Article

Not peer-reviewed version

Machine Learning as a Strategic Tool for Helping Cocoa Farmers in Côte d'Ivoire

Stefano Ferraris [‡] , Rosa Meo [*,‡] , Stefano Pinardi [‡] , Matteo Salis , Gabriele Sartor

Posted Date: 4 August 2023

doi: 10.20944/preprints202308.0416.v1

Keywords: cocoa farmers; low-cost smart agriculture; remote sensors monitoring; water resources forecasting; YOLO; U-NET; deforestation; drought prevision; socio-technical transition

 doi:10.20944/preprints202308.0416.v1

Article

Machine Learning as a Strategic Tool for Helping Cocoa Farmers in Côte d'Ivoire

Stefano Ferraris [1,†,‡], Rosa Meo [2,†,‡], Stefano Pinardi [3,‡], Matteo Salis [2] and Gabriele Sartor [2,*]

[1] Interuniversity Department of Regional and Urban Studies and Planning, Politecnico di Torino and University of Turin, 10125 Turin, Italy; stefano.ferraris@polito.it

[2] Department of Computer Science, University of Turin, 10149 Turin, Italy; rosa.meo@unito.it, gabriele.sartor@unito.it, matteo.salis@unito.it

[3] Department of Foreign Languages, Literatures and Modern Cultures, University of Turin, 10124 Turin, Italy; stefano.pinardi@unito.it

* Correspondence: rosa.meo@unito.it

‡ Meo and Pinardi had the original idea. Salis and Sartor performed all the experiments. All the authors contributed with discussions to the writing of the article.

Abstract: Machine Learning can be used for social good. In this paper, we discuss how it can be used to combat climate change and facilitate land management and farming in developing countries and in particular in Côte d'Ivoire. This paper explores models that improve land and water management and agricultural farming cultivation to contrast climate change. Côte d'Ivoire is the largest producer of cocoa beans (43%) in the world, but deforestation, lack of rainfall, drought, and climate change threaten crops and the already fragile economy of Ivorian farmers. It is important to combat climate change with methods and techniques that are affordable to the local farmers and also induce positive effects in production. We discuss the use of low-cost sensors to collect data on the soil and open data and open source software to develop AI tools. We show that using deep neural networks (YOLOv5m) is effective for detecting healthy plants and pods of cocoa from damaged ones only using mobile phone images. Focusing on a single land is not enough to combat climate change, which has different causes and involves also knowledge at a higher scale. We propose a new method of forecasting for the analysis of remote sensors. Remote sensor data come from GRACE NASA Mission and ERA5 produced by the Copernicus Climate Change Service at ECMWF. We implement a new deep neural network architecture named CIWA-net. It is based on a Fully Convolutional Neural Network (FCN) [1] and it is a U-net like architecture [2]. The aim of CIWA-net is to forecast Total Water Storage Anomalies (TWSA). We show the quality of our model with a comparison to a *vanilla* Convolutional Neural Network. CIWA-net could be used also for the detection of lands that interfere with agricultural work and yields, such as deserted areas, water-soaking soil areas, zones at risk of desertification, and poor land use. The employment of AI at the service of agriculture can decrease crop losses and waste, lower the inputs onto the soil of fertilizers, responsible for the increase of Greenhouse Gases. It could be useful to help the small farmers (at a local scale) and also the policy-makers and farmers' cooperatives (at the regional scale) to take valid and coordinated countermeasures to improve the correct use of the lands, helping to contrast and adapt to climate change.

Keywords: cocoa farmers; low-cost smart agriculture; remote sensors monitoring; water resources forecasting; YOLO; U-NET; deforestation; drought prevision; socio-technical transition

---

1. Introduction

The Anthropocene is the geological era that began when human activities have had a global and evident effect on the ecosphere of lands, oceans, and water, all over the world. Human actions are driving anthropological climate change (ACC) which is an effect of the human socio-economical activities.

 doi:10.20944/preprints202308.0416.v1

1.1. The complex system that causes and is implied in the ACC

Climate change includes global warming: Earth's weather patterns are caused by the emission of greenhouse gases (GHGs), carbon dioxide (CO2), and methane which are mostly emitted from the use of fossil fuels for energy use. Also, agricultural practices and forest loss are relevant additional inducers of climate change. 23% of the GHGs are produced by agriculture farming, coastal economies, and the erroneous management of forests and land.

The 2019 UN "Climate Change and Land Report" [3] draft by the Intergovernmental Panel on Climate Change (IPCC) states:

- The ACC will increase drought in some areas and extreme rainfall in other areas of the world, affecting agricultural production, the ocean economy, and the security of food supplies around the world;
- it will have effects on the water level by deeply involving the coastal areas and their cities;
- it will raise the temperature averages 1-4 degrees upward, "shifting" warm climatic zones northward, altering marine habitat, and coastal economies, changing land needs and local habits, inducing tropical rains, dramatically reducing the extent of glaciers, lake levels, and the natural water reserves;
- it will have an effect in terms of the transmigration of animals and insects to areas where they were not normally present.

The IPCC 2023 literally states that human activities, mainly through emissions of greenhouse gases, "have unequivocally caused global warming, with the global average surface temperature in 2011–2020 reaching 1.1°C above the average temperature in 1850–1900" [4]. Global greenhouse gas emissions have continued to increase, with unequal historical and ongoing contributions arising from unsustainable energy use, land use and land-use change, lifestyles, and patterns of consumption and production across regions, between and within countries, and among individuals.

Heat waves, intense storms, and weather extremes are other visible effects of the ACC. It affects everyone and any area on the planet, no one can consider themselves exempt, but the consequences will weigh most heavily on the weakest and most vulnerable populations and areas. The coastal zone is considered at risk of flooding, intense storm, and rising water levels. The Mediterranean is at high risk of desertification and weather extremes. The savage intensive agriculture in Africa is destroying the forests with the effects of desertification of entire regions (Côte d'Ivoire has used 90% of their forest to produce Cocoa, mainly exported to western countries). This area will be subject to relevant climate changes with severe economic consequences: wars or mass migrations are expected with effects on the economy and lifestyle. Econometric models indicate that the ACC has reduced global agricultural Total Factor Productivity (TFP) by about 21% since 1961, a slowdown that is equivalent to losing seven years of productivity growth [5,6]. The effect is substantially more severe in warmer regions, like Africa, Latin America, and Asia where it is expected a reduction of ~26-34% [5]. The agriculture and ocean food industries have grown more vulnerable to ongoing climate change [4].

There are three categories of causes for GHGs:

- the anthropogenic activity that changes the land cover and land management;
- indirect effects of anthropogenic activity, such as carbon dioxide (CO2), fertilization, and nitrogen deposition;
- natural climate variability and natural disturbances (e.g., wildfires, windrow, disease).

Deforestation makes cultivation areas more exposed. It decreases the variety of species, weakens the ecosystem, and makes soils drier and more arid, pushing local farmers to compensate by using even more fertilizers and the already scarce waters. These significantly increase GHGs, water dispersion, and the exploitation of the territory.

 doi:10.20944/preprints202308.0416.v1

1.2. The motivations of the case study

Côte d'Ivoire is the largest producer of cocoa beans in the world accounting for 43% of the global production [7]. However, the working conditions in the region are far from ideal, with farmers earning about 1.00 Euro per day. The combination of deforestation, inadequate rainfall, the drain of underground waters, and the impacts of climate change poses significant threats to crops and the already fragile cocoa economy of Ivorian farmers. To address this pressing issue in a developing country, it is crucial to combat climate change using methods and techniques that are affordable and accessible to the local economy.

Modeling terrestrial water content is crucial for pursuing Sustainable Development Goals issued by the 2030 Agenda of the United Nations, especially for the 6^{th} Goal: Ensure Availability and Sustainable Management of Water and Sanitation for all. Extreme hydrological events (i.e., droughts and floods) have a severe impact on a region, making environmental resource policies very relevant to the development of a country. All over the world over the last decade, nearly 1.43 billion people have been affected by droughts and .65 billion by floods [8]. Furthermore, global warming is going to make droughts more likely and severe [9]. Especially for Côte d'Ivoire, the likelihood of occurrence of droughts will increase by 7.5% in the future (2051-2100) and 6.6 million people (22% of the population) will be exposed every year to these events [10]. Of course, the agricultural sector will suffer from this scenario: it implies catastrophic consequences on the already fragile economy and society of the country because agriculture is one of its driving economic sectors[1].

Many studies already analyzed the distribution of rainfall in space and time in Côte d'Ivoire and they agree on a decreasing trend below the usual average causing water shortage [12–14]. This makes Côte d'Ivoire an area highly vulnerable [12,15]. For these reasons, it is essential to develop supporting instruments for mitigation and adaptation policies. However, in many Developing Countries, there could be a lack of resources to build such instruments, for example, a collaborative network for monitoring surface water or groundwater [8]. In this context, open data and Machine Learning techniques can provide support to the sustainable management of environmental resources. For example, Gravity Recovery And Climate Experiment (GRACE) NASA mission data [16,17] inform about the Total Water Storage Anomalies (TWSA) which could be used for determining the water shortage period, the anthropogenic drought, and the water resource depletion, in general, at a regional or catchment level [18–20].

Terrain sensors, satellite data analysis, neural models, and digitalization can improve the total factor productivity, by producing more with fewer inputs (less fertilizers, water, energy, capitals [21]) reducing land misuse, deforestation, and GHGs. The research community must use the capabilities at its disposal to introduce instruments of control and methods of continuous monitoring and techniques of forecasting that will help small farmers to produce more with a reduced environmental impact. Also, it is relevant to give the political decision-makers the best information and a set of tools to control and fight climate change at a larger level. Finally, a better perception of the green impacts of their production can induce the local farmers to follow more correct habits and spread new agronomic approaches more ecologically attentive.

1.3. Technology and Machine Learning methods at the service of the problem

Low-cost sensors, remote sensing, Machine Learning methods such as predictive models and more specifically deep neural networks, and cooperative approaches can make an enormous contribution to fighting climate change in agriculture in developing countries. The research must be addressed affording the problem as a global phenomenon that has multiple causes. Collaboration among experts, researchers from different areas, non-governmental organizations (NGOs), and local authorities is

---

1 Agriculture accounts for 22% of the Gross Domestic Product (GDP), about $50 - 70\%$ of the total export earning and employs nearly 50% of the labor force [11]

 doi:10.20944/preprints202308.0416.v1

essential to bring about lasting changes. The ability to manage territories, soil, waters, crops, and countering extreme weather, also depends on how we act within communities that are part of the projects and may have limited economic resources.

This article illustrates a method of crop monitoring to support the identification of cocoa health status with low-cost imagery to help farmers in their production and increase their efficiency and yield. It then shows a new model for TWSA forecasting to identify water availability increase or reduction in dry areas at a regional scale, which can help to implement policies of intervention to prevent phenomena of desertification and land management. We also discuss how the same methods could be useful to local policymakers, and cooperatives of farmers, to identify land that is transforming without planning or control, leading to forest depletion, deforestation, and land misuse.

