The Impact of Investment on Irrigated Rice, Dryland Agriculture and Afforestation in Senegal using SAR and Optical Time-Series Imagery in Data Fusion Approaches

Source: Gray Tappan, collaborator

NASA JSW
LCLUC Science Team Meeting
May 8, 2023

Project Team

Principal Investigator: **Dr. Christopher Neigh** - NASA GSFC - Greenbelt, MD

Science PI: **Dr. Konrad Wessels** - George Mason University - Fairfax, VA

Co-Investigators:
- **Dr. Molly Brown** - University of Maryland - College Park, MD
- **Dr. Mark Carroll** - NASA GSFC - Greenbelt, MD
- **Dr. Nathan Thomas** - NASA GSFC/University of Maryland - College Park, MD
- **Dr. Akiko Elders** - NASA GSFC/Morgan State University - Greenbelt, MD

Collaborators:
- **Dr. Pete Bunting** - Aberystwyth University - Wales
- **Dr. Gray Tappan** - USGS - Sioux Falls, SD
- **Dr. Renaud Mathieu** - International Rice Research Institute - Philippines
- **Dr. Paul Montesano** - NASA GSFC/ADNET Systems - Greenbelt, MD

Technical Staff:
- **Margaret Wooten** - NASA GSFC/SSAI - Greenbelt, MD
- **Jordan Caraballo-Vega** - NASA GSFC - Greenbelt, MD
- **Minh Tri Le** - George Mason University - Fairfax, VA
- **William Wagner** - NASA GSFC/SSAI - Greenbelt, MD

Outline

- Introduction & Project Objectives
- Model Development and Experimentation
- Current Results
- Next Steps: Model and technology transfer, 1D CNN

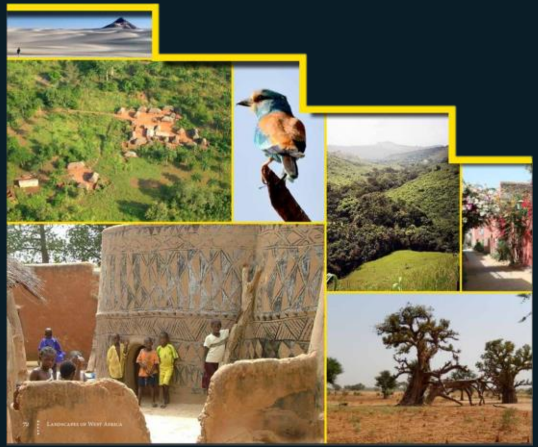
- HLS update
- Landsat Next update

Introduction

Over the past 30 years, West Africa's (WA) population has more than doubled and is an epicenter of land cover change.

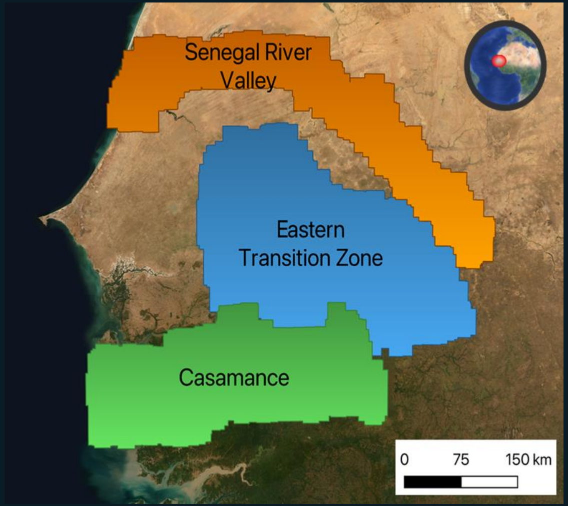
- By 2030, 490 million people are projected to live in WA, which will stress food production and natural resources.
- This hotspot has not been sufficiently documented, due to strong phenology changes in the Sahel from wet and dry seasons and the inability of multi-temporal moderate resolution observations to adequately resolve sub-hectare agriculture change.
- Senegal has experienced widespread agriculture extensification in dryland regions, expansion of irrigated rice cultivation, and forest loss and conversion to agriculture and agroforestry.
- Changes have been reported but not sufficiently mapped and quantified with current standard remote sensing methodologies.

We contend that a multi-sensor data fusion approach utilizing Deep Learning (DL) techniques, including Convolutional Neural Networks (CNNs), in NASA's High-Performance Computing (HPC) environment can significantly improve policy-relevant land cover and land use change information in Senegal over the past decade.



eros.usgs.gov/westafrica

Project Objectives



Our study is focused on three hotspots in Senegal:

(i) Senegal River Valley (SRV; 4.6K km^2)
(ii) Expanded Eastern Transition Zone (ETZ; 46.2K km^2)
(iii) Casamance (CAS; 23.8K km^2);

With the goals of:

1. Quantifying changes in the extent and intensity of irrigated rice and dryland agriculture;
2. Testing CNNs on VHR data for extracting croplands and individual trees at regional scales; and
3. Assess agroforestry and reforestation in degraded fields using time-series SAR (2015 to the present) and VHR data (2010 to the present).

NASA MUSLI Program Broad Question - Multiresolution/Multimodal data fusion

Can we reliably detect LCLUC with VHR CNNs, Landsat, S1 & S2 time-series data in a highly dynamic ecosystem?

Current LCLUC Mapping Challenges in Senegal

- Can 2m VHR imagery enhance our understanding of changes in the extent, intensity and land use of agriculture and forestry in Senegal?
- How can we better take advantage of NASA's HEC resources to apply Deep learning for land cover change monitoring?
- Can we scale-up Unet CNNs to compensate for the diversity of landscapes and images to map regional land cover?

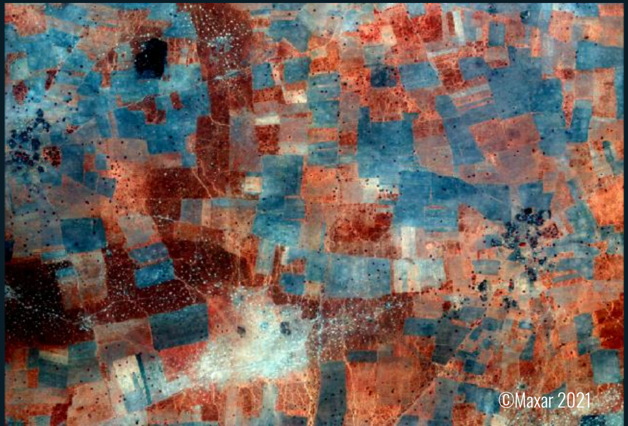

Wet and dry seasonality for cluster of typical fields in Senegal, Photos from Collaborator Gray Tappan

Challenges: highly variable image conditions

- Abundant burnt areas of various ages in croplands and natural veg
- Phenology, active crops and other natural vegetation spectrally indistinguishable in wet season images
- Fallow fields revert to “natural” herbaceous cover and back to active croplands - problem for change monitoring.

07/14/2012

03/14/2021



04/26/2021

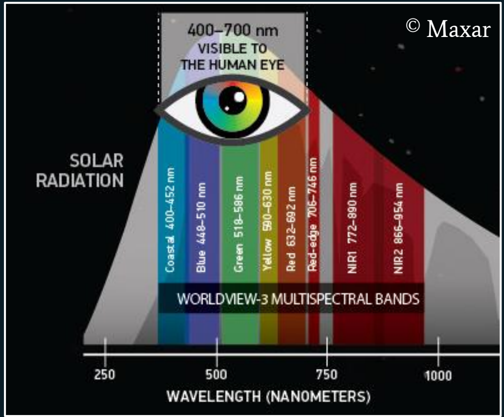
WorldView Data Volume

Number of Observations

Region	N Scenes	N Strips
SRV	1992	714
ETZ	3356	611
CAS	2713	711
Total	8061	2036

X

8 Multispec Bands



- Coastal Blue
- Blue
- Green
- Yellow
- Red
- Red-edge
- NIR-1
- NIR-2

Data

Intermediate: 10.9 TB | 13.4 TB | 8.4 TB → 32.7 TB

Final: 2.6 TB | 3.2 TB | 2 TB → 7.8 TB

CAS
ETZ
SRV

Total Volume = over 40 TB of data!

