Communication

From Open Data to Open Analyses—New Opportunities for Environmental Applications?

Thomas Klein [1,*], Mats Nilsson [2], Anders Persson [3] and Bertil Håkansson [1]

[1]   Science Affairs Department, Swedish Agency for Marine and Water Management, SE-404 39 Göteborg, Sweden; bertil.hakansson@havochvatten.se

[2]   Department of Forest Resource Management, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden; mats.nilsson@slu.se

[3]   Geographic Information Division, Swedish Forest Agency, SE-551 83 Jönköping, Sweden; anders.persson@skogsstyrelsen.se

*   Correspondence: thomas.klein@havochvatten.se; Tel.: +46-765-38-6049

Abstract: In this study, we explore the potential of open and accessible data in combination with interactive cloud-processing capabilities for applications in environmental monitoring and policies. During the last few years, the amount of open Earth observation and open national data has increased substantially. In parallel, access to analysis capabilities for such data has improved. The search and extraction of data from larger Earth observations archives and the processing of larger amounts of data have hitherto been an obstacle for many potential users. With the availability of new cloud solutions such as Google Earth Engine, NASA Earth Exchange, or ESA Cloud Toolbox, accessing and processing of large datasets have become easier for a wider range of users. In this communication, we briefly summarize these recent trends and illustrate their potential by four application showcases from terrestrial and aquatic environmental monitoring. We accessed and processed data from US and European Earth observation satellite archives with Google Earth Engine. As a complement, we also used open Swedish national data and open source desktop tools. We hope that our positive user experiences can encourage other environmental data users to further explore the new opportunities for easy access to open data and cloud-based processing capabilities.

Keywords: open data; Earth observation; open analyses; cloud-based data access and processing; Google Earth Engine; terrestrial; aquatic; environmental monitoring and policies

1. Introduction

In this study, we explore the potential of open and accessible data in combination with interactive cloud-processing capabilities for applications in environmental monitoring and policies. The last two decades have seen substantial efforts to better coordinate environmental data and improve their accessibility. Already in 1998, the Arhus-convention [1] established a number of public rights with regard to access to environmental information. Following the Johannesburg World summit on sustainable development, the Group on Earth Observation (GEO) [2] was founded with a specific objective to create a Global Earth Observation System of Systems (GEOSS). Since then, GEO has been working on the coordination of Earth observation and to improve access to data, both for research and environmental decision making. GEO has over many years advocated for a more open sharing of environmental data, in line with increasing evidence that environmental data are more widely used if they are openly accessible [3]. A benchmark was set in 2008 when US Geological Survey (USGS) and National Aeronautics and Space Administration (NASA) decided to make the Landsat data archive openly accessible [4]. The Research Data Alliance (RDA) is a recent international initiative on data

sharing. RDA aims to support researchers and innovators to openly share data across technologies, disciplines, and countries to address the grand challenges of society [5].

In parallel with international initiatives, data sharing and access have also improved at the European level. Regarding data collected by public agencies, one driver has been European legislation. Directive 2003/98/EC and its revision 2013/37/EU have the objective to further the reuse of public sector information (PSI) [6] on a non-discriminatory basis. Directive 2007/2/EC aims to make Europe's geodata more widely interoperable and accessible and to establish an infrastructure for spatial information in the European Community (INSPIRE) [7]. INSPIRE provides legal obligations for open viewing services and metadata and addresses technical challenges related to exchange of metadata and data. Furthermore, the European Community is fostering open access of scientific publications and data in its research programmes [8] with a view to create better science and innovation. Since 1998, Europe had also been working on its own Earth observation programme, the Copernicus-initiative (previously known as Global Monitoring for Environment and Security, GMES). Copernicus became operational in 2014 and comprises six families of satellite missions called Sentinels and six thematic services targeting different applications [9,10]. Copernicus applies a data policy prescribing full, free, and open access to satellite data and products created by the programme itself. Sentinel missions are generally based on constellations of two satellites, which allow for improved revisit time and coverage. Copernicus also plans for long-term continuity of observations with the goal to support environmental monitoring over longer time horizons and to enable the detection of climate trends. Data from Copernicus satellites such as Sentinel 1A and 2A are already available and used in the showcases described in Section 2.

Parallel to the international and European developments many countries, including Sweden, are now moving towards open data. This regards access to environmental data collected by public agencies [11] but also initiatives to improve open access to data collected by scientists [12].

Advantages with open access range from innovative and interdisciplinary research, decision making support, accountability, and public participation to employment and associated tax income (e.g., [3,13,14]). At the same time, data access arrangements need to respect the legal rights and legitimate interests of different stakeholders. As a consequence, access to, or use of, certain data may be restricted for reasons of national security and law enforcement, intellectual property rights, personal privacy and confidentiality, indigenous rights, or protection of sensitive ecological, natural, archaeological, or cultural resources (e.g., [3,13]). However, the observed trend towards open environment data at different levels suggest that open access to publicly funded environmental data is becoming a default, and that necessary restrictions are handled on a case by case basis.

An additional data trend worth mentioning is citizen science, crowd-sourced science, community-based monitoring, or participatory monitoring (e.g., [15]). The increasing abundance of smart devices equipped with geospatially aware utilities can support the collection of observations for scientific research. At the same time, the systematic involvement of non-professional observers also enables public participation in scientific research and support scientific communication. Citizen science has been used in different scientific domains. However, it is important to assess the reliability and account for potential biases of the collected data [15]. An example from natural sciences with reference to Earth observation is the GEO-Wiki [16], where crowdsourcing is used to improve remote sensing-based global land cover products. A recent Swedish example uses citizen observations to improve the monitoring of selected aquatic species [17].

The rapidly increasing amount of environmental data also creates challenges. Firstly, there is a need for long-term management of the data throughout its lifecycle. This is often referred to as data curation and includes topics such as data documentation, discovery, formats, exchange, and long-term preservation. Examples of research infrastructures that work on these challenges are the US National Science Foundation's initiative DataOne [18] or the European Research project OpenAire [19]. The European Commission has also launched the European Open Science Cloud with a view to create a trusted environment for hosting and processing research data [8].

Secondly, there is a need for data dissemination and processing capabilities to handle increasing amounts of data. Data are often made available to users through portals. A global example is the GEOSS-portal [20], which provides an entry point to an extensive amount of Earth observations provided collectively by GEO member states and participating organisations. The GEOSS-portal allows for the search of data according to a variety of criteria, including if the data are part of the GEO Data Collection of Open Resources for Everyone (DataCORE). A European example is the Copernicus Open Access Hub (previously known as Sentinels Scientific Data Hub), which allows registered users to visualize and download Sentinel data [21]. Some countries (e.g., Sweden [22]) also work on the implementation of national solutions for the distribution of Earth observation data. Users can download data from these portals and process them locally. In addition to commercial desktop tools, a variety of geospatial open source software is available for such tasks (e.g., [23]).

