Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed April 24, 2020 (MN LATEX style file v2.2)

The Urban Observatory: Quantification of the Rebound Effect

Priyanshi Singh[1]
[1]*New York University (NYU)*

April 24, 2020

1 ABSTRACT

The rebound effect is described as the circumstance where gains through implementation of energy efficient technologies are potentially offset by increased usage patterns. Such an effect may have implications on energy policy, business or building management and consumer behavior. Consequently, this has been an important subject of scientific inquiry. Previous research has examined whether there is an effect and what methodologies best measure it. We seek to quantify the rebound effect by introducing a new empirical methodology for observing and analyzing a particular energy end use, lighting. We do so by leveraging two areas of research at the NYU Center for Urban Science and Progress Urban Observatory, particularly identifying patterns in light usage with broadband imaging and classifying light bulb types by their spectral signature using hyperspectral imaging. We bring these work streams together and integrate urban lightscape data with geospatial data for the first time to address a research hypothesis as to whether there is a significant variation in total lights on duration for efficient versus non-efficient light technology. While a statistically significant rebound effect was not found, we were able to establish an empirical constraint on the amplitude of the rebound effect upon which to base future scientific inquiry. We were also able to quantify light usage for several light types as well as establish an analytical pipeline that shows promise as a new avenue for empirical research into the rebound effect.

2 INTRODUCTION

New York City is known as the city that never sleeps, which perhaps is a reason why energy consumption for residents in the New York State topped 1,115 trillion British Thermal Units (Btu) in 2015 Administration (2015) – enough to produce approximately 327 billion kilowatt hours (kWh) of electricity 201 (2017). Energy consumption is a global concern, as worldwide consumption is projected to increase by 48% in 2040 compared to the 2012 levels Singer and Peterson (2016). It is therefore no surprise that significant focus is placed on energy efficiency with substantial investment in both the public and private sectors. In 2016 the United States alone invested $231 billion in energy efficiency 201 (2017). Such investments may include retail-focused strategies for efficient lighting technologies, smart sensors for HVAC, and better insulation materials for roofing, among others.

It is also no surprise that much attention has been given to understanding the dynamics of energy efficiency and consumption, which can be difficult to characterize. In as early as 1865, Jevon's Paradox stated "when technological progress increases the efficiency with which a resource is used, the rate of consumption of that resource rises because of increased consumption and demand" Polimeni (2012). The modern term for this is the "rebound effect."

Unfortunately, not only are the dynamics of energy efficiency and consumption complicated, but empirical data is sparse and existing research can contain potential biases Sorrell and Dimitropoulos (2007), Sorrell et al. (2009). Critics of research on the rebound effect claim that it is widely overstated and can be a distraction from the cause of conservation Gillingham et al. (2013). Indeed the field of energy efficiency and economics has advanced greatly since Jevon first made his claim in 1865. For reasons explored in more detail later in this paper, it is clear that the definitions of the rebound effect and the approaches to measure it are quite varied and in need of more data as well as robust constraints with which to frame scientific research, such as hypothesis testing.

Fortunately, just as technological advancements are made in energy efficiency, so too are advancements in scientific inquiry. Progress with remote sensing instrumentation and technologies for collecting data, greater computational resources, and statistical learning methods can be brought to bear on issues such as the rebound effect. Indeed, the field of urban informatics, defined as "the study, design, and practice of urban experiences across different urban contexts that are created by new opportunities of real-time, ubiquitous technology and the augmentation that mediates the physical and digital layers of people networks and urban infrastructures," Hinds et al. (2011) shows tremendous promise for just this sort of inquiry.

We leverage advanced lightscape imaging survey methods pioneered by the Urban Observatory (UO) at the New York University Center for Urban Science and Progress (CUSP) to carry out the first empirical observation of the rebound in Brooklyn, New York. Specifically, we determine if significant variance in light use exists between users of

conventional and energy efficient technology. We developed and tested the methods in the analytical pipeline required to carry out this analysis and have collected 29 nights' worth of raw image data. We contribute to technical methods of light source identification, classification, and integration with existing records data as well as constructs an analytical pipeline that can be used in other aspects of UO research.