Counteracting climate change requires working concurrently at different scales, long and short. The short scale entails working at the local level, employing terrain sensors and images taken from mobile devices to help single farmers' crop activities (see task 1). The long scale is at the national and sub-national level and consists of using satellite data (NASA's GRACE mission [16] and ERA5 Copernicus Climate Change Service produced at the European Center Medium Weather Forecast - ECMWF [22]) to detect changes and depletion of water resources in the terrestrial ecosystem at the 0.25° scale, harnessing and soliciting a cooperative approach among small farmers and policymakers (see task 2).

With the results, we target the local smallholder cocoa producers [23] the NGOs operating in the Abidjan area (e.g. Communauté Abel, Grand Bassam, Côte d'Ivoire [24] connected with the Gruppo Abele Foundation's Choco+ initiative [25]) working with the local farmers and the local authorities in the context of Euro-African cooperation (cf. Pinardi et al. 2023 paragraph 3.2 [26]).

2. Related works in Smart Agriculture and Terrain Monitoring

Cocoa plantations in Côte d'Ivoire are one of the main drivers of degradation and deforestation [27,28]. For this reason, tools for detecting the depletion of forestry resources are highly needed and relevant to implement policies for sustainable management. Some studies already tried to detect cocoa plantations in forest areas in Côte d'Ivoire [29,30]; and some global tools already exist to detect tree losses or Tropical Moist Forest (TMF) degradation and deforestation [31,32]. In Africa, several factors contribute to GHGs emissions. Deforestation, driven by activities like logging and the expansion of agricultural lands, releases significant amounts of CO2 into the atmosphere [33,34]. Additionally, the agricultural sector emits methane – a potent GHG [34,35]. Furthermore, the use of synthetic fertilizers in agriculture and waste management practices release N2O, another powerful GHGs [34,35]. Africa contributed 11% of GHG emissions growth since 1990 (2.3 GtCO2-eq) and 10% (0.7 GtCO2-eq) since 2010 [36]. Vast tropical rainforests and other ecosystems play a crucial role in mitigating the impacts of GHGs emissions. Unfortunately, deforestation and land degradation reduce their capacity [33,34,37].

Fundamental research is essential to explore new hypotheses and gain a deeper understanding of ecological processes and their interactions [37].

Also, the determination of reasonable scales and the selection of appropriate explanatory and response variables is dependent on an understanding of the context and systems under study [38]. Fortunately, some initiatives to analyze these phenomena already exist. For instance, ECMWF gives access to the reanalysis data set of atmospheric composition (AC) produced by the Copernicus Atmosphere Monitoring Service (CAMS): data are currently covering the period 2003-June 2022 [39] with a resolution of approximately 80 km with a sub-daily and monthly frequency. The separate CAMS global greenhouse gas reanalysis (EGG4) currently covers the period 2003-2020. In Section 5 we discuss a data analysis of ERA5, from the Copernicus reanalysis database [22].

Today it is widespread the idea that AI can be employed in Precision Agriculture [40,41] where operators mainly follow a cost-benefit analysis, focusing on ROI (return on investment) [42]. Little emphasis is placed on the social impacts of the processes. Instead, our concern is how the transformation can take into account the social dimensions that may influence development. In this

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 August 2023 doi:10.20944/preprints202308.0416.v1

context, it is also important to address the digital costs, which are particularly relevant when working in developing countries, such as Côte d'Ivoire, where, according to the World Economic Forum [43], farmer wages average around 1.00 Euro per day, and these workers contribute significantly to national domestic production.

To address the economic constraint, we have adopted a cost-effective approach by leveraging open-source software and accessing freely available data from sources such as GRACE, and ERA5. Furthermore, we have implemented low-cost ground sensors (see 3.1). Taking advantage of the widespread use of mobile phones in Côte d'Ivoire, farmers capture close-up images of cocoa plants and pods to assess the plants' health. This method allows for easy identification of health issues. By employing open-source methods, we can significantly reduce the financial burden. This approach ensures that even in regions with limited financial resources, it is possible to give a meaningful contribution to environmental conservation and sustainable agricultural practices. This calls for equal attention to these constraints and opens up a new area of research related to socio-technical transition [26], where machine learning and sensor monitoring are not only instruments to increment yield production, or to generate new business or markets [42] but are pivotal to changing the social and economical landscape of an entire region.

With these constraints in mind, the article focuses on YOLOv5m [44] for the identification of healthy/unhealthy cocoa pod (cf. 4), Convolutional Neural Network (CNN), Fully Convolutional Neural Network (FCN) [1] U-net like architecture [2] for remote sensor analysis for water resources monitoring (cf. 5) and land management (cf. section future works 7) in the context of a social-technical transition [26].

3. Open-Source Strategic Tools

An open-source approach to precision agriculture and more in general big data management and stream data processing for prediction comes with a large potential for innovation capacity thanks to the ability to freely reuse the software under open-source licenses [45].

The capability to deploy existing technology, digital platforms, and open data collections facilitates innovation by leveraging innovative services and organizational models. In particular, the evolution of open-source software communities that support software development, sharing, and reuse (like Github [46], Joinup [47], Apache [48], GNU and the Free Software Foundation [49]) increase the diversity of potential users, that act as testers and expert sources. For businesses that have less expertise in programming, open-source offers visibility into how developers manage datasets and software and helps them to cut costs because they do not need an in-house software development company. Open-source platforms, communities, and initiatives provide accessibility to a myriad of AI tools, libraries, and documentation that facilitate the development of new tools in different business sectors, using AI techniques while they are receiving feedback from experts within the community.

Open collaboration environments and open-source software (OSS) are the results of collaborative projects that speed up the reproducibility of the research (like Wikipedia and Open data repositories, such as the University of California Irvine [50] and Kaggle [51]). In turn, the presence of large volumes of data and pre-trained models accelerated advancements in deep learning with software libraries like TensorFlow [52], DSSTNE (Deep Scalable Sparse Tensor Network Engine) [53] and Keras [54]. Other frameworks are based on High-Performance Computing [55,56] and cloud environments with virtual machines and software containers, like ML-Ops [57]: they assist the developer and machine learning expert in the pipeline of data analysis, provide memory space and computing power for storage and computing resources. For specific applications such as image recognition, there are web services that let users perform deep learning and even prediction without programming [58,59].

Finally, open data licenses such as Community Data License Agreement (CDLA) have begun to *commoditize* training data. These license terms will help "democratize" the overall AI marketplace by lowering the barriers to entry in the market of AI. Proprietary datasets could continue to exist, but in

 doi:10.20944/preprints202308.0416.v1

two versions (one of them under the CDLA license) and this solution could allow everyone, including smaller players, to build credible products [60].

3.1. Pervasive IoT systems

The implementation of pervasive IoT systems needs components that communicate data and perform efficient data ingestion from a stream coming from sensors. In these tasks, many open source software projects exist that allow data ingestion, like Apache Kafka [61], and data storage in specialized databases for time series such as InfluxDB [62].

Among the communication protocols in a distributed system, there are many protocols with different characteristics. There is Message Queue Telemetry Transport (MQTT), which is lightweight and can work in very low bandwidth networks, HyperText Transfer Protocol (HTTP) and WebSocket which establish a TCP connection and are based on the request-response scheme. Constrained Application Protocol (CoAP) is based on a web transfer protocol and should be used with limited networks, with low bandwidth and low availability. Data Distribution Service (DDS) adopts a publish-subscribe methodology, Advanced Message Queue Protocol (AMQP) is TCP-based and guarantees delivery and acknowledgment and has two levels of quality of service. Extensible Messaging and Presence Protocol (XMPP) are based on Extensible Markup Language (XML). OPC Unified Architecture (OPC UA) is a transport-agnostic protocol and supports both request/response and publish/subscribe methods.

As regards sensors, emitting the measured data, we refer to the open sensors. Arduino sensors play a vital role in modern agriculture, enabling farmers to monitor and optimize various aspects of their crops and environment. These sensors are relatively affordable, easy to use, and can be integrated into automated systems. Here are some key applications of Arduino sensors in agriculture:

Soil Moisture Sensors: These sensors measure the moisture content in the soil, allowing farmers to determine the optimal time for irrigation. By ensuring the right amount of water is provided to the plants, farmers can prevent overwatering or under watering, leading to better crop yield and water conservation [63].
Temperature and Humidity Sensors: Monitoring temperature and humidity levels is crucial for crop health. Arduino's sensors can help farmers assess the environmental conditions and make adjustments accordingly, such as turning on irrigation systems or activating ventilation in greenhouses [64].
Light Sensors: Light sensors help farmers analyze the intensity of sunlight reaching the crops. This information is valuable in determining suitable planting locations, optimizing crop layouts, and even deciding the best time for harvesting [65].
Weather Stations: Arduino-based weather stations can collect data on various weather parameters such as temperature, humidity, wind speed, and precipitation. Farmers can use this data to anticipate weather changes and prepare for potential adverse conditions [66].
Crop Health Monitoring: Sensors like pH sensors and nutrient level sensors can provide insights into the health of the crops and soil. Farmers can adjust fertilization and nutrient application based on real-time data, leading to healthier plants and better yields [67].
Pest Detection: Some Arduino sensors can identify pests and diseases early on by detecting specific patterns or changes in the environment caused by these issues. This helps farmers implement targeted pest control measures, reducing the need for excessive pesticide use [68].
Automated Irrigation Systems: By integrating Arduino sensors with irrigation systems, farmers can create automated setups that respond to real-time data. These systems can turn on or off the irrigation based on soil moisture levels, weather conditions, and crop requirements [69].
Crop Growth Monitoring: Sensors like ultrasonic distance sensors or infrared sensors can measure crop height and growth rate. This information allows farmers to track the development of their crops and make timely decisions regarding pruning or harvesting [67].
Livestock Monitoring: In addition to crop-related applications, Arduino sensors can also be used to monitor the health and behavior of livestock. For example, sensors can track the body temperature of animals, detect estrus in cattle, or monitor feeding and drinking habits [70].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 August 2023 doi:10.20944/preprints202308.0416.v1

Automated Greenhouse Systems: Arduino sensors can be integrated into smart greenhouse systems, controlling temperature, humidity, and ventilation automatically to create an optimal environment for plant growth [71].

Overall, Arduino sensors offer an affordable and accessible way for farmers to gather valuable data, optimize their farming practices, and make informed decisions to enhance productivity and sustainability in agriculture.

4. Task 1: cocoa pods classification model

To guarantee good quality beans, cocoa plants have to be continuously monitored from the ripening to the harvesting phase. The accurate selection of cocoa beans is a crucial undertaking that has a significant impact on the subsequent activities and, consequently, on the final product's quality. Recognizing ripeness and the absence of anomalies in the bean is still a manual activity and assumes the presence of expert operators employed in the field. Recently, to provide partial support to operators, some AI tools have been proposed [72–74]. In this paper, we suggest a possible approach to support farmers to recognize good-quality cocoa beans using state-of-the-art AI tools, such as the neural net architecture YOLO [44].

4.1. Data

In our setting, we employed an open-source labeled dateset[2] identifying healthy and damaged beans affected by Monilia and Phytophthora diseases, in three different classes (Healthy, Monilla, Fito). More specifically, we have 312 images: 107 Fito, 105 Monilla, and 100 Healthy. These data are already labeled and ready to use, then no pre-processing steps are necessary.