With the increasing amount of environment data, cloud access and processing solutions have emerged as a new paradigm. Cloud solutions provide users with access to shared computer processing resources and data. Instead of downloading massive amounts of data and processing them locally, users can access and process data in the cloud and just download the results of the processing. Cloud solutions also allow users to collaborate in the cloud or to share results directly from the cloud. Figure 1 illustrates different ways for users to access environmental data and processing capabilities. Examples for such cloud solutions are European Space Agency (ESA) Research and Service Support (RSS) Cloud Toolbox [24] and NASA Earth Exchange (NEX) [25], which provide access to Earth Observation data in combination with tools and workspace for analyses. In addition, commercial companies have started to offer services combining both data access and processing facilities. For the examples described in this article we used one such initiative, namely Google Earth Engine (GEE) [26]. The use of GEE is free for research, education, and non-profit after an application procedure. GEE provides access to a huge archive of Earth observation data and allows the endorsed user to use a quota of its analysis capabilities. Examples of scientific applications of GEE include the Global Forest Change (GFC) database [27] and the Global Surface Water Explorer [28].
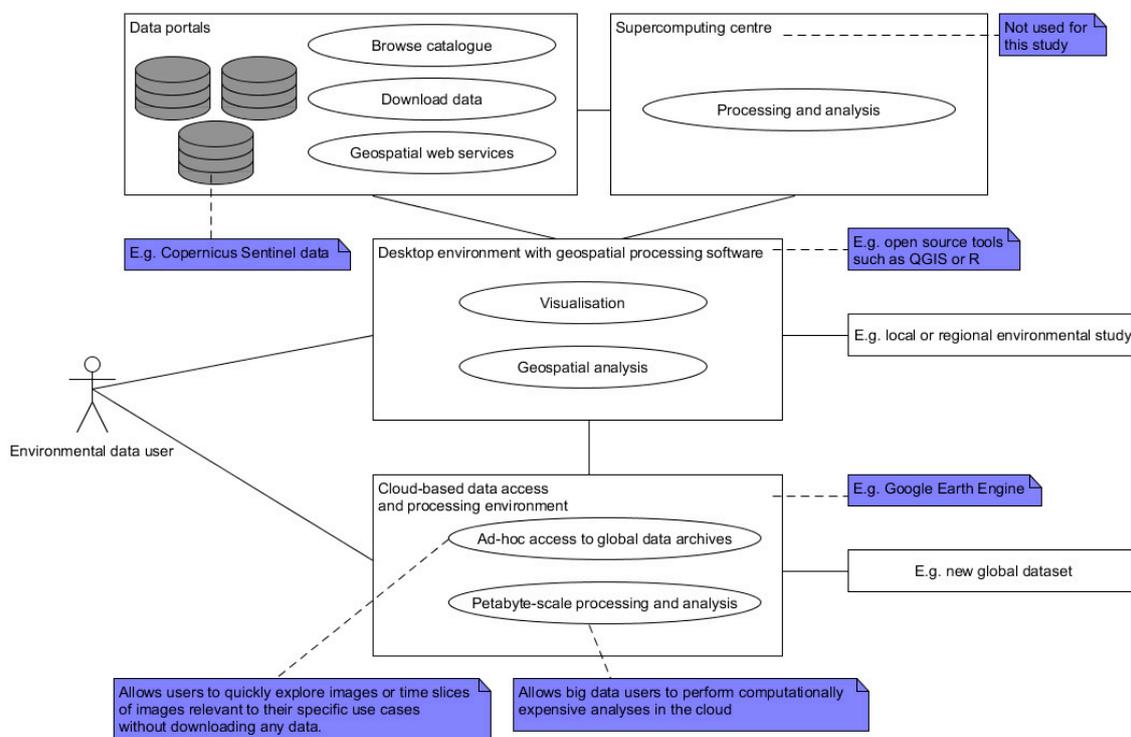

Figure 1. Illustration of different options for environmental users to access data and processing capabilities.

In this study, we explore four terrestrial and aquatic showcases to test the potential of open and accessible data in combination with interactive cloud-processing capabilities for applications in environmental monitoring and policies. The showcases use established methodologies and address environmental user needs rather than new research questions. We used open data from US and European Earth observation satellites and applied GEE to access, extract, and process the data. We hope that the communication of our user experiences can inspire other scientists and environmental data users in different fields to explore these new opportunities further.

2. Materials and Methods

2.1. Description of Showcases

The four showcases comprise user experiences from terrestrial and aquatic environmental monitoring and management. Table 1 provides an overview of the showcases and the data sources used. An overview of the location of the study areas is given in Figure 2a with more detailed views of three study areas in Figure 2b–d.
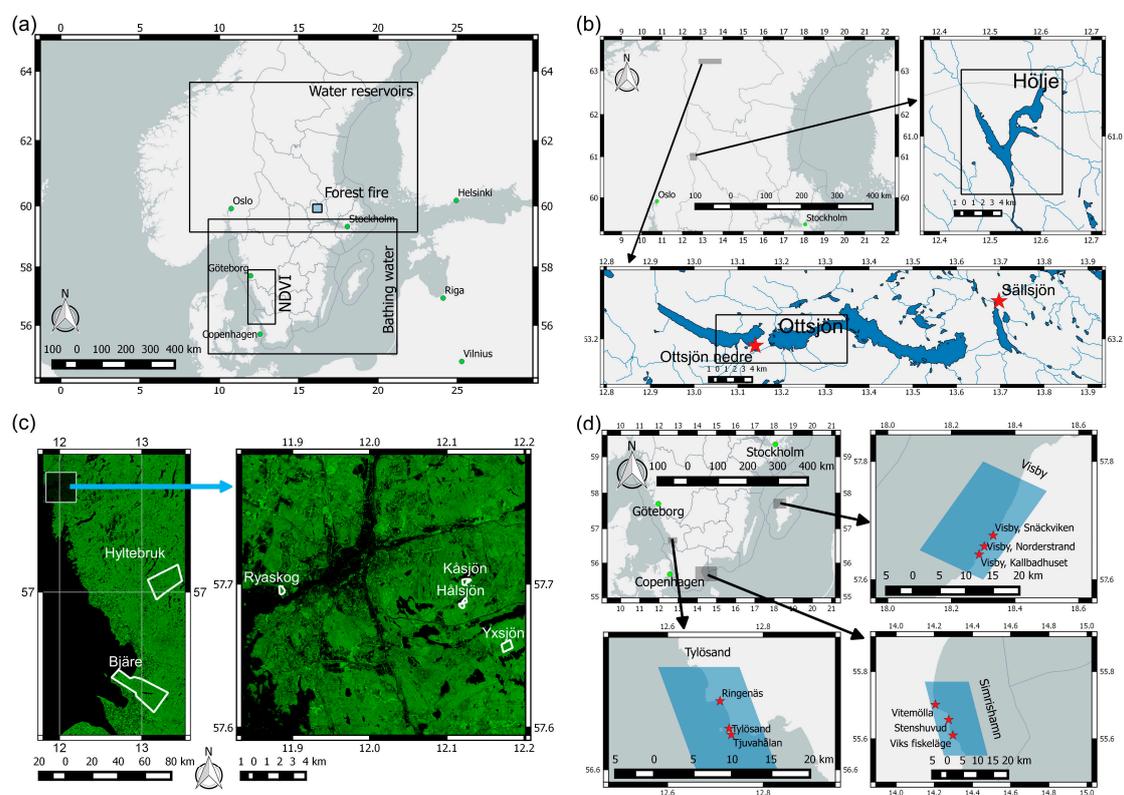

Figure 2. (**a**) Overview map showing the locations of the four study areas. The smallest rectangle (shaded grey) marks the study area for the showcase forest fire in Sweden in 2014. The three larger rectangles mark the positions of study areas shown in detail in (**b**–**d**). (**b**) Study area for the showcase water reservoir monitoring. The two enlargements show the areas of Ottsjön and Hölje, including the extent of water bodies and streams according to [29]. The rectangles in the enlargements correspond to the areas Ottsjön and Hölje used for the comparison of satellite images from different seasons. Background map ©OpenStreetMap Contributors, Swedish county boundaries (grey lines) ©Lantmäteriet. (**c**) Study area and test sites for the showcase Normalized Difference Vegetation Index (NDVI). Background image: NDVI on 5 May 2016 computed from Sentinel 2A. (**d**) Study area for the showcase bathing water monitoring. The three enlargement maps show the location of the in-situ bathing water measurements and the polygons for which water temperatures were calculated from Advanced Very High Resolution Radiometer (AVHRR) data. Background map as in (**b**).

Table 1. Description of showcases.

