

Una nuova era nell'uso del telerilevamento a supporto della pianificazione sostenibile del territorio

Big Data e intelligenza artificiale a portata di mano

Gherardo Chirici

DAGRI, Università degli Studi di Firenze
gherardo.chirici@unifi.it

© 2020 The Author(s)
This article is published
with Creative Commons
license CC BY-SA 4.0
Firenze University Press.
DOI: 10.13128/contest-11988
www.fupress.net/index.php/contesti/

keywords
remote sensing
big data
artificial intelligence
copernicus

Since the seventies of the last century, remote sensing technologies for Earth observation have generated a constant flow of data from different platforms, in different formats and with different purposes. From these, through successive steps, spatial information useful to support the monitoring and planning of territorial resources is generated. The remotely sensed information is recognized as a classic example of big data ante litteram. Today, new cloud computing technologies

Introduzione

Il telerilevamento

Il telerilevamento è definibile come la tecnologia atta a misurare a distanza le caratteristiche di un oggetto o di una superficie (Sabins, 1978).

L'osservazione della Terra (*Earth Observation*

- EO) si riferisce all'uso di tecnologie di telerilevamento per monitorare la terra, le acque (mari, fiumi, laghi) e l'atmosfera. L'EO è basata su satelliti equipaggiati con sensori per raccogliere dati della superficie terrestre. Questi vengono quindi elaborati e analizzati al fine di estrarre diversi tipi di informazioni che possono servire per una vasta gamma di applicazioni commerciali e di ricerca.

(such as Google Earth Engine) make it possible to tackle the complex problem of managing and processing big data from remote sensing with new strategies that have revolutionized the way of understanding the use of these data sources. Also thanks to the advent of artificial intelligence technologies. From experiments on small study areas, today we work globally thanks to the possibility of operationally processing vast multidimensional and multitemporal datasets. The greater availability of information from space is exemplified by the numerous services offered by the European Copernicus program that allow the generation of important information useful to support urban planning, agriculture and sustainable management of forest resources.

Le tecnologie di EO utilizzano diversi tipi di sensori:

- Telerilevamento passivo: sensori ottici o termici che rilevano l'energia ricevuta dalla Terra a causa della riflessione e riemissione dell'energia del Sole dalla superficie o dall'atmosfera terrestre. Operano nelle lunghezze d'onda comprese tra il visibile e l'infrarosso.
- Telerilevamento attivo: i sensori radar (*Radio Detection and Ranging*)

operano nella parte inferiore dello spettro (microonde) inviando energia alla Terra e monitorando l'energia ricevuta di ritorno dalla superficie terrestre o dall'atmosfera, consentendo il monitoraggio anche notturno e in tutte le condizioni meteorologiche; i sensori lidar (*Light Detection and Ranging*) inviano invece impulsi laser nel campo dell'infrarosso e calcolano la posizione nello spazio dei bersagli colpiti, permettendo la ricostruzione tridimensionale degli oggetti presenti sulla superficie terrestre.

Sulla base della risoluzione spaziale (o geometrica) del sensore, ovvero della dimensione dei pixel analizzati dai sensori, i dati da telerilevamento si classificano in:

- Bassa e media risoluzione: più di 30 metri per pixel;
- Alta risoluzione (HR): tra 1 e 30 metri per pixel;
- Risoluzione molto alta (VHR): meno di 1 metro per pixel, chiamate anche immagini sub-metriche.

La risoluzione temporale definisce la frequenza con cui i dati vengono acquisiti, le orbite della maggior parte dei satelliti per l'osservazione terrestre a medio-alta risoluzione (pixel di 10-30 metri) sono dell'ordine delle poche settimane, per aumentare la frequenza di acquisizione è però sufficiente porre satelliti gemelli sulla stessa

orbita. Con questa tecnica è possibile acquisire un'immagine completa di tutta la Terra una o più volte al giorno.

Nel caso dei sensori ottici un altro parametro importante è la risoluzione spettrale. Questa è definita dall'ampiezza delle bande dello spettro che possono essere distinte dal sensore. In genere maggiore è la risoluzione spaziale e minore sarà la risoluzione spettrale. Le applicazioni e le tecnologie di osservazione della terra sono utilizzate in un'ampia gamma di applicazioni che riguardano il monitoraggio terrestre, marino e atmosferico.

Secondo le stime fornite dal *Copernicus Market Report (2017)* nel 2017, l'economia globale dell'EO è stimata tra 9,6 e 9,8 miliardi di EUR, diviso tra le vendite di satelliti (la sezione più a monte della catena di fornitura), l'acquisizione, l'elaborazione e la trasformazione dei dati in prodotti informativi per gli utenti finali (la sezione più a valle della catena). Il mercato globale è guidato dalla sezione a monte, che costituisce circa il 70% dei ricavi totali. Il mercato a valle dell'EO globale è stimato essere compreso tra 2,6 e 2,8 miliardi di Euro, principalmente guidato dalle applicazioni in campo pubblico, che rappresentano tra il 50% e il 60% dei ricavi. Mentre i ricavi a monte tendono ad oscillare nel corso degli anni a seconda delle fluttuazioni per la necessità di mettere in orbita grandi satelliti, il mercato a valle mostra una crescita costante, a un CAGR (Compound Annual Growth Rate) previsto del 7% fino al 2022.

Il mercato dei prodotti da EO ha subito un forte impatto derivante dal processo di democratizzazione dei dati, ovvero l'apertura verso tutti dei dati da telerilevamento,

prima riservati a una strettissima fascia di utilizzatori. Oggi recuperare i dati da telerilevamento non è più un problema e per una vasta gamma di questi l'accesso è gratuito.

A seguito di questa politica in favore dei dati open access, anche i dati forniti da ditte private a pagamento stanno subendo un graduale abbassamento del prezzo di mercato. Compensata però dall'espansione della base dei clienti, che è in costante aumento.

La disponibilità a pagare è diversa a seconda del mercato, con attori della difesa e dell'intelligence disposti a pagare di più per immagini ad alta risoluzione, rispetto ad altri settori (come agricoltura e foreste) dove la diffusione dei dati da EO, seppur in crescita, è frenata dalla forte sensibilità al prezzo dei prodotti.

Big data

Una delle molte definizioni che viene data ai big data è di un insieme di dati che per dimensione o tipo vanno oltre la capacità dei database relazionali tradizionali di acquisirli, gestirli ed elaborarli con bassa latenza.

Kitchin (2013) dettaglia meglio le caratteristiche dei Big Data:

- enormi in volume, costituiti almeno da terabyte o petabyte di dati;
- la cui generazione avviene ad alta velocità, creati in tempo reale o quasi;
- diversi per varietà, essendo strutturati e non strutturati in natura;
- di portata esaustiva, catturano intere popolazioni o sistemi (n = tutti);
- a grana fine nella risoluzione ma univoci nell'identificazione;

Generazione	Quando	Approccio scientifico	Metodo
Prima	pre-rinascimento	Scienza sperimentale	Empirico, osservazione dei fenomeni naturali
Seconda	pre-computer	Scienza teorica	Modellistico, generalizzazione
Terza	pre-big data	Scienza computazionale	Simulazione di fenomeni complessi
Quarta	oggi	Scienza esplorativa	Data mining, data intensive, intelligenza artificiale

Le quattro fasi o paradigmi della scienza, modificato

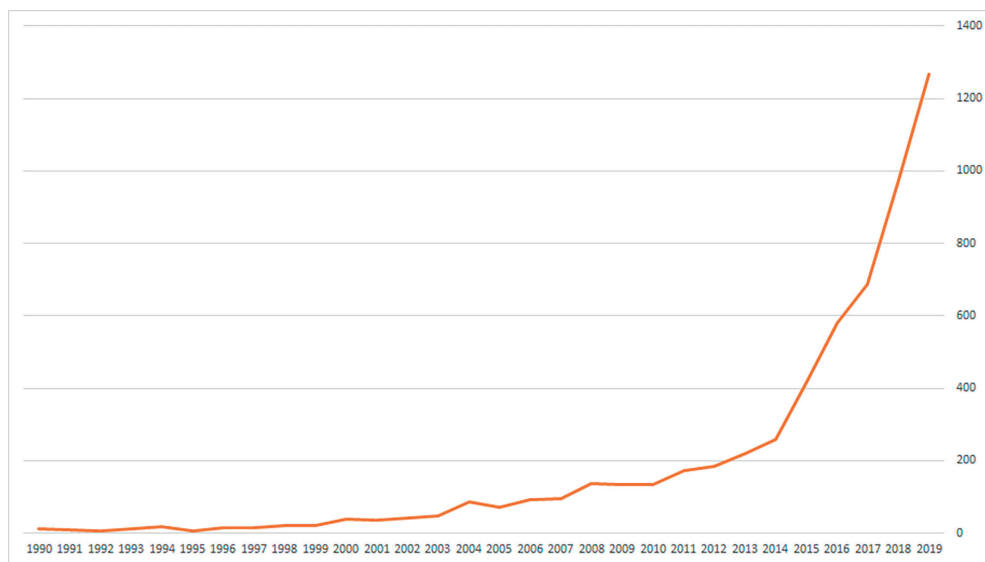
Tabella 1
da Kitchin (2014)

- di natura relazionale, contenente campi comuni che consentono l'unione di diversi set di dati;
- flessibile, mantenendo i tratti di estensionalità (si possono aggiungere facilmente nuovi campi) e scalabilità (possono espandersi rapidamente di dimensione).

