Comparative Land Cover Analysis and Evaluation of Remote Sensing Tool Accuracy Using GLOBE Observer Photos

Buheis, Lucas; Farber, Sophia; Karvir, Aakash; Mdinaradze, Sophia; Ramprashad, Jamie; Shindano, Rinnah; Talwar Tvisha; Wheeler, James

Abstract

The applications of land cover data are numerous, as land cover includes both anthropogenic and natural material, from buildings to bodies of water. Accordingly, land cover change tells the story of a location in many ways because land cover is interconnected with human and animal activity, climate and weather, and natural disasters. Our team collected photos of the land cover in eight locations around the continental United States in order to explore local land cover trends. We compared our field observation photos with several land cover datasets from remote sensing sources, identifying consistencies between datasets and our observations, as well as land cover types where disagreement occurred between sources. Our diverse locations and broad range of land cover types, coupled with additional remote sensing data related to human activity and environmental conditions, appeared well suited to an analysis of the influence of land cover on wildfire risk potential. As a result of our land cover investigation, we seek to answer the question: How can we utilize artificial intelligence, remote sensing, and land cover data to classify the susceptibility of a region to wildfire spread once one begins, to assist in safety procedures and containment efforts? Our goal is to create FIRECAST: a Fire Index Risk Estimator using Climate, Anthropogenic, and Soil Trends. This exploratory tool will utilize machine learning to provide timely wildfire risk assessments to aid local communities in resource allocation and preventative measures. Key in development of FIRECAST will be our land cover photos, taken using the GLOBE Observer App, to validate land cover classification. As we continue our research, we hope to incorporate more input datasets, including more social aspects, to improve the accuracy of FIRECAST and ensure greater consistency with established fire risk indices. *Keywords*: land cover, GLOBE Observer, remote sensing, wildfire, risk assessment

Introduction

In recent years, climate change has resulted in the increasing frequency of wildfires, which not only devastate local ecosystems, but also endanger human lives and critical infrastructure. Unfortunately, many areas undergoing rapid urbanization have not properly assessed fire risk resulting in ineffective resource allocation and policies for fire mitigation. Thus, analyzing factors contributing to wildfire and being able to predict the change of wildfire is crucial in assisting cities to devise their response strategies.

There are a few crucial factors to take into consideration when looking at fires. Land cover is a critical factor to consider, as it reflects both natural and human-build features of the environment that might influence the risk and impact of a wildfire. Though factors such as vegetation, infrastructure, impervious surfaces, and water bodies impact the chances of wildfire spread, remote sensing data often falls short here as its datasets don't evolve with changing landscape. As a result, this project uses the GLOBE Observer app to collect land cover data in order to validate satellite imagery. By synthesizing this data with other metrics known to increase fire risk (such as soil moisture, precipitation levels, and average temperature), the accuracy of the model increases.

The objective of this study is to propose an effective ML-based methodology for assessing the risk of a wildfire emerging in a specific area. This led to the development of FIRECAST: Fire Index Risk Estimator using Climate, Anthropogenic, and Soil Trends, which is a machine learning-based tool designed to evaluate regional wildfire susceptibility. By integrating citizen science data with satellite imagery and climate variables, FIRECAST aims to offer timely, location-specific fire risk assessments to aid community preparedness, and hopefully improve cities' abilities to respond to wildfires when they break out. Though this is just exploratory research, this approach showcases how technologies can help communities better understand and prepare for wildfire threats in a changing world.

Methods and Materials

Case studies were conducted using NASA's adopt-a-pixel method, where we investigated 3 kilometer areas in our hometowns. The Areas of Interest (AOIs) included 37 locations laid out in a 6x6 grid, with a centroid, that were evenly spanned out over 3 kilometers. The points are spaced 500 meters apart, with 100x100 meter areas surrounding each point. The projected sampling grid was created using NESEC (Wp_Admin, n.d.) that provided a CSV file with each of 37 longitude and latitude coordinates along with a GeoJSON file. To increase efficiency in field investigations, the CSV file was uploaded to a web-based mapping service that provided directions to each of the points in the AOI.

