data

MDPI

Data Descriptor

A Sentinel-2 Dataset for Uganda

Jonas Ardö

Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, 22362 Lund, Sweden; jonas.ardo@nateko.lu.se

Abstract: Earth observation data provide useful information for the monitoring and management of vegetation- and land-related resources. The Framework for Operational Radiometric Correction for Environmental monitoring (FORCE) was used to download, process and composite Sentinel-2 data from 2018–2020 for Uganda. Over 16,500 Sentinel-2 data granules were downloaded and processed from top of the atmosphere reflectance to bottom of the atmosphere reflectance and higher-level products, totalling > 9 TB of input data. The output data include the number of clear sky observations per year, the best available pixel composite per year and vegetation indices (mean of EVI and NDVI) per quarter. The study intention was to provide analysis-ready data for all of Uganda from Sentinel-2 at 10 m spatial resolution, allowing users to bypass some basic processing and, hence, facilitate environmental monitoring.

Dataset: DOI: 10.5878/bc12-w579.

Dataset License: CC-BY-NC

Keywords: Sentinel-2; Uganda; FORCE; EVI; NDVI; earth observation; environmental monitoring

Citation: Ardö, J. A Sentinel-2 Data Set for Uganda. *Data* **2021**, *6*, 35. https://doi.org/10.3390/data6040035

Academic Editor: Jamal Jokar Arsanjani

Received: 15 February 2021
Accepted: 18 March 2021
Published: 30 March 2021

1. Introduction

1.1. Uganda

Uganda covers about 241,500 km^2 whereof ~197,000 km^2 is land and the rest is water and permanent wetlands (Figures 1 and 2). The current (2021) population of about 44 million is expected to reach about 90 million in the year 2050 and about 137 million in the year 2100 [1], indicating an increasing pressure on resources based on ecosystem services such as food, feed, fuel, and fibre [2]. This stresses the importance of land resource monitoring, supporting efficient and sustainable land use related to forestry and agriculture [3,4], natural conservation, and the preservation of biodiversity and ecosystem services.

A range of environmental changes in Uganda have recently been reported, including land use/land cover changes [5,6], vegetation changes [7,8], drought [9], and negative soil changes [10]. Flooding, land slides, mud flows, flash floods [11–13] and related disasters are also significant risks in Uganda's mountainous regions. These phenomena can partly be effected by vegetation changes due to forest fires [12], land cover changes and deforestation [14].

1.2. Earth Observation

Earth observation (EO) is an efficient tool for the monitoring of land use, land cover and vegetation, as well as related changes (e.g., drought, flooding, deforestation), as it is the only tool with full global cover at high temporal (daily-weekly) and spatial (10–30 m) resolution. The Sentinel-2 mission comprises two polar-orbiting satellites, Sentinel-2A and Sentinel-2B, launched in 2015 and 2017, respectively. Each satellite has a return time of about 10 days at the equator and 5 days for the combination of both missions. It is equipped with a multispectral instrument (MSI) configured as a push-broom scanner. It includes 13 spectral bands in three spatial resolutions (four bands at 10 m spatial resolution,

six bands at 20 m, and three bands at 60 m, Table 1, [15]). Sentinel-2 is suitable for mapping vegetation, land cover, and phenology, as well as associated changes over time [16].

All Sentinel-2 data are freely available from ESA [17] or from other data repositories [18,19] and about 4 TB of Sentinel-2 data are published daily [20]. The handling and processing of such large datasets can be facilitated by access to high-performance computing resources or cloud-based computing resources. Such handling and processing may be slow and inconvenient using standard desktop PCs. An intention here is to provide analysis-ready data (ARD) that can help to avoid some of this basic processing.

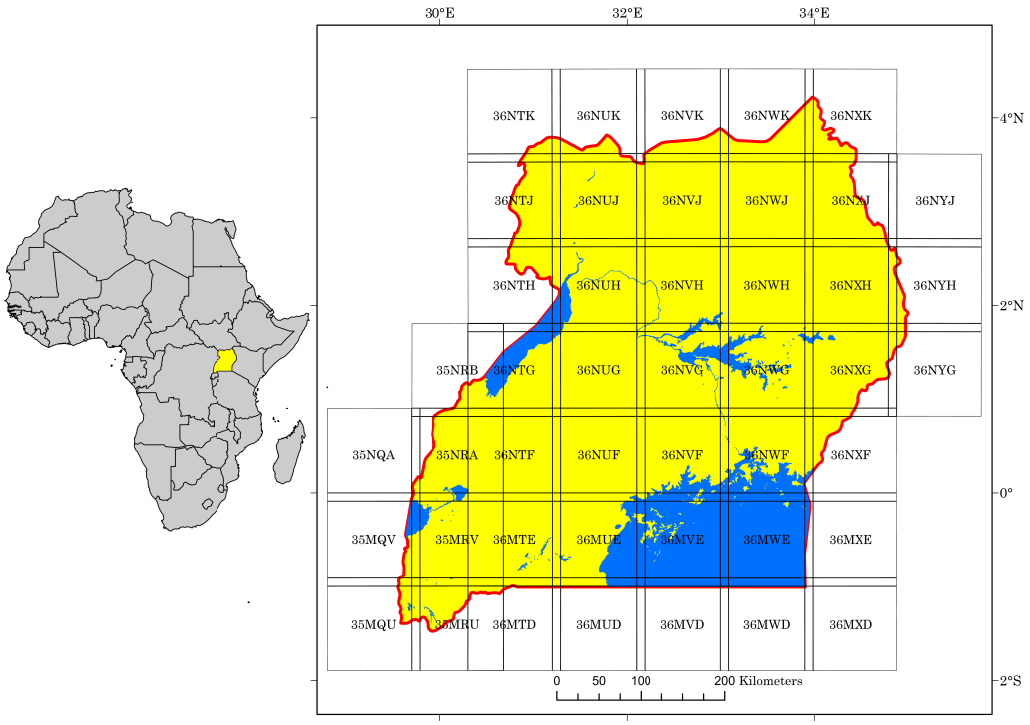

Figure 1. Africa, Uganda and the Sentinel-2 tiles used, water in blue.

Table 1. Spectral region, ESA Sentinel-2 band from MSI (Sentinel-2 A/B MultiSpectral Instrument (MSI)), and the corresponding FORCE level-2 (L2) output band, wavelengths, and original spatial resolution (all FORCE output is resampled to 10 m spatial resolution).

