PEOPLE
Ecosystem Restoration

D2A: STATE-OF-THE-ART REVIEW

January 2023

esa · Hatfield · UBC THE UNIVERSITY OF BRITISH COLUMBIA · VTT · ICAS

PIONEER EARTH OBSERVATION APPLICATIONS FOR THE ENVIRONMENT – ECOSYSTEM RESTORATION (PEOPLE-ER)

D2A: STATE-OF-THE-ART REVIEW

Prepared for:

ESA/ESRIN
LARGO GALILEO GALILEI 1
00044 FRASCATI (ROMA)
ITALY

Prepared by:

HATFIELD CONSULTANTS LLP
#200 - 850 HARBOURSIDE DRIVE
NORTH VANCOUVER, BC
CANADA V7P 0A3
TEL: 1.604.926.3261 • WWW.HATFIELDGROUP.COM

JANUARY 2023

ESA11339
VERSION 1

TABLE OF CONTENTS

LIST OF ACRONYMS

ADB	Asian Development Bank
ADES	Application Deployment & Execution Service
AI	Artificial Intelligence
AMA	Analytical Mechanics Associates
ANR	Assisted Natural Regeneration
APN	African Parks Network
API	Application Programming Interface
AusAid	Australian Agency for International Development
AVI	Advanced Vegetation Index
AWS	Amazon Web Services
BAP	Best Available Pixel
BRDF	Bidirectional Reflectance Distribution Function
BSI	Bare Soil Index
C3S	Copernicus Climate Change Service
CCD	Continuous Change Detection
CCDC	Continuous Change Detection and Classification
CEOS	NASA/Committee on Earth Observation Satellite
CHS	Cloud Hosting Solutions
CLMS	Copernicus Land Monitoring Service
CSIRO	Commonwealth Scientific and Industrial Research Organisation
COLD	COntinuous monitoring of Land Disturbance
DAACs	Distributed Active Archive Centers
DEA	Digital Earth Australia
DIAS	Data and Information Access Services
dNBR	Differenced Normalized Burn Ratio
DOY	Day Of Year
EBV	Essential Biodiversity Variables
EC	European Commission
ECMWF	European Centre for Medium-Range Weather Forecasts
EDC	Euro Data Cube
EEA	European Economic Area
ENET	Elastic NET
EO	Earth Observation
EOEPCA	EO Exploitation Platform Common Architecture
ER	Ecosystem Restoration
EROS	Earth Resources Observation and Science
ESA	European Space Agency
EU	European Union
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites

FAO	Food and Agriculture Organization of the United Nations
FCC	Forest Canopy Cover
FLR	Forest and Landscape Restoration
FLRM	Forest and Landscape Restoration Mechanism
FP	Framework Programs
F-TEP	Forestry TEP
GEDI	Global Ecosystem Dynamics Investigation
GEO	Group on Earth Observations
GFOI	Global Forest Observations Initiative
GHCN	Global Historical Climatology Network
GHRC	Global Hydrology Research Center
GPFLR	Global Partnership on Forest and Landscape Restoration
G-POD	Grid Processing On Demand
GUI	Graphical User Interface
HadISD	Hadley Centre Integrated Surface Database
HANTS	Harmonic ANalysis of Time Series
HAPO	Harmonic Adaptive Penalty Operator
HLS	Harmonized Landsat Sentinel
HPC	High Performance Computing
IDL	Interactive Data Language
INCDS	National Institute for Research and Development in Forestry
IPCC	Intergovernmental Panel on Climate Change
IPT-Poland	EO Innovation Platform Testbed Poland
IRSS	Integrated Remote Sensing Studio
IUCN	International Union for Conservation of Nature
IW	Interferometric Wide
LASSO	Least Absolute Shrinkage and Selection Operator
LUKE	Natural Resources Institute (Finland)
MK	Mann-Kendall
MEPs	Mission Exploitation Platforms
NASA	National Aeronautics and Space Administration (USA)
NBR	Normalized Burn Ratio
NBS	Nature-Based Solutions
NDVI	Normalized Difference Vegetation Index (NDVI)
NOAA	National Oceanic and Atmospheric Administration
OGC	Open Geospatial Consortium
OLS	Ordinary Least Squares
PEOPLE	Pioneer Earth Observation apPlications for the Environment
R&D	Research and Development
RF	Random Forest
RFI	Regional Flyway Initiative

REPs	Regional Exploitation Platforms
RI	Recovery Indicator
ROMSILVA	National Forest Administration (Romania)
RS	Remote Sensing
RTO	Research and Technology Organizations (RTO)
R-VSPI	RADAR-Vegetation Structural Perpendicular Index
SAR	Synthetic-Aperture Radar
SDG	Sustainable Development Goals
SERNbc	Society for Ecosystem Restoration in northern British Columbia
SI	Shadow Index
SMEs	Small and Medium Enterprises
SMOSAR	Soil MOisture retrieval from multi-temporal SAR
SSEP	Super Site Exploitation Platform
SSI	Scaled Shadow Index
SSM	Surface Soil Moisture
STAC	SpatioTemporal Asset Catalogs
SVM	Support Vector Machine
TCC	Tree Canopy Density
TEPs	Thematic Exploitation Platforms
UKAid	United Kingdom Agency for International Development
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme
UNFCCC	United Nations Framework Convention on Climate Change
UN-REDD	United Nations Programme on Reducing Emissions from Deforestation and Forest Degradation
US	United States
USAID	United States Agency for International Development
USGS	US Geological Survey
UTM	Universal Transverse Mercator
VHR	Very High Resolution
VSPI	Vegetation Structural Perpendicular Index
WMO	World Meteorological Organization
WRI	World Resources Institute
WWF	World Wildlife Fund
XGBoost	Extreme Gradient Boosting
Y2R	Years to Recovery

DISTRIBUTION LIST

The following individuals/firms have received this document:

Name	Firm	Email	SharePoint	Project Website
Frank Martin Seifert	ESA	✓	-	-
Early Adopters representative	IUCN Vietnam, African Parks, Luke, INCDS, SERNbc	✓	✓	-
	Public		--	✓

AMENDMENT RECORD

This report has been issued and amended as follows:

Issue	Description	Date	Approved by	
1	First version of D2 – State of the Art Review	20230131	Andy Dean Project Manager	Marcos Kavlin Assistant Project Manager

1.0   INTRODUCTION

Ecosystem Restoration (ER) is important to reverse biodiversity loss and is a critical element of nature-based solutions (NBS) for climate change mitigation and adaptation, food security, and disaster risk reduction. ER is needed on a large scale to achieve the United Nations (UN) sustainable development agenda and as part of the UN Decade on Ecosystem Restoration (2021−2030). At the Convention on Biological Diversity (CBD) COP 15 in Montreal in December 2022, nations adopted a target to "Ensure that by 2030 at least 30 percent of areas of degraded terrestrial, inland water, and coastal and marine ecosystems are under effective restoration, in order to enhance biodiversity and ecosystem functions and services, ecological integrity and connectivity."[1]

Effective planning, monitoring, and assessment of ER is required to evaluate ecosystem functions and to determine whether ER is having the desired impact. ER investments must be data-driven, requiring historical information on ecosystem disturbance and degradation, to enable planning of interventions, which are then monitored for their impact. There is a huge opportunity for satellite Earth Observation (EO) applications for ER, to meet the needs for regular, repeat measures of ER processes over long time periods covering large, often remote, areas.

To support ER investments, innovative methods are required to deliver high-quality EO-based products and indicators targeting high-priority biodiversity variables.

The Pioneer Earth Observation apPlications for the Environment (PEOPLE) ER project financed by the European Space Agency (ESA) is a trailblazer project to develop innovative high-quality EO-based application products, indicators, and methods, targeting ER research and development (R&D) priorities.

PEOPLE-ER is led by Hatfield Consultants – a science-driven service-oriented company that builds solutions to complex environmental challenges, with a depth of experience in ER projects in Canada and around the world. Hatfield is a trusted partner for the development of cutting-edge and practical EO technologies. The PEOPLE-ER consortium includes:

- VTT – the remote sensing team at VTT Technical Research Centre of Finland produces EO data processing chains for domestic and international users. The team is internationally known, particularly for its forest monitoring applications and the Forestry TEP cloud processing platform. VTT is ranked among the leading European Research and Technology Organisations (RTO).

- University of British Columbia, Faculty of Forestry – Dr. Nicholas Coops leads the Integrated Remote Sensing Studio (IRSS) and is a leading international research scientist in the application of EO technologies for forest ecosystem assessment and monitoring, including ER and the prioritization of methods and products for remote sensing essential biodiversity variables (RS-EBVs).

The Early Adopters are:

- **National Institute for Research and Development in Forestry (INCDS)** (Romania) – formally a member of the consortium, INCDS is the main organisation of research and development in forestry from Romania. INCDS is in charge for the forest resources assessment and monitoring

---

[1] HYPERLINK "https://www.cbd.int/article/cop15-cbd-press-release-final-19dec2022"https://www.cbd.int/article/cop15-cbd-press-release-final-19dec2022

in Romania through National Forest Inventory. INCDS has also secured the support of two Romanian NGOs as documented in letters of support: Forestry Society Association and Fundatia Grupul Verde Oradea.

- **IUCN** (Vietnam) – established in 1948, IUCN is an international authority working on a wide range of themes related to nature conservation, forests, ecosystem management, protected areas, global policy and governance and rights.

- **African Parks Network** – APN is a leading non-profit conservation organisation that takes on the complete responsibility for the rehabilitation and long-term management of national parks across Africa in partnership with governments and local communities.

- **Society for Ecosystem Restoration in northern British Columbia (SERNbc) (Canada)** – a key enabler for ER in forested ecosystems affected by cumulative disturbances from forest operations, energy exploration, wildfires, and forest pests/diseases.

- **Natural Resources Institute (Luke) (Finland)** – as one of the biggest clusters of bioeconomy expertise in Europe, Luke develops knowledge-based solution models and services for renewable natural resources management and supports decision-making in society.

1.1   SCOPE

This is document is part of Deliverable 2 (D2) and addresses the State-of-the-art review of relevant EO algorithms, methods, models, non-EO data, information technology with direct importance to ecosystem restoration.

2.0   ECOSYSTEM RESTORATION

ER is the process of assisting the recovery of an ecosystem that has been degraded, damaged, or destroyed to restore habitat and ecosystem functions. Restoration activities often initiate a trajectory of ecosystem recovery, which can take years, decades, or longer (Society for Ecological Restoration 2022).

The are many approaches to ER, from passive approaches which allow vegetation natural regeneration (e.g., following a wildfire) to active restoration involving planting or full or partial removal of vegetation (planting, removing invasive species). Assisted natural regeneration (ANR) is a more recently recognized approach with a blend of active and passive approaches with the goal of eliminating barriers and threats to natural recovery.

Examples of ANR include:

- Preventing degradation (e.g., fire breaks, reducing domestic animal grazing)

- Restoring natural functions (e.g., re-opening an area to flooding or natural occurrence of fire)

- Improving ecosystem structure / configuration (e.g., reducing fragmentation, improving forest structural diversity)

ANR is receiving attention due to it being cost-effective and the potential of NBS to contribute to mitigation of climate change through carbon sequestration. Through ANR and taking a landscape approach, improvement of the conditions of the ecosystem as well as local communities is an important goal and expected to improve permanence and resilience of ecosystem improvements.

Forest and Landscape Restoration (FLR) changes the scope and the dimension of restoration. Rather than taking a site-level approach, FLR moves towards a landscape view, aiming to restore ecosystem services (Laestadius et al. 2015) and improve the livelihoods of humans along the way (WRI 2011). Nevertheless, this novel landscape-scale approach does have its complications. Restoring an entire landscape implies a diverse suite of approaches and navigating complex relationships with the people who live there.

This increase in scale implies greater focus on the cost associated with restoration. Thus, lead to much debate as to whether to include passive regeneration into the equation and count it as part of FLR approaches, in part due to the unpredictability of passive regeneration (Chazdon 2008). However, it was found that natural regeneration could obtain greater success than human-led interventions, all while costing less (Chazdon 2008). Natural regeneration was therefore included as a potential approach when conducting FLR (Adams et al. 2016; Lazos-Chavero et al. 2016).

Both natural regeneration and active restoration are important when it comes to the climate crisis, however, it is important to note that both approaches are long-term processes and need to persist across the landscape to achieve their goals (Chazdon 2008).

EO data and remote sensing approaches allow us to monitor and learn from the successes and failures of different ER and FLR initiatives, thus allowing practitioners to make better decisions with more information to back them up.

Monitoring the progress of restoration initiatives relies on understanding the practitioner's goals, and whether they fall under culture, food & products, water, energy, biodiversity, soil, or climate themes (FAO and WRI 2019). The FAO and WRI have developed a framework for identifying the priorities and indicators for the monitoring of forest and landscape restoration initiatives. It can be summarized in three steps (FAO and WRI 2019):

- Determine goals, land-use, and barriers

- Filter by constraints, priorities, and data availability

- Set up a system based on indicators, metrics, and optional indexes

EO tools can be designed within the context of this framework to maximize the applicability and efficacy of the monitoring program, granting valuable context, and expanding the possible indicators with which to evaluate restoration initiatives.

