

# **Comparison of Land Cover Satellite Data to GLOBE Observer Utilizing Adopt-a-Pixel Methodology Toward the Development of a Novel Citizen Science Observation System**

Samuel Bawden, Nandini Khaneja, Neev Tamboli, Jordan Rodriguez

## **Abstract**

Accurate land cover classification is critical for environmental monitoring, urban planning, and climate research, yet commonly used satellite products often struggle to represent heterogeneous landscapes at fine spatial resolutions. This study evaluates the agreement between remotely sensed land cover datasets and ground-based citizen science observations collected using NASA's GLOBE Observer application. Utilizing the Adopt-a-Pixel 3km methodology (Low et al 2021), standardized areas of interest were observed across four diverse regions in the United States: Washington D.C.; New Hyde Park, New York; Fremont, California; and Weslaco, Texas. Ground observations were supplemented with high-resolution reference data generated using Collect Earth Online and compared against multiple satellite land cover products, including WorldCover, Dynamic World, ESRI, Landsat Time Series, and global tree canopy datasets. Results indicate that satellite products frequently over-generalize land cover in mixed urban and suburban environments, with WorldCover showing the highest overall agreement with ground-truthed observations. Limitations in the GLOBE Observer application, specifically reduced image resolution and inconsistent data retention, motivate the creation of a novel supplemental system designed to allow citizen scientists to generate usable observation data.

## **Introduction**

Satellite-derived land cover data are widely used to study environmental change, ecosystem dynamics, and urban expansion. However, classification accuracy is often limited by spatial resolution, frequency of observation, and generalization, particularly in landscapes containing a mix of developed and vegetated land cover types or where shadows are present. Ground-based validation is therefore essential to assess and improve the reliability of these datasets.

Citizen science is the practice of allowing common, non-scientifically affiliated citizens to contribute to the research landscape and utilizing these contributions to fuel public scientific literacy and increase the scalability of scientific observations and claims. Citizen science is most commonly practiced through public data collection/reporting, surveys, and observations of local conditions. The practice is incredibly important to the scientific community because, by involving more people in the scientific process, scientists gain access to vast amounts of data (improved and more informed analysis), are able to tackle research questions on greater scales, and can overcome logistical constraints such as geographic or political limitations.

NASA's GLOBE Observer App is a citizen science initiative that uses a mobile app to allow citizens to collect environmental data, primarily focusing on land cover, mosquito, tree

canopy, and cloud observations to support NASA's Earth science research. The app also allows citizens to upload their collected data to large GLOBE databases from which researchers can select data to be used in studies and, in recent years, data-driven AI models. GLOBE is one of the world's most used citizen science platforms, and its use in increasing the availability and reliability of environmental data is used by scientists all over the world.

## **Methods**

### **Adopt-a-Pixel Methodology**

The GLOBE Observer mobile app tracks a user's location and guides them to capture images facing North, South, East, West, Up, and Down. By recording both the user's position and the directional context of each photo, GLOBE observations provide valuable metadata that researchers can compare with remotely sensed data. However, while GLOBE includes geolocation services, it lacks a built-in framework to define larger areas of study beyond the 50-meter scale recommended for land cover classification.

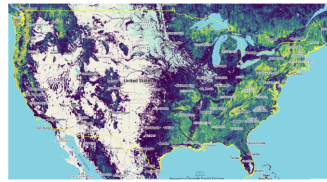
The Adopt a Pixel 3 km methodology (Low et al. 2021) proposes a solution to this issue. In this approach, researchers define a standardized Area of Interest (AOI) measuring 3 km by 3 km. Each AOI is divided into a 6 by 6 grid, creating 36 coordinate points evenly spaced 500 m apart. At each point, a 100 m by 100 m square, called a primary sampling unit (PSU), is centered. A 37th square, known as the centroid, is added at the exact center of the AOI, bringing the total number of sampling units to 37.

Once applied, the Adopt a Pixel 3 km methodology clearly delineates a study area and enables methodical, repeatable observations across defined spatial units. Its structured yet accessible format makes it well suited for both researchers and citizen scientists. In addition to the ease of comparison between in-person observed data and remotely sensed data that the methodology offers, the short distances between sampling points encourage local participation, helping to democratize environmental monitoring and foster broader engagement in Earth sciences.