Showcase	Type	Areas	Data Sources
Normalized Difference Vegetation Index (NDVI)	Seasonal change	Ryaskog, Kåsjön, Hålsjön, Yxsjön, Bjäre, Hyltebruk	Sentinel 2A
Forest fire	Descriptive analysis	Virsbo	Landsat 8, Sentinel 1A, Sentinel 2A, catchment areas [29], valuable waters [30],
Water reservoirs	Change monitoring	Hölje, Ottsjön	Landsat 8, water flow measurements [31]
Bathing water	Seasonal time series	Tylösand, Simrishamn, Visby	AVHRR Pathfinder SST, bathing water measurements [32]

2.1.1. Normalized Difference Vegetation Index (NDVI) Time Series

In this showcase we use time series data from Sentinel 2A for four small study areas (20–70 ha) in southern Sweden: Ryaskog, Kåsjön, Yxsjön, and Hålsjön (Figure 2c). All four areas consist of middle-aged and mature forests. Ryaskog and Kåsjön are dominated by deciduous trees (~80% of the standing volume) while Yxsjön and Hålsjön are dominated by coniferous trees (<15% deciduous) [33]. We also selected two larger study areas: Bjäre and Hyltebruk (Figure 2c). Hyltebruk is dominated by coniferous trees but still contains 24% of deciduous trees and in Bjäre 55% of the standing volume originates from deciduous trees [33]. Both Hyltebruk and Bjäre are dominated by middle-aged and mature forest. Only 19% of the forest area in Hyltebruk and 13% of the forest area in Bjäre are covered by forests younger than 15 years according to GFC (version 1.2, Department of Geographical Sciences, University of Maryland, USA) [27]. A compilation of the forest characteristics for the NDVI study areas is provided in Table 2.

Table 2. Forest characteristics of NDVI study areas.

Area	Type	Deciduous Area (% of Forest)	Forested Area	Area Where Age <15 Years
Ryaskog	Deciduous	No data	100%	0%
Kåsjön	Deciduous	79%	100%	0%
Hålsjön	Coniferous	11%	100%	0%
Yxjön	Coniferous	13%	100%	0%
Bjäre	Deciduous	55%	62%	13%
Hyltebruk	Coniferous	24%	78%	19%

With both Sentinel 2A and 2B operational, global coverage of five days will be provided. At the latitude of Sweden new images will be available every second or third day which offers new possibilities to assess differences between tree species. In this showcase, we compare seasonal changes in NDVI for forest areas with a priori known proportions of deciduous and coniferous trees, respectively. To avoid influence and possible bias from agricultural land the forest land was masked out using the Swedish land surveys classification of forest land. NDVI was calculated from Copernicus Sentinel 2A in GEE data as

$$NDVI = (NIR - VIR)/(NIR + VIR) \tag{1}$$

where NIR and VIR refer to near infrared and visible red bands of the respective sensor. For NIR and VIR we chose bands 8 and 4, respectively, in line with ESA's procedure for Sentinel 2 products [34]. To avoid cloud impacts on NDVI we performed an a priori inspection of Sentinel 2A images available in GEE and selected only cloud-free images.

2.1.2. Forest Fire in Sweden 2014

This showcase addresses a forest fire in Västmanland, Sweden, in August 2014. The fire covered an area of approximately 13,000 ha (Figure 2a) of which 9500 ha were productive forest land [35]. The forested area is dominated by coniferous trees, primarily pine (*Pinus sylvestris*) and spruce forest (*Picea abies*). The event, its consequences, and the emergency management of the crisis have been evaluated in various national reports and investigations (e.g., [36,37].

Two Landsat 8 images, one registered seven weeks prior to the fire event (2014-06-10) and one registered five weeks after the fire event (14 October 2014), and a Sentinel 2A image from 5 May 2016 were used to map the burnt area. In addition, the mean backscattering intensity according to four Sentinel 1A radar images (single co-polarization vertical transmit/vertical receive) acquired during 5 September 2014 to 24 September 2014 was used as an alternative resource for mapping.

The difference of the normalized burn ratio (NBR) computed from the Landsat 8 images registered prior and after the fire was calculated as

$$\text{dNBR} = \text{NBR(10 June 2014)} - \text{NBR(14 June 2014)}, \tag{2}$$

$$\text{where NBR} = (\text{NIR} - \text{SWIR})/(\text{NIR} + \text{SWIR}). \tag{3}$$

For NIR and SWIR (short-wave infrared) we chose Landsat 8 bands 5 and 7, respectively. The resulting dNBR image was masked using a manual delineation of the burnt area in the Landsat 8 image from 10 June 2014. The dNBR was then classified according to categories defined by the USGS FIREMON programme [38] with a view to illustrate the severity of the burn.

2.1.3. Water Reservoir Monitoring

In Scandinavia, energy consumption peaks during the winter season, leading to low water levels in dams and downstream during late winter and early spring. Hydropower dams close or reduce the outflow for filling during snowmelt season and, if necessary, also during summer to reach full capacity for the coming winter [39]. This showcase aims to demonstrate the potential of remote sensing data for the mapping of water surface area variations. For this purpose, we selected the Swedish hydropower dams Ottsjön and Hölje (Figure 2b). Two 8-day Landsat 8 composites with starting dates 23 April 2014 and 14 September 2014 were used to map the water levels in Ottsjön and Hölje.

Water flow data for the two stations Ottsjön Nedre (upstream from the basin) and Sällsjön (downstream) were available from the water web [31] of the Swedish Meteorological and Hydrological Institute (SMHI). The two red stars in Figure 2b show the locations of the two stations. The water flow data were used to explain the surface water area changes at Ottsjön. In order to filter out short-term variability from the daily time series and to highlight variations at the weekly to monthly scale, the cumulative water discharge was calculated for both stations.

2.1.4. Bathing Water Monitoring

The European Bathing Water Directive 2006/7/EC [40] aims to provide information on the bathing water quality to European citizens and requires the regular monitoring of several water quality parameters, including water temperature, during the bathing season. In most cases monitoring uses in-situ measurements on a monthly basis. In this showcase, we investigate the usefulness of sea surface temperature (SST) obtained from satellite measurements as a potential complement to the relatively sparse in-situ observations. Three major Swedish coastal bathing water sites (Tylösand, Simrishamn, and Visby) were used as test areas (Figure 2d). For each test area, we compared in-situ observations with the mean SST from the US National Oceanic and Atmospheric Administration (NOAA) AVHRR Pathfinder dataset [41]. Mean SST was calculated using polygons of about 10 pixels (individual pixel size: 4 × 4 km) close to the coast (Figure 2d). In-situ bathing water measurements were obtained from

the Swedish Bathing Water Information System [32]. We chose the Swedish bathing season (June to August) of 2012 as the time period for this showcase.

