Land consumption estimation using Landsat satellite data and change detection techniques in the Google Earth Engine cloud environment

Abstract
Land consumption is the permanent or reversible transformation of non-artificial into artificial land cover, that leads to urban sprawl and densification, as well as the loss of the soil resource. The impervious cover of the land surface causes land degradation, as it contributes to increasing environmental and territorial risks. Mapping and monitoring the urban growth is therefore essential for sustainable development. Earth observation remote sensing is a powerful way to provide continuous information on land cover changes. Among various satellite datasets, the multi-decadal archive of Landsat images is particularly suitable for many urban change detection studies. The aim of this thesis is to extract land consumption information from open Landsat satellite imagery by applying different algorithms, also with some innovative elements, in the Google Earth Engine (GEE) cloud environment. The first change detection method implied the development and application of a novel index-based image classification approach, involving multiple decision rules, from Landsat 8 data. The “post-classification comparison” between binary “Urban/Non-urban” maps, followed by the “image differencing” between multi-date land surface albedo maps, was successfully tested for the study area of Bitritto and subsequently applied to the Bari territory (2015-2023), with very good results. The second change detection method consisted in the implementation of the Continuous Change Detection and Classification algorithm in GEE, involving a stack of multitemporal cloud-free Landsat images to identify temporal spectral behavior, pixel-by-pixel. The multitemporal “Urban/Non-urban” maps, produced using per-pixel model coefficients as input variables for Random Forest classifications, subsequently underwent “post-classification comparison” and allowed to retrieve urban growth information, with very satisfactory results, for all the capitals of Italian regions (2006-2023).
Both methods slightly underestimated the changes that really occurred, due to the “mixed pixels” issue and the filtering techniques, while GEE allowed relatively rapid processing of remote sensing data.

Cover: the land consumption map for the city of Bari (2006-2023)

DICATECh

2025

Carlo Barletta

Land consumption estimation using Landsat satellite data and change detection techniques in the Google Earth Engine cloud environment

01

D.R.S.A.T.E.

Doctorate in Risk And Environmental, Territorial And Building Development

Coordinator: Prof. Vito Iacobellis

XXXVII CYCLE
ICAR/06 – Topography and Cartography

DICATECh
Department of Civil, Environmental, Building Engineering and Chemistry

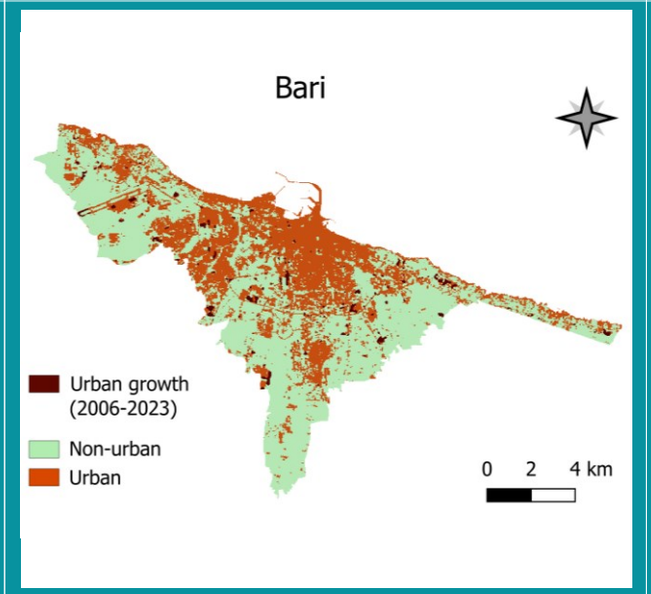
POLITECNICO DI BARI

01
2025

Carlo Barletta

Land consumption estimation using Landsat satellite data and change detection techniques in the Google Earth Engine cloud environment

Prof. Eufemia Tarantino
DICATECh
Politecnico di Bari



Politecnico di Bari

LIBERATORIA PER L'ARCHIVIAZIONE DELLA TESI DI DOTTORATO

Al Magnifico Rettore
del Politecnico di Bari

Il sottoscritto BARLETTA CARLO nato a BARI (BA) il 17/04/1989 residente a BARI (BA) in via AMENDOLA 199/C e-mail carlobarletta4@gmail.com iscritto al 3° anno di Corso di Dottorato di Ricerca in RISCHIO E SVILUPPO AMBIENTALE, TERRITORIALE ED EDILIZIO ciclo 37° ed essendo stato ammesso a sostenere l'esame finale con la prevista discussione della tesi dal titolo:

"Land consumption estimation using Landsat satellite data and change detection techniques in the Google Earth Engine cloud environment"

DICHIARA

1) di essere consapevole che, ai sensi del D.P.R. n. 445 del 28.12.2000, le dichiarazioni mendaci, la falsità negli atti e l'uso di atti falsi sono puniti ai sensi del codice penale e delle Leggi speciali in materia, e che nel caso ricorressero dette ipotesi, decade fin dall'inizio e senza necessità di nessuna formalità dai benefici conseguenti al provvedimento emanato sulla base di tali dichiarazioni;
2) di essere iscritto al Corso di Dottorato di ricerca RISCHIO E SVILUPPO AMBIENTALE, TERRITORIALE ED EDILIZIO ciclo 37°, corso attivato ai sensi del "*Regolamento dei Corsi di Dottorato di ricerca del Politecnico di Bari"*, emanato con D.R. n.286 del 01.07.2013;
3) di essere pienamente a conoscenza delle disposizioni contenute nel predetto Regolamento in merito alla procedura di deposito, pubblicazione e autoarchiviazione della tesi di dottorato nell'Archivio Istituzionale ad accesso aperto alla letteratura scientifica;
4) di essere consapevole che attraverso l'autoarchiviazione delle tesi nell'Archivio Istituzionale ad accesso aperto alla letteratura scientifica del Politecnico di Bari (IRIS-POLIBA), l'Ateneo archivierà e renderà consultabile in rete (nel rispetto della Policy di Ateneo di cui al D.R. 642 del 13.11.2015) il testo completo della tesi di dottorato, fatta salva la possibilità di sottoscrizione di apposite licenze per le relative condizioni di utilizzo (di cui al sito http://www.creativecommons.it/Licenze), e fatte salve, altresì, le eventuali esigenze di "embargo", legate a strette considerazioni sulla tutelabilità e sfruttamento industriale/commerciale dei contenuti della tesi, da rappresentarsi mediante compilazione e sottoscrizione del modulo in calce (Richiesta di embargo);
5) che la tesi da depositare in IRIS-POLIBA, in formato digitale (PDF/A) sarà del tutto identica a quelle **consegnate**/inviate/da inviarsi ai componenti della commissione per l'esame finale e a qualsiasi altra copia depositata presso gli Uffici del Politecnico di Bari in forma cartacea o digitale, ovvero a quella da discutere in sede di esame finale, a quella da depositare, a cura dell'Ateneo, presso le Biblioteche Nazionali Centrali di Roma e Firenze e presso tutti gli Uffici competenti per legge al momento del deposito stesso, e che di conseguenza va esclusa qualsiasi responsabilità del Politecnico di Bari per quanto riguarda eventuali errori, imprecisioni o omissioni nei contenuti della tesi;
6) che il contenuto e l'organizzazione della tesi è opera originale realizzata dal sottoscritto e non compromette in alcun modo i diritti di terzi, ivi compresi quelli relativi alla sicurezza dei dati personali; che pertanto il Politecnico di Bari ed i suoi funzionari sono in ogni caso esenti da responsabilità di qualsivoglia natura: civile, amministrativa e penale e saranno dal sottoscritto tenuti indenni da qualsiasi richiesta o rivendicazione da parte di terzi;
7) che il contenuto della tesi non infrange in alcun modo il diritto d'Autore né gli obblighi connessi alla salvaguardia di diritti morali od economici di altri autori o di altri aventi diritto, sia per testi, immagini, foto, tabelle, o altre parti di cui la tesi è composta.

Luogo e data Bari, 12/12/2024 Firma Carlo Barletta

Il/La sottoscritto, con l'autoarchiviazione della propria tesi di dottorato nell'Archivio Istituzionale ad accesso aperto del Politecnico di Bari (POLIBA-IRIS), pur mantenendo su di essa tutti i diritti d'autore, morali ed economici, ai sensi della normativa vigente (Legge 633/1941 e ss.mm.ii.),

CONCEDE

- al Politecnico di Bari il permesso di trasferire l'opera su qualsiasi supporto e di convertirla in qualsiasi formato al fine di una corretta conservazione nel tempo. Il Politecnico di Bari garantisce che non verrà effettuata alcuna modifica al contenuto e alla struttura dell'opera.

- al Politecnico di Bari la possibilità di riprodurre l'opera in più di una copia per fini di sicurezza, back-up e conservazione.

Luogo e data Bari, 12/12/2024 Firma Carlo Barletta

D.R.S.A.T.E.

POLITECNICO DI BARI

01

2025

Doctorate in Risk And Environmental, Territorial And Building Development

Coordinator: Prof. Vito Iacobellis

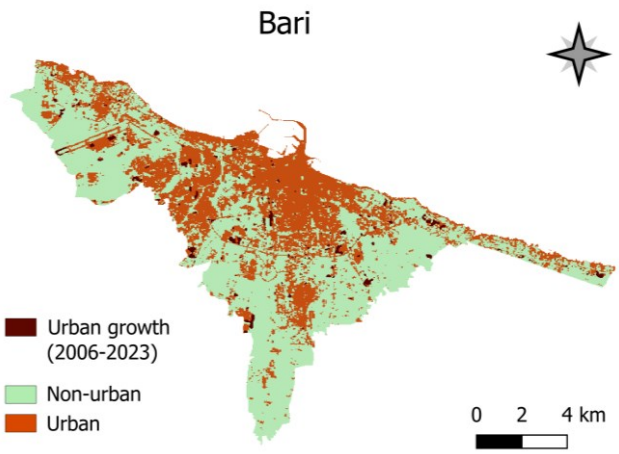
XXXVII CYCLE
ICAR/06 – Topography and Cartography

DICATECh
Department of Civil, Environmental, Building Engineering and Chemistry

Carlo Barletta

Land consumption estimation using Landsat satellite data and change detection techniques in the Google Earth Engine cloud environment

Prof. Eufemia Tarantino
DICATECh
Politecnico di Bari



D.R.S.A.T.E.

POLITECNICO DI BARI

01

2025

Dottorato in Rischio E Sviluppo Ambientale, Territoriale Ed Edilizio

Coordinatore: Prof. Vito Iacobellis

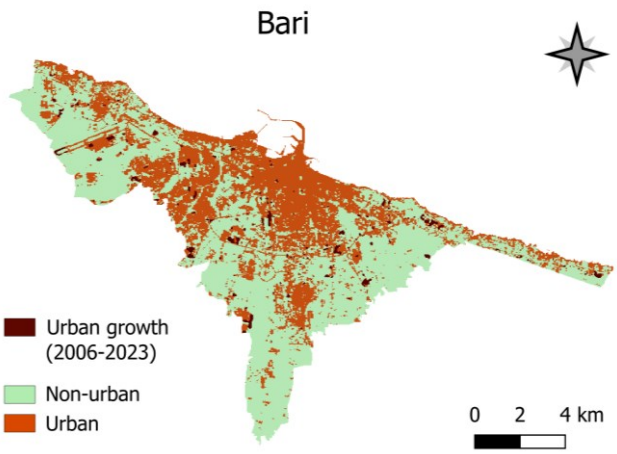
XXXVII CICLO
ICAR/06 – Topografia e Cartografia

DICATECh
Dipartimento di Ingegneria Civile, Ambientale, Edile e di Chimica

Carlo Barletta

Stima del consumo di suolo mediante utilizzo di dati satellitari Landsat e tecniche di change detection in ambiente cloud Google Earth Engine

Prof.ssa Eufemia Tarantino
DICATECh
Politecnico di Bari



EXTENDED ABSTRACT *(eng)*

Land consumption is the permanent or reversible transformation of non-artificial into artificial land cover, linked to settlements and infrastructural dynamics, that leads to urban sprawl and densification, as well as the loss of the soil resource. The permanent covering of the soil with impervious materials (defined as “soil sealing”) is a major cause of land degradation, as it compromises soil ecosystem services and contributes to increasing environmental and territorial risks.

Mapping and monitoring of urban growth aims at identifying the amount and location of land cover changes, in order to develop and analyze different strategies for sustainable development. Earth observation remote sensing is a powerful way to provide continuous information on land cover changes at different geographical scales of analysis, offering unique advantages in terms of spatio-temporal coverage and cost. Among various satellite datasets, the multi-decadal archive of medium-resolution Landsat images is considered particularly suitable for many urban change detection studies.

Several techniques have been developed and implemented over the years to detect changes occurred on the Earth’s surface from remotely sensed satellite data. Nevertheless, the opening of the Landsat archive in 2008 allowed the formulation of new approaches (such as the “Continuous Change Detection and Classification” algorithm) that use a stack of all available images over a given study area.

However, such approaches require huge amount of data storage and computations. To overcome this issue, cloud-based platforms (e.g., Google Earth Engine)

are actually very used for research purposes since they allow access to multiple image processing algorithms and numerous public satellite datasets.

The objective of this thesis is to extract land consumption information from open Landsat satellite imagery by applying different algorithms, also with some innovative elements, in the Google Earth Engine cloud environment, where appropriate JavaScript codes were written and implemented.

The first change detection method used implied the development and application of a novel index-based image classification method, involving multiple decision rules, to retrieve multitemporal "Urban/Non-urban" maps from Landsat 8 data. These maps underwent a "post-classification comparison", pixel-by-pixel, followed by an "image differencing" between multi-date land surface albedo maps and a neighborhood filtering technique. This approach, successfully tested for the study area of Bitritto (a town proximal to the municipality of Bari, the capital city of Apulia Region in southern Italy), was subsequently applied to obtain urban growth map for the Bari territory, from 2015 to 2023. The results were satisfactory since high accuracies were found for the change maps for both Bitritto and Bari case studies. Further analysis consisted in quantifying the absolute, relative and annual average land loss (i.e., urban growth metrics) for the Bari area.

The second change detection method consisted in the implementation of the Continuous Change Detection and Classification algorithm, involving a stack of multitemporal cloud-free Landsat images to identify temporal spectral behavior on a pixel-by-pixel basis, obtained by comparing satellite observations with observation-fitted model predictions. This approach allowed to extract multitemporal "Urban/Non-urban" maps using per-pixel model coefficients as input variables for Random Forest classifications. The maps underwent "post-classification comparison" and a mode filtering technique; then, urban growth maps and metrics were retrieved, with very satisfactory result, for Bari (2015-2023) and for all the twenty capitals of Italian regions (2006-2023). Further analysis, concerning the study area of Bari, consisted in overlapping the urban

growth map (2006-2023) with the map of hydraulically hazardous areas (delimitated by the Basin Authority for the hydro-geological asset plan of the Apulia Region) in order to verify the effectiveness of land planning in avoiding urban sprawl and densification in protected areas and the consequent increase of hydro-geological risk.

Both the change detection approaches slightly underestimated the changes that really occurred, due to the 30 m geometric resolution of the Landsat images (causing the “mixed pixels issue, which limited the ability of the methods to detect small changes), and the filtering techniques.

Lastly, the Google Earth Engine platform, free-to-use for research purposes, allowed relatively rapid processing of remote sensing data, particularly with respect to the Continuous Change Detection and Classification method, which requires high computational resources.

Keywords: *Remote sensing, open satellite data, urban growth, land cover change, soil sealing, cloud computing.*

EXTENDED ABSTRACT *(ita)*

Il consumo di suolo rappresenta la trasformazione, permanente o reversibile, della copertura del suolo da non artificiale ad artificiale, legata a dinamiche insediative ed infrastrutturali, che comporta l'espansione e la densificazione urbana, così come la perdita della risorsa suolo. La copertura permanente del suolo con materiali impermeabili (definita "soil sealing") è una delle cause principali di degradazione del suolo, poiché ne compromette i servizi ecosistemici e contribuisce ad incrementare fenomeni di rischio ambientale e territoriale. La mappatura e il monitoraggio della crescita urbana mira ad individuare la quantità e la localizzazione dei cambiamenti di copertura del suolo, al fine di sviluppare ed analizzare differenti strategie per lo sviluppo sostenibile. Il telerilevamento per l'osservazione della Terra è un potente strumento per fornire continue informazioni sui cambiamenti di copertura del suolo a differenti scale geografiche di analisi, offrendo vantaggi unici in termini di copertura spazio-temporale e costi. Tra vari datasets satellitari, l'archivio multi-decennale di immagini Landsat a media risoluzione è considerato particolarmente adatto per molti studi di change detection a scala urbana.

Diverse tecniche sono state sviluppate e implementate negli anni per individuare i cambiamenti avvenuti sulla superficie terrestre da immagini satellitari telerilevate. Tuttavia, l'apertura dell'archivio Landsat nel 2008 ha permesso lo sviluppo di nuovi approcci (come l'algoritmo di "Continuous Change Detection and Classification") che usano uno stack di tutte le immagini disponibili su una data area. Tuttavia, tali approcci richiedono un'enorme quantità di archiviazione dati e di calcolo. Al fine di superare questo problema, piattaforme cloud (come Goo-

gle Earth Engine) sono oggi molto utilizzate per scopi di ricerca poiché permettono l'accesso a diversi algoritmi di processamento delle immagini e numerosi datasets satellitari pubblici.

L'obiettivo del presente lavoro di tesi è di estrarre informazioni sul consumo di suolo mediante utilizzo di immagini satellitari open Landsat applicando differenti algoritmi, anche con alcuni elementi innovativi, in ambiente cloud Google Earth Engine, dove appropriati codici JavaScript sono stati sviluppati e implementati.

Il primo metodo di change detection utilizzato ha implicato lo sviluppo e l'applicazione di un nuovo metodo di classificazione delle immagini index-based, che utilizza multiple regole decisionali al fine di ricavare mappe "Urban/Non-urban" multi-temporali da dati Landsat 8. Queste mappe sono state sottoposte a "post-classification comparison", pixel per pixel, seguito da "image differencing" tra mappe di albedo superficiale multi-temporali e da una tecnica di filtraggio di vicinanza. Questo approccio, testato con successo sull'area di studio di Bitritto (un paese prossimo al Comune di Bari, il capoluogo della Regione Puglia, nel sud Italia), è stato in seguito applicato al fine di ottenere la mappa di crescita urbana per il territorio di Bari, avvenuta tra il 2015 e il 2023. I risultati sono stati soddisfacenti poiché sono stati ottenuti alti valori di accuratezza riguardo le mappe di cambiamento su entrambi i casi studio di Bitritto e di Bari. Ulteriori analisi hanno riguardato la quantificazione del consumo di suolo in termini assoluto, relativo e annuale medio (ovvero le metriche di crescita urbana) per l'area di Bari.

Il secondo metodo di change detection ha riguardato l'implementazione dell'algoritmo di Continuous Change Detection and Classification, che usa uno stack di immagini multi-temporali prive di nuvole Landsat per individuare il comportamento spettrale nel tempo di ogni pixel, ottenuto mediante il confronto tra le osservazioni satellitari e i valori predetti di un modello fittato sulle osservazioni. Questo approccio ha permesso di estrarre mappe "Urban/Non-urban" multi-temporali utilizzando i coefficienti del modello, pixel per pixel,

come variabili di input per classificazioni Random Forest. Le mappe sono state sottoposte a “post-classification comparison” e ad una tecnica di filtraggio modale; in seguito, le mappe e le metriche di crescita urbana sono state ricavate, con risultati abbastanza soddisfacenti, su Bari (2015-2023) e su tutti i venti capoluoghi di regione italiani (2006-2023). Ulteriori analisi, riguardo l'area di studio di Bari, hanno riguardato la sovrapposizione della mappa di crescita urbana (2006-2023) con la mappa delle aree di pericolosità idraulica (delimitate dall’Autorità di Bacino per il piano di assetto idro-geologico della Regione Puglia) al fine di verificare l’efficacia della pianificazione territoriale nell’evitare l’espansione e la densificazione urbana in aree protette e il conseguente aumento del rischio idro-geologico.

Entrambi i metodi di change detection hanno leggermente sottostimato i cambiamenti realmente avvenuti, a causa della risoluzione geometrica di 30 m delle immagini Landsat (che comporta il problema dei “pixel misti”, che ha limitato la capacità dei metodi di individuare piccoli cambiamenti) e delle tecniche di filtraggio.

Infine, la piattaforma Google Earth Engine, gratuita per fini di ricerca, ha permesso una elaborazione relativamente rapida dei dati telerilevati, in particolar modo riguardo il metodo di Continuous Change Detection and Classification che richiede alte risorse computazionali.

Keywords: *Remote sensing, open satellite data, urban growth, land cover change, soil sealing, cloud computing.*

INDEX

INTRODUCTION

Urbanization is one of the main triggering factors for land use/land cover (LU/LC) changes in cities worldwide. Urban growth, i.e. the increase in artificial surfaces in urban areas as a result of the construction of new buildings, roads and other impervious[1] surfaces, has significant environmental and territorial impacts. Soil is an essential, multifunctional and vital ecosystem. It provides food, biomass and raw materials; it is a key element of the landscape and cultural heritage; and it plays a fundamental role in maintaining biodiversity, providing habitats for many living species. Soil stores, filters and transforms many substances, including water, nutrients and carbon. It is therefore highly valuable from both an environmental and socio-economic perspective. As a limited and practically non-renewable resource, soil needs to be protected and preserved for future generations. Among other causes of soil degradation (such as erosion, pollution and salinization), soil sealing, meaning the permanent covering of soil with impermeable man-made materials such as concrete and asphalt, is one of the most serious issues because of the negative impact it can generate on soil ecosystem services [1-5].

In an urban system, the natural ecosystem component is uniquely interconnected with the social, economic, cultural and political components. Indeed, all cities face today the challenge of reconciling economic activity and growth with environmental considerations in order to achieve sustainable development.

Land consumption (Figure 0.1), defined as the conversion of non-artificial cover (i.e. agricultural, natural or semi-natural cover) to an artificial cover (buildings, infrastructures, mining areas, construction sites, photovoltaic plants or other paved areas) of the soil, in a permanent or reversible manner, is a process associated with the loss of the soil resource, and is linked to settlement and infrastructural dynamics resulting from urban sprawl, densification or the conversion of land in urban context [3-4].

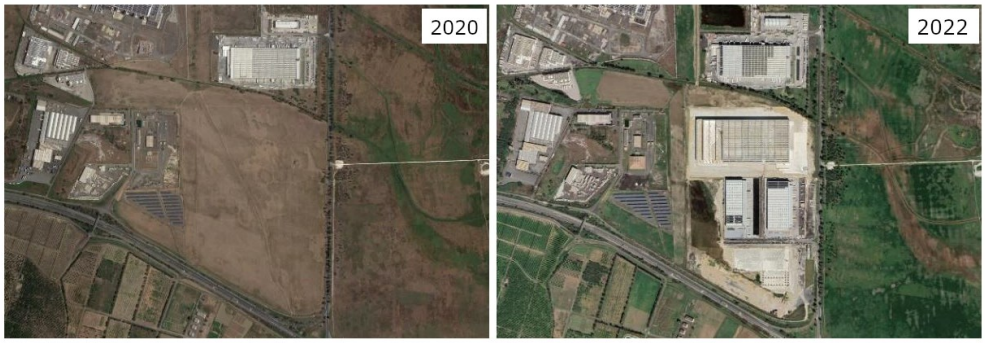


Fig. 0.1 - Example of land consumption in the city of Catania (2020-2022). Source: [4]

Several factors are driving this process, such as:

- the population growth;
- the demand for a better quality of life and standard of living;
- the demand for greener and more attractive housing outside the city center;
- the mass tourism;
- the lack of appreciation of the value of soil as a limited resource [3].

As mentioned above, the permanent covering of soil with impervious materials (soil sealing) is one of the main causes of land degradation because it compromises the ecosystem functions of soil. In particular, soil sealing has an impact on (not exhaustive list [3]):

- soil's water retention capability: sealing reduces the amount of precipitation that can be absorbed by the soil. Soil-water dynamics depend on various mechanisms that control the hydrological balance. Indeed, not all the water that reaches a basin infiltrate into the soil. According to a conceptual scheme, part of the precipitation runs off the surface and part returns to the atmosphere through evapo-transpiration [6-7]. The reduced capacity of the soil to absorb water (Figure 0.2) increases runoff and the flooding risk. Furthermore, most of the water stored in the soil is available to the vegetation; therefore, reduced water retention capac-

ity increases also the risk of drought, thereby increasing the need for irrigation and exacerbating salinization problems in agriculture;

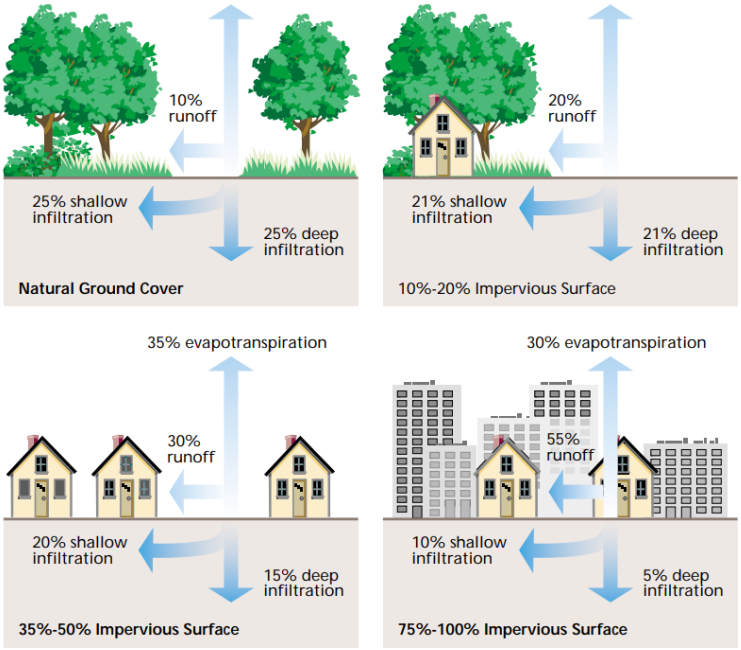


Fig. 0.2 - Relationship between impervious cover and hydrological balance components. Source: [8]

- microclimate in urban environments: sealing reduces the evapotranspiration (due to the loss of soil and vegetation) and increases the absorption of solar energy (due to asphalt and concrete) in urban areas, contributing to the urban heat island effect, which could be particularly dangerous for human health during heat waves; the presence of green spaces in a city contributes also to temperature regulation and has a positive effect on air humidity;
- carbon cycle: the soil removed during construction activities (i.e. the topsoil) contains a huge amount of organic carbon, which is partly irretrievably lost even if it is reused;

- air pollution: the loss of trees and vegetation (which can trap particles and pollutant gases, in addition to influencing wind turbulence) has also a negative impact on air quality in urban areas;
- biodiversity: sealing reduces the presence of below-ground micro-organisms (which are fundamental for breaking down organic matter in the soil, recycling nutrients and sequestering and storing carbon) and compromises the survival of many above-ground animal species, threatening their habitats;
- landscape: sealing could contribute to landscape degradation, which has a negative impact on some economic activities such as tourism [3].

Monitoring and evaluating the soil sealing process is very important to understand and tackle the problem of soil degradation. Studies on land consumption usually focus on mapping the phenomenon in a geographical context and detecting temporal changes. Detailed visualization and quantification of impervious surfaces and monitoring of urban expansion provide potentially valuable information for understanding urban patterns and dynamics, for developing new sustainable land management strategies or for evaluating the effectiveness of actual policies in addressing the adverse impacts of soil sealing. The monitoring of urban growth is mainly aimed at identifying the type, amount and location of land transformation [1, 9-11].

Ground-based data collection is generally time consuming, labour-intensive and costly. In contrast, Earth Observation (EO) remote sensing techniques are well suited to continuously acquire information on LU/LC changes at different geographical scales of analysis (from small areas to national or global scales), offering unique advantages in terms of spatio-temporal coverage and cost.

EO remote sensing is a technique that uses sensors, at a certain distance from the Earth, to measure the amount of electromagnetic energy coming from an area on the land surface, and then uses digital image processing to gather information from this data.

The main challenges that need to be addressed to obtain information on LU/LC changes from EO satellite data are:

- the geometric resolution of the sensor should be sufficient to spatially differentiate features;
- the spectral and radiometric resolutions of the sensor should allow objects to be separated thematically;
- the temporal resolution should be sufficient to provide information on changes occurred over time [1, 9-14].