Region	ESA Band	FORCE L2 Band	Wavelength [µm]	Original Resolution [m]
Blue	1	-	0.433–0.453	60
Blue	2	1	0.440–0.538	10
Green	3	2	0.537–0.582	10
Red	4	3	0.646–0.684	10
Rededge1	5	4	0.694–0.713	20
Rededge2	6	5	0.731–0.749	20
Rededge3	7	6	0.769–0.797	20
BroadNIR	8	7	0.760–0.908	10
NIR	8B	8	0.848–0.881	20
NIR	9	-	0.935–0.955	60
NIR	10	-	1.360–1.390	60
SWIR1	11	9	1.539–1.682	20
SWIR2	12	10	2.078–2.320	20

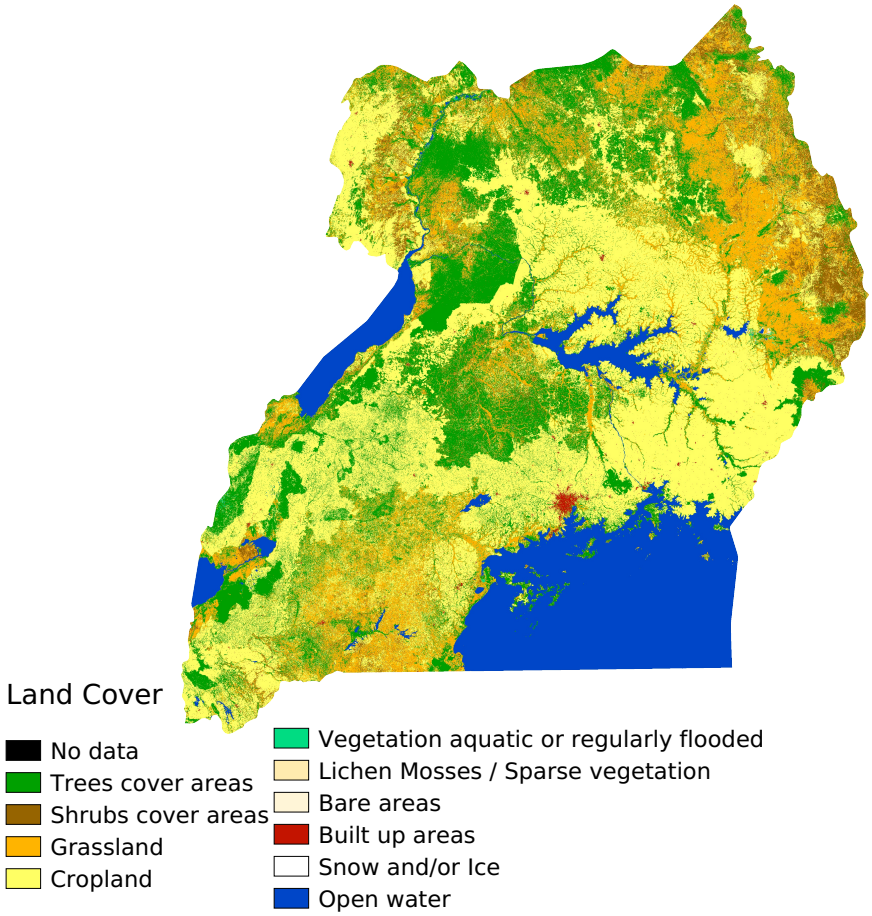

Figure 2. Land cover from ESA climate change initiative derived from Sentinel-2 for 2016 [21].

1.3. Aim

This study aims to present, describe and provide access to analysis-ready Sentinel-2 data covering Uganda for the period from 2018 to 2020.

2. Materials and Methods

2.1. Data

2.1.1. Sentinel-2

Sentinel-2 data (Level 1C, top of the atmosphere reflectance) for 2018 (4900 granules), 2019 (4944 granules) and 2020 (4929 granules) for 45 tiles over Uganda (Figure 1) were downloaded from Google Cloud Platform (https://console.cloud.google.com/storage/browser/gcp-public-data-sentinel-2 (accessed on 1 November 2020)). Each tile covers approximately 100 × 100 km and about 600 MB of data in the original UTM/WGS84 projection. The tiling definition follows the US-MGRS (Military Grid Reference System) approach [22]. It is based on the standard 6° longitude by 8° latitude UTM zones, divided into 100 × 100 km tiles. The two first characters of the tile name (e.g., 36NTK) indicate the UTM zone (35 and 36 for Uganda); the third character indicates the 8° latitude band (M and N for Uganda). Characters 4 and 5 denote the 100 × 100 km tile within the grid zone [23].

2.1.2. Land Cover

As an additional dataset (not used during the Sentinel-2 processing presented here) the 20 m spatial resolution S2 prototype land cover map for Africa was used, with reference year 2016 [21] (Figure 2). The land cover map was resampled to 10 m spatial resolution using nearest neighbour, and with identical tiling and map projection parameters to the Sentinel-2 data (Figure 3). This additional dataset may be useful in further analysis of the Sentinel-2 data for Uganda presented here.

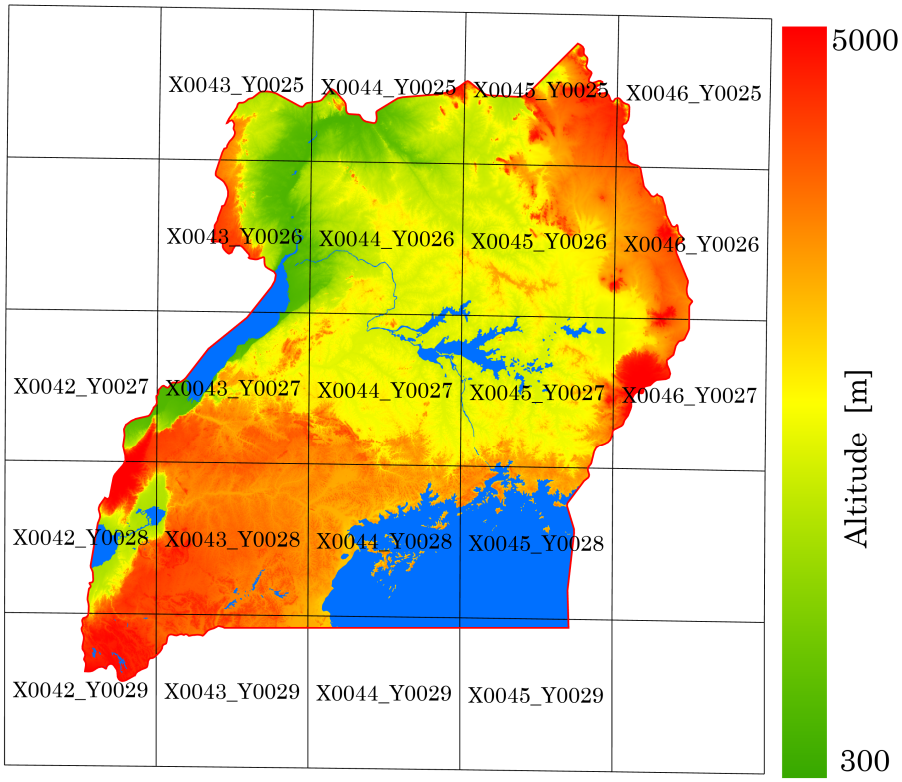

Figure 3. The SRTM 30 m spatial resolution elevation data used in the topographic calibration of the Sentinel-2 data. Overlaid is the non-overlapping output grid, which is part of the GLANCE grid for Africa [24]. Water areas in blue.

2.1.3. Digital Elevation Data

Digital elevation data are required for the topographic correction of EO data in Framework for Operational Radiometric Correction for Environmental monitoring (FORCE). The 1-arc-second (about 30 m spatial resolution) digital elevation model (DEM) derived from the shuttle radar topography mission (SRTM) was downloaded from http://e4ftl01.cr.usgs.gov/ (accessed on 3 November 2020) using the *30-Meter SRTM Tile Downloader* [25] and mosaiced into a virtual mosaic using the Geospatial Data Abstraction Library (GDAL) [26] (Figure 3).

