

## Data Acquisition and Processing for GeoAI Models to Support Sustainable Agricultural Practices

Agustin Garcia Pereira  
Insight Centre for Data  
Analytics NUI Galway  
[agustin.garciapereira@insight-centre.org](mailto:agustin.garciapereira@insight-centre.org)

Adegboyega Ojo  
Insight Centre for Data Analytics NUI  
Galway- Dept. of Applied Informatics  
Fac. of Mgt and Economic  
Gdańsk Univ. of Tech  
[adeboyega.ojo@insight-centre.org](mailto:adeboyega.ojo@insight-centre.org)

Edward Curry  
Insight Centre for  
Data Analytics NUI  
Galway  
[edward.curry@insight-centre.org](mailto:edward.curry@insight-centre.org)

Lukasz Porwol  
Insight Centre for  
Data Analytics NUI  
Galway  
[lukasz.porwol@insight-centre.org](mailto:lukasz.porwol@insight-centre.org)

### Abstract

*There are growing opportunities to leverage new technologies and data sources to address global problems related to sustainability, climate change, and biodiversity loss. The emerging discipline of GeoAI resulting from the convergence of AI and Geospatial science (Geo-AI) is enabling the possibility to harness the increasingly available open Earth Observation data collected from different constellations of satellites and sensors with high spatial, spectral and temporal resolutions. However, transforming these raw data into high-quality datasets that could be used for training AI and specifically deep learning models are technically challenging. This paper describes the process and results of synthesizing labelled-datasets that could be used for training AI (specifically Convolutional Neural Networks) models for determining agricultural land use pattern to support decisions for sustainable farming. In our opinion, this work is a significant step forward in addressing the paucity of usable datasets for developing scalable GeoAI models for sustainable agriculture.*

### 1. Introduction

The need for spatial information about agricultural practices is expected to proliferate due to environmental, agronomic, and economic factors. Governments are increasingly concerned about food quality, creating the need for tracing the production and verifying agricultural practices [1].

Modern agricultural practices are a significant cause of environmental pollution [2]. The increasing use of fertilizers, water, and arable land is largely responsible for tonnes of reactive nitrogen and phosphorus ending up in the environment, polluting waterways and coastal zones, accumulating in land

systems, and being released to the atmosphere [2][3]. Meanwhile, the amount of land being used for agricultural purposes is increasing [4]. The Food and Agriculture Organization (FAO) estimates that 11% of the globe's land surface (13.4 billion ha) is used for producing crops. Specifically, in Europe, this percentage is expected to rise to 53% of the total land surface [5]. FAO estimates that by 2050 food production will double due to the increase of the global population and the advent of biofuels [6]. There is an urgent need for more sustainable agricultural practices that are able to generate higher crop yields while minimizing the environmental footprint. Historically, crop rotation has been one of the first sustainable agricultural practices [7]. W. Claupein et al. [8] defines crops rotation as the *regular and orderly alternation in the cultivation of different plant species in a temporal sequence in a given area*. The National Organic Program under the U.S. Code of Federal Regulations requires the adoption of crop rotation practices for farms to receive organic certification [9]. The list of the benefits of crop rotation includes: a) interrupting harmful insect life-cycles, b) suppressing soilborne plant diseases, c) preventing soil erosion, d) building organic matter, e) fixing nitrogen and f) increasing biodiversity of the farm.

In comparison with crops grown in monocropping, crops grown in rotations often yield more due to soil fertility improvements [10], while reducing the external dependency on agrochemicals [11], [12]. Despite the fact that crop rotation's importance has been acknowledged for a long time, such a simple practice has been displaced with the widespread adoption of industrial agriculture. In the last couple of decades, the interest in crop rotations has been renewed due to its benefits, the current climate change situation and its promotion by many agroecological movements across the world [13], [14]. If we are to meet the sustainability challenges presented, new incentives, policies, and technologies are required to monitor how agricultural

practices shape, fortunately not irreversibly, the surface of the Earth [15]. The situation described earlier highlights the importance of mapping and characterizing cropping practices. To this end, the literature suggests that remote sensing has been an effective tool for monitoring the land surface properties resulting from human practices. Despite significant efforts made in this area, an extensive literature review shows that only 9% of the total remote sensing and agriculture publications focus on cropping practices [16]. Moreover, due to the wide variety of agricultural practices and the complexity for describing them over large areas using satellite data, studies have been mostly limited to case studies [17]. Earth is continuously monitored by satellites, drones and different types of sensors as never before, providing rich data at different dimensions' resolutions, while the internet has made it widely available and accessible. For instance, the two Sentinel-2 satellites provide worldwide and freely distributed images every five days, while Landsat-7 and Landsat-8 missions offer an eight-day revisit time [18]. Despite this fact, most recent AI models or classifiers used in operational mapping generally use single date spectral data for classification, due to the technical challenge in leveraging the rich time-series data for characterizing land cover dynamics [19]. Recently, deep learning has shown outstanding results in many fields including remote sensing. Harnessing the ability of these models to learn feature representations exclusively from raw data without the need of domain-specific knowledge, deep neural networks have been used in tasks including image classification, object detection, semantic segmentation, classification from time series and anomaly detection [20], [21] in remotely sensed imagery. This emerging intersection of geospatial science and artificial intelligence is referred to in the literature and in the industry as GeoAI.

In this work, we present a process (pipeline) that consumes satellite imagery and creates temporally-sampled and labelled datasets that could be used for training Convolutional Neural Networks or deep learning models for mapping and identifying cropping practices.