Change detection is the process of identifying differences in the state of specific characteristics of an object or phenomenon by measuring, analyzing and comparing it at different times. For LU/LC change detection, the basic premise is that LU/LC variations of an area should result in a change in the energy measured by the sensor (the radiance) and that this change should be greater than the radiance change caused by other factors (such as differences in atmospheric conditions, in the solar angle or in soil moisture conditions).

Various procedures have been proposed over the years to perform LU/LC change detection analyses from single multi-temporal satellite imagery, including “image classification and post-classification comparison”, “image differencing”, “image ratioing”, “change vector analysis”, “image regression”, and so on [15-17]. Among them, “image classification and post-classification comparison”, consisting in comparing multi-date LU/LC classified maps on a pixel-by-pixel level, is a simple, effective, and widely used method but it is prone to error propagation from the independently classified maps [17].

A wide range of methods have been developed in the literature for LU/LC classification of satellite images such as supervised, unsupervised, object-based, index-based, sub-pixel, spectral mixture analysis etc. [1, 9, 18]. Each method has its pros. and cons., and the selection of the most suitable classifier depends on many factors including algorithm performance, analyst experience, and computational resources. In general, the classification of remotely sensed images is a complex process that involves choosing an appropriate classification scheme,

selecting representative training samples (from the field, or using higher resolution aerial orthophotos and satellite images), preprocessing the images, selecting a suitable classifier to extract features, possible post-classification processing, and assessing accuracy of the LU/LC maps [19].

The “Image differencing” method between multitemporal images, instead, is effective for rapid and quantitative estimation of changes, but is limited by the difficulty of interpreting different types of LU/LC transformations characterized by similar reflectance change. However, useful information can be retrieved by measuring the temporal shift of spectral indices or biophysical variables (such as the land surface albedo) [11, 17, 20-22].

Different EO satellite missions provide multi-temporal data that are adapt to estimate the amount of artificial surfaces and detect LU/LC and urban growth at various spatial and temporal resolutions.

Nowadays, open data policies allow to freely download and use remotely sensed images of the land surface provided by missions such as Landsat (United States Geological Survey, USGS / National Aeronautics and Space Administration, NASA) and Sentinel (European Copernicus Programme). Data in open format, which are easily accessible by everyone, are high valuable for the role they can play in investigating and monitoring natural resources as well as phenomena that occurs on the land surface. Open geospatial and EO data can produce, indeed, a powerful impact on Earth science and environmental studies since they can foster, for example, the development of new methods and services based on them [23].

For urban change detection analysis, the multispectral Landsat images, that have been acquired for more than 50 years (Figure 0.3), result to be reasonably adapt for their geometric (30 m), spectral, radiometric and temporal resolution [12, 24-27].

Moreover, the opening of the multi-decadal Landsat archive in 2008 paved the way for new opportunities to use satellite imagery for change detection purposes.

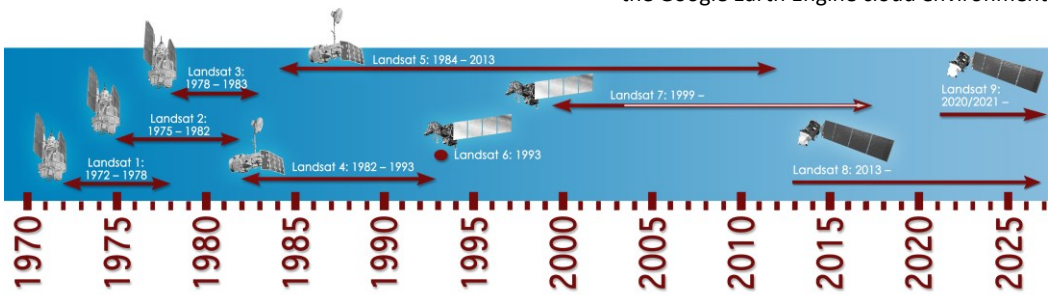


Fig. 0.3 - Landsat mission: operational life span of each Landsat satellite. Source: [28]

In recent years, new approaches that exploit all available observations, such as Continuous Change Detection and Classification (CCDC) [29], have been proposed in the literature, in addition to conventional methods that rely on single multi-temporal images. CCDC employs a per-pixel model fitting approach that captures the temporal trend of the surface reflectances for each pixel, allowing to detect changes in LU/LC when the fitted model breaks down. By reducing the reliance on single observations, which can't capture intra-annual dynamics, and by being able to obtain a LU/LC classified map for every point in time, this method has been shown to provide higher accuracy than conventional techniques. However, the CCDC approach is computationally expensive because it requires a large amount of data storage [29-31].

Today, to handle a large volume of EO data, cloud-based processing platforms (such as Google Earth Engine, GEE) are becoming increasingly popular. GEE is made up of a high-performance computing service linked to an interactive programming interface and a multi-petabyte data catalog containing a massive number of open satellite data (such as Landsat and Sentinel imagery) [27, 32-33].

The aim of this thesis work is to estimate land consumption from Landsat satellite data by developing and implementing different algorithms, also including some innovative elements. For this purpose, two change detection techniques have been applied. The first method (described in Chapter 3), developed in this research, consists of an integrated approach involving "image classification and

post-classification comparison" (including a novel decision-tree and index-based classification method for Landsat 8 data) and "image differencing" (between multitemporal land surface albedo maps) change detection techniques. The second method (explained in Chapter 4) consists of the implementation of the CCDC method [29] (which makes use of a cube of multitemporal and cloud-free Landsat images).

The analyses were carried out in the GEE cloud environment, where two appropriate JavaScript codes were written.

The thesis is organized as follows: in Chapter 1 is briefly presented the GEE cloud platform and its main functionalities for EO data processing; Chapter 2 explores the land surface albedo variable and its role for environmental and territorial analyses; in Chapter 3 is described the method used to detect land consumption for the areas of Bitritto and Bari (2015-2023), in southern Italy, using the novel index-based image classification method, and the "post-classification comparison and albedo change analysis" technique; Chapter 4, instead, illustrates the CCDC algorithm, which was implemented in GEE (using all available Landsat 4, 5, 7, 8 and 9 images) to estimate the land consumption for the area of Bari (2015-2023) and for all the twenty Italian regional capital cities (2006-2023); lastly, in Chapter 5 is reported the discussion of the results, followed by the conclusions of the study.

1. GOOGLE EARTH ENGINE CLOUD PLATFORM FOR EARTH OBSERVATION REMOTE SENSING DATA ANALYSIS

Google Earth Engine is a cloud computing platform launched by Google in 2010 that enables parallelized processing of EO and geospatial data using Google's computational infrastructure. Given the massive volume of this kind of data ("geo-big data"), acquired from remote sensing sensors, the cloud technology is today considered the best choice for their storage, processing, analysis and visualization. The GEE high-performance computational service is connected to a continuously updated multi-petabyte catalog (the "Earth Engine data catalog", https://developers.google.com/earth-engine/datasets/) of free and open raw and preprocessed data (such as Landsat and Sentinel images, as well as digital elevation models, land cover products, climate, meteorological and geophysical datasets) [33-35].

GEE can be accessed through an internet-based Application Programming Interface (API) with an associated web-based Interactive Development Environment (IDE), which supports fast prototyping and visualization of results. The available APIs are JavaScript, Python and REST. In particular, the Code Editor (https://code.earthengine.google.com/) is the IDE designed for writing, developing and running even complex scripts written in JavaScript programming language. It contains the following five elements (Figure 1.1):

- the search bar (for finding datasets in the GEE data catalog);
- the left panel, including:
  - the "Scripts" tab (for scripts management);
  - the "Docs" tab (with API documentation);
  - the "Asset" tab (for asset management);
- the center panel, including:
  - the JavaScript code editor (for writing the script);

  - the “Get Link”, “Save”, “Run”, “Reset”, “Apps” and “Settings” buttons;
- the right panel, including:
  - the “Inspector” tab (for interactive map queries);
  - the “Console” tab (for displaying printed text, objects and charts);
  - the “Tasks” tab (to handle long-running tasks);
- the “map” panel (for visualizing generated maps) [36].

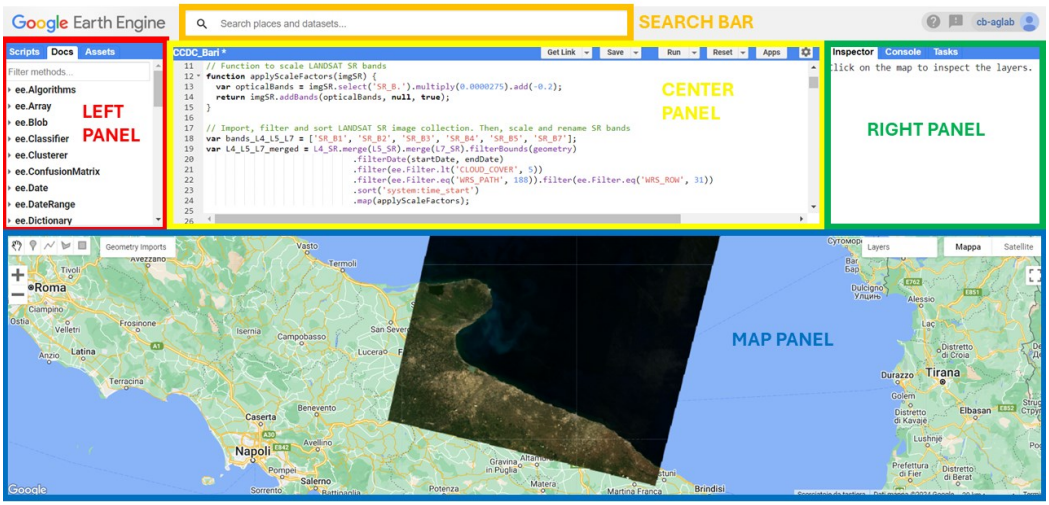


Fig. 1.1 – The Earth Engine Code Editor

Moreover, the API consists of Earth Engine objects representing data types such as rasters, vectors, strings and numbers. Each of these objects belongs to a specific class (with its own set of available functions). The most common classes of objects in GEE are:

- image (a raster data type);
- image collection (a set of images);
- feature (a vector data type);
- feature collection (a set of features);
- reducer (an object which compute statistics or perform aggregations);
- join (an object which combines datasets);

- array (an object for multidimensional analyses);
- chart (an object for charting properties) [37].

The EE objects and their functions can be combined in different ways to create customizable scripts. The code, written by the user in the client library, is converted into a JSON request object, which is then sent to the Google's servers. On the server, the data is loaded, the computation is performed, and the result is sent back to the client. This process allows very large computations to be prototyped without relying on the computing capabilities of the user's workstation.
The Code Editor, which runs in the user's browser, contains a library of hundreds of built-in algorithms (Table 1.1), ranging from simple mathematical functions to more complex machine learning algorithms, to perform basic or advanced remote sensing image processing computations. GEE users don't need to download a GEE dataset to process it or install other software to execute tasks. The GEE architecture, in fact, allows its users to query directly from the GEE library.
Multiple operations can be performed on a single or a batch of images. For example, machine learning algorithms (such as Random Forest, Support Vector Machine and Classification and Regression Tree, CART) can be applied for supervised image classification [32-35, 38]. Multi-temporal analysis of image stacks can be performed, instead, using functions such as CCDC, Landsat-based detection of Trends (LandTrendr) or Exponentially Weighted Moving Average Change Detection (EWMACD) [34].
Moreover, users can implement many approaches for image processing and analysis. Code development in GEE is usually done in a structured way. It starts with data acquisition, including satellite images and auxiliary data. For data selection and preparation, the user-friendly interface helps to create filtered datasets, calculate spectral indices or band combinations, and perform statistical operations, among other things [39-40]. The user can compose multiple functions in the script according to his needs, building a computational description that takes the form of a Directed Acyclic Graph (DAG), that assembles the chain

of operations and expresses the complete computation. The DAG is then sent to the servers to be evaluated.

Tab. 1.1 – Summary of common operations available in GEE. Source: [33-34]

Algorithm Package	**Operations**
Machine Learning	Supervised classification and regression, unsupervised classification, TensorFlow
Image	Visualization, masking, mosaicking, registration, clipping, resampling, reducing resolution, band manipulation, spectral operations, mathematical operations, Boolean operations, morphological operations, logical operations, edge detection, terrain operations, connected components, texture, convolution, etc.
Image Collection	Filtering, mapping, reducing, composing, mosaicking, etc.
Geometry, Feature, Feature Collections	Filtering, mapping, reducing, measurement operations, intersection, union, difference, buffer, spatial interpolation, etc.
Reducer	Image reduction, image collection reduction, statistical operations, raster to vector conversion, vector to raster conversion, zonal statistics, correlation, regression, etc.
Array/Matrix	Array construction, product, determinant, transpose, inverse, pseudoinverse, etc.
Charts	Time-series charts, histograms, image regions charts, time series in image regions, etc.

GEE is today one of the most used platforms from researchers worldwide addressing environmental and territorial issues (it is free for academic and non-commercial purposes). The main applications of research studies conducted in the GEE environment regard LU/LC mapping and monitoring, forest mapping, drought monitoring, natural hazard management, urban analysis, crop yield es-

timation, snow mapping, soil condition monitoring, flood mapping, species habitat monitoring, etc. [33].

One of the many advantages of GEE is that it stores and analyzes satellite images in the form of pyramids and tiles. Each image, sliced into tiles (tile size: 256x256) which conserve its original projection and resolution, has its pyramid at different pixel resolutions (each zoom level is created by downsampling the previous level by a factor of two) to allow rapid visualization during algorithm development. Indeed, Google Earth Engine is designed to provide rapid, interactive exploration and analysis of geospatial data, allowing users to pan and zoom the results to investigate a subset of the image at a time. To achieve this, GEE uses a lazy computation model that enables it to compute only those parts of the output that are needed to fulfill the current query [32, 34]. For example, time-series analysis can be performed in an efficient way, since a statistical aggregation can be applied temporally over an entire image stack to compute per-pixel statistics. This operation is performed using both tiling and aggregations, so that each tile is computed in parallel using the lazy computation model and, within each tile, the aggregation is performed for each pixel. Furthermore, fast filtering and sorting of image collections is provided. As a result, users can select their data of interest from millions of images on the basis of various spatial and temporal constraints.

Another advantage of GEE is that a number of pre-processing steps have already been applied to the datasets available in the GEE catalog (Figure 1.2), providing users with corrected data (e.g., the orthorectified and calibrated Top of Atmosphere, TOA, and the orthorectified and atmospherically corrected Surface Reflectance, SR, Landsat datasets) alongside the raw data [34].

Because of the benefits that were mentioned above, GEE is a powerful tool for researchers, with more functionalities than a traditional Geographic Information System (GIS) software [27].

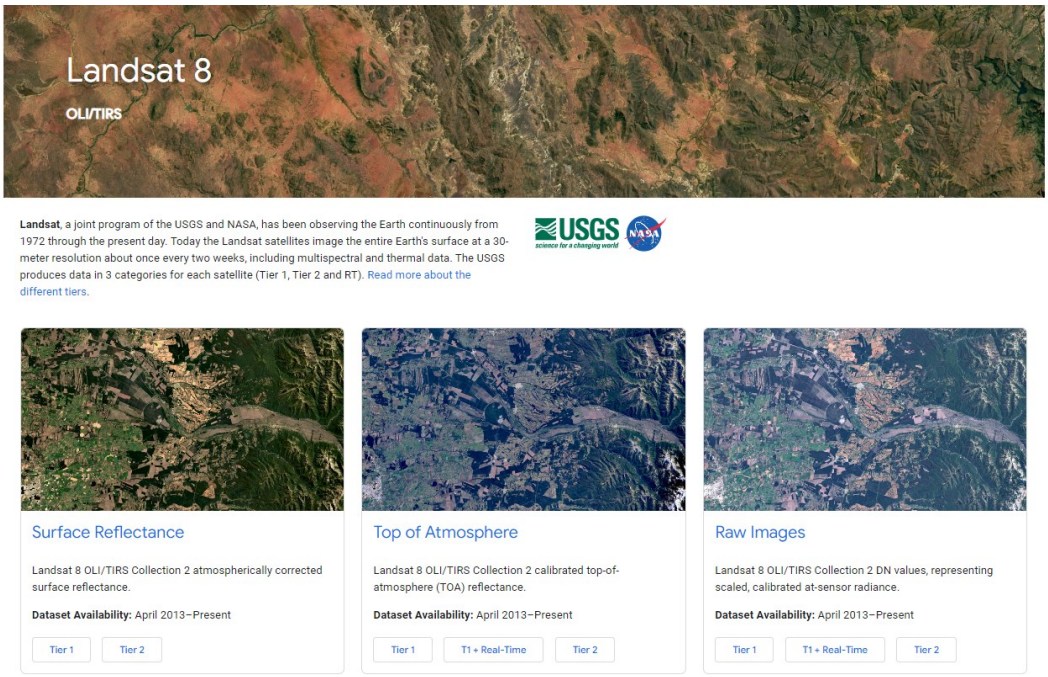


Fig. 1.2 – The Landsat 8 collection 2 preprocessed and raw data available in the GEE data catalog

However, GEE does have some limitations. For example, GEE currently has only a few image classification and regression algorithms, and the large training datasets required for some machine or deep learning methods cannot be used due to computational restrictions, necessitating implementation of these algorithms in external environments. GEE is, in fact, a shared resource, and limits for users are imposed. For example, in interactive sessions, the maximum durations of requests per user is 270 s, the maximum number of simultaneous requests is 40, and the maximum number of simultaneous executions for some operations is 25 [32].

Another limitation is that the data generated by the user are stored on the servers of a private company (Google), which is not appropriate for governments or other private companies. Moreover, if it is necessary to download the processed data for further analysis in another software environment, the user can encounter a time-consuming process due to the huge size of the map and the limitations of Internet speed [34].

2. LAND SURFACE ALBEDO TO SUPPORT TERRITORIAL ANALYSES

2.1. Overview

The land surface albedo is a key bio-geophysical parameter involved in various types of environmental, climatic and territorial studies. Defined as the ratio of the surface-reflected flux density (W/m^2) to the incoming solar irradiance, it quantifies the ability of a surface to reflect the incident solar radiation. More specifically, surface albedo values range from 0 (for the ideal case of a "black body") to 1 (for full reflectivity). Since surface albedo affects the net radiation, its measurement is helpful for studying the energy transfer between the Earth's surface and the atmosphere.

Monitoring surface albedo, which is considered an essential climate variable (ECV) by the Global Climate Observing System (GCOS) [41], is crucial for assessing the impacts of global issues such as climate change and urbanization, which are defined in the United Nations (UN) Sustainable Development Goals (SDGs) (https://sdgs.un.org/goals) to protect the Earth from global environmental threats [27, 42-44].

Albedo depends on the physical properties (i.e. roughness, density, type, darkness, impurities etc.) and geometry of the material. However, its spatial and temporal variability is influenced by atmospheric conditions and the angular distribution of the incoming solar radiation (i.e., the change in the solar zenith angle Z, during the day), as well as by other factors such as soil moisture, vegetation dynamics, snow cover and LU/LC changes. The albedo can also vary with the season: for example, a surface covered with snow in winter and with grass in spring will have a different albedo.

The surface albedo should not be confused with the planetary albedo (or TOA albedo), which is the fraction of solar radiation scattered by the Earth and constituted by both the surface albedo and the atmospheric albedo (Figure 2.1).

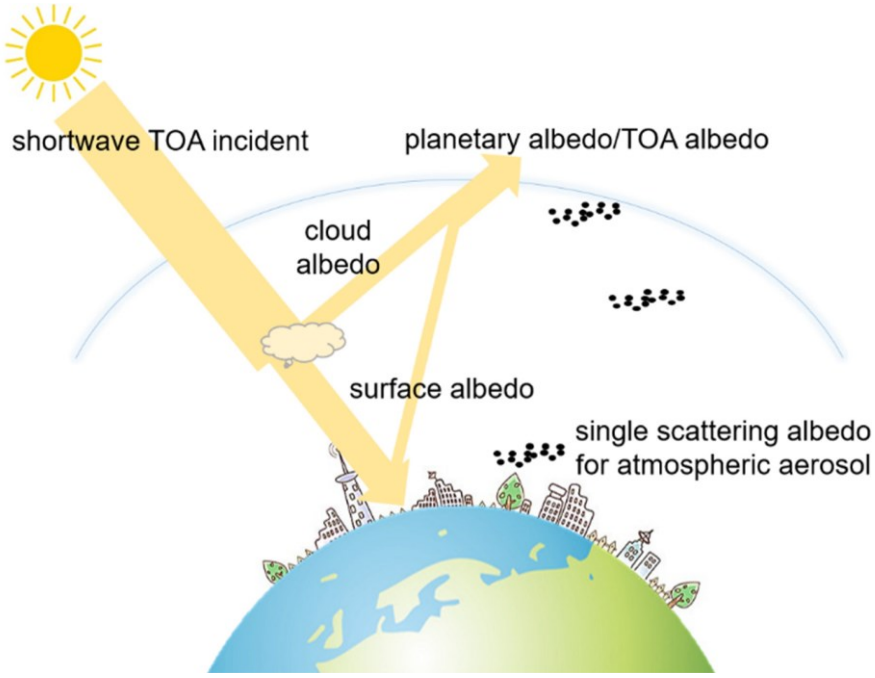


Fig. 2.1 - Various definitions of albedo. Source: [43]

Furthermore, another clarification is needed: when the albedo refers to the full spectrum (Figure 2.2) of the solar radiation (0.3-3 µm), it is called "broadband" albedo. The term "narrowband" albedo, on the other hand, refers to a narrow range of wavelengths (i.e., a specific spectral channel of a sensor) [42-46].
Regarding the phenomenon of reflection, two main types of surfaces behaviors exist (Figure 2.3):

- the specular reflection (in the case of very smooth surfaces);
- the diffuse reflection (in the case of surfaces with a certain roughness).

In the first case, the outgoing radiance is regulated by the anisotropic model, in the second case by the isotropic model. Real surfaces behave in a composite manner. The ideal case of a rough surface, used to approximate real situations,

is the Lambertian surface, where the reflection is diffuse and the radiance is constant for every observed geometry.

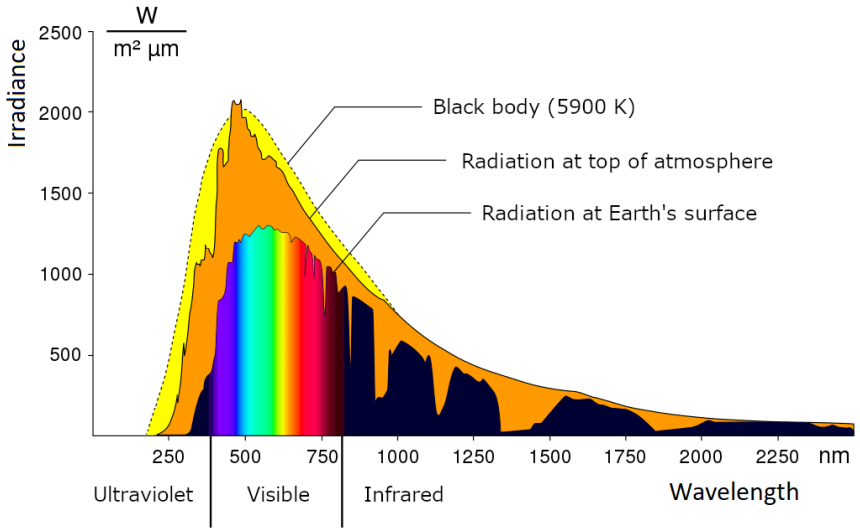


Fig. 2.2 - Theoretical black body spectrum and full solar spectrum. Source: [Sonne Strahlungsintensitaet.svg, Wikipedia, CC BY-SA 2.0 de] (modified)

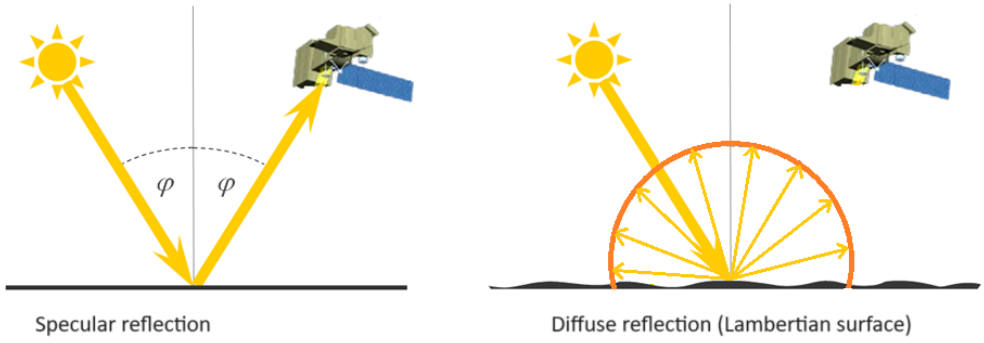


Fig. 2.3 - Different mechanisms of reflection. Source: [47] (modified)

The reflectivity of a natural body is related to both the incident geometry and the observed geometry of the radiation. The incident geometries of the Sun's radiation most commonly considered in the literature are:

- the natural irradiance;
- the ideal direct radiation;

- the ideal diffuse radiation.

There are also three main sub-classifications of the albedo, depending on the type of incident solar radiation considered:

- the “black-sky” albedo (the albedo in the absence of diffuse radiation, i.e. considering only direct radiation);
- the “white-sky” albedo (the albedo in the absence of direct radiation, i.e. considering only isotropic diffuse radiation);
- the “blue-sky” albedo (the albedo in real conditions, i.e. under natural irradiance conditions).

The knowledge of the broadband surface “blue-sky” albedo, i.e. the albedo under natural conditions of direct and diffuse solar radiation, is the objective of many literature studies. Its estimation for different land cover types is not an easy task, as it usually requires knowledge of atmospheric conditions and surface characteristics [13, 43-44, 48].

In general, the surface albedo varies with natural (soil, vegetation, water, snow and ice) and artificial (urban areas such as buildings, roads, etc.) land cover types.

Soil albedo shows different patterns according to different characteristics such as soil irregularities, moisture, darkness and colour. For example, soil moisture is associated with a reduction in the albedo of bare soil and canopy, as well as the presence of a biocrust layer over bare soil also contributes to albedo reduction. On the other hand, in crop fields, greenhouse cladding materials have been found to potentially increase the albedo of the area in which they are installed.

The albedo of the vegetation canopy, on the other hand, is related to factors such as species, structure, greenness and fractional vegetation cover.

In addition, water surfaces generally have lower albedos compared to the surrounding landscape, while snow and ice albedos typically vary with type, age, density, concentration, roughness and thickness, among other factors.

The albedo for urban surfaces depends on the complexity of urban morphology, the type and age of the materials, as well as latitude and time. Furthermore, complex albedo variations in urban areas are also due to the presence of soil-vegetation plots [43, 45].

Previous literature studies have provided characteristic albedo values for several LU/LC categories. Taha [49] reported that the albedo of many European urbanized areas ranges from 0.15 to 0.20. According to Allen et al. [50], agricultural areas have an albedo between 0.14 and 0.22, while water bodies have an albedo between 0.025 and 0.348. Brivio et al. [13] reported that the albedo of roads is between 0.10 to 0.28, the albedo of bare soil ranges from 0.05 to 0.31, while the albedo of vegetation ranges from 0.14 to 0.45 (depending on the phenological stage).

Albedo variations can be directly related to LU/LC transformations that have occurred over time in urban and/or rural environments, and are thus useful for change detection analyses. For instance, substantial increases in albedo can result from the transition from dense to sparse vegetation, from vegetation to bare soil, and from bare soil or vegetation to built-up land [51].

There are two main methods of calculating the surface albedo:

- The direct measurement;
- The indirect estimation.

The direct measurement is carried out on the ground using instruments called albedometers, which usually consist of two pyranometers (the upper one measuring the incoming radiation and the lower one measuring the reflected radiation).

However, the indirect estimation from multispectral images acquired by satellite platforms allows the analysis to be extended from local to regional or global scales. The main methods for retrieving broadband surface albedo from satellite data are:

- Narrow-to-broadband conversion algorithms;
- Bidirectional reflectance distribution function (BRDF) angular modeling;

- Direct estimation algorithms;
- Geostationary satellite data algorithms.

The first two methods are the most commonly used in the literature: the former computes the broadband albedo from a linear combination of narrowband albedos (with different weight coefficients) and is suitable when the reflectance is assumed to be isotropic (Lambertian surfaces). In this case, the conversion coefficients are calculated from ground measurements or model simulations. The latter, instead, relies on multi-angle satellite observations and describes the reflectance anisotropy of the surfaces.

This study focuses only on broadband surface albedo (also referred to in this text as “land surface albedo”, “surface albedo” or, merely, “albedo”) and narrow-to-broadband conversion methods [43-46, 52].