Secondo Laney (2020) non importa se si tratti di dati generati da progetti governativi, applicazioni commerciali o per ricerca accademica, i big data devono essere caratterizzati dalle 3V: ovvero un elevato Volume, Velocità di generazione e Varietà nelle caratteristiche del dato.

La rappresentazione spaziale in forma digitale della superficie terrestre utilizzata dal telerilevamento è un ottimo esempio *ante litteram* di big data, big data ancora prima che il termine venisse coniato. Nonostante il recente sviluppo di alcune tecnologie di telerilevamento laser (lidar) che producono dati grezzi in formato vettoriale, la vasta maggioranza dei prodotti da telerilevamento è in formato raster, ovvero di matrice. Facciamo un esempio. Se vogliamo rappresentare tutte le terre emerse della Terra con una risoluzione di 10 m generiamo un'immagine contenente circa 1,5 bilioni di pixel. Considerando anche soltanto l'ormai

desueta profondità di 8 bit ogni immagine 'peserebbe' 1,53 terabyte, considerando le consuete frequenze di acquisizione, un solo satellite in un'anno genererebbe approssimativamente tra i 30 e i 50 terabyte per ognuno dei canali di acquisizione. La creazione di un semplice indice di attività fotosintetica globale con 32 bit di profondità peserebbe poco meno di 6 terabyte (almeno in un formato di dati non compresso). Sebbene dunque la produzione di big data sia esistita in alcuni domini, come il telerilevamento, le previsioni meteorologiche e i mercati finanziari, ormai da svariati decenni, solo molto recentemente una serie di sviluppi tecnologici, come l'*ubiquitous computing*, il funzionamento diffuso e ad alta velocità di Internet e i nuovi design usati nella gestione delle banche dati permettono non solo l'archiviazione ma anche un processamento dei big data in tempi ragionevoli. Questo passaggio secondo Kitchin (2014) costituisce un vero punto di svolta per la generazione di routine di analisi specificatamente progettate per far fronte a una tale abbondanza di dati. Queste nuove tecniche di processamento sono radicate nella ricerca sull'intelligenza artificiale e sui sistemi esperti che hanno cercato di produrre sistemi di apprendimento automatico in grado rilevare pattern significativi nei big data attraverso modelli computazionali atti alla generazione di modelli predittivi, basati, il più delle volte,



Il trend di pubblicazioni su telerilevamento e big data.

Fig. 1
Fonte: SCOPUS.

sulle architetture delle reti neurali (Han et al., 2011; Hastie et al., 2009).

Come stabilito da Kuhn (1962), un paradigma costituisce un modo accettato di interrogare il mondo e sintetizzare la conoscenza comune a una proporzione sostanziale di ricercatori in una disciplina, in un determinato momento nel tempo. Periodicamente, sostiene Kuhn (1962), emerge un nuovo modo di pensare che sfida teorie e approcci precedentemente accettati.

È ipotizzabile che l'avanzamento tecnologico che ha reso maggiormente disponibili in forma organizzata e accessibile al processamento i big data possa aver generato il passaggio a una nuova generazione della ricerca scientifica basata su approcci esplorativi ai dati derivati dall'applicazione dell'approccio all'intelligenza artificiale e non più meramente orientato alla simulazione dei fenomeni (Tabella 1)

È quanto molto probabilmente è accaduto negli ultimi anni anche nel campo del telerilevamento. Facciamo un esempio. Con i metodi scientifici di terza generazione tipicamente una ricerca scientifica basata su telerilevamento veniva realizzata su un certo ambito spaziale di studio (l'area di studio) e in un certo ambito temporale di studio (il periodo indagato). Se da un lato la ricerca poteva essere realizzata in un periodo relativamente breve e con risorse limitate, i risultati che se ne ottenevano soffrivano inevitabilmente di un limite implicito: la mancanza di generalizzabilità. I risultati sarebbero stati veri anche fuori dall'area di studio e in un altro periodo?

Oggi, con un approccio di quarta generazione, la possibilità di gestire e manipolare i big data permette di affrontare il problema alla radice: l'area di studio sarà quindi tutta la Terra, il periodo di studio è da oggi indietro fino alla data di acquisizione della prima immagine telerilevata disponibile.

Big data e telerilevamento

Il volume di dati prodotto dalle principali agenzie come NASA (National Aeronautics and Space Administration USA), JAXA (Japan Aerospace Exploration Agency), ESA (European Space Agency) (in particolare grazie al rinnovato vigore offerto dal programma Copernicus) e a tutti gli altri gestori di piattaforme di EO, ha introdotto una sfida per lo sviluppo di piattaforme on-line atte a gestire, elaborare e diffondere questi dati. Si è in tal modo creato un nuovo mercato e quindi un'opportunità per i grandi attori del settore Information & Communication, come Google e Amazon, che avevano tecnologie già sviluppate e quindi si sono trovati meglio posizionati per affrontare la questione dei big data.

Il nuovo mercato dei *Big Data Analytics* (BDA) nel campo EO si basa sull'elaborazione, l'analisi e la fusione di più immagini con altre fonti di dati al fine di creare informazioni non precedentemente disponibili. Si sono quindi affermati modelli di business per l'EO orientati a offrire servizi digitali basati su una forte potenza di calcolo, piattaforme basate su *cloud* e la fusione di un numero sempre maggiore di fonti di dati.

Anche nel campo della ricerca è evidente l'impegno della comunità scientifica a esplorare le opportunità offerte dall'approccio di quarta generazione. Lo testimoniano i numeri speciali dedicati a questo argomento da molte delle principali riviste di telerilevamento.

Si inizia nel 2015 con *IEEE JSTARS* che pubblica il numero speciale su "*Big Data in Remote Sensing*" e il *Journal of Applied Remote Sensing* con "*Management and Analytics of Remotely Sensed Big Data*". Seguono nel 2016 *IEEE Geoscience and Remote Sensing Magazine* con lo special issue su "*Big Data from Space and Geoinformatics*" e di Springer in *Geoinformatica* con "*Big Spatial and Spatiotemporal Data Management and Analytics*". Nel 2017 è la volta di *Remote Sensing of Environment* con "*Big Remotely Sensed Data: tools, applications and experiences*". In casa MDPI a partire dal 2018 sulla rivista *Remote Sensing* si aprono diversi special issue "*Advanced Machine Learning and Big Data Analytics in Remote Sensing for Natural Hazards Management*", "*SAR in the Big Data Era*" e "*Analysis of Big Data in Remote Sensing*". Dello stesso anno lo speciale "*Social Sensing and Big Data Computing for Disaster Management*" nella rivista *International Journal of Digital Earth*. In SCOPUS usando una query del tipo "*remote AND sensing AND big AND data*" al 2019 risultano 5852 documenti, con un evidente trend di crescita (Figura 1).