## 2. Background

The availability of Earth Observation (EO) data is continuously increasing thanks to the contributions of Landsat and Sentinel archives [19]. At the same time, community-based geospatial databases collect a vast amount of crowdsourced georeferenced data that contribute to describing objects and things based on their location. It is estimated that 80% of the data

created every day is georeferenced [22] indicating its relevance for discovering knowledge. The increasing volume and types variety of collected geospatial big data creates new opportunities but also poses additional challenges in storing, managing, processing, analyzing and visualizing data [22]. With the advent of cloud computing and GPUs and decreasing computing cost, hardware limitations and software barriers hindering large data processing have been largely lowered. The advances in computing power and data availability comes in parallel to significant developments in the field of artificial intelligence (AI) algorithms, in particular, deep learning. The availability of a plethora of AI solutions and strong industry support, offering affordable data processing services, has opened up opportunities for new applications. The term "GeoAI" has also been already widely utilized by the Industry and practitioners ahead of the majority of academic researchers and scientists in the domain who yet rarely refer to that term even though their works are strongly related. Microsoft and Esri joined their efforts to present the GeoAI data science virtual machine, integrating the popular ArcGIS Pro capabilities, such as geospatial analytics and visualization, with AI cloud technology and infrastructure <sup>1</sup>. In this context, GeoAI is defined by T. Vopham et al. [23] as "*an emerging scientific discipline that combines innovations in spatial science, artificial intelligence methods in machine learning (e.g., deep learning), data mining, and high-performance computing to extract knowledge from spatial big data*" and [24] states that "*the moulding together of artificial intelligence (AI) and the geographic/geographic information systems (GIS) dimension creates GeoAI*".

### 2.1. GeoAI applications

GeoAI models have been developed for different applications. In this section, we describe some specific contributions using GeoAI, and in section 2.2 we describe how these models can be utilized for advancing the mapping and classification of agricultural practices. The publications analyzed were the result of a search in the Scopus database using the keyword "geoai". We also included three articles from the platform Medium<sup>2</sup> that we considered valuable to showcase GeoAI potential. We clustered the applications areas in four different categories: *geospatial modeling*, *imagery processing*, *navigation* and *governance and societal* and we briefly describe them next.

<sup>1</sup> <https://azure.microsoft.com/en-us/blog/microsoft-and-esri-launch-geospatial-ai-on-azure/>

<sup>2</sup> <https://medium.com/>



### 2.1.1. Geospatial modeling

The ability to infer the behavior of different variables in the spatial dimension is key to understand how different phenomena and events occur and the interrelation between them. S. Law et al. [25] used Google Street View images and an abstract 3D model of a city created with Esri City Engine to train a Convolutional Neural Network (CNN) able to predict the probability of having an active frontage on every single street segment in London, providing important insights for urban design and security in public spaces. K. Elgarroussi et al. [26] created a spatiotemporal emotion change analysis framework that monitors and summarizes the change of positive and negative emotions over time and space. Authors used geolocated tweets collected in the state of New York in June 2014, a contour-based spatial clustering algorithm and VADER (a sentiment analysis tool) to assign an emotional score to each tweet. The study presents a novel approach for sentiment change storytelling. In the contribution of G. Xi et al. [27], authors used a geolocated dataset of influenza-like illnesses (ILI) activities in Shenzhen City, China to train a CNN with Residual learning able to predict influenza trends by integrating the spatial-temporal properties of influenza at an intra-urban scale. The experiments show that the model outperformed other four baselines models for one-week-ahead and two-week-ahead predictions. S. Srivastava et al. [28] trained a CNN using Google Street View imagery and a dataset of Amsterdam's buildings to predict multiple co-occurring building function classes per building. The proposed model contemplates several zoom levels of the acquired images and outperformed other baseline CNN models. T. Vopham et al. [23] presents an overview of the use of GeoAI in environmental epidemiology. Authors state that the use of spatial and temporal big data integrated with high-performance computing, data mining, deep learning, and big data infrastructures can help predict the amount of an environmental factor at a particular time and location and produce high-resolution exposure models that represent a particular environmental variable. They point out the potential use of GeoAI for addressing human health-related problems. In this way, [24] describes the importance of location in population and individual health and poses the increasing importance of GeoAI models and new sources of data for improving human health. Despite publicly-available mobility trajectory datasets are fundamental for algorithms' training and performance assessment, genuine privacy concerns restrict their use. To address this problem, V.

Kulkarni [29] used a Nokia Mobile mobility traces dataset of 191 users collected in Switzerland over two years to train a Long Short-term Memory (LSTM) recurrent neural network (RNN) that extracts substantive behavioral patterns of users from actual mobility traces datasets. This work aims to create new and larger datasets, that simulate the actual characteristics of users from a given dataset.

### 2.1.2. Remotely sensed imagery processing

Effective large aerial imagery datasets processing is of fundamental importance for many applications, including maps creation, land use mapping, geological processes, navigation, and place-based studies. W. Li et al. [30], [31] utilized CNN to detect terrain features such as craters, lakes, volcanos, and sand dunes. The model was trained with remote sensing imagery and image augmentation and ensemble learning techniques were used for training. A. Tayyebi [32] used data augmentation techniques on 1-meter resolution remote imagery from National Agriculture Imagery Program (NAIP) from Alabama State to train a CNN able to segment land cover in buildings, roads or parking lots, water, harvested, open land or bare land, forest and planted or dark cropland. On the other hand, D. Jha et al. [33] used satellite imagery from the Esri World Imagery base map to train a CNN model that can detect swimming pools and classify them in neglect or not. In the study, authors were able to map all pools in the City of Redlands in 10 minutes using Google Cloud Services. C. B. Collins et al. [34] trained a super-resolution convolutional neural network (SRCNN) using low resolution images from the Advanced Wide Field Sensor (AWiFS) sensors and high-resolution images from the Linear Imaging Self Scanner (LISS-III) sensors aboard the Indian Space Research Organisation's (ISRO) Resources at-1 and -2 missions to enhance low resolution imagery. Authors showed that the model performs well in places where no high-resolution imagery is available.