2.2. Workflow and Tools

For this study, we used a combination of different ways to access and process environmental data as illustrated in Figure 1. Firstly, we used the cloud-based data access and processing environment GEE. GEE is a computing platform that allows users to run geospatial analysis on Google's infrastructure. Access to GEE is limited and subject to explicit user requests via a web form. An in-depth discussion of the technical details, tutorials, and more is available at GEE's website [26]. GEE provides several ways to interact with its platform. The GEE Code Editor is a web-based Integrated Development Environment (IDE) for writing and running scripts. The GEE Explorer is a lightweight web application for exploring the GEE data catalogue and running simple analyses. In addition, GEE client libraries provide Python and JavaScript wrappers around the GEE Application Programming Interface (API). The GEE data catalogue contains an extensive amount of large public data sets [42], including prominent Earth science raster datasets from major programmes such as Landsat and Copernicus. The GEE web interface provides for easy search in the data catalogue and access to basic documentation of each dataset. Links to the original data source allow the user to access more extensive information. Datasets can either be visualized in the GEE Explorer or imported into the Code Editor with a single click. The Code Editor is a web-based IDE for the GEE JavaScript API with rich data analysis capabilities and numerous example scripts. The extensive documentation of GEE comprises reference manuals, tutorials, and case studies. Both functionality and API are still experimental. For the showcases in this article, we used the GEE web-based IDE for writing and running scripts in JavaScript. Our scripts are relatively simple and use only a subset of the extensive datasets and processing capabilities provided by GEE. Code examples for the four showcases are available at https://code.earthengine.google.com/da3fb00f4f441747f3149cedae02b77a (requires a GEE account) and in the supplementary material. We chose GEE for this study because of several reasons. At the time when the project was started, GEE offered the most extensive data catalogue compared to other solutions such as NASA Earth Exchange or ESA Cloud Toolbox. In addition, the JavaScript API, with its rich documentation, facilitated learning and using GEE. This allowed us to quickly explore large Earth observation datasets and to write scripts to select and process images or time slices of images. The response of GEE in our showcases was very fast (from almost instantaneous to a few seconds), which supported interactive work with large data times slices and aggregations as well as algorithms. However, a detailed benchmarking of GEE-performance and in-depth comparisons between different cloud- and desktop-based tools is outside the scope of this paper. A major benefit with the use of a cloud-based tool such as GEE was that we could access and process data from large data archives without downloading and administrating these data locally. Instead, download could be limited to analyses results. GEE allows the user to save analyses results and export these to different formats, which facilitates further analyses or visualization or results using other tools. For the showcases forest fire and water reservoir monitoring, our analyses resulted in images that were exported to the standard raster format GeoTIFF. For the showcases NDVI and bathing water, the processing results were time series of data aggregations for regions of interest. For the export of time series, we chose the CSV-format.

Secondly, we also used a desktop environment with open source geospatial software for the visualization of GEE analyses results or the analyses of complementary data not available in GEE. While GEE allows the user to upload complementary raster data or vector data for private use, a desktop solution felt more convenient to handle comparably small complementary data and to perform more advanced data visualization. Open data from national portals (e.g., Swedish bathing water measurements [32]) were downloaded and processed locally on the desktop computer. For the plot of time series (showcases NDVI, water reservoir monitoring, and bathing water) we used the free software environment for statistical computing and graphics R [43]. The desktop open source tool QGIS [44] was

used for the production of overview maps and for the combination of the GEE-computed assessment of the burnt area with national data layers in the showcase forest fire. In the latter case, national datasets where included using Open Geospatial Consortium (OGC) [45] web map services (WMS).

3. Results and Discussion

3.1. NDVI Time Series

Figure 3a displays the mean NDVI for the four smaller areas with clearly pronounced forest characteristics. NDVI for the deciduous-dominated areas Ryaskog and Kåsjön shows a stronger seasonal cycle than for the coniferous-dominated areas Yxsjön and Hålsjön. NDVI at Ryaskog and Kåsjön increases substantially starting in May with maximum values of about 0.75 compared to values of about 0.6 at the coniferous sites. We assume that this difference is a result of the leafing of deciduous trees. Figure 3b shows an in-situ photo of the vegetation status at Ryaskog on 8 May 2016, corresponding to the third date in Figure 3a. This is also in agreement with voluntary observations of leafing of trees from the Swedish National Phenology Network (SWE-NPN). According to open maps from SWE-NPN [46], there is a substantial increase of leafing observations in Southern Sweden for most Swedish tree species starting in May 2016 (not shown).
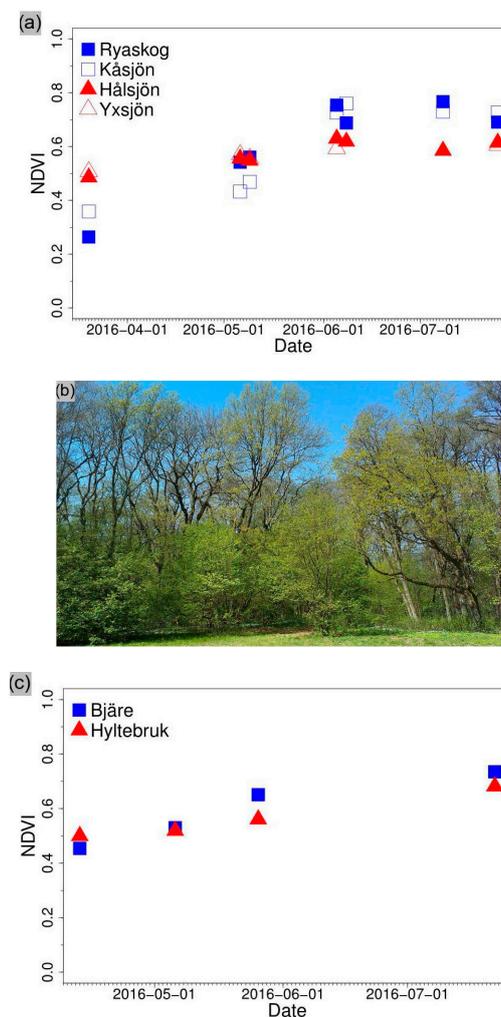

Figure 3. Showcase NDVI. (**a**) NDVI from selected cloud-free Sentinel 2A images for Ryaskog, Kåsjön, Hålsjön, and Yxsjön. Blue and red symbols correspond to deciduous- and coniferous-dominated areas, respectively. (**b**) In-situ photo of vegetation at Ryaskog, 8 May 2016. (**c**) Same as (**a**) but for Bjäre and Hyltebruk.

For Bjäre and Hyltebruk (Figure 3c), the difference of the seasonal cycle of NDVI shows a similar but less pronounced pattern. NDVI at Hyltebruk exceeds NDVI at Bjäre in April. With the leafing of deciduous trees the difference is reversed with NDVI amounting to 0.74 and 0.68 in July at Bjäre and Hyltebruk, respectively. Despite differences of standing volumes of deciduous trees of about 30%, the differences in the seasonal cycle of NDVI are rather small for these two areas. We assume that this is partly a result of the large fraction of relatively young forest in Hyltebruk. This is in line with results by Jin and Eklundh [47] that show that NDVI is rather insensitive for boreal biomes.

Sentinel 2 provides multi-spectral data with a large swath width (290 km) and short revisit time (five days at the equator when both Sentinel 2A and 2B are operational). This in combination with a resolution of 10 m in the four visible and near-infrared bands and 20 m in the six red-edge and shortwave-infrared (SWIR) bands provide new opportunities for, e.g., vegetation mapping and tree species classification. It has been shown that the red-edge and SWIR bands are important for vegetation mapping (e.g., [48,49]). However, the use of time series data in vegetation classification is not trivial. One problem is that data might be lacking due to weather conditions such as clouds. This problem could be addressed by doing a trend reconstruction, e.g., using the Harmonic ANalysis of Time Series (HANTS) method [50].

Although the time series explored for this showcase were rather limited, we found interactive cloud-based infrastructure to be a useful tool to quickly extract relevant satellite images and compute average NDVI for the different study areas. Such access and processing capabilities can be expected to be even more important for larger datasets and computationally expensive analyses.

3.2. Forest Fire in SWEDEN 2014

Figure 4a shows a Landsat 8 image registered after the fire where the area affected by the fire can be clearly identified (appears darker than the surroundings). The burnt area can also be seen in a radar image from Sentinel 1A (Figure 4b). In Figure 4b, the burnt area appears lighter than its surroundings due a stronger ground reflection in the area cleared of vegetation by the fire. In this particular case, the surroundings of the affected area are of mostly rural character, which facilitates the identification of the burnt area. It should be noted that other areas (e.g., urban areas) can have very similar signatures as the burnt area.

Figure 4c provides a more recent view of the burnt area obtained by the multispectral instrument of Copernicus satellite Sentinel 2A, which has a better spatial resolution (pixel size: 10 m) compared to Landsat 8 (pixel size: 30 m). The Sentinel 2A image was registered almost two years after the event, and the area affected by the fire is still clearly visible. It is also possible to get various images of the active fire (not shown) (e.g., directly from Landsat scenes taking during the event), as well as from the extensive archive of open maps produced by the Copernicus Emergency Response Service to support the management of the acute crises [51]. This illustrates the increasing amount of openly available satellite images and products that can help users to identify forest fires and map their impact areas.