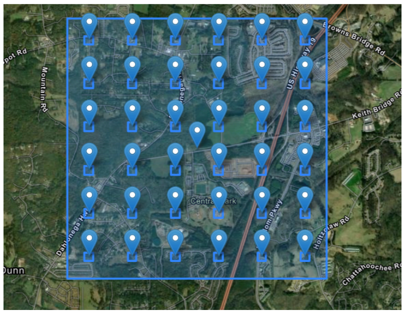

Figure 1: 6x6 Sampling Grid (Wp_Admin, n.d.)

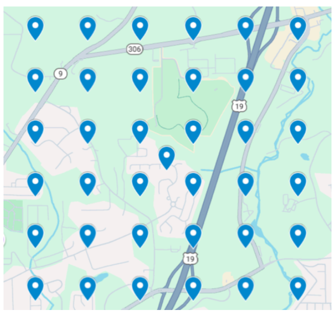

Figure 2: Web mapping service (Google Maps, n.d.)

Field Investigations included traveling to each of the 37 locations and photographing them in the GLOBE Observer application (*GLOBE Observer - GLOBE.gov*, n.d.). The GLOBE application provided a method of observing land cover by taking photos at each site in all directions (up, down, north, east, south and west). Once the site was captured, we analyzed each image by calculating the percentage of each land cover type present in the photo. Once observations were completed at each of the 37 sites, the data was uploaded to the GLOBE Visualization System for further analysis.

Fig 3: Up, Down, West, South, East, North Images(In order). Authors' own work.

After all the data was collected, we used Earth Map (*Earth Map*, n.d.) to interpret and extract new meaning from our results. Earth Map provided the ability to contextualize each of our AOIs by providing a wide breadth of data resources to pull from. The Earth Map datasets we utilized included the Dynamic World 10m, ESRI 10m, and World Cover 10m land cover layers, which were created using satellite data from the Sentinel1 and 2 missions, and the WRI/Meta 1m Tree Canopy layer using satellite data from the WorldView-4 mission. This data made it possible to come up with the idea of FIRECAST. Though still in the development process, FIRECAST will be trained using data from GHSL (*Global Human Settlement - GHSL Homepage - European Commission*, 2016), NDVI (*Normalized Difference Vegetation Index - Didan K,* 2021), and water deficit from MODIS along with soil moisture, maximum temperature, precipitation, and wind speed. To further refine the effectiveness of FIRECAST, we are utilizing the images collected through the GLOBE Application to visually validate land cover conditions and enhance the spatial relevance of our model inputs. To produce this tool, we plan to train FIRECAST with XGBOOST, which has the ability to handle nonlinear environmental data.

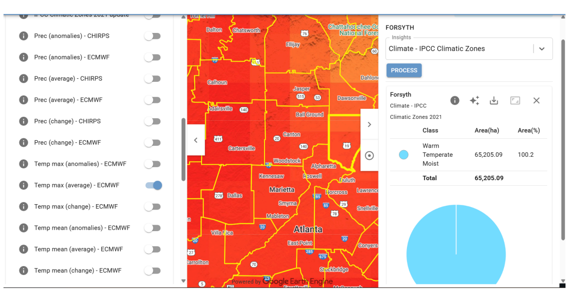

Figure 4: Earth Map data site (Earth Map, n.d.)

Results

These case studies provided information about the similarities and differences between the land cover in different locations and how it changed over time. In recent years, the majority of these locations have become more urbanized and have exhibited an increase in population, which in some cases was reflected by our times series and land cover data.

The GLOBE Observer photos were particularly useful when verifying the accuracy of the satellite data and images. Since the photos were more precise, they were able to show the many different types of land cover in a small region. Even though the photos often did not show the entire 50m grid square around our AOI point, often due to vertical obstructions such as trees and buildings, their land cover is consistent with the rest of the square due to its small area.