3.0   STATE-OF-THE-ART ANALYSIS

EO data and associated analytical approaches have demonstrated consistent and transparent capabilities for spatially explicit characterization of ecosystem disturbance and recovery processes. EO offers a means to assess ecosystem characteristics across larges areas and to provide quantitative and transparent data for guiding ER activities.

This state-of-the-art analysis is structured based on scientific methods and models, algorithms, and information technology / platforms, with a focus on the capabilities in the context of datasets and methods that ecologists use and value to understand ecosystem condition.

The **essential biodiversity variables (EBVs)** provide a common framework of complementary biological measurements for capturing considerable biodiversity change and are produced by

integrating primary observations from in situ monitoring with remote sensing. A prioritized list of remote sensing biodiversity products was developed to improve the monitoring of geospatial biodiversity patterns, enhancing the EBV framework and its applicability. **Ecosystem structure** and **ecosystem function** EBV classes, which capture the biological effects of disturbance as well as habitat structure, were shown to be the most relevant, feasible, accurate and mature for direct monitoring of biodiversity using EO (Skidmore et al. 2021).

Our review of the state of the art aims to identify the ER methods in context of priority RS-EBVs. Our approach is to identify those methods with a sufficient application readiness level to be integrated into the platforms, while simultaneously advancing the state of the art for the PEOPLE-ER early adopters.

3.1 METHODS AND MODELS

3.1.1 Compositing and Time Series Harmonization

Time series analysis is fundamental for ER and may be based on annual or seasonal composites of multi-spectral or radar data. Creating cloud-free, seasonally coherent composite multi-spectral images is a fundamental requirement and often a major challenge for ER practitioners.

A composite requires two key steps – cloud-detection on a single-scene pixel-basis followed by a selection of the best pixel to be part of a composite image.

Most cloud detection methodologies rely on physical rules and use spectral indices to separate clear sky pixels from clouds (Tarrio et al. 2020). Within the optical ranges of the spectrum, one can distinguish cloudy pixels from clear pixels by their higher reflectance among most wavelengths. Additionally, the thermal band can be useful in the detection of clouds due to their lower temperature when compared to clear pixels. Another band which can help improve detection of cirrus clouds is the SWIR (shortwave infrared) (Zhu et al. 2015; Zhu et al. 2019).

Some of the most common cloud detection algorithms include: SEN2COR, FMASK, LASRC, MAJA, TMASK, and most recently s2cloudless. These algorithms are explored in more detail in section 3.2.2.

Median and Mean Composite with Temporal Metrics

An effective approach to generate a composite is to use statistical measure to select a pixel. Common approaches include median and mean. In the case of a median composite, each pixel is selected from a range of images to have the median (or middle) value out of all possible values value. An advantage is that the median is relatively robust to outliers and the value is a real, observed measurement. Mean composites involve taking the average value for each pixel. Unlike the median, the mean composite can contain pixel values that were not part of the original dataset.

Statistical analysis applied to image data from a specific period (e.g., a year or season) can support understanding of vegetation seasonality and phenology, which can be useful for ER analysis. For example, 75th percentile or max and min values of a vegetation index.

Best Available Pixel (BAP)

Scientists at Natural Resources Canada pioneered the utilization of the full time series of Landsat satellite data for forest and land cover change monitoring in Canada. Utilizing the full Landsat archive, standardized calibration of image products, and increasing computer processing and storage capabilities, method have been established to produce large-area, cloud-free, surface reflectance pixel-

based image composites (White et al. 2014). With BAP composites, change monitoring no longer relies on scene-based analysis. Annual BAP composites are surface reflectance composites that use the best available pixel observation (from the target year) for any given pixel location, selecting from a temporal series of candidate images. Annual composites are produced using a set of specified rules that are defined according to the information need. For example, an annual composite may be designed to capture a specific time or a limited phenological window (White et al. 2014). The BAP approach integrates scores for pixel selection: sensor score, day of year (DOY) score, distance to cloud or cloud shadow score, and opacity score.

The sensor and DOY score are calculated at the image level (i.e., all pixels within the image receive the same score), whilst the cloud/cloud shadow and opacity scores are unique to each pixel. All scores were then summed to provide a total score for each pixel, and the pixel with the largest score (i.e., the BAP) was used in the image composite.

- Sensor score – Pixels from Landsat TM images are assigned a score of 1; pixels in Landsat ETM+ images are assigned a score of 0.5. Pre-2003, both sensors receive a score of 1.

- DOY – A score is assigned to all pixels in an image according to the DOY the image was acquired relative to the target DOY.

- Distance to Cloud or Cloud Shadow Score – Using the outputs from Fmask[2], a distance to cloud or cloud shadow score is assigned, whereby pixels identified as clouds or cloud shadows are assigned "no data" value and any pixel located at a distance greater than 50 pixels from an identified cloud or cloud shadow pixel is assigned a score of 1. Pixels that are not identified as clouds or cloud shadows and that are less than 50 pixels away from clouds and cloud shadows are assigned a score between 0 and 1.

- Opacity Score – Since hazy images can confound the generation of quality image composites, an opacity score was calculated using the atmospheric opacity band output by LEDAPS[3]. Pixels with an opacity value < 0.2 were assigned a score of 1 and pixels with an opacity value > 0.3 were labelled as "no data". Pixels with opacity values ≥ 0.2 and < 0.3 were assigned a score between 0 and 1.

The BAP composites have been used in a variety of applications related to land cover and land use change and recovery of ecosystems following disturbance. A notable example was the National Terrestrial Ecosystem Monitoring System and the "Composite2Change" approach, which applied spectral trend analysis to the BAP composites to identify disturbance and recovery of forests (Hermosilla et al. 2015).

Sentinel-2 and Landsat Harmonization

An increase in the temporal revisit of multi-spectral satellite can increase the likelihood of obtaining cloud- and shadow-free observations as well as to improve mapping of rapidly- or seasonally changing features.

---

[2] C implementation of Fmask as supplied with Landsat imagery: https://www.usgs.gov/core-science-systems/nli/landsat/cfmask-algorithm

[3] https://www.usgs.gov/media/files/landsat-4-7-collection-1-surface-reflectance-code-ledaps-product-guide

The USGS Harmonized Landsat Sentinel (HLS)[4] uses a processing chain involving several separate radiometric and geometric adjustments, with a goal of eliminating differences in retrieved surface reflectance arising solely from differences in instrumentation. Input data products from Landsat 8 (Collection 2 Level 1T top-of-atmosphere reflectance or top-of-atmosphere apparent temperature) and Sentinel-2 (L1C top-of-atmosphere reflectance) are ingested for HLS processing. A series of radiometric and geometric corrections are applied to convert data to surface reflectance, adjust for BRDF (Bidirectional Reflectance Distribution Function) differences, and adjust for spectral bandpass differences. For more information on HLS products see Section 3.2.1.

The Copernicus Sen2Like method provides another solution for harmonizing and fusing Landsat 8/Landsat 9 data with Sentinel-2 data (Saunier et al. 2022). For details on the Sen2Like algorithm see Section 3.2.1.

3.1.2   Vegetation Recovery / Trends

In forested ecosystems, the analysis of time series satellite data enables monitoring of multiple aspects of forest recovery over time: the extent and type of forest disturbance (Senf et al. 2015; White et al. 2017; Häme et al. 2020; Antropov et al. 2021), the return of vegetation, the re-establishment of trees (White et al. 2018), the return of forest structure (Senf et al. 2019), and the return of pre-disturbance forest spatial patterns (Hermosilla et al. 2019).

Inputs to recovery or trend analysis may be annual or seasonal composites or complete time series of multi-spectral or radar datasets.

EBV Class: Ecosystem structure and function

RS-EBV products:

- (4) Land cover (vegetation type) (recovery is used to classify change in land cover, e.g., bare -> shrub -> tree)

- (20) Land surface peak (maximum of season)

Spectral Recovery

Time series of remotely sensed imagery can be a useful source of information when seeking to understand forest recovery over entire landscapes (White et al. 2018). Data sources such as the Landsat time series have allowed researchers study disturbances across forested landscapes and the ensuing recovery at diverse scales (regions, countries, and the globe) (Schroeder et al. 2006; Hansen et al. 2013; White et al. 2017). To evaluate the recovery of an ecosystem using this data we rely on the relationship between spectral indicators and the actual recovery in forest structure (White et al. 2018). These indicators are often derived from fitted time-series that use different spectral indexes such as the Normalized Burn Ratio (NBR), Normalized Difference Vegetation Index (NDVI), and others.

Some of these indicators include:

- Years to Recovery (Y2R) – Years to recovery refers to the number of years it takes a pixel to recover to 80% of its pre-disturbance value (Pickell et al. 2016; White et al. 2017; White et al.

---

[4] https://lpdaac.usgs.gov/data/get-started-data/collection-overview/missions/harmonized-landsat-sentinel-2-hls-overview/

2018). To calculate the pre-disturbance value, one uses the average value of the pre-disturbance value of the previous 2 years for the pixel. This indicator is used as a long-term tool to assess forest regeneration, which enables it link spectral recovery with structural and compositional recovery of a pixel (White et al. 2018).

- Δ (Index) – The change (Δ) in an index (e.g., NBR) at a specific point in time after recovery is another metric that can be used to assess short-term recovery in a pixel (White et al. 2018).

- Recovery Indicator (RI) – The recovery indicator shows the spectral change (Δ) in an index over a recovery period as relative to the spectral change (Δ) of the same index during the disturbance event (White et al. 2018).

- Relative Spectral Magnitude – Characterized as the magnitude of the recovery relative to the spectral value post disturbance (Nguyen et al. 2018).

Although we have listed four metrics, there are many more metrics available. It is important to note that the best indicator of recovery can vary in each landscape, and it is important to tailor one's approach for each use case (Liu et al. 2021).

Temporal fitting – e.g., LandTrendr

LandTrendr is a method for land cover change detection that is based on long time series of Landsat satellite data, rather than relying on pairwise comparison between two dates (Kennedy et al. 2010). The algorithm aims to resolve relevant change of the time series while eliminating noise introduced by ephemeral changes in illumination, phenology, atmospheric condition, and geometric registration. LandTrendr is based on annual image composites based on multiple images per year (median composite approach) followed by the extraction of temporal trajectories of spectral data on a pixel-by-pixel basis. Temporal segmentation allows capture of slowly evolving processes, such as regrowth, and abrupt events, such as forest harvest. A range of parameters and threshold-based filtering is used to reduce the role of false positive detections.

The temporal trajectory of data values is a sequence of connected linear segments bounded by nodes or breakpoints referred to as "vertices". The incoming data values can be spectral bands, derived spectral indices, or even other metrics that themselves capture yearly behavior. The only constraint is that there can be one value per year. Two phases are used to find vertices. In a **forward** phase, candidate vertices are identified through an iterative anomaly detection criterion, and in a **reverse** phase those candidate vertices are culled using an angle-of-change criterion. Once a user-defined maximum number of vertices is identified, straight-line segments are fit to the observed spectral metric values, working from the first year in the time series to the last. Fitting uses either simple regression or point-to-point fitting, constrained to ensure that the starting vertex of a subsequent segment is anchored to the ending vertex of the prior segment. From this best-fitting model with the maximum number of vertices, an iterative process of vertex-removal and re-fitting is conducted to find successively simpler renditions of the time series. At each step in the iteration, goodness of fit statistics are calculated that penalize more complex fits, and fitting statistics are compared against a user-defined threshold. Finally, from these successively simpler renditions of the time series, the best model is chosen based on the best fitting statistic, with a user-defined parameter allowing more complex fits to be chosen if they are within a user-defined proportion of the best fit statistic (Kennedy et al. 2018).

LandTrendr utilizes pre-processed, calibrated, and cloud masked Landsat time series data where a single data value per year must be identified (i.e., annual composite). Subsequently, LandTrendr applies

a despiking algorithm and a process to address gaps or no data values each year, representing a crucial feature of the algorithm.

Harmonic Modelling

To understand the effects of natural and anthropogenic disturbances, and the subsequent regeneration of a landscape we must first be able to model the seasonality of the vegetation (Zhou et al. 2022).

Continuous Change Detection and Classification (CCDC) is an example of a temporal segmentation algorithm that uses harmonic modelling to distinguish intra-annual change in vegetation due to seasonal phenology from gradual trends or abrupt changes (Zhu and Woodcock 2014; Zhu et al. 2020). See Section 3.2.2 for details on CCDC.

A widely recognized method to model time-series observations using harmonic components is the Harmonic ANalysis of Time Series (HANTS) model (Menenti et al. 1993; Verhoef, W. 1996; Roerink et al. 2000). This model deals with missing data from the time-series using by filling these gaps using an Ordinary Least Squares (OLS) approach to calculate significant frequencies which are expected to be in the time-series (Zhou et al. 2022).