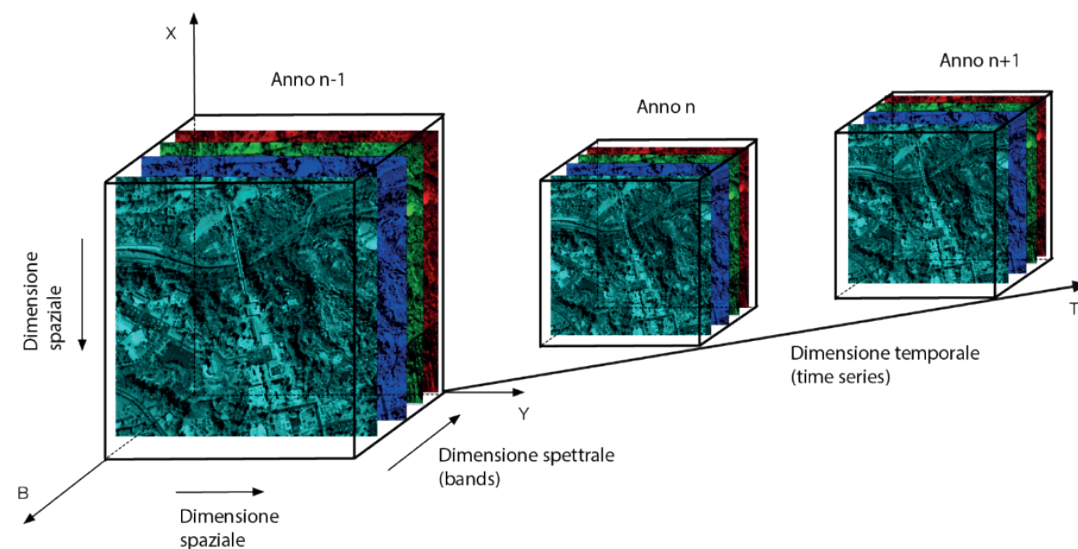
L'incremento della mole di dati da telerilevamento è dovuto alla loro multidimensionalità

Nel telerilevamento il termine big data non si riferisce solo al volume e alla velocità di produzione dei dati (lo *streaming*), ci sono infatti anche altri elementi che

La multidimensionalità dei dati telerilevati

Fig. 2

Da Ma et al., 2015, modificato

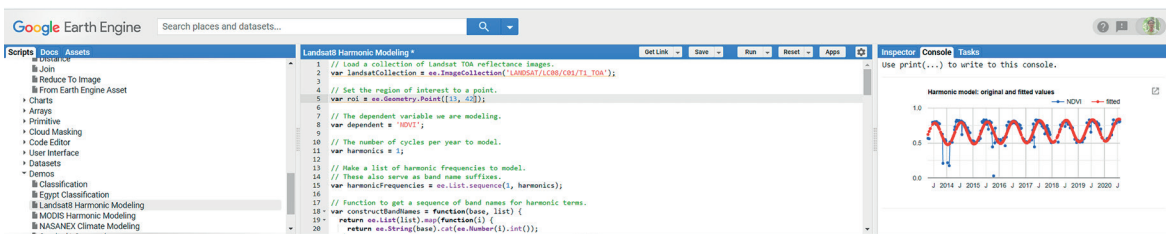


mettono in crisi le capacità di archiviazione e di elaborazione di questi dati quali la loro varietà e complessità. In altre parole la loro multidimensionalità, che porta inevitabilmente a una maggiore complessità. I dati da telerilevamento normalmente servono molte discipline, dalle scienze della terra, al monitoraggio ambientale, i processi terrestri, l'atmosfera, l'idrologia e l'oceanografia, dallo studio delle aree urbane, delle coste, delle aree agricole, alle foreste e agli altri biomi. L'ampia gamma delle discipline coinvolte dà luogo anche a diverse forme di pre-trattamento dei dati grezzi che crea spesso la duplicazione degli archivi. Per esempio per i dati ottici si troveranno le immagini grezze, quelle normalizzate radiometricamente, quelle con diversi livelli di ortocorrezione, quelle con valori *Bottom of Atmosphere* (BOA) e quelle *Top of Atmosphere* (TOA). Senza contare i mosaici creati con diverse tecniche per la eliminazione delle coperture nuvolose. Ma et al. (2015) riporta per gli archivi della NASA circa 7000 tipi di set di dati. Nella maggior parte dei casi, i set sono memorizzati in file strutturati utilizzando vari formati standard, inclusi HDF, netCDF, GeoTIFF, FAST, ASCII, JPG2000 e così via. Oggi il progetto ESDIS (*Earth Science Data and Information System*) della NASA raccoglie metriche da tutto il sistema di dati e informazioni del sistema di osservazione della terra (EOSDIS) la cui dimensione è ad oggi di 33.6 petabyte, in crescita giornaliera

di circa 20 terabyte, i dati distribuiti on-line agli utenti sono circa 103 terabyte al giorno (<https://earthdata.nasa.gov/eosdis/system-performance>). Il programma COPERNICUS riporta invece una crescita annua di circa 8 terabyte al giorno, con una consistenza complessiva stimata in circa 130 petabyte, che lo rende il data provider geografico più grande del mondo (<https://www.copernicus.eu/en/what-can-you-do-130-petabytes-data>). Le ragioni per il costante trend di aumento (cioè l'aumento della velocità di crescita dei dati da EO) è da ricercarsi in diverse ragioni:

1. Per il proliferare del numero di missioni e del numero di satelliti per missione. Attualmente sulla base della banca dati OSCAR (*Observing Systems Capability Analysis and Review Tool*) disponibile open access all'indirizzo <https://www.wmo-sat.info/oscar/satellites>, il numero di piattaforme di osservazione della terra operative sono 240 (ognuna talvolta con più satelliti come i Sentinel, i MODIS, Landsat, NOAA AVHRR, Planet, ecc.), includendo anche le missioni meteorologiche il totale tra missioni passate, presenti e previste per il prossimo futuro è pari a 779. Come conseguenza del diffondersi delle missioni con satelliti gemelli o con micro-satelliti si è assistito alla nascita di un telerilevamento in near-real-time con immagini ad alta risoluzione giornaliere. Vale la pena ricordare il caso forse più clamoroso, ovvero la nuova costellazione

2. Per l'aumento della risoluzione geometrica. I satelliti con risoluzione metrica o sub-metrica sono in costante proliferazione ma anche la risoluzione dei normali satelliti ottici come i Sentinel-2 è nove volte superiore alla precedente generazione di satelliti NASA Landsat.
3. L'aumento della profondità, ovvero della memoria occupata da ogni pixel di ogni immagine. Fino a dieci anni fa lo standard era di 8 bit (2^8 ovvero un range di 256 valori o tonalità di grigio). Oggi ci si spinge frequentemente a 11 o 12 bit.
4. Per l'aumento del numero di bande dei sensori ottici. Si moltiplicano le missioni da satellite iperspettrali come Hyperion da 220 bande spettrali con una risoluzione di 10 nm, lo strumento WIS con 812 bande, o le 128 bande a bordo del HJ-1^o, fino alle 250 bande dei sensori a bordo della missione Italiana PRISMA.
5. L'avvento di nuove tecnologie. Alle tradizionali tecnologie di acquisizione su base ottica passiva e radar a microonde si aggiunge da qualche anno la tecnologia lidar. Con alcune missioni per il monitoraggio della consistenza dei ghiacci (ICESAT) ma oggi in corso di sperimentazione in forma di batterie di sensori con acquisizioni in *fullwave form* dal programma GEDI (*Global Ecosystem Dynamics Investigation*) a bordo della Stazione Spaziale Internazionale.

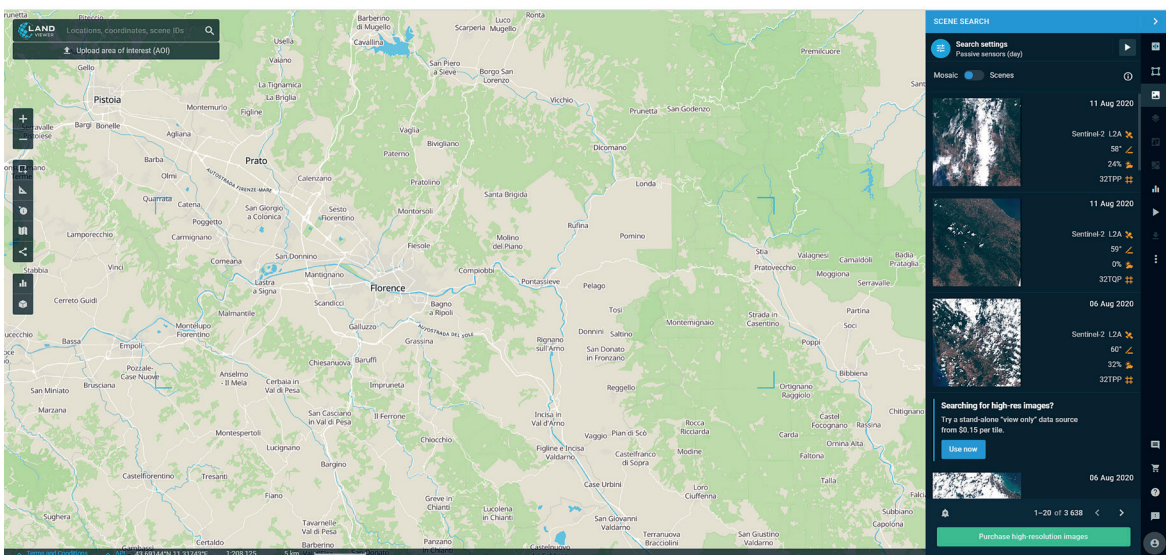
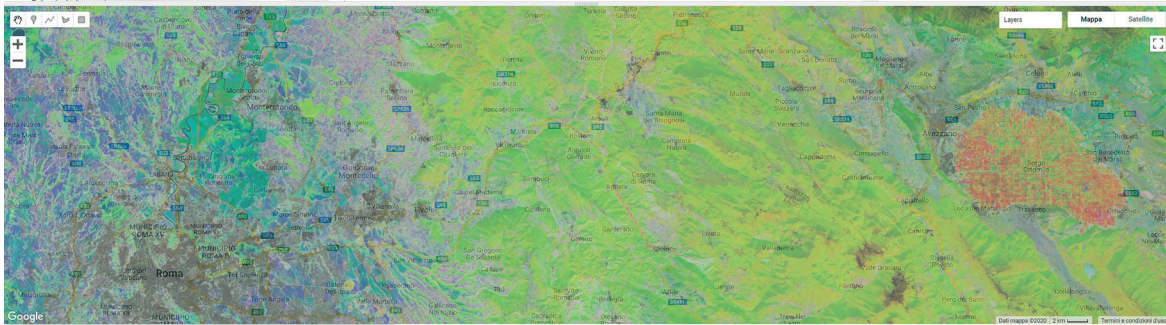