2.2. Processing

All processing of Sentinel-2 data was performed using FORCE (v 3.4-3.6) [27,28]. FORCE "*is an all-in-one solution for the mass-processing and analysis of medium-resolution satellite image archives for large area + time series applications*" and an open source software under the terms of the GNU General Public License [29]. FORCE can be downloaded from GitHub [30]. Figure 4 gives an overview of the processing flow.

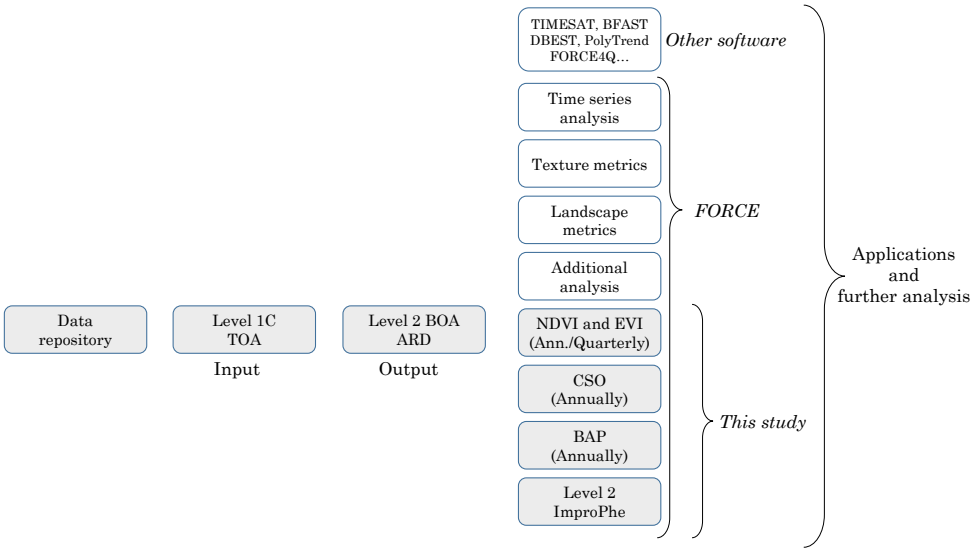

Figure 4. Process flow for Sentinel-2 data covering Uganda. Data and information flow from the left, where level 1C input data (top of the atmosphere reflectance) are retrieved from the repository, to the right, where analysis-ready (ARD) data in the form of level 2 output data (bottom of the atmosphere reflectance) are calculated. A range of options for further analysis exists, and these are labelled either as *This study*, *FORCE* or *Other software*. *This study* (grey shaded boxes) include a set of measures of NDVI and EVI (min., max., average, standard deviation, 5%, 50% and 95% quantiles), calculated annually, and average NDVI and EVI, calculated quarterly. Clear sky observations (CSO) and best available pixel (BAP) are produced annually. The 20 m resolutions bands are resampled to 10 m spatial resolution with the ImproPhe method. Additionally available (but unused in the this study) FORCE process sub-models utilizing the FORCE data cube are labelled *FORCE*. *Other software* tools suitable for time series analysis (e.g., TIMESAT [31]), trend analysis (e.g., Poly_Trend [32]) and change detection in time series (e.g., DBEST [33] and BFAST [34]) can be utilized, but these normally require a time series of level 2 data.

The hardware used was a Dell PowerEdge R730 with two sockets carrying Intel Xeon (2.6 GHz) CPUs, allowing up to 72 processes to be executed in parallel. Two Redundant Array of Independent Disks (RAID-6) systems, of 73 TB each, were available for the storing of data. In the CENTOS-7-based multiuser system, resources were shared with other users, and the number of parallel processes was limited to about 50–70% of the available capacity.

2.2.1. Radiometric, Atmospheric and Topographic Corrections

All downloaded Sentinel-2 data were level 1C, i.e., top of the atmosphere (TOA) reflectance [35] and processed to level 2, bottom of atmosphere reflectance (BOA) using FORCE. The basic processing scheme is described by Frantz et al [36]. Cloud masking is based on a modified version of Fmask, as detailed in [36]. The radiometric correction includes a radiative transfer atmospheric correction and the aerosol optical depth is estimated over dark objects. Topography normalization in FORCE is performed by a modified C-correction [36,37]. The spatial resolution of the Sentinel-2 20 m bands was improved using the original 10 m bands as targets.

The atmospheric correction performance of FORCE was evaluated and compared to other processors in a recent inter-comparison exercise [38], which stated, *'The results of the APU (accuracy, precision, and uncertainty) analysis . . . indicating that LaSRC, FORCE, and MACCS provided accurate and robust surface reflectance estimates for all the cases'*, (LaSRC = Landsat 8 Collection 1 Land Surface Reflectance Code, MACCS = multi-sensor atmospheric correction and cloud screening spectro-temporal processor). Further details are available in [28] and references therein. In total, 16,773 level 1C (TOA) Sentinel-2 granules were downloaded and processed to level 2 (BOA).

2.2.2. Geometric Correction

FORCE uses a data cube concept, where the level 2 and higher levels of output are reprojected into a common coordinate system and organized in non-overlapping tiles (Figure 3). Each output tile is defined as a square of 15,000 rows (lines) by 15,000 columns (pixels), corresponding to 150,000 by 150,000 m. Tiles (X0042_Y0025, X0042_Y0025, X0046_Y0028 and X0046_Y0029) outside the Uganda national border were not processed. All output data are created in a Lambert Azimuthal Equal Area projection corresponding to the *Global LANd Cover mapping and Estimation* (GLANCE) grid for Africa [24]). The spatial resolution of the 20 m Sentinel-2 bands (Table 2) was improved to 10 m using the native 10 m bands and the ImproPhe method [39]. Hence, all output bands were produced with a 10 × 10 m spatial resolution.

2.2.3. Clear Sky Observations

For each year, the number of clear sky observations (CSO) were calculated [40]. The number of CSOs may, for example, influence the possibility of detecting the time of harvest and similar events, and can hence be used when planning EO studies. CSO may also be useful in planning time series analysis for the assessment of phenology or other phenomena, as well as to decide if a multi-sensor dataset is needed, for example, Sentinel-2 in combination with Landsat [41].

2.2.4. Best Available Pixel Composite

FORCE Level 3 compositing was used to generate seamless and gap-free composites of reflectance from temporal aggregations of level 2 data using the best available pixel (BAP) compositing [42]. BAP composites were produced using static target dates [43] for each year.

For each composite, two meta data files were produced, in addition to the reflectance data in the BAP composites: one file with compositing information (INF) and one file with a compositing score (SRC). The INF product contains information about the selected observation in the BAP product, as specified in Table 2. It is a multi-band image including pixel-wise information on data quality, and the day and year of included observations and sensor (Sentinel-2 A/B) of the best observation.

The Quality Assurance Information (QAI) is a score of the quality indicators used to create the BAP by selecting the pixel with the highest score for each location. All available observations for a specific pixel are assessed in terms of their suitability for the composite [42,44] using seven quality indicators. These indicators concern missing data, clouds (several indicators), cloud shadows, snow, and where the spectral reflectance is estimated to be <0 or >1.