The implementation steps for estimating the broadband albedo using the narrow-to-broadband conversion algorithm are illustrated in Figure 2.4. The conversion coefficients are first derived using spectral reflectance data (measured in situ) and solar fluxes simulated by radiative transfer models. The coefficients are then applied to the narrowband albedos to obtain the broadband surface albedo.

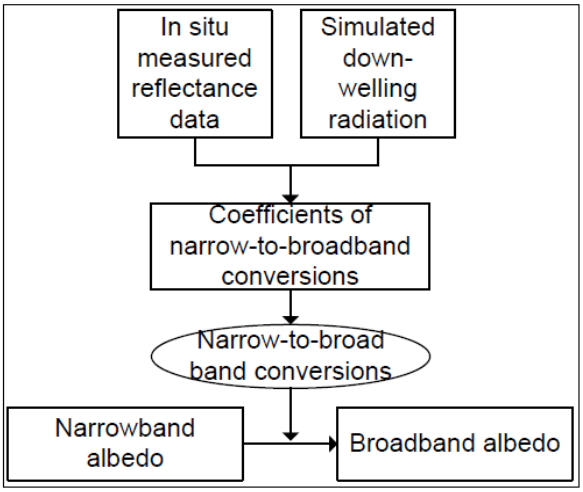


Fig. 2.4 - Flowchart of the narrow-to-broadband conversion algorithm for broadband land surface albedo estimation from satellite data. Source: [52]

EO remote sensing is an effective way to calculate and monitor surface albedo at different geometric and temporal resolutions. Since the end of the 20th century, several narrow-to-broadband algorithms have been developed in the literature for estimating land surface albedo from Landsat, Advance Very High-Resolution Radiometer (AVHRR) and Meteosat satellite data. Liang (2000) [53] improved these studies and provided simple conversion coefficients from various narrowband satellite sensors (including Landsat 5 Thematic Mapper, TM) under common atmospheric and surface conditions. The coefficients were obtained from extensive radiative transfer model simulations and a spectral reflectance library. Tasumi et al. [54] in 2008 suggested a more complex strategy to estimate the broadband albedo from Landsat 5 (TM) and 7 Enhanced Thematic Mapper+ (ETM+) data based on band-to-band atmospheric correction through a radiative transfer model. However, several studies calculate the albedo from Landsat data by first combining narrowband reflectances without atmospheric correction, and then applying the atmospheric adjustment based on the method proposed by Zhong and Li in 1998 [55] and adopted by Bastiaanssen et al. (1998) [56]. In this way, the relative contribution of each spectral band to the broadband albedo is determined, since the solar radiation spectrum is well known. This approach was followed by da Silva et al. (2016) [57], who presented the first algorithm developed specifically for Landsat 8 data. This method has been shown to produce maps with high level of detail for the various LU/LC types, as well as accurate estimates of broadband albedo, although it is more time-consuming than other easy-to-use algorithms (such as the Liang approach) due to the need to assess the atmospheric transmissivity in order to apply the atmospheric correction [27, 43-46, 52, 57].

2.2. da Silva et al. algorithm (2016) for Landsat 8 data

The da Silva et al. algorithm (2016) [57], for land surface albedo estimation using Landsat 8 data, is expressed by the following equation:

$$\alpha = \frac{(\alpha_{TOA} - \alpha_{ATM})}{\tau^2} \quad (1)$$

Where α_{TOA}, α_{ATM} and τ are the TOA albedo, the path radiance albedo and the atmospheric transmissivity, respectively. As reported by [50], α_{ATM} ranges from 0.025 and 0.04. A value of 0.03 is usually chosen, as recommended by [58]. The value of τ is, then, estimated by applying the following expression [59]:

$$\tau = 0.35 + 0.627\exp\left[-\frac{0.00146P_o}{K_t \cos Z} - 0.075\left(\frac{W}{\cos Z}\right)^{0.4}\right] \quad (2)$$

in which P_0, W, K_t and Z are the local atmospheric pressure [kPa], the precipitable water [mm], the air turbidity coefficient and the solar zenith angle [°], respectively.

The TOA albedo is estimated by linearly combining the TOA reflectances (r_i) for each Landsat 8 spectral band (from band 2 to band 7), as follows:

$$\alpha_{\mathrm{TOA}} = p_2 \cdot r_2 + p_3 \cdot r_3 + p_4 \cdot r_4 + p_5 \cdot r_5 + p_6 \cdot r_6 + p_7 \cdot r_7 \quad (3)$$

Where p_i (Table 2.1) are weighting coefficients obtained from the ratio between the solar constant of each band and the sum of the solar constants of the bands used in the albedo calculation [27, 57, 60].

Tab. 2.1 – Weighting coefficients of the da Silva et al. algorithm [57].

Weighting coefficients (p_i)	**Value**
p_2	0.300
p_3	0.277
p_4	0.233
p_5	0.143
p_6	0.036
p_7	0.012

2.3. Application of da Silva et al. algorithm (2016) for estimating albedo from Landsat 8 and 9 data

In this study, the da Silva et al. algorithm (2016) was applied to Landsat 8 and 9 images in order to evaluate new Landsat 9 sensor (launched in 2021) capabilities for measuring albedo and assure continuity in environmental monitoring through the Landsat mission.

The sensors on board Landsat 8 and Landsat 9 platforms have, in fact, comparable characteristics, such as geometric, temporal, and spectral resolutions. This responds to the need to keep satellites monitoring natural resources and processes, thereby identifying long-term changes on the Earth's surface. Nonetheless, the Landsat 9 instrument has a higher radiometric resolution (14 bit-resolution) than Landsat 8 (12 bits).

The method was assessed using Landsat 8 and 9 images from two cities in southern Italy: Palermo (Sicily), and Cagliari (Sardinia).

2.3.1 Study areas

The Mediterranean coastal cities of Palermo and Cagliari (Figure 2.5) were chosen as study areas because, for these sites, images with comparable acquisition data were identified for both the Landsat 8 and Landsat 9 missions. Additionally, these zones have comparable geographical, geomorphological, and climate characteristics, with hot and dry summers and mild winters [62-63]. Palermo is located on the northern coast of Sicily Island on a densely urbanized wide plain surrounded by mountains to the south and west. Its elevation is about 150 m above sea level (a.s.l.) [64]. Meanwhile, Cagliari is in the southern part of Sardinia Island and stretches across various hills (with an elevation of around 80–100 m a.s.l.) It is surrounded by wetlands (such as lagoons and ponds) to the west, north, and east [65].

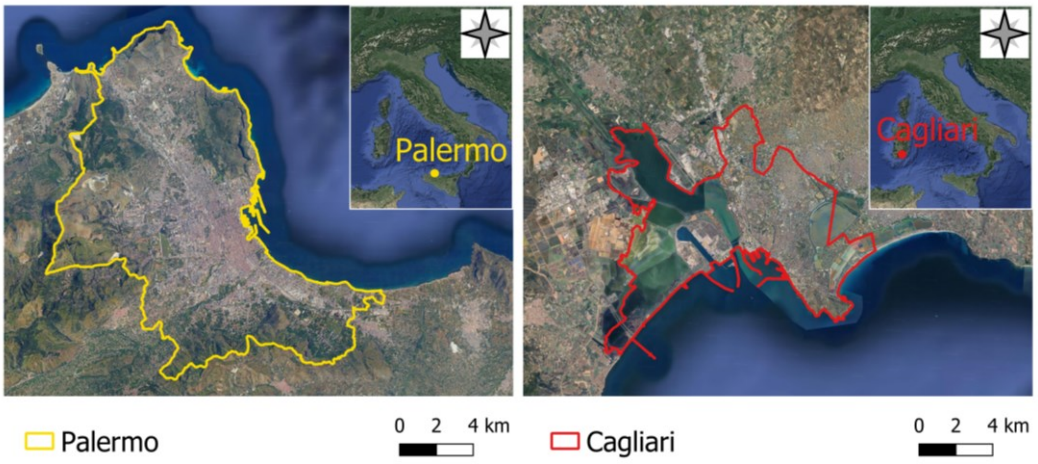


Fig. 2.5 – Study areas of Palermo and Cagliari

2.3.2 Input data and albedo estimation

The input satellite data was chosen based on two key criteria: less-than-10% cloud coverage, and the same collecting day and hour. The first criterion assured that the clouds had the least influence on the processing output, while the second provided equivalent illumination conditions and, therefore, a similar performance comparison. This last criterion was satisfied since an underfed maneuver was completed between November 11 and 17, 2021 (the Landsat 9 satellite traveled beneath the Landsat 8 spacecraft [66]). Thus, the images that were captured on November 15 and November 13, 2021, for Palermo and Cagliari, respectively, matched both requirements and were selected. Their characteristics are described in Table 2.2. Then, cloud detection and masking were performed for every image.

To estimate the albedos using the da Silva et al. method (2016), TOA images, already orthorectified and calibrated, were used.

In equation (2), P_o was obtained from measurements at 10:00 on the days of the satellite images at the Palermo and Cagliari stations of the Italian National Tidegauge network (https://mareografico.it/). W was, instead, retrieved from the "National Center for Environmental Prediction/National Center for Atmos-

pheric Research (NCEP/NCAR) Reanalysis data, Water Vapor" dataset [61] (available in the GEE data catalog), which provides W values at six-hour temporal resolutions (00:00, 6:00, 12:00, and 18:00). The average of the two closest-in-time values of W (6:00 and 12:00) was computed to select a value of W that was representative of the Landsat acquisition time (around 10:00) for each satellite image (as described in [60]). A K_t value of 1 is commonly used for pure air, while a value of 0.5 is used for polluted air [27, 59]. Because the regional authorities detected air pollution levels at both sites that were lower than the legal limits in November 2021, a value of 1 was chosen in this case. Lastly, the Sun's zenith angle (Z) was retrieved from the image metadata.

Tab. 2.2 – Selected image features for Palermo and Cagliari areas.

ID	Study area	Mission	Acquisition date [dd/mm/yyyy]	Acquisition time	Cloud cover [%]
1	Palermo	Landsat 8	15/11/2021	09:48	0.70
2	Palermo	Landsat 9	15/11/2021	09:46	0.98
3	Cagliari	Landsat 8	13/11/2021	10:00	5.89
4	Cagliari	Landsat 9	13/11/2021	10:05	1.12

2.3.3 Statistical analysis and outliers' extraction

The resultant albedo maps were sampled and statistically investigated in order to evaluate the albedo distribution and variability within the study areas using base statistics metrics, and then to investigate the Landsat 9 data potentialities in estimating the albedos using algorithms that were optimized for Landsat 8 and comparing their results to those that were obtained by processing the Landsat 8 images.

To meet the first objective, the mean (μ), standard deviations (SD), median (m), distribution trends, kurtosis (κ), and skewness (sk) were computed. For the second one, scatterplots that compared the Landsat 8 and Landsat 9 estimates were calculated, along with the correlation coefficient (ρ) and root mean square error (RMSE).

Since many outliers were present in albedo distributions, the Tukey's filter [67] was applied to detect and analyze them across both territories. According to equation (4), an observation is labeled as an outlier when its value exceeds the lower (Q_1) or upper quartiles (Q_3) [67]:

$$[Q_1 - 1.5(Q_3 - Q_1), Q_3 + 1.5(Q_3 - Q_1)] \quad (4)$$

2.3.4 Results for Palermo and Cagliari

The albedo maps (Figure 2.6), obtained in this work, highlight the variability of the albedo within the case study territories.

In Figure 2.6, the darker regions show lower values, whereas the yellow-orange areas have higher values. Table 2.3 presents the statistical metrics that were calculated for the entire territories of Palermo and Cagliari.

Based on the results of da Silva et al.'s technique, the μ value was around 0.2 in each map; moreover, the SD in Palermo and Cagliari was around 0.07 for the Landsat 8 and Landsat 9 estimations. Furthermore, ρ was quite high in both areas (reaching 0.97 for Palermo and 0.94 for Cagliari), while the RMSE values were low (totaling 0.02 for Palermo and Cagliari). Histograms were created to determine how the data was distributed (Figure 2.7).

These distributions were visually congruent with the previously reported metrics values (Table 2.3). Figure 2.8 shows the boxplots of the albedo values, which illustrate that many outliers were present in the upper section of the graph.

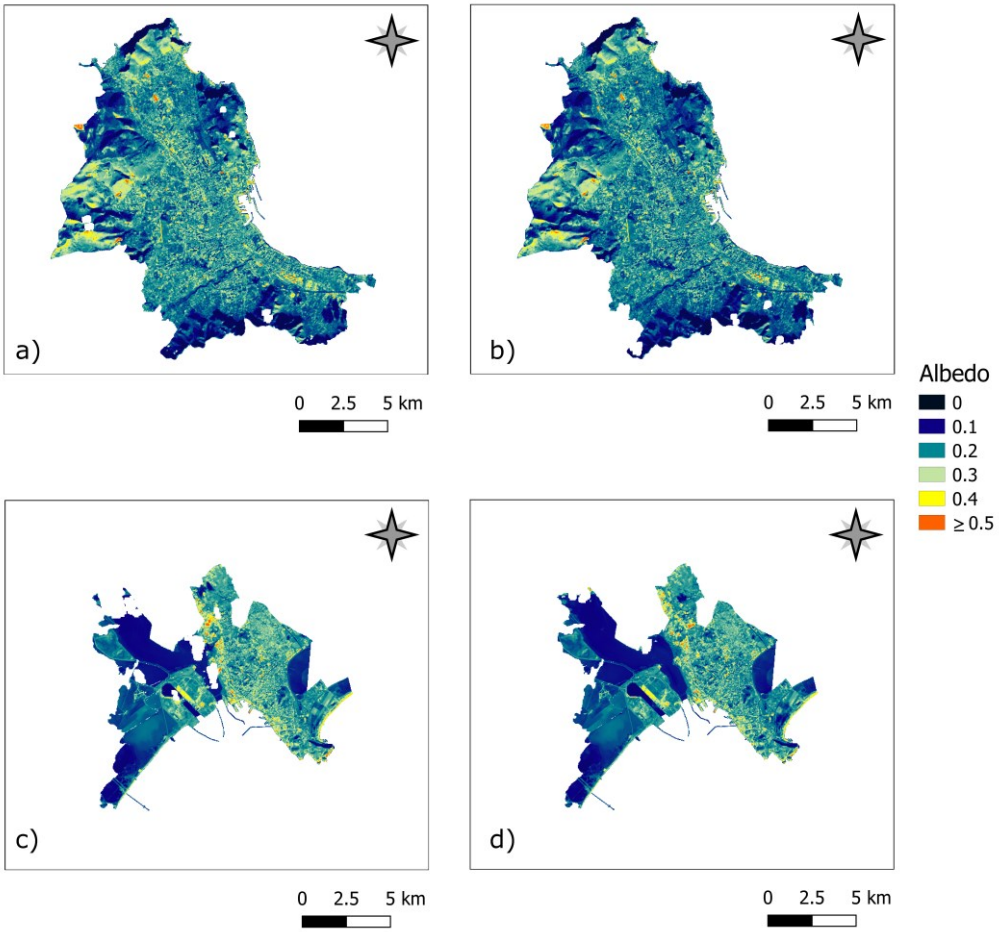


Fig. 2.6 – Albedo maps obtained from Palermo Landsat 8 (a) and Landsat 9 (b), and Cagliari Landsat 8 (c) and Landsat 9 (d) images using the da Silva et al. algorithm

Tab. 2.3 – Statistical metrics of albedo estimation for Palermo and Cagliari.

Study area	Mission	μ	ρ	SD	m	RMSE	κ	sk
Palermo	Landsat 8	0.20	0.97	0.07	0.20	0.02	8.24	0.83
	Landsat 9	0.19		0.07	0.19		8.69	0.87
Cagliari	Landsat 8	0.20	0.94	0.07	0.20	0.02	6.78	1.03
	Landsat 9	0.19		0.07	0.19		5.74	0.93

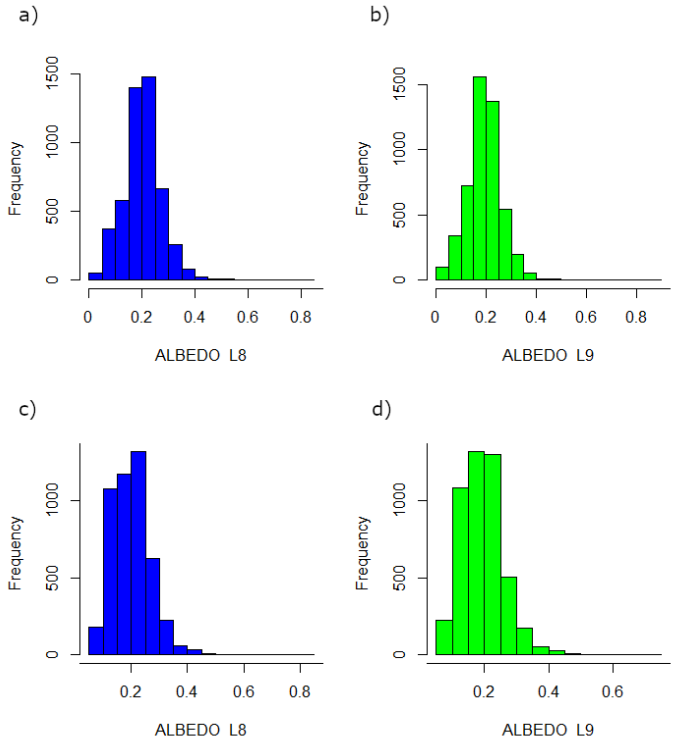


Fig. 2.7 – Histograms of albedo values generated for Palermo Landsat 8 (a) and Landsat 9 (b), and Cagliari Landsat 8 (c) and Landsat 9 (d)

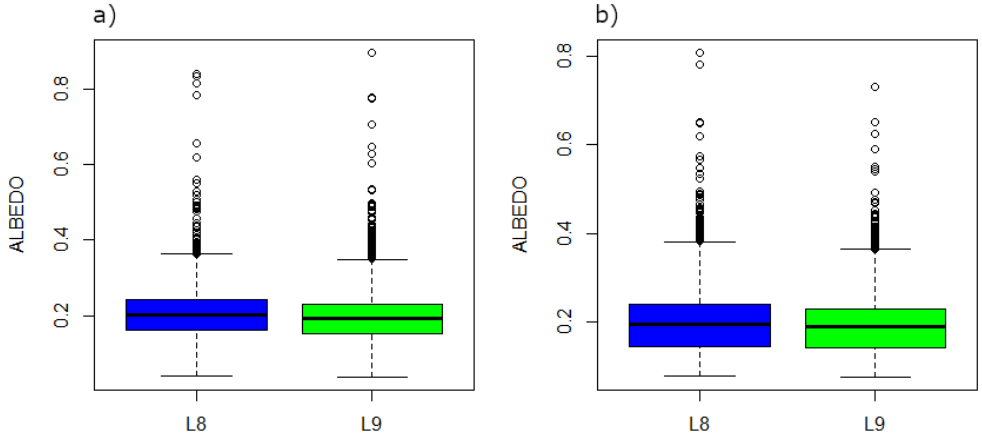


Fig. 2.8 – Boxplots of albedo values for Palermo (a) and Cagliari (b)

Moreover, the scatterplots between the Landsat 8 and Landsat 9 outcomes for both case studies are illustrated in Figure 2.9. Most of the spots in Palermo (Figure 2.9a) were scattered along the diagram's bisector line; this indicated that many of the albedo values from Landsat 8 and Landsat 9 were linearly correlated. However, there were some inaccurate values – particularly in the lower-left corner of the image. Cagliari's pattern (Figure 2.9b) was similar to Palermo's, but there were more mismatched spots across the plot. This is verified by Table 2.3, which indicates that ρ was lower in Cagliari than in Palermo and that RMSE was greater in Cagliari than in the corresponding Palermo area.

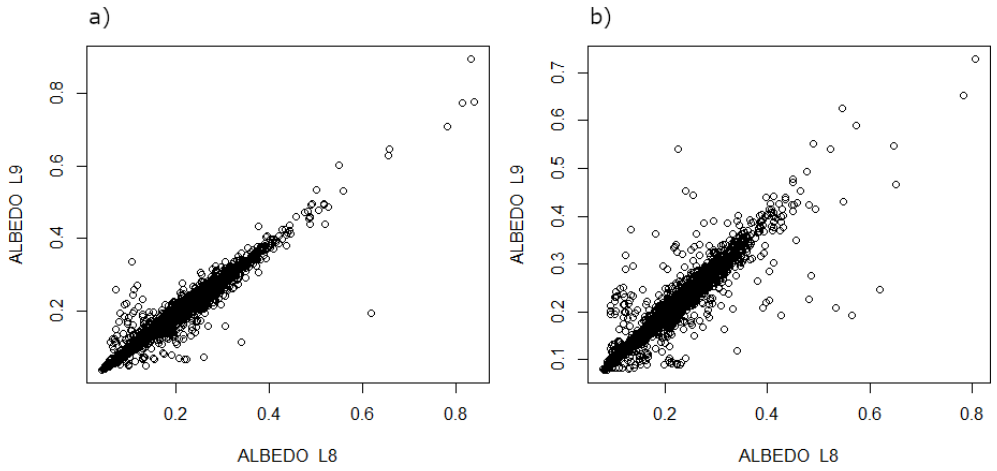


Fig. 2.9 – Scatterplots between albedo values estimated from Landsat 8 and Landsat 9 data for Palermo (a) and Cagliari (b)

According to the statistical metrics reported in Table 2.3, the Landsat 8 data produced slightly higher μ, SD, and m values than the Landsat 9 data did (with a difference that was less than 10^{-2}), for both case studies. Moreover, the mean albedo values that were determined were quite close to 0.20 for both sites. This information, when combined with the SD values, corresponds to those provided in previous research publications: Taha [49] indicated that the albedo values in urban regions range between 0.10 and 0.20 for several European cities, while

Brivio et al. [13] reported that the albedo values in urbanized areas range between 0.12 and 0.21.

The ρ values were very high, whereas the RMSE values were low for both study areas. This is supported by the scatterplots in Figure 2.9, which indicate an optimal correlation of the albedos for both Palermo and Cagliari (despite the fact that Cagliari had a higher number of mis-corresponding values). The Palermo region has, in fact, a greater correlation and a smaller RMSE between the Landsat 8 and Landsat 9 estimations.

In addition, the scatterplots (Fig. 2.9) and the boxplots (Fig. 2.8) demonstrate the presence of multiple outliers in the albedo distributions for both territories. To detect and extract them, Tukey's filter and an outlier analysis were employed.

Figure 2.10 represents the location of the outliers, extracted from Landsat 8 and Landsat 9 albedo estimations, for Palermo and Cagliari regions. Most of the Landsat 8 and 9 observed outliers overlapped in both study areas (Table 2.4). In Palermo, overlapping outliers represented more than 89% of the totals; in Cagliari, instead, these accounted for more than 76% of the totals.

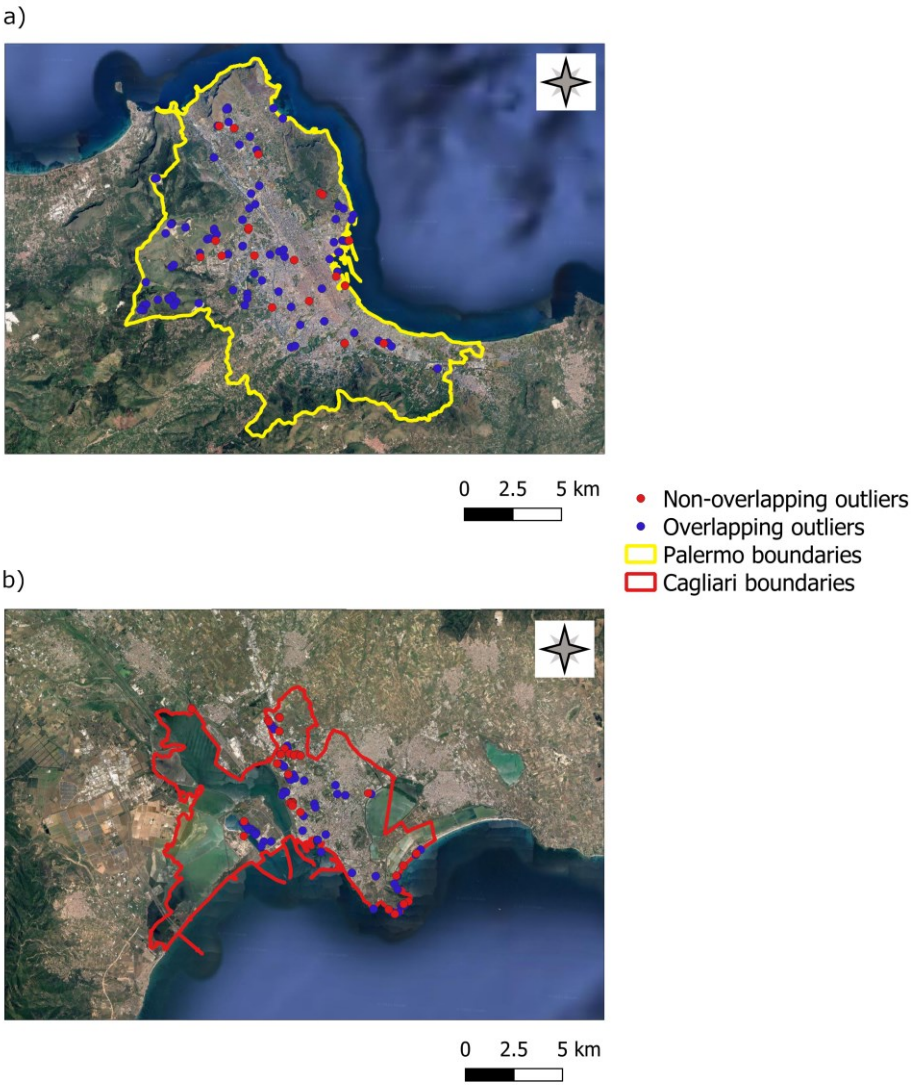


Fig. 2.10 – Outliers' positions in Palermo (a) and Cagliari (b)

Tab. 2.4 – Percentage of overlapping and non-overlapping outliers for Palermo and Cagliari study areas

Study area	**Image**	**Overlapping outliers [%]**	**Non-overlapping outliers [%]**
Palermo	Landsat 8	89	11
	Landsat 9	91	9
Cagliari	Landsat 8	79	21
	Landsat 9	76	24

As a result of the statistical comparison, it is possible to affirm that the da Silva et al. method [57] is appliable to Landsat 9 images with a good approximation, even if some radiometric differences (resulting in a few mismatched albedo values) exist between the Landsat 8 and 9 computations. The comparability of the outputs implies that such procedure may be therefore extended to Landsat 9 data, which can be utilized in the Landsat 8 continuity [27].

3. ESTIMATING LAND CONSUMPTION FROM LANDSAT 8 DATA USING POST-CLASSIFICATION AND ALBEDO CHANGE ANALYSIS

This chapter describes a change detection method, consisting of a hierarchical model that integrates a novel index-based classification method and albedo change analysis, developed in this research to extract information on land consumption (the "urban growth") from Landsat 8 (L8) satellite imagery. A custom JavaScript code was created and implemented in GEE to perform the analysis [11].

3.1. Data

Table 3.1 depicts the main characteristics of the Landsat 8 satellite. Table 3.2 shows instead the spectral bands of the Landsat 8 sensors. The instruments on board the Landsat 8 platform (the Operational Land Imager, OLI, and Thermal Infrared Sensor, TIRS, sensors) are capable of viewing a swath 185 Km wide and collecting images along that swath as the satellite moves through its orbit. The orbit is maintained so that it is possible to see the entire surface of the Earth once every 16 days [68-70].

Tab. 3.1 – Main characteristics of the Landsat 8 satellite.

Landsat 8	
Launch date	February 11, 2013
Sensors	OLI, TIRS
Altitude	705 Km
Inclination	98.2°
Orbit	Polar, sun-synchronous
Scene size	185 Km x 180 Km
Repeat coverage	16 days
Radiometric resolution	12-bit

Tab. 3.2 – Spectral bands of the Landsat 8 satellite sensors.