L'API on-line di GEE

Fig. 3

LandViewer un servizio on-line per la ricerca di immagini telerilevate basato sulla piattaforma AWS

Fig. 4



decennio passato, ha però il notevole fattore limitante, oltre al costo di implementazione e manutenzione, di dover creare una copia dei dati da telerilevamento in locale. Visto l'aumentato flusso dello streaming appare oggi molto più vantaggioso il secondo approccio.

Infatti nell'ultimo decennio, il cloud computing e gli altri sistemi HPC sono diventati disponibili per il ricercatore e il consumatore medio con schemi pay-per-use (Barr, 2006). Il concetto alla base del cloud computing è l'idea di offrire il computing come servizio: i clienti hanno una serie sempre più diversificata di esigenze e spesso desiderano poter personalizzare i propri sistemi computazionali a un livello piuttosto basso. Prima della *cloud computing*, i prodotti software generalmente presupponevano (1) la proprietà dell'hardware da parte dell'utente finale o (2) specificità di dominio piuttosto rigide. Con il cloud computing, l'idea è di fornire risorse computazionali gestite come un prodotto commerciale, consentendo all'utente finale di scegliere e distribuire il software sopra i livelli sottostanti, senza soluzione di continuità, hardware / rete / sistema operativo (Hegeman et al., 2014). Vediamo ora i principali sistemi di *cloud computing* partendo dai primi due che offrono servizi specifici per l'EO.

Google Earth Engine (GEE) è una piattaforma basata su cloud per l'analisi geospaziale su

scala planetaria che consente di accedere alle enormi capacità di calcolo di Google. GEE permette di sostenere una varietà di programmi/servizi di monitoraggio su diverse questioni sociali ad alto impatto tra cui deforestazione, siccità, disastri, malattie, sicurezza alimentare, gestione dell'acqua, monitoraggio del clima e protezione ambientale. È unico nel campo come piattaforma integrata progettata per potenziare non solo gli scienziati del telerilevamento tradizionale, ma anche un pubblico molto più ampio che non ha la capacità tecnica necessaria per utilizzare i supercomputer tradizionali o le risorse di cloud computing su larga scala (Gorelick et al., 2017)

GEE è costituito da un catalogo di dati pronto per l'analisi multi-petabyte insieme a un servizio di calcolo intrinsecamente parallelo ad alte prestazioni. Vi si accede e si controlla tramite un'API (*Application Programming Interface*) accessibile da Internet e un ambiente di sviluppo interattivo (IDE) associato basato sul web che consente la prototipazione rapida e la visualizzazione dei risultati.

L'uso di GEE è gratuito per gli utenti Google purchè non si richiedano risorse particolarmente elevate (necessarie per esempio per analisi globali multitemporali con dati a media o alta risoluzione. In questi casi è necessario stipulare accordi specifici

Big Data Analytics: dai dati alle informazioni
 Questa proliferazione di dati senza precedenti ha posto sfide significative nella gestione, elaborazione e interpretazione dei big data da EO. Sono stati compiuti grandi sforzi per incorporare il paradigma del calcolo ad alte prestazioni (*High Performance Computing* - HPC) nelle applicazioni di telerilevamento (Plaa et al., 2007). Gli approcci HPC stanno diventando dominanti per poter affrontare

gli enormi requisiti computazionali introdotti dalla necessità di processare grandi moli di dati da EO. L'approccio HPC all'analisi dei big data nel campo del telerilevamento si basa sulla scelta di due strategie alternative: 1) avere a disposizione un centro di calcolo ad alte prestazioni; 2) avvalersi di servizi di cloud computing sviluppati da terze parti. Il primo approccio, dominante fino al

con Google. Per lo storage dei risultati sono necessari account *Google Drive* di dimensioni adeguate.

Tra le poche alternative a GEE l'*Amazon Web Services* (AWS), una sussidiaria di Amazon che fornisce API e piattaforme di cloud computing su richiesta a privati, aziende e governi, con pagamento in base al consumo. Questi servizi web di cloud computing forniscono una varietà di infrastrutture tecniche astratte di base e blocchi e strumenti informatici distribuiti. Uno di questi servizi è *Amazon Elastic Compute Cloud* (EC2), che consente agli utenti di avere a disposizione un cluster virtuale di computer, sempre disponibile, tramite Internet. La versione dei computer virtuali di AWS emula la maggior parte degli attributi di un computer reale, comprese le unità di elaborazione centrale hardware (CPU) e le unità di elaborazione grafica (GPU); memoria locale / RAM; archiviazione su disco rigido / SSD; una scelta di sistemi operativi; networking; e software applicativo precaricato come server web. La tecnologia AWS è implementata nelle server farm di tutto il mondo e gestita dalla filiale di Amazon.

Simile nella filosofia ad AWS l'IBM Compute è un server cloud ad alte prestazioni che permette la virtualizzazione di sistemi HPC ottimizzando le configurazioni di base per soddisfare le esigenze di carico di lavoro in termini di RAM, SSD, GPU e altro. E'

organizzato in 60 data center. Altri servizi simili di cloud computing sono offerti da importanti multinazionali quali Oracle e Microsoft.

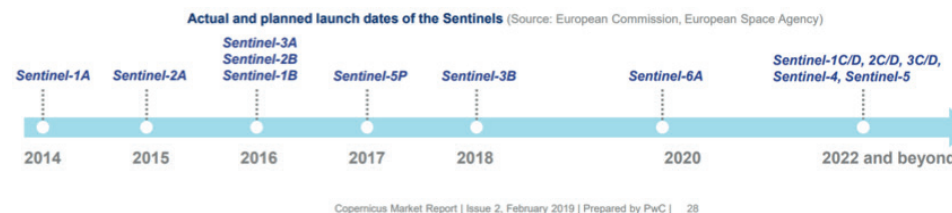
In questo quadro emerge chiaramente il ruolo di leader del programma Europeo Copernicus, attualmente il principale provider di informazioni da EO e quindi uno dei principali motori del mercato Big Data Analytics.

Il programma Europeo Copernicus

Copernicus è il programma dell'Unione Europea per l'osservazione della Terra e i servizi associati.

Dal 2008 al 2020, si prevede che gli investimenti totali nel programma Copernicus raggiungeranno gli 8,2 miliardi di euro. Nello stesso periodo, questo investimento genererà benefici economici compresi tra 16,2 e 21,3 miliardi di euro (esclusi i benefici non monetari). Questo valore economico è generato attraverso il valore aggiunto creato nell'industria spaziale a monte, e le vendite di applicazioni basate su Copernicus a valle.

Il programma Copernicus è cofinanziato dalla Commissione Europea e dall'ESA. L'ESA, insieme a EUMETSAT (*European Organisation for the Exploitation of Meteorological Satellites*) su aspetti specifici, è prevalentemente responsabile dello sviluppo delle componenti spaziali, denominate anche infrastrutture in Copernicus, mentre la Commissione europea è responsabile



Evoluzione temporale dei lanci delle Sentinelle Copernicus

Fig. 5

dell'implementazione e della gestione di tutte e tre le componenti: spaziale, in-situ, servizi. Copernicus riceve, elaborare e offre set di dati da varie fonti - i satelliti di EO (le proprie Sentinelle, e le altre missioni partner) e una moltitudine di sensori in situ a terra, in mare, o in volo.

Copernicus dispone di sei servizi tematici (Terra, Mare, Atmosfera, Clima, Emergenza e Sicurezza) supportando lo sviluppo di molte applicazioni. I servizi Copernicus elaborano e analizzano i dati, li integrano con altre fonti, offrendo *Geographic Information Systems* (GIS) ai propri utenti e al servizio delle autorità pubbliche e delle imprese commerciali. A partire dal 2017 la Commissione europea ha lanciato l'iniziativa di servizi di accesso on-line (DIAS - *Data and Information Access Services*) per sviluppare meglio la diffusione dei dati e delle informazioni Copernicus.

