What's Up On The Roof: Tracking Cool Roofs in India with Satellite Imaging

VARCHITA LALWANI, ANUPAM SOBTI, and VISHAL GARG, Plaksha University, India
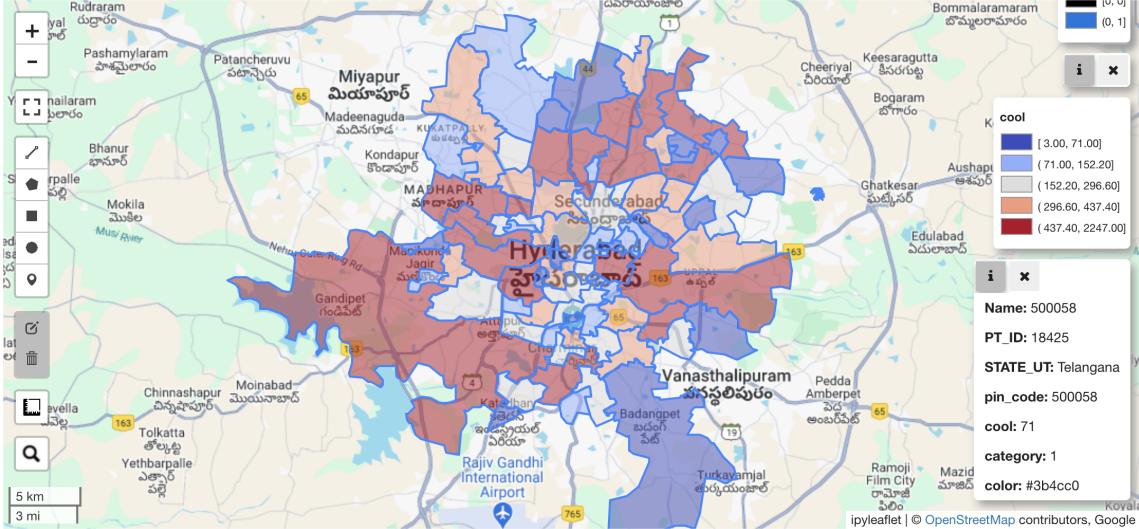
Fig. 1. We've developed a satellite image based tracker for identifying building-level roof reflectance information using machine learning. A cool roof tracker provides an eagle eye view to a city administrator on the progress of cool roof implementation in the city. The available information at the building level could also be utilized for regulatory and compliance purposes. This tracking also lays the foundation for real-time urban heat island monitoring in the city.

With rapid pace of urbanization, the increase in built up area is leading to urban heat islands which absorb a lot of the heat received through solar radiation. To reduce this effect, reflective coatings on top of building roofs, known as cool roofs are installed. Cool roofs reflect light back through the atmosphere to minimize the heat radiation within the urban space. In this paper, we introduce a cool roof tracker. The tracker provides building level information of whether a roof is cool or not, thus indicating if an urban heat island is likely to develop in a region. This information helps policymakers identify the vulnerable regions in a city and track the progress of cool roof implementation in all areas.

There are three key components of the system: building polygon extraction, reflectance features extraction from roof images; and classification into cool vs non cool buildings. There are a number of satellites operational today which offer different strengths - high resolution, high revisit time, multiple spectral bands, etc. In this work, we do a mix of quantitative and qualitative evaluation of the efficacy of different satellite modalities for each of the system components, build robust models with extensive experimentation on color space, augmentations, and generalization achieving an F1-score of 97% on our test set. We also develop a prototype dashboard as

Authors' address: Varchita Lalwani, varchita.lalwani@plaksha.edu.in; Anupam Sobti, anupam.sobti@plaksha.edu.in; Vishal Garg, vishal.garg@plaksha.edu.in, Plaksha University, Mohali, Punjab, India.

a demonstration for supporting data-driven decision making for state level heat mitigation plans as well as regulatory and compliance tool for measuring roof reflectance.

Additional Key Words and Phrases: Cool roof tracker for India, roof reflectivity, satellite imagery, sustainable building practices, remote sensing, machine learning techniques, OpenBuilding dataset, albedo values, cool roof classification, cool roofs

1 INTRODUCTION

Climate change is raising temperatures globally. March 2022 was the hottest month recorded in India in the past 122 years [15]. In this period, twelve people died from heatstroke at an outdoor event in Mumbai. Scientists from ten different countries found that climate change [7] made it 30 times more likely for India to experience the April heatwave this year. The default behavior to mitigate urban heat is the use of additional cooling equipment (and thus, electricity) in the city leading to even more heat being emitted in the environment causing an urban heat island effect. This increases the vulnerability of the urban poor and poses an undue demand on the electric grid. To prevent this catastrophic situation, there must be passive solutions to tackle heat accumulating in urban areas. While planting trees, preserving water bodies are important steps; even when we build concrete structures, cool roofs are a way to reflect the heat back into the atmosphere [37]. A cool roof is a reflective coating/material on the roof designed to reflect heat carried by the sunlight back through the atmosphere.

The sun emits energy across a wide range of wavelengths. Most of this energy falls between 300 nanometers (nm) and 2500 nm. Within this range, about 5% is in the ultraviolet (UV) range (300-400 nm), about 43% is in the visible range (400-700 nm), and about 52% is in the near-infrared (NIR) part (700-2500 nm). When sunlight hits a roof, some of it bounces off or scatters, giving the roof its color and gloss. The rest is absorbed by the roof. This absorbed energy heats the roof which then further radiates into the atmosphere. The roof spreads heat through convection - with air in the atmosphere, conduction - through the layers within the roof such as insulation, and, radiation - through the release of electromagnetic waves, which can heat anything solid that absorbs them, including nearby air or objects. A cool roof is designed to reflect more sunlight and absorb less heat compared to regular roofs. This is through specialized materials with higher reflectivity and less absorbption in the infrared spectrum. This can save energy by keeping indoor temperatures lower, and hence reduce the need for air conditioning. Cool roofs have proven to be [22] an effective mechanism even for rural communities for economically building resilience to heat. In summary, cool roofs are high in both reflectivity and thermal emissivity. Therefore, it's a cost-effective green technology, especially effective for places with high temperatures.

Indian government as well as non-governmental organizations are actively working towards validation and adoption of cool roof solutions. A study from 2011 [8] shows that installing cool roofs could save 10 to 19 percent energy in the top floor of buildings in Hyderabad, potentially reducing citywide air temperature by $2°C$ ($3.6°F$) and could save five billions rupees in electricity over 10 years across India. The cool roof deployment has also been studied internationally. The effect of painting roofs with special reflective coatings on energy use in Florida homes is evaluated in [24]. The study was done on nine houses from 1991 to 1994, comparing how much electricity they used before and after making the roofs reflective in the middle of summer. The houses reduced electricity consumption for air conditioning after the roofs were painted white by upto 42% with an average of 19%. The amount of cooling savings also depend on factors

like ceiling insulation, reflectivity of roofs, presence of air ducts, sizing of the AC, etc. Pilots in Boston[28], NYC[10] have also advocated widespread deployment of cool roofs. In 2013, New York City in the U.S. had cooled rooftops covering an area over 45 acres. This helped offset over 831 tons of CO_2e (carbon dioxide equivalent)[22].

To track the need and state of implementation of the cool roofs program, various cities collect high resolution orthorectified imagery of the cities and also label building footprints [12]. However, in the Indian context, with just the sheer scale of cities and limited resource allocation to this activity, it is imperative to device a cost effective and scalable solution. Therefore, we do not assume availability of such data and rely completely on satellite sources. There are three components to our system: i) **building footprint extraction** using high resolution images from the maxar satellite ii) **reflectance feature extraction** from roof images, and, iii) **classification** of cool vs non cool roof. Extracting building footprints requires high resolution (<50 cm) imagery typically acquired on demand through satellite maneuvering[2]. Due to this on demand nature, it is costly to acquire. However, services such as Google Maps API[1] allow access to images captured in the past for almost all locations on earth at a fraction of the cost. Since building construction is a slower process, this data suffices for polygon extraction. For extraction of reflectance features and specifically infrared reflectance, however, multiple spectral bands are required which are not available at such high resolution. Hence, we evaluate various satellites - Sentinel2[11], Planet[13], etc. for this purpose. A review of satellites used in this work is shared in Table 1. We have explored different machine learning classification methods for cool/nooncool classification. We also compare the two paradigms for cool roof classification - direct classification (cool vs non-cool) and estimation of reflectance (derived albedo), thereafter, using a threshold to classify cool/non cool buildings. To learn the latter in high resolution, we also propose a method to learn reflectivity (albedo) with 3m resolution planetscope bands as inputs and low resolution albedo measurements from Landsat[33] as target values. We focus our evaluation on methods which are explainable and interpretable and also demonstrate this understanding through our analysis.

Our key contributions are as follows:

(1) We develop and evaluate an entire pipeline for identification and classification of building roofs as cool/non-cool roofs. The evaluation uses a mix of quantitative and qualitative evaluation to assess various satellites and modalities for their efficacy in cool roof tracking.

(2) We device a method for leveraging low resolution albedo estimates from Landsat[33] images with high resolution satellites and thus provide high resolution reflectance (derived albedo) features for roof images.

(3) We build robust classifiers for cool vs non cool building classification that generalize to different cities, sizes, etc.

(4) We build a prototype dashboard to demonstrate data-driven decision making at a city level, sub-city (pin code) level and individual building level.

2 RELATED WORK

Practical deployments of this nature are highly interdisciplinary. The decision making is complex due to various aspects that need to be understood - materials for cool roofs, urban heat islands, technology for sensing, socio-economic conditions, etc. We cover related work in this broad range to facilitate broader understanding of the ecosystem.