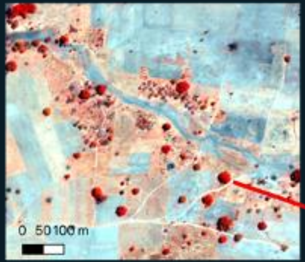
Outline

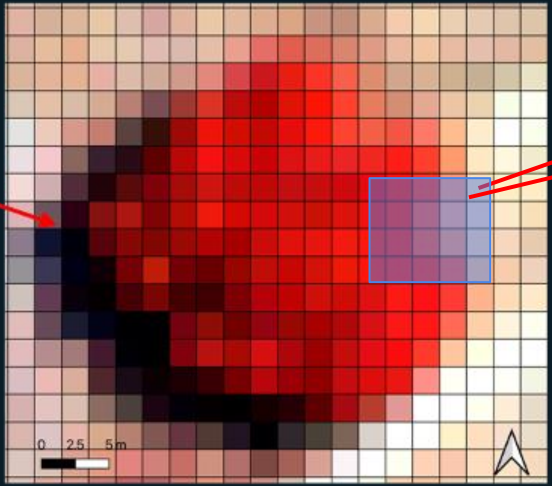
- Introduction & Project Objectives
- Model Development and Experimentation
- Current Results
- Next Steps: Model and technology transfer, 1D CNN

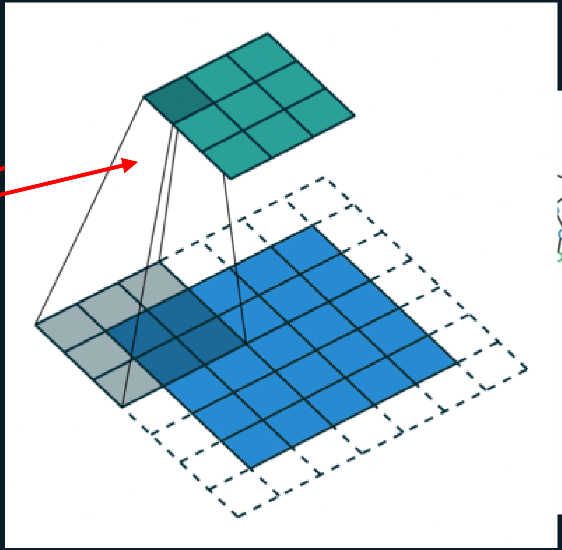
- HLS update
- Landsat Next update

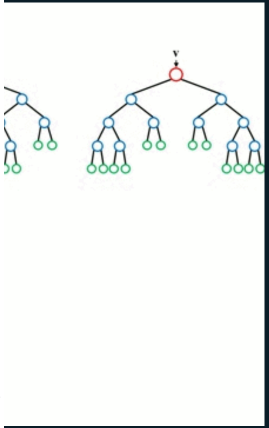
Deep Learning Advantage

- Deep learning is in a mature state and is now feasible with existing computational resources.
- Unlike traditional methods such as Random Forest, spatial features are considered, and patterns are extracted from the data.
- There is a reiteration process that improves the generalization of the algorithms.
- Convolutional Neural Networks (CNNs) combine many layers to extract these patterns.

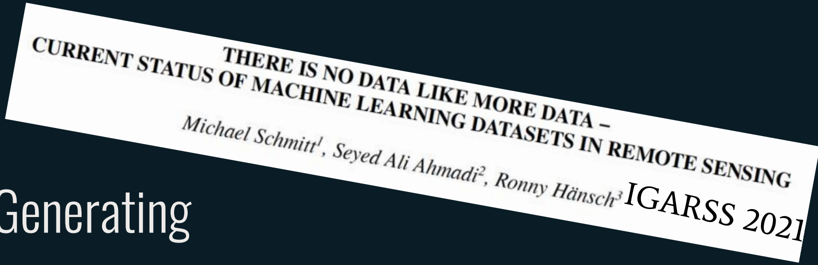








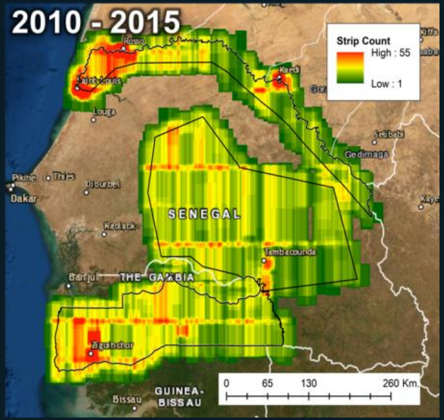
Challenges and Research questions

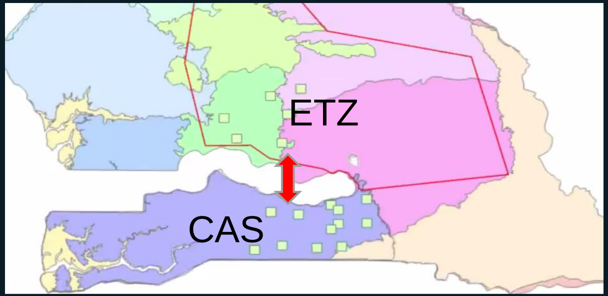

THERE IS NO DATA LIKE MORE DATA –
CURRENT STATUS OF MACHINE LEARNING DATASETS IN REMOTE SENSING

Michael Schmitt[1], Seyed Ali Ahmadi[2], Ronny Hänsch[3]

IGARSS 2021

- Unet CNN requires fully classified image as training. Generating large amounts of training data is laborious and expensive, especially manual annotation.
- Machine learning requires large amounts of representative training data. How much data is enough - when does accuracy stop improving with additional training data?
- Can a Unet model trained in one area be transferred to another region with minimal extra training data from new area? - Transfer learning, where only the final layers of the U-Net architecture are train with additional data





CNN Model Development: negotiating space and time

Model Training

Time – Epochs/Years

Time – Seasonality

Location

Generalized Model

Model Implementation

All available data

- A generalized model uses training from any scenario for model implementation
- Can this model provide reliable accuracy in time and space?

VS.

- Locally optimized models which require a different stratification approach of training/model development.