Figure 4d shows the difference of the normalized burn ratio (dNBR) calculated using the two Landsat 8 images registered before and after the fire. A visual evaluation indicated that the calculated dNBR reflects the fire severity quite well [52]. Figure 4d also shows water bodies identified as particularly valuable with regard to the Swedish environmental objective of flourishing lakes and streams [30]. The extent of sub-catchment areas is illustrated using the open Swedish Water Archive [29]. Several of the water bodies marked as particularly valuable receive water input from catchments located at least partially in the burnt areas. The Swedish Agency for Marine and Water Management is funding an ongoing investigation that uses in-situ measurements to assess the fire's potential impacts on the water quality in water bodies and streams within the affected catchments.

This showcase benefited substantially from open access to different Earth observation datasets and national open data. Again, interactive cloud-based infrastructure proved useful as a tool to quickly access, explore, and process satellite images. However, for the combination of satellite images with

national GIS-data, we relied on QGIS because of its versatile plotting functionalities and the ability to easily include GIS-layers using OGC's WMS.
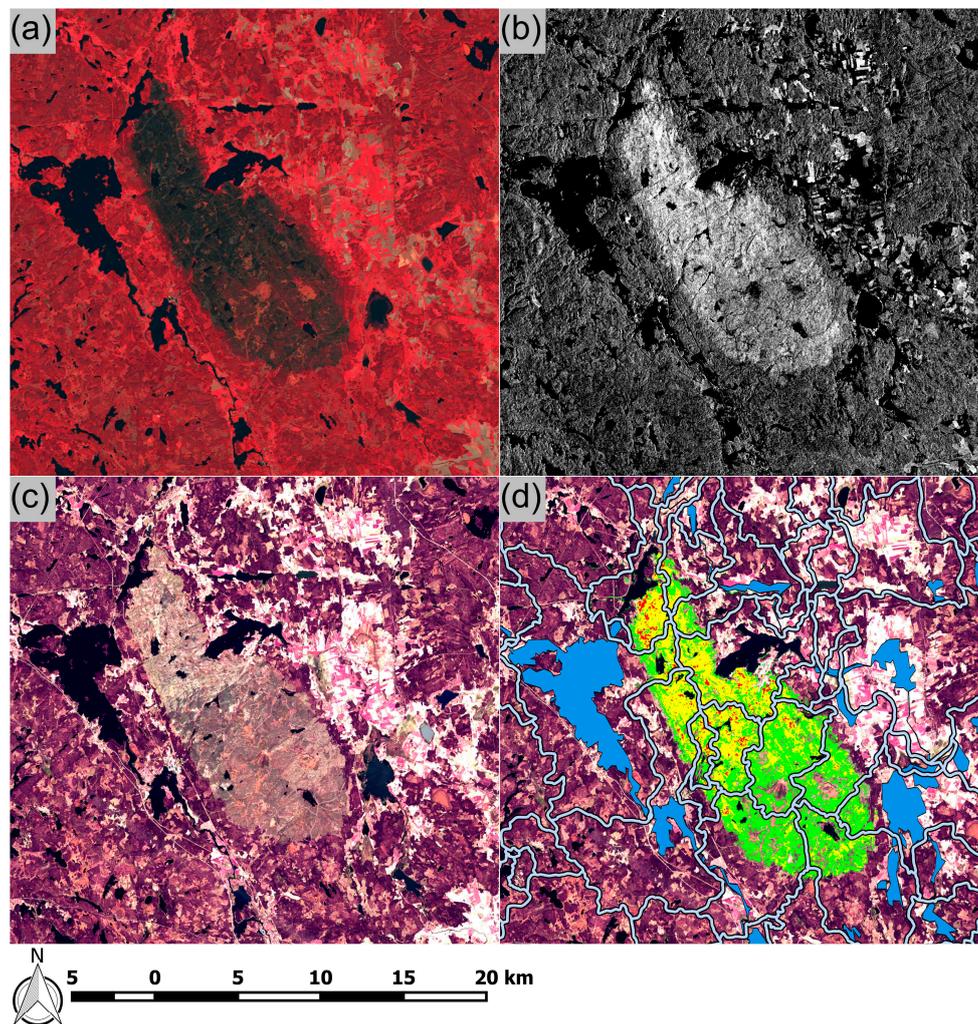

Figure 4. Showcase forest fire. (**a**) Landsat 8 image acquired on 14 September 2014 (about seven weeks after the fire). False colour composite (bands 5, 4, and 3), resolution 30 m. (**b**) Mean of four Sentinel 1A images acquired during 5 September 2014 and 24 September 2014 (ground range detected backscattering intensity, descending orbit, single co-polarization vertical transmit/vertical receive, resolution 10 m). (**c**) Sentinel 2A image acquired 5 May 2016, false colour composite (bands 5, 4, and 3), resolution 10 m. (**d**) Same as (**c**) but with dNBR-layer superimposed. Green, yellow, and red correspond to the severity of the burn (most severely burnt areas in red). In addition, valuable waters according to [30] are shown in blue. Light blue lines with black boundaries denote catchments (sub-basins) according to Swedish Meteorological and Hydrological Institute (SMHI) [29].

3.3. Water Reservoir Monitoring

Figure 5a,b show two Landsat 8 composite images for April 2014 and September 2014. A visual inter-comparison of these two images shows a substantial change in areal extent of Ottsjö dam. The flow dynamics leading to these areal changes can be demonstrated using SMHI flow measurements for stations Ottsjö Nedre (upstream from the basin) and Sällsjön (downstream). In order to filter out short-term variability from the daily time series and to highlight variations at the weekly to monthly scale, we computed the cumulative water discharge for both stations (Figure 5c). Until May 2014 water flow at Ottsjön Nedre is low while water is being discharged at Sällsjön, resulting in the low areal

extent of Ottsjön dam. Starting in early May 2014, the dam is being refilled. In September 2014, Ottsjö dam has been refilled, and the water flows at both stations are roughly in balance. The fact that the cumulative flow at Sällsjön is generally slightly larger than at Ottsjö is due to Sällsjön receiving some additional water inflow from its catchment area. Still, the major inflow comes from Ottsjö, and thus the key dynamics result from dam regulation connected to hydropower needs.
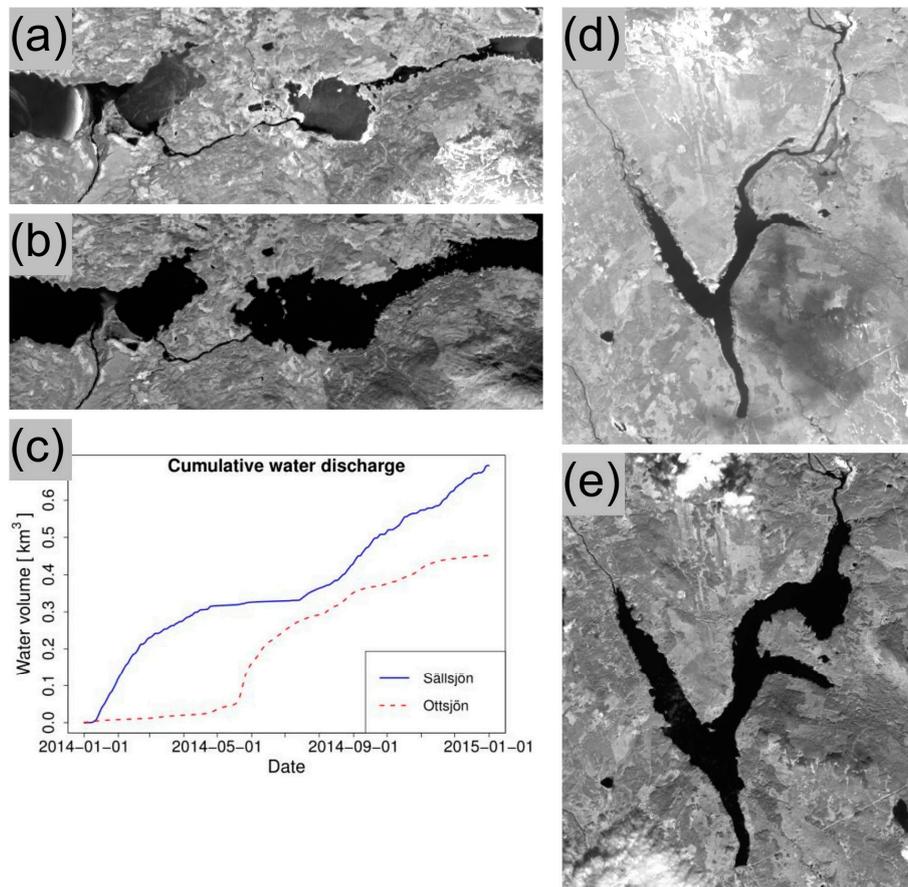