However, the HANTS model, when using OLS can be prone to overfitting. To address this limitation other regression approaches such as LASSO (Least Absolute Shrinkage and Selection Operator) (Tibshirani 1996), and Ridge regression (Hoerl and Kennard 1970) have been developed. The use of these approaches often improves the accuracy of the HANTS model for harmonic analysis, although it can still struggle with overfitting when dealing with high frequency cycles (Zhou et al. 2022).

The novel Harmonic Adaptive Penalty Operator (HAPO) (Zhou et al. 2022) was found to show highly accurate HANTS model results and outperform OLS, LASSO and Ridge in multi-year time-series experiments (Zhou et al. 2022). This regression method can be used with Sentinel-2, Landsat, and harmonized multi-sensor time-series data (Zhou et al. 2022).

Accounting for Climate Variability

Climate inputs exert considerable influences on ecosystem characteristics and functions. Ongoing climate changes have been gradually transforming thermal and hydrological regimes of terrestrial ecosystems (Denissen et al. 2022). Conducting a robust climate trend analysis will help to identify potential key drivers of ecological changes and understand the mechanisms of ecosystem degradation and restoration.

The Mann-Kendall (MK) statistical test has been commonly applied to detect monotonic trends in meteorological and hydrological time series (Zhang et al. 2001; Burn and Hag Elnur 2002; Ford et al. 2019; Luhunga and Songoro 2020). As a rank-based non-parametric test, it has no requirements of homoscedasticity or prior assumptions on the distribution of the data sample (Önöz and Bayazit 2003) and is tolerable to missing data (Ford et al. 2019). Original MK test was found subject to serial correlation (autocorrelation) within time series, leading to over-rejections of the null hypothesis of no trend (Yue and Wang 2002). A modified version of MK test is the Trend-Free Pre-Whitening (TFPW) (Yue et al. 2002), which was proposed to eliminate the influence of serial correlation while maintaining primary trend structure in the time series. TFPW algorithm is available in a Python package named *pyMannKendall* (Hussain and Mahmud 2019).

Although MK test indicates significance level of trends, it does not provide information regarding rate or magnitude of changes. As a supplement to MK test, Theil-Sen estimator can be used for estimating slope of trends. It is considered more accurate than ordinary linear regression for skewed and heteroskedastic data and less sensitive to outliers (Wilcox 1998). Due to its robustness and efficiency, Theil-Sen estimator has become one of the most popular non-parametric techniques to estimate linear trends for long-term time series of hydrometeorological variables (Luhunga and Songoro 2020; Yao et al. 2021 Jun 25) and vegetation indices (Myers-Smith et al. 2020).

3.1.3 Wetness / Water Variability and Dynamics

For wetland ecosystems, the location and persistence of surface water (inland and coastal) is a key driver of biological diversity and ecosystems services. EO time series have proven capability in the detection of surface water location and seasonality, with recent innovation in cloud computing and algorithms producing long-term datasets at high resolution.

SAR sensors are capable of distinguishing open water from other land cover classes relatively well due to shallow scattering signature and thus strong contrast with other land cover classes. C-band SAR image time series are particularly suitable for mapping water bodies with high accuracy (Lamarche et al. 2017). The extent of permanent water bodies and dynamics can be monitored. Recent land cover classification experiments with deep learning semantic segmentation models suggest very high accuracies can be obtained with dual-pol IW mode Sentinel-1 data. Further improvements can be expected from combining SAR and optical datasets for delineating inland water bodies, for example with Copernicus datasets (Gao et al. 2018).

Another relevant EBV in the context of wetness and water variability is surface soil moisture (SSM). Monitoring of SSM can improve flood and drought predictions since it affects the amount of water available for vegetation growth. Microwave based SSM retrieval (with passive and active microwave sensors) is based on the contrast between the dielectric coefficient of water and dry soil. Operational soil moisture products rely on radiometer based semiempirical-model predictions at coarse resolution with high revisit rate (Kerr et al. 2012), as well as on time series of SAR data for higher resolution products at 1 km pixel spacing (Song et al. 2021). One of the central SAR time series mapping approaches is based on relating observed change in backscatter intensity to soil moisture changes. (Wagner et al. 1999; Bauer-Marschallinger et al. 2019), while also several others are known with considerably lower technology readiness level requiring interferometric SAR or polarimetric SAR image datasets.

> EBV Class: Ecosystem function and structure
>
> RS-EBV product: (2) Biological effects of irregular inundation

Global Surface Water Permanence

Two key initiatives to produce global surface water permanence data are the Joint Research Centre (JRC) global surface water permanence dataset based on Landsat time series (Pekel et al. 2016) and the ESA financed WorldWater project[5].

WorldWater is developing scientifically robust methods that exploit the full time series of Sentinel-1, Sentinel-2 and Landsat satellite imagery to improve the capture the seasonal changes of surface waters,

---

[5] https://worldwater.earth/

and to complement these observations with radar altimetry measurements of water levels. Initial results of a round robin exercise organized within the WorldWater project confirm that high accuracies in surface water mapping can be achieved with optical and SAR data. Particularly, optical images are better at capturing spatial detail while SAR data provide a better seasonal characterization when looking at the classification performance across several study sites as well as through time. Based on these findings sensor fused approaches are recommended, with both supervised and unsupervised learning can provide very good results. Software implementations are expected to be freely available at the end of the project (at this point, the work is unpublished, and software is not available).

3.1.4   Canopy Density / Cover

The spatial location of tree canopy and the proportional tree crown coverage per unit area (pixel) is an important forest structure parameter. Calculating tree canopy cover before and after intervention, we can determine areas of stable, increasing, or decreasing tree canopy and determine the success of ER. Important methodological advances account for the seasonality of vegetation phenology, especially deciduous trees. An important methodological requirement for tree or canopy cover models is to account for the seasonality of vegetation phenology especially deciduous trees. This can be accomplished by including optical data captured at different times of year.

Supervised models are calibrated on manually labelled reference data. The EO4SD-FM project[6] describes a two-step procedure. First, a tree cover mask is produced by pixel-based classification of dense time series EO data and then a model is used to correlate EO data to the tree cover density values which are represented in a reference data. The reference data could be generated from forest inventory plot data (if available), polygons or points derived from visual image interpretation of VHR data, canopy metrics generated from the Global Ecosystem Dynamics Investigation (GEDI). Various supervised methods for classifying Forest Canopy Cover (FCC) using Sentinel-2 and derived indices (NDVI, NDVI-A, NDRE, and NDI45) (Nasiri et al. 2022). They tested and compared Random Forest (RF), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Elastic Net (ENET) algorithms, as well as evaluated their variable importance. They determined that the best performing algorithm for their use case was the RF model, and that the most important predictor variable was NDVI. This approach has the potential to tailor and improve FCC classifications on a site-specific level (Nasiri et al. 2022).

Unsupervised models such as the forest canopy density index that is derived from combination of spectral indexes that can be generated from EO data, such as the Advanced Vegetation Index (AVI), Bare Soil Index (BI) and Shadow Index (SI) or Scaled Shadow Index (SSI) (Rikimaru et al. 2002; Joshi et al. 2006; Loi et al. 2017).

The benefit of a supervised models is that it is trained used measurements of canopy cover obtained in the AOI to be mapped and can integrate multiple EO datasets. However, visual interpretation of canopy cover from VHR images is time-consuming and a limited amount of training data can be obtained, which will result in models that do not generalize well. Forest plots may be limited, and their availability should be confirmed by with the relevant forestry agencies. Whereas, using GEDI forest canopy metrics, a model can train more quickly, but GEDI data is likely of lower precision than plot measurements.

---

[6] https://www.eo4sd-forest.info/portfolio/

The benefit of unsupervised models is that they are easier to implement and can be readily used over large areas to generate wall-to-wall maps. However, even using unsupervised methods, the outputs need to be validated to assess performance and calibrated.

> EBV Class: Ecosystem structure
>
> RS-EBV product: 15 – Plant area index profile (canopy cover)

3.1.5 Canopy Height

Consistent, large-scale operational monitoring of forest canopy height is an important task for forest-related carbon emissions and quantifying the effectiveness of forest restoration. Recent research into modelling of canopy height using satellite EO data includes using ICESat-2 to develop machine learning models for Landsat and Sentinel-1 and Sentinel-2 data (Li et al. 2020) and various machine learning and physics model-based methods utilizing Landsat, ALOS PALSAR, Sentinel-1, RADARSAT-2 and TerraSAR-X/TanDEM-X imagery (Praks et al. 2012; Olesk et al. 2016; García et al. 2018; Astola et al. 2019; Antropov et al. 2021). Since April 2019, the GEDI lidar instrument onboard the International Space Station has collected vegetation structure data in a systematic sampling approach. GEDI represents a unique global dataset using which researchers can model canopy height using synoptic EO data. For example, Potapov et al. (2021) employed global Landsat analysis-ready data to extrapolate GEDI footprint-level forest canopy height measurements, creating a 30 m spatial resolution global forest canopy height map for the year 2019[7].

Combined SAR and optical datasets provide not only for interoperability, but also complement each other in mapping forest variables (Antropov et al. 2022; Ge et al. 2022). More advanced interferometric SAR datasets, such as bistatic interferometric TanDEM-X can be useful in delivering knowledge on forest vertical structure, however not readily available for wide-area routine forest mapping. The potential of 6-12 days apart repeat pass interferometric Sentinel-1 signatures in forest mapping is limited to the relatively dry or winter-season forest (Jacobs et al. 2021; Cartus et al. 2022).

The state of the art related to analytical methods is rapidly developing with the trend to move beyond the widespread application of shallow learning and statistical methods to advanced machine learning and deep learning. Recent developments include use of convolutional and recurrent deep learning methodologies for mapping forest variables, such as forest tree height, using Sentinel-2 imagery (Mottus et al. 2021), Sentinel-1 time series (Ge et al. 2022) and combined satellite SAR and optical datasets (Antropov et al. 2022; Ge et al. 2022). Another example is the global canopy height – ETH Zurich – probabilistic deep learning model.[8]

> EBV Class: Ecosystem structure
>
> RS-EBV product: 14 – vegetation height

3.1.6 Wildfire

Fire regime within forest, shrub, grassland, and peatland ecosystems are a key driver of ecosystem change, especially given interactions between fire and other disturbances, such as drought and insect

---

[7] Global Forest Canopy Height, 2019 | GLAD (umd.edu)

[8] A high-resolution canopy height model of the Earth (langnico.github.io)

outbreaks, and climate change. Fire has effects on tree regeneration, species composition, and the spatial configuration of habitat and biodiversity. ER may include mapping of burned area and ecological impacts, monitoring of vegetation recovery following fire events, and planning and assessing prescribed burning.

Multi-spectral sensors measure the reflectance of surface features covering the short and near infrared wavelengths thus providing information on the impact of wildfire on healthy vegetation when pre- and post-fire data are compared and analyzed. Well-established, simple methods include the Differenced Normalized Burn Ratio (dNBR), which is based on comparing pre- and post-fire NBR images or composite images to identify burned areas because of the loss in vegetation and structural reflectance. (Cocke et al. 2005; White and Long 2019; Coops et al. 2020). The dNBR method is commonly applied to Landsat and Sentinel-2 datasets, producing high spatial resolution data, but frequency of update can be a challenge given cloud or smoke. The Sentinel-3 SYN Burned Area product integrates Sentinel-3 and VIIRS and provides a new global burned area dataset at 300 m. Published under the ESA CCI, the new product improves temporal reporting and accuracy compared to other regional scale products (Lizundia-Loiola et al. 2022).

Radar time series may provide additional data to assess wildfire impact through the trajectory of backscatter and its linkage to burn severity and forest structure. A Sentinel-1 based index that is derived from the Vegetation Structural Perpendicular Index (VSPI) called the R-VSPI (Chhabra et al. 2022). The R-VSPI can identify wildfires as well as estimate the impact on a forested landscape by wildfires as reliably as the optical VSPI. Using the R-VSPI one can create a denser time-series due to radar's ability to pierce through clouds than the optical VSPI method (Chhabra et al. 2022). Radar's higher sensitivity for vegetation structure also led to greater detail in the assessment of wildfire impact and recovery (Minchella et al. 2009; Chhabra et al. 2022).

Both optical and radar-based approaches have complementary strengths. A mixed radar/optical based approach would provide a robust analysis of a study site allowing us to characterize the effects of wildfires pre and post disturbance reliably (Chhabra et al. 2022).

It is important to note that wildfires and their effects both have site specific effects, as well as landscape level effects. At a landscape level different fire regimes may contribute to landscape heterogeneity which in turn may be a desired restoration outcome for a project (Chia et al. 2016). It is therefore important, when monitoring a fire-prone landscape to consider both site-level changes as well as the overall landscape's context.

> EBV Class: Ecosystem structure
>
> RS-EBV product: (1) Biological effects of fire disturbance (direction, duration, abruptness, magnitude, extent and frequency)

3.1.7 Climate and Other Ancillary Data

A range of modelling and derived datasets can support assessment and monitoring for ER using satellite EO data. The most important datasets related to climate given that trends and short-term variability in climate variables can have a large effect on observed and modeled EO parameters.

