A Multi-source, End-to-End Solution for Tracking Climate Change Adaptation in Agriculture

Alejandro Coca-Castro
King's College London
alejandro.coca_castro@kcl.ac.uk

Aaron Golden
NUI Galway
aaron.golden@nuigalway.ie

Louis Reymondin
The Alliance of Bioversity International and the International Center for Tropical Agriculture
l.reymondin@cgiar.org

Abstract

The impact of climate change on tropical agri-food systems will depend on both the direction and magnitude of climate change, and the agricultural sector's adaptive capacity, the latter being affected by the chosen adaptation strategies. By extending SEIRS, a Satellite Remote Sensing (SRS) based system originally developed by the International Center for Tropical Agriculture - CIAT for monitoring U.S. Government-funded development programs across cropping areas in Africa, this research proposes the development and deployment of a scalable AI-based platform exploiting free-of-charge SRS data that will enable the agri-food sector to monitor a wide range of climate change adaptation (CCA) interventions in a timely, evidence-driven and comparable manner. The main contributions of the platform are i) ingesting and processing variety sources of SRS data with a considerable record (> 5 years) of vegetation greenness and precipitation (input data); ii) operating an end-to-end system by exploiting AI-based models suited to time series analysis such as Seq2Seq and Transformers; iii) providing customised proxies informing the success or failure of a given local CCA intervention(s).

1 Introduction

Adaptation in agriculture refers to the set of strategies used by individual farmers and land managers, sectors, industries and governments to minimise risk and reduce exposure to external perturbations [7]. Whilst there have been increases in global investment on climate change adaptation (CCA), these adaptation strategies can have both beneficial and unintended detrimental consequences when a wider context or longer time frames are considered [5]. Consequently, it remains key to the development of rigorous analytical methods to fully assess the impact of adaptation measures, particularly across countries in the developing world that are the most vulnerable to climate change.

Amongst the available data sources that would form the basis for such methodologies, Satellite Remote Sensing (SRS) data can be considered the most suited due to their large-area coverage and relatively easy access [4]. SRS instruments such as the Moderate Resolution Spectroradiometer (MODIS) and the Advanced Very High Resolution Radiometer (AVHRR) have been mainly used to assess the impacts of drought on crop productivity [2]. These particular instruments have served as early warning systems that fundamentally compare current vegetation indices (VIs, measurements of greenness) to the long-term average [4]. The normalized difference vegetation index (NDVI), the ratio between near infrared and red light within the electromagnetic spectrum, is one of the most widely used VIs.

Tackling Climate Change with Machine Learning workshop at NeurIPS 2020.

Following the plethora of SRS-based early warning systems for decision making, TAPAS "Tracking Adaptation Progress in Agriculture and Food Security Using an AI-powered Satellite Remote Sensing Platform" is a trans-disciplinary initiative aiming to enable countries to develop evidence-based means of measuring, reporting and verifying CCA in the agri-food sector with a particular focus on countries in the developing world that are the most vulnerable to climate change. TAPAS builds on the previous work of the System for Evaluation of Impact Using Remote Sensing (SEIRS) infrastructure. Since its inception, SEIRS has monitored U.S. Government-funded development programs particularly those involving interventions across cropping areas in Sub-Saharan Africa.

Contributions. We will plan to extend the SEIRS platform so that a) it ingests and processes multi-source satellite time series data b) it will utilise AI-based models that can handle corrupt and incomplete data. We also propose new NDVI-derived products (proxies of crop yield) suited to indicate the status of a local CCA intervention.

2 Tracking Adaptation in Agriculture Using Satellite Remote Sensing

2.1 Background and Related Work

SRS data streams have been key resources to monitor crop systems, their growing conditions, status, and agro-climatic conditions likely to impact them at national, regional and global scales [1]. Whilst SRS along with AI-based techniques have been widely explored to propose operational systems to measure and monitor the impact of climatic extremes e.g. droughts or floods on agri-food systems, few examples exist to inform and monitor the status of CCA strategies implemented on the ground.