Model Training

Specific Epochs & Seasonality

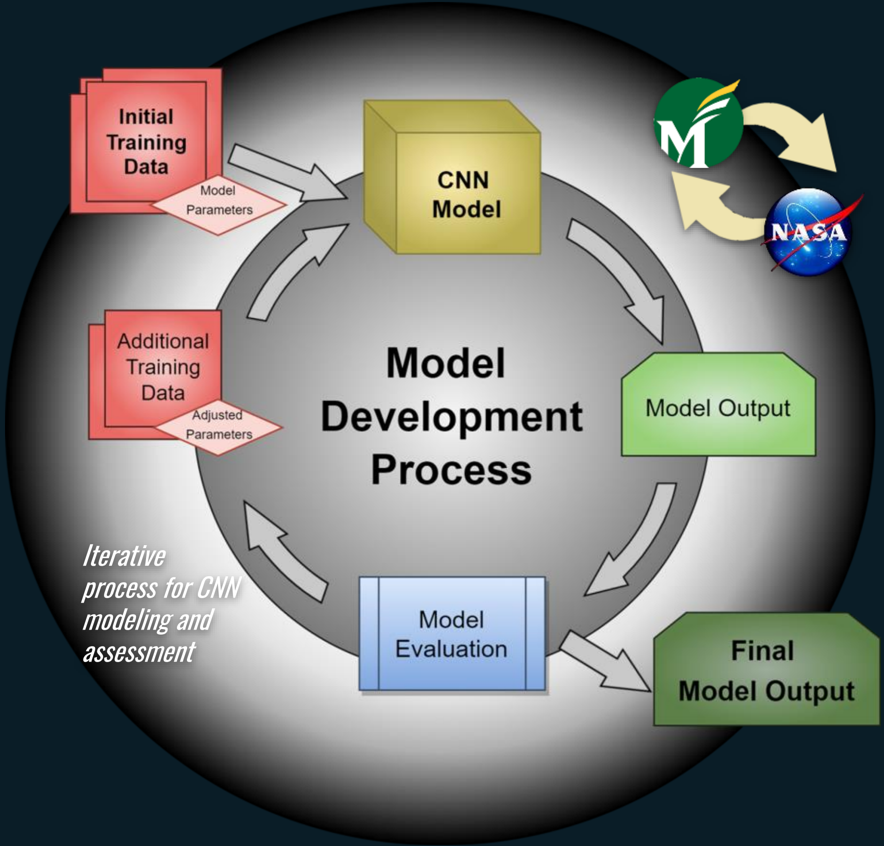
Specific Location

Many Localized Models

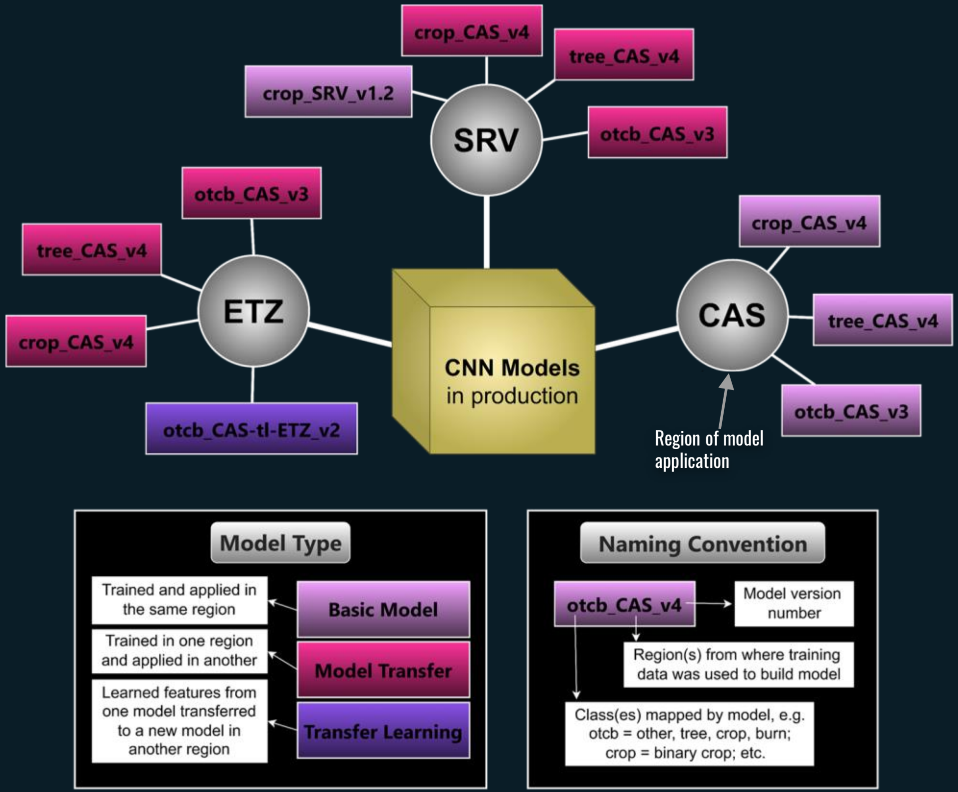
Model Implementation

Specific Data

CNN Model Development = Many, Many Experiments...



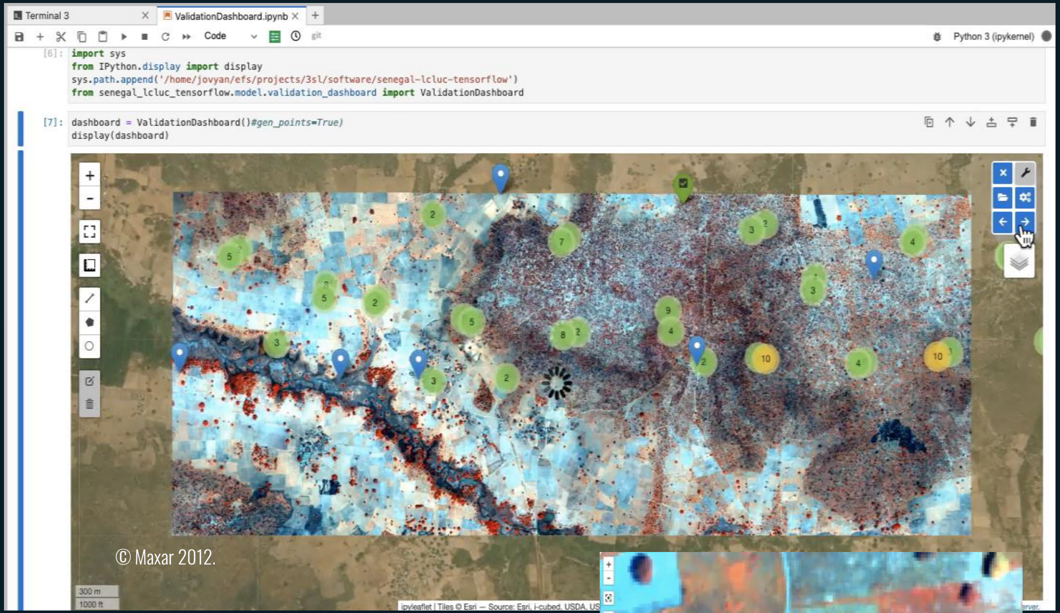
Iterative process for CNN modeling and assessment



Multiclass vs. Binary Ensemble

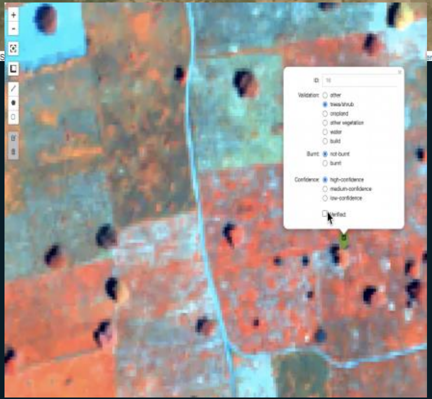
GMU Remote Sensing Students Training & Validation Tool

- NASA developed a cloud-native system for the visualization and validation of WV imagery.
- The system allows multiple operators to simultaneously validate given points following Oloffson (2014) guidelines.
- The system was deployed in AWS as a Jupyter Notebook.



- 200 points per image ~ 35 000 records
- 3 independent observers
- confidence and inter-observer variability captured
- 437 person hours on 59 images
- 12 operators, GMU students / waged employees
- 63 more images to be validated

Validation tool: Cloud Optimized GeoTIFFs (COGs) for faster rendering



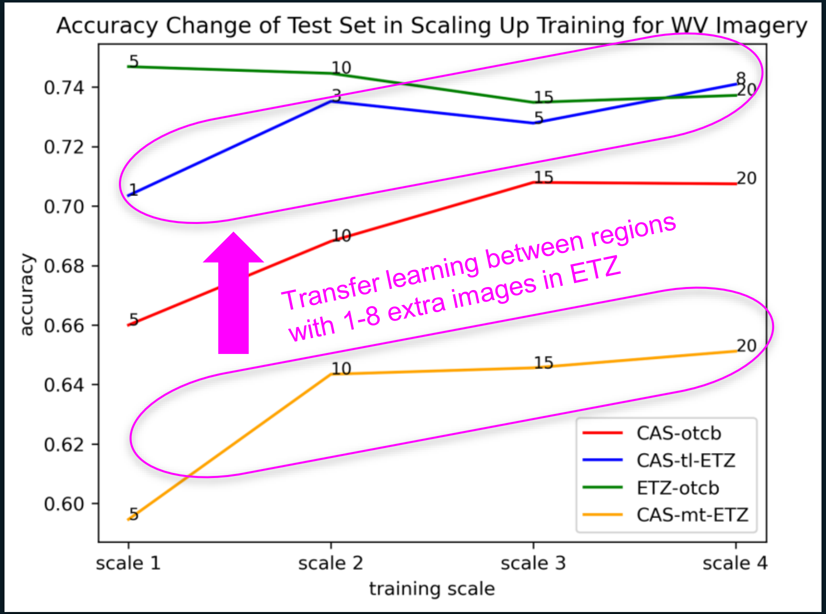
Outline

- Introduction & Project Objectives
- Model Development and Experimentation
- Current Results
- Next Steps: Model and technology transfer, 1D CNN