La maggior parte di questi consente solo il download e talvolta lo sfruttamento delle informazioni a cui si accede tramite un servizio di mappatura web (WMS). È da notare che attraverso appositi accordi di collaborazione molte delle più importanti

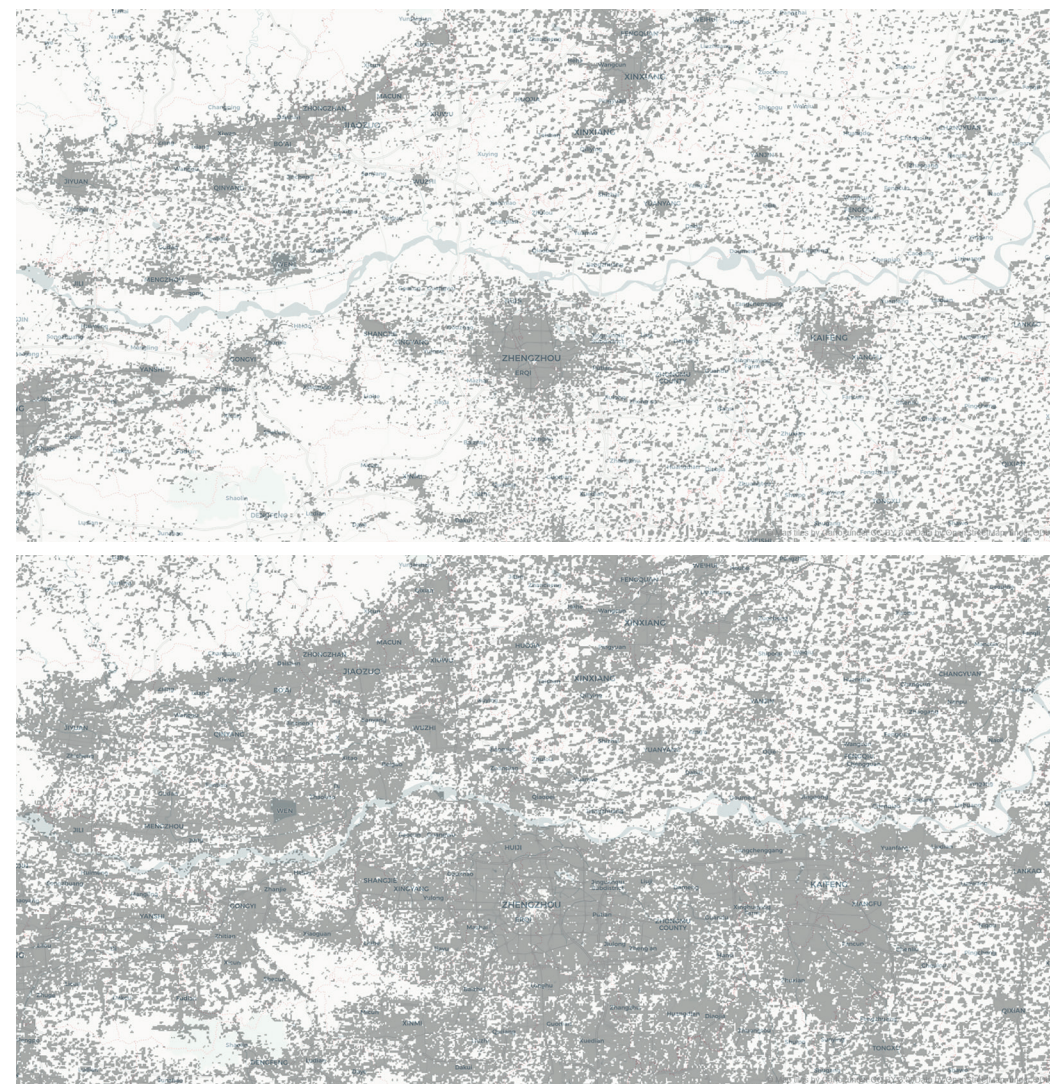
banche dati Copernicus sono oggi accessibili attraverso i due più comuni ambienti di sviluppo per il cloud computing: ovvero i già citati GEE e, in misura più ridotta, AWS. Sul piano nazionale è attivo il Forum Nazionale degli Utenti del Programma Copernicus, coordinato dall'Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA). Il Forum Nazionale consiste in uno strumento preposto alla raccolta dei requisiti degli utenti finali ed intermedi con l'obiettivo di giungere ad una posizione nazionale concordata ed armonizzata sia per gli sviluppi nazionali in seno alla Space Economy e sia verso il Forum e Comitato Europeo del Copernicus, coerentemente con le esigenze istituzionali, della ricerca, dell'impresa e dell'industria del Paese. Le attività del progetto Copernicus e del Forum sono state recentemente presentate in una giornata di studi svoltasi il 19/06/2020 presso l'Università degli Studi di Firenze (<https://www.isprambiente.gov.it/news/copernicus-il-programma-europeo-di-osservazione-della-terra-e-le-sue-applicazioni>).

Esempio dei dati disponibili sul TEP

Fig. 6a-6b

<https://urban-tep.eu/puma/tool/>

Per l'area di Zhengzhou nella Provincia di Henan in Cina l'area urbanizzata nel 1975 (in alto) e quella al 2014 (in basso) secondo il dataset GHSL.



La componente spaziale

La componente spaziale include lo sviluppo, il lancio e la gestione dei satelliti Sentinel e di una serie di missioni operate da partner.

In primo luogo, i Sentinel sono una famiglia di satelliti di proprietà dell'UE, che sono stati specificatamente progettati per soddisfare le esigenze dei servizi Copernicus e dei loro utenti. Attualmente, Sentinel-1, 2, 3 e 5p sono già stati lanciati (Figura 5). Sentinel-4, 5 e 6 sono attesi nei prossimi anni. Sentinel-1, 2, 3, 5p e 6 sono dedicati all'osservazione della Terra attraverso vari sensori (ad esempio radar, ottico, altimetrico), mentre Sentinel 4 e 5 sono satelliti meteorologici operati da EUMETSAT.

Le Sentinelle sono integrate da altre missioni, che esistono indipendentemente da Copernicus. Queste missioni sono gestite da ESA (ad esempio ERS), EUMETSAT (ad esempio METEOSAT), gli Stati membri dell'UE, altri paesi terzi (es. Landsat o MODIS negli USA) o fornitori commerciali (es. Deimos, Pleiadi). Queste missioni sono importanti perché forniscono a Copernicus dati ad altissima risoluzione che non possono essere ottenuti tramite le Sentinelle o compensare le Sentinelle che non sono state ancora lanciate.

La componente in situ

Copernicus si basa anche su dati in situ acquisiti a terra, per lo più appartenenti agli Stati membri dell'UE o a infrastrutture di ricerca internazionali, e che vengono messe

a disposizione di Copernicus tramite specifici accordi.

Sono incluse le osservazioni a terra, dal mare e dall'aria, così come i riferimenti geospaziali e i dati ausiliari concessi in licenza o forniti per l'uso nel programma Copernicus. Più recentemente, i dati in situ hanno trovato nuove fonti nei sensori e nelle immagini raccolte da veicoli aerei senza pilota (UAV) o da crowdsourcing (ad esempio OpenStreetMap). I dati in situ hanno due ruoli:

- Calibrare, integrare e convalidare i dati ottenuti dalle missioni satellitari per garantirne l'affidabilità nel tempo.
- Integrare le informazioni Copernicus dallo spazio per produrre prodotti e fornire servizi più vicini alle esigenze degli utenti finali.

La componente dei servizi

I servizi sfruttano i dati Sentinel, delle missioni partner e dei dati in situ, per fornire prodotti e informazioni a supporto commerciale, applicazioni istituzionali e di ricerca.

Il progetto Copernicus ha permesso di ottenere notevoli benefici, qui di seguito se ne analizzano i principali nel settore dell'urbanistica, e per i comparti agricoli e forestali. I dati sono ottenuti dall'ultimo Copernicus Market Report 2019 disponibile on-line https://www.copernicus.eu/sites/default/files/2019-02/PwC_Copernicus_Market_Report_2019_PDF_version.pdf

Urbanistica

Gli utenti intermedi dei dati forniti da Copernicus sono principalmente società di servizi a valore aggiunto, startup e PMI, che preparano poi prodotti e servizi per gli utenti finali, che sono prevalentemente autorità locali, regionali e nazionali. Tra i dati forniti da Copernicus quelli più utilizzati in campo urbanistico sono i dati radar Sentinel-1, utilizzati per il rilevamento delle modifiche all'edificato e per lo sviluppo di modelli

3D. I dati ottici Sentinel-2 forniscono dati rilevanti per il monitoraggio della crescita urbana grazie alla elevata risoluzione spaziale. A partire da queste fonti di dati sono stati sviluppati importanti dataset globali come il *Global Human Settlement Layer* (GHSL) coordinato dal *Joint Research Centre* della Commissione Europea (<https://ghsl.jrc.ec.europa.eu/>) e il *World Settlement Footprint* 2015 coordinato dall'European Space Agency (<https://www.esa.int/>

[Applications/Observing_the_Earth/Mapping_our_global_human_footprint](#)). Entrambi i progetti hanno portato alla produzione di un vasto spettro di strati informativi inerenti lo studio dell'evoluzione temporale delle aree urbanizzate, tutti consultabili, insieme ad altre banche dati tematiche inerenti, attraverso la piattaforma *Urban Thematic Exploitation Platform* (TEP) <https://urban-tep.eu/puma/tool/>.