4.2. Model

For this task, we adopt the pre-trained YOLOv5m model[3], an improvement of the original YOLO model [44]. We train the last layer of the model on the Cocoa Diseases dataset using Google Colab[4]. Previous systems like Region-based CNN (R-CNN) first generate potential bounding boxes and then run a classifier for detecting objects inside images. However, post-processing is needed to eliminate duplicates and produce a valuable output. YOLO approach is different: it is made by a unique Convolutional Neural Network (CNN) that simultaneously predicts different bounding boxes and the associated probabilities. YOLO does not look at inputs locally but globally, and it uses information from the entire image for each simultaneous box prediction; not by chance, YOLO stands for You Look Only Once. This structure makes YOLO very efficient and effective such that it can be used for object detection on smartphone applications [75–77].

4.3. Results

We trained the model on the training-set for 50 epochs using a batch size of 16 images. Figure 2 shows the loss reduction on the training set with the number of epochs. While Figure 1 shows the confusion matrix on test-set with 5-fold cross validation. It represents the performance of the implemented models in discriminating different classes. Our model detects almost every time (98%) a healthy fruit, while it has some difficulties in differentiating between diseases. A more complex model could be implemented to discriminate better. However, given we are proving the feasibility of an instrument that should be used on smartphones, it could be worth giving up a little bit of performance in favor of a faster and lighter model. Furthermore, detecting a specific disease could be difficult also

---

[2] Contains information from https://www.kaggle.com/datasets/serranosebas/enfermedades-cacao-yolov4, which is made available here under the Open Database License (ODbL).
[3] https://docs.ultralytics.com/yolov5/
[4] https://colab.research.google.com/

for experts. A free, fast, and open tool that could tell farmers if fruits are healthy or not, could be already a tremendous help for helping in crop monitoring and yield forecasting.

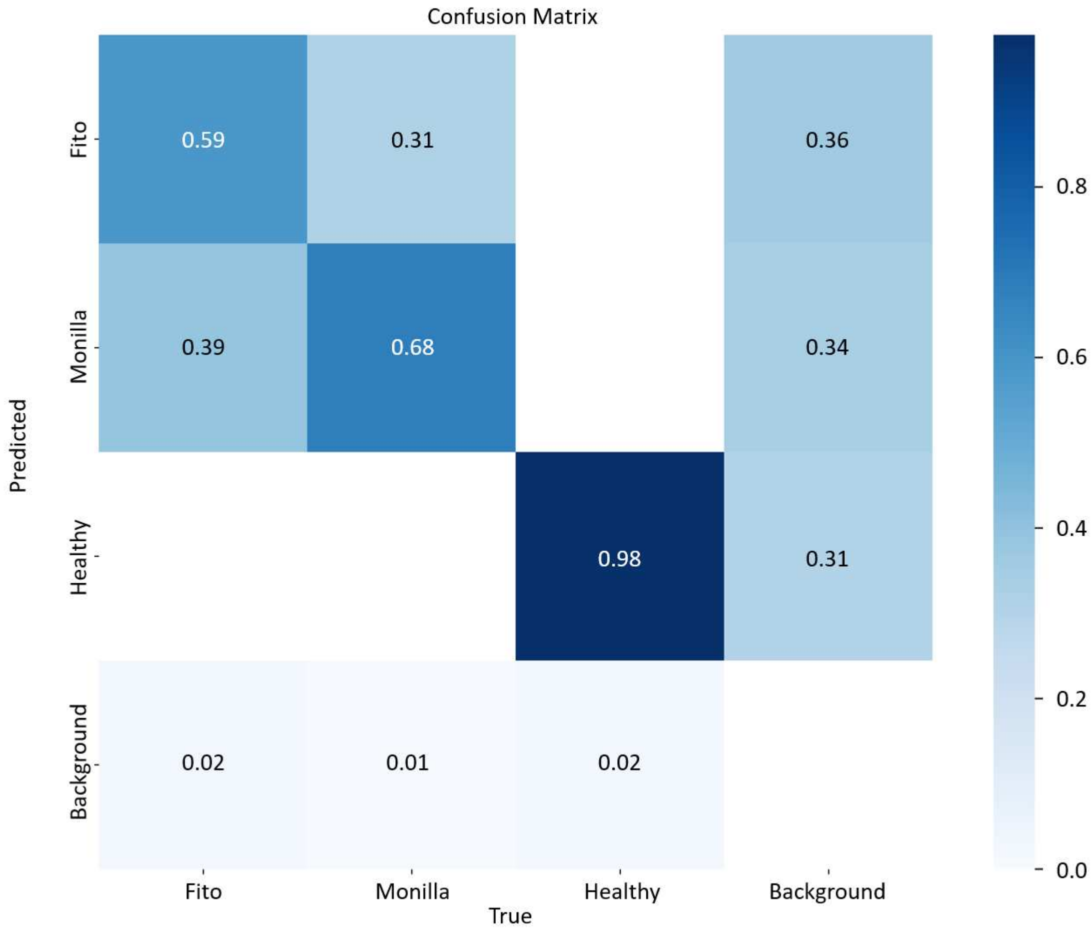


Figure 1. Confusion matrix for the YOLOv5m model trained on the Cocoa Diseases dataset

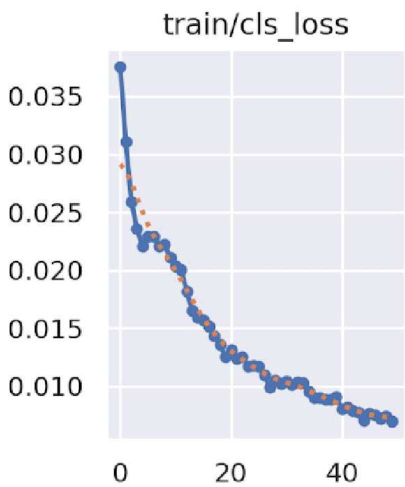


Figure 2. The loss over the training set for each epoch.

Some predictions for a validation set are in Figure 3. As the confusion matrix revealed, our model detects healthy fruits very accurately, while it confuses Monilla and Fito classes. When the model is more confused, and the chance of making a mistake grows, the probability associated with each

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 August 2023 doi:10.20944/preprints202308.0416.v1

predicted bounding box and its class is lower (like in the first images on the left in Figure 3). This should be considered as a warning and more accurate analyses should be carried on the fruit.

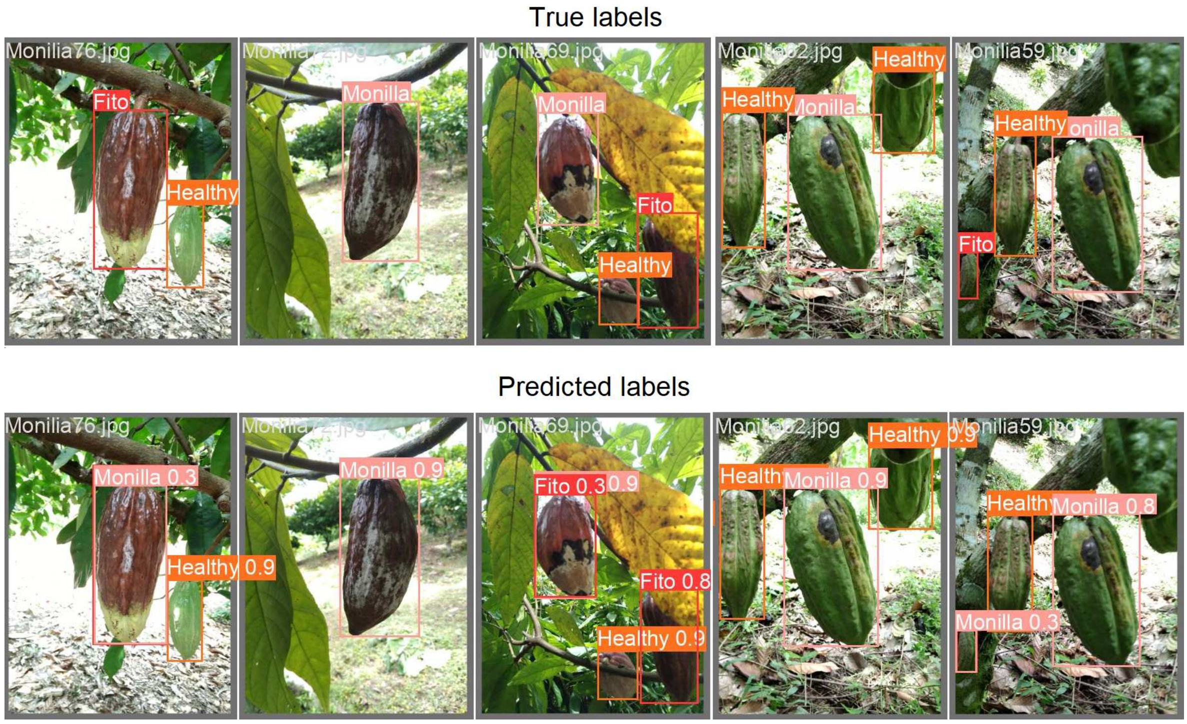


Figure 3. Examples of labeled images in the training set and the corresponding prediction with probabilities

5. Task 2: GRACE prediction Model

Previous studies already showed the effectiveness of using Deep Learning for Remote Sensing data and in particular for Satellite Image Time Series (SITS) [78–80]. With this task, we develop an instrument to monitor and predict the anomalies of water resources. In particular, our aim is to model the Total Water Storage Anomalies (TWSA), predicting the value of the next month (at time $t+1$) considering the values of meteorological and land variables at previous time t. We used as a baseline a *vanilla* Convolutional Neural Network (CNN) and implemented a Fully Convolutional Neural Network (FCN) [1] with a U-net like architecture [2], named CIWA-net (Côte d'Ivoire Water Anomalies network). We chose these architectures for the neural networks because previous studies already reported their good performance in modeling spatio-temporal phenomena and SITS [81–84]. On the contrary, FCNs are more suited for pixel-per-pixel tasks [2,85]. In this work, we test these two different architectures – vanilla CNN and CIWA-net – in different implementations. In each implementation we integrate a different autoregressive component that consists in the delayed target variable TWSA, at a different time lag, as an additional input. In this way, we test if a *vanilla* CNN or CIWA-net could enhance their performances by integrating the temporal information of the target, without a deep modification of their architecture.

5.1. Data

We use GRACE Mascons Solution[5] [86] monthly data. GRACE satellite measures the variation in the earth's gravity field for each month and then estimates the changes in the Equivalent Water Thickness (EWT) at a spatial resolution of 0.25° (25 km). EWT is related to the total amount of water

[5] Downloaded from https://www2.csr.utexas.edu/grace/

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 August 2023 doi:10.20944/preprints202308.0416.v1

stored and available for a unit volume. However, the native cell grid resolution of GRACE data is larger than 25 km – about 120 km – because the earth's gravity field estimations vary slowly in space. For this reason, it is advised to be careful in interpreting GRACE data in basins smaller than 200.000 km^2. Notwithstanding these limitations, our aim is to build a model able to predict GRACE images that represent the target variable in a spatial map. Therefore, we should not attribute these limits due to the low resolution of the input images to the neural network models used for prediction.