Landsat 8 band	Sensor	Geometric resolution	Wavelength range
Band 1 – Coastal aerosol	OLI	30 m	0.435 – 0.451 µm
Band 2 – BLUE	OLI	30 m	0.452 – 0.512 µm
Band 3 – GREEN	OLI	30 m	0.533 – 0.590 µm
Band 4 – RED	OLI	30 m	0.636 – 0.673 µm
Band 5 – NIR	OLI	30 m	0.851 – 0.879 µm
Band 6 – SWIR 1	OLI	30 m	1.566 – 1.651 µm
Band 7 – SWIR 2	OLI	30 m	2.107 – 2.294 µm
Band 8 – Panchromatic	OLI	15 m	0.503 – 0.676 µm
Band 9 – Cirrus	OLI	30 m	1.363 – 1.384 µm
Band 10 – TIR 1	TIRS	100 m	10.60 – 11.19 µm
Band 11 – TIR 2	TIRS	100 m	11.50 – 12.51 µm

The SR imagery, derived from TOA reflectance data (corrected for the varying scattering and absorbing effects of atmospheric gases and aerosols), is required to properly monitor land surface changes. The L8 SR data are generated using the Land Surface Reflectance Code (LaSRC) (version 1.5.0), which uses atmospheric ancillary data, a digital elevation model and a unique radiative transfer model [68, 71].

3.2. Development of the method

The method consists of a combined approach using two widely recognized change detection techniques:

- "image classification and post-classification comparison";
- "image differencing".

Synthetically, the proposed methodology consists of four main steps:

1. first, cloud-free L8 SR images for each reference year undergo an innovative index-based LU/LC classification;

2. the resulting classified maps, aggregated into a binary scheme (“Urban” and “Non-urban” areas), are then subjected to the “post-classification comparison”;
3. for areas identified as potentially changing from "Non-urban" to "Urban", further analysis using the "image differencing" method is necessary. This involves comparing two land surface albedo maps derived from single multi-temporal winter L8 images;
4. finally, a neighborhood filter is applied to remove isolated "changed" pixels and refine the results.

This methodological approach integrates both LU/LC classification and change detection techniques, providing a comprehensive analysis of urban growth dynamics over time [11].

3.2.1 Study area

The territory of the municipality of Bitritto (Figure 3.1), in southern Italy, was selected as pilot site to test such a methodology. The analysis was carried out in order to extract the urban growth occurred from 2015 to 2023.

Located in the center of the Apulia region, the study area (approximately 18 km^2) is part of the Metropolitan City of Bari. The area is predominantly flat, and its land cover consists largely of agricultural fields or other green spaces and relatively minor artificial or built-up land. Nevertheless, like many other areas of Apulia, it has experienced significant land consumption over the last few decades. The climate of Bitritto is Mediterranean, with hot and dry summers, and mild winters. The rainfall is mainly concentrated in the autumn and winter period, while the summers are usually relatively dry, with no precipitation even for long periods of time [4, 72-73].

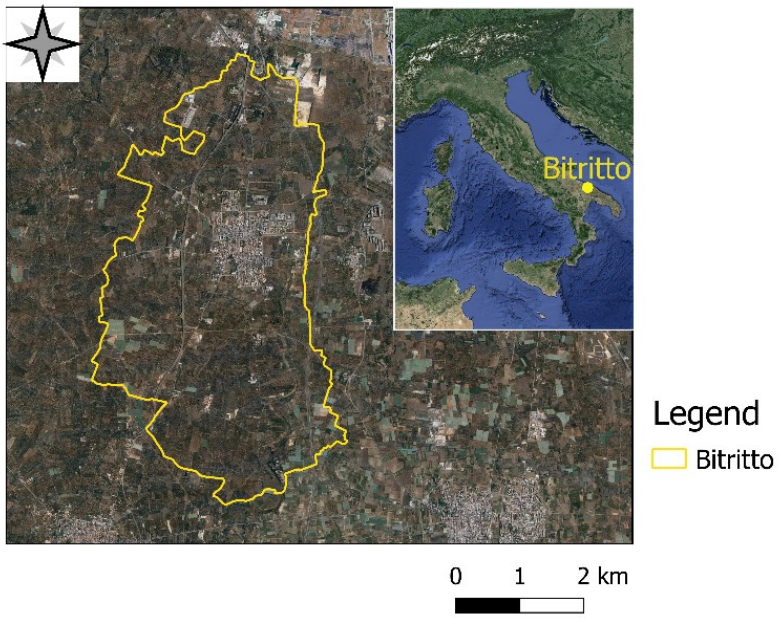


Fig. 3.1 – Study area of Bitritto

3.2.2. Workflow

The methodological approach, implemented in the GEE Code Editor by writing a proper JavaScript code, is illustrated in Figure 3.2.

After selecting single cloud-free L8 SR and raw satellite images, acquired in summer period (dated 29 August 2015 and 19 August 2023) from the GEE data catalog, the LU/LC maps were obtained using a novel classification method (the decision tree algorithm illustrated in Figure 3.3). This innovative classification, developed in this study, involves a hierarchical method using SwirTirRed (STRed) and SwiRed spectral indices, the near infrared (NIR)/shortwave infrared band 2 (SWIR2) ratio, and the NIR band.

The extracted LU/LC classes were then aggregated into binary maps to delineate "Urban" and "Non-urban" areas. Then, the "post-classification comparison" method was applied pixel-by pixel.

Subsequently, the "image differencing" method was applied to the resulting potential "changed" pixels (in particular, for those transitioning from "Non-urban"

to “Urban”) to refine the change detection, thus eliminating spurious "changed" areas due to misclassifications present in the 2015 and 2023 LU/LC maps. For this purpose, two late winter L8 TOA images, acquired on 18 February 2015 and 24 February 2023, were used to retrieve the land surface albedo maps with the da Silva et al. (2016) algorithm [57].

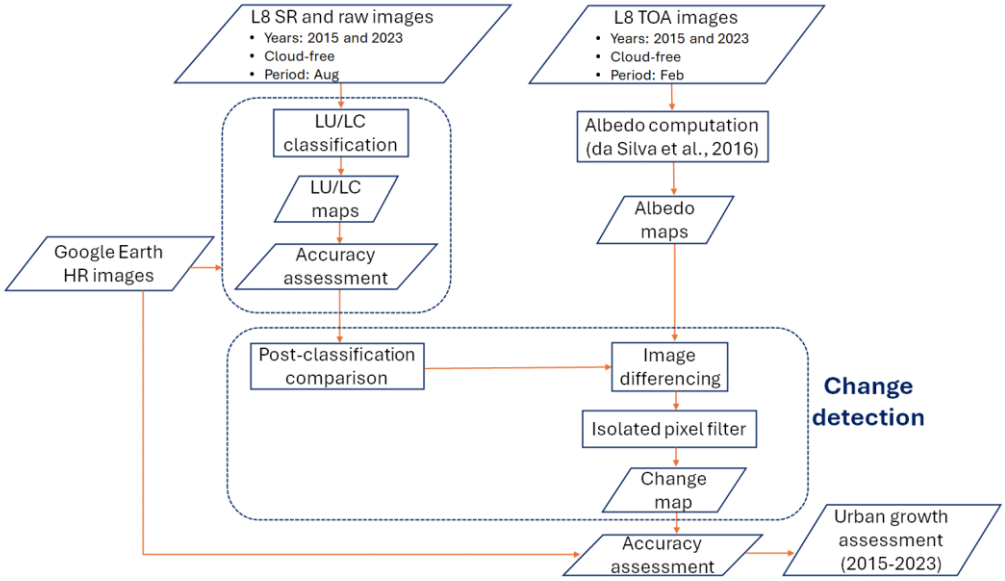


Fig. 3.2 – Workflow of the “post-classification and albedo change” method. HR: High-resolution

Therefore, after subtracting the 2015 from the 2023 albedo maps and defining an appropriate threshold representing significant changes in LU/LC, this stage effectively identified and removed several incorrect "changed" areas.

Lastly, a connected-neighbors filter to remove isolated pixels was applied to improve the final change map [11].

3.2.3. The novel index-based classification method for Landsat 8 data

In general, an appropriate classification system is a prerequisite for a good LU/LC classification. It is necessary that such a scheme (which mainly depends on the user’s need, sensor resolution and time constraints) is informative, ex-

haustive and separable. Moreover, a hierarchical model, which combines more sub-classes into a higher-level class, is often required [19, 74].

The chosen classification scheme for this index-based classification method includes four main LU/LC categories to be distinguished: “Vegetation”, “Plastic greenhouses”, “Urban areas” and “Bare soil”. Index-based approach was preferred to other methods since it is considered an effective technique for quickly extracting LU/LC classes from satellite data, since it exploits land surface features capabilities in absorbing, reflecting and transmitting the radiation in different spectral bands of the electromagnetic spectrum [75].

The proposed method (Figure 3.3) involves the sequential application of three binary decision rules including the STRed and SwiRed spectral indices (which were introduced for the Landsat Image Classification Algorithm [LICA] [75] in 2020]), the NIR/SWIR2 ratio and the NIR band.

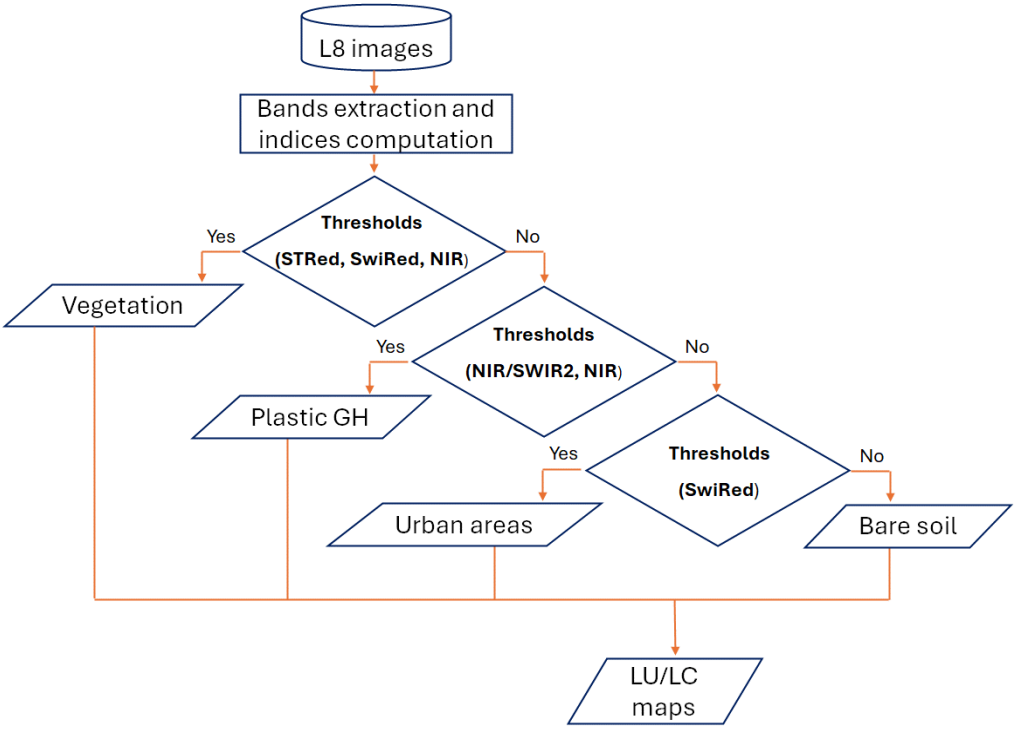


Fig. 3.3 – Flowchart of the novel decision tree algorithm for the LU/LC classification

Therefore, the LU/LC classes were distinguished in the following order:

1. “Vegetation” (using STRed, SwiRed and NIR band thresholds);
2. “Plastic greenhouses” (using NIR/SWIR2 ratio and NIR band thresholds);
3. “Urban areas” and “Bare soil” (using SwiRed thresholds).

The STRed and SwiRed indices are expressed as follows:

$$STRed = \frac{SWIR1+Red-TIR1}{SWIR1+Red+TIR1} \quad , \tag{1}$$

$$SwiRed = \frac{SWIR1 - Red}{SWIR1 + Red} \tag{2}$$

where SWIR1, Red and TIR1 are Landsat 8 shortwave infrared band 1, red band and thermal infrared band 1, respectively [75]. SWIR 1, Red and TIR 1 bands are commonly applied in literature for LU/LC classification. Indeed, SWIR 1 band is particularly sensitive to the water content in plants and soil, Red band depends on the energy absorbed by the chlorophyll, while TIR 1 is linked to the energy absorbed and emitted by the surfaces [14, 68].

The training stage of the classification allowed to manually select a consistent number of training pixels representative of each of the four LU/LC classes on L8 SR images. Visualization of high-resolution images, available in Google Earth Pro (https://www.google.com/intl/it/earth/about/versions/#earth-pro), was very useful to correctly distinguish each category on the ground. As recommended by [13], the training pixels were selected within homogeneous areas (to adequately represent the class model), in different zones of the study area (to account for the intrinsic variability of the class), and in sufficient numbers (to achieve meaningful results).

Graphical representation of the spectral response patterns for these training areas, useful to analyze the spectral separability of each thematic class, were

obtained in the GEE platform. Average spectral signature plots for the four LU/LC categories are shown in Figure 3.4.

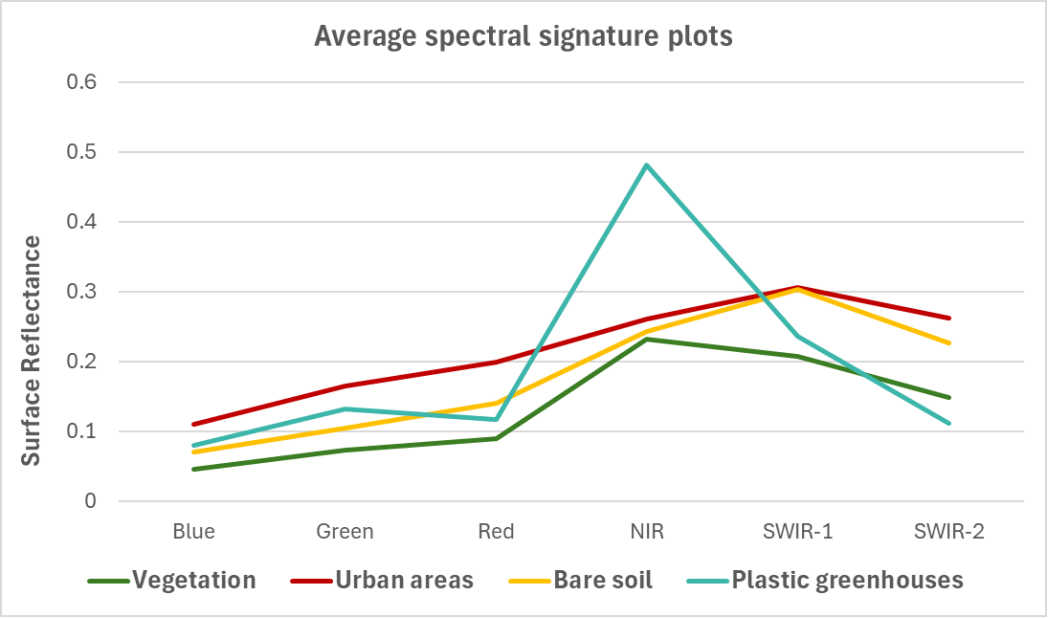


Fig. 3.4 – Average spectral signature plots obtained from training pixels selected for each LU/LC class (Bitritto)

Moreover, a scatter diagram between STRed and SwiRed indices (Figure 3.5) was obtained from the training pixels also to select appropriate thresholds for the decision rules. The analysis suggested the selection of SwiRed thresholds (adopted in LICA [75] only for the extraction of built-up land) to better separate the "Vegetation" class (including both sparse and dense vegetated areas) from other categories, in addition to the STRed index (already used in LICA [75] for the extraction of vegetation).

Furthermore, in order to detect the "Plastic greenhouses", the analysis of Figure 3.4 suggested that the NIR and SWIR2 bands might be the most suitable to identify them with good accuracy, due to the relative reversal of their spectral response for these bands compared to those of the other three categories [11, 14]. The scatter plot between NIR and SWIR2 reflectances confirmed this hy-

pothesis, as shown in Figure 3.6, where the spectral separability of plastics from the other classes is very clear.

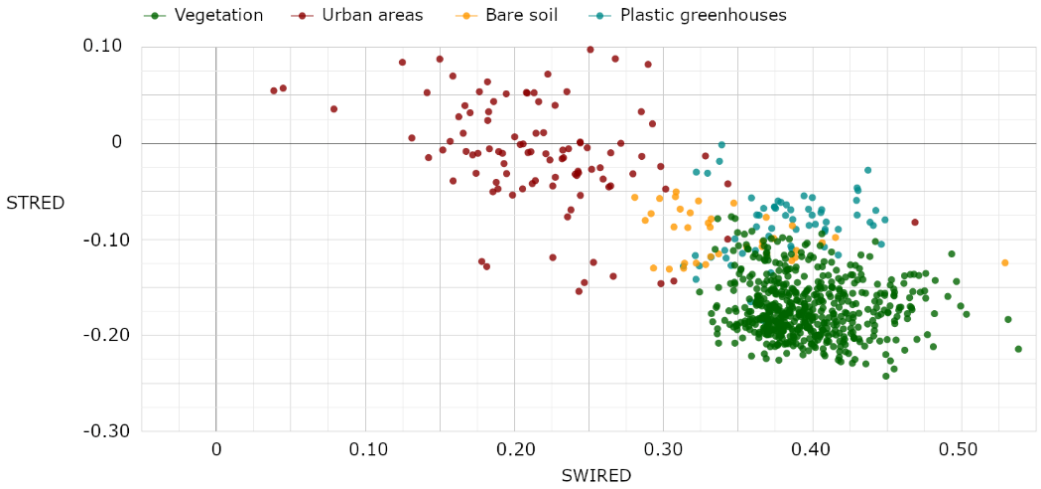


Fig. 3.5 – Scatter diagram (STRed/SwiRed) for training pixels of “Vegetation”, “Urban areas”, “Bare soil”, and “Plastic greenhouses” classes obtained from L8 SR images for Bitritto

Hence, thresholds of NIR and NIR/SWIR2 ratio were defined from Landsat 8 SR images to separate the “Plastic greenhouses” LU/LC class.

After distinguishing “Vegetation” and “Plastic greenhouses”, the SwiRed index was applied to L8 SR images to extract the “Urban areas” class (inclusive of the LICA “Built-up” and “Mining areas” categories [75]). More specifically, the “Urban areas” class includes land covered by man-made structures (i.e. residential, commercial and industrial buildings, transportation networks, other paved areas, construction sites etc.), as well as mines and quarries.

After the extraction of “Urban areas”, the remaining part of the images was classified as “Bare soil” [11].

The resulting classified maps were aggregated into binary classifications consisting only of the classes “Urban areas” and “Non-urban areas” (the latter being the result of the aggregation of “Vegetation”, “Plastic greenhouses” and “Bare soil”), in order to perform the change detection and the estimation of the urban growth (2015-2023).

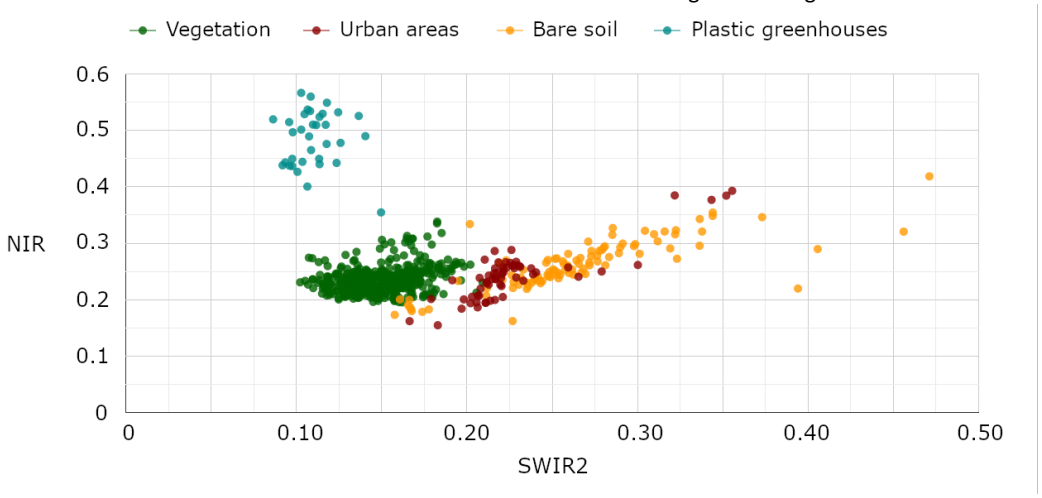


Fig. 3.6 – Scatter diagram (NIR/SWIR2) for training pixels of "Vegetation", "Urban areas", "Bare soil", and "Plastic greenhouses" classes obtained from L8 SR images for Bitritto

3.2.4 Change detection

After aggregating the 2015 and 2023 LU/LC maps into a binary scheme ("Urban areas" and "Non-urban areas"), the two resulting maps (the "Urban/Non-urban maps") were compared pixel-by-pixel. In this manner, the possible outputs were "no change" (for the same classification in both the 2015 and 2023 maps) or "change" (for pixels classified differently in the 2015 and 2023 maps).

Once the post-classification comparison was performed, the "image differencing" step was applied to the resulting "changed" pixels (specifically, for pixels changed from "Non-urban" in 2015 to "Urban" in 2023), by computing the difference image between the two land surface albedo maps, estimated with the da Silva et al. (2016) algorithm [57], from single cloud-free L8 TOA winter images (captured on 18 February 2015 and on 24 February for the reference years 2015 and 2023, respectively).

A suitable threshold of albedo change, potentially indicative of meaningful LU/LC changes, from "Non-urban" to "Urban", was established. Late winter imagery was preferred over other seasons because changes from bare or sparsely vegetated land to built-up areas may be most evident due to the higher mois-

ture content and green cover (leading to a lower albedo value) of the former in this season, as higher rainfall in the autumn/winter period (leading to higher soil moisture at the end of winter) is characteristic of the climate of the study area [11, 73, 76-81]. Consequently, an increase in albedo, from 2015 to 2023 images, could be a sign of a land cover conversion from non-artificialized land to urbanized land ("Urban growth") [11, 82-83]. An albedo increase greater than 0.055 between the 2015 and 2023 maps was set as the threshold for detecting urban growth.

After performing the "image differencing" phase, a connected-neighbors filter to eliminate isolated "changed" pixels was implemented in GEE to further clean the change map from residual misclassifications, and to produce the "urban change map" (2015-2023) depicting two distinct categories: "Urban growth" and "No growth" [11].

3.2.5 Albedo estimation with da Silva et al. (2016) algorithm

The land surface albedo maps for the reference years 2015 and 2023 were obtained by using the da Silva et al. algorithm (2016) [57], as described in Chapter 2.

The value of P_0 for the calculation of the atmospheric transmissivity was obtained from the measurement (at 10:00 on the days of the L8 satellite images) taken at the Bari station of the Italian National Tidegauge network (https://mareografico.it). W was instead provided by the "NCEP/NCAR Reanalysis data, Water Vapor" dataset [61], while K_t was set equal to 1 (pure air). Finally, the value of Z was calculated from the sun elevation angle [°], reported in the image metadata.

3.2.6 Accuracy assessment

The estimation of the accuracy of a classified map is usually done by comparing information from reference sites with information on the map for a number of test areas, through the construction of the "error matrix". The error

matrix is a square array composed of rows and columns that express the labels assigned to a given class in the "reference" map relative to the labels assigned to a given class in the classified map [84].

To assess the accuracy of the LU/LC maps, classified with the novel index-based method, a multi-temporal dataset of test pixels was generated using the stratified random sampling[2] method. A total number of 245 "pure" pixels were chosen, according to the extent of each LU/LC class, in the following proportions: 140 for "Vegetation", 55 for "Urban areas", 30 for "Bare soil" and 20 for "Plastic greenhouses". A visual interpretation of the high-resolution Google Earth images, chosen as "reference" data, was carried out to assign each sample pixel to the correct category, based on its location. The metrics[3] calculated from the error matrix to evaluate the accuracy were overall accuracy (O.A.), producer's accuracy (P.A.), and user's accuracy (U.A.) [13, 85-86].

In order to assess, instead, the accuracy of the change map, depicting the soil consumption that occurred in Bitritto for the period 2015-2023, a dataset of 100 test pixels, corresponding to "Urban growth" and "No growth" categories, was obtained using the stratified random sampling method, after visual inspection of multitemporal high-resolution Google Earth images. According to the extent of each class, pixels were selected in the following proportions: 20 for the "Urban growth" category, and 80 for the "No growth" class. Even in this case, the accuracy metrics calculated were O.A., P.A. and U.A. [11].

3.3. Results for Bitritto (2015-2023)

The LU/LC maps, produced in GEE for the reference years 2015 and 2023 after implementing the index-based approach described previously, are displayed in Figure 3.7. The accuracy metrics obtained for the validation of the thematic maps are instead shown in Table 3.4.

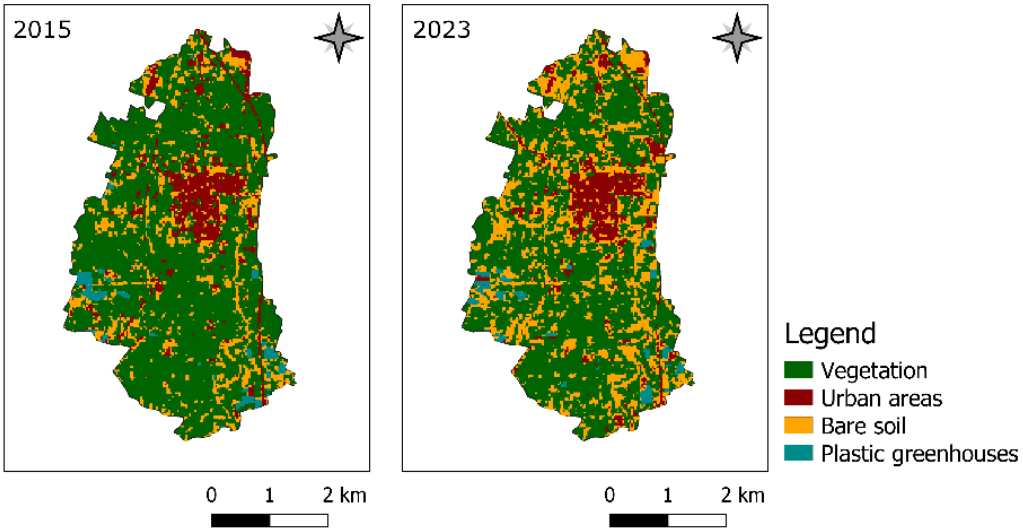


Fig. 3.7 – LU/LC maps for the years 2015 and 2023 for Bitritto, obtained using the novel index-based classification method

Tab. 3.4 – Accuracy metrics for the LU/LC maps (Bitritto).

	2015		**2023**	
LU/LC class	**P.A. [%]**	**U.A. [%]**	**P.A. [%]**	**U.A. [%]**
Vegetation	95	92	93	94
Urban areas	85	98	84	96
Bare soil	70	64	73	56
Plastic greenhouses	100	100	95	100
O.A. [%]	90		89	

The analysis of the accuracy metrics reveals that the O.A. was quite high for both classification maps (around 90%), while “Plastic greenhouses” and “Vegetation” were the classes with both O.A. and U.A. greater than 90%. The extraction of “Urban areas” was well achieved with P.A. accuracies of 85% and 84% for 2015 and 2023, respectively, while the U.A. accuracies were very high totaling 98% in the 2015 map and 96% for the 2023 classification. Lower accuracies, instead, were found for the bare soil extraction (P.A. > 70% and U.A. > 55%).

The “Urban/Non-urban maps” for the years 2015 and 2023 obtained after the aggregation of the LU/LC classes are reported in Figure 3.8. The visual analysis of these classifications suggests that, probably, only minor changes are likely to have occurred in the period 2015-2023.