Roof Materials. There are several materials where commercial products for cool roofing are available now, e.g., precoated concrete, asphalt, brick, timber, canvas, PVC, polycarbonate, fiberglass, spray, paint, membrane, and metal sheet. While white-colored cool roofs are most common, coloured cool roofs are also available for aesthetic purposes. The colored roofs still have a high reflectivity in the infrared spectrum, making it an effective solution. Cool colored materials can be defined as spectrally selective substances exhibiting moderate or low visible solar reflectance but

Satellite Name	Resolution	Total Bands	Band Names	Collection Frequency
Maxar	0.5m	3	R, G, B	On-demand
PlanetScope	3m	8	R, G, B, NIR, RE, Y, G1, CB	Daily
Basemaps	3m	4	R, G, B, NIR	Continuous
Sentinel-2	10-60m	13	CA (60), B (10), G (10), R(10), VRE (20), VRE (20), VRE (20), NIR (10), VRE (20), WV (60), SWIR-C (60), SWIR (20), SWIR (20)	Every 5 days
Landsat 8	30m	11	CA, B, G, R, NIR, SWIR1, SWIR2, PAN, C	Every 16 days

Table 1. Satellite Information - Total bands and frequency of collection. The full forms of bands names are, R: Red, G: Green, B: Blue, RE: Red Edge, Y: Yellow, G1: Green I, CB: Coastal Blue, CA: Coastal Aerosol, NIR: Near-Infrared, SWIR: Shortwave Infrared, SWIR: Shortwave Infrared, PAN: Panchromatic, C: Cirrus, VRE: Vegetation Red Edge, WV: Water Vapour. For sentinel the number in brackets are resolution for that band. If the name of the two bands are same or differentiated with numerics such as 1, 2 etc, it means that their wavelengths are different

high near-infrared reflectance, leading to moderate solar reflectance and high thermal emittance [32]. The reflectance characteristics of different cool roof materials are shown in [32]. There have also been efforts in low-income communities to use local material such as tires and white tarp on metal or asbestos roofs for making temporary cool roofs. This helps in informal housing, however this solution was discarded being prone to water pooling and diseases on the roofs. In 2017 Ahmedabad and Hyderabad piloted cool roof programs [22]. Picking the right material for a cool roof is essential but non-trivial. There are lots of choices, and they differ in how much heat they let out, how well they reflect sunlight, how long they last, and how much they cost upfront. The best material also depends on the weather, the kind of roof we have, and the HVAC system we use. In some countries, there are groups like the Cool Roof Rating Council that test and label roofing materials to see how good they are at reflecting sunlight and releasing heat. They have a list of over 1,000 materials, mostly for the U.S. market. In Europe, the EU Cool Roofs Council rates different roof materials using something called the Solar Reflectance Index (SRI). In India, institutions like IIIT-H have testing centers to see how well different materials work in the Indian climate.

Monitoring technology. Kim *et. al.*[17] propose a method for identifying the location of buildings using image processing techniques applied to digital surface model (DSM) images. Subsequently, it uses this information to determine the materials present on the roofs of these buildings on DSM images. The approach involves detecting buildings in satellite images by analyzing road patterns. Then, a Convolutional Neural Network (CNN) is employed to classify the roofing materials, which include concrete, healthy metal, unfinished roofs, and other types. The Open AI Caribbean Challenge [4] focuses on predicting the type of roof in drone imagery. The dataset includes overhead images from seven locations across three countries, featuring labeled building footprints and images with approximately 4cm resolution.

The images have 3 channels of RGB and a 4th mask channel of opacity. The goal is to classify each building footprint based on the roof material type. Solovyeh *et. al.*[29] utilize a neural network model adapting ImageNet weights to classify roof materials into categories such as concrete cement, healthy metal, incomplete roofs, and other types. The method primarily focuses on the analysis of multi-channel images. Concentrating on identifying roof tiles containing asbestos, Osińska *et. al.*[23] address the challenges of distinguishing roofing materials using high-resolution satellite images. Roofs exhibit a wide range of shapes and textures influenced by varying illumination conditions. The proposed solution involves applying topographic correction to satellite images, aiming to alleviate the impact of lighting variations in airborne laser scanning (ALS) and WorldView-2 multispectral satellite images. The research assesses the effectiveness of this approach in classifying roofing materials, specifically utilizing an 8-channel WorldView-2 satellite image. The results demonstrate that implementing topographic correction enhances the accuracy of material classification. A strategy to reduce urban temperatures and energy consumption is the widespread application of cool materials with high albedo and high emissivity in the urban envelope. The objective of [9] is to design a procedure that allows the discrimination of different roofing materials within the Metropolitan Area of Mendoza, Argentina, to determine the possible energy savings and improvements in urban micro-climatic conditions associated with the increase of albedo in the roofs of the city. They work with Sentinel 2a imagery and figure out what materials are on the roofs and 3 series of spectral signatures were obtained from built urban areas. Big cities growing quickly change the natural environment a lot. The use of concrete and asphalt has had direct impacts on the urban environment such as land-surface temperature (LST) and atmospheric dynamics. Using lots of concrete and asphalt makes cities hotter and affects the weather. These changes to the climatic variables have long-term and short-term consequences.

Urban Heat Island Understanding. A study was conducted on Auburn University's campus to accurately measure Land Surface Temperature (LST), solar irradiance, and albedo using remotely sensed images and geographic models [14]. The objective was to comprehend how environmental changes might impact energy consumption. They employed a statistical model to reveal the connections between natural changes and urban settings, providing insights into potential future scenarios. Identifying these connections enables strategic planning, including the consideration of cool roofs that help lower temperatures by reflecting sunlight. The study utilized data sources such as Landsat 8 OLI (Operational Land Imager sensor) with a resolution of 30 meters, OpenStreetMap, Google Earth, and LiDAR imagery at approximately 3 meters resolution. Understanding how much sunlight a single roof reflects or how much surface albedo at individual roof scale is also studied. A method is presented by Kalantar *et. al.*[16] for quantifying surface albedo of individual roofs in a complex urban area using the integration of Landsat 8 and airborne LiDAR data. First, individual roofs were extracted from airborne LiDAR data and orthophotos using optimized segmentation and supervised object based image analysis (OBIA). Support vector machine (SVM) was used as a classifier in OBIA process for extracting individual roofs. After that, surface albedo was calculated for each individual roof from Landsat images and generated thematic maps of mean surface albedo of individual roofs. To assess the impact of greenery and photovoltaic panels, Wu *et. al.*[34] introduced Roofpedia, a system for tracking solar and green roofs across 17 cities. The authors present a set of three contributions: (i) Utilizing satellite imagery for the automatic identification of pertinent urban roof types; (ii) Establishing an accessible roof registry to map out the spatial layout and size of solar and green roofs among over one million buildings spanning 17 cities; and (iii) Introducing the Roofpedia Index, derived from the registry data, to assess and compare cities based on their adoption of sustainable roof practices, including solar and green roof implementation. They utilized satellite images with a resolution of 50 centimeters through the Mapbox static tiles API. Additionally, information about building footprints was extracted from OpenStreetMap. Using image processing techniques on

WorldView-3 sensor images [31] identifies buildings with asbestos roofs near Prato, Italy. [30] leverages deep learning techniques for the automated classification of roof characteristics from very high-resolution ortho-photos and airborne LiDAR data obtained in Dominica following Hurricane Maria in 2017.

Policy/regulations on cool roofs. Indian cities of New Delhi, Ahmedabad, and Hyderabad are leading with initiatives to adopt cool roofs on public and government buildings. Ahmedabad and Hyderabad are working with Natural Resources Defence Council (NRDC), Administrative Staff College of India (ASCI), Public Health Foundation of India (PHFI) and Indian Institute of Public Health Gandhinagar (IIPH-G) to develop their cool roof programs. After successfully trying out cool roofs in Hyderabad in 2017 [22], the Telangana Cool Roofs Policy 2023-2028 suggests using cool roofs to handle extreme heat, make people more comfortable, and reduce hot spots in Telangana's cities. The policy aims to install at least 300 square kilometers of cool roofs by 2028. Even internationally, the research conducted by [12] examines the effectiveness of New York City's cool roof program in mitigating the urban heat island effect. The study tries to answer two questions: 1) To what extent do cool roofs contribute to reducing land temperatures? 2) Is there a trend of increased usage of cool roofs in neighborhoods experiencing the highest levels of heata? The goal is to see if the cool roof program is being done fairly across the city and helping the areas that need it the most. In 2020, New York City had 158.52 km^2 of rooftops, and about 38% of them had cool roofs. To create a cool roof dataset for New York City in 2020, they combined Landsat OLI Level 2 Science Products images from 2018 and 2016 with New York State orthoimagery from 2020. The Landsat images were used to train an algorithm to measure roof reflectivity based on methods by Liang (2001) and Smith (2010) [20] for Landsat 5 and 8 bands, respectively. The resulting dataset provides roof reflectivity values on a 0-1 scale, identifying cool roofs as those with reflectivity values above 0.6. The details of the method is descibed in Section 4.2.2. Vector datasets, including building footprints, borough boundaries, and the Heat Vulnerability Index (HVI) were integrated. Statistical analysis, including descriptive statistics, correlation, and multiple linear regression, was used to assess the effectiveness of cool roofs in mitigating land surface temperatures. A two-way ANOVA evaluated their implementation in heat-vulnerable neighborhoods.

Our work in context. The translation of these works to the Indian setting had a few special considerations: i) We wanted to build a scalable method with limited financial allocation to the activity, thereby restricting us to the use of satellite imagery ii) We'd still like to capture the reflectance of a diverse set of buildings - in terms of their sizes, architectures and locations. Smaller buildings pose a challenge compared to larger ones due to several reasons: satellite images have limited resolution, meaning they can only capture details down to a certain size. Smaller buildings may not be clearly visible or distinguishable in the images, especially if the resolution is not high enough (refer Figure 2). This limited resolution can make it difficult to precisely identify the boundaries and characteristics of smaller buildings. Also small buildings may be overshadowed by larger surrounding structures or objects, making them difficult to identify or analyze accurately in the satellite images. Additionally, smaller buildings may blend into their surroundings more easily, further complicating their detection. This can make it challenging to extract meaningful information about the roofs of smaller buildings from the images alone. Factors like cloud cover, shadows, or vegetation can obscure smaller buildings in satellite imagery, further complicating the analysis process. These factors can reduce the visibility of small buildings and make it harder to accurately assess their roofs.