The changes, or anomalies, in EWT are calculated with respect to the 2004-2009 time mean baseline and they represent the Total terrestrial Water Storage Anomalies (TWSA) from soil, snow, surface water, groundwater, and aquifers. GRACE mission started in April 2002 and ended in June 2017, but from March 2018 the new GRACE-Follow-on (GRACE-FO) mission has started. Given the missing values from October 2017 to March 2018, we get monthly data from April 2002 to June 2017. Given our interest in the temporal relationship of the phenomenon, we consider the period from April 2002 to April 2015 as our training set, roughly 85% of the available observations; and used the period from May 2015 to June 2017 as our test set, nearly 15% of the available observations.

We use meteorological and land data from the fifth-generation ECMWF reanalysis for the global climate and weather European dataset (ERA5) [22]. ERA5 contains a large number of atmospheric, land, and oceanic climate variables, combining model data with observations from the year 1940 up to the present time. The spatial resolution of the data is 0.25° with an hourly frequency. However, given the time resolution (monthly) of the dependent variable, we use the monthly average ERA5 data. Furthermore, we adopt the same time domain used for GRACE data (i.e., from April 2002 to June 2017) and we split the data for training and testing accordingly. Among all the ERA5 variables, we select the 10 features that are likely to be associated with our target variable (listed in Table 1). All these features are considered as the different channels of an image available at a time t. Hence, for each time step, we have an ERA5 image made by 10 channels.

Table 1. ERA5 features selected for the neural network inputs

Feature	Unit
Surface net solar radiation	$J \cdot m^2$
Skin temperature	K
Evaporation	m of water equivalent
Total precipitation	m
Leaf area index, high vegetation	$m^2 \cdot m^{-2}$
Leaf area index, low vegetation	$m^2 \cdot m^{-2}$
Volumetric soil water layer 1	$m^3 \cdot m^{-3}$
Volumetric soil water layer 2	$m^3 \cdot m^{-3}$
Volumetric soil water layer 3	$m^3 \cdot m^{-3}$
Volumetric soil water layer 4	$m^3 \cdot m^{-3}$

5.1.1. Data preprocessing

Focusing on Côte d'Ivoire, we cropped both GRACE and ERA5 to a square of coordinate -8.875° and -2.125° of longitude, and 4.125° and 10.88° of latitude. Given the difference between ERA5 and GRACE reference grid, we reproject ERA5 data on the GRACE reference grid. In this way, we obtained 28x28 resolution images with all pixels referenced to the same geographical coordinates.

GRACE data exhibit some monthly missing values (NA) in the selected time window due to technical reasons [87]. For filling the missing monthly data we adopted a linear interpolation over the time dimension. In other words, for each month $t \in T_{missing}$ without observations, a new sample y_t has been created linearly interpolating the samples y_{t-1} and y_{t+1} (where y denotes an entire image).

Given the different ranges of the variable, input and output were normalized. More precisely, each feature of input was standardized such that $\frac{x_f - \mu_f}{\sigma_f}$, where $f \in F$ represents a specific feature, and the mean and the variance (μ_f, σ_f) were calculated considering the pixels of all the images in the entire

 doi:10.20944/preprints202308.0416.v1

observed period. Finally, the output was scaled between 0 and 1 using the *min-max feature scaling* formula $\frac{Y-Y_{min}}{Y_{max}-Y_{min}}$.

5.2. Model

As already said, we implemented and compared two different models, a *vanilla* CNN, and CIWA-net using Google Colab. The best model is adopted to prove the feasibility of the task 2. Mean absolute error loss is used for both architectures, and ReLU function is used for the activation of every layer. Every convolutional layer is performed with zero-padding and strides equal to one [6], while strides equal to two are used for downsampling the images of the CIWA-net model because it needs to reduce the size of the input. Transposed convolutions are performed for the reverse operation in order to obtain back images of the original size. Therefore, by upsampling images we augmented again the resolution and, thanks also to the skip connections present in the CIWA-net model, we restored the original input size. The implemented architectures are depicted in Figure 4 and Figure 5. Both models take as input images of dimension $(28, 28, n\_features)$, where we recall that $28x28$ are the dimensions obtained after preprocessing GRACE and ERA5 images, and $n\_features$ is the number of features selected for the experiments and is also the number of channels. The output of the neural network models is an image of dimension $(28, 28, 1)$ because for each image pixel it estimates one feature only: the target variable. To integrate an "autoregressive" component we developed different scenarios:

- $n\_features = 10$, when models take in input only the 10 features of ERA5, listed in Table 1 at time t, i.e. an image with 10 channels;
- $n\_features = 10 + \delta$, where δ stands for the number of additional channels, each of them made by a delayed GRACE data image. Hence, for $\delta = 2$ our input image has 12 channels, 10 of which are ERA5 variables at time t, one is GRACE data at time t and the last is GRACE data at time $t-1$, all trying to predict GRACE at time $t+1$.

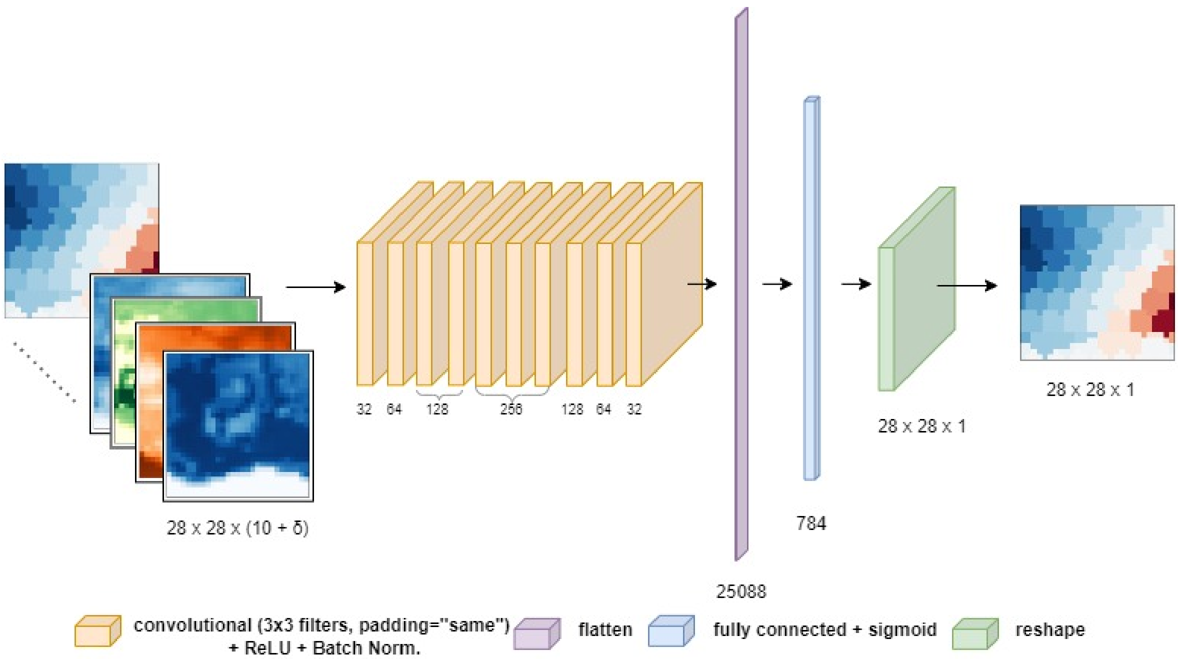


Figure 4. Architecture of the 'vanilla' CNN. The numbers under each layer are the number of filters. The numbers under the images or the reshape layers are their dimensions

[6] Zero-padding is used to fill the pixels in the contours to make input data of the same size of output; strides is the unit shift between one window and the next one

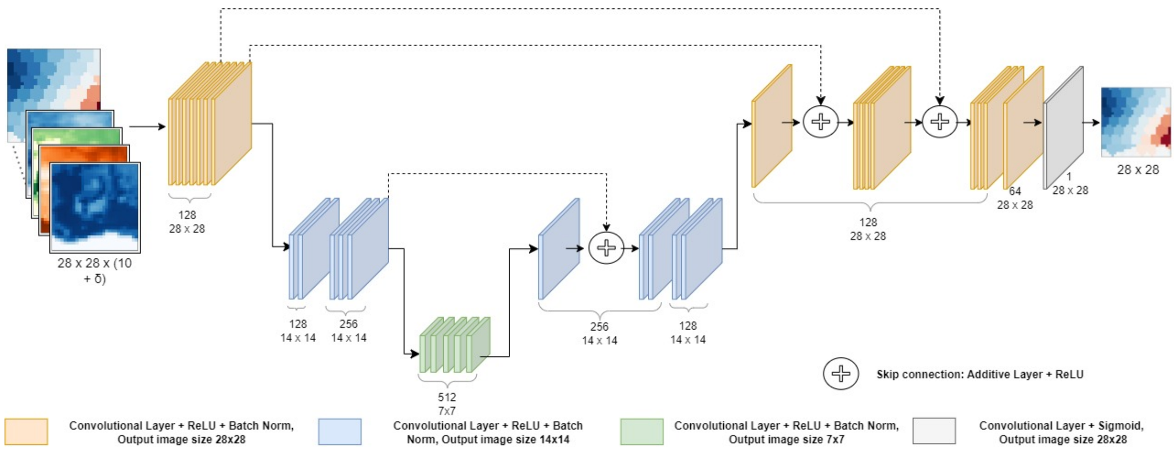


Figure 5. Architecture of CIWA-net. The numbers under each layer are the number of filters and the output image resolutions; curly brackets are used for layers with the same numbers

5.3. Results

Table 2 shows Mean Squared Error (MSE) and Mean Absolute Error (MAE) for different scenarios of the inputs in the training set and in the test set comparing the *vanilla* CNN and our implemented CIWA-net models.

Table 2. Training and test loss using MAE and MSE for the *vanilla* CNN and CIWA-net models; we vary the number of delayed GRACE data taken as additional input channels (δ). Bold numbers are the best test scores for the CNN model. Underlined bold numbers are the overall best test scores

		$\delta=0$	$\delta=1$	$\delta=2$	$\delta=3$	$\delta=4$	$\delta=5$
***vanilla* CNN**	*Train MAE*	0.00440	0.00435	0.00328	0.00743	0.00547	0.00669
	Train MSE	0.00004	0.00004	0.00002	0.00012	0.00007	0.00011
	Test MAE	0.03940	**0.03426**	0.03593	0.03836	0.03585	0.03569
	Test MSE	0.00304	**0.00242**	0.00267	0.00278	0.00260	0.00261
CIWA-net	*Train MAE*	0.01932	0.01948	0.01258	0.01490	0.02425	0.01639
	Train MSE	0.00074	0.00078	0.00033	0.00046	0.00146	0.00055
	Test MAE	0.04461	0.03479	<u>**0.03189**</u>	0.03273	0.03861	0.03435
	Test MSE	0.00407	0.00218	<u>**0.00192**</u>	0.00200	0.00281	0.00217

The results exhibit promising outcomes. We can observe that the prediction errors are in the order of some fraction of percent and this is generally very low.