We were unable to obtain images of several of our AOI points due to difficulties in locating and inability to access certain areas, leading to some data loss. For some team members, AOI points were on private property, while others had difficulties with AOI points surrounded by protected areas or dense vegetation, rendering them inaccessible. In addition, we faced several technical difficulties while using GLOBE Observer that prevented images from uploading or led to our images being lost from their database. We were able to recover much of that data by revisiting our sites and making new observations that were successfully saved.

In summary, the majority of the case study locations had an abundance of vegetation and tree cover with portions of built-up areas. This is demonstrated by Figure 4, which shows the full grid areas of three locations from our case study. These examples show similarities between how this vegetation is distributed through our areas of interest; tree cover and grassland tended to be more dispersed throughout the area, but the built-up areas were more concentrated. An exception to this was in the third image, where the tree cover and built-up areas were scattered around each other, likely due to many large pockets of trees and vegetation in cities. Other areas, such as Van Nuys, are more urban and suburban, with little vegetation. However, some vegetation is present around residential areas on sidewalks and lawns.

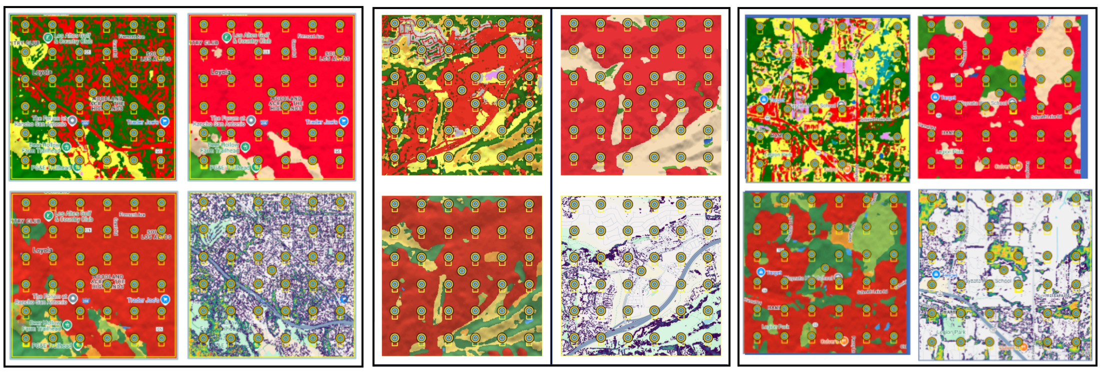

Figure 4: The full grid areas of Los Altos, California; San Diego, California; and Plymouth, Minnesota, from left to right. Each location shows a grid with data from World Cover (top left), ESRI (top right), Dynamic World (bottom left), and Tree Canopy Meta (bottom right). (Earth Map, n.d.).

Though our data was able to capture the general trends of each region's land cover, there were many inconsistencies between datasets. ESRI and Dynamic World present each region as very built-up, while Meta Tree Canopy and World Cover show more vegetation and tree cover.

Some more examples of these discrepancies are in Figure 5, AOI points 22, 23, and 24 in the case study of Cumming, Georgia. In Point 22, Dynamic World completely overrepresents the amount of built-up land cover, while ESRI shows the density and location of the tree canopy incorrectly. On the other hand, Dynamic World's representation in point 23 is correct, while ESRI misrepresents the area as being built-up. Finally, point 24 contains accurate data from Dynamic World and ESRI, and contrary to most points, the Tree Canopy Meta and World Cover data is incorrect due to a misrepresentation of the location of tree cover.