3 LITERATURE REVIEW AND PREVIOUS RESEARCH

3.1 Rebound effect

A key premise to the rebound effect is the implication that energy efficiency has a price content, such that "we may expect the energy price reduction that increased efficiency entails to exert an upward pressure on the demand for energy. This pressure will partly offset, and may more than offset, the energy saving that results from improved appliance efficiency" Khazzoom (1980).

Berkhout et al. (2000) provides a more thorough breakdown of the effect:

• technological progress makes equipment more energy efficient, therefore

• less energy is needed to produce the same amount of product, using the same amount of equipment; however, not everything stays the same, therefore

• the equipment has become more energy efficient, which causes the cost per unit of services of the equipment to fall, i.e. a price decrease, and

• a price decrease normally leads to an increase in consumption, therefore

• an increase in energy consumption may occur due to increased efficiency, which quantifies the rebound effect.

If the net usage of the product, such as lighting, is large enough then the increase in energy consumption may actually exceed the energy saved through efficiency completely, leading to an overall net increase in consumption compared to non-efficient technology.

The rebound effect can be broadened to include macroeconomic as well as microeconomic effects. These might include market-clearing price and quantity adjustments and even transformational effects in energy production and social conventions for energy use Greening et al. (2000). We remain focused on micro-economic behavior; however, it is worth noting that even if direct rebound effects are small, an aggregated impact could be non-trivial.

Empirical studies of the rebound effect generally fall into two categories: quasi-experimental and econometric Sorrell et al. (2009). Quasi-experimental studies attempt to directly measure the rebound effect by comparing energy demand before and after adopted energy efficient technology. This is done either by measuring the direct output of the technology, such as heating or lighting, or the energy input, such as fuel or electricity. In either case, obtaining direct measurements is exceedingly difficult. Most studies surveyed used within-group before-and-after comparisons instead of including a control group Sorrell et al. (2009), which makes them prone to bias Frondel and Schmidt (2005), Meyer (1995). Since households must opt-in to these studies, there is also a selection bias Hartman (1988). This is particularly problematic given two key counterfactuals that must be assessed: what the energy use would have been had the technology not been put in place, which is a measure of energy savings; and subsequently, what the energy use would have been had there not been a change in behavior, which is the rebound effect.

While engineering methods can help determine the energy saved through efficient technology, without the ability to approximate a randomly assigned controlled study or generate enough data, it is unlikely such studies could produce a statistically valid result for the rebound effect. For these reasons, econometric approaches are more common. They use secondary data sources to estimate the rebound effect, typically by estimating elasticities. Elasticity is the change in a particular variable, such as energy demand, following a change in another, such as energy efficiency, when all other variables are held constant Sorrell et al. (2009). Several common definitions of the rebound effect leverage energy efficiency, energy cost, and energy price elasticity with regard to useful work and demand for energy Sorrell and Dimitropoulos (2007).

Elasticity by definition requires that all other variables be held constant, opening these econometric approaches to bias and confoundedness. For example, while a general assumption may be that consumers react to a decrease in price by increasing consumption, it could also be true that energy-conscious consumers would not increase consumption despite cheaper energy. Alternatively, it is possible that the perception of saving energy leads consumers using efficient technology to feel more comfortable using more energy, independent of cost or even demand for useful work. While investments in energy efficiency do not necessarily translate into decreased energy consumption (or vice versa), it is an indication that this is not a straightforward matter and favors a more direct empirical inquiry into actual usage behavior.