Table 2. Compositing information related to the BAP composites. The information is stored in a GeoTIFF file (`*_INF.tif`) with 6 bands.

Band	Description
1	Quality Assurance Information of best observation
2	Number of cloud-free observations within compositing period
3	Acquisition Day Of Year (DOY) of best observation
4	Acquisition Year of best observation
5	Difference between band 3 and Target DOY
6	Sensor ID of best observation (Sentinel-2A or 2B)

2.2.5. Vegetation Indices

The enhanced vegetation index (EVI) [45]

$$2.5 \times \frac{(NIR - RED)}{NIR + 6 \times RED - 7.5 \times BLUE + 1} \quad (1)$$

and the normalized difference vegetation index (NDVI) [46]

$$\frac{(NIR - RED)}{(NIR + RED)} \quad (2)$$

were calculated from level-2 BOA data. Near infrared (NIR), RED and BLUE in Equations (1) and (2) denote spectral reflectance in those spectral regions. For each year (2018, 2019, and 2020), the minimum (MIN), the average (AVG), the 5% quantile (Q05), the 50% quantile (Q50, i.e., the median) the 95% quantile (Q95), the maximum (MAX) and the standard deviation (STD) were calculated. For each quarter (Jan-Mar, Apr-Jun, Jul-Sep, and Oct-Dec) of the year, average EVI and NDVI were calculated.

2.2.6. Output Format and Naming Convention

All level 2 data were produced in the ENVI format, including a binary data file (*.dat) and a separate file with header information/metadata (*.hdr) [47]. Due to the size of the level-2 data (>50 TB), these were not made available in any repository.

All level-3 or higher output data are stored as 16-bit GeoTIFF files, each covering one tile of 15,000 rows ×, 15,000 columns, with 10 m spatial resolution and 10 bands (Figure 3, Table 1). All Sentinel-2 output data are scaled with a factor 10,000 (a reflectance of 0.5 is represented by 5000). Missing values are represented by −9999. Output files were organized as described in Appendix A and named according to the FORCE naming conventions described in Appendix B.

2.2.7. Dataset Structure and Reproduction

The dataset is spatially structured as a data cube with non-overlapping tiles, allowing one or a set of tiles to be processed when a sub-national study area is desired. Temporally the dataset is structured into quarterly (average NDVI and EVI) and annually integrated variables (CSO, BAP and descriptive measures of NDVI and EVI; see Section 2.2.5), allowing flexible temporal sub-setting independently from, or integrated with, the spatial sub-setting.

The completed level 2 dataset can be reproduced from the original level 1C data using the FORCE parameter files that determine processing details (available on request from the author).

3. Results

The output of the level-2 and level-3 processing is tiled (Figure 3), where each tile is 150 × 150 km, corresponding to 15,000 rows by 15,000 columns for the 10 × 10 m Sentinel-2 pixels.

3.1. Level-2 Data

Level-2 data contain BOA reflectance for all ingested level 1C data. Each level-2 output file includes 10 wavelength bands (Table 1). Further specifications for the Sentinel-2 bands and radiometric properties can be found in [15] and in Table 1.

3.2. Level 3 and Higher Level Data

3.2.1. Vegetation Indices

Average EVI and NDVI were created for each quarter of each year, as well as average EVI and NVDI for each year, as exemplified in Figure 5.

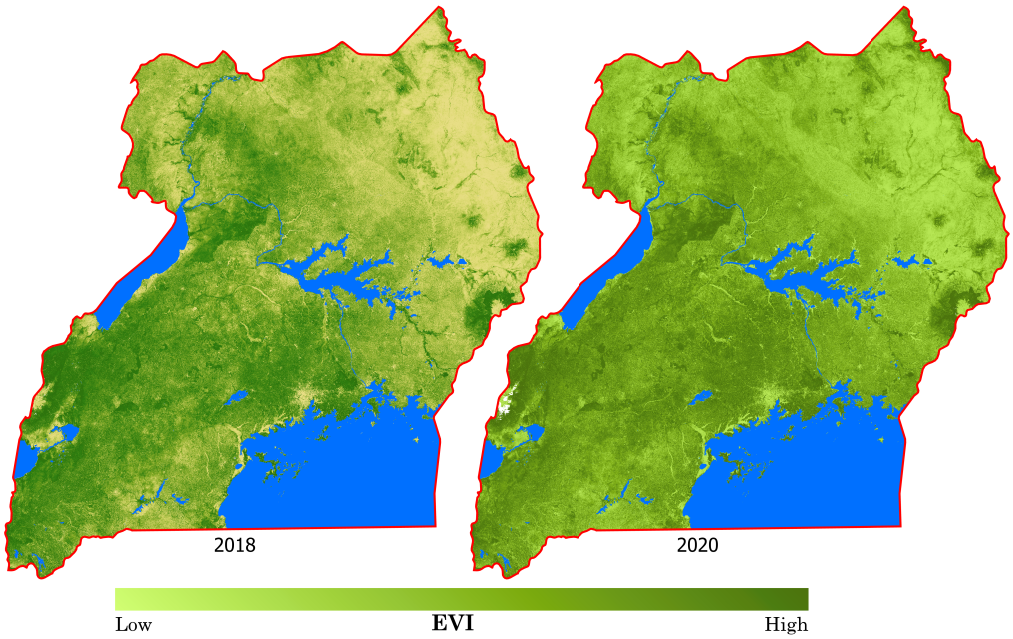

Figure 5. Average EVI from Sentinel-2 for 2018 and 2020 at 10 m spatial resolution for Uganda. See Figure 8 for illustrations of the differences in EVI (2018 vs. 2020) for a spatial subset.

3.2.2. Best Available Pixel Composites (BAP)

A BAP composite, including 10 bands (Table 1), was created for each year 2018–2020 (e.g., Figure 6).

Figure 6. Example of a cloud-free mosaic of Uganda for 2019 based on best available pixels (BAP). False color composite with Sentinel band 2,3,4 (RGB) .

3.2.3. Clear Sky Observations (CSO)

The number of available clear-sky observations (CSO) vary strongly spatially but show a similar pattern from year to year (Figure 7).

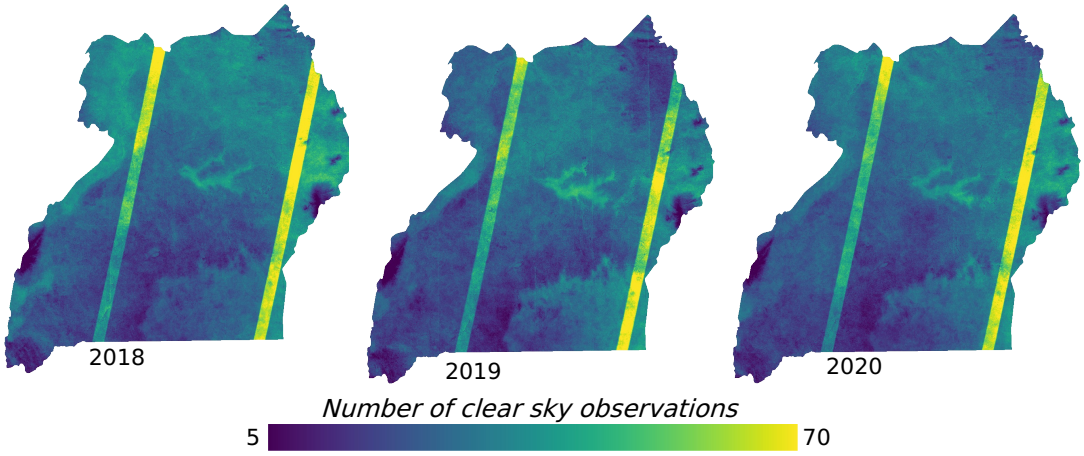

Figure 7. Number of clear-sky observations for 2018–2020. The stripes are caused by higher data availability due to Sentinel-2A/B orbit overlaps.