### 2.1.3. Navigation

Q. Li [35] trained CNN using sampled images from smartphone's camera videos of indoor locations. Authors used Hidden Markov Model, Viterbi algorithm and a topological map to obtain location information by recognizing landmarks in the indoor environment. The trained CNN correctly recognized the landmarks in the scene. To minimize distance error between prediction and ground-truth traces of an actual car ride J. Murphy [36] trained a CNN that achieves human-level performance when classifying noise level of Global Positioning System (GPS) input data on a given route.

The algorithm conditionally selects between using the raw GPS data and the map-matched route as the best estimate of a driving path. In [37] the authors analyze several Unmanned Aerial Vehicles (UAV) image localization research papers. Authors describe traditional visual localization systems based on image retrieval and image localization based on deep learning approaches. The purpose of visual localization systems is to find the highest matching image between the reference image database and the UAVs aerial image, and output position information in an acceptable time. T. Sun et al. [38] trained a CNN model that combines satellite imagery and GPS data to improve road extraction quality. The study used 120 satellite images from Beijing, paved roads data from OpenStreetMap as ground truth, and GPS data of 65 taxis in Beijing (a total 192 hours of driving) and showed improved performance over other alternatives. N. Pourebrahim et al. [39] compared the performance of Neural Networks and Gravity Models to predict human mobility between cities using Twitter data. Authors used the Origin-Destination Employment Statistics (LODES) for New York City from the U.S. Census Bureau as ground-truth data, containing workers home and employment locations with other characteristics such as age, earnings, industry distributions, and local workforce indicators. More than two million geolocated tweets posted in New York City from June 2015 to May 2016 were also used for the study. Results proved that adding Twitter data to both models improved the performance.

#### 2.1.4. Governance and societal

Governments are responsible for collecting and managing a vast amount of data, but their primary responsibility is to improve the quality of life of the citizens by making data-driven decisions. The previously described contribution of D. Jha et al. [33] for swimming pool detection, also suggests its utilization for vector-borne disease prevention, since the model can classify pools in neglect and non-neglect, preventing the spread of viruses carried by mosquitoes. At the same time, governments can benefit from this model to update the records of swimming pools for taxes assessment. In the last time, 3D city models have become a vital tool for decision making. D. Kudinov et al. [40] used 200 square miles of aerial LiDAR with 213,000 roof segments manually digitized by human editors including types such as flat, gable, hip, shed, dome, vault, and mansard to train a CNN that automatically generates segment polygons of roofs from a raster image. The model contributes to boosting human

editors' productivity and reduce the cost of creating and maintaining 3D city models. T. O. [41] proposed a deep learning semantic segmentation algorithm for post-disaster damage mapping, taking as a case study the 2011 Tohoku Earthquake-Tsunami and aiming at accelerating operational disaster response practice. High-resolution Worldview-2 images were collected before and after the earthquake, and ground truth data were collected from damage inventories for the study area. Although the model showed a good performance, the authors pointed out the limitation of the optical remote sensor in identifying below-roof damage. This study was carried out using the Microsoft Azure resources provided by AI for Earth grant program.

As we can see from the applications described before, GeoAI models are generally supervised learning algorithms that rely heavily on the availability of labeled data for achieving acceptable performances.

## 2.2. GeoAI models for sustainable agriculture

Maged N. et al. [24] points out the potential applicability of GeoAI in capturing and modeling location-based features at a high spatiotemporal resolution. However, GeoAI solutions reviewed have not addressed the mapping and characterization of cropping practices by harnessing the temporal dynamics of land use. GeoAI models, especially CNN's with convolutions applied in the temporal dimension, have the potential to harness significant amounts of data, including spectral and temporal characteristics of remotely sensed images, to classify agricultural practices. That would include unsustainable practices classes such as *three years of soybean monocropping* or *three years of corn monocropping*; and sustainable practices based on crop rotation schemas, such as *alfalfa – fallow – corn* rotation. A similar approach made use of single-date high spatial resolution imagery for classifying land use in six classes [32], as described before. Despite that one author considered using bi-temporal (before and after disaster) satellite imagery for post-disaster damage mapping [41], the solution does not entirely rely on the pre-disaster imagery.

Scarce examples on the application of deep learning techniques on remotely sensed time-series can be found out of the GeoAI term scope. Zhong, Liheng et al. [42] have exploited the intrinsic characteristics of time-series data to describe seasonal patterns and sequential relationships for classifying summer crops. They developed different deep neural network architectures and used Enhanced Vegetation Index (EVI) calculated from Landsat Level 2 product imagery bands and ground in-situ data from California Department of Water Resources (see Table 1). Their results, based on an architecture that includes one-dimension convolution





and an inception module, outperformed traditional algorithms for land use classification including XGBoost, Random Forest, Support Vector Machine and recurrent deep neural networks. Pelletier et al. [43] proposed a temporal convolutional neural network constructed with three convolutional layers, a dense layer and finally, a Softmax layer. Different to [42], authors of this study used three spectral bands of the available satellite imagery. Results show that the proposed architecture outperformed Random Forest algorithm by 2 to 3 % and based on the evidence gathered they point out the importance of using both spectral and temporal dimensions when computing the convolutions.

### 3. Research Challenges

Creating GeoAI models for mapping and characterizing sustainable agricultural practices (and also not sustainable practices) encompasses several challenges. Next, we list and classify them in data-related (a-g) and model-related challenges (h-i). In our contribution, we address the challenges a-g related to time-series data acquisition and processing.