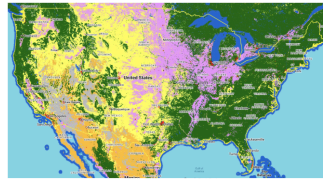
This methodology has been particularly useful in educational and citizen science contexts. As part of the NASA SEES internship, students implemented the Adopt a Pixel 3 km Methodology in conjunction with the GLOBE Observer app to define AOIs and investigate local land cover conditions. Our team's field sites sampled diverse environments across the United States, including Washington, D.C.; New Hyde Park, NY; Weslaco, TX; and Fremont, CA. This implementation demonstrated the feasibility of the methodology for collecting consistent, scalable data across multiple geographic regions.

In addition to improving spatial coverage, citizen science observations capture lived landscape complexity that is often invisible to algorithmic classification. App users must navigate across private property boundaries and interactions with community members in the field, shaping where and how the data are collected. These human factors are not necessarily sources of error but intrinsic components of real-world environmental monitoring, providing essential context for interpreting discrepancies between satellite and ground data.





Meta/WRI Global Canopy Height



WorldCover 10m 2020/2021 (ESA)



Dynamics World



ESRI 2017/2024 Land Cover

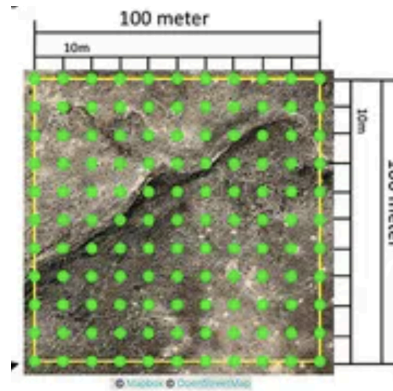
\*Figures 4-6: Land cover and canopy height visualizations generated using EarthMap (Google Earth Engine–based platform), incorporating datasets including ESA WorldCover, Meta/WRI Global Canopy Height, Dynamic World, and ESRI Land Cover.

- **Google Earth Engine Apps: LandSat Time Series.** Also a Google Earth Engine-based tool, LandSat Time series draws from the Landsat archives, and provides annual images, time-lapse animations, and interactive graphs to visualize and track trends in vegetation, water, urbanization, and land cover change over decades.

While Earth Map and LandSat Time Series provided access to remotely sensed data, further reference data was required to more critically analyze the accuracy of satellite observations. To create additional reference data, we used the following sources:

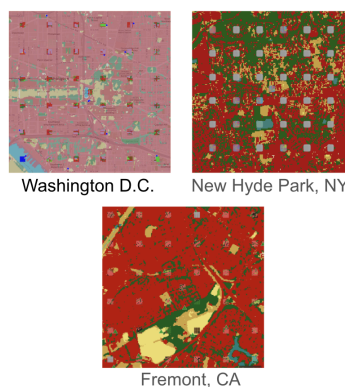
- **Collect Earth Online (CEO)** is a free, open-source web platform that allows users to analyze high-resolution satellite imagery directly in their browser. One powerful feature is its ability to divide each 100 m × 100 m primary sampling unit into a 10 × 10 grid of dots (100 per unit) where users can manually assign land cover classifications to each individual dot. This fine-scale classification process significantly increases the amount of data available to researchers, adding over 3,700 additional land cover observations to each Area of Interest (AOI), making CEO an especially valuable tool for detailed land cover observation, and the generation of reference data.





\*Figure 7: Example of 10 x 10 dot grid generated in CEO, light green indicates that the predominant land cover classification is grass. (Low et al 2021)

**ArcGIS Online** is a geographic information system (GIS) tool used to create, manage, and analyze spatial data. It allows users to work with maps and satellite imagery to perform geospatial analysis, identify patterns, and model spatial relationships. The platform supports a wide range of data formats and tools for tasks like land cover classification, terrain modeling, spatial statistics, and data overlays.



\*Figure 8: ArcGIS maps created of three AOIs layered over WorldCover 10m 2020/2021 (ESA). Figure generated by researchers.

## Comparison

To analyze the level of agreement between reference and remotely sensed data in the context of each AOI, results from each data collection method are placed side by side in a comprehensive table. The table is organized so that moving from top to bottom, each column represents all of the collected data from one specific source or data set, and moving from right to left, each row contains all of the collected data from one specific sampling unit. In this way, we can easily compare accuracies, features, resolutions, and classifications provided by different data sources. This table was included as part of a larger poster on the topic of the same research. The poster also included additional representations of collected data, including our entire AOIs mapped onto each of the satellite data sets, as well as a side by side series of image chips pulled from the LandSat Time Series data. To demonstrate the process behind evaluating the extent to which remotely sensed data and reference data agree with one another, data is pulled from rows, or primary sampling units, where the reference data and remotely

sensed data agreed, partially agreed, and completely disagreed. All AOIs were analyzed using identical sampling frameworks, classification categories, and comparison criteria to ensure consistency across regions.