Figure 5. Showcase hydropower dams. (**a**) Satellite image of area Ottsjön. Band 5 of a Landsat 8 8-day composite with start date 23 April 2014. The width of the depicted area is approximately 12 km. For a scale bar see Figure 2b. (**b**) Same as (**a**) but for 8-day composite with start date 14 September 2014. (**c**) Cumulative water flow at stations Ottsjön and Sällsjön according to [29]). (**d**) Same as (**a**) but for Hölje. The width of the depicted area is approximately 10 km. For a scale bar see Figure 2b). (**e**) Same as (**b**) but for Hölje.

The dynamics in Hölje dam are very similar to Ottsjön, with Hölje dam showing low areal extent in April 2014 and large extent in September 2014 after the dam has been refilled (Figure 5d,e). In dams with large shallow areas included, like for the case of Ottsjö and Hölje, these large water level differences imply negative impacts for the shallow water zone and put high stress on the ecosystem, both on the habitats as well as on biotopes in question [39].

Our experiences from this showcase suggest that timely, open, and easily accessible satellite data can be an effective tool to monitor and assess areal changes of man-made dams. For example, the European Water Framework Directive 2000/60/EC [53] requires information on hydromorphological changes of rivers and lakes, e.g., to assess whether there are heavy modifications of a water body.

The two major benefits of using interactive cloud-based infrastructure for this showcase were, firstly, the comprehensive archive of readily available satellite images and, secondly, the versatility of GEE's IDE and tools that allowed us to quickly access and explore satellite images. In principle, such capabilities could support many environmental users that need to monitor water bodies.

3.4. Bathing Water Monitoring

In the final showcase, the potential of using SST from the NOAA AVHRR Pathfinder dataset for monitoring water temperature is demonstrated. The agreement is moderate for Simrishamn, where both the in-situ measurements and the satellite derived SST show relatively large variations (Figure 6a). For Tylösand and Visby, the SST from NOAA AVHRR and the in-situ data are in good agreement (Figure 6b,c). The poorer agreement for Simrishamn might be a result of upwelling, which is a common phenomenon in that area. Thus, the bathing water temperature can vary substantially both between relatively close times and places depending on the prevailing meteorological conditions and associated oceanic currents. A user determining the conditions for a bath in Simrishamn from the monthly in-situ data only might be surprised by the actual bathing water temperatures. A better choice would be to also consider remote sensing products available at higher frequency or possibly model products assimilating in-situ and satellite observations (e.g., from the Copernicus Marine Environmental Monitoring Service) [54].

Besides more representative information on the bathing water and the marine environment, there could also be potential for increased cost effectiveness of the monitoring. Shutler et al. [55] provide a recent review of the use of satellite sensors for the study of physical oceanography processes. In addition, there is progress on the monitoring of water quality aspects using satellite sensors. The newly launched Copernicus satellite Sentinel 3A was still under calibration at the time of writing. Once data from Sentinel 3A's Ocean and Land Colour Instrument (OLCI) become available, there will be possibilities to investigate water quality parameters and derive a variety of biogeochemical products in the marine and freshwater domain, including chlorophyll-a, total suspended matter, and coloured dissolved organic matter. Several of these products will also be relevant for the European Water Framework Directive [53] and the European Marine Strategy Framework Directive [56]. Pilot work in this area is already going on, e.g., by the European research project Global Lakes Sentinel Services (GLASS) [57].

Our experiences from this showcase support our findings from the previous three examples. It is not surprising that open global Earth observation data benefit a wide range of environmental users. However, the value of environmental data is in their application. Thus, there is a need to ensure their accessibility and to provide capabilities that allow both professional and non-expert users to easily apply Earth observation data. As illustrated by the use of GEE in our showcases, cloud-based access and processing environments can indeed deliver such capabilities and complement desktop-based geospatial data processing.
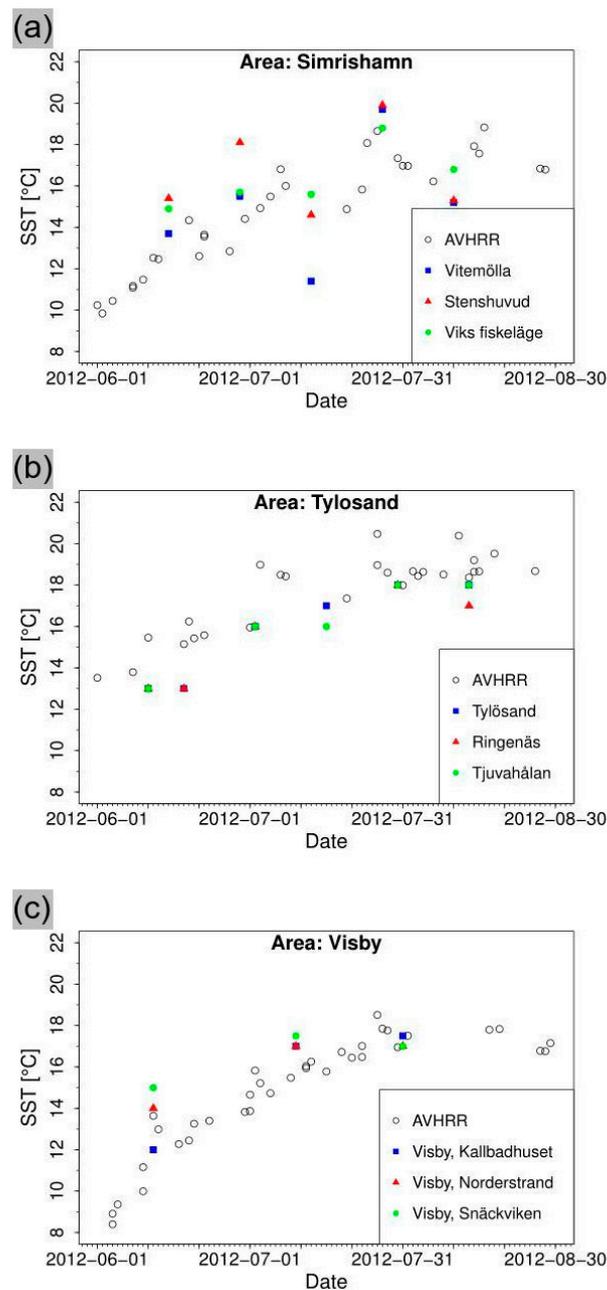
Figure 6. Showcase bathing water. (**a**) Time series of sea surface temperature (SST) for the test area Simrishamn. SST averaged from AVHRR-pixels is shown by open points, while the full symbols correspond to bathing water measurements at the sites Vitemölla, Stenshuvud, and Viks fiskeläge. (**b**) Same as (**a**) but for area Tylösand and observation sites Tylösand, Ringenäs, and Tjuvahålan. (**c**) Same as (**b**) but for area Visby and observation sites Visby Kallbadhuset, Visby Norderstrand, and Visby Snäckviken.