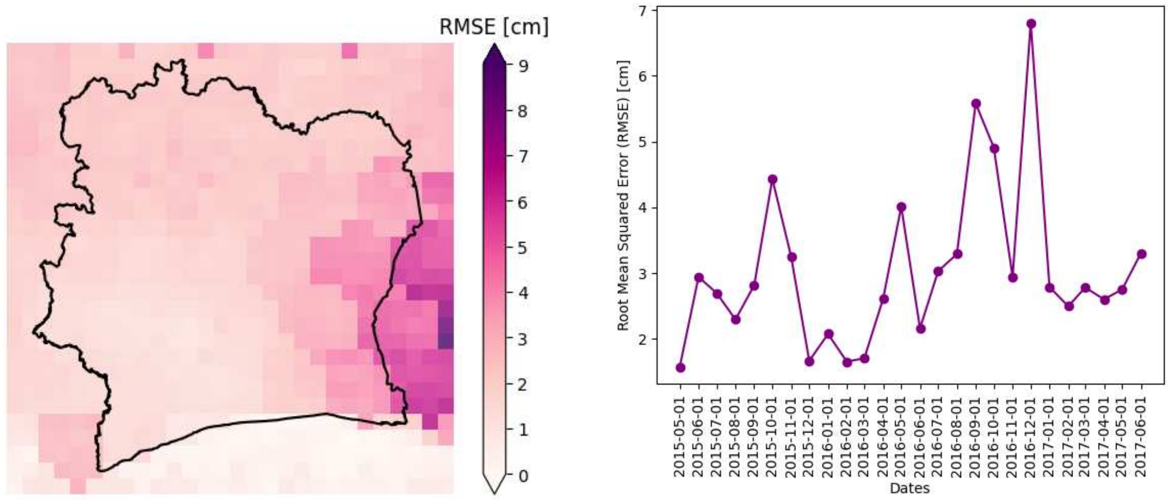
Comparing the *vanilla* CNN and the CIWA-net models we can see that CNN models overfit the data: in fact, they achieve much fewer errors in training-set but similar or higher errors on the test-set than CIWA-net. For $\delta = 0$ CNN model performs better than CIWA-net in terms of test-set, while for $\delta > 0$ CIWA-net models seem to take more advantage from the additional input represented by the GRACE delayed channels, and succeed in achieving smaller test errors (except for $\delta = 4$). For $\delta = 1$ CNN model outperforms CIWA-net only if we consider MAE in the test-set. However, it has an MSE in the test-set higher than the respective CIWA-net: this means that CIWA-net penalizes higher errors more.

Apart from the slightly better performance among neural networks with $\delta > 0$, it is evident that introducing one or more ($\delta > 0$) GRACE delays as additional input channels for both CNN and CIWA-net seems to enhance their performance. Even if the differences are quite small (10^-3 order of magnitude), introducing too many GRACE delays appears to be not worth it because it increases the test errors or does not significantly improve performances. The optimal number of delays appears to be 1 or 2. Maybe this could be due to the fact that input images have already a high number of channels (10) and adding too many channels seems to condense too much information in a single image. The best results are obtained using $\delta = 1$ for the CNN model, and $\delta = 2$ for the CIWA-net

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 August 2023 doi:10.20944/preprints202308.0416.v1

model, the latter achieves the overall best scores (in underlined bold in Table 2) and it is selected as the final model for task 2. Hence, we estimate the spatial and temporal errors committed by the final model below.

As spatial error quantification, we compute the Root Mean Squared Error (RMSE) for each pixel (in Figure 6a) in the test set (from May 2015 to June 2017). It is evident that in the southern-east part the model makes more errors. This could be due to the fact that GRACE data exhibit the highest variance in this area while ERA5 data have the lowest variance. This makes the task of explaining TWSA variability using ERA5 data more difficult in this area.



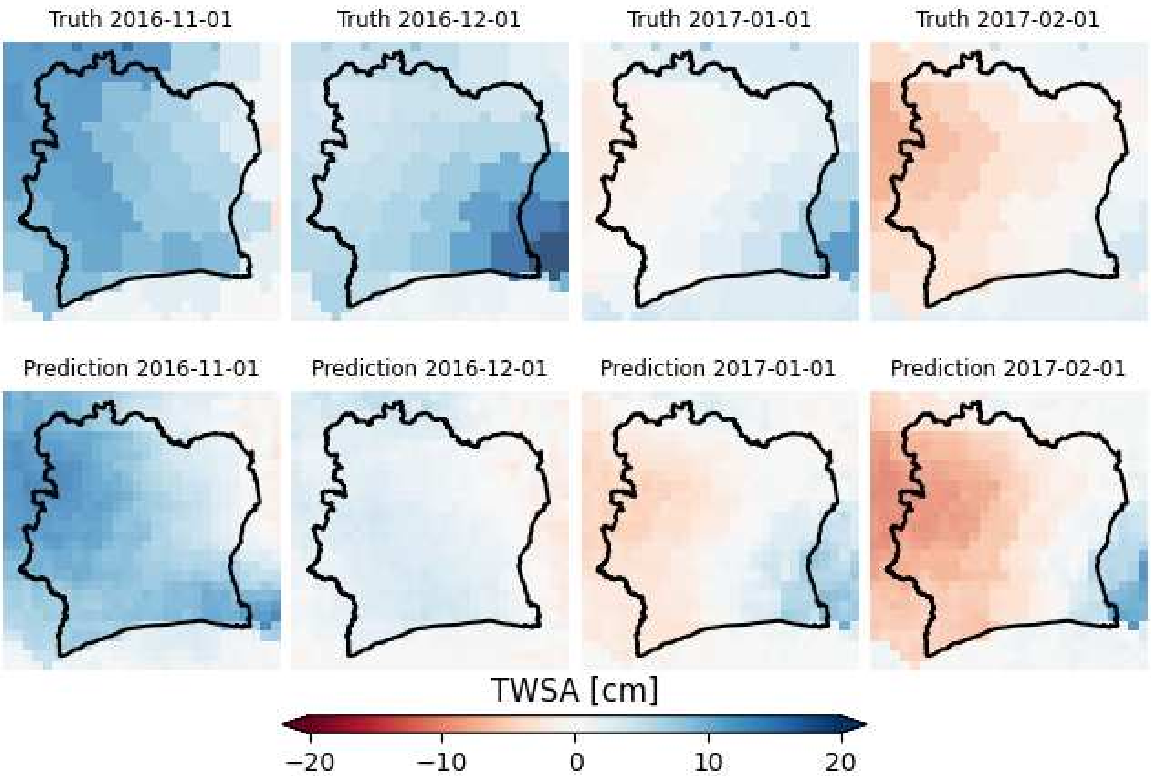
(a) **(b)**

Figure 6. a) Root Mean Squared Error (RMSE) for each 0.25° pixel computed on the test set (from May 2015 to June 2017). **b)** Pixel average Root Mean Squared Error (RMSE) computed for every single image (i.e. month) in the test set. Errors are computed using CIWA-net with $\delta = 2$ predictions with respect to the true TWSA GRACE values in cm

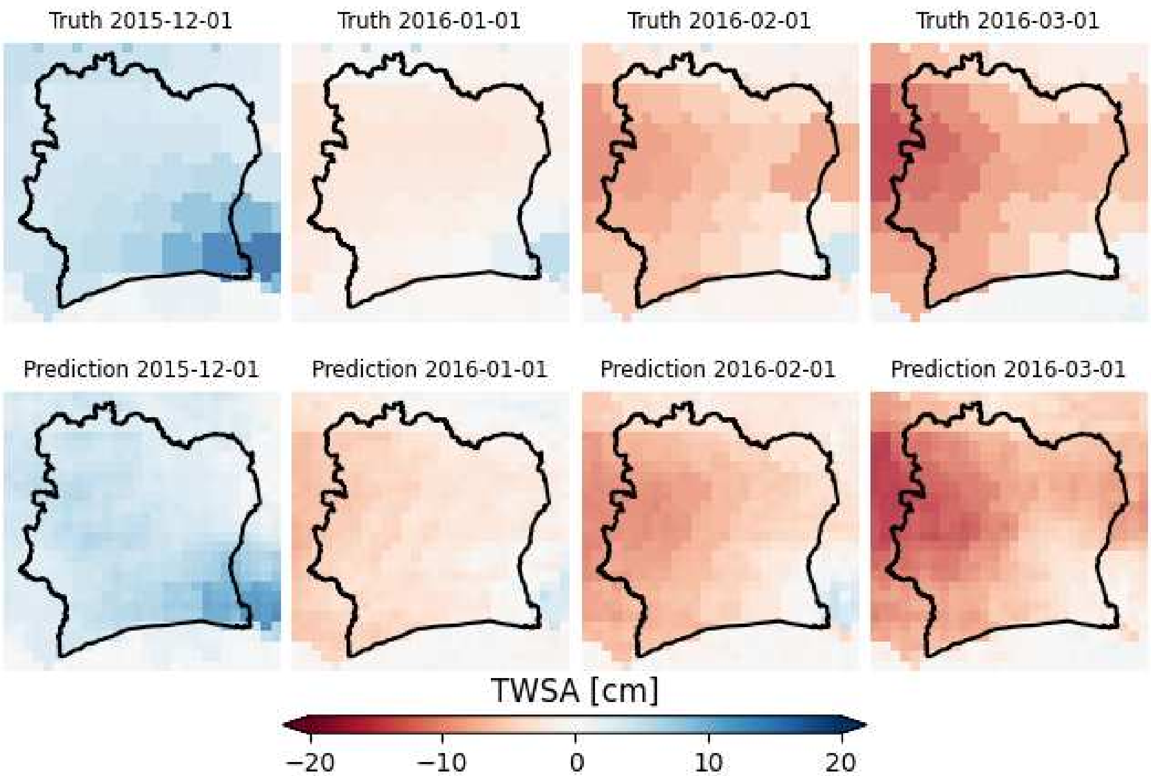
However, it is worth noting the highest errors are outside our region of interest (Côte d'Ivoire). Given the overall standard deviation of 9.4cm for all the available GRACE data from April 2002 to June 2017, RMSE values inside Côte d'Ivoire (in which the maximum RMSE is 6.15cm) could be regarded as acceptable, even if model improvements are needed.