## **Community and Environmental Context**

### **Washington D.C**

Washington, D.C.'s environmental history is complicated, and defined by its unique geography, industrial past, and continued commitment to urban sustainability.

Originally built on marshland at the confluence of the Potomac and Anacostia Rivers, D.C. was prone to flooding and vector borne disease. The growing population of the city viewed the area's natural environment as barriers to urbanization and growth, which led to large scale drainage and land reclamation projects in the late 19th and early 20th centuries. On this altered land, central neighborhoods, The National Mall, and The Tidal Basin were constructed, all of which are integral to the city's urban plan. While these efforts made development possible, they also drastically altered the natural landscape, reducing vital wetland ecosystems and setting the stage for future environmental challenges.

As the city continued to grow and modernize, D.C. built power plants, reservoirs, and industrial districts, but in the mid 20th century, there was very little attention to detail or consequence when it came to environmentally impactful construction or human activity. This expansion led to the extreme pollution of the Potomac and Anacostia Rivers, to the point where exposure to the water was often unsafe.

Moving later into the 20th century, as national attention to worsening environmental conditions as a result of unsustainable human action increased, residents of D.C. began to speak up about the unhealthy conditions of the city's waterways. The citizen outcry and protest increased, and with the formation of the Environmental Protection Agency (EPA), efforts to clean D.C. moved into full swing. Water cleanup efforts, including the construction of water treatment facilities and regulations towards citizen and industrial waste management, worked to clean the Rivers and turn them into the picturesque healthy local treasures they are today.

By the late 20th century and into the 21st century, D.C. invested heavily towards becoming a national leader in urban sustainability. The city has adopted a wide range of sustainability measures, including ambitious green building standards, significant investments in public transportation and bike infrastructure, and the creation of green roofs, rain gardens, and park space. Today, D.C.'s story represents a broader shift from environmental degradation to preservation and innovation. The city's commitment to sustainability is not just about correcting past mistakes, it is also about setting an example for other urban areas facing similar environmental challenges.

This layered environmental history contributes to a highly heterogeneous land cover mosaic, which directly affects the accuracy of satellite classification and provides critical context for interpreting agreement and disagreement between remotely sensed and ground-based observation.

## **New Hyde Park, NY**

New Hyde Park, formerly called Hyde Park, is one of the earliest settlements in the United States. Dutch settlers arrived in the 1620s, followed by English settlers in the 1640s. Originally a part of the Town of Hempstead, the land was first used as a racecourse and later became farmland. The arrival of the Long Island Railroad in 1837 and later trolley and bus lines transformed the area into a commuter hub, spurring waves of immigration from German, Irish, Polish, Italian, and Jewish communities. By the 1920s, large-scale housing developments replaced much of the farmland, leading to rapid suburbanization and the need for modern infrastructure and local governance.

The transition from agricultural land to mid-20th-century suburban development produced a fragmented canopy structure, characterized by mature street trees, private backyards, and local parks, a land cover pattern that is difficult for moderate-resolution satellite products to resolve.

## **Fremont, CA**

Fremont, California, is a large suburban city in the heart of the San Francisco Bay Area and roughly thirty miles away from San Francisco itself. The natural landscape of the area is marshlands and plains, and it remained this way until the mid-1700s, which was when the Spanish settled in the area for the construction of their religious “missions.” However, apart from their central churches and town halls, the Spanish didn’t alter the landscape drastically, and the Fremont area stayed in roughly its natural state until the mid-1900s. It was in the mid-1900s that Fremont, along with the rest of the Bay Area, saw its population boom as Silicon Valley grew with the rise of electronics and the computer age. In these last 50-60 years, the land has gone from green marsh and farmlands to mainly suburban housing and commercial centers.

This transition has created a primarily developed landscape with some remaining greenery, also creating a dense, heterogeneous environment in which current land cover observation methods may fail to capture sufficient granularity.