In our research, we utilize 50cm resolution RGB satellite images from Maxar for building polygon extraction but utilizes Planetscope[13] satellite images with a resolution of 3 meters for capturing the refletance information in different bands. The 50cm resolution images are limited to RGB bands only. In contrast, our study investigates the utilization of multiple bands for classification, which offers more comprehensive data compared to traditional RGB images. However,

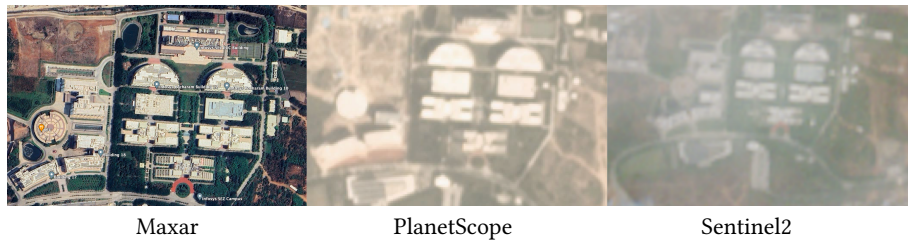|		
Maxar	PlanetScope	Sentinel2

Fig. 2. Example images of the same region from Maxar, Planet and Sentinel satellites. The resolution of images decreases from 50 cm in Maxar to 3 cm in Planetscope and 10 m in Sentinel2. However, the range of spectrums captured increases through the images with RGB in Maxar, 8 bands in Planetscope (in visible and near infrared spectrums) and 13 bands in Sentinel2 (including long wave infra-red spectrum which can directly sense heat).

even with this additional information, feature engineering becomes crucial, especially when dealing with the extraction of satellite building polygons, which tend to have a diverse range of polygon sizes. We also developed a specialized dashboard for cool roofs to aid urban planners in making policy decisions. This dashboard offers comprehensive analysis at multiple levels, including city-wide, pin-code level, and individual building insights. In our Hyderabad city dashboard (Figure 1), cooler regions are represented by shades of blue. Density information is provided for each pin code, allowing us to visualize the distribution of cool roofs and identify the coolest areas based on density. This visualization not only highlights the distribution of cool roofs but also identifies the coolest areas based on density. These detailed insights can inform urban planners about the effectiveness of cool roof initiatives and guide strategic interventions to mitigate urban heat island effects and enhance urban resilience. Further validation of this dashboard is outside the scope of this work.

The paper is organized into sections as follows: In Section 3, we present the qualitative analysis for building polygon and reflectance image extraction. Section 4 delves into the machine learning choices for classification of cool roofs. Our results are presented in Section 5, followed by a further analysis of intuition behind features learnt in Section 6. We present the details of our first dashboard prototype in Section 7 and conclude with Section 8.

3 BUILDING POLYGON AND IMAGE EXTRACTION

In Figure 2, we show views of the same buildings through different satellite images. As is evident from Figure 2, high resolution imagery is required for accurate building polygon extraction. In this section, we discuss how building polygons, and thereafter, multispectral signatures of roof reflectance, are extracted from satellite images. We qualitatively evaluate how different choices, i.e., which satellites and resolutions are appropriate for this application. This section also elaborates on the collection of ground truth data used for training the classifier.

3.1 Building footprint extraction

Here, we evaluate four distinct methodologies, each contributing unique insights into the extraction process. These are state of the art methods in computer vision which use supervised/semi-supervised methods for building polygon extraction. A typical pipeline consists of a segmentation model that identifies building pixels from the overall image. This extracted blob is then polygonized with the help of a post-processing step to extract the exact polygon for the building. Since the focus of the paper is towards roof reflectance features and classifier design, our evaluation of these methods was qualitatively done. There is also lack of annotated datasets for building polygons in the Indian setting.
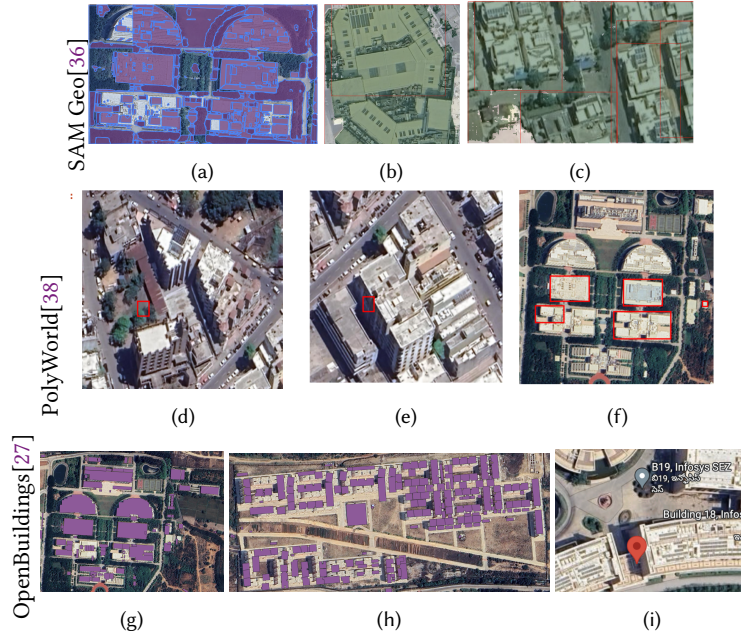
Fig. 3. Qualitative comparison of extracting building polygons using pre-trained models. (a), (b), (c) Prompt SAM (prompt used: buildings) [36] on tifs gives us low precision results, a huge amount of soil, land, greenery or background area is also getting extracted with buildings. Also, it is not able to extract seperate buildgs. (d), (e), (f) Using PolyWorldPretrainNetwork [38] - Buildings are exracted with very low recall. (g), A qualitative comparison of extracting building polygons using pre-trained models. Using OpenBuildingDataset - Even the small buildings are extracted. (h) Offset Imagery with OpenBuildingDatset - The extracted polygons have a little bit of a shift. In some images, this shift is quite big, but in most images, it's just a small shift. These are the biggest shifts found in the dataset. (i) Building in shadow - images lying in shadows so we remove them while pre-processing.

(1) **PolyWorldPretrainedNetwork** This neural network directly extracts building vertices from images, forming accurate polygons by predicting connection strengths between vertices using a graph neural network [38]. The assignments are determined by solving a differentiable optimal transport problem, and the vertex positions are optimized by minimizing a combined segmentation and polygonal angle difference loss. This network is trained on CrowdAI Mapping Challenge dataset. As depicted in Figures 3a, 3b, and 3c, we observe a minimal number of accurately extracted building images from a vast pool of existing buildings.

(2) **GlobalMLBuildingFootprint** This dataset employs a two-stage process [21] for building extraction. The initial stage involves semantic segmentation, wherein building pixels are recognized on aerial images using deep neural networks. Subsequently, the pixel detections are converted into polygons through a process known as polygonization. They provide geometry for different parts of the world. However, the details they provide for buildings in India aren't as useful for our specific project.

(3) **Segment Anything Model (SAM)** We also explored Segment Anything Model[18], which involves a new task, a model, and a dataset for image segmentation. This is a foundation model for image segmentation, trained in self-supervised fashion. The anticipated segments encompass 5 to 6 buildings within each segment. With lack of large scale annotations, this is not a feasible model and also it is trained on non-geospatial data. We also evaluated segment-geospatial [36], built on SAM for geospatial data. However, it doesn't give promising results

Model	Method	Trained on	Number of Samples
PolyWorldPretrainedNetwork	Graph Neural Network	CrowdAI Mapping Challenge dataset	>280k
GlobalMLBuildingFootprint	2-stage process (semantic segmentation and polygonization)	-	-
Segment Anything Model	Self-supervised	RGB Images	1 billion masks
SAMGeo	Python package for segmenting geospatial data with the SAM	-	-
OpenBuildingDataset	Instance segmentation and self-training	50 cm Maxar imagery	100k

Table 2. Models and Methods for Polygon Extraction: pre-trained methods for polygon extraction, including information on the satellite/image source they were trained on and the number of samples used.

with zero-shot learning on our data. As shown in Figures 3d, 3e, and 3f, when using text as input to identify buildings, it captures a lot of background noise and misses many buildings.

(4) **OpenBuildingDataset** We found results from continental-scale building detection from high-resolution satellite imagery [27], with a focus on the entire continent of Africa, using 50 cm satellite imagery. This approach utilizes a model training pipeline based on the U-Net architecture. The study delves into variations in architecture, loss functions, regularization, pre-training, self-training and post-processing, aiming to enhance instance segmentation performance. The experiments involve a dataset comprising 100k satellite images across Africa, with 1.75M manually labeled building instances. Examples are shown in Figure 3g.

Among these methods, as summarized in Table 2, **OpenBuildingDataset** exhibited superior performance across various scenarios (as assessed qualitatively), including the number of buildings extracted, the accuracy of extracted buildings with minimal offset, and the ability to extract the least amount of background noise. Consequently, it was selected over the other methods. To capture limitations of this model on our data, we want to highlight that despite the inclusion of a diverse range of training data, the OpenBuildingDataset faces challenges in certain scenarios – geological or vegetation features that may be mistaken for built structures, settlements with numerous contiguous buildings lacking clear delineations, areas characterized by small buildings which may appear only a few pixels wide at the given image resolution, rural or desert areas where buildings constructed with natural materials visually blend into the surroundings. Regions with high-rise buildings pose challenges, as the model is trained to detect building rooftops rather than bases. As a result, depending on the satellite image's viewing angle, the roofs of high-rise buildings may shift, making it difficult to track buildings consistently across a stack of imagery in a given location. Figure 3g shows that even the tiniest buildings are captured well. However, in Figure 3h, some images have a noticeable offset. The paper [27] suggests that this might be due to past imagery being used for prediction of polygons. The offset leads to a number of non-building pixels (shadows or other materials around the building) to be captured Figure 3i.

3.2 Building image extraction

3.2.1 Planet daily imagery. The planetscope satellite covers (almost) daily images of most regions on earth through a constellation of low earth orbit satellites. These are made available soon after using the planetscope explorer/API. The Planet (8 bands) images were obtained using the PSB.SD Super DOVE instrument type featuring Red, Green, Blue, NIR, Blue, Red Edge, Coastal, Yellow, and Panchromatic bands. Figure 4 shows some samples. Note that the values in these

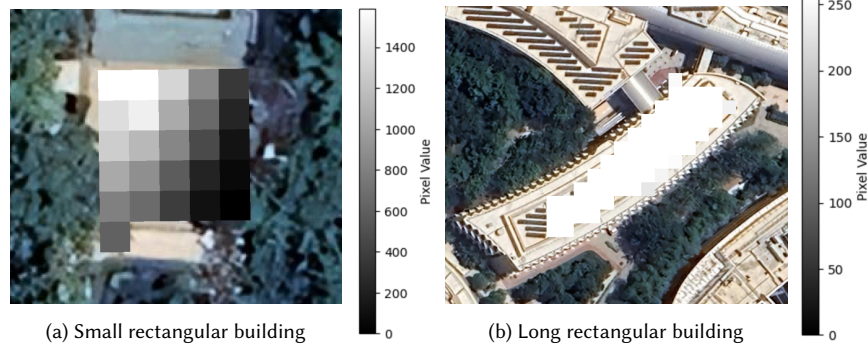(a) Small rectangular building   (b) Long rectangular building

Fig. 4. We used the extracted building polygons using OpenBuildings and extracted red channel from the planet 8 band (surface reflectance product) and 4 band (basemap) images. One can observe that depending on the shape of the building, different number of pixels might be returned.

bands are in very different numerical range. Thus, normalization is required before visualization. QGIS[26] has built in visualization functions for such data as well.