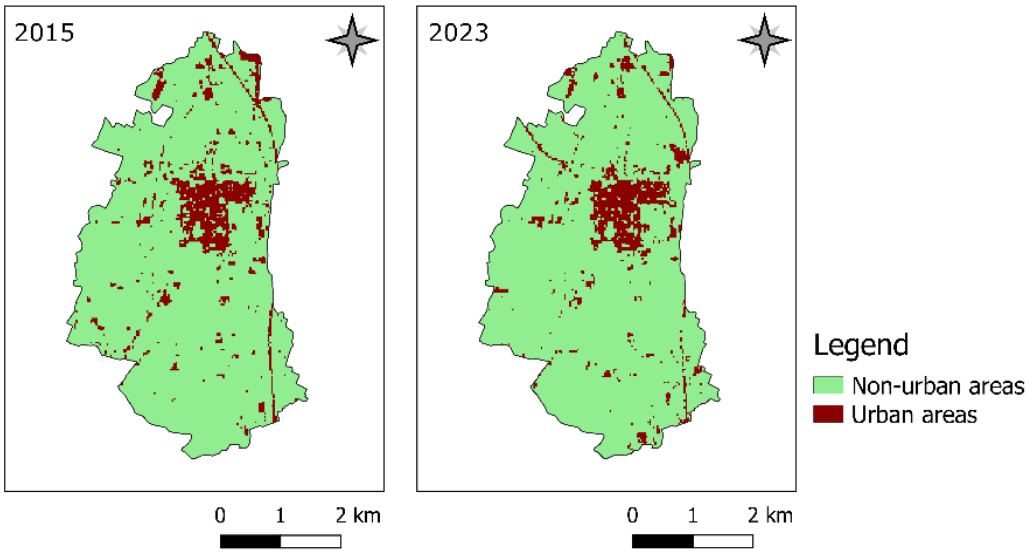


Fig. 3.8 – “Urban/Non-urban” maps for the years 2015 and 2023 for Bitritto, obtained after aggregating the LU/LC maps

The land surface albedo maps produced in GEE by applying the da Silva et al. (2016) algorithm [57] are instead illustrated in Figure 3.9, while the final “urban change map” (2015-2023), obtained using the change detection approach described above, is shown in Figure 3.10. The accuracy metrics for the change map are shown in Table 3.5.

Table 3.5 shows that the O.A. for the urban change map was high (92%). Regarding the estimation of the “Urban growth” (2015-2023), the results were satisfactory results, with a P.A. of 85% and a U.A. of 77%.

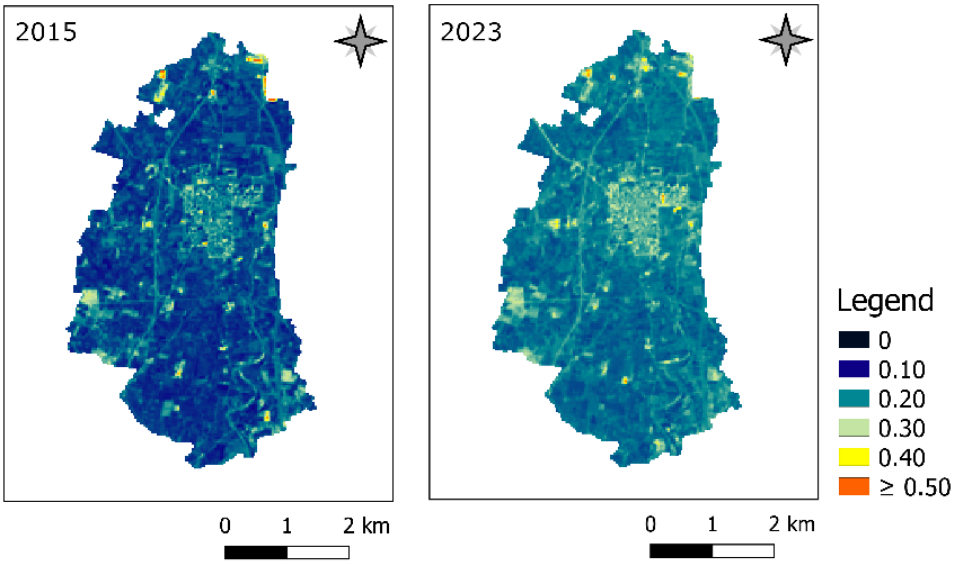


Fig. 3.9 – Land surface albedo maps for the years 2015 and 2023 for Bitritto area

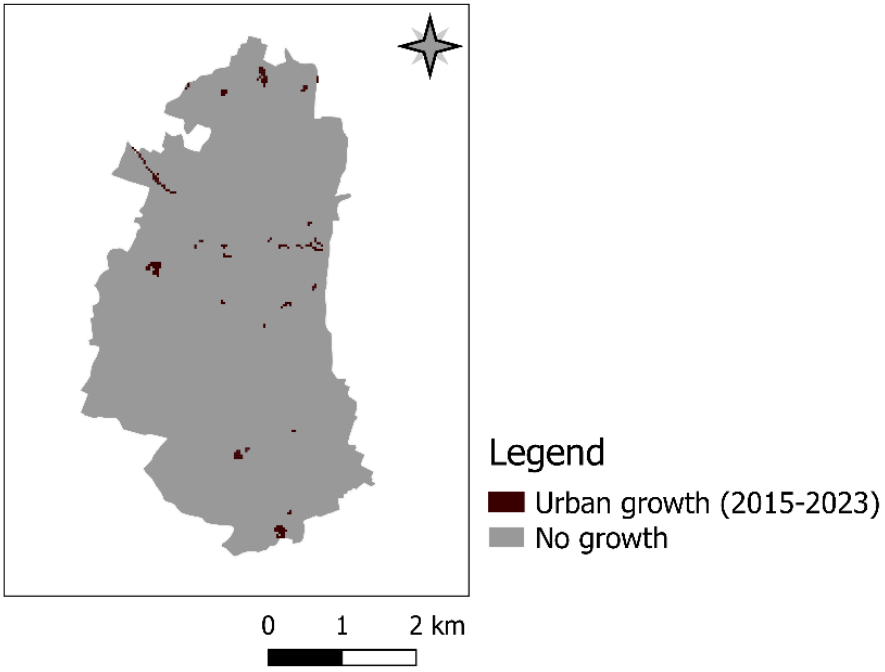


Fig. 3.10 – Urban change map for the period 2015-2023 for Bitritto area

Tab. 3.5 – Accuracy metrics for the urban change map (Bitritto).

	2015-2023	
Class	**P.A. [%]**	**U.A. [%]**
Urban growth	85	77
No growth	94	96
O.A. [%]	92	

Furthermore, the analysis of the map confirms that the land consumption for the Bitritto territory is moderate for the period 2015-2023. However, there are still a few mis-changed areas in the map, probably related to the problem of "mixed" pixels (which is common in the classification of medium-resolution images [18, 87], such as Landsat data) or to the low spectral confusion that persists between certain classes (in particular "Urban areas" and "Bare soil") [11]. The results are promising, regardless of the applicability of this algorithm also to other study areas in the Apulia region, in order to automatically extract information on land consumption occurring at the local scale and to rapidly monitor urban development, thanks to the relative simplicity and practicality of this method.

3.4. Application of the method for Bari (2015-2023)

The change detection method developed in this research study (Figure 3.2) for the Bitritto site, consisting of the integration of the post-classification comparison and the land surface albedo change analysis (described in paragraph 3.2), was subsequently applied to the territory of the municipality of Bari in order to extract information on the land consumption occurred at the urban scale between 2015 and 2023.

The same input data (L8 raw and SR images acquired on 29 August 2015 and 19 August 2023 for the LU/LC classification, and L8 TOA images collected on 18 February 2015 and 24 February 2023 for the albedo estimation) as those used for the Bitritto case study were chosen because they cover the entire territory of Bari.

3.4.1 Study area

Bari (Figure 3.11) is a Mediterranean coastal city and the capital city of the Apulia region, in southern Italy.

Its territory covers an area of about 116 km^2, characterized by a flat topography. Surface water bodies are almost absent, although erosive karstic grooves (called “lame”) cross the territory and convey water in case of very heavy rainfall. The land cover consists of built-up land with a mosaic of agricultural fields in the peri-urban area (which extends into the city along the “lame”), while green spaces are scarce.

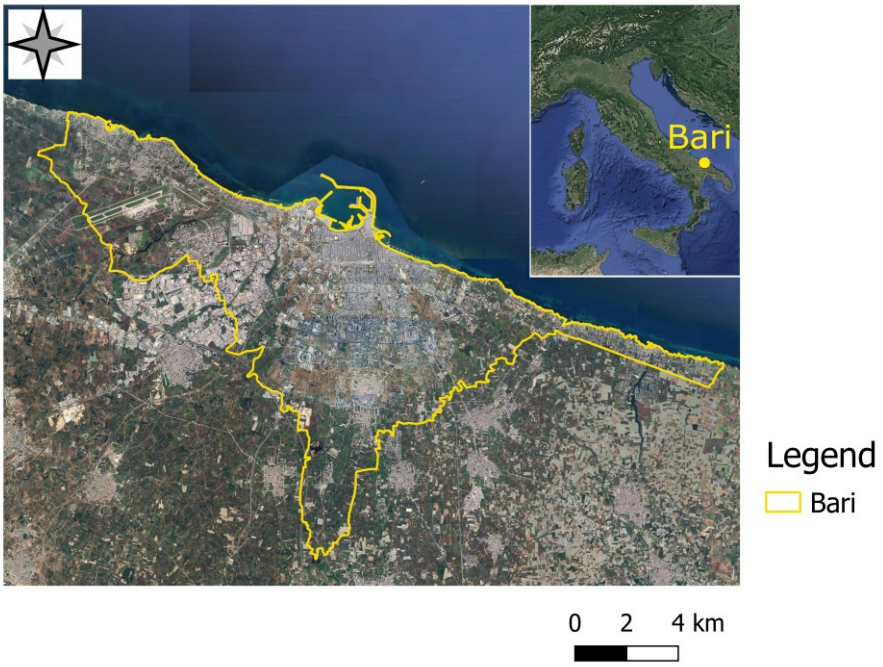


Fig. 3.11 – Study area of Bari

In the second half of the mid twentieth century, Bari underwent an urban sprawl that led to the expansion of its inner core along preferential directions or nuclei. Like other areas of the Apulia region, Bari has experienced significant land consumption in recent decades, with a tendency towards dispersion and diffusion in suburban and rural areas [4, 38, 73, 88-89].

3.4.2 Accuracy assessment

For the accuracy assessment of the change map, obtained for the site of Bari after the application of the change detection method, a dataset of 220 test pixels, corresponding to “Urban growth” and “No growth” classes, was generated using the stratified random sampling method, after visual inspection of multitemporal high-resolution Google Earth images. According to the extent of each class, pixels were selected in the following proportions: 20 for the “Urban growth” category, and 200 for the “No growth” class. The accuracy metrics calculated were O.A., P.A. and U.A.

3.4.3 Urban growth analysis

Moreover, in order to quantify and analyze the increase of urban areas, occurred in Bari between 2015 and 2023, some metrics (explained in Table 3.6) were calculated [2].

It is important to remind and underline that in this study for “urban areas” is intended the land covered by artificial structures (i.e. residential, commercial and industrial buildings, transportation infrastructures, construction sites and other paved surfaces), as well as mines and quarries.

Tab. 3.6 – Description of the urban growth metrics calculated for the study area of Bari (2015-2023).

Metric	Description
Growth of Urban areas ($GU_{2015\text{-}2023}$) [Km^2]	Urban growth from 2015 to 2023 (in Km^2) estimated with the change detection method
Rate of Urban Growth ($RGU_{2015\text{-}2023}$) [%]	Rate of Urban growth (in percentage), related to urban areas present in the 2015 map $RGU_{2015\text{-}2023} = ((GU_{2015\text{-}2023})/ \text{Urban area}_{2015}) *100$

Average annual Urban Growth (AaGU) [Km^2/year]	$AaGU = (GU_{2015\text{-}2023})/N_{YEARS}$

3.4.4 Results

The “Urban/Non-urban maps” of Bari produced for the reference years 2015 and 2023 are shown in Figure 3.12.

In Figure 3.13 are illustrated the land surface albedo maps calculated with the da Silva et al. method [57], while Figure 3.14 shows the urban change maps which highlights the soil consumption occurred between 2015 and 2023, as retrieved by the application of the change detection method.

In Table 3.7 are reported the accuracy metrics relative to the urban change map (2015-2023).

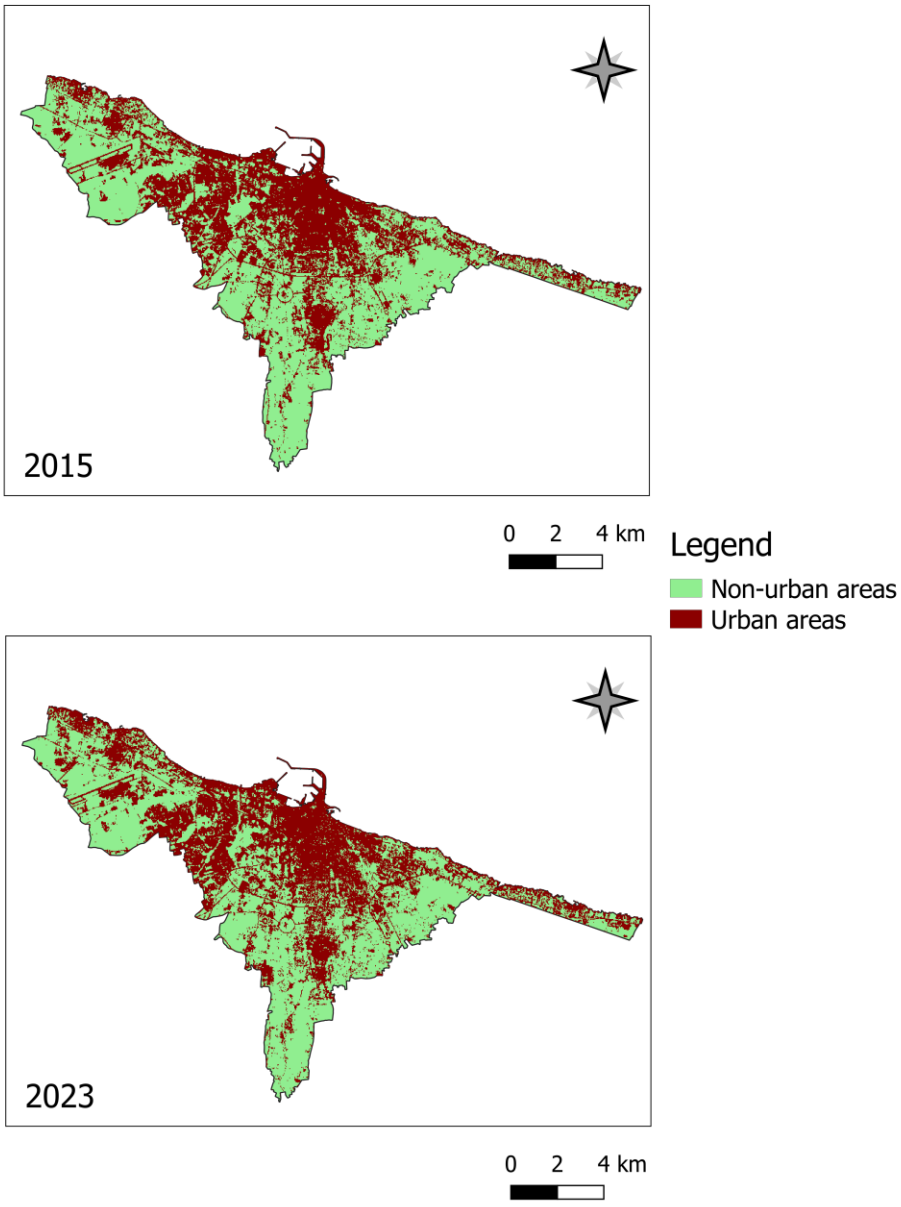


Fig. 3.12 – “Urban/Non-urban” maps produced for the study area of Bari for the reference years 2015 and 2023 using the novel index-based classification method

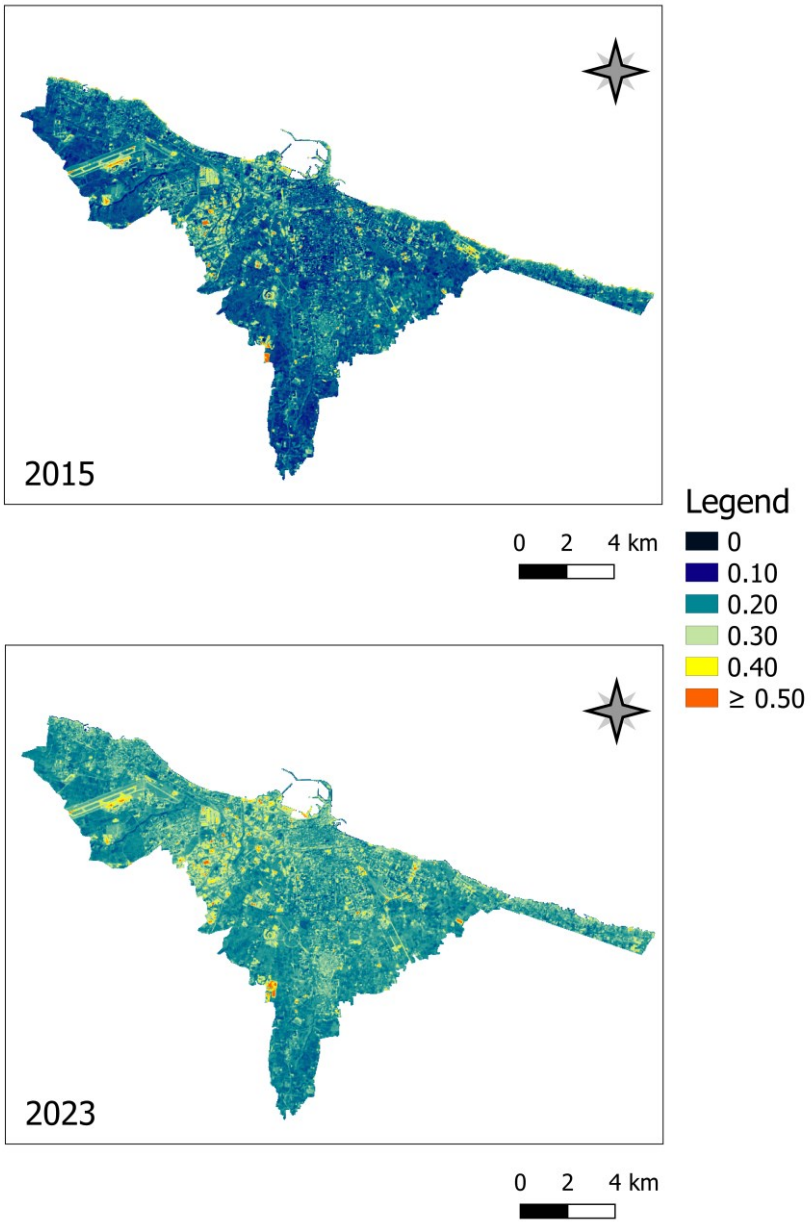


Fig. 3.13 – Land surface albedo maps produced for the study area of Bari for the reference years 2015 and 2023

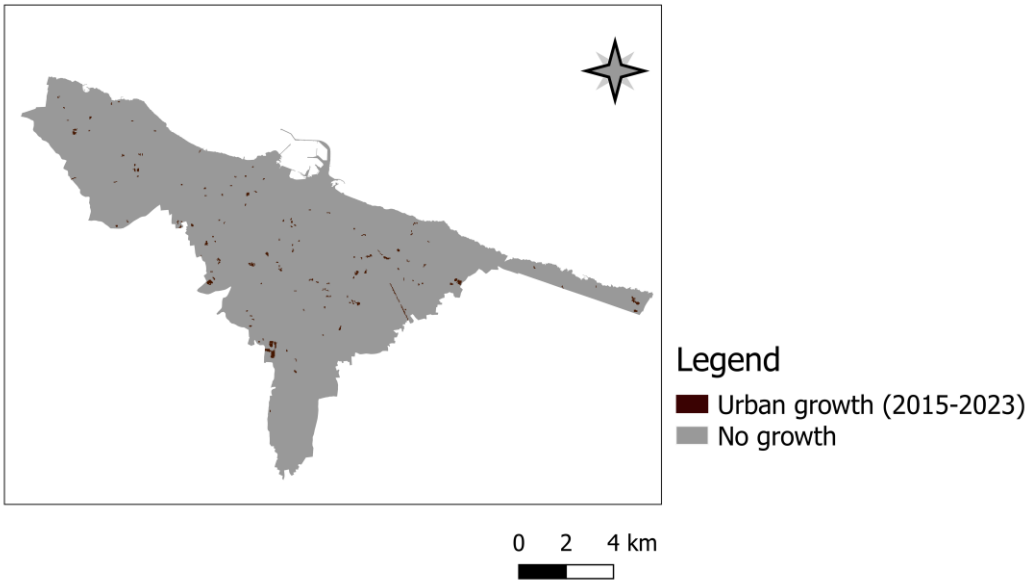


Fig. 3.14 – Urban change map estimated for the period 2015-2023 for the study area of Bari using the "post-classification and albedo change analysis" method

Tab. 3.7 – Accuracy metrics for the urban change map (Bari).

	2015-2023	
Class	**P.A. [%]**	**U.A. [%]**
Urban growth	80	94
No growth	99	98
O.A. [%]	98	

Table 3.7 highlights the very high accuracies were found for the urban change map obtained through the application of this change detection method. The O.A., in fact, totals 98% while the P.A. and the U.A. for the "Urban growth" category are 80% and 94%, respectively. This result confirms the applicability of this method for the area of Bari, already highlighted in paragraph 3.3, where it was demonstrated that very good results were obtained for the test site of Bitritto, which is a municipality bordering with Bari and having very similar climatic and geomorphological characteristics (they belong indeed to the same geographical context). Table 3.8 shows the urban growth metrics (explained in Ta-

ble 3.6) estimated for the study site of Bari, which are useful to quantify the land consumption and analyze urban dynamics.

Tab. 3.8 – Urban growth metrics calculated for Bari (2015-2023) using the post-classification and albedo change method.

Metric	Value
$GU_{2015\text{-}2023}$ [Km^2]	1.1
$RGU_{2015\text{-}2023}$ [%]	+2.3
AaGU [Km^2/year]	0.1

From Table 3.8 it is possible to notice that the soil consumption (the increase of urban areas between 2015 and 2023, highlighted in Figure 3.14) estimated with the post-classification comparison and albedo change method for Bari is equal to 1.1 Km^2. The growth of urban areas in this period (8 years), in relation to the consumed land in 2015, was found to be +2.3%, with an average annual urban growth of around 0.14 Km^2 per year. Therefore, as for Bitritto, a moderate land consumption was found to be occurred for the city of Bari. However, the visual inspection of Figure 3.14 confirms the tendency to a densification of urban land with a sparse dispersion and diffusion of impervious surfaces also within suburban and rural areas.

4. *ESTIMATING LAND CONSUMPTION FROM LANDSAT DATA USING CONTINUOUS CHANGE DETECTION AND CLASSIFICATION ALGORITHM*

This chapter describes the Continuous Change Detection and Classification method, implemented in the GEE cloud environment, and used in this work to estimate land consumption from Landsat satellite data. The CCDC method adopts a per-pixel model fitting approach to capture seasonal and inter-annual surface variations from all available satellite observations. The changes are detected for each individual pixel as the fitted model breaks down over time, and a classification of the different models allows a LU/LC map to be produced for each point in the time series. This approach therefore accounts for the temporal trend of each pixel rather than the spectral signature at a given point in time, potentially achieving higher classification accuracies [29, 31].

In general, the changes occurring on the land surface fall into three categories:

1. intra-annual changes;
2. gradual inter-annual changes;
3. abrupt changes.

The intra-annual changes (e.g. changes resulting from vegetation phenology) are driven by seasonal patterns of environmental variables (such as temperature and precipitation), while the gradual inter-annual changes are determined by factors such as climate variability or vegetation growth. The abrupt changes, on the other hand, are due to phenomena such as deforestation, fires or urbanization. All of these categories can be captured by the time series model coefficients of the CCDC algorithm [26, 29-31].

The advantage of using all available satellite observations is that, in this way, it is possible to analyze also intra-annual surface dynamics that can't be detected from conventional single-date methods. Moreover, the term "continuous" re-

fers also to the capability to extract LU/LC maps for any data time within the period covered by the images [29].

4.1. Data

In this study, all available SR images from the Landsat 4, 5, 7, 8 and 9 missions [69, 90-93], acquired from 1 January 1984 to 31 January 2024, and with a cloud cover less than 5%, were selected from the GEE data catalog. The Landsat bands involved in the CCDC method (Table 4.1) are the BLUE, GREEN, RED, NIR, SWIR 1 and SWIR 2 of the TM (Landsat 4 and 5), ETM+ (Landsat 7), OLI (Landsat 8) and OLI-2 (Landsat 9) sensors, which have a geometric resolution of 30 m.

Tab. 4.1 – Characteristics of the Landsat spectral bands selected in this study for the CCDC method.

Sensor	Band	Geometric resolution	Wavelength range
L4 TM	Band 1 – BLUE	30 m	0.45 – 0.52 µm
	Band 2 – GREEN		0.52 – 0.60 µm
	Band 3 – RED		0.63 – 0.69 µm
	Band 4 – NIR		0.76 – 0.90 µm
	Band 5 – SWIR 1		1.55 – 1.75 µm
	Band 7 – SWIR 2		2.08 – 2.35 µm
L5 TM	Band 1 – BLUE	30 m	0.45 – 0.52 µm
	Band 2 – GREEN		0.52 – 0.60 µm
	Band 3 – RED		0.63 – 0.69 µm
	Band 4 – NIR		0.76 – 0.90 µm
	Band 5 – SWIR 1		1.55 – 1.75 µm
	Band 7 – SWIR 2		2.08 – 2.35 µm
L7 ETM+	Band 1 – BLUE	30 m	0.45 – 0.52 µm
	Band 2 – GREEN		0.52 – 0.60 µm
	Band 3 – RED		0.63 – 0.69 µm

	Band 4 – NIR		0.77 – 0.90 µm
	Band 5 – SWIR 1		1.55 – 1.75 µm
	Band 7 – SWIR 2		2.08 – 2.35 µm
L8 OLI	Band 2 – BLUE	30 m	0.45 – 0.51 µm
	Band 3 – GREEN		0.53 – 0.59 µm
	Band 4 – RED		0.64 – 0.67 µm
	Band 5 – NIR		0.85 – 0.88 µm
	Band 6 – SWIR 1		1.57 – 1.65 µm
	Band 7 – SWIR 2		2.11 – 2.29 µm
L9 OLI-2	Band 2 – BLUE	30 m	0.45 – 0.51 µm
	Band 3 – GREEN		0.53 – 0.59 µm
	Band 4 – RED		0.64 – 0.67 µm
	Band 5 – NIR		0.85 – 0.88 µm
	Band 6 – SWIR 1		1.57 – 1.65 µm
	Band 7 – SWIR 2		2.11 – 2.29 µm

4.2. Method

Synthetically, the approach followed in this study to estimate urban growth using the CCDC algorithm consists of the following steps (Figure 4.1):

- first, after selecting all available SR Landsat images (1984-2024) with cloud cover less than 5% and defining the spectral bands to be included in the algorithm, the *TemporalSegmentation.Ccdc()* GEE function is implemented and the CCDC output (consisting of a change multiband raster) is obtained and stored in GEE;
- from the CCDC output multiband raster, and after defining the time period of interest (to perform the change analysis), per-pixel model coefficients for the involved time segments are extracted and set as input predictors for the supervised LU/LC classification;

- the Random Forest classifier is used to produce multitemporal "Urban/Non-urban" land cover maps and the "post-classification comparison" method is applied pixel-by-pixel;
- a mode filter is applied to refine the results and remove the "salt and pepper" noise (typical effect of pixel-based classifications);
- then, the final "urban change map" for the period of interest is obtained.

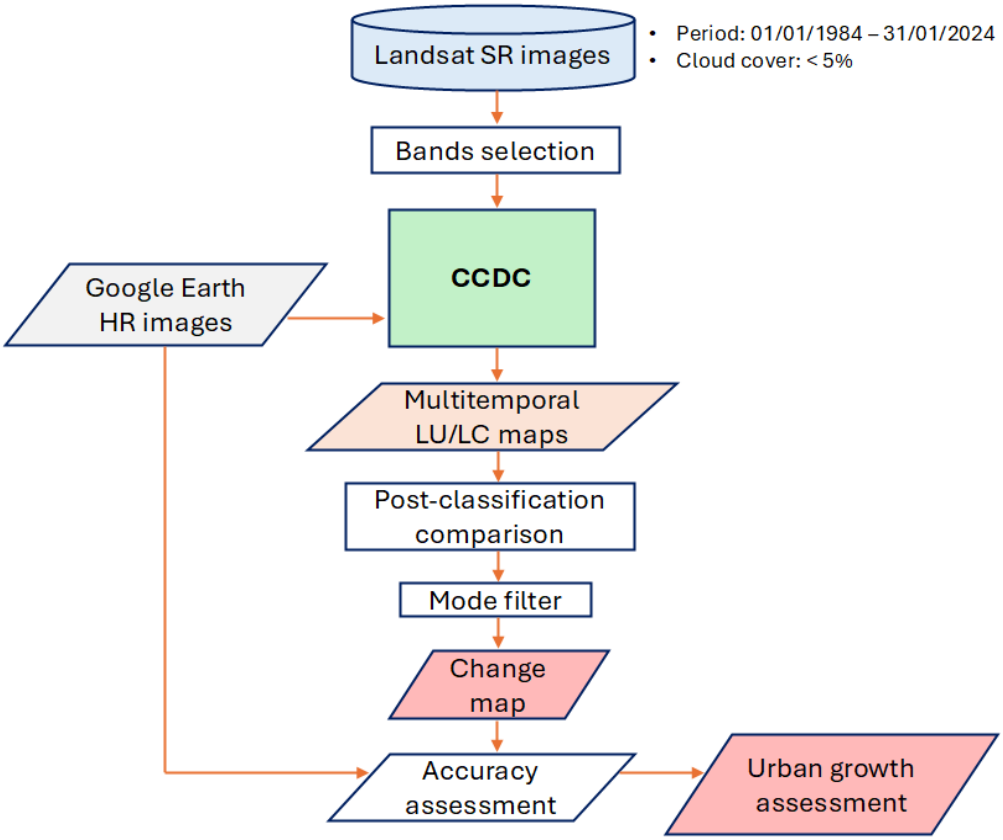


Fig. 4.1 – Operative workflow followed in this study. HR: High-Resolution

4.2.1. Study areas

The analysis was carried out for the study area of Bari (Figure 3.11), which was already investigated in Chapter 3. As in the case of the change detection method described in the previous Chapter, land consumption was also estimated for

the period 2015-2023, in order to compare the performance of the two methods.