## **Weslaco, TX**

Weslaco, Texas is located in the heart of the Rio Grande Valley, a region whose environmental history has recently been shaped by water scarcity, agriculture, and rapid population growth. Originally, the land surrounding present-day Weslaco consisted of semi-arid plains, native grasslands, and *resaca* systems (former river channels formed by shifts in the Rio Grande). These natural waterways played a critical role in sustaining local ecosystems and indigenous communities long before large-scale settlement.

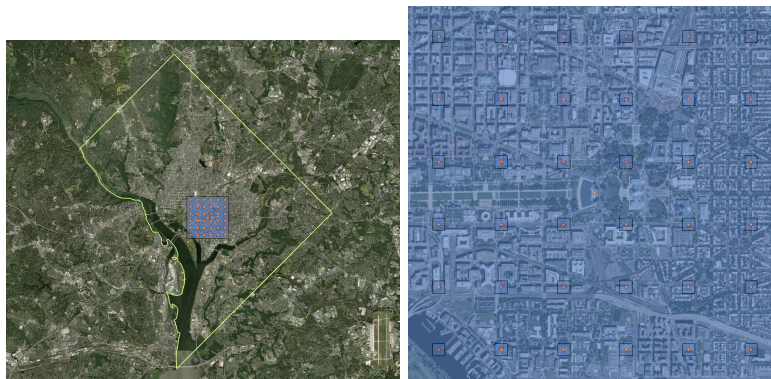
In the late 20th and 21st centuries, awareness of these environmental challenges prompted renewed efforts toward conservation and sustainable management. Weslaco became home to important environmental research and educational centers, including Estero Llano Grande State Park. Local and regional initiatives now focus on water efficiency, habitat restoration, and environmental education, particularly as the Rio Grande Valley faces increasing stress from climate change and population growth.

Today, the city's future sustainability depends on responsible water management, conservation of remaining natural habitats, and continued public engagement in protecting the region's fragile environment.

## Results

### Washington D.C

The following section describes the findings by Samuel Bawden in his D.C. AOI. The Washington D.C. AOI is centered on the Ulysses S. Grant Memorial in front of the U.S. Capitol Building, and encompasses the majority of central Washington D.C.



\* Figures 9-10: The left image shows AOI in blue in context of the Districts limits. The right image zooms in on the AOI, overlaying the primary sampling units onto the geography of the city.

### Washington D.C. Data Comparisons

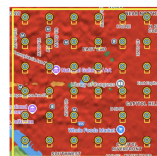
Platform	LandSat 5-9	WorldView w-4	Sentinel-1/2			GLOBE Observer						Collect Earth Online
Primary Sample Unit	LandSat Time Series Graph	2m Tree Canopy MHA	World Cover 10m	Dynamic World 10m	ESRI 10m	Up	Down	west	south	east	north	high resolution in image interpretation
0												
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												
26												
27												
28												
29												
30												
31												
32												
33												
34												
35												
36												



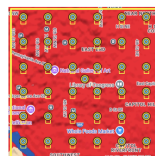
Meta/WRI Global  
Canopy Height



WorldCover 10m  
2020/2021 (ESA)

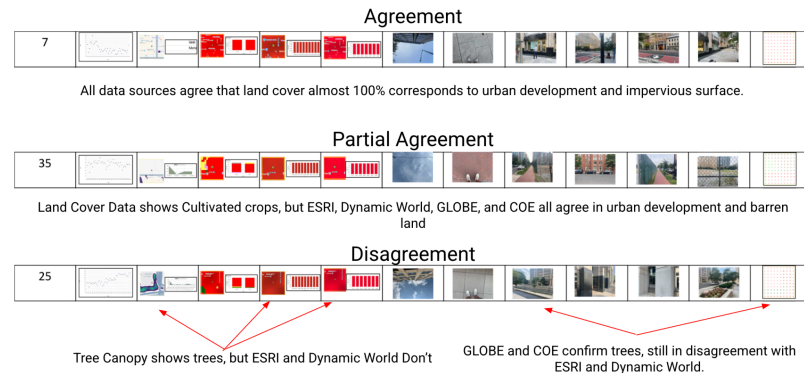


Dynamic World



ESRI 2017/2024  
Land Cover

Figures 11-13: Washington D.C AOI Table, Satellite Imagery Visual Representation (Meta/WRI, WorldCover, Dynamic World, ESRI), Agreement/Partial Agreement/Disagreement Row Comparison



Through these representations, some clear conclusions can be drawn. GLOBE Observations and CEO data do not often fully align with remotely sensed data. Upon closer inspection of the satellite data sources, it is shown that the WorldCover 10m 2020/2021 (ESA) data is the most consistent with the reference data. Overall, these types of comparisons are extremely helpful in validating the efficacy of ground-truthing methods, and makes a strong case for the usefulness of the additional data that citizen science methods offer researchers.