3.2.2 Planet Basemaps. Planetscope uses the daily satellite images and combine them for images that are consistent for scientific use. Both daily images and basemaps are ortho-rectified. Additionally, basemaps are spatially accurate data with minimized haze, light, and topographic effects, also seasonally color-balanced, seamless, and cloud-free. Basemaps have 4 bands instead of 8. Figure 4b. The data sourced for non-cool roofs is same as sourced for Planet (Table 3). Here also to enhance the diversity of our dataset, we extract images from three months (Table 4). Also, we ensured here that all the collected tiles had a minimal cloud cover, specifically less than 3%.

Region	Cool/Noncool	Min/Max Confidence	Min/Max Area (m^2)	Count
Hyderabad	Cool	0.6504, 0.9408	10.394, 7914.757	185
Noida	Cool	0.6512, 0.8905	10.1001, 3697.3973	41
Tamil Nadu	Cool	0.6524, 0.966	5.7126, 8144.7421	181
Noida	Noncool	0.6502, 0.9415	6.2636, 5996.127	718
Hyderabad	Noncool	0.6506, 0.9208	6.2657, 1707.1257	199
Other	Noncool	0.6508, 0.9434	6.3303, 1978.695	915

Table 3. Polygons from OpenBuildingDataset for our dataset. The table has information about the region, confidence of being a building (min/max), area information and total number of buildings extracted in that region

3.2.3 Sentinel 2. We also extracted images from Sentinel 2A satellite. We used all the bands at 10m and 20m resolutions. The building polygons often intersect with 1-2 pixels of the satellite image. These values are used as is. However, for a better estimation of reflectance, we derive a number of features from these images as discussed in Section 4.2.

3.2.4 Landsat. In the literature, one common method for estimating reflectance is to measure albedo. Albedo measurement can only be done through 5 bands (NIR, Shortwave and C-bands) from the landsat satellite at 30m resolution. In this work, we also derive a high-resolution albedo using Planetscope pixels as input and albedo as output. For obtaining

	Hyderabad (cool)	Tamil Nadu (cool)	Noida (noncool)	Noida (cool)	Anand (noncool)	Hyderabad (noncool)
Months (Basemaps)	July June May	August June May	August June July	August June July	August June July	July June May
Dates (Planet 8 bands)	10 August 31 May 24 May	21 August 29 May 24 May	17 August 11 June 14 June	12 August 14 June 11 June	1 June 31 May 24 May	10 August 31 May 24 May

Table 4. Basemaps and Planet imagery extraction information - To incorporate any temporal fluctuations, we extracted information from three months/dates in 2023.

high-resolution albedo data, we utilized Landsat 8 imagery at a resolution of 30 meters and 3m resolution Planetscope imagery.

4 PROPOSED METHODOLOGY

4.1 Ground truth

In our data collection process, we obtained ground truth data for both cool and non-cool roofs by searching for public announcements from companies and buildings regarding the installation of cool roofs. We conducted this search through news articles related to cool roofs. Additionally, for some samples, we contacted manufacturers involved in installing cool roofs to gather relevant data. We visually validated the roofs, as depicted in Figure 5. Cool roofs appear brighter than non-cool roofs even when visualized in RGB spectrum (although we found that reflectance was a better measure). Specifically, we identified cool roofs in Hyderabad, DLF Noida, and Tamil Nadu. For non-cool roofs, we sourced data from areas including Noida, Hyderabad and Anand, Gujarat making our **first set**. In total, as shown in Table 3 our dataset comprises 407 cool roof buildings and 1832 non-cool roof buildings from single date imagery. To enhance the diversity of our dataset and incorporate temporal robustness, we extracted information from three dates for the same buildings as shown in Table 4. Additionally, we ensured that all the collected tiles had a minimal cloud cover, specifically less than 3%. In total (for 3 dates), our dataset comprises of 1221 cool roof buildings and 5496 non-cool roof buildings, making our **second set**. We were also able to obtain 139 images through information shared by some cool roof manufacturers about their deployments. These images show different materials used for making roofs cooler, making our **third set**. The data encompasses different places, and variations in both area range and confidence of being a building from OpenDatasetBuilding as shown in Table 3. The information for all three sets is in Table 5.

4.2 Roof Feature Extraction

The primary objective is to create a binary classification system that can accurately determine whether a given roof is a "cool roof" or a "non-cool roof". We have experimented with 5 machine learning models namely, decision tree, random forest, logistic regression, multi-layer perceptron and gradient boosting with F1-score, precision and recall as metric for binary classification with losses as described in Table 7. All the classification results are F1 scores at an empirically derived confidence threshold of 0.5.

For images with 8 bands, we've extracted a total of 132 features for each color space. These features include statistics like mean, variance, entropy, minimum, maximum, median, range, 25th and 75th percentiles, coefficient of variance, skewness, kurtosis, and local binary pattern, making a total of 104 features across all bands. Additionally, we've

Fig. 5. Visual validation of our dataset, where the top row showcases regions with cool roofs and the bottom row showcases regions with non-cool roofs. From left to right in the top row, buildings are located in Hyderabad, Tamil Nadu, and Noida. In the bottom row, from left to right, the locations are Hyderabad and Anand. Cool roofs should appear brighter compared to non-cool roofs
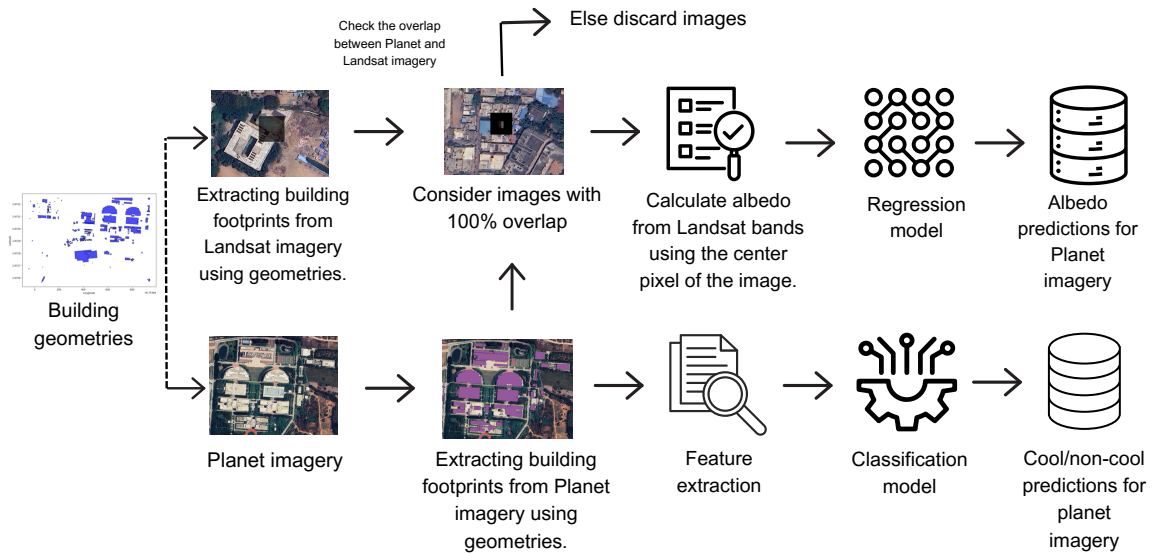

Fig. 6. We evaluate two approaches to cool roof classification: one focuses on feature extraction and cool roof classification methods while the other employs albedo extraction and regression for predicting high resolution albedo values (derived albedo) on planet imagery. Pass the geometry of the buildings (from OpenBuildings) to planet satellite imagery, extract roofs from the satellite images. Once we have the extracted satellite images for the buildings, extract useful features for each building. For the other branch in the pipeline, we pass the geometry of the buildings to landsat satellite imagery to extract roofs. To ensure correct mapping of planet and landsat pixels, we filter buildings where there is 100% overlap between the planet and landsat pixels. We calculate center pixel albedo or median albedo using landsat bands. Apply regression with planet band values as input and albedo values as output and we will have high resolution derived albedo predictions from planet imagery.

calculated the correlation between all color channels, such as red-green, red-blue, red-NIR, and so on, making a total of 132 features. This involves examining how these color channels relate to each other, giving us insights into the relationships within the image data. Table 6 displays features and explains why each feature is chosen.

Manuscript submitted to ACM

Region	Cool/Noncool	First Set	Second Set	Third Set
Hyderabad	Cool	185	555	555
Noida	Cool	41	123	123
Tamil Nadu	Cool	181	543	543
miscellaneous	Cool			139
		407	1221	1360
Noida	Noncool	718	2154	2154
Hyderabad	Noncool	199	597	597
Other	Noncool	915	2745	2745
		1832	5496	5496

Table 5. Description of all three sets used in this paper for cool/non-cool classification.

Feature	Intuition
Mean	Average pixel intensity, higher for reflective cool roofs
Minimum	Minimum pixel intensity, lower for potentially non-cool roofs
Variance	Spread of pixel intensities, higher for diverse materials
Maximum	Maximum pixel intensity, higher for reflective cool roofs
Entropy	Image complexity, lower for repetitive patterns in non-cool roofs
Median	Middle pixel intensity, less sensitive to extreme values
Range	Difference between maximum and minimum intensities
25th Percentile	Value below which 25% of pixel intensities fall
75th Percentile	Value below which 75% of pixel intensities fall
Coefficient of Variance	Relative variability of pixel intensities, higher for diverse materials
Skewness	Asymmetry of pixel intensity distribution, indicating certain patterns
Kurtosis	Sharpness of pixel intensity distribution, higher for concentrated values
Local Binary Pattern (LBP)	Captures local texture patterns, useful for identifying variations
Color Channels Correlation	Correlation between color channels, specific to reflective cool roofs

Table 6. Features extracted from images with 8 bands and their intuition for cool roof classification.

The process for cool rool classification is as follows and architecture is as shown in Figure 6 We downloaded high-resolution satellite imagery of the area of interest. From OpenBuildingDataset we acquired geometry of each building. Using the geometry information we extracted building polygons from the satellite imagery. This involves rasterio, geopandas operations to overlay the building footprints on the satellite imagery, essentially cropping or masking the imagery to include only the regions where the buildings are located. Then we extracted relevant features from these polygons for use in the classification model. These features served as input variables for the machine learning model. To prevent overfitting, we used k-fold validation with 5 folds. Once the model is trained and validated, we made predictions on the cities. This involves applying the model to additional satellite imagery to classify cool roofs across a larger area or for a different time period. The frequency and band information for every satellite is described in Table 1.