3.3. Data Availability and Download

Data can be downloaded from the URLs in Table 3.

Table 3. Overview of available data. Data level, time interval of acquisition or product, and URL where data can be downloaded from. For each vegetation index (EVI, NDVI) is the minimum, average, maximum, standard deviation, the 5% quantile, the median (50% quantile), 95% quantile calculated. TOA = Top Of the Atmosphere reflectance, BOA = Bottom Of the Atmosphere reflectance, BAP = Best Available Pixel, EVI = Enhanced Vegetation Index, NDVI = Normalized Difference Vegetation Index, CSO = Clear Sky Observations.

Level	Interval	Description	URL
1C	5 day	TOA-reflectance	https://console.cloud.google.com/storage/browser/gcp-public-data-sentinel-2 (accessed on 22 March 2021)
2	5 day	BOA-reflectance	Not shared in repository (too large)
3	Annual	BAP	*
Higher	Quaterly	EVI	*
Higher	Quaterly	NDVI	*
Higher	Annual	CSO	*
Extra	-	Elevation	*
Extra	2016	Land Cover	*

* URL for download of data = https://snd.gu.se/en/catalogue/study/2021-50 (accessed on 29 March 2021).

4. Discussion

Information derived from EO data can support scientific and applied investigations, as well as the monitoring and planning of land-use/land-cover-related activities, and hence contribute to the sound management of ecosystem services [2,48]. The provision of analysis-ready data can be a suitable starting point for deriving such information.

FORCE, the free software used here, is highly suitable for the mass processing of EO data such as Sentinel-2 or Landsat. As the Sentinel-2 data and the other data used are also free, basic processing, in principle, can be carried out by users. However, the Sentinel-2 data presented here comprise several TBs, and the processing can be demanding and time-consuming when using standard computing. FORCE also includes a range of tools for higher-level processing [49], including time series analysis, machine learning, and landscape metrics, allowing for further analysis of the analysis-ready data.

The idea behind and aim of this study was to provide analysis-ready data from Sentinel-2 for Uganda, allowing users to avoid some basic data handling and processing. This aim has been fulfilled. These data provide a potential basis for further national and regional studies based on EO in Uganda, and can potentially increase the use and applications of Sentinel-2 data. The current data, provided here, can be expanded for use of, for example, older Landsat data, allowing retrospective studies of land cover and land cover changes. They can also serve as a basis for future assessment efforts or near real-time monitoring [50].

Figure 8 shows the difference (increase or decrease >0.1) in mean EVI from 2018 to 2020 for a part of Uganda (Mount Elgon) as an example of how the output data can be used to identify environmental changes. EVI are strongly related to carbon assimilation [51,52] and provide an estimate of the productivity of the target vegetation studied. Therefore, regions of increasing and decreasing vegetation productivity can be detected either using time integrated measures, as in Figure 8, or using trend analysis [32,53] and detection of changes [33]. A recent review of land-use change and land-cover change in Uganda provides an overview of the leading driver of such changes [54]. EO based monitoring of biomass [55] and terrestrial carbon stocks [5] further show potential applications of the dataset for the quantification of resources related to crucial ecosystem services [2,56].

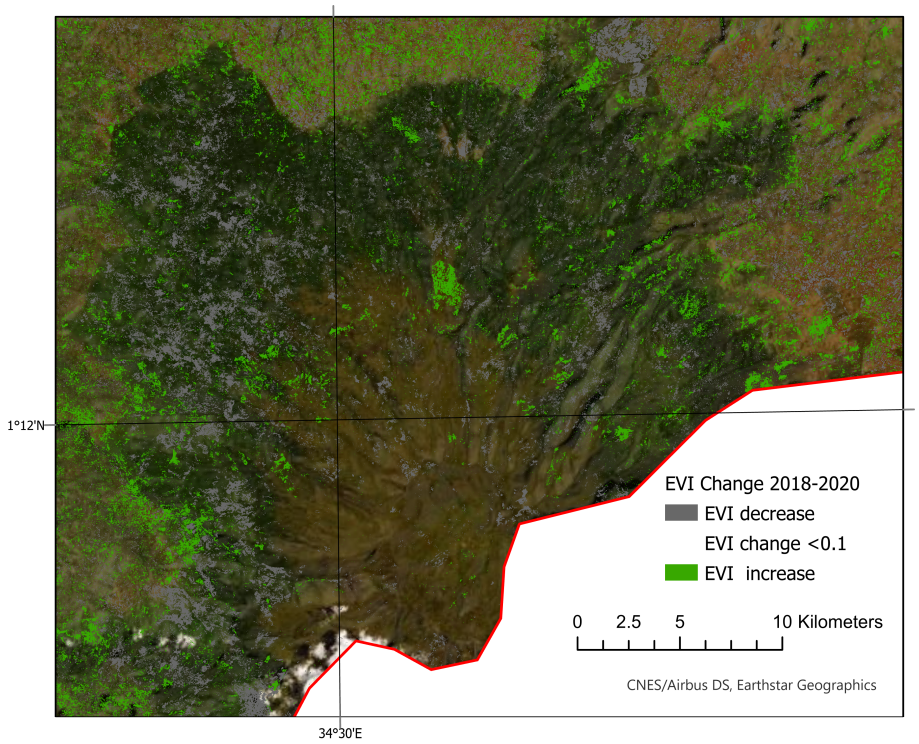

Figure 8. EVI annual average difference for the Mount Elgon region, 2018–2020. Grey areas indicate a decrease, green indicate an increase, of at least 0.1 from 2018 to 2020. Changes <0.1 are not shown.

The data sets presented are suitable for the classification of land use, land cover and related vegetation properties, especially properties associated with NDVI and EVI (as quarterly data are provided), but any vegetation index or band combination can be computed from the annual BAP composites, as they include all ten wavelength bands (Table 1). Change detection, on as annual basis, can be performed in any spectral band combination covered by the data and on quarterly basis using NDVI and EVI. A further outline of applications of this dataset is outside the scope of the paper, but the author is open for collaboration and suggestions of case studies.

The uncertainties of this dataset are related to the ESA's processing up to level 1C and to processing within FORCE. ESA reports monthly on geometric and radiometric performance [57], stating absolute geolocation without ground control points to be <11 m (95%

confidence) and absolute radiometric uncertainty for from band 1 to band 12 (excluding band 10) $< 3\% \pm 2\%$ [58]. Uncertainties originating from FORCE related to radiometry are comparable to other similar methods [38]. Hence, the quality of the Sentinel-2 data set is good and ESA will further increase the geometric quality when the planned processing baseline 03.00 is introduced [58], covering both Europe and Africa.