According to the above premise, the following sections aim to introduce SEIRS and describe the main contributions derived from this system to build a new one suited to the needs of TAPAS, and in general to the climate change community.

The SEIRS Platform. SEIRS operates by modelling and analysing two SRS products, MODIS NDVI product (MOD13) and Global Precipitation Measurement (GPM), to predict differences in vegetation greenness according to computer model generated synthetic data and observed data. Five major steps are involved (see Figure 1) i) delimit cropping areas according to a crop mask extracted from existing land cover maps ii) reconstruct MODIS raw NDVI time series by using the Harmonic Analysis of NDVI Time Series (HANTS) algorithm [8]; iii) model both reconstructed NDVI and precipitation data by using a Convolutional Neural Network (CNN) architecture; iv) evaluate the model according to generated synthetic data and orbit-derived observed data; and v) provide sets of NDVI-derived products suited to the analysis of U.S. Government-funded development programs.

Following the logic behind SEIRS, the proposed system aims to contribute in three main aspects i) explore multi-source input data (Section 2.2), ii) implement new modelling methods suited to extract information from raw and noise time series (Section 2.3), and iii) design customised NDVI-derived products (Section 2.4) according to the need of TAPAS.

2.2 The need for Multi-source Input Data

Whilst MODIS data provides consistent and regular observations (every 16-days), other SRS instruments such as Landsat and Sentinel can be complementary. These particular instruments allow capturing with more accuracy changes in vegetation greenness across small-scale farming systems. Moreover, MODIS instruments have exceeded their design lifetime. In this regard, by considering all possible SRS products in which vegetation greenness can be derived such as NDVI is expected to increase the robustness of the TAPAS platform. Table 1 depicts the proposed input SRS to be considered.

2.3 Exploring Sequential and Transformer Architectures

Under SEIRS, using only historical, pre-development intervention SRS data, a Convolutional Neural Network (CNN) was built to learn vegetation patterns and their relationship to growing season conditions. The CNN model predicts an entire stream of synthetic vegetation data (23 MODIS observations or images per year), for hundreds of thousands of agricultural fields, with each individual

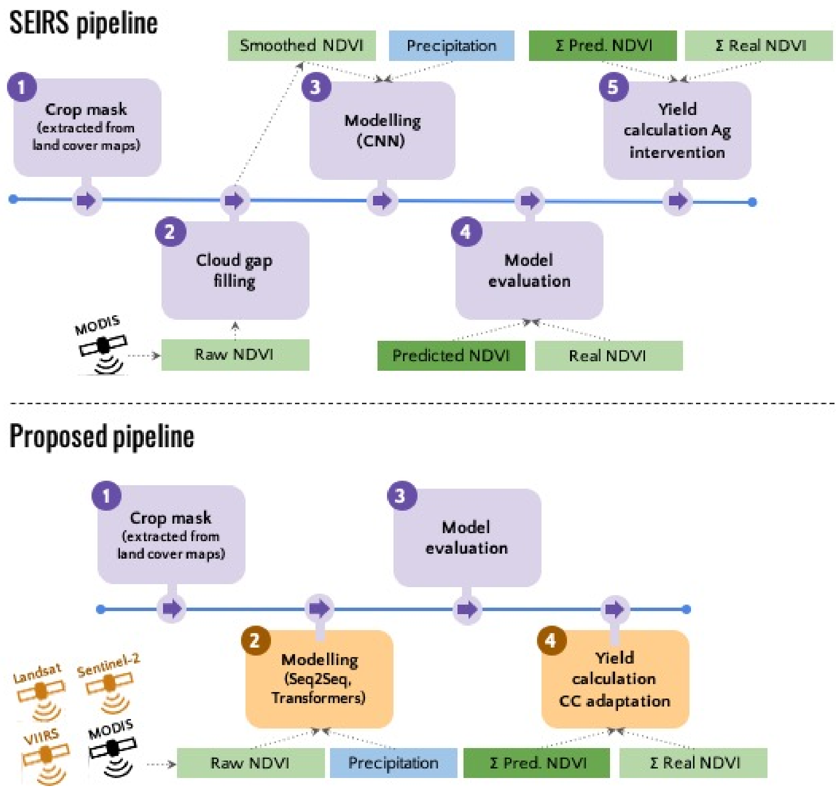

Figure 1: Comparison of key steps in SEIRS (*Top*) and proposed implementation for TAPAS (*Bottom*). Those proposed components different to SEIRS are highlighted in a different colour.

