8

Satellite Analysis Ready Data for the Sustainable Development Goals

Brian D. Killough

ABSTRACT

Satellite data efficiently prepared into analysis ready format have the potential to exploit the growing volumes of free and open data to address the United Nations Sustainable Development Goals (SDG). Combining these data with modern data management technologies, such as the Open Data Cube, allows users to rapidly produce analytical products and utilize time series data sets for change detection and interoperable use of diverse data sets. To demonstrate the potential of these data and systems to address the SDGs, several examples will be shown from the Africa Regional Data Cube. It is these use cases that demonstrate the value and impact of Earth observation data for a sustainable future.

8.1. INTRODUCTION

With each passing year, new generations of Earth observation (EO) satellites are creating increasingly significant volumes of free and open data with comprehensive global coverage, such that the lack of data is no longer an issue. In addition to these data, research and development activities have delivered new applications that offer significant potential to deliver great impact to important environmental, economic, and social challenges, including at the local, regional, and global scales. These applications highlight the value of EO, though the challenge is in providing the proper connections between data, applications, and users. Even today, much of the archived EO satellite data is underutilized despite modern computing and analysis infrastructures.

Addressing this challenge is difficult for advanced economies and even more challenging for developing countries with an interest in using EO satellite data. It is simply not technically feasible or financially affordable to consider traditional local processing and data distribution methods (e.g., scene-based file download over the Internet) to address this scaling challenge in many economies, as the size of the data and complexities in preparation, handling, storage, and analysis remain significant obstacles. Fortunately, just as satellite Earth observation technology has advanced significantly, so too has information technology. The data management and analysis challenges arising from the huge increase in free and open data volumes can be overcome with new computing technologies (e.g., distributed computing clouds) and data architectures (e.g., data cubes). Solutions such as the Open Data Cube (ODC) have a great potential to streamline data distribution and management for providers while simultaneously lowering the technical barriers for users to exploit the data to their full potential.

8.2. ANALYSIS READY DATA

Many satellite data users lack the expertise, infrastructure, and Internet bandwidth to efficiently and effectively access, preprocess, and utilize the growing volume of

NASA Langley Research Center, The College of William and Mary, Hampton, Virginia, USA

Earth Observation Applications and Global Policy Frameworks, Geophysical Monograph 274, First Edition.
Edited by Argyro Kavvada, Douglas Cripe, and Lawrence Friedl.

DOI:10.1002/9781119536789.ch08

space-based data for local, regional, and national decision making. Even sophisticated users of EO data typically invest a large proportion of their effort into data preparation. This is a major barrier to full and successful utilization of space-based data and threatens the success of major global and regional initiatives, such as the United Nations SDG initiative and the goals of the Group on Earth Observations (GEO). As data volumes grow, this barrier is becoming more significant for all users.

Countries and international organizations have expressed a desire for support from the Committee on Earth Observation Satellites (CEOS) to facilitate access to and processing of satellite data into Analysis Ready Data (ARD) products. Systematic and regular provision of ARD will greatly reduce the burden on global satellite data users. The provision of these data is possible through many options including systematic processing and distribution, processing on hosted platforms, and processing via toolkits provided to users.

ARD products are intended to be flexible and accessible products suitable for a wide range of users for a wide variety of applications, including time series analysis and multisensor interoperability. They are also intended to support rapid ingestion and exploitation via high-performance computing, cloud computing, and other future data architectures. However, ARD products may not be suitable for all purposes, and are not intended as a replacement for other types of satellite products. As an example, ARD is the cornerstone of the ODC initiative, as this architecture depends on ARD to allow efficient creation of data cubes and subsequent analyses. The result of this effort will be improved interoperability among data sets, facilitating time series analyses and enhanced global use and scientific value of satellite data.

The call for ARD across many satellite agencies, users (Giuliani et al., 2017), and in other communities, is being driven by practical requirements to reduce the demand for limited resources and expertise in the preparation of data, and to ensure that those preparatory steps are fully accounted and understood. These drivers are becoming stronger as the community seeks to apply rapidly growing data volumes from a range of sensors in new ways, especially as time series. Effective implementation of ARD requires that the definition that follows and product specifications are widely accepted and understood so that the ARD products deliver their intended benefit to end users and can address the SDGs (Kavvada & Held, 2018).

By definition, Analysis Ready Data (ARD) are satellite data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and interoperability both through time and with other data sets. This definition of ARD is not exclusive or prescriptive. It is expected that a range of data products will be produced by satellite data providers to meet the needs of a diverse user community, and that many of those products will be fully ready for their users. However, the definition of ARD reflects the attributes of fundamental measurement products for the majority of global remote sensing users and are the minimum level required to support time series analysis and data interoperability. ARD are therefore geophysical measurements that are comparable in space and time, with sufficient per-pixel (or observation-level) metadata to enable users to select observations of interest for their analyses.

CEOS is currently developing the definition and requirements for three land-based ARD products. These include optical moderate resolution (e.g., Landsat), surface reflectance and surface temperature, and radar moderate resolution (e.g., Sentinel-1) normalized backscatter. Though the details are left in other documents, each of these ARD products includes minimum requirements for general metadata, per-pixel metadata, radiometric correction, and geometric correction. Optical measurements require additional corrections for atmospheric absorption and scattering and solar and viewing angle corrections. Radar measurements require additional corrections for terrain topography and incidence angle.

The CEOS ARD framework presents two levels of requirements, threshold and target. Products that meet all threshold requirements should be immediately useful for scientific analysis or decision making. Products that meet target requirements will reduce the overall product uncertainties and enhance broad-scale applications. For example, the products may enhance interoperability or provide increased accuracy through additional corrections that are not reasonable at the threshold level. Target requirements anticipate continuous improvement of methods and evolution of community expectations, which are both normal and inevitable in a developing field.

CEOS Analysis Ready Data for Land (CARD4L) are satellite data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and interoperability both through time and with other data sets. This CEOS version of ARD (CARD4L) maintains consistency among many different methods of ARD production and makes it easier for users to progress toward decision-making products and maintain the accuracy of these products in space and time. Without CARD4L, there may be issues with spatial pixel alignment in time series or pixel geolocation as well as a lack of critical metadata. That said, CARD4L is a critical step in data production and may not be possible for all data providers given the additional work and documentation needed to meet the specifications.

