Citizen science to assess light pollution with mobile phones

Gorka Muñoz-Gil,1, 2, * Alexandre Dauphin,1, * Federica A. Beduini,1, † and Alejandro Sánchez de Miguel3, 4, 5

1ICFO – Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain
2Institute for Theoretical Physics, University of Innsbruck, Technikerstr. 21a, A-6020 Innsbruck, Austria
3Depto. Física de la Tierra y Astrofísica, Instituto de Física de Partículas y del Cosmos (IPARCOS), Universidad Complutense, Madrid, Spain
4Environment and Sustainability Institute, University of Exeter, Penryn, Cornwall TR10 9FE, U.K.
5Instituto de Astrofísica de Andalucía, Glorieta de la Astronomía, s/n, 18008, Granada, Spain

The analysis of the colour of artificial lights at night has impact on diverse fields, but current data sources have either limited resolution or scarce availability of images for a specific region. In this work, we propose crowdsourced photos of streetlights as an alternative data source: for this, we designed NightUp Castelldefels, a pilot for a citizen-science experiment aimed at collecting data about the colour of streetlights. In particular, we extract the colour from the collected images and compare it to an official database, showing that it is possible to classify streetlights according to their colour from photos taken by untrained citizens with their own smartphones. We also compare our findings to the results obtained from one of the current sources for this kind of studies. The comparison highlights how the two approaches give complementary information about artificial lights at night in the area. This work opens a new avenue in the study of the colour of artificial lights at night with the possibility of accurate, massive and cheap data collection.

I. INTRODUCTION

Artificial lights at night are an important indicator of human activity and of their impact on the environment. Their study allows for an estimation of many complex global parameters, with applications ranging from economy - allowing for predictions of gross domestic product [1], population density [2] and energy consumption [3; 4] - to environmental sciences - helping in the estimation of carbon dioxide emissions [5; 6], sky brightness [7; 8], and landscape connectivity [9]. Most of the above-mentioned studies focus on light intensity, with data collected from panchromatic remote sensing instruments placed on satellites, such as the DMSP-OLS [3] and the VIIRS-DNB from the Suomi North Polar Partnership (SNPP) [10].

However, important information can also be extracted from the spectra of light sources. Indeed, recent works have demonstrated that different ranges of the visible spectrum impact differently a plethora of biological processes, affecting profoundly the flora [11], the fauna [12; 13; 14] and human communities [15; 16]. In particular, according to various studies on ecological and health indicators [16; 17; 18], it is the blue range of the spectrum that affects most the living beings.

Currently, there is not enough field surveys that relates color ambient light with health impacts. [19] did a field survey to compare the intensity but not the resolution. And [20] showed the limitations of comparing uncalibrated and low resolution ISS images with field data, again without color information. A summary of all health and environmental impacts ca be found on [21].

At the same time, we have poor estimations of the light intensity in this range, due to the blindness to blue light of the instruments on satellites mentioned above. For example, current estimations of the growth of blue light global emission in the 1992-2017 period varies between 49% to 270% depending on the different scenarios. Even, recent publications show how European Union light emission have increased 13% on green light and 25% in blue light [22]. It is thus evident that it is crucial to find other data sources to estimate the blue contribution to light emissions, even with limited accuracy, in order to reduce significantly this error and obtain better estimations of the radiance [23; 24] that would lead to more accurate economical and ecological indicators[25]. It has been demonstrated that obtaining at least an approximation of the colour temperature of the light sources is important[26]: while it does not give as complete information as a spectrum, it is a sufficient parameter to determine the potential environmental impact.

Various systems have been tested to determine at least the colour temperature of the lighting technologies used in different regions of the world. Photo-metric techniques have limitations, either in spatial and spectral resolution,
that make it difficult and expensive to obtain. Images are thus the main data source for this kind of analysis, no matter if they have been taken from airplanes, cars, the International Space Station (ISS) or other smaller artificial satellites [25]. Regarding satellite images, colored data are currently available only from two sources: the photos taken by astronauts from the International Space Station (ISS) with a commercial digital single-lens reflex camera [27; 28] and the ones generated by the commercial satellite JL1-03B [29]. Nonetheless, both sources have serious limitations. For example, JL1-03B images are mainly accessible on request and cover few regions of the Earth. The ISS images coverage of the Earth surface is higher, but their resolution varies widely depending on the location. While low resolution images may be sufficient to represent large areas in some cases, they are not useful in densely populated regions, where outdoor lighting is usually extremely variable: this highlights the importance of collecting data with the highest possible resolution, especially in urban areas. Moreover, as most streetlamps emit light towards the ground, the spectra extracted from the images taken from satellites may be distorted, depending on the characteristics of the reflecting surfaces [30].

It is also possible to infer blue light emissions from streetlight databases, which usually contain information about the location and the lighting technology (which determines the emission spectrum). However, it is difficult to obtain such data, as public lighting databases are usually managed by local authorities and their accessibility may vary depending on the local government. Moreover, an important contribution to light pollution comes from private illumination, for which no database exists. To circumvent these problems, researchers have used satellite images to assess and characterize light in the visible spectrum for specific regions.

In this work, we propose the use of widely available light sensors - smartphone cameras - as a way to characterize the colour of artificial lights at night. This approach is a low-cost solution to complement and enrich satellite data while addressing some of the challenges of the other data sources. With people collecting photos of streetlights with their own smartphone, it is possible to collect information about the colour of artificial lights at night with unprecedented spatial resolution, as well as to reach urban areas that are not covered by currently available data sets. Furthermore, as the photos aim directly at the source of the light, they are not affected by the colour aberrations that affect satellite images due to reflection. To this aim, we have designed a citizen science experiment (NightUp, see Figure 1), in which participants photograph artificial lights with their mobile phones. In this work, we analyse the data collected during the pilot NightUp campaign in the city of Castelldefels (Barcelona, Spain), showing the potential of mobile phones and citizen participation as an important source of data to study artificial lights at night.