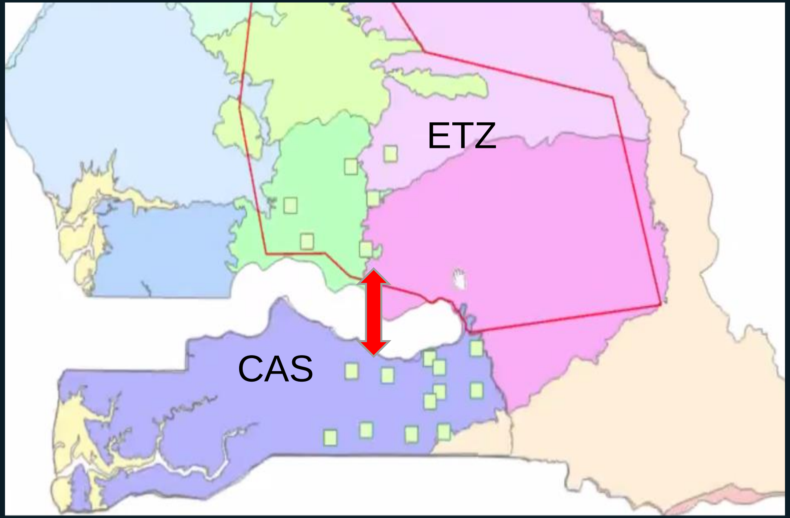
- HLS update
- Landsat Next update

Preliminary Results: multiclass model transfer (mt) vs. transfer learning (tl)

- Scaling up with more training: 5, 10, 15, 20, 30 images, improves accuracy, but peaks at 70-74% (green line below)
- Transfer learning (CAS-tl-ETZ) achieves high accuracy after adding just 3-8 images from ETZ to CAS model (yellow to blue line below)
- Independent validation is currently underway.



Accuracy on test images (never included in training)