Moreover, the study was extended to evaluate also the urban growth for all the twenty regional capital cities of Italy (Figure 4.2). In this case, the time period investigated was 2006-2023.

Fig. 4.2 – The Regional Italian capital cities

Table 4.2 shows the area (in Km^2) of the municipalities territory, the Worldwide Reference System-2 (WRS-2)[4] PATH and RAW [94] numbers of the scene chosen for each case study, and the total number of images founded.

Tab. 4.2 – Area, WRS-2 Path and Row, and number of Landsat images used for each city.

City	Region	Area [Km^2]	Path	Row	Number of images
Ancona	Marche	124.6	190	30	282
Aosta	Valle d'Aosta	21.4	195	28	83
Bari	Puglia	116.3	188	31	335
Bologna	Emilia-Romagna	140.8	192	29	211
Cagliari	Sardegna	84.6	192	33	315
Campobasso	Molise	55.8	189	31	257
Catanzaro	Calabria	111.6	187	33	320
Firenze	Toscana	102.3	192	30	263
Genova	Liguria	240.5	194	29	171
L'Aquila	Abruzzo	472.7	190	31	223
Milano	Lombardia	181.8	194	28	99
Napoli	Campania	118.5	190	32	303
Palermo	Sicilia	160.1	189	34	323
Perugia	Umbria	449.1	191	30	242
Potenza	Basilicata	174.1	188	32	262
Roma	Lazio	1285.6	191	31	323
Torino	Piemonte	130.1	195	29	147
Trento	Trentino Alto-Adige	157.9	192	28	154
Trieste	Friuli Venezia-Giulia	84.9	191	28	102
Venezia	Veneto	415.5	192	28	154

4.2.2. CCDC algorithm

The CCDC algorithm consists of three main steps (Figure 4.3):

1. image preprocessing;
2. continuous change detection;
3. continuous LU/LC classification.

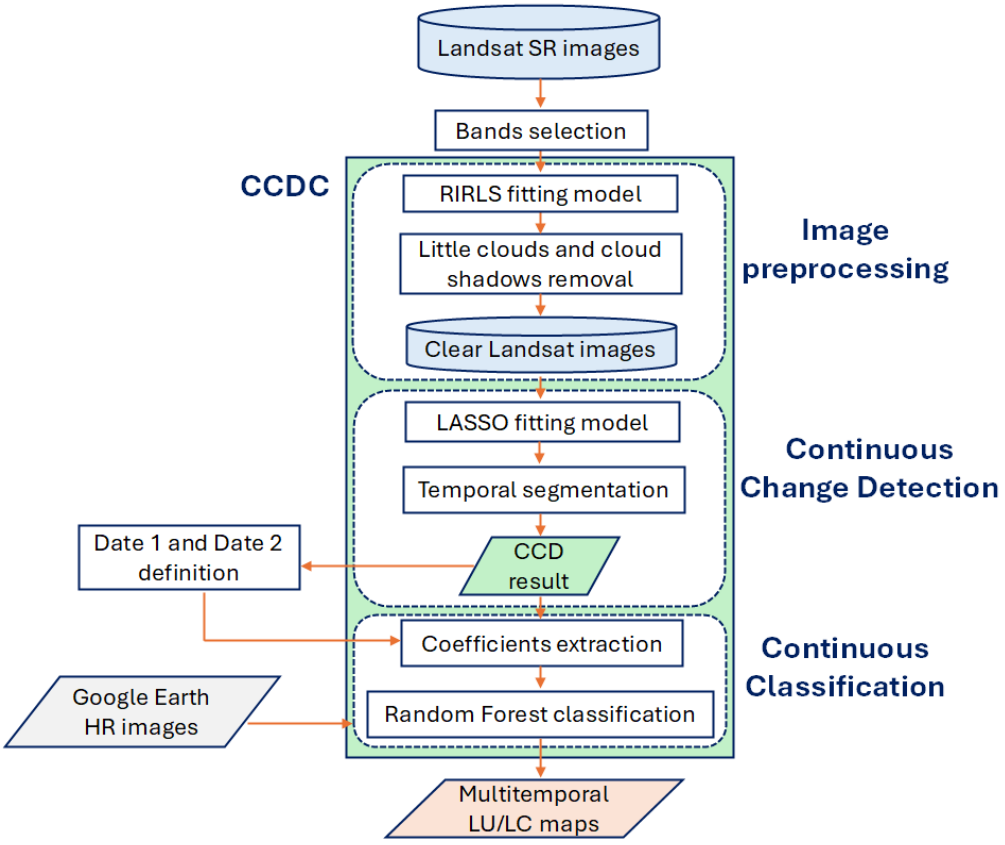


Fig. 4.3 – Flowchart of the CCDC algorithm followed in this study. HR: High-Resolution

In this work, the *TemporalSegmentation.Ccdc()* algorithm (available in the GEE library) was used to implement the CCDC method in GEE.

The first step was to remove small clouds and cloud shadows from the Landsat SR image stack. This is necessary because aerosols and smoke can determine the detection of ephemeral changes, which can easily be confused with real LU/LC changes. For this phase, a time series model was estimated using the Robust Iteratively Reweighted Least Squares (RIRLS) method [95], then outliers are detected and removed from the time series by comparing the model estimations with Landsat observations [29]. The RIRLS fitting model is only estimated using observations in the GREEN and SWIR 1 bands, since clouds are brighter in

the GREEN band, while cloud shadows are darker in the SWIR 1 band, as reported in [29].

Subsequently, the continuous change detection phase identifies breaks in a time series regression model, fitted to the Landsat data for each spectral band (BLUE, GREEN, RED, NIR, SWIR 1 and SWIR 2), by comparing model predictions with satellite observations. For this step, the time series fitting models were obtained from Landsat observations using the Least Absolute Shrinkage and Selection Operator (LASSO) regression [96] (with the value of the parameter λ set equal to 20 to avoid model overfitting and spurious changes detection [31, 97]).

In its general form, the fitting model is expressed by the following equation [31]:

$$\rho^{LASSO}(i,t) = a_{0,i} + a_{1,i}\cos\left(\frac{2\pi}{T}t\right) + b_{1,i}\sin\left(\frac{2\pi}{T}t\right) + c_{1,i}t \quad (6)$$

for $\{\beta_{k-1} < t \leq \beta_k\}$,

where:

$\rho^{LASSO}(i,t)$ is the predicted value for band i at time t;

$a_{0,i}$ is the intercept for the band i, which captures the average reflectance;

$a_{1,i}$ and $b_{1,i}$ are the amplitude and phase for the band i, useful for capturing the intra-annual changes;

$c_{1,i}$ is the slope for the band i, useful for capturing the inter-annual changes;

T is the number of days in a year (365);

β_k is the kth break point.

Model fitting starts at the beginning of the time series and moves forward in time. More specifically, a pixel is flagged as changing if its value is outside a defined range (predicted value ± 3RMSE) for three successive observations.

In particular, the following condition has to be satisfied:

$$\frac{1}{j}\sum_{i=1}^{j}\frac{\left|\rho(i,t)-\rho^{LASSO}(i,t)\right|}{3RMSE_i} > 1\ (three\ times\ consecutively) \quad (7)$$

where:

$\rho(i,t)$ is the observed value of band i at time t;

$RMSE_i$ is the root mean square error computed for band i;

j is the number of Landsat bands (7).

Once a break is found in the time series of a band for a pixel, a new regression model is initialized (Figure 4.4). The process then carries on until the end of the time series.

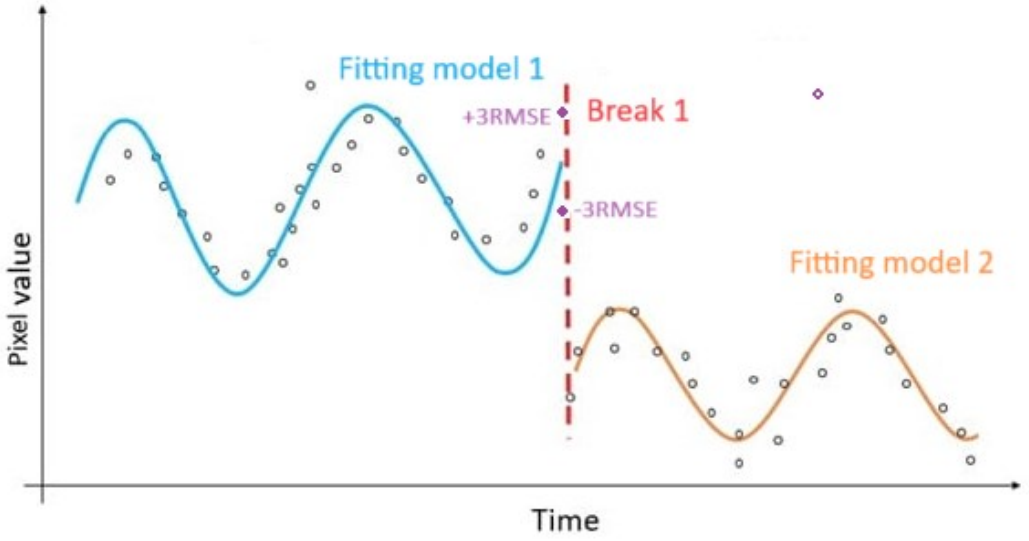


Fig. 4.4 – Illustration of the fitting models (related to two time segments) in their general form for a pixel when a change occurs. Source: [98] (modified)

After the second stage, each pixel has its own time series models before and after each break (i.e. for each time segment). The output of the CCD stage is, in fact, a multiband raster (with multidimensional array bands) that contains several information including the per-pixel model coefficients for each band, as well as the number of time segments, and the start and end dates of each time segment.

In the third step, classified LU/LC maps can be obtained for a defined point in the time series (each pixel has its own model coefficients for the segment covering the chosen time point) by using the model coefficients as inputs for a supervised classification. In this study, the Random Forest classifier (with 100 decision trees) was used to obtain LU/LC maps to be compared for the change detection. The per-pixel intercept coefficients of the model for each spectral band were selected as input data for the classification. Several training sample pixels were selected over the study area, with the aid of high-resolution Google Earth satellite images, belonging to two LU/LC categories: "Urban" (which includes all artificial surfaces such as buildings, roads, parking lots etc.) and "Non-urban" areas (including soil, vegetation and water).

4.2.3. Random Forest classification

Random Forest [99] is a machine learning[5] ensemble classifier formed by a collection of tree-structured classifiers. Given a sample of training data as input, Random Forest uses random subsets of that data through replacement to build multiple decision trees.

More specifically, each tree depends on the values of the random vector (sampled independently), and independently classifies the data, casting a unit vote for the most popular class. For each tree built, data is subsampled from the original full dataset (bagging technique); the samples not used for training (called "out-of-bag" samples) are used for evaluating the classifier's performance, providing an unbiased estimate of the generalization error.

The user-defined parameters are the number of trees and the number of variables per split. The generalization error always converges as the number of trees increases, so there is no overfitting. Therefore, the number of trees can be as large as possible, although beyond a certain value, additional trees won't improve the classification accuracy.

Random Forest is considered one of the best classifiers for LU/LC in urban landscapes because of its robustness to noise and outliers and its high accuracy [99-102].

In this study, 100 number of trees and the square root of the number of variables (as number of variables per split) were set as parameters to perform the Random Forest classification with the *ee.Classifier.smileRandomForest()* function (available in the GEE library), as suggested by [100].

The per-pixel intercept fitted-model coefficients for each spectral band used in the CCDC algorithm, and extracted from the dates of interest time segments, were selected as predictive variables for the Random Forest classifier and the "Urban/Non-urban maps" for the reference years were produced for each study site.

To construct city-specific Random Forest models for the regional capital cities of Italy, a consistent number of training sample pixels were selected in each study area (totaling 17390 pixels considering all cities), with the help of Google Earth high resolution historical images. These pixels were picked up in a sufficient number, within homogeneous areas and in different zones [13-14] of each territory. In the adopted classification scheme, "Urban areas" includes all artificial surfaces, while "Non-urban areas" includes soil, vegetation and water.

4.2.4. Change detection

After obtaining two binary LU/LC maps ("Urban/Non-urban maps") for the two dates of interest with the Random Forest classification, they were compared pixel-by-pixel. In this manner, the possible outputs were "no change" (for the same classification in both dates) or "change" (for pixels classified differently in the two maps).

Once the post-classification comparison was performed, for pixels changed from "Non-urban" to "Urban" (representing the urban growth) a mode filter to refine the change maps from the "salt and pepper" noise was applied and a fi-

nal “urban change map” (depicting two distinct categories: “Urban growth” and “No growth”) was produced for each case study.
The mode filter is a post-classification spatial filtering technique applied to every pixel of a map using a moving window (named “kernel”) of a certain size. The kernel is moved throughout the original classified map and the categorical value (i.e. the label of the LU/LC class) of the center pixel of the window is changed to the most frequent class within that window [14, 103]. In this work, a kernel of 3x3 pixels was chosen to apply the mode filter.

4.2.5. Accuracy assessment

For the accuracy assessment of the change map obtained for the site of Bari (2015-2023), the same datasets of 220 test pixels, generated with the stratified random sampling method and used in Chapter 3 (20 pixels for the “Urban Growth” class, and 200 pixels for the “No growth” category), was used.
For the accuracy assessment of the change maps produced for the Italian regional capital cities (2006-2023), instead, a multitemporal datasets of several test pixels (in proportion 1: 11 between “Urban growth” and “No growth” classes) were chosen for each municipality with the stratified random sampling method, after visual inspection of multitemporal high-resolution Google Earth images.
The accuracy metrics calculated were O.A., P.A. and U.A.

4.2.6. Urban growth analysis

In order to quantify and analyze the increase of urban areas, occurred in Bari between 2015 and 2023, the same metrics reported in Table 3.6 were calculated.
Regarding the urban growth analysis for the regional capital cities of Italy (2006-2023), instead, the following metrics (Table 4.3) were computed:

Tab. 4.3 – Description of the urban growth metrics calculated for the Italian regional capital cities (2006-2023).

Metric	Description
Growth of Urban areas ($GU_{2006\text{-}2023}$) [Km^2]	Urban growth from 2006 to 2023 (in Km^2) estimated with the CCDC method
Rate of Urban Growth ($RGU_{2006\text{-}2023}$) [%]	Rate of Urban growth (in percentage), related to urban areas present in the 2015 map (obtained with the CCDC method) $RGU_{2006\text{-}2023} = ((GU_{2006\text{-}2023})/ \text{Urban area}_{2015}) *100$
Average annual Urban Growth (AaGU) [Km^2/year]	$AaGU = (GU_{2006\text{-}2023})/N_{YEARS}$

4.2.7. Overlay between urban growth map and hydraulic hazard map for the Bari territory

For the study area of Bari, after extracting the urban growth map (2006-2023) with the CCDC method, it was overlaid with the map of the hydraulically hazardous areas, as delimited by the Basin Authority of Apulia Region (updated 2020) and provided by the Italian National Institute for Environmental Protection and Research (ISPRA) [104]. This was done in order to verify whether urban development has taken place in protected areas according to the current planning rules.

Hazard (P), understood as a characteristic of a territory that makes it vulnerable to phenomena of instability, and which is defined as the probability of the occurrence of a calamitous event with a defined temporal frequency and specific intensity, is one of the elements that contribute to the definition of environmental risk (R).

More specifically, risk is defined as the magnitude of the expected damage following the occurrence of a given event, in a given time interval and in a given area. In addition to hazard, risk is a function of vulnerability (V), which is the expected degree of loss (ranging from 0 to 1) for a given element depending on the intensity of the event, and exposure value (E), which is the number of people or the value of natural or economic resources exposed to a given hazard [73].

With regard to hydraulic hazard, three types of protected areas are defined [73]:

- Areas with high hydraulic hazard: areas subjected to flooding due to flood events with a return period ≤ 30 years;
- Areas with medium hydraulic hazard: areas subjected to flooding due to flood events with a return period of between 30 and 200 years;
- Areas with low hydraulic hazard: areas subjected to flooding due to flood events with a return period of between 200 and 500 years.

4.3. Results for Bari (2015-2023)

The "Urban/Non-urban" LU/LC maps generated using the CCDC approach for the area of Bari (2015-2023) are reported in Figure 4.5.

The urban growth map retrieved after the "post-classification comparison" of these maps and the mode filter, as explained before, is shown in Figure 4.6. Table 4.4, instead, reports the accuracy metrics of the error matrix relative to the urban change map (2015-2023) (Fig. 4.6).

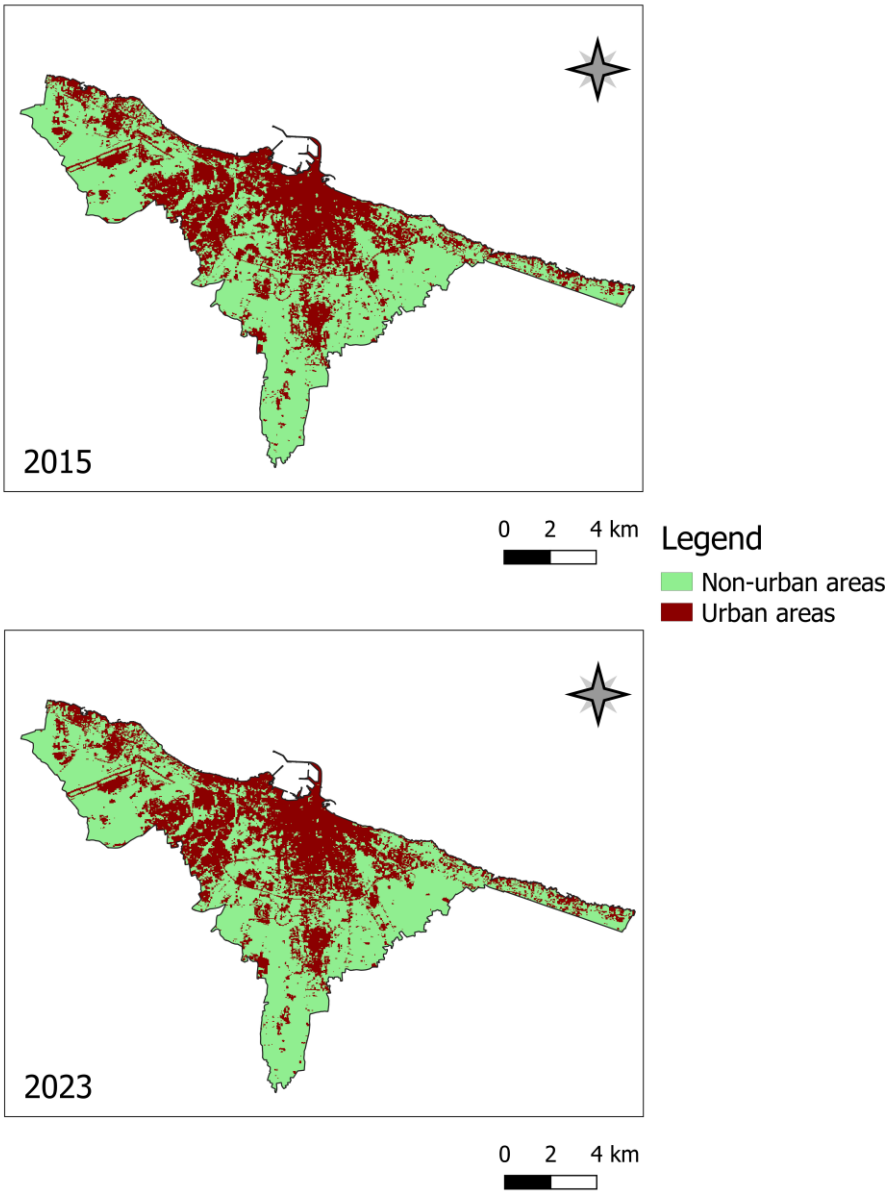


Fig. 4.5 – “Urban/Non-urban” maps produced with the CCDC method for the study area of Bari for the reference years 2015 and 2023

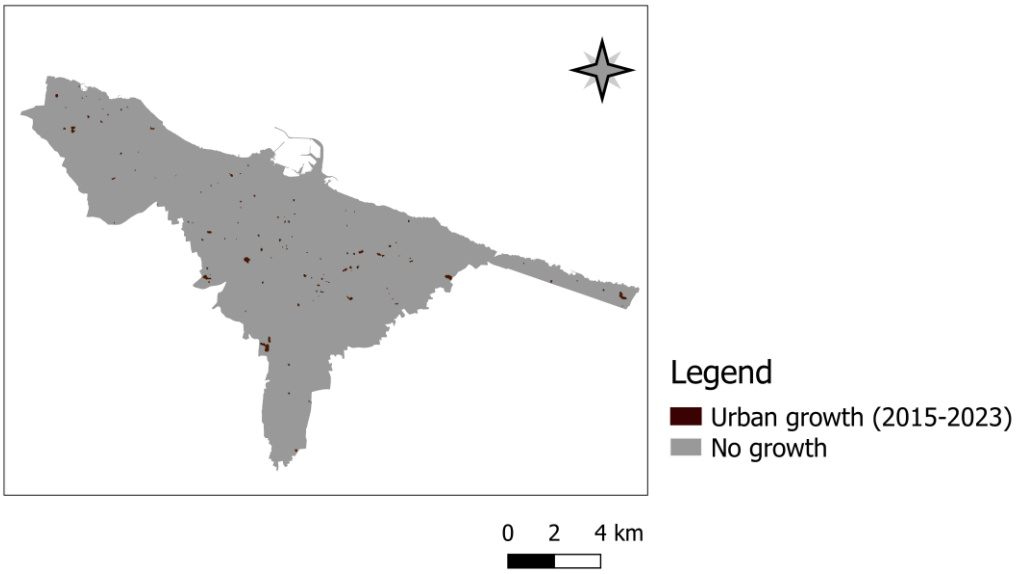


Fig. 4.6 – Urban change map estimated with the CCDC method for the period 2015-2023 for the study area of Bari

Tab. 4.4 – Accuracy metrics for the urban change map (2015-2023) for the area of Bari obtained with the CCDC method.

	2015-2023	
Class	**P.A. [%]**	**U.A. [%]**
Urban growth	80	100
No growth	100	98
O.A. [%]	98	

From Table 4.4 it is possible to notice that very high accuracies characterize the urban growth map obtained through the application of the CCDC method. The O.A. is very high (totaling 98%) while the P.A. and the U.A. for the "Urban growth" category are 80% and 100%, respectively.

Table 4.5 shows the urban growth metrics estimated for Bari, which quantify the land consumption estimated with this method.

Tab. 4.5 – Urban growth metrics calculated for Bari (2015-2023) using the CCDC method.

Metric	Value
$GU_{2015\text{-}2023}$ [Km^2]	0.8
$RGU_{2015\text{-}2023}$ [%]	+1.8
AaGU [Km^2/year]	0.1

From Table 4.5 it is possible to notice that the land consumption from 2015 to 2023 (Figure 4.6) estimated using the CCDC method for Bari is equal to 0.8 Km^2. The rate of urban growth in this period (8 years), instead, was found to be +1.8%, with an average annual urban growth of 0.1 Km^2 per year.

4.4. Results for the Italian regional capital cities (2006-2023)

The urban growth maps (2006-2023) for the twenty regional capital cities of Italy, produced using the CCDC method in GEE environment, are illustrated in Figures from 4.7 to 4.13, in which cities were grouped according to their size.