Non mancano d'altra parte gli strati informativi sviluppati direttamente dal *Copernicus Land Monitoring Service* (CLMS) quali le immagini telerilevate che possono essere utilizzate per la creazione di prodotti o i dati cartografici già elaborati come il *Corine Land Cover*, gli *High Resolution Layers*, o il *Land Surface Temperature*, tutti disponibili on-line <https://land.copernicus.eu/>.

I risultati delle ricerche inerenti lo studio globale dei trend di popolazione dipingono un quadro estremamente grave. Secondo il *World Urbanization Prospects 2018* (<https://population.un.org/wup/>) con grandi concentrazioni di persone, proprietà e ricchezza, le città si stanno espandendo come mai prima d'ora, con la metà della popolazione globale che vive nelle aree urbane. Il mondo si sta urbanizzando con un tasso di crescita mai così alto: oltre il 50% della popolazione mondiale vive nelle città, che crescerà al 66% entro il 2050, su un totale di 9,8 miliardi di persone. A fianco della moltitudine di

opportunità in rapida espansione, le città devono affrontare anche un certo numero di sfide. Disagi economici, conflitti sociali e disastri ambientali si verificano sempre più frequentemente proprio nelle città che si stanno espandendo più velocemente. La popolazione urbana mondiale sta crescendo quattro volte più velocemente della popolazione rurale. Il 90% di questa crescita si sta verificando in regioni in via di sviluppo e l'Africa attualmente ha il tasso di crescita urbana più veloce, a più del 5% all'anno. Evidentemente l'espansione urbana deve essere monitorata per poter garantire che proceda in forma sostenibile, senza compromettere le risorse ambientali e territoriali, né peggiorare la qualità della vita e la sicurezza della popolazione urbana. L'osservazione della Terra, in questo senso, rappresenta uno strumento fondamentale per la generazione di informazioni utili a supportare le più opportune scelte in ambito urbanistico. Oltre ai già citati servizi finalizzati alla mappatura multitemporale delle aree urbane si annoverano il monitoraggio delle temperature e delle isole di calore, il monitoraggio del verde urbano, analisi dell'*urban sprawl*, il supporto alla gestione e alla pianificazione delle aree di trasporto, alla ricostruzione tridimensionale dell'edificato, e le deformazioni derivanti da fenomeni come terremoti, frane e alluvioni.

Agricoltura

L'agricoltura è stato uno dei primi mercati del telerilevamento e continua a esserci un crescente interesse per il modo in cui i prodotti da telerilevamento possono supportare le sfide affrontate dal settore. Le immagini da satellite possono dare un contributo significativo all'efficienza di moderne pratiche agricole come la *precision agriculture*. Spesso i dati da EO vengono combinati con i dati in situ (Di Biase et al., 2018) al fine di produrre risultati più mirati e localizzati per l'agricoltore, mentre i dati aerei o da droni possono fungere da dati complementari.



















Nonostante la lunga tradizione di uso, il mercato del telerilevamento soffre di una forte frammentazione con utenti intermedi che includono una varietà di attori: start-up, piccole e medie imprese (PMI), grandi società e attori puramente scientifici come organismi di ricerca e università. Gli utenti finali sono abbastanza bilanciati tra attori pubblici (come le autorità governative che necessitano di precise informazioni per politiche e regolamenti specifici, o per monitorare questioni specifiche) e attori privati (cooperative agricole, aziende alimentari, consulenti agronomici, ecc.). Gli agricoltori, che dovrebbero essere i principali utenti finali dei prodotti agricoli basati sui dati EO, incontrano invece una certa difficoltà di accesso ai servizi Copernicus a causa della mancanza di infrastrutture (ad esempio

una buona connessione a banda larga) o di conoscenza digitale.

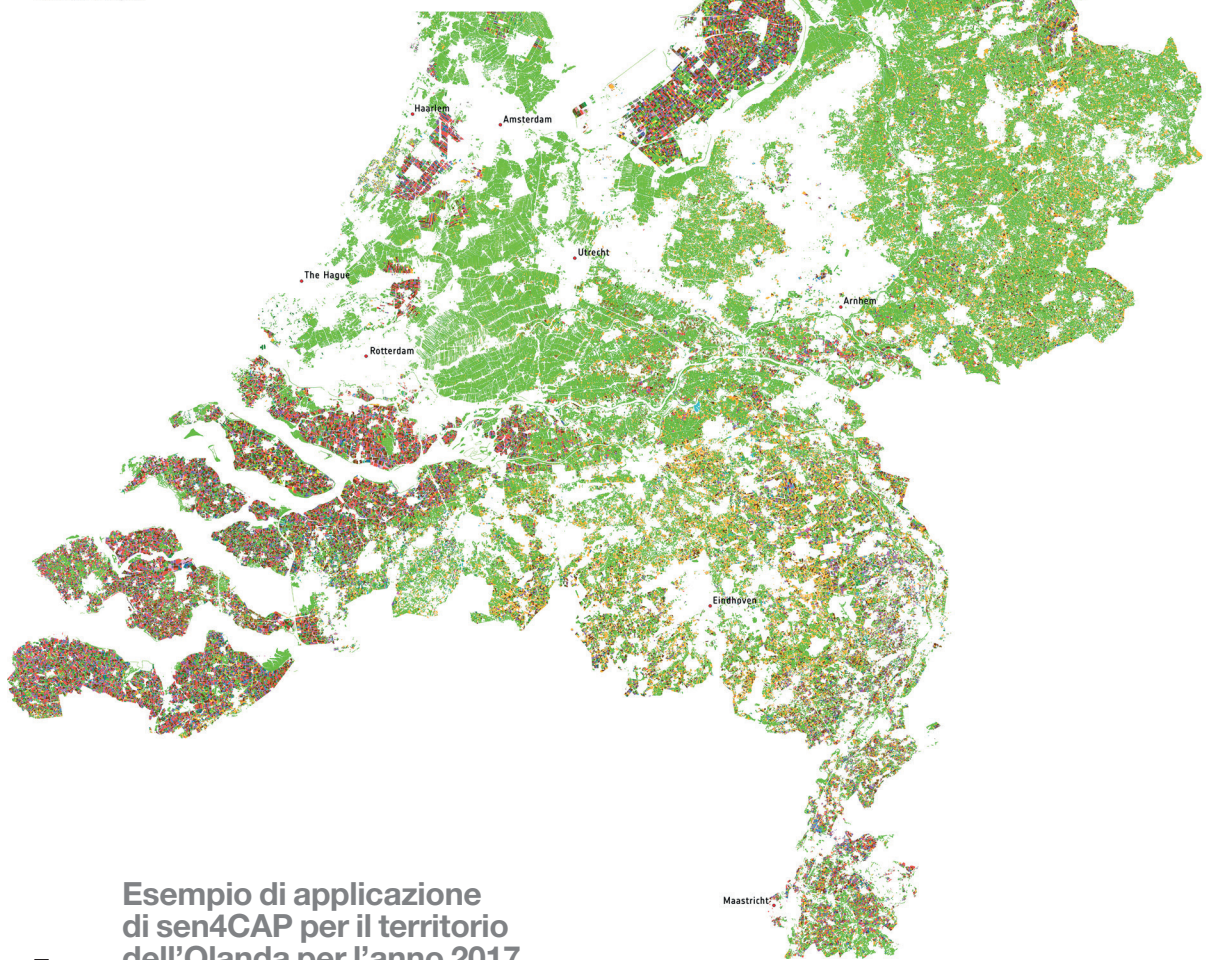
Il settore agricolo gioca un ruolo cruciale nell'economia globale, ed è in rapida evoluzione a causa dell'emergere e lo sviluppo di numerose innovazioni tecniche, in un quadro di rilevanti dinamiche ambientali in rapida evoluzione. Negli anni futuri, l'agricoltura dovrà affrontare diverse pressioni significative. La popolazione mondiale dovrebbe crescere dagli attuali 7,4 miliardi di 2018 a 9,8 miliardi entro il 2050, causando la necessità di accrescere la produzione di cibo dagli attuali 8,4 miliardi di tonnellate a circa 13,5 miliardi di tonnellate all'anno. I cambiamenti climatici e l'intensificazione nella concorrenza per terra, acqua ed energia imporranno scelte strategiche per assicurare una corretta ma sostenibile produzione agricola.