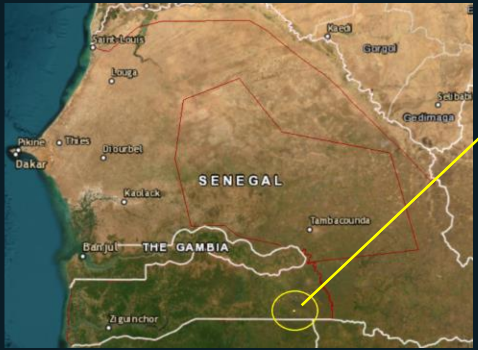


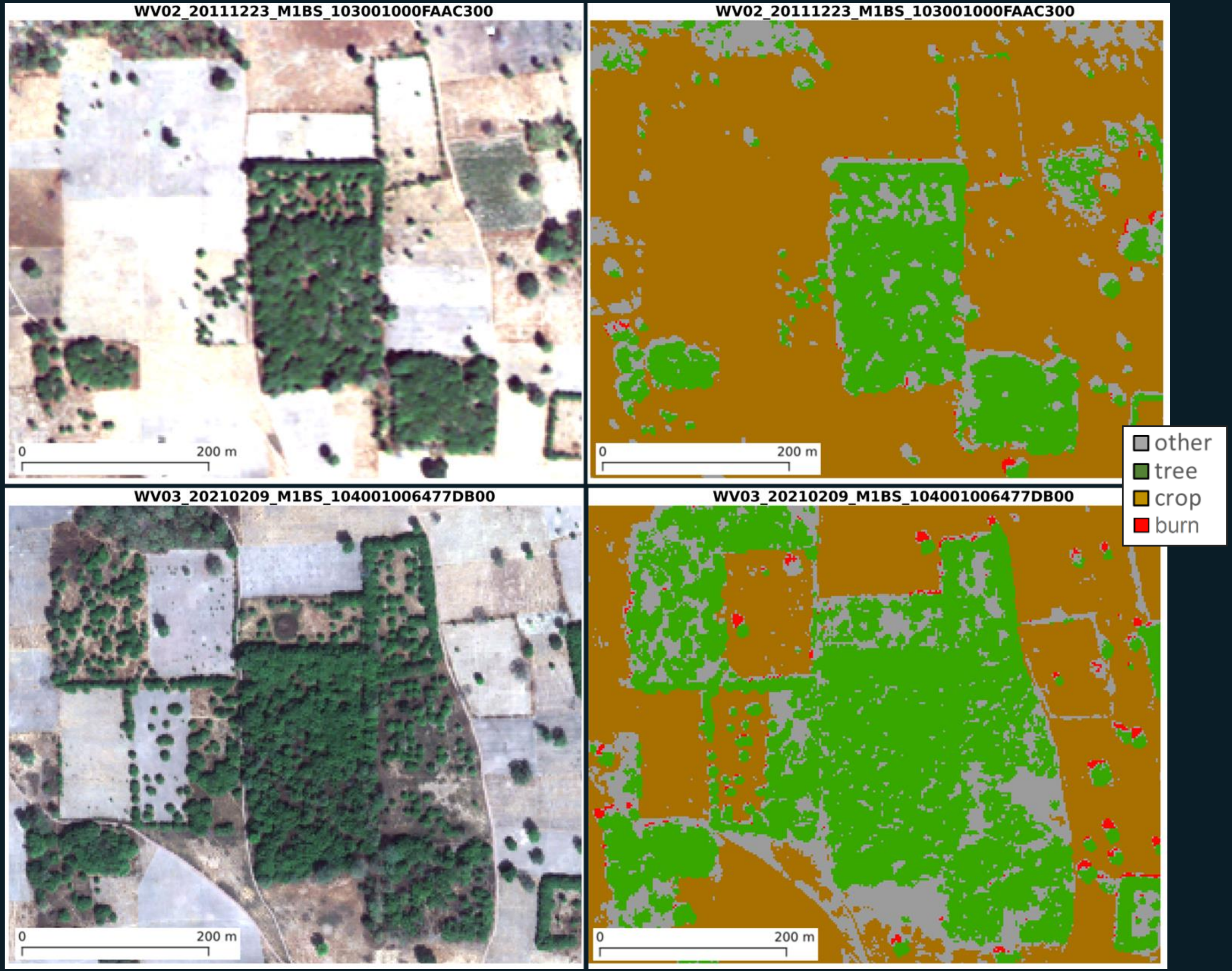
Current Results: CAS example

Basic Multiclass Model:
Trained and applied in the same region

December 2011

February 2021



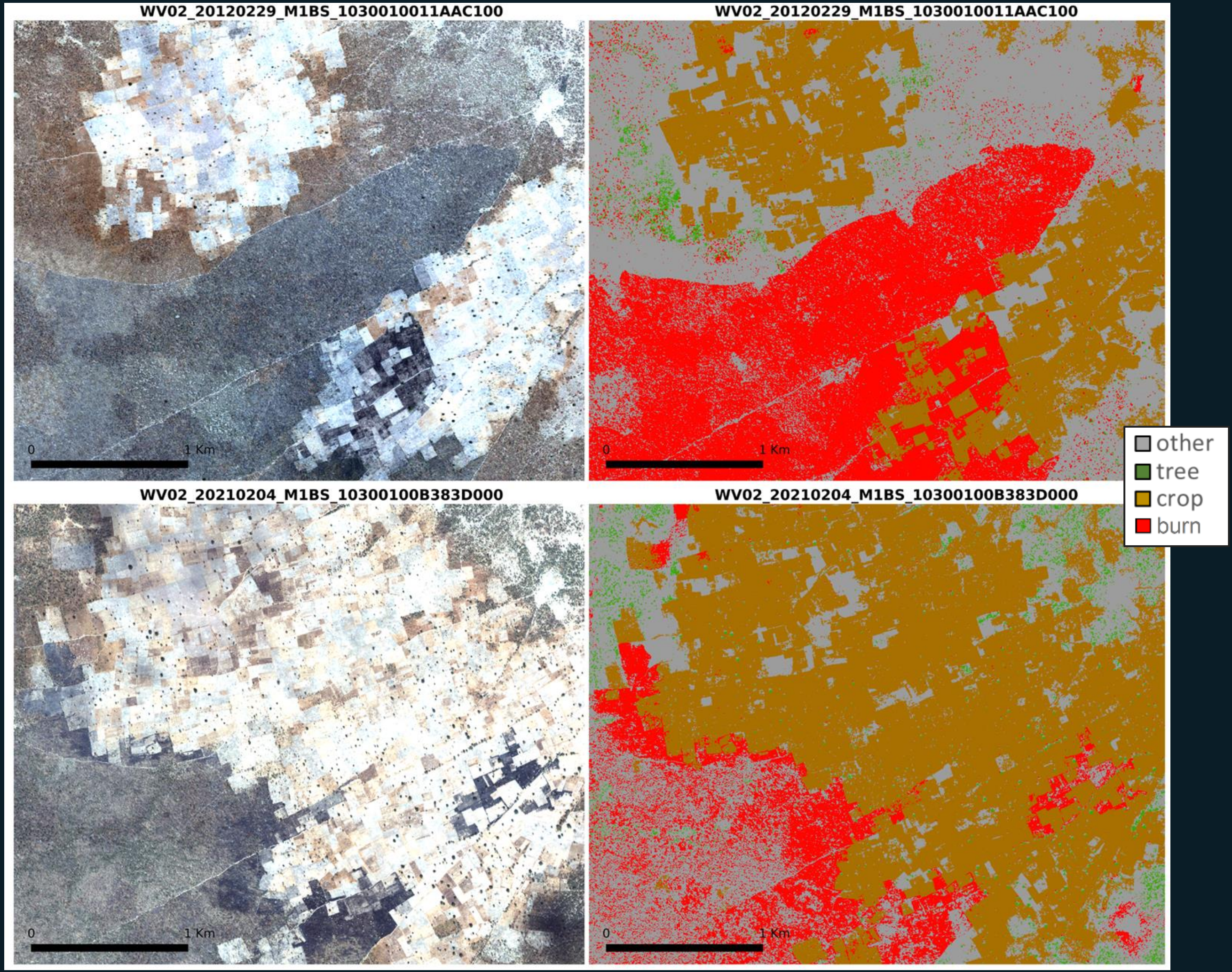


Current Results: CAS example

Basic Multiclass Model:
Trained and applied in the same region

February 2011

February 2021

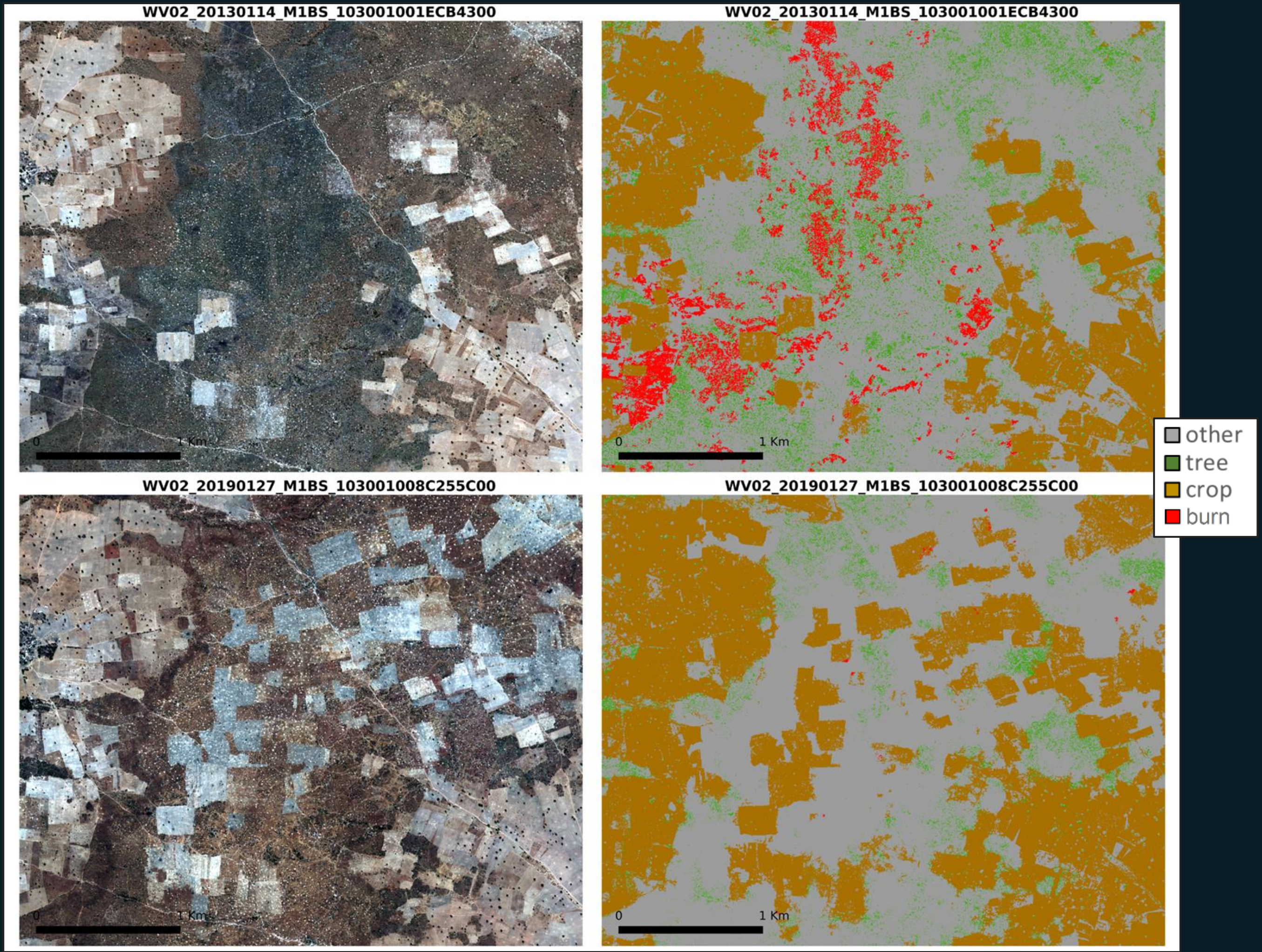


Current Results: ETZ example

Basic Multiclass Model:
Trained and applied in the same region

January 2013

January 2019



Current Results: Overall Validation

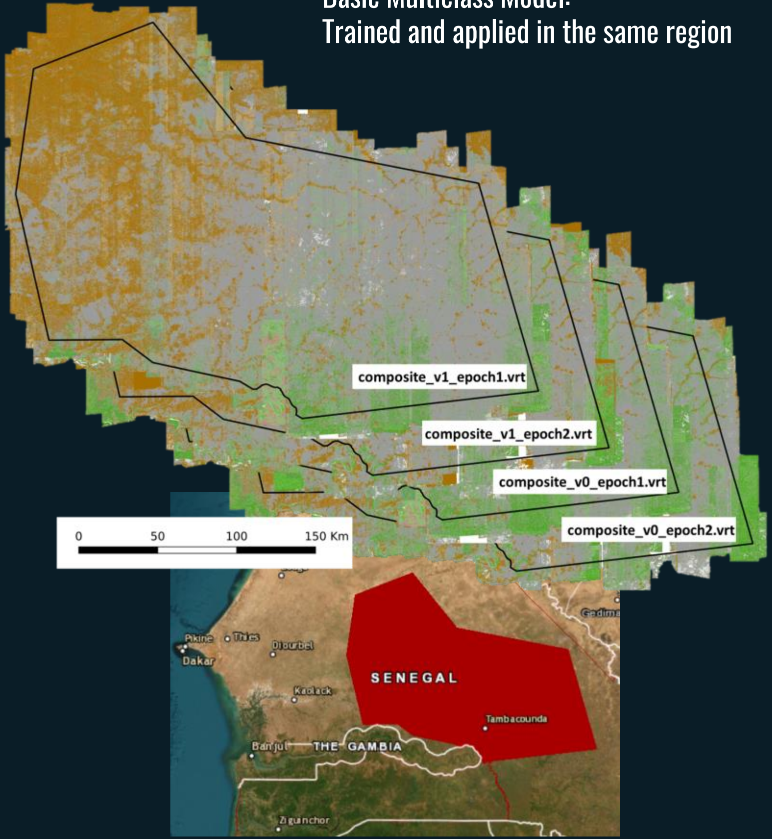
- Our validation statistics are taken from validation points where all observers agreed on.
- We compared the output of a multiclass model, with the output of an ensemble of a multiclass and binary tree (task specific) model.
- Output metrics suggest a multistep approach of binary ensemble is beneficial to improve overall accuracy.

