# **Towards a Temporal Deep Learning Model to Support Sustainable Agricultural Practices**

Agustin Garcia Pereira<sup>1</sup>, Lukasz Porwol<sup>1</sup>, Adegboyega Ojo<sup>12</sup>, and Edward Curry<sup>1</sup>

<sup>1</sup> Insight Centre for Data Analytics, NUI Galway, Ireland.
<sup>2</sup> Department of Applied Informatics, Faculty of Management and Economics, Gdańsk University of Technology, Poland. {name}.{surname}@insight-centre.org

Abstract. The impressive results of deep learning in many different fields, specifically in remote sensing, together with the growing availability of open Earth Observation data creates new opportunities to address global problems. One such global problem is associated with the simplification and intensification of agricultural systems which threatens the worldwide sustainability of crop production. Despite the fact that a plethora of satellite images describe a given location on earth every year, very few deep learning-based solutions have harnessed the temporal and sequential dynamics of land use to map sustainable and unsustainable cropping practices. In this paper, we present the preliminary results of a set of experiments conducted using one-dimensional Convolutional Neural Networks (CNN) for classifying multispectral time series derived from Landsat satellites constellation. The experimental data is related to agricultural practices in Sacramento County, California, United States of America. We discuss the applicability of this approach for mapping sustainable crop rotationbased practices which have been proven to mitigate the environmental impact of agricultural land use dynamics.

Keywords: AI, Deep Learning, Satellite Images, Sustainable Agriculture

# 1 Introduction

Confronted with important global problems related to agriculture sustainability, food security, climate change, and biodiversity loss, new ecological movements across the world are promoting a set of "ecological intensification" principles, as an alternative paradigm to mainstream agricultural practices [1–3]. Practices such as intercropping, double cropping, crop rotations and the use of cover crops have been shown to increase agriculture sustainability [4]. There is an increasing tendency among farmers, decision-makers, and society in general to establish cropping systems that allow, not only the maximization of crop yield but also the provision of ecosystem benefits [4]. In this regard, the need for spatial information about agricultural practices is expected to grow rapidly [5], and remote sensing has been shown to be an effective tool for monitoring the land surface properties resulting from human practices. Despite signif-

Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

icant 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 [6].

Recent applications of deep learning in many fields, including remote sensing, together with the increasing availability of free satellite images with higher spectral, spatial and temporal resolutions, creates new opportunities to tackle global challenges. Deep learning-based models have the ability to learn feature representations exclusively from raw data without the need for domain-specific knowledge. This fact, together with the advances in computational power, has encouraged the use of deep neural networks for many tasks, including image classification, object detection, semantic segmentation and anomaly detection [7, 8] in remotely sensed imagery. However, most recent AI models or classification [9] and do not harness the temporal resolution of remotely sensed time-series images.

In this work, we present a set of experiments using Convolutional Neural Networks (CNNs) with convolutions in the temporal dimension and more than 400,000 remotely sensed time series data to classify land use and agricultural practices.

# 2 Related Work

Zhong, Liheng et al. [10] 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. 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. [11] proposed a temporal convolutional neural network constructed with three convolutional layers, a dense layer and finally, a SoftMax layer. Different to [10], the 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. Cai et al. [12] developed a deep learning architecture to train a model able to classify corn and soybean fields. They used a combination of Landsat-5, Landsat-7 and Landsat-8 satellite images time-series covering a period of sixteen years. They report an overall accuracy of 97%.

### 3 Data

### 3.1 Study Area

The setting of the study is a surface of 4466 km<sup>2</sup> or 1724 square miles in Sacramento County, in the west part of the United States of America and encompasses a one-year

period ranging from January 2015 to December 2015. **Fig. 1** shows the delimited surface. The annual mean temperature in this area is 16.1 °C with a monthly daily average temperature ranging from 8.0 °C in December and 24.2 °C in July. The wet season period extends from October to April. The region of Sacramento has a strong agricultural tradition and remains an important economic force not only in California but also at a national level [13].



Fig. 1. Study Area, Sacramento County, United States of America. The red thick line delimits the region of interest for this study and the green polygons represent the agricultural fields our experiments and analysis will be based on.

#### 3.2 Ground Truth Data

In this study, the 2015 Sacramento County land use survey was used as a source of "ground truth" data. The Survey was developed by the State of California, Department of Water Resources (DWR). The main goal of this survey is to map agricultural fields. In this regard, a surveyor visited almost every delineated field, providing a high land use assessment accuracy.