Figure 5: AOI points 22, 23, and 24 in the case study of Cumming, Georgia. The datasets shown are (from left to right) Landsat Time Series, Tree Canopy Meta, World Cover, Dynamic World, and ESRI. (author's own work)

Similar issues were reported in our case study of Houston, Texas. World Cover only showed instances of built-up land cover in regions with very little vegetation, such as large parking lots, while ESRI and Dynamic World only showed tree cover in areas where the vegetation was incredibly dense. Figure 6 shows AOI points 19 and 20, which can show this discrepancy. Point 19 accurately represents the area's land cover because of the area's urban nature. However, in point 20, World Cover overrepresents the tree cover, while Dynamic World and ESRI show it to be more built-up than it truly is.

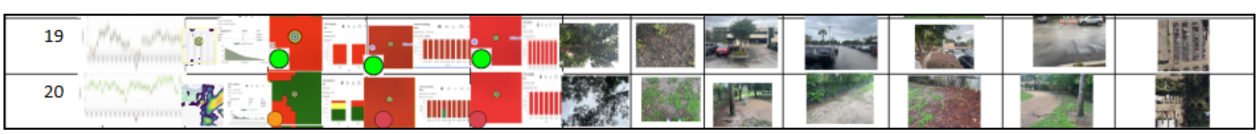

Figure 6: AOI points 19 and 20 in the case study of Houston, Texas. The datasets shown are (from left to right) Landsat Time Series, Tree Canopy Meta, World Cover, Dynamic World, and ESRI. (author's own work)

We noticed another error in the data, specifically in the LandSat Time Series, where the reliability of these images is inconsistent between the different locations. The cases of Van Nuys, California, and San Diego, California, both show the area gradually becoming more urbanized. The more recent images have more instances of purple, representing asphalt, and white, representing buildings. Figures 7 and 8 show both of these comparisons.

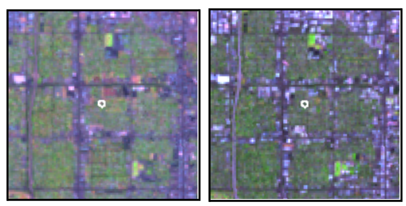

Figure 7: Images of Van Nuys, California in 1985 (left) and 2025 (right). (Landsat Time Series, Google Earth Engine, 1985 and 2025).

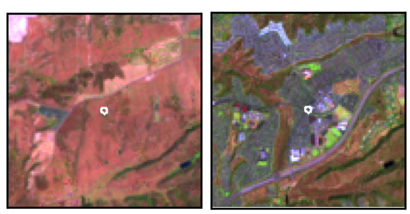

Figure 8: Images of San Diego, California in 1985 (left) and 2025 (right). (Landsat Time Series, Google Earth Engine, 1985 and 2025).

However, the landsat time series data was not always consistent with our knowledge about change in our communities. An example of this is the case study of Los Altos, California. The area has undergone a significant increase in population and urbanization in recent years, but this is not reflected in the satellite images. Rather, the amount of vegetation in the area has seemingly increased since 1985. This is shown in Figure 9 below.

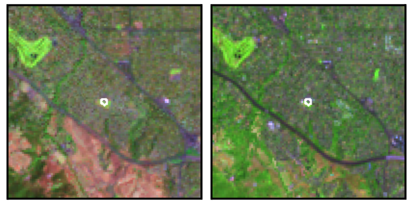

Figure 9: Images of Los Altos, California in 1985 (left) and 2025 (right). (Landsat Time Series, Google Earth Engine, 1985 and 2025).

Discussion

The results of these case studies provide insight on both the regions being examined and the datasets being used in our analysis.

Our data has a large margin of error due to the inconsistencies between our GLOBE Observer ground photos, the Collect Earth Online satellite images, and the datasets available on Earth Map. While our Meta Tree Canopy and World Cover data tended to match when comparing tree cover, they tend to differ from the results given by the Dynamic World and ESRI datasets. After comparing them to our GLOBE observations, we determined all datasets misrepresented the land cover; they only considered the broader land cover and not the variations at a smaller scale. While World Cover and Tree Canopy Meta overrepresent the tree density in areas, Dynamic World and ESRI overrepresent how built-up an area is. This high error value is surprising, since all of the datasets have a relatively high resolution of ten meters.