- a) **Lack of ground in-situ data:** Learning deep models for remote sensing images time-series requires geospatial and temporal labeled land use data, rich enough to understand the practices carried out in a specific period and at the pixel level (space). This type of data is usually collected by governmental organizations through land use surveys. Maged N. et al. [24] identified the challenge of scarce availability of labeled training data for GeoAI models.
- b) **Prerequisite ground in-situ data GIS processing:** land use ground in-situ data is usually presented in a single year temporal resolution, while cropping patterns usually involve several years analysis. In this way, GIS processing tasks must be done to generate labels for each minimal homogeneous land parcel. Figure 1 shows an example of ground in-situ data GIS processing step for the years 2017-2018.
- c) **Satellite data acquisition:** high temporal and spatial resolution satellite imagery is needed to create time-series observational data of required length, consistency and continuity to capture land use dynamics features for the study area.
- d) **Temporal mismatch of available ground in-situ and satellite data:** satellites have been launched in different dates. Several ground in-situ data sources available do not match, or partially match, the temporal availability of remotely sensed data for that place. Table 1 presents some identified sources of ground in-situ data together with their date range of availability.

- e) Table 2 lists available satellites providing free imagery together with the launch date, revision time and the number of bands.
- f) **Pixel-level sampling:** Time series can be created following a pixel-based or an object-based approach. In the first, the classification algorithms exploit the temporal variability of spectral characteristics between classes to classify a pixel. On the other hand, an object-based approach also exploits the spatial and textural information of a group of pixels grouped in a meaningful way, relying heavily on a previous perfect segmentation process [44]. Despite an object-based approach might be a valid option for a single year land use classification, where the agricultural fields are well defined [45], in our study we focus on the creation of labelled time-series representing agricultural practices across many years, thus, making the segmentation process dependent on the assumption that agricultural fields won't change their boundaries during that period. The pixel-level sampling process must ensure that the pixel data collected over time, always refer to the same point in the space, as shown in Figure 2.
- g) **Clouds coverage and missing data handling:** in areas of persistent clouds presence, or failure of the sensors, the availability of periodical observations can be compromised. The Best Available Pixel (BAP) strategy, used to create cloud-free and spatially-contiguous image composites over large areas [46], is not applicable for time-series creation since each observation in time at the pixel level is important.
- h) **Data format:** time-series data should be converted to a tabular format indexing the bands' information at the pixel and date level, enabling its consumption using the state-of-the-art deep learning frameworks.
- i) **Lack of temporal network architectures:** popular deep learning architectures currently used in the GeoAI domain exploit only the spatial and spectral data of single images. Land use practices analysis involves harnessing temporal information, together with the spatial and spectral information from satellite images. This shift from single image analysis to time-series analysis creates the need for developing novel architectures [21], [47].
- j) **Spatial non-transferability of the models created:** ground truth data is scarce or inexistent for many places on Earth. The differences in climate, soil and agricultural practices across geographically separated regions compromise the performance of the developed models.

### 4. Methodology

In our study, we focused on addressing the challenges a-g related to time-series data creation,



identified in section 3. Our methodology combines desk research and the quantitative experimental method of analysis to construct a pipeline that enables the consumption of satellite imagery and the creation of temporal-sampled and labeled time-series datasets. In particular, we investigated the most common approaches and tools for image acquisition and processing used by the geospatial community.

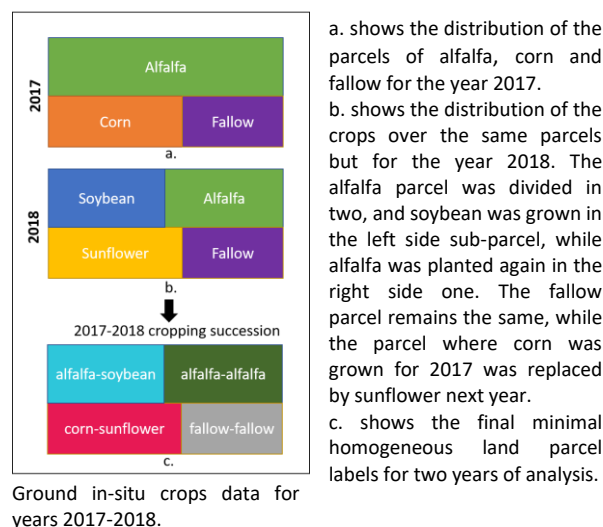


Figure 1. GIS ground in-situ data processing

Table 1. Ground in-situ datasets identified

Dataset	Description	Date range	Link
California Land Use Survey	Yearly land use survey conducted by the California Department of Water Resources. More than 95% of the developed agricultural areas are visited in each survey, and more than 70 crops categories are mapped.	1976, 1986, 1988-1991, 1993-2015	<a href="https://water.ca.gov/Programs/Water-Use-And-Efficiency/Land-And-Water-Use/Land-Use-Surveys">https://water.ca.gov/Programs/Water-Use-And-Efficiency/Land-And-Water-Use/Land-Use-Surveys</a>
Crop-Scape	The Cropland Data Layer is a raster, geo-referenced, crop-specific land cover data layer created annually for the continental United States using moderate resolution satellite imagery and extensive agricultural ground truth	1997-2018	<a href="https://nassgeodat.a.gmu.edu/CropScape/">https://nassgeodat.a.gmu.edu/CropScape/</a>
CEH Land Cover@plus: Crops data	Record of all land registered for agricultural use and are submitted annually by farmers	2015-2018	<a href="https://www.ceh.ac.uk/ceh-land-cover-plus-crop-map-quality-assurance">https://www.ceh.ac.uk/ceh-land-cover-plus-crop-map-quality-assurance</a>

Table 2. Free access observation satellites

Satellite	Launch	Revisit time	Bands №	Resolution
Landsat 7	15/04/1999	8 days	8	30 m to 60 m

Landsat 8	11/02/2013		11	30 m to 100 m
Sentinel-2A	22/06/2015	5 days	13	10 m to 60 m
Sentinel-2B	07/03/2017		13	10 m to 60 m

## 4.1. Context

The context of this study is based on the characterization of agricultural practices in an area of San Jose, California, U.S.A. The region of interest was delimited by human observation at satellite-based maps and aiming at including an extended area of both agricultural activity and ground in-situ data availability. For the purposes of this study, we have a particular interest in analyzing how the Normalized Difference Vegetation Index (NDVI) varies depending on the different types of crops grown in a season. The practice of harvesting more than one crop in the same season is called sequential cropping and, as same as crop rotations, is also done as an ecological practice [17]. Thus, we will also analyze how the one-single crop a year and double-crop a year practices are reflected in the NDVI time-series.