The Copernicus Climate Change Service (C3S) provides historical, current, and predicted global climate data. **ERA5-Land** is a state-of-the-art global reanalysis dataset that describes evolution of the

water and energy cycles over land. It was generated by the European Center for Medium-Range Weather Forecast (ECMWF) as a part of C3S. This dataset includes a total of 50 variables closely related to terrestrial ecosystem functions and characteristics, including solar radiation, air temperature, precipitation, evaporation, and soil moisture. The period of reanalysis spans from 1950 to present at a temporal resolution of an hour. The lengthy, continuous historical data produced by ERA5 in a consistent manner are particularly useful for analysis of long-term trends and anomalies of climate states (Chiaravalloti et al. 2022; Rolle et al. 2022). ERA5-Land has a global coverage of land area at a spatial resolution of 9km. It could be an alternative data source for regions where archived observation records are absent. High degree of consistency between ERA5-Land and field observations has been reported for air temperature (Yilmaz 2023), precipitation (Lavers et al. 2022), although moderate bias may exist for soil moisture in humid regions (Wu et al. 2021). The older version, ERA-Interim, has been routinely used for WMO annual assessment and the IPCC assessment of climate changes. ERA5 provides even more accurate reanalysis hydroclimate variables than ERA-Interim (Betts et al. 2019; Gleixner et al. 2020), so it could be a reliable data source for trend analysis of climate variables.

Another useful climatology dataset is the Global Historical Climatology Network (GHCN) and UK Met Office Hadley Centre **Integrated Surface Database (HadISD)**, which provides long-term observed records of temperature and precipitation at daily scale. HadISD contains a selection of 9555 stations that passed through a standard QC test (Dunn et al. 2022), and GHCN integrated more than 100,000 stations in 180 countries and territories (Menne et al. 2012). Although they both provide valuable station-based records to analyze long-term climate trends, the data availability and completeness show great continental dependency. There is high density of stations distributed across North America and Europe, but observations in Africa are geographically sparse, especially in central Africa. Also, observation periods vary considerably among stations, ranging from several years to more than one century. Lack of consistent spatial-temporal coverage to construct location-specific time series leads to great challenges for climate trend analysis in data-scarcity regions (Mistry et al. 2022).

The **Copernicus Land Monitoring Service (CLMS)** provides geospatial information on land cover and land use change, vegetation state and water cycle. It provides systematic monitoring of biophysical parameters globally every ten days and land cover and land use mapping within Europe and globally.

Two other prominent land cover datasets are **ESA WorldCover** and **WRI Dynamic World**. WorldCover is a global land cover map for 2020 and 2021 at 10m resolution, based on Sentinel-1 and Sentinel-2 data. Dynamic World is a near real-time land use/landcover dataset made with deep learning, based on Sentinel-2 Top of Atmosphere imagery. The dataset is produced with Google Earth Engine and AI Platform at 10m resolution.

The **CopernicusDEM** (Global and European Digital Elevation Model) is another critical dataset for ER related analysis and modelling. The Copernicus DEM is provided at three different levels:

- A 10 m resolution European extent DEM

- A 30 m resolution Global extent DEM

- A 90 m resolution Global extent DEM

3.2 ALGORITHMS AND LIBRARIES

3.2.1 Compositing and Time Series Harmonization

A variety of proprietary/commercial and open-source compositing methods exist to create seasonal or annual image composites of Sentinel-2 or Landsat data. A few notable examples are:

- GEE-BAP[9] – in 2021, the Natural Resources Canada BAP algorithm was implemented in GEE. This GEE application enables the generation of annual BAP image composites for large areas combining multiple Landsat sensors and images.

- Sentinel Hub S2cloudless – Sentinel Hub's s2cloudless is a Python package that can automate cloud detection in Sentinel-2 imagery. It uses a LightGMB machine learning model to identify clouds in Sentinel-2 imagery with great accuracy.

- GEE compositing – Compositing in GEE is generically done using reducers which can be applied to a temporal image stack. Earth Engine can pick the median, maximum or minimum value of the stack for each band.

- LT-GEE buildSRcollection – This Landtrendr GEE function builds an annual cloud and cloud shadow masked yearly medoid composites of Landsat surface reflectance imagery.

- Batch pixel-based composition algorithm – A batch pixel-based compositing algorithm for Landsat was developed and published on the GEE platform (Li et al. 2022). This algorithm uses all valid pixels from a reference image for the main portion of the composite and then substitutes the missing data with batches of pixels using a priority coefficient model (Li et al. 2022). This algorithm requires fewer observations, has less bias, and lowers pixel dispersion and improves the future analysis of Landsat images, although it can also be used for the compositing of other optical satellite imagery such as Sentinel-2 (Li et al. 2022).

- SEN2COR – The Sentinel 2 Correction algorithm is used by ESA in Sentinel-2 surface reflectance products (Tarrio et al. 2020). The algorithm uses a series of spectral thresholds, ratios, and indices with the goal of computing cloud probabilities for each pixel (Richter et al. 2012). To identify cirrus clouds this algorithm relies on two thresholds applied to the cirrus band (Tarrio et al. 2020). Cloud shadows are predicted from two probability layers; one geometric layer derived from the sun position, zenith angle, sun elevation, and cloud height; the other layer is derived from a neural network dark area classification.

- FMASK – The function of mask algorithm was developed as a cloud and cloud shadow detection algorithm for Landsat (Zhu and Woodcock 2012). FMASK version 4.0 has integrated Global Surface Water Occurrence (GSWO) and a Digital Elevation Model (DEM) to further improve detection of cloud and shadows for Sentinel-2 data (Qiu et al. 2023).

- MAJA – MACCS-ATCOR Joint Algorithm is a spectro-temporal method of cloud detection and atmospheric correction, meant to be used with both Landsat and Sentinel-2 imagery (Hagolle et al. 2010). It is. It used by the French Theia Land Data Center to produce Sentinel-2 surface reflectance products (Tarrio et al. 2020). This algorithm classifies low cloud pixels based on their spectral differences relative to a reference composite image that contains the most recent

---

[9] GitHub – saveriofrancini/bap: Best Available Pixel calculation using Google Earth Engine

BAP. High clouds are detected using a cirrus band threshold that varies with altitude (Hagolle et al. 2010; Baetens et al. 2019; Tarrio et al. 2020).

▪ Tmask – Tmask is a multi-temporal mask (Zhu and Woodcock 2014). It detects cloud and cloud-shadow by using Fmask applied to all available images for cloud screening, followed by a thresholding in the green band to detect previously misidentified clouds. The expected surface reflectance is then estimated for the remaining clear-sky pixels, using a Robust Iteratively Reweighted Least Squares (RIRLS) model. Pixels that deviate from the predicted surface reflectance are then flagged as either clouds or cloud shadows using thresholding rules.

▪ LaSRC – Land Surface Reflectance Code is a surface reflectance algorithm for Landsat-8 (Vermote et al. 2016; Skakun et al. 2019). It has a cloud mask in its quality assurance layer, that is produced during the atmospheric correction process. This code uses an implementation of Fmask in the C programming language, to identify clouds before estimating reflectance. Clouds are classified by calculating and aerosol optical thickness residual using ratios of Landsat-8's blue, red and SWIR bands (Tarrio et al. 2020).

▪ s2cloudless – Sentinel Hub uses s2cloudless for generation of cloud masks. It relies on machine learning techniques and is a single-scene cloud detection algorithm (Zupanc 2017). This method assigns a probability of being cloud to each pixel based on the pixel's spectral response. It then uses convolution to consider adjacent pixels when constructing a cloud mask of a given scene.

When s2cloudless was compared with FMASK and Sen2Cor (both single-scene algorithms), using a hand labeled dataset as validation (Hollstein et al. 2016), it showed better performance. It had better cloud detection rates and lower misclassification rates. The MAJA algorithm was also compared to s2cloudless on a small dataset and had a lower cloud and cirrus detection rate. However, it did have a lower misclassification rate for land classes as cloud (Zupanc 2017).

USGS HLS comprises three types of products: "S10" products – atmospherically corrected Sentinel-2 images in their native resolution and geometry; and the harmonized products "HLSS30" and "HLSL30." These products have been radiometrically harmonized to the maximum extent and then gridded to a common 30 m UTM basis using the Sentinel-2 tile system. The S30 and L30 products are resampled as needed to a common 30-meter resolution UTM projection and tiled using the Sentinel-2 Military Grid Reference System (MGRS) UTM grid. The data generated are available through NASA Earthdata Search but note that **S10 products are not normally archived**. An assessment of spectral correspondence between S30 and L30 found high agreement (r= 0.87–0.96) for spectral channels and an r = 0.99 for NBR with low relative root-mean-square difference values (1.7%–3.3%) (Wulder 2016; Wulder et al. 2021).

The Sen2Like framework is a scientific and open-source software[10]. Sen2Like ingests a stack of L1/L2 S2/LS8 and LS9 products and generates L2H and L2F product types; in the L2H products, the native spatial resolutions of input images are preserved, whilst in the L2F products, the resolutions of the LS8/LS9 image data are upsampled to the pixel spacing of the relevant S2 band. The validation of the ARD shows that the quality of input data is preserved and the inter-calibration between different sensors is improved (Saunier et al. 2022).

---

[10] https://github.com/senbox-org/sen2like

3.2.2 Spectral Recovery / Trends

There are several algorithms available and hosted in GEE and a few open-source Python modules or software packages that are relevant to vegetation recovery and trend analysis.

CCDC/COLD is a temporal segmentation algorithm that uses harmonic modelling to distinguish intra-annual change in vegetation due to seasonal phenology from gradual trends or abrupt changes (Zhu and Woodcock 2014; Zhu et al. 2020). A series of consecutive values that diverge from the harmonic is identified as an abrupt disturbance, which would not be expected by the intra- and inter-annual change modelled in the harmonic. The harmonic is generated during an initialization period, where a minimum of 12 clear pixel values (no cloud or snow contaminated observations) occurring over the period of at least one year are used to fit eight model coefficients using Least Absolute Shrinkage and Selection Operator (LASSO). The CCDC/COLD algorithm predicts reflectance values for a given band and identifies when a series of new data points diverge from this prediction, triggering a temporal break (Zhu et al., 2020). The predicted reflectance values account for both intra-annual and inter-annual changes, providing a decomposition between seasonal and long-term changes in the reflectance of a band or index. The best-known implementation of CCDC/COLD is within GEE. **pyCCD** is a Python implementation of CCDC developed by USGS – The CCD component of the CCDC algorithm is converted to create the Python-based CCD (PyCCD) library. This library was created during the implementation of CCDC to produce the LCMAP collection1.0 (Xian et al. 2022). This library is a per-pixel algorithm that outputs the spectral change segments of the time-series input data.

LandTrendr was originally written in Interactive Data Language (IDL), which limited accessibility and scalability of the application. In 2018, LandTrendr was implemented in GEE (Kennedy et al. 2018). The GEE platform simplifies pre-processing steps, allowing focus on the translation of the core temporal segmentation algorithm.

The Python module **statsmodels** provides classes and functions for many statistical models. Among them are two functions that can be used for seasonal trend decomposition of time-series data: *seasonal_decompose* (a naïve decomposition method) and *Season-Trend Decomposition using LOESS* (STL) (Cleveland et al. 1990). The STL function automatically detects the data's frequency and returns four plots:

- A curve plotting the observed data

- A curve depicting the data's trend once the seasonality has been removed

- A curve depicting the seasonality itself

- And the residuals from the model

Tools such as these allow us to visualize our data before and after the seasonality has been removed. This in turn gives a useful view of the importance seasonality may or may not play in the trend underlying the data. This kind of insight can be of value when seeking to monitor or plan restoration initiatives. HAPO has been made open-source and published in a GIT repository[11].

---

[11] https://code.usgs.gov/lcmap/research/Harmonic-Adaptive-Penalty-Operator-HAPO-/-/tree/main

3.2.3 Wetness / Water Variability and Dynamics

Data and software availability situation is complicated. For example, the JRC global surface water permanence data is publicly available, but the software code is proprietary. ESA WorldWater project is supposed to make the software available to potential users at the time of project completion. Several general-purpose semantic segmentation deep learning models that can be used for delineating open water bodies are publicly available.

Regarding surface soil moisture mapping, "Soil MOisture retrieval from multi-temporal SAR data" (SMOSAR) code (v2.0) is available. Some code can be provided by authors, e.g., MULESME, but requites commercially owned software as its components (Pulvirenti et al. 2018).

3.2.4 Canopy Density / Cover

There are multiple canopy density/cover datasets openly available such as the global Tree Canopy Cover (TCC) dataset produced by NASA's Land-Cover and Land-Use Change Program. However, the source code for many of the datasets isn't available.

GEDI L2B Canopy Cover Vertical Profile – The Canopy Cover Vertical Profile is a dataset derived from the GEDI instrument with an average footprint of 25 meters. This product is based on the directional gap probability derived from the L1B waveform (Dubayah et al. 2020).

SentinelHub – The Pseudo Forest Canopy Density (Pseudo-FCD) index is an index derived by calculating an Advanced vegetation index (AVI), Bare soil index (BSI), Canopy shadow index (SI), a Thermal index and subsequently combining them to create an indicator of Forest Canopy Density. This indicator does on occasion struggle with water bodies in the landscape. The source code has been shared as a Sentinel-Hub custom script[12].