ARD can contribute and support the generation of SDG indicators by reducing the complexity of data preparation and allowing a high level of spatial and temporal consistency. The case studies in this document demonstrate how these data might be used to support SDGs, but users should consult the specific SDG indicator documents to understand the established methodologies and standards required to develop products suitable for U.N. submission. This is ultimately the responsibility of country-level statistical agencies, which might have access to any number of technical tools and data.

8.3. OPEN DATA CUBE

As the world develops, so does its knowledge of, and demand for, EO satellite data. The primary problems for users are data access, data preparation, and efficient analyses to support user applications. CEOS, through its network of global connections, has determined that global users share many common needs including (1) minimizing time and expertise required to access and prepare satellite data; (2) access to free and open EO satellite data and application algorithms; (3) access to open-source software solutions that are advanced through community contributions; (4) use of consistent data architectures that allow sharing of code, tools, and algorithms; (5) use of efficient time series analyses to support land change applications; (6) use of multiple data sets together (e.g., interoperability and complementarity); (7) use of common geographic information system (GIS) tools; (8) flexible data architectures (e.g., local, cloud) that avoid commercial and Internet dependence; and (9) sustained training and community support.

It is these common needs that led the CEOS organization toward the initiation of the Open Data Cube (ODC) initiative (Killough, 2018; Ross et al., 2017), which seeks to provide a free and open data architecture solution that has value to its global users and increases the impact of EO satellite data. The ODC is based on the implementation approach used by the Australian Geoscience Data Cube (AGDC; Lewis et al., 2017) and the Digital Earth Australia program (Gavin et al., 2018), but is modified to allow globalization for a diverse set of users, data sets, and deployment options. Though the ODC is a single approach to data management, it is not the only valid approach for managing data in cube formats. For example, Google Earth Engine (GEE) uses a similar data management approach and has transformed the EO satellite data user community. More recently, the European Commission (EC) has launched an initiative to develop Copernicus Data and Information Access Services (DIAS) and EarthServer, the European Union Big Data initiative, uses an OGC-compliant RASDAMAN array database. Each of these data exploitation methods is quite similar to the ODC, but offers alternative options for implementation.

In response to user demand, such technological solutions remove the burden of data preparation, yield rapid results, and foster an active and engaged global community of contributors. Hence, CEOS is committed to stewarding and contributing to the ODC architecture as part of the ODC community. They seek to encourage others to join the initiative with a goal to meet the targeted needs of users, similar to the objectives of the AGDC, GEE, DIAS, and EarthServer, but differing in implementation.

The objective of the ODC is to increase the impact of satellite data by providing an open and freely accessible exploitation tool (Giuliani, 2019), and to foster a community to develop, sustain, and grow the breadth and depth of applications. This solution intends to support key objectives, which include building the capacity of users to apply EO satellite data and to support global priority agendas, such as those found in the United Nations Sustainable Development Goals (UN-SDG) and the Paris and Sendai Agreements. To ensure success, the ODC plans to develop and foster an open-source ODC community that is actively engaged and contributes to the core code, shares algorithms, and provides support to each other for the resolution of problems.

Successful ODC implementations exist in over 100 countries including: Australia, Colombia, Switzerland, Taiwan, and Africa. In the case of Africa, the entire continent has adoped the ODC as the core infrastructure for the Digital Earth Africa program, initiated in 2019. Since its inception in early 2017, the ODC has made significant progress in the advancement of open-source software tools and algorithms that support the deployment and operation of data cubes on local or cloud computing systems around the world. In addition, all of the current and future data cubes depend on Analysis Ready Data (ARD) to ensure efficient time series analyses and data interoperability (Killough et al., 2020).

The ODC community has made substantial progress in the development of tools and algorithms to support country-level deployment prototypes. For example, NASA's CEOS Systems Engineering Office developed a Web-based user interface that includes 19 sample data cubes from around the world (for example, Bangladesh, Cameroon, Colombia, Ghana, Honduras, Kenya, Samoa Islands, South Africa, Togo, Tonga, Vietnam, Uruguay) and 9 common applications (custom mosaics, water extent, water quality, coastal change, fractional cover, urbanization, NDVI anomaly, landslide risk, spectral indices) that give users experience with data cubes and allow exploration of potential application products. For more advanced users, the ODC community has developed Jupyter notebooks that use Python programming code to run additional application algorithms (e.g., change detection, land classification clustering, machine learning water detection with radar).

As new generations of EO data create increasingly larger volumes of data, global users will struggle to prepare and manage those data to address their local, regional, and national decision-making needs. The ODC initiative provides an innovative, free, and open-data architecture solution that lowers the technical barrier for global users to exploit satellite data and optimize its societal benefit. As data cubes expand around the world and become operational, there will be an increased potential to address local, regional, and global needs.

8.4. AFRICA REGIONAL DATA CUBE

In May 2018, the Global Partnership for Sustainable Development Data (GPSDD), the Committee on Earth Observation Satellite (CEOS), Amazon, and Strathmore University announced the release of the Africa Regional Data Cube (ARDC) to support five countries: Kenya, Senegal, Sierra Leone, Ghana, and Tanzania (Killough, 2019). The ARDC is focused on building the capacity of users in this region to apply Earth observation satellite data to address local and national needs as well as the objectives of the Group on Earth Observations (GEO) and the United Nations Sustainable Development Goals (UN-SDG). The ARDC supports several key users, including government ministries, national statistical agencies, geographic institutes, and research scientists.

The ARDC has a data volume of 11 TB and is hosted on the Amazon Web Services (AWS) cloud. To date, the ARDC includes only Landsat analysis-ready data since the year 2000. In addition to the Landsat data, the ARDC is working to process Copernicus data from the Sentinel-1 and Sentinel-2 missions. In the case of Sentinel-2, the surface reflectance ARD requirements are the same as those for Landsat. In the case of Sentinel-1, CEOS has recently developed new product family specifications for radar backscatter intensity, which is commonly accepted as the simplest form of ARD from radar missions. In late 2019, the ARDC will be expanded to include this Sentinel data, which will greatly enhance the output products for SDGs.