3.2 Urban informatics and remote light sensing

Observing light remotely can improve upon previous methodologies, such as by reducing selection bias and obtaining a sufficiently large and varied sample size of both energy efficient lighting technology users (treatment group) and non-users (control group) to allow for more robust statistical assumptions, rather than before-and-after comparisons. Integrating remote observations with the

existing data can further improve statistical assumptions by conditioning on these covariates, such as income or building type.

Remote sensing allows direct independent observation of precisely when and with what type of technology lights are used. The benefits of measuring lighting use as it relates to urban characteristics have been demonstrated, although some key challenges include lack of data, particularly with more precise spatial resolution Hale et al. (2013).

A report on energy use in New York City suggested that improved lighting technologies could greatly reduce overall energy consumption, estimating that 13% of the energy use within benchmarked buildings of all types in New York City was used for lighting. However, that same study acknowledged that this estimate was based on a survey of professional auditors' assessments of lighting use by floor size and not directly measured data Council et al. (2016). Identifying a way to measure its use directly could have a great benefit to better characterizing the impact of investments in energy efficiency lighting technology. Previous research by the UO has laid a foundation for such an inquiry.

The UO first provided light curve extraction from images of a city lightscape Dobler et al. (2015), identifying unique light sources from broadband images of Manhattan and analyzing their light curves for major "on" and "off" transitions with a modified version of Canny's edge detection approach Canny (1986). The UO's subsequent survey of hyperspectral images of Manhattan used advanced imaging technology, a single slit spectrograph with 872 spectral channels from 0.4–1.0 μm Dobler et al. (2016). The resulting spectra were clustered into groups with similar features in the wavelength domain and correlated with spectra of known light sources. This method successfully classified 13 types of lighting used in the observation field and even proposed new spectral templates of light classification. These efforts demonstrate how an approach for quantifying the light use for specific lighting technologies can be developed by calculating total "on time duration" for light sources based on their on and off transitions and comparing these on times across lighting types.

We extend the previous UO studies with respect to urban lightscape analysis in several ways. First, we add to the corpus of data with an observation field of view that has less density and lower building structure than Manhattan. This more common building typology facilitates replication in other urban areas without Manhattan's unique profile. Second, we integrate both broadband images and hyperspectral images, which were previously analyzed separately, using these independent imaging techniques to facilitate source identification and classification. Third, we map these images with a three-dimensional projection of the built-environment by leveraging photogrammetry techniques. This allows for integration with public records, such as building and socioeconomic and demographic (SED) data. Finally, this research enhances, automates, and further validates several techniques for image processing, particularly by:

Figure 1. Figure 1. The hyperspectral camera (on the left) and the broadband camera (on the right) used to collect observed data. The cameras are shown here in the approximate orientation in which they captured data from a field of view south of the CUSP facilities at 1 Metro Tech Center, Brooklyn.

- improving upon and automating methods for light source extraction and interpretation of light curves to isolate on and off transitions and lighting use behavior;
- integrating both broadband images and hyperspectral images; and
- projecting light sources onto a 3-dimensional LiDAR model of the cityscape to enable integration with Primary Land Use Tax Lot Output (PLUTO), which could be used to differentiate between residential and commercial buildings

4 DATA COLLECTION

This project is founded upon primary data collection, which represents an important contribution for the field of urban informatics as well as provides insight into methods of instrumentation and raw image data processing.

4.1 Broadband imaging

We began broadband image collection on the evening of June 25, 2017 and continued through the evening of July 23,

Figure 2. Figure 2. Approximate broadband imaging field of view over Southern Brooklyn A range of residential building types is visible, including the Gowanus Houses public housing complex in the center of the view.