For the temporal error, we computed the RMSE for each image of the test set (i.e. for each month of the test set) considering all the pixels of a particular month as observations of the same instance. Figure 6b shows the RMSE series. A peak for the date 2016-12-01 is evident, and it is an explanation for the previously mentioned spatial estimation difficulties. In fact, we notice that the spatial distribution of the target is very different in the two consecutive months (see Figure 7a). Similar explanations could be given for the two other peaks (2015-10-01 and 2016-09-01).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 4 August 2023 doi:10.20944/preprints202308.0416.v1



(a)



(b)

Figure 7. Ground truth (first row) and prediction (second row) of GRACE TWSA from 2016-11-01 to 2017-02-01 (**a**), and from 2015-12-01 to 2016-03-01 **b**). Predictions are made using CIWA-net model with $\delta = 2$. In **a)** 2016-12-01 is the date for which CIWA-net performs the worst. In **b)** 2015-12-01 and 2016-02-01 are dates for which CIWA-net performs the best.

 doi:10.20944/preprints202308.0416.v1

Figure 7a and Figure 7b show some predictions made by the selected model in task 2 for some dates of the test set, including the one for which the model commits the highest RMSE (2016-12-01) and some of the best performance model dates (2015-12-01 and 2016-02-01). In general, the model is able to correctly predict the evolution of TWSA, apart from some anomalous and abrupt changes (2016-12-01) with respect to the previous months, for which the model needs to be improved. Differences between the resolutions of ERA5 and the native one in GRACE make model outputs vary more than the true ones. In fact, true GRACE values remain constant also in the neighbor pixels, which does not occur in ERA5 data and model predictions. However, this is not a relevant problem for the final aim of the research: we should consider that the model should be used by policymakers for a national or sub-national forecast about the main TWSA, and not necessarily at a 0.25° resolution. Considering the above mentioned resolution limitations, for this task we consider that our model succeeds: it is able to detect major changes and their approximate locations.

6. Discussion

In the paper, we demonstrate that different types of low-cost sensors can be employed to help developing countries in managing their resources. On one hand, policymakers can have a regional level perspective analyzing satellite data and, on the other hand, local farmers can be supported by tools for on-site decision-making or monitoring.

As demonstrated in Section 4, with a relatively simple dataset for classification purposes, it is possible to build classification apps to support local farmers in their activities. Nowadays, local farmers endowed with a smartphone could be able to collect on-site pictures of their crops and make them labelling by an expert. With such a dataset, it is feasible to create AI classifiers able to help farmers in identifying particular features, which inexpert workers could not detect. Locals can be instructed on how to build classification datasets correctly for training machine learning models.

For instance, depending on the case, it could be necessary to collect more than a certain number of samples, take pictures respecting some characteristics (e.g., with a variable or constant distance from the target, one or several features to detect, avoiding or allowing redundant object samples, constant sensor's type for acquiring the image, etc.). Consequently, without employing expensive instruments, a dataset is easily producible through the use of a low-cost smartphone or camera and a local expert could label it once and for all. Then, using a free GPU usage service (e.g., Google Colab, etc.), a standard pre-trained machine learning model can be easily adapted to the chosen task. Finally, a model implemented using libraries such as TensorFlow, Keras or PyTorch can be easily converted into a format running on a smartphone, making it usable by the local farmers who can benefit in their activities.

The accuracy of our model based on YOLOv5m image segmentation, as depicted in Figure 1, is 98% for detecting healthy fruits, while the other two classes are often confused. However, if the task of the farmer is to detect healthy and damaged fruits, the tool is pretty accurate. YOLOv5m is not the best model for object detection in terms of accuracy, but it revealed to be particularly suitable for real-time tasks, like in our case. Furthermore, the model has been trained over a small dataset, given the scarcity of this kind of data (only 312 images). Nevertheless, even though the dataset needs to be extended, the result should be considered interesting since the model internally implements a data augmentation step that makes it more general and robust with new images.

Other cheap sensors can be employed on-site (e.g., based on Arduino) to measure various aspects of the environment (e.g., soil moisture, air humidity, air quality, etc). In this way, it would be feasible to gather specific information impossible to be collected by other means with a similar local precision (e.g., satellites, drones, etc).

In addition, the huge amount of satellite public images and data present online enables several terrestrial analyses, from shorter (e.g., centimeters or meters of ground resolution) to longer scale (e.g., kilometers of ground resolution). NASA and ESA space agencies offer for free access to their satellites datasets and encourage their usage. In particular, Earth Observing System (EOS) and Copernicus programs, implemented respectively by NASA and ESA, supply continuously different images of the Earth in the visible and non-visible spectral range. Visible and non-visible bands enable everyone to perform several analyses for free: for instance to detect deforestation or degradation of the land with a few meters of ground resolution.

In our case, we demonstrated in Section 5 the possibility to use GRACE and ERA5 datasets to build a model based on deep neural networks of prediction for drought a month in advance, which could represent a decisive improvement for the developing countries and would allow them to manage water resources in the short-mid term. This proved the feasibility of building such a tool to support policymakers who have to monitor and manage resources at the regional level. However, this kind of dataset is quite limited by its low resolution, which is approximately 25km/pixel, which is already extrapolated by an original granularity of 120km as explained in Section 5.1. Therefore, this dataset could be improved significantly in the future in order to detect smaller basins and bigger ones more accurately. Anyway, over the years, satellite missions are increasing the acquisition frequency and ground resolution of their products which will contribute to building more precise predictive models from a temporal and a resolution perspective. Additionally, when it is feasible, this information could be also improved by integrating data acquired by sensors placed on the ground, resorting to the so-called *down-scaling* methods.

In summary, our article shows the potential of cost-effective sensors and freely available satellite data to empower developing countries in managing their resources effectively. By providing tools and models accessible to both policymakers and local farmers, we can make significant strides in resource management and environmental monitoring without relying on expensive instruments.

7. Future works

Task 1 of Section 4 proposed a tool for local farmers to detect healthy and unhealthy fruits. This tool could be used for smartphone images and a smartphone application could be implemented for user-friendly usage. A more complex model could be implemented with the aim of better discriminating among diseases. Anyway, fast and light implementation should be preferred given the smartphone application context. Integrating more diseases and introducing additional classes could be worth it. Experts should join the design and training phase of the model, bringing domain knowledge. In this way, a more exhaustive and reliable tool could be implemented and furnished to local farmers to improve their farming activities.

In task 2, we tested and proved the feasibility of supporting natural resource management using open data and publicly available software. Developing countries will benefit the most from such low-cost instruments to tackle climate change issues. One element that will improve CIWA-net performances is the training set dimension, which is now limited to the Côte d'Ivoire. Future analyses could take into consideration also neighboring countries with similar morphology, economy, and climate for the same period. In this way, we could offer more reliable instruments capable of making more precise predictions. Furthermore, it should be worth retrieving proxy data of anthropogenic pressure on water resources: in this way, a model could predict not only meteorological droughts but also those induced by human over-consumption. Other types of CNN-like architectures could be very suited for this task, for example, tempCNN, TCN, and convLSTM [82,88,89] could be used considering our SITS as a video. These architectures may be useful for additional comparisons to our CIWA-net selected model or to implement a new mixed architecture specifically designed to capture spatio-temporal relations proper of SITS.

It is possible to implement models for detecting degradation and deforestation using satellite images of Sentinel-1 and Sentinel-2 European Space Agency (ESA) missions (available at: https://dataspace.copernicus.eu/). Sentinel-2 data are multi-spectral (RGB, NIR, VNIR, and SWIR) images of different resolutions up to 10m, and continental land coverage is about every 5 days[7]. Some researches already showed the feasibility of using Sentinel-2 images and neural networks for deforestation detection [90–93]. Differently, Sentinel-1 furnishes C-band synthetic aperture radar images of the entire globe every 12 days[8]. This type of data appears to be relevant for forest change detection, especially if combined with Sentinel-2 data to overcome cloud coverage and adverse meteorological conditions[94–97]. CIWA-net model implemented for our task 2 in Section 5 could be restructured for forest monitoring, i.e. a Fully Convolutional Network for classification purposes taking in input Sentinel-1 and Sentinel-2 images at the beginning and at the end of a monitoring period, producing in output a classification map detecting deforestation and degradation.

8. Conclusions

Our study focuses on the urgent need for climate change mitigation and adaptation in developing countries like Côte d'Ivoire, where the economy heavily depends on agriculture, and farmers face challenges due to deforestation, lack of rainfall, and climate change. By using cost-effective AI models and open-source software, this work demonstrates the potential to provide valuable support to farmers and policymakers, aiding in sustainable land and water management practices. The study emphasizes the importance of collaboration among experts, researchers, NGOs, and policymakers to bring about lasting changes and achieve sustainable development goals.

The article shows the application of machine learning for sensors analysis for social good, specifically in combating climate change and facilitating land management and farming in developing countries, focusing on Côte d'Ivoire, which is the largest producer of cocoa beans in the world. To contrast climate change it is necessary to work at different scales from the single pod to the regional scale. Furthermore, considering that cocoa cultivation is done by a number of small farmers that earn around 1.0 Euro per day from their labor, it is important to maintain costs low. The paper proposes the use of deep neural networks (YOLOv5m) to distinguish healthy cocoa plants and pods from unhealthy ones using mobile phone images, thus helping local farmers improve cocoa production at low costs. Additionally, the article suggests a new method of forecasting TWSA using Fully Convolutional Neural Networks (FCN) U-net like architecture, called CIWA-net. These approaches can help combat climate change by decreasing crop losses, and waste, reducing the inputs on the soil (fertilizers) which are responsible for greenhouse gas emissions, and helping water resources forecasting to counter extreme water events, desertification and growth of arid areas.

The new architecture CIWA-net is discussed and shown through comparison with a *vanilla* CNN model. For the YOLOv5m accuracy is shown using a confusion matrix. The proposed models for cocoa pod classification and water resources forecasting show promising results, with the potential to help farmers, NGOs, and local authorities make informed decisions to address climate change impacts and improve agricultural practices.

Overall, the article highlights the significant role of machine learning and open-source technologies in addressing climate change and promoting a just social-technological transitions in developing countries to benefit both small farmers and the larger community. The article underscores the substantial role of machine learning and open source technologies in addressing climate change and facilitating equitable socio-technological transitions in developing countries. It can bring advantages to both small-scale farmers and the wider community.

We stress the urgent need to address the negative consequences of cocoa cultivation and habitat loss in Côte d'Ivoire, and the nearby countries (Republic of Ghana, Togolese Republic, also producers of cocoa), emphasizing the significance of biodiversity conservation, deforestation control, habitat and water management, for effective sustainability measures. Our classifiers and sensor forecasting have

---

[7] For more details, please visit: https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi.
[8] For more details, please visit: https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-1-sar.

the potential of mitigating the environmental impact at low costs and promoting social outcomes at affordable costs.