Since the release of the ARDC, there have been several interactions with local Africa users, including countries outside the ARDC, to explore use cases focused on UN-SDGs. The following case studies will demonstrate how satellite ARD was used with the ODC to address these use cases and provide significant impact for local Africa users. These ARD data sets were compiled in time series stacks to allow valuable assessments of changing and land and water resources, which would be nearly impossible, or quite difficult, using traditional scene-based analysis methods. In addition, the relevant SDGs are noted for each use case to highlight the potential impact of these data.

8.4.1. Flood Risk Map for Strategic Supply of Polio Vaccines Near Lake Chad

The following use case addresses two SDGs including SDG 3.B.1 (proportion of the population with access to affordable medicines and vaccines on a sustainable basis) and SDG 6.6.1 (change in the extent of water related ecosystems over time).

Lake Chad is an historically large, shallow (10 m maximum), lake in central Africa, which has varied in size over the centuries. According to the United Nations, it shrank by as much as 95% from 1963 to 1998 primarily due to increased demand from the local population. Lake Chad is economically important, providing water to more than 68 million people living in the four countries surrounding it (Chad, Cameroon, Niger, and Nigeria) on the edge of the Sahara Desert. The lake follows a complex annual water cycle. The beginning of the rainy season in the upper basin (May–June) gives rise to floods (August–September), which fill up Lake Chad (October–January) before evaporation associated with the end of the flow leads to a drop in the water level.

The Gates Foundation approached the ODC team in 2017 to obtain water extent time series data over Lake Chad to compare with their population density maps to strategically supply polio vaccines to small villages around the lake. Many of these villages become isolated during the rainy season and are accessible only for a small portion of the year. Knowledge of the flood risk for any given area and the times that water exists in each area will be extremely helpful for this project.

The CEOS Systems Engineering Office (SEO) used a Lake Chad Data Cube and the Australian Water Detection from Space (WOFS) algorithm (Mueller et al., 2016) to examine the water extent for this region. WOFS is a 23-step band-based decision tree that classifies pixels as water or nonwater with 97% overall accuracy. This water management product provides critical insight into the behavior of surface water over time and, in particular, the extent of flooding.

The pixel-level results (Fig. 8.1) show the percentage of observations detected as water over the 17 year time series (January 2000 through December 2016) in a southern portion of Lake Chad. The percentage calculation is based on the total number of water observations divided by the number of clear observations in the time series. These results demonstrate the variation in lake extent with significant change in the western region near the border of Nigeria. The Chari River can be seen originating from the middle-right of the image and then producing forks of entry into the larger lake. Though much of the lake is persistent water over the time period, there are low plain regions in the south where water was infrequently detected (2% to 20% of the time), likely due to annual rainy seasons.

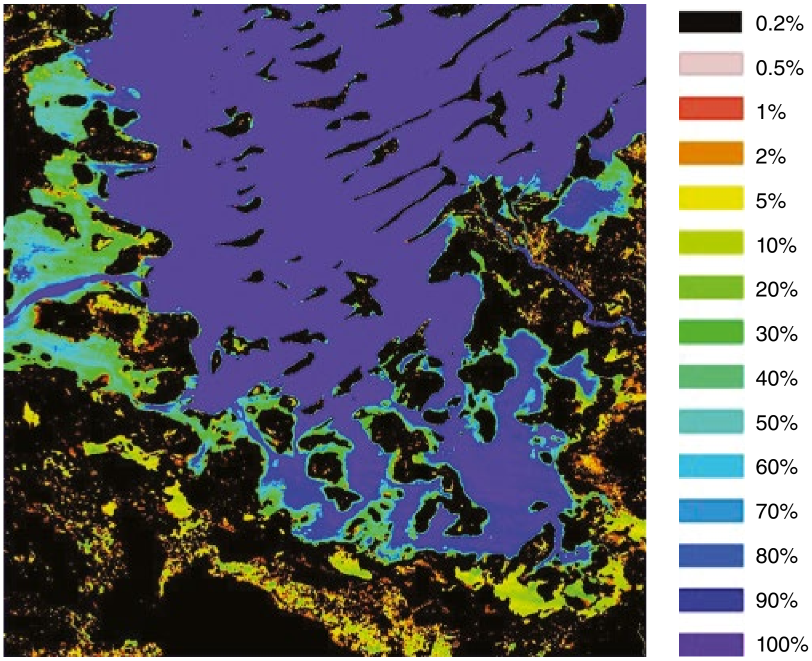

Figure 8.1 WOFS water detection results show the percentage of observations detected as water over the 17 year (2000 through 2016) time series. Frequent or permanent water is blue. Infrequent water or low flood risk is red/yellow. These maps can be used to determine flood risk and examine water extent history in specific locations for more accurate polio vaccine routing.

Knowledge of the water variability spatially and temporally is critical to the Gates Foundation as it plans its polio vaccine administration route. Figure 8.2 shows a view of the water variability for the entire Lake Chad region. The southern portion of the lake is the focus of the polio project where the WOFS data will be studied to understand spatial and temporal constraints for vaccine delivery.

8.4.2. Deforestation in the Tanzania East Chenene Forest Reserve

The following use case addresses two SDGs: SDG 15.1.1 (forest area as a proportion of total land area) and SDG 15.3.1 (proportion of land that is degraded over total land area).

The East Chenene Forest Reserve is in Dodoma, Tanzania. This protected region is of high interest to the Tanzanian government, as recent deforestation and new settlements along the perimeter of the forest reserve have impacted the region. An analysis was completed using the ARDC to detect land change and, specifically, deforestation. Decision makers from the government and national statistical agency desire to monitor the change in forest cover and assess land degradation. This information will allow improved land management practices and allow targeted investigations of recent change.