## 4.2. Imagery acquisition

Optical satellite imagery can be consumed using several different tools and platforms. Lately, several online platforms started to provide access to EO data online [48]. During our study, we analyzed two of the most common ones<sup>3,4</sup> and we concluded that the main advantage put forward by those tools is online processing on the infrastructure provided while extracting data from this environment is challenging. Therefore, we dismissed this source. If we consider the research challenge d) we can observe that Landsat constellation provides a lower revisit time than Sentinel, but the overlapped period with ground in-situ data available is much higher. Due to this fact, Landsat products were preferred for this study. Among the available Landsat products, Landsat Level 2 is a research-quality, application-ready science product derived from Landsat Level 1 data [49] and can be downloaded, on-demand, from USGS webpage<sup>5</sup>. Because these images are radiometric-calibrated and atmospheric-corrected, the imagery preprocessing task is significantly simplified, and thus, we decided to use this type of data in our study. Although the QGIS tool provides a popular plugin called “Semi-Automatic

<sup>3</sup> <https://www.onda-dias.eu/cms/>

<sup>4</sup> <https://earthengine.google.com/>

<sup>5</sup> <https://earthexplorer.usgs.gov/>

Classification”, which allows different products downloads, the new U.S. Geological Survey (USGS) Landsat Level 2 product is not available.

The download process requires the user to log in, select a Region of Interest (ROI), by either providing a shapefile or selecting a region from the map, and select the desired time range and image products. The metadata results are sent to the user email in a spreadsheet format. Last, the user needs to submit a list of Landsat Product Identifiers chosen from the received spreadsheet using the interface in Land Satellites Data System Science Research and Development webpage<sup>6</sup>. The contents available for request include Surface Reflectance for all multispectral bands and spectral indexes like NDVI and EVI. Once the order is processed, and the imagery is ready, the data can be downloaded using the bulk downloader script provided by this service.

### 4.3. Ground in-situ data acquisition and processing

Some identified and useful ground in-situ data sources are listed in Table 1. This type of geospatial data is usually distributed in vector or raster files. GIS software such as QGIS and ArcGIS are designed to handle and process this type of geospatial data and help create appropriate labels for the agricultural fields.

### 4.4. Pixel sampling and extraction

Once the images acquired are decompressed and the ground in-situ data that we want to use as labels for the time-series are also downloaded and pre-processed, we face challenge f). This challenge addresses the need for taking pixel samples for each class in the ground in-situ data across all the satellite images, as depicted in Figure 2. The first step for tackling this challenge is to be able to manage all satellite images as a single file containing multiple bands. Virtual Rasters help solve this issue, managing all the images as a virtual single file. Geospatial Data Abstraction Library (GDAL) provides support for its creation<sup>7</sup>. Orfeo ToolBox (OTB)<sup>8</sup> is an open-source project developed by the geospatial community for processing remote sensing images [50]. Despite that this library is very powerful in terms of scalability and versatility and it is widely used in the geospatial

community, it still does not provide a well-defined pipeline for automatically creating ordered time-series from several image sources. In this way, OTB provides support for *Polygon Class Statistics*, *Sample Selection* and *Sample Extraction* at the pixel level.

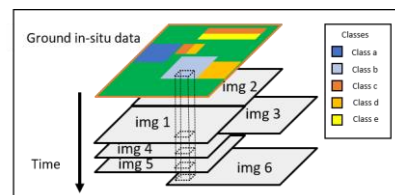


Figure 2. Ground in-situ data and satellite images

Despite that the tool allows the use of several bands during the sample extraction phase, the tool does not provide support for:

1) automatically indexing the captured values at the pixel level by capture date; 2) performing temporal sampling; 3) filling the gaps caused by clouds or missing data; 4) handling the availability of more than one image in the area of interest, for a given date (as depicted in the example of **img 1** and **img 2** in Figure 2). Furthermore, the tool exports its results in SQLite format, creating an additional gap towards fulfilling challenge h). Thus, challenges f), g), and h) remain unsolved. After identifying these issues, we leveraged the OTB processing module in Python addressing these needs.

## 5. Results

We developed an end to end pipeline that can consume a collection of satellite images and a ground in-situ shapefile dataset to create labeled, temporal sampled and linearly interpolated time series at the pixel level. Figure 3 presents a general overview of the pipeline. We have made our code available at <https://github.com/agustinp/remoteSensingTimeSeries> so that others build on our work.