Table 1: Input SRS products to derive vegetation greeness.

Product	Ground sample distance	Temporal coverage	Frequency
MODIS (MOD13Q1)	250-m	2000-to present[1]	16-day composite
VIIRS (VNP13)VNP13	500-m	2012-to present	16-day composite
Landsat legacy (1-8)	30-m (red, NIR)	1975-to present	Revisit, 8-day
Sentinel-2	10-m (red, NIR)	2015-to present	Revisit, 10-day

[1] The MODIS data will be discontinued as the instrument has exceeded their design lifetime (6 years).

prediction adjusted to current and preceding growing conditions. Whilst SEIRS' CNN model has showed reasonable results, it has some limitations that reduce its feasibility to fully operate on an end-to-end fashion in that it requires some level of input data preprocessing to handle irregular time-series observations and variable sequence length as Landsat and Sentinel-2 are.

To tackle the above limitations seen in SEIRS, we are aiming to explore and assess more sophisticated methods suited to analyse long-term SRS time series data with minimal preprocessing and region-specific expert knowledge. Following the previous work by Rußwurm and Körner [10], who assessed multiple state-of-art traditional and deep neural network methods for classifying multi-temporal SRS data, the following approaches are proposed:

Recurrent approach: In contrast with CNNs, which contain different types of layers that perform different functions (e.g., convolutional, pooling and nonlinear layers), Recurrent Neural Networks (RNNs) architectures consist only of recurrent layers. Basic RNNs are prone to vanishing and exploding gradients through time, inhibiting the extraction of features from long-term temporal contexts. To overcome such gradient issues, specialized memory units using a gate-based system were proposed, initially Long Short-Term Memory (LSTM) memory cells, and later, Gated Recurrent Units

(GRUs). These adapted RNNs are commonly used in encode-decoder architectures for generative prediction of words. For remote sensing applications, the encoder model has been utilized mainly for classification [9], regression [3] and change detection [6] tasks.

Attention-Based Approach: Following the adoption of self-attention in the NLP literature as an efficient alternative to RNNs, Rußwurm and Körner [10] proposed to apply a self-attention Transformer network to pixel-based classification using Sentinel-2 time series data. The authors demonstrated the Transformer network excels and presents the same robustness to handle noisy observations (e.g. clouds) as RNN-based models. We, therefore, propose to extend self-attention mechanisms for NDVI forecasting.

2.4 Crop Yield Proxies for Tracking Climate Change Adaptation

For SEIRS, a series of NDVI-derived products were developed according to the needs of a given project, in in particular, to identify known areas where agriculture interventions had resulted in clear deviations from what would have happened had no intervention taken place. With TAPAS, models will be again trained to predict normal NDVI patterns of behaviour over time, with significantly enhanced precision. The resulting differences between predicted and observed NDVI estimates obtained from SRS will be used again to flag locations where external factors clearly have perturbed normal growth cycles. However, having previously studied the properties of NDVI variations at curated geospatial locations for which associated crop estimates have been simultaneously compiled, this information will be used to interrogate the TAPAS predicted/counterfactual data space, and in so doing, place constraints on predicted crop yield/losses at those locations associated with the most extreme departures from baseline NDVI temporal behaviour. In this way not only would specific regions be flagged - as was the case with SEIRS - but an in addition, an estimate of the likely crop yield/loss associated with the observed perturbation at those locations, with an associated monetary cost/gain. The ability to track changes in crop yield as determined from counterfactual analysis of both predicted and observed NDVI timeseries using TAPAS offers a powerful, tractable, evidence-based proxy for assessing climate change adaptation going forward.