NASA's CEOS Systems Engineering Office (SEO) developed a simple Normalized Difference Vegetation Index (NDVI) Anomaly algorithm to detect deforestation. The notebook compares NDVI between two time periods to detect land change. In the case of deforestation, the NDVI values will reduce from stable high values (0.6 to 0.9, typical for forests) to lower values (< 0.6). This change can be detected and used to investigate deforestation or monitor the extent of the land change. The performance of this algorithm is extremely fast, as NDVI is merely a simple equation using two Landsat bands. The results of this analysis approach are shown in Figures 8.3 and 8.4. Due to extreme cloudiness in this region and issues with the Landsat Scan-Line Corrector (SLC) issue, a baseline NDVI image required 3 years of data (2004 to 2007). This baseline period was compared with a later period (2013) to identify land change areas, likely coincident with deforestation (Figs. 8.3 and 8.4).

In addition to the NDVI Anomaly algorithm, an additional analysis was conducted using the Python Continuous Change Detection (PyCCD) algorithm (Zhu & Woodcock, 2014) to detect land change over the time series. This output was compared with the well-known Global Forest Watch product (Hansen et al., 2013). The results are quite consistent among the two products (Fig. 8.5). Running this more complex algorithm, in addition to the NDVI Anomaly algorithm, can provide more information and confirmation of deforestation locations.

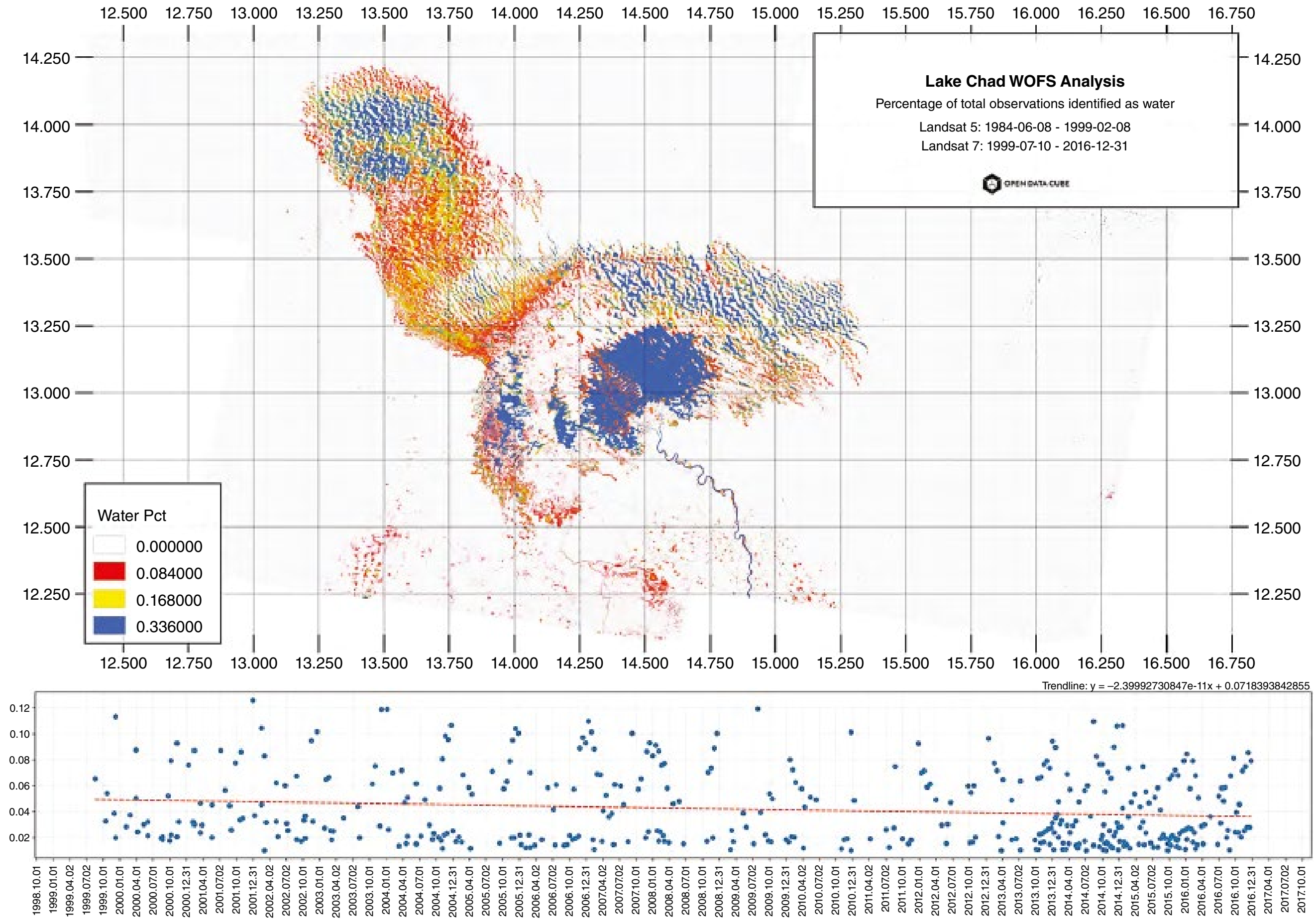

Figure 8.2 WOFS analysis results (a) from 1984 through 2016 (32 years) show wide variability in water extent. Regions to the north (yellow/red) only experienced water before year 2000. The southern portion of the lake (blue) is the area of current water and the focus of the polio project. Graph (b) shows the steady decline in lake levels from 1999 through 2016.