3.2.5 Canopy Height

There are several publicly open datasets, often accompanied by ATBDs or published papers, however, code availability is limited.

F-TEP services are proprietary, possibly some of them can be made open, e.g., kNN method (Antropov et al. 2017), and several planned methodologies that will be an outcome of ESA Forest Carbon Monitoring (Antropov et al. 2022) and ESA RepreSent projects. Within ESA RepreSent, several semi-supervised and semi-supervised deep learning methodologies will be made openly available during late spring 2023.

Additionally, several general-purpose semantic segmentation or regression deep learning models can be used for producing forest variables such as forest tree height, but need representative training data, or are not yet validated.

3.3 INFORMATION TECHNOLOGY / PLATFORMS

As the volume and velocity of satellite earth observation datasets continues to grow, the technology required to manage and process EO datasets also needs to keep pace. Further, advanced EO data

---

[12] https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-2/pseudo_forest_canopy_density/

analysis techniques exceeds the capacity and capability of personal computers requiring new technologies for EO data analysis.

To efficiently process large EO datasets, information technology solutions are required that minimize data movement and maximize processing capabilities. These solutions are characterized as "EO data exploitation platforms" as these computational platforms seek to exploit the value stored in EO datasets for various use cases.

3.3.1 European Approach to EO Data Exploitation Platforms

In Europe, most EO data exploitation platforms are funded, and to a great extent directed, by the public sector, but with substantial input and participation by the private and academic sectors. There is a great deal of emphasis put on building capacity in the private and academic sectors, and frequently projects will be duplicated across several initiatives to foster this. This public-private business model for EO data resource development has been through innovation and industry stimulus projects that typically bring together pan-European consortiums that are inclusive of academic institutions and small and medium enterprises (SMEs). Many current European EO data exploitation platforms can trace their origins to European Commission (EC) research and innovation framework programs (FP), such as Rasdaman[13] which received €1.5M in EC FP4 funding[14] between 1995 and 1998. Other European EO data exploitation platforms can trace their origin to commercialization activities conducted by ESA, such as Terradue[15] which traces its origins to the commercialization of ESA's Grid Processing On Demand[16] (G-POD) system in 2006[17]. National governments in Europe have also sponsored the creation of EO data exploitation platforms.

Thematic Exploitation Platforms

In 2012 ESA started to develop a variety of EO Exploitation Platforms, starting with the "Super Site Exploitation Platform" (SSEP) and the "Exploitation Platform for Soil Moisture". In 2013 ESA launched a program to develop a suite of Thematic Exploitation Platforms (TEPs) which are thematically targeted EO data exploitation platforms that provide "collocation of data, processing capabilities, and ICT infrastructure, thus providing a complete work-environment for users performing scientific exploitation of EO data"[18]. ESA supported the development of seven TEPs targeting different user communities: Coastal, Forestry, Hydrology, Geohazards, Polar, Urban and Food Security. TEPs are just one type of EO exploitation platform supported by ESA; Regional Exploitation Platforms (REPs) and Mission Exploitation Platforms (MEPs) also exist to serve targeted geographic regions and specific EO missions respectively.

The Forestry Thematic Exploitation Platform (Forestry TEP) is operated by VTT. Target user groups for the platform include value-adding service industry, research and academia, large forest owners, forest industry and non-governmental organizations (NGOs). Key user needs include up-to-date information on forest resources and forest carbon balance, facilitating responsible forest management and the

---

[13] https://rasdaman.org/

[14] https://cordis.europa.eu/project/id/20073/results

[15] https://www.terradue.com/

[16] https://gpod.eo.esa.int/

[17] https://gsaw.org/wp-content/uploads/2014/10/2010s11d_brito.pdf

[18] "Implementation of Thematic Exploitation Platforms" Request for Information, ESA, Sept 2013.

carbon trade. Value-adding service providers and researchers can benefit from the development features and the API interfaces of the platform to develop and publish remote sensing and forestry insight services, enabling to reach new audiences and serve customers globally.

Forestry TEP – also known as F-TEP – offers access to satellite data, a collection of value adding processing services and popular tools. The platform features and processing services can also be accessed from external systems via a REST API interface. Importantly, the users can create their own processing services using the development environment offered on the platform. The created services can optionally be made available to colleagues and partners or all platform users either freely or on a licensing basis. Recently, Forestry TEP has been successfully exploited in research projects that are building new processing tools and chains on the platform. These projects include Forest Digital Twin Earth Precursor (ESA)[19], Forest Flux (EU Horizon 2020)[20], and Forest Carbon Monitoring (ESA)[21].

To integrate its many Exploitation Platforms, ESA is promoting the EO Exploitation Platform Common Architecture (EOEPCA)[22]. This architecture envisions these platforms as a Network of Resources[23] that integrates the various exploitation platforms using standardized APIs. In particular, the EOEPCA presents an implementation of the recently published OGC Best Practice for Earth Observation Application Package by the Open Geospatial Consortium (OGC), facilitating distribution of EO applications to various exploitation platforms. In EOEPCA, a key operational element is the Application Deployment & Execution Service (ADES).

The Forestry TEP is currently undergoing major architectural revision to implement the EOEPCA in a new Kubernetes based environment. A new processing service concept and developer tooling will be introduced, with Docker remaining as the container technology for the services. Release is targeted for early 2023, with a plan to maintain the current offering in parallel until all features have been implemented. Support and guidance will be provided to users in the transition.

Copernicus DIAS

In 2016 ESA announced the "EO Innovation Platform Testbed Poland" (IPT-Poland) initiative to validate data distribution to a cloud-based infrastructure to exploit data from the Copernicus Sentinel satellites. The successful operation of the IPT-Poland resulted in a 2017 tender for the creation of Copernicus Data and Information Access Services (DIAS). A DIAS was envisioned as a "cloud-based one-stop shop for all Copernicus satellite data and imagery as well as information from the six Copernicus services, that also give access to sophisticated processing tools and resources"[24]. The DIAS tender resulted in five contracts, each valued at €10-15M over 4 years (2018-2021), to consortiums required by contract to include SMEs. In the second cycle of Copernicus (2021-2027), only two of the DIAS were selected to continue: CREODIAS and WEkEO.

CREODIAS is a cloud computing system implemented by CloudFerro that provides[25]:

---

[19] foresttwin.org

[20] forestflux.eu

[21] forestcarbonplatform.org

[22] https://eoepca.github.io/

[23] https://eo4society.esa.int/2019/06/07/network-of-resources/

[24] https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Accessing_Copernicus_data_made_easier

[25] https://creodias.eu/what-is-creodias

- Scalable on-demand OpenStack cloud computing platform for data processing;

- Over 34 PB of Earth Observation data (Copernicus Sentinels, Landsat, Envisat and others); and

- Access to array of Platform as a Service applications.

In the second cycle of Copernicus (2021-2027), CREODIAS together with T-Systems, Sinergise, VITO, DLR, ACRI-ST and RHEA will operate a revised "Copernicus Data Access Service"[26]. In this service, T-Systems and CloudFerro will provide the cloud infrastructure, Sinergise and VITO will integrate Sentinel Hub and OpenEO data discovery and processing tools, while DLR, ACRI-ST and RHEA will provide on-demand processing and access to Copernicus Contributing Missions. The first version of this service is expected to be fully operational in June 2023.

WEkEO is implemented by ECMWF, EUMETSAT, EEA, and Mercator Ocean international. WEkEO implements a federated architecture that brings together computing and storage systems from CREODIAS, EUMETSAT, ECMWF and Mercator Ocean into a single platform. For data processing, dedicated virtual machines of various sizes (2-128 CPUs and 16-4 TB RAM) are available. In 2021 the EUMETSAT Council approved a four-party arrangement between EUMETSAT, the European Centre for Medium-Range Weather Forecasts (ECMWF), Mercator Ocean International and the European Environment Agency to continue to cooperate on WEkEO from 2022 to 2027[27]. Planned improvements to WEkEO under the second cycle of Copernicus (2021-2027) have not been released publicly.

Euro Data Cube

Euro Data Cube[28] (EDC) is a platform integrating EO data access systems, tools for access and analysis, and a marketplace for solutions. EDC is funded by ESA and has its roots in the 2017-2020 DataCube Services for Copernicus (DCS4COP) project funded under Horizon 2020 innovation funding[29]. Data available through the EDC includes all open missions (e.g., Sentinel, Landsat, MODIS, etc.), commercial satellites (PlanetScope, Pleiades, SPOT, WorldView, etc.) as well as Level 3 products (Copernicus Land Monitoring Services, C3S, etc.). Access and analysis tools include cloud workspaces, access to the Sentinel Hub cloud API, batch processing systems and the creation of arbitrary data cubes through the xcube open source Python library.

The Polar Thematic Exploitation Platform (P-TEP) has recently transitioned[30] to use the EDC infrastructure and services to modernize its' offerings and to reduce ongoing development costs.

3.3.2   United States Approach to EO Data Exploitation Platforms

The United States has long been the world leader in EO data exploitation systems due to its leadership in developing space-based systems. In general, the United States' approach to developing EO data exploitation platforms is to increasingly develop these resources on large commercial public cloud

---

[26] https://medium.com/sentinel-hub/new-copernicus-data-access-service-to-support-the-ecosystem-for-earth-observation-412f829355a3

[27] Page 45 of https://www-cdn.eumetsat.int/files/2022-07/EUMETSAT%20Annual%20Report%202021.pdf

[28] https://eurodatacube.com/

[29] https://cordis.europa.eu/project/id/776342 and http://www.dcs4cop.eu/

[30] https://eo4society.esa.int/2021/12/14/polar-tep-evolution-benefits-from-euro-data-cube/

operators such as AWS, Google Cloud and Microsoft Azure. Cloud migration projects by NASA, USGS and NOAA are indicative of this approach.

To manage the wealth of EO data being generated, NASA operates a network of 12 Distributed Active Archive Centers (DAACs) to host and make available EO data. These DAACs have annual budgets of over US$79 million[31] and host data and computation systems in data centers owned and operated by NASA. In 2016 NASA began a system evolution project called "Earthdata Cloud". The goal of the Earthdata Cloud project is to "develop, test, and deploy commercial cloud environments to realize storage, processing, and operations efficiencies; improve cross-DAAC collaboration; and provide new data access and services."[32] In 2018, NASA entered into a 5-year, $65 million task order with AWS to support the Agency's evolution to the public cloud. As part of this move to the public cloud, the Global Hydrology Resource Center DAAC (GHRC DAAC) was selected as the first DAAC to migrate to AWS and is now operating (as of March 2020) in the cloud in parallel with on-premises systems. While NASA primarily uses AWS, it also has partnerships with Google and Microsoft to make NASA science datasets available on the Google Cloud, GEE and Azure.

NASA has sponsored the development of a number of open source software systems to assist with EO data ingestion, management and analysis. Two notable NASA sponsored projects are Project Cumulus and Pangeo:

- Project Cumulus[33] uses AWS specific technologies to provide: data acquisition from data providers; data ingest; harvest, creation, and publication of dataset metadata to the Common Metadata Repository; storage and distribution of data, including disaster recovery; and publication of metrics to a federated Metrics System. Using Project Cumulus the GHRC DAAC is now operating cloud-native functions on AWS.

- Pangeo[34] is a project that is developing and supporting a suite of interconnected software packages that enable scalable geoscience data analytics. Pangeo maintains an infrastructure neutral approach (it supports most commercial cloud vendors along with high performance computing [HPC] systems) and aims to develop a cloud-native open architecture that can conduct distributed processing across HPC and cloud systems. Although primarily funded and developed by US institutions, the Pangeo project has several international collaborations including with CNES and the UK Met Office's Informatics Lab.

In 2015, the US National Oceanic and Atmospheric Administration (NOAA) initiated a project called the "Big Data Project". The goal of this project is to provide public access to NOAA's open data on commercial cloud platforms through public-private partnerships. Initially this program partnered with five commercial entities (AWS, Google Cloud, IBM, Microsoft Azure, and the Open Common Consortium)[35], and was modified in late 2019 to be with 3 entities (AWS, Google Cloud, and Microsoft Azure)[36]. Through this program, the public can access NOAA data through the selected cloud providers for free while only paying for the compute and private storage resources.

---

[31] Page 10 of https://www.oversight.gov/sites/default/files/oig-reports/IG-20-011.pdf

[32] https://www.oversight.gov/sites/default/files/oig-reports/IG-20-011.pdf

[33] https://github.com/nasa/cumulus/ and https://nasa.github.io/cumulus/

[34] https://pangeo.io

[35] https://www.noaa.gov/organization/information-technology/evolution-of-big-data-program

[36] https://www.noaa.gov/media-release/cloud-platforms-unleash-full-potential-of-noaa-s-environmental-data

In 2017 the US Geological Survey (USGS) began the process of transitioning Landsat data storage and processing to take advantage of cloud architectures[37]. Using the cloud the USGS seeks to change the way that Landsat data is used – transitioning away from a download model to a model that uses the full Landsat archive. To minimize risks, the USGS will "be implementing Landsat's cloud architecture using a hybrid approach that includes the USGS Cloud Hosting Solutions (CHS) program and existing Earth Resources Observation and Science (EROS) Center private cloud capabilities, including network, compute, and storage resources"[38].