2017 for a total of 29 consecutive nights of data. For each night during the observation period, the broadband camera scanned the view shown in Figure 2 every 10 seconds from 9 pm to about 6 am, for a total of approximately 3,200 images per night. The raw images stored in the CUSP data facility are in 3-color (RGB) format with dimensions of 3072 x 4096 pixels. Each image is a composite of 20 exposures of approximately 34 milliseconds (ms) each for total exposure of approximately 680 ms per image. Each image scan is approximately 12 Megabytes (Mb). If 3,200 scans are collected in a given night at 12 Mbs per scan, about 38.4 Gigabytes (Gb) of data are produced that night alone. Approximately 1.1 Terabytes (Tb) of data have been collected from broadband image scans in the 29 night collection period.

Aside from the challenges of collecting, processing and manipulating such large datasets, the nighttime images have a low signal to noise (S/N) ratio that makes analysis and classification of particular sources difficult. Furthermore, the classification must rely upon unsupervised techniques as labeling data through human-annotated "ground truth" methods is not practical. This not only makes light source identification, classification, and on/off transition measurements challenging, but renders traditional methods of algorithm evaluation inadequate.

The broadband images in this research were used primarily to identify specific light sources within the image plane, which were registered with a shared frame of reference for the hyperspectral images and the three-dimensional LiDAR model, and to measure the duration of lighting use, particularly the total on time duration of a given source, which constitutes the dependent variable of the research hypothesis.

4.2 Hyperspectral imaging

Hyperspectral imaging (HSI) is a remote sensing technique through which we collect high resolution spectral data that can be used for various detection purposes. Several scientific studies Elvidge et al. (2010), Dobler et al. (2016) relating to urban context in the past used HSI for lighting type detection. The aforementioned Dobler et al. (2016) describes a novel approach adapted from astronomical techniques to detect lighting types in an urban landscape using a hyperspectral sensor. Dobler et al. (2016) also described the nuances associated with the post-processing of the

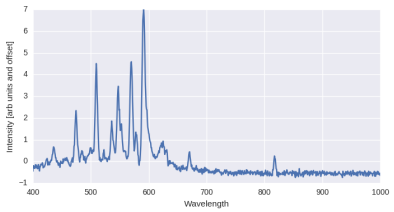

Figure 3. Figure 3. Relative Intensities of a light source across 848 wavelength channels between 400-1000nm

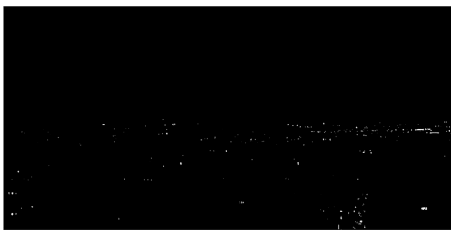

Figure 4. Figure 4. 2D representation of a hyperspectral scan with light sources

hyperspectral scans to extract the clean spectral information of various sources in the field and subsequently the identification of lighting types. Adopting a similar pipeline, this project uses the hyperspectral scans to match the usage information of sources obtained through broadband imaging with the corresponding lighting type, to measure the rebound characteristics across the lighting types.

The UO has a VNIR (Visible Near Infra Red) Hyperspectral Imager with a spectral range of 395-1008 nm. It is a vertical slit aperture instrument that collects a narrow beam of light that gets split by a prism into 848 hyperspectral channels with a spectral resolution of less than 1 nm (Figure 3). It uses a horizontal pan mechanism to sweep through the field of view and in our case takes ˜160 seconds to capture one complete scan of our subject field. The digital output from the camera is a raw file with 848 (number of spectral channels) x 1600 (pixel rows) x 3194 (pixel columns) dimensions. Each hyperspectral scan is about ˜8.2 GB. The camera captured a set of 15 observations on the night of May 26, 2017, producing ˜123 GB of data in one night. These scans were used to identify the spectra of various light sources within the scene (Figure 4).