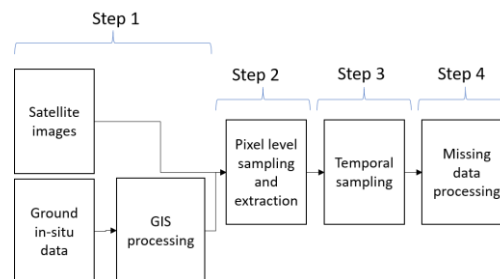


Figure 3. Pipeline schematic

For showcasing the effectivity of our results, we used the pipeline for a collection of satellites images from San Jose, California, U.S.A using ground in-situ data gathered from California Land Use Survey (see

<sup>6</sup> <https://espa.cr.usgs.gov/>

<sup>7</sup> <https://gdal.org/>

<sup>8</sup> <https://www.orfeo-toolbox.org>

Table 1). Next, we describe the results for each of the steps. During *Step 1*, we downloaded 74 satellite images from *Landsat 8 OLI/TIRS C1 Level-2* and *Landsat 7 ETM+ C1 Level-2* products intersecting the area of interest (described previously in the context definition) corresponding to the year 2014 and presenting less than 20% of cloud cover. The images were decompressed in a general directory containing a unique directory for each image product. Each image product consists of several bands, as depicted in Figure 4.

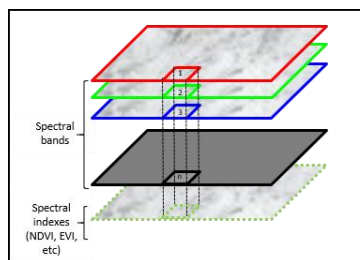
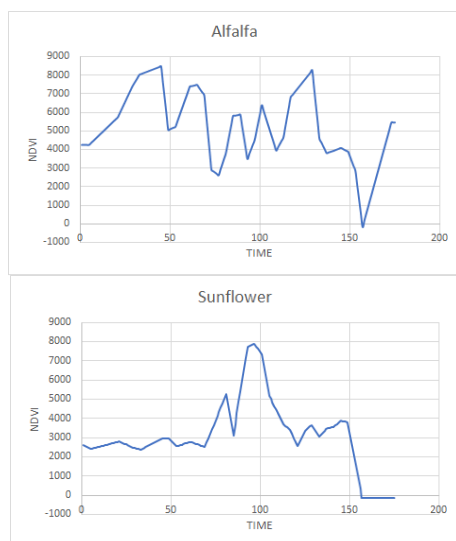


Figure 4. Satellite image product structure

In this case, as we are using Landsat Level-2 processed data, each image product also has its correspondent Surface Reflectance, NDVI, and EVI layer. The ground in-situ data consists of a vector shapefile dataset for the year 2014 containing polygons mapping each agricultural parcel and providing different information at the polygon level.

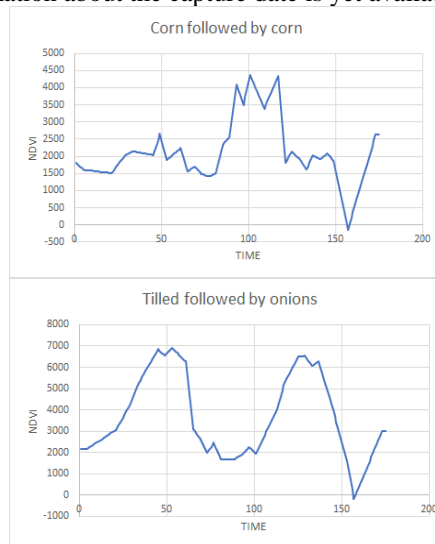


One-year pixels NDVI time-series for single-crop practices. The time dimension is represented using a 2-days grid.

Figure 5. Single crops NDVI

In this case, as we are working with one-year temporal resolution ground in-situ data, there is no need to perform GIS processing to calculate the agricultural practices labels (as was the case for the example depicted in Figure 1, Section 3).

In *Step 2*, pixel-level sampling and extraction are supported by the provided script named *sampling\_extraction.py*. At this stage, the pipeline lets the user select the sampling strategy to be followed and the product band from which to extract the information. By default, all the samples available are extracted, and the pipeline uses NDVI product band, but multiple bands information can be extracted in parallel. The main output of this step is a tabulated structure containing, for each row, the geospatial information of the sample pixel and the bands (in this case NDVI) for each image of the collection. Missing values may occur, and no information about the capture date is yet available.



One-year pixels NDVI time-series for double crop-practices. The time dimension is represented using a 2-days grid.

Figure 6. Double crops NDVI

During *Step 3*, temporal sampling is supported by the provided script named *temporal\_sampling.py*. At this stage, the user can select the starting date for sampling. In this case, we selected “2014-01-01” and a frequency of two days. Considering that for our study area the minimum time distance between images acquired by Landsat 7 and Landsat 8 is one day (due to different orbits overlying the study area), we decided to establish a 2-day time grid, in order to minimize the distortion of this time relation while minimizing the overall time-series length. A 2-day grid will decrease the length of a single year time-series by half, while the worst possible time relation distortion will be of one day. For one year, the pipeline will create a grid of 183 observations and will try to complete it using the available information from the images collection. A look ahead strategy is followed to consider the availability of images on the dates that do not appear in the created grid. Missing values are still present.

Missing data processing is supported by the provided script *interpolation.py* during *Step 4*. The gaps



produced by the absence of imagery, sensors failure, or clouds are filled using linear interpolation. Linear interpolation has demonstrated to be an effective method for remote sensing time series gap-filling [51], [52]. The tabulated structure is then saved in a CSV file, by using one row for each pixel and adding pixel-class information columns as required by the user, together with the time-indexed band values. Different bands values can be extracted for the same pixel, leveraging the satellite images spectral dimension. In Figure 5 and Figure 6 we present some temporal-sampled and labeled NDVI time-series for our region of interest.

## 6. Discussion

The need for high-quality geospatial datasets labeling and characterizing land-use dynamics traverse many areas of applications, including growth urbanization analysis, soil erosion and desertification assessment, floods risk assessment, rural unpaved paths flooding prediction, shore shrinking assessment, city vegetation mapping, agriculture, among others. We believe this paper contributes to advancing GeoAI science in many of the areas mentioned and helps to reduce the gap between geospatial sciences and the AI community. The limitations of this work include lack of interpolation errors assessment for missing data processing and lack of multi-satellite sensor support. Future work will address these issues and will include a segmentation step in the pipeline for enabling object-based analysis approach complementing the pixel-based approach followed in this publication, as well as a pixel-neighbourhood approach to also leverage the spatial dimension at the pixel level.