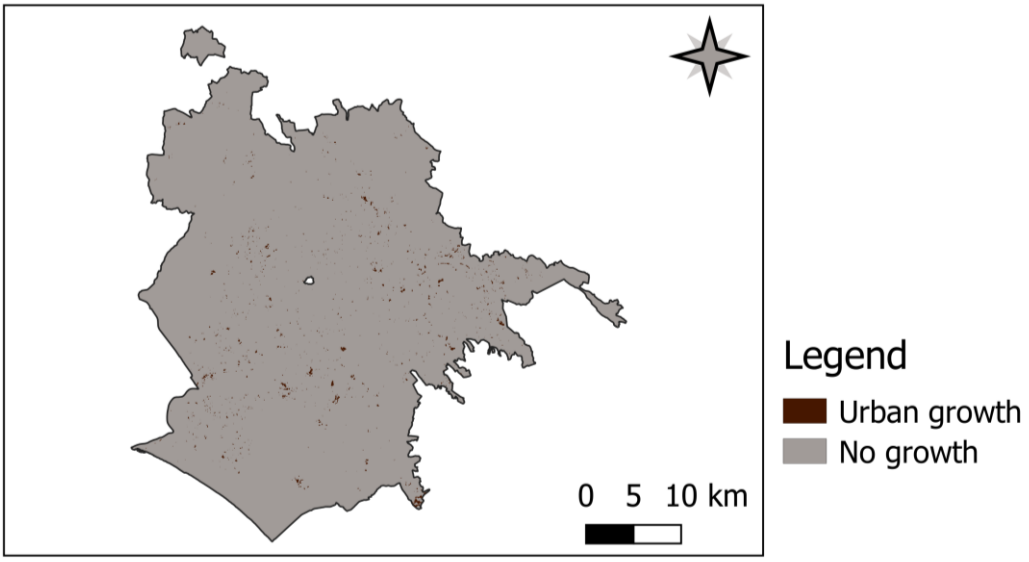


Fig. 4.7 – Urban change map estimated with the CCDC method for the period 2006-2023 for the study area of Roma

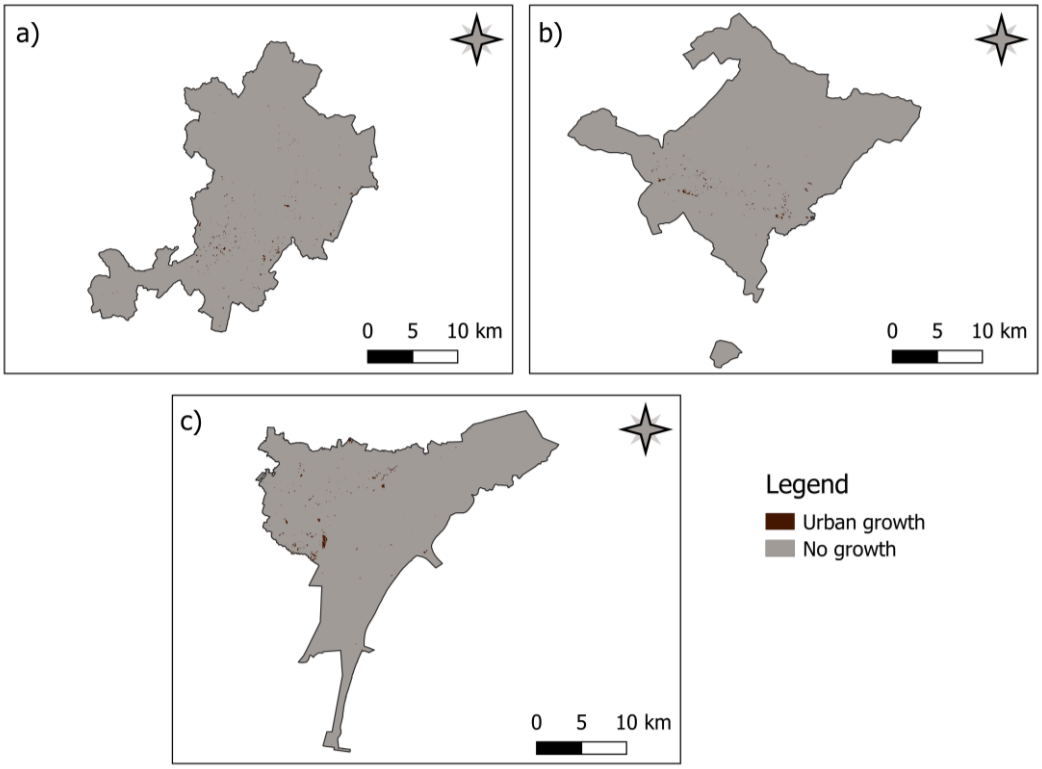


Fig. 4.8 – Urban change maps estimated with the CCDC method for the period 2006-2023 for the study areas of Perugia (a), L'Aquila (b) and Venezia (c)

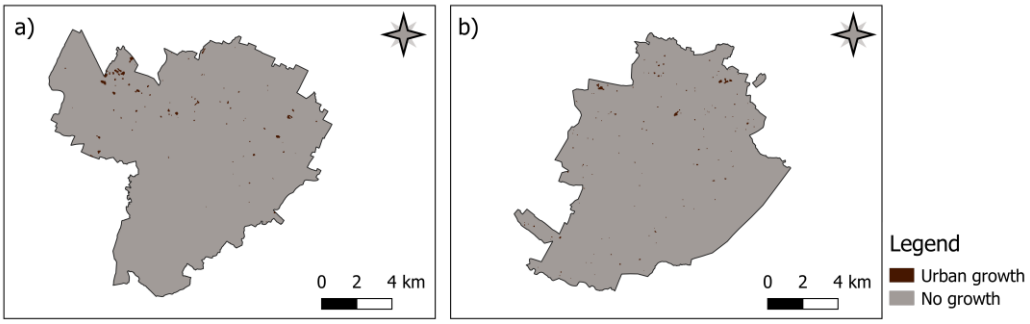


Fig. 4.9 – Urban change maps estimated with the CCDC method for the period 2006-2023 for the study areas of Bologna (a) and Torino (b)

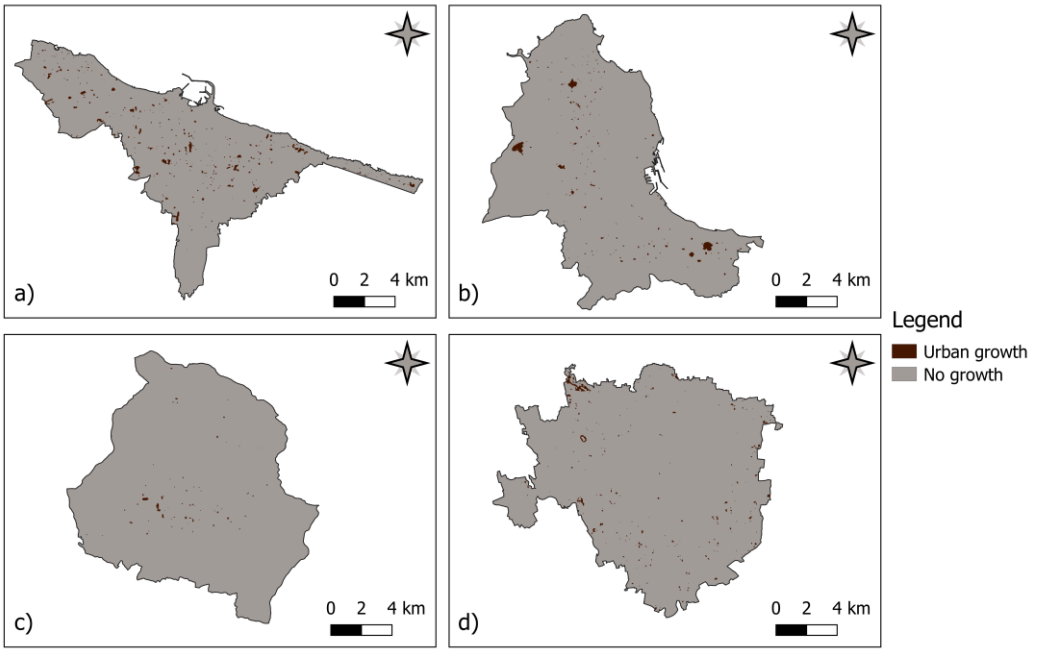


Fig. 4.10 – Urban change maps estimated with the CCDC method for the period 2006-2023 for the study areas of Bari (a), Palermo (b), Potenza (c) and Milano (d)

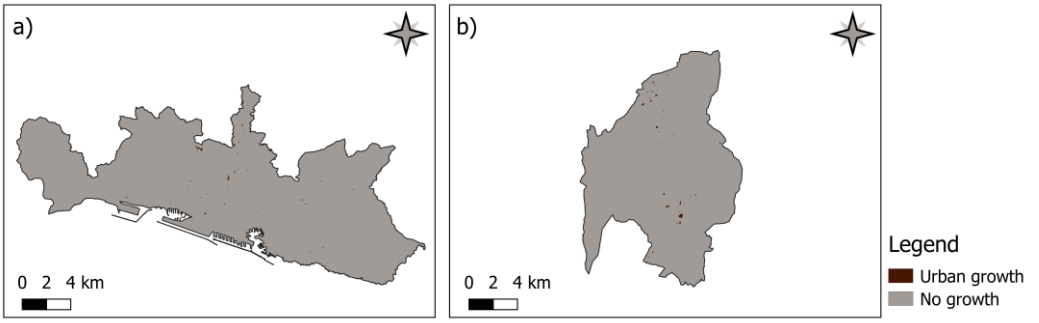


Fig. 4.11 – Urban change maps estimated with the CCDC method for the period 2006-2023 for the study areas of Genova (a) and Trento (b)

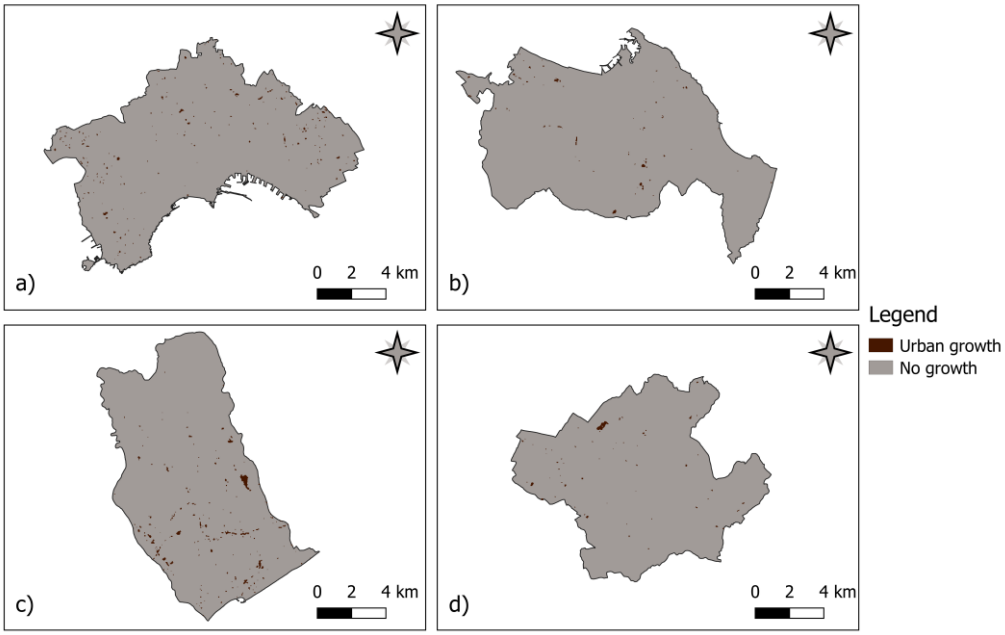


Fig. 4.12 – Urban change maps estimated with the CCDC method for the period 2006-2023 for the study areas of Napoli (a), Ancona (b), Catanzaro (c) and Firenze (d)

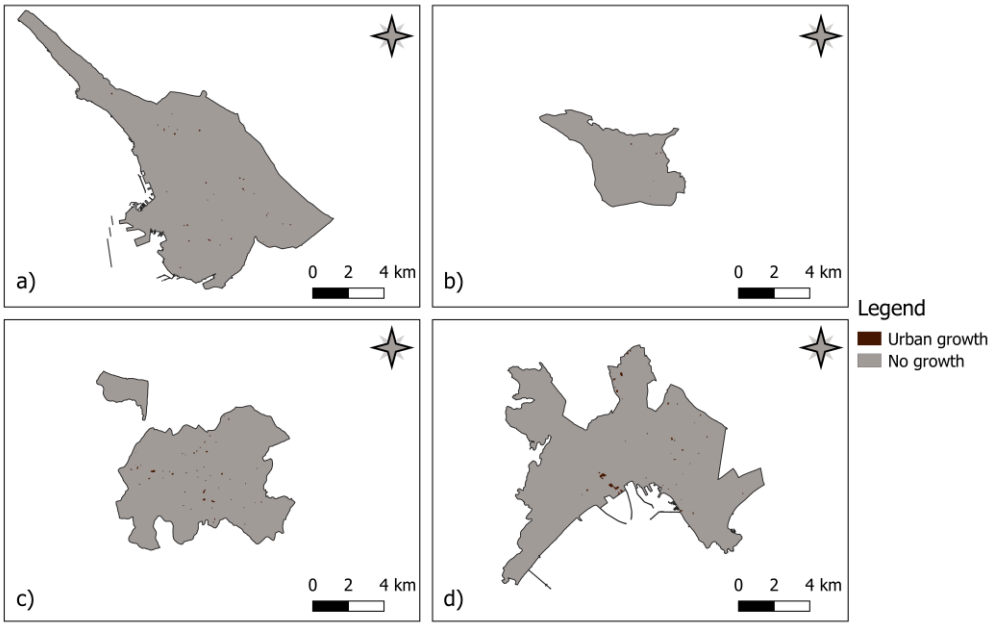


Fig. 4.13 – Urban change maps estimated with the CCDC method for the period 2006-2023 for the study areas of Trieste (a), Aosta (b), Campobasso (c) and Cagliari (d)

The accuracy metrics obtained for the maps in Figures from 4.7 to 4.13 are reported in Table 4.6, while the urban growth metrics for each city (2006-2023), which allow to analyze the outcomes and the urban expansion dynamics, are instead highlighted in Table 4.7.

Tab. 4.6 – Accuracy metrics for the urban change maps (2006-2023) obtained for the regional capital cities of Italy with the CCDC method.

City	**O.A. [%]**	**PA [%]**		**UA [%]**	
		Urban growth	No growth	Urban growth	No growth
Ancona	96	70	99	82	97
Aosta	90	43	95	43	95
Bari	98	84	99	89	99
Bologna	97	80	99	84	98
Cagliari	96	70	99	82	97
Campobasso	96	70	99	82	97
Catanzaro	97	73	99	85	98
Firenze	98	85	99	85	99
Genova	95	60	98	75	96
L'Aquila	97	80	99	86	98
Milano	96	67	98	77	97
Napoli	97	83	98	81	98
Palermo	98	88	99	85	99
Perugia	96	70	99	85	97
Potenza	96	80	98	76	98
Roma	97	88	98	79	99
Torino	94	65	97	65	97
Trento	96	80	98	76	98
Trieste	94	60	97	67	96
Venezia	96	73	98	76	98

From Table 4.6 it is possible to notice that very high accuracies were achieved for the change maps. In fact, the O.A. is higher than 94% for 19 out of 20 cases (only the map of Aosta exhibits an accuracy of 90%, probably due to the very low land consumption occurred for this territory).

Tab. 4.7 – Urban growth metrics calculated for the Italian regional capital cities (2006-2023) using the CCDC method.

City	$GU_{2006\text{-}2023}$ [Km^2]	$RGU_{2006\text{-}2023}$ [%]	AaGU [Km^2/year]
Ancona	0.4	+2.75	< 0.1
Aosta	< 0.1	+0.40	< 0.1
Bari	2.6	+6.35	0.2
Bologna	0.8	+2.26	< 0.1
Cagliari	0.4	+2.09	< 0.1
Campobasso	0.2	+2.64	< 0.1
Catanzaro	1.4	+11.38	0.1
Firenze	0.5	+1.50	< 0.1
Genova	0.3	+0.78	< 0.1
L'Aquila	1.6	+8.51	0.1
Milano	1.9	+2.13	0.1
Napoli	1.0	+1.34	0.1
Palermo	1.8	+3.31	0.1
Perugia	2.0	+6.55	0.1
Potenza	0.4	+4.26	< 0.1
Roma	10.1	+4.26	0.6
Torino	0.6	+0.85	< 0.1
Trento	0.4	+2.42	< 0.1
Trieste	0.2	+0.67	< 0.1
Venezia	2.6	+5.48	0.2

The highest O.A. (98%) were found for the sites of Bari, Firenze and Palermo. The P.A. were also high (always ≥ 95% for the "No growth" category), with a value ranging from 65% to 88% for the "Urban growth" class for almost all case studies (the highest values of 88% were obtained for Palermo and Roma). On the contrary, the lowest values of P.A. for the "Urban growth" class were found for Aosta (43%), Trieste (60%) and Torino (65%). The U.A. were quite high, with values ≥ 95% for each study areas for the "No growth" category, while for the "Urban growth" class, values ≥ 75% were obtained for almost all cities, where the highest values characterizing the maps of Bari (89%) and L'Aquila (86%).

Table 4.7 shows that, throughout the investigated period, the city with the highest land consumption is Roma (around 10 Km^2), which is also the largest metropolitan area in Italy (Table 4.2). Other municipalities with higher urban growth are Bari (2.6 Km^2), Venezia (2.6 Km^2), Perugia (2.0 Km^2) and Milano (1.9 Km^2). On the other hand, the cities with the lowest land take are Aosta (< 0.1 Km^2), Trieste (0.2 Km^2), Campobasso (0.2 Km^2) and Cagliari (0.4 Km^2).

The average annual growth of artificial areas was found to be ≤ 0.2 Km^2/year for all cities, with the exception of Rome for which an increase of 0.6 Km^2 per year was detected.

From Table 4.7 it is possible to notice that the municipality with the highest rate of urban growth is Catanzaro (+11.38%) followed by L'Aquila (+8.51%), Perugia (+6.55%) and Bari (+6.35%). Very low rates were detected instead for Aosta (+0.40%), Trieste (+0.67%) and Genova (+0.78%).

4.4.1. Further analysis for Bari

As reported previously, for the area of Bari the urban growth map (2006-2023), obtained with the CCDC method, was overlaid with the map of hydraulic hazards [104], as delimited by the Basin Authority of the Apulia Region. The result of this overlay is illustrated in Figure 4.14.

From this figure, it can be seen that most of the new artificial areas were built outside the protected areas. Only one new impervious surface in the eastern part of the city seems to overlap with vulnerable areas (in particular with high and medium hydraulic hazard areas): it is highlighted in Figure 4.15, where a Google Earth high-resolution image was used to inspect this area and assess whether it corresponds to a new built-up area. From Figure 4.15 it can be noticed that this site doesn't correspond to a new building but appears to be a temporary working site where reversible land consumption has occurred (it is visible the removal of vegetation and compaction of the soil).

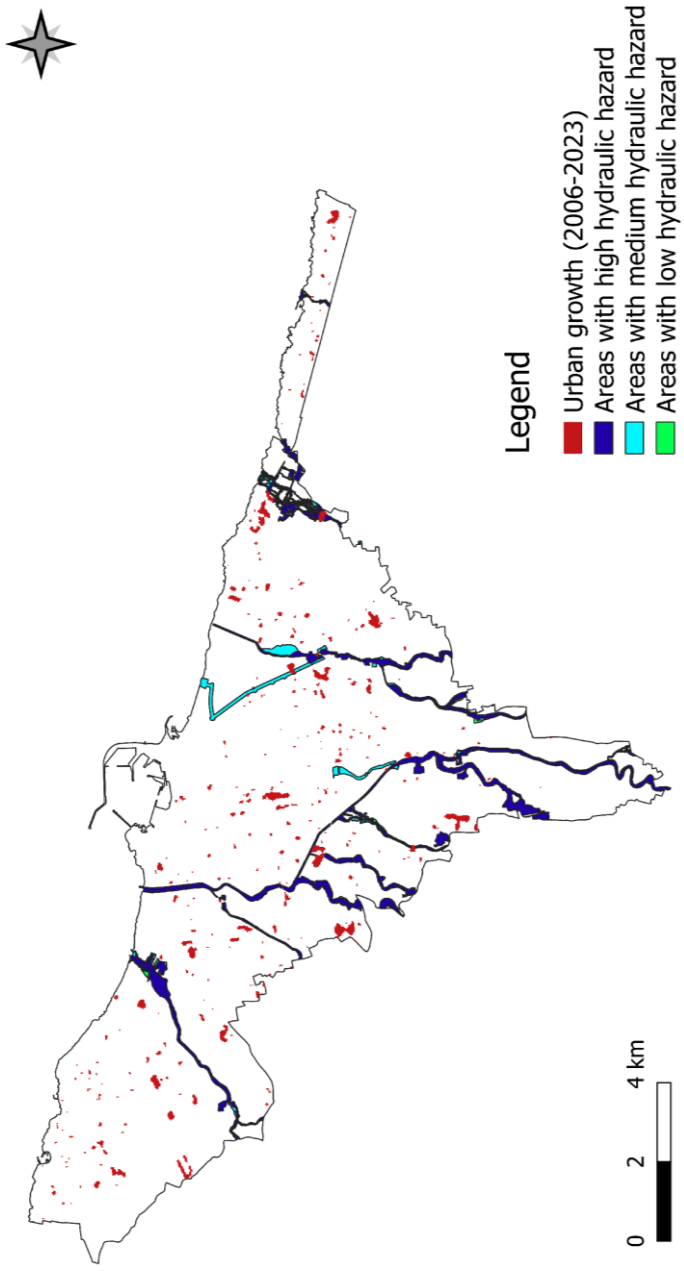

Legend
Urban growth (2006-2023)
Areas with high hydraulic hazard
Areas with medium hydraulic hazard
Areas with low hydraulic hazard
0
2
4 km

Fig. 4.14 – Urban growth map (estimated with the CCDC method for the period 2006-2023) overlaid with the hydraulic hazard map [104] (updated 2020) for Bari

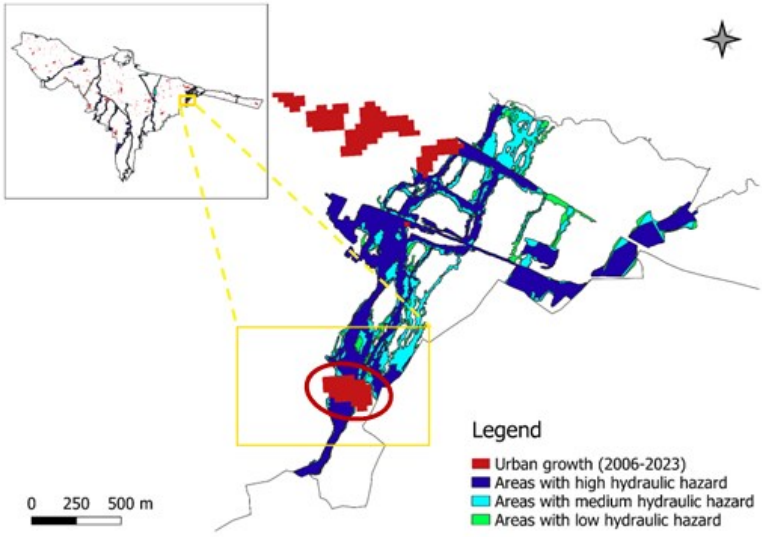


Fig. 4.15 – Detail of a new artificial surface that overlap with hydraulic hazard areas in Bari

5. DISCUSSION

The aim of this study is to estimate land consumption in different urban contexts using multiple change detection techniques applied to Landsat satellite data. For this purpose, two algorithms have been implemented in the GEE cloud platform. The first (explained in Chapter 3), applied to Landsat 8 imagery, involves a post-classification comparison between multidate LU/LC maps (produced with a novel index-based classification algorithm developed for this research, which adopts LICA's STRed and SwiRed indices [75], NIR band and NIR/SWIR2 ratio) and land surface albedo change analysis. The second (described in Chapter 4) is related to the implementation of the CCDC algorithm [29], which makes use of all available satellite images and allows to obtain multidate LU/LC maps by using per-pixel model coefficients as predictive variables for a defined classifier (in this study, the "Random Forest" classifier).

Both methods gave very satisfactory results in terms of the accuracy of the "Urban growth" maps. The "post-classification and albedo change analysis" method, proposed for the test area of Bitritto (Apulia region in southern Italy) [11], also yielded very good results when applied to the case study of Bari (2015-2023), the capital city of Apulia region and belonging to the same geographical, climatic and geomorphologic context as Bitritto. The application of this approach allowed to estimate the land consumption map for Bari (2015-2023) with a P.A. and a U.A. for the "Urban growth" category of 80% and 94% respectively, and an O.A. of 98%. Furthermore, the urban growth of Bari was quantified at about 1.1 Km^2 (with an average growth of 0.1 Km^2 per year), with a growth rate of artificial surfaces of +2.3% between 2015 and 2023.

On the other hand, the CCDC algorithm was used to estimate land consumption for Bari (from 2015 to 2023, in order to compare the results with those of the first method) and for all twenty Italian regional capitals from 2006 to 2023.

The results for Bari (2015-2023) were very promising. Indeed, a P.A. of 80% (same results as the first method) and a U.A. of 100% (higher performance compared to the first method) were found for the "Urban growth" class, with an O.A. of 98%. Moreover, this approach allowed a slightly lower estimation of the land consumption occurred (0.8 Km^2, and +1.8% of new impervious surfaces between 2015 and 2023) compared to the "post-classification comparison and albedo change analysis" algorithm, which could be due to the different technique or the post-classification filter (the isolated pixel filtering in the first method and the mode filter in the second).

However, both approaches underestimated the changes that actually occurred for two main reasons: i) the geometric resolution of 30 m of the Landsat imagery hampered the ability of the methods to detect minor changes present in some areas (e.g. newly built small residential complexes) due to the "mixed pixels" issue, typical of pixel-based medium-resolution image classification; ii) the filtering technique used to refine the results from noise (the "salt and pepper" effect, which is also typical of change detection from two pixel-based LU/LC classifications). From the results, it can be inferred that the mode filter used for the CCDC method, which replaces the label of the central pixel of the kernel with the most frequent class in the same kernel, could have led to an estimate of 0.3 Km^2 less land loss than the other technique.

For the Italian regional capitals (2006-2023), the CCDC methods produced accurate results. In particular, the "Urban growth" class in the change maps showed P.A. values higher than 65% for most of the study sites, while its U.A. values were found to be higher than 75% in many cases. This outcome indicates the good applicability of this method for the entire Italian territory.

In absolute terms, the city of Roma experienced the highest urban growth in Italy (estimated about 10 km^2 and 0.6 Km^2 per year), while Aosta registered the lowest growth (less than 0.1 Km^2). In particular, the result for Roma is in line with other studies [4, 104-106], such as those carried out by ISPRA as part of

the Soil4Life project (https://soil4life.eu/) which underlined the highest land grab in Rome compared to other Italian cities (about 90 hectares per year).
In relative terms, instead, the highest rate of land consumption was found in the municipality of Catanzaro (about +11% of new urban land areas compared to those existing in 2006, partly due to a new quarry in the central-eastern part of the territory), followed by L'Aquila (about +8.5%, where new residential areas were built after the earthquake in 2009 [107]) and Perugia (about +6.5%, as underlined also by ISPRA [4, 108]). The lowest rate (+0.40%) was recorded in Aosta.
With regard to the city of Bari, it resulted one of the Italian cities with the highest growth in built-up areas in Italy during the period investigated, as highlighted by other studies [4, 109-110], which underlined its high land consumption rate despite the city's lack of demographic growth and the emerging need for urban regeneration (which would increase the number of parks and other green areas) in order to counteract the adverse effects of climate change (such as the increase in the urban heat island effect and extreme events) and to achieve the goals of the UN Agenda 2030 (https://sdgs.un.org/2030agenda).
Furthermore, concerning the case study of Bari, the overlap between the urban growth map (2006-2023) and the map of the hydraulically hazardous areas, which were delimited for the Hydro-geological Asset Plan of the Apulia Region (adopted for the first time in 2005 [73]), allowed to verify the effectiveness of the institution of the protected areas in avoiding the increase of the flood risk.
Given the possibility of extracting the urban growth maps for different areas in a relatively short time, the CCDC method could be considered more flexible and faster than the "post-classification and albedo change method", which requires the calculation of multidate albedo maps for the specific region of interest (depending on local variables such as P_0, W and Z, which are used to estimate atmospheric transmissivity) and has been tested in this study only for the areas of Bari and Bitritto (which are in the same geographical context) and using only Landsat 8 data. Further analysis is needed to verify the applicability of the

method to other study areas in Italy and using data from other Landsat missions. In this perspective, it should be underlined that the algorithm of da Silva et al. (2016), used for the albedo estimation, was developed specifically for Landsat 8 data, which have some differences in the spectral ranges of the different bands compared to previous Landsat missions (see Table 4.1). However, the promising results described in Chapter 3 for the case studies of Palermo and Cagliari, where the algorithm of da Silva et al. (2016) was applied to retrieve albedo maps also from Landsat 9 images, providing very similar estimates compared to those obtained from Landsat 8 data, suggest that the "post-classification and albedo change method" could be reasonably applied in the future to contribute to urban growth assessment, at least also using Landsat 9 data.

It is also important to add that the high accuracies obtained by the two employed methods, which allowed a detailed mapping of the impervious surfaces and a valuable detection of the location of the land consumption for each case study, suggest the inclusion of further analyses, to be carried out in the future, that could be useful to effectively evaluate the possible consequences of land consumption on the environment and territory. In particular, accurate information on urban growth may be essential for estimating its negative impacts on soil ecosystem services and biodiversity (by calculating metrics such as the "potential impact area", the "degree of landscape fragmentation" or the "loss of agricultural land", among others [4]) or for studies on the potential increase of the urban heat island effect (which consists of the difference between the temperature of urban areas and the temperature of surrounding non-urban areas) caused by the expansion of urban areas, as Landsat data also provide information on land surface temperature (LST) at 30 m resolution [4, 60, 88].

The analyses carried out in this study were possible thanks to the GEE cloud environment, a free-to-use environment for research purposes, where different JavaScript codes were written in the Code Editor. GEE allowed the processing of satellite data with a reduced operational and computational time compared to

traditional desktop software, thanks to the presence of the Landsat archive in the GEE data catalog and to the cloud technology, in particular with respect to the CCDC method, which involves the processing of a stack of images and requires a large number of computations to be performed.

CONCLUSIONS

Mapping and monitoring the growth of impervious areas is a useful way to pursue sustainable development and to measure the effectiveness of current planning regulations in preventing an increase of environmental risks. EO remote sensing techniques play a crucial role in providing continuous information on changes in the land surface at large scales.

The aim of this study was to employ different techniques, including some innovative elements, to estimate land consumption occurred in different urban contexts using Landsat satellite data and the GEE cloud platform. Both methods (the "post-classification comparison and albedo change analysis" and the "Continuous Change Detection and Classification") provided very satisfactory results, for each case studies, in terms of the accuracy of the estimated urban growth maps and level of detail. Nevertheless, the geometric resolution of the Landsat images (30 m) and the filtering techniques, used to refine the maps from noise, affected the ability of these approaches to retrieve small changes, due to the "mixed pixels" issue or the removal of some correctly detected changes, respectively. Therefore, the results presented in this work, when compared with the analysis of high-resolution images or other literature studies, slightly underestimate the landscape variations that really occurred over time in the investigated areas.

Furthermore, the mapping and monitoring of urban growth using EO remote sensing techniques, when combined with the analysis with other informative layers, such as hazard or risk maps, proved to be a valuable tool for the evaluation of current sustainable development strategies and environmental and territorial risk management.

Lastly, the GEE platform allowed for rapid processing of EO big data, paving the way for new possibilities in environmental studies requiring high computational capacities.

NOTES

1. The impervious surfaces are constructed surfaces such as sidewalks, roads, runways, rooftops and parking lots, which are covered by impermeable materials such as concrete, asphalt and stone. These materials seal surfaces, preventing precipitation and meltwater from infiltrate into the soil (Cfr. Bhatta, B., 2010a).

2. The stratified random sampling method ensures that a minimum number of samples are randomly allocated to each land cover class after the classified map has been produced. A “stratum” is obtained by extracting only pixels associated with a specific class, and the sample pixels are, then, randomly distributed throughout the strata (Cfr. Jensen, J.R., 2015).