References

1. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431–3440. doi:10.1109/CVPR.2015.7298965.
2. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015; Navab, N.; Hornegger, J.; Wells, W.M.; Frangi, A.F., Eds.; Springer International Publishing: Cham, 2015; pp. 234–241.
3. Shukla, P.R.; Skea, J.; Calvo Buendia, E.; Masson-Delmotte, V.; Pörtner, H.O.; Roberts, D.; Zhai, P.; Slade, R.; Connors, S.; Van Diemen, R.; others. IPCC, 2019: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems **2019**.
4. IPCC-2023. Summary for Policymakers. In: Climate Change 2023: Synthesis Report. https://www.ipcc.ch/report/ar6/syr/downloads/report/IPCC_AR6_SYR_SPM.pdf, 2023. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. Geneva, Switzerland, pp. 1-34, doi:10.59327/IPCC/AR6-9789291691647.001.
5. Ortiz-Bobea, A.; Ault, T.R.; Carrillo, C.M.; Chambers, R.G.; Lobell, D.B. Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change* **2021**, *11*, 306–312.
6. Sickles, R.; Zelenyuk, V. *Measurement of productivity and efficiency*; Cambridge University Press, 2019. doi:https://doi.org/10.1017/9781139565981.
7. Kakaoplattform. https://www.kakaoplattform.ch/about-cocoa/cocoa-facts-and-figures, 2023. Available online: accessed on 29 July 2023.
8. Gosset, M.; Dibi-Anoh, P.A.; Schumann, G.; Hostache, R.; Paris, A.; Zahiri, E.P.; Kacou, M.; Gal, L. Hydrometeorological Extreme Events in Africa: The Role of Satellite Observations for Monitoring Pluvial and Fluvial Flood Risk. *Surveys in Geophysics* **2023**, *44*, 197–223. doi:10.1007/s10712-022-09749-6.
9. Caretta, M.; Mukherji, A.; Arfanuzzaman, M.; Betts, R.; Gelfan, A.; Hirabayashi, Y.; Lissner, T.; Liu, J.; Lopez Gunn, E.; Morgan, R.; Mwanga, S.; Supratid, S. Water. In *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK and New York, NY, USA, 2022; pp. 551—-712. doi:10.1017/9781009325844.006.
10. UNDRR.; CIMA. *Côte d'Ivoire Disaster Risk Profile*; United Nations Office for Disaster Risk Reduction and CIMA Research Foundation, 2019.
11. Bank, W. *Côte d'Ivoire - Country Economic Memorandum: Sustaining the Growth Acceleration (English)*; 2021.
12. Dibi-Anoh, P.A.; Koné, M.; Gerdener, H.; Kusche, J.; N'Da, C.K. Hydrometeorological Extreme Events in West Africa: Droughts. *Surveys in Geophysics* **2023**, *44*, 173–195. doi:10.1007/s10712-022-09748-7.
13. Vaissaire, J. Événements climatiques extrêmes Réduire les vulnérabilités des systèmes écologiques et sociaux. Sous la direction de Henri Décamps, Institut de France-Académie des Sciences-2010. *Bulletin de l'Académie Vétérinaire de France* **2011**, *164*, 74–75.
14. Masih, I.; Maskey, S.; Mussá, F.E.F.; Trambauer, P. A review of droughts on the African continent: a geospatial and long-term perspective. *Hydrology and Earth System Sciences* **2014**, *18*, 3635–3649. doi:10.5194/hess-18-3635-2014.
15. Quesada-Montano, B.; Wetterhall, F.; Westerberg, I.K.; Hidalgo, H.G.; Halldin, S. Characterising droughts in Central America with uncertain hydro-meteorological data. *Theoretical and Applied Climatology* **2019**, *137*, 2125–2138. doi:10.1007/s00704-018-2730-z.
16. Landerer, F.W.; Flechtner, F.M.; Save, H.; Webb, F.H.; Bandikova, T.; Bertiger, W.I.; Bettadpur, S.V.; Byun, S.H.; Dahle, C.; Dobslaw, H.; Fahnestock, E.; Harvey, N.; Kang, Z.; Kruizinga, G.L.H.; Loomis, B.D.; McCullough, C.; Murböck, M.; Nagel, P.; Paik, M.; Pie, N.; Poole, S.; Strekalov, D.; Tamisiea, M.E.; Wang, F.; Watkins, M.M.; Wen, H.Y.; Wiese, D.N.; Yuan, D.N. Extending the Global Mass Change Data Record: GRACE Follow-On Instrument and Science Data Performance. *Geophysical Research Letters*

2020, *47*, e2020GL088306, [https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL088306]. e2020GL088306 2020GL088306, doi:https://doi.org/10.1029/2020GL088306.

17. Frappart, F.; Ramillien, G. Monitoring groundwater storage changes using the Gravity Recovery and Climate Experiment (GRACE) satellite mission: A review. *Remote Sensing* **2018**, *10*, 829.
18. Humphrey, V.; Gudmundsson, L. GRACE-REC: a reconstruction of climate-driven water storage changes over the last century. *Earth System Science Data* **2019**, *11*, 1153–1170. doi:10.5194/essd-11-1153-2019.
19. Hosseini-Moghari, S.M.; Araghinejad, S.; Ebrahimi, K.; Tang, Q.; AghaKouchak, A. Using GRACE satellite observations for separating meteorological variability from anthropogenic impacts on water availability. *Scientific Reports* **2020**, *10*, 15098. doi:10.1038/s41598-020-71837-7.
20. Solander, K.C.; Reager, J.T.; Wada, Y.; Famiglietti, J.S.; Middleton, R.S. GRACE satellite observations reveal the severity of recent water over-consumption in the United States. *Scientific Reports* **2017**, *7*, 8723. doi:10.1038/s41598-017-07450-y.
21. J., P., Agricultural sustainability: concepts, principles and evidence. In *Philosophical Transactions of the Royal Society*; The Royal Society, 2008; pp. 447–465.
22. Hersbach, H.; Bell, B.; Berrisford, P.; Biavati, G.; Horányi, A.; Muñoz Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Rozum, I.; Schepers, D.; Simmons, A.; Soci, C.; Dee, D.; Thépaut, J.N. ERA5 monthly averaged data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (accessed on 26-07-2023). Technical report, 2023. doi:10.24381/cds.f17050d7.
23. www.chocofair.org - accessed on 31/07/2023). https://www.chocofair.org, 2023.
24. Communauté Abel, Grand Bassam, Côte d'Ivoire - accessed on 31/07/2023). http://www.communauteabel.org/, 2023.
25. Gruppo Abele, Turin, Italy, Choco+ Intiative - accessed on 31/07/2023). https://www.gruppoabele.org/it-schede-1639-choco, 2023.
26. Pinardi, S.; Salis, M.; Sartor, G.; Meo, R. EU-Africa: Digital and Social Questions in a Multicultural Agroecological Transition for the Cocoa Production in Africa. *Social Sciences* **2023**, *12*. doi:10.3390/socsci12070398.
27. Ruf, F.; Schroth, G.; Doffangui, K. Climate change, cocoa migrations and deforestation in West Africa: What does the past tell us about the future? *Sustainability Science* **2015**, *10*, 101–111. doi:10.1007/s11625-014-0282-4.
28. Barima, Y.S.S.; Kouakou, A.T.M.; Bamba, I.; Sangne, Y.C.; Godron, M.; Andrieu, J.; Bogaert, J. Cocoa crops are destroying the forest reserves of the classified forest of Haut-Sassandra (Ivory Coast). *Global Ecology and Conservation* **2016**, *8*, 85–98. doi:https://doi.org/10.1016/j.gecco.2016.08.009.
29. Abu, I.O.; Szantoi, Z.; Brink, A.; Robuchon, M.; Thiel, M. Detecting cocoa plantations in Côte d'Ivoire and Ghana and their implications on protected areas. *Ecological Indicators* **2021**, *129*, 107863. doi:https://doi.org/10.1016/j.ecolind.2021.107863.
30. Kalischek, N.; Lang, N.; Renier, C.; Daudt, R.C.; Addoah, T.; Thompson, W.; Blaser-Hart, W.J.; Garrett, R.; Schindler, K.; Wegner, J.D. Cocoa plantations are associated with deforestation in Côte d'Ivoire and Ghana. *Nature Food* **2023**, *4*, 384–393. doi:10.1038/s43016-023-00751-8.
31. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; Kommareddy, A.; Egorov, A.; Chini, L.; Justice, C.O.; Townshend, J.R.G. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **2013**, *342*, 850–853. doi:10.1126/science.1244693.
32. Vancutsem, C.; Achard, F.; Pekel, J.F.; Vieilledent, G.; Carboni, S.; Simonetti, D.; Gallego, J.; Aragão, L.E.O.C.; Nasi, R. Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. *Science Advances* **2021**, *7*, eabe1603. doi:10.1126/sciadv.abe1603.
33. Achard, F.; Beuchle, R.; Mayaux, P.; Stibig, H.J.; Bodart, C.; Brink, A.; Carboni, S.; Desclée, B.; Donnay, F.; Eva, H.D.; others. Determination of tropical deforestation rates and related carbon losses from 1990 to 2010. *Global change biology* **2014**, *20*, 2540–2554. doi:https://doi.org/10.1111/gcb.12605.
34. FAO-2020. The State of Food and Agriculture 2020. Overcoming Water Challenges in Agriculture **2020**. p. 289pp. doi:https://doi.org/10.4060/ca9692en.
35. IPCC-2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Technical report, Geneva, Switzerland, 2014.

36. Dhakal, S.; Minx, J.; Toth, F.L. Climate Change 2022 - Mitigation of Climate Change Working Group III Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Technical report, Cambridge, UK, 2023. doi:https://doi.org/10.1017/9781009157926.004.
37. Hong, C.; Burney, J.A.; Pongratz, J.; Nabel, J.E.M.S.; Mueller, N.D.; Jackson, R.B.; Davis, S.J. Global and regional drivers of land-use emissions in 1961–2017. *Nature* **2021**, *589*, 554–561. doi:10.1038/s41586-020-03138-y.
38. Dalgaard, T.; Hutchings, N.; Porter, J. Agroecology, scaling and interdisciplinarity. *Agriculture, Ecosystems & Environment* **2003**, *100*, 39–51. doi:https://doi.org/10.1016/S0167-8809(03)00152-X.
39. https://www.ecmwf.int/en/forecasts/dataset/cams-global-reanalysis, 2023. CAMS reanalysis data set of atmospheric composition (AC) produced by the Copernicus Atmosphere Monitoring Service. Period 2003-June 2022.
40. Pierce, F, J.; Nowak, P. Aspects of Precision Agriculture; Academic Press, 1999; Vol. 67, *Advances in Agronomy*, pp. 1–85. doi:https://doi.org/10.1016/S0065-2113(08)60513-1.
41. Wezel, A.; Herren, B.G.; Kerr, R.B.; Barrios, E.; Gonçalves, A.; Sinclair, F. Agroecological principles and elements and their implications for transitioning to sustainable food systems. A review. *Agronomy for Sustainable Development* **2020**, *40*, 40. doi:10.1007/s13593-020-00646-z.
42. Duncan, E.; Glaros, A.; Ross, D.Z.; Nost, E. New but for whom? Discourses of innovation in precision agriculture. *Agriculture and Human Values* **2021**, *38*, 1181–1199. doi:10.1007/s10460-021-10244-8.
43. 2022., A. Ivory Coast battles chocolate companies to improve farmers' lives. https://www.aljazeera.com/features/2022/12/22/ivory-coast-battles-chocolate-companies-to-improve-farmers-lives, 2022. Available online: accessed on 29 July 2023.
44. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788. doi:10.1109/CVPR.2016.91.
45. Theben, Alexandra and Gunderson, Laura and López-Forées, Laura and Misuraca, Gianluca and Lipiáñez-Villaneuva, Francisco. Challenges and limits of an open source approach to Artificial Intelligence. https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662908/IPOL_STU(2021)662908_EN.pdf, 2021. Policy Department for Economic, Scientific and Quality of Life Policies Directorate-General for Internal Policies, Special Committee on Artificial Intelligence in a Digital Age, EU Parliament.
46. GITHUB platform, with an open source section and social good projects, today equipped with Copilot, an AI-based tool in support to developers. https://github.blog/category/open-source/.
47. Joinup project, for Interoperable Europe and Interoperable solutions. https://joinup.ec.europa.eu.
48. Apache community. https://httpd.apache.org.
49. Free Software Foundation. https://www.gnu.org.
50. UC Irvine Machine Learning Repository. https://archive.ics.uci.edu/.
51. Kaggle: Level up with the largest AI & ML community. https://www.kaggle.com/.
52. Tensorflow platform. https://www.tensorflow.org.
53. Deep Scalable Sparse Tensor Network Engine: an Amazon developed library for building Deep Learning Machine Learning models. https://github.com/amazon-archives/amazon-dsstne.
54. Keras: deep learning API, in Python, running on top of the machine learning platform TensorFlow. https://keras.io/.
55. OpenStack: The Most Widely Deployed Open Source Cloud Software in the World, a set of software components that provide common services for cloud infrastructure. https://www.openstack.org/.
56. OpenHPC: A Community-based effort provides tools, documentation, and more to accelerate innovation and broaden access to HPC. https://opensource.com/article/17/11/openhpc.
57. ML-Ops: The Best Open-Source MLOps Tools You Should Know. https://neptune.ai/blog/best-open-source-mlops-tools.
58. Curate better data, build better models with FiftyOne. https://voxel51.com/.
59. OpenCV Library: the world's largest resource of Computer Vision. https://opencv.org/.
60. Open Source Innovation in Artificial Intelligence and Data. https://lfaidata.foundation.
61. Apache Kafka: an open-source distributed event streaming platform for high-performance data pipelines, streaming analytics, data integration, and mission-critical applications. https://kafka.apache.org.