## New Hyde Park, NY

This following section describes the findings by Nandini Khaneja in her New Hyde Park AOI. The New Hyde Park AOI is centered in a suburban area within New Hyde Park itself but it encompasses portions of the neighboring areas North New Hyde Park, Manhasset Hills, and Garden City Park.

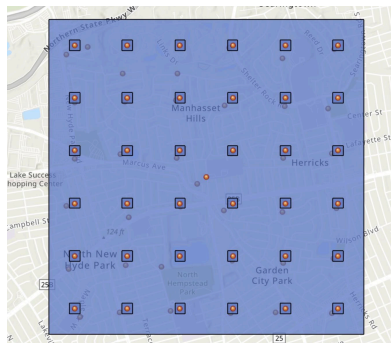
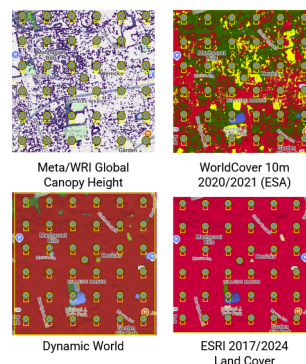


Figure 14: Magnified 3km AOI centered in New Hyde Park

## New Hyde Park Data Comparison

Index	Landmarks S-R	WorldView-4	Sentinel-1/2	GLOBE Observer	Call out Earth +
Index	Image Date	Image Date	Image Date	Image Date	Image Date
0					
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					

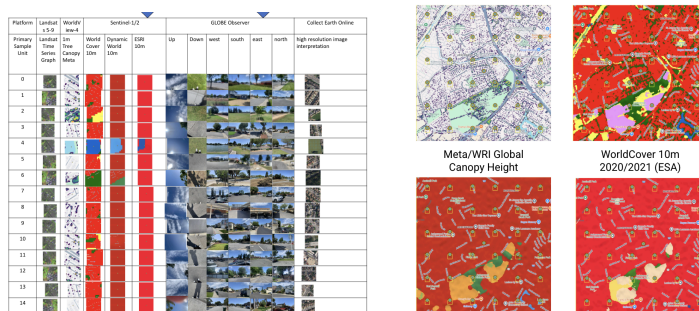


Figures 15-17: New Hyde Park, NY AOI Table, Satellite Imagery Visual Representation (Meta/WRI,



## Fremont, CA

## Fremont Data Comparison

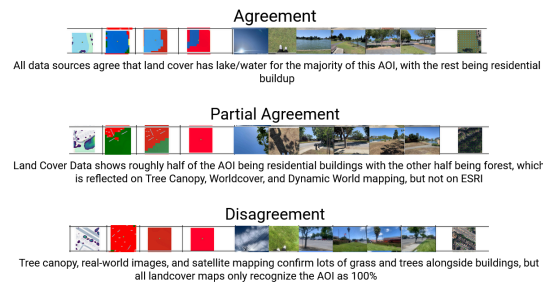


Figures 17-19: Fremont, CA AOI Table, Satellite Imagery Visual Representation

## (Meta/WRI, WorldCover, Dynamic World, ESRI), Agreement/Partial Agreement/Disagreement Row Comparison

Upon analyzing the land cover data by satellites over the area of interest, it that most of the Fremont area is buildings and construction (indicated the satellite maps). Despite the of human-made land cover, a large

land just below the center of the area of interest is reported as cropland and grass, and this is due to the Ardenwood Historic Farm—a municipality and state-protected farm attempting to preserve the ambiance and lifestyle of the pre-Silicon Valley Bay Area. The farm is covered with mostly wheat crops, grass, trees, and rolling open fields.