4.3 LiDAR point cloud data

In addition to collecting broadband and hyperspectral data, we used existing external data sources to identify physical locations of each identified light source. Specifically, a LiDAR (Light Detection and Ranging) scan collected by the private mapping company Sanborn for the City University of New Yorkserved as the data source to map the topography of the observed scene to image pixels. The LiDAR dataset is from 2010 and thus may not include buildings that were built in recent years. However, establishing an analytical pipeline using the LiDAR data ensures that the work can be easily replicated once the next dataset becomes avail-

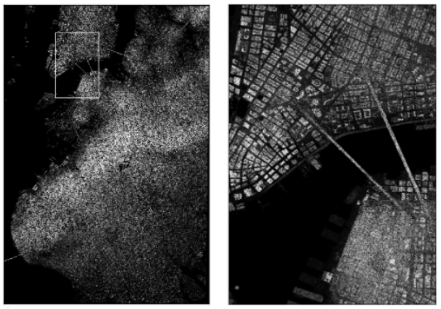

Figure 5. Figure 4. Rasterized LiDAR point cloud data of Brooklyn (left) and a zoomed-in view of Brooklyn and Manhattan bridges (right)

able (expected in fall 2017). The LiDAR data is segmented into tiles that cover all five boroughs of NYC at a roughly 1 foot spatial resolution. We selected the tiles containing the Brooklyn view and rasterized them into a spatial grid, a matrix of cells organized into rows and columns, where each cell contains elevation information (Figure 4).

5 METHODOLOGY

We propose a testable research hypothesis that there are observable differences in the lighting use of users of energy efficient technologies compared to users of non-efficient technology. The null hypothesis states that there is no significant difference in total light use—measured as the on time duration of light—between the users of energy efficient lighting and users of non-energy efficient lighting. Rejecting the null hypothesis would indicate a possible effect.

5.1 Source extraction

Prior to identifying light sources, we removed daylight images from the dataset by taking the average luminosity of an entire image and omitting any images with average light intensity greater than 0.5 (arbitrary) light units above the mean of the entire night, which resulted in omitting approximately 1,000 images, or 167 minutes of data from each night for the final analysis. The resulted in 2,600 images per night over a time period approximately from 9:16 pm in the evening to 4:30 am the following morning.

Previous UO research relied on a manual method for extracting light sources Dobler et al. (2015). To expand upon this, we used the time dependence of individual pixels to determine which pixels can be associated to the same source. The intuition is that a light turning on or off would cause higher correlation in contiguous pixels, so highly correlated neighboring pixels were grouped together as a light source. While factors such as camera drift may increase correlation over time, a sufficient threshold would filter these pixels out. Additionally, integration with pixel spectra data will further filter out data not likely to be a unique light source. As such, not only would this technique

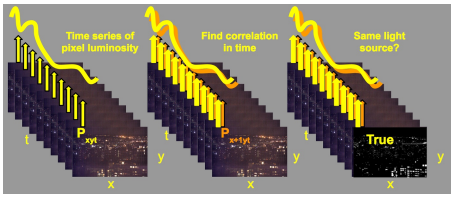

Figure 6. Figure 5. Temporal correlation of pixels (Left) A datacube represents time series of light intensity values for a pixel p at location (x, y) throughout the images in time (t); (center) correlation coefficients of neighboring pixel time series' are calculated; (right) highly correlated pixels are grouped and labeled as a unique light source and their location is preserved in a 2D boolean mask.

automate the light extraction process, but reduces preprocessing needed to clean noise.

To identify sources, we loaded images as a data cube of light intensity values for each pixel over all images in time at 1 minute increments. Each pixel time series was standardized by subtracting the mean from each intensity value and dividing by the standard deviation. The standardization was done for each night over 7 nights, June 25, 2017 through July 1, 2017, to produce an aggregated data cube of 3,031 images (7 nights, 433 min a night) x 3072 pixel rows x 4096 pixel columns. Future work should also apply a Gaussian filter over these time series to smooth patterns in the light curves.