4.2.1 Effect of Color Space. We tried out different features taken from various color spaces [19]. The idea behind this is that each color space has its own special qualities or advantages.

RGB (Red, Green, Blue): This is the familiar color model based on how we humans see and understand colors. It's great basic color perception, but it might struggle to capture subtle color differences, especially in different lighting

Model	Loss Function	Formula	Description
Multi-layer Perceptron (MLP), Logistic Regression, Gradient Boosting	Cross-Entropy Loss	$-\frac{1}{N}\sum_{i=1}^{N}[\,y_i\log(p_i)\; +(1-y_i)\log(1-p_i)\,]$	Measures the dissimilarity between predicted probabilities and true binary labels. Penalizes the model more for confidently incorrect predictions.
Random Forest, Decision Tree	Gini Impurity	$1-\sum_{k=1}^{K}p_k^2$	Measures the impurity or disorder of a set of labels. The Gini impurity is minimized by splitting the data based on the feature that best separates the classes.

Table 7. Loss Functions used for machine learning models. y_i represents the true binary label of the i-th instance. In the context of binary classification, y_i is either 0 or 1, indicating the true class of the instance. p_i represents the predicted probability of the positive class (class 1) for the i-th instance. In logistic regression and similar models, p_i is the model's estimate of the probability that the i-th instance belongs to the positive class. k: is often used as an index variable in summations. In the context of the Gini impurity formula, k represents the index of a particular class in the set of classes. N represents the total number of instances in the dataset or the total number of samples.

conditions. HSV (Hue, Saturation, Value): This model represents color itself, along with intensity and brightness. It's handy for detecting changes in color intensity or brightness, which can be important when dealing with features like cool roofs. LAB: This is a uniform color space, meaning that a similar numerical change in color values corresponds to a similar perceived change in color. It's useful for capturing subtle color differences accurately. This is especially important when precise color information is important, such as identifying delicate variations in roof colors.

The findings indicate that only using the RGB color space is sufficient to capture all the details mentioned earlier, and the outcomes are at their best when only RGB colors are used.

4.2.2 Measuring direct reflectivity. The New York City dashboard Flohr *et. al.* [12] is created using a mix of satellite images and geographic data. The work created roof albedo measurements for the year 2020 using Landsat-8 (Landsat OLI Level 2 Science Products images from 2018 and 2016) and ortho imagery for New York State. Landsat images were used to train an algorithm to measure reflectivity using ortho images. First, they used the reflectance of Landsat images to measure the narrow band surface albedo using the method developed by Liang [20] for Landsat 5 images. To use similar wavelengths for Landsat 8 bands they modified Liang's Landsat formula to calculate Landsat shortwave albedo using Smith's (2010) [20] normalization method below:

$$\text{albedo} = \frac{0.356 \cdot b1 + 0.130 \cdot b3 + 0.373 \cdot b4 + 0.085 \cdot b5 + 0.072 \cdot b7 - 0.0018}{0.365 + 0.130 + 0.373 + 0.085 + 0.072} \tag{1}$$

where $b1$ is band 1, $b3$ is band 3, $b4$ is band 4, $b5$ is band 5, and $b7$ is band 7.

As we can observe, this uses 8 band information from the satellite Landsat including bands that are not present on planetscope images. However, landsat has a poor resolution of 30m per pixel. Therefore, it cannot be used directly for roof reflectivity estimation. We therefore learn a regression model with output as albedo and input and the bands from planetscope satellite. We extract albedo values from Landsat images specifically at the center pixels of buildings [12] where there is complete overlap between Planet and Landsat images. Subsequently, using the center pixel information from Planet images, we predict the albedo values obtained from Landsat images. For each building shape, we extracted

Landsat images in TIFF format. Even if the building shape covers only a small fraction of a pixel in the Landsat image, the intersection operation identifies all the pixels it overlaps with as shown in Figure 7a. Since raster data is pixel-based, the result includes entire pixels intersected by the building shape, even if only a part of the pixel is covered. To ensure compatibility, we reprojected the Landsat image to match the resolution of the Planet image. Then, we created a bounding box that encompasses both images and calculate the percentage of overlapping area within this common extent. We only considered buildings with a 100% overlap. We also checked additional area percentages. If both "Y beyond X" and "X beyond Y" were 0, we kept these data points as Figure 7b. We removed data points where all band values needed for calculating albedo are zero or NaN. With data from a single day, we captured around 200 buildings based on these constraints, focusing specifically on Hyderabad. In total, 5724 buildings are extracted using multiple days. Flohr *et. al.* considered a roof reflectivity less than 0.6, a dark roof and otherwise, a cool roof. We verify the recall using this criterion on our dataset. Results are described in Section 5.2.

5 CLASSIFICATION RESULTS

5.1 Choice of imagery for classification

As described in Section 3.2, there are a number of satellite options when classifying roof reflectance depending on the number of bands and the frequency/cost of data collection. Since Maxar is impractical from a cost point of view, we explore sentinel and planet based classifications. The Sentinel satellite imagery provides details at a resolution of 20 meters, capturing the Earth's surface with moderate spatial precision.

The evaluation of the cool roof classification models trained on Sentinel-20m data is presented in Table 9. The outcomes, as shown in the table, do not exhibit highly encouraging results. The model performance metrics, including accuracy for random forest are reported for various scenarios such as considering all bands, individual bands (Red, Green, Blue, NIR), and more. These results provide insights into the challenges and limitations associated with using Sentinel-20m imagery for cool roof classification. Table 8 shows the classifier performance with 8-band planet images as well as 4-band planet basemap images. One can observe that the 8-band images consistently provide better classifier performance. We conducted experiments using Basemaps for the same dates and locations as our Planet-8 bands dataset.

Choice of color space. Across all classifiers (Table 8), RGB and LAB performed consistently well as compared to HSV or a combination of the three. It is also observable that the rest of the 5 bands are also important for the classifier since taking only the RGB/LAB colorspaces doesn't result in high accuracy classifiers.

Dimensionality reduction. To reduce any overfitting that might be happening due to large number of features, we ran PCA on the features reducing the space to 30 dimensions. the right side table represents the results after PCA.

Sanity Test. We tested the top 2 contenders - RGB* and LAB with a sanity test where we ran the model over the entire city of Hyderabad and reported % of cool buildings. We found that LAB resulted in as high as 25% cool buildings while RGB* resulted in 20% cool buildings. This suggested that the model may have overfit. We therefore chose RGB* as the color space for our classification. The MLP model with 30 features from PCA dimension reduction works best with a recall of approximately 0.8944 means that the model is able to correctly identify about 89.44% of the actual positive instances and a precision of approximately 0.8571 means that when the model predicts a positive outcome, it is correct about 85.71% of the time.

BASEMAPS	DT	RF	LR	MLP	GB
RGB	82.35	78.99	69.74	76.05	77.73
HSV	**83.19**	78.99	71	78.57	79.83
LAB	80.67	80.25	74.36	84.4	79.9
RGB-HSV-LAB	76.05	80.67	78.15	81.51	77.77

SR	DT	RF	LR	MLP	GB
RGB	83.29	87.17	83.23	87.74	83.19
HSV	80.72	81.67	78.72	81.97	81.96
LAB	83.15	84.76	80.85	**89.08**	83.21
RGB-HSV-LAB	78.78	83.6	82.73	84.35	81.98
RGB (3 bands)	72.99	77.58	71.06	73.26	72.02
HSV (3 bands)	58.4	60.21	58.79	55.74	51.13
LAB (3 bands)	55.21	60.52	61.57	65.94	59.93

SR	DT	RF	GB	Log Reg	MLP
RGB*	82.77	80.25	76.05	70	84.03
HSV*	78.99	81.09	76	71.84	83.61
LAB*	79.41	85.71	77.31	73.94	83.61
RGB-HSV-LAB*	82.35	82.77	79.41	76	**86.55**

SR	DT	RF	LR	MLP	GB
RGB*	76.54	86.02	82.38	**87.26**	83.50
HSV*	71.25	73.87	78.41	81.05	75.43
LAB*	70.49	80.12	80.55	86.94	78.56
RGB-HSV-LAB*	75.54	75.75	81.78	81.97	81.3

Table 8. Roof classification F1 scores on test set of "second set" across Planet Basemap and Surface Reflectance (SR) images. * represents that PCA was used before classification. Observe that - 8 band images perform better than basemaps or 3 band images; RGB* and LAB are the top performers; MLP performs best out of all models.

Sentinel-10m	RF
All	67.74
R	64.71
G	61.29
B	64.51
NIR	63.53

Table 9. Roof classification F1 scores on the test set of "second set" using the Sentinel2 data. Compared to Table 8, we note that planet based classification has much better accuracy.

SR	Train Region	No of Samples	MLP
RGB	Hyd	1110	94.85
RGB	Tamil Nadu	1086	94.99
RGB	Hyd + aug	1510	**97.37**
RGB	Tamil Nadu + aug	1486	95.12
RGB	all + aug	3054	92.51

Table 10. This table shows the effect of adding augmentations (gaussian noise, smoothing, scaling and noise) to the dataset containing Planetscope 8-band surface reflectance (SR) images. We observe that the augmented dataset with hyderabad region images provide the best accuracy. The numbers are reported on a test region containing roof samples from all cities (non-overlapping with train region).