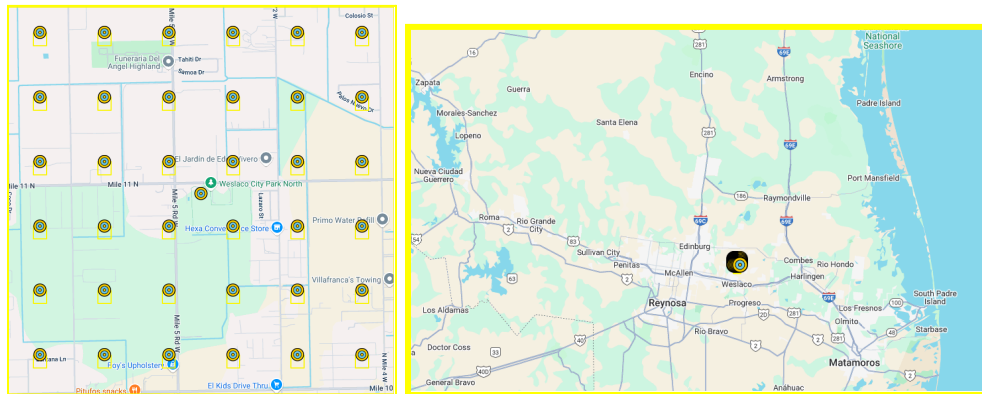
captured was clear covered in by red on prevalence chunk of

However, when comparing satellite-reported land cover with the field data collected by the researchers, there were many inaccuracies present. The low resolution of satellite imagery often renders over-generalizations of the land, such as overlooking patches of grass, trees, and soil, and passing them off as construction. From the 37 primary sampling units within the area of interest, there was a spectrum of inaccuracies, from completely inaccurate to almost completely accurate, as shown below.

Analyzing these agreement and disagreement patterns is important because it reveals not only what data is right or wrong, but also the types and locations of data that are commonly misrepresented. This analysis can provide insight as to what parts of satellite imagery can be changed to increase accuracy or potentially taught to AI models to automatically analyze such land cover areas with improved accuracy and resolve errors.

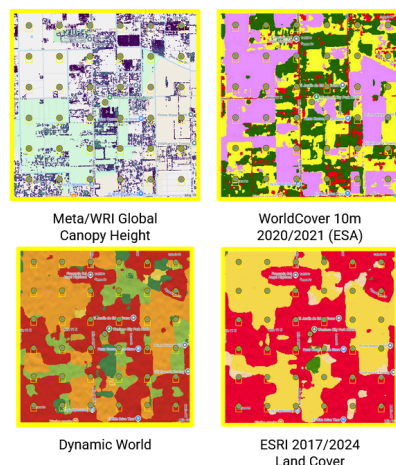
## Weslaco, TX

The following section describes the findings of Jordan Rodriguez for the Weslaco AOI. The Weslaco AOI is centered within the northernmost part of the city of Weslaco in the Rio Grande Valley. The AOI is situated in predominantly suburban and agricultural regions and encompasses residential neighborhoods, commercial corridors, and adjacent farmland, providing a representative view of land use patterns and environmental conditions in South Texas.



Figures 20-21: The left image shows the magnified 3km AOI centered in Weslaco. The right image shows the AOI in context of the Rio Grande Valley.

## Weslaco Data Comparisons



Figures 22-24: Weslaco, TX  
AOI Table, Satellite Imagery  
Visual Representation  
(Meta/WRI, WorldCover,  
Dynamic World, ESRI),  
Agreement/Partial  
Agreement/Disagreement  
Row Comparison



## Agreement



The field photos show bare soil, farmland, and natural vegetation, which all four datasets consistently recognized.

## Partial Agreement



The land appears semi-developed, leading some datasets to label it as grassland, while others marked it as cropland or built up. Differences in resolution and/or image date likely caused the slight mismatches.

## Disagreement



The photos show sidewalks, homes, and paved areas, but some datasets still classify them as vegetation or cropland. This is likely due to outdated satellite data or sensors missing narrow urban features.

After analyzing satellite-derived land-cover data across the Weslaco AOI, it became evident that much of the region is characterized by a mix of urban development and agricultural land use. Satellite classifications, primarily from Sentinel-2 and WorldView imagery, identified extensive residential and commercial areas, along with large tracts of cropland and grassland surrounding and interwoven throughout the city. This pattern reflects Weslaco's position as a growing suburban center embedded within an agriculturally intensive region of the Rio Grande Valley, where farmland and urban infrastructure coexist closely,

A technical failure within the GLOBE Observer platform prevented successful photo uploads for most AOIs, limiting usable ground data to only 4 locations. As a result, discrepancies between satellite-reported land cover and actual surface conditions could not be fully assessed across the study area. While satellite imagery offers a comprehensive, large-scale view of Weslaco's landscape, its resolution can lead to generalized classifications that overlook smaller features such as isolated vegetation and bare soil patches. Without sufficient ground truthing, these classifications remain largely unverified, reducing the accuracy and reliability of environmental assessments.