The aggregated time-series vector of each pixel was multiplied and averaged by the pixel to the right, $< p_{ij} \cdot p_{ij+1} >$, and below, $< p_{ij} \cdot p_{i+1j} >$, in order to identify correlation of light intensity in time. Highly correlated pixels were considered components of the same light source. We calculated the mean light intensity across pixels in the light source, which we then used to extract the light curve of the given light source. Figure 5 visualizes the steps to create the boolean mask that identifies light sources.

Each group of contiguous pixels in this boolean mask were given unique labels that would be used to index individual light sources. With a threshold of 0.60 correlation over observations from June 25 to July 1, this method identified 6,870 possible light sources.

This technique performed well and indicates a promising new approach for future research, though we recommend further improvements such as including parallelization of the correlation algorithm to more efficiently handle the large memory footprint of multiple weeks worth of broadband imaging. Most important would be the introduction of a method for ground-truthing, such as randomizing pixels to identify a minimum noise threshold, or employing human-annotated labels.

5.2 Light curve extraction and on/off transition detection

The source extraction process preserved pixel location for light sources in the original raw broadband images so that

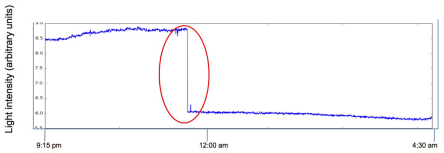

Figure 7. Figure 6. A single source light curve for a single night This light curve features a prominent "off" transition before midnight. Not all light curves will have such clean transitions.

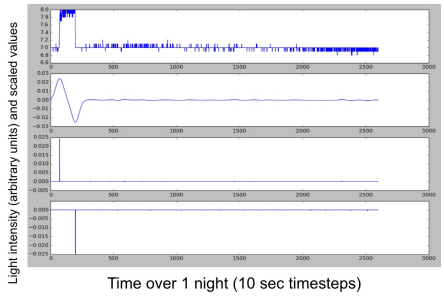

Figure 8. Figure 7. Key steps to detect on and off transitions The original light curve (top) is smoothed with a Gaussian operator, upon which the derivative is taken (2nd from top), from which extrema are found in positive directions to indicate "on" (3rd from the top), and negative directions to indicate "off" (bottom). Not pictured are the intermediary steps to amplify the transition and increase the signal / noise ratio.

the light curves could be extracted over all 29 nights of the observation period.

On and off transitions were detected following the adopted Canny edge detection method Canny (1986) that was used in earlier UO work Dobler et al. (2015). Noise was suppressed by convolving each light curve with a Gaussian filter and taking the first derivative of the result. The extrema of these derivatives indicate on and off transitions. Following Dobler et al. (2015), we conducted robustness tests to ensure a clear signal. First, the mean and variance of the signal were tightened by removing values outside 2σ, re-calculating the mean and variance to "clip" the distribution; this was iterated ten times. This adjusted distribution was used to identify values outside 10σ which indicated possible transitions. Second, the mean values for the 5 minute period before and after each edge was calculated and correlated with the change in light intensity, with a high correlation indicating that the transition exceed the surrounding noise. We used only the on and off transitions that passed these robustness tests in the final analysis. It is worth noting that of the 196 light sources classified, only 7 failed to produce adequate enough on and off transitions to calculate on time duration.

We seek to understand light usage behavior, particular the lights on duration. A quick examination of a particular set of light sources, the Gowanus Houses, a New York City Housing Authority (NYCHA) complex, revealed possible patterns in residential light usage over the course of the 29 night observation period. The Gowanus Houses were visually identified in the broadband image and sources within

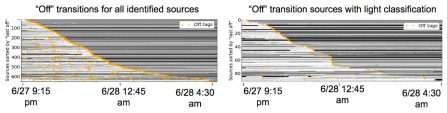

Figure 9. Figure 8. Light curves from June 27, 2017 The left plot shows light curve intensities (normalized) as sorted by the last off transition of the night. The right plot shows only light sources classified by type (n = 196) on the same night, where some steep drop-offs are visible indicating several sources turned off in around the same time.