La *Food and Agriculture Organization* delle Nazioni Unite (FAO) riporta che per il 2018 il numero di persone con alimentazione insufficiente è aumentata dai circa 804 milioni del 2016 agli 821 milioni del 2017. In Europa il settore è sostenuto dalla Politica Agricola Comunitaria (un investimento di 60 miliardi di euro all'anno) che sostiene gli agricoltori Europei nel difficile lavoro di mantenere l'attività assicurando una vantaggiosa produzione di alimenti sicuri, riducendo l'impatto sull'ambiente e sul clima, e proteggendo la biodiversità.

Agricultural landscape of the Netherlands

	Oat		Tulip
	Sugar beet		Winter carrot
	Clover root		Winter wheat
	Winter barley		Maize
	Summer barley		Permanent crops
	Lilium		Grass
	Medicago		Summer wheat
	Onion		Fallow
	Potato		Other crops

Credits: Contains modified Copernicus Sentinel data (2017) and Land Parcel Information System (LPIS) of Netherlands, processed by ESA-Sen4CAP (led by UCLouvain with CS-Romania, e-GEOS, GISAT and Sinergise)



Esempio di applicazione di sen4CAP per il territorio dell'Olanda per l'anno 2017

Fig. 7

Da https://www.esa.int/ESA_Multimedia/Images/2018/05/Crop_map.

Le applicazioni basate su EO possono aggiungere valore a una serie di sottosectori agricoli, comprese le attività relative alle colture (ad es. monitoraggio della salute delle colture, inventari delle colture, monitoraggio del tipo di colture), gestione delle risorse idriche (es. umidità del suolo, irrigazione delle colture, monitoraggio della siccità) e ottimizzazione e gestione dell'uso del suolo (per esempio per il monitoraggio della biodiversità). Alcuni esempi significativi dell'utilizzo di big data è il riconoscimento e la mappatura automatica delle principali tipologie di colture agricole basate sull'analisi della firma fenologica di scene Sentinel2. L'applicazione, sviluppata da ESA è disponibile all'indirizzo <http://esa-sen4cap.org> e il suo utilizzo operativo è previsto per il supporto ai controlli nei pagamenti dei contributi di supporto alla PAC. In Figura 7 l'applicazione di sen4CAP su tutto il territorio dell'Olanda.

I prodotti derivanti da sen4CAP dovrebbero essere a breve disponibili sulla piattaforma Copernicus sia in termini di serie multitemporali di osservazioni della fenologia ad alta risoluzione (10 m) (*High Resolution Vegetation Phenology and Productivity*) sia

in termini di mappe annuali dei principali tipi di coltivazione (*High Resolution Layer Crop Types*).

Foreste

Le foreste in Europa coprono una superficie di circa 215 milioni di ettari (circa il 33% delle terre emerse), questa superficie è in costante aumento. Lo stato di salute delle foreste è considerato in modo positivo, le fonti di stress che producono danni (incendi, danni da vento, attacchi di patogeni) riguardano solo circa il 3% delle foreste Europee. Preleviamo legno e legname dalle nostre foreste ma a un ritmo che, se pur in crescita, è ancora ampiamente sostenibile.

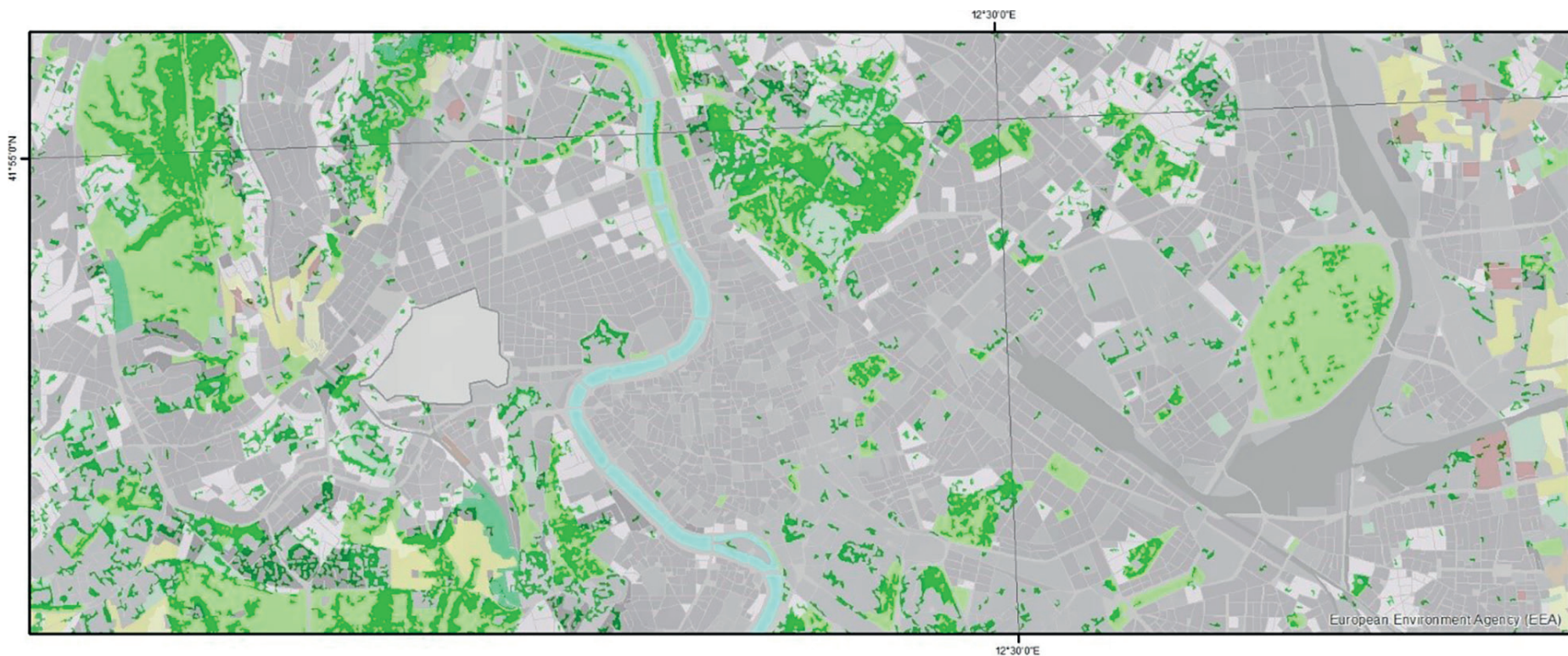
Le foreste in Europa assimilano ogni anno 719 milioni di tonnellate di CO₂, circa il 9% delle nostre emissioni. Le nuove tecnologie impiegate nella trasformazione del legno permettono oggi la creazione di prodotti che possono sostituire materiali tradizionali (come cemento e acciaio) la cui produzione determina enormi emissioni di CO₂ (Forest Europe, 2015)

Anche il settore forestale, come quello agricolo, è riconosciuto come uno degli utilizzatori storici delle informazioni derivanti dal telerilevamento. Iniziative globali,

Per il centro di Roma, sullo sfondo dello Urban Atlas, in verde le informazioni dello Street Tree Layer.

Fig. 8

Entrambi i prodotti sono vettoriali e distribuiti dal servizio Copernicus Land.



come il programma delle Nazioni Unite sulla riduzione delle emissioni dovute alla deforestazione e al degrado delle foreste (*Reducing Emissions from Deforestation and Forest Degradation - REDD*), i programmi permanenti di monitoraggio delle risorse forestali a livello globale (FAO - *Global Forest Resource Monitoring*) ed Europeo (*Forest Europe - Interministerial Conference on the Protection of Forests in Europe*) richiedono ai paesi partecipanti la disponibilità di dati estremamente accurati e precisi sulle foreste. Le foreste producono una vasta gamma di servizi ecosistemici: produzione di prodotti come legno, legname e cellulosa, protezione dal rischio idrogeologico (alluvioni, frane e valanghe), protezione della biodiversità,

capacità di contrastare il cambiamento climatico grazie all'assimilazione della CO₂ atmosferica, socio culturale (sostegno alle popolazioni che vivono grazie al reddito derivante dalla filiera foresta-legno ma anche per la fruizione ricreativa dei boschi e al loro valore culturale, estetico e paesaggistico). Una gestione sostenibile delle risorse forestali ha per obiettivo il mantenimento e la conservazione per le future generazioni della capacità di produrre servizi ecosistemici. Questo evidentemente si può raggiungere solo attraverso la quantificazione dei servizi ecosistemici attraverso indicatori spazializzati. Per molti di questi indicatori i dati da telerilevamento possono essere utilizzati con successo.