Field photos showed strong agreement for bare soil, farmland, and natural vegetation, which were consistently identified across the provided datasets. Partial agreement and disagreement occurred in semi-developed and urban areas, where differences in image resolution, capture date, or outdated satellite data led to misclassification of built surfaces as vegetation or cropland.

## General Trends

Analyzing some overarching trends across all of this data, it was found that almost all areas of interest saw increases in building and urban development over the last few decades. This was paired, in all areas of interests, with a steady decrease in tree and grass cover, which could be seen as concerning from an environmental preservation perspective. It can be assumed that the urban and suburban nature of the studied areas of interests and their proximity to some of the largest cities in the nation caused the prevalence of buildings and construction in the studied land cover.

In comparing the different data sources, it was noticed that the European Space Agency's Worldcover mapping was the most detailed and accurate, especially compared to Dynamic World and ESRI, which often over-generalized entire areas and failed to see the diversity of land cover that was truly present. For instance, there were many instances in which Dynamic World and ESRI would miss patches of trees and backyards, flagging them as construction, whereas Worldcover would pick them up and mark them as their respective natural land cover.

Lastly, there were differences of large magnitudes between the high resolution analysis of images and lower resolution classifications. This shows that the lower resolution satellite imagery is over generalized and cannot capture minute details, rendering those sources less suitable for scientific use for land cover monitoring and potential training of AI models.

## **Future Research**

Insights from this work revealed both the value of ground-based citizen observations and the limitations of existing platforms in supporting high-resolution, scalable land cover analysis. While the Adopt-a-Pixel methodology enabled structured comparisons between field observations and satellite products, practical constraints within GLOBE, most notably reduced image fidelity, fragmented visualization tools, and limited support for downstream analysis, restricted the scientific usability of collected data. These gaps motivated the development of EarthLens as a next-generation citizen science observation system designed to preserve data quality while expanding analytical capability.

EarthLens is intended to supplement existing programs like GLOBE by addressing participant-observed limitations, a drone and app-based platform that integrates high-resolution aerial imagery, real-time AI analysis, citizen photo uploads, satellite overlays, and collaborative visualization into a single workflow. The system combines onboard imaging and spectrometer hardware with native software to enable geotagged observations, automated land cover classification, and direct comparison between drone, satellite, and GLOBE datasets. By addressing the technical barriers identified through this research, EarthLens provides a scalable framework that democratizes environmental observation while generating more actionable data for assessing land cover change, disaster impacts, and broader environmental dynamics.

## **Conclusion**

This study demonstrates that structured citizen science observations, when combined with the Adopt-a-Pixel methodology, provide critical insight into the strengths and limitations of widely used satellite land cover products. Across four geographically and socially distinct regions, comparisons indicate discrepancies between the ground and satellite observations, with the most accurate tools being the Meta/WRI tree canopy and WorldCover 10m map. This work highlights the role of community context and participant experience in shaping environmental data collection. Local land-use history, access constraints, and data acquisition tool limitations directly influenced how and where observations were made. Limitations identified in the GLOBE Observer platform, particularly reduced image fidelity and data loss, emphasize the need for tools that address these user issues in order to generate usable, citizen science based land-cover data for researchers. Together, these results affirm that integrating citizen involvement, community knowledge and remote sensing allows for a multifaceted and more accurate assessment of a region's environmental conditions.

## **Acknowledgements**

We would like to sincerely thank our incredible mentors — Dr. Russanne Low, Mr. Peder Nelson, Dr. Cassie Soeffing, Mr. Andrew Clark, and Dr. Brianna Lind — for their guidance, support, and encouragement throughout this experience. We are also deeply grateful to the SEES program, NASA, UT Austin, and the Texas Space Grant Consortium for making this opportunity possible.

Finally, thank you to our fellow interns — your collaboration, creativity, and energy make this program the unique and enriching experience it is.

## **References**

Low RD, Nelson PV, Soeffing C, Clark A and SEES 2020 Mosquito Mappers Research Team (2021) Adopt a Pixel 3 km: A Multiscale Data Set Linking Remotely Sensed Land Cover Imagery With Field Based Citizen Science Observation. *Front. Clim.* 3:658063. doi: 10.3389/fclim.2021.658063