The dataset is distributed in a Shapefile format and consists of a total of 40205 polygons, each of them containing 29 different attributes that describe the agricultural land use practices at the field level.

#### 3.3 Satellite Imagery

Different satellite constellations provide freely distributed images of the world surface in a continues manner and at different spatial, spectral and time resolutions. For instance, the two Sentinel-2 satellites provide 10m resolution imagery of the planet surface every five days, whereas a combination of Landsat-7 and Landsat-8 satellites offer an eight days revisit time with a 30m spatial resolution [14]. Due to the fact that satellites have been launched at different dates, a match between the ground truth data to be used and the availability of remotely sensed data represents a strong limitation at the time of selecting a satellite product. Considering that the land use survey described before is based on the agricultural fields for the year 2015, a combination of Landsat-7 and Landsat-8 was deemed as the best option. It offers an eight-days revisit time since 2013, whereas Sentinel-2 images only offer five-days revisit time since the year 2017.

Among the available Landsat products, Landsat Level 2 is a research-quality, application-ready science product derived from Landsat Level 1 data [15] and can be downloaded, on-demand, from USGS webpage<sup>1</sup>. The selection of this source of remotely sensed data is motivated by the fact that these images are radiometric-calibrated and atmospheric-corrected.

A total of 178 Surface Reflectance image products were downloaded for the region of interest delimited with red in **Fig. 1** for the year 2015. From this set, 88 images correspond to Landsat-8 and a total of 90 images correspond to Landsat-7. Three bands were selected for the application of this study. The green band emphasizes peak vegetation, which is useful for assessing plant vigor. Red Band discriminates vegetation slopes while Near Infrared (NIR) emphasizes biomass<sup>2</sup>. **1** summarizes the bands' information per satellite.

Landsat 7			Landsat 8			
Band #	Name	Wavelength (µm)	Band #	Name	Wavelength (µm)	
Band 2	Green	0.52-0.60	Band 3	Green	0.53-0.59	
Band 3	Red	0.63-0.69	Band 4	Red	0.64-0.67	
Band 4	NIR	0.77–0.90	Band 5	NIR	0.85–0.88	

Table 1. Spectral bands' information.

As we can see from **Table 1**, the spectral ranges of the different bands are slightly different between Landsat-7 and Landsat-8. These differences have been studied in [16] suggesting that their impact on a model depends on the sensitivity of the model in question. Studies have shown the insignificant impact of these differences on classification models [10, 12].

#### 3.4 Data preprocessing

In order to create pixel-level labeled time-series we followed the pipeline we proposed in a previous paper. In that contribution, we created a set of tools leveraging the open-source Orfeo ToolBox (OTB)<sup>3</sup> tool and we presented an end to end pipeline that

<sup>&</sup>lt;sup>1</sup> https://earthexplorer.usgs.gov/

<sup>&</sup>lt;sup>2</sup> https://www.usgs.gov/faqs/what-are-best-landsat-spectral-bands-use-my-research?qtnews science products=7#qt-news science products

<sup>&</sup>lt;sup>3</sup> https://www.orfeo-toolbox.org

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. The code assets created in that study where made available for others to reuse and can be found in https://github.com/agustingp/remoteSensingTimeSeries.

# 4 Methodology

During this study, we designed two experiments to assess the performance of CNNs for classifying land-use practices that can inform decision making to achieve a more sustainable agriculture. Below, we explain the aim of each experiment and describe the process for creating the labeled time-series dataset.

## 4.1 Experiment 1

The first experiment was designed aiming at classifying two agricultural practices: single cropping and double cropping, within the same year. Double cropping practice is an important sustainable practice that aims at reducing the fallow periods of the land, exploiting solar energy to enhance the quality of the soil and preventing soil erosion [17]. When introducing the so-called "Cover Crops" into annual crop rotations, double cropping has been shown to improve the provision of multiple ecosystem services in time, such as biomass production, N supply, soil C storage, NO3 retention, erosion control, weed suppression, and beneficial insect conservation [4].

For the experiment, we first filtered the ground truth data removing classes that were not representing agricultural fields. Then we identified the fields where double cropping practice and single cropping practice took place during the year. **Table 2** presents the total amount of polygons and the total amount of pixels sampled. For each pixel, a labeled time series was created using the three spectral bands from the satellite images. This process is explained in Section 3.4. In this case, as the double cropping pixels as the limit to be sampled from the single cropping polygons. The sampling was done randomly, maximizing the diversity of single cropping polygons and not exceeding the limit of 29596 pixels. In the end, the total number of pixels sampled was 29596 for both classes.