The CSO (Figure 7) indicates the number of clear-sky observations for various loci, hence supporting the planning of what type of analysis could be suitable in various areas. In mountain regions of Uganda, it can be difficult to obtain cloud-free observations annually, even with satellites passing as often as every fifth day. Hence, there can still be missing data due to cloud and cloud shadows in the BAP composites (Figure 6). The vegetation indices are functions (average, quantiles, etc.) of several observations and, hence, less directly influenced by clouds and missing data.

There are currently no direct plans to continue to maintain the level 2 and level 3 analysis-ready data presented beyond 2020.

Whether the future handling and analysis of EO data belong to cloud-based computing and storage services or local services is still uncertain [59,60]. The answer is probably that both strategies will be valid, for the near future at least, even if the produced datasets grow faster than the corresponding ability to analyse the data [61], especially on local platforms. Platforms like Digital Earth Africa https://www.digitalearthafrica.org/ (accessed on 1 March 2021) may provide support for the transition from local-based to cloud-based, assisting in handling the ever-growing data that are produced. Using FORCE and similar tools in cloud-based computing services could further increase data-handling efficiency, avoiding the shuffling of TBs of data, and support environmental monitoring.

Funding: This research was funded by Swedish International Development Agency (SIDA) through the Building Resilient Ecosystems and Livelihoods to Climate Change Disasters (BREAD) project, grant no. 331. Ardö was supported by an infrastructure grant from the Faculty of Science, Lund University.

Data Availability Statement: All level-3 and higher-level data can be downloaded from https://snd.gu.se/en/catalogue/study/2021-50 (accessed on 29 March 2021). The level-2 data (>50 TB) and parameter files used with FORCE are available from the author upon request as long as it is stored on disk.

Acknowledgments: Thanks to David Frantz and co-workers, Humbolt University, Berlin of for making the FORCE software freely available. The SRTM elevation data provided by NASA and the ESA climate change initiative S2 prototype LC 20m map of Africa 2016 provided by ESA, are both acknowledge as useful contributions.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. File Organization

All level 3 or higher-level products (CSO, BAP, and vegetation indices) are organized in one directory per FORCE-tile (Figure 3) and one directory with virtual mosaics [26]. Hence, each directory include files with the same file names, separated by the directory they are stored in and corresponding geo-location. Each directory is compressed with `tar -zcvf archive-name.tar.gz directory-name` where the archive-name is the same as directory-name. This means that all files belonging to directory `./X0042_Y0027` are stored in the file `X0042_Y0027.tar.gz`. Extract the files with `tar -zxvf archive-name.tar.gz` or equivalent functionality. A list of directories (and compressed files) is given below. In order to access the virtual mosaics in the `./mosaic` folder, all directories must be downloaded and uncompressed, or a new virtual mosaic created [26]. A definition of the data cube projections and setting used is available in a plain-text file called `datacube-definition.prj` with the standard format used by FORCE.

```
./X0042_Y0027
./X0042_Y0028
./X0042_Y0029
```

```
./X0043_Y0025
./X0043_Y0026
./X0043_Y0027
./X0043_Y0028
./X0043_Y0029
./X0044_Y0025
./X0044_Y0026
./X0044_Y0027
./X0044_Y0028
./X0044_Y0029
./X0045_Y0025
./X0045_Y0026
./X0045_Y0027
./X0045_Y0028
./X0045_Y0029
./X0046_Y0025
./X0046_Y0026
./X0046_Y0027
./mosaic
```

The data are delivered in 22 compressed files with names and sizes as listed below. In total, 403 GB in compressed form. Below K = kilobyte and G = Gigabyte and denoting file size.

```
60K     mosaic.tar.gz
2.9G    X0042_Y0027.tar.gz
15G     X0042_Y0028.tar.gz
7.3G    X0042_Y0029.tar.gz
8.8G    X0043_Y0025.tar.gz
19G     X0043_Y0026.tar.gz
31G     X0043_Y0027.tar.gz
38G     X0043_Y0028.tar.gz
4.4G    X0043_Y0029.tar.gz
16G     X0044_Y0025.tar.gz
39G     X0044_Y0026.tar.gz
39G     X0044_Y0027.tar.gz
35G     X0044_Y0028.tar.gz
2.6G    X0044_Y0029.tar.gz
22G     X0045_Y0025.tar.gz
38G     X0045_Y0026.tar.gz
38G     X0045_Y0027.tar.gz
26G     X0045_Y0028.tar.gz
1.7G    X0045_Y0029.tar.gz
1.5G    X0046_Y0025.tar.gz
15G     X0046_Y0026.tar.gz
8.6G    X0046_Y0027.tar.gz
403G~total
```

Appendix B. File Naming Convention

Appendix B.1. BAP

The BAP files are named according to the FORCE 29-digit naming convention [62]. For example `20190629_LEVEL3_SEN2L_BAP.tif` denoting a level 3 BAP product from Sentinel-2 with 26 June 2019 as the target date (Table A1).

Table A1. Naming convention for best available pixel (BAP) composites.

Digits	Description
1–8	Target date as YYYYMMDD
10–15	Product Level
17–21	Sensor ID (SEN2L, Sentinel-2 land bands)
23–25	Product type (BAP, INF, SCR)
27–29	File type (GeoTiff)

Appendix B.2. Elevation and Land Cover Data

All digital elevation data files (one file for each FORCE tile) are named `Uganda_DEM.tif` and the virtual mosaic file is named `Uganda_DEM.vrt` All land cover files (one file for each FORCE tile) are named `ESACCI-LC_uganda.tif` virtual mosaic file is named `ESACCI-LC_uganda.vrt`.

Appendix B.3. CSO

The CSO files are named according to the FORCE 37-digit naming convention for CSO [63]. For example `2018-2018_001-365-12_HL_CSO_SEN2L_NUM.tif` is the number of CSO for the period DOY 1–DOY 365 in the year 2018 in GeoTiff format (Table A2).

Table A2. Naming convention for clear sky observations (CSO).

Digits	Description
1–9	Temporal range for the years as YYYY–YYYY
11–17	Temporal binning in DOY as DDD–DDD
19–20	Temporal binning in months
22–23	Product level
25–27	Product
29–33	Sensor ID (SEN2L, Sentinel-2 land bands)
35–37	Product type (NUM = number of observations)
39–41	File type (GeoTiff)

Appendix B.4. Vegetation indices

The vegetation indices files are named according to the FORCE 42 to 46 digit naming convention [64]. For example is `2020-2020_001-365_HL_TSA_SEN2L_EVI_FBQ_QUARTER-4.tif` average EVI for year 2020 quarter 4 (October-December) in GeoTiff format. HL stands for higher level and TSA for time series analysis, the FORCE sub-module used. (Table A3).

Table A3. Naming convention for vegetation indices.