3.3.3   Australian Approach to EO Data Exploitation Systems

In Australia, the development of EO data exploitation platforms are tied into the ongoing investment and development of Digital Earth Australia (DEA). DEA is an Australian government investment focused on increasing the utilization of Earth Observation data across Australia and is one of seven priorities stated in the Australian Space Agencies' 2019-2028 Space Strategy[39]. It is a program covering access to data, along with outreach and education activities.

Using high performance computing power provided by the Australian National Computational Infrastructure and commercial cloud computing platforms (a hybrid-cloud approach), DEA organizes and prepares EO data into stacks of consistent, time-stamped observations that can be quickly manipulated and analyzed to provide information about a range of environmental factors such as water availability, crop health and ground cover.

A key goal of DEA is to enable the Australian spatial industry to exploit the full value of EO information to enhance their business and be competitive in global markets[40]. DEA is expected to enable Australian small businesses and industry to more readily access near real time satellite data and derived information to innovate and create new products. DEA regularly publishes[41] updates to its program's roadmap allowing projects, product improvements and user engagement activities to be easily monitored.

The technology underpinning DEA is the open source Open Data Cube[42] system. Open Data Cube originally was developed purely by GeoScience Australia but is now supported by five other institutional partners: NASA/Committee on Earth Observation Satellite (CEOS), United States Geological Survey (USGS), Commonwealth Scientific and Industrial Research Organisation (CSIRO), Catapult Satellite Applications, and Analytical Mechanics Associates (AMA). Initially, Open Data Cube could only be used on the Australian National Computational Infrastructure HPC system, but recent advances have allowed the system to be easily installed on AWS using the "Cube in a box" system[43].

---

[37] http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-49/1.%20Tuesday%20April%2021/20200421T1150_Data%20Interoperability-Landsat%20in%20Cloud.pdf

[38] https://www.usgs.gov/news/landsat-data-moving-public-cloud-early-2020

[39] https://publications.industry.gov.au/publications/advancing-space-australian-civil-space-strategy-2019-2028.pdf

[40] https://www.ga.gov.au/__data/assets/pdf_file/0006/94056/DEA-Program-Roadmap-May-2020.pdf

[41] https://www.dea.ga.gov.au/news/

[42] http://www.opendatacube.org/

[43] https://www.opendatacube.org/ciab . The Cube in a Box system notably provides a "magic link" to launch an Open Data Cube instance on your own AWS resources with a single click: https://github.com/opendatacube/cube-in-a-box#magic-link

DEA stores a range of data products on Amazon Web Service's S3 with free public access[44]. Tools available to end-users of DEA include a Jupyter-based sandbox, STAC metadata services, a web-map explorer interface, and OGC data access services.

3.3.4 Public Cloud Provider EO Platforms

Public cloud providers have seen the opportunity to use their cloud computing platforms for EO data storage and analysis. Systems created by the major public cloud providers include the Planetary Computer from Microsoft, GEE, and Earth on AWS.

Microsoft Planetary Computer is a platform created by Microsoft that brings together an extensive EO data archive together with Jupyter lab interfaces linked to scalable EO data analytics using the Dask python library. The Planetary Computer provides direct access in object storage to Sentinel-1 GRD and RTC processed data, Sentinel-2, MODIS, Landsat 5 to 8, and Planet data (NICFI), among numerous other datasets. Microsoft has embraced the ideas and techniques pioneered by Pangeo, which includes the use of STAC for cataloging and querying data, storing analysis-ready data in blob storage, and providing Jupyter-based computation and analysis environments linked to Dask clusters for scalable data analysis close to the data. While currently in a public preview that provides free access to all users (commercial and non-commercial), it is expected that the Planetary Computer will eventually require payment to access computational resources. All systems used to run the Planetary Computer are open source[45].

Google Earth Engine (GEE) is the most established EO cloud platform developed by Google. Users interact with the platform via the Earth Engine Javascript Code Editor or Javascript and Python API using the Earth Engine data catalog. GEE is evolving from a non-commercial platform mainly used by research scientists to a commercial subscription platform. GEE is a proprietary platform that is not open source resulting in users being locked-into only using GEE.

AWS has developed **Earth on AWS**, a registry of open geospatial data that supports developers to build applications using AWS cloud computing infrastructure. The Earth on AWS program provides free hosting (on AWS S3 object storage buckets) of satellite EO datasets which can then be exploited by implementing cloud-computing systems next to these existing data stores. The implementation of these computational systems is not provided by AWS and must be implemented from scratch by the user.

3.3.4.1 Related Public Cloud EO Data Exploitation Platforms

Several notable EO Data Exploitation Platforms have been created based on the above noted public cloud platforms. These include SEPAL and the Sentinel Hub.

SEPAL is a free, open-source cloud-computing platform created and operated by the Forestry Department of the United Nations Food and Agriculture Organization (FAO). The SEPAL platform provides users access to geospatial data resources (either via the GEE data catalogue or user-provided data) and resources to process these data to produce useful information. SEPAL currently "resides" in the Amazon Web Services (AWS) ecosystem and makes use of AWS cloud instances that can be used to process data using pre-programmed SEPAL applications or your own code (in R, Python or in the

---

[44] https://docs.dea.ga.gov.au/setup/AWS/data_and_metadata.html

[45] https://github.com/microsoft/PlanetaryComputer

terminal). SEPAL is also closely linked to GEE and largely relies on GEE for scalable data processing[46]. Due to its links with FAO and the OpenForis suite of tools, many government agencies[47] around the world actively use SEPAL.

Sentinel Hub is a flagship European example of a cloud platform built on AWS, leveraging Earth on AWS datasets. Created by Sinergise (of Ljubljana, Slovenia) the Sentinel Hub provides a cloud-based data API that removes the complexity of processing large volumes of satellite data. The Sentinel Hub website also provides two web applications, a "Sentinel Playground" and an "EO Browser" both of which are built on top of their EO data API. The Sentinel Hub API implements the WMS, WCS and WMTS OGC protocols, but prefers that its' proprietary APIs be used by most users. Data processing commands sent to the Sentinel Hub API must use custom "eval scripts" which use the javascript language but use structures only available to the Sentinel Hub. The Sentinel Hub API forms key parts of other platforms, such as the EDC and can be found on most DIAS, most notably CREODIAS.

3.3.5   Standardization of EO Analytic Applications

Recent initiatives aim to allow EO application developers to create their application in such a way that it is not tied to a singular EO exploitation platform, but instead can be distributed to a variety of platforms. By implementing the algorithms or applications according to an open standard, the hope is that the applications will not be locked into a single EO exploitation platform but instead be portable to a variety of competing platforms.

Beyond portability, the use of open standards also allows for a future of distributed processing, where a processing workflow can constitute piece solutions on multiple platforms, working in sequence (or even in parallel in complex workflows) to achieve the end result.

OGC Earth Observation Application Package

The Open Geospatial Consortium (OGC) published in 2021 the Best Practice for Earth Observation Application Package (OGC 2021)[48]. This document describes the approach taken by ESA's Earth Observation Exploitation Platform Common Architecture (EOEPCA) project to package EO applications and deploy them to remote platforms. Inside the EOEPCA project, application packages are handled by the Execution Management Service (EMS) and Application Deployment & Execution Service (ADES). Technically, EO application packages rely on the Common Workflow Language (CWL) and the SpatioTemporal Asset Catalog (STAC) specifications, docker containers and the OGC API – Processes standard.

The use of the EO Application Package standard was tested in the OGC's Earth Observation Applications Pilot which was conducted between December 2019 and September 2020. Results of this pilot[49] showed the EO Application Package system to be functional, although further work was needed to make systems operational. User uptake of the EO Application Package standard appears to be limited, potentially due to the complexity of defining EO application packages.

---

[46] https://docs.sepal.io/en/latest/setup/gee.html

[47] For a partial list of OpenForis users, see https://openforis.org/collaborators/

[48] https://docs.ogc.org/bp/20-089r1.html

[49] Full engineering reports from the pilot study are available online: https://www.ogc.org/projects/initiatives/eoa-pilot

VTT is developing an EO Application Package compliant execution environment on the Forestry TEP. This service, along with other EOEPCA capabilities, is expected to be operational in 2023.

OpenEO

The genesis of the OpenEO project started in 2016 with the idea that a common standard interface could be defined to allow EO analysis programs to run across a wide variety of cloud computing systems. Initially funded by a Horizon2020 grant, the OpenEO project has designed a standard server API which is interacted with by R, Python or Javascript libraries. By implementing the server API on multiple cloud computing systems (including systems like CREODIAS, EODC, GEE and VITO's Terrascope), cloud system independent analysis has been showcased.

User uptake of the OpenEO standard appears to be growing, although is potentially limited by the need for users to learn a new way of accessing and processing EO data. Use of OpenEO will be increasing due to its inclusion in the Copernicus Data Access Service.

3.3.6   Scalable/Parallel Processing

Just as the volume of satellite data continues to grow, so does the need to process data in a performant manner. To process a time series of satellite data over a large area, scalable computational systems are required.

In computational systems, there are two methods to scale resource: vertically or horizontally. Scaling vertically typically means adding more resources (CPU, RAM, storage, network) to a single machine. Scaling horizontally adds multiple machines that are then tasked to run analysis in parallel.

GEE and Rasdaman provide abstractions on top of scalable parallel processing systems such that analysts and users do not need to understand the complexities of parallel processing being undertaken. HPC systems use job submission/management systems such as SLURM, SGE, TORQUE, LSF, DRMAA and PBS to standardize the distribution processing tasks to parallel computing systems.

The Pangeo initiative advocates a combination of Kubernetes for container orchestration and horizontal scaling, Dask for distributed compute, and Jupyter notebooks to provide a user interface for performing analyses. Kubernetes is a popular open-source system for automating deployment, horizontal scaling, and management of containerized applications. Dask is a distributed system which can scale efficiently from a single computer to hundreds of servers using horizontal scaling. Dask uses regular Python code and Python APIs to scale the work for existing Python structures (like NumPy arrays and Pandas dataframes) concurrently. Microsoft has embraced the Pangeo initiative for its Planetary Computer platform.

Commercial cloud providers also provide specialized interfaces to horizontal scaling infrastructure, all of which can be used to analyze satellite datasets in a scalable manner, while hiding the complexity of parallel processing:

- Microsoft Azure has Azure Databricks, Batch, Azure Machine Learning, HDInsight, and Azure Analysis Services;

- Google Cloud has BigQuery, Batch, Dataflow and Dataproc; and

- AWS has Amazon Athena, Amazon EMR and Amazon SageMaker.

The Forestry TEP supports parallel processing through dynamic assignment of processing jobs to virtual machines (VM), with a single VM potentially sharing multiple jobs, depending on the configuration. Processing jobs are executed in parallel using the configured resource pool and queued if the full capacity of the pool is reached.

3.3.7 User Interfaces and Integrations

Industry and government usually have pre-existing applications for visualization and manipulation of geospatial datasets. Common applications include ArcGIS Online, ArcGIS Desktop, and QGIS. Web-based data visualization dashboards are often created to provide users with the interface to explore complex, large datasets, using tools such as with Tableau, ArcGIS web apps, Microsoft PowerBI and other specialized enterprise resource management systems. Many of these applications seek to provide a common operating picture by integrating data from multiple sources, visualizing datasets, and providing action tracking.

To integrate datasets from multiple data sources, application programming interfaces (APIs) are commonly used. Previously APIs used Extensible Markup Language (XML) formats to interchange data, while modern interfaces use JavaScript Object Notation (JSON). To structure the APIs, XML-based interchanges typically were structured using the Simple Object Access Protocol (SOAP) standard, while modern JSON-based interchanges typically use the Representational state transfer (REST) architecture. To achieve higher performance, modern data interchange systems implement the gRPC Remote Procedure Call (gRPC) framework which enables low latency, highly scalable, distributed systems.

For geospatial systems, the OGC sets standards for implementing geospatial data-interchange systems. In recent years OGC standards have been modernized moving from XML based interchanges to modern REST/JSON interchanges. For example, to send raster data between systems the OGC previously developed the Web Coverage Service (WCS) specification. Recently the OGC has worked to modernize this WCS protocol which is now called OGC API-Coverages.

4.0 REFERENCES

Adams R, Jeanrenaud S, Bessant J, Denyer D, Overy P. 2016. Sustainability-oriented Innovation: A Systematic Review: Sustainability-oriented Innovation. International Journal of Management Reviews. 18(2):180–205. doi:10.1111/ijmr.12068.

Antropov O, Miettinen J, Hame T, Yrjo R, Seitsonen L, McRoberts RE, Santoro M, Cartus O, Duran NM, Herold M, et al. 2022. Intercomparison of Earth Observation Data and Methods for Forest Mapping in the Context of Forest Carbon Monitoring. In: IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium. Kuala Lumpur, Malaysia: IEEE. p. 5777–5780. [accessed 2023 Jan 4]. https://ieeexplore.ieee.org/document/9884618/.