An example of this is shown in Figure 5, the case study of Van Nuys, California. The ESRI and Dynamic World datasets represent its land cover in alignment with its general trend as an urban, built-up area. However, the increased precision of the Meta Tree Canopy and World Cover datasets allow them to show the instances of tree cover and grass throughout the residential areas. Due to ESRI and Dynamic World's inability to recognize the sporadic placement of trees, the World Cover and Tree Canopy data more accurately represent the land cover.

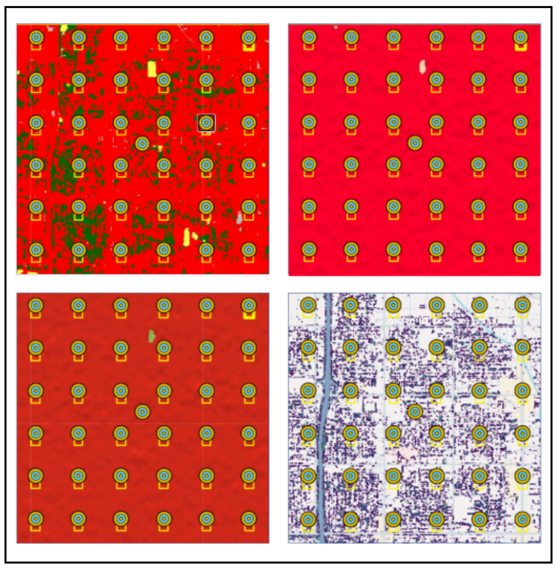

Figure 10: the land cover of the AOI grid of Van Nuys, California using World Cover (top left), ESRI (top right), Dynamic World (bottom left), and Tree Canopy Meta (bottom right). (Earth Map, n.d.)

There are a multitude of variables that can explain the discrepancies. For example, each dataset may use different criteria to classify the land cover of an area. Other explanations are obstructions in the images and the resolution not being sufficient for this purpose. Finally, inaccuracies between the data sourced from Earth Map can also be explained by the fact that it was collected several years earlier than our GLOBE observations and Collect Earth Online satellite images, which were used to verify them. It is possible that there were significant changes in vegetation over time, leading to the Earth Map information being outdated.

The inaccuracies in the Landsat Time Series's representation of Los Altos's change over time may have similarly been caused by the satellite's limited resolution. It may have not been able to pick up data that was detailed enough to show the known trends in urbanization and growth the city had exhibited.

Due to the discrepancies we faced while analyzing our data, it would be useful to consider datasets using satellites with a higher resolution in the future. In addition, our research highlighted the importance of confirming the accuracy of all data used in an experiment to prevent incorrect conclusions from being made.

Conclusion

Our case studies investigated land cover trends across the United States and demonstrated the limitations of remote-sensing data. While satellite-based sources provide valuable large-scale perspectives though comprehensive datasets, our on-site observations revealed consistent misclassifications between datasets, typically involving areas with mixed environments, such as urban-natural environments and areas containing assorted vegetation.

The idea for our proposed fire index risk estimator—FIRECAST—emerged from these findings. Utilizing machine-learning, FIRECAST aims to provide accurate data on wildfire susceptibility to support the prevention and containment of wildfires. Though still an exploratory project, FIRECAST highlights the need for multi-source datasets and adaptive machine learning to address the challenges posed by the dynamic nature of landscape and climate conditions.

To further improve model accuracy, future research will focus on integrating higher-resolution imagery along with recently-gathered data, addressing the problems arising from the utilization of outdated data sources, especially as land cover is constantly evolving. Nonetheless, our research underscores the importance of gathering data through field observations and on-site data sources to circumvent data discrepancies and inconsistencies arising from the dependence on satellite-based sources, in addition to highlighting the value and importance of community-driven data collection.