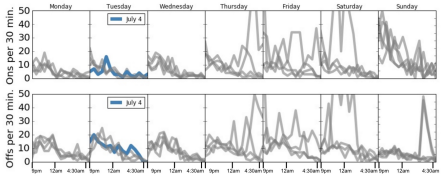

Figure 10. Figure 9. Nightly on and off transitions per 30 minutes for Gowanus Houses The plots above show the patterns of on transitions for nights of the week (top row) and off transitions (bottom row). The difference in on distributions and off distributions is apparent on typical weekdays. Distinct weekday vs weekend activity is visible as is anomalous activity leading up to and during the Independence Day holiday on Tuesday, July 4.

that area were selected (n = 166).

5.3 Lights on duration

Our dependent variable is the total lights on duration. This was calculated for each unique light extracted from the broadband mask using the time index of the on and off transitions. For a given light source, for each night $n \in N$ (N = 29), for each on or off tag per night $t \in T$, (T = the total on and off tags of a particular night), the total duration for the source is:

$$= \sum_n^N \sum_t^T n_t - n_{t-1}, \text{if } n_t = \text{off and } n_{t-1} = \text{on}$$

Some light curves may have multiple consecutive on transitions or multiple consecutive off transitions, so a logical framework was incorporated to preserve the previous state such that the above formula aggregated the total on duration across such on and off transitions, as illustrated in Figure 10. The assumption for this work being that intermediate on and off transitions represented changes in light intensity, but not a complete on and off transition. Such an example may be a shade drawn, etc. Such edge cases can be complicated, and warrant further improvements to the duration method.

5.4 Image registration of broadband and hyperspectral sources

Image registration is a process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. The aim is to integrate information from multiple observations of

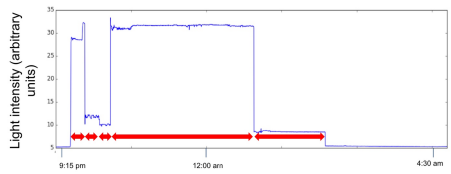

Figure 11. Figure 10. Single light curve with multiple consecutive on and off transitions Challenging cases appear when a light source has an off transition followed by another off transition. Such transitions were assumed to be variations of the same light source and the algorithm was designed to aggregated all intermediary transitions as a total on time, as indicated by the red arrows.

Figure 12. Figure 11. The figure shows the input image transparently overlaid on the reference image. The displayed image is taken from Loomis, John. (2009)

the same system. For this project, the imaging devices were deployed in interior locations. Thus, because the cameras were perfectly stationary (e.g., there were no vibrations due to the wind as there were in Dobler 2015) registration between images from the same camera was not required. We verified that such offsets were <1 pixel from the start time to end of our broadband and hyperspectral observations. However, registration between the broadband and hyperspectral images was necessary to associate individual sources.

In order to identify corresponding sources in both hyperspectral and broadband imaging, for each labeled pixel in the broadband source mask, the corresponding hyperspectral pixel was found by co-registering the broadband and hyperspectral images. Since the broadband images have a higher spatial resolution (3072 x 4096 pixels) than hyperspectral images (1600 x 3194 pixels), the pixel centroid locations were rescaled prior to registration. The scaling factor was calculated by comparing the distances between seven scattered sources in broadband and hyperspectral images.

The process of image registration typically includes two images: the first, called the reference image, is considered the reference to which the other images, called input images, are compared (Figure 11). The objective of image registration is to bring the input image into alignment with the reference image by applying a spatial transformation to the input image. A spatial transformation maps locations in the reference image to new locations in the input image. Determining the spatial transformation parameters, mainly row offset, column offset and vector rotation to bring the input image into alignment is the key to the image registration process (Figure 12).

Image registration is a sequential process and it begins with extracting sources from the input image and calculating quad of sources (4 sub-groups of sources) from the extracted sources by comparing the distances between sources

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