Best Practice

Evaluating Satellite Land Cover Classification Accuracy Using Participatory Remote Sensing

Aren Finelt[1,*], Russanne Low[2], and Peder Nelson[2]

[1] 12th Grade, Saint John's School, San Juan, Puerto Rico, USA; https://orcid.org/0009-0004-7841-5086
[2] Institute for Global Environmental Strategies; low.russanne@gmail.com; https://orcid.org/0000-0002-7912-4350
[3] Oregon State University; peder.nelson@oregonstate.edu; https://orcid.org/0000-0003-3979-9051
[*] Corresponding Author: finelt.aren@gmail.com

JOURNAL OF EARTH OBSERVATION & GEOSPATIAL APPLICATIONS

Academic Editor: FirstName LastName
Received: 1 March 2026
Revised: 25 April 2026
Accepted: 27 April 2026
Published: 30 April 2026

Abstract: Satellite remote sensing has become a foundational tool for monitoring land cover change and urban development. However, classification accuracy remains limited in complex urban environments where vegetation, built infrastructure, and water coexist within small spatial scales. This study evaluated the accuracy of multiple satellite land cover datasets within a 3 km × 3 km area of interest in Hato Rey, Puerto Rico, using the National Aeronautics and Space Administration (NASA) Adopt-a-Pixel participatory remote sensing framework. Ground observations were collected at 23 sampling locations using the Global Learning and Observations to Benefit the Environment (GLOBE) Observer application and compared with land cover classifications derived from European Space Agency (ESA) WorldCover, Dynamic World, ESRI Land Cover, Meta Tree Canopy, and Landsat imagery. Results showed that several satellite datasets frequently overgeneralized land cover as entirely urban, failing to detect localized vegetation and water features observed on the ground. Dynamic World and ESRI Land Cover achieved higher overall accuracy (91.3%) due to consistent classification of urban areas, while Meta Tree Canopy and ESA WorldCover demonstrated improved detection of vegetation in heterogeneous environments, despite slightly lower overall accuracy (87.0%). Landsat time series analysis revealed mixed vegetation trends over time, reflecting both urban expansion and persistent vegetation. These findings demonstrate the importance of integrating participatory ground observations with satellite data to improve land cover classification accuracy and enhance environmental monitoring in rapidly developing urban environments.

Keywords: remote sensing, land cover, urban environment, Puerto Rico, participatory science

1. Introduction

Remote sensing provides one of the most powerful tools available for understanding Earth's surface. Satellites continuously collect data that allow scientists to monitor environmental change across spatial scales ranging from individual cities to the entire planet. These observations have become essential for studying urban development, vegetation change, and environmental processes that affect both ecosystems and human populations (Gorelick et al., 2017). Land cover classification derived from satellite imagery plays a critical role in this effort. These classifications allow researchers to quantify the extent of built infrastructure, vegetation, water, and other land cover types. Such information is used to evaluate changes in urban development and the environmental effect of said developments. However, despite advances in satellite technology, classification accuracy remains limited, particularly in heterogeneous urban environments where multiple land cover types exist within small areas (Brown et al., 2022). Satellite classification relies on spectral reflectance, which measures how surfaces reflect electromagnetic radiation; while this method is highly effective for identifying large, uniform land cover types such as forests or open water, it becomes less reliable when applied to mixed landscapes. In urban environments, individual satellite pixels often contain a combination of buildings, vegetation, and pavement, creating classification uncertainty and can result in

Citation: Finelt, A., Low, R., and Nelson, P. (2026). Evaluating Satellite Land Cover Classification Accuracy Using Participatory Remote Sensing. Journal of Earth Observation and Geospatial Applications, 2(1), 69–77. https://doi.org/10.65372/3w1fbk03