Digits	Description
1–9	Temporal range for the years as YYYY–YYYY
11–17	Temporal range for the DOY as DDD–DDD
19–20	Product level (HL)
22–24	Submodel (TSA)
26–30	Sensor ID (SEN2L, Sentinel-2 land bands)
32–34	Index Short Name (EVI, NDV)
36–38	Product type (FBQ = folded by quarter)
40–42	File type (GeoTiff)

References

1. UN. *World Population Prospects 2019: Volume II: Demographic Profiles*; Technical Report; United Nations, Department of Economic and Social Affairs Population Division: New York, NY, USA, 2019.
2. Lunyolo, L.D.; Khalifa, M.; Ribbe, L. Assessing the interaction of land cover/land use dynamics, climate extremes and food systems in Uganda. *Sci. Total Environ.* **2021**, *753*, 142549. [CrossRef]
3. Lobell, D.B.; Azzari, G.; Burke, M.; Gourlay, S.; Jin, Z.; Kilic, T.; Murray, S. *Eyes in the Sky, Boots on the Ground: Assessing Satellite- and Ground-Based Approaches to Crop Yield Measurement and Analysis in Uganda*; The World Bank: Washington, DC, USA, 2018. Available online: https://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-8374 (accessed on 15 October 2020). [CrossRef]
4. Barbosa, H.; Tote, C.; Kumar, L.; Bamutaze, Y. Harnessing Earth Observation and Satellite Information for Monitoring Desertification, Drought and Agricultural Activities in Developing Countries. In *Environmental Change and Sustainability*; Silvern, S., Young, S., Eds.; IntechOpen: Rijeka, Croatia, 2013; Chapter 4. [CrossRef]
5. Zhang, F.; Zhan, J.; Zhang, Q.; Yao, L.; Liu, W. Impacts of land use/cover change on terrestrial carbon stocks in Uganda. *Phys. Chem. Earth Parts A/B/C* **2017**, *101*, 195–203.
6. Nakalembe, C.; Dempewolf, J.; Justice, C. Agricultural land use change in Karamoja Region, Uganda. *Land Use Policy* **2017**, *62*, 2–12. [CrossRef]
7. Morgan, B.; Awange, J.; Saleem, A.; Kexiang, H. Understanding vegetation variability and their "hotspots" within Lake Victoria Basin (LVB: 2003–2018). *Appl. Geogr.* **2020**, *122*, 102238. [CrossRef]
8. Bernard, B.; Aron, M.; Loy, T.; Muhamud, N.W.; Benard, S. The impact of refugee settlements on land use changes and vegetation degradation in West Nile Sub-region, Uganda. *Geocarto Int.* **2020**. [CrossRef]
9. Kyatengerwa, C.; Kim, D.; Choi, M. A national-scale drought assessment in Uganda based on evapotranspiration deficits from the Bouchet hypothesis. *J. Hydrol.* **2020**, *580*, 124348. [CrossRef]
10. Mugizi, F.M.; Matsumoto, T. A curse or a blessing? Population pressure and soil quality in Sub-Saharan Africa: Evidence from rural Uganda. *Ecol. Econ.* **2021**, *179*, 106851. [CrossRef]
11. Habonimana, H.V. Integrated Flood Modeling in Lubigi Catchment Kampala. Master's Thesis, Faculty of Geo-Information Science and Earth Observation of the University of Twente, Enschede, The Netherlands, 2014.
12. Jacobs, L.; Maes, J.; Mertens, K.; Sekajugo, J.; Thiery, W.; van Lipzig, N.; Poesen, J.; Kervyn, M.; Dewitte, O. Reconstruction of a flash flood event through a multi-hazard approach: Focus on the Rwenzori Mountains, Uganda. *Nat. Hazards* **2016**, *84*, 851–876. [CrossRef]
13. Nseka, D.; Bamutaze, Y.; Mugagga, F.; Nakileza, B. The Fragility of Agricultural Landscapes and Resilience of Communities to Landslide Occurrence in the Tropical Humid Environments of Kigezi Highlands in South Western Uganda. In *Agriculture and Ecosystem Resilience in Sub Saharan Africa: Livelihood Pathways Under Changing Climate*; Bamutaze, Y., Kyamanywa, S., Singh, B.R., Nabanoga, G., Lal, R., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 279–305. [CrossRef]
14. Nakileza, B.R.; Nedala, S. Topographic influence on landslides characteristics and implication for risk management in upper Manafwa catchment, Mt Elgon Uganda. *Geoenviron. Disasters* **2020**, *7*, 27. [CrossRef]
15. ESA. Sentinel-2 MSI Spatial Resolution. Available online: https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial (accessed on 3 September 2020).
16. Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.R.; Murayama, Y.; Ranagalage, M. Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sens.* **2020**, *12*, 2291. [CrossRef]
17. ESA. Copernicus Open Access Hub. Available online: https://scihub.copernicus.eu/ (accessed on 3 September 2020).
18. CNES. French Access to the Sentinel Products. Available online: https://peps.cnes.fr/rocket/#/home (accessed on 1 September 2020).
19. Google. Sentinel-2 Data. Available online: https://cloud.google.com/storage/docs/public-datasets/sentinel-2 (accessed on 12 December 2020).
20. ESA. Sentinel-2 Images the Globe Every 5 Days. Available online: https://earth.esa.int/web/sentinel/missions/sentinel-2/news/-/asset_publisher/Ac0d/content/sentinel-2-images-the-globe-every-5-days (accessed on 21 July 2020).
21. ESA. CCI LAND COVER—S2 Prototype Land Cover 20 m Map of Africa 2016. Available online: http://2016africalandcover20m.esrin.esa.int/ (accessed on 1 February 2021).
22. ESA. *Sentinel-2 User Handbook*; Technical Report; European Space Agency. 2015. Available online: https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/-/asset_publisher/Wk0TKajiISaR/content/sentinel-2-user-handbook (accessed on 24 March 2021).
23. Hagolle, O. The Sentinel-2 Tiles, How They Work? Available online: https://labo.obs-mip.fr/multitemp/the-sentinel-2-tiles-how-they-work/ (accessed on 30 January 2021).
24. Anon. Africa Global LANd Cover Mapping and Estimation (GLANCE) Grids. Available online: https://measures-glance.github.io/glance-grids/grids/af (accessed on 25 May 2020).
25. Anon. 30-Meter SRTM Tile Downloader. Available online: https://dwtkns.com/srtm30m/ (accessed on 3 May 2020).
26. Anon. Geospatial Data Abstraction Library. Available online: https://gdal.org/programs/gdalbuildvrt.html (accessed on 4 April 2020).
27. Frantz, D. FORCE. Available online: https://force-eo.readthedocs.io/en/latest/ (accessed on 5 September 2020).
28. Frantz, D. FORCE—Landsat + Sentinel-2 Analysis Ready Data and Beyond. *Remote Sens.* **2019**, *11*, 1124. [CrossRef]