Adding augmentations. In our effort to enhance the effectiveness of our model for cool roof classification, we applied several data augmentation techniques to our relatively small dataset. These augmentations aim to artificially expand the dataset and introduce variability to help the model generalize better. We implemented four types of augmentations: mixing, gaussian smoothing, adding noise, and scaling. For cropping, we randomly selected some indices in two feature vectors and swapped their values, effectively altering a third of the feature values. This helps the model learn from different perspectives and focuses on specific details, preventing it from relying too heavily on the exact position of features. Gaussian smoothing involved applying a Gaussian filter with a sigma value of 0.4 to the 132-feature vector. It helps reduce noise and fine details in a vector, making the model more robust to minor variations and ensuring it focuses on essential features. Additionally, we introduced random uniform noise to the feature vector, with a noise level set to 0.1. By simulating real-world imperfections, the model becomes more resilient to noise and small fluctuations, improving its performance on less pristine data. Lastly, we scaled random features in the vector within a specified

Name	Values	Name	Values
Satellite	Planet	Bands	8
Features	132	Color Space	RGB
PCA	30	Model	MLP
Iterations	1000	Layers	2
Activation	ReLu	Optimizer	Adam
Learning rate	0.001	L2 regularization	0.0001
Neurons	[64, 8]	Scaling	Scaling factor range (0.1, 0.4)
Cropping	30	Gaussian Smoothing	Sigma = 0.4
Adding Noise	Noise level = 0.1		

Table 11. Information about Model and Data Augmentations for Cool Roof Classification

(a) Partial overlap between planet and landsat images
(b) Full overlap between planet and landsat images

Fig. 7. Landsat image extraction - (a) The intersection operation identifies all pixels that overlap with the building shape, even if it covers only a small fraction of a Landsat pixel. (b) We match the resolution of Landsat with Planet images and buildings with 100% overlap are considered, we keep data points where both "Planet beyond Landsat" and "Landsat beyond Planet" are 0

scaling factor range of (0.1, 0.4). This variation helps the model adapt to different scales and understand features at various sizes, promoting better generalization to diverse inputs. These augmentations proved beneficial for our cool roof classification task, as they introduced diverse variations in the data, helping the model to better capture and understand patterns, ultimately improving its overall performance.

With augmentations on Hyderabad data we achieved 97.37 F1-score with a precision of approximately 96.3% means that out of all the instances the model predicted as positive, about 96.3% were correct, and only 3.7% were false alarms (incorrectly predicted positives) and a recall of approximately 94.6% means that the model identified about 94.6% of all the actual positive instances, and 5.4% of positive instances were missed (false negatives). The outcomes of training our model using both Noida data and the entire dataset, along with augmentations, are presented in Table 10. The model on Hyderabad data along with augmentation performs well possibly due to a larger number of samples from the city. The joint model doesn't work as well possibly due to the difference in the distribution of samples from the two cities and hence, a difficulty in model convergence.

From our experiments, we found that using Planet's 8-band RGB images, applying PCA (Principal Component Analysis), and incorporating augmentations yielded the best results for classifying cool roofs. Moreover, this approach demonstrated good generalizability when applied to other cities. Also the parameters for augmentations along with best model details are summarized in Table 11.

5.2 Reflectivity based classification

Flohr *et. al.*[12] suggested the use of albedo as a measure of reflectivity (Section 4.2.2). To validate this use, we find the classification accuracy based on the reflectivity using a threshold of 0.5. We focused on buildings with more than 4 pixels (a total of 2889 samples) and calculated albedo by taking the median of raw albedo values from 5 pixels. Additionally, we used Planet pixels as input and a linear regression model to predict albedo for the 5 pixels. However, we couldn't distinguish between cool and non-cool values with high accuracy (Table 12). This lack of distinction may be due to our imagery having a 3m resolution. It's possible that we need higher resolution data or additional pre-processing to improve the results. Therefore, direct reflectivity based classification is not a feasible approach for our use case.

Model	Bands	RMSE	F1 score
LR	Basemap (4)	0.14	69.45
RF	Basemap (4)	0.13	68.75
MLP	Basemap (4)	0.145	60.63
LR	Planet SR (8)	0.138	50.36
RF	Planet SR (8)	0.133	59.37
MLP	Planet SR (8)	0.14	51.99

Table 12. Using the derived albedo reflectivity, we classify roofs into cool/noncool using a threshold of 0.5 on our "second set". Note that the best model on Planetscope SR (MLP) had an F1-score of 97% after augmentation and 89% before augmentation. Models: Linear regression (LR), Random forest (RF), Multilayer perceptron (MLP). This suggests that using albedo for classification into cool/non-cool is not a reliable approach.

Place	cool %	non-cool %	cool	non-cool	total	Date	cool area %	BCQ
Hyderabad	4.58	95.41	44936	930879	1002825	10-8-23	7.18	6
Chennai	4.07	95.92	14592	343933	358525	14-12-23	3.89	8
Chandigarh	0.53	99.46	889	166942	167831	14-12-23	63.22	1
Kolkata	0.99	99.01	1589	306153	307742	9-12-23	8.4	5
Ahmedabad	7.21	92.78	155507	2001320	2156827	13-12-23	20.9	4
Mumbai	1.49	98.5	12237	809063	821300	03-12-23	4.41	7
Delhi	0.66	99.33	201	29900	30101	17-12-23	48.87	2
Bangalore	23.58	76.41	893744	2896523	3790267	20-11-23	37.24	3

Table 13. Planet 8 bands - city prediction. The table displays the number of cool and non-cool buildings, along with their percentages, the total number of buildings in each city, and the date when the imagery was captured. The prediction are made on model trained with Planet-8Bands-Hyderabad-RGB-PCA-Augmentations. Building Coolness Quotient (BCQ) ranking is based on the prevalence of cool roofs among buildings.

5.3 Scaling it to multiple cities

To capture the performance of cities with respect to cool roof deployment we introduce **Building Coolness Quotient (BCQ)** which estimates the coolness of a region via the proportion of buildings which are cool. We estimated the BCQ of 8 cities using our best model (MLP on 8 band planet imagery with 132 features in RGB color space and reducing dimensionality to 30 using PCA).

$$BCQ = \frac{\text{total number of buildings with cool roofs}}{\text{total number of buildings}} \tag{2}$$

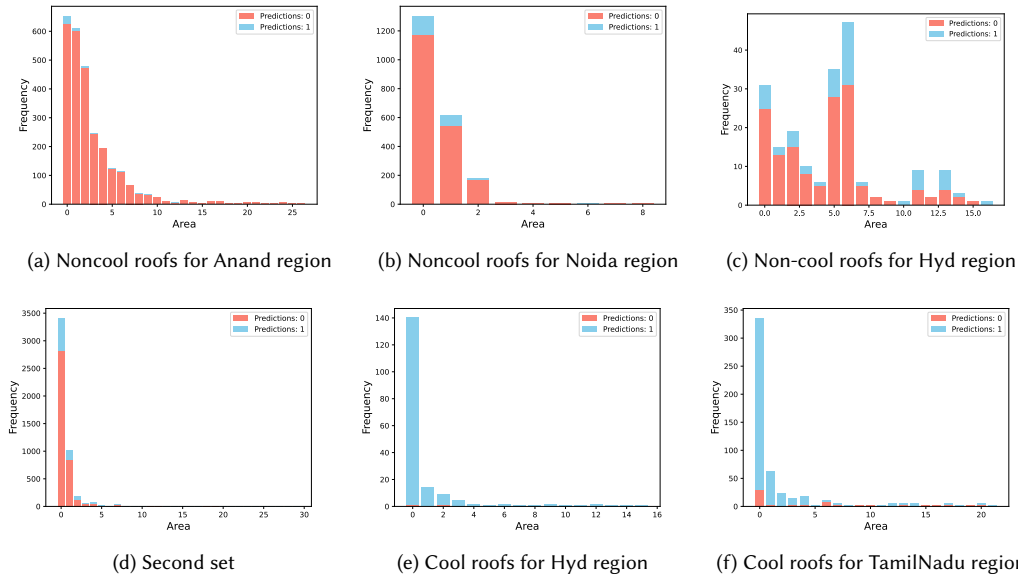|		
(a) Noncool roofs for Anand region	(b) Noncool roofs for Noida region	(c) Non-cool roofs for Hyd region
(d) Second set	(e) Cool roofs for Hyd region	(f) Cool roofs for TamilNadu region

Fig. 8. Classification results by area. We've considered a total of 50 bins in our analysis. Orange denotes predictions indicating cool areas, while blue signifies predictions indicating non-cool areas. The dataset shows a bias towards smaller areas, encompassing both cool and non-cool regions across various cities and within individual cities. Hyd represents Hyderabad

This ratio tells us the prevalence of cool roofs among all the buildings in the city. Table 13 shows cool and non-cool roof prediction for 8 different cities, with their cool/non-cool percentages, total buildings and date of the imagery captured. Note that for most cities we observe small percentages of cool buildings. In some cities, the cool area percentage is quite high indicating that the buildings with larger areas are cool. This might be a feature of larger corporates opting to use cool roofs. We plan to release building-wise coolness in a public dashboard.

5.4 Generalization to different building sizes

The Figure 8 displays classification outcomes for different bins on the area of buildings in meters. Our assessment utilizes 50 bins on building areas. We observe that our dataset has a bias towards smaller regions. Figure 8(d). This trend is observable in both cool and non-cool areas within and across cities. In this framework, blue denotes predictions for cool areas, while orange represents predictions for non-cool areas. Despite the dataset's bias towards smaller areas, our model consistently showcases robust generalization capabilities across all categories.

6 FEATURE ANALYSIS - PCA

To justify and understand the hyperparameters used in the classifier, we further investigate the choices through a number of experiments.
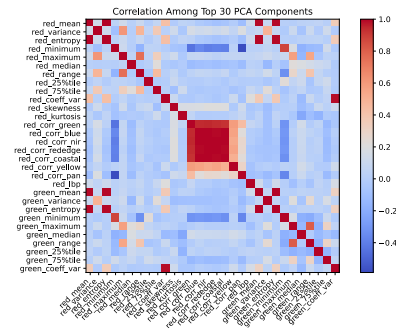
(1) Cumulative Explained Variance: This experiment is used to find the right number of components to be used for PCA. The cumulative explained variance plot in Figure 9(a) provides insights into how much of the total variance in our original dataset is retained as we include more principal components. The x-axis represents the principal components. In this case, it ranges from 1 to 30, indicating the 30 principal components obtained from
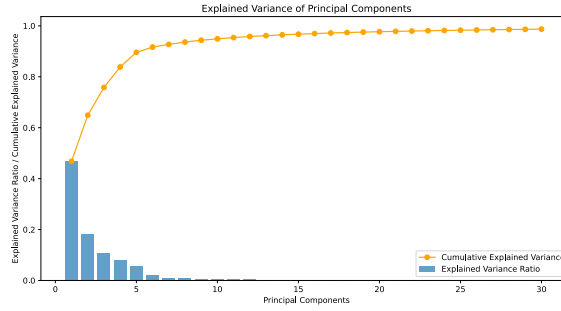
(a) Explained variance of principal components. Therefore we chose dimensionality as 30
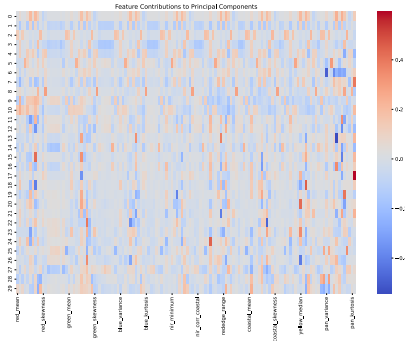
(b) Feature Influence on Principal Components

(c) Correlation among top 30 PCA components

Fig. 9. a) The plot explores how much variance each principal component captures individually and how the cumulative explained variance increases as we include more components. (b) Feature Influence on Principal Components Features with values close to zero have minimal influence on the respective component. (c) Correlation among top 30 PCA components

the PCA. The left y-axis represents the explained variance ratio for each individual principal component. Each bar in the bar plot corresponds to the explained variance ratio of a specific principal component. As we move along the x-axis (increasing the number of principal components), the cumulative explained variance increases. The point where the curve starts to flatten or reaches an elbow is crucial. It indicates the number of principal components needed to capture the majority of the dataset's variability. The goal is to strike a balance between reducing dimensionality and retaining enough information to represent the original data adequately. Fewer principal components mean a more compact representation of the data, but it comes at the cost of losing some information. We experimented with more than 30 components but best results were achieved at 30 components.