misrepresentation of actual land cover (Karra et al., 2021). Recent global land cover products such as European Space Agency (ESA) WorldCover, Dynamic World, and ESRI Land Cover have improved spatial resolution and classification methods, achieving overall accuracies of approximately 70–85% depending on land cover type and region (ESA, 2021; Brown et al., 2022). However, classification accuracy remains highly variable in dense urban environments, where shadows, vegetation canopy, and complex infrastructure interfere with classification algorithms. Participatory remote sensing offers an effective method for addressing these limitations. The National Aeronautics and Space Administration (NASA) Adopt-a-Pixel framework combines satellite observations with ground-based citizen science data to evaluate land cover classification accuracy (Low et al., 2021). Using the Global Learning and Observations to Benefit the Environment (GLOBE) Observer mobile application (GLOBE Observer, 2023), participants collect ground photographs that provide direct evidence of land cover conditions. These observations allow researchers to compare satellite classification with real-world conditions and identify classification errors. This approach provides critical validation data that cannot be obtained through satellite observation alone. Ground observations capture fine-scale land cover variability that is often invisible to satellite sensors. By integrating satellite and ground data, participatory remote sensing improves classification accuracy and enhances understanding of land cover patterns. Hato Rey, Puerto Rico, provides an ideal study area for evaluating satellite classification accuracy. Located in San Juan, Hato Rey is the island's primary financial district and has experienced significant urban development. Despite extensive construction, vegetation remains present throughout the area, creating a heterogeneous land cover environment that challenges satellite classification. The objective of this study was to evaluate the accuracy of satellite land cover classification in Hato Rey using participatory remote sensing methods. It was hypothesized that satellite datasets would frequently overgeneralize land cover classification, particularly by misclassifying heterogeneous urban environments as entirely built-up.

2. Study Area and Methods

2.1. Study Area

The study area (Figure 1) encompasses the district of Hato Rey within the municipality of San Juan, Puerto Rico. Hato Rey is located in the north-central portion of the San Juan metropolitan region and serves as the island's primary financial and commercial center. Often referred to as the "Wall Street of the Caribbean," the district contains a dense concentration of high-rise office buildings, residential complexes, transportation infrastructure, and commercial areas. In addition to heavily developed urban land, the landscape also includes smaller parks, roadside vegetation, and water features associated with the Río Piedras river and engineered drainage systems. This combination of built infrastructure and fragmented vegetation creates a heterogeneous land cover environment, making it a suitable location for evaluating the accuracy of satellite-based land cover classification in a dense urban setting.

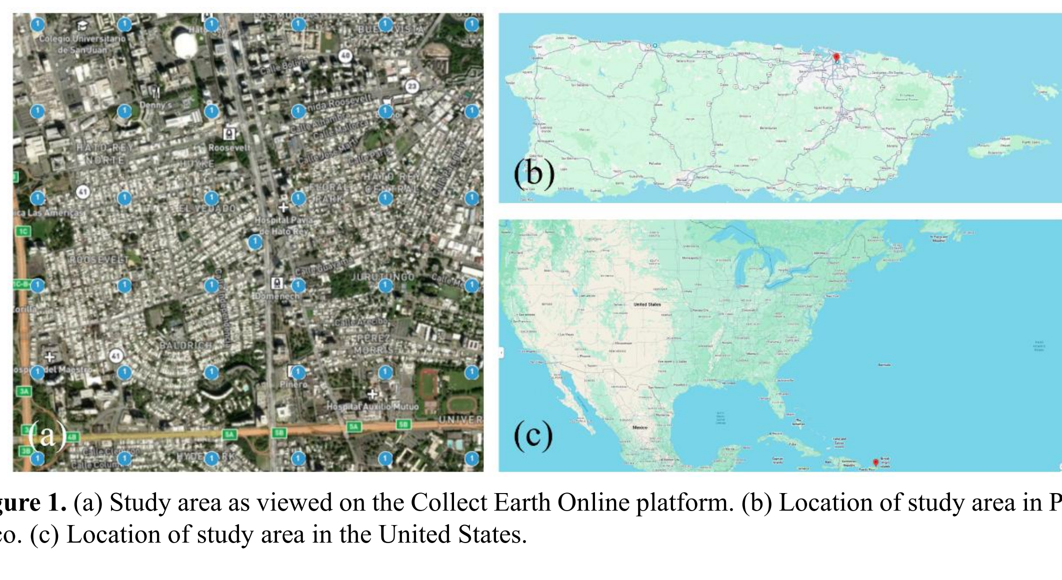


Figure 1. (a) Study area as viewed on the Collect Earth Online platform. (b) Location of study area in Puerto Rico. (c) Location of study area in the United States.