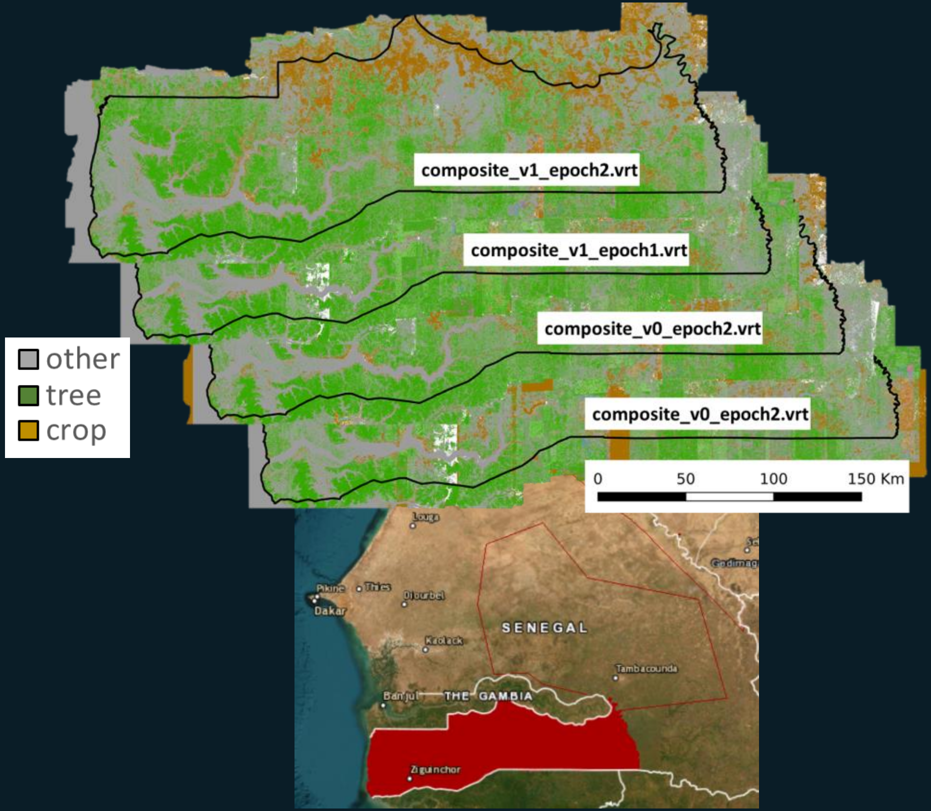
Individual Model		Reference				
		Other	Tree	Crop	Total	User's Acc.
Map	Other	394	14	291	699	56%
	Tree	611	1119	23	1753	64%
	Crop	61	5	660	726	91%
	Total	1066	1138	974	Overall Acc.	68%
	Prod. Acc.	40%	98%	68%	Overall Prec.	78%

With binary tree model ensemble		Reference				
		Other	Tree	Crop	Total	User's Acc.
Map	Other	388	21	290	699	56%
	Tree	443	1294	19	1756	74%
	Crop	61	5	660	726	91%
	Total	892	1320	969	Overall Acc.	74%
	Prod. Acc.	43%	98%	68%	Overall Prec.	79%

Preliminary Composites

Basic Multiclass Model:
Trained and applied in the same region



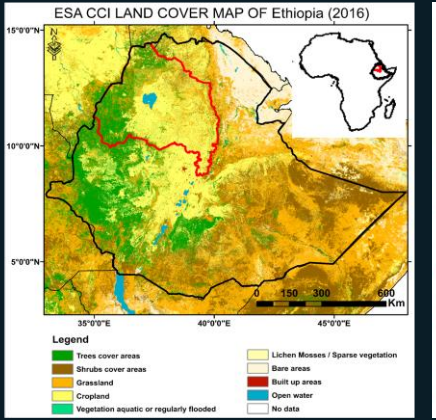


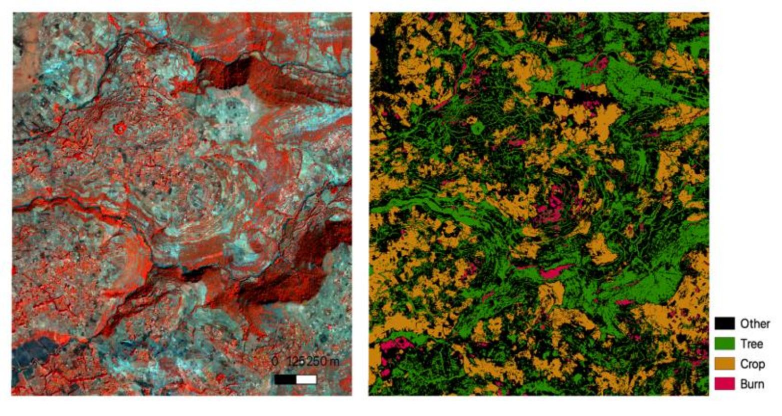
Outline

- Introduction & Project Objectives
- Model Development and Experimentation
- Current Results
- Next Steps: Model and technology transfer, 1D CNN

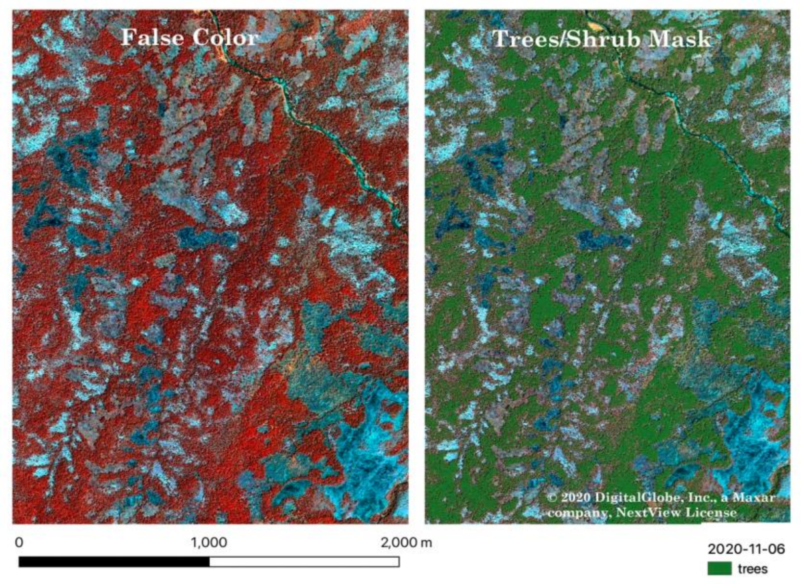
- HLS update
- Landsat Next update

Current Results: Zero-shot Transfer to Downstream Tasks in Ethiopia





- Located on the northwestern highlands of Ethiopia, the Amhara Region occupies one-fifth of the area of the country (> 160,000 km^2), whole supporting more than one-third of the national cereal crop production.
- We performed zero-shot learning using our models trained in Senegal across the Amhara Region.
- Next steps include using pre-trained models from Senegal, and perform transfer learning to train the model with additional classes of interest in the Ahmara.