3 Experimental setup and anticipated results

3.1 Study area

The project will initially target the Senegal River Valley for which previous benchmark work has been implemented using SEIRS. This region has multiple governmental and private CCA strategies have taken place and where the TAPAS project has access to comprehensive curated ground thruth data. The initial plan is to run multiple extensive experiments across this area mainly to show the drawbacks and/or benefits of the proposed adaptations in comparison with the baseline (SEIRS).

3.2 Implementation

The whole framework of SEIRS, from data acquisition to model production, was programmed in Java. The CNN models were deployed in the Eclipse Deeplearning4j library. This particular library also offers a wide variety of models, including those targeted in this proposal, Seq2Seq and Transformers. Therefore, the same language and library will be used and experiments will be implemented by using Azure cloud credits granted through the TAPAS project.

4 Conclusions

This proposal highlights the role of SRS data along with AI-based methods can play as powerful tools to enable the agri-food sector to monitor a wide range of climate change adaptation (CCA) interventions in a timely, evidence-driven and comparable manner. Building on previous work, it is expected that this proposal will contribute to the global climate change community's efforts with a multi-source, end-to-end solution for tracking CCA in agriculture, in particular following the needs captured by the TAPAS initiative. When climate change professionals can tie pixels to their programs, these differences can also offer insights into where, or to what extent, specific CCA programs or policies might be having an effect across multiple adaptation sites in agriculture settings.

Acknowledgments. The authors acknowledge Dr Yannis Kalantidis who kindly assisted and guided writing this proposal as part of the NeurIPS 2020 Workshop Tackling Climate Change with Machine Learning mentorship program.

References

[1] Inbal Becker-Reshef, Brian Barker, Michael Humber, Estefania Puricelli, Antonio Sanchez, Ritvik Sahajpal, Katie McGaughey, Christopher Justice, Bettina Baruth, Bingfang Wu, Adam Prakash, Abbassian Abdolreza, and Ian Jarvis. The GEOGLAM crop monitor for AMIS: Assessing crop conditions in the context of global markets. *Global Food Security*, 2019.

[2] Lingtong Du, Qingjiu Tian, Tao Yu, Qingyan Meng, Tamas Jancso, Peter Udvardy, and Yan Huang. A comprehensive drought monitoring method integrating MODIS and TRMM data. *International Journal of Applied Earth Observation and Geoinformation*, 2013.

[3] Natalia Efremova, Dmitry Zausaev, and Gleb Antipov. Prediction of soil moisture content based on satellite data and sequence-to-sequence networks, 2019.

[4] Emmanuel Letouzé, Patrick Vinck, Bessie Schwarz, Simone Sala, David Sangokoya, and Tellman Tellman. DATA-POP ALLIANCE Big Data for Climate Resilience Inputs for World Development Report 2016 Big Data for Climate Resilience. Technical Report October 2015, Data-pop Alliance, 2015.

[5] Achut Manandhar, Alex Fischer, David J. Bradley, Mashfiqus Salehin, M. Sirajul Islam, Rob Hope, and David A. Clifton. Machine Learning to Evaluate Impacts of Flood Protection in Bangladesh, 1983–2014. *Water*, 12(2):483, feb 2020.

[6] Lichao Mou, Lorenzo Bruzzone, and Xiao Xiang Zhu. Learning spectral-spatialoral features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 2019.

[7] Gigi Owen. What makes climate change adaptation effective? a systematic review of the literature. *Global Environmental Change*, 62:102071, 2020.

[8] G. J. Roerink, M. Menenti, and W. Verhoef. Reconstructing cloudfree NDVI composites using Fourier analysis of time series. *International Journal of Remote Sensing*, 2000.

[9] Marc Rußwurm and Marco K rner. Multi-temporal land cover classification with sequential recurrent encoders. *ISPRS International Journal of Geo-Information*, 2018.

[10] Marc Rußwurm and Marco Körner. Self-attention for raw optical satellite time series classification, 2020.