62. InfluxDB: a scalable datastore for metrics, events, and real-time analytics. https://github.com/influxdata/influxdb.
63. Hamoodi, S.A.; Hamoodi, A.N.; Haydar, G.M. Automated irrigation system based on soil moisture using arduino board. *Bulletin of Electrical Engineering and Informatics* **2020**, *9*, 870–876.
64. Mylonas, A.; Kazanci, O.B.; Andersen, R.K.; Olesen, B.W. Capabilities and limitations of wireless CO2, temperature and relative humidity sensors. *Building and Environment* **2019**, *154*, 362–374.
65. Rosell-Polo, J.R.; Cheein, F.A.; Gregorio, E.; Andújar, D.; Puigdomènech, L.; Masip, J.; Escolà, A. Advances in structured light sensors applications in precision agriculture and livestock farming. *Advances in agronomy* **2015**, *133*, 71–112.
66. Tanner, B.D. Automated weather stations. *Remote Sensing Reviews* **1990**, *5*, 73–98.
67. Pascual, R.L.; Sanchez, D.M.R.; Naces, D.L.E.; Nuñez, W.A. A Wireless Sensor Network using XBee for precision agriculture of sweet potatoes (Ipomoea batatas). 2015 International Conference on Humanoid, Nanotechnology, Information Technology,Communication and Control, Environment and Management (HNICEM), 2015, pp. 1–4. doi:10.1109/HNICEM.2015.7393212.
68. Romero, C.; Teologo, A.; Nuesca, A.L. Arduino Rice Pest Trap using Laser Sensor. 2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM). IEEE, 2021, pp. 1–6.
69. Kriti, T.; Bhatia, S. Automatic irrigation system using Arduino UNO. 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), 2017.
70. Frost, A.; Schofield, C.; Beaulah, S.; Mottram, T.; Lines, J.; Wathes, C. A review of livestock monitoring and the need for integrated systems. *Computers and electronics in agriculture* **1997**, *17*, 139–159.
71. Li, H.; Guo, Y.; Zhao, H.; Wang, Y.; Chow, D. Towards automated greenhouse: A state of the art review on greenhouse monitoring methods and technologies based on internet of things. *Computers and Electronics in Agriculture* **2021**, *191*, 106558.
72. Bastidas-Alva, R.A.; Cardenas, J.A.P.; Espinoza, K.S.B.; Nuñez, V.K.P.; Rivera, M.E.Q.; Huaytalla, J. Recognition and classification system for trinitario cocoa fruits according to their ripening stage based on the Yolo v5 algorithm. 2022 Asia Conference on Advanced Robotics, Automation, and Control Engineering (ARACE). IEEE, 2022, pp. 138–142.
73. Ko, K.; Jang, I.; Choi, J.H.; Lim, J.H.; Lee, D.U. Stochastic Decision Fusion of Convolutional Neural Networks for Tomato Ripeness Detection in Agricultural Sorting Systems. *Sensors* **2021**, *21*, 917. doi:10.3390/s21030917.
74. MacEachern, C.B.; Esau, T.J.; Schumann, A.W.; Hennessy, P.J.; Zaman, Q.U. Detection of fruit maturity stage and yield estimation in wild blueberry using deep learning convolutional neural networks. *Smart Agricultural Technology* **2023**, *3*, 100099. doi:https://doi.org/10.1016/j.atech.2022.100099.
75. Chen, J.W.; Lin, W.J.; Cheng, H.J.; Hung, C.L.; Lin, C.Y.; Chen, S.P. A Smartphone-Based Application for Scale Pest Detection Using Multiple-Object Detection Methods. *Electronics* **2021**, *10*, 372. doi:10.3390/electronics10040372.
76. Degu, M.Z.; Simegn, G.L. Smartphone based detection and classification of poultry diseases from chicken fecal images using deep learning techniques. *Smart Agricultural Technology* **2023**, *4*, 100221. doi:https://doi.org/10.1016/j.atech.2023.100221.
77. Zhou, C.; Lee, W.S.; Liburd, O.E.; Aygun, I.; Zhou, X.; Pourreza, A.; Schueller, J.K.; Ampatzidis, Y. Detecting two-spotted spider mites and predatory mites in strawberry using deep learning. *Smart Agricultural Technology* **2023**, *4*, 100229. doi:https://doi.org/10.1016/j.atech.2023.100229.
78. Moskolaï, W.R.; Abdou, W.; Dipanda, A.; Kolyang. Application of Deep Learning Architectures for Satellite Image Time Series Prediction: A Review. *Remote Sensing* **2021**, *13*. doi:10.3390/rs13234822.
79. Yu, Q.; Wang, S.; He, H.; Yang, K.; Ma, L.; Li, J. Reconstructing GRACE-like TWS anomalies for the Canadian landmass using deep learning and land surface model. *International Journal of Applied Earth Observation and Geoinformation* **2021**, *102*, 102404. doi:https://doi.org/10.1016/j.jag.2021.102404.
80. Ahmed, M.; Sultan, M.; Elbayoumi, T.; Tissot, P. Forecasting GRACE Data over the African Watersheds Using Artificial Neural Networks. *Remote Sensing* **2019**, *11*. doi:10.3390/rs11151769.
81. Brock, J.; Abdallah, Z.S. Investigating Temporal Convolutional Neural Networks for Satellite Image Time Series Classification: A survey, 2023, [arXiv:cs.CV/2204.08461].
82. Pelletier, C.; Webb, G.I.; Petitjean, F. Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series. *Remote Sensing* **2019**, *11*. doi:10.3390/rs11050523.

83. Mboga, N.; Georganos, S.; Grippa, T.; Lennert, M.; Vanhuysse, S.; Wolff, E. Fully Convolutional Networks and Geographic Object-Based Image Analysis for the Classification of VHR Imagery. *Remote Sensing* **2019**, *11*. doi:10.3390/rs11050597.
84. Mohammadi, S.; Belgiu, M.; Stein, A. Improvement in crop mapping from satellite image time series by effectively supervising deep neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing* **2023**, *198*, 272–283. doi:https://doi.org/10.1016/j.isprsjprs.2023.03.007.
85. Shelhamer, E.; Long, J.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2017**, *39*, 640–651. doi:10.1109/TPAMI.2016.2572683.
86. Himanshu, S. CSR GRACE and GRACE-FO RL06 Mascon Solutions v02. doi: 10.15781/cgq9-nh24.
87. Yi, S.; Sneeuw, N. Filling the data gaps within GRACE missions using singular spectrum analysis. *Journal of Geophysical Research: Solid Earth* **2021**, *126*, e2020JB021227.
88. Lea, C.; Vidal, R.; Reiter, A.; Hager, G.D. Temporal Convolutional Networks: A Unified Approach to Action Segmentation. *CoRR* **2016**, *abs/1608.08242*, [1608.08242].
89. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.Y.; Wong, W.k.; Woo, W.c. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1; MIT Press: Cambridge, MA, USA, 2015; NIPS'15, p. 802–810.
90. Torres, D.L.; Turnes, J.N.; Soto Vega, P.J.; Feitosa, R.Q.; Silva, D.E.; Marcato Junior, J.; Almeida, C. Deforestation Detection with Fully Convolutional Networks in the Amazon Forest from Landsat-8 and Sentinel-2 Images. *Remote Sensing* **2021**, *13*. doi:10.3390/rs13245084.
91. Moncrieff, G.R. Continuous Land Cover Change Detection in a Critically Endangered Shrubland Ecosystem Using Neural Networks. *Remote Sensing* **2022**, *14*. doi:10.3390/rs14122766.
92. Ortega Adarme, M.X.; Costa, G.A.O.P.; Feitosa, R.Q. MULTI-ATTENTION GHOSTNET FOR DEFORESTATION DETECTION IN THE AMAZON RAINFOREST. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2022**, *V-3-2022*, 657–664. doi:10.5194/isprs-annals-V-3-2022-657-2022.
93. Ashiagbor, G.; Asare-Ansah, A.O.; Amoah, E.B.; Asante, W.A.; Mensah, Y.A. Assessment of machine learning classifiers in mapping the cocoa-forest mosaic landscape of Ghana. *Scientific African* **2023**, *20*, e01718. doi:https://doi.org/10.1016/j.sciaf.2023.e01718.
94. Ortega Adarme, M.; Doblas Prieto, J.; Queiroz Feitosa, R.; De Almeida, C.A. Improving Deforestation Detection on Tropical Rainforests Using Sentinel-1 Data and Convolutional Neural Networks. *Remote Sensing* **2022**, *14*. doi:10.3390/rs14143290.
95. Dascălu, A.; Catalão, J.; Navarro, A. Detecting Deforestation Using Logistic Analysis and Sentinel-1 Multitemporal Backscatter Data. *Remote Sensing* **2023**, *15*. doi:10.3390/rs15020290.
96. Slagter, B.; Reiche, J.; Marcos, D.; Mullissa, A.; Lossou, E.; Peña-Claros, M.; Herold, M. Monitoring direct drivers of small-scale tropical forest disturbance in near real-time with Sentinel-1 and -2 data. *Remote Sensing of Environment* **2023**, *295*, 113655. doi:https://doi.org/10.1016/j.rse.2023.113655.
97. Solórzano, J.V.; Mas, J.F.; Gallardo-Cruz, J.A.; Gao, Y.; Fernández-Montes de Oca, A. Deforestation detection using a spatio-temporal deep learning approach with synthetic aperture radar and multispectral images. *ISPRS Journal of Photogrammetry and Remote Sensing* **2023**, *199*, 87–101. doi:https://doi.org/10.1016/j.isprsjprs.2023.03.017.