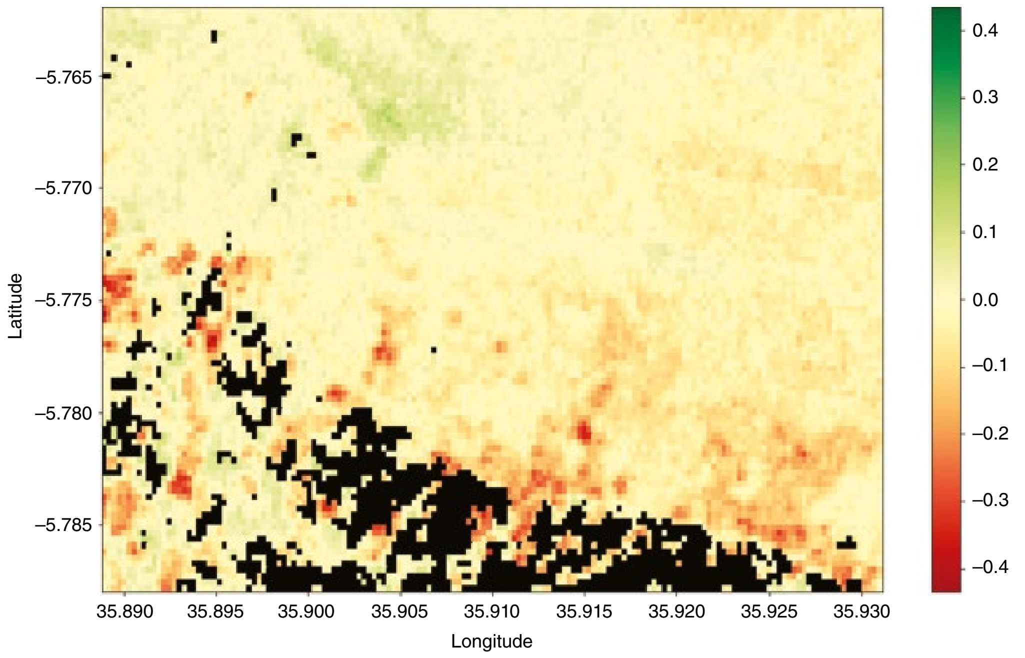

Figure 8.3 NDVI anomaly product for a portion of the East Chenene Forest Reserve. The product compares the NDVI from the baseline maximum NDVI mosaic (2004 to 2007) to a later analysis period (2013). Areas of deforestation appear in orange/red as the NDVI has been reduced along the perimeter of the forest reserve. The black pixes are nonforest areas and not relevant.

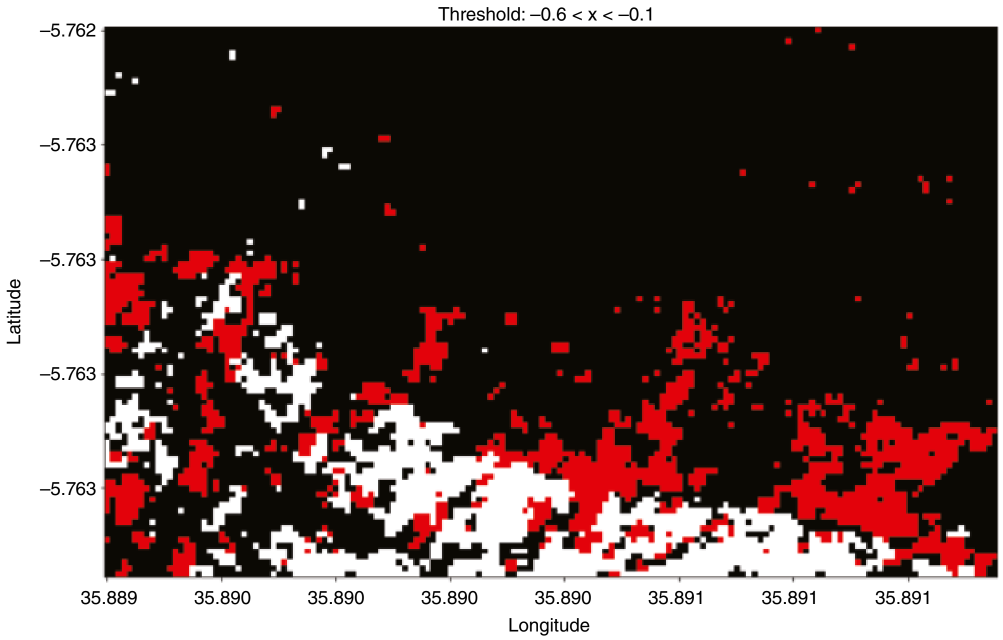

Figure 8.4 NDVI anomaly threshold product for a portion of the East Chenene Forest Reserve. The results show the pixels (red) that have experienced a loss of 0.1 to 0.6 in NDVI between the baseline and analysis time period. Since these pixels are likely forest in the baseline image (stable NDVI of 0.6 to 0.9), they are likely deforestation. The white pixels are nonforest areas. These results can be used to target and investigate large areas of deforestation.

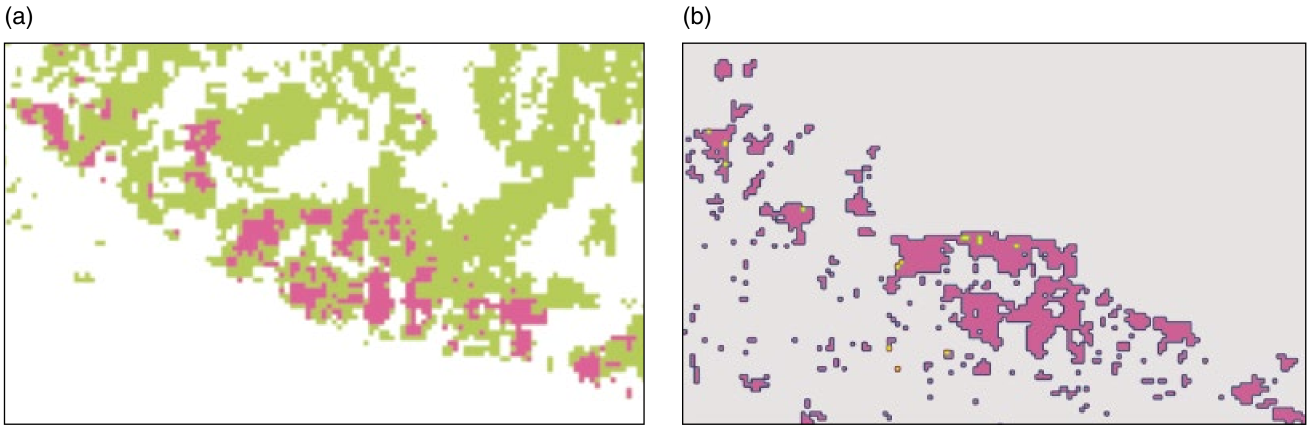

Figure 8.5 Deforestation analysis products for 2004 thru 2013 from two sources. (a) Global Forest Watch deforestation results are shown. The forest mask is green and areas of deforestation are pink. The details of this algorithm are not known, but the results are commonly used across the globe. (b) The data cube PyCCD algorithm deforestation results are shown with areas of land change (mostly deforestation) in pink with small areas of multiple land changes in yellow. The deforestation results are quite consistent between these two examples.