## 7. Conclusions

In this paper, we presented the challenges associated with the creation of high-quality geospatial temporally-sampled and labeled datasets for sustainable agriculture GeoAI models development. We designed a pipeline for addressing these challenges, and we provided the actual implementation of the scripts supporting this process. Finally, we discussed other areas of application of the presented solution as well as limitations and future work.

"This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289\_P2, co-funded by the European Regional Development Fund".

## 8. References

- [1] J. De Baerdemaeker, *Precision agriculture technology and robotics for good agricultural practices*, vol. 1, no. PART 1. IFAC, 2013.
- [2] K. J. Noone and C. Folke, "A safe operating space for humanity," no. May 2014, 2013.
- [3] D. Malakoff, "Death by Suffocation in the Gulf of Mexico," *Science* (80-. ), vol. 281, no. 5374, pp. 190–192, 1998.
- [4] Millennium Ecosystem Assessment (Program), *Ecosystems and human well-being: our human planet: summary for decision-makers*, no. 5. 2005.
- [5] M. D. A. Rounsevell, J. E. Annetts, E. Audsley, T. Mayr, and I. Reginster, "Modelling the spatial distribution of agricultural land use at the regional scale," *Agric. Ecosyst. Environ.*, vol. 95, no. 2–3, pp. 465–479, 2003.
- [6] FAO, "The challenge of Global Agriculture towards 2050," *How to Feed World 2050*, pp. 1–4, 2009.
- [7] A. H. Society, "Fallowing , Crop Rotation , and Crop Yields in Roman Times Author ( s ): Kenneth D . White Published by: Agricultural History Society Stable URL: <http://www.jstor.org/stable/3741455>," vol. 44, no. 3, pp. 281–290, 2016.
- [8] W. Claupein, "Geisler, G.: Pflanzenbau. Ein Lehrbuch - Biologische Grundlagen und Technik der Pflanzenproduktion. Zweite, neubearb. und erw. Aufl., Verlag Paul Parey, Berlin und Hamburg, 1988; 530 Seiten, 201 Abbildungen und 282 Tabellen, kartoniert DM 128,-; ISBN 3-48," *J. plant Nutr. soil Sci.*, vol. 151, no. 4, pp. 273–274, 1988.
- [9] P. Coleman, *Guide for Organic Crop Producers*. 2012.
- [10] B. G. Ellis, W. L. Hargrove, F. J. Pierce, and C. W. Rice, "Crop Rotation and Its Impact on Efficiency of Water and Nitrogen Use," no. 12542, 2013.
- [11] S. J. Sarandón and Claudia Cecilia Flores, *Agroecología: bases teóricas para el diseño y manejo de Agroecosistemas sustentables*. Universidad Nacional de La Plata – Editorial de la Universidad de La Plata, 2014.
- [12] T. Dias, A. Dukes, and P. M. Antunes, "Accounting for soil biotic effects on soil health and crop productivity in the design of crop rotations," *J. Sci. Food Agric.*, vol. 95, no. 3, pp. 447–454, 2014.
- [13] R. Sahajpal, X. Zhang, R. C. Izaurralde, I. Gelfand, and G. C. Hurtt, "Identifying representative crop rotation patterns and grassland loss in the US Western Corn Belt," *Comput. Electron. Agric.*, vol. 108, pp. 173–182, 2014.
- [14] TWN, *Agroecology: Key Concepts and Practices*. 2015.
- [15] D. Tilman, K. G. Cassman, P. A. Matson, R. Naylor, and S. Polasky, "Agricultural sustainability and intensive production practices," *Nature*, vol. 418, no. August, pp. 671–677, 2002.
- [16] A. Bégué *et al.*, "Agricultural Systems Studies using Remote Sensing To cite this version: HAL Id: hal-02098284," 2019.
- [17] A. Bégué *et al.*, "Remote sensing and cropping practices: A review," *Remote Sens.*, vol. 10, no. 1, pp. 1–32, 2018.
- [18] FAO, "Handbook on remote sensing for agricultural statistics," 2016.
- [19] C. Gómez, J. C. White, and M. A. Wulder, "Optical remotely sensed time series data for land cover classification: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 116, pp.



55–72, 2016.

[20] R. Cresson, “A Framework for Remote Sensing Images Processing Using Deep Learning Techniques,” *IEEE Geosci. Remote Sens. Lett.*, 2018.

[21] X. X. Zhu *et al.*, “Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources,” *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 8–36, 2017.

[22] S. Li *et al.*, “Geospatial big data handling theory and methods: A review and research challenges,” *ISPRS J. Photogramm. Remote Sens.*, vol. 115, pp. 119–133, 2016.

[23] T. Vopham, J. E. Hart, F. Laden, and Y. Y. Chiang, “Emerging trends in geospatial artificial intelligence (geoAI): Potential applications for environmental epidemiology,” *Environ. Heal. A Glob. Access Sci. Source*, vol. 17, no. 1, pp. 1–6, 2018.

[24] M. N. Kamel Boulos, G. Peng, and T. Vopham, “An overview of GeoAI applications in health and healthcare,” *Int. J. Health Geogr.*, vol. 18, no. 1, pp. 1–9, 2019.

[25] S. Law, Y. Shen, and C. Seresinhe, “An application of convolutional neural network in street image classification,” *Proc. 1st Work. Artif. Intell. Deep Learn. Geogr. Knowl. Discov. - GeoAI '17*, pp. 5–9, 2017.