(2) Feature Influence on Principal Components: In this heatmap Figure 9(b) - rows correspond to principal components, columns correspond to original features and color intensity represents the magnitude and direction of the feature contributions. Positive values indicate a positive correlation between the original feature and the principal component. Negative values indicate a negative correlation. The features represent sparse loading means that only a few features have significant contributions to a principal component, suggesting that the

component captures specific patterns not the general patters. The heatmaps shows the discrimination is not extremely positively or negatively correlated and hence the classification is difficult.

(3) Dominant Features: We got repeated names in the list of dominant features (Table 14) for different principal components, it suggests that those features are influential across multiple components. Features such as "pan 25 percentile," "red variance," "yellow median," etc., are contributing to the patterns captured by multiple principal components. These features exhibit a multifaceted influence on the variability in our data, affecting different aspects captured by different principal components. Almost all the bands are present in domainant features implying why planet-8 bands work better than Basemaps with 4 bands.

(4) Top-5 features: Table 15 displays the top 5 features for top 3 principal components. Loading represents the weight or coefficient assigned to each feature in a principal component. It indicates the strength and direction of the relationship between the original features and the principal component. Positive loadings suggest a positive relationship, while negative loadings suggest a negative relationship. Contribution percentage represents the proportion of the total variance in the dataset that is explained by a particular feature in the principal component. Higher contribution percentages indicate that the corresponding feature has a more significant impact on the variability captured by the principal component. The sum of contribution percentages across all features in a principal component adds up to 100%. The loading vector is used to transform the original features into the new set of principal components.

The contribution of each feature to a principal component is proportional to the square of its loading. Therefore, features with higher loadings contribute more to the overall variability captured by the principal component. PC1 is mainly formed by positive contributions from features related to the 25th percentile of pan values, median pan values, and the 25th percentile of yellow values. These features contribute positively to the variability captured by PC1. PC2 is primarily formed by negative contributions from features related to the variance of rededge values, the range of red values, and the variance of coastal values. These features contribute negatively to the variability captured by PC2, suggesting an inverse relationship. PC3 is predominantly formed by positive contributions from features related to the coefficient of variation in pan, yellow, and rededge values. These features contribute positively to the variability captured by PC3. Our feature extraction has presence of both positive and negative loadings in principal components enhances the ability of the model to discern meaningful patterns in the dataset, making it more effective for classification tasks. It allows the model to consider a broader range of relationships between features, leading to a more comprehensive understanding of the underlying data structure.

(5) Correlation among top 30 PCA components: Each principal component is a linear combination of all the original features (in our case, 132 features). The coefficients of this linear combination are determined during the PCA process. When we select the top 30 principal components and calculate their correlation matrix (Figure 9(c)), we are examining how these components are related to each other. High Positive Correlations: If we observe clusters of warmer (reddish) colors, it suggests that the corresponding components are positively correlated. This indicates that those components tend to move together, capturing similar information. In other words, these components are redundant in terms of the information they represent. High Negative Correlations: If we observe clusters of cooler (bluish) colors, it suggests negative correlations between components. Negative correlations imply that the components capture complementary or opposite information. In this case, each component might represent different aspects of the data, and they are not redundant. Low Correlations: If we observe mostly neutral (close to white) or faintly colored cells, it indicates low or no correlation between the

components. This suggests that the components are relatively independent of each other, meaning they capture distinct and non-redundant information.

PC	Dominant Feature	PC	Dominant Feature	PC	Dominant Feature
1	pan_25%tile	11	red_variance	21	yellow_median
2	rededge_variance	12	pan_25%tile	22	rededge_median
3	pan_coeff_var	13	pan_25%tile	23	coastal_25%tile
4	red_75%tile	14	pan_median	24	yellow_median
5	pan_75%tile	15	pan_25%tile	25	rededge_variance
6	pan_minimum	16	red_25%tile	26	pan_coeff_var
7	pan_variance	17	pan_median	27	yellow_median
8	pan_lbp	18	pan_lbp	28	coastal_25%tile
9	red_kurtosis	19	yellow_25%tile	29	pan_entropy
10	red_variance	20	pan_75%tile	30	pan_entropy

Table 14. Dominant Features in Each Principal Component by Magnitude. Almost all the bands are present in domainant features implying why planet-8 bands work better than Basemaps with 4 bands.

Principal Component	Feature	Loading	Contribution (%)
1	pan_25%tile	0.2416	3.03
	pan_median	0.2395	3.00
	yellow_25%tile	0.2107	2.64
	yellow_median	0.2096	2.63
	rededge_25%tile	0.2058	2.58
2	rededge_variance	-0.1522	1.62
	red_range	-0.1477	1.57
	coastal_variance	-0.1463	1.56
	rededge_range	-0.1448	1.54
	pan_median	0.1434	1.53
3	pan_coeff_var	0.2237	3.16
	yellow_coeff_var	0.2108	2.97
	rededge_coeff_var	0.2102	2.96
	coastal_coeff_var	0.2082	2.94
	nir_coeff_var	0.2002	2.82

Table 15. Top 5 Features for top 3 Principal Components. Columns starting from left: Component number, the name of the feature, the weight assigned to the feature in the principal component and the percentage of the total variance explained by the given feature in the principal component.

7 DASHBOARD DESIGN

We designed our dashboard using open source tools in Python. We used the mapping library leafmap Wu *et. al.* [35] in the solara framework [5] to enable interaction. For rendering all polygons of a city together, we used deckGL [3] along with Duckdb [6] for reading parquet files. Further, we will be integrating this with a PostGIS database [25] for fetching polygons on demand while zooming in. We're showcasing the outcomes for Hyderabad on the dashboard. Figure 1 and Figure 10 show dashboard images. Figure 10a is an example region from Hyderabad, the upper part is non-cool while the lower cluster is of cool roofs. Our model can accurately tell the difference even when it's challenging for the naked

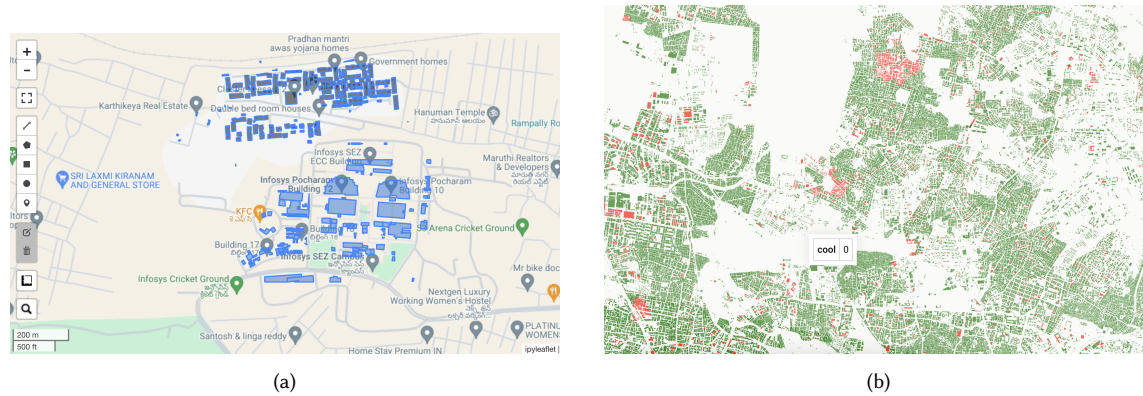(a)   (b)

Fig. 10. Cool roof tracker - Dashboard. (a) region in Hyderabad, the upper part is non-cool while the lower cluster is of cool roofs. (b) Predictions for the entire city of Hyderabad. The red clusters represent cool areas, while the green clusters represent non-cool areas.

eye, and it works well with small buildings. In Figure 10b the red clusters represent cool areas, while the green clusters represent non-cool areas. This figure predicts the entire city of Hyderabad.

8 CONCLUSION AND FUTURE WORK

We presented a scalable and accurate tracker for cool roofs which can work purely through satellite images. This tool combines satellite images and machine learning models. We evaluate eight major cities with our system, with performance outlined in our paper. Our approach involves analyzing satellite data such as PlanetScope images, Planet Basemaps and Sentinel images, extracting features, and employing various machine learning models for classification. Our experiments suggest that PlanetScope imagery is more effective than Sentinel and Planet Basemaps for classification. Our most successful model utilizes MLP with RGB color space, extracting 132 features and applying PCA to reduce dimensionality to 30. Our work also involves a comprehensive analysis of the extracted features, along with providing the rationale behind why we chose to extract each particular feature. We compared our method to the Albedo extraction method, where we calculated albedo from Landsat8 images and trained regression between Planet images and extracted albedo values. Our results indicate that our model outperforms this method. We also qualitatively evaluated different methods for extracting buildings from images, finding that OpenBuildings performs the best, despite some offset issues that need addressing in future research. Although our model is tailored for Indian cities, it would be interesting to assess its performance on roofs in other countries. Our dataset mainly comprises smaller roofs, yet our method shows adaptability with the existing data. Future efforts will involve acquiring additional data encompassing various roof sizes and conducting quantitative analysis of our model's performance.

We've developed a prototype dashboard to support data-driven decision-making at various levels: city, sub-city (pin code), and individual building. This dashboard serves as a valuable tool for policymakers, offering insights into urban heat clusters and the extent of cool areas. By visualizing data on urban heat clusters and cool areas, policymakers can make informed decisions to address heat mitigation strategies effectively. Our future works also include to enable public feedback on the dashboard and use aggregated data to further improve models since the current number of ground truth samples were limited. Additionally, incorporating unlabeled data to enhance our model's reliability and generality, ensuring wider applicability. To deepen our understanding, we plan to gather more domain knowledge, investigating if

seasonal changes impact roof reflectivity and incorporating these findings into our system. Temporal analysis will be an important aspect of our future work, providing insights into how roof conditions evolve over time. Currently the data has less than 3% of cloud coverage, we aim to assess our model's performance in different cloud coverage and other weather conditions.

In terms of effectiveness, future work can be, investigating reflectivity on different cool roof materials and how to incorporate it in model. Also to see how cool roofs connect with things like the Urban Heat Island (UHI) effect and how much people pay for energy. Geospatial analysis can be utilized to understand the distribution of cool and non-cool roofs on a map, potentially revealing geographical patterns.

9 ACKNOWLEDGMENTS

REFERENCES

[1] 2005. Google Maps. https://www.google.com/maps. Accessed: January 5, 2024.

[2] 2005. Maxar Satellite Imagery. https://www.maxar.com/products/satellite-imagery. Accessed: April 13, 2024.

[3] 2017. deck.gl Documentation. https://deck.gl/docs. Accessed January 3, 2024.

[4] 2019. DrivenData - Disaster Response - Roof Type. https://www.drivendata.org/competitions/58/disaster-response-roof-type/page/143. Accessed January 3, 2024.

[5] 2019. Solara.dev. https://solara.dev. Accessed: Januaray 3, 2024.

[6] 2020. DuckDB Documentation. https://duckdb.org/docs/archive/0.9.2/. Accessed January 3, 2024.

[7] 2023. Extreme humid heat in South Asia in April 2023, largely driven by climate change, detrimental to vulnerable and disadvantaged communities. https://www.worldweatherattribution.org/extreme-humid-heat-in-south-asia-in-april-2023-largely-driven-by-climate-change-detrimental-to-vulnerable-and-disadvantaged-communities/. Accessed January 3, 2024.

[8] Hashem Akbari, Linda Gartland, and Steven Konopacki. 1998. Measured energy savings from light colored roofs: Results from three California demonstration sites. Lawrence Berkeley National Laboratory, Environmental Energy Technologies Division. https://www.osti.gov/biblio/674928 Accessed January 1, 2024.

[9] Noelia Liliana Alchapar, María Florencia Colli, and EN Correa. 2020. Albedo quantification using remote sensing techniques. Cool roof in the metropolitan area of Mendoza-Argentina. In IOP Conference Series: Earth and Environmental Science, Vol. 503. IOP Publishing, 012035. https://etd.auburn.edu/bitstream/handle/10415/5984/A%20Multiple%20Regression%20Model%20of%20Cool%20Roofs_Thesis_Final.pdf?sequence=2

[10] ArcGIS. 2022. Title of the Story Map. https://storymaps.arcgis.com/stories/0cdc24592f85480ebaa094037b47a767. Accessed January 2, 2024.

[11] Matthias Drusch, Umberto Del Bello, Sébastien Carlier, Olivier Colin, Veronica Fernandez, Ferran Gascon, Bianca Hoersch, Claudia Isola, Paolo Laberinti, Philippe Martimort, et al. 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote sensing of Environment 120 (2012), 25–36. https://www.sciencedirect.com/science/article/abs/pii/S0034425712000636 Accessed January 3, 2024.

[12] Travis Flohr, Mehdi Heris, Rosy George, and Andrea Avila. 2023. Investigating New York City's Cool Roof Program Implementation Using Remote Sensing Through an Environmental Justice Lens. Journal of Digital Landscape Architecture 2023, 8 (2023), 65–74. https://gispoint.de/fileadmin/user_upload/paper_gis_open/DLA_2023/537740008.pdf

[13] Amy E Frazier and Benjamin L Hemingway. 2021. A technical review of planet smallsat data: Practical considerations for processing and using planetscope imagery. Remote Sensing 13, 19 (2021), 3930. https://www.mdpi.com/2072-4292/13/19/3930

[14] Seth Alexander Greer. 2017. Are Cool Roofs Really Cool? A Predictive Regression Model for Buildings on Auburn University's Campus. Ph. D. Dissertation. Auburn University.

[15] Vandana K. 2023. The white roofs cooling women's homes in Indian slums. https://www.bbc.com/future/article/20230628-the-white-roofs-cooling-womens-homes-in-indian-slums. Accessed January 1, 2024.

[16] Bahareh Kalantar, S Mansor, Z Khuzaimah, M Ibrahim Sameen, and B Pradhan. 2017. Modelling mean albedo of individual roofs in complex urban areas using satellite images and airborne laser scanning point clouds. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-2/W7 (2017), 237–240. https://isprs-archives.copernicus.org/articles/XLII-2-W7/237/2017/isprs-archives-XLII-2-W7-237-2017.pdf

[17] Jonguk Kim, Hyansu Bae, Hyunwoo Kang, and Suk Gyu Lee. 2021. CNN algorithm for roof detection and material classification in satellite images. Electronics 10, 13 (2021), 1592.

[18] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. 2023. Segment anything. arXiv preprint arXiv:2304.02643 (2023). https://arxiv.org/abs/2304.02643

[19] VH Kondekar and SKA Bodhe. 2018. Comprehensive Investigation of Color Models used in Image Processing. *International Journal of Computer Applications* 180, 22 (2018), 19–24. https://ijcaonline.org/archives/volume180/number22/vipul-2018-ijca-916507.pdf Accessed January 3, 2024.

[20] Shunlin Liang. 2001. Narrowband to broadband conversions of land surface albedo I: Algorithms. *Remote sensing of environment* 76, 2 (2001), 213–238. https://www.sciencedirect.com/science/article/abs/pii/S0034425700002054 Accessed January 3, 2024.

[21] Microsoft. 2023. GlobalMLBuildingFootprints. https://github.com/microsoft/GlobalMLBuildingFootprints. Accessed December 1, 2023.

[22] Natural Resources Defense Council. 2022. *Issue Brief - Cool Roofs: Protecting Local Communities and Saving Energy.* https://www.nrdc.org/sites/default/files/ib_-_cool_roofs_-_hyd_workshop.pdf Accessed January 3, 2024.

[23] Katarzyna Osińska-Skotak and Wojciech Ostrowski. 2015. Use of satellite and ALS data for classification of roofing materials on the example of asbestos roof tile identification. *Technical Sciences* 18(4) (01 2015). Accessed December 1, 2023.

[24] Danny S Parker, Stephen Barkaszi, Subrato Chandra, and David J Beal. 1995. Measured cooling energy savings from reflective roofing systems in Florida: field and laboratory research results. *Proceeding of the Thermal Performance of the Exterior Envelopes of Buildings* (1995), 105–115. http://www.fsec.ucf.edu/en/publications/html/fsec-pf-293-95/index.htm Accessed December 3, 2023.

[25] PostGIS. 2001. PostGIS Documentation. https://postgis.net/documentation/. Accessed January 3, 2024.

[26] QGIS Development Team. 2024. *QGIS Geographic Information System.* QGIS Association. https://www.qgis.org Accessed January 3, 2024.

[27] Wojciech Sirko, Sergii Kashubin, Marvin Ritter, Abigail Annkah, Yasser Salah Eddine Bouchareb, Yann Dauphin, Daniel Keysers, Maxim Neumann, Moustapha Cisse, and John Quinn. 2021. Continental-scale building detection from high resolution satellite imagery. *arXiv preprint arXiv:2107.12283* (2021). https://arxiv.org/abs/2107.12283

[28] Ian A Smith, Katharine Lusk, and Lucy R Hutyra. 2022. On the use of cool roofs to reduce residential heat exposure disparities in Boston, MA. (2022). https://www.mdpi.com/2079-9292/10/13/1592 Accessed January 1, 2024.

[29] Roman A. Solovyev. 2020. Roof material classification from aerial imagery. *CoRR* abs/2004.11482 (2020).

[30] Isabelle Tingzon, Nuala Cowan, and Pierre Chrzanowski. 2023. Fusing VHR Post-disaster Aerial Imagery and LiDAR Data for Roof Classification in the Caribbean. 3742–3749. https://doi.org/10.1109/ICCVW60793.2023.00402

[31] Maurizio Tommasini, Alessandro Bacciottini, and Monica Gherardelli. 2019. A QGIS tool for automatically identifying asbestos roofing. *ISPRS International Journal of Geo-Information* 8, 3 (2019), 131.

[32] University of New South Wales. 2022. Cool Roofs Cost Benefit Analysis. https://www.unsw.edu.au/content/dam/pdfs/unsw-adobe-websites/arts-design-architecture/built-environment/our-research/high-performance-architecture-research-cluster/2022-04-Volume-1.pdf. Accessed January 2, 2024.

[33] Darrel L Williams, Samuel Goward, and Terry Arvidson. 2006. Landsat. *Photogrammetric Engineering & Remote Sensing* 72, 10 (2006), 1171–1178.

[34] Abraham Noah Wu and Filip Biljecki. 2021. Roofpedia: Automatic mapping of green and solar roofs for an open roofscape registry and evaluation of urban sustainability. *Landscape and Urban Planning* 214 (2021), 104167. https://doi.org/10.1016/j.landurbplan.2021.104167

[35] Qiusheng Wu. 2021. Leafmap: A Python package for interactive mapping and geospatial analysis with minimal coding in a Jupyter environment. *Journal of Open Source Software* 6, 63 (2021), 3414. https://doi.org/10.21105/joss.03414

[36] Q. Wu and L. Osco. 2023. samgeo: A Python package for segmenting geospatial data with the Segment Anything Model (SAM). *Journal of Open Source Software* 8, 89 (2023), 5663. https://doi.org/10.21105/joss.05663

[37] Tengfang Xu, Jayant Sathaye, Hashem Akbari, Vishal Garg, and Surekha Tetali. 2012. Quantifying the direct benefits of cool roofs in an urban setting: Reduced cooling energy use and lowered greenhouse gas emissions. *Building and Environment* 48 (2012), 1–6. https://doi.org/10.1016/j.buildenv.2011.08.011

[38] Stefano Zorzi, Shabab Bazrafkan, Stefan Habenschuss, and Friedrich Fraundorfer. 2022. Polyworld: Polygonal building extraction with graph neural networks in satellite images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* 1848–1857. https://arxiv.org/abs/2111.15491