3. The accuracy metrics of the error matrix are defined as follows: overall accuracy is the sum of the correctly classified sample units to the total number of sample units in the error matrix. Producer’s accuracy is obtained by dividing the total number of correctly classified sample units in a category to the total number of sample units in that category as indicated by the reference data. User’s accuracy is, instead, obtained by dividing the total number of correctly classified sample units in a category to the total number of sample units classified as belonging to that category in the map (Cfr. Congalton, J.R. and Green, K., 2009).

4. The Worldwide Reference System-2 is a global notation system for Landsat data. It allows a user to search satellite imagery over any part of the world by specifying a nominal scene center, designated by PATH and ROW numbers. The WRS-2 PATH number is sequentially assigned to 233 nominal satellite orbital tracks, from east to west, while the WRS-2 ROW number is the latitudinal center line of an image frame (there are 248 row intervals per complete orbit). The combination of a PATH number and a ROW number

uniquely identifies a nominal scene center (Cfr. National Aeronautics and Space Administration, 2024b).

5. Machine Learning is a branch of Artificial Intelligence that consists of an empirical approach to both classification and regression of non-linear systems that can involve from a few to thousands of variables. Machine Learning algorithms learn the behavior of a system from a set of training data without any prior knowledge of the nature of the relationships between the data. Machine Learning is ideal for solving those problems for which our theoretical knowledge is still incomplete, but for which we have a significant number of observations and other data at our disposal (Cfr. Lary, D.J., Alavi, A.H., Gandomi, A.H. and Walker, A.L., 2016).

ACKNOWLEDGEMENTS

I am grateful to my tutor, Prof. Eufemia Tarantino, for providing me with the opportunity to pursue a doctoral degree in a field of study that I am truly passionate about, and for her timely guidance and supervision. I would also like to express my gratitude to Dr. Alessandra Capolupo, Researcher at Polytechnic of Bari, for her valuable suggestions and for assisting me in maintaining an organized approach to my work.

I am greatly thankful also to Prof. Ionut Sandric of the University of Bucharest – Faculty of Geography, for his kind welcome and support during my period of study and research in Romania. His guidance and inspiration have been pivotal to my academic pursuits and instilling confidence in my abilities. His valuable suggestion to employ the "CCDC" method has proved to be a significant contribution to my thesis.

Lastly, I would like to thank Dr. Nikos Papadopoulos and Dr. Nasos Argyriou, researchers at the Institute for Mediterranean Studies – Foundation for Research and Technology Hellas, for kindly welcoming me during my secondment at their institution under the EU-funded COASTLINE project and for encouraging the development of new ideas in my work.

LIST OF FIGURES

LIST OF TABLES

BIBLIOGRAPHY

1. Peroni, F. *et al.* (2022) 'How to map soil sealing, land take and impervious surfaces? A systematic review'*, Environmental Research Letters*, 17(5), 053005.
2. Valera Lozano, A., Añó Vidal, C. and Sánchez Díaz, J. (2019) 'Urban growth (1956-2012) and soil sealing in the metropolitan area of Valencia (Eastern Spain)', *Spanish Journal of Soil Science*, 9(2), pp. 88-104.
3. European Commission (2012) *Guidelines on best practice to limit, mitigate or compensate soil sealing*. Publications Office of the European Union. Available at: https://data.europa.eu/doi/10.2779/75498.
4. Munafò, M. (2022) *Consumo di suolo, dinamiche territoriali e servizi ecosistemici*. Edizione 2022. Report SNPA 32/22.
5. Bhatta, B. (2010a) 'Urban Growth and Sprawl' in Bhatta, B. (ed.) *Analysis of Urban Growth and Sprawl from Remote Sensing Data*, Advances in Geographic Information Science. Springer-Verlag Berlin Heidelberg, pp. 1-16.
6. Barletta, C., Capolupo, A. and Tarantino, E. (2023) 'Integration of Copernicus Data and Services to Assess Local Aridity Conditions in the Apulian Context: The Case of Marina di Ginosa' in Gervasi, O. *et al.* (eds.) *Computational Science and Its Applications – ICCSA 2023 Workshops*, LNCS 14107. Springer Cham., pp. 201-216.
7. Manfreda, S., Iacobellis, V. and Fiorentino, M. (2010) *Appunti di idrologia superficiale*. 1st edn. Aracne editrice, Roma.
8. FISRWG (Federal Interagency Stream Restoration Working Group) (1998) *Stream Corridor Restoration: Principles, Processes, and Practices*.
9. Weng, Q. (2008) 'Remote Sensing of Impervious Surfaces: An Overview' in Weng, Q. (ed.) *Remote Sensing of Impervious Surfaces*, Taylor & Francis Series in Remote Sensing Applications. CRC Press, pp. xv-xxvii.

10. Bhatta, B. (2010b) 'Remote Sensing, GIS, and Urban Analysis' in Bhatta, B. (ed.) *Analysis of Urban Growth and Sprawl from Remote Sensing Data*, Advances in Geographic Information Science. Springer-Verlag Berlin Heidelberg, pp. 49-63.
11. Barletta, C., Capolupo, A. and Tarantino, E. (2024) 'Estimating Urban Growth from Landsat 8 Data Using Post-classification and Albedo Change Analysis in GEE Environment' in Gervasi, O. *et al.* (eds.) *Computational Science and Its Applications – ICCSA 2024 Workshops*, LNCS 14819. Springer Cham., pp. 185-200.
12. Kabisch, N., Selsam, P., Kirsten, T., Lausch, A. and Bumberger, J. (2019) 'A multi-sensor and multitemporal remote sensing approach to detect land cover change dynamics in heterogeneous urban landscapes', *Ecological Indicators*, 99, pp. 273–282.
13. Brivio, P., Lechi, G. and Zilioli, E. (2006) *Principi e metodi di telerilevamento*. CittaStudi.
14. Lillesand, T.M., Kiefer, R.W. and Chipman, J.W. (2015) *Remote sensing and image interpretation*. 7th Edn. Wiley.
15. Singh, A. (1989) 'Review Article Digital change detection techniques using remotely-sensed data', *International Journal of Remote Sensing*, 10(6), pp. 989-1003.
16. Vázquez-Jiménez, R., Ramos-Bernal, R.N., Romero-Calcerrada, R., Arrogante-Funes, P., Sanchez Tizapa, S. and Novillo, C.J. (2018) 'Thresholding Algorithm Optimization for Change Detection to Satellite Imagery' in Travieso-Gonzalez, C.M. (ed.) *Colorimetry and Image Processing*. InTech. Available at: http://dx.doi.org/10.5772/intechopen.71002.
17. Bhatta, B. (2010c) 'Mapping and Monitoring Urban Growth' in Bhatta, B. (ed.) *Analysis of Urban Growth and Sprawl from Remote Sensing Data*, Advances in Geographic Information Science. Springer-Verlag Berlin Heidelberg, pp. 65-83.

18. Wang, Y. and Li, M. (2019) 'Urban impervious surface detection from remote sensing images: a review of the methods and challenges', *IEEE Geoscience and Remote Sensing Magazine*, 7(3), pp. 64–93.
19. Lu, D. and Weng, Q. (2007) 'A survey of image classification methods and techniques for improving classification performance', *International Journal of Remote Sensing*, 28(5), pp. 823-870.
20. Münch, Z., Gibson, L. and Palmer, A. (2019) 'Monitoring effects of land cover change on biophysical drivers in rangelands using albedo', *Land*, 8(2), 33.
21. Cunha, J., Nóbrega, R.L.B., Rufino, I., Erasmi, S., Galvão, C. and Valente, F. (2020) 'Surface albedo as a proxy for land-cover clearing in seasonally dry forests: evidence from the Brazilian Caatinga', *Remote Sensing of Environment*, 238, 111250.
22. Houspanossian, J., Giménez, R., Jobbágy, E. and Nosetto, M. (2017) 'Surface albedo raise in the south American Chaco: combined effects of deforestation and agricultural changes', *Agricultural and Forest Meteorology*, 232, pp. 118–127.
23. Barletta, C., Capolupo, A. and Tarantino, E. (2023) *Open multitemporal Earth Observation data for land surface albedo estimation in urban areas*. Available at: https://talks.osgeo.org/foss4g-it-2023/talk/9UWGEF/.
24. Wulder, M.A. *et al.* (2022) 'Fifty years of Landsat science and impacts', *Remote Sensing of Environment*, 280, 113195.
25. Masek, J.G., Wulder, M.A., Markham, B., McCorkel, J., Crawford, C.J., Storey, J. and Jenstrom D.T. (2020) 'Landsat 9: Empowering open science and applications through continuity', *Remote Sensing of Environment*, 248, 111968.
26. Zhu, Z. (2017) 'Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications', *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, pp. 370–384.

27. Barletta, C., Capolupo, A. and Tarantino, E. (2023) 'Extracting Land Surface Albedo from Landsat 9 Data in GEE Platform to Support Climate Change Analysis', *Geomatics and Environmental Engineering*, 17(6), pp. 35-75.
28. National Aeronautics and Space Administration (2013) *A Landsat Timeline*. Available at: https://landsat.gsfc.nasa.gov/article/a-landsat-timeline/.
29. Zhu, Z. and Woodcock, C.E. (2014) 'Continuous change detection and classification of land cover using all available Landsat data', *Remote Sensing of Environment*, 144, pp. 152–171.
30. Woodcock, C.E., Loveland, T.R., Herold, M. and Bauer, M.E. (2020) 'Transitioning from change detection to monitoring with remote sensing: A paradigm shift', *Remote Sensing of Environment*, 238, 111558.
31. Awty-Carroll, K., Bunting, P., Hardy, A. and Bell, G. (2019) 'Using Continuous Change Detection and Classification of Landsat Data to Investigate Long-Term Mangrove Dynamics in the Sundarbans Region', *Remote Sensing*, 11, 2833.
32. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R. (2017) 'Google earth engine: planetary-scale geospatial analysis for everyone', *Remote Sensing of Environment*, 202, pp. 18–27.
33. Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S. and Brisco, B. (2020) 'Google earth engine for geo-big data applications: a meta-analysis and systematic review', *ISPRS Journal of Photogrammetry and Remote Sensing*, 164, pp. 152–170.
34. Amani, M. *et al.* (2020) 'Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, pp. 5326-5350.
35. Pérez-Cutillas, P., Pérez-Navarro, A., Conesa-García, C., Zema, D.A. and Amado-Álvarez, J.P. (2023) 'What is going on within google earth en-

gine? A systematic review and meta-analysis', *Remote sensing applications: Society and environment*, 29, 100907.

36. Google Developers (2024) *Earth Engine Code Editor*. Available at: https://developers.google.com/earth-engine/guides/playground.
37. Google Developers (2021) *Objects and Methods Overview*. Available at: https://developers.google.com/earth-engine/guides/objects_methods_overview.
38. Barletta, C., Capolupo, A. and Tarantino, E. (2022) 'Exploring the Potentialities of Landsat 8 and Sentinel-2 Satellite Data for Estimating the Land Surface Albedo in Urban Areas Using GEE Platform' in Gervasi, O. *et al.* (eds.) *Computational Science and Its Applications – ICCSA 2022 Workshops*, LNCS 13379. Springer Cham., pp. 435-449.
39. Tassi, A., Gigante, D., Modica, G., Di Martino, L. and Vizzari, M. (2021) 'Pixel- vs. Object-Based Landsat 8 Data Classification in Google Earth Engine Using Random Forest: The Case Study of Maiella National Park', *Remote Sensing*, 13, 2299.
40. Capolupo, A., Barletta, C., Esposito, D. and Tarantino, E. (2024) 'Earth Observation Data for Sustainable Management of Water Resources to Inform Spatial Planning Strategies' in: Marucci, A. *et al.* (eds.) *Innovation in Urban and Regional Planning - INPUT 2023*, LNCE 467. Springer Cham., pp. 24-35.
41. Global Climate Observing System (2024) *Essential Climate Variables*. Available at: https://gcos.wmo.int/en/essential-climate-variables/table.
42. Liu, Q., Wen, J., Qu, Y., He, T., Zhang, X. and Wang, L. (2012) 'Broadband albedo' in Liang, S., Li, X. and Wang, J. (eds.) *Advanced Remote Sensing: Terrestrial Information Extraction and Applications*. Academic Press Cambridge, pp. 173–231.
43. Zhang, X. *et al.* (2022) 'Review of land surface albedo: variance characteristics, climate effect and management strategy', *Remote Sensing*, 14(6), 1382.

44. Andres-Anaya, P., Sanchez-Aparicio, M., Del Pozo, S., Lagüela, S., Hernández-López, D. and Gonzalez-Aguilera, D. (2024) 'A New Methodology for Estimating Surface Albedo in Heterogeneous Areas from Satellite Imagery', *Applied Sciences*, 14, 75.
45. Baldinelli, G., Bonafoni, S. and Rotili, A. (2017) 'Albedo Retrieval From Multispectral Landsat 8 Observation in Urban Environment: Algorithm Validation by in situ Measurements', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(10), pp. 4504-4511.
46. Bonafoni, S. and Sekertekin, A. (2020) 'Albedo Retrieval From Sentinel-2 by New Narrow-to-Broadband Conversion Coefficients', *IEEE Geoscience and Remote Sensing Letters*, 17(9), pp. 1618-1622.
47. SEOS project (2024) *Interaction of radiation with the earth surface*. Available at: https://seos-project.eu/remotesensing/remotesensing-c01-p05.html.
48. Hofierka, J. and Onačillová, K. (2022) 'Estimating Visible Band Albedo from Aerial Orthophotographs in Urban Areas', *Remote Sensing*, 14, 164.
49. Taha, H. (1997) 'Urban climates and heat islands: albedo, evapotranspiration, and anthropogenic heat', *Energy and Buildings*, 25(2), pp. 99–103.
50. Allen, R.G., Waters, R., Tasumi, M., Trezza, R. and Bastiaanssen, W. (2002) *SEBAL: Surface Energy Balance Algorithms for Land: Idaho Implementation. Advanced Training and Users Manual*. Version 1.0.
51. Capolupo, A., Monterisi, C., Sonnessa, A., Caporusso, G. and Tarantino, E. (2021) 'Modeling Land Cover Impact on Albedo Changes in Google Earth Engine Environment' in Gervasi, O. *et al.* (eds.) *Computational Science and Its Applications – ICCSA 2021 Workshops*, LNCS 12955. Springer Cham., pp. 89-101.
52. Qu, Y., Liang, S., Liu, Q., He, T., Liu, S. and Li, X. (2015) 'Mapping surface broadband albedo from satellite observations: a review of literatures on algorithms and products', *Remote Sensing*, 7(1), pp. 990–1020.

53. Liang, S. (2000) 'Narrowband to broadband conversions of land surface albedo I Algorithms', *Remote Sensing of Environment*, 76, pp. 213-238.
54. Tasumi, M., Allen, R.G., Trezza, R. (2008) 'At-surface reflectance and albedo from satellite for operational calculation of land surface energy balance', *Journal of Hydrological Engineering*, 13(2), pp. 51–63.
55. Zhong, Q. and Li, Y.H. (1988) 'Satellite observation of surface albedo over the Qinghai-Xizang plateau region', *Advances in Atmospheric Science*, 5, pp. 57–65.
56. Bastiaanssen, W., Menenti, M., Feddes, R.A. and Holslag, A.A.M.A. (1998) 'Remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation', *Journal of Hydrology*, 212–213, pp. 198–212.
57. da Silva, B.B., Braga, A.C., Braga, C.C., de Oliveira, L.M.M., Montenegro, S.M.G.L. and Barbosa, B. Jr (2016) 'Procedures for calculation of the albedo with OLI-Landsat 8 images: Application to the Brazilian semi-arid', *Revista Brasileira de Engenharia Agrícola e Ambiental*, 20(1), pp. 3–8.
58. Bastiaanssen, W. (2000) 'SEBAL-based sensible and latent heat fluxes in the irrigated Gediz Basin, Turkey', *Journal of Hydrology*, 229(1–2), pp. 87–100.
59. Allen, R.G., Tasumi, M. and Trezza, R. (2007) 'Satellite-based energy balance for mapping evapotranspiration with internalized calibration (METRIC) – model', *Journal of Irrigation and Drainage Engineering*, 133(4), pp. 380–394.
60. Parastatidis, D., Mitraka, Z., Chrysoulakis, N. and Abrams, M. (2017) 'Online global land surface temperature estimation from Landsat', *Remote Sensing*, 9(12), 1208.
61. Kalnay, E. *et al.* (1996) 'The NCEP/NCAR 40-Year Reanalysis Project', *Bulletin of the American Meteorological Society*, 77(3), pp. 437–472.
62. Regione Siciliana - Assessorato Territorio e Ambiente (2004) *Piano Stralcio di Bacino per l'Assetto Idrogeologico della Regione Siciliana: Relazione Generale*. Available at: https://www.sitr.regione.sicilia.it/pai/.

63. Cristiano, E., Urru, S., Farris, S., Ruggiu, D., Deidda, R. and Viola, F. (2020) 'Analysis of potential benefits on flood mitigation of a CAM green roof in Mediterranean urban areas', *Building and Environment*, 183, 107179.
64. Cappadonia, C., Di Maggio, C., Agate, M. and Agnesi, V. (2020) 'Geomorphology of the urban area of Palermo (Italy)', *Journal of Maps*, 16(2), pp. 274–284.
65. Porta, M., Buosi, C., Trogu, D., Ibba, A. and De Muro S. (2021) 'An integrated sea-land approach for analyzing forms, processes, deposits and the evolution of the urban coastal belt of Cagliari', *Journal of Maps*, 17, pp. 65–74.
66. Gross, G., Helder, D., Begeman, C., Leigh, L., Kaewmanee, M. and Shah, R. (2022) 'Initial cross-calibration of Landsat 8 and Landsat 9 using the simultaneous underfly event', *Remote Sensing*, 14(10), 2418.
67. Tukey, J.W. (1977) 'Easy summaries – numerical and graphical' in Tukey, J.W. (ed.) *Exploratory Data Analysis*. Addison-Wesley Reading, pp. 27–56.
68. Roy, D.P. et al. (2014) 'Landsat-8: Science and product vision for terrestrial global change research', *Remote Sensing of Environment*, 145, pp. 154–172.
69. United States Geological Survey (2024a) *Landsat 8*. Available at: https://www.usgs.gov/landsat-missions/landsat-8.
70. National Aeronautics and Space Administration (2024a) *Landsat 8*. Available at: https://landsat.gsfc.nasa.gov/satellites/landsat-8/.
71. United States Geological Survey (2024b) *Landsat Collection 2 Surface Reflectance*. Available at: https://www.usgs.gov/landsat-missions/landsat-collection-2-surface-reflectance.
72. Labadessa, R. and Ancillotto, L. (2023) 'Small but irreplaceable: The conservation value of landscape remnants for urban plant diversity', *Journal of Environmental Management*, 339, 117907.

73. Autorità di Bacino della Puglia (2004) *Piano di Bacino Stralcio dell'Assetto Idrogeologico (PAI)*.
74. Jensen, J.R. (2015) *Introductory digital image processing: a remote sensing perspective*. 4th Edn. Pearson.
75. Capolupo, A., Monterisi, C. and Tarantino, E. (2020) 'Landsat images classification algorithm (LICA) to automatically extract land cover information in Google earth engine environment', *Remote Sensing*, 12(7), 1201.
76. Polemio, M. and Lonigro, T. (2015) 'Trends in climate, short-duration rainfall, and damaging hydrogeological events (Apulia, Southern Italy)'. *Natural Hazards*, 75(1), pp. 515–540.
77. Cotecchia, V., Simeone, V. and Gabriele, S. (2014) 'Caratteri climatici' in Le a*cque sotterranee e l'intrusione marina in Puglia: dalla ricerca all'emergenza nella salvaguardia della risorsa*, Memorie descrittive della carta geologica d'Italia 92. ISPRA Servizio Geologico d'Italia, pp. 338–369.
78. Guan, X., Huang, J., Guo, N., Bi, J. and Wang, G. (2009) 'Variability of soil moisture and its relationship with surface albedo and soil thermal parameters over the Loess Plateau', *Advances in Atmospheric Science*, 26, pp. 692–700.
79. Sanchez-Mejia, Z.M., Papuga, S.A., Swetish, J.B., van Leeuwen, W.J.D., Szutu, D. and Hartfield, K. (2014) 'Quantifying the influence of deep soil moisture on ecosystem albedo: the role of vegetation', *Water Resources Research*, 50(5), pp. 4038–4053.
80. Yang, J., Li, Z., Zhai, P., Zhao, Y. and Gao, X. (2020) 'The influence of soil moisture and solar altitude on surface spectral albedo in arid area', *Environmental Research Letters*, 15(3), 035010.
81. Liu, S., Chadwick, O.A., Roberts, D.A. and Still, C.J. (2011) 'Relationships between GPP, satellite measures of greenness and canopy water con-

tent with soil moisture in Mediterranean-climate grassland and oak savanna', *Applied and Environmental Soil Science*, 2011.

82. de Oliveira, M.L., Rufino, I.A.A., Cunha, J.E.D.B.L., Vasconcelos, R.S. and de Brito, H.C. (2020) 'Urban growth dynamics based on surface Albedo changes in Petrolina, Brazil', *Acta Scientiarum. Technology*, 42, e46270.
83. Hou, M., Hu, Y. and He, Y. (2014) 'Modifications in vegetation cover and surface albedo during rapid urbanization: a case study from South China', *Environmental Earth Sciences*, 72, pp. 1659–1666.
84. Congalton, R.G. and Green, K. (2009) *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. 2nd edn. CRC Press.
85. Congalton, R.G. (1991) 'A review of assessing the accuracy of classifications of remotely sensed data', *Remote Sensing of Environment*, 37(1), pp. 35–46.
86. Congalton, R.G. (2001) 'Accuracy assessment and validation of remotely sensed and other spatial information', *International Journal of Wildland Fire*, 10(4), pp. 321–328.
87. Weng, Q. (2012) 'Remote sensing of impervious surfaces in the urban areas: requirements, methods, and trends', *Remote Sensing of Environment*, 117, pp. 34–49.
88. Leone, A., Gobattoni, F., Pelorosso, R. and Calace, F. (2020) 'Nature-based climate adaptation for compact cities: green courtyards as urban cool islands', *Plurimondi*, 18, pp. 83-110.
89. Pelorosso, R., Gobattoni, F., Geri, F., Monaco, R. and Leone, A. (2016) 'Evaluation of Ecosystem Services related to Bio-Energy Landscape Connectivity (BELC) for land use decision making across different planning scales', *Ecological Indicators*, 61(1), pp. 114-129.
90. United States Geological Survey (2024c) *Landsat 4*. Available at: https://www.usgs.gov/landsat-missions/landsat-4.
91. United States Geological Survey (2024d) *Landsat 5*. Available at: https://www.usgs.gov/landsat-missions/landsat-5.

92. United States Geological Survey (2024e) *Landsat 7*. Available at: https://www.usgs.gov/landsat-missions/landsat-7.
93. United States Geological Survey (2024f) *Landsat 9*. Available at: https://www.usgs.gov/landsat-missions/landsat-9.
94. National Aeronautics and Space Administration (2024b) *The Worldwide Reference System*. Available at: https://landsat.gsfc.nasa.gov/about/the-worldwide-reference-system/.
95. Holland, P.W. and Welsch, R.E. (1977) 'Robust regression using iteratively reweighted least-squares', *Communications in Statistics - Theory and Methods*, 6(9), pp. 813–827.
96. Tibshirani, R. (1996) 'Regression Shrinkage and Selection via the Lasso', *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), pp. 267–288.
97. Deng, C. and Zhu, Z. (2020) 'Continuous subpixel monitoring of urban impervious surface using Landsat time series', *Remote Sensing of Environment*, 238, 110929.
98. ESRI (2024) *How Analyze Changes Using CCDC works*. Available at: https://pro.arcgis.com/en/pro-app/latest/tool-reference/image-analyst/how-analyze-changes-using-ccdc-works.htm.
99. Breiman, L. (2001) 'Random Forest', *Machine Learning*, 45, pp. 5-32.
100. Shetty, S. (2019) *Analysis of Machine Learning Classifiers for LULC Classification on Google Earth Engine*. Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente (Supervisors: Gupta, P.K., Belgiu, M. and Srivastav, S.K.).
101. Belgiu, M. and Drăguţ, L. (2016) 'Random forest in remote sensing: A review of applications and future directions', *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, pp. 24–31.
102. Lary, D.J., Alavi, A.H., Gandomi, A.H. and Walker, A.L. (2016) 'Machine learning in geosciences and remote sensing', *Geoscience Frontiers*, 7(1), pp. 3-10.

103. Thomas, I.L. (1980) 'Spatial post-processing of spectrally-classified Landsat data', *Photogrammetric Engineering and Remote Sensing*, 46, pp. 1201–1206.

104. Trigila, A., Iadanza, C., Lastoria, B., Bussettini, M. and Barbano, A. (2021) *Dissesto idrogeologico in Italia: pericolosità e indicatori di rischio*. Edizione 2021. Report ISPRA 356/2021.

105. Italian National Institute for Environmental Protection and Research (2021) *Consumo di suolo: Roma perde terreno al ritmo di 100 ettari all'anno, ben 150 campi di calcio costruiti ogni 365 giorni*. Available at: https://www.isprambiente.gov.it/it/archivio/notizie-e-novita-normative/notizie-ispra/2021/05/consumo-di-suolo-roma-perde-terreno-al-ritmo-di-100-ettari-all2019anno-ben-150-campi-di-calcio-costruiti-ogni-365-giorni.

106. Rainews (2022) *Roma prima città per consumo di suolo*. Available at: https://www.rainews.it/tgr/lazio/articoli/2022/10/roma-prima-citt-per-consumo-di-suolo-a26c02d9-d913-4da3-95f5-549aa825a64e.html.

107. Mannella, A., Di Ludovico, M., Sabino, A., Prota, A., Dolce, M. and Manfredi, G. (2017) 'Analysis of the Population Assistance and Returning Home in the Reconstruction Process of the 2009 L'Aquila Earthquake', *Sustainability*, 9(8), 1395.

108. Legambiente Umbria (2022) *Dai dati Ispra il consumo di suolo dell'Umbria*. Available at: https://www.legambienteumbria.it/dai-dati-ispra-il-consumo-di-suolo-dellumbria/.

109. La Gazzetta del Mezzogiorno (2022) *Bari, sempre più cemento la città sta soffocando*. Available at: https://www.lagazzettadelmezzogiorno.it/news/bari/1352330/bari-sempre-piu-cemento-la-citta-sta-soffocando.html.

110. Ambient&Ambienti (2023) *A Bari si continua a costruire: consumo di suolo o rigenerazione urbana?* Available at:

https://www.ambienteambienti.com/a-bari-si-continua-a-costruire-consumo-di-suolo-o-rigenerazione-urbana/.

CURRICULUM

PERSONAL INFORMATION

Name	**BARLETTA, Carlo**
Address	**199 C, Via Amendola, 70126, Bari, Italy**
E- mail	**carlo.barletta@poliba.it**
Nationality	**Italian**
Google Scholar	**[Link]**

EDUCATION

• Date	Nov 2021 - Oct 2024
• Name and type of education Institution	Polytechnic of Bari University, Italy
• Field of study	Remote sensing and G.I.S.
• Role	PhD student
• Date	Sep 2024 - Oct 2024
• Name and type of education Institution	GeoSat Research laboratory, Institute of Mediterranean Studies (IMS) – Foundation for Research and Technology Hellas (FORTH), Greece
• Field of study	Remote sensing and G.I.S.
• Role	Secondee (under the EU-funded “COASTLINE” project)
• Date	Mar 2024 - Jun 2024
• Name and type of education Institution	University of Bucharest – Faculty of Geography, Romania
• Field of study	Remote sensing and G.I.S.
• Role	Visiting PhD student

• Date	Mar 2023
• Name and type of education Institution	Polytechnic of Bari University, Italy
• Qualification	License for Civil and Environmental engineering profession (Section A)
• Date	Jul 2021
• Name and type of education Institution	Polytechnic of Bari University, Italy
• Field of study	Environmental and land engineering
• Qualification	Master's degree

WORK EXPERIENCE

• Date	Jan 2021 - Jun 2021
• Name of employer	Polytechnic of Bari University, Italy – Applied Geomatics laboratory (AG lab)
• Type of employment	Internship

LANGUAGE SKILLS

NATIVE LANGUAGE	Italian
OTHER LANGUAGES	English Level B2 Spanish Level B1

DIGITAL SKILLS

High knowledge of Microsoft Office Package (Word, Excel, Powerpoint), QGIS software, R software, Google Earth Engine platform, JavaScript programming