Antropov O, Rauste Y, Häme T, Praks J. 2017. Polarimetric ALOS PALSAR Time Series in Mapping Biomass of Boreal Forests. Remote Sensing. 9(10):999. doi:10.3390/rs9100999.

Antropov O, Rauste Y, Praks J, Seifert FM, Häme T. 2021. Mapping Forest Disturbance Due to Selective Logging in the Congo Basin with RADARSAT-2 Time Series. Remote Sensing. 13(4):740. doi:10.3390/rs13040740.

Astola H, Häme T, Sirro L, Molinier M, Kilpi J. 2019. Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region. Remote Sensing of Environment. 223:257–273.

Baetens L, Desjardins C, Hagolle O. 2019. Validation of Copernicus Sentinel-2 Cloud Masks Obtained from MAJA, Sen2Cor, and FMask Processors Using Reference Cloud Masks Generated with a Supervised Active Learning Procedure. Remote Sensing. 11(4):433. doi:10.3390/rs11040433.

Bauer-Marschallinger B, Freeman V, Cao S, Paulik C, Schaufler S, Stachl T, Modanesi S, Massari C, Ciabatta L, Brocca L, et al. 2019. Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and Overcoming Obstacles. IEEE Trans Geosci Remote Sensing. 57(1):520–539. doi:10.1109/TGRS.2018.2858004.

Betts AK, Chan DZ, Desjardins RL. 2019. Near-Surface Biases in ERA5 Over the Canadian Prairies. Frontiers in Environmental Science. 7. [accessed 2022 Dec 20]. https://www.frontiersin.org/articles/10.3389/fenvs.2019.00129.

Burn DH, Hag Elnur MA. 2002. Detection of hydrologic trends and variability. Journal of Hydrology. 255(1):107–122. doi:10.1016/S0022-1694(01)00514-5.

Cartus O, Santoro M, Wegmuller U, Labriere N, Chave J. 2022. Sentinel-1 Coherence for Mapping Above-Ground Biomass in Semiarid Forest Areas. IEEE Geosci Remote Sensing Lett. 19:1–5. doi:10.1109/LGRS.2021.3071949.

Chazdon RL. 2008. Beyond Deforestation: Restoring Forests and Ecosystem Services on Degraded Lands. Science. 320(5882):1458–1460. doi:10.1126/science.1155365.

Chhabra A, Rüdiger C, Yebra M, Jagdhuber T, Hilton J. 2022. RADAR-Vegetation Structural Perpendicular Index (R-VSPI) for the Quantification of Wildfire Impact and Post-Fire Vegetation Recovery. Remote Sensing. 14(13):3132. doi:10.3390/rs14133132.

Chia EK, Bassett M, Leonard SWJ, Holland GJ, Ritchie EG, Clarke MF, Bennett AF. 2016. Effects of the fire regime on mammal occurrence after wildfire: Site effects vs landscape context in fire-prone forests. Forest Ecology and Management. 363:130–139. doi:10.1016/j.foreco.2015.12.008.

Chiaravalloti F, Caloiero T, Coscarelli R. 2022. The Long-Term ERA5 Data Series for Trend Analysis of Rainfall in Italy. Hydrology. 9(2):18. doi:10.3390/hydrology9020018.

Cleveland R, Cleveland W, McRae J, Terpenning I. 1990. STL: A seasonal-trend decomposition. J Off Stat. 6(1):3–73.

Cocke AE, Fulé PZ, Crouse JE. 2005. Comparison of burn severity assessments using Differenced Normalized Burn Ratio and ground data. Int J Wildland Fire. 14(2):189. doi:10.1071/WF04010.

Coops NC, Shang C, Wulder MA, White JC, Hermosilla T. 2020. Change in forest condition: Characterizing non-stand replacing disturbances using time series satellite imagery. Forest Ecology and Management. 474:118370. doi:10.1016/j.foreco.2020.118370.

Denissen JMC, Teuling AJ, Pitman AJ, Koirala S, Migliavacca M, Li W, Reichstein M, Winkler AJ, Zhan C, Orth R. 2022. Widespread shift from ecosystem energy to water limitation with climate change. Nat Clim Chang. 12(7):677–684. doi:10.1038/s41558-022-01403-8.

Dubayah R, Blair JB, Goetz S, Fatoyinbo L, Hansen M, Healey S, Hofton M, Hurtt G, Kellner J, Luthcke S, et al. 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. Science of Remote Sensing. 1:100002. doi:10.1016/j.srs.2020.100002.

Dunn RJH, Azorin-Molina C, Menne MJ, Zeng Z, Casey NW, Shen C. 2022. Reduction in reversal of global stilling arising from correction to encoding of calm periods *. Environ Res Commun. 4(6):061003. doi:10.1088/2515-7620/ac770a.

European Commission. 2019. COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Stepping up EU Action to Protect and Restore the World's Forests COM/2019/352 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1565272554103&uri=CELEX:52019DC0352.

European Commission. 2020. COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS EU Biodiversity Strategy for 2030 Bringing nature back into our lives COM/2020/380 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1590574123338&uri=CELEX:52020DC0380.

European Commission. 2021. COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS New EU Forest Strategy for 2030 COM/2021/572 final. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021DC0572.

FAO, WRI. 2019. The Road to Restoration – A Guide to Identifying Priorities and Indicators for Monitoring Forest and Landscape Restoration.

Ford JD, Clark D, Pearce T, Berrang-Ford L, Copland L, Dawson J, New M, Harper SL. 2019. Changing access to ice, land and water in Arctic communities. Nat Clim Chang. 9(4):335–339. doi:10.1038/s41558-019-0435-7.

Gao F, Ma F, Wang J, Sun J, Yang E, Zhou H. 2018. Visual Saliency Modeling for River Detection in High-Resolution SAR Imagery. IEEE Access. 6:1000–1014. doi:10.1109/ACCESS.2017.2777444.

García M, Saatchi S, Ustin S, Balzter H. 2018. Modelling forest canopy height by integrating airborne LiDAR samples with satellite Radar and multispectral imagery. International Journal of Applied Earth Observation and Geoinformation. 66:159–173. doi:10.1016/j.jag.2017.11.017.

Ge S, Su W, Gu H, Rauste Y, Praks J, Antropov O. 2022. Improved LSTM Model for Boreal Forest Height Mapping Using Sentinel-1 Time Series. Remote Sensing. 14(21):5560. doi:10.3390/rs14215560.

Gleixner S, Demissie T, Diro GT. 2020. Did ERA5 Improve Temperature and Precipitation Reanalysis over East Africa? Atmosphere. 11(9):996. doi:10.3390/atmos11090996.

Hagolle O, Huc M, Pascual DV, Dedieu G. 2010. A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENμS, LANDSAT and SENTINEL-2 images. Remote Sensing of Environment. 114(8):1747–1755. doi:10.1016/j.rse.2010.03.002.

Häme T, Sirro L, Kilpi J, Seitsonen L, Andersson K, Melkas T. 2020. A Hierarchical Clustering Method for Land Cover Change Detection and Identification. Remote Sensing. 12(11):1751. doi:10.3390/rs12111751.

Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, et al. 2013. High-resolution global maps of 21st-century forest cover change. Science. 342(6160):850–853. doi:10.1126/science.1244693.

Hermosilla T, Wulder MA, White JC, Coops NC, Hobart GW. 2015. An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. Remote Sensing of Environment. 158:220–234. doi:10.1016/j.rse.2014.11.005.

Hermosilla T, Wulder MA, White JC, Coops NC, Pickell PD, Bolton DK. 2019. Impact of time on interpretations of forest fragmentation: Three-decades of fragmentation dynamics over Canada. Remote Sensing of Environment. 222:65–77. doi:10.1016/j.rse.2018.12.027.

Hoerl AE, Kennard RW. 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics. 12(1):55–67. doi:10.1080/00401706.1970.10488634.

Hollstein A, Segl K, Guanter L, Brell M, Enesco M. 2016. Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images. Remote Sensing. 8(8):666. doi:10.3390/rs8080666.

Hussain M, Mahmud I. 2019. pyMannKendall: a python package for non parametric Mann Kendall family of trend tests. Journal of Open Source Software. 4(39).

Jacobs JM, Hunsaker AG, Sullivan FB, Palace M, Burakowski EA, Herrick C, Cho E. 2021. Snow depth mapping with unpiloted aerial system lidar observations: a case study in Durham, New Hampshire, United States. The Cryosphere. 15(3):1485–1500. doi:10.5194/tc-15-1485-2021.

Joshi C, Leeuw JD, Skidmore AK, Duren IC van, van Oosten H. 2006. Remotely sensed estimation of forest canopy density: A comparison of the performance of four methods. International Journal of Applied Earth Observation and Geoinformation. 8(2):84–95. doi:10.1016/j.jag.2005.08.004.

Kennedy RE, Yang Z, Cohen WB. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. Remote Sensing of Environment. 114(12):2897–2910.

Kennedy RE, Yang Z, Gorelick N, Braaten J, Cavalcante L, Cohen WB, Healey S. 2018. Implementation of the LandTrendr Algorithm on Google Earth Engine. Remote Sensing. 10(5):691. doi:10.3390/rs10050691.

Kerr YH, Waldteufel P, Richaume P, Wigneron JP, Ferrazzoli P, Mahmoodi A, Al Bitar A, Cabot F, Gruhier C, Juglea SE, et al. 2012. The SMOS Soil Moisture Retrieval Algorithm. IEEE Trans Geosci Remote Sensing. 50(5):1384–1403. doi:10.1109/TGRS.2012.2184548.

Laestadius L, Buckingham K, Maginnis S, Saint-Laurent C. 2015. Before Bonn and beyond: the history and future of forest landscape restoration. Unasylva. 66(245):11–18.

Lamarche C, Santoro M, Bontemps S, d'Andrimont R, Radoux J, Giustarini L, Brockmann C, Wevers J, Defourny P, Arino O. 2017. Compilation and Validation of SAR and Optical Data Products for a Complete and Global Map of Inland/Ocean Water Tailored to the Climate Modeling Community. Remote Sensing. 9(1):36. doi:10.3390/rs9010036.

Lavers DA, Simmons A, Vamborg F, Rodwell MJ. 2022. An evaluation of ERA5 precipitation for climate monitoring. Quart J Royal Meteoro Soc. 148(748):3152–3165. doi:10.1002/qj.4351.

Lazos-Chavero E, Zinda J, Bennett-Curry A, Balvanera P, Bloomfield G, Lindell C, Negra C. 2016. Stakeholders and tropical reforestation: challenges, trade-offs, and strategies in dynamic environments. Biotropica. 48(6):900–914. doi:10.1111/btp.12391.

Li W, Niu Z, Shang R, Qin Y, Wang L, Chen H. 2020. High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data. International Journal of Applied Earth Observation and Geoinformation. 92:102163. doi:10.1016/j.jag.2020.102163.

Li Y, Duan Y, Kuang Z, Chen Y, Zhang W, Li X. 2022. Uncertainty Estimation via Response Scaling for Pseudo-Mask Noise Mitigation in Weakly-Supervised Semantic Segmentation. AAAI. 36(2):1447–1455. doi:10.1609/aaai.v36i2.20034.

Liu M, Liu X, Wu L, Tang Y, Li Y, Zhang Y, Ye L, Zhang B. 2021. Establishing forest resilience indicators in the hilly red soil region of southern China from vegetation greenness and landscape metrics using dense Landsat time series. Ecological Indicators. 121:106985. doi:10.1016/j.ecolind.2020.106985.

Lizundia-Loiola J, Franquesa M, Khairoun A, Chuvieco E. 2022. Global burned area mapping from Sentinel-3 Synergy and VIIRS active fires. Remote Sensing of Environment. 282:113298. doi:10.1016/j.rse.2022.113298.

Loi DT, Chou T-Y, Fang Y-M. 2017. Integration of GIS and Remote Sensing for Evaluating Forest Canopy Density Index in Thai Nguyen Province, Vietnam. IJESD. 8(8):539–542. doi:10.18178/ijesd.2017.8.8.1012.

Luhunga PM, Songoro AE. 2020. Analysis of Climate Change and Extreme Climatic Events in the Lake Victoria Region of Tanzania. Frontiers in Climate. 2. [accessed 2022 Dec 20]. https://www.frontiersin.org/articles/10.3389/fclim.2020.559584.

Menenti M, Azzali S, Verhoef W, van Swol R. 1993. Mapping agroecological zones and time lag in vegetation growth by means of fourier analysis of time series of NDVI images. Advances in Space Research. 13(5):233–237. doi:10.1016/0273-1177(93)90550-U.

Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG. 2012. An Overview of the Global Historical Climatology Network-Daily Database. Journal of Atmospheric and Oceanic Technology. 29(7):897–910. doi:10.1175/JTECH-D-11-00103.1.

Minchella A, Del Frate F, Capogna F, Anselmi S, Manes F. 2009. Use of multitemporal SAR data for monitoring vegetation recovery of Mediterranean burned areas. Remote Sensing of Environment. 113(3):588–597. doi:10.1016/j.rse.2008.11.004.