Badge Descriptions

Below we explain how our research meets the criteria for three badges for the International Virtual Science Symposium.

I WORK WITH SATELLITE DATA

As explained in our methods and results, we compared our field observations from GLOBE Observer to land cover classifications generated using data from the Sentinel 1 and 2 missions. Data from these satellite missions were used to create the Dynamic World 10m, ESRI 10m, and World Cover 10m land cover classification layers accessed through Earthmap.org. These datasets were key in our analysis of the consistency between our observed land cover and land cover data layers. Additionally, our investigations of land cover change over time directly involved satellite images of our areas of interest, captured by NASA's Landsat missions and accessed through the Google Earth Engine Landsat Time Series Explorer. Finally, in the construction of our FIRECAST model, we seek to use data accessed through Earthmap.org based on satellite data. The Global Human Settlement (GHSL), Normalized Difference Vegetation Index (NDVI), and water deficit layers accessible in Earth Map are derived from satellite data from the MODIS mission, again demonstrating our dependence on satellite mission data in our continuing research in addition to the work we have already done. These layers are key to the FIRECAST tool, since human populations are a risk factor as well as an indication of the need for protection in fire-prone areas, vegetation conditions are highly influential in the potential for spread of a wildfire, and water deficit is related to vegetation, especially in causing vegetation to dry out and turn into better wildfire fuel. Because satellite data for these factors can be accessed on a large-scale with detail, these data are key to FIRECAST development.

I AM A DATA SCIENTIST

Each of the eight members of our research team collected land cover photos using the GLOBE Observer mobile app, and downloaded these photos from the GLOBE database in order to better understand land cover trends in their local area of interest. Original data collection was vital to developing our research question, as our GLOBE Observer photos were analyzed for consistency with existing land cover datasets. The GLOBE Observer photos will continue to be used in the training and construction of the FIRECAST model in order to validate land cover classifications that represent a significant risk factor in wildfire risk potential.

I AM A PROBLEM SOLVER

While doing research into our topic and analyzing different environmental issues that exist today, our group noticed that there were some major issues with the way cities responded to fires, and with increased wildfires happening in the past year, it is important now more than ever that fires are quickly responded to. In order to solve the issue of slow response time, and engineer a way to cut times, we brainstormed an indicator that would be able to assess a certain region's risk of having a wildfire that takes into consideration population and urban development. This would allow cities to better allocate resources to respond to fires, as they would be able to see which regions in specific have a higher risk of wildfires emerging. GLOBE images were used in this process to help identify which regions/areas have higher urbanization rates, as these regions are the ones in which fires could cause the most damage. Thus, our solution aims to provide a comprehensive solution to late wildfire responses amidst an urbanizing world.

“References”

Earth Map. (n.d.). https://earthmap.org/

CEO Collect Earth Online. (2023, August 30). *Collect Earth online*. Collect Earth Online - Satellite Image Viewing & Interpretation System. https://www.collect.earth/

Didan, K. (2021). MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V061 [Data set]. NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MODIS/MOD13Q1.061

Global Human Settlement - GHSL homepage - European Commission. (2016, July 6). https://human-settlement.emergency.copernicus.eu/

GLOBE Observer - GLOBE Observer - GLOBE.gov. (n.d.). https://observer.globe.gov/

Google Earth Timelapse. (n.d.). https://earthengine.google.com/timelapse/

Google Maps. (n.d.). Google Maps. https://maps.google.com/

Multi-Resolution Land Characteristics (MRLC) Consortium | Multi-Resolution Land Characteristics (MRLC) Consortium. (2025, July 28). https://www.mrlc.gov/

Wp_Admin. (n.d.). *NASA Earth Science Education Collaborative*. NASA Earth Science Education Collaborative. https://nesec.strategies.org/

Appendix

The data that we compiled from these sources is made available in posters for each case study, contained in the following github: https://github.com/sophia-mery/infernointelResources.git

Acknowledgments

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).