NASA NPP: woubet.g.alemu@nasa.gov

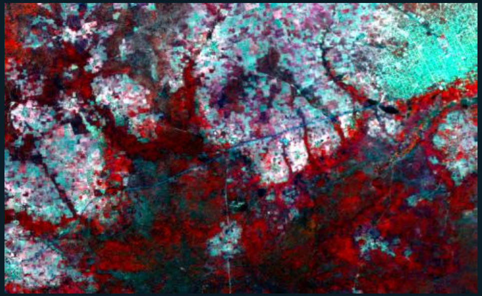
Technology Transfer & Outreach

Institut Sénégalais de Recherches Agricoles

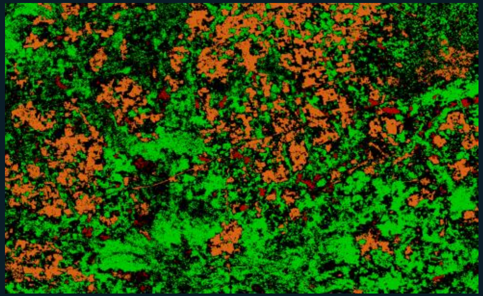
EARTH OBSERVATIONS FOR ECOSYSTEM ACCOUNTING

US Partnerships Initiative funding opportunity proposal might fund collaboration with ISRA. The proposal would fund:

- Downscaling CNN model to moderate resolution PlanetScope and Harmonized Landsat-Sentinel data
- Exploration of use of CNN model on UAV data generated by ISRA personnel
- Exchange of personnel from US and Senegal to extend knowledge
- Training of 20 Senegalese scientists on CNN methods and their use in land use classification relevant to Senegal agriculture and forest



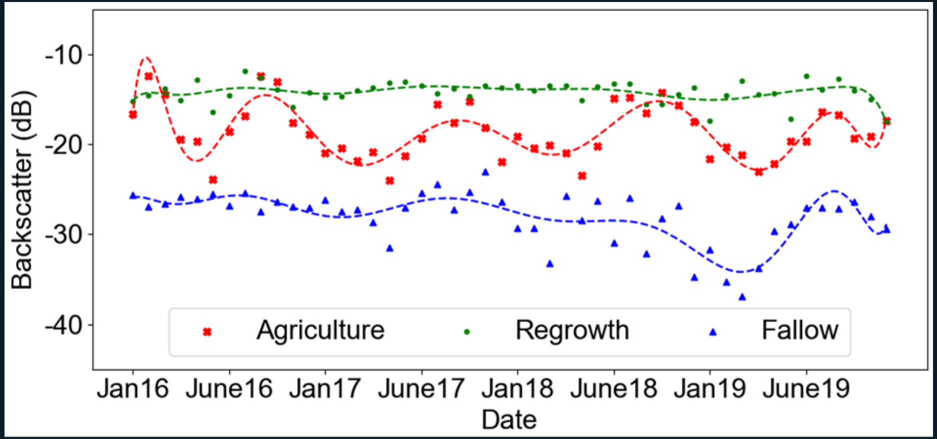
Planet (Open Access Monthly VNIR Mosaics)
https://www.planet.com/nicfi/



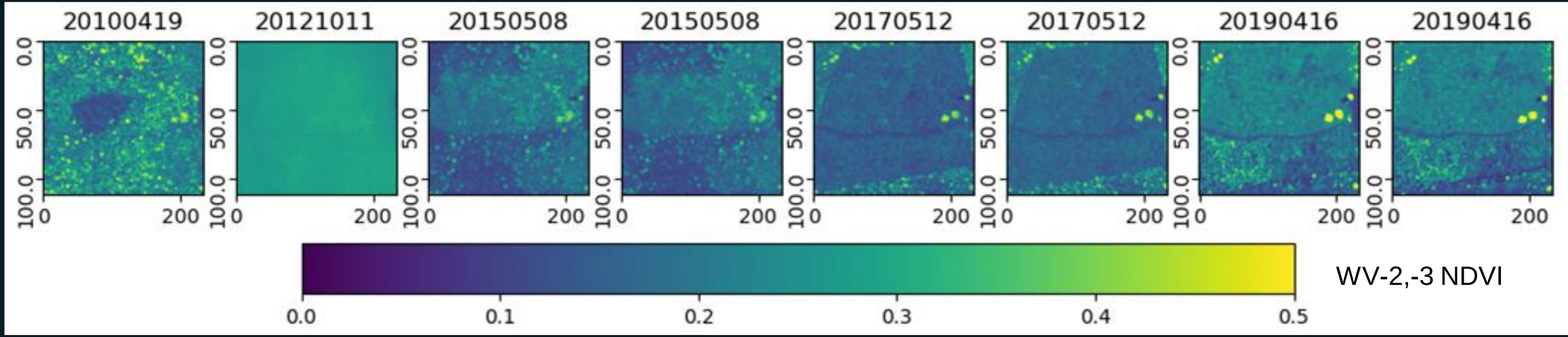
Preliminary Landcover output (CNN model transfer)

Classifying potential Land-Use signals with VHR/SAR Fusion

1D CNN to classify drivers of change with thematic VHR objects and S1 time-series backscatter



Maxar/Google Earth with Example of Hansen et al. Forest Loss polygons



Take-Aways from Preliminary Results

- Large area VHR land cover change mapping at 2 m (Currently 70-74% overall accuracy)
- We have built a Research Platform to test CNNs and scale up 2m change mapping:
    - Methods and systems have successfully been developed to scale-up Unet image segmentation on many (1,000s) VHR images.
        - *Transferred learning experiments between regions are promising and require limited additional training.*
    - Training accuracy 88%, testing accuracy 74%, additional experiments are currently underway.
    - Independent validation protocols have been scaled-up and run in the cloud by external users.
    - Optical/SAR Data Fusion with a 1-D CNN being implemented to understand land use change and intensity.
    - 2m mapping captures subtle LCC from smallholder agriculture that is difficult to distinguish with moderate resolution data.
    - Some unexplainable low accuracy of CNN classification exists on some images, and challenges remain.
- ISRA Tech Transfer

Project Deliverables

Presentations:

Caraballo-Vega, J.A., et al. (2023, January 24). Utilizing GPUs in Machine Learning for Earth Sciences. Earth Science Information Partners (ESIP) Meeting [Presentation]. Online.

Caraballo-Vega, J.A., et al. (2023, February 1). Advances in Vegetation Height Estimates from WorldView Multispectral Imagery using Convolutional Neural Networks for Regression [Poster]. Artificial Intelligence Center of Excellence Seminar [Presentation]. Greenbelt, MD, United States.

Neigh, Christopher SR, et al. (2023, February 17). The Impact of Investment on Irrigated Rice, Dryland Agriculture and Afforestation in Senegal using SAR and Optical Time-Series Imagery in Data Fusion Approaches. LCLUC Webinar Series [Presentation]. Online.

Caraballo-Vega, J. A. (2023, March 29). GPU Accelerated Deep Learning Pipelines for Large-scale Very High-Resolution Imagery Analysis [Lightning Talk]. Joint Agency Commercial Imagery Evaluation (JACIE) Workshop. Reston, VA, United States.

Caraballo-Vega, J.A., et al. (2023, March 21). Advances in Vegetation Height Estimates from WorldView Multispectral Imagery using Convolutional Neural Networks for Regression [Poster]. 3rd SMD and ETD Workshop on A.I. and Data Science: Leaping Toward Our Future Goals. Greenbelt, MD, United States.

Wooten, Margaret, Christopher SR Neigh, Jordan Caraballo-Vega, Mark Carroll, Nathan Marc Thomas, Minh Tri Tri Le, and Konrad J. Wessels. Mapping Land Cover Change Across Senegal with Convolutional Neural Networks and Very High Resolution Data [Poster]. In Fall Meeting 2022. AGU, 2022.

Le, Minh Tri, Konrad J. Wessels, Jordan Caraballo-Vega, Nathan Marc Thomas, Margaret Wooten, Mark Carroll, and Christopher SR Neigh. Training Strategies of CNN for Land Cover Mapping with High Resolution Multi-spectral Imagery in Senegal [Poster]. In Fall Meeting 2022. AGU, 2022.

Alemu, Woubet G., et al. An ensemble of Convolutional Neural Networks for Land Use Land Cover Classification in the Amhara Region of Northwest Ethiopia using Very High-Resolution Commercial Satellite Imageries [Online Poster]. Fall Meeting 2022. AGU, 2022.

Caraballo-Vega, Jordan, et al. (2022, November 13-18). Advancing Geospatial Data Structures on GPUs: Mapping the Earth at Fine Scale. The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC22). Dallas, TX, United States. https://www.nas.nasa.gov/SC22/research/project14.html.

Caraballo-Vega, Jordan, Paul Montesano, Mark Carroll, Christopher SR Neigh, Minh Tri Tri Le, Konrad Wessels, Margaret Wooten, and Zachary Williams. Advances in Vegetation Height Estimates from WorldView Multispectral Imagery Using Deep Learning [Online Poster]. In Fall Meeting 2022. AGU, 2022.