Class name	Total polygons	Total pixels
Double cropping	256	29596
Single cropping	6763	29596
Total	7019	59192

Table 2. Experiment 1 data details.

#### 4.2 Experiment 2

The second experiment consists of the classification of 20 different agricultural land use classes. During this practical experience, we focused on the classification of different crops that were grown in a "Single Cropping" approach. The rotation of crops across different years often leads to better yields due to soil fertility improvements [18], while also reducing the external dependency on agrochemicals [19, 20]. As we described in Section 1, crop rotations should follow a set of rules and criteria in order to be efficient. In this way, we aim to develop a model that classifies a comprehensive list of crops harnessing their temporal growing patterns.

For the experiment, we first filtered the ground truth data removing all the classes that did not represent an agricultural field. Following this, we removed the doublecropping practice ones, to focus on the fields where only one crop was grown during the year. From a list of 47 different crops (also including agricultural classes such as "fallow"), we selected the 20 classes that were best represented in terms of the number of pixels available. However, class imbalances are present in our dataset with the highest number of pixels for "Mixed Pasture" class, and the lowest number of pixels for "Grain Sorghum" class. **Table 3** presents the total amount of polygons and the total amount of pixels sampled. Finally, we followed the same criteria for time series creation as for Experiment 1.

Class name	Total polygons	Total pixels
Eucalyptus	77	5169
Walnuts	57	4812
Pears	418	23511
Almonds	13	2339
Cherries	85	5258
Safflower	65	7051
Corn	774	90805
Grain sorghum	17	2065
Sudan	35	4587
Beans	25	2775
Hay	26	2246
Rest	233	11489
Alfalfa	523	62949
Clover	38	4576
Mixed Pasture	1338	101675
Melons	50	2678
Potatoes	16	2854
Tomatoes	102	18203
Flowers	41	2485

Table 3. Experiment 2 data details.

6

Mixed (4+)	221	2758
Total	4154	360285

#### 4.3 CNN Architecture

In this study, we employed the CNN architecture proposed by Pelletier et al. [11] and implemented using Keras framework. **Fig. 2** depicts a general view of the architecture. For simplicity, we excluded from the diagram the Batch Normalization, Activation, and Dropout layers. This sequence is followed after each 1 Dimension Convolution and after the Dense layer, as well. **Table 4** presents a list of parameters and values used for the network configuration.



Fig. 2. CNN architecture.

Table 4. Architecture parameters.

Batch size	32	Loss function	categorical_crossentropy
Epochs	20	Kernel regulirizer	12
Optimizer	Adam	l2 rate	1.00E-06
Learning rate	0.001	Dropout rate	0.5
Beta_1	0.9	Kernel/filter size	5
Beta_2	0.999	Activation	ReLu

#### 4.4 Models Training

The models for both experiments were trained using the Azure cloud infrastructure provided by Microsoft AI for Earth grant program. The virtual machine uses an NVIDIA Tesla K80 GPU card. Each dataset created for Experiment 1 and Experiment 2, respectively, was split in two, 80% for the training set and 20% for the testing set. The training set was also split in runtime to separate some data for validation. We used a validation rate of 0.05, which means that 5% of the training set was used to validate the model performance during training. During each partitioning step, the

division of data was done at the polygon level, meaning that no pixels from the same polygon are in the training, testing or validation set at the same time.

## 5 Results

In this section, we present the results of the data preprocessing pipeline as well as the trained models' performance.

### 5.1 Time-series profiles

After the preprocessing process described in Section 3.4, we obtained 419,477 pixels' time-series. A sample of this time-series was plotted for both experiments. Fig. 3 presents the temporal and multispectral timeseries data for two different pixels, the left-hand side one corresponds to the "Single cropping" class, whereas the right-hand side corresponds to the "Double cropping" class. Fig. 4 presents this information for other two pixels classes, "Corn" and "Alfalfa", respectively.



**Fig. 3.** The chart on the left presents a pixel-level time series for the class "Single cropping". The right-hand side chart is a representation of a pixel-level time series for the class "Double cropping". In different colors each of the three different spectral bands: Red, Green, and NIR.