2.2. Data

The primary data for this study consisted of ground-based observations and satellite-derived land cover datasets. Ground observations were collected using the GLOBE Observer Land Cover mobile application at 23 accessible points within a 3 km × 3 km Adopt-a-Pixel area of interest containing 37 total sampling locations arranged in a grid pattern. The remaining sampling points were inaccessible due primarily to their location within private property, reflecting the highly developed nature of the study area. Observations were conducted over several weeks, and each accessible sampling location was documented using the standard GLOBE Observer protocol, which includes photographs taken in six directions: north, south, east, west, up, and down. These ground observations provided a reference dataset for evaluating satellite land cover classification accuracy. Figure 2 shows the location of the study area within Puerto Rico and the Adopt-a-Pixel sampling grid overlay.

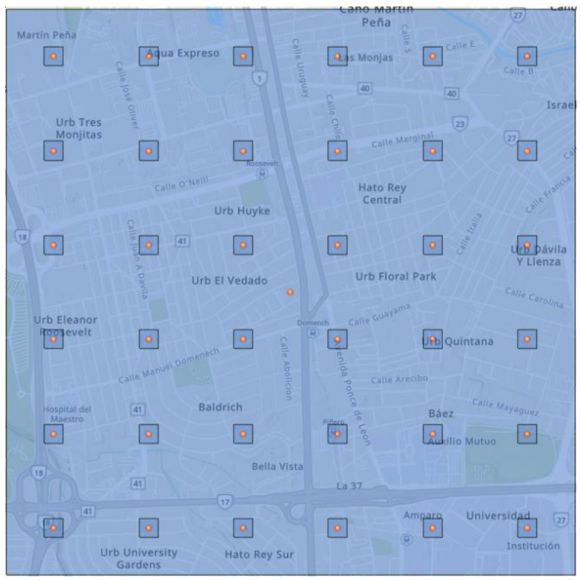


Figure 2. Location of the study area in Hato Rey, Puerto Rico, with the Adopt-a-Pixel 3 km × 3 km sampling grid overlay, Courtesy of ArcGIS.

Ground observations were collected at 23 accessible sampling locations. Satellite datasets provided spatially continuous land cover information for the study area and included the Meta 1 m Tree Canopy dataset (Tolan et al., 2023), ESA WorldCover (10 m resolution), Dynamic World (10 m resolution), and ESRI 10 m Land Cover datasets. These datasets were accessed and analyzed using the EarthMap platform, which allows visualization and comparison of multiple satellite-derived land cover products. Additional data were obtained from Landsat satellite imagery (USGS, 2020) spanning from 1984 to 2025 to evaluate long-term vegetation and land cover trends. Collect Earth Online (Saah et al., 2019) was also used to manually classify land cover based on high-resolution aerial imagery. All data, including satellite classification, ground observations, and manual classification results, were exported in CSV and GeoJSON formats to allow integration across platforms and facilitate comparison between datasets.

2.3. Methods

Land cover classification within the area of interest was evaluated through comparison of satellite-derived datasets, manual image interpretation, and ground observations. Satellite land cover products, including ESA WorldCover, Dynamic World, ESRI Land Cover, and Meta Tree Canopy, were examined using the EarthMap platform. Each dataset was visually analyzed across the area of interest and at individual sampling locations to identify differences in classification, particularly between built and vegetated land cover. Manual classification was conducted using Collect Earth Online to provide an additional reference for comparison. High-resolution imagery was used to visually assess land cover within each primary sample unit, allowing satellite classifications to be evaluated against observed land cover patterns.

Ground observations collected through the GLOBE Observer application were used to validate satellite classification at selected sampling points. Each sampling location was assigned a dominant land cover class (Urban or Vegetation) based on ground photographs taken in six directions. Satellite-derived classifications were then compared directly with these ground-based labels to determine agreement.

A confusion matrix was constructed for each dataset to quantify classification performance. Overall accuracy, producer’s accuracy, and user’s accuracy were calculated following standard remote sensing accuracy assessment methods. Agreement was defined as a match between the dominant ground-observed land cover and the satellite classification for the corresponding location. In cases where both built and vegetated features were present, classification was based on the dominant observed land cover.