29. Frantz, D. Gnu General Public License. Available online: https://force-eo.readthedocs.io/en/latest/policy/license.html?highlight=GNU#gnu-general-public-license (accessed on 5 September 2020).
30. Frantz, D. GITHUB. Available online: https://github.com/davidfrantz/force (accessed on 5 May 2020).
31. Eklundh, L.; Jönsson, P. TIMESAT: A Software Package for Time-Series Processing and Assessment of Vegetation Dynamics. In *Remote Sensing Time Series. Remote Sensing and Digital Image Processing*; Kuenzer, C., Dech, S., Wagner, W., Eds.; Springer: Cham, Switzerland, 2015; Volume 22.
32. Jamali, S.; Seaquist, J.; Eklundh, L.; Ardö, J. Automated mapping of vegetation trends with polynomials using NDVI imagery over the Sahel. *Remote Sens. Environ.* **2014**, *141*, 79–89. [CrossRef]
33. Jamali, S.; Jönsson, P.; Eklundh, L.; Ardö, J.; Seaquist, J. Detecting changes in vegetation trends using time series segmentation. *Remote Sens. Environ.* **2015**, *156*, 182–195. [CrossRef]
34. Verbesselt, J.; Hyndman, R.; Zeileis, A.; Culvenor, D. Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. *Remote Sens. Environ.* **2010**, *114*, 2970–2980. [CrossRef]
35. ESA. Level-1C Processing. Available online: https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-1c-processing (accessed on 30 August 2020).
36. Frantz, D.; Röder, A.; Stellmes, M.; Hill, J. An Operational Radiometric Landsat Preprocessing Framework for Large-Area Time Series Applications. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 3928–3943. [CrossRef]
37. Frantz, D.; Röder, A.; Udelhoven, T.; Schmidt, M. Enhancing the Detectability of Clouds and Their Shadows in Multitemporal Dryland Landsat Imagery: Extending Fmask. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1242–1246. [CrossRef]
38. Doxani, G.; Vermote, E.; Roger, J.C.; Gascon, F.; Adriaensen, S.; Frantz, D.; Hagolle, O.; Hollstein, A.; Kirches, G.; Li, F.; et al. Atmospheric Correction Inter-Comparison Exercise. *Remote Sens.* **2018**, *10*, 352. [CrossRef]
39. Frantz, D.; Stellmes, M.; Röder, A.; Udelhoven, T.; Mader, S.; Hill, J. Improving the Spatial Resolution of Land Surface Phenology by Fusing Medium- and Coarse-Resolution Inputs. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4153–4164. [CrossRef]
40. Frantz, D. FORCE. Available online: https://force-eo.readthedocs.io/en/latest/components/higher-level/cso/index.html#cso (accessed on 29 January 2021).
41. Jönsson, P.; Cai, Z.; Melaas, E.; Friedl, M.A.; Eklundh, L. A Method for Robust Estimation of Vegetation Seasonality from Landsat and Sentinel-2 Time Series Data. *Remote Sens.* **2018**, *10*, 635. [CrossRef]
42. Griffiths, P.; van der Linden, S.; Kuemmerle, T.; Hostert, P. A Pixel-Based Landsat Compositing Algorithm for Large Area Land Cover Mapping. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 2088–2101. [CrossRef]
43. Phenology-adaptive pixel-based compositing using optical earth observation imagery. *Remote Sens. Environ.* **2017**, *190*, 331–347. [CrossRef]
44. Frantz, D. FORCE. Available online: https://force-eo.readthedocs.io/en/latest/components/higher-level/l3/index.html#level3 (accessed on 29 January 2021).
45. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.; Gao, X.; Ferreira, L. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
46. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]
47. L3HARRIS-Geospatial. ENVI Image Files. Available online: https://www.l3harrisgeospatial.com/docs/ENVIImageFiles.html (accessed on 2 February 2021)
48. Abdi, A.M.; Seaquist, J.; Tenenbaum, D.E.; Eklundh, L.; Ardö, J. The supply and demand of net primary production in the Sahel. *Environ. Res. Lett.* **2014**, *9*. [CrossRef]
49. Frantz, D. FORCE—Higher Level. Available online: https://force-eo.readthedocs.io/en/latest/components/higher-level/index.html (accessed on 22 March 2021).
50. Defourny, P.; Bontemps, S.; Bellemans, N.; Cara, C.; Dedieu, G.; Guzzonato, E.; Hagolle, O.; Inglada, J.; Nicola, L.; Rabaute, T.; et al. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sens. Environ.* **2019**, *221*, 551–568. [CrossRef]
51. Sjöström, M.; Ardö, J.; Arneth, A.; Boulain, N.; Cappelaere, B.; Eklundh, L.; de Grandcourt, A.; Kutsch, W.L.; Merbold, L.; Nouvellon, Y.; et al. Exploring the potential of MODIS EVI for modeling gross primary production across African ecosystems. *Remote Sens. Environ.* **2011**, *115*, 1081–1089. [CrossRef]
52. Ma, X.; Huete, A.; Yu, Q.; Restrepo-Coupe, N.; Beringer, J.; Hutley, L.B.; Kanniah, K.D.; Cleverly, J.; Eamus, D. Parameterization of an ecosystem light-use-efficiency model for predicting savanna GPP using MODIS EVI. *Remote Sens. Environ.* **2014**, *154*, 253–271. [CrossRef]
53. Verbesselt, J.; Hyndman, R.; Newnham, G.; Culvenor, D. Detecting trend and seasonal changes in satellite image time series. *Remote Sens. Environ.* **2010**, *114*, 106–115. [CrossRef]
54. Kilama Luwa, J.; Bamutaze, Y.; Majaliwa Mwanjalolo, J.G.; Waiswa, D.; Pilesjö, P.; Mukengere, E.B. Impacts of land use and land cover change in response to different driving forces in Uganda: Evidence from a review. *Afr. Geogr. Rev.* **2020**, 1–17. [CrossRef]
55. Avitabile, V.; Baccini, A.; Friedl, M.A.; Schmullius, C. Capabilities and limitations of Landsat and land cover data for aboveground woody biomass estimation of Uganda. *Remote Sens. Environ.* **2012**, *117*, 366–380. [CrossRef]
56. Anonymous. *Towards Ecosystem Accounts for Uganda—Uganda Natural Capital Accounting Program*; Report; Uganda Bureau of Statistics: Kampala, Uganda, 2020.

57. ESA. Data Product Quality Reports. Available online: https://sentinels.copernicus.eu/web/sentinel/data-product-quality-reports (accessed on 17 March 2021).
58. ESA. L1C Data Quality Report Ref. S2-PDGS-MPC-DQR. 2021. Available online: https://sentinel.esa.int/documents/247904/685211/Sentinel-2+Data+Quality+Report+(DQR)/f42497d3-611f-4165-bcc1-2f81421c646a (accessed on 22 March 2021).
59. Bauer, P.; Stevens, B.; Hazeleger, W. A digital twin of Earth for the green transition. *Nat. Clim. Chang.* **2021**, *11*, 80–83. [CrossRef]
60. Bauer, P.; Dueben, P.D.; Hoefler, T.; Quintino, T.; Schulthess, T.C.; Wedi, N.P. The digital revolution of Earth-system science. *Nat. Comput. Sci.* **2021**, *1*, 104–113. [CrossRef]
61. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; Prabhat. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204. [CrossRef] [PubMed]
62. Frantz, D. FORCE—L3 Naming Convention. Available online: https://force-eo.readthedocs.io/en/latest/components/higher-level/l3/format.html#naming-convention (accessed on 1 September 2020).
63. Frantz, D. FORCE—CSO Naming Convention. Available online: https://force-eo.readthedocs.io/en/latest/components/higher-level/cso/format.html#naming-convention (accessed on 5 January 2021).
64. Frantz, D. FORCE—TSA Naming Convention. Available online: https://force-eo.readthedocs.io/en/latest/components/higher-level/tsa/format.html#naming-convention (accessed on 5 January 2021).