Fig. 4. The chart on the left-hand side represents a pixel-level time series for the class "Corn". The right-hand side chart represents a pixel-level time series for the class "Alfalfa". In different colors each of the three different spectral bands: Red, Green, and NIR.

#### 5.2 Models Evaluation

The overall classification accuracy for Experiment 1 was 88% and Experiment 2 was 89%. To achieve an extra performance verification of the model developed in Experiment 1, we used it to classify the entire dataset from Experiment 2. As we explained before, Experiment 2 dataset contains only pixels that belongs to single cropping practice. The results of the classification show 97% for this dataset. The highly accurate results can be explained because the model in Experiment 1 was trained with a sampling of pixels that maximized the diversity of polygons represented in the dataset. Then, most of the polygons in Experiment 2 where sampled in the dataset of Experiment 1. These results show that even the use of a small number of pixels coming from the same polygon are representative enough for the network to learn the specific time-patterns of the polygon.

In Fig. 5 and Fig. 6 we present the recall, precision and F-score for each experiment classes.

	Single cropping	Double cropping
Recall	0.975117	0.773945
Precision	0.835777	0.963454
<b>F-Score</b>	0.900086	0.858364

	Eucalyptus	Walnuts	Pears	Almonds	Cherries	Safflower	Corn	Grain sorghum	Sudan	Beans
Recall	0.972042	0.873541	0.977685	0	0.605696	0.940824	0.973111	0.35	0.814614	0.558704
Precision	0.981588	0.971861	0.927635	0	0.953448	0.685964	0.948505	0.96	0.858773	0.61745
F-Score	0.976792	0.920082	0.952003	0	0.740791	0.79343	0.96065	0.512977	0.836111	0.58661
	Hay	Rest	Alfalfa	Clover	Mixed Pasture	Melons	Potatoes	Tomatoes	Flowers	Mixed (4+)
Recall	Hay 0.230337	Rest 0.308461	Alfalfa 0.811866	<b>Clover</b> 0.966599	Mixed Pasture 0.979955	Melons 0.156609	Potatoes 0	0.899607	Flowers 0.972973	Mixed (4+) 0.670404
Recall Precision										. ,

Fig. 5. Experiment 1 statistics.

Fig. 6. Experiment 2 statistics.



Fig. 7. Experiment 1 confusion matrix.



Fig. 8. Experiment 2 confusion matrix

# 6 Discussion

The paucity of up-to-date ground truth data presents a problem for utilizing newer satellite imagery with a higher spatial, spectral and temporal resolution for training supervised AI models, as it is the European Sentinel constellation. On the other hand, initiatives like the Harmonized Landsat and Sentinel-2 [21] by NASA aiming at creating a virtual constellation of surface reflectance data coming from different satellites, should be strongly supported. Currently, this product is not available worldwide, creating another mismatch with ground truth data. Comparing our experiments with other

er related studies, ours have made use of publicly available satellite imagery, making a transfer learning approach that would make the process of fitting the models for other geographical locations, viable. We have also utilized convolution layers to learn temporal patterns from land-use dynamics. While some studies have only focused on the classification of a few agricultural types, we have trained a single model that is able to classify 20 agricultural classes with 89% accuracy. None of the studies analyzed before had classified pure temporal characteristics, as we did in Experiment 1. The confusion matrix for Experiment 2 shows that the network is making mistakes in classifying classes with similar characteristics. For instance, Grain Sorghum is being confused with Corn most of the times. The physiological and developmental similar characteristics between this two crops have been well documented [22] and can explain the network confusion between these two classes. In the case of Mixed Pastures, greatest confusion occurs with Alfalfa class, and vice versa (confusion of Alfalfa with Mixed Pastures). Alfalfa is a type of pasture, and the confusion can be explained because of a high concentration of Alfalfa in these pixels, or the presence of other types of pastures in Alfalfa fields, respectively. Future work includes finetuning the model parameters to improve precision and recall, extending the number of classes to learn, and create a more general architecture that is able to handle multiple years of data. Further work will also involve evaluation of the transferability of the models learned on the Sacramento data to different geographic locations.

# 7 Conclusion

In this paper, we presented a novel approach to processing satellite imagery validated in a set of two distinct experiments using Convolutional Neural Networks with convolutions in the temporal dimension. Most of the recent AI models or classifiers used in operational mapping use single date spectral data for classification and do not harness the temporal resolution of remotely sensed time-series images. Therefore, we argue that our solution provides an important contribution to the domain. Our evaluation showed promising results with our models achieving 88% and 89% accuracy for experiments 1 and 2 respectively. Our future work will focus on improving these results, covering more classes and with a more general architecture.