Inaccessible sampling points were excluded from quantitative analysis, resulting in a total of 23 evaluated locations. Differences in spatial resolution between datasets (e.g., 1 m Meta Tree Canopy vs. 10 m land cover products) were considered when interpreting classification discrepancies, particularly in areas with fragmented vegetation.

Landsat imagery was also examined to assess recent land cover stability within the study area. Satellite images from multiple years were visually compared to identify any major changes in land cover. This combined qualitative and quantitative approach allowed differences between satellite datasets and actual land cover conditions to be systematically identified and evaluated.

3. Results

3.1. Comparison Between Dynamic World and ESRI Land Cover

Land cover classification derived from the Dynamic World and ESRI datasets showed strong agreement at Point 31, where both datasets classified the area as built environment. As shown in Figure 3, both classification products identified the entire primary sample unit as developed land, with no vegetation detected. This result is consistent with the highly urbanized nature of Hato Rey, where large areas consist of buildings, roads, and impervious surfaces.

The agreement between these two independent datasets suggests that both classification methods are effective at identifying heavily developed urban areas. However, the uniform classification also indicates a lack of sensitivity to more detailed land cover features such as individual trees or fragmented vegetation, which may be present but not detected at the spatial resolution of the datasets. Table 1 shows that both datasets classified all 23 points as Urban, yielding a producer's accuracy of 0% for vegetation. Overall accuracy = 91.3%, inflated by the dominance of urban cover in the study area.

Figure 3. Comparison of Dynamic World (a) and ESRI (b) land cover classification at Point 31. Both datasets classify the area as built environment, demonstrating agreement in highly urbanized areas.

Table 1. Confusion matrix for Dynamic World and ESRI vs. GLOBE Observer (n = 23).

GLOBE \ Satellite	Urban	Vegetation	Row Total
Urban	21	0	21
Vegetation	2	0	2
Column Total	23	0	23

3.2. Comparison Between ESA WorldCover and Meta Tree Canopy

Comparison between the ESA WorldCover and Meta Tree Canopy datasets revealed both agreement and disagreement depending on location. At Point 20, both datasets indicated the presence of vegetation. As shown in Figure 4, the Meta Tree Canopy dataset detected tree canopy in areas where ESA WorldCover also identified vegetation. This agreement suggests that both datasets accurately captured vegetation at this location.

In contrast, disagreement was observed at Point 10. While the Meta Tree Canopy dataset detected tree canopy, ESA WorldCover classified the same area as built environment. This discrepancy suggests that ESA WorldCover may underrepresent vegetation in dense urban environments where tree canopy exists alongside built infrastructure. The higher spatial resolution of the Meta Tree Canopy dataset allows it to detect smaller vegetation features that may not be captured in lower-resolution land cover products.

(a)

(b)

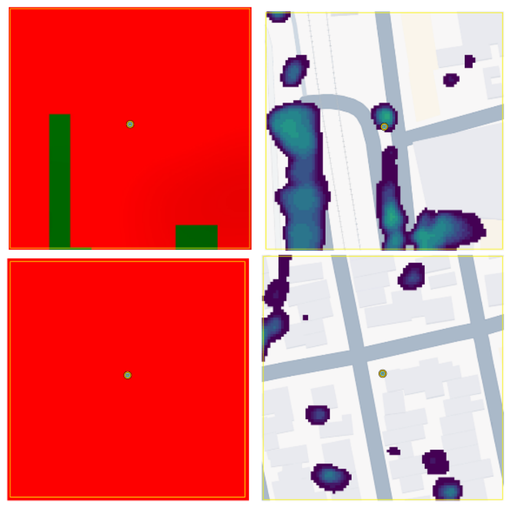


Figure 4. Comparison between ESA WorldCover (Left) and Meta Tree Canopy datasets (Right). (a) Agreement at Point 20 shows both datasets detecting vegetation, while (b) disagreement at Point 10 demonstrates ESA WorldCover failing to identify tree canopy detected by the Meta dataset.