8.4.3. Detection of Illegal Mining in Ghana

The following use case addresses three SDGs: SDG 6.3.2 (proportion of bodies of water with good ambient water quality), SDG 6.6.1 (change in the extent of water related ecosystems over time), and SDG 15.3.1 (proportion of land that is degraded over total land area).

Illegal mining in Ghana has a significant social, economic, and environmental impact (Attuquayefio et al., 2017). The typical process begins with deforestation of the region followed by excavation and drilling into the surface. Following the extraction of the desired minerals, the site is left with small, contaminated water bodies with high sediment levels due to underground aquifer exploitation. Since these mining regions are often located near rivers and other natural water bodies, the impact of the contaminated water (from excavation) and soil runoff (from deforestation) is significant. In addition, it is known that temporary villages often exist near the mining sites and use this water for daily activities.

The Ghana government, namely the Ministry of the Environment and the National Statistical Office, is interested in detecting these illegal mining sites to target the investigation of the activity, measure the extent of the damage, and measure the impacts of management programs. The ARDC, with its time series of land imaging data, is ideal for this purpose. Though it is often difficult to detect these high sediment water bodies (as they appear to be bare soil from space), it is possible to see the impact of deforestation and the patterns of bare soil and water surrounding these regions.

The images below (Figs. 8.6 and 8.7) show fractional cover (FC) in (a) 2000 and (b) 2017. The third image (c) shows the loss of dense vegetation over the same time period (black pixels). FC is an iterative algorithm (Guerschman et al., 2015) that classifies every pixel as a fraction of bare soil (red), photosynthetic vegetation (green), and nonphotosynthetic vegetation (blue). FC can be used to identify areas that have moved from vegetated (green) to nonvegetated (blue or red). Mining areas are easy to locate as the dense forest (green) is replaced with bare soil or contaminated water (red) or nonphotosynthetic vegetation (blue). This change is evident in the two images from 2000 to 2017.

8.4.4. Urbanization Growth and Deforestation in Freetown, Sierra Leone

The following use case addresses two SDGs: SDG 11.3.1 (ratio of land consumption rate to population growth rate) and SDG 15.1.1 (forest area as a proportion of total land area).

Many governments, including Sierra Leone, are interested in tracking urbanization to understand the changes in land resources and corresponding population growth rates. These government agencies include national statistical offices, urban planning managers, and the ministries of agriculture and environment. It is known that increases in urbanization have an impact on the environment and the health of a population. With urbanization products from the ARDC, government decision makers can measure the extent and location of urban growth to help planning of water and land use.

A study was conducted over the city of Freetown (Fig. 8.8) using the ARDC. A common method to identify urban areas is to use the Normalized Difference Built Index (NDBI). The NDBI uses the Landsat shortwave infrared (SWIR-1) band and the near infrared (NIR) band to distinguish urban areas from surrounding vegetation or bare soil. This index is well documented and can be used to measure the amount of urban area by using simple thresholds. In addition to using NDBI, the NASA SEO team has investigated the use of Normalized Difference Vegetation Index (NDVI) and the bare soil component of a fractional cover product.

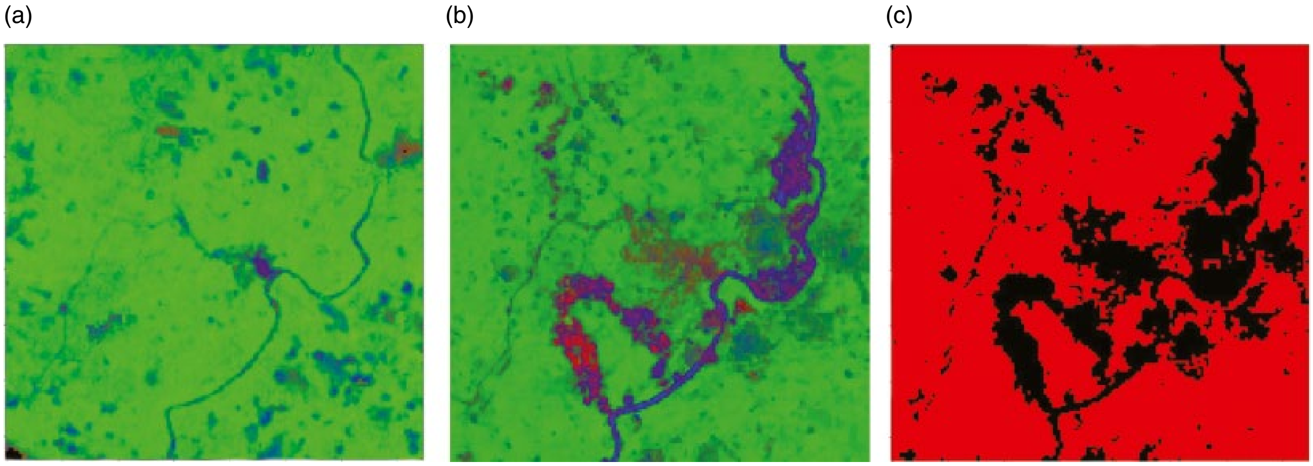

Figure 8.6 Fractional cover (FC) in (a) 2000 and (b) 2017 over the south Ankobra River in Ghana. FC is an iterative algorithm that classifies every pixel as a fraction of bare soil (red), photosynthetic vegetation (green), and nonphotosynthetic vegetation (blue). Mining areas are easy to locate, as the dense forest (green) is replaced with bare soil or contaminated water (red) or nonphotosynthetic vegetation (blue) between (a) 2000 and (b) 2017. The image (c) far right is a dense vegetation mask using NDVI (Normalized Difference Vegetation Index) where black pixels are likely deforestation. The region experienced a 13% decrease in dense vegetation (likely deforestation) over the time period.