Entrambi i dati Sentinel-1 e 2 sono fonti riconosciute di preziose informazioni per la mappatura della superficie forestale e della copertura arborea (i così detti *Trees Outside Forests* o *Small Woody Features*), la stima della biomassa e del carbonio atmosferico assimilato (Puletti et al., 2018). Le immagini satellitari permettono anche l'individuazione delle principali forme di disturbo derivanti da fonti biotiche (insetti e patogeni) e abiotiche (utilizzazioni, incendi, danni da vento). Gli utenti intermedi dei dati da EO nel dominio forestale includono diversi attori privati (dalle microimprese ai più grandi attori), istituti di ricerca pubblici e organizzazioni di gestione forestale.

I prodotti EO finalizzati alla silvicoltura sono attualmente utilizzati principalmente da utenti finali pubblici (90%), piuttosto che da utenti privati (10%). Tra i servizi Copernicus più utilizzati in campo forestale vi sono gli *High Resolution Layers* per la mappatura della densità della copertura arborea (*Tree cover density*), dei *Dominant leaf type* (latifoglie vs. conifere), e del *Forest type product* specificatamente sviluppato per coincidere il più precisamente possibile con la definizione internazionale della FAO utilizzata in campo inventariale. A queste informazioni si aggiungono quelle relative alla copertura di alberi fuori foresta: gli *Small Woody Features* e lo *Street Tree Layer* in ambiente urbano (Figura 8).

Bibliografia

Conclusioni

Le tecnologie di osservazione della Terra costituiscono un chiaro caso di big data *ante litteram*. In questo contributo abbiamo cercato di presentare le principali motivazioni che hanno portato all'aumento della velocità di produzione di questi dati: più missioni, più satelliti per missione, maggiore risoluzione geometrica, maggiore risoluzione spettrale e radiometrica, internet più veloce, democratizzazione nell'accesso ai dati (oggi spesso gratuito).

In questo scenario spicca il programma Copernicus che, tramite uno sforzo economico senza pari della Commissione Europea e dell'Agenzia Spaziale Europea (ESA), è ad oggi il principale distributore globale di dati e informazioni derivanti, o comunque collegate, alle tecnologie di EO.

La proliferazione dei dati dà luogo a una ulteriore crescente complessità dei dati da EO intrinsecamente caratterizzati già da una elevata multidimensionalità.

L'elaborazione di questi dati per la produzione di informazioni (*Big Data Analytics*) non può basarsi sugli approcci modellistici tipici di un periodo storico nel quale la mole di dati era infinitamente inferiore. Ma deve invece poter esplorare più in profondità la complessità tipica dei big data, con l'uso di approcci di data mining basati sull'intelligenza artificiale.

Tali approcci sono però ad alta intensità computazionale e per essere operati in tempi ragionevoli richiedono sistemi HPC che devono

includere: (1) cluster di supercomputer e piattaforme di cloud computing ottimizzati per carichi ad alta intensità di dati; (2) file system e database paralleli che si basino sulla disponibilità di accesso veloce ai dati; (3) strumenti di gestione dei dati per il controllo del posizionamento dei dati in memoria locale dei dati multilivello; (4) pianificazione delle attività di elaborazione in parallelo. Operare con tali requisiti attraverso infrastrutture di calcolo locali appare oggi poco efficace.

In questo scenario si sono sviluppati diversi servizi commerciali che permettono di operare attraverso sistemi HPC virtuali residenti su servizi di cloud esterni. Nel panorama dei competitor più importanti dell'IC quali Amazon, Oracle, IBM e Microsoft spicca Google Earth Engine, ad oggi la piattaforma più utilizzata, specie nel settore della ricerca. È probabile che nel prossimo futuro possano emergere alcuni nuovi problemi con l'ulteriore aumento della quantità e dell'applicazione diffusa dei dati da telerilevamento. Tra queste la crescente domanda di capacità di elaborazione in tempo reale o in tempo quasi reale da parte di molte applicazioni EO per fornire informazioni ritenute critiche dal punto di vista temporale (per motivi di sicurezza o per emergenze ambientali). In questi casi il trattamento dei dati dovrà essere fatto in tempi brevissimi e quindi saranno necessarie infrastrutture di calcolo parallelo specificatamente destinate.

Barr J. 2006, *Amazon EC2 Beta*, Amazon Web Services Blog, <<http://aws.typepad.com/aws/2006/08/amazon-ec2-beta.html>> (05/20)

Boyd D., Crawford K. 2012, *Critical questions for big data*, «Information, Communication and Society», vol. 15, n. 5, pp. 662-679.

Di Biase R.M., Fattorini L., Marchi M. 2018, *Statistical inferential techniques for approaching forest mapping. A review of methods*, «Annals of Silvicultural Research», vol. 42, n. 2, pp. 46-58.

FOREST EUROPE 2015, *State of Europe's Forests*, Ministerial Conference on the Protection of Forests in Europe, <<https://www.forest-europe.org/docs/fullsoef2015.pdf>> (07/20)

Francini S., McRoberts R.E., Giannetti F., Mencucci M., Marchetti M., Scarascia Mugnozza G., Chirici G. 2020, *Near-real time forest change detection using PlanetScope imagery*, «European Journal of Remote Sensing», in press.

Gorelick N., Hancher M., Dixon M., Ilyushchenko S., Thau D., Moore R. 2017, *Google Earth Engine: Planetary-scale geospatial analysis for everyone*, «Remote Sensing of Environment», vol. 202, pp. 18-27.

Hegeman J.W., Sardeshmukh V.B., Sugumaran R., Armstrong M.P. 2014, *Distributed LiDAR data processing in a high-memory cloud-computing environment*, «Annals of GIS», vol. 20, pp. 255-264.

Kitchin, R. 2013, *Big data and human geography: Opportunities, challenges and risks*, «Dialogues in Human Geography», vol. 3, n. 3, pp. 262-267.

Laney, D. 2012, *3D Data Management: Controlling Data Volume, Velocity and Variety*, <<https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-DataVolume-Velocity-and-Variety.pdf>> (08/20).

Leach N., Coops N.C., Obrknezev N. 2019, *Normalization method for multi-sensor high spatial and temporal resolution satellite imagery with radiometric inconsistencies*, «Computer and Electronics in Agriculture», vol. 164 <<https://doi.org/10.1016/j.compag.2019.104893>> (08/20).

Liu P., Di L., Du Q., Wang L. 2018, *Remote Sensing Big Data: Theory, Methods and Applications*, «Remote Sensing», vol. 10, pp. 711.

Ma Y., Wu H., Wang L., Huang B., Ranjan R., Zomaya A., Jie W. 2015, «*Remote sensing big data computing: Challenges and opportunities*, Future Generation Computer Systems», vol. 51, pp. 47-60.

Mayer-Schonberger V., Cukier K. 2013, *Big Data: A Revolution that Will Change How We Live, Work and Think*, John Murray, London.

Mashey J.R. 1998, *Big Data ... and the Next Wave of InfraStress*, Slides from invited talk, Usenix. <https://static.usenix.org/event/usenix99/invited_talks/mashey.pdf> (08/20).

Antonio J., Plaa Chein I. 2007, *Chang High Performance Computing in Remote Sensing*, Chapman & Hall/CRC.

Puletti N., Chianucci F., Castaldi C. 2018, *Use of Sentinel-2 for forest classification in Mediterranean environments*, «Annals of Silvicultural Research», vol. 42, n. 1, pp. 32-39.

Rathore M.M., Ahmad A., Paul A., Rho S. 2016, *Urban planning and building smart cities based on the Internet of Things using Big Data analytics*, «Computer Networks», vol. 101, pp. 63-80.

Skytland N. 2012, *Big data: What is NASA doing with big data today*, Open. Gov open access article <<https://open.nasa.gov/blog/what-is-nasa-doing-with-big-data-today/>> (09/20).

Jin J., Gubbi J., Marusic S., Palaniswami M. 2014, *An information framework for creating a smart city through Internet of things*, «IEEE Internet Things J.», vol. 1, n. 2, pp. 112-121.

Kitchin A. 2014, *The real-time city? Big data and smart urbanism*, «Geojournal», vol. 79, pp. 1-14.

Kuhn T. 1962, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.

Han J., Kamber M.P. 2011, *Data Mining: Concepts and Techniques*, 3rd ed., Morgan Kaufmann, Waltham.

Hastie T., Tibshirani R., Friedman J. 2009, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., Springer, New York.

Sabins F.F. 1978, *Remote Sensing: Principles and Interpretation*, Freeman.