Table 2. Confusion matrix for WorldCover and Tree Canopy Meta vs. GLOBE Observer (n = 23). Both datasets produced identical classification results. Overall accuracy = 87.0%. The 3 Urban points classified as Vegetation likely reflect sub-pixel green cover present in ground imagery.

GLOBE \ WorldCover	Urban	Vegetation	Row Total
Urban	18	3	21
Vegetation	0	2	2
Column Total	18	5	23

3.3. Comparison Between Satellite Classification and Ground Observations

Comparison between satellite classification and ground observations revealed varying levels of agreement depending on location. At Point 5, ground photographs showed a fully developed urban roadway with no visible vegetation, and both ESA WorldCover and ESRI datasets correctly classified the area as built environment (Figure 5). This result demonstrates that satellite classification can accurately represent land cover in fully urbanized areas.

At Point 11, partial agreement was observed. Ground photographs showed a developed area with nearby vegetation, and while both satellite datasets classified the majority of the area as built environment, ESA WorldCover identified a small portion of vegetation. This suggests that satellite datasets can partially detect vegetation in mixed environments but may underestimate its extent.

In contrast, disagreement was observed at the third ground observation point, where ground photographs showed visible vegetation, but satellite datasets classified the area as entirely built environment. This misclassification demonstrates a limitation of satellite land cover datasets in detecting small or fragmented vegetation in dense urban areas. These findings highlight the importance of ground observations in validating satellite classification accuracy.

Table 3 shows the accuracy by data source. The overall accuracy is misleading here. Dynamic World and ESRI score higher (91.3%) purely by defaulting every point to Urban. Tree Canopy Meta and WorldCover score lower (87.0%) but are the only datasets that detected any vegetation at all.

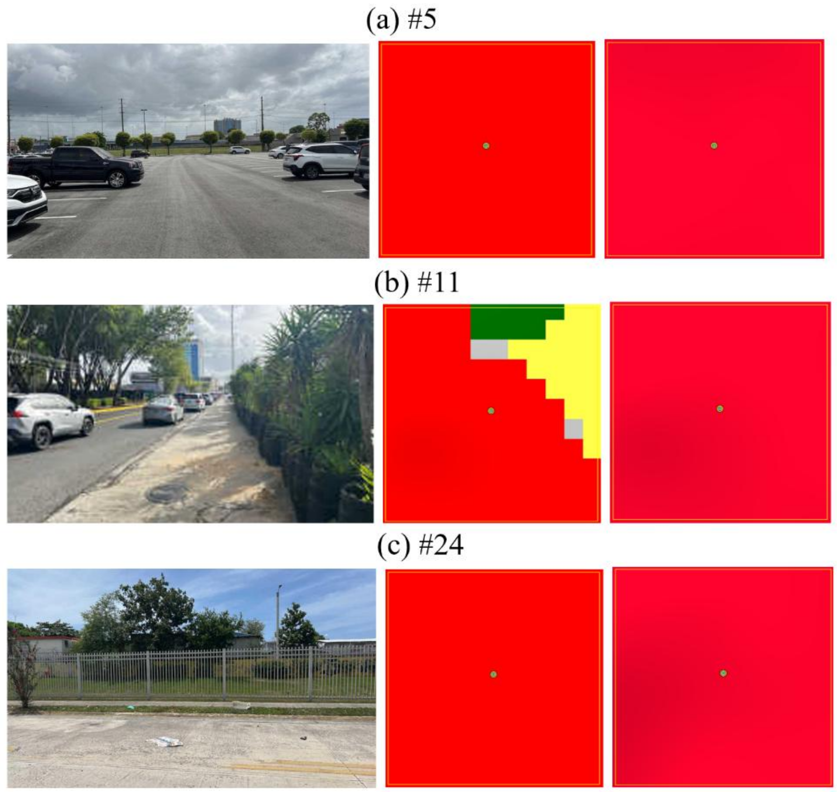


Figure 5. Comparison between ground observations and satellite classification at #5 (a), #11 (b), and #24 (c) observation points. Satellite datasets show varying levels of agreement with actual land cover conditions.

Table 3. Comparison of four satellite land cover datasets against GLOBE Observer ground truth (n = 23). Datasets ranked by ability to detect sub-pixel vegetation in a predominantly urban landscape.