Alemu, Woubet, et al. "Land Use Land Cover Classification in the Amhara Region, Northwest Ethiopia, Using Convolutional Neural Networks." AGU Fall Meeting Abstracts. Vol. 2021. 2021.

Caraballo-Vega, Jordan, et al. "Towards Scalable & GPU Accelerated Earth Science Imagery Processing: An AI/ML Case Study." AGU Fall Meeting Abstracts. Vol. 2021. 2021.

Caraballo-Vega, Jordan, et al. "Ensemble Learning Methods & Deep Learning for the Task of Cloud Detection." AGU Fall Meeting Abstracts. Vol. 2021. 2021.

Publications:

Caraballo-Vega, J. A., Carroll, M. L., Neigh, C. S. R., Wooten, M., Lee, B., Weis, A., Aronne, M., Alemu, W. G., & Williams, Z. (2023). Optimizing WorldView-2, -3 cloud masking using machine learning approaches. Remote Sensing of Environment, 284, 113332. https://doi.org/10.1016/j.rse.2022.113332

Alemu, W. G., & Neigh, C. S. (2022). Desert Locust Cropland Damage Differentiated from Drought, with Multi-Source Remote Sensing in Ethiopia. Remote Sensing, 14, 1723. https://doi.org/10.3390/rs14071723

Caraballo-Vega, Jordan A., et al. (2022). Remote Sensing Powered Containers for Big Data and AI/ML Analysis: Accelerating Science, Standardizing Operations. IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium. IEEE. https://doi.org/10.1109/IGARSS46834.2022.9883436

Outline

- Introduction & Project Objectives
- Model Development and Experimentation
- Current Results
- Next Steps: Model and technology transfer, 1D CNN

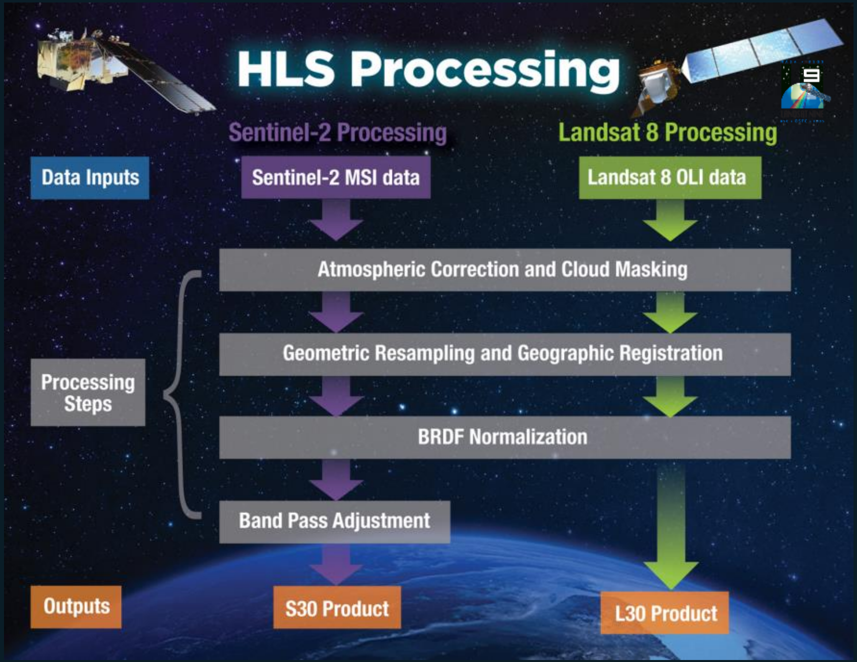
- HLS update
- Landsat Next update

HLS Update

- Global 2020 -> 2015 processing to July of 2018 of S2 has been completed and is available.
- Global 2018 -> 2015 processing of all the S2 be completed by this summer.
- Google has ingested 30% of HLS to date and it will hopefully be available this Fall in GEE.
- V2.0 publication is in prep and details about updates to the processing stream will be available on

https://hls.gsfc.nasa.gov

https://www.earthdata.nasa.gov/esds/harmonized-landsat-sentinel-2



Outline

- Introduction & Project Objectives
- Model Development and Experimentation
- Current Results
- Next Steps: Model and technology transfer, 1D CNN

- HLS update
- Landsat Next update

NASA’s Landsat Project Scientists

*Retired

Landsat Next Update

Mission Architecture concept definition is complete, and

Landsat Next is officially in Phase A!

Pre-Phase A - user surveys; architectural studies to determine science mission requirements

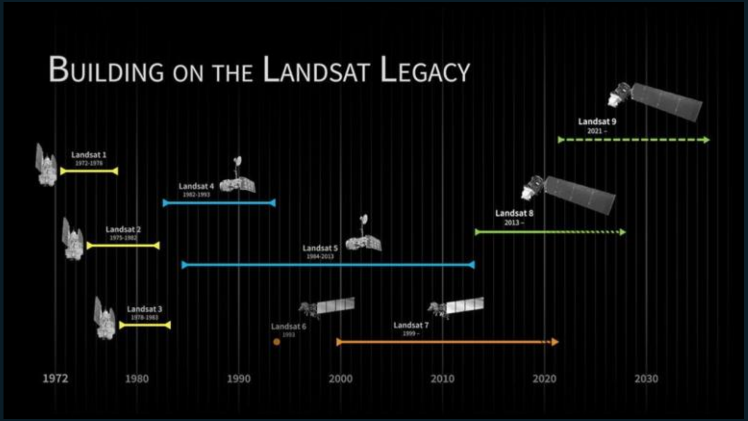
Phase A – science flow down to hardware requirements; architecture credibility and refinement

Phase B – preliminary design and technology completion

Phase C – final design and fabrication

Phase D –system assembly, integration/testing, and launch readiness;

Phase E – starts after on-orbit operational checkout and ends at the mission's operational end.



Ongoing Phase A Work:

- **Landsat Next Instrument Suite (LandIS)** is on the Mission's critical path the final RFP will be out this spring; expect to have a vendor on contract by early 2024.
- **NASA Spacecraft and Observatory Integration studies** to assess data management; instrument accommodation; and launch vehicle packaging.
- **USGS Ground System studies** to assess ground stations, data compression, constellation mission operations.

Landsat Next Triplet Constellation & Mission Overview

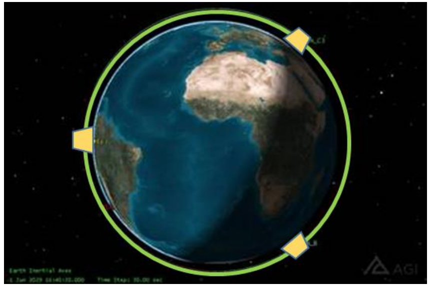
“Triplet” constellation consists of **three, identical observatories with equal orbit spacing.**

- Each observatory images a swath that will provide full global coverage.
- Three satellites improve revisit frequency and mission resiliency.
- Launch Readiness Date: **November 2030.**

Mission architecture is endorsed by Landsat user communities, including more than 12 Federal Agencies and Landsat Science Team.

- ❑ Temporal, spectral, and spatial **capabilities are increased approximately two fold**, while maintaining **historical Landsat data continuity.**
- ❑ Satisfies **primary user need of increased revisit frequency** with maximum resiliency.
    - ❑ 6 day revisit at equator with three observatories, and 2-3 d revisit at higher latitudes.
    - ❑ Provides resiliency at observatory level, and enables affordable sustainment option.

Parameter	Value
Mission Category	Category 2, Class B
Mission Life	5-Years
Altitude	653 km Sun-synchronous
Inclination	~98-degrees
Orbital Separation	120-degrees
Mean Local Time (MLT)	10:10 AM +/- 5-min
Repeating Ground Track	18-day
Constellation Aggregate	6-day
Swath Width	164 km
Half Angle FOV	7.2-degrees



Landsat Next improved spectral, spatial and temporal resolution enables numerous improvements in user applications.
https://landsat.gsfc.nasa.gov/satellites/landsat-next/

Thank You!

Ten-year comparison of a landscape near Keur Mandiaye, Saloum. Above: 15 January 1994. Below: 6 December 2004

PI:
christopher.s.neigh@nasa.gov