4. Conclusions

The last few years have seen a rapid increase of open environment data both at the global, the European, and the Swedish national levels. At the same time, practical access to, and use of, big data such as multi-year satellite data archives have often been limited to expert users based at organizations with substantial in-house data processing capabilities or access to supercomputing centres. Cloud-based platforms such as GEE establish a new paradigm by making big data archives easily accessible and by providing easy to use processing capabilities to a wider range of users. This

paradigm not only supports petabyte-scale scientific analyses but also facilitates smaller applications. Even users that are only interested in a small fraction of the available data benefit in several ways. There is no longer a need to download and administrate extensive amounts of data just to find the limited portion of data relevant to a specific use case. Instead, the search for, and browsing of, data can easily be performed in the cloud-based archives. In addition, the available processing tools allow the user to easily process and aggregate data and just download relevant results. This new paradigm does not replace but complements traditional desktop-based analyses. We found the combination of these two particularly useful, as it facilitates the use of data from global Earth observation archives for rather local environmental applications and in combination with available open national data.

The user experiences described in this article benefited substantially from open and easy access to Earth observation, national data, and analyses capabilities. The continued increase in open data sources (e.g., forthcoming Earth observation satellites from the European Copernicus programme) will provide additional opportunities for users of environmental data in science and management. Our user experiences so far support the hypothesis that providing open access to large data archives in combination with open access to new environmental analytical capabilities such as interactive cloud-processing could be a key enabling factor for more widespread application of Earth observation data. The findings of our limited study are also in line with studies such as [27,28]. However, it will take years of research to assess the impact of these new opportunities. Based on our experiences from this study, we also conclude that cloud-based data access and processing is not necessarily a replacement but rather a complement to traditional desktop-based analyses and visualization.

In general, the trends towards open access to data, software, standards, and processing capabilities clearly support basic scientific principles such as openness, accessibility, and reproducibility. However, additional challenges will need to be addressed in order to move from open data to fully open and reproducible analyses. Firstly, there is a need to ensure equitable access to interactive cloud-based processing capabilities for a wide spectrum of users. Secondly, fully reproducible analyses will also require open access to the source code of algorithms used in the cloud-based processing environments and the documentation of the specific software versions applied in the respective analyses.

Looking ahead, we hope that the showcases in this article can inspire scientists and environmental data users to explore the new paradigm of cloud-based environmental data access and processing. The global nature of Earth observation and the increasing amount of open data should make it easy to adapt and extend our showcases and to develop applications for different geographic regions and other environmental questions. As an example, we intend to perform an intercomparison of global datasets such as the International Geosphere–Biosphere Programme (IGBP) forest classification [58] and the GFC database with existing national data. Open forest data for Sweden are available from the Swedish National Forest Inventory [33] and the Swedish Forest Agency's data portal [59]. The latter also comprises annual forest fellings. While there can be methodological challenges in comparing global databases with more local datasets, such exercises are important for mutual quality checks and improvements of databases. The latter is of particular importance as global databases may be used to assess environmental changes or even compliance with international agreements in cases where other data sources are not available. Thus, easy access to open data and cloud-based processing capabilities also has a potential to support the consistency and quality of data for monitoring, reporting, and verification of environmental objectives at different levels.

Conflicts of Interest: The authors declare no conflict of interest.

References

1.　The United Nations Economic Commission for Europe (UNECE) Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters. Available online: http://www.unece.org/fileadmin/DAM/env/pp/documents/cep43e.pdf (accessed on 27 January 2017).

2.　Homepage of the Group on Earth Observation (GEO). Available online: http://www.earthobservations.org/ (accessed on 27 January 2017).

3.　Uhlir, P.F.; Chen, R.S.; Gabrynowicz, J.I.; Janssen, K. Toward Implementation of the Global Earth Observation System of Systems Data Sharing Principles. *J. Space Law* **2009**, *35*, 201–290. [CrossRef]

4.　Woodcock, C.E.; Allen, R.; Anderson, M.; Belward, A.; Bindschadler, R.; Cohen, W.B.; Gao, F.; Goward, S.N.; Helder, D.; Helmer, E.; et al. Free access to Landsat imagery. *Science* **2008**, *320*, 1011. [CrossRef] [PubMed]

5.　Research Data Alliance (RDA). Available online: https://www.rd-alliance.org/ (accessed on 20 February 2017).

6.　Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the Re-Use of Public Sector Information. Available online: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:en:PDF (accessed on 27 January 2017).

7.　Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Available online: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32007L0002 (accessed on 27 January 2017).

8.　European Commission Website on European Research & Innovation. Available online: http://ec.europa.eu/research/openscience/ (accessed on 27 January 2017).

9.　ESA Fact Sheet. Available online: http://esamultimedia.esa.int/docs/EarthObservation/sentinels_facts_2015.pdf (accessed on 27 January 2017).

10.　Homepage of the European Earth Observation Programme Copernicus, Previously Known as GMES (Global Monitoring for Environment and Security). Available online: http://www.copernicus.eu/ (accessed on 27 January 2017).

11.　Joint Strategy for Environmental Data Management of the Swedish Environmental Protection Agency, the Swedish Agency for Marine and Water Management, the County Administrative Boards and the Regional Water Authorities. Available online: https://www.naturvardsverket.se/upload/sa-mar-miljon/oppna-data/miljodatastrategi/strategi-for-%20miljodatahantering-20150928-%20version-1-01-eng.pdf (accessed on 27 January 2017).

12.　Swedish Research Council. Proposal for National Guidelines for Open Access to Scientific Information. Available online: https://publikationer.vr.se/en/product/proposal-for-national-guidelines-for-open-access-to-scientific-information/ (accessed on 27 January 2017).

13.　Principles and Guidelines for Access to Research Data from Public Funding. Available online: http://www.oecd.org/dataoecd/9/61/38500813.pdf (accessed on 27 July 2017).

14.　About GMES and Data: Geese and Golden Eggs. A Study on the Economic Benefits of a Free and Open Data Policy for Sentinel Satellite data. Available online: http://esamultimedia.esa.int/docs/EarthObservation/Open_Data_Study_Final_Report.pdf (accessed on 27 January 2017).

15.　Conrad, C.C.; Hilchey, K.G. A review of citizen science and community-based environmental monitoring: Issues and opportunities. *Environ. Monit. Assess.* **2011**, *176*, 273–291. [CrossRef] [PubMed]

16.　Fritz, S.; McCallum, I.; Schill, C.; Perger, C.; Grillmayer, R.; Achard, F.; Kraxner, F.; Obersteiner, M. Geo-Wiki.Org: The Use of Crowdsourcing to Improve Global Land Cover. *Remote Sens.* **2009**, *1*, 345–354. [CrossRef]

17.　Swedish Agency for Marine and Water Management. Reporting App for Voluntary Aquatic Species Observations. Available online: https://www.havochvatten.se/hav/fiske--fritid/arter/frammande-arter/rappen---rapportering-av-vattenorganismer.html (accessed on 27 January 2017).