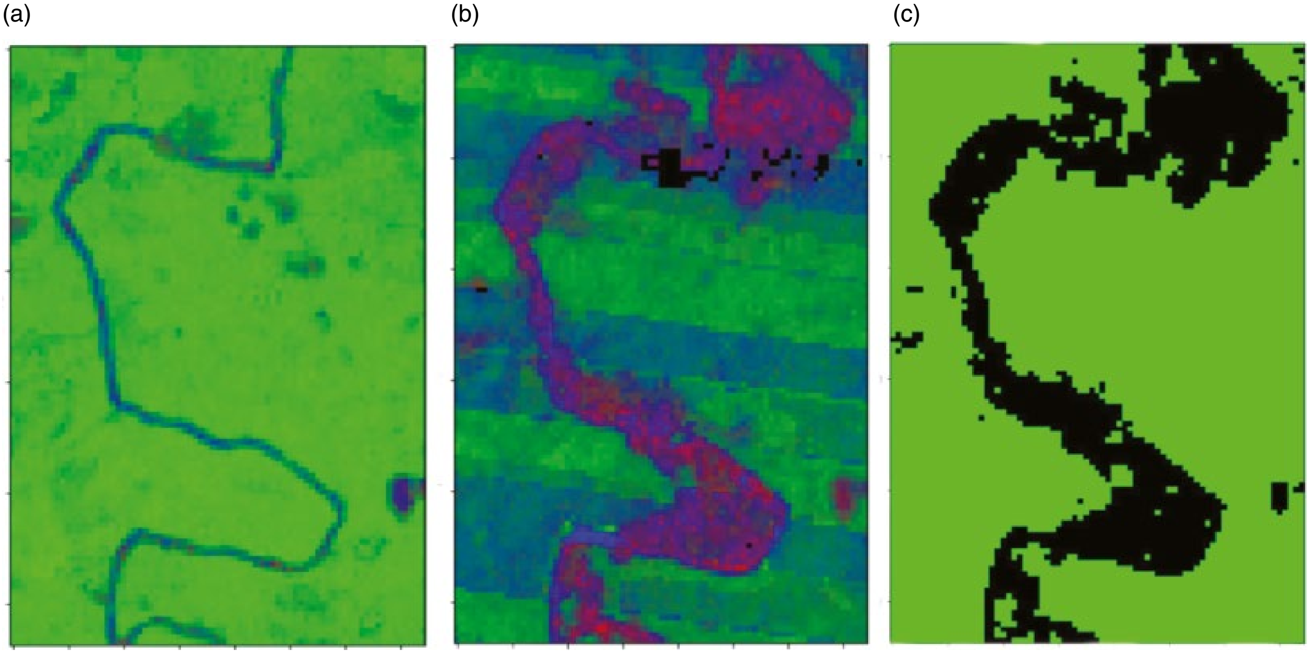

Figure 8.7 Fractional cover (FC) in (a) 2000 and (b) 2017 over the north Ankobra River in Ghana. FC is an iterative algorithm that classifies every pixel as a fraction of bare soil (red), photosynthetic vegetation (green), and nonphotosynthetic vegetation (blue). Mining areas are easy to locate, as the dense forest (green) is replaced with bare soil or contaminated water (red) or nonphotosynthetic vegetation (blue) between (a) 2000 and (b) 2017. The banding in (b) is due to the Landsat Scan Line Corrector (SLC) error and cloud variation among the scenes. The image in (c) is an FC threshold mask of the 2017 image (b) where dense vegetation is shown in green (PV of 0.3 to 1.0) and deforestation is shown as black pixels. The region experienced a 23% decrease in dense vegetation (likely deforestation) over the time period.

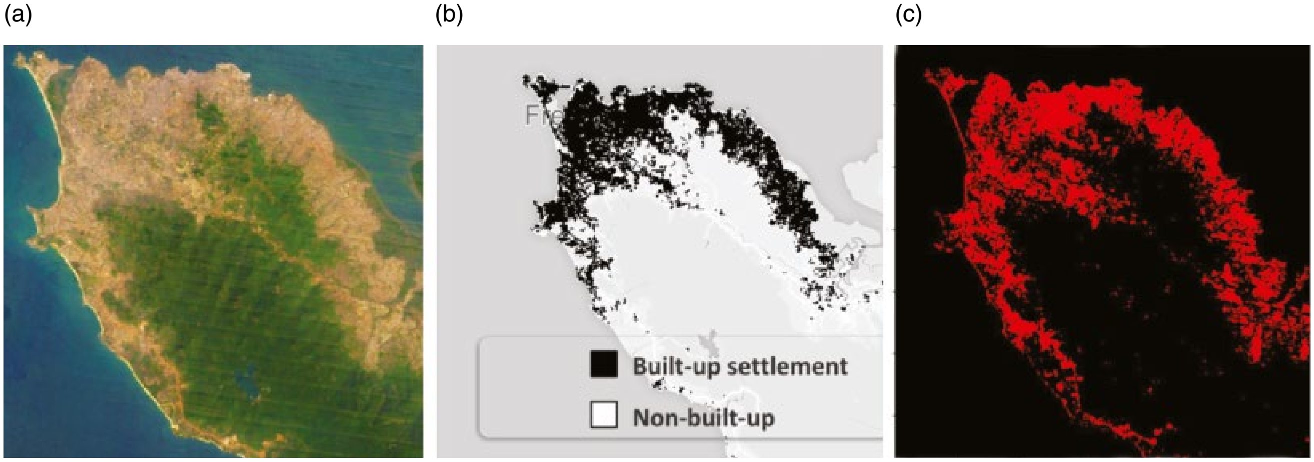

Figure 8.8 (a) A Landsat RGB image of Freetown in 2012 is shown. (b) An estimation of the built-up area from the European Space Agency Urban Thematic Exploitation Program (TEP) using radar data from the TerraSAR-X and TanDEM-X missions is shown. (c) An estimation of the urban area using the NDBI algorithm (0.0 to 0.3 threshold) and the ARDC is shown. The results of these two analyses are quite similar, though the urban areas from (c) the ARDC are more extensive.