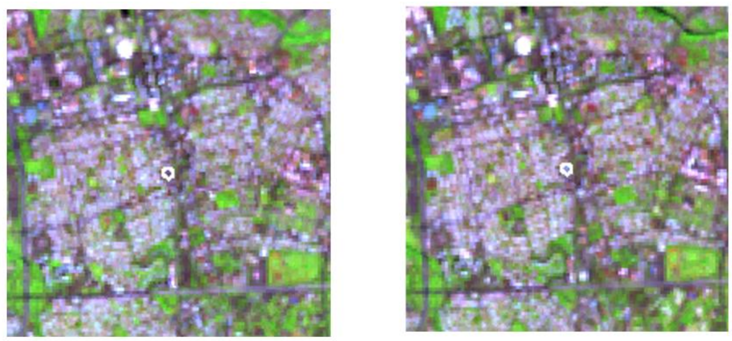
Dataset	Overall Accuracy	Agreement (n=23)	Urban Detection	Vegetation Detection	Green Pocket Sensitivity	Rank
Tree Canopy Meta	87.0%	20 / 23	Strong	Strong	Best	1st
WorldCover (ESA)	87.0%	20 / 23	Strong	Moderate	2nd best	2nd
Dynamic World	91.3%	21 / 23	Strong	None	None	3rd
ESRI	91.3%	21 / 23	Strong	None	None	3rd

3.4 Landsat Time-Series Comparison

Comparison of Landsat imagery between 2020 and 2025 revealed minimal visible change in land cover within the study area. As shown in Figure 6, the overall spatial distribution of built infrastructure and vegetation remained largely consistent between the five years. This stability suggests that the study area has already undergone significant urban development and that land cover patterns have remained relatively stable in recent years. Minor variations in vegetation may be present, but no major changes in land cover were observed.

4. Discussion

The results show that satellite land cover classification accuracy in Hato Rey varied depending on land cover type and dataset resolution. Dynamic World and ESRI Land Cover showed strong agreement in fully developed areas, correctly classifying locations dominated by buildings and impervious surfaces. This indicates that satellite classification performs reliably where land cover is uniform and clearly defined. However, both ESA WorldCover and Meta Tree Canopy revealed differences in vegetation detection. The Meta dataset identified tree canopy in locations classified as built environment by WorldCover, indicating that lower-resolution land cover products may underestimate vegetation in dense urban environments. This is likely due to mixed pixels, where vegetation and built surfaces occur within the same satellite pixel. Comparison with ground observations confirmed this limitation. Satellite datasets accurately classified fully developed areas but failed to detect vegetation at some mixed land cover locations. This demonstrates that classification accuracy decreases where vegetation is fragmented or located near infrastructure. Landsat imagery showed minimal recent land cover change, suggesting that classification differences between datasets were primarily due to methodological differences rather than actual environmental change. Overall, the Meta Tree Canopy dataset provided the most accurate representation of vegetation, while Dynamic World and ESRI Land Cover were most consistent in identifying built areas. These results highlight the importance of combining satellite datasets with ground observations to improve interpretation of urban land cover. In Table 4, the producer's accuracy of 0% for Vegetation in Dynamic World and ESRI indicates complete failure to detect green cover despite its presence in ground-level imagery. Tree Canopy Meta and WorldCover both achieved 100% producer's accuracy for Vegetation, though user's accuracy (40.0%) reflects over-prediction of vegetated cover at urban-dominated points.



(a) 2020 (b) 2025

Figure 6. Landsat imagery comparison between 2020 and 2025 showing stable land cover patterns within the study area.

Table 4. Producer's and user's accuracy by class and dataset.