18.　Data Observation Network for Earth. Available online: https://www.dataone.org/ (accessed on 27 January 2017).

19.　Homepage of the OpenAIRE2020 Project. Available online: https://www.openaire.eu/ (accessed on 27 January 2017).

20. GEOSS-Portal. Available online: http://www.geoportal.org/ (accessed on 27 January 2017).
21. Sentinel Scientific Data Hub. Available online: https://scihub.copernicus.eu/ (accessed on 27 January 2017).
22. Swedish National Space Board. Swea—Swedish Earth Data Access. Available online: http://swea. rymdstyrelsen.se/ (accessed on 27 January 2017).
23. The Open Source Geospatial Foundation. Available online: http://www.osgeo.org/ (accessed on 27 January 2017).
24. ESA RSS Cloud Toolbox. Available online: http://eogrid.esrin.esa.int/cloudtoolbox/ (accessed on 27 January 2017).
25. Home page of the NASA Earth Exchange—NEX. Available online: https://nex.nasa.gov/nex/ (accessed on 27 January 2017).
26. Google Earth Engine: A Planetary-Scale Geospatial Analysis Platform. Available online: https://earthengine. google.com (accessed on 27 January 2017).
27. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.S.; Turubanova, A.; Tyukavina, A.; Thau, D.; Stehman, S.V. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **2013**, *342*, 850–853. [CrossRef] [PubMed]
28. Pekel, J.F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [CrossRef] [PubMed]
29. Swedish Water Archive. Database Version SVAR2012_2. Available online: http://www.smhi.se/ klimatdata/hydrologi/sjoar-och-vattendrag/ladda-ner-data-fran-svenskt-vattenarkiv-1.20127 (accessed on 27 January 2017). (In Swedish)
30. Swedish Agency for Marine and Water Management. En Sammanställning av Sveriges Mest värdefulla Sötvattensmiljöer för Miljökvalitetsmålet Levande sjöar och Vattendrag. Available online: https://www.havochvatten.se/hav/samordning--fakta/kartor--gis/karttjanster/karttjanster/ vardefulla-vatten.html (accessed on 27 January 2017). (In Swedish)
31. Water Flow Observations from SMHI Water Web. Available online: http://vattenwebb.smhi.se/station/ (accessed on 27 January 2017). (In Swedish)
32. Badplatsen—Badplatser och Badvattnets Kvalitet. Available online: https://www.havochvatten.se/hav/ fiske--fritid/badvatten.html (accessed on 27 January 2017). (In Swedish)
33. The Swedish National Forest Inventory of the Swedish University of Agricultural Sciences. Available online: http://www.slu.se/en/Collaborative-Centres-and-Projects/the-swedish-national-forest-inventory/ (accessed on 27 January 2017).
34. Level-2A Algorithm Overview. Available online: https://sentinel.esa.int/web/sentinel/technical-guides/ sentinel-2-msi/level-2a/algorithm (accessed on 27 January 2017).
35. Nilsson, B.; Tyboni, M.; Pettersson, A.; Granström, A.; Olsson, H. *Punktgittertolkning av brandområdet i Västmanland. Redovisning av uppdrag från Skogsstyrelsen*; Work report No. 433; Sveriges lantbruksuniversitet, Institutionen för skoglig resurshushållning: Umeå, Sweden, 2014. (In Swedish)
36. Rapport från Skogsbrandsutredningen. Available online: http://www.regeringen.se/rapporter/2015/03/ rapport-fran-skogsbrandsutredningen/ (accessed on 28 March 2017). (In Swedish)
37. Ansvar, Samverkan, Handling: Åtgärder för Stärkt Krisberedskap Utifrån Erfarenheterna från Skogsbranden i Västmanland 2014. Available online: https://www.msb.se/sv/Produkter--tjanster/Publikationer/ Publikationer-fran-MSB/Ansvar-samverkan-handling--atgarder-for-starkt-krisberedskap-utifran- erfarenheterna-fran-skogsbranden-i-Vastmanland-2014/ (accessed on 27 January 2017). (In Swedish)
38. FIREMON BR Cheat Sheet V4, June 2004. Available online: http://burnseverity.cr.usgs.gov/pdfs/LAv4_ BR_CheatSheet.pdf (accessed on 20 July 2016).
39. Strömberg, M.; Borg, C.; Degerman, E.; Friberg, S.; Jonzon, G.; Jougda, L.; Norström, L.; Sers, B. *Ångermanälvsprojektet—Förslag Till Miljöförbättrande Åtgärder i Mellersta Ångermanälven och Nedre Fjällsjöälven*; Report No. 9; Skogsstyrelsen: Jönköping, Sweden, 2015. (In Swedish)
40. Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 Concerning the Management of Bathing Water Quality and Repealing Directive 76/160/EEC. Available online: http: //eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32006L0007 (accessed on 27 January 2017).

41. Casey, K.S.; Evans, R.H.; Baringer, W.; Kilpatrick, K.A.; Podesta, G.P.; Walsh, S.; Williams, E.; Brandon, T.B. AVHRR Pathfinder Version 5.2 Level 3 Collated (L3C) Global 4 km Sea Surface Temperature for 1981–2012. National Oceanographic Data Center, NOAA. Available online: http://dx.doi.org/10.7289/V5WD3XHB (accessed on 27 January 2017).

42. Google Earth Engine's Public Data Catalogue. Available online: https://earthengine.google.com/datasets (accessed on 21 March 2017).

43. R Foundation for Statistical Computing. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017. Available online: http://www.R-project.org/ (accessed on 27 January 2017).

44. QGIS Geographic Information System. Open Source Geospatial Foundation Project 2016. Available online: http://www.qgis.org/ (accessed on 27 January 2017).

45. Open Geospatial Consortium. Available online: http://www.opengeospatial.org/ (accessed on 27 January 2017).

46. Homepage of the Swedish National Phenology Network. Available online: http://www.naturenskalender.se/ (accessed on 27 January 2017). (In Swedish)

47. Jin, H.; Eklundh, L. A physically based vegetation index for improved monitoring of plant phenology. *Remote Sens. Environ.* **2014**, *152*, 512–525. [CrossRef]

48. Schuster, C.; Förster, M.; Kleinschmit, B. Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. *Int. J. Remote Sens.* **2012**, *33*, 5583–5599. [CrossRef]

49. Immitzer, M.; Vuolo, F.; Atzberger, C. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sens.* **2016**, *8*, 166. [CrossRef]

50. Roerink, G.J.; Menenti, M.; Verhoef, W. Reconstructing cloudfree NDVI composites using Fourier analysis of time series. *Int. J. Remote Sens.* **2000**, *21*, 1911–1917. [CrossRef]

51. Copernicus Emergency Management Service. Activation EMSR096: Fire in Sweden. Available online: http://emergency.copernicus.eu/mapping/list-of-components/EMSR096 (accessed on 27 January 2017).

52. Inka Bohlin; Swedish University of Agricultural Sciences, Umeå, Sweden. Personal communication, 2017.

53. Water Framework Directive. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 Establishing a Framework for Community Action in the Field of Water Policy 2000. Available online: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32000L0060 (accessed on 27 January 2017).

54. Homepage of the Copernicus Marine Environment Monitoring Service. Available online: http://marine.copernicus.eu/ (accessed on 27 January 2017).

55. Shutler, J.D.; Quartly, G.D.; Donlon, C.J.; Sathyendranath, S.; Platt, T.; Chapron, B.; Johannessen, J.A.; Girard-Ardhuin, F.; Nightingale, P.D.; Woolf, D.K.; et al. Progress in satellite remote sensing for studying physical processes at the ocean surface and its borders with the atmosphere and sea ice. *Prog. Phys. Geogr.* **2016**, *40*, 215–246. [CrossRef]

56. Marine Strategy Framework Directive. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 Establishing a Framework for Community Action in the Field of Marine Environmental Policy. 2008. Available online: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32008L0056 (accessed on 27 January 2017).

57. Homepage of the Global Lakes Sentinel Services (GLASS) Project Funded by the European Union under the 7th Framework Programme (Project Number: 313256). Available online: http://www.glass-project.eu/ (accessed on 27 January 2017).

58. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182. [CrossRef]

59. Swedish Forest Agency Data Portal. Available online: http://skogsdataportalen.skogsstyrelsen.se/Skogsdataportalen/ (accessed on 28 March 2017). (In Swedish)