[26] K. Elgarroussi, S. Wang, R. Banerjee, and C. F. Eick, “Aconagua: A Novel Spatiotemporal Emotion Change Analysis Framework,” *Proc. 2nd ACM SIGSPATIAL Int. Work. AI Geogr. Kwl. Discov.*, no. Ccdm, pp. 54–61, 2018.

[27] G. Xi and S. Mei, “A Deep Residual Network Integrating Spatial-temporal Properties to Predict Influenza Trends at an Intra-urban Scale,” 2009.

[28] S. Srivastava, J. E. Vargas-Muñoz, D. Swinkels, and D. Tuia, “Multilabel Building Functions Classification from Ground Pictures using Convolutional Neural Networks,” *Proc. 2nd ACM SIGSPATIAL Int. Work. AI Geogr. Knowl. Discov.*, pp. 43–46, 2018.

[29] V. Kulkarni, “Generating Synthetic Mobility Traffic Using RNNs,” pp. 1–4.

[30] W. Li and B. Zhou, “Recognizing terrain features on terrestrial surface using a deep learning model -- An example with crater detection.”

[31] W. Li and C. Y. Hsu, “Automated terrain feature identification from remote sensing imagery: a deep learning approach,” *International Journal of Geographical Information Science*, vol. 00, no. 00, Taylor & Francis, pp. 1–24, 2018.

[32] A. Tayyebi, “High Resolution Land Cover Mapping using Deep learning.” [Online]. Available: <https://medium.com/geoai/high-resolution-land-cover-mapping-using-deep-learning-7126fee571dd>. [Accessed: 23-Feb-2019].

[33] D. Jha and R. Singh, “Swimming pool detection and classification using deep learning.” [Online]. Available: <https://medium.com/geoai/swimming-pool-detection-and-classification-using-deep-learning-aaf4a3a5e652>. [Accessed: 23-Feb-2019].

[34] C. B. Collins, J. M. Beck, S. M. Bridges, J. A. Rushing, and S. J. Graves, “Deep Learning for Multisens or Image Resolution Enhancement,” 2017.

[35] Q. Li, “Visual Landmark Sequence-based Indoor Localization,” no. 1, 2017.

[36] J. Murphy, “Image-based Classification of GPS Noise Level using Convolutional Neural Networks for Accurate

Distance Estimation.”

[37] Y. Xu, L. Pan, C. Du, J. Li, N. Jing, and J. Wu, “Vision-based UAVs Aerial Image Localization: A Survey,” *Proc. 2nd ACM SIGSPATIAL Int. Work. AI Geogr. Knowl. Discov.*, pp. 9–18, 2018.

[38] T. Sun, Z. Di, and Y. Wang, “Combining Satellite Imagery and GPS Data for Road Extraction,” vol. 3281550, no. c, pp. 4–7, 2018.

[39] N. Pourebrahim, J.-C. Thill, S. Sultana, and S. Mohanty, “Enhancing trip distribution prediction with twitter data: Comparison of neural network and gravity models,” *Proc. 2nd ACM SIGSPATIAL Int. Work. AI Geogr. Knowl. Discov. GeoAI 2018*, 2018.

[40] D. Kudinov, D. Hedges, and O. Maher, “Reconstructing 3D buildings from aerial LiDAR with AI: details.” [Online]. Available: <https://medium.com/geoai/reconstructing-3d-buildings-from-aerial-lidar-with-ai-details-6a81cb3079c0>. [Accessed: 23-Feb-2019].

[41] T. O. Satellite-based and D. U. U. Convolutional, “Towards Operational Satellite-Based Damage-Mapping Using U-Net Convolutional Network: A Case Study of 2011 Tohoku,” 2018.

[42] L. Zhong, L. Hu, and H. Zhou, “Deep learning based multi-temporal crop classification,” *Remote Sens. Environ.*, vol. 221, no. November 2018, pp. 430–443, 2019.

[43] C. Pelletier, G. I. Webb, and F. Petitjean, “Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series,” no. March, 2018.

[44] N. Zerrouki and D. Bouchaffra, “Pixel-based or object-based: Which approach is more appropriate for remote sensing image classification?,” *Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern.*, vol. 2014–January, no. January, pp. 864–869, 2014.

[45] Z. Zhou, S. Li, and Y. Shao, “Object-oriented crops classification for remote sensing images based on convolutional neural network,” *Image Signal Process. Remote Sens. XXIV*, vol. 1078922, no. October 2018, p. 78, 2018.

[46] J. C. White *et al.*, “Pixel-based image compositing for large-area dense time series applications and science,” *Can. J. Remote Sens.*, vol. 40, no. 3, pp. 192–212, 2014.

[47] A. Kamlaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Comput. Electron. Agric.*, vol. 147, no. February, pp. 70–90, 2018.

[48] ESA, “Satellite Earth Observations in Support of the Sustainable Development Goals,” *Report*, 2018.

[49] “Landsat Science Products.” [Online]. Available: <https://www.usgs.gov/land-resources/nli/landsat/landsat-science-products>. [Accessed: 14-Jun-2019].

[50] M. Grizonnet, J. Michel, V. Poughon, J. Inglada, M. Savinaud, and R. Cresson, “Orfeo ToolBox: open source processing of remote sensing images,” *Open Geospatial Data, Softw. Stand.*, vol. 2, no. 1, pp. 0–7, 2017.

[51] J. Chen, P. Jönsson, M. Tamura, Z. Gu, B. Matsushita, and L. Eklundh, “A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter,” *Remote Sens. Environ.*, vol. 91, no. 3–4, pp. 332–344, 2004.

[52] C. L. M. de O. Santos *et al.*, “Classification of crops, pastures, and tree plantations along the season with multi-sensor image time series in a subtropical agricultural region,” *Remote Sens.*, vol. 11, no. 3, 2019.