Mistry MN, Schneider R, Masselot P, Royé D, Armstrong B, Kyselý J, Orru H, Sera F, Tong S, Lavigne É, et al. 2022. Comparison of weather station and climate reanalysis data for modelling temperature-related mortality. Sci Rep. 12(1):5178. doi:10.1038/s41598-022-09049-4.

Mottus M, Molinier M, Halme E, Cu H, Laaksonen J. 2021. Patch Size Selection for Analysis of Sub-Meter Resolution Hyperspectral Imagery of Forests. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Brussels, Belgium: IEEE. p. 2035–2038. [accessed 2023 Jan 17]. https://ieeexplore.ieee.org/document/9554257/.

Myers-Smith IH, Kerby JT, Phoenix GK, Bjerke JW, Epstein HE, Assmann JJ, John C, Andreu-Hayles L, Angers-Blondin S, Beck PSA, et al. 2020. Complexity revealed in the greening of the Arctic. Nat Clim Chang. 10(2):106–117. doi:10.1038/s41558-019-0688-1.

Nasiri V, Darvishsefat AA, Arefi H, Griess VC, Sadeghi SMM, Borz SA. 2022. Modeling Forest Canopy Cover: A Synergistic Use of Sentinel-2, Aerial Photogrammetry Data, and Machine Learning. Remote Sensing. 14(6):1453. doi:10.3390/rs14061453.

Nguyen TH, Jones SD, Soto-Berelov M, Haywood A, Hislop S. 2018. A spatial and temporal analysis of forest dynamics using Landsat time-series. Remote Sensing of Environment. 217:461–475. doi:10.1016/j.rse.2018.08.028.

OGC. 2021. OGC Best Practice for Earth Observation Application Package. Report No.: 20-089r1. https://docs.ogc.org/bp/20-089r1.html#toc0.

Olesk A, Praks J, Antropov O, Zalite K, Arumäe T, Voormansik K. 2016. Interferometric SAR coherence models for characterization of hemiboreal forests using TanDEM-X data. Remote Sensing. 8(9):700.

Önöz B, Bayazit M. 2003. The Power of Statistical Tests for Trend Detection. Turkish Journal of Engineering and Environmental Sciences. 27(4):247–251. doi:10.3906/sag-1205-120.

Pekel J-F, Cottam A, Gorelick N, Belward AS. 2016. High-resolution mapping of global surface water and its long-term changes. Nature. 540(7633):418–422. doi:10.1038/nature20584.

Pickell PD, Hermosilla T, Frazier RJ, Coops NC, Wulder MA. 2016. Forest recovery trends derived from Landsat time series for North American boreal forests. International Journal of Remote Sensing. 37(1):138–149. doi:10.1080/2150704X.2015.1126375.

Potapov P, Li X, Hernandez-Serna A, Tyukavina A, Hansen MC, Kommareddy A, Pickens A, Turubanova S, Tang H, Silva CE, et al. 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. Remote Sensing of Environment. 253:112165. doi:10.1016/j.rse.2020.112165.

Praks J, Antropov O, Hallikainen MT. 2012. LIDAR-aided SAR interferometry studies in boreal forest: Scattering phase center and extinction coefficient at X-and L-band. IEEE transactions on geoscience and remote sensing. 50(10):3831–3843.

Pulvirenti L, Squicciarino G, Cenci L, Boni G, Pierdicca N, Chini M, Versace C, Campanella P. 2018. A surface soil moisture mapping service at national (Italian) scale based on Sentinel-1 data. Environmental Modelling & Software. 102:13–28. doi:10.1016/j.envsoft.2017.12.022.

Qiu S, Zhu Z, Olofsson P, Woodcock CE, Jin S. 2023. Evaluation of Landsat image compositing algorithms. Remote Sensing of Environment. 285:113375. doi:10.1016/j.rse.2022.113375.

Richter R, Louis J, Muller-Wilm. U. 2012. Sentinel-2 MSI—level 2A products algorithm. European Space Agency Report No.: SP49. [accessed 2020 Dec 22]. https://scholar.google.com/scholar_lookup?title=Sentinel-2+MSI%E2%80%94Level+2A+Products+Algorithm+Theoretical+Basis+Document&author=Muller-Wilm,+U.&publication_year=2012.

Rikimaru A, Roy P, Miyatake S. 2002. Tropical forest cover density mapping. Tropical ecology. 43(1):39–47.

Roerink GJ, Menenti M, Verhoef W. 2000. Reconstructing cloudfree NDVI composites using Fourier analysis of time series. International Journal of Remote Sensing. 21(9):1911–1917. doi:10.1080/014311600209814.

Rolle M, Tamea S, Claps P. 2022. Climate-driven trends in agricultural water requirement: an ERA5-based assessment at daily scale over 50 years. Environ Res Lett. 17(4):044017. doi:10.1088/1748-9326/ac57e4.

Saunier S, Pflug B, Lobos IM, Franch B, Louis J, De Los Reyes R, Debaecker V, Cadau EG, Boccia V, Gascon F, et al. 2022. Sen2Like: Paving the Way towards Harmonization and Fusion of Optical Data. Remote Sensing. 14(16):3855. doi:10.3390/rs14163855.

Schroeder TA, Cohen WB, Song C, Canty MJ, Yang Z. 2006. Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon. Remote Sensing of Environment. 103(1):16–26. doi:10.1016/j.rse.2006.03.008.

Senf C, Müller J, Seidl R. 2019. Post-disturbance recovery of forest cover and tree height differ with management in Central Europe. Landscape Ecol. 34(12):2837–2850. doi:10.1007/s10980-019-00921-9.

Senf C, Pflugmacher D, Wulder MA, Hostert P. 2015. Characterizing spectral–temporal patterns of defoliator and bark beetle disturbances using Landsat time series. Remote Sensing of Environment. 170:166–177. doi:10.1016/j.rse.2015.09.019.

Skakun S, Vermote EF, Roger J-C, Justice CO, Masek JG. 2019. Validation of the LaSRC Cloud Detection Algorithm for Landsat 8 Images. IEEE J Sel Top Appl Earth Observations Remote Sensing. 12(7):2439–2446. doi:10.1109/JSTARS.2019.2894553.

Skidmore AK, Coops NC, Neinavaz E, Ali A, Schaepman ME, Paganini M, Kissling WD, Vihervaara P, Darvishzadeh R, Feilhauer H, et al. 2021. Priority list of biodiversity metrics to observe from space. Nat Ecol Evol. 5(7):896–906. doi:10.1038/s41559-021-01451-x.

Society for Ecological Restoration. 2022. What is Ecological Restoration? https://www.ser-rrc.org/what-is-ecological-restoration/.

Song W, Li M, Gao W, Huang D, Ma Z, Liotta A, Perra C. 2021. Automatic Sea-Ice Classification of SAR Images Based on Spatial and Temporal Features Learning. IEEE Trans Geosci Remote Sensing. 59(12):9887–9901. doi:10.1109/TGRS.2020.3049031.

Tarrio K, Tang X, Masek JG, Claverie M, Ju J, Qiu S, Zhu Z, Woodcock CE. 2020. Comparison of cloud detection algorithms for Sentinel-2 imagery. Science of Remote Sensing. 2:100010. doi:10.1016/j.srs.2020.100010.

Tibshirani R. 1996. Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological). 58(1):267–288. doi:10.1111/j.2517-6161.1996.tb02080.x.

Verhoef, W. 1996. Application of harmonic analysis of NDVI time series (HANTS). Fourier analysis of temporal NDVI in the Southern African and American continents. 108:19–24.

Vermote E, Justice C, Claverie M, Franch B. 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. Remote Sensing of Environment. 185:46–56. doi:10.1016/j.rse.2016.04.008.

Wagner W, Lemoine G, Rott H. 1999. A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data. Remote Sensing of Environment. 70(2):191–207. doi:10.1016/S0034-4257(99)00036-X.

White AM, Long JW. 2019. Understanding ecological contexts for active reforestation following wildfires. New Forests. 50(1):41–56. doi:10.1007/s11056-018-9675-z.

White JC, Saarinen N, Kankare V, Wulder MA, Hermosilla T, Coops NC, Pickell PD, Holopainen M, Hyyppä J, Vastaranta M. 2018. Confirmation of post-harvest spectral recovery from Landsat time series using measures of forest cover and height derived from airborne laser scanning data. Remote Sensing of Environment. 216:262–275. doi:10.1016/j.rse.2018.07.004.

White JC, Wulder MA, Hermosilla T, Coops NC, Hobart GW. 2017. A nationwide annual characterization of 25years of forest disturbance and recovery for Canada using Landsat time series. Remote Sensing of Environment. 194:303–321. doi:10.1016/j.rse.2017.03.035.

White JC, Wulder MA, Hobart GW, Luther JE, Hermosilla T, Griffiths P, Coops NC, Hall RJ, Hostert P, Dyk A, et al. 2014. Pixel-Based Image Compositing for Large-Area Dense Time Series Applications and Science. Canadian Journal of Remote Sensing. 40(3):192–212. doi:10.1080/07038992.2014.945827.

Wilcox R. 1998. A Note on the Theil-Sen Regression Estimator When the Regressor Is Random and the Error Term Is Heteroscedastic. Biom J. 40(3):261–268. doi:10.1002/(SICI)1521-4036(199807)40:3<261::AID-BIMJ261>3.0.CO;2-V.

WRI. 2011. Making Big Ideas Happen.

Wu Z, Feng H, He H, Zhou J, Zhang Y. 2021. Evaluation of Soil Moisture Climatology and Anomaly Components Derived From ERA5-Land and GLDAS-2.1 in China. Water Resour Manage. 35(2):629–643. doi:10.1007/s11269-020-02743-w.

Wulder M. 2016. Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters: Progress in Physical Geography. doi:10.1177/030913339802200402. [accessed 2020 Feb 4]. https://journals.sagepub.com/doi/10.1177/030913339802200402.

Wulder MA, Hermosilla T, White JC, Hobart G, Masek JG. 2021. Augmenting Landsat time series with Harmonized Landsat Sentinel-2 data products: Assessment of spectral correspondence. Science of Remote Sensing. 4:100031. doi:10.1016/j.srs.2021.100031.

Xian GZ, Smith K, Wellington D, Horton J, Zhou Q, Li C, Auch R, Brown JF, Zhu Z, Reker RR. 2022. Implementation of the CCDC algorithm to produce the LCMAP Collection 1.0 annual land surface change product. Earth Syst Sci Data. 14(1):143–162. doi:10.5194/essd-14-143-2022.

Yao S, Jiang D, Zhang Z. 2021 Jun 25. Moisture sources of heavy precipitation in Xinjiang characterized by meteorological patterns. Journal of Hydrometeorology. doi:10.1175/JHM-D-20-0236.1. [accessed 2023 Jan 17]. https://journals.ametsoc.org/view/journals/hydr/aop/JHM-D-20-0236.1/JHM-D-20-0236.1.xml.

Yilmaz M. 2023. Accuracy assessment of temperature trends from ERA5 and ERA5-Land. Science of The Total Environment. 856:159182. doi:10.1016/j.scitotenv.2022.159182.

Yue S, Pilon P, Phinney B, Cavadias G. 2002. The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrol Process. 16(9):1807–1829. doi:10.1002/hyp.1095.

Yue S, Wang CY. 2002. Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test: TECHNICAL NOTE. Water Resour Res. 38(6):4-1-4–7. doi:10.1029/2001WR000861.

Zhang X, Harvey KD, Hogg WD, Yuzyk TR. 2001. Trends in Canadian streamflow. Water Resour Res. 37(4):987–998. doi:10.1029/2000WR900357.

Zhou Q, Zhu Z, Xian G, Li C. 2022. A novel regression method for harmonic analysis of time series. ISPRS Journal of Photogrammetry and Remote Sensing. 185:48–61. doi:10.1016/j.isprsjprs.2022.01.006.

Zhu Z, Woodcock CE. 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote sensing of environment. 118:83–94.

Zhu Z, Woodcock CE. 2014. Continuous change detection and classification of land cover using all available Landsat data. Remote sensing of Environment. 144:152–171.

Zhu Z, Woodcock CE, Holden C, Yang Z. 2015. Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. Remote Sensing of Environment. 162:67–83. doi:10.1016/j.rse.2015.02.009.

Zhu Z, Wulder MA, Roy DP, Woodcock CE, Hansen MC, Radeloff VC, Healey SP, Schaaf C, Hostert P, Strobl P, et al. 2019. Benefits of the free and open Landsat data policy. Remote Sensing of Environment 224: 382-385. 224:382–385. doi:10.1016/j.rse.2019.02.016.

Zhu Z, Zhang J, Yang Z, Aljaddani AH, Cohen WB, Qiu S, Zhou C. 2020. Continuous monitoring of land disturbance based on Landsat time series. Remote Sensing of Environment. 238:111116. doi:10.1016/j.rse.2019.03.009.

Zupanc A. 2017. Improving Cloud Detection with Machine Learning. Sentinel Hub Blog. [accessed 2023 Jan 4]. https://medium.com/sentinel-hub/improving-cloud-detection-with-machine-learning-c09dc5d7cf13.