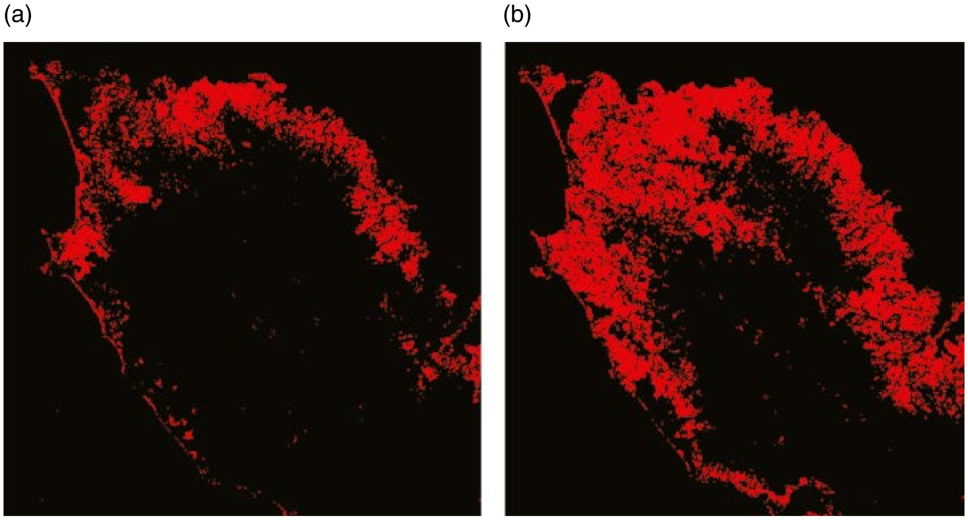

Figure 8.9 An estimation of the urban area using the ARDC and the NDBI algorithm (0.0 to 0.3 threshold) is shown for (a) 2000 and (b) 2017. The urban area increased 11.4% over this 17 year time period. With knowledge of the corresponding population increase for the same time period, it is possible to directly calculate the SDG 11.3.1 indicator.

To address concerns with population growth compared with land consumption rate (SDG 11.3.1), it is necessary to compare urbanization across a time series. Data cubes, such as the ARDC, are a unique platform to accomplish this analysis. The results shown in Figure 8.9 are a typical product that can be used to compare urban extent between two time periods.

In late 2020, the ARDC transitioned into the Digital Earth Africa program. This move was a planned evolution to increase scale, sustainability, and functionality through a continent wide operational service for Earth observation data. During the period of 2018 through 2020, the ARDC provided access to data for five countries and built invaluable insights into which satellite data and derived products are of most value, and lessons learned on the application of this technology. The success of this venture was a critical factor in gaining acceptance from African users and securing funding to initiate Digital Earth Africa.

8.5. CONCLUSIONS

Satellite missions will continue to provide increasingly larger volumes of free and open data for global users, so efficiently preparing this data into analysis ready format is critical to achieving progress on the UN-SDGs. Such data preparation will ensure consistency of time series and interoperability between different data sets to develop value-added products. With free and open-source innovative solutions such as the ODC, it is possible for all global users to exploit this information, address the SDGs, and optimize societal benefit and decision making. Though merely a few examples have been shown in this chapter, there are many more SDGs that can benefit from satellite data. Over time, it is expected that data access and data use will become easier and faster so that everyone across the globe can benefit.

ACKNOWLEDGMENTS

Funding for this work was provided by NASA's Earth Science Division. The work was coordinated through the Committee on Earth Observation Satellites (CEOS) Systems Engineering Office (SEO) with critical support from Analytical Mechanics and Associates, Inc. as the primary subcontractor.

REFERENCES

Attuquayefio, D., Owusu, E., & Ofori, B. (2017). Impact of mining and forest regeneration on small mammal biodiversity in the Western Region of Ghana. *Environmental Monitoring and Assessment*, *189*, 5.

Gavin, D., Dhu, T., Sagar, S., Mueller, N., Dunn, B., Lewis, A., at al. (2018). *Digital Earth Australia: From satellite data to better decisions*. Paper presented at the 2018 International Geoscience and Remote Sensing Symposium (IGARSS).

Giuliani, G., Camara, G., Killough, B., & Minchin, S. (2019). Earth observation open science: Enhancing reproducible science using data cubes. *Data* 2019, *4*, 147.

Giuliani, G., Chatenoux, B., De Bono, A., Rodila, D., Richard, J., Allenbach, R., et al. (2017). Building an Earth observations data cube: Lessons learned from the Swiss data cube (SDC) on generating analysis ready data (ARD). *Big Earth Data*, *1*, 1–18.

Guerschman, J., Scarth, P., McVicar, T., Renzullo, L., Malthus, T., Stewart, J., at al. (2015). Assessing the effects of site heterogeneity and soil properties when unmixing photosynthetic vegetation, non-photosynthetic vegetation and bare soil fractions from Landsat and MODIS data. *Remote Sensing of Environment*, *161*, 12–26.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*, 850–853.

Kavvada, A., & Held, A. (2018). *Analysis-ready Earth observation data and the United Nations sustainable development goals*. Paper presented at the 2018 International Geoscience and Remote Sensing Symposium (IGARSS).

Killough, B. (2018). *Overview of the Open Data Cube Initiative*. Paper presented at the 2018 International Geoscience and Remote Sensing Symposium (IGARSS).

Killough, B. (2019). *The impact of analysis ready data in the Africa Regional data cube*. Paper presented at the 2019 International Geoscience and Remote Sensing Symposium (IGARSS).

Killough, B., Siqueira, A., & Dyke, G. (2020). *Advancements in the open data cube and analysis ready data: Past, present, and future*. Paper presented at the 2020 International Geoscience and Remote Sensing Symposium (IGARSS).

Lewis, A., Oliver, S., Lymburner, L., Evans, B., Wyborn, L., Mueller, N., at al. (2017). The Australian geoscience data cube: Foundations and lessons learned. *Remote Sensing of the Environment*, *202*, 276–292.

Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., et al. (2016). Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote Sensing of the Environment*, *174*, 341–352.

Ross, J., Killough, B., Dhu, T., & Paget, M. (2017). *Open data cube and the Committee on Earth Observation Satellites Data Cube Initiative*. Paper presented at the 2017 International Astronautical Congress (IAC).

Zhu, Z., & Woodcock, C. (2014). Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of the Environment*, *144*, 152–171.