Dataset	Class	Correct	Actual	Predicted	Producer's Accuracy	User's Accuracy
Tree Canopy Meta	Urban	18	21	18	85.7%	100.0%
	Vegetation	2	2	5	100.0%	40.0%
WorldCover (ESA)	Urban	18	21	18	85.7%	100.0%
	Vegetation	2	2	5	100.0%	40.0%
Dynamic World	Urban	21	21	23	100.0%	91.3%
	Vegetation	0	2	0	0.0%	N/A
ESRI	Urban	21	21	23	100.0%	91.3%
	Vegetation	0	2	0	0.0%	N/A

5. Conclusions

This study evaluated the accuracy of satellite-derived land cover classification within Hato Rey, Puerto Rico, through comparison of ESA WorldCover, Dynamic World, ESRI Land Cover, and Meta Tree Canopy datasets with ground observations collected using the GLOBE Observer application. The results showed that Dynamic World and ESRI Land Cover consistently and accurately classified fully developed urban areas but were less effective at identifying vegetation in mixed land cover environments. The Meta Tree Canopy dataset detected vegetation that was not represented in the ESA WorldCover dataset, demonstrating the importance of higher spatial resolution for identifying tree canopy in dense urban settings.

Comparison with ground observations confirmed that vegetation was present in some locations classified as built environment by satellite land cover datasets, indicating limitations in the ability of moderate-resolution classification products to detect fragmented urban vegetation. Analysis of Landsat imagery showed minimal land cover change in the study area in recent years, suggesting that observed differences between datasets were primarily the result of classification limitations rather than actual environmental change.

These findings demonstrate that while satellite land cover datasets such as Dynamic World, ESRI Land Cover, and ESA WorldCover provide reliable identification of large-scale urban development, higher-resolution datasets such as Meta Tree Canopy and ground-based observations are necessary to accurately characterize vegetation in complex urban environments. Integrating satellite datasets with participatory ground observations improves the accuracy and interpretation of urban land cover classification.

Funding: This research was funded by NASA CSR SEES.

Data Availability Statement: Data are available at https://zenodo.org/records/19747877

Acknowledgment: The authors would like to acknowledge the support of the 2025 Earth System Explorers (ESE) Team, NASA. Science Mentors, and ESE peer mentors. NASA STEM Enhancement in the Earth Sciences (SEES) Virtual High School Internship program. The NASA Earth Science Education Collaborative leads Earth Explorers through an award to the Institute for Global Environmental Strategies, Arlington, VA (NASA Award NNX6AE28A). The SEES High School Summer Intern Program is led by the Texas Space Grant Consortium at the University of Texas at Austin (NASA Award NNX16AB89A0).

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., Czerwinski, W., Pasquarella, V. J., Haertel, R., Ilyushchenko, S., Schwehr, K., & Weisse, M. (2022). Dynamic World, near real-time global 10 m land use land cover mapping. *Scientific Data*, *9*(1), 251. https://doi.org/10.1038/s41597-022-01307-4

European Space Agency (ESA). (2021). *ESA WorldCover 10 m 2020 v100 product user manual.* https://worldcover2020.esa.int/

GLOBE Observer. (2023). *GLOBE Observer land cover protocol.* https://observer.globe.gov/

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J. C., Mathis, M., & Brumby, S. P. (2021). Global land use / land cover mapping with Sentinel-2 and deep learning. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 4704–4707. https://doi.org/10.1109/IGARSS47720.2021.9553499

Low, R. D., Nelson, P. V., Soeffing, C., Clark, A., & 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. *Frontiers in Climate*, *3*, Article 658063. https://doi.org/10.3389/fclim.2021.658063

National Oceanic and Atmospheric Administration (NOAA). (2023). *Climate data for Puerto Rico*. https://www.noaa.gov/

Saah, D., Johnson, G., Ashmall, W., Tondapu, G., Tenneson, K., Patterson, M., & Poortinga, A. (2019). Collect Earth Online: A web-based tool for land cover and land use reference data collection. *Remote Sensing*, *11*(21), 2535. https://doi.org/10.3390/rs11212535

Tolan, J., Yang, H. I., Nosarzewski, B., Couairon, G., Vo, H. V., Brandt, J., Spore, J., Majumdar, S., Haziza, D., & Vamaraju, J. (2023). Global tree canopy height map from high-resolution satellite imagery. *Science*, *380*(6655), 1245–1249. https://doi.org/10.1126/science.ade7957

U.S. Environmental Protection Agency (US EPA). (2022). *Climate change indicators: Puerto Rico.* https://www.epa.gov/

U.S. Geological Survey (USGS). (2020). *Landsat 8 data users handbook*. https://www.usgs.gov/