### References

1. Sahajpal R, Zhang X, Izaurralde RC, et al (2014) Identifying representative crop rotation patterns and grassland loss in the US Western Corn Belt. Comput Electron Agric 108:173–182. https://doi.org/10.1016/j.compag.2014.08.005

2. TWN (2015) Agroecology: Key Concepts and Practices

3. Doré T, Makowski D, Malézieux E, et al (2011) Facing up to the paradigm of ecological intensification in agronomy: Revisiting methods, concepts and knowledge. Eur J Agron 34:197–210. https://doi.org/10.1016/j.eja.2011.02.006

4. Schipanski ME, Barbercheck M, Douglas MR, et al (2014) A framework for evaluating ecosystem services provided by cover crops in agroecosystems. Agric Syst 125:12–22.

https://doi.org/10.1016/j.agsy.2013.11.004

5. De Baerdemaeker J (2013) Precision agriculture technology and robotics for good agricultural practices. IFAC

6. Bégué A, Arvor D, Lelong C, et al (2019) Agricultural Systems Studies using Remote Sensing

7. Cresson R (2018) A Framework for Remote Sensing Images Processing Using Deep Learning Techniques. IEEE Geosci Remote Sens Lett. https://doi.org/10.1109/LGRS.2018.2867949

8. Zhu XX, Tuia D, Mou L, et al (2017) Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. IEEE Geosci Remote Sens Mag 5:8–36. https://doi.org/10.1109/MGRS.2017.2762307

9. Gómez C, White JC, Wulder MA (2016) Optical remotely sensed time series data for land cover classification: A review. ISPRS J Photogramm Remote Sens 116:55–72. https://doi.org/10.1016/j.isprsjprs.2016.03.008

10. Zhong L, Hu L, Zhou H (2019) Deep learning based multi-temporal crop classification. Remote Sens Environ 221:430–443. https://doi.org/10.1016/j.rse.2018.11.032

11. Pelletier C, Webb GI, Petitjean F (2018) Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series. https://doi.org/10.3390/rs11050523

12. Cai Y, Guan K, Peng J, et al (2018) A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. Remote Sens Environ 210:35–47. https://doi.org/10.1016/j.rse.2018.02.045

13. Coppock R Agriculture in the Sacramento Region Trends and Prospects

14. FAO (2016) Handbook on remote sensing for agricultural statistics. https://doi.org/10.13140/RG.2.2.13259.69920

15. Landsat Science Products. https://www.usgs.gov/land-resources/nli/landsat/landsat-science-products. Accessed 14 Jun 2019

16. Flood N (2014) Continuity of reflectance data between landsat-7 ETM+ and landsat-8 OLI, for both top-of-atmosphere and surface reflectance: A study in the australian landscape. Remote Sens 6:7952–7970. https://doi.org/10.3390/rs6097952

17. Pinto P, Fernández Long ME, Piñeiro G (2017) Including cover crops during fallow periods for increasing ecosystem services: Is it possible in croplands of Southern South America? Agric Ecosyst Environ 248:48–57. https://doi.org/10.1016/j.agee.2017.07.028

18. Ellis BG, Hargrove WL, Pierce FJ, Rice CW (2013) Crop Rotation and Its Impact on Efficiency of Water and Nitrogen Use. https://doi.org/10.2134/asaspecpub51.c3

19. Sarandón SJ, Claudia Cecilia Flores (2014) Agroecología: bases teóricas para el diseño y manejo de Agroecosistemas sustentables. Universidad Nacional de La Plata – Editorial de la Universidad de La Plata

20. Dias T, Dukes A, Antunes PM (2014) Accounting for soil biotic effects on soil health and crop productivity in the design of crop rotations. J Sci Food Agric 95:447–454. https://doi.org/10.1002/jsfa.6565

21. Claverie M, Ju J, Masek JG, et al (2018) The Harmonized Landsat and Sentinel-2 surface reflectance data set. Remote Sens Environ 219:145–161. https://doi.org/10.1016/j.rse.2018.09.002

22. Assefa Y, Roozeboom K, Thompson C, et al (2014) Corn and Grain Sorghum Comparison: All Things Considered. Corn Grain Sorghum Comp All Things Consid 1–116