II. MATERIALS AND METHODS

A. Data acquisition by citizen science

Even though there are many definitions of citizen science [31], they converge on the fact that it involves crowdsourcing part of a scientific experiment (e.g. data collection, pattern recognition, data analysis, ...) to voluntary citizens outside the academic environment. This practice is very valuable, because it strengthens the relationship between
FIG. 2: Top row: Three examples of light source location by means of our algorithm. Bottom row: The area in the rectangle of the image above is cropped and the circle in the bottom right corner shows the colour assigned to each lamp.

the public and the academic world and transforms citizens from passive recipients of dissemination activities into active elements of the research process. At the same time, citizen science experiments are a very relevant tool for the scientific community, that can lead to solid scientific results that would be difficult without a collective effort due, for example, to the huge amount of data to be collected and/or classified [32; 33; 34; 35].

In the last twenty years, the advent of smartphones has become a game changer in citizen science. Indeed, their collections of sensors (microphones, location sensors, cameras, etc.) allows citizens to have a small laboratory in their pockets [36; 37; 38]. More importantly, scientists now can potentially reach every corner of the world for their data collection campaigns. With this in mind, we have designed NightUp, a citizen science experiment that uses these widespread cameras to collect data missing in the light pollution scientific community: the spatial distribution of colour for artificial light at night.

The success of a citizen science experiment like NightUp depends on two main factors: effective engagement strategies to gather and retain volunteers and a simple but rigorous method to collect the data to ensure that the citizen science practitioners deliver good quality data with minimum training. For the pilot phase, we have focused mainly on the latter factor, developing a mobile phone application (app) with an intuitive user experience and visual tutorials, allowing participants to collect data without special equipment, nor specific training. The app has been designed with a single and simple task for the participants: they only had to point their smartphones at the streetlights and take a photo through the app. The app would then upload it automatically to a server together with the geolocation of the device, its accuracy and a timestamp. Users could also answer optional questions about the light direction and the distance between similar lamps. The app collected the data anonymously, assigning an alphanumeric ID to each device, without gathering any other personal information.

The idea of the NightUp method is that any person with a smartphone around the world should be able to contribute to the experiment with few simple steps. In order to compare results, the devices must be calibrated, but it is unrealistic to assume that each user could calibrate their own device or that we could do it for each model available on the market. We thus assumed that the calibration that smartphone cameras undergo to meet the ISO standards would be sufficient to obtain satisfactory results. Moreover, the app set the white balance of the smartphone camera to a colour temperature of 3000 K, when allowed by the device.

The data presented in this paper comes from the NightUp pilot campaign (NightUp Castelldefels), that was designed to test the hypothesis that no further calibration is needed to distinguish warmer light sources from colder ones. The pilot campaign lasted from November 2019 to March 2020, involving 74 unique users that contributed with 1112 photos. This first campaign was limited to the Castelldefels area, a mid-sized city that shows a variety of natural (beach, hills, plain) and artificial (industrial, commercial and residential areas) landscapes. Additional data from Castelldefels, Barcelona and El Prat de Llobregat areas were added in Fall 2020 and 2021, reaching a total of 1372
B. Image analysis

In this section, we describe the three steps for the analysis of the photos taken with the NightUp application. The first step consists in manually cleaning the dataset from spurious images. The goal of the second step is to extract the colour of a given streetlight. To this end, we develop an algorithm to locate the region where the lamp is in the image. Then, we extract the average pixel colour to identify the colour temperature of the light. Finally, we associate all the extracted colours to their corresponding geolocalizations and plot their spatial distribution on a map of Castelldefels.

1. Manual verification

In citizen science experiments, we often face the existence of incorrect samples, due for example to possible miscommunication between organizers and participants. This often generates invalid data, which potentially lead to inaccurate results. In our case, before starting the analysis process on the data set of the pilot campaign (November 2019-March 2020), we discarded the following types of photos (number and percentage w.r.t. the total pilot dataset in parenthesis): i) photos where the background is not completely dark (e.g. taken at dusk) that can lead to an overestimation of the blue component in our analysis (235 photos, 21.1 %); ii) photos that - despite the instructions given - are not taken pointing directly at a light source (113 photos, 10.2 %); iii) photos which are not streetlights (409 photos, 36.8 %). Note that a photo can fail on various of the previous points.

After the manual verification, we have been left with 698 photos, a 62.8% of the total photos taken by untrained citizens. The photos taken after March 2020, used in Section III.D.3, were taken by users personally trained by the organizers, so that none of their photos have been discarded during the manual validation process.

2. Colour extraction

In order to find the light source in the uploaded photos and extract their colours, we developed an algorithm that works as follows: i) the image is transformed into its gray-scaled version; ii) we apply a Gaussian blur, a common procedure in image analysis to filter noisy images and enhance their segmentation [39]; iii) we set an intensity threshold to highlight only the brightest regions of the photos; iv) we select the largest region arising from the threshold, assuming that it corresponds to the lamp. It is a fair assumption, as we instructed citizens to point directly at the light source when shooting the photo.

We crop the region where the lamp is with a rectangular box large enough to contain both the light source and its aura. In fact, the lamp pixels are often saturated, so the aura - which shares the spectral properties of the lamp - is especially important to extract the colour information. To do this, we exclude all pixels in the cropped box that
are saturated in at least one of the Red (R), Green (G) and Blue (B) channels of the digital image. We average the value in each channel over the remaining pixels, obtaining the average colour of the box, which we assign to each image. Figure 2 shows the results of this procedure for three different photos, where each image has an assigned colour, depicted in the coloured circles.

Human experts have checked manually that our algorithm extracts the correct colour from all the photos (698) that passed the manual verification described in II.B.1, proving our algorithm as a strong and robust method. If we apply the same procedure to the whole dataset (1389 photos), the segmentation was correct in 87.97% of the photos, while the colour was correctly extracted in 79.4%. This percentage could be improved in future campaigns, for example adding to the app automatic quality checks and feedback for the user before the submission of the photos. Moreover, the photos processed here may now serve as a valid training dataset for upcoming works, in which machine-learning-based segmentation algorithms [40] may enhance the accuracy of the analysis of the photos.

For the following analysis, we consider only manually verified photos, which have been segmented correctly. We perform a change of basis from RGB to the Hue (H), Saturation (S) and Value (V) model [41], which is particularly suitable to easily distinguish between blue and red. Figure 3 shows the extracted colours of the lights in a H-S diagram. For our discussion, we focus on the H and S variables, because the value of V (associated to lightness) gives no information relative to the light spectrum. Notably, we see that two main regions arise, one with $0 \leq H \leq 0.2$, corresponding to warm colours, and a second region with Hue values around $0.6$, corresponding to cold ones.

An important assumption for the NightUp project is that smartphones cameras can successfully distinguish between different colours, even without additional calibration. To test the validity of such statement, we choose the three devices that took the most photos (59, 57 and 52, respectively) and plot their H-S diagram (see Figure 4). We observe that each device reproduces the separation of collected data into two main regions, as in Figure 3, even if with less statistics, as expected.

3. Spatial distribution of the colour of artificial light at night

The NightUp app records the location of the smartphone when the photo is taken, so that we can represent the collected data on a map as in Figure 5(a), obtaining information about the spatial distribution of the colour of artificial light at night in Castelldefels. We observe that the data points are not homogeneously distributed: they appear mainly concentrated around the commercial area. This shows the importance of the collaboration with local entities (such as city council, library and schools) and of focused engagement campaigns for the data collection in order to cover wider territories. For instance, a data collection activity was organized in collaboration with the school Col·legi Frangoal (Latitude: 41.288°, Longitude: 1.984°), resulting in almost 30 new devices and more than 580 photos in one night, increasing notably the density of photos in the north east area and along the central part of the northern highway (upper red road), as depicted in Figure 5(a).

Maps like the ones in Figure 5(a) have the potential to help the light pollution scientific community to estimate the impact of the colour of light in living beings and the blue contribution to global light emission.

FIG. 4: Scatter plots of the colour extracted from the photos taken by the three devices with more uploads to the NightUp application.
III. RESULTS

The NightUp dataset is composed of images taken with different devices with no specific calibration: it is therefore important to show that the colours extracted with the method described in Section II.B provide valid information for scientific studies. To this aim, in this section, we compare our results with the streetlight database from the city council of Castelldefels, identifying regions in the H-S space that correspond to different colour categories that can be used for the characterization of unlabelled databases. Moreover, we compare the colours extracted from NightUp data to the ones derived from an image shot from the ISS, which is currently used as one of the main sources for this type of information.

A. Comparing NightUp data to the city council streetlight database

Thanks to the collaboration of Castelldefels city council in the NightUp project, we have access to a list of verified streetlights [see Figure 5(b)], including their spatial coordinates and the lamp type. This information allows us to associate each image with a potential real streetlight and validate the colour information extracted from the photo. It is important to note that such information might not be easy to obtain for many cities or regions: for this reason, an experiment like NightUp, once validated, could be an important source of information for the light pollution scientific community. Moreover, city councils' databases only cover public lightning managed by the municipality, thus not including private lights and the ones who might be managed by other kinds of local government.

B. Geolocalization

The city council database distinguishes between five different lamp types, based on the technology used for light generation, with huge disparity in number (percentage in parenthesis): high-pressure Sodium (HPS, 76.0%), light emitting diode (LED, 17.4%), plug-in fluorescent (PL, 2.4%), metal-halide (MH, 2.4%) and fluorescent (F, 1.8%) lamps. The reported colour temperature for the different technologies are 3000-3500 K for LED, 4000 K for PL, 3500-4000 K for MH and 3500 K for F lamps.

Associating an image from the NightUp dataset to a lamp from the city council database is not trivial in some cases. In fact, the locations in our dataset have limited precision (15 m on average) and accuracy (users may have shot the photos from a certain distance of the streetlight). This may lead to mistakes in associating a NightUp image to a streetlight from the city council database, especially in areas with high density of streetlights of different types. To minimize such error source, we restrict our analysis to the photos for which the three closest lamps are of the same type.
C. Colour extraction with different devices

In order to check whether smartphone cameras are able to provide consistent information about the colour of a light source, we investigate the Hue-Saturation distribution of the extracted colours from photos taken in a small area (Latitude $\in [41.2884, 41.2888]$, Longitude $\in [1.9799, 1.981]$) in which only HPS lamps are present, according to the city council database. The assumption is that in a small area the lamps would be as similar as possible, allowing us to compare how different smartphone cameras respond to the same light source even without prior calibration. In Figure 6 we highlight the colours corresponding to the photos taken by four different devices in that area. Even though the different points are not exactly in the same region of the H-S diagram, showing some dispersion among the different devices, all photos cluster in the top left of the diagram, corresponding to the warm colours associated with HPS lamps. This certifies that it is possible to identify colours within a certain precision with the NightUp method, even with different devices that have not been calibrated for this task.

1. Lamp colour clusterization and automatic labelling

The previous figure suggests that - even though different devices may assign slightly different colours to the same light source - the colours extracted with our procedure for the same kind of lamp will cluster in the same region of the Hue-Saturation space.

In fact, thanks to the comparison of our human-verified database with the city council, we observe that the majority of the extracted colours associated to HPS lamps via the method described in paragraph III.B fall in the region marked with the orange border in Figure 7. This region contains 82% of the photos associated to HPS lamps and only 3.2% of non-HPS lamps. For this reason, we can assume that a future photo falling in this region will be most likely a lamp that emits warm light, like a HPS one or a LED with color temperature smaller than 3000 K.

Using only the average colour as information, it is difficult to distinguish between different light technologies that emit light in overlapping regions of the visible spectrum [27]. For this reason, we cannot assign a specific light technology to an area of the H-S diagram as we did for the HPS lamps. Of course, having access to the lighting technology (and hence to the exact light spectrum) would be optimal. However, in many light pollution studies it is sufficient to be able to distinguish in which region of the light spectrum the lamp emits more light. For this reason, we focus on demonstrating that the NightUp method can successfully distinguish warm lamps, such as the HPS ones, with almost no emissions in the blue region of the spectrum, from other light sources with colour temperature higher then 3000 K that show more emissions in the blue region of the spectrum.

We decide to create three separated regions, as shown in Figure 7 assigned to cold (blue box, 33.3% of total non-HPS lamps), neutral (gray box, 33.3%) and warm (yellow box, 25%) to classify the non-HPS lamps according to their amount of blue light in their average colour. The last region contains all the rest of the H-S space, that contains only a few outliers, usually related to poor quality images, in which phones wrongly capture the colours.
FIG. 7: Scatter plot of the whole NightUp database as a function of saturation (distance from the center) and hue (angle). The coloured lines separate the regions of the H-S pace corresponding to different blue-light intensity levels. The boundaries of these regions are indicated in the legend.

With this classification, it is easy to assign a new unlabelled photo to one of these categories, making it easy to distinguish between HPS lamps and the other sources of light and to characterize the light source according to its blue light content.

D. Comparing NightUp data to ISS images

As a further validation of our method, we compare our results to the ones obtained from the analysis of high resolution images taken by astronauts from the International Space Station (ISS).

1. ISS image analysis

We focus our study on the image ISS062-E-102578, which is the sharpest image of the area taken from the ISS in the period of the NightUp experiment. It is important to note that the ISS images are radiance calibrated [28], and their pixel value is given in radiance units (nW · cm⁻² · sr⁻¹ · Å⁻¹) (where nW refer to nanoWatts, cm to centimeters, sr to steradians and Å to angstroms). On the other hand, the NightUp images are displayed in JPEG format, which means that their pixel values have arbitrary units calibrated according to the ISO standards [42] and the correlated colour temperature (CCT) definition [43]. In order to be able to compare the images from the two different sources, we apply a sequence of manipulations to the ISS image which effectively generate pixel values comparable to those of a JPEG images (see Appendix for the detailed procedure).

2. Comparison between the NightUp and ISS results for the city of Casteldefells

Figure 8(a) depicts the ISS image ISS062-E-102578 centered in the Barcelona area, after the transformation described in the previous paragraph. In this image, we identify the pixels corresponding to the geolocations given by: i) the Nightup photos; ii) the Casteldefels city council database. We then extract the colour for each one of these pixels and plot the corresponding H-S values in Figure 8(b-c) for each set of localizations. As opposed to previous plots, the colour of each point here represents the kernel density estimation w.r.t. to a Gaussian kernel, \( \rho(H, S) \) [44]. As the number of data points is much larger than in previous cases, with this procedure we can show an accurate approximation of the density distribution of lamps in the H-S space. Here, in contrast to the NightUp analysis (see Figure 3), the distribution of colours is heavily concentrated around values of H~ 0 and does not yield a clear distinction between different lamp colours.

It is important to note that the ISS images have much lower spatial resolution than the NightUp data. For instance, the ISS sampling is 25 meters per pixel, with a Point Spread Function (PSF) of 75 meters. Moreover, the ISS image include all sources of artificial light at night, including the ones that are not public streetlights. While the NightUp dataset also contains non-streetlights sources, the users were asked to make photos only from the street, hence lowering the contribution of other light sources (e.g. lights in private gardens or soccer fields). Another important source of
error is the sensitivity towards obstacles for the ISS images: for example, the ISS cannot detect the streetlights under dense canopies of trees. Moreover, streetlights are usually directed towards the ground, as their main objective is illuminating the streets. In this case, the ISS images can capture their reflection, which affects the measured colours [30] and in some cases the direct light. On the other hand, NightUp photos capture the light source directly, without obstacles or reflections that may change the measured colour.

Finally, we observe in both Figure 8(b-c) that many points show saturation values $S = 1$ (outer ring in red). All these points correspond to colours with very low value $V$, meaning that the associated pixels were mainly dark (black). This could arise mostly from two phenomena: i) the presence of an obstacle between the light emitter and the satellite; ii) the ISS images are not automatically assigned to a location, but they are georeferenced with the procedure described in [28]; eventual errors in georeferencing the ISS images lead to extract the colour of pixel where no lamp is present. These two effects are not present in NightUp, as i) the photos are directly aimed at the light source, and hence no obstacle is expected; ii) the geolocation is automatically assigned to the photo by the NightUp app, even if with some smaller errors, as shown in Section III.B and Figure 5.

In Figure 9 we show the map of Castelldefels and highlight: a) the lamps of the city council database, as well as their type (orange for HPS, blue for the rest); b) the colour of the pixels associated to the city council database lamps as extracted from the ISS image; c) the colour of the images of the NightUp dataset. This figure highlights some differences between the information that can be extracted from the two sources. For instance, the ISS image allows
FIG. 10: **Top row:** Comparison between the colours from the ISS (b) and from NightUp (c) for a region labelled by the city council of Casteldefells as LEDs (a). **Bottom row:** Comparison between the colours from the ISS (e) and from NightUp (f) for a region labelled by the city council of Casteldefells as HPS (d). The colour of each point in (b,c) represent its kernel density estimation value w.r.t. a Gaussian kernel $\rho(H,S)$.

for a complete characterization of all the lamps in the municipality with just one image. However, it is difficult to distinguish regions with different lamp types, as we have shown in Figure 8. On the contrary, the NightUp approach allows for a much better characterization of the colour, allowing us to distinguish HPS lamps from the rest. However, data are not uniformly distributed and mainly concentrated in the most populated areas. Nevertheless, this can be addressed with more focused advertising and the collaboration with local entities in future NightUp campaigns. We note that the data presented in this paper was acquired at different times: the NightUp dataset was gathered from November 2019 to March 2020, the city council’s database was issued in Summer 2020 while the ISS image dates of March 2020. While some changes to the public streetlights may have happened in between, those are minimal and we don’t expect them to affect our analysis.

To further extend our comparison, we focus now on two areas of Castelldefels: each one contains lamps from the same lamp category as given by the city council database. Both are marked in Figure 9 with white boxes and further highlighted in Figure 10. The top-right area of the maps contains only HPS lamps, while the bottom-left contains only LED lamps. The extracted colours from the ISS image and NightUp are shown in Figure 10(b,e) and (c,f), respectively. We first focus on the LED region (Figure 10(a-c)), where 24 photos were acquired with NightUp and 174 points extracted from the ISS image. All the NightUp points cluster in the Neutral and Warm regions, which in both cases refer to non-HPS images, as shown in Section III.C.1 and Figure 7. On the other hand, the colours extracted from the ISS show a huge dispersion, without a clear clusterization. In the HPS area, 86 points in H-S diagram come from NightUp and 411 from the ISS. Importantly, 94.2% of the NightUp photos are located in the HPS colour region (Figure 10(f)). The colours extracted from the ISS showcase here more a pronounced cluster, compared to plot the LED area. However, the colours extracted from the ISS have low V values and concentrate in similar regions of the H-S diagram for the LED and HPS areas, making it difficult to distinguish between lamp categories. This further proves that the NightUp approach is an important tool for colour characterization in artificial lights at night.
We finally study how our method (see Sec. III.A) can classify unlabelled images according to their colour. To this end, in Fall 2020, we performed additional data acquisition in the cities of Barcelona and El Prat de Llobregat, for which we did not have any official database to compare with. Figure 11(a, d) show the colours extracted from the ISS ISS062-E-102578 image and the images gathered with the NightUp app, respectively. The latter were taken by trained citizens which additionally reported the visual colour of each streetlight photographed. While in this case we don’t have access to an official database of lamp types, we use the visual report of the trained citizens in order to validate the following analysis: in the Barcelona and El Prat de Llobregat urban areas, they report mainly yellowish lamps, while the photos in the industrial area of Barcelona were of whitish streetlights (boxed region in Figure 11(a, d)). This allowed us to apply our method with more non-HPS lamps, that were scarce in Castelldefels, but abundant in the boxed region. Figure 11(b,e) show the H-S diagram for the whole unlabelled dataset both for the ISS and for the NightUp images, respectively. We can see how NightUp data are much more clustered than the ISS ones: we can distinguish various groups of colours, which, after inspection by the trained citizens, are directly related to the type of lamp that was pictured. To showcase this, we show in Figure 11(c,f) an excerpt of the previous dataset only containing the images taken in the boxed region. The trained citizens reported that in that area there were only whitish lamps, except for one that was clearly yellow. This is reflected in Figure 11(f), where one can see that our analysis classifies these images in the neutral region, clearly separated from the rest of images of the unlabelled dataset. Importantly, the automatic colour analysis matches the colours described by the humans experts, hence further highlighting the validity of the proposed method.
IV. DISCUSSION

The NightUp citizen science experiment aims at collecting information about the spatial distribution of colour of artificial lights at night, a powerful tool to understand light pollution, as the effects of light on the behaviour and health of humans, animals and plants varies enormously depending of the range of the visible spectrum considered. By means of a cross-platform mobile application, users were asked to take photos of streetlights. Then, we devised an algorithm to detect the lamps in the photos and extract their colour. With the geolocation of the NightUp data, we created a map that displays the colour of lamps in different regions. In order to validate the method before deploying it at large scale, we focused a first pilot NightUp campaign in the Castelldefels area, near Barcelona (Spain), for which we had an updated streetlight database provided by the city council, containing the geolocation and the lamp technology (e.g. high-pressure sodium or LED) for each streetlight in the city. With this information, we were able to test the NightUp experiment in various conditions, testing both the accuracy of the data acquisition as well as the colour extraction algorithm. In particular, we show that our method gives a estimation of the color of artificial light at night, sufficiently precise to distinguish warmer light sources from colder ones without a specific calibration of the device. Thanks to the comparison with the city council database, we identify the area of the H-S diagram associated to lamps emitting warm light. We define other four areas associated to different classes of lamp colours, and with it create an automatic classification algorithm of new photos, which we will use in future NightUp campaigns to automatically classify the photos from users.

As final benchmark, and in order to directly test the method w.r.t state-of-the-art approaches, we compare our results with those obtained by analyzing satellite images taken from the ISS. The satellite approach has different benefits, like being able to characterize the colour of night illumination of large regions with very few, or even a single image shot. However, this comes with a smaller spatial resolution, that makes characterizing densely illuminated regions very hard. Moreover, the presence of obstacles or the fact that streetlights are pointed towards the ground heavily affect the colour analysis with this images. When compared to NightUp, we clearly see that colour extraction with the latter leads to clear distinctions between lamp colour categories, allowing to distinguish the HPS lamps from the other technologies, while the colours extracted from the ISS images can be hardly differentiated.

In general there is no other field work similar on concept to what NightUp propose and only other light pollution programs like Globe at Night[45] and Nachtlichter similar[46] and the ongoing Street Spectra project[47] could reach similar outcomes with much more resources.

To summarize, NightUp allows for an efficient and accurate estimation of the colours of streetlights in densely populated regions. The information extracted from this analysis could be then used by local governments to optimize outdoor lighting and address light pollution problems in their regions as well as by scientist in the light pollution community to expand the studies on the effect of the colour of light. For example, as an important outlook for this project, one can use the colours extracted from the NightUp data to efficiently calibrate satellite images, as for example the one arising from the Visible and Infrared Scanner (VIRS), to improve the precision of ecological and economical parameters extracted from them.

Furthermore, because of its minimal technical requirements and the simple storytelling, we believe that NightUp can potentially expand to a global scale with the support of successful engagement strategies, allowing scientists to access and then characterize vast regions of the world with very few measurements.

AUTHORS CONTRIBUTION

Conceptualization, Federica Beduini and Alejandro Sánchez De Miguel; Data curation, Gorka Muñoz-Gil, Alexandre Dauphin and Alejandro Sánchez De Miguel; Methodology, Gorka Muñoz-Gil, Alexandre Dauphin and Alejandro Sánchez De Miguel; Project administration, Federica Beduini; Software, Gorka Muñoz-Gil and Alexandre Dauphin; Validation, Alejandro Sánchez De Miguel; Visualization, Gorka Muñoz-Gil and Alexandre Dauphin; Writing – original draft, Gorka Muñoz-Gil, Alexandre Dauphin, Federica Beduini and Alejandro Sánchez De Miguel; Writing – review & editing, Gorka Muñoz-Gil, Alexandre Dauphin, Federica Beduini and Alejandro Sánchez De Miguel.

DATA AVAILABILITY

The necessary code to reproduce the results shown in this paper, as well as a representative sample of the images analyzed can be found in the repository: https://github.com/gorkamunoz/NightUp
ACKNOWLEDGMENTS

Thanks to the "Cities at Night" Program for the ISS images. The images from the ISS are courtesy of the Earth Science and Remote Sensing Unit, NASA Johnson Space Center. We acknowledge Scifabric for the design of the NightUp web application. Thanks to Lluís Torner and Silvia Carrasco for the institutional support and valuable discussions and advice. Thanks to David Paredes-Barato for his technical support and Lydia Sansmartí-Vila and Silvia Tognetti for participating in the data collection and engagement activities. We thank Castelldefels city council, that provided the streetlight database that allowed us to validate our data. Special thanks go to Alfonso López Borgoño, who helped us in different phases of the citizen science project. The Ramón Fernández Jurado library (especially its director Marta Granel), the Aula Senior project and the Col·legi Frangoal were crucial in helping us engaging with the people of Castelldefels. Last, but not least, all the anonymous citizens who contributed to NightUp deserve a special mention here because none of these results would have been possible without them.

FUNDING

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 847635 (UNA4CAREER). RALAN map project. This work was supported by the EMISSI@N project (NERC grant NE/P01156X/1). G.M.G. acknowledges support from the Austrian Science Fund (FWF) through SFB BeyondC F7102 and from Fundació LaCaixa. A.D. ERC AdG NOQIA; Ministerio de Ciencia y Innovación Agencia Estatal de Investigaciones (PGC2018-097027-B-I00/10.13039/501100011033, CEX2019-000910-S/10.13039/501100011033, Plan National FIDEUA PID2019-106901GB-B-I00, FPI, QUANTERA MAQS PCI2019-111828-2, QUANTERA DYNAMITE PCI2022-132919, Proyectos de I+D+i “Retos Colaboración” QUSPIN RTC2019-007196-7); European Union NextGenerationEU (PRTR); Fundació Celllex; Fundació Mir-Puig; Generalitat de Catalunya (European Social Fund FEDER and CERCA program (AGAUR Grant No. 2017 SGR 134, QuantumCAT U16-011424, co-funded by ERDF Operational Program of Catalonia 2014-2020); Barcelona Supercomputing Center MareNostrum (FI-2022-1-0042); EU Horizon 2020 FET-OPEN OPTOlogic (Grant No 899794); National Science Centre, Poland (Symfonia Grant No. 2016/20/W/ST4/00314); European Union’s Horizon 2020 research and innovation programme under the Marie-Skłodowska-Curie grant agreement No 101029393 (STREDCH) and No 847648 (“La Caixa” Junior Leaders fellowships ID100010434: LCF/BQ/PI19/11690013, LCF/BQ/PI20/11760031, LCF/BQ/PR20/11770012, LCF/BQ/PR21/11840013). A.D. further acknowledges the financial support from a fellowship granted by the La Caixa Foundation (ID 100010434, fellowship code LCF/BQ/PR20/11770012). Project cofinanced by the Diputació de Barcelona through the BiblioLab program (21296).

CONFLICT OF INTERESTS

A. Sánchez de Miguel discloses that he does consulting work sporadically for Savestars Consulting S.L. and is a member of the board of Cel Fosc. The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this article.

Appendix A: ISS image transformation procedure

In this paper we use the ISS image ISS062-E-102578, which we first process following the procedure proposed in Ref.[28] using synthetic photometry from high resolution spectra. As we stated in section III.D.1, the ISS image values are in physical units (radiance values), whereas the NightUp ones are in the JPEG format, with arbitrary units.

In order to transform from radiance values to the HS values used in this paper, we use the following sequence of manipulations, which effectively generate image values comparable to the JPEG ones. The ISS plots do not use the same CCT, contrast and saturation adjustments of the NightUp images, so they are not directly comparable, although there is a homomorphic transformation between them. In order to compare the different datasets, we set the center of the HS plot on known areas of pure white sources: this allows us to compare the two datasets, but it may also create a rotation on the HS plane and stretching effect. We then rescale the three channels following \( x = (x - x_{\text{min}}) / x_{\text{max}} \) with \( r_{\text{min}}, g_{\text{min}}, b_{\text{min}} = 0.01, 0.01, 0.003 \) and \( r_{\text{max}}, g_{\text{max}}, b_{\text{max}} = 0.681, 0.603, 0.558 \). The \( r_{\text{min}}, g_{\text{min}}, b_{\text{min}} \) values have been selected from the background of the delta of the Llobregat river, a non-illuminated area in the ISS image. With this operation, some negative values that do not have physical meaning may appear: they come from the noise distribution of the background. In order to avoid these artifacts, we set all the negative values to 0. We then apply the so-called
gamma correction $x \rightarrow x^{\gamma}$, with $\gamma_r, \gamma_g, \gamma_b = 1/1.8, 1/2.8, 1/3.9$. Finally, we transform the new RGB amplitudes to HSV [41]. As the JPEG images are in arbitrary units, we can choose the values $x_{\text{max}}$ and $\gamma_x$ ad hoc, so that the visual comparison between the two data sets becomes easier.

**REFERENCES**

[1] D. J. Forbes, GİScience & remote sensing **50**, 483 (2013).
[2] T. Ghosh, R. L. Powell, C. D Elvidge, K. E Baugh, P. C Sutton, and S. Anderson, The Open Geography Journal **3** (2010).
[3] C. D. Elvidge, K. E. Baugh, E. A. Kihn, H. W. Kroehl, E. R. Davis, and C. W. Davis, International Journal of Remote Sensing **18**, 1373 (1997).
[4] A. Sánchez de Miguel, J. Zamorano, J. G. Castaño, and S. Pascual, Journal of Quantitative Spectroscopy and Radiative Transfer **139**, 109 (2014).
[5] C. N. Doll, J.-P. Muller, and C. D. Elvidge, Ambio , 157 (2000).
[6] S. Asefi-Najafabady, P. Rayner, K. Gurney, A. McRobert, Y. Song, K. Coltin, J. Huang, C. Elvidge, and K. Baugh, Journal of Geophysical Research: Atmospheres **119**, 10 (2014).
[7] A. Sánchez de Miguel, C. C. Kyba, J. Zamorano, J. Gallego, and K. J. Gaston, Scientific reports **10**, 1 (2020).
[8] P. Falchi, P. Cinzano, D. Duriscoe, C. C. Kyba, C. D. Elvidge, K. Baugh, B. A. Portnov, N. A. Rybnikova, and R. Furgoni, Science advances **2**, e1600377 (2016).
[9] A. Laforge, J. Pauwels, B. Faure, Y. Bas, C. Kerbiriou, J. Fonderflick, and A. Besnard, Landscape Ecology **34**, 793 (2019).
[10] M. O. Román, Z. Wang, Q. Sun, V. Kalb, S. D. Miller, A. Molthan, L. Schultz, J. Bell, E. C. Stokes, B. Pandey, et al., Remote Sensing of Environment **210**, 113 (2018).
[11] X. Lian, L. Jiao, J. Zhong, Q. Jia, J. Liu, and Z. Liu, Environmental Pollution , 118110 (2021).
[12] J. J. Foster, C. Tocco, J. Smolka, L. Khaldy, E. Baird, M. J. Byrne, D.-E. Nilsson, and M. Dacke, Current Biology (2021).
[13] J. Pauwels, I. Le Viol, C. Azam, N. Valet, J.-F. Julien, Y. Bas, C. Lemarchand, A. S. de Miguel, and C. Kerbiriou, Landscape and Urban planning **183**, 12 (2019).
[14] L. Colmanstedt, J. I. Gillen, and L. E. Munstermann, Journal of the American Mosquito Control Association **24**, 331 (2008).
[15] R. Chepesiuk, Missing the dark: health effects of light pollution (2009).
[16] A. García-Saenz, A. Sánchez de Miguel, A. Espinosa, A. Valentín, N. Aragonés, J. Llorca, P. Amiano, V. Martin Sánchez, M. Guevara, R. Capelo, et al., Environmental health perspectives **126**, 047011 (2018).
[17] M. Aubé, J. Roby, and M. Kocifaj, PLoS one **8**, e67798 (2013).
[18] R. Gray, S. A. Perkins, R. Suryakumar, B. Neuman, and W. A. Maxwell, Journal of Cataract & Refractive Surgery **37**, 38 (2011).
[19] M. A. McIsaac, E. Sanders, T. Kuester, K. J. Aronson, and C. C. Kyba, Environmental epidemiology **5** (2021).
[20] A. Huss, L. van Wel, L. Bogaards, T. Vrijkotte, L. Wolf, G. Hoek, and R. Vermeulen, Environmental health perspectives **127**, 067001 (2019).
[21] K. J. Gaston and A. S. de Miguel, Annual Review of Environment and Resources **47** (2022).
[22] A. Sánchez de Miguel, J. Bennie, E. Rosenfeld, S. Dzurjak, and K. J. Gaston, Science Advances **8**, eabl6891 (2022).
[23] C. C. Kyba, T. Kuester, A. S. De Miguel, K. Baugh, A. Jechow, F. Hölker, J. Bennie, C. D. Elvidge, K. J. Gaston, and L. Guanter, Science advances **6**, e170528 (2017).
[24] A. Sánchez de Miguel, J. Bennie, E. Rosenfeld, S. Dzurjak, and K. J. Gaston, Remote Sensing **13**, 3311 (2021).
[25] N. Levin, C. C. Kyba, Q. Zhang, A. S. de Miguel, M. O. Román, X. Li, B. A. Portnov, A. L. Molthan, A. Jechow, S. D. Miller, et al., Remote Sensing of Environment **237**, 114443 (2020).
[26] T. Longcore, A. Rodriguez, B. Witherington, L. Penniman, L. Herf, and M. Herf, Journal of Experimental Zoology Part A: Ecological and Integrative Physiology **329**, 511 (2018).
[27] A. Sánchez de Miguel, C. C. Kyba, M. Aubé, J. Zamorano, N. Cardiel, C. Tapia, J. Bennie, and K. J. Gaston, Remote sensing of environment **224**, 92 (2019).
[28] A. Sánchez de Miguel, J. Zamorano, M. Aubé, J. Bennie, J. Gallego, F. Ocaña, D. R. Pettit, W. L. Stefanov, and K. J. Gaston, Remote Sensing of Environment **264**, 112611 (2021).
[29] Q. Zheng, Q. Weng, L. Huang, K. Wang, J. Deng, R. Jiang, Z. Ye, and M. Gan, Remote sensing of environment **215**, 300 (2018).
[30] J. Shepherd and J. Dymond, International Journal of Remote Sensing **24**, 3503 (2003).
[31] M. M. Haklay, D. Dörler, F. Heigl, M. Manzoni, S. Hecker, and K. Vohland, What is citizen science? the challenges of definition, in *The Science of Citizen Science*, edited by K. Vohland, A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perelló, M. Ponti, R. Samson, and K. Wagenknecht (Springer International Publishing, Cham, 2021) pp. 13–33.
[32] S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović, et al., Nature **466**, 756 (2010).
[33] C. Lintott, K. Schawinski, S. Bamford, A. Slosar, K. Land, D. Thomas, E. Edmondson, K. Masters, R. C. Nichol, M. J. Raddick, et al., Monthly Notices of the Royal Astronomical Society **410**, 166 (2011).
[34] B. T. Collaboration et al., Nature **557**, 212 (2018).
[35] J. H. M. Jensen, M. Gajdacz, S. Z. Ahmed, J. H. Czarkowski, C. Weidner, J. Rafner, J. J. Sørensen, K. Molmer, and J. F. Sherson, Phys. Rev. Research **3**, 013057 (2021).
[36] A. M. Land-Zandstra, J. L. Devilee, F. Snik, F. Buurmeijer, and J. M. van den Broek, Public Understanding of Science 25, 45 (2016).

[37] J. R. Palmer, A. Oltra, F. Collantes, J. A. Delgado, J. Lucientes, S. Delacour, M. Bengoa, R. Eritja, and F. Bartumeus, Nature Communications 8, 1 (2017).

[38] S. Odenwald, Citizen Science: Theory and Practice 5 (2020).

[39] E. S. Gedraite and M. Hadad, in Proceedings ELMAR-2011 (IEEE, 2011) pp. 393–396.

[40] O. Ronneberger, P. Fischer, and T. Brox, in International Conference on Medical image computing and computer-assisted intervention (Springer, 2015) pp. 234–241.

[41] W. Burger and M. J. Burge, Principles of Digital Image Processing: Fundamental Techniques. Undergraduate Topics in Computer Science (Springer-Verlag, 2009).

[42] ISO 12232:2006 (E), photography–digital still cameras–determination of exposure index, iso speed ratings, standard output sensitivity, and recommended exposure index (2006).

[43] D. B. Judd, JOSA 25, 24 (1935).

[44] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, The elements of statistical learning: data mining, inference, and prediction, Vol. 2 (Springer, 2009).

[45] C. E. Walker, S. M. Pompea, and D. Isbell, in Epo and a changing world: Creating linkages and expanding partnerships, Vol. 389 (2008) p. 423.

[46] N. Weinberger, S. Woll, C. C. M. Kyba, and N. Schulte-Römer, Sustainability 13, 11613 (2021).

[47] R. G. E. G. L. Zamorano, J., González and C. Tapia, Street Spectra, https://streetspectra.actionproject.eu/ (2022), [Online; accessed 20-September-2022].