files. In order to get the seasonality parameters, NDVI time series coming from AVHRR or MODIS are considered. Nevertheless, the program does not fully account for new generation sensors, which spatial resolution is higher. Neither does it account for irregularly sampled time series. This leads to the need of new adaptations or new methods to analyze such data.

2.3.3 Trend Analysis

Two general approaches to trend analysis can be identified in the literature: i) pixel based; and ii) context-based. As for bi-temporal change detection, also for trend analysis, standardized PCA of time series can be used. In particular, it has proved to be a powerful technique to separate changes that took place at different time frequencies, e.g., decadal time-scale changes in productivity, seasonal changes or sensor-related value drifts [168]. In [212], a neural-network based methodology was developed for the detection of burned areas from time series of SPOT-VGT data. In particular, the use of a Multi-Layer Perceptron (MLP) neural network allows one to exploit not only the spectral information of the observed targets, but also the spatial and temporal relationship of the phenomenon. Also the Vegetation Cover Conversion (VCC) product of MODIS detects changes in the land-cover by analyzing the movement and trajectory of pixels within a time series [213]. In particular, the red and near-infrared channels are considered. In [214], a preliminary work for time series analysis based on the use of the Kalman filter [215] was proposed. In [216], long-term change-detection was addressed by using a regression technique based on a tree algorithm. In [217], standardized principal components analysis (SPCA) was performed on a NDVI time series to monitor global vegetation dynamics. NDVI was used also in [218] and [219] on a time series of three images, change-detection was performed by manually labeling classes after clustering and thresholding the three NDVI images according to methods based on statistical models. In [220] and [221] authors proposed a method that considers a number of time series made for instance of subsequent n pixels. The method can detect linear relations between the pixel values. NDVI has also been used together with albedo [222]–[224] to analyze long-term trends in vegetation productivity (integrated NDVI) and albedo (mean anniversaries annual values from the same period).

Attempts to compare or contrast two times series are reported in [225] and [226]. In [225], the image data were compared with the output of a temperature accumulation model to identify critical plant phenological parameters driven by temperature. In [226], a number of local plant growth models were compared with image data to investigate the extendibility of the models to larger areas. Other works compare two or more time series in order to consider the natural behavior of land covers. Such behaviors result in varying yearly time series, because of the high land cover variability. Therefore, methods have been developed in order to cluster similar temporal behaviors of pixels or objects: e.g., K-means clustering with Euclidean distance

[227], [228] or dynamic time warping (DTW) [181]. The latter is a more suitable measure for time series similarity since DTW measures the cost of alignment between two temporal sequences considering nonlinear warping in the time domain.

One of the most widely studied application of trend analysis is the seasonal development of vegetation canopies in medium and coarse resolution optical data such as NOAA/AVHRR data (e.g., [229]–[232]), SPOT/VGT data (e.g., [233], [234]), and MODIS data [204]. In addition, Landsat data has been also used [235]. Many of these studies concentrated on precise timing of spring greening up, progress of autumn colors and the snowmelt and springtime greening up. Figure 2.7 shows an example of the snow cover evolution in the Trentino region (North of Italy) from 2006-2017 [236], [237]. The method takes advantage of the multiple resolution bands of MODIS and integrates them by multi-resolution fusion in order to have 250 m resolution images. A SVM classifier is used in conjunction with ancillary data and indices such as NDVI and Normalized Difference Snow Index (NDSI) for calculating snow cover. A comparison with the historical average snow cover (red line) is carried on over the time series to understand the snow cover trend.

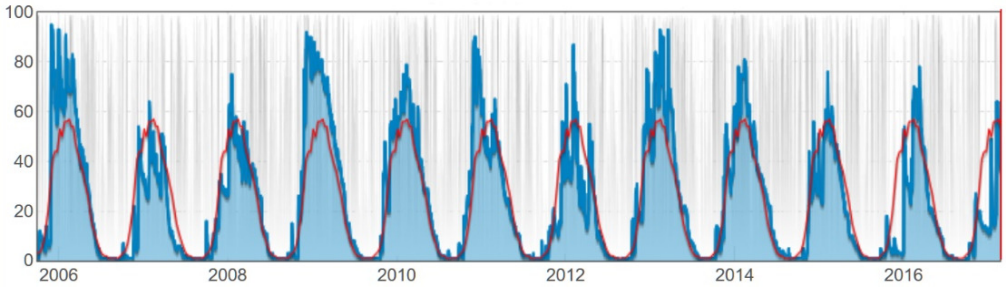


Figure 2.7. Evolution of the snow cover in the Trentino region (North of Italy) from 2006 until 2017. The information is generated by means of MODIS data [237].

Although NDVI is perhaps the most widely recognized optically-based vegetation index, other indices can be linked with crop/vegetation condition, and offer some advantages over NDVI. The Soil Adjusted Vegetation Index (SAVI) is of interest for multitemporal monitoring of crop condition as this index is linked directly with LAI [238], [239]. SAVI has an advantage over other optical indices like NDVI since SAVI minimizes soil effects. Other applications can be also found in the literature, i.e., urban expansion monitoring [170], [183], [240], Land Surface Temperature (LST) and Lake Surface Temperature Monitoring (LSWT) [241]–[245], among others.

Spatial context information has been also considered in trend analysis. In [40] a method was proposed where a pixel is represented with a set of complex features that captures temporal behaviors with both pixel- and spatial-context-based information. The set of features includes: i) pixel-based features such as spectral information; ii) spatial-context features that capture the information from neighboring pixels like probability of belonging to a given class based on kriging indicator; and iii) temporal correlation information like class conditional transition probabilities. The integration of this complex information results in an accurate change-detection map able to identify: when, where, and what land-cover changes occurred. An alternate approach, which also considers spatio-context information, is presented in [246]. The unsupervised change-detection method is based on a three-dimensional clustering performed by means of segmentation technique. Any type of temporal change is detected as a boundary perpendicular to the time axis and sequential application of a supervised algorithm provides an efficient and accurate change analysis. In [184] a technique is presented that compares or contrasts two time series considering spatial correlation between neighboring pixels. Complex spatio–temporal patterns present in two long time series of data are compared for studying the seasonal cycles of vegetation in a large part of the global land surface.

Chapter 3

3. An Approach for Unsupervised Change Detection in Multitemporal VHR Images Acquired by Different Multispectral Sensors[2]

This chapter proposes an approach for the detection of changes in multitemporal VHR optical images acquired by different multispectral sensors. The proposed approach, which is inspired by a recent framework developed to support the design of change-detection systems for single-sensor VHR remote sensing images, addresses and integrates in the general approach a strategy to effectively deal with multisensor information, i.e., to perform change detection between VHR images acquired by different multispectral sensors at two dates. This is achieved by the definition of procedures for the homogenization of radiometric, spectral and geometric image properties. These procedures map images into a common feature space where the information acquired by different multispectral sensors becomes comparable across time. Although the approach is general, here we optimize it for the detection of changes in vegetation and urban areas by employing features based on linear transformations (Tasseled Caps and Orthogonal Equations), which are shown to be effective for representing the multisensor information in a homogeneous physical way irrespectively of the considered sensor. Experiments on multitemporal images acquired by different VHR satellite systems (i.e., QuickBird, WorldView-2 and GeoEye-1) confirm the effectiveness of the proposed approach.

3.1. Introduction

The use of Remote Sensing (RS) in the analysis and evaluation of environmental processes evolution is a valuable tool which relevance increased in conjunction with the use of digital image processing techniques. Due to the improvement of both acquisition sensor technology and data processing algorithms, it is possible to get an accurate and automatic identification and extraction of features for understanding the environmental changes occurring on the ground due to natural and anthropic interactions. The technological evolution resulted in the availability of multitemporal and multispectral satellite images with Very High spatial Resolution (VHR) acquired by passive sensors (e.g., QuickBird, WorlView-2, GeoEye, Pleiades). These images allow a detailed geometrical analysis when compared to medium or high spatial resolution data [31]–[34]. When considering VHR satellite systems, it is difficult to define Time Series (TS) made of images from one single sensor that satisfy the application temporal resolution constraints and show homogeneous acquisition conditions characteristics (e.g., similar light conditions, similar acquisition angle). This is mainly due to the satellite revisit period, the possible competing orders of different users on the satellite pointing, the limited life of a satellite mission, and weather conditions. However, a considerable number of satellites have been launched in the last decades, thus, to mitigate the abovementioned limitations, it is possible to build TS by considering images acquired by different multispectral VHR sensors.

The definition of multisensor TS increases the probability to have sequences of multitemporal images with a proper time sampling but poses some challenges. In addition to real changes occurred on the ground, multisensor multitemporal images are affected by differences induced by the acquisition conditions (e.g., atmospheric conditions and acquisition system). On the one hand, some of the differences in atmospheric

[2] Part of this chapter appears in:
Journal paper: Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "An approach for unsupervised change detection in multitemporal VHR images acquired by different multispectral sensors," under revision in *Remote Sensing, MDPI*.
Conference paper: Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "Change detection in Very High Resolution multisensor optical images," 2014, vol. 9244, pp. 924410-924410–13.

conditions (e.g., cloud cover), and the differences in acquisition system (e.g., view angle and seasonal effects) affect single-sensor multitemporal image processing as well [30]. On the other hand, multisensor TS poses the big challenge of having system induced differences due to the type of sensor and the sensor acquisition modes. They result in spectral and geometric differences. Such differences make change detection state-of-the-art methods less effective since they usually assume that multitemporal images are acquired by the same sensor and under similar acquisition conditions. Similar observations hold for the use of quantities like Digital Numbers (DN) that become non-comparable, thus their use may increase Change Detection (CD) error. Moreover, standard methods for converting DN in reflectance can be insufficient when using images acquired by different sensors [39].

An appropriate multitemporal image homogenization is therefore required to reduce both spectral and geometrical differences, and to ensure that differences in the multitemporal images can be associated to real changes occurred on the ground. In the literature, multitemporal image homogenization methods are available for medium and low geometrical resolution data [247]–[252]. However, their adaptation to VHR images is still an open issue because of the higher within-class spectral variability induced by the very high spatial resolution, when compared to lower spatial resolution sensors. In [49] and [36], authors made a first attempt to adapt existing methods to radiometric homogenization of a pair of VHR images (IKONOS and QuickBird). Spatial resolution differences were mitigated by resampling to the images with the lowest spatial resolution [36], [49]. Other works transform DNs to physical information [39], [47], [48]. Pacifici *et al.* [39] showed the importance of working with physical quantities that are homogenous across time when using VHR optical images acquired by the same sensor. They pointed out the improvement achieved while increasing the level of information abstraction in the feature space, and its consequence for the final image homogenization. However, despite some work on adaptation of homogenization methods has been done in the literature, and as per the best knowledge of the authors, no formal approach exists that properly accounts for the complexity of data acquired by similar, yet different VHR multispectral sensors [47], [48], [253], [254], and guide the user from data pre-processing until the final CD process.

In the literature, methods are available for multitemporal VHR optical images information extraction in the context of CD [30], [255]–[257]. Both supervised and unsupervised CD techniques have been widely used in several RS applications (e.g., flood detection, damage assessment, environmental monitoring). The main drawback of supervised methods lies in the need of collecting and constructing ground reference data for the system-training phase. However, they lead in general to more accurate results. On the other hand, unsupervised techniques have the advantage of not requiring any ground reference information. This mainly results in a faster and more operational processing, but lower accuracy in part due to the use of non-homogenous features. Accordingly, in the literature a large attention has been devoted to improve the accuracy of unsupervised approaches [30], [39], [50], [256]–[258]. Among CD techniques, the most widely used are the ones based on Principal Component Analysis (PCA), Iteratively Reweighted MAD and Change Vector Analysis (CVA) [52], [90], [256], [259].

CVA has been widely applied to the original spectral bands of multitemporal images from low to very high spatial resolution [50], [258], [260]–[262]. Other authors used CVA on the vegetation index or the Tasseled Cap (TC) feature space, but with low and medium spatial resolution images only, and/or for the detection of specific/single vegetation changes [261], [262]. In the case of VHR images, original spectral bands are traditionally used, and the CD mainly focuses on the separation of change and no change classes, or the identification of a single kind of change, without considering the nature (kind) of the changes [30], [100], [102]–[104]. In low, medium, high and VHR cases, features are mostly selected according to the possible changes occurred on the ground. But less attention is devoted to explicitly mitigate the differences induced by the acquisition system, both from the homogenization and feature selection perspectives [100], [102]–[104], which in turn would lead to an improvement on the CD accuracy. In the specific case of Iteratively Reweighted MAD, little attention is devoted to the possibility of using these features for CD in

multisensor images [102]. Mostly qualitative visual analysis about the possibility to detect changes is conducted, but no quantitative change information extraction approach is provided [90]. In this context, the need arises of defining proper techniques and operational strategies for homogenization of multitemporal images acquired by different VHR multispectral sensors that mitigates differences in both atmospheric conditions and acquisition system parameters, thus reducing their impact on the multitemporal information extraction, and on the CD accuracy. The technique should also be able to extract changes automatically and distinguish among different kinds.

This chapter presents an approach for unsupervised CD in multitemporal VHR images acquired by different multispectral sensors, which is inspired from the framework for the design of CD systems for VHR images presented in [30]. The proposed approach exploits some of the concepts in [30] and extends and integrates them with a strategy for mitigating the effects of the non-homogeneous properties of multitemporal images acquired by different VHR multispectral satellite systems at both pre-processing and feature extraction/change detection level. The main steps of the proposed approach are: i) mitigation of differences induced by the use of VHR multitemporal images acquired by different sensors; and ii) detection of multiple changes occurred on the ground by means of high level physical features. The first step is conducted by defining homogenization procedures that address radiometric, spectral and geometrical differences. Thus, multitemporal multisensor images become more comparable (i.e., more homogeneous) across time. Homogenization is further improved by extracting proper features from multisensor images that allow for an effective multitemporal comparison across sensors at a given level of abstraction. Features are designed to detect multiple changes relevant to the user. Although the approach is general, here we concentrate on the selection of high level features suitable to detect changes in vegetation and urban areas. Linear/Orthogonal transformation features are selected with no loss of general validity for the detection step. They are shown to be effective for representing the multisensor information in a coherent physical way versus the considered sensor, but other features can be considered as well. The second step is conducted by means of CVA in spherical coordinates, where the changes are represented by magnitude and direction variables. Separation among the changes is carried out in an automatic way. In the case of the magnitude variable, changed and non-changed pixels are separated by means of a Bayesian decision rule [44]. Whereas along direction variables, an adaptation of the Two-Stage Multithreshold Otsu (TSMO) method [45] is used. Experiments carried out on multitemporal VHR image pairs acquired by different sensors confirm the effectiveness of the proposed approach.

The remainder of this chapter is structured as follows. Section 3.2 presents an overview and an analysis of the properties of a CD system for VHR RS images. Section 3.3 illustrates the proposed approach for the mitigation of differences induced by the use of VHR multisensor multitemporal images and the CD process. Section 3.4 introduces the multisensor datasets, describes the design of experiments and illustrates the experimental results. Finally, Section 3.5 draws the conclusions and future works.

3.2. CD Systems for VHR Remote Sensing Images

When considering VHR RS images, the CD problem becomes complex due to the heterogeneous spatial and spectral characteristics. Further, standard CD techniques often do not account for the semantic meaning of changes of interest. Standard CD methods often assume that unchanged pixels have similar signatures at the two dates, whereas changed ones do not. Unfortunately, this assumption is often not satisfied when we consider multitemporal VHR images, since additional differences may appear due to spectral and spatial heterogeneity. This becomes more critical when multisensor images are considered [20], [30]. In accordance with [30], a proper definition and modeling of the concept of change is fundamental for the development of effective techniques that can mitigate the intrinsic differences in multitemporal VHR data, and accurately detect multiple changes. Further, most of the current methods for CD in VHR images focus on: i) handling images acquired by the same sensor [263], [264], or ii) detecting specific kind of changes by using multisensor images (i.e., deforestation, burned areas, buildings detection) [100], [102]–[104]. Thus,

their applicability to CD in multisensor image pairs is limited. Thus, their applicability to CD in multisensor image pairs is limited.

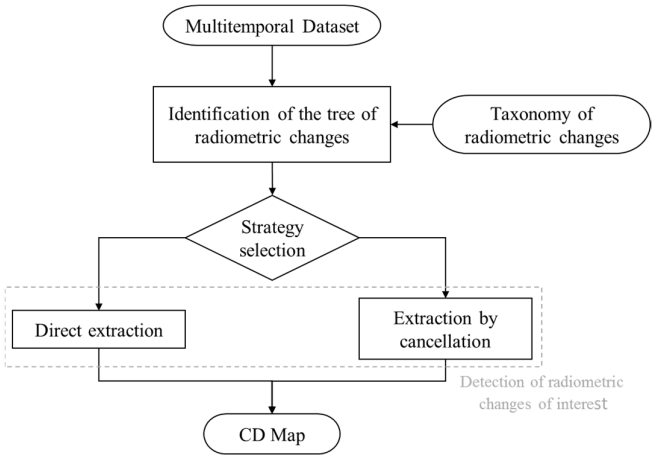


Figure 3.1. Flowchart of the procedure for the definition of novel CD methods [30].

In this chapter, we aim at developing a CD approach for multisensor VHR optical images. To this purpose, the framework in [30] is used as a baseline. Figure 3.1 presents the general flowchart proposed in [30] that consists of two main steps: i) definition of the tree of radiometric changes and; ii) detection of changes, which can be conducted by two different strategies. In the first step, possible classes of radiometric changes, are analyzed and their taxonomy is defined. The resulting tree of radiometric changes is specific for the considered CD problem. To this end, a categorization of the different possible radiometric changes that may be present in a multitemporal VHR dataset is required.

Figure 3.2 shows a tree that models radiometric changes for a generic CD problem in multitemporal multisensor VHR images. According to [30], two main kinds of radiometric changes (Ω_{Rad}) originate because of the complexity of VHR images: i) changes due to acquisition conditions (Ω_{Acq}) and; ii) changes occurred on the ground (Ω_{Grd}). The former corresponds usually to changes of no interest for the end user. The latter includes the changes relevant from the user's viewpoint. Ω_{Acq} changes are the ones associated to differences in atmospheric conditions (Ω_{Atm}) and in the acquisition system (Ω_{Sys}). The latter results in the appearance of undesired change patterns, differences in the geometry and in shadows, and thus errors in the final CD map. Ω_{Sys} changes, such as the ones due to the type of sensor (Ω_{Sen}), can be mitigated by working with proper homogeneous features.

Among Ω_{Sen}, we find changes due to differences in the spectral resolution (Ω_{Spe}) and spatial resolution (Ω_{Spa}). Ω_{Sys} includes changes due to the sensor view angle (Ω_{Ang}) or the solar rays incidence angle (Ω_{Sun}). Ω_{Ang} can be related to the topography (Ω_{Top}) and the relief (Ω_{Rel}). For example, they may generate changes in shadows. Whereas Ω_{Sun} refers to effects such the ones induced by seasonal variations of the solar ray incidence angle, which generates shadows differences that are not associated to changes on the ground. On the other hand, Ω_{Grd} changes can be divided into four main categories: natural disasters (Ω_{Dis} – e.g., earthquake damages), vegetation phenology (Ω_{Veg}- e.g., leaves lost during winter), environmental conditions (Ω_{Env} – e.g., variation in soil moisture levels) and anthropogenic activities (Ω_{Ant} – e.g., harvested crop fields, new buildings). The tree structure illustrated in Figure 3.2 has a wide validity and fits to most of the CD problems. However, depending on the specific CD problem, the tree can be adapted and optimized. On the one hand, some leaves/nodes might be irrelevant and thus can be removed. On the other hand, some leaves may require to be further specified in accordance with the specific study case (see Sec. 3.4 for an example of how to define the tree of radiometric changes for a concrete CD problem). To extract the changes of interest, it is necessary to select effective features and to count on prior knowledge about the

problem.

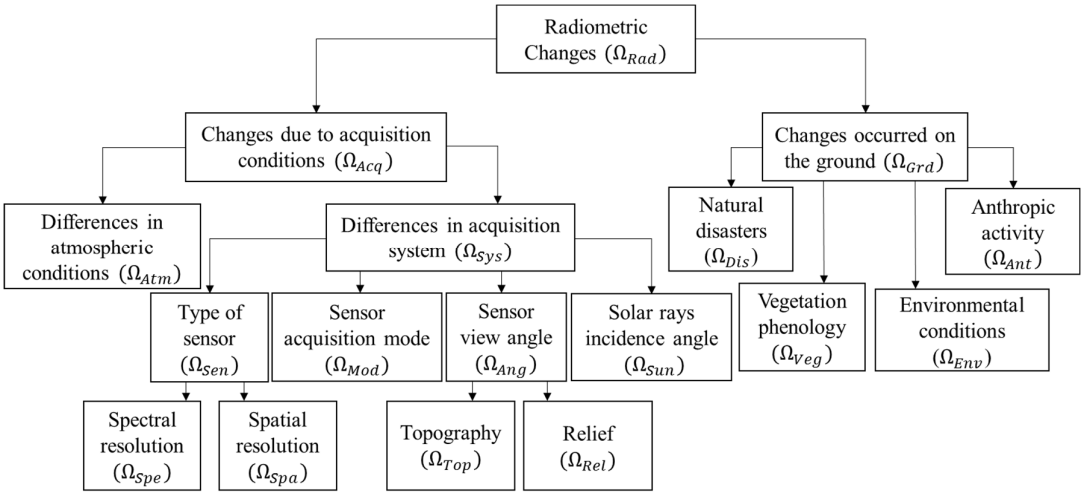


Figure 3.2. General tree of radiometric changes for the multisensor VHR images case [30].

The second step detects the kind of changes identified in the previous step, by selecting the strategy for the design of the CD method. According to Figure 3.1, two possible strategies can be adopted based on: i) direct extraction of the radiometric changes of interest, or ii) detection by cancellation of non-interesting radiometric changes. Most of the current available CD methods for VHR RS images make use of the direct extraction strategy since their goal is to extract a specific kind of change. Nevertheless, sometimes it is easier to detect the radiometric changes that are of no-interest, and therefore to detect relevant changes by cancellation.

3.3. Proposed Approach to Unsupervised CD in VHR Multispectral Images Acquired by Different Sensors

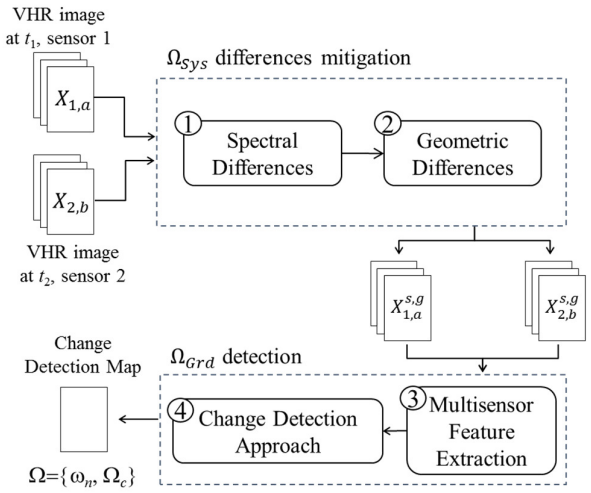


Figure 3.3. Block scheme of the proposed approach to CD in multitemporal multisensor VHR optical images.

In this section, details on the proposed approach to solve the CD problem in VHR multisensor optical images are given. For handling the problem, we focus on two issues: i) mitigation of multisensory induced changes Ω_{Sys} by the homogenization of multispectral data acquired by different VHR sensors and; ii) detection of Ω_{Grd} changes by mitigation of residual Ω_{Sys} at the level of feature extraction. Figure 3.3 depicts the block scheme of the proposed approach. In order to accomplish the Ω_{Sys} mitigation, two main steps are

considered: 1) spectral and; 2) geometric differences mitigation. Ω_{Grd} detection is accomplished in two steps: 3) multisensor feature extraction and; 4) change detection.

Let us consider two VHR optical images, $X_{1,a}$ and $X_{2,b}$, acquired by different sensors S_1 and S_2, where $a = 1,2,\dots,A$ and $b = 1,2,\dots,B$ represent the multispectral bands for S_1 and S_2. The number A and B of spectral channels in S_1 and S_2 can be equal or different depending on the sensor. Given the use of different VHR sensors, $X_{1,a}$ and $X_{2,b}$ are likely to show different number of acquisition bands with slightly different bandwidth and spatial resolution, and/or different view angle. In other words, different spectral and geometrical properties. Let us assume that $X_{1,a}$ and $X_{2,b}$ sizes are $I_1 \times J_1$ and $I_2 \times J_2$, respectively, and that the images are acquired over the same geographical area at different times t_1 and t_2.

3.3.1 Ω_{Sys} Differences Mitigation

When dealing with multitemporal images acquired by different sensors S_1 and S_2, one of the critical issues is to identify and remove acquisition system induced changes (Ω_{Sys}). Handling the differences due to Ω_{Sys}, contributes to mitigate issues on the left side of the tree of radiometric changes (Figure 3.2). In single sensor VHR images, Ω_{Sys} are mainly due to differences given by the sensor view angle (Ω_{Ang}), and are accentuated by the topography (Ω_{Top}) and relief (Ω_{Rel}). All of them contribute to the geometrical differences and result in radiometric distortions. When multisensor VHR images are considered, additional problems arise due to the type of sensor (Ω_{Sen}) and thus the differences in the spectral (Ω_{Spe}) and spatial (Ω_{Spa}) resolution. Ω_{Spe} can be mitigated by performing a radiometric normalization of the images. Whereas Ω_{Spa} should be managed by means of geometric corrections, since they contribute to geometric differences.

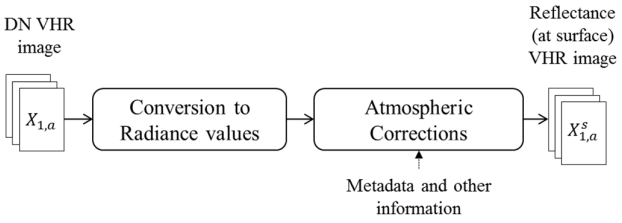


Figure 3.4. Block scheme for mitigation of $\boldsymbol{\Omega_{Spe}}$ in multisensor VHR images.

To mitigate Ω_{Spe}, two macro-categories of normalization methods exist in the literature: absolute and relative methods. The former involves the conversion of the DN values into the corresponding ground reflectance ones [39]. While the latter performs image-to-image adaptation in the histogram feature space [36], [49], [249]–[252]. When data from two different sensors S_1 and S_2 are considered, the spectral information in the multisensor images is not comparable. Thus, absolute normalization is preferred with respect to the relative one, though not limited.

Absolute normalization estimates surface reflectance values providing information at physical level and mitigating spectral differences. The steps for spectral differences mitigation are shown in Figure 3.4, where $X_{1,a}$ and $X_{2,b}$ are converted from DN to surface reflectance (ρ^{TOA} [unitless]) images, $X^s_{1,a}$ and $X^s_{2,b}$ (where s stands for spectrally corrected). Although this step might result obvious in the CD process, several works can be found in the literature that use DN in the comparison step. Further, the mitigation of Ω_{Spe} becomes more critical when multisensor data are considered. To get ρ^{TOA}, the digitalization process performed at the sensor during image formation must be inverted [39]. Parameters such as the mean exoatmospheric solar irradiance, solar zenith angle, Earth-Sun distance, radiance value and others are required. They can be retrieved from the metadata or from user guides, and are specific for each satellite. The resulting $X^s_{1,a}$ and $X^s_{2,b}$ have the same physical meaning. However, some differences cannot be mitigated. Thus, in addition to physical driven strategies, some data driven ones (feature extraction) are required (see next section).

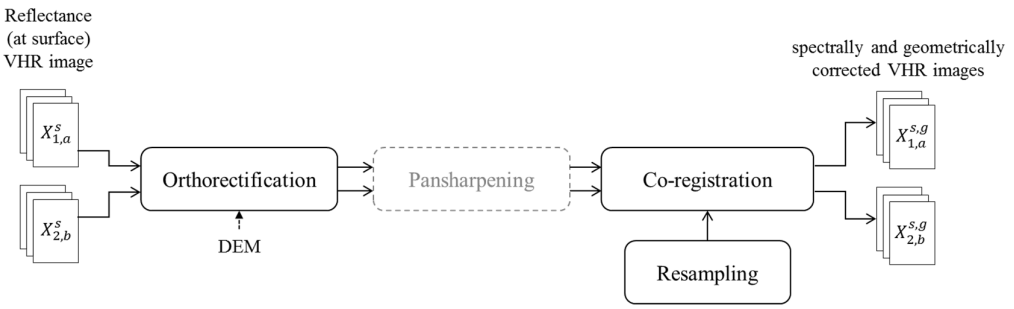


Figure 3.5. Block scheme followed for the mitigation of geometric differences in multisensor VHR images.

Satellites carrying on-board VHR sensors have the capability to acquire images with different view angles; this increases the probability of having multitemporal images with the required time resolution and cloud free on a given area. However, when multitemporal images are considered, differences in the acquisition view angle can induce misalignments because of the impact of topography (Ω_{Top}), small changes in relief of the terrain or the presence of buildings (Ω_{Rel}) [265]–[268]. Further, when $X_{1,a}$ and $X_{2,b}$ are acquired by different VHR sensors, small differences in the spatial resolution (Ω_{Spa}) are also expected. To achieve the geometric differences mitigation (step 2, Figure 3.3), the block scheme shown in Figure 3.5 is followed. $X^s_{1,a}$ and $X^s_{2,b}$ are the input images and $X^{s,g}_{1,a}$ and $X^{s,g}_{2,b}$ are the spectrally and geometrically mitigated ones (g stands for geometrically corrected).

Geometric distortions due to the joint effect of topography, relief and sensor view angle, are mitigated by applying orthorectification with a high resolution Digital Elevation Model (DEM). In this way, most of the misalignments between multitemporal images due to Ω_{Top} and Ω_{Rel} are corrected. However, additional issues remain because of differences in spatial resolutions, thus co-registration should be applied so that pixels with the same coordinates in the images may be associated with the same area on the ground. This step is very critical since a poor co-registration may result in an unreliable final CD map [15]. On the other hand, it is important to clarify that neither orthorectification, nor co-registration solve the problems derived by the presence of vertical structures. This kind of changes are usually considered as sources of noise and are not of interest, thus they can be removed/mitigated by some feature extraction strategy applied during the CD stage (e.g., shadow detection and removal).

Pansharpening (PS) could be applied between orthorectification and co-registration as an optional step. It is meant to improve the spatial information by integrating the high spectral and low spatial resolution bands with the high spatial and low spectral resolution panchromatic band. Several PS methods exist in the literature, e.g., Intensity-Hue-Saturation (IHS), PCA, wavelets and Gram-Schmidt [269]. While co-registering, a resampling of the image with highest spatial resolution is performed in order to get the same common spatial resolution of the one with lower spatial resolution. The spatial resolution of VHR images is metric to sub-metric. Looking at existing satellite missions, the spatial resolution of VHR optical images ranges from the 0.3m of WV-3 and 4 to the 1m of Kompsat-2 in the panchromatic channel. Thus, when considering VHR multisensor image pairs the maximum resolution difference may rise up to about 0.7m. Therefore, VHR images spatial resolutions are different yet similar and comparable. The outputs are multisensor VHR images, $X^{s,g}_{1,a}$ and $X^{s,g}_{2,b}$, showing the same spatial resolution and the same physical information, where Ω_{Sys} have been mitigated.

3.3.2 Ω_{Grd} detection

Once the Ω_{Sys} have been mitigated, the proposed approach performs Ω_{Grd} detection. It extracts the changes of interest by selecting and extracting significant features for specific changes present in the study area. Standard CD approaches like Univariate Difference Image (UID) [53] and CVA [270] perform multitemporal comparison by means of the difference operator. The multispectral (or single spectral, in the case

of UID) difference image X_D is composed by Spectral Change Vectors (SCV).

The rationale behind the use of the difference operator is that unchanged samples show similar spectral signatures and thus result in SCVs with almost all zero components, whereas changed samples show SCVs with components far from zero. However, when multisensor images are considered, such an assumption is seldom satisfied, even after Ω_{Sys} mitigation. Thus, further homogenization is required in order to satisfy the a priori assumption for a successful employment of simple methods like UID and CVA. When multisensor images are considered, a proper feature space should be explicitly identified where pixel based comparison is meaningful. Here we propose the use of higher-level physical features, derived from the ρ^{TOA} ones. While reducing residual sensor-induced differences in the unchanged areas, thus improving the homogenization level, a better highlightening of the changes of interest is achieved. In other words, working with higher-level physical quantities improves the level of abstraction while increasing the probability to detect the changes of interest and reducing the number of false alarms [39].

Since the proposed approach is general, any feature with high-level physical meaning can be used (e.g., radiometric indices). Further, since the approach is designed for VHR images, it may benefit from the use of spatial context information [30], [67]. However here we are interested in understanding the performance of the approach for CD and thus of the multisensor homogenization procedure in mitigating the effects of sensor differences on the CD map. Accordingly, pixel-based features are considered such as radiometric indices. As an example, if changes due to vegetation phenology (Ω_{Veg}) are present, a radiometric index to detect vegetation can be used. In the case of natural disasters (Ω_{Dis}) in urban areas, a building index plus vegetation or soil index could be considered. Radiometric indices suitable to detect most of the relevant kind of changes can be found in the literature.

Selection of a proper index becomes more and more complex when the Ω_{Grd} are coming from different sources. Since here we focus on vegetation and urban changes, we select features based on linear transformations such as TC or Orthogonal Equations (OrE) among the others [271], [272]. TC features were designed originally as a linear transformation for the agricultural analysis on single date images [271], but they were further analyzed and used for CD analysis in multitemporal images. The literature works mainly used TCs in medium resolution and single sensor images (e.g., Landsat) [261], [262]. Three main TCs have been studied (i.e., Brightness, Greenness and Wetness) because of their ability to detect and monitor soil content or transitions, vegetation and canopy moisture. Since TC is an invariant transformation in the physical feature space, its features are consistent between different scenes in a multitemporal time series [271] and therefore could be invariant between multisensor multitemporal images. Similarly, the OrE were derived following the TC philosophy and highlight information on crops, vegetation and soil. Additionally, OrE were designed as a substitution for the sensors which TC coefficients have not been derived in the literature, yet.

Given the above discussion, we use TC or OrE for the Ω_{Grd} detection process. These transformations compute a linear combination of the spectral bands. The number of features derived from TC is the same as the number of input features (in our case A and B, respectively), but only 3 TCs are generally used for CD. In the case of OrE, only 3 features are derived. Equation (3.1) shows the general equation to calculate TC or OrE features (F), where j represents each of the F features and $C_{j,a}$ are the coefficients calculated for each X_j^F. The same equation applies for TC and OrE, though only the red, blue, green and NIR bands of the sensors are used for the latter. The following analysis applies for both TC and OrE.

$$X_j^F = \sum_{a=1}^{A} C_{j,a} X_{1,a}^{s,g} \tag{3.1}$$

Once physical level features have been extracted from different sensors, step 4 (Figure 3.3) applies CD.

As mentioned above, CVA is employed. Since three features are considered, a 3-D representation is obtained [273]. Each SCV of X_D^F is defined as in equation (3.2).

$$X_D^F = X_2^F - X_1^F \tag{3.2}$$

Each SCV component captures the multitemporal behavior of the corresponding feature (either TC or OrE). SCV components tend to assume small values when no change occurred. Whereas if changes occurred, components assume large values (either negative or positive) depending on the kind of change. This is true, even if single date features may have different ranges across each other since the difference operator (3.2) is applied to corresponding features computed at t_1 and t_2 showing similar range. To effectively perform CD in the multidimensional space defined by X_D^F, the information in X_D^F vector is represented in spherical coordinates by computing its magnitude (ρ), azimuth angle (θ), and elevation angle (φ). The relationship between X_D^F in Cartesian coordinates and Spherical coordinates is described by equations (3.3)-(3.5).

$$\rho = \sqrt{{X_{D,1}^F}^2 + {X_{D,2}^F}^2 + {X_{D,3}^F}^2} \tag{3.3}$$

$$\theta = atan2\left(\frac{X_{D,2}^F}{X_{D,1}^F}\right)^3 \tag{3.4}$$

$$\varphi = arcos\left(\frac{X_{D,3}^F}{\rho}\right) \tag{3.5}$$

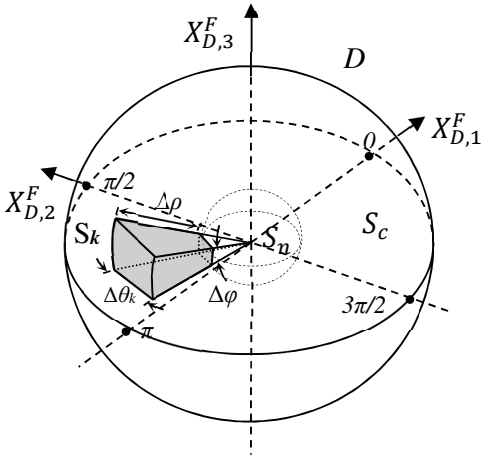


Figure 3.6. Regions of interest for CVA in spherical coordinates: domain D of SCVs in X_D^F, sphere S_n of no-changes, spherical shell S_c including changes and solid truncated cone S_k associated to a generic change k [273].

In the spherical representation, unchanged samples having small values in all the X_D^F components assume small magnitude (ρ) values; whereas changed samples assume large values in one or more X_D^F components thus showing a large magnitude (ρ) and a direction along θ and φ variables that depends on the ratios among values of the X_D^F components (i.e., on the kind of change). Therefore, the magnitude variable carries information about the presence/absence of changes, whereas the direction variables carry information about the possible kind of changes [50], [270], [273]. According to these observations, a magnitude-direction domain (D) (Figure 3.6) can be defined as in equation (3.6) that includes all SCVs.

$$D = \{\rho \in [0, \rho_{max}], \theta \in [0,2\pi) \; and \; \varphi \in [0,\pi)\} \tag{3.6}$$

[3] $atan2$ is a common variation on the $arctan$ function that solves the problem of $arctan$ to distinguish between diametrically opposite directions. $atan2$ range is confined in the interval $(-\pi, \pi]$.

ρ_{max} is the maximum magnitude of X_D^{TC}.

Three subsets of D are of interest. i) the sphere (S_n) that includes unchanged pixels, i.e. the ones with small magnitude values; and ii) the spherical shell (S_c) that includes changed pixels, i.e., the ones with large magnitude values. S_n and S_c are complementary, the radius T of S_n is the inner radius of S_c, and their union provides D. T separates changed from unchanged samples along the magnitude. Since the magnitude is a compressed 1-dimensional representation of the change problem, T is obtained as a trade-off among the effects of the kinds of change. Yet, the definition of T along the magnitude only is a simple and effective solution [50], [273]; and iii) truncated cone sectors (S_k) of changed pixels associated to preferred directions (θ_k, φ_k) in S_c (gray shaded truncated cone in Figure 3.6). Each preferred direction is associated to a specific kind of change Ω_{Grd} (see Figure 3.2). The volume S_k associated with the *k-th* change is defined as:

$$S_k = \{\rho, \theta, \varphi : T \leq \rho < \rho_{max}, \theta_{k_1} \leq \theta < \theta_{k_2}, \varphi_{k_1} \leq \varphi < \varphi_{k_2}\} \tag{3.7}$$

The upper and lower bounds θ_{k_1}, θ_{k_2}, φ_{k_1} and φ_{k_2}, as well as T, can be calculated manually or automatically [117]. Once the angular thresholds have been estimated, the magnitude threshold T can be refined for S_k to account for the behavior of each specific kind of change [57]. Finally, the change detection map (CD_{map}) is built by including the following labels $\Omega = \{\omega_n, \Omega_{Grd}\}$, with $\Omega_{Grd} = \{\Omega_{Dis}, \Omega_{Veg}, \Omega_{Env}, \dots, \Omega_{Ant}\}$, where ω_n refers to unchanged areas. As last step, non-relevant changes (e.g., misregistration, shadows), usually associated with the left side of the tree of radiometric changes shown in Figure 3.2, are removed from the CD_{map}.

3.4. Experimental Results

3.4.1 Dataset Description and Design of Experiments

The proposed approach was validated over different areas located in the Trentino region in the north of Italy between 45°57'45.21"- 46°07'28.04"N and 10°58'35.37"-11°14'29.27"E (Figure 3.7). These areas show interesting properties from the point of view of orographic conformation and environmental variety. Over a relatively small region it is possible to find: i) flat regions including precious apple and vineyard fields, and urban, sub-urban and industrial areas with different density and structure; and ii) hill and mountain environments with a variety of tree species.

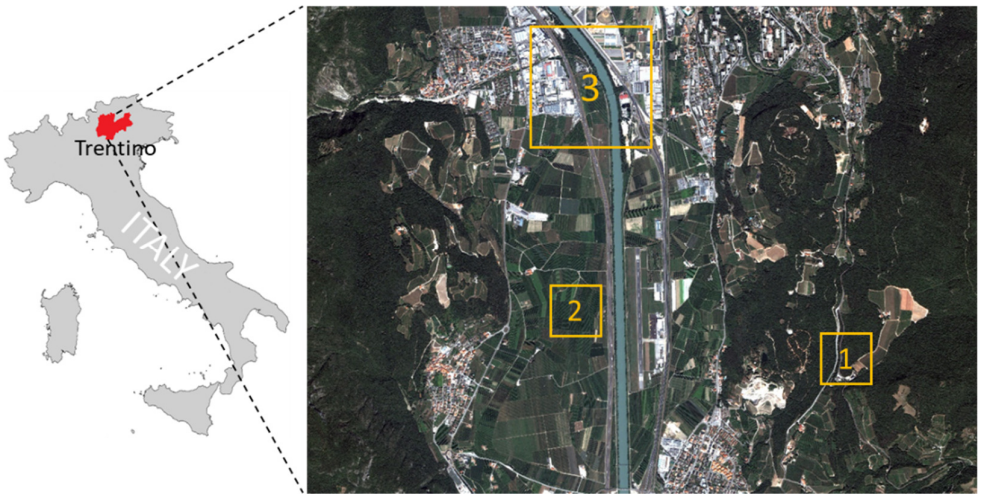


Figure 3.7. Area of interest, Trentino region in the North of Italy.

Three multitemporal data sets made up of a QuickBird (QB), two Worldview-2 (WV-2) and one GeoEye-

1 (GE-1) images were constructed over the sample areas (yellow squares in Figure 3.7). The three datasets were selected such that different kinds of change are represented. Therefore, dataset 1 shows the transition from forest area to several kinds of vegetation, dataset 2 shows transitions among different phenological states of crop areas; and dataset 3 shows transitions from vegetation and bare soil (and vice versa) and changes in roofs and roads around an urban area. These datasets allow us to evaluate the complexity of working with multisensor VHR optical images. For datasets 1 and 2, X_1 is a QB image, acquired in July 2006 with 14.1° off-nadir angle. Whereas X_2 is a WV-2 image acquired in August 2010 with 19.3° off-nadir angle. For dataset 3, X_1 is a WV-2 image, acquired in May 2011 with 7.8° off-nadir angle. Whereas X_2 is a GE-1 image acquired in September 2011 with a 14.4° off-nadir angle. The three satellites show some remarkable differences due to differences in the view angle, the spectral resolution, the number of bands and the spatial resolution.

Table 3.1. Main characteristics of QuickBird, WorldView-2 and GeoEye-1 optical sensors **[274]**.

Satellite	QuickBird	WorldView-2	GeoEye-1
Bands (nm)	445-900 (pan)	450-800 (pan)	450-800 (pan)
		400-450 (coastal)	
	450-520 (blue)	450-510 (blue)	450-510 (blue)
	520-600 (green)	510-580 (green)	510-580 (green)
		585-626 (yellow)	
	630-690 (red)	630-690 (red)	655-690 (red)
		705-745 (red edge)	
	760-900 (NIR)	770-895 (NIR 1)	780-920 (NIR)
		860-1040 (NIR 2)	
Spatial Resolution (m)	0.61	0.46	0.41
	2.44	1.84	1.65

The QB and GE-1 images have four multispectral bands, whereas WV-2 has eight. The spatial resolution of the QB image is 0.6m for the panchromatic band and 2.4m for multispectral bands, whereas WV-2 and GE-1 offer a higher spatial resolution in both panchromatic and multispectral bands with 0.5m and 2m, respectively. Table 3.1 summarizes the characteristics of QB, WV-2 and GE-1 images from the spectral and spatial point of view. The spatial resolution differences imply that the sizes $I_1 \times J_1$ and $I_2 \times J_2$ of X_1 and X_2 images, respectively, are different despite they cover the same surface. The size of QB image in Figure 3.7 is 8674×6300 pixels, whereas the size of WV-2 and GE-1 images is 10297×7139 pixels. Thus, pixel-by-pixel comparison cannot be directly applied since the same pixel coordinates in the two images do not correspond to the same position on the ground. Concerning spectral resolution, we can observe that the four primary multi-spectral bands of QB, GE-1 and WV-2 are acquired over similar spectral ranges (e.g., red), but not fully identical (e.g., blue). Similar considerations hold for green and NIR bands.

In order to apply the proposed approach, we define the tree of radiometric changes (Ω_{Rad}) specific for the considered study areas (one single joint tree is provided for the three datasets), apply mitigation, extract suitable features and perform Ω_{Grd} detection. As a first step, we define the specific tree of radiometric changes (Ω_{Rad}) for the considered problem by starting from the general tree given in Figure 3.2 [30]. The three datasets show changes due to acquisition conditions (Ω_{Acq}) and changes occurred on the ground (Ω_{Grd}). With regard to Ω_{Grd}, there are changes in the phenological state of the vegetation (Ω_{Veg}) (e.g., radiometry of some crops yards, trees, small roads between crop yards (Ω_{Cro}), re-vegetation) and changes due to anthropic activities (Ω_{Ant}) (e.g., changes in road Ω_{Roa}, deforestation Ω_{Def}, roofs Ω_{Bui} and crop planting). It is important to clarify that even though the kinds of changes can be visually separated by photointerpretation, we do not have enough information to give a precise "from-to" label to them.

Concerning Ω_{Acq}, both atmospheric conditions (Ω_{Atm}) and acquisition system (Ω_{Sys}) effects are present. Ω_{Atm} is mitigated by means of atmospheric corrections [39]. Ω_{Sys} is related to the type of sensor (Ω_{Sen}) and to the sensor view angle (Ω_{Ang}). The Ω_{Ang} changes generate small differences in the appearance of objects, leading to geometric distortions, thus to residual misregistration even after proper alignment of images, and to spectral differences when tall buildings are present. We can also see some differences in shadows, which become more critical when high buildings, structures or reliefs are present. Ω_{Ang} and Ω_{Sen} changes are non-relevant from the application viewpoint. Therefore, they are explicitly handled before proceeding to detect Ω_{Grd} and tuning the final CD map. Ω_{Sys} such as the ones due to sensor acquisition mode (Ω_{Mod}) are not considered since we are working with passive sensors. Ω_{Grd} like the ones due to natural disasters (Ω_{Dis}) or environmental conditions (Ω_{Env}) are neglected, since such events did not occur in the considered study area. According to this analysis and to the general taxonomy in Section 3.2, the tree of radiometric changes for the considered problem becomes the one in Figure 3.8.

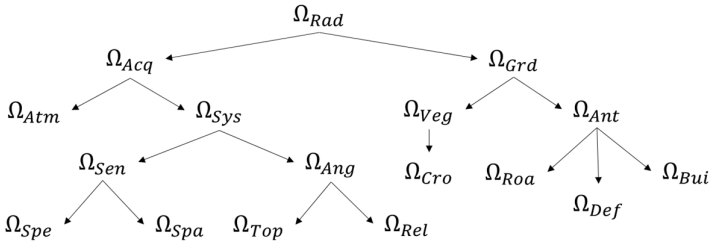


Figure 3.8. Tree of radiometric changes for the considered problem.

Once the radiometric tree was defined, we performed spectral and geometric differences mitigation. All images were provided by DigitalGlobe Foundation in the context of the "MS-TS – Analysis of Multisensor VHR image Time Series" project [275]. Conversion from DNs to surface reflectance (TOA) was conducted before delivery by means of the Atmospheric Compensation (AComp) algorithm [39], [276], [277]. Given the orography of the study area and the possible distortions, we applied orthorectification by using a DEM obtained from LiDAR data [278]. Further distortions appear in dataset 1, since it is located in mountain area, and dataset 3 because of the presence of buildings. Additional pixel-to-pixel problems are also observed due to Ω_{Ang}, and co-registration should be applied.

In order to achieve a better co-registration, PS was applied by means of the Gram-Schmidt method. Here ENVI software package was employed [279]. After PS, the spatial resolution for QB is 0.6m, and 0.5m for GE-1 and WV-2 multispectral bands. Co-registration was carried out over the QB - WV-2 and WV-2 - GE-1 pairs, covering the whole study area in Figure 3.7, by using a polynomial function of second order. For the QB 2006 and WV-2 2010 couple, 79 uniformly distributed Ground Control Points (GCP) were selected. Whereas 68 uniformly distributed GCP were selected for the WV-2 2011 and GE-1 2011 couple. The WV-2 2010 image was resampled during co-registration. Resampling was performed by means of the nearest neighbor interpolation. Figure 3.9 shows the pansharpened multisensor VHR QB, GE-1 and WV-2 images after applying spectral and geometric mitigation in the first and second column, respectively.

Datasets 1 and 2 show a common spatial resolution of 0.6m and a size of 640x640 pixels. Whereas dataset 3 shows a common spatial resolution of 0.5m and a size of 1800x1800 pixels. Datasets 1, 2 and 3 appear in first, second and third row of Figure 3.9, respectively. In order to perform qualitative and quantitative analysis, a reference map for datasets 1 and 2 was defined by photointerpretation and exploiting prior knowledge on the scene as no ground truth was available (see Figure 3.9 (c) and (f), showing 332414 and 280928 unchanged pixels (white color) and 77186 and 128672 changed pixels (different colors), respectively). For dataset 3, considering the extent of the area and the fact that we have no complete knowledge of the changes occurred on the ground, it was not possible to derive a complete reference map. Thus, quantitative analysis was based on 62808 pixels marked as changed, and 6263 as unchanged, selected by photointerpretation.

For comparison purposes, a false color composition of the two acquisitions is provided, green and fuchsia areas represent changes (Figure 3.9 (i)). Changed pixels in the reference map include Ω_{Grd}, only. For dataset 1, changes from i) forest to bare soil, ii) forest to grass, iii) base soil to grass and iv) bare soil to some road are identified. For dataset 2, changes from i) dense vegetation to sparse or light vegetation and vice versa, and ii) bare soil to vegetation are present. And for dataset 3 changes are from i) bare soil to vegetation, both dense and sparse, ii) one to another color of the roofs, and iii) old to new roads.

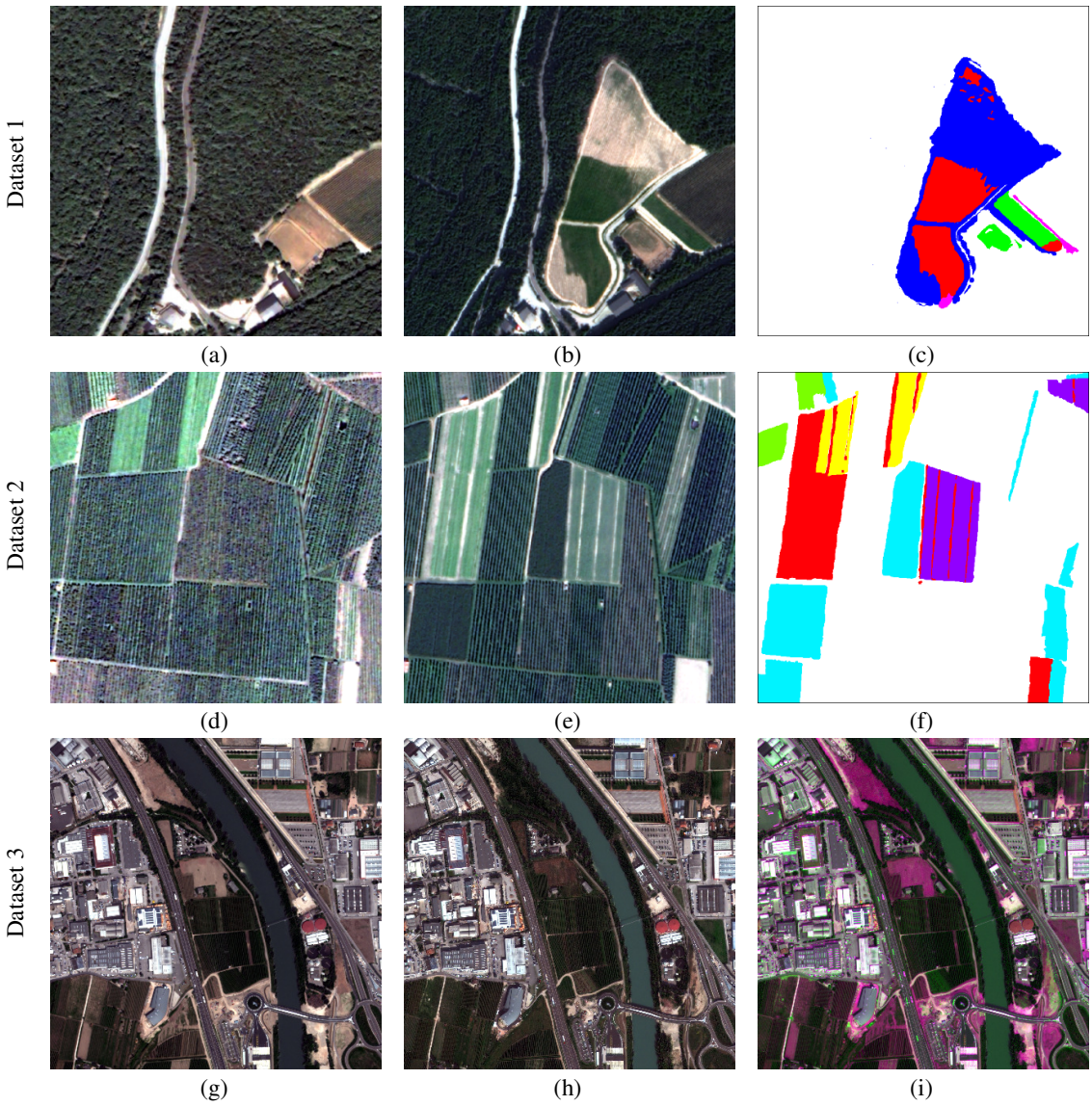


Figure 3.9. True color composition of the pansharpened multispectral multisensor VHR datasets: (a), (d) QB image acquired in July 2006; (b), (e) WV-2 image acquired in August 2010, (g) WV-2 image acquired in May 2011 and; (h) GE-1 image acquired in September 2011. (c) and (f) Reference maps and; (i) false color composition for dataset 3 (magenta and green shades highlight changes).

Once spectral and geometric mitigation was achieved, mitigation of residual Ω_{Sys} was performed at the level of feature extraction. The selection of the features is therefore designed to detect the residual Ω_{Sys} and the Ω_{Grd}. According to the tree of radiometric changes (Figure 3.8), residual Ω_{Sys} might be related to Ω_{Ang}, resulting in possible shadows and/or registration noise, whereas Ω_{Grd} includes three kind of changes: Ω_{Cro}, Ω_{Roa} and Ω_{Def}. Residual Ω_{Sys} due to shadows, due to vegetation or buildings, were detected by means of

the method in [280], whereas Ω_{Sys} due to registration noise are negligible. Ω_{Grd} were detected by employing TC and OrE features.

Three main TCs (i.e., Brightness, Greenness and Wetness) and three OrE (Crop mark, Vegetation and Soil) have been studied because of their sensibility to soil content or transitions from-to soil, vegetation, canopy moisture and anthropogenic activities have been studied. Thus, we expect them to properly detect the different changes in the study area, with the exception of some transitions between green areas that do not show up in TC features. These are the cases of datasets 1 and 3, where transitions from forest to grass and crop to grass are misdetected. This is due to the fact that the difference is more in texture rather than in TC features. In order to evaluate and compare the proposed approach, experiments carried on features such as TOA, are also conducted. The detection of changes is obtained by applying CVA to a 3D TC and OrE feature space. However, other features such as IR-MAD [90] could be also used, despite no approaches for automatic detection with these features in multisensor VHR images exist yet. The drawback with IR-MAD is that it requires an end-user interaction to select the most specific components that represent the specific change of interest and to separate among changed and un-changed samples. This is time consuming and makes the approach not fully automatic.

Two experiments were designed: i) experiment 1 (exp. 1) applies CVA to the transformed TOA bands features; and ii) experiment 2 (exp. 2) applies CVA to TC (for datasets 1 and 2) or OrE (for dataset 3 which has a large presence or urban areas) features. In exp. 1, the first two selected bands are the Near-IR (NIR) and Red (R) given their high spectral sensibility in the analysis of vegetation and anthropic activities. The R bands of QB, GE-1 and WV-2 have quite similar spectral range, whereas the NIR ones do not (see Table 3.1). Therefore, between NIR1 and NIR2 WV-2 bands, NIR1 (770-995nm) was selected given that its spectral range better matches to the QB NIR (760-900nm) and GE-1 NIR (757-853nm) band spectral range. Another TOA feature to be selected could be the Green or Blue band. Empirical experiments showed a slightly improvement in the final CD accuracy while using Green band instead of Blue one. Therefore, Green band was selected as third feature. In exp. 2, for datasets 1 and 2, TC features were selected based on: i) the maximum number of TC features that can be derived for each specific sensor, ii) the radiometric tree of changes; and iii) the possibility to compare between multisensor TC features. Thus, 4 TC features were derived, bounded by QB properties. In accordance with the radiometric tree, features should be selected that are able to highlight Ω_{Veg} and Ω_{Ant}. Based on the level of comparison between the multisensor TC coefficients, and according to the state-of-the-art, only the first three TC features of each sensor show similar physical meaning. For dataset 3, only three OrE features exist in the literature and are derived by means of the 4 main spectral bands of the sensors.

TC and OrE features were derived directly from the spectrally mitigated data and by using the coefficients in Table 3.2, Table 3.3 and Table 3.4. For TC, only coefficients corresponding to the first three TC feature are present. Here the set of coefficients in [281] was applied to the QB image. The coefficients are derived from the DN feature space (Table 3.2). There are no TC coefficients derived from TOA values for QB images in the literature. Thus, we applied the QB TC coefficients over the QB DN features, and compared the derived TC features as a higher level primitive. For the WV-2 image, coefficients are applied as given in [282]. They are derived for the TOA features (Table 3.3). The OrE features were derived by means of the coefficients shown in Table 3.4 and as per equation (3.1).

Table 3.2. TC coefficients for QuickBird DN values [281].

Bands	Brightness (TC1)	Greenness (TC2)	Wetness (TC3)
B1	0.319	-0.121	0.652
B2	0.542	-0.331	0.375
B3	0.490	-0.517	-0.639
B4	0.604	0.780	-0.163

Table 3.3. TC coefficients for WorldView-2 TOA values [282].

Bands	Brightness (TC1)	Greenness (TC2)	Wetness and shadows (TC3)
B1	-0.060436	-0.140191	-0.270951
B2	0.012147	-0.206224	-0.315708
B3	0.125846	-0.215854	-0.317263
B4	0.313039	-0.314441	-0.242544
B5	0.412175	-0.410892	-0.256463
B6	0.482758	0.095786	-0.096550
B7	-0.160654	0.600549	-0.742535
B8	0.673510	0.503672	0.202430

Table 3.4. Orthogonal coefficients for WorldView-2 and GeoEye-1 [272].

Orthogonal Component		Blue	Green	Red	NIR
WV-2	Crop Mark	-0.38	-0.71	0.20	-0.56
	Vegetation	-0.37	-0.39	-0.67	0.52
	Soil	0.09	0.27	-0.71	-0.65
GE-1	Crop Mark	-0.39	-0.73	0.17	-0.54
	Vegetation	-0.35	-0.37	-0.68	0.54
	Soil	0.08	0.27	-0.71	-0.65

3.4.2 Experimental Results

In order to assess the effectiveness of the Ω_{Sys} mitigation and the Ω_{Grd} detection approach based on higher-level physical features, CVA was applied by considering the 3D feature space defined above and by means of equations (3.2)-(3.5). We first extracted all the areas that correspond to radiometric changes in the image, by thresholding the magnitude variable. The selection of a threshold T over ρ showed to be a simple and fast way to separate among changed and non-changed pixels. A good separation, among the three dimensions of CVA was guaranteed in average. T was automatically selected by applying a Bayesian decision rule [44]. The T values for each of the datasets in the two experiments are shown in Table 3.5.

Table 3.5. Magnitude (T) threshold values for the three datasets in Exp.1 and 2.

Dataset	Exp.	T
1	1	0.060
	2	0.080
2	1	0.025
	2	0.030
3	1	0.090
	2	0.060

Figure 3.10 shows the multispectral difference image 3D histogram for the dataset 1. In Figure 3.10 (a) and (c), it is possible to see how X_D^{TOA} and X_D^{TC} are distributed in a coplanar and sparse way, respectively. This kind of distribution leads to an easier visual interpretation of the X_D^{TC}, if compared to X_D^{TOA}. In fact, from the 3D histograms in Figure 3.10, we can see that the coplanar distribution of TOA features does not allow good separation between changes of interest and changes of no-interest. A similar behavior is observed between the TOA and OrE features. Right column of Figure 3.10, presents the changed samples after removing the unchanged ones. Different clusters fit to different S_k sectors and are associated with different changes. As we move from X_D^{TOA} to X_D^{TC}, or X_D^{TOA} to X_D^{OrE}, it becomes evident how different clusters locate around preferred directions and how the number of clusters increases, making their detection and separation more effective. Separation among different clusters can be performed by automatic or manual methods.

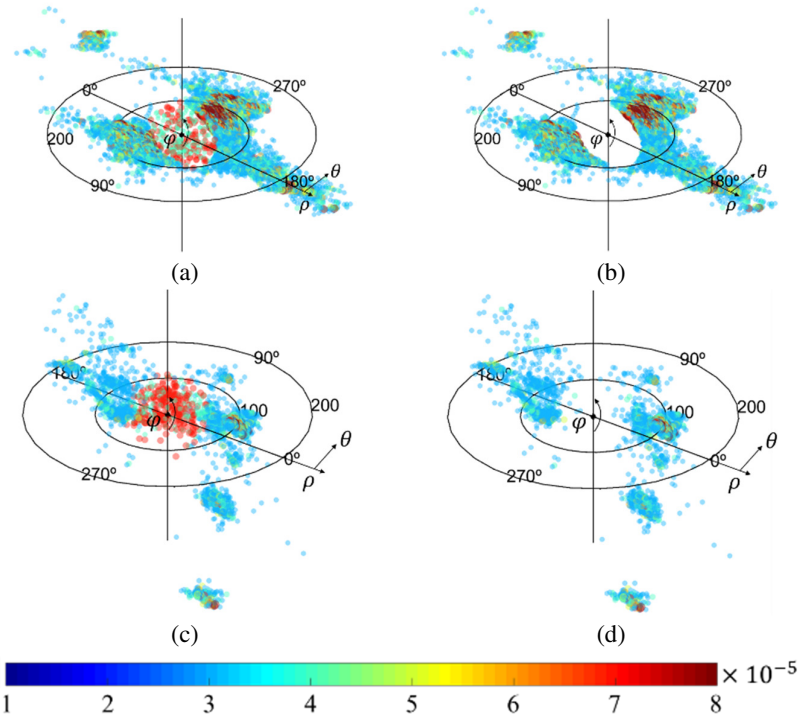
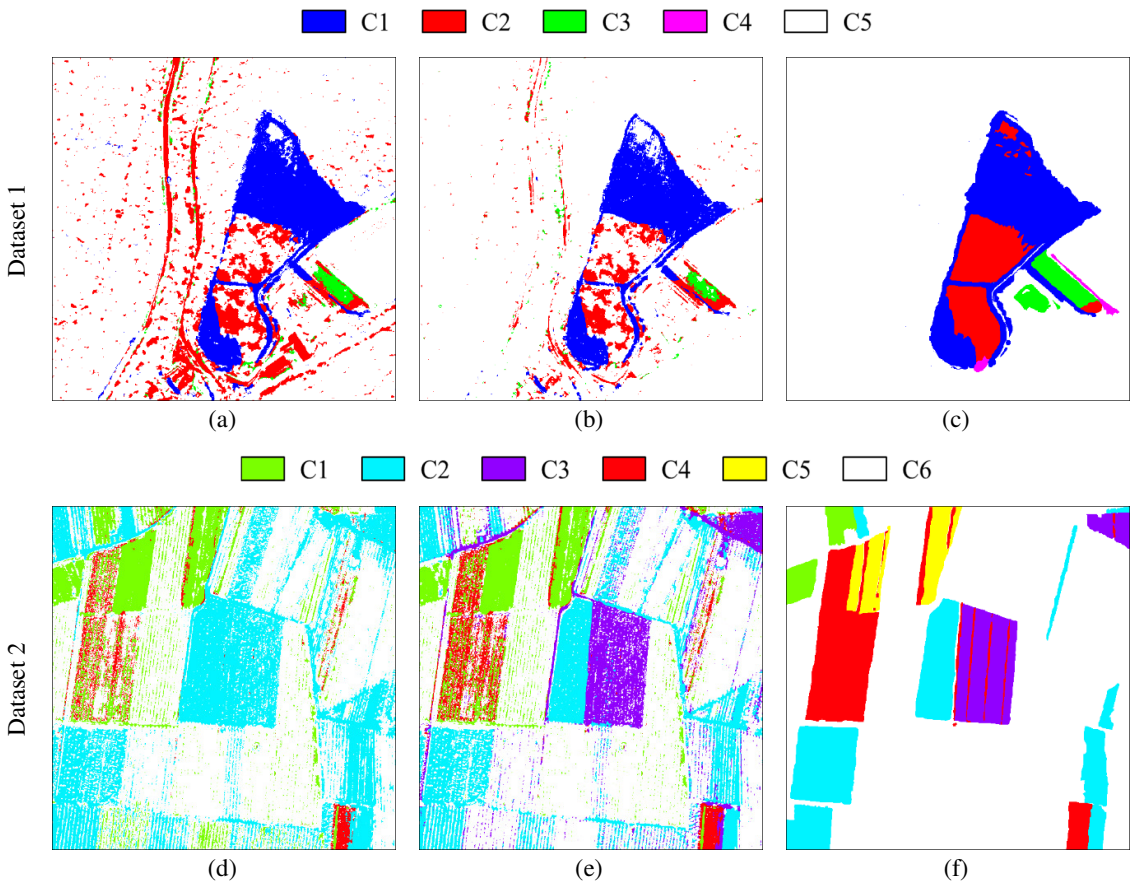


Figure 3.10. Histogram in spherical coordinates of (a) X_D^{TOA}, (c) X_D^{TC}, and changed samples after thresholding ρ for (b) TOA and (d) TC features.



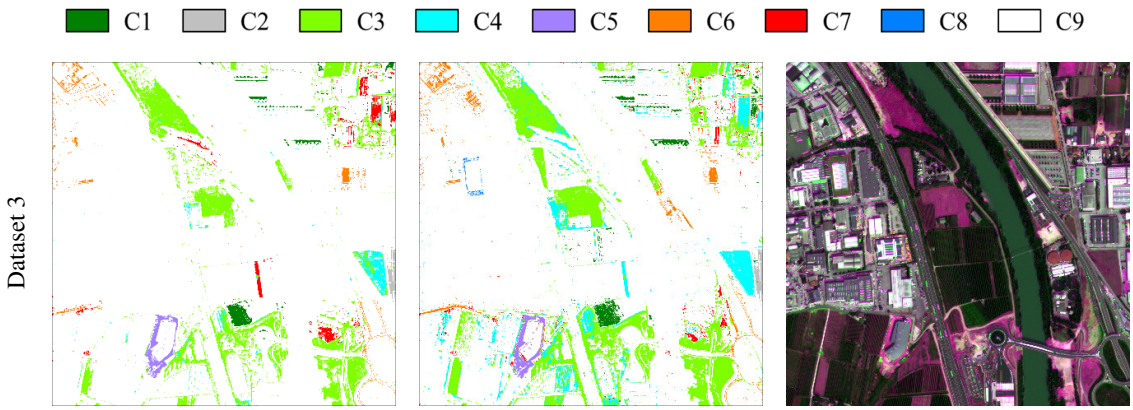


Figure 3.11. Change detection maps obtained by CVA in 3D applied to the three datasets in: (a), (d), (g) TOA features; (b), (e) TC features; and (h) Orthogonal features. (c), (f) Reference map; and (i) false color composition for dataset 3 (magenta and green shades highlight changes).

Last step builds the multiple CD_{map} by means of the extraction by cancellation strategy. To this end, selection of each of the clusters in the changed region was conducted automatically by means of an adaptation of the TSMO proposed in [45]. Final multiple CD maps were built by cancelling the remaining Ω_{Sys} and the Ω_{Grd} that are out of interest (e.g., cars in road and parking areas, Figure 3.11). As expected, some vegetation changes that affect more texture rather than TC features are misdetected in datasets 1 and 3. In other words, the selected higher-level physical features are not optimized for those changes, but different higher-level physical features could be selected to properly model most of the kinds of changes.

In order to perform a quantitative analysis, a comparison of the CD maps with their reference maps (for datasets 1 and 2) and the set of points collected by photointerpretation (dataset 3) was carried out. The comparison pointed out the improvement achieved when working with higher-level physical quantities, specifically for transitions from and to bare soil and different types of crops. This is confirmed when we analyze dataset 1 where changes are mainly from-to vegetation and bare soil. TC features outperform the results obtained when using TOA features (Table 3.6). As expected, both TC and TOA features have similar problems to identify the change from forest to vegetation. An analogous situation occurs on dataset 2, with changes from-to different types of crops. In this case, TC outperforms TOA being able to separate among changes C2 and C3 (which cannot be discriminated when using TOA features - see Figure 3.11 (d), (e) and (f)).

For datasets 1 and 2, the major improvement is related to the decrease of False Alarms (FA) (see Table 3.6 and Table 3.7). In dataset 1, the FA correspond to the main road passing through the area and to some of the remaining shadows generated by the tree lines. Even though an index was used to remove the shadows, a small percentage of them remained. In dataset 2, the FA correspond mainly to the linear structures like roads in between the different crops. Moreover, it is possible to observe improvements in terms of detection and separation of the different kinds of changes. The number of Missed Alarms (MA) decreased as well when working with TC, leading to a better detection of changes, especially when a higher number of changes exist, such as the case of dataset 2. From the quantitative viewpoint, the reduction in both FA and MA reflects in the Overall Accuracy (OA) that increases of about 5% and 7% for dataset 1 and 2, respectively (see Table 3.6 and Table 3.7).

In the case of dataset 3, on the one hand, we can see how TOA features have problems to properly separate among C3 and C4 classes that correspond to transitions from bare soil to sparse and dense vegetation, respectively. On the other hand, the number of MA for the class C3 by TOA features is clearly larger than for OrE features (see Table 3.8). Moreover, OrE features are able to better detect the classes C2 and C7 than TOA features. C2 and C7 correspond to changes occurred on building roofs and roads renewed in the

studied period. Finally, TOA features do not detect class C8 (change in roof color of a building), whereas OrE features do. Other transitions from-to vegetation and bare soil can be seen in the change classes with less MA and FA when the OrE features are used. It is worth noting that changes due to cars on roads or in parking areas were considered as non-relevant for this study and thus neglected. From the quantitative perspective, the OA obtained by OrE features increased the overall accuracy of about 2% over that of TOA features. This improvement can be considered relevant given the complexity of the scene with the presence of more kinds of changes. Proper higher-level physical features, such as some radiometric indexes or texture features, would provide better results. Table 3.6, Table 3.7 and Table 3.8 show the confusion matrices obtained for each dataset in the two experiments.

Table 3.6. Confusion matrices for dataset 1 in Exp. 1 and Exp. 2.

Exp.	Changes Found	Actual Changes					
		C1	C2	C3	C4	C5	Reliability
1	C1	**39775**	596	91	0	2129	93.38%
	C2	2206	**9500**	1501	44	28428	22.79%
	C3	9	5	**2427**	128	2077	52.24%
	C4	0	0	0	**0**	0	0.00%
	C5	3765	13090	2718	1327	**299780**	93.48%
	Accuracy	86.93%	40.96%	36.02%	0.00%	90.18%	
Overall Accuracy		85.81%					
2	C1	**36558**	250	92	0	910	96.69%
	C2	2412	**8153**	1355	94	6571	43.87%
	C3	0	2	**1464**	42	1152	55.03%
	C4	0	0	0	**0**	0	0.00%
	C5	6785	14786	3829	1363	**323781**	92.36%
	Accuracy	79.9%	35.15%	21.72%	0.00%	97.40%	
Overall Accuracy		90.32%					

Table 3.7. Confusion matrices for dataset 2 in Exp. 1 and Exp. 2.

Exp.	Changes Found	Actual Changes						
		C1	C2	C3	C4	C5	C6	Reliability
1	C1	**5381**	6	1	10387	12286	16408	12.10%
	C2	142	**40242**	19330	3560	4	47406	36.35%
	C3	0	0	**0**	0	0	0	0.00%
	C4	18	0	50	**9472**	10	2688	77.40%
	C5	0	0	0	0	**0**	0	0.00%
	C6	1198	7973	3643	13800	1169	**214426**	88.53%
	Accuracy	79.85%	83.45%	0.00%	25.48%	0.00%	76.33%	
Overall Accuracy		65.80%						
2	C1	**3269**	8	3	10447	12423	12197	8.52%
	C2	66	**37971**	291	0	22	29301	56.13%
	C3	116	773	**19534**	3713	2	15080	49.81%
	C4	59	3	196	**15476**	51	4156	77.61%
	C5	0	0	0	0	**0**	0	0.00%
	C6	3229	9466	3000	7583	971	**220194**	90.08%
	Accuracy	48.51%	78.74%	84.84%	41.58%	0.00%	78.38%	
Overall Accuracy		72.37%						

Table 3.8. Confusion matrices for dataset 3 in Exp. 1 and Exp. 2.

Exp.	Changes Found	Actual Changes									
		C1	C2	C3	C4	C5	C6	C7	C8	C9	Reliability
1	C1	**905**	0	0	0	0	0	0	0	0	100%
	C2	0	**206**	0	0	0	0	0	0	0	100%
	C3	0	0	**1642**	334	0	0	0	0	11	82.64%
	C4	0	0	27	**450**	0	0	16	0	0	91.28%
	C5	0	0	0	0	**240**	0	0	0	0	100%
	C6	0	0	0	0	0	**400**	0	0	12	97.09%
	C7	3	0	0	68	0	0	**195**	0	0	73.31%
	C8	0	0	0	0	0	0	0	**0**	0	0.0%
	C9	725	0	203	392	0	198	89	213	**62785**	97.18%
	Accuracy	55.42%	100%	87.71%	36.17%	100%	66.89%	65.00%	0.00%	99.96%	
Overall Accuracy		96.75%									
2	C1	**1294**	0	0	0	0	0	0	0	91	93.43%
	C2	0	**206**	0	0	0	0	0	0	0	100%
	C3	0	0	**1792**	134	0	0	0	0	22	91.99%
	C4	0	0	27	**1052**	0	0	16	0	0	96.07%
	C5	0	0	0	0	**240**	0	0	0	0	100%
	C6	0	0	0	0	0	**582**	0	0	143	80.27%
	C7	0	0	0	0	0	0	**158**	0	205	43.53%
	C8	0	0	0	0	0	0	0	**66**	0	100%
	C9	339	0	26	58	0	16	110	147	**62347**	98.90%
	Accuracy	79.24%	100%	97.13%	84.57%	100%	97.32%	55.63%	30.98%	99.27%	
Overall Accuracy		98.07%									

3.5. Conclusion

In this chapter, an approach for CD in VHR multispectral multisensor optical images has been proposed. The proposed approach aims at defining and illustrating a data flow for effectively handling differences due to acquisition sensors. It is based on a general framework for the design of CD systems for VHR multitemporal images presented in [30]. In order to deal with multispectral and multitemporal images acquired by different sensors, it integrates in the general approach the following two concepts: i) spectral, radiometric and geometric homogenization between images acquired by different sensors; and ii) detection of multiple changes by means of features that guarantees homogeneity over time and across sensors. Experimental results on real datasets, made-up of VHR bi-temporal and multisensor optical images, confirmed the effectiveness of the proposed block scheme and the improvement achieved by the use of higher-level physical features (i.e., TC and OrE) over the traditional features (i.e., TOA). A major improvement is observed when changes from-to vegetation and bare soil, and different types of crops are considered since TC were selected to highlight such kind of changes. The use of OrE for the detection of changes described above, as well as for changes from-to vegetation and bare soil (i.e., forest to grass, crop to grass, bare soil to grass/forest) and small changes on roads and roofs, resulted in better CD accuracy than that obtained by using TOA. In general, both TC and OrE features allow a better separation and interpretation of Ω_{Grd} by guaranteeing that these changes are distributed in compact and well separated clusters. These improvements are related to two facts: i) the magnitude of the differences present between the two multisensor images and, ii) the robustness of the proposed method to the size of the changed areas in the scene. It is clear that when higher magnitude differences are present, the proposed method works better. We can also see that independently of the size of changed areas (~ 19% for dataset 1 and ~32% for dataset 2), the proposed methods, by using higher-level physical features, outperforms the cases where traditional features are used.

Further analysis should be carried out to determine which cluster is representing a specific kind of change, and to define appropriate features for other kind of changes. For the mitigation of remaining Ω_{Sys} and the

better detection of the Ω_{Grd}, the use of additional features, either in the physical feature space or in the spatial feature space, could help to make the separation and distinction better, thus improving the final OA. Additional improvements from the Ω_{Sys}mitigation process point of view in both, spectral and geometric perspective should be considered. For the spectral differences, and given that some of the VHR multisensor optical images have different number of bands and different spectral ranges, the use of regression methods for predicting bands that match from the spectral viewpoint could be considered. For the geometric differences, improvements on the co-registration process by the use of co-registration methods designed for multisensor images could be also explored. Improvements on the detection of building changes could be also integrated.

Chapter 4

4. Generation of Homogeneous VHR Time Series by non-Parametric Regression of Multisensor Multitemporal Images[4]

This chapter presents an approach for the generation of homogeneous VHR TS by means of a non-parametric regression of multisensor multitemporal images. Advantages and disadvantages of using ARN and RRN methods, available in the literature, to perform homogenization of VHR multisensor images are presented. The proposed approach takes advantage of both ARN and RRN methods to perform the homogenization, but introduces a novel RRN method based on a non-parametric regression and CD-driven approach. By focusing on the spectral domain, the proposed approach transforms a VHR image from a multisensor TS into the spectral domain of another image in the same multisensor TS, but acquired by a different sensor. To this end, a prediction-based approach relying on a non-parametric regression method is employed to mitigate sensor-dependent spectral differences. The impact of possible changes occurred on the ground is mitigated by training the prediction model on un-changed samples, only. Experimental results obtained on VHR optical multisensor images confirm the effectiveness of the proposed approach.

4.1. Introduction

A generation of satellite sensors (e.g., IKONOS, QuickBird, GeoEye, WorldView-2, Pleiades) exists able to acquire multitemporal images with Very High spatial Resolution (VHR). Such sensors open to a large set of new applications in the field of Multi-Temporal (MT) analysis (e.g., Change Detection (CD)) that require spatial detail information. Nevertheless, when a single VHR optical passive sensor is considered, Time Series (TS) are likely to show a poor temporal resolution and to include images with non-homogeneous acquisition conditions (e.g., lack of similar light conditions, different acquisition angle). This is mainly due to the satellite revisit period, the possible competing orders of different users on the satellite pointing, the limited life of a satellite mission, and weather conditions [30], [35], [39]. In order to mitigate the above-mentioned limitations and to have TS showing a (very) high resolution, in both space and time, multisensor optical acquisitions can be considered for TS construction. In this way, the probability of having frequent and good images over the same geographical area of interest increases. However, the use of multisensor multitemporal images shows several issues. The main one is related to images homogeneity or consistency. Lack of consistency in MT images applications impacts on the outcomes accuracy and reliability [39]. On the one side, images in multisensor TS are affected by differences induced by the acquisition conditions (e.g., atmospheric conditions and acquisition system). Some of the differences in atmospheric conditions (e.g., cloud cover), and acquisition system (e.g., view angle and seasonal effects) affect single-sensor multitemporal image as well [30], [35], [39]. On the other side, multisensor TS poses the big challenge of having system intrinsic differences due to the type of sensor and the sensor acquisition modes. The above-mentioned issues are mainly related to differences in: i) the geometrical resolution; ii) the radiometric resolution; and iii) the spectral resolution, range and bandwidth of the sensors [35], [48], [49]. In conclusion, among the various aspects of image pre-processing for CD when considering multisensor heterogeneous

[4] Part of this chapter appears in:
Journal paper: Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "Generation of homogeneous VHR time series by non-parametric regression of multisensor multitemporal images," ready to be submitted to *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*.
Conference paper: Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "VHR time-series generation by prediction and fusion of multi-sensor images," in 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2015, pp. 3298–3301.

images, there are two outstanding requirements: i) geometric homogenization and, ii) spectral homogenization.

The issues related to geometric homogenization have been widely studied in the literature for low to VHR images and for multisensor images as well [14], [15], [17], [35], [283]–[285]. Whereas the issues arising in the spectral domain are less investigated. VHR available sensors may show significant differences in this domain since sensors are available acquiring images over different spectral ranges (e.g., IKONOS acquires bands in the range 445-853nm, while Pleiades acquires in the range 430-950nm [274]). In the case where similar spectral ranges are considered, the number of bands and/or their width may differ (e.g., QuickBird has 4 bands whereas WorldView-2 has 8 bands over the same range). Thus, the effective usage of multisensor TS poses the issue of how to perform multisensor data spectral homogenization. In the literature, two kinds of approaches have been proposed for addressing radiometric differences in single-sensor: i) Absolute Radiometric Normalization (ARN) [48] and; ii) Relative Radiometric Normalization (RRN) [49]. ARN refers to the use of physical parameters and makes it possible to relate the Digital Number (DN) with the radiance and reflectance at the Earth surface. A considerable amount of work has been carried out in the literature to design ARN correction methods and specifically to perform the atmospheric corrections [90], [251], [252], [286]–[288]. Other works have shown the importance of working with physical quantities for improving the CD accuracy [39], [289], [290]. Nevertheless, the use of ARN methods is not that common, even for single-sensor images, since their application requires knowledge of both the sensor spectral profile and atmospheric conditions at the time of acquisition. When data from two sensors S_1 and S_2 are considered, their spectral information is less comparable from the physical viewpoint with respect to data acquired from the same sensor. Thus, requiring the ARN methods application [35], [39]. Regarding RRN, it refers to transformations at pixel level for image-to-image adaptation. RRN methods use one image as a reference and adjust the radiometric properties of the other image to match the reference [48], [248], [288]. Thus, normalized images appear as being acquired with the reference image sensor and under similar atmospheric and illumination conditions [248]. However, RRN methods do not remove the differences introduced by atmospheric conditions in MT images, whereas ARN ones do. A variety of RRN methods have been developed in the literature for single-sensor images. Hall *et al* [248] developed a radiometric rectification technique that corrects or rectifies images of the same area by using landscape elements which reflectance is nearly constant over time. Others have used similar procedures [287], [291], [292]. The drawback of these kind of approaches is that the landscape elements are selected by visual inspection, which could result in a subjective radiometric normalization. Thus, further methods perform the correction by scene-to-scene histogram normalization [24], [49], [250], scene-to-scene correction using dark and bright targets [248] or Pseudo Invariant Features (PIF) [247]. Most RRN methods assume that a linear relationship exists among times t_1 and t_2. Thus, the mathematical model describing standard RRN for MT images involves a linear regression [90]. Such regression models are built according to invariant samples or target points known as Radiometric Control Set Samples (RCSS). Burns and Joyce [293], and Singh [294] developed different techniques to select RCSS; however, those techniques still lead to low accuracy. Heo and Fitzhugh [295] suggested a method for obtaining the optimal linear equation with a given set of target points, but the results depend on the subjective selection of RCSS. Many other RRN methods, based on a linear relationship, can be found in the literature [36], [47], [90], [251], [252], [288], [296], [297]. Most of them assume images with the same spectral and geometric characteristics (single-sensor), and acquired by medium or high spatial resolution sensors, devoting poor attention to multisensor and VHR optical images [36], [48], [49]. This linear relation influences the normalization results and is usually assumed because of simplification of modeling [297]. However, and as shown in Figure 4.1, when multisensor VHR images are considered, the assumption of a linear model is seldom satisfied, leading to critical limitations. In fact, a non-linear non-parametric regression model would better adjust to the multisensor VHR problem. The problem remains how to automatically select the proper RCSS to derive the model itself. Thus allowing to perform the homogenization process among multisensor MT VHR images. Therefore, development of new

homogenization methods suitable for multisensor MT VHR images, become of great relevance where VHR intrinsic characteristics are considered, and both ARN and non-parametric RRN methods are combined to exploit their advantages and reduce their drawbacks.

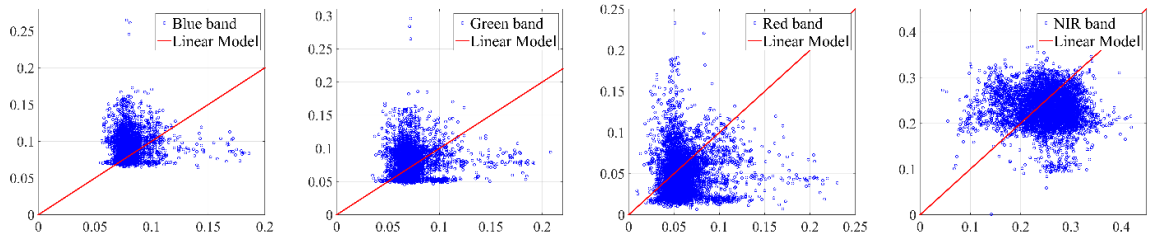


Figure 4.1. Band by band scatterograms of QuickBird 2006 versus WorldView-2 2010 images in unchanged pixels.

This chapter presents a novel method for the generation of homogeneous VHR TS focused on the mitigation of intrinsic spectral induced differences. It is based on non-parametric regression and aims at generating consistent multisensor TS showing a homogeneous spectral representation. The proposed method jointly exploits the capabilities of ARN and RRN approaches, by adapting them to the complexity of multisensor and VHR images. A first homogenization is carried on by transforming DN to physical values (ARN) and by transforming the multisensor images into a common spatial resolution, by means of state-of-the-art methods [14], [15], [17], [35], [39], [285]. The second homogenization step is based on RRN approaches and performs a non-parametric regression (prediction) to derive a model that represents the relationship between S_1 and S_2. The model is derived by means of a machine learning algorithm, i.e., Artificial Neural Networks (ANN) or Support Vector Regression (SVR), and by introducing a novel approach for the RCSS selection. The selection of reliable RCSS is based on a novel CD-driven approach that takes as input invariant features. Because of the lack of one-to-one correspondence among VHR multisensor bands, invariant features cannot be directly derived from S_1 and S_2 bands. Therefore, a fusion of S_1 and S_2 bands is conducted. Such fusion guarantees the selection of radiometrically reliable/invariant features (derived from the fused bands), which in turn will result in an easier selection of RCSS. To demonstrate the effectiveness of the homogenization procedure, MT information is extracted, specifically by applying CD by means of Change Vector Analysis (CVA) [50]. Experiments were carried out on MT multisensor VHR image pairs, including images acquired by 2 sensors.

The remainder of this chapter is structured as follows. Section 4.2 presents an overview of the most common radiometric normalization methods for remote sensing images. Section 4.3 illustrates the proposed method for generation of homogeneous VHR TS by non-parametric regression of multisensor MT images. Sections 4.4 and 4.5 explain in detail the steps of the proposed approach. Section 4.6 presents the MT information extraction as an evaluation mean. Section 4.7 introduces the multisensor datasets used in the experiments, describes the design of experiments and illustrates the experimental results. Finally, Section 4.8 draws the conclusions and future works.

4.2. Radiometric Normalization Methods for Remote Sensing Images

There are several reasons why a sensor response to a given target varies over the time. Among the more relevant ones, we have: i) changes in satellite sensor calibration, ii) differences in illumination and viewing angle, iii) variation in atmospheric effects, and iv) real changes on the ground [30], [35]. The goal of radiometric normalization methods is to remove or compensate for the above-mentioned effects, but the actual changes on the ground. In other words, the goal is to normalize the DN of images acquired under different conditions and report them to a common scale. Two radiometric normalization approaches can be found in the literature, ARN and RRN. The former refers to compensation by means of physical parameters, whereas the latter refers to compensation by adaptation at radiometric values level. Rather ARN or RRN methods

are generally used in the literature to perform normalization [14], [15], [17], [35], [283]–[285], but they are seldom used in combination to get the best from the two [35], [297]. Since ARN methods have been widely standardized in the literature, in this section we focus the attention on the description of RRN ones. RRN methods can be divided into three categories: i) statistical methods (e.g., standard deviation based method), ii) histogram matching methods and, iii) regression methods [49], [288]. Given the image-to-image relationship complexity, RRN methods mostly used in the literature are based on regression, and are the ones described in here. RRN regression methods are based on 3 steps: i) selection of the model, ii) selection of RCSS and iii) estimation of normalization coefficients. For the model, literature mostly rely on linear regression due to its simplicity [297]. The assumption is that linear effects on data are greater than nonlinear effects. Casseles and Garcia [287] showed that for low spatial resolution and single-sensor images, the relationship between the reference and subject image can be linear:

$$Y_m = c_m X_m + d_m \tag{4.1}$$

where Y_m is the observed response in a given band of the reference image Y, X_m is the corresponding vector of observed predictors of the subject image X and, c_m and d_m are the normalization coefficients. Thus, in equation (4.1), the subject image X is normalized by the reference image Y by means of a linear regression. Nevertheless, the assumptions for applying a linear regression are seldom satisfied (especially if we consider multisensor VHR images) and the specific regression model is often unknown. A non-parametric regression analysis relaxes the assumptions of linearity, substituting it by a weaker assumption of a smooth population regression function of the form:

$$Y_m = f(X_m) + \varepsilon_m \tag{4.2}$$

where $f \in \mathcal{F}$, $\mathcal{F}$ is some class of regression functions, and ε is the additive error with zero mean and constant variance. The main advantage here is that $\mathcal{F}$ is a rich enough class such as it is possible to approximate a very large set of regression functions. The cost of relaxing the linearity assumption implies higher computation burden, but with the gain of a more accurate estimate of the regression function. Examples of non-parametric regression methods are ANN [298] and SVR [299]. The second step is the selection of ideal RCSS. Ideal RCSS should be [49], [300]: i) preferably at the same elevation, ii) contain minimal amount of vegetation (when possible), iii) invariant over the acquisition time, and iv) distributed over the spectrum such as the regression model can be reliable. Several methods have been introduced in the literature for the selection of RCSS [247]–[249]. In the next, some of the state of the art methods for selection of RCSS and further estimation of normalization coefficients, in low/medium resolution images, are briefly described together with their limits.

- Simple Regression (SR) [249]: Uses all pixels in both images to calculate normalization coefficients throughout least-squares. Because of this, it works well only when the considered images are stable over time, which is seldom satisfied in VHR images.
- Pseudo Invariant Features (PIF) [247]: Elements such as concrete, asphalt and rooftops are assumed statistically stable between the acquisition dates. Differences in the gray-level distribution of invariant objects are assumed to be linear and are corrected statistically to perform the normalization. It needs human intervention to extract the PIF set. Both linear assumption and human intervention are difficult while working with VHR images.
- Dark-Bright (DB) [248]: the average of a set of dark and bright pixels, extracted from the subject and reference image through Tasseled-Cap (TC) greenness-brightness transformation, is used to derive the normalization coefficients. Appropriate threshold values are required to obtain the dark and bright pixel sets. However, it is difficult to obtain them automatically. No TC coefficients are available for all VHR sensors.
- No Change set (NC) [249]: locates the statistical centers for stable land and stable water data clusters using the near-infrared (NIR) date 1 versus date 2 scatterograms to establish an initial regression

line. At these wavelengths a distinct axis of ‘‘no-change’’ can be observed. Pixels falling within the NC region are used in the regression analysis of each band to compute normalization coefficients. Some limitations are: i) presence of both land and water areas is required and ii) two NIR bands are necessary. The latter issue is critical in VHR images, since often one NIR channel is available.

- Iteratively Re-weighted Multivariate Alteration Detection (IR-MAD) [252]: uses MAD transformation to select no-change pixels in bi-temporal images by assuming a Gaussian distribution of the difference image. It is based on the linear combination of the DN of all the bands in the images, where the normalization coefficients are determined by applying standard Canonical Correlation Analysis (CCA). The final normalization is carried out by means of orthogonal linear regression. As it has already been pointed out, multisensor MT VHR images do not follow a linear relationship, neither does the difference image follow a Gaussian distribution [56].

If single sensor VHR MT images are considered, few of the above methods can be applied [247], [252]. For multisensor VHR images, none of the existing methods can be applied for normalization without modifications on: i) the regression model and/or ii) the selection of RCSS. This is due to the fact that multisensor VHR image pairs may show stronger radiometric dissimilarities than single-sensor ones, even if change did not occur. This results in a lower number of spectrally invariant objects [14].

4.3. Proposed Method for Generation of Homogeneous VHR TS by Non-parametric Regression of Multisensor MT Images

When a single VHR optical passive sensor is considered, TS are likely to show a poor temporal resolution and to include images with non-homogeneous acquisition conditions. In order to have TS showing a (very) high resolution, in both space and time, multisensor optical acquisitions can be considered. Despite increasing the probability of having frequent and good images over the same geographical area, the use of multisensor MT images for information extraction requires reliable homogenization methods. Thus, we propose a method to generate homogeneous TS from multisensor MT VHR images by jointly exploiting the advantage of ARN and RRN methods. First, normalization of multisensor optical data by means of ARN and geometric normalization is applied. Second, RRN is conducted by means of the non-parametric regression of already absolutely corrected VHR multisensor images. The former step guarantees the multisensor comparison from both physical and geometrical view point [35]. Whereas the latter guarantees the comparison from the spectral view point at pixel level. The method is based on non-parametric regression and uses the information provided by image at time t_2 acquired by sensor S_2 to effectively predict how spectral bands of the image acquired at t_1 by sensor S_1 would behave if acquired at t_2 by sensor S_1. Figure 4.2 depicts the block scheme of the proposed method.

Let us consider two VHR optical images, acquired by sensors S_1 and S_2 over the same geographical area, at times t_1 and t_2 and with sizes $I_1 \times J_1$ and $I_2 \times J_2$, respectively. Let a $(a = 1,2, \dots, A)$ and b $(b = 1,2, \dots, B)$ represent the generic multispectral bands of S_1 and S_2, respectively. Given the use of different sensors, the two images are likely to show different bandwidths, spatial resolutions and/or view angles. In other words, we assume that S_1 and S_2 may show: i) slightly different spatial resolution, ii) the same spectral range, iii) a different number of spectral bands A and B, with $B > A$, and iv) different spectral resolution. Thus, images acquired by the two sensors are similar, but not homogenous, and therefore not fully equivalent in the context of MT information extraction. Nevertheless, given the acquisition of images over the same spectral range, more than one band from S_2 acquires information over the same range of a single band from S_1. Because of this, more than one band from S_2 can be used to predict the corresponding bands of S_1, deriving the homogenization process.

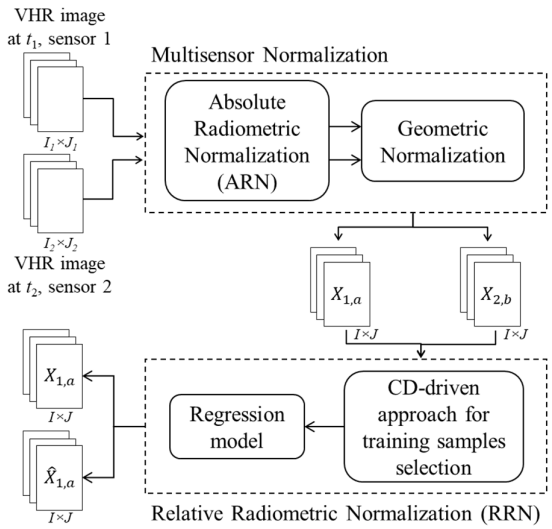


Figure 4.2. Block scheme of the proposed method for generation of homogeneous VHR TS from multisensor MT images.

4.4. Multisensor Normalization

A preliminary homogenization process is carried on over the VHR multisensor images in order to mitigate differences induced by the use of multisensor acquisitions that result in both spectral and geometrical distortions [30]. In doing this, we guarantee a similarity, from the physical and geometrical viewpoints. To achieve this goal, two main steps are followed: i) ARN mitigation and ii) geometric normalization (see Figure 4.2).

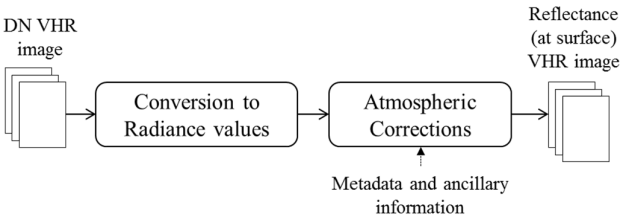


Figure 4.3. Block scheme for ARN mitigation process.

ARN mitigation process is shown in Figure 4.3. The original VHR images are first converted from DN to radiance values and then to Top Of Atmosphere (TOA) reflectance, known as at-surface reflectance. This is achieved by applying atmospheric corrections, resulting in images with the same physical meaning (further details can be found in [35] and [39]). At the end of this step, some of the radiometrical differences among multisensor MT VHR images are corrected. Nevertheless, some differences remain that cannot be corrected or mitigated by ARN. Thus, a data-driven mechanism is later used to compensate for them.

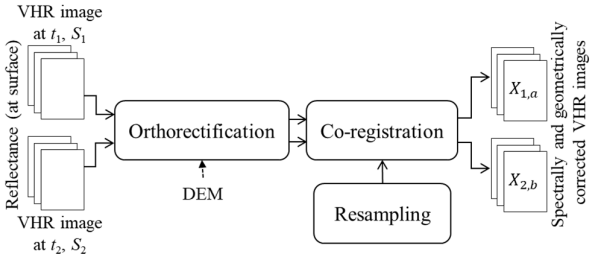


Figure 4.4. Block scheme for the geometric normalization of multisensor VHR images.

Concerning geometrical differences, two kinds of sources can be identified: i) differences in the acquisition view angle, and ii) differences in the type of sensor because of multisensor acquisitions. The former can induce differences in the images when there are small changes in the topography and relief of the terrain. Whereas the latter results in differences in the spatial resolution and therefore in the correspondence of the same spatial position in the images. To achieve the geometric normalization, the block scheme shown in Figure 4.4 is followed. Once again, this process does not fully correct for all the possible geometrical differences, but does help to mitigate for them. Surface reflectance images coming from both sensors are used as the input images for this step. They are orthorectified separately and then co-registered in order to guarantee a correspondence of each position on the ground in the MT images. The preliminary homogenized images are $X_{1,a}$ and $X_{2,b}$. Further details can be found in [35].

4.5. Relative Radiometric Normalization

With the preliminary homogenization step, we guarantee that $X_{1,a}$ and $X_{2,b}$ have common: i) spatial resolution, ii) image size $I \times J$, and iii) radiometric representation from the physical viewpoint. Yet, sensor induced differences in terms of radiometry and number of bands remain. To mitigate for these issues we perform further normalization by means of a novel RRN, based on a non-parametric regression that guarantees the finding of a model suitable for multisensor VHR images. This model jointly accounts for the number of issues arising and remaining from the spectral, radiometrical and geometrical differences. To find the model, we need stable RCSS that ensure the modeling of the complex relationship existing between S_1 and S_2. Opposite to state-of-the-art methods, RCSS are automatically selected. To this aim, a CD-driven approach is considered.

The block scheme for the proposed RRN approach is based on 4 steps (Figure 4.5): i) selection of invariant features from $X_{1,a}$ and $X_{2,b}$; ii) selection of unchanged samples, iii) sub-sample of unchanged samples; and iv) non-parametric regression based on prediction and fusion with two phases, training and recalling. Most of these steps have been separately used in literature for single-sensor and low resolution images. But they have been used in a different way.

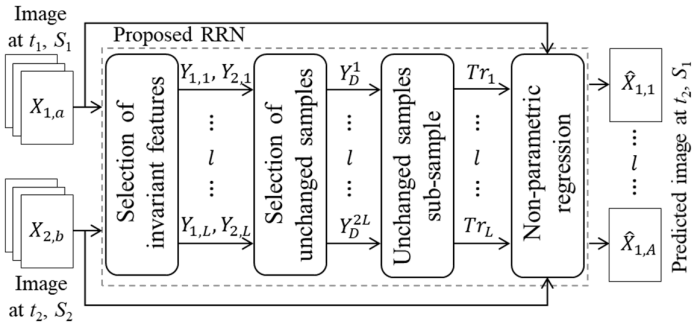


Figure 4.5. Block scheme followed for the proposed RRN process.

4.5.1 Selection of Invariant Features

In order to properly model the radiometric relationship between S_1 and S_2, only un-changed (radiometrically invariant) samples should be considered. To select them, spectral information is commonly used in literature. Nevertheless, the higher spatial details given by the use of VHR images, makes spectral information alone not enough to decide if a change has occurred or not. Spatial correlation should be used as well to decide is a pixel has indeed changed or not. In order to extract the spectral and spatial information, invariant spectral and spatial features are considered. The decision about which sample has changed or not can rely on a CD approach, where a pixel-by-pixel comparison is applied. Nevertheless, given that $X_{1,a}$ and $X_{2,b}$ show different spectral behaviors, $B > A$, the correspondence between the ath and bth bands is not

one-to-one. However, since S_1 and S_2 acquire images over the same spectral range, L couples of bands ($X_{1,l}$ and $X_{2,l}$) can be identified that show the most similar central wavelength and bandwidth, with $L \leq A < B$. From this set of most similar bands, a set of invariant features are derived for unchanged samples detection. Even though atmospheric corrections have been applied, residual differences exist among spectral channels. This results in L bands histogram shape difference and relative shift (see Figure 4.6), even in case of no-change. Thus, compensation of spectral differences is carried out on the L pairs in terms of shift. This pushes the peak values ($P_{max_{S_1}}$ and $P_{max_{S_2}}$) to the same position in the spectral domain. The shift value is extracted as $P = \left|P_{max_{S_2}} - P_{max_{S_1}}\right|$. The set of spectral ($spe$) invariant features is then extracted from the histogram compensated bands $Y_{1,l,spe}$ and $Y_{2,l,spe}$, with $l \in [1, L]$.

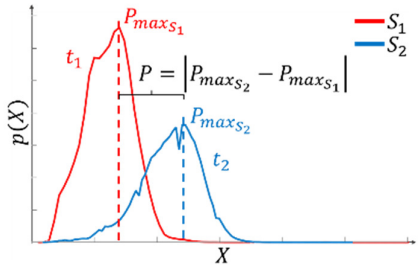


Figure 4.6. Example of histogram shift of a pair of most similar bands in two different VHR sensors.

The high spatial variability of spectral signatures in VHR images, results in the increase of single date image interclass variability and thus in the variability of both change and no change classes. This phenomenon is enhanced by the temporal spectral variability induced by the use of multisensor VHR images. This results in a higher complexity in finding un-changed training samples. The effect can be mitigated by using other features that exploit the spatial details offered by the VHR images. In the proposed method, textural features extracted from the L most similar bands are employed as a complement to select the invariant training samples. Thus, a spatial invariant features set is defined by the textural (txt) features extracted from the L most similar bands as $Y_{1,l,txt}$ and $Y_{2,l,txt}$, where $l \in [1, L]$.

4.5.2 Unchanged Samples Selection

In order to model the temporal and spectral relationship between $X_{1,a}$ and $X_{2,b}$, we select samples that are i) likely unchanged (ω_n) and ii) as much as possible representative of the spectral variability of unchanged information over the bands. The reliability of these samples is guaranteed by a CD-driven approach based on Univariate Image Difference (UID) and a conservative threshold. UID is applied to each pair of invariant features ($Y_{1,l}$,$Y_{2,l}$) as:

$$Y_D^l = \left|Y_{2,l} - Y_{1,l}\right|, l = 1, \dots, 2L \tag{4.3}$$

where Y_D^l is the absolute difference image anb $Y_{1,l} = \{Y_{1,l,spe}, Y_{1,l,txt}\}$ and $Y_{2,l} = \{Y_{2,l,spe}, Y_{2,l,txt}\}$. The most reliable unchanged samples are selected as the ones showing a small difference value, i.e., below a certain threshold value T^l. For each pair of invariant features, a different T^l is calculated. The selection criteria for the threshold should account for the greater spectral variability of the ω_n class. Thus, a proper threshold selection requires the relaxation of the criteria of the state-of-the-art methods. In greater detail, T^l should be selected in order to guarantee the presence of unchanged pixels covering the whole spectral range of samples from the study area. To this end, a combination of two stages is used: i) calculation of a reliable threshold value by means of a state-of-the-art method, and ii) definition of a more relaxed threshold. First, the optimal decision threshold (T_{RR}^l) according to the Bayesian decision theory is applied, considering a Rayleigh-Rice (RR) model [56], for each lth most similar band in the feature space. Where RR model is a

good approximation. Samples in Y_D^l close to T_{RR}^l (see Figure 4.7) are uncertain. Whereas pure unchanged pixels are needed to model the spectral relationship between $X_{1,a}$ and $X_{2,b}$ properly. Therefore, T_{RR}^l is refined to limit the presence of changed samples. The refinement process is conducted to satisfy the tradeoff between removing unchanged samples (small threshold values) and preserving the spectral variability (large threshold values) as:

$$T^l = \frac{max_{Y_D^l} - T_{RR}^l}{2}, l = 1, \dots, 2L \tag{4.4}$$

where $max_{Y_D^l}$ is the first maximum peak on the left of T_{RR}^l in the invariant feature space. Once T^l is selected, Y_D^l is thresholded to get a map of candidate unchanged samples for the training set.

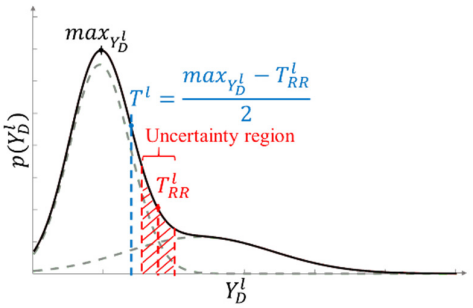


Figure 4.7. Example of a $\boldsymbol{Y_D^l}$ histogram with the selection of thresholds $\boldsymbol{T_{RR}^l}$ and $\boldsymbol{T^l}$.

4.5.3 Unchanged Samples Sub-sampling

In order to avoid over-fitting and overload while performing the non-parametric regression, detected unchanged samples are sub-sampled by following two guidelines: i) select unchanged samples common to both spectral and textural features, and ii) select samples by preserving their statistical distribution. The previous guidelines are applied separately to each pair of invariant features. The sets of unchanged samples represent the variety of spectral relationships existing among the L most similar bands from the spectral and textural viewpoint. The first guideline assures that samples are unchanged both from the spectral and textural point of view. Nevertheless, the number of training samples remains high. Thus, the second guideline aims at further reducing this number by preserving the spatial distribution of the samples over the scene uniform, as well as the statistical distribution of the original $X_{1,a}$ and $X_{2,b}$. For each pair of invariant features, the entire image is first divided into small blocks of size $R \times R$, with $R < I < J$. Then a fraction of samples from each block is selected. On top of these samples, a maximum number of pixels (Q), over all possible spectral values, are selected.

4.5.4 Non-parametric Regression

The last step of the proposed RRN approach is the non-parametric regression. Where the exact form of the nonlinear function f in equation (4.1) does not need to be known explicitly prior to model training. In order to perform the non-parametric regression, sets of training samples Tr_l ($l = 1, \cdots, L$) must be defined for each spectral channel to be predicted. Each of the Tr_l sample sets is built by considering the positions given by the sub-sample step (previous section), but in this case invariant features are extracted from all the channels in S_1 and S_2. This given the fact that several bands in $X_{2,b}$ may contribute to the spectral information in each of the bands in $X_{1,a}$. Accordingly, our set of both spectral and spatial invariant features is defined as $X_1 = \{X_{1,a,spe}, X_{1,a,txt}\}$ and $X_2 = \{X_{2,b,spe}, X_{2,b,txt}\}$.

Given the sets of Tr_l, as described above, and the sets of features X_1 and X_2, we model the temporal and spectral relationship between each $X_{1,a}$ and X_2. X_2 represents the input (also known as predictor or feature)

and $X_{1,a}$ the target (also known as output or response). The complexity of the relationship between $X_{1,a}$ and X_2 is captured by estimating f with a non-parametric regression method. Possible choices are ANN or SVR, however any other method can be used. Here we propose a configuration for the input and target training samples for predicting $\hat{X}_{1,a}$ (image acquired at t_2 by S_1). $\hat{X}_{1,a}$ contains only spectral channels, though textural information is also available, but not predicted in the proposed configuration. A total of A non-parametric regression models have to be trained, one for each of the L most similar bands, between S_1 and S_2. The target for each regression model is one of the spectral bands $X_{1,a,spe}$ to be predicted. Since several bands in $X_{2,b}$ contribute to the spectral information in each of the bands in $X_{1,a}$, all the bands in $X_{2,b}$ are used as input for the A regression models, i.e., $X_{2,b,spe}$. As per the selection of training samples, the use of spectral information alone ($X_{2,b,spe}$) is not enough to model the relationship among $X_{1,a}$ and $X_{2,b}$. Thus, textural features are also used, where $X_{2,b,txt}$ are extracted from $X_{2,b,spe}$. Thus, a total of $2 \times B$ inputs are finally used for the derivation of the A regression models. These inputs are always the same, whereas the target varies. Equation (4.5) shows the non-parametric regression model used for the training process in the case of ANN. In (4.5), H and O stand for hidden and output layers, respectively. $\varphi_H(\cdot)$ and $\varphi_O(\cdot)$ represent the activation functions. w_{bj} and w_{jk} represent the weights of the network connection in the hidden and output layers, respectively. And β_H and β_O represent the bias introduced to the transfer functions. Once the model has been obtained, all the channels in $\hat{X}_{1,a}$ are predicted.

$$X_{1,a} = \varphi_O(\beta_O + \sum_{j=1}^{J} w_{jk}\,(\varphi_H(\beta_H + \sum_{b=1}^{2B} w_{bj}X_2))_j) \tag{4.5}$$

4.6. Multitemporal Information Extraction

Several methods for MT information extraction can be found in the literature that may benefit of the proposed approach for the generation of homogeneous TS from MT multisensor VHR images. The selection of the method to derive MT information depends on the application itself. However, here we consider CD as an option to demonstrate the effectiveness of the proposed approach. According to the-state-of-the-art, the CVA technique in a polar domain framework as proposed in [50] is used. To this end, the multispectral difference images X_D is computed as:

$$X_D = \hat{X}_{1,a} - X_{1,a} \tag{4.6}$$

where X_D is the magnitude of the change vectors. $\rho \in [0, \rho_{max}]$ (eq. (4.7)) carries information about presence/absence of changes. Small magnitude values are associated to no changes, whereas large values are associated to change. A threshold T is commonly applied to separates changed from unchanged samples [50]. Here, the optimal T is obtained according to the Bayesian decision theory by considering a Rayleigh-Rice (RR) model [56].

$$\rho = \sqrt{\sum_{a=1}^{A} (X_{D,a})^2} \tag{4.7}$$

4.7. Experimental Results

4.7.1 Dataset Description and Multisensor Normalization

In order to validate the proposed approach, pairs of VHR optical images acquired over an area located in the Trentino region in the north of Italy (Figure 4.8), were selected. This area shows interesting properties from the point of view of its orographic conformation and environmental variety. Over a relatively small

area it is possible to find: i) precious apple and vineyard fields; and ii) urban, sub-urban and industrial areas with different density and structure. Two multitemporal data sets made up of two QuickBird (QB) and one Worldview-2 (WV-2) images, were built over the sample area in Figure 4.8 (orange and blue squares). The datasets were selected such that: i) different kinds of land covers and changes are represented and ii) the performance of the proposed approach is demonstrated under constrained situations. Therefore, the datasets show a small amount of water (dataset 1 only) from the river passing nearby, transitions among different phenological states in crop areas and transitions from vegetation to bare soil and to new man-made objects (i.e., construction of a new road and a new building). Details about the 2 datasets are given in Table 4.1. Dataset 2 falls inside the area of dataset 1, but t_1 image is acquired in 2005 instead of 2006. The reasons to make such selection are to show the robustness of the method: i) when less spectral information is available and ii) when the images are acquired in different seasons. The former decreases the probability to model the relationship between t_1 and t_2. And the latter increases the spectral variability among images.

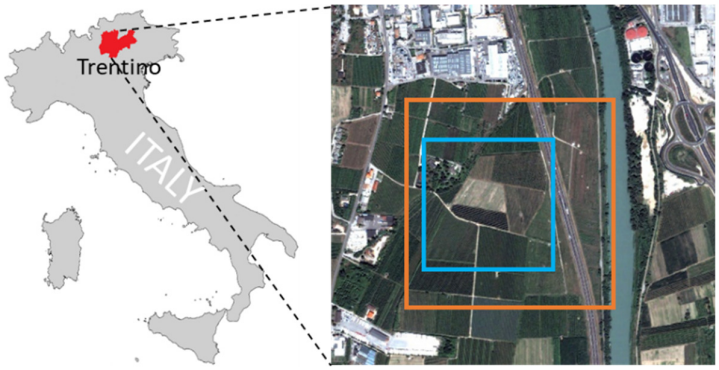


Figure 4.8. Area of interest, Trentino region in the North of Italy.

Table 4.1. Dataset Description.

	Dataset 1		**Dataset 2**	
	t_1	t_2	t_1	t_2
Sensor	QB	WV-2	QB	WV-2
Acquisition date	July 2006	August 2010	October 2005	August 2010
Off-nadir angle	14.1°	19.3°	9.8°	19.3°

Table 4.2. Main characteristics of QuickBird and WorldView-2 optical sensors [274].

Satellite	**QuickBird**	**WorldView-2**
Bands (nm)	445-900 (pan)	450-800 (pan)
		400-450 (coastal)
	450-520 (blue)	450-510 (blue)
	520-600 (green)	510-580 (green)
		585-626 (yellow)
	630-690 (red)	630-690 (red)
		705-745 (red edge)
	760-900 (NIR)	770-895 (NIR 1)
		860-1040 (NIR 2)
Spatial Resolution (m)	0.61	0.46
	2.44	1.84

The QB image has four multispectral bands ($A = 4$), whereas WV-2 has eight bands ($B = 8$). The spatial resolution of the QB image is 0.6m for the panchromatic band and 2.4m for multispectral bands, whereas

WV-2 offers a higher spatial resolution in both panchromatic and multispectral bands with 0.5m and 2m, respectively. Table 4.2 summarizes the characteristics of QB and WV-2 images from the spectral and spatial point of view. The spatial resolution differences imply that the sizes $I_1 \times J_1$ and $I_2 \times J_2$ of QB and WV-2 images, respectively, are different despite they cover the same surface. Thus, pixel-by-pixel comparison cannot be directly applied since the same pixel coordinates in the two images do not correspond to the same position on the ground. Concerning spectral domain, we can observe that the four primary multi-spectral bands of QB and WV-2 are acquired over similar spectral ranges (e.g., red), but not identical (e.g., blue). Similar considerations hold for bands green and NIR. Given that bands for both sensors fall into the same ranges, they can be compared and used to perform the regression. In order to apply the proposed method for generation of homogeneous VHR TS, multisensor homogenization and RRN were applied.

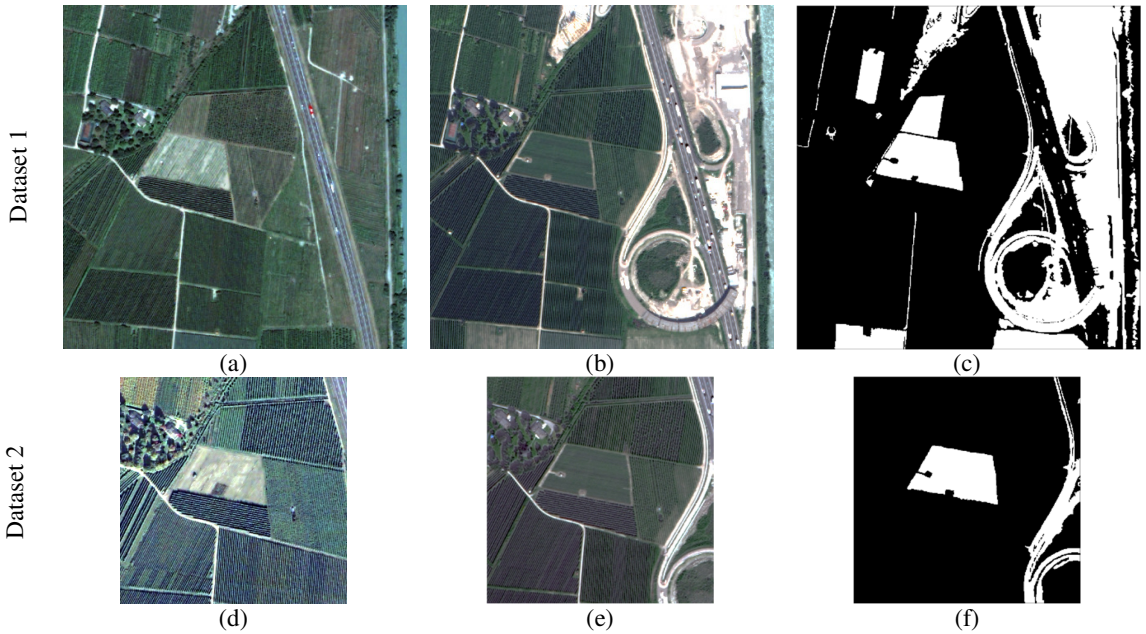


Figure 4.9. True color composition of the pansharpened multispectral multisensor VHR datasets: (a) QB image acquired in July 2006 and (d) October 2005 and; (b), (e) WV-2 image acquired in August 2010. (c), (f) Reference maps.

Multisensor homogenization was performed by: i) absolute radiometric normalization and ii) geometric normalization. For ARN step, all images were provided by DigitalGlobe Foundation in the context of the "MS-TS – Analysis of Multisensor VHR image Time Series" project [275]. Conversion from DNs to TOA was conducted before delivery by means of their Atmospheric Compensation (AComp) algorithm, which corrects for the scattering and absorption effects in the atmosphere [30], [276], [277]. Given the orography of the study area and the possible distortions, we applied orthorectification by using a DEM (1m) obtained from LiDAR data [278]. Additional pixel-to-pixel problems are also observed due to differences in the view angle, and co-registration should be applied. In order to achieve a better co-registration, PanSharpening (PS) was applied by means of the Gram-Schmidt method. Here ENVI software package was employed [279]. After PS, the spatial resolution for QB and WV-2 multispectral bands is 0.6m and 0.5m, respectively. Co-registration of the two QB and WV-2 pairs, covering the whole study area in Figure 4.8 was conducted by using a polynomial function of second order. For the QB 2006 and WV-2 2010 couple (dataset 1), 79 uniformly distributed Ground Control Points (GCP) were selected. Whereas 51 uniformly distributed GCP were selected for the QB 2005 and WV-2 2010 couple (dataset 2). The WV-2 image was resampled during co-registration in order to be able to extract MT information. Resampling was performed by means of the nearest neighbor interpolation. Figure 4.9 shows the pansharpened multisensor VHR QB and WV-2 images after preliminary normalization. Both datasets show a common spatial resolution of 0.6m, but dataset 1 is larger (1024x1024 pixels) than dataset 2 (640x640 pixels). In order to perform qualitative and quantitative

analysis, a reference map, for each dataset, was defined by photointerpretation and a priori knowledge on the scene (Figure 4.9 third column), showing 706364 and 357630 unchanged pixels (black color) and 334469 and 51970 changed pixels (white color), respectively.

4.7.2 Relative Radiometric Normalization and Design of Experiments

Once normalization is achieved, we move to the proposed RRN. The first step is to select invariant features among the two sensors. To this end, a total of $L = 4$ most similar bands among S_1 and S_2 are first identified (see Table III). Then, spectral and textural features are extracted. For spectral features, the $L=4$ pairs of most similar bands were compensated based on histogram shift value P (see Table 4.3). WV-2 image is compensated to QB one in each dataset. In the case of textural features, well-known state-of-the-art features, such as Gray Level Co-occurrence Matrix (GLCM) [301] and Gabor ones [302], were tested under different configurations. GLCM-contrast showed the best results in terms of normalization and CD results, and was thus selected as the invariant feature.

Table 4.3. Pair of most similar bands for QB and WV-2 and histogram compensation value (P) for datasets 1 and 2.

QB (nm)	WV-2 (nm)	Dataset 1	Dataset 2
450-520 (B1)	450-510 (B2)	0.020	0.023
520-600 (B2)	510-580 (B3)	0.016	0.016
630-690 (B3)	630-690 (B5)	0.023	0.020
760-900 (B4)	770-895 (B7)	0.145	0.090

The next steps are the selection of unchanged samples and their sub-sampling. The first step applies the absolute UID as in (4.3), and then T^l is calculated as in (4.4). The T^l values for each invariant features (spectral and textural), as well as the number of training samples before ($\left|Y_D^l < T^l\right|$) and after ($|Tr_l|$) the sub-sampling step are given in Table 4.4 for both datasets. The size of the blocks for the spatial reduction was $R = 64$, with a fraction of training samples of 30% in both datasets. When no samples are found in the block, no operation is carried on. A maximum of $Q = 100$ samples per reflectance value (with a precision of 3), over the whole spectral range were selected. The selection of these parameters is related to: i) the computational power and ii) the conservation of spectral variability to model the relationship between S_1 and S_2. Bigger numbers result in the selection of more samples, which in turn results in a higher computational power and a higher probability of overfitting while applying the non-parametric regression. Lower numbers result in the selection of less samples, which in turn results in lower computational power and lower spectral representation to model the S_1 and S_2 relationship.

Table 4.4. Threshold value (T^l) and number of training samples before ($\left|Y_D^l < T^l\right|$) and after ($|Tr_l|$) sub-sampling step.

Band	**Feature**	**Dataset 1**			**Dataset 2**		
		$\boldsymbol{T^l}$	$\boldsymbol{\left|Y_D^l < T^l\right|}$	$\boldsymbol{\|Tr_l\|}$	$\boldsymbol{T^l}$	$\boldsymbol{\left|X_{D,l} < T_l\right|}$	$\boldsymbol{\|Tr_l\|}$
Blue	Spectral	0.012	616739	4998	0.015	258524	4151
	Texture	0.040	242034		0.040	112858	
Green	Spectral	0.013	560094	5436	0.019	255086	5323
	Texture	0.030	158496		0.038	113545	
Red	Spectral	0.014	528475	6155	0.015	227125	6196
	Texture	0.030	161157		0.038	113635	
NIR	Spectral	0.025	265267	8980	0.021	56017	5280
	Texture	0.030	157491		0.039	112884	

The last step of the proposed RRN is the non-parametric regression, which was carried on by means of two methods, ANN and SVR. A total of 4 regression models were derived, one for each couple of most similar bands. The input data corresponded to the spectral and textural features of the WV-2 images, for a

total of 16 input variables, whereas the target corresponded to the specific spectral band to be predicted. Final configuration for ANN and SVR were the same for both datasets.

1. *For ANN*, i) one input layer with 16 neurons (one for each WV-2 spectral and textural bands), ii) three hidden layers with 15, 8 and 10 neurons having tan-sigmoid, log-sigmoid and pure-line activation functions, respectively, and iii) one output layer with one neuron having a linear activation function, were used. The number of neurons in the input and output layers are the same as the number of input and output features. The best configuration for the number of hidden layers and their neurons was selected based on the Mean Squared Error (MSE). Number of hidden layers varied from 1 to 5, whereas the number of neurons varied from 8-20.
2. *For SVR*, a similar configuration as for ANN was held, with 16 variables for the input and one variable for the target. An RBF kernel was used and the corresponding parameters for the training process were gotten by cross-validation process. Ranges for the parameters were: $C = [10{,}500]$, $gamma = [0.1{,}3.0]$ and $eps = [0.01{,}0.1]$.

The final MSE and computational times, calculated between the QB image $X_{1,a}(t_1)$ and the predicted QB $\hat{X}_{1,a}(t_2)$ by the 4 ANNs and the 4 SVRs, are provided in Table 4.5. In the specific case of SVR, the training time includes both training and cross-validation. The computational times were obtained by using MATLAB® on a standard workstation with hardware Intel(R) Xeon(R) CPU @3.40 GHz, 16.00 GB RAM. The MSE is quite similar for the two non-parametric regression methods, but the time required for the training process differs significantly. In the case of SVR, it depends on the number of training samples. Comparing $|Tr_l|$ shown in Table 4.4 with the training times in Table 4.5, we can easily conclude that the training time increases with $|Tr_l|$ in a nonlinear way.

Table 4.5. Final MSE and computational time of ANN and SVR for datasets 1 and 2.

Spectral Band	Dataset 1				Dataset 2			
	ANN		SVR		ANN		SVR	
	MSE (10^{-5})	Time	MSE (10^{-5})	Time	MSE (10^{-5})	Time	MSE (10^{-5})	Time
Blue	2.2190	1-4 s	3.6822	0.073 h	6.0812	1-4 s	6.4367	1.099 h
Green	3.2567		4.5511	0.452 h	8.9224		10.3898	2.811 h
Red	4.9059		6.0398	1.446 h	12.4701		13.9383	4.076 h
NIR	15.192		16.721	9.006 h	7.1271		7.4896	2.474 h

The performance of the proposed approach was evaluated by qualitative and quantitative analysis. In the qualitative case, visual comparison of the original and predicted images was carried on and histograms were extracted over unchanged areas to assess the similarity of original ($X_{1,a}$) and predicted ($\hat{X}_{1,a}$) images. In the quantitative case, two approaches are used: i) calculation of Kullback-Leibler (KL) distance between the histograms of unchanged areas [303] and; ii) MT information extraction, change detection by means of CVA-magnitude thresholding. In the KL distance, we expect histograms from the predicted images ($\hat{X}_{1,a}$) to be closer to those of the QB ones ($X_{1,a}$). For the CD based validation, three experiments were designed: i) experiment 1 (exp. 1) applies CVA to the preliminary normalized images ($X_{1,a}$ and $X_{2,b}$), ii) experiment 2 (exp. 2) applies CVA to the original image ($X_{1,a}$) and predicted image obtained with ANN ($\hat{X}_{1,a_{ANN}}$); and iii) experiment 3 (exp. 3) applies CVA to the original image ($X_{1,a}$) and predicted image obtained with SVR ($\hat{X}_{1,a_{SVR}}$). In exp. 2 and 3, the four spectral bands are used for computing the magnitude variable (see eq. (4.7)). Whereas in the case of exp. 1 the most similar bands are used.

4.7.3 MT Information Extraction and Experimental Results

For the qualitative assessment, the original QB images are first compared, to the ANN and SVR predicted ones for 2010. Figure 4.10 shows true color composition of the original QB ((a) and (d)) and the predicted

QB images by ANN ((b) and (e)) and SVR ((c) and (f)). The comparison between the predicted QB image and the original QB one, points out that they like each other (except for changed areas). On the other hand, and despite changed areas are not represented in the training sets, a visual comparison of the predicted images ($\hat{X}_{1,a}$) with the original WV-2 image ($X_{2,b}$), shows that the proposed approach allows for a proper prediction the latter ones. Even though atmospheric corrections are applied, some small areas show less similarity due to the saturation induced by highly reflective surfaces, present in the original images.

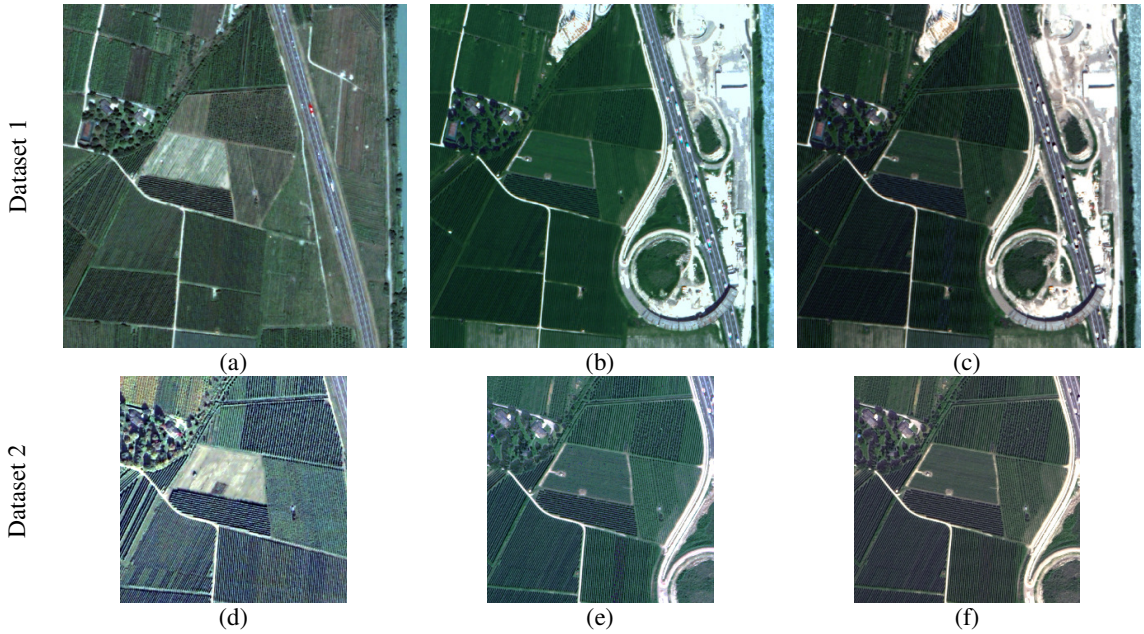


Figure 4.10. True color composition of (a) QB 2006 and (d) QB 2005 images; and QB 2010 predicted images by (b), (e) ANN; and (c), (f) SVR for datasets 1 and 2, respectively.

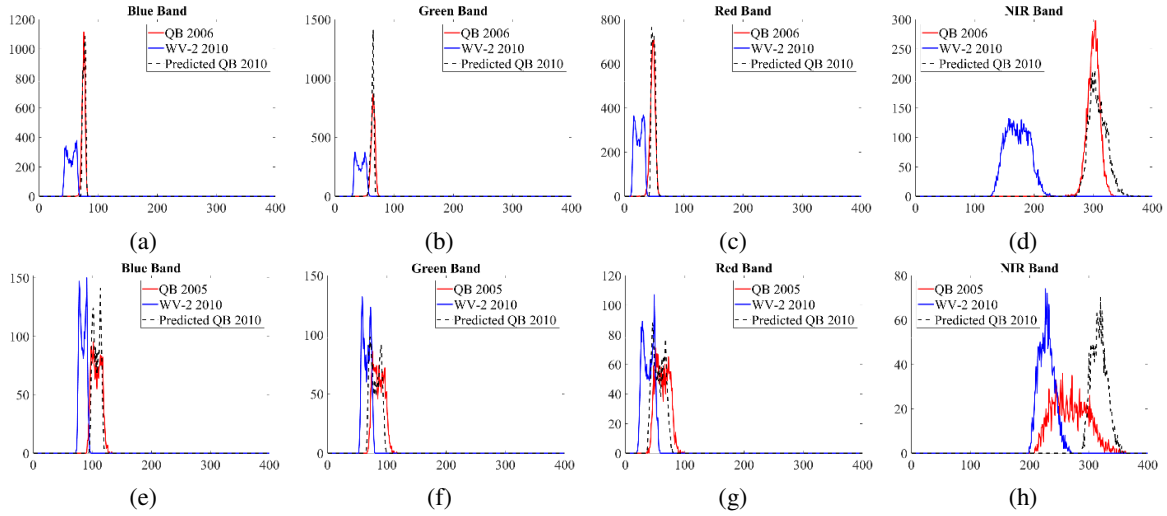


Figure 4.11. Histograms of unchanged areas for $X_{1,l}$, $X_{2,l}$ and $\hat{X}_{1,a_{ANN}}$ for datasets 1 (top) and 2 (bottom).

In order to further corroborate the previous observation, histograms of unchanged areas were compared. It is expected that after prediction, $\hat{X}_{1,a}$ histograms become closer to $X_{1,a}$ than to $X_{2,b}$, both in shape and position. On the other hand, we also expect to have a better performance in the case of dataset 1, since it has less seasonal differences. The histograms in Figure 4.11 and Figure 4.12 confirm this, as well as the goodness of the proposed method. As expected, for each most similar spectral pair, the histograms of the

$\hat{X}_{1,a}$ spectral bands predicted from the WV-2 image by ANN and SVR (dashed black lines in Figure 4.11 and Figure 4.12) have behaviors more similar to the original QB images (red line in Figure 4.11 and Figure 4.12) rather than to the WV-2 ones (blue line in Figure 4.11 and Figure 4.12).

The above is quantitatively corroborated by the KL distance shown in Table 4.6. In the specific case of ANN histograms, they are more similar to the original QB images, both in shape and position, with the exception of NIR band in dataset 2. In the case of SVR histograms, one can see that the predicted images have changed their histogram in shape and position, but not as much as in the ANN case. The problem with NIR band in dataset 2 remains as for ANN case. The main difference between ANN and SVR results relies in the fact that SVR is not fully able to model the shape of the histogram, which remains still more similar to that of the original WV-2 image. Even though the proposed approach is able to mitigate for most of the intrinsic differences introduced by the complexity of working with MT multisensor VHR images itself, some differences in shape and position can be still seen. Regarding the dataset 2, Table 4.6 allows us to see how similar the results between ANN and SVR are. As expected from the beginning, the use of images acquired at different seasons introduces further spectral variability among the images resulting in the lack of spectral information to model the relationship between the two sensors. This is reflected in the poor performance of the NIR band that has a higher response to vegetation.

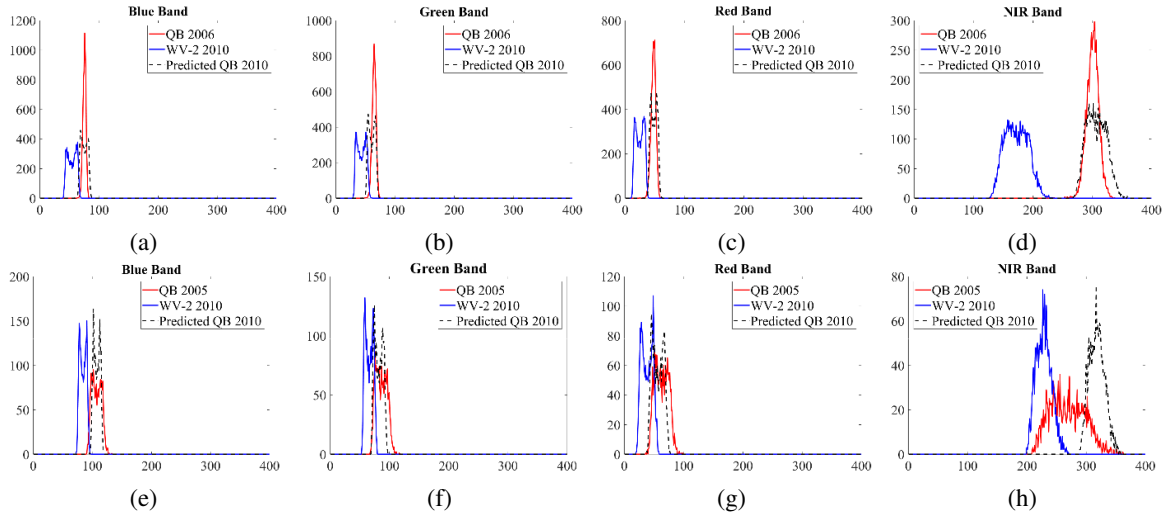


Figure 4.12. Histograms of unchanged areas for $X_{1,l}$, $X_{2,l}$ and $\hat{X}_{1,a_{SVR}}$ for datasets 1 (top) and 2 (bottom).

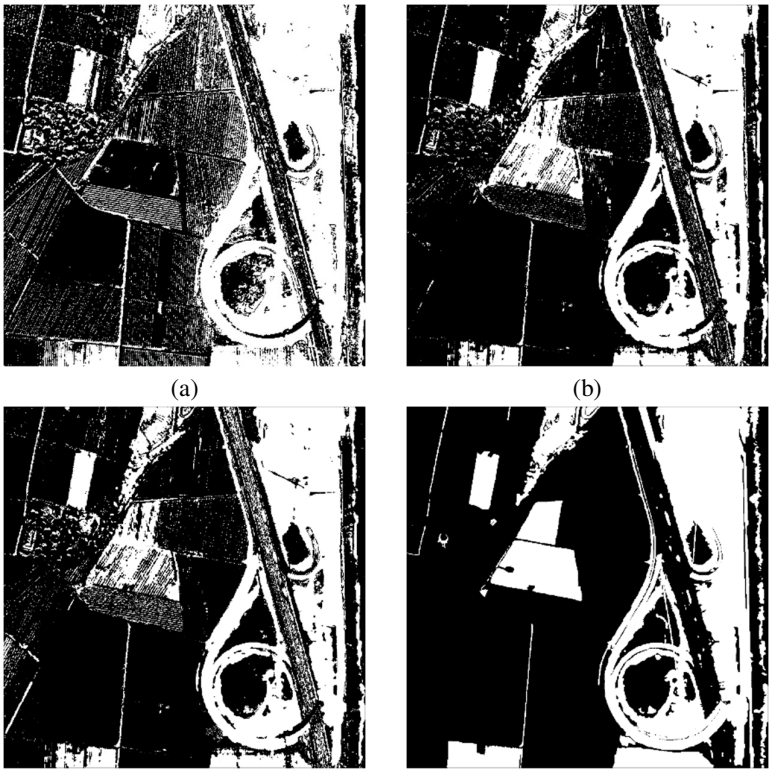
Table 4.6. KL distance between the histograms of unchanged areas for datasets 1 and 2.

Spectral Band	Dataset 1				Dataset 2			
	$\hat{X}_{1,a_{ANN}}$		$\hat{X}_{1,a_{SVR}}$		$\hat{X}_{1,a_{ANN}}$		$\hat{X}_{1,a_{SVR}}$	
	$X_{1,l}$	$X_{2,l}$	$X_{1,l}$	$X_{2,l}$	$X_{1,l}$	$X_{2,l}$	$X_{1,l}$	$X_{2,l}$
Blue	2.65	83.43	11.87	69.05	0.75	18.14	1.64	17.82
Green	3.73	79.76	9.55	55.20	1.70	9.18	1.66	11.24
Red	3.36	81.66	3.44	77.94	1.43	8.39	1.67	7.80
NIR	3.78	63.26	5.16	62.50	7.41	14.41	7.42	14.40

For the qualitative analysis, CD maps are derived by means of CVA and as defined in exp. 1, 2 and 3, by means of equation (4.7). Areas corresponding to radiometric changes were extracted by thresholding the magnitude variable. *T* was automatically selected by applying the method in [56]. The *T* values for exp. 1, 2 and 3 for dataset 1 were 0.1350, 0.0939 and 0.1260, and for dataset 2 where 0.0826, 0.1220 and 0.1294, respectively. A comparison of the CD maps with the reference maps (see Figure 4.13 and Figure 4.14) pointed out the improvement achieved when working with the proposed approach, both in ANN and SVR regression cases, specifically on the reduction of False Alarms (FA) and Missed Alarms (MA) for dataset

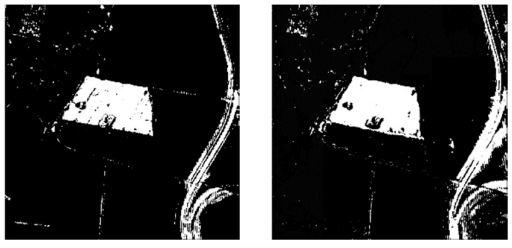
1 and the reduction of MA for dataset 2. In the case of dataset 1, these improvements are related to the identification of the two specific changes located in the central part of the study area, which were not identified by exp. 1. Whereas in the case of dataset 2, the improvements are related to a better detection of the change due to the new road.

Table 4.7 shows the detailed results, where for dataset 1 the number of FA identified by exp. 1 decreased almost of 67% when ANN predicted image was used and about 61% when SVR one was used. These results are reflected in the OA, where exp. 2 gained about 11% over exp. 1 and 1% over exp. 3; and exp. 3 gained about 10% over exp. 1. For dataset 2 the number of FA decreased of about 67% in exp. 2 and 71% in exp. 3. This results in the increase of the OA for both experiments in about 2%. Here the OA increase from exp.1 to exp. 2 and 3 is lower than in dataset 1 due to differences in the seasonal acquisition. Yet the proposed approach is able to improve the CD results. The previous results prove the effectiveness of the proposed approach for the two non-parametric regression methods used to derive the models for the homogenization of multisensor MT VHR images. Furthermore, based on the qualitative and quantitative analysis, as well as the time required for the training process in both datasets, we can see that for the considered images it is slightly better to use ANN as the non-parametric regression method. However, other regression methods can be considered.

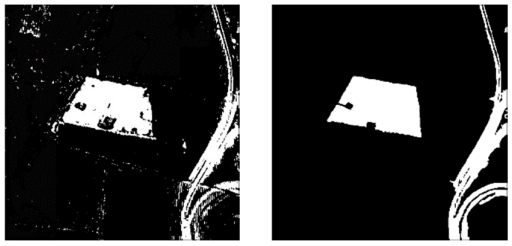


(a) (b) (c) (d)

Figure 4.13. Binary CD maps obtained by CVA for a) exp. 1, b) exp. 2 and c) exp. 3. d) Reference map (dataset 1).



(a) (b)



(c) (d)

Figure 4.14. Binary CD maps obtained by CVA for a) exp. 1, b) exp. 2 and c) exp. 3. d) Reference map (dataset 2).

Table 4.7. False Alarms (FA), Missed Alarms (MA), Overall Error (OE) And Overall Accuracy (OA) for the proposed approach over the three experiments (datasets 1 and 2).

Exp.	Dataset 1				Dataset 2			
	FA	**MA**	**OE**	**OA (%)**	**FA**	**MA**	**OE**	**OA (%)**
1	145477	68659	214136	79.58	4432	12339	16771	95.90
2	47210	52046	99256	90.53	1536	6713	8249	97.99
3	56383	57454	113837	89.14	1394	8645	10039	97.55

4.8. Conclusion

In this chapter, a method for generation of homogenous VHR TS based on non-parametric regression of multisensor MT optical images has been proposed. The proposed method achieves the goal while effectively normalizing the multisensor images by combining a preliminary normalization (based on ARN) with a proposed RRN method. The proposed RRN method is based on a CD-driven approach and a non-parametric regression model. In RRN, a prediction is achieved by deriving single band models with non-parametric regressions (ANN and SVR) that allows to capture the spatial/spectral variability of VHR multisensor images. The models allowed us to predict all the spectral bands for the $X_{1,a}$ image, as if they would have been acquired at t_2, based on the original $X_{2,b}$ image bands and the reference un-changed pixels. The proposed CD-driven approach for training samples selection showed to be effective from the MSE view point, as well as from the MT information extraction one. Experimental results on real datasets, made-up of VHR bi-temporal and multisensor optical images, confirmed the effectiveness of the proposed approach and the improvement in multitemporal analysis that can be achieved by using the predicted images over the preliminary normalized ones. A major improvement is observed when ANN predicted image is used, but a high improvement is visible while using SVR predicted one as well. The previously mention improvements are independent of the size of the changed areas in the scene. In fact, dataset 1 shows a 32% of changed areas against a 13% in dataset 2. Yet, improvements on the CD process can be seen in both cases, showing the robustness of the proposed approach.

The selection of other spectral and/or textural features for the training stage could be considered to further improve the current CD OA. Future works consider the use of more VHR multisensor optical pairs with more complex scenarios. A further analysis can be carried on by using Deep Neural Network non-parametric regression to achieve both the normalization and CD processes at the same time. A possible extension of the proposed approach to generate homogeneous multisensor VHR time series is foreseeing. Here, the reference image remains the one acquired at t_1, whereas each subject image is the one acquired at a given time t_n with $n = 1, \dots, N$. Since the same combination of multisensor pairs may happen several times over the whole time series, this information should be saved in order to used such points to improve the training process. This will also help in the cases where certain land covers are not present from one time to the other.

Chapter 5

5. An Approach to Multiple Change Detection in VHR Optical Images Based on Iterative Clustering and Adaptive Thresholding[5]

This chapter describes an approach for multiple change detection in VHR optical images (single- and multi-sensor) based on iterative clustering and adaptive thresholding. Problems arising from assumptions on the number of classes and their distribution are discussed. The proposed approach is developed in 3 steps: (1) separation among changes and no-changes by means of a threshold, (2) detection of multiple changes by means of an iterative clustering and, (3) tuning of the separation of changes and no-changes by tuning the threshold for each type of change found in the step 2. The step 2 is distribution free and the whole method is automatic.

5.1. Introduction

When dealing with VHR optical images, one of the most common approaches to multiple Change Detection (CD) is the computation of the multispectral difference image by means of Change Vector Analysis (CVA) [50], [52], [273]. CVA uses two or more spectral channels to map both the magnitude and the direction of change. To do this, samples statistical distribution is usually analyzed in (hyper-) spherical coordinates where, in the simplest case, the magnitude is represented by ρ and the direction by θ [50], [52], [273]. Figure 5.1 shows a 2D-CVA toy example, where the main characteristics of changed and unchanged samples in a polar domain (magnitude and direction) can be observed:

- Un-changed samples tend to cluster near the origin (light gray circle);
- Pixels of the same kind of change tend to cluster and pixels of different kind of changes tend to locate in different portions of the polar domain (colored clusters);
- Different kinds of change may show stronger or weaker magnitude (green vs red cluster);
- Different kinds of change may overlap (orange and blue clusters).

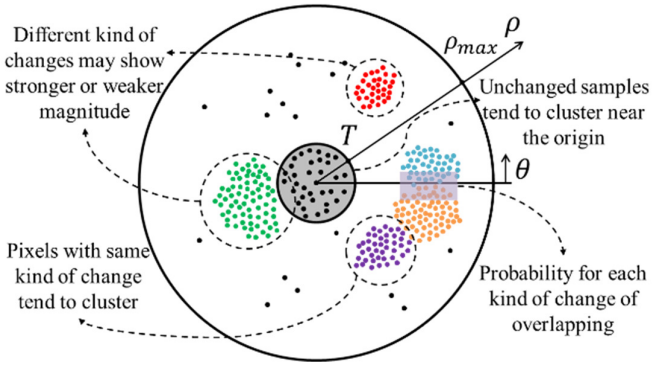


Figure 5.1. 2D-CVA toy example representation of a difference image in polar coordinates.

[5] Part of this chapter appears in:
Journal paper: Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "An approach to multiple change detection in VHR optical images based on iterative clustering and adaptive thresholding," under revision in *IEEE Geosci. Remote Sens. Lett.*.
Conference paper: Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "An approach to multiple Change Detection in multisensor VHR optical images based on iterative clustering," in 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 5149–5152.

In order to separate changed from unchanged samples, sub-optimal solutions are often employed that identify decision thresholds along the magnitude and direction variables independently [50]–[52]. A threshold over the magnitude variable provides separation between changed and unchanged samples. Whereas multiple thresholds along the direction variable provide separation among different kinds of change. However, such thresholds are complex to define. The accuracy of these sub-optimal methods depends on the a priori knowledge of both class statistical models and expected number of changes in order to define the thresholds [50]–[52]. Moreover, when the number of dimensions increases, sub-optimal solutions become less effective as they do not fully exploit the correlation among variables. Two main limitations arise from the above-mentioned assumptions. The first one is given by the definition of the threshold value along the magnitude, which is defined by considering all kinds of change as a single large metaclass [50], [53]–[57]. This implies that the threshold along the magnitude is defined according to the average properties of the different kinds of change and does not gather the peculiarities of each of them. The second one is given by the definition of different thresholds along the direction variable, which is usually defined by assuming a known statistical distribution and number of classes [50]–[52].

In order to overcome the two main limitations, methods based on adaptive thresholding or clustering have been designed [52], [57], [58], [95], [124], [304], [305]. According to the literature, and under the simplifying assumption of independence between magnitude and direction variables, two main steps are usually applied [50], [56]. The first step aims at distinguishing changed from unchanged pixels by defining a single threshold value along the magnitude variable. The second one separates different kinds of change from each other by defining threshold values along the direction(s). Both steps can be addressed by using thresholding techniques available in the literature [52], but specific solutions are required depending on the type of data (e.g., optical, SAR) and application (e.g., bi-temporal or multitemporal analysis, short or long term analysis). Sometimes [57], a third step is considered that adapts the threshold along the magnitude to the characteristics of different kinds of change identified in step 2. Nevertheless, the problem remains about how to properly detect the number of different kinds of change and how to separate them without prior information on their number and statistical distribution.

To overcome these problems, methods based on clustering can be considered [54], [95], [124], [304]. There are mainly two kinds of clustering algorithms [305]: i) partitioning; and ii) hierarchical methods. The former rely on the centroid-based model and partition the data into k clusters (e.g., k-means). Such methods often exploit a predefined statistical model for the classes and assume the knowledge of the number of clusters. Thus, they show similar drawbacks to the sub-optimal thresholding solutions. Whereas the latter (e.g., Density Based Spatial Clustering of Applications with Noise (DBSCAN) [306]) build a cluster hierarchy based on a tree structure. Tree construction may follow: i) bottom-up; or ii) top-down paradigms. The top-down approach requires at every step the identification of which clusters to split and how to do it. The bottom-up approach requires the definition of a cluster similarity measure for merging. Hierarchical clustering is able to identify clusters with different size and densities, and is sensitive to outliers.

One of the most well-known hierarchical clustering is the Ordering Points to Identify the Clustering Structure (OPTICS) [307]. OPTICS was proposed in the literature in order to overcome two specific drawbacks: i) to allow to search for clusters with different densities among the same set of points; and ii) to be able to deal with sparse data. OPTICS performs well when there are large differences in cluster densities and therefore is able to deal with sparse data. Accordingly, the use of OPTICS for solving the problem of multiple CD in VHR images in the (hyper-) spherical domain is promising. However, the overlapping of change clusters reduces OPTICS performance and increases computational time. This leads to the need of defining novel efficient methods in which threshold values (in the magnitude domain) and change sectors (in the direction domain) are adaptively defined.

In this chapter we propose an approach to multiple CD that jointly exploits the histogram distribution of the magnitude and direction variables in (hyper-) spherical coordinates. The approach is distribution free

(in the direction domain) and thus particularly suitable for VHR images. It first performs a rough discrimination between unchanged and changed samples according to the properties of the magnitude variable. Then, it automatically identifies the sectors along direction variables associated to different kinds of change by iteratively performing density-based clustering and region growing [58], [59], while adding samples from around the joint magnitude-directions histogram peaks. Finally, it tunes the magnitude threshold by adapting it according to the characteristics of different kinds of change. CD is achieved in an unsupervised and application-independent way.

The rest of this chapter is organized as follows. Section 5.2 describes the proposed method giving details on each of its three steps. Section 5.3 presents the datasets and experimental results. Finally, Section 5.4 draws the conclusions and future developments of this work.

5.2. Proposed Method for Multiple Change Detection by Iterative Clustering and Adaptive Thresholding

Let us consider two VHR optical images, X_1 and X_2, acquired over the same geographical area by sensors S_1 and S_2 at time t_1 and t_2, respectively. S_1 and S_2 can be the same or different sensors (e.g., GeoEye, QuickBird, WorldView, Pleiades). Let $X_{1,a}$ and $X_{2,b}$ be the images associated to spectral bands a ($a = 1,2,\dots,A$) and b ($b = 1,2,\dots,B$) for S_1 and S_2. a and b may be associated to the same wavelength or not depending on the sensor properties. Let $I_1 \times J_1$ and $I_2 \times J_2$ be the size of X_1 and X_2, respectively. Let X_D be the multispectral difference image obtained by subtracting on a pixel base the spectral feature vectors. Let ω_n be the class of unchanged pixels and $\Omega_c = \{\omega_{c_1}, \omega_{c_2}, \dots, \omega_{c_K}\}$ the metaclass of K possible kinds of change. Thus, $\Omega = \omega_n \cup \Omega_c$ is the set of classes to be identified by the proposed approach.

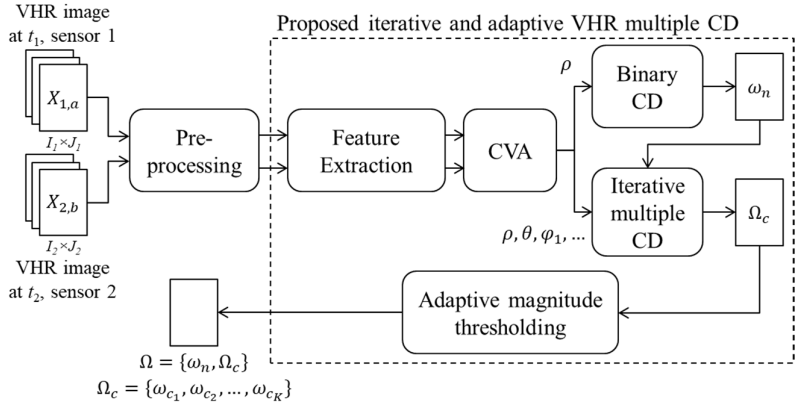


Figure 5.2. Block scheme of the proposed approach to multiple CD in VHR images.

Figure 5.2 depicts the block scheme of the proposed approach to multiple CD in VHR images based on iterative clustering and adaptive thresholding. The pre-processing step includes radiometric and geometric corrections to mitigate differences among the images not related to changes [5], [39], [62]. This step is particularly critical if multisensor optical sensors are considered. Feature extraction is then carried out in order to improve the representation and the separation of the multiple changes present in the study area [5]. The well-known CVA is used to highlight the presence/absence of changes. For sake of simplicity and visualization purposes, CVA is often applied to a 2D space [50], [56]. Employing a higher number of dimensions increases the change representation capabilities, but also the complexity due to density and sparsity of clusters. Here we consider a 3D problem; however, the extension to a higher number of dimensions is straightforward. The representation of X_D in spherical coordinates is given by equations (5.1)-(5.3) [273].

$$\rho = \sqrt{X_{D,1}^2 + X_{D,2}^2 + X_{D,3}^2} \tag{5.1}$$

$$\theta = atan2\left(\frac{X_{D,2}}{X_{D,1}}\right)^{6} \tag{5.2}$$

$$\varphi = arcos\left(\frac{X_{D,3}}{\rho}\right) \tag{5.3}$$

where ρ represents the magnitude variable, and θ and φ the direction variables (the number of direction variables increases as the space dimensionality does). Once the (hyper-) spherical space is defined, 3 steps are considered (see Figure 5.3). Step 1 separates unchanged from changed samples by means of the magnitude variable (binary CD). Step 2 iteratively identifies multiple changes by means of a hierarchical algorithm applied to the direction variables (iterative multiple CD). And step 3 adaptively adjust the magnitude threshold values by considering each change class separately (adaptive magnitude thresholding).

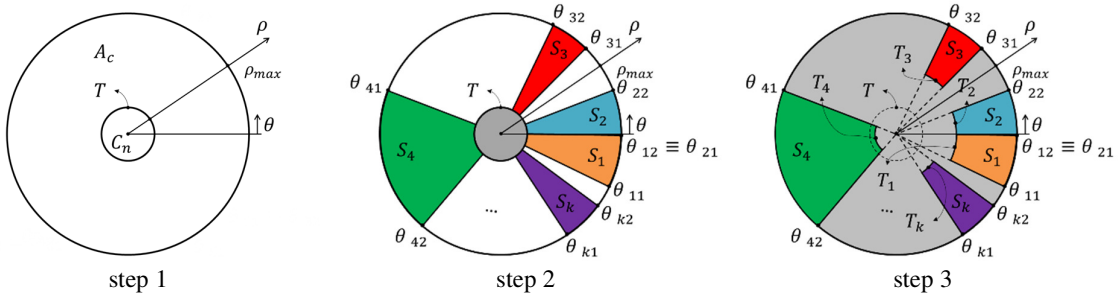


Figure 5.3. Illustration of the 3-step procedure for the detection of multiple changes in a CVA polar coordinate system (zenithal view of the azimuth and magnitude variables) [57].

5.2.1 Step 1: Binary CD

In order to distinguish among different kinds of change, unchanged samples must be first removed. In the literature, it has been demonstrated that in a (hyper-) spherical domain unchanged pixels tend to cluster around the origin of the coordinate system, i.e., they show a small magnitude $\rho \leq T$, where T is a decision threshold separating changed from unchanged samples [50]. T can be calculated by means of the Bayesian decision theory by estimating class statistical parameters using the Expectation-Maximization (EM) algorithm [308]. This can be done by following a Gaussian ([53], [308]), Rayleigh-Rice ([56]) or a compound multiclass statistical model ([309]). In step 1, the statistical distribution of ρ is exploited to separate the feature space of multispectral change vectors in two portions (see Figure 5.3, step 1): 1) the circle of unchanged samples (C_n) and 2) the annulus of changed samples (A_c).

5.2.2 Step 2: Iterative Multiple CD

In step 2, we focus on A_c, where: i) changed pixels are located far from the origin, ii) different kinds of change may show stronger or weaker magnitude (ρ) and specific preferred direction (θ, φ); and iii) pixels affected by the same kind of change tend to cluster. Thus for each kind of change a cluster exists with a peak surrounded by a spread and irregular cloud of points. Change clusters may overlap. In order to distinguish among cluster corresponding to a change, an iterative clustering is applied.

The proposed method jointly exploits the capabilities of OPTICS and the well-known region growing algorithm [59] in an iterative configuration, such as that drawbacks are addressed. In the specific case of OPTICS, main problems are related to the overlap of different clusters in a cloud of points. Region growing requires the definition of seed points (from the user) in order to determine whether a neighbor pixel belongs to a given region. Once the seeds are defined, it starts to add samples from surrounding pixels in the cloud

[6] $atan2$ is a common variation on the $arctan$ function that solves the problem of $arctan$ to distinguish between diametrically opposite directions. $atan2$ range is confined in the interval $(-\pi, \pi]$.

of points, which results in problems with overlapping clusters. Here our set of cloud points is defined by the ρ, θ, φ histogram resulted from the CVA (hyper-) spherical representation of the feature space (see Figure 5.4 (a)). In our iterative clustering, instead of applying the OPTICS/region growing directly to the whole cloud of points (as traditionally done in the literature), we exploit the different frequency levels (n – see Figure 5.4 (b)) of the magnitude-direction variables. This configuration allows us to reduce the computational time, by starting from a reduced number of samples (green dots/line in Figure 5.4), and adding them as we move to lower frequency levels (pink, blue and purple dots/lines in Figure 5.4).

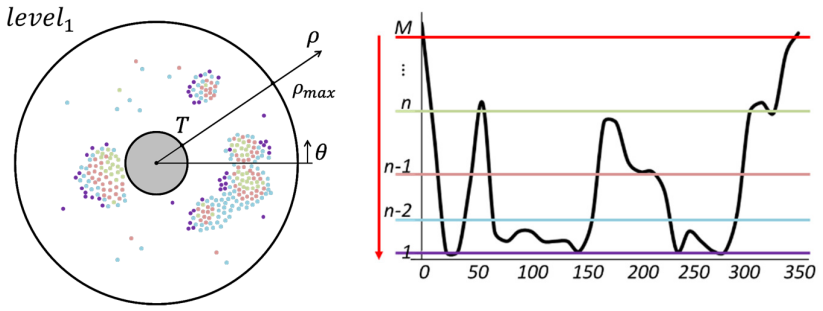


Figure 5.4. Illustration of (a) the cloud points histogram in a CVA polar coordinate system (zenithal view of the azimuth and magnitude variables) and (b) its corresponding azimuth histogram.

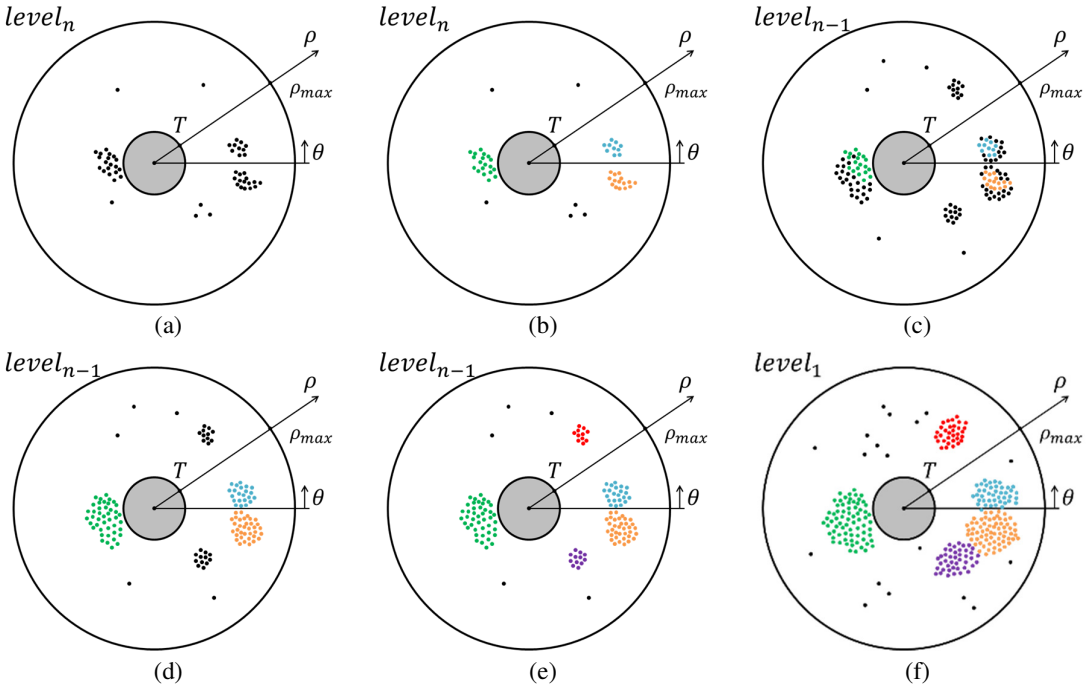


Figure 5.5. Illustration of how the proposed step-2 works (zenithal view of the azimuth and magnitude variables).

Let M be the maximum number of samples in a bin of the histogram in the (hyper-) spherical domain. First, iterative clustering finds a set of bins showing a number of samples higher than M-n ($n = 1, \ldots, M$). In this way at least one cloud of samples can be defined which is dense and isolated enough to be considered as a cluster. Whether the cluster is dense enough or not is decided by the initial conditions of the density-based clustering algorithm. Let us assume that these conditions are reached at the level n of the histogram (where $0 < n < M$), and that k=3 ($k \geq 1$) clusters are differentiated (Figure 5.5 (a)). Each of the clusters is expected to represent a different kind of change. The most populated bins are used to apply the OPTICS

density-based algorithm (Figure 5.5 (b)). OPTICS requires some initial parameters: i) the minimum number of points (*MinPts*) to form a cluster, and ii) the reachability (*Eps*), i.e., the maximum distance (radius) to which a point can be considered of the same cluster. Once the first k clusters are found, the algorithm evolves to the next frequency level (*level n*-1), i.e., to histogram bins with less number of samples (see Figure 5.4 (b)). Thus, new samples are added (black dots in Figure 5.5 (c)). Each of the existing k clusters is grown by means of the region growing [58], [59] (Figure 5.5 (d)). In order to apply the region growing algorithm, the number of cycles and the neighborhood connectivity need to be defined. Then the OPTICS algorithm is applied (with the same parameters) to search for new possible clusters (Figure 5.5 (e)). The same process is iterated until the algorithm reaches the *level* 1 and all the histogram bins are processed (k=*5,* Figure 5.5 (f)). At the end, the optimal clusters are found, each of them corresponding to one specific kind of change, together with outliers. S_k sectors are defined according to the region associated to the kth cluster (Figure 5.3 , step 2). Figure 5.5 (a)-(f) illustrates the working process of the proposed step 2 by considering a 2D zenithal view of the spherical space. Given the nature of the proposed iterative multiple CD step, we name it as Growing OPTICS (GO).

5.2.3 Step 3: Adaptive Magnitude Thresholding

In step 1 a single global threshold value T along ρ is computed that exploits the properties of the meta-class Ω_c. However, each $\omega_{c_k} \in \Omega_c$ may have a different overlapping grade with the distribution of ω_n. Thus, the optimal threshold T_k along the magnitude variable that separates a given kind of change ω_{c_k} from ω_n might be different from the global threshold T defined in step 1. Thresholds T_k $(k = 1, \dots, K)$ can be different from each other. To properly exploit the peculiarities of each ω_{c_k}, the last step tunes the threshold value along ρ by considering only the pixels in each sector S_k, $k = 1, \dots, K$. Therefore, K binary problems are defined that differ from the problem described in step 1 because of the different balance between change and no-change class prior probabilities and conditional distributions. Each optimal threshold value T_k can be computed by applying the same methods as for step 1. The residual gray light area in Figure 5.3, step 3 is associated to unchanged pixels, whereas the sectors $S_1, S_2, S_3, S_4, \dots, S_k$ are associated to changed areas.

5.3. Experimental Results

5.3.1 Dataset Description and Design of Experiments

In order to validate the proposed approach, two datasets made up of both single-sensor and multisensor VHR optical images were selected. The study area is located in the Trentino region in the north of Italy. For dataset 1, images were acquired by the QuickBird (QB) and the WorldView-2 (WV-2) satellites in 2006 and 2010, respectively. Whereas for dataset 2, images were acquired by the QB satellite in 2005 and 2006. The images were pre-processed by absolute radiometric and geometric corrections and integration, as per [5]. Pre-processed images (Figure 5.6) show the same spatial resolution (i.e., 0.6m) and size (i.e., 640×640 pixels).

The selected areas of interest contain mainly agricultural fields, plus a new road for dataset 2 (Figure 5.6). In order to properly highlight these kinds of change, Tasseled-Cap (TC) features were extracted [5]. CVA is applied to the brightness, wetness and greenness features and the information is represented in spherical coordinates [5], [273]. A reference map of each study area was built by photointerpretation (Figure 5.6 (c) and (f)), and 5 and 3 kinds of change were found for datasets 1 and 2, respectively. Table 5.1 (a) and (b) show the number of pixels for each of the 5 and 3 change classes ($C_{1,1}$-$C_{1,5}$ and $C_{2,1}$-$C_{2,3}$) and no change class ($C_{1,6}$ and $C_{1,4}$) for both datasets.

In order to validate the accuracy of the proposed approach, performance of GO is compared to that of OPTICS in the 3-step configuration.

Table 5.1. Number of changed and unchanged pixels for (a) dataset 1 and (b) dataset 2.

(a)

	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{1,5}$	$C_{1,6}$
Dataset 1	6739	48221	23024	37219	13469	280928

(b)

	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$
Dataset 2	27542	23781	647	357630

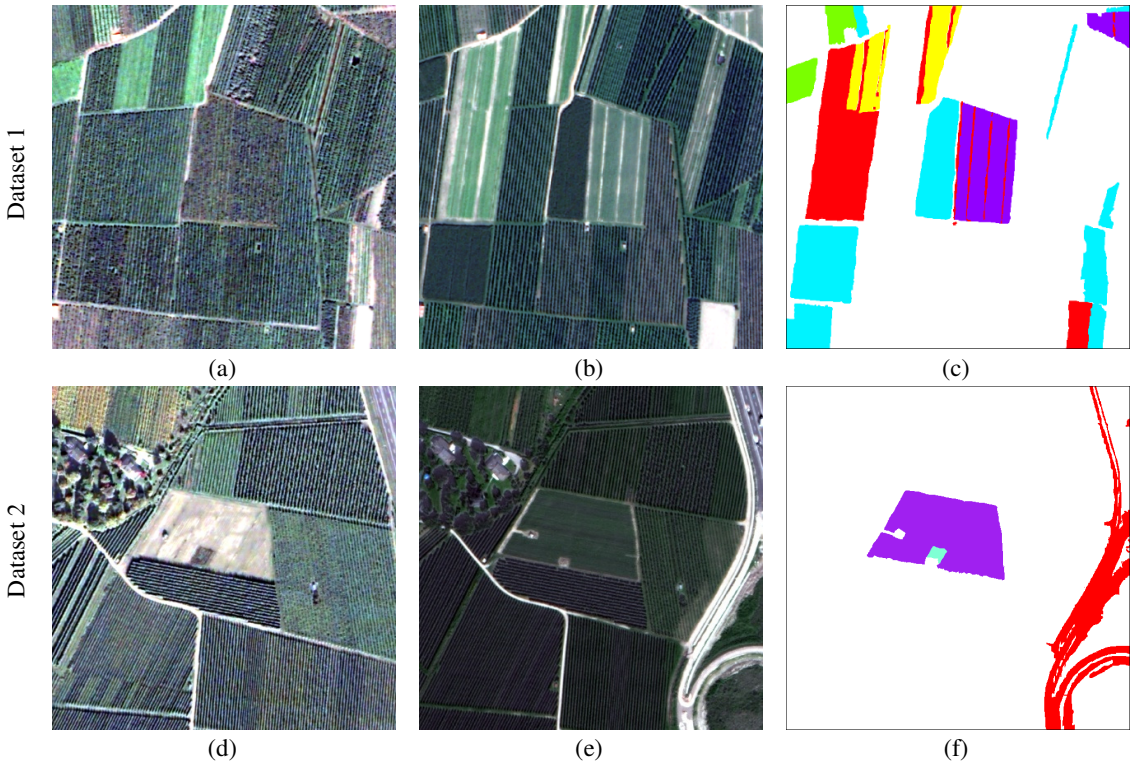


Figure 5.6. True color composition of dataset 1: (a) QB 2006 and (b) WV-2 2010 images; and dataset 2: (d) QB 2005 and (e) QB 2006 images. (c) and (f) Reference maps.

5.3.2 Experimental Results

Once pre-processing and feature selection have been carried out, we proceed with the multiple CD. In step 1 the threshold T is calculated by means of [56]. Results are reported in Table II for both datasets. Step 2 is applied only after masking unchanged samples. For both datasets, the minimum number of clusters k for the initial step of GO was set to 2, based on the density of the clusters and the initial parameters of OPTICS. *MinPts* was set to 20 and the number of iterations of the region growing algorithm was set to 5 with a neighborhood of connectivity equal to 26. The selection of these parameters is critical for the performance of the standard and proposed approaches, but are easy to be fixed. In this set up GO was able to detect all the changes in 11.9s for dataset 1 and 8.8s for datasets 2, using MATLAB® on a standard workstation. Hardware is Intel(R) Core(TM) i7-3630QM CPU @2.40 GHz, 16.00GB RAM. Whereas the standard OPTICS algorithm required 140s for dataset 1 and 125s for dataset 2. Step 3 identifies T_k values (Table 5.2). As expected T_k threshold values are different among them and from T.

Table 5.2. Magnitude threshold values for Datasets 1 and 2.

Dataset	T	T_1	T_2	T_3	T_4	T_5
1	0.030	0.027	0.035	0.047	0.032	0.026
2	0.090	0.088	0.097	0.079	-	-

The CD maps obtained are shown in Figure 5.7. It is possible to see that the proposed approach was able to properly detect all the kinds of change. A visual comparison of the two CD maps with the reference map (Figure 5.7 (c) and (f)), points out the improvement achieved while working with the GO algorithm. In dataset 1, GO algorithm detects $C_{1,1}$ and $C_{1,4}$ properly, whereas OPTICS tends to confuse those classes with other changes, resulting in several false alarms. In both experiments, one part of $C_{1,1}$ is misdetected, maybe because of TC features. GO is further able to detect and separate class $C_{1,5}$ from $C_{1,1}$ (which correspond to overlapped clusters in the spherical domain), whereas OPTICS confuses it with class $C_{1,1}$. Concerning $C_{1,2}$ and $C_{1,3}$ the proposed approach results in a slightly better detection (Table 5.3). In dataset 2, improvements are mainly achieved in the better detection of $C_{2,2}$, as well as the reduction of false alarms with a better detection of unchanged samples (Table 5.4). In this specific dataset, no overlapping clusters were present. The proposed approach outperforms the standard one of about 7% and 2% for datasets 1 and 2, respectively, without using any prior knowledge about the number of changes.

Table 5.3. Confusion matrix for standard OPTICS and proposed GO in dataset 1.

	Changes Found	Actual Changes					
		$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{1,5}$	$C_{1,6}$
Standard OPTICS	$C_{1,1}$	**3243**	8	3	10249	11976	8143
	$C_{1,2}$	24	**41087**	291	0	12	4023
	$C_{1,3}$	102	1011	**21300**	4782	2	5445
	$C_{1,4}$	39	3	196	**18069**	41	1569
	$C_{1,5}$	0	0	0	0	**0**	0
	$C_{1,6}$	3331	6112	1234	4119	1438	**261748**
Overall Accuracy		84.34%					
Proposed GO	$C_{1,1}$	**3893**	23	3	995	60	8754
	$C_{1,2}$	24	**44671**	83	0	10	12
	$C_{1,3}$	5	358	**16992**	480	0	11
	$C_{1,4}$	41	6	658	**29885**	1034	1480
	$C_{1,5}$	0	0	0	781	**11242**	660
	$C_{1,6}$	2776	3163	5288	5078	1123	**270011**
Overall Accuracy		91.97%					

Table 5.4. Confusion matrix for standard OPTICS and proposed GO in dataset 2.

	Changes Found	Actual Changes			
		$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$
Standard OPTICS	$C_{2,1}$	**24870**	0	4	1139
	$C_{2,2}$	0	**14769**	0	1641
	$C_{2,3}$	49	0	**429**	896
	$C_{2,4}$	2623	8981	214	**342860**
Overall Accuracy		96.10%			
Proposed GO	$C_{2,1}$	**24827**	0	6	981
	$C_{2,2}$	0	**22413**	0	3218
	$C_{2,3}$	80	0	**480**	165
	$C_{2,4}$	2635	1368	161	**353266**
Overall Accuracy		97.90%			

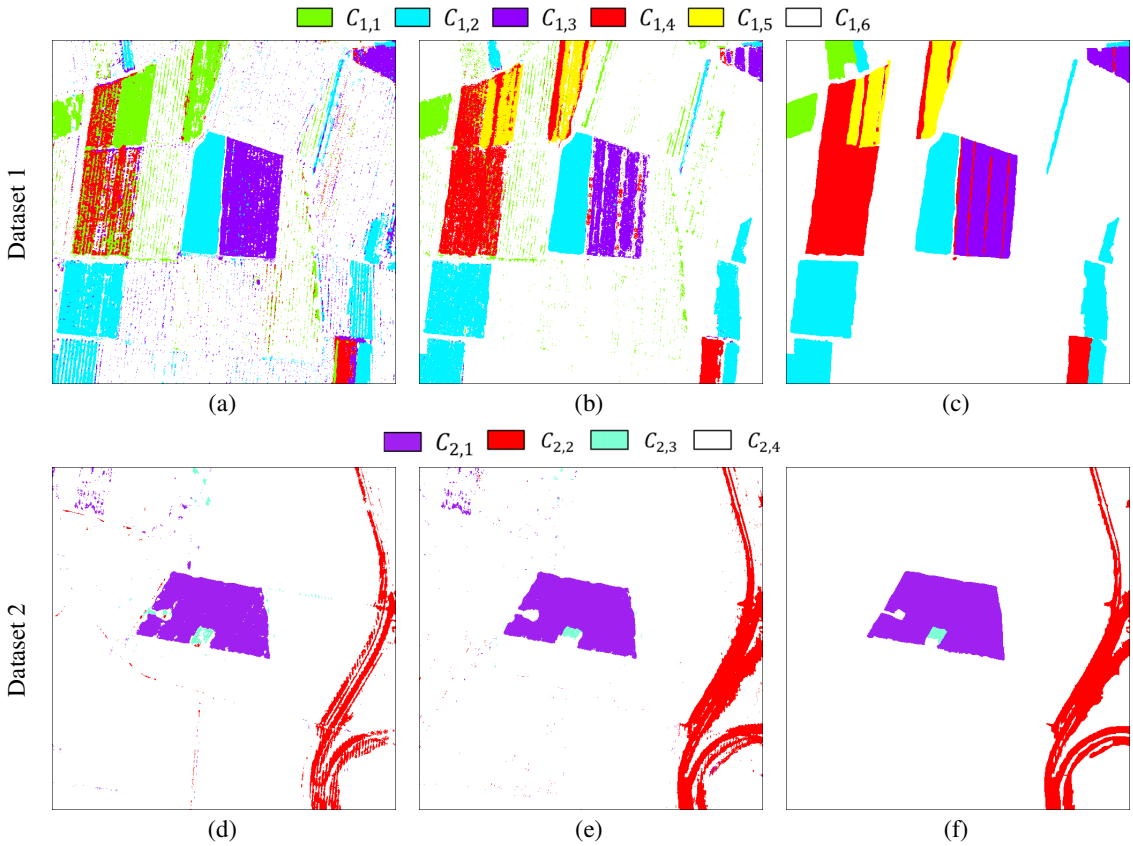


Figure 5.7. CD maps obtained by applying the step 2 with: (a) and (d) standard OPTICS approach; (b) and (e) proposed GO approach; and (c) and (f) Reference Maps.

5.4. Conclusions

In this chapter, an approach to multiple CD in VHR optical images based on iterative clustering and adaptive thresholding has been proposed. The adaptive GO works in (hyper-) spherical coordinates, is distribution free (in the direction domain), unsupervised and automatically identifies the number of changes. The effectiveness of the adaptive GO, with respect to standard thresholding methods, has been validated in two pairs of VHR single-sensor and multi-sensor images. Such datasets showed complex and noisy set-ups. GO demonstrated to be robust in handling noisy samples and overlapping clusters and required a reduced computational time when compared to OPTICS. The adaptive thresholding improved the overall accuracy, especially in the case of dataset 1, because of overlapping clusters. Open issues regarding the selection of initial clustering algorithm parameters remain (for both OPTICS and region growing). Though the parameters are easy to setup (correlated to data spatial resolution and density of cloud points), they are no fixed in an automatic way. Future works will consider the applicability and adaptation of the method for CD in images with lower spatial resolution as well as possible extensions to other applications (e.g., classification).

Chapter 6

6. Spatio-temporal evolution of crop fields in Sentinel-2 Satellite Image Time Series[7]

This chapter introduces an approach for the spatio-temporal evolution of crop fields in Sentinel-2 satellite image time series. A state-of-the-art review is made to show the relevance of developing new algorithms that account for the characteristics of sensors such as S2, as well as the importance for agricultural applications. The approach is divided in different stages that work in an automatic way and without need of in-situ data. Spatial, spectral and temporal information are jointly exploited to derive products of interest in the agricultural analysis.

6.1. Introduction

As the Earth population and the demand for food continue to increase, the need for precise and timely information about the properties and dynamics of global agricultural systems is becoming increasingly important. Since decades, Earth Observation (EO) satellites provide a unique way to observe our living planet from space. Thanks to the revisit property of the EO satellites, a huge amount of multitemporal images are now available in archives and continue to be acquired. The spectral, temporal, and spatial information contained in such satellite image time series offers unique opportunities for monitoring and managing land cover dynamics from local to global spatial scales [310]. Satellite image time series provides temporal and spatial consistent spectral measurements of Earth surface that correlate directly with annual/inter-annual cycles of vegetation growth (i.e., phenology). Therefore SITS offers a significant insight into the response of vegetation at short and long terms, which is suitable for precision agriculture [38], [310], [311]. Nevertheless, high spatial/temporal resolution cropland cover maps are not available yet that provide information regarding: i) geographic distribution, ii) areal extent, iii) cropping intensity; and iv) crops state.

Mapping such information from remote sensing is challenging, especially for areas with high density of small crop fields, where in turn only aggregated analysis is usually provided. In order to create such cropland cover maps, some requirements must be fulfilled. Among them, the most critical ones are to have: i) continuous and regular acquisitions that allow for the derivation of phenological parameters and the mapping of agricultural areas, ii) Satellite Image Time Series (SITS) with a temporal resolution high enough to follow the fast evolution of crops (which can vary from few weeks to few months); and iii) an updated map containing the field boundaries and the fields that accurately separates single/small fields and thus allows the monitoring of single agricultural units. Regarding the first requirement, state-of-the-art methods are able to generate continuous SITS, but require to have regular acquisitions in time [29], [38], [198], [312], [313]. In the second requirement, unlike other kinds of vegetation, crops often reach canopy maturation quickly after plantation and have a relatively short growing period [60], [314], [315]. Up to date, data from a single satellite sensor were not capable of simultaneously capturing the fast-changing agricultural activities on the ground in terms of both spatial and temporal resolutions [316]. The third requirement is difficult

[7] Part of this chapter appears in:
Journal paper: Y. T. Solano-Correa, F. Bovolo, L. Bruzzone, and D. Fernández-Prieto, "Spatio-temporal evolution of crop fields in Sentinel-2 Satellite Image Time Series," ready to be submitted to *Remote Sens. of Env.*.
Conference paper 1: Y. T. Solano-Correa, F. Bovolo, L. Bruzzone, and D. Fernández-Prieto, "Spatio-temporal evolution of crop fields in Sentinel-2 Satellite Image Time Series," in 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 2017, pp. 1–4.
Conference paper 2: Y. T. Solano-Correa, F. Bovolo, L. Bruzzone, and D. Fernández-Prieto, "Derivation of cropland phenological parameters by NDVI smoothing of Sentinel-2 Satellite Image Time Series," accepted in 2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS-2018).

to handle and is indeed a recurrent problem in Remote Sensing applications (RS) [60], [317], [318]. Agricultural areas usually contain small fields [317], [319], when compared to the sensor spatial resolution (e.g., MODIS, Landsat) and thus they are difficult to characterize [60]. In essence, producing detailed cropland cover maps requires SITS with high temporal and spatial resolutions acquired with regular time spacing.

The spatial and temporal resolutions of up to date satellite sensors are compromised. High spatial resolution sensors like Landsat do not have enough temporal resolution, whereas coarse spatial resolution sensors like MODIS have high temporal resolution [316]. Since 2015, both high temporal and high spatial resolution SITS became available within the European Space Agency (ESA) Copernicus program, more specifically due to the Sentinel-2 (S2) constellation. S2 constellation is made up of two twin satellites named: S2A (launched in 2015) and S2B (launched in 2017). Combining both satellites gives the capacity to cover Earth every five days (at the Equator) with 10m to 60m resolution and 13 spectral bands. As per Landsat constellation, the revisit period of a single S2 satellite increases as we move from the Equator to the North hemisphere as a result of overlapping orbits [317], [320]–[322]. The overlapping orbits together with the high temporal resolution and the width swath of S2 (290Km), result in a clear increase of the number of acquisitions over a specific geographical area (which satisfies the first requirement). At the same time, this results in a higher probability of acquiring free cloud images, but tends to generate irregularly sample data in time. By considering two overlapping orbits, we have that: i) for Landsat, we move from a regular acquisition of 16 days to an irregular acquisition in a sequence of 7-8-7-8 days, ii) for a single S2 satellite, we move from a regular acquisition of 10 days, to an irregular acquisition in a sequence of 3-7-3-7 days; and iii) for the two S2, we move from a regular acquisition of 5 days, to an irregular acquisition in a sequence of 2-3-2-3 days. Thus, S2-SITS makes it possible to generate products with a better compromise in terms of spatial and temporal resolutions, allowing for the analysis at single crop field level (highly relevant in precision agriculture). But S2-SITS also introduce the challenge of dealing with irregularly acquired data.

Identifying each single crop field in a given agricultural area, allows for the temporal analysis at individual field level, whereas the widespread practice is to perform an analysis over several aggregated fields. This is because the availability of crop field boundary maps is uncommon. In the literature, the crop fields boundaries are mostly obtained by segmentation applied to the original spectral bands or to some vegetation indices extracted from one or few multitemporal remote sensing images [61]. Thus, they account only for crops in a single or few stages of the vegetative cycle. However, given the crop phenology and the presence of several types of crop, it is expected that each crop shows different maturation instants in the vegetative cycle. This means that at a certain time, crops will be at various phenological stages (i.e., some will be at the last stage of the vegetation cycle, whereas others will not). Thus, single or few dates analysis tends to underestimate crop fields since the ones in non-vegetative stage are likely to be not identified. Thus, creating a crop field map that accounts for every single field at distinct stages of the vegetative cycle is relevant. Dense SITS can be used to reduce the impacts of ambiguities due to the phenological stage and the spatial arrangement of field boundaries (irrigation ditches, tracks and roads, fences and hedges, weed and grass swards, trees and shrubs) that in single date satellite images may not be spectrally separable from field interiors [60], [323]–[325]. Further, and importantly, dense time series reduce the influence of missing, shadowed and atmospherically contaminated observations [317], [318], [326] and enables specific crop and non-crop phenologies to be considered as part of the algorithm implementation. Therefore, in this chapter we exploit the temporal evolution of the spectral and spatial information in dense S2-SITS to build a multitemporal crop field map.

Once a multitemporal map of single crop fields and their boundaries is available, temporal analysis can be carried out at field level. Nonetheless, to effectively track continuous intra-annual dynamics, such as vegetation phenology, noise-free remotely sensed data acquired at appropriate time instants is required. S2-SITS correspond to discrete non-uniformly and irregularly sampled measurements of spectral variations, which, moreover, can be corrupted by atmospheric, geometric, and radiometric disturbances [327]. Therefore, to increase accuracy in phenological parameters estimation, we resort to denoising procedures and fill the gaps between satellite image acquisition dates with a pertinent continuous interpolating function [327].

The literature contains reference to a broad variety of strategies designed to reduce the impacts of missing (noisy) data, based on smoothing and function-fitting algorithms [29], [37], [38], [313], [315], [328], but: i) none of them considers irregular sample data, and ii) data are assumed to show a harmonic behavior, given by one (or two) cropping cycles. These algorithms increase the spatial and temporal consistency of SITS used to derive phenological parameters [204], but are not suitable when referred to S2-SITS. This is due to the fact that up-to date, analysis were performed at an aggregated level (e.g., forest, grass), where a single harmonic behavior was assumed over the year. When a single crop field is analyzed, more than one harmonic is expected, especially if an intensively cultivated area is observed. Furthermore, these algorithms are also used to implicitly derive phenological parameters such as the number of cropping cycles, length of season, middle of season and maximum value. Since this is achieved by relying on the assumption of a single harmonic behavior of the data, it results in poor estimation when applied to S2-SITS. Other methods derive the phenological parameters independently from the SITS smoothing, but require user defined thresholds [29], [313]. Therefore, there is a clear need to develop and/or adapt state-of-the-art algorithms/methods to the analysis of dense and irregularly sampled S2-SITS in the context of precision agriculture.

In this chapter, we propose an approach for spatio-temporal evolution of crop fields in S2-SITS. The approach automatically: i) identifies and separates all the crop fields cultivated at least once over a given area, ii) reconstructs continuous and regular sampled S2-SITS by means of an adaptive non-parametric regression model; and iii) extracts spatio-temporal information from the previous derived data. The proposed approach is automatic, can easily account for the new upcoming images and is able to deal with irregularly sampled SITS.

The reminder of this chapter is structured as follows. Section 6.2 presents the proposed approach for the spatio-temporal evolution of crop fields in S2-SITS in detail. Section 6.3 describes the dataset used to validate the proposed approach, presents the results for its different stages and compares the results with state-of-the-art methods. Finally, Section 6.4 draws the conclusions and illustrates future works.

6.2. Proposed Approach for Spatio-Temporal Evolution of Crop Fields in S2-SITS

In this section, details of the proposed approach for studying the spatio-temporal evolution of crop fields through the use of S2-SITS are given. To study the spatio-temporal evolution, we jointly exploit the spectral, spatial and temporal information in the S2-SITS, and perform a detailed analysis at the level of single crop field. Figure 6.1 depicts the block scheme of the proposed approach that is based on four main steps: i) time series pre-processing, ii) spatio-temporal fusion, iii) daily time series reconstruction and; iv) spatio-temporal information extraction. First, the S2 time series is pre-processed in order to guarantee similar spectral and spatial characteristics among images and to filter out cloudy images. Second, a multitemporal vegetation map is built by fusing spectral, spatial and temporal evolution of NDVI time series. Third, the multitemporal vegetation map is used as a mask in order to build regular and continuous NDVI time series at the level of single objects/crop fields. Finally, the multitemporal vegetation map and the daily NDVI-TS are used to extract spatio-temporal information through the analysis of different radiometric indices. Of particular interest is the derivation of products such as phenological parameters and cumulative indices which are later used to map the area and perform statistical analysis.

Let us consider a dense sampled time series, $SITS = \{X_1, X_2, X_3, \dots, X_N\}$, acquired by S2 over the same agricultural area in a period $[t_1, t_N]$. Let us assume that the area is observed by two S2 orbits, such that images are acquired with non-uniform time sampling (irregularly). Let Ω_f ($f \in [1, F]$) be the set of F existing fields in X_n. Ω_f corresponds to crop fields that have been cultivated and to crop fields left to rest (fallow). Let us finally assume that different crop fields in X_n ($n \in [1, N]$), are cultivated with different types of crops, thus they show different phenological cycles, and that intensive crop rotation is practiced in the area (i.e., some fields are cultivated with different crops over the seasons).

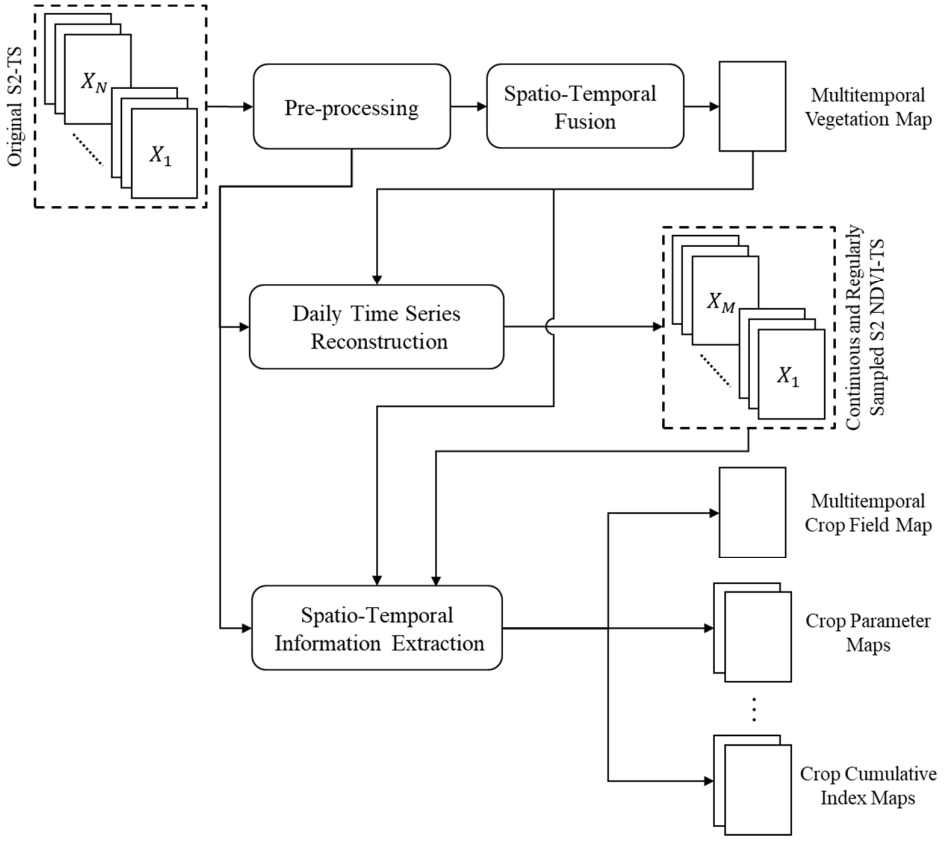


Figure 6.1. Block scheme of the proposed approach for spatio-temporal evolution of crop fields in S2-SITS.

6.2.1 Pre-processing

The pre-processing phase seeks to prepare the S2-SITS images for the next steps. First, the images are atmospherically corrected (converted from L1C to L2A level) by using the sen2cor tool provided by ESA [329]. Second, the 6 bands at 20m are up-scaled to 10m by means of an optimized High Pass Filter (HPF) [330]. Band 4 is used as panchromatic band for bands 5, 11 and 12, whereas band 8 is used for bands 6, 7 and 8A. The three 60m bands are not up-scaled since they do not provide relevant information regarding vegetation. Thus, here they are used for atmospheric correction and cloud screening by sen2cor, only. Third, each image in the S2-TS is cut to the size of the study area. This is performed only when the study area is smaller/bigger than a S2 tile size. When the study area is smaller than a tile, this stage may help to increase the number of free cloud images in the TS, given that clouds do not necessarily cover the whole tile [331]. Fourth, images with high cloud percentage are filtered out from the SITS by means of the SLC map obtained after applying the sen2cor correction. At this stage of the approach, all the images containing any pixel marked as: cloud shadows, cloud low probability, cloud medium probability, cloud high probability and thin cirrus are removed. Finally, images acquired by overlapping orbit swaths are co-registered. As shown by [321], overlapping S2 swaths have been found to be mis-registered by typically 10m up to 100m (depending on the study area). Such issue is highly critical in time series analysis and even more in the spatio-temporal evolution of crop fields where spatial variability is high. To perform the co-registration, here we use as reference image the one acquired at first in the whole time series. Whereas target image/s will be the one/s acquired after. Any state-of-the-art method can be considered to perform the co-registration. Here we consider the area-based least squares matching approach that provides sub-pixel precision suitable for the S2 spatial resolution and the task of crop fields analysis [332].

6.2.2 Spatio-Temporal Fusion

The spatio-temporal fusion takes advantage of both the availability of S2-SITS and the dynamic vegetative cycle of the crops, to detect and separate the F crop fields in X_n by considering their multitemporal evolution. Figure 6.2 depicts the block scheme of the proposed approach, where the NDVI is first extracted from each of the images in S2-SITS. Then, spectral and spatial analyses are carried on over the S2 NDVI-SITS to get a multitemporal object mask. Depending on the study area, these objects may correspond only to crop fields (Ω_f), or to other kinds of vegetation (ω_{ov}) such as forest or wetlands (which may have single date NDVI response similar to that of a crop, but shows different temporal evolution in SITS). In the last step, a connected component labeling is applied to build a multitemporal vegetation map, such that each possible crop field can be later analyzed separately. The multitemporal vegetation map is represented by $\{\omega_v, \omega_{nv}, \omega_e\}$. Here ω_v represents the areas that have had vegetation at least once over the studied period, with $\omega_v = \{\Omega_f, \omega_{ov}\}$. $\Omega_f = \{\omega_1, \omega_2, \dots, \omega_F\}$ are the different crop fields (cultivated and fallow) labeled separately; and ω_{ov} represents other vegetation areas such as forest or grass. Finally, ω_{nv} represents all the pixels that never showed vegetation, such as buildings and roads; and ω_e represents the crop field borders. In order to separate among ω_v in the multitemporal vegetation map, the map is filter out in a latter step. In literature, this is usually done by considering a land cover map containing the class "agricultural areas" (if available). The problem of this strategy is that maps are usually generated by considering one (or few) acquisition dates. Thus, crop multitemporal variability is not fully captured. Because of this, it is difficult to rely on them. Finally, considering the worldwide average minimum size of a crop field and the S2 spatial resolution, only crop fields with areas bigger than 1ha are considered in the process [317], [319].

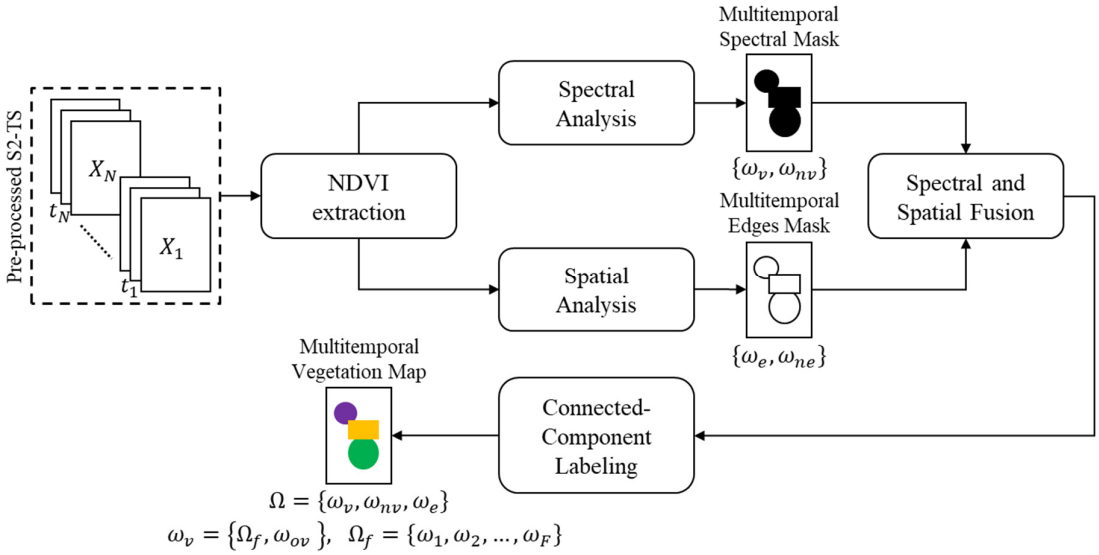


Figure 6.2. Block scheme of the proposed approach to spatio-temporal fusion.

6.2.2.1 NDVI Extraction

In the literature, the separation or classification of different crops or land covers is carried out by using features such as the original spectral bands, radiometric indices, etc. The selection of each feature is determined by its ability in highlighting certain land covers over the others (i.e. ensure separability of the different classes in the feature space). In remote sensing, many radiometric indices have been developed in the literature that allow to highlight different properties of vegetation. These radiometric indices exploit the spectral information acquired by the different sensor bands and analyze their physical response in order to measure biomass or vigor of vegetation. Table 6.1 shows the list of the most common radiometric indices used in remote sensing. Here the spectral bands are shown as per S2 numbering [333].

In the specific case of vegetation and agricultural fields, the NDVI is widely used [61], [197], [334]. NDVI takes advantage of the contrast between the Red and Near-InfraRed (NIR) spectra to assess whether an area contains green vegetation or not. NDVI equation is given in (6.1), where the NIR and Red bands of X_n are used to get X_{NDVI_n}. Equation (6.1) is applied over the entire S2-SITS and a NDVI-SITS is obtained. Each NDVI image ranges from -1 to 1.

$$X_{NDVI_n} = \frac{X_{NIR_n} - X_{Red_n}}{X_{NIR_n} + X_{Red_n}} \tag{6.1}$$

Generally, healthy green vegetation absorbs most of the visible light, and reflects a large portion of the NIR light, thus NDVI tends to assume large positive values. Unhealthy or sparse vegetation and water behave in the opposite way, thus NDVI tends to assume extreme negative values. Bare soil reflects moderately in both the Red and NIR, thus NDVI is close to zero.

Table 6.1. List of the most common radiometric indices used in remote sensing [333].

Index	Equation
Normalized Difference Vegetation Index	$NDVI = \frac{B8a - B4}{B8a + B4}$
Normalized Difference Water Index	$NDWI = \frac{B12 - B4}{B12 + B4}$
Normalized Difference IR Index	$NDII = \frac{B8 - B11}{B8 + B11}$
Soil Adjusted Total Vegetation Index	$SATVI = 1.5 \cdot \left(\frac{B11 - B4}{B11 + B4 + 0.5}\right) - \left(\frac{B12}{2}\right)$
Plant Senescence Reflectance Index	$PSRI = \frac{B4 - B3}{B6}$
Cellulose Absorption Index	$CAI = \frac{B12}{B11}$

6.2.2.2 Spectral Analysis

According to the NDVI characteristics, a threshold $0 < T < 1$ is applied to X_{NDVI_n} ($n \in [1, N]$) to identify areas with vegetation. This is done for each available image in the NDVI-SITS. Once vegetation is identified for each acquisition, multitemporal analysis is conducted to select the areas that have had vegetation at least once over the studied period. This is obtained, by defining a multitemporal binary spectral mask $X_{Spectral_{mask}}$ based on NDVI response as in (6.2). Since all the images available in the NDVI-SITS are considered, the high multitemporal variability of the different crop fields is mapped in $X_{Spectral_{mask}}$. When a new X_{N+1} image is acquired, the current $X_{Spectral_{mask}}$ can be easily updated by including $X_{NDVI_{N+1}}$ in (6.2).

$$X_{Spectral_{mask}} = \begin{cases} \omega_v, & if \left(\sum_{n=1}^{n=N} (X_{NDVI_n} > T)\right) \geq 1 \\ \omega_{nv}, & otherwise \end{cases} \tag{6.2}$$

$X_{Spectral_{mask}}$ will contain all the areas with vegetation (ω_v) and the areas with non-vegetation (ω_{nv}) over the time, ω_v inlcudes crop field (Ω_f) and other vegetation (ω_{ov}) as well. Finally, $X_{Spectral_{mask}}$ does not provide spatial distinction among vegetation. In other words, if neighboring fields show the same NDVI behavior over time, they cannot be distinguished from each other. Neither does $X_{Spectral_{mask}}$ provide spectral distinction with other vegetation classes such as forest or grass. Thus, to help in the separation among different types of vegetation, spatial analysis is required. Whereas to help in the separation among crop fields and other vegetation, other strategies are implemented.

6.2.2.3 Spatial Analysis

To achieve the spatial separation among the areas with vegetation (ω_v) present in $X_{Spectral_{mask}}$, we take advantage of the spatial information in the NDVI-SITS. Because of the phenological behavior of the crops, there exist a time t_{MoS} representing the mature stage in the vegetative cycle. MoS stands for Middle of Season and $t_{MoS} \in [t_1, t_n]$. Since t_{MoS} varies crop by crop and thus field by field, a gradient exists such that we can employ it to separate the vegetated areas identified in the previous step according to the phenological behaviors. Edges of mature crops can be detected for each X_n in NDVI-SITS (X_{edges_n}) and used as a reference for separating the crop fields. The accuracy of this step depends on whether two neighboring crop fields have different t_{MoS} or not. An edge is defined as the locality of connected components where the image intensity varies rapidly. The gradient is commonly used in the literature to extract edges (e.g., Laplacian and Sobel operators) [335]. The edge detection can be applied to NDVI or other spectral feature able to highlight the vegetative stages of a crop. Single date edge maps are jointly used to detect areas with edges (ω_e) and separate them from no edges (ω_{ne}). To ensure the selection of edges that possibly correspond to a specific crop field, only areas where strong edges exist are considered as being reliable. A strong edge is defined if an edge has been detected at least twice over the NDVI-SITS (6.3). Therefore, $X_{Edges_{mask}}$ contains the boundaries between the areas with vegetation (ω_v). A low probability exists that some areas corresponding to other vegetation (ω_{ov}) exhibit strong gradients in different times of the year resulting in strong edges. $X_{Edges_{mask}}$ can be easily updated as for $X_{Spectral_{mask}}$.

$$X_{Edges_{mask}} = \begin{cases} \omega_e, & if \left(\sum_{n=1}^{n=N} X_{edges_n} \right) > 1 \\ \omega_{ne}, & otherwise \end{cases} \tag{6.3}$$

6.2.2.4 Spectral and Spatial Fusion

By combining the $X_{Spectral_{mask}}$ and $X_{Edges_{mask}}$ as in (6.4), we ensure that all the areas in which there was vegetation at least once in the NDVI-SITS, are identified and separated. Thus, X_{mask} is a binary mask where "0" values represent the areas with no vegetation and the crop field borders, whereas "1" corresponds to the areas with vegetation. The borders identified by $X_{Edges_{mask}}$ are separated from to the class of vegetation because they contain mixed pixels between vegetation and bare soil.

$$X_{mask} = X_{Spectral_{mask}} \cdot X_{Edges_{mask}} \tag{6.4}$$

6.2.2.5 Connected-component Labeling

X_{mask} accounts for all the regions in which there was vegetation at least once. Now labeling is required such that each area is unambiguous. To this aim, the well-known connected-component labeling method is used [336]. This method is used in computer vision to detect connected regions in binary images, although color images and data with higher dimensionality can be processed. To decide if two pixels or regions are connected, 4-connected or 8-connected medium image graphs are used. As a result, we get a multitemporal vegetation map $X_{veg_{map}}$ that is built by considering the spectral, spatial and temporal evolution of a NDVI-TS. Pixels in this map are labelled as $\{\Omega_f, \overline{\Omega_f}\}$, with $\overline{\Omega_f} = \{\omega_{ov}, \omega_{nv}, \omega_e\}$. From $X_{veg_{map}}$, we can now perform TS analysis at the single object level.

6.2.3 Daily Time Series Reconstruction

To analyze the spatio-temporal evolution of agricultural areas, modelling the presence of the vegetation cycles among different crops is usually done in literature. To this end, continuous and regularly spaced SITS are needed. To obtain SITS with such characteristics, we propose a novel approach for daily TS

reconstruction that makes use of a non-parametric regression model derived from NDVI-SITS. The reconstruction is performed at single object level, by means of $X_{veg_{map}}$, is automatic and is able to deal with irregularly sampled images. One of the many problems with state-of-the-art methods is the proper reconstruction of the starting and end day of the given SITS. This is traditionally solved by: i) including extra images acquired before and after the considered period; or ii) by replacing the before and after acquisitions with the first and last images available in the SITS [29], [337]. This becomes an extra constrain for S2, since up to date only 3 years of acquisitions are available. Thus, the proposed approach is developed without such constrain. Three main steps are followed: i) definition and extraction of NDVI-SITS sets, ii) building regularly sampled TS; and iii) adaptive non-parametric regression of NDVI-SITS.

6.2.3.1 Definition and Extraction of NDVI-SITS Sets

In order to perform the TS reconstruction, a set of NDVI training samples is required. Given that no crop field maps are usually available in literature the selection of these training samples is performed for every single pixel available in an image (including no-crop ones). Once the selection process has been carried out for all the images in the SITS, areas or pixels of interest (where the user is sure to find a crop field) are selected to be further analyzed. In the proposed method the sets are extracted from the original NDVI-SITS acquisitions and for each of the ω_v in $X_{veg_{map}}$. In order to mitigate for noisy components the mean NDVI value of each ω_v in $X_{veg_{map}}$ is used as training samples. Even though we partly reduce the effects of noisy data, issues remain due to the irregularly acquired data and the missing acquisitions due to presence of clouds.

6.2.3.2 Building Regularly Sampled NDVI-SITS

The acquisition of data with two overlapping orbits, results in irregularly sampled data in time, independently of the presence of clouds. This adds an extra level of complexity to the derivation of phenological parameters, where equally spaced data (in time) and regular shapes (bell like) are usually assumed. Thus, the irregularity of data needs to be handled before applying any further step. In other words, a preliminary symmetrical and uniformly sampled NDVI-SITS needs to be guaranteed. To this aim, the simplest approach in literature is to replace each data value in the TS by a linear combination of nearby values in a window [29]. We propose an alternative approach, where a smoothed time series is simulated generating daily acquisitions by applying a cubic interpolation. Then, we further smooth the smallest noisy samples by means of a low pass filter with a given window width w. Both, the interpolation degree and the window width, determine the degree of smoothing, but also affect the ability of the method to follow rapid NDVI evolution [29].

6.2.3.3 Adaptive Non-parametric Regression of NDVI-SITS

Because of the annual phenologies of vegetation, state-of-the-art methods for smoothing NDVI-SITS usually make two assumptions: i) a crop field is cultivated once, maximum twice, per year (showing one, maximum two peaks) and; ii) the temporal behavior of crops can be considered as a single harmonic. As a consequence of the above mentioned assumptions, harmonic functions are used with an a priori user defined number of cropping cycles. Thus, simple harmonic oscillators and Fourier series are among the most commonly used models to fit satellite image time series [184], [311]. However, simple harmonic oscillators fail at reproducing more than one harmonic and Fourier series approaches assume regularly spaced TS and underlying components that are truly harmonic [184], [198], [311]. Thus, methods are not fully automatic and do not account for cultivations with number of cycles greater than two, neither can they account for the variability introduced by crop rotation practices. To overcome these limitations, the derivation of a function by means of a non-parametric regression is proposed. Opposite to literature, where the process is done at pixel level, this process is carried out at single object level, and for each of the ω_v object in $X_{veg_{map}}$. This characteristic gives an advantage to the method, which is able to model the behavior of each object (crop

field) in a separate and homogeneous way. Here, we use a MLP-NN with one hidden layer to perform the non-parametric regression. A novel adaptive configuration is introduced for better performance. The number of neurons in the hidden layer (n_{HL_i}) is determined in accordance with a Mean Square Error (MSE) threshold (T_{MSE}). A given $n_{HL_i} = n_{HL_1}$ (lower bound) is first established by the user to train the MLP-NN. After the first iteration ($i = 1$), if the MSE is lower than T_{MSE}, then n_{HL_1} is kept as well as the generated model (6.5). Otherwise, the MLP-NN is re-trained in a second iteration with an increased number of neurons. n_{HL_1} is increased by 1. The process is repeated until the condition on MSE is satisfied ($MSE \leq T_{MSE}$) or until the number of iterations reaches a pre-defined upper bound. When the upper bound is reached, the MLP-NN stops increasing n_{HL_i} in the hidden layer and gives a warning of non-convergence.

$$n_{HL_i} = \begin{cases} n_{HL_1} \\ n_{HL_{i-1}} + 1, \quad if\ MSE > T_{MSE}\ and\ i < I \end{cases} \tag{6.5}$$

6.2.4 Spatio-Temporal Information Extraction

In the last step, we focus the attention on the extraction of relevant information that effectively exploits the spectral, spatial and temporal characteristics of S2-SITS. This information is further used to build a multitemporal crop field map ($X_{crop_{map}} = \{\Omega_f, \overline{\Omega_f}\}$) where Ω_f is finally separated from ω_{ov}.To this aim, we use both the $X_{veg_{map}}$ and the continuous and regularly sample NDVI-SITS. Spatio-temporal information is then extracted to: i) estimate and build phenological parameter maps; ii) apply a set of rules to separate Ω_f from ω_{ov}; and iii) study the cumulative behavior of different radiometric indices for each Ω_f in $X_{crop_{map}}$.

6.2.4.1 Estimating Phenological Parameters

Traditional ground truth measures of phenology involve the observation of individual plants or groups of plants of one species to identify specific events such as budburst, leaf-out, or flowering. Similar measurements can be derived by means of satellite imagery and by studying the NDVI temporal evolution. The precision of this derivation increases as the accuracy of the NDVI-SITS regression increases and is traditionally performed in a semi-automatic way. In this work, the phenological parameters listed below (but others can be considered as well) are estimated in an automatic way and from the NDVI-SITS profiles after regression:

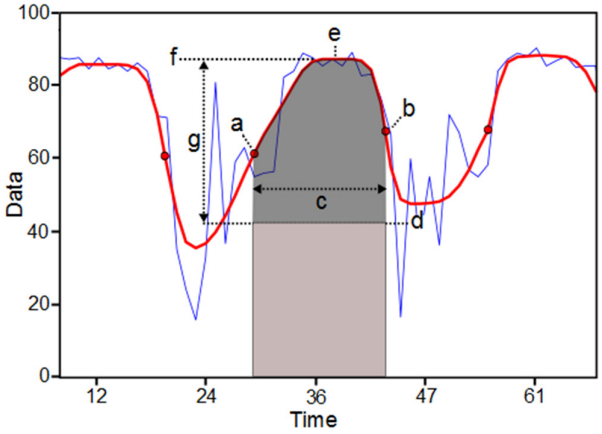


Figure 6.3. Block Phenological parameters: (a) beginning of season, (b) end of season, (c) length of season, (d) base value, (e) time of middle of season, (f) maximum value and (g) amplitude (modified from [29]).

- Beginning of Season (BoS - Figure 6.3 point a): is calculated by means of the second derivative, where the descendent point of inflexion corresponds to the BoS.

- End of Season (EoS - Figure 6.3 point b): is calculated by means of the second derivative, where the ascendant point of inflexion corresponds to the EoS.
- Length of Season (LoS - Figure 6.3 point c): is the time from BoS to EoS and is calculated by subtracting (b) and (a).
- Base Value (BV - Figure 6.3 point d): is calculated by finding the minimum values, from the first derivative, located just before and after the maximum value.
- Middle of Season (MoS - Figure 6.3 point e): is located where the maximum value of the season is found.
- Maximum Value (MV - Figure 6.3 point f): is calculated by finding the maximum values, from the first derivative, that are between the BoS and EoS.
- Amplitude (Figure 6.3 point g): is calculated by subtracting the maximum and base value.
- Number of cropping cycles: corresponds to the number of times that a set of BoS, EoS and MoS metrics are found.

Once the phenological parameters have been derived, we can proceed to build out multitemporal crop field map, where crop fields (Ω_f) are finally separated from other type of vegetation (ω_{ov} – grass and forest). Inside Ω_f we find cultivated crop fields and fallow crop fields. The latter ones areas are not considered to generate further products. ω_{ov} are assigned to $\overline{\Omega_f}$. According to literature, a cultivated area presents a NDVI profile similar to that of a bell, with values that go from 0 to 100 (for a scaled NDVI). Non-cultivated ones (fallowed area) can present a similar bell behavior as a cultivated one, but with a NDVI value lower than a given threshold [315]. If a set of BoS, EoS and MoS falls below the fallow threshold, then the field is not considered for further analysis and the number of cropping cycles is set to zero. A further rule is considered in order to separate grass and forest from crop fields. Grass and forest might show bell shape profile similar to crop ones, but the LoS parameter will be larger than 3 months. Thus, ω_v with a LoS>3 months are assigned to $\overline{\Omega_f}$. In this way, we build our multitemporal crop field map ($X_{crop_{map}}$) where areas corresponding to crop fields (cultivated and fallow) are separated from other classes. Finally, different crop parameter maps can be built: i) winter/summer crop maps; and ii) detailed phenology retrieval maps.

6.2.4.2 Cumulative Radiometric Indices Analysis

Starting from $X_{crop_{map}}$, other information can be extracted. A cumulative index is thus computed from the entire S2-SITS as follows:

$$X_D = \sum_{n=1}^{N} (X_{Radiometric\ Index_n}) \tag{6.6}$$

The information in X_D can be analyzed according to different ranges of the radiometric index. For example, NDWI ranges in [0, 1] and is used for understanding water content in crop fields. The NDWI works in a similar way to the NDVI, but uses the Short Wave Infra-Red (SWIR) band instead of the NIR for the detection of water content. In this case, and according to literature, we define ranges as: i) $0 < X_D < 0.25$, ii) $0.25 \leq X_D < 0.5$, iii) $0.5 \leq X_D < 0.75$; and $0.75 \leq X_D < 1$. The higher the value of X_D, the higher the content of water. If different indices are studied, the meaning of the cumulative index changes. Here, the short list of radiometric indices shown in Table 6.1 is considered. Once the different cumulative indices have been computed, statistical analyses is carried out for crop fields that might exhibit a particular problem. This decision must be taken by an expert.

6.3. Results and Discussion

6.3.1 Dataset Description, Pre-processing and Ground Truth

To validate the proposed approach, a SITS acquired by S2A over the area of Barrax in the South-East part of Spain was selected. Barrax area falls in the tile T30SWJ that is acquired by two overlapping swath S2A orbits: orbit 094 and orbit 051. If a given tile is acquired by a single orbit, a single S2 satellite and under clear sky conditions, we expect an image every 10 days. By considering two orbits, we expect an image every 3-7-3-7 days. The above conditions increase the probability of having cloud free images and guarantee denser SITS for agricultural analysis, but provides irregular samples in time. By using the two orbits, and from the starting acquisition date of S2A until November 30, 2016, 76 images were acquired. Following the steps in section 6.2.1, we first carried out atmospheric corrections by means of sen2cor [329]. After this, bands at 20m were up-scaled to 10m by means of a modified HPF (see sec. 6.2.1) and later on cut to the area of interest. In the Barrax area, a total of 49 images were found to be cloud free. The selection of the cloud free images is carried out by means of the cloud mask generated from sen2cor. The temporal distribution of the free cloud images is shown in Figure 6.4, where red dots represent images acquired on orbit 051 and blue dots the ones on orbit 094. By considering the two orbits, we can see that there is at least one free cloud image per month. For October there were no images provided by ESA. The reference image to perform co-registration was acquired in the orbit 051.

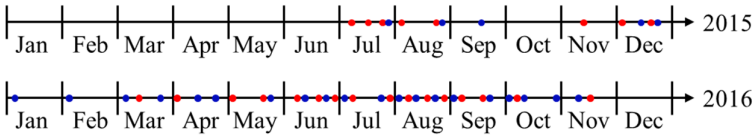


Figure 6.4. Cloud free images timelines for 2015 and 2016 (red: orbit 051 and blue: orbit 094).

The area around Barrax (Figure 6.5) has been used intensively for agricultural research since many years given: i) the different rotation practices, ii) the presence of crops with more than one crop cycle per year, iii) the presence of crops such as the Alfalfa which is harvested several times per year because of its fast growing rate, and iv) the presence of same type of crop under irrigated and non-irrigated conditions [338]. This area locates in an arid terrain with low annual precipitation and requires artificial irrigation. Figure 6.5 shows the location of the Barrax area in Spain, as well as a true color composition of a S2A image acquired on April 14, 2016. Images evaluated for this chapter have a common size of 726px by 954px.

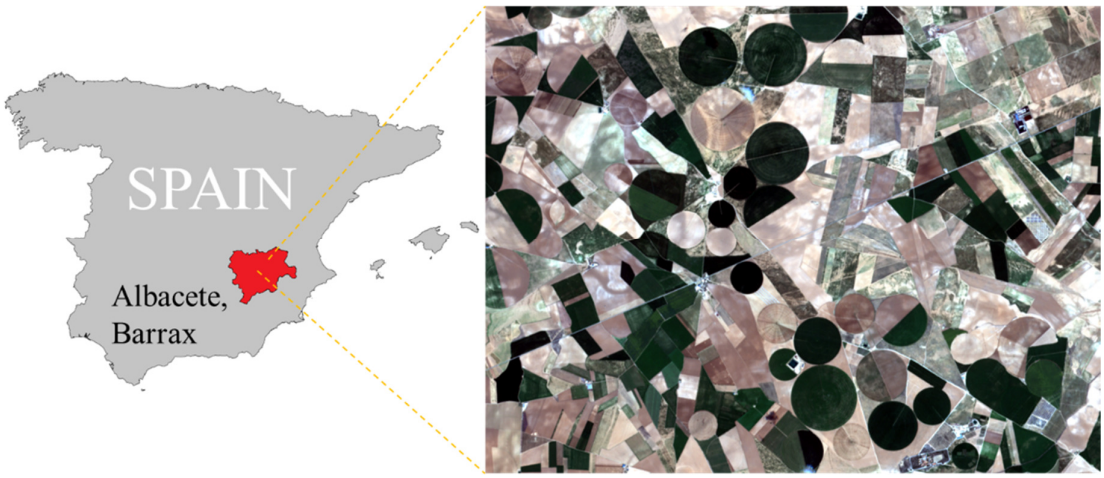


Figure 6.5. Study area location in Barrax, Spain.

In order to validate the proposed approach, two kinds of reference information are used in this chapter: i) a land cover map available for the study area and built by considering a single date acquisition; and ii)

cropping cycles information collected by a survey carried on by experts. In the case of the study area, the most updated land cover map available (up to the date in which we have started the study) is from 2015-2016, and covers a smaller area only (see Figure 6.6). This cadastral map is built yearly and by considering the information reported by farmers at a given date of the year. To build the reference information about the cropping cycles, a survey was implemented by using the original NDVI-SITS and asking to different experts to identify the number of cropping cycles. This was done for all the crop fields in $X_{crop_{map}}$.

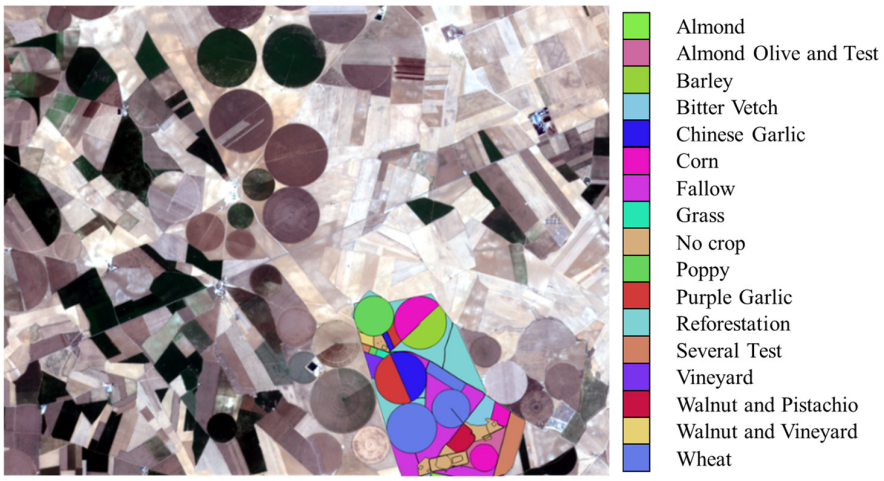


Figure 6.6. Land cover map of 2015-2016 falling inside the study area in Barrax, Spain.

6.3.2 Multitemporal Vegetation Map

Once the S2A-SITS has been defined and pre-processed, we create a NDVI-SITS based on equation (6.1). Then, we built the $X_{Spectral_{mask}}$ by setting $T = 0.1$ in equation (6.2). $X_{Edges_{mask}}$ was computed by applying a Sobel filter to each of the images in the NDVI-SITS [335] and as indicated by equation (6.3). The equation (6.4) was used to combine $X_{Spectral_{mask}}$ and $X_{Edges_{mask}}$ and get X_{mask}. Finally, a connected-component labeling [336] was applied with a connectivity 8, allowing us to obtain the multitemporal vegetation map. Figure 6.7 shows $X_{veg_{map}}$, with a total of 307 objects. Black color represents the areas with no-vegetation and edges, and the other colors represent the vegetated areas. From a qualitative analysis to Figure 6.5, the accuracy of the result is good.

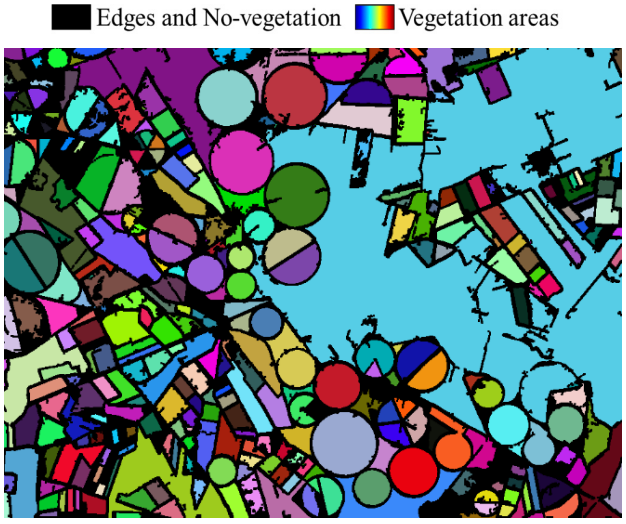


Figure 6.7. Multitemporal vegetation map obtained after applying the spatio-temporal fusion step. Vegetated areas are colored randomly to illustrate that they are separate objects.

Full validation of the proposed spatio-temporal fusion step cannot be performed at this stage due to the lack of similar vegetation maps, built by considering the multitemporal information and the rotation practices. Thus, qualitative comparison is carried out between the multitemporal vegetation map obtained from the spatio-temporal fusion step and the cadastral map shown in Figure 6.6. Administrative boundaries were extracted from this map to help for a better visualization. Figure 6.8 shows (a) the 2015-2016 ground truth (closer look of the cadastral map shown in Figure 6.6) and (d) the $X_{veg_{map}}$ for the same area. As a further comparison with the proposed approach, vegetation maps were computed by using one single image at t_1 and t_{45} (Figure 6.8 (b) and (c)). When a single image is used, underestimation occurs: by using t_1 few areas are detected, and while using t_{45} many areas are detected as a single one (e.g., light green and blue ones). Similar situations can be seen for other images in the NDVI-SITS, but when the proposed spatio-temporal fusion step is considered, detection and separation of the areas is accurate. Some issues are found when neighboring crops have a similar vegetation cycle over the whole period (e.g., areas enclosed in the magenta box in Figure 6.8 (a) and (d)), resulting in a poor separation of the areas. Other differences can be seen while looking at other areas, where the shape and use of some of the areas is different to that indicated by the administrative boundaries and the ground truth (e.g., light-green boxes in Figure 6.8 (d)). This is because the administrative boundaries do not account for the real way in which the fields/areas are used over the time. Thus, the relevance of the proposed approach in creating a multitemporal crop field map where the different crop fields are automatically detected and separated, while accounting for rotation practices, becomes clear.

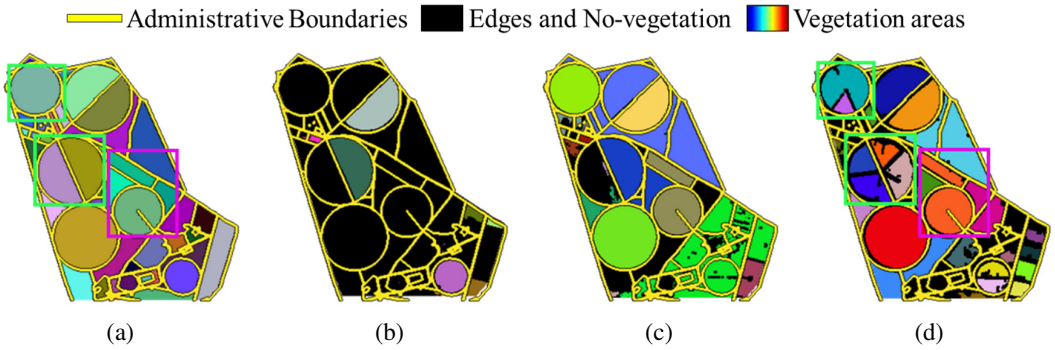


Figure 6.8. (a) 2015-2016 ground truth, vegetation map of (b) $\boldsymbol{X_1}$ and (c) $\boldsymbol{X_{45}}$; and (b) $\boldsymbol{X_{veg_{map}}}$.

6.3.3 NDVI-SITS Reconstruction

The NDVI-SITS are reconstructed with the result of having continuous and regular acquisitions in time. Here the accuracy assessment is performed at both quantitative (based on Mean Square Error - MSE) and qualitative levels and at single object level. In literature, one of the methods giving the best performance for NDVI-SITS smoothing is based on Fourier series algorithms and known as Harmonic ANalysis of Time Series (HANTS) [185]. HANTS expects: i) a harmonic behavior from the data, ii) a known number of cropping cycles and iii) regularly sampled data. Here, it was modified from its original setup in order to assume a near-daily acquisition of the data and render it comparable with our proposed approach. One important parameter to be setup for HANTS is the number of frequencies. This number is usually defined by the user and is related to the number of expected cropping cycles in the study area. Since 2 cropping cycles are usually expected, it is traditionally setup as 3. In our case, and since no a priori information about the rotation practices is available, and given the heterogeneous characteristics of the area, only a trade-off value can be selected. Following some trial and error tests, the number of frequencies in HANTS was setup to 4 for all the crop fields. In the case of the proposed approach, a couple of trials were enough to establish the setup parameters as: $w = 20$, $n_{HL_1} = 12$, $T_{MSE} = 0.01$ and $I = 100$. The computational time for both methods is low and similar and the average MSE value among the 307 objects in $X_{veg_{map}}$ was of 9.7 ×

10^{-3} for HANTS and 3.6×10^{-3} for the proposed approach. From the average MSE value, we can see that the proposed approach performs better than HANTS. If we take a closer look to the MSE value for each ω_v in $X_{veg_{map}}$ (Figure 6.9), we can see the improvement in a more clear way. As shown by the box plot in Figure 6.9, the MSE of HANTS never overpasses that of the proposed approach and its variability is also higher.

Figure 6.10 shows the smoothed result from HANTS and the proposed approach for two of the objects in ω_v (NDVI values have been scaled by 100 for visualization purposes). Where it is possible to appreciate: i) the complexity and variability of one object with respect to the other (one cropping cycle against four), and ii) how the harmonic characteristic of HANTS makes it fail in reconstructing the real behavior of the vegetation. More in detail, the harmonic behavior of HANTS allows it to work relatively well when the object shows only one cropping cycle (Figure 6.10 (a)), but it tends to fail as the number of cropping cycles increases (Figure 6.10 (b)). This kind of behavior affects the derivation of phenological parameters since it depends on the smoothing step precision.

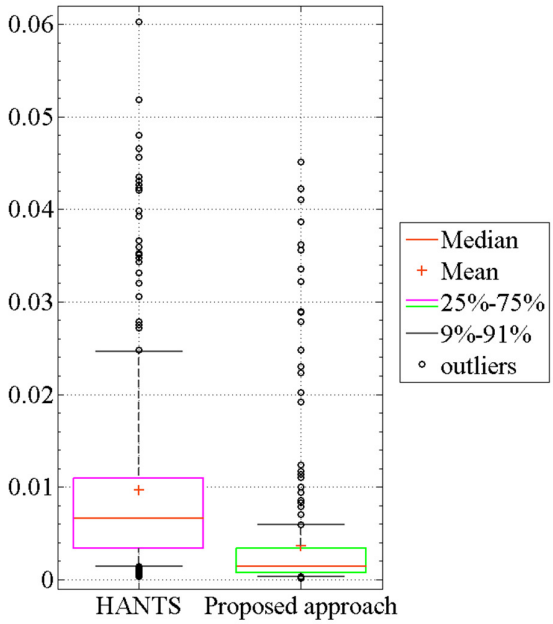
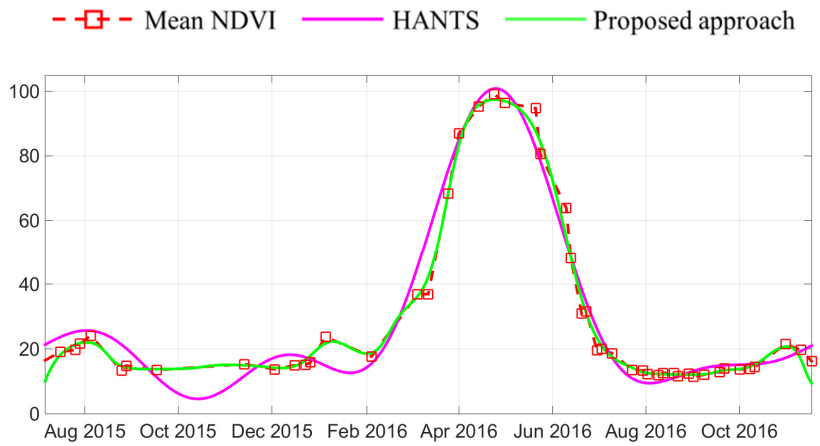


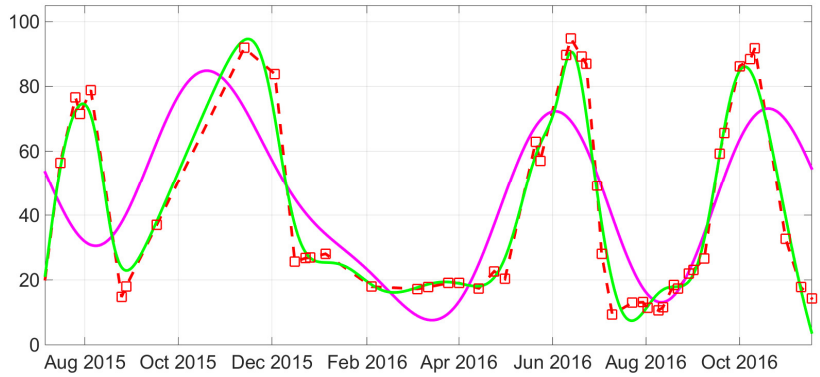
Figure 6.9. Box plot of the MSE values for HANTS and the proposed approach for each $\boldsymbol{\omega_v}$ in $\boldsymbol{X_{veg_{map}}}$.

6.3.4 Phenological Parameters Maps

Once the NDVI-SITS has been reconstructed on a daily basis, spatio-temporal information can be extracted. Here we present the results for phenological parameters derivation. No prior information was available to perform a complete accuracy assessment of all the phenological parameters. Nevertheless, given that the number of cropping cycles can be easily identified by visual inspection, a survey was carried out by considering this parameter. The number of cropping cycles of 10 crop fields was selected for evaluation. These crop fields were selected as representative of the whole study area. It was found that the proposed approach correctly detected the number of cycles for all of them, whereas HANTS was only able to detect them well in 6 out of 10 cases. Table 6.2 shows the detailed results for these 10 crop fields (the two first lines refer to the examples in Figure 6.10). The MSE for the reconstruction step is also shown. We can see how the proposed approach outperforms HANTS. The improvement is larger when more than one cropping cycle was present.



(a)



(b)

Figure 6.10. Smoothed NDVI by HANTS and proposed approach for two different crop fields: (a) single cropping cycle field; and (b) 4-cropping cycle field.

Table 6.2. MSE (10^{-3}) and detected # of cycles for HANTS and the proposed approach

# of cycles	Detection by HANTS	Detection by proposed approach	HANTS MSE	Proposed approach MSE
1	1	1	2.7	0.7
4	3	4	45.7	3.2
1	1	1	8.1	4.9
2	2	2	12.3	1.8
3	4	3	18.2	3.6
2	2	2	15.9	4.4
3	4	3	14.4	11.1
1	1	1	2.9	1.4
3	3	3	8.9	2.2
2	4	2	14.7	3.9

Once the phenological parameters are estimated for all the objects in $X_{veg_{map}}$, we proceed to analyze the LoS parameter in order to separate crop fields from other types of vegetation and to build $X_{crop_{map}}$. Such was the case of the large light blue area in Figure 6.7. By exploring the S2-SITS, it was possible to see that most of the time this area had a low NDVI value. Occasionally, vegetation was seen appearing and disappearing from one image to the other. Finally, the phenological parameter maps are built by color-coding

the different crop fields in accordance with the time (day, week or month) of phenological events. In the specific case of MV, the crop fields are color-coded in different NDVI ranges. Examples for BoS, MoS and EoS for a period of 12 months and coded per month (July 2015–June 2016) are shown in Figure 6.11.

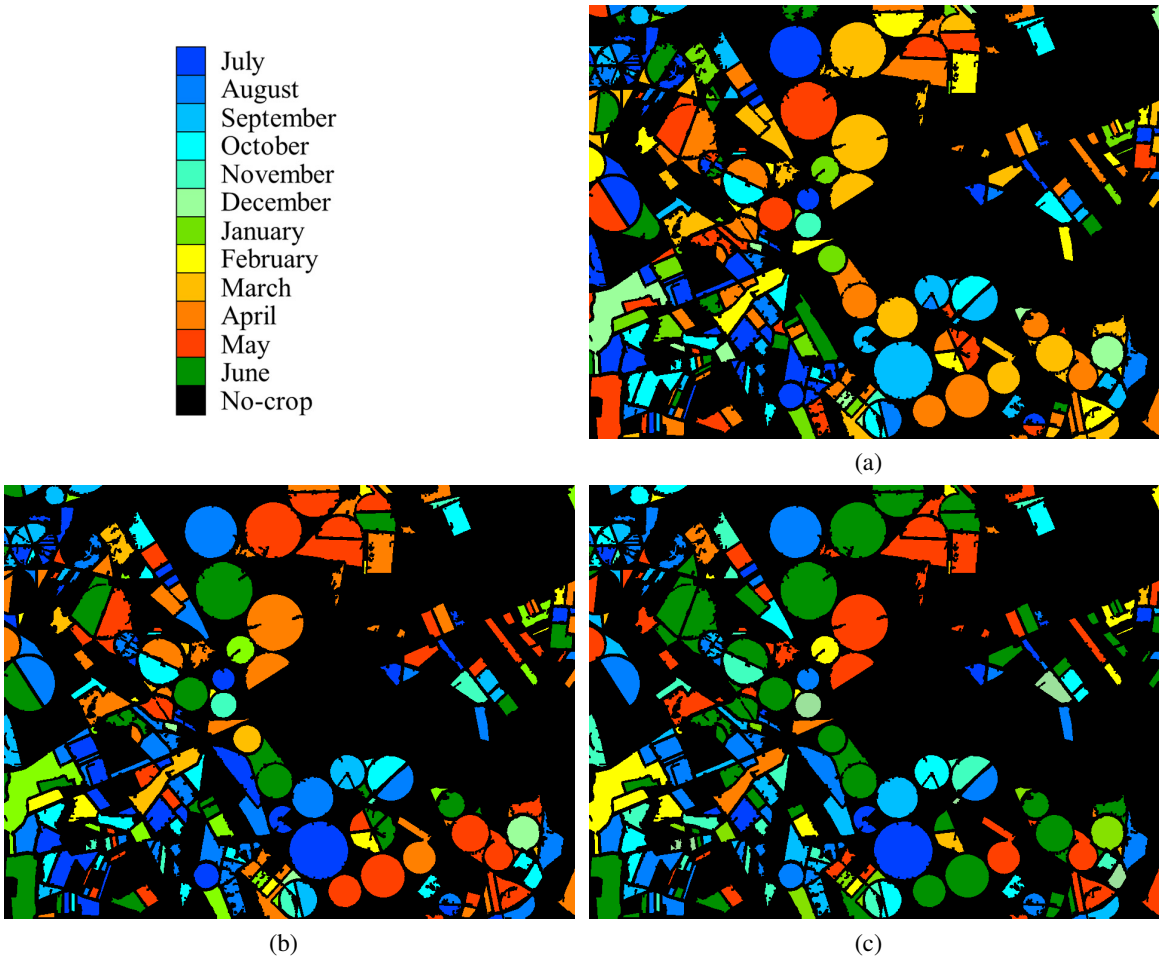


Figure 6.11. Phenology estimation maps over 12 months period for (a) BoS, (b) MoS; and (c) EoS.

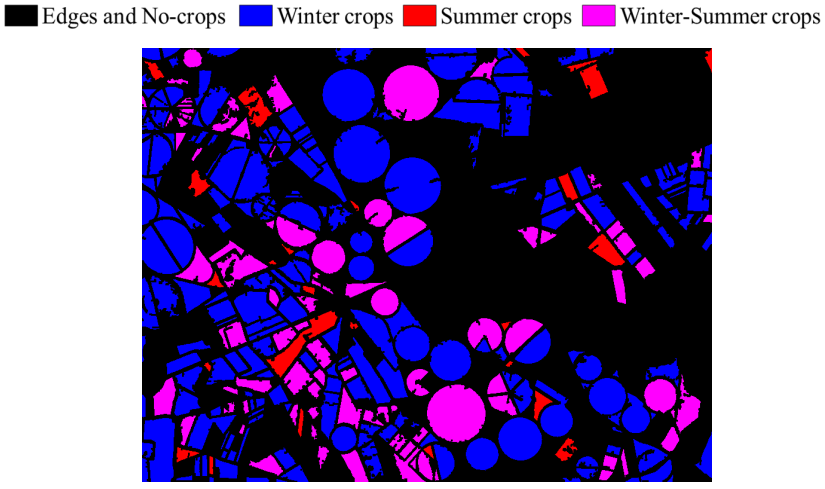


Figure 6.12. Winter-Summer map for the Barrax area in the period October 2015 to September 2016.

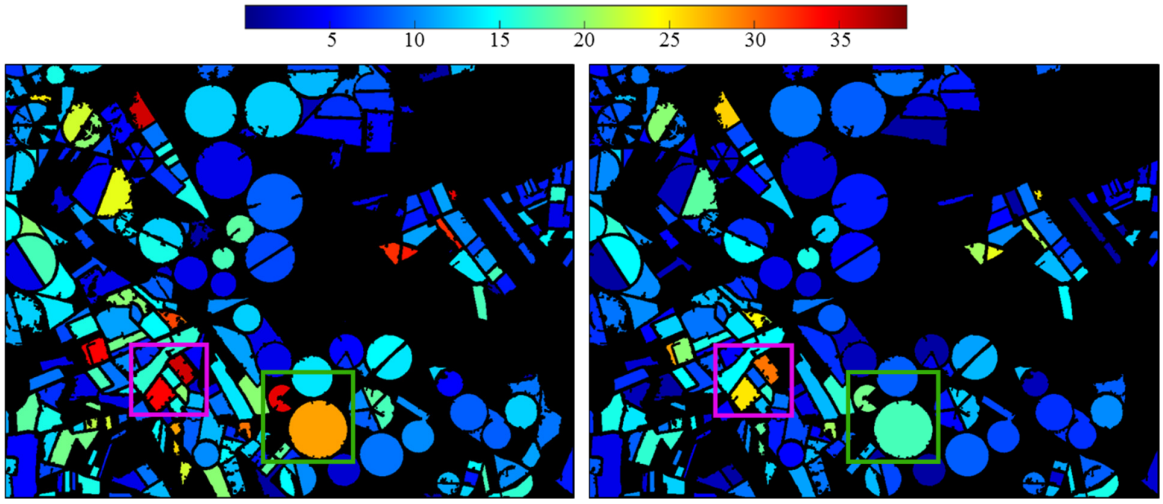
Another interesting map is the one showing the separation among winter and summer crops (Figure 6.12). Winter crops are defined as the ones with BoS, MoS and EoS in the period October-2015 to March-2016.

If the same metrics fall inside the period April to September (2016), the crops are considered as summer ones. Crop fields showing cycles in both periods are considered as winter and summer ones. The crop label in the 2 periods might or might not be the same.

6.3.5 Cumulative Indices Maps

In here we performed the spatio-temporal information extraction by means of a cumulative index (defined in (6.6)) applied over different radiometric indices. We applied all the radiometric indices described in Table 6.1, but a detailed analysis is only provided for two of them (NDVI and NDWI). The reason is that NDVI and NDWI are of easier interpretation given the great amount of studies available in literature [197], [334], [339], [340]. Therefore, let us analyze the intensity of crop field use over time. This is achieved by analyzing: i) the presence of vegetation, and ii) the content of water. The former is studied by accumulating the number of times that a certain field had vegetation (i.e., when $X_{D_{NDVI}} > 0.5$ [334]). The latter by accumulating the number of times that $X_{D_{NDWI}} > 0.5$. Cumulative indices can be studied at pixel level or at crop level by extracting the mean cumulative value for each crop field. Figure 6.13 (a) and (b), show the mean cumulative NDVI and NDWI over the available S2-SITS. Dark blue areas show NDVI and NDWI lower than 0.5 on every t_n $(n = 1, \dots, 49)$. Whereas blue and red shaded areas represent the fields with a higher vegetation greenness and thus more intensive use. When analyzing the results across crop fields, non-uniform response in time and space can be seen. Furthermore, it seems there is no direct correlation between frequently vegetated and frequently irrigated areas. In other words, fields with similar values on mean cumulative NDVI, do not show similar values of mean cumulative NDWI.

If we take a closer look to the crop fields in the magenta rectangle in Figure 6.13, we can see that they have a similar red and blue shade in the mean cumulative NDVI (i.e., they were vegetated for a similar number of days during the period), but they have a different behavior in the mean cumulative NDWI. Vice-versa, crop fields in the green rectangle show similar light blue shades for the NDWI, but different NDVI. This kind of information can become relevant for the identification of lack of irrigation, or the lack of any nutrient, etc. On the other hand, it can also be an indicator of different types of crops being cultivated in the area, though used with the same intensity. A proper validation and interpretation would require ground truth and agricultural expert analysis.



(a) (b)

Figure 6.13. Mean cumulative (a) NDVI and (b) NDWI with a value higher than 0.5.

Let us take a closer look to the blue shade crop fields inside the magenta square by extracting some statistical information such as minimum, maximum and mean values for the NDVI and NDWI SITS. For understanding purposes, let us define the largest crop field as crop 1 and the smallest one as crop 2. Figure

6.14 shows the temporal evolution of each statistical parameter for the two crop fields for NDVI (Figure 6.14 (a) and (c)) and NDWI (Figure 6.14 (b) and (d)). By observing Figure 6.14 (a) and (c), we can conclude that crop 1 and crop 2 most probably correspond to two types of crops. Since they show very similar intensity in terms of the mean cumulative NDVI, the difference of the mean cumulative NDWI and the temporal evolution of NDVI and NDWI may be associated to the fact that different types of crop have different irrigation requirements.

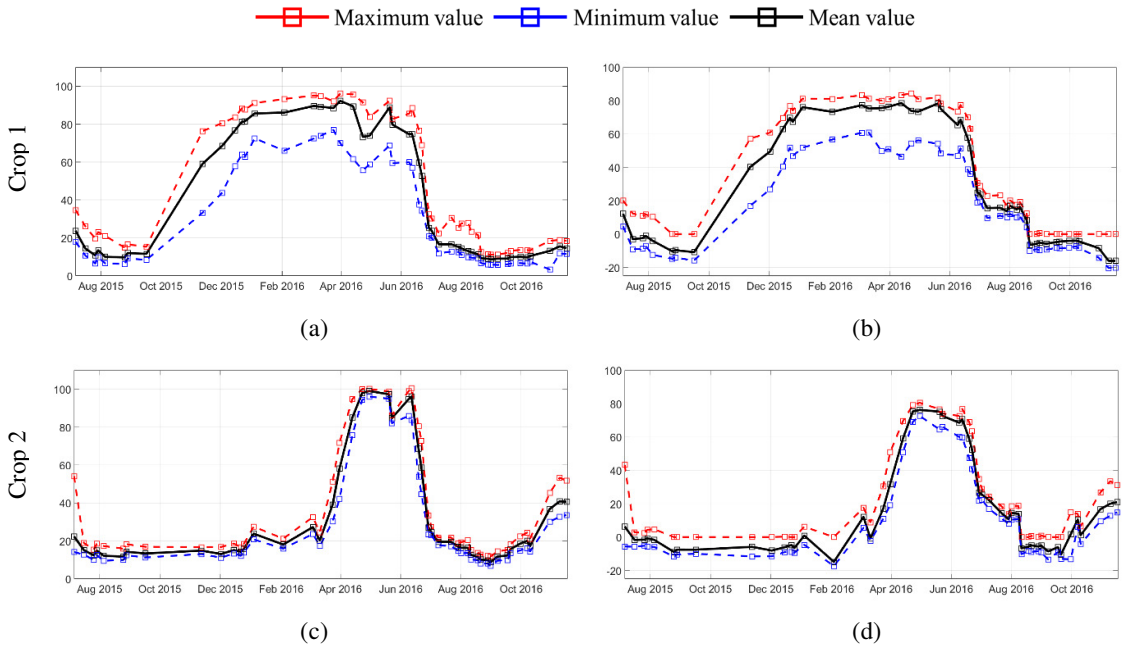


Figure 6.14. Statistical NDVI ((a) and (c)) and NDWI ((b) and (d)) analysis for crop fields 1 and 2.

6.4. Conclusions

An approach to study the spatio-temporal evolution of crop fields by jointly exploiting the spectral, spatial and temporal information of S2-SITS has been presented in this chapter. The proposed approach is composed by four main steps: i) pre-processing of S2-SITS, ii) spatio-temporal fusion, iii) daily time series reconstruction; and iv) spatio-temporal information extraction. The combination of these steps resulted in an important tool which is relevant and suitable for precision agriculture. In the pre-processing step, state-of-the-art methods were used to render the S2-SITS homogeneous and suitable for information extraction. In the second step, an approach to crop field mapping that detects, separates and labels the crop fields that have been cultivated at least once in a given period was proposed. The relevance of this step is in the capability to build a multitemporal crop field map that accounts for the temporal evolution of crops, but also for the dynamic change that a single crop field can experience from one year to the other. This step was validated by qualitatively comparing the results with a portion of ground truth available for the study area. This analysis pointed out the relevance of considering both the spatial and temporal information. Because of the same reason, the approach allows the possibility to update the multitemporal crop field map when a new acquisition is available. In the third step, a novel approach for daily S2 NDVI-SITS reconstruction was proposed. The approach is fully automatic and is able to deal with data acquired with irregular temporal spacing. To deal with irregularly sample data, the proposed approach introduces two novel stages: i) preliminary smoothing and construction of daily acquisitions and ii) a novel adaptive non-parametric regression for irregular temporal behavior at single crop field level. The results of the reconstruction were compared to those of HANTS, showing better performance both from the quantitative and qualitative viewpoints. Finally, in the fourth step, spatio-temporal information was extracted by deriving phenological parameters and by means of a cumulative radiometric index. For the phenological parameters, the number of

cropping cycles for a number of selected crop fields was verified and different thematic maps were built. From the cumulative radiometric index, the temporal evolution and behavior of different crop fields in terms of a specific radiometric index were mapped. Examples were shown for the specific cases of NDVI and NDWI indices, were a statistical analysis was performed for two crop fields. As future developments, we consider the use of other radiometric indices, as well as the integration of other edge and shape detection methods for the spatio-temporal fusion step. Possibility to include a land cover map containing information regarding agricultural areas will be also considered. Other radiometric indices will be also considered in the spatio-temporal evolution analysis step in order to improve the derivation of phenological parameters and to account for extra information relevant in the understanding of precision agriculture.

Chapter 7

7. Conclusions

This chapter draws the conclusions of the thesis by presenting a general discussion of the work and by providing a summary of the novel contributions illustrated along the document. The remaining open issues and further developments of the research activities are also discussed.

Summary and Discussion

In this thesis we have presented novel methods for unsupervised and automatic analysis of multitemporal multispectral satellite images. In particular, we focused the attention on data acquired by rather: i) multisensor VHR images (e.g., IKONOS, QuickBird, GeoEye, WorldView); or ii) high spatial and very high temporal resolution sensors (e.g., Sentinel-2). The proposed methods represent a valuable contribution to the state-of-the-art, where most of the multitemporal analysis techniques have been designed to work with: i) multitemporal VHR images acquired by single sensors, and ii) multispectral images acquired by high spatial resolution sensors, but with low temporal resolution or very high temporal resolution, but low spatial resolution. This reduces the effectiveness of existing techniques when applied to the complex multitemporal problems in both VHR multisensor and high spatial and very high temporal resolution images.

An intensive review of the literature in multitemporal analysis of remotely sensed image data was presented that allowed us to highlight the different open issues in the topic. From there, the two main problems described above were identified and addressed. Accordingly, four main novel contributions to the state-of-the-art were presented in this thesis. The first three contributions addressed the problems arising from the analysis of multisensor VHR multispectral images and were explored in detail. Whereas the fourth one dealt with the problems faced while working with dense long time series acquired by sensors with high spatial and very high temporal resolutions. The fourth one represents a starting point for further developments. Nevertheless, the contribution is relevant to the literature.

The first contribution presents an approach for unsupervised CD in multisensor multitemporal VHR images, where the possible sources of noise/changes were studied in detail and a strategy to mitigate them at the level of pre-processing and feature extraction was presented. The proposed approach defined and illustrated a data flow for effectively handling differences due to acquisition sensors. In order to deal with multispectral and multitemporal images acquired by different sensors, it integrates in the general approach the following two concepts: i) spectral, radiometric and geometric homogenization between images acquired by different sensors; and ii) detection of multiple changes by means of higher-level physical features that guarantee homogeneity over time and across sensors. Experimental results obtained on real datasets, confirmed the effectiveness of the proposed block scheme and the improvement achieved by the use of higher-level physical features (i.e., Tasseled Caps (TC) and Orthogonal Equations (OrE)) over the traditional features (i.e., TOA). In general, both TC and OrE features allowed a better separation and interpretation of changes on the ground (Ω_{Grd}) by guaranteeing that these changes were distributed in compact and well separated clusters.

In the second contribution, further attention was paid to the homogenization step and a method to generate homogeneous VHR time series focused on the homogenization of intrinsic spectral induced differences was presented. The method is based on the generation of homogenous VHR TS by means of a non-parametric regression of multisensor MT optical images. The goal was achieved while effectively normalizing the multisensor images by combining a preliminary normalization (based on ARN) with a proposed RRN method. The proposed RRN method is based on a CD-driven approach and a non-parametric regression model. In RRN, a prediction is achieved by deriving single band models with non-parametric regressions

(ANN and SVR) that allows to capture the spatial/spectral variability of VHR multisensor images. The models allowed us to predict all the spectral bands for the X_1 image, as if it would have been acquired at t_2, based on the original X_2 image and the reference un-changed pixels. Experimental results on real datasets confirmed the effectiveness of the proposed approach and the improvement in multitemporal analysis that can be achieved by using the predicted images over the preliminary normalized ones.

The third contribution further focuses on the detection of multiple changes, while relaxing the hypothesis on the knowledge of the statistical distribution of the classes. To this aim, a method to multiple CD in VHR optical images based on iterative clustering and adaptive thresholding was implemented. The iterative clustering was achieved by combining two hierarchical methods: OPTICS and region growing. Because of this and the nature of the method, it was named as Growing OPTICS (GO). The adaptive GO works in (hyper-) spherical coordinates, is distribution free (in the direction domain), unsupervised and automatically identifies the number of changes. The effectiveness of the adaptive GO, with respect to standard thresholding methods, was validated on 2 pairs of VHR single-sensor and multi-sensor images. Such datasets showed complex and noisy set-ups. GO demonstrated to be robust in handling noisy samples and overlapping clusters and reduced the computational time, when compared to OPTICS. The adaptive thresholding improved the overall accuracy, especially in the case were overlapping clusters were present in the data.

In the fourth contribution, an approach to handle images acquired by both high spatial and very high temporal resolution sensors was presented. To this aim, spectral, spatial and temporal information of Sentinel-2 satellite images time series was exploited in four different phases: i) pre-processing of S2-SITS, ii) spatio-temporal fusion, iii) daily time series reconstruction; and iv) spatio-temporal information extraction. The combination of these steps resulted in an important tool which is fully automatic and allows for the derivation of different relevant products in the precision agriculture field. In the pre-processing step, well-known state-of-the-art methods were used to render the S2-SITS homogeneous and suitable for information extraction. In the second step, an approach to crop field mapping that detects, separates and labels the crop fields that have been cultivated at least once in a given period was proposed. In the third step, a novel approach for daily S2 NDVI-SITS reconstruction was proposed. The approach is fully automatic and is able to deal with data acquired with irregularly temporal spacing. Finally, in the fourth step, spatio-temporal information was extracted by deriving phenological parameters and by means of a cumulative radiometric index. In both cases, different thematic maps were built.

Future Developments

In the different research activities, we defined and developed methods suitable for addressing the challenging problems emerging while working with: i) multisensor VHR images; and ii) high spatial and very high temporal resolution images. On the basis of the developed methods, the analysis and the experimental results carried out in the framework of this thesis, we identified some interesting direction of research as future developments.

Regarding the CD problem in multisensor multispectral VHR images, two main areas remain open to be further explored: i) mitigation of remaining Ω_{Sys}; and ii) detection of the Ω_{Grd}. In the former, selection of other spectral and/or textural features for the training stage of the homogenization process based on non-parametric regression should be considered. In this same area, future works should consider the use of more VHR multisensor optical pairs with more complex scenarios, to make the approach fully independent of the non-parametric regression configuration. Integration of ARN and harmonization methods should be also considered. In the latter case, open issues regarding the selection of initial clustering algorithm parameters in GO remain. Thus, a detailed analysis should be carried out in accordance with the spatial resolution of the sensors and the density of cloud points in the different feature spaces, to help for an easier and intuitive selection of the parameters. Future works could also consider the applicability and adaptation of the GO method for CD in images with lower spatial resolution as well as possible extensions to other applications (e.g., classification). A further analysis can be carried out by using Deep Neural Network

(DNN) non-parametric regression to achieve both the normalization and CD processes at the same time. Finally, DNN can be also used to determine which cluster is representing a specific kind of change, and to define appropriate features to identify different kinds of change along multisensors.

Regarding the exploitation of spectral, spatial and temporal information acquired by sensors with high spatial and very high temporal resolutions, the future developments are wide and not even fully known. Nevertheless, of particular attention would be the applications derived from the daily TS reconstruction method. In a first scenario, it could be applied to all the spectral bands and to all the pixels in the study area, such as that a whole TS reconstruction is achieved and analysis such as classification or CD can be performed. On a second scenario, it could be adapted as a gap-filling method where few clouds or noisy pixels are present in a given image of the TS. Finally, it can be used in the context of classification, where a crop classification map that considers multitemporal variability could be built. Such classification together with phenological parameters can be used to accurately map agricultural areas over a semester, year or decades. Interaction with agricultural experts should be considered for the design of other products relevant in agriculture precision.

List of Publications

BOOK CHAPTER

[B1]. F. Bovolo, L. Bruzzone, and **Y. T. Solano-Correa**, "Multitemporal Analysis of Remotely Sensed Image Data," in Comprehensive Remote Sensing, 1st Edition, Elsevier, vol. 2, pp. 156-185, Nov. 2017.

INTERNATIONAL JOURNALS

[J1]. E. L. Pencue-Fierro, **Y. T. Solano-Correa**, J. C. Corrales-Muñoz, and A. Figueroa-Casas, "A Semi-Supervised Hybrid Approach for Multitemporal Multi-Region Multisensor Landsat Data Classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 9, no. 12, pp. 5424–5435, Dec. 2016.

[J2]. **Y. T. Solano-Correa**, F. Bovolo, and L. Bruzzone, "An approach for unsupervised change detection in multitemporal VHR images acquired by different multispectral sensors," *Remote Sensing, MDPI*, **under revision**.

[J3]. **Y. T. Solano-Correa**, F. Bovolo, and L. Bruzzone, "Generation of homogeneous VHR time series by non-parametric regression of multisensor multitemporal images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **ready to submit**.

[J4]. **Y. T. Solano-Correa**, F. Bovolo, and L. Bruzzone, "An approach to multiple change detection in VHR optical images based on iterative clustering and adaptive thresholding," *IEEE Geosci. Remote Sens. Lett.*, **under revision**.

[J5]. **Y. T. Solano-Correa**, F. Bovolo, L. Bruzzone, and D. Fernández-Prieto, "Spatio-temporal evolution of crop fields in Sentinel-2 Satellite Image Time Series," *Remote Sens. of Environ.*, **ready to submit**.

INTERNATIONAL CONFERENCES

[C1]. **Y. T. Solano-Correa**, F. Bovolo, and L. Bruzzone, "Change detection in very high resolution multisensor optical images," 2014, vol. 9244, pp. 924410-924410–13.

[C2]. **Y. T. Solano-Correa**, L. Pencue-Fierro, and A. Figueroa-Casas, "Determining the effects of ENSO phenomena on Andean areas by applying radiometric indices on long time series," in 2015 8th International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multi-Temp), 2015, pp. 1–4.

[C3]. **Y. T. Solano-Correa**, F. Bovolo, and L. Bruzzone, "VHR time-series generation by prediction and fusion of multi-sensor images," in 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2015, pp. 3298–3301.

[C4]. **Y. T. Solano-Correa**, F. Bovolo, and L. Bruzzone, "An approach to multiple Change Detection in multisensor VHR optical images based on iterative clustering," in 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 5149–5152.

[C5]. **Y. T. Solano-Correa**, F. Bovolo, L. Bruzzone, and D. Fernández-Prieto, "Spatio-temporal evolution of crop fields in Sentinel-2 Satellite Image Time Series," in 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 2017, pp. 1–4.

[C6]. L. Bruzzone, F. Bovolo, C. Paris, **Y. T. Solano-Correa**, M. Zanetti, and D. Fernández-Prieto, "Analysis of multitemporal Sentinel-2 images in the framework of the ESA Scientific Exploitation of Operational Missions," in 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 2017, pp. 1–4.

[C7]. E. L. Pencue-Fierro, **Y. T. Solano-Correa**, J. C. Corrales-Muñoz, and A. Figueroa-Casas, "Analysis of Riparian forest buffers dynamics in Colombian basins by Landsat Time Series," in 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 2017, pp. 1–4.

[C8]. **Y. T. Solano-Correa**, F. Bovolo, L. Bruzzone, and D. Fernández-Prieto, "Derivation of cropland phenological parameters by NDVI smoothing of Sentinel-2 Satellite Image Time Series," in 2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS-2018), **accepted**.

Bibliography

[1] F. J. Ahern *et al.*, "Review Article Radiometric correction of visible and infrared remote sensing data at the Canada Centre for Remote Sensing," *Int. J. Remote Sens.*, vol. 8, no. 9, pp. 1349–1376, Sep. 1987.

[2] J. F. Moreno, "Radiometric corrections of visible/infrared satellite data over terrestrial environments: angular, atmospheric and topographic effects," in *Geoscience and Remote Sensing Symposium, 1996. IGARSS '96. "Remote Sensing for a Sustainable Future.", International*, 1996, vol. 3, pp. 1823–1825 vol.3.

[3] M. M. Rahman, G. J. Hay, I. Couloigner, B. Hemachandran, and J. Bailin, "A comparison of four relative radiometric normalization (RRN) techniques for mosaicing H-res multi-temporal thermal infrared (TIR) flight-lines of a complex urban scene," *ISPRS J. Photogramm. Remote Sens.*, vol. 106, pp. 82–94, Aug. 2015.

[4] Y. Zhang, L. Yu, M. Sun, and X. Zhu, "A Mixed Radiometric Normalization Method for Mosaicking of High-Resolution Satellite Imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. PP, no. 99, pp. 1–13, 2017.

[5] Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "Change detection in very high resolution multisensor optical images," in *Image and Signal Processing for Remote Sensing XX (SPIE)*, 2014, vol. 9244, pp. 924410-924410–13.

[6] Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "VHR time-series generation by prediction and fusion of multi-sensor images," in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 3298–3301.

[7] A. Bannari, D. Morin, G. B. Bénié, and F. J. Bonn, "A theoretical review of different mathematical models of geometric corrections applied to remote sensing images," *Remote Sens. Rev.*, vol. 13, no. 1–2, pp. 27–47, Aug. 1995.

[8] J. Moreno and J. G. Webster, "Remote Sensing Geometric Corrections," in *Wiley Encyclopedia of Electrical and Electronics Engineering*, John Wiley & Sons, Inc., 1999.

[9] T. Toutin, "Review article: Geometric processing of remote sensing images: models, algorithms and methods," *Int. J. Remote Sens.*, vol. 25, no. 10, pp. 1893–1924, May 2004.

[10] T. Toutin, "Geometric Correction of Remotely Sensed Images," in *Remote Sensing of Forest Environments*, M. A. Wulder and S. E. Franklin, Eds. Springer US, 2003, pp. 143–180.

[11] A. Marsetič, K. Oštir, and M. K. Fras, "Automatic Orthorectification of High-Resolution Optical Satellite Images Using Vector Roads," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 6035–6047, Nov. 2015.

[12] S. Dawn, V. Saxena, and B. Sharma, "Remote Sensing Image Registration Techniques: A Survey," in *Image and Signal Processing*, 2010, pp. 103–112.

[13] Y. Han, F. Bovolo, and L. Bruzzone, "Segmentation-Based Fine Registration of Very High Resolution Multitemporal Images," *IEEE Trans. Geosci. Remote Sens.*, vol. PP, no. 99, pp. 1–14, 2017.

[14] Y. Han, F. Bovolo, and L. Bruzzone, "Edge-Based Registration-Noise Estimation in VHR Multitemporal and Multisensor Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 9, pp. 1231–1235, Sep. 2016.

[15] Y. Han, F. Bovolo, and L. Bruzzone, "An Approach to Fine Coregistration Between Very High Resolution Multispectral Images Based on Registration Noise Distribution," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6650–6662, Dec. 2015.

[16] J. Le moigne, N. S. Netanyahu, and R. D. Eastman, *Image Registration for Remote Sensing*. Cambridge University Press Cambridge, 2011.

[17] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[18] L. Bruzzone and R. Cossu, "An adaptive approach to reducing registration noise effects in unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2455–2465, Nov. 2003.

[19] X. Dai and S. Khorram, "The effects of image misregistration on the accuracy of remotely sensed change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1566–1577, Sep. 1998.

[20] S. Marchesi, F. Bovolo, and L. Bruzzone, "A Context-Sensitive Technique Robust to Registration Noise for Change Detection in VHR Multispectral Images," *IEEE Trans. Image Process.*, vol. 19, no. 7, pp. 1877–1889, Jul. 2010.

[21] J. R. G. Townshend, C. O. Justice, C. Gurney, and J. McManus, "The impact of misregistration on change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 5, pp. 1054–1060, Sep. 1992.

[22] A. Singh, "Review Article Digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, Jun. 1989.

[23] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, "Detecting trend and seasonal changes in satellite image time series," *Remote Sens. Environ.*, vol. 114, no. 1, pp. 106–115, Jan. 2010.

[24] P. S. Chavez and D. J. Mackinnon, "Automatic Detection of Vegetation Changes in the Southwestern United States Using Remotely Sensed Images," *Photogramm. Eng. Remote Sens.*, vol. 60, no. 5, May 1994.

[25] T. Fung, "An Assessment Of TM Imagery For Land-cover Change Detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 4, pp. 681–684, Jul. 1990.

[26] T. Fung and E. Ledrew, "Application of principal components analysis to change detection," *Photogramm. Eng. Remote Sens.*, vol. 53, no. 12, pp. 1649–1658, 1987.

[27] D. M. Muchoney and B. N. Haack, "Change detection for monitoring forest defoliation," *Photogramm. Eng. Remote Sens.*, vol. 60, no. 10, pp. 1243–1251, 1994.

[28] J. R. G. Townshend and C. O. Justice, "Spatial variability of images and the monitoring of changes in the Normalized Difference Vegetation Index," *Int. J. Remote Sens.*, vol. 16, no. 12, pp. 2187–2195, Aug. 1995.

[29] L. Eklundh and P. Jönsson, "TIMESAT for Processing Time-Series Data from Satellite Sensors for Land Surface Monitoring," in *Multitemporal Remote Sensing*, Y. Ban, Ed. Springer International Publishing, 2016, pp. 177–194.

[30] L. Bruzzone and F. Bovolo, "A Novel Framework for the Design of Change-Detection Systems for Very-High-Resolution Remote Sensing Images," *Proc. IEEE*, vol. 101, no. 3, pp. 609–630, Mar. 2013.

[31] M. K. Ridd and J. Liu, "A Comparison of Four Algorithms for Change Detection in an Urban Environment," *Remote Sens. Environ.*, vol. 63, no. 2, pp. 95–100, Feb. 1998.

[32] H. Nemmour and Y. Chibani, "Multiple support vector machines for land cover change detection: An application for mapping urban extensions," *ISPRS J. Photogramm. Remote Sens.*, vol. 61, no. 2, pp. 125–133, Nov. 2006.

[33] B. Desclée, P. Bogaert, and P. Defourny, "Forest change detection by statistical object-based method," *Remote Sens. Environ.*, vol. 102, no. 1, pp. 1–11, May 2006.

[34] X. Yang and L. Chen, "Using multi-temporal remote sensor imagery to detect earthquake-triggered landslides," *Int. J. Appl. Earth Obs. Geoinformation*, vol. 12, no. 6, pp. 487–495, Dec. 2010.

[35] Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "Change detection in very high resolution multisensor optical images," 2014, vol. 9244, pp. 924410-924410–13.

[36] H. Zhang, J. Chen, and Z. Mao, "The research on relative radiometric normalization for change detection of multitemporal images," presented at the Image and Signal Processing for Remote Sensing XV, 2009, vol. 7477, p. 747714.

[37] P. M. Atkinson, C. Jeganathan, J. Dash, and C. Atzberger, "Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology," *Remote Sens. Environ.*, vol. 123, pp. 400–417, Aug. 2012.

[38] B. A. Bradley, R. W. Jacob, J. F. Hermance, and J. F. Mustard, "A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data," *Remote Sens. Environ.*, vol. 106, no. 2, pp. 137–145, Jan. 2007.

[39] F. Pacifici, N. Longbotham, and W. J. Emery, "The Importance of Physical Quantities for the Analysis of Multitemporal and Multiangular Optical Very High Spatial Resolution Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6241–6256, Oct. 2014.

[40] A. Boucher, K. C. Seto, and A. G. Journel, "A Novel Method for Mapping Land Cover Changes: Incorporating Time and Space With Geostatistics," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3427–3435, Nov. 2006.

[41] J. M. B. Carreiras, J. M. C. Pereira, Y. E. Shimabukuro, and D. Stroppiana, "Evaluation of compositing algorithms over the Brazilian Amazon using SPOT-4 VEGETATION data," *Int. J. Remote Sens.*, vol. 24, no. 17, pp. 3427–3440, Jan. 2003.

[42] B. Holben and R. S. Fraser, "Red and near-infrared sensor response to off-nadiir viewing," *Int. J. Remote Sens.*, vol. 5, no. 1, pp. 145–160, Jan. 1984.

[43] H. Kobayashi and D. G. Dye, "Atmospheric conditions for monitoring the long-term vegetation dynamics in the Amazon using normalized difference vegetation index," *Remote Sens. Environ.*, vol. 97, no. 4, pp. 519–525, Sep. 2005.

[44] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: image segmentation using expectation-maximization and its application to image querying," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 8, pp. 1026–1038, Aug. 2002.

[45] D. Y. Huang, T. W. Lin, and W. C. Hu, "Automatic multilevel thresholding based on two-stage Otsu's method with cluster determination by valley estimation," in *International Journal of Innovative Computing, Information and Control*, 2011, vol. 7, pp. 5631–5644.

[46] A. Sedaghat, M. Mokhtarzade, and H. Ebadi, "Uniform Robust Scale-Invariant Feature Matching for Optical Remote Sensing Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4516–4527, Nov. 2011.

[47] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "Change detection for high resolution satellite images, based on SIFT descriptors and an a contrario approach," in *2014 IEEE Geoscience and Remote Sensing Symposium*, 2014, pp. 1281–1284.

[48] M. N. Klaric *et al.*, "GeoCDX: An Automated Change Detection and Exploitation System for High-Resolution Satellite Imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2067–2086, Apr. 2013.

[49] G. Hong and Y. Zhang, "A comparative study on radiometric normalization using high resolution satellite images," *Int. J. Remote Sens.*, vol. 29, no. 2, pp. 425–438, Jan. 2008.

[50] F. Bovolo and L. Bruzzone, "A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in the Polar Domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 218–236, Jan. 2007.

[51] F. Bovolo, S. Marchesi, and L. Bruzzone, "A Framework for Automatic and Unsupervised Detection of Multiple Changes in Multitemporal Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2196–2212, Jun. 2012.

[52] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.

[53] L. Bruzzone and D. Fernández-Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1171–1182, May 2000.

[54] Z. Yetgin, "Unsupervised Change Detection of Satellite Images Using Local Gradual Descent," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1919–1929, May 2012.

[55] L. Bruzzone and D. Fernández-Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 452–466, Apr. 2002.

[56] M. Zanetti, F. Bovolo, and L. Bruzzone, "Rayleigh-Rice Mixture Parameter Estimation via EM Algorithm for Change Detection in Multispectral Images," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5004–5016, Dec. 2015.

[57] F. Bovolo and L. Bruzzone, "An adaptive thresholding approach to multiple-change detection in multispectral images," in *2011 IEEE International Geoscience and Remote Sensing Symposium*, 2011, pp. 233–236.

[58] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Pearson Education, 2009.

[59] M. Petrou and C. Petrou, *Image Processing: The Fundamentals*. John Wiley & Sons, 2010.

[60] Y. T. Solano-Correa, F. Bovolo, L. Bruzzone, and D. Fernández-Prieto, "Spatio-temporal evolution of crop fields in Sentinel-2 Satellite Image Time Series," in *2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, 2017, pp. 1–4.

[61] Y. Xie, Z. Sha, and M. Yu, "Remote sensing imagery in vegetation mapping: a review," *J. Plant Ecol.*, vol. 1, no. 1, pp. 9–23, Mar. 2008.

[62] F. Bovolo, L. Bruzzone, and Y. T. Solano-Correa, "Multitemporal Analysis of Remotely Sensed Image Data," in *Comprehensive Remote Sensing*, S. Liang, Ed. Oxford: Elsevier, 2018, pp. 156–185.

[63] G. Moser, E. Angiati, and S. B. Serpico, "Multiscale Unsupervised Change Detection on Optical Images by Markov Random Fields and Wavelets," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 725–729, Jul. 2011.

[64] F. Bovolo and L. Bruzzone, "An Adaptive Technique based on Similarity Measures for Change Detection in Very High Resolution SAR Images," in *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*, 2008, vol. 3, p. III – 158–III – 161.

[65] J. G. Liu, A. Black, H. Lee, H. Hanaizumi, and J. M. Moore, "Land surface change detection in a desert area in Algeria using multi-temporal ERS SAR coherence images," *Int. J. Remote Sens.*, vol. 22, no. 13, pp. 2463–2477, Jan. 2001.

[66] S. K. Pal, A. Ghosh, and B. U. Shankar, "Segmentation of remotely sensed images with fuzzy thresholding, and quantitative evaluation," *Int. J. Remote Sens.*, vol. 21, no. 11, pp. 2269–2300, Jan. 2000.

[67] F. Bovolo, "A Multilevel Parcel-Based Approach to Change Detection in Very High Resolution Multitemporal Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 1, pp. 33–37, Jan. 2009.

[68] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2963–2972, Dec. 2005.

[69] T. Celik, "A Bayesian approach to unsupervised multiscale change detection in synthetic aperture radar images," *Signal Process.*, vol. 90, no. 5, pp. 1471–1485, May 2010.

[70] T. Celik and K. K. Ma, "Multitemporal Image Change Detection Using Undecimated Discrete Wavelet Transform and Active Contours," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 706–716, Feb. 2011.

[71] M. Gong, Z. Zhou, and J. Ma, "Change Detection in Synthetic Aperture Radar Images based on Image Fusion and Fuzzy Clustering," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2141–2151, Apr. 2012.

[72] B. Hou, Q. Wei, Y. Zheng, and S. Wang, "Unsupervised Change Detection in SAR Image Based on Gauss-Log Ratio Image Fusion and Compressed Projection," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 8, pp. 3297–3317, Aug. 2014.

[73] W. Dierking and H. Skriver, "Change detection for thematic mapping by means of airborne multitemporal polarimetric SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 3, pp. 618–636, Mar. 2002.

[74] S. Quegan, T. L. Toan, J. J. Yu, F. Ribbes, and N. Floury, "Multitemporal ERS SAR analysis applied to forest mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 741–753, Mar. 2000.

[75] Y. Bazi, L. Bruzzone, and F. Melgani, "An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 874–887, Apr. 2005.

[76] C. Carincotte, S. Derrode, and S. Bourennane, "Unsupervised change detection on SAR images using fuzzy hidden Markov chains," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp. 432–441, Feb. 2006.

[77] J. Inglada and G. Mercier, "A New Statistical Similarity Measure for Change Detection in Multitemporal SAR Images and Its Extension to Multiscale Change Analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1432–1445, May 2007.

[78] L. Bruzzone and D. Fernández-Prieto, "An adaptive parcel-based technique for unsupervised change detection," *Int. J. Remote Sens.*, vol. 21, no. 4, pp. 817–822, Jan. 2000.

[79] G. Mercier, G. Moser, and S. B. Serpico, "Conditional Copulas for Change Detection in Heterogeneous Remote Sensing Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1428–1441, May 2008.

[80] M. dalla Mura, J. A. Benediktsson, F. Bovolo, and L. Bruzzone, "An Unsupervised Technique Based on Morphological Filters for Change Detection in Very High Resolution Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 433–437, Jul. 2008.

[81] P. Du, S. Liu, P. Gamba, K. Tan, and J. Xia, "Fusion of Difference Images for Change Detection Over Urban Areas," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 5, no. 4, pp. 1076–1086, Aug. 2012.

[82] N. Falco, M. D. Mura, F. Bovolo, J. A. Benediktsson, and L. Bruzzone, "Change Detection in VHR Images Based on Morphological Attribute Profiles," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 636–640, May 2013.

[83] S. Liu, L. Bruzzone, F. Bovolo, and P. Du, "Hierarchical Unsupervised Change Detection in Multitemporal Hyperspectral Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 244–260, Jan. 2015.

[84] S. Liu, L. Bruzzone, F. Bovolo, M. Zanetti, and P. Du, "Sequential Spectral Change Vector Analysis for Iteratively Discovering and Detecting Multiple Changes in Hyperspectral Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4363–4378, Aug. 2015.

[85] S. Liu, L. Bruzzone, F. Bovolo, and P. Du, "A novel sequential spectral change vector analysis for representing and detecting multiple changes in hyperspectral images," in *2014 IEEE Geoscience and Remote Sensing Symposium*, 2014, pp. 4656–4659.

[86] M. Zanetti and L. Bruzzone, "A generalized statistical model for binary change detection in multispectral images," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 3378–3381.

[87] J. Chen, X. Chen, X. Cui, and J. Chen, "Change Vector Analysis in Posterior Probability Space: A New Method for Land Cover Change Detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 2, pp. 317–321, Mar. 2011.

[88] P. R. Marpu, P. Gamba, and M. J. Canty, "Improving Change Detection Results of IR-MAD by Eliminating Strong Changes," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 799–803, Jul. 2011.

[89] A. A. Nielsen, "Kernel Maximum Autocorrelation Factor and Minimum Noise Fraction Transformations," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 612–624, Mar. 2011.

[90] A. A. Nielsen, "The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 463–478, Feb. 2007.

[91] A. A. Nielsen and K. Conradsen, "Multivariate alteration detection (MAD) in multispectral, bi-temporal image data: A new approach to change detction studies," Informatics and Mathematical Modelling (IMM), Technical University of Denmark, Report, 1997.

[92] G. F. Byrne, P. F. Crapper, and K. K. Mayo, "Monitoring land-cover change by principal component analysis of multitemporal landsat data," *Remote Sens. Environ.*, vol. 10, no. 3, pp. 175–184, Nov. 1980.

[93] J. A. Richards and A. K. Milne, "MAPPING FIRE BURNS AND VEGETATION REGENERATION USING PRINCIPAL COMPONENTS ANALYSIS.," 1983, vol. 2, p. 5. 1–5. 6.

[94] J. B. Collins and C. E. Woodcock, "Change detection using the Gramm-Schmidt transformation applied to mapping forest mortality," *Remote Sens. Environ.*, vol. 50, no. 3, pp. 267–279, Dec. 1994.

[95] Y. T. Solano-Correa, F. Bovolo, and L. Bruzzone, "An approach to multiple Change Detection in multisensor VHR optical images based on iterative clustering," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 5149–5152.

[96] T. Celik, "Multiscale Change Detection in Multitemporal Satellite Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 820–824, Oct. 2009.

[97] S. Li, L. Fang, and H. Yin, "Multitemporal Image Change Detection Using a Detail-Enhancing Approach With Nonsubsampled Contourlet Transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 5, pp. 836–840, Sep. 2012.

[98] Y. Bazi, F. Melgani, and H. D. Al-Sharari, "Unsupervised Change Detection in Multispectral Remotely Sensed Imagery With Level Set Methods," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 8, pp. 3178–3187, Aug. 2010.

[99] L. Huo, X. Feng, C. Huo, Z. Zhou, and C. Pan, "Change Field: A New Change Measure for VHR Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 10, pp. 1812–1816, Oct. 2014.

[100] K. Karantzalos, "Recent Advances on 2D and 3D Change Detection in Urban Environments from Remote Sensing Data," in *Computational Approaches for Urban Environments*, Springer, Cham, 2015, pp. 237–272.

[101] Y. Tang and L. Zhang, "Urban Change Analysis with Multi-Sensor Multispectral Imagery," *Remote Sens.*, vol. 9, no. 3, p. 252, Mar. 2017.

[102] A. Argyridis and D. P. Argialas, "Building change detection through multi-scale GEOBIA approach by integrating deep belief networks with fuzzy ontologies," *Int. J. Image Data Fusion*, vol. 7, no. 2, pp. 148–171, Apr. 2016.

[103] J. Tian, S. Cui, and P. Reinartz, "Building Change Detection Based on Satellite Stereo Imagery and Digital Surface Models," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 406–417, Jan. 2014.

[104] B. Kozhikkodan-Veettil and R. Pereira-Zanardi, "A comparative study of various urban change detection techniques using high spatial resolution commercial satellite images: Quickbird and Worldview-2," *Int. J. Adv. Remote Sens. GIS*, vol. 1, no. 1, pp. 76–84, Aug. 2012.

[105] J. Inglada and G. Alain, "On the possibility of automatic multisensor image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2104–2120, Oct. 2004.

[106] Z. g Liu, G. Mercier, J. Dezert, and Q. Pan, "Change Detection in Heterogeneous Remote Sensing Images Based on Multidimensional Evidential Reasoning," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 168–172, Jan. 2014.

[107] F. Orsomando, P. Lombardo, M. Zavagli, and M. Costantini, "SAR and Optical Data Fusion for Change Detection," in *2007 Urban Remote Sensing Joint Event*, 2007, pp. 1–9.

[108] D. Brunner, G. Lemoine, and L. Bruzzone, "Earthquake Damage Assessment of Buildings Using VHR Optical and SAR Imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2403–2420, May 2010.

[109] L. Bruzzone and D. Fernández-Prieto, "Unsupervised change detection in multisource and multisensor remote sensing images," in *IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No.00CH37120)*, 2000, vol. 6, pp. 2441–2443 vol.6.

[110] Y. Ban, O. Yousif, and H. Hu, "Fusion of SAR and Optical Data for Urban Land Cover Mapping and Change Detection," in *Global Urban Monitoring and Assessment through Earth Observation*, CRC Press, 2014, pp. 353–386.

[111] B. Mishra and J. Susaki, "Optical and SAR data integration for automatic change pattern detection," in *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2014, vol. II-7, pp. 39–46.

[112] J. Ma, M. Gong, and Z. Zhou, "Wavelet Fusion on Ratio Images for Change Detection in SAR Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 6, pp. 1122–1126, Nov. 2012.

[113] P. Du, S. Liu, J. Xia, and Y. Zhao, "Information fusion techniques for change detection from multitemporal remote sensing images," *Inf. Fusion*, vol. 14, no. 1, pp. 19–27, Jan. 2013.

[114] T. Fung and E. Ledrew, "The determination of optimal threshold levels for change detection using various accuracy indices," *Photogramm. Eng. Remote Sens.*, vol. 54, no. 10, pp. 1449–1454, 1988.

[115] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," 1976.

[116] L. Bruzzone and R. Cossu, "Analysis of Multitemporal Remote-Sensing Images for Change Detection: Bayesian Thresholding Approaches," University of Trento, Departmental Technical Report, 2002.

[117] L. Bruzzone and D. Fernández-Prieto, "A minimum-cost thresholding technique for unsupervised change detection," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3539–3544, Jan. 2000.

[118] M. Hao, W. Shi, H. Zhang, and C. Li, "Unsupervised Change Detection With Expectation-Maximization-Based Level Set," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 210–214, Jan. 2014.

[119] L. Bruzzone and S. B. Serpico, "Detection of changes in remotely-sensed images by the selective use of multi-spectral information," *Int. J. Remote Sens.*, vol. 18, no. 18, pp. 3883–3888, Dec. 1997.

[120] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.

[121] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognit.*, vol. 19, no. 1, pp. 41–47, Jan. 1986.

[122] L.-K. Huang and M.-J. J. Wang, "Image thresholding by minimizing the measures of fuzziness," *Pattern Recognit.*, vol. 28, no. 1, pp. 41–51, Jan. 1995.

[123] F. Bovolo, L. Bruzzone, and M. Marconcini, "A Novel Approach to Unsupervised Change Detection Based on a Semisupervised SVM and a Similarity Measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2070–2082, Jul. 2008.

[124] T. Celik, "Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and -Means Clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 772–776, Oct. 2009.

[125] A. Ghosh, B. N. Subudhi, and L. Bruzzone, "Integration of Gibbs Markov Random Field and Hopfield-Type Neural Networks for Unsupervised Change Detection in Remotely Sensed Multitemporal Images," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3087–3096, Aug. 2013.

[126] S. U. Lee, S. Yoon Chung, and R. H. Park, "A comparative performance study of several global thresholding techniques for segmentation," *Comput. Vis. Graph. Image Process.*, vol. 52, no. 2, pp. 171–190, Nov. 1990.

[127] F. Melgani, G. Moser, and S. B. Serpico, "Unsupervised change-detection methods for remote-sensing images," *Opt. Eng.*, vol. 41, no. 12, pp. 3288–3297, 2002.

[128] J. Muñoz-Mari, F. Bovolo, L. Gomez-Chova, L. Bruzzone, and G. Camp-Valls, "Semisupervised One-Class Support Vector Machines for Classification of Remote Sensing Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 8, pp. 3188–3197, Aug. 2010.

[129] P. L. Rosin and E. Ioannidis, "Evaluation of global image thresholding for change detection," *Pattern Recognit. Lett.*, vol. 24, no. 14, pp. 2345–2356, Oct. 2003.

[130] M. Volpi, D. Tuia, G. Camps-Valls, and M. Kanevski, "Unsupervised Change Detection With Kernels," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 6, pp. 1026–1030, Nov. 2012.

[131] Y. Bazi, L. Bruzzone, and F. Melgani, "Automatic identification of the number and values of decision thresholds in the log-ratio image for change detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 3, pp. 349–353, Jul. 2006.

[132] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and C. Zoppetti, "Nonparametric Change Detection in Multitemporal SAR Images Based on Mean-Shift Clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2022–2031, Apr. 2013.

[133] S. Ghosh, L. Bruzzone, S. Patra, F. Bovolo, and A. Ghosh, "A Context-Sensitive Technique for Unsupervised Change Detection Based on Hopfield-Type Neural Networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 778–789, Mar. 2007.

[134] M. D. Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Morphological Attribute Profiles for the Analysis of Very High Resolution Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3747–3762, Oct. 2010.

[135] F. Bovolo and L. Bruzzone, "An adaptive multiscale random field technique for unsupervised change detection in VHR multitemporal images," in *2009 IEEE International Geoscience and Remote Sensing Symposium*, 2009, vol. 4, p. IV-777-IV-780.

[136] T. Celik and K. K. Ma, "Unsupervised Change Detection for Satellite Images Using Dual-Tree Complex Wavelet Transform," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1199–1210, Mar. 2010.

[137] B. Jeon and D. A. Landgrebe, "Classification with spatio-temporal interpixel class dependency contexts," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 4, pp. 663–672, Jul. 1992.

[138] L. Bruzzone, D. Fernández-Prieto, and S. B. Serpico, "A neural-statistical approach to multitemporal and multisource remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1350–1359, May 1999.

[139] L. Bruzzone and D. Fernández-Prieto, "Unsupervised retraining of a maximum likelihood classifier for the analysis of multitemporal remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 456–460, Feb. 2001.

[140] L. Bruzzone and S. B. Serpico, "An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 858–867, Jul. 1997.

[141] A. H. S. Solberg, T. Taxt, and A. K. Jain, "A Markov random field model for classification of multisource satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 1, pp. 100–113, Jan. 1996.

[142] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1992.

[143] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines*. Cambridge University Press Cambridge, 2000.

[144] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 2013.

[145] M. Gong, T. Zhan, P. Zhang, and Q. Miao, "Superpixel-Based Difference Representation Learning for Change Detection in Multispectral Remote Sensing Images," *IEEE Trans. Geosci. Remote Sens.*, vol. PP, no. 99, pp. 1–16, 2017.

[146] H. Zhang, M. Gong, P. Zhang, L. Su, and J. Shi, "Feature-Level Change Detection Using Deep Representation and Feature Change Analysis for Multispectral Imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 11, pp. 1666–1670, Nov. 2016.

[147] P. J. Howarth and G. M. Wickware, "Procedures for change detection using Landsat digital data," *Int. J. Remote Sens.*, vol. 2, no. 3, pp. 277–291, 1981.

[148] F. Ling, W. Li, Y. Du, and X. Li, "Land Cover Change Mapping at the Subpixel Scale With Different Spatial-Resolution Remotely Sensed Imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 1, pp. 182–186, Jan. 2011.

[149] F. Pacifici, F. D. Frate, C. Solimini, and W. J. Emery, "An Innovative Neural-Net Method to Detect Temporal Changes in High-Resolution Optical Satellite Imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 9, pp. 2940–2952, Sep. 2007.

[150] P. Serra, X. Pons, and D. Saurí, "Post-classification change detection with data from different sensors: Some accuracy considerations," *Int. J. Remote Sens.*, vol. 24, no. 16, pp. 3311–3340, Jan. 2003.

[151] P. C. Smits and A. Annoni, "Updating land-cover maps by using texture information from very high-resolution space-borne imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1244–1254, May 1999.

[152] P. Kempeneers, F. Sedano, P. Strobl, D. O. McInerney, and J. San-Miguel-Ayanz, "Increasing Robustness of Postclassification Change Detection Using Time Series of Land Cover Maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3327–3339, Sep. 2012.

[153] M. Volpi, D. Tuia, F. Bovolo, M. Kanevski, and L. Bruzzone, "Supervised change detection in VHR images using contextual information and support vector machines," *Int. J. Appl. Earth Obs. Geoinformation*, vol. 20, pp. 77–85, Feb. 2013.

[154] R. A. Schowengerdt, *Techniques for Image Processing and Classifications in Remote Sensing*. Academic Press, 2012.

[155] L. Bruzzone, R. Cossu, and G. Vernazza, "Detection of land-cover transitions by combining multidate classifiers," *Pattern Recognit. Lett.*, vol. 25, no. 13, pp. 1491–1500, Oct. 2004.

[156] D. Fernández-Prieto and O. Arino, "A partially supervised change-detection technique," in *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217)*, 2001, vol. 1, pp. 196–198 vol.1.

[157] R. Cossu, S. Chaudhuri, and L. Bruzzone, "A context-sensitive Bayesian technique for the partially supervised classification of multitemporal images," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 3, pp. 352–356, Jul. 2005.

[158] T. M. Pellizzeri, P. Lombardo, P. Gamba, and F. dell'Acqua, "Multisource urban classification: joint processing of optical and SAR data for land cover mapping," in *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477)*, 2003, vol. 2, pp. 1044–1046 vol.2.

[159] B. Demir, F. Bovolo, and L. Bruzzone, "Classification of Time Series of Multispectral Images With Limited Training Data," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3219–3233, Aug. 2013.

[160] B. Demir, F. Bovolo, and L. Bruzzone, "Updating Land-Cover Maps by Classification of Image Time Series: A Novel Change-Detection-Driven Transfer Learning Approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 300–312, Jan. 2013.

[161] B. Demir, F. Bovolo, and L. Bruzzone, "Detection of Land-Cover Transitions in Multitemporal Remote Sensing Images With Active-Learning-Based Compound Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1930–1941, May 2012.

[162] A. Bastarrika, E. Chuvieco, and M. P. Martin, "Automatic Burned Land Mapping From MODIS Time Series Images: Assessment in Mediterranean Ecosystems," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3401–3413, Sep. 2011.

[163] A. A. Gitelson, "Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation," *J. Plant Physiol.*, vol. 161, no. 2, pp. 165–173, 2004.

[164] E. F. Lambin and M. Linderman, "Time series of remote sensing data for land change science," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 7, pp. 1926–1928, Jul. 2006.

[165] A. Viña, G. M. Henebry, and A. A. Gitelson, "Satellite monitoring of vegetation dynamics: Sensitivity enhancement by the wide dynamic range vegetation index," *Geophys Res Lett*, vol. 31, no. 4, p. L04503, Feb. 2004.

[166] J. E. Vogelmann *et al.*, "Monitoring Landscape Change for LANDFIRE Using Multi-Temporal Satellite Imagery and Ancillary Data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 4, no. 2, pp. 252–264, Jun. 2011.

[167] L. Bruzzone, M. Marconcini, U. Wegmuller, and A. Wiesmann, "An advanced system for the automatic classification of multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1321–1334, Jun. 2004.

[168] J. R. Eastman and M. Filk, "Long sequence time series evaluation using standardized principal components," *Photogramm. Eng. Remote Sens.*, vol. 59, no. 6, pp. 991–996, 1993.

[169] B. P. Salmon, J. C. Olivier, K. J. Wessels, W. Kleynhans, F. Van den Bergh, and K. C. Steenkamp, "Unsupervised land cover change detection: Meaningful sequential time series analysis," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 4, no. 2, pp. 327–335, 2011.

[170] K. S. Chen, H. W. Wang, C. T. Wang, and W. Y. Chang, "A Study of Decadal Coastal Changes on Western Taiwan Using a Time Series of ERS Satellite SAR Images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 4, no. 4, pp. 826–835, Dec. 2011.

[171] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 1998, vol. 454, pp. 903–995.

[172] S. Jamali, P. Jönsson, L. Eklundh, J. Ardö, and J. Seaquist, "Detecting changes in vegetation trends using time series segmentation," *Remote Sens. Environ.*, vol. 156, pp. 182–195, Jan. 2015.

[173] D. P. Ward *et al.*, "Floodplain inundation and vegetation dynamics in the Alligator Rivers region (Kakadu) of northern Australia assessed using optical and radar remote sensing," *Remote Sens. Environ.*, vol. 147, pp. 43–55, May 2014.

[174] D. C. Morton *et al.*, "Mapping canopy damage from understory fires in Amazon forests using annual time series of Landsat and MODIS data," *Remote Sens. Environ.*, vol. 115, no. 7, pp. 1706–1720, Jul. 2011.

[175] "Sentinel Online - ESA," 2017. [Online]. Available: https://sentinels.copernicus.eu/web/sentinel/home. [Accessed: 15-Mar-2017].

[176] M. Immitzer, F. Vuolo, and C. Atzberger, "First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe," *Remote Sens.*, vol. 8, no. 3, p. 166, Feb. 2016.

[177] E. Mandanici and G. Bitelli, "Preliminary Comparison of Sentinel-2 and Landsat 8 Imagery for a Combined Use," *Remote Sens.*, vol. 8, no. 12, p. 1014, Dec. 2016.

[178] A. Verhegghen *et al.*, "The Potential of Sentinel Satellites for Burnt Area Mapping and Monitoring in the Congo Basin Forests," *Remote Sens.*, vol. 8, no. 12, p. 986, Nov. 2016.

[179] T. Arantes, M. Chaves, R. Bastos, L. M. Carvalho, and M. Oliveira, "EFFECTIVENESS OF BFAST ALGORITHM TO CHARACTERIZE TIME SERIES OF DENSE FOREST, AGRICULTURE AND PASTURE IN THE AMAZON REGION," *Theor. Appl. Eng.*, vol. 1, no. 1, Sep. 2017.

[180] F. Petitjean and J. Weber, "Efficient Satellite Image Time Series Analysis Under Time Warping," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 6, pp. 1143–1147, Jun. 2014.

[181] F. Petitjean, J. Inglada, and P. Gançarski, "Satellite image time series analysis under time warping," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 8, pp. 3081–3095, 2012.

[182] R. S. DeFries, M. Hansen, M. Steininger, R. Dubayah, R. Sohlberg, and J. R. Townshend, "Subpixel forest cover in central Africa from multisensor, multitemporal data," *Remote Sens. Environ.*, vol. 60, no. 3, pp. 228–246, Jun. 1997.

[183] S. Li, Y. Wang, P. Chen, X. Xu, C. Cheng, and B. Chen, "Spatiotemporal Fuzzy Clustering Strategy for Urban Expansion Monitoring Based on Time Series of Pixel-Level Optical and SAR Images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. PP, no. 99, pp. 1–11, 2017.

[184] K. R. McCloy and W. Lucht, "Comparative evaluation of seasonal patterns in long time series of satellite image data and simulations of a global vegetation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 1, pp. 140–153, Jan. 2004.

[185] G. J. Roerink, M. Menenti, and W. Verhoef, "Reconstructing cloudfree NDVI composites using Fourier analysis of time series," *Int. J. Remote Sens.*, vol. 21, no. 9, pp. 1911–1917, Jan. 2000.

[186] S. N. Goward, B. Markham, D. G. Dye, W. Dulaney, and J. Yang, "Normalized difference vegetation index measurements from the advanced very high resolution radiometer," *Remote Sens. Environ.*, vol. 35, no. 2, pp. 257–277, Feb. 1991.

[187] B. N. Holben, "Characteristics of maximum-value composite images from temporal AVHRR data," *Int. J. Remote Sens.*, vol. 7, no. 11, pp. 1417–1434, Nov. 1986.

[188] G. R. Arce, "Median and Weighted Median Smoothers," in *Nonlinear Signal Processing*, John Wiley & Sons, Inc., 2004, pp. 80–138.

[189] N. Viovy, O. Arino, and A. S. Belward, "The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series," *Int. J. Remote Sens.*, vol. 13, no. 8, pp. 1585–1590, May 1992.

[190] J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh, "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter," *Remote Sens. Environ.*, vol. 91, no. 3–4, pp. 332–344, Jun. 2004.

[191] A. Van Dijk, S. L. Callis, C. M. Sakamoto, and W. L. Decker, "Smoothing vegetation index profiles: an alternative method for reducing radiometric disturbance in NOAA/AVHRR data," *Photogramm. Eng. Remote Sens. USA*, 1987.

[192] M. Berterretche, A. Hudak, W. Cohen, T. Maiersperger, S. Gower, and J. Dungan, "Comparison of regression and geostatistical methods for mapping Leaf Area Index (LAI) with Landsat ETM data over a boreal forest," *USDA For. Serv. UNL Fac. Publ.*, Jan. 2005.

[193] P. Goovaerts, *Geostatistics for Natural Resources Evaluation*. Oxford University Press, 1997.

[194] G. Wang, D. Garcia, Y. Liu, R. de Jeu, and A. Johannes Dolman, "A three-dimensional gap filling method for large geophysical datasets: Application to global satellite soil moisture observations," *Environ. Model. Softw.*, vol. 30, pp. 139–142, Apr. 2012.

[195] H. Fang, S. Liang, J. R. Townshend, and R. E. Dickinson, "Spatially and temporally continuous LAI data sets based on an integrated filtering method: Examples from North America," *Remote Sens. Environ.*, vol. 112, no. 1, pp. 75–93, Jan. 2008.

[196] F. Gao *et al.*, "An algorithm to produce temporally and spatially continuous MODIS-LAI time series," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 60–64, 2008.

[197] B. Martínez and M. A. Gilabert, "Vegetation dynamics from NDVI time series analysis using the wavelet transform," *Remote Sens. Environ.*, vol. 113, no. 9, pp. 1823–1842, Sep. 2009.

[198] P. Jonsson and L. Eklundh, "Seasonality extraction by function fitting to time-series of satellite sensor data," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1824–1832, Aug. 2002.

[199] A. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures.," *Anal Chem*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964.

[200] X. Zhang *et al.*, "Monitoring vegetation phenology using MODIS," *Remote Sens. Environ.*, vol. 84, no. 3, pp. 471–475, Mar. 2003.

[201] J. N. Hird and G. J. McDermid, "Noise reduction of NDVI time series: An empirical comparison of selected techniques," *Remote Sens. Environ.*, vol. 113, no. 1, pp. 248–258, Jan. 2009.

[202] M. Menenti, S. Azzali, W. Verhoef, and R. van Swol, "Mapping agroecological zones and time lag in vegetation growth by means of fourier analysis of time series of NDVI images," *Adv. Space Res.*, vol. 13, no. 5, pp. 233–237, May 1993.

[203] L. Olsson and L. Eklundh, "Fourier Series for analysis of temporal sequences of satellite sensor imagery," *Int. J. Remote Sens.*, vol. 15, no. 18, pp. 3735–3741, Dec. 1994.

[204] P. S. A. Beck, C. Atzberger, K. A. Høgda, B. Johansen, and A. K. Skidmore, "Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI," *Remote Sens. Environ.*, vol. 100, no. 3, pp. 321–334, Feb. 2006.

[205] B. Jiang, S. Liang, J. Wang, and Z. Xiao, "Modeling MODIS LAI time series using three statistical methods," *Remote Sens. Environ.*, vol. 114, no. 7, pp. 1432–1444, Jul. 2010.

[206] S. Kandasamy, F. Baret, A. Verger, P. Neveux, and M. Weiss, "A comparison of methods for smoothing and gap filling time series of remote sensing observations – application to MODIS LAI products," *Biogeosciences*, vol. 10, no. 6, pp. 4055–4071, Jun. 2013.

[207] F. Gao, "Integrating Landsat with MODIS Products for Vegetation Monitoring," in *Satellite-based Applications on Climate Change*, J. Qu, A. Powell, and M. V. K. Sivakumar, Eds. Springer Netherlands, 2013, pp. 247–261.

[208] F. Gao *et al.*, "Fusing Landsat and MODIS Data for Vegetation Monitoring," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 47–60, Sep. 2015.

[209] M. C. Hansen, D. P. Roy, E. Lindquist, B. Adusei, C. O. Justice, and A. Altstatt, "A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin," *Remote Sens. Environ.*, vol. 112, no. 5, pp. 2495–2513, May 2008.

[210] C. Hüttich, M. Herold, B. J. Strohbach, and S. Dech, "Integrating in-situ, Landsat, and MODIS data for mapping in Southern African savannas: experiences of LCCS-based land-cover mapping in the Kalahari in Namibia," *Environ. Monit. Assess.*, vol. 176, no. 1–4, pp. 531–547, May 2011.

[211] X. Zhu, E. H. Helmer, F. Gao, D. Liu, J. Chen, and M. A. Lefsky, "A flexible spatiotemporal method for fusing satellite images with different resolutions," *Remote Sens. Environ.*, vol. 172, pp. 165–177, Jan. 2016.

[212] P. A. Brivio, M. Maggi, E. Binaghi, I. Gallo, and J. M. Gregoire, "Exploiting spatial and temporal information for extracting burned areas from time series of SPOT-VGT data," in *Analysis of Multi-temporal Remote Sensing Images*, World Scientific, 2002, pp. 132–139.

[213] J. R. G. Townshend, R. S. DeFries, X. Zhan, M. Hansen, C. Dimiceli, and J. M. Gregoire, "MODIS 250m and 500m time series data for change detection and continuous representation of vegetation characteristics," in *Analysis of Multi-temporal Remote Sensing Images*, World Scientific, 2002, pp. 233–240.

[214] S. Joyce, "An application of Kalman filtering for monitoring forest growth aided by satellite image time series," in *Analysis of Multi-temporal Remote Sensing Images*, World Scientific, 2002, pp. 371–378.

[215] R. E. Kalman and others, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960.

[216] M. Hansen, J. C. W. Chan, J. Pagis, R. S. DeFries, and D. Luo, "Long term change detection using continuous fields of tree cover from 8km AVHRR data for the years 1982-2000," in *Analysis of Multi-temporal Remote Sensing Images*, World Scientific, 2002, pp. 363–370.

[217] M. Hall-Beyer, "Comparison of single year and multi year NDVI time series Principal Components over a grassland-forest ecotone," in *Analysis of Multi-temporal Remote Sensing Images*, World Scientific, 2002, pp. 355–362.

[218] D. J. Hayes and S. A. Sader, "Comparison of change-detection techniques for monitoring tropical forest clearing and vegetation regrowth in a time series," *Photogram. Eng. Rem. Sens.*, vol. 67, pp. 1067–1075, 2001.

[219] E. H. Wilson and S. A. Sader, "Detection of forest harvest type using multiple dates of Landsat TM imagery," *Remote Sens. Environ.*, vol. 80, no. 3, pp. 385–396, Jun. 2002.

[220] R. F. Engle and C. W. Granger, "Co-integration and error correction: representation, estimation, and testing," *Econom. J. Econom. Soc.*, pp. 251–276, 1987.

[221] S. Johansen, *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press on Demand, 1995.

[222] D. O. Fuller, "Trends in NDVI time series and their relation to rangeland and crop production in Senegal, 1987-1993," *Int. J. Remote Sens.*, vol. 19, no. 10, pp. 2013–2018, 1998.

[223] R. R. Nemani *et al.*, "Climate-driven increases in global terrestrial net primary production from 1982 to 1999," *science*, vol. 300, no. 5625, pp. 1560–1563, 2003.

[224] O. Rigina and M. S. Rasmussen, "Using trend line and principal component analysis to study vegetation changes in Senegal 1986—1999 from AVHRR NDVI 8 km data," *Geogr. Tidsskr.-Dan. J. Geogr.*, vol. 103, no. 1, pp. 31–42, Jan. 2003.

[225] J. P. Jenkins, B. H. Braswell, S. E. Frolking, and J. D. Aber, "Detecting and predicting spatial and interannual patterns of temperate forest springtime phenology in the eastern U.S.," *Geophys Res Lett*, vol. 29, no. 24, p. 2201, Dec. 2002.

[226] A. Botta, N. Viovy, P. Ciais, P. Friedlingstein, and P. Monfray, "A global prognostic scheme of leaf onset using satellite data," *Glob. Change Biol.*, vol. 6, no. 7, pp. 709–725, 2000.

[227] F. Petitjean, C. Kurtz, N. Passat, and P. Gançarski, "Spatio-temporal reasoning for the classification of satellite image time series," *Pattern Recognit. Lett.*, vol. 33, no. 13, pp. 1805–1815, Oct. 2012.

[228] C. Vaduva, T. Costachioiu, C. Patrascu, I. Gavat, V. Lazarescu, and M. Datcu, "A Latent Analysis of Earth Surface Dynamic Evolution Using Change Map Time Series," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2105–2118, Apr. 2013.

[229] C. O. Justice, J. R. G. Townshend, B. N. Holben, and et C. Tucker, "Analysis of the phenology of global vegetation using meteorological satellite data," *Int. J. Remote Sens.*, vol. 6, no. 8, pp. 1271–1318, 1985.

[230] A. Kawabata, K. Ichii, and Y. Yamaguchi, "Global monitoring of interannual changes in vegetation activities using NDVI and its relationships to temperature and precipitation," *Int. J. Remote Sens.*, vol. 22, no. 7, pp. 1377–1382, Jan. 2001.

[231] R. B. Myneni, C. D. Keeling, C. J. Tucker, G. Asrar, and R. R. Nemani, "Increased plant growth in the northern high latitudes from 1981 to 1991," *Nature*, vol. 386, no. 6626, p. 698, 1997.

[232] R. Suzuki, T. Nomaki, and T. Yasunari, "West–east contrast of phenology and climate in northern Asia revealed using a remotely sensed vegetation index," *Int J Biometeorol*, vol. 47, no. 3, pp. 126–138, May 2003.

[233] N. Delbart, T. Le Toan, L. Kergoat, and V. Fedotova, "Remote sensing of spring phenology in boreal regions: A free of snow-effect method using NOAA-AVHRR and SPOT-VGT data (1982–2004)," *Remote Sens. Environ.*, vol. 101, no. 1, pp. 52–62, Mar. 2006.

[234] L. Telesca and R. Lasaponara, "Quantifying intra-annual persistent behaviour in SPOT-VEGETATION NDVI data for Mediterranean ecosystems of southern Italy," *Remote Sens. Environ.*, vol. 101, no. 1, pp. 95–103, 2006.

[235] J. I. Fisher, J. F. Mustard, and M. A. Vadeboncoeur, "Green leaf phenology at Landsat resolution: Scaling from the field to the satellite," *Remote Sens. Environ.*, vol. 100, no. 2, pp. 265–279, Jan. 2006.

[236] L. Bruzzone, C. Marin, M. Santoni, F. Bovolo, A. Trenti, and W. Beozzo, "Mappe neve del Trentino con CRIOPAT," *Neve e Valanghe*, vol. 83, pp. 4–11, Dec-2014.

[237] "Meteotrentino - CRIOPAT," 2017. [Online]. Available: http://www.meteotrentino.it/neve-ghiacci/criopat/criopat.aspx?ID=237. [Accessed: 15-Mar-2017].

[238] B. J. Choudhury, N. U. Ahmed, S. B. Idso, R. J. Reginato, and C. S. T. Daughtry, "Relations between evaporation coefficients and vegetation indices studied by model simulations," *Remote Sens. Environ.*, vol. 50, no. 1, pp. 1–17, Oct. 1994.

[239] A. R. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote Sens. Environ.*, vol. 25, no. 3, pp. 295–309, Aug. 1988.

[240] C. Sun, Z. Wu, Z. Lv, N. Yao, and J. Wei, "Quantifying different types of urban growth and the change dynamic in Guangzhou using multi-temporal remote sensing data," *Int. J. Appl. Earth Obs. Geoinformation*, vol. 21, pp. 409–417, Apr. 2013.

[241] M. T. Dokulil *et al.*, "Twenty years of spatially coherent deepwater warming in lakes across Europe related to the North Atlantic Oscillation," *Limnol. Oceanogr.*, vol. 51, no. 6, pp. 2787–2793, Nov. 2006.

[242] Z.-L. Li *et al.*, "Satellite-derived land surface temperature: Current status and perspectives," *Remote Sens. Environ.*, vol. 131, pp. 14–37, Apr. 2013.

[243] S. Pareeth *et al.*, "Warming trends of perialpine lakes from homogenised time series of historical satellite and in-situ data," *Sci. Total Environ.*, vol. 578, pp. 417–426, Feb. 2017.

[244] S. Pareeth, N. Salmaso, R. Adrian, and M. Neteler, "Homogenised daily lake surface water temperature data generated from multiple satellite sensors: A long-term case study of a large sub-Alpine lake," *Sci. Rep.*, vol. 6, p. 31251, Aug. 2016.

[245] S. Pareeth *et al.*, "New Automated Method to Develop Geometrically Corrected Time Series of Brightness Temperatures from Historical AVHRR LAC Data," *Remote Sens.*, vol. 8, no. 3, p. 169, Feb. 2016.

[246] T. Yamamoto, H. Hanaizumi, and S. Chino, "A change detection method for remotely sensed multispectral and multitemporal images using 3-D segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 5, pp. 976–985, 2001.

[247] J. R. Schott, C. Salvaggio, and W. J. Volchok, "Radiometric scene normalization using pseudoinvariant features," *Remote Sens. Environ.*, vol. 26, no. 1, pp. 1–16, Oct. 1988.

[248] F. G. Hall, D. E. Strebel, J. E. Nickeson, and S. J. Goetz, "Radiometric rectification: Toward a common radiometric response among multidate, multisensor images," *Remote Sens. Environ.*, vol. 35, no. 1, pp. 11–27, Jan. 1991.

[249] C. D. Elvidge, D. Yuan, R. D. Werackoon, and R. S. Lunetta, "Relative Radiometric Normalization of Landsat Multispectral Scanner (MSS) Data Using an Automated Scattergram Controlled Regression," *Photogramm. Eng. Remote Sens.*, vol. 61, no. 10, pp. 1255–1260, Oct. 1995.

[250] X. J. Yang and C. P. Lo, "Relative radiometric normalization performance for change detection from multi-date satellite images," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 8, pp. 967–980, Aug. 2000.

[251] M. J. Canty, A. A. Nielsen, and M. Schmidt, "Automatic radiometric normalization of multitemporal satellite imagery," *Remote Sens. Environ.*, vol. 91, no. 3, pp. 441–451, Jun. 2004.

[252] M. J. Canty and A. A. Nielsen, "Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted MAD transformation," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 1025–1036, Mar. 2008.

[253] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. L. Rojo-Alvarez, and M. Martinez-Ramon, "Kernel-Based Framework for Multitemporal and Multisource Remote Sensing Data Classification and Change Detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1822–1835, Jun. 2008.

[254] D. Lu, G. Li, and E. Moran, "Current situation and needs of change detection techniques," *Int. J. Image Data Fusion*, vol. 5, no. 1, pp. 13–38, Jan. 2014.

[255] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens.*, vol. 28, no. 5, pp. 823–870, Mar. 2007.

[256] M. Hussain, D. Chen, A. Cheng, H. Wei, and D. Stanley, "Change detection from remotely sensed images: From pixel-based to object-based approaches," *ISPRS J. Photogramm. Remote Sens.*, vol. 80, pp. 91–106, Jun. 2013.

[257] F. Bovolo and L. Bruzzone, "The Time Variable in Data Fusion: A Change Detection Perspective," *IEEE Geosci. Remote Sens. Mag.*, vol. 3, no. 3, pp. 8–26, Sep. 2015.

[258] J. Chen, P. Gong, C. He, R. Pu, and P. Shi, "Land-use/land-cover change detection using improved change-vector analysis," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 4, pp. 369–379, Apr. 2003.

[259] D. Lu, P. Mausel, E. Brondízio, and E. Moran, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2401, Jun. 2004.

[260] Y. Bayarjargal, A. Karnieli, M. Bayasgalan, S. Khudulmur, C. Gandush, and C. J. Tucker, "A comparative study of NOAA–AVHRR derived drought indices using change vector analysis," *Remote Sens. Environ.*, vol. 105, no. 1, pp. 9–22, Nov. 2006.

[261] R. Lorena, J. R. Dos Santos, Y. E. Shimabukuro, I. F. Brown, and H. Johann, "A Change Vector Analysis Technique to Monitor Land Use/land Cover in Sw Brazilian Amazon: Acre State - Semantic Scholar," in *Integrating Remote Sensing at the Global, Regional and Local Scale*, 2002, pp. 8–15.

[262] K. Kuzera, J. Rogan, and J. R. Eastman, "Monitoring Vegetation Regeneration and Deforestation Using Change Vector Analysis: Mt. St. Helens Study Area," in *ASPRS 2005 Annual Conference*, Baltimore, USA, 2005.

[263] P. Lu, A. Stumpf, N. Kerle, and N. Casagli, "Object-Oriented Change Detection for Landslide Rapid Mapping," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 701–705, Jul. 2011.

[264] M. Molinier, J. Laaksonen, and T. Hame, "Detecting Man-Made Structures and Changes in Satellite Imagery With a Content-Based Information Retrieval System Built on Self-Organizing Maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 861–874, Apr. 2007.

[265] N. Longbotham, C. Chaapel, L. Bleiler, C. Padwick, W. J. Emery, and F. Pacifici, "Very High Resolution Multiangle Urban Classification Analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1155–1170, Apr. 2012.

[266] G. Matasci, N. Longbotham, F. Pacifici, M. Kanevski, and D. Tuia, "Understanding angular effects in VHR imagery and their significance for urban land-cover model portability: A study of two multi-angle in-track image sequences," *ISPRS J. Photogramm. Remote Sens.*, vol. 107, pp. 99–111, Sep. 2015.

[267] R. Duca and F. D. Frate, "Hyperspectral and Multiangle CHRIS-PROBA Images for the Generation of Land Cover Maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2857–2866, Oct. 2008.

[268] M. J. Barnsley, J. J. Settle, M. A. Cutter, D. R. Lobb, and F. Teston, "The PROBA/CHRIS mission: a low-cost smallsat for hyperspectral multiangle observations of the Earth surface and atmosphere," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1512–1520, Jul. 2004.

[269] Y. Zhang and G. Hong, "An IHS and wavelet integrated approach to improve pan-sharpening visual quality of natural colour IKONOS and QuickBird images," *Inf. Fusion*, vol. 6, no. 3, pp. 225–234, Sep. 2005.

[270] W. A. Malila, "Change vector analysis: an approach for detecting forest changes with Landsat [Remote sensing; Idaho]," presented at the LARS Symposium on Machine Processing of Remotely Sensed Data, 1980, pp. 326–335.

[271] R. J. Kauth and G. S. Thomas, "The tasselled cap - A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat," in *The Symposium on Machine Processing of Remotely Sensed Data*, West Lafayette, Indiana, 1976, p. 4B–41–4B–50.

[272] A. Agapiou, D. D. Alexakis, A. Sarris, and D. G. Hadjimitsis, "Orthogonal Equations of Multi-Spectral Satellite Imagery for the Identification of Un-Excavated Archaeological Sites," *Remote Sens.*, vol. 5, no. 12, pp. 6560–6586, Dec. 2013.

[273] D. C. Zanotta, L. Bruzzone, F. Bovolo, and Y. E. Shimabukuro, "An Adaptive Semisupervised Approach to the Detection of User-Defined Recurrent Changes in Image Time Series," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3707–3719, Jul. 2015.

[274] C. Persello, "Advanced Techniques for the Classification of Very High Resolution and Hyperspectral Remote Sensing Images," PhD, University of Trento, Trento, 2010.

[275] "DigitalGlobe Foundation." [Online]. Available: http://foundation.digitalglobe.com/. [Accessed: 15-Jan-2018].

[276] F. Pacifici, "Atmospheric compensation in satellite imagery," US9396528 B2, 19-Jul-2016.

[277] "DigitalGlobe Atmospheric Compensation." [Online]. Available: http://explore.digitalglobe.com/AComp.html?utm_source=blog&utm_medium=website&utm_campaign=AComp. [Accessed: 15-Jan-2018].

[278] C. Paris and L. Bruzzone, "A Three-Dimensional Model-Based Approach to the Estimation of the Tree Top Height by Fusing Low-Density LiDAR Data and Very High Resolution Optical Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 467–480, Jan. 2015.

[279] "ENVI - The Leading Geospatial Analytics Software | Harris Geospatial." [Online]. Available: http://www.harrisgeospatial.com/SoftwareTechnology/ENVI.aspx. [Accessed: 15-Jan-2018].

[280] V. J. D. Tsai, "A comparative study on shadow compensation of color aerial images in invariant color models," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1661–1671, Jun. 2006.

[281] L. D. Yarbrough, G. Easson, and J. S. Kuszmaul, "QuickBird 2 Tasseled Cap transform coefficients: a comparison of derivation methods," presented at the PECORA 16 "Global Priorities in Land Remote Sensing," Sioux Falls, SD, USA, 2005.

[282] L. D. Yarbrough, K. Navulur, and R. Ravi, "Presentation of the Kauth–Thomas transform for WorldView-2 reflectance data," *Remote Sens. Lett.*, vol. 5, no. 2, pp. 131–138, Feb. 2014.

[283] J. Inglada and A. Giros, "On the possibility of automatic multisensor image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2104–2120, Oct. 2004.

[284] A. Wong and D. A. Clausi, "ARRSI: Automatic Registration of Remote-Sensing Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1483–1493, May 2007.

[285] Y. Han, J. Choi, Y. Byun, and Y. Kim, "Parameter Optimization for the Extraction of Matching Points Between High-Resolution Multisensor Images in Urban Areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5612–5621, Sep. 2014.

[286] Y. J. Kaufman, "Atmospheric effect on spectral signature-measurements and corrections," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 4, pp. 441–450, Jul. 1988.

[287] V. CASELLES and M. J. L. GARCÍA, "An alternative simple approach to estimate atmospheric correction in multitemporal studies," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 1127–1134, Jun. 1989.

[288] Y. Du, P. M. Teillet, and J. Cihlar, "Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection," *Remote Sens. Environ.*, vol. 82, no. 1, pp. 123–134, Sep. 2002.

[289] A. Baraldi, "Impact of Radiometric Calibration and Specifications of Spaceborne Optical Imaging Sensors on the Development of Operational Automatic Remote Sensing Image Understanding Systems," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 2, no. 2, pp. 104–134, Jun. 2009.

[290] G. Schaepman-Strub, M. E. Schaepman, T. H. Painter, S. Dangel, and J. V. Martonchik, "Reflectance quantities in optical remote sensing—definitions and case studies," *Remote Sens. Environ.*, vol. 103, no. 1, pp. 27–42, Jul. 2006.

[291] J. E. CONEL, "Determination of surface reflectance and estimates of atmospheric optical depth and single scattering albedo from Landsat Thematic Mapper data," *Int. J. Remote Sens.*, vol. 11, no. 5, pp. 783–828, May 1990.

[292] P. R. Coppin and M. E. Bauer, "Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 918–927, Jul. 1994.

[293] G. S. Burns and A. T. Joyce, "Evaluation of Land Cover Change Detection Techniques Using Landsat Mss Data," in *Proceeding of the 7th Pecora Symposium*, Sioux Falls, SD, USA, 1981, pp. 252–260.

[294] A. Sinch, *Change detection in the tropical forest environment of northeastern India using Landsat. In: "Remote Sensing and Tropical Land Management*. London: John Wiley and Sons, 1986.

[295] J. Heo and T. W. FitzHugh, "A standardized radiometric normalization method for change detection using remotely sensed imagery," *Photogramm. Eng. Remote Sens.*, vol. 66, no. 2, pp. 173–181, Feb. 2000.

[296] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[297] V. Sadeghi, H. Ebadi, and F. Farnood-Ahmadi, "A new model for automatic normalization of multitemporal satellite images using Artificial Neural Network and mathematical methods," *Appl. Math. Model.*, vol. 37, no. 9, pp. 6437–6445, May 2013.

[298] D. R. Insua and P. Müller, "Feedforward Neural Networks for Nonparametric Regression," in *Practical Nonparametric and Semiparametric Bayesian Statistics*, Springer, New York, NY, 1998, pp. 181–193.

[299] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, "Support Vector Regression Machines," in *Advances in Neural Information Processing Systems 9*, M. C. Mozer, M. I. Jordan, and T. Petsche, Eds. MIT Press, 1997, pp. 155–161.

[300] J. R. Jensen, "Urban/Suburban Land Use Analysis," in *Manual of Remote Sensing*, 2nd ed., Falls Church, VA: American Society of Photogrammetry, 1983, pp. 1571–1666.

[301] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

[302] S. E. Grigorescu, N. Petkov, and P. Kruizinga, "Comparison of texture features based on Gabor filters," *IEEE Trans. Image Process.*, vol. 11, no. 10, pp. 1160–1167, Oct. 2002.

[303] S. Kullback and R. A. Leiber, "On Information and Sufficiency," *Ann. Math. Stat.*, vol. 22, no. 1, pp. 79–86, Mar. 1951.

[304] T. Leichtle, C. Geiß, M. Wurm, T. Lakes, and H. Taubenböck, "Unsupervised change detection in VHR remote sensing imagery – an object-based clustering approach in a dynamic urban environment," *Int. J. Appl. Earth Obs. Geoinformation*, vol. 54, no. Supplement C, pp. 15–27, Feb. 2017.

[305] L. Rokach and O. Maimon, "Clustering Methods," in *Data Mining and Knowledge Discovery Handbook*, Springer, Boston, MA, 2005, pp. 321–352.

[306] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A Density-based Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon, 1996, pp. 226–231.

[307] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "OPTICS: Ordering Points to Identify the Clustering Structure," in *Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data*, New York, NY, USA, 1999, pp. 49–60.

[308] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 47–60, Nov. 1996.

[309] M. Zanetti and L. Bruzzone, "A Theoretical Framework for Change Detection Based on a Compound Multiclass Statistical Model of the Difference Image," *IEEE Trans. Geosci. Remote Sens.*, vol. PP, no. 99, pp. 1–15, 2017.

[310] J. Pinter Paul J. *et al.*, "Remote Sensing for Crop Management," *Photogramm. Eng. Remote Sens.*, vol. 69, no. 6, pp. 647–664, Jun. 2003.

[311] J. F. Hermance, R. W. Jacob, B. A. Bradley, and J. F. Mustard, "Extracting Phenological Signals From Multiyear AVHRR NDVI Time Series: Framework for Applying High-Order Annual Splines With Roughness Damping," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3264–3276, Oct. 2007.

[312] H. Carrao, P. Gonçalves, and M. Caetano, "A Nonlinear Harmonic Model for Fitting Satellite Image Time Series: Analysis and Prediction of Land Cover Dynamics," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1919–1930, Apr. 2010.

[313] A. Palacios-Orueta *et al.*, "Derivation of phenological metrics by function fitting to time-series of Spectral Shape Indexes AS1 and AS2: Mapping cotton phenological stages using MODIS time series," *Remote Sens. Environ.*, vol. 126, pp. 148–159, Nov. 2012.

[314] S. Estel, T. Kuemmerle, C. Levers, M. Baumann, and P. Hostert, "Mapping cropland-use intensity across Europe using MODIS NDVI time series," *Environ. Res. Lett.*, vol. 11, no. 2, p. 024015, Feb. 2016.

[315] L. Li, M. A. Friedl, Q. Xin, J. Gray, Y. Pan, and S. Frolking, "Mapping Crop Cycles in China Using MODIS-EVI Time Series," *Remote Sens.*, vol. 6, no. 3, pp. 2473–2493, Mar. 2014.

[316] Q. Xin, P. Olofsson, Z. Zhu, B. Tan, and C. E. Woodcock, "Toward near real-time monitoring of forest disturbance by fusion of MODIS and Landsat data," *Remote Sens. Environ.*, vol. 135, no. Supplement C, pp. 234–247, Aug. 2013.

[317] L. Yan and D. P. Roy, "Conterminous United States crop field size quantification from multi-temporal Landsat data," *Remote Sens. Environ.*, vol. 172, pp. 67–86, Jan. 2016.

[318] L. Yan and D. P. Roy, "Automated crop field extraction from multi-temporal Web Enabled Landsat Data," *Remote Sens. Environ.*, vol. 144, pp. 42–64, Mar. 2014.

[319] C. J. Weissteiner, C. García-Feced, and M. L. Paracchini, "A new view on EU agricultural landscapes: Quantifying patchiness to assess farmland heterogeneity," *Ecol. Indic.*, vol. 61, pp. 317–327, Feb. 2016.

[320] J. Li and D. P. Roy, "A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring," *Remote Sens.*, vol. 9, no. 9, p. 902, Aug. 2017.

[321] D. P. Roy, J. Li, H. K. Zhang, L. Yan, H. Huang, and Z. Li, "Examination of Sentinel-2A multi-spectral instrument (MSI) reflectance anisotropy and the suitability of a general method to normalize MSI reflectance to nadir BRDF adjusted reflectance," *Remote Sens. Environ.*, vol. 199, pp. 25–38, Sep. 2017.

[322] D. P. Roy, J. Li, H. K. Zhang, and L. Yan, "Best practices for the reprojection and resampling of Sentinel-2 Multi Spectral Instrument Level 1C data," *Remote Sens. Lett.*, vol. 7, no. 11, pp. 1023–1032, Nov. 2016.

[323] G. Duveiller and P. Defourny, "A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing," *Remote Sens. Environ.*, vol. 114, no. 11, pp. 2637–2650, Nov. 2010.

[324] M. Ozdogan, "The spatial distribution of crop types from MODIS data: Temporal unmixing using Independent Component Analysis," *Remote Sens. Environ.*, vol. 114, no. 6, pp. 1190–1204, Jun. 2010.

[325] A. Rydberg and G. Borgefors, "Integrated method for boundary delineation of agricultural fields in multispectral satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2514–2520, Nov. 2001.

[326] Z. Zhu and C. E. Woodcock, "Continuous change detection and classification of land cover using all available Landsat data," *Remote Sens. Environ.*, vol. 144, no. Supplement C, pp. 152–171, Mar. 2014.

[327] S. Moulin, A. Bondeau, and R. Delecolle, "Combining agricultural crop models and satellite observations: From field to regional scales," *Int. J. Remote Sens.*, vol. 19, no. 6, pp. 1021–1036, Jan. 1998.

[328] H. Carrão, P. Gonçalves, and M. Caetano, "Contribution of multispectral and multitemporal information from MODIS images to land cover classification," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 986–997, Mar. 2008.

[329] ESA, "ESA - Sen2Cor Configuration and User Manual," Dec-2017. [Online]. Available: http://step.esa.int/main/third-party-plugins-2/sen2cor/. [Accessed: 02-Jan-2018].

[330] Q. Wang, W. Shi, Z. Li, and P. M. Atkinson, "Fusion of Sentinel-2 images," *Remote Sens. Environ.*, vol. 187, pp. 241–252, Dec. 2016.

[331] E. L. Pencue-Fierro, Y. T. Solano-Correa, J. C. Corrales-Muñoz, and A. Figueroa-Casas, "A Semi-Supervised Hybrid Approach for Multitemporal Multi-Region Multisensor Landsat Data Classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 9, no. 12, pp. 5424–5435, Dec. 2016.

[332] L. Yan, D. P. Roy, H. Zhang, J. Li, and H. Huang, "An Automated Approach for Sub-Pixel Registration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) Imagery," *Remote Sens.*, vol. 8, no. 6, p. 520, Jun. 2016.

[333] M. J. Hill, "Vegetation index suites as indicators of vegetation state in grassland and savanna: An analysis with simulated SENTINEL 2 data for a North American transect," *Remote Sens. Environ.*, vol. 137, pp. 94–111, Oct. 2013.

[334] Y. Gu, J. F. Brown, J. P. Verdin, and B. Wardlow, "A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States," *Geophys. Res. Lett.*, vol. 34, no. 6, p. L06407, Mar. 2007.

[335] D. Ziou and S. Tabbone, "Edge Detection Techniques - An Overview," *Int. J. Pattern Recognit. Image Anal.*, vol. 8, pp. 537–559, 1998.

[336] L. He, X. Ren, Q. Gao, X. Zhao, B. Yao, and Y. Chao, "The connected-component labeling problem: A review of state-of-the-art algorithms," *Pattern Recognit.*, vol. 70, pp. 25–43, Oct. 2017.

[337] P. Jönsson and L. Eklundh, "TIMESAT—a program for analyzing time-series of satellite sensor data," *Comput. Geosci.*, vol. 30, no. 8, pp. 833–845, Oct. 2004.

[338] J. Delegido, J. Verrelst, L. Alonso, and J. Moreno, "Evaluation of Sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content," *Sensors*, vol. 11, no. 7, pp. 7063–7081, 2011.

[339] A. Anyamba and J. R. Eastman, "Interannual variability of NDVI over Africa and its relation to El Niño/Southern Oscillation," *Int. J. Remote Sens.*, vol. 17, no. 13, pp. 2533–2548, Sep. 1996.

[340] S. Azzali and M. Menenti, "Mapping vegetation-soil-climate complexes in southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data," *Int. J. Remote Sens.*, vol. 21, no. 5, pp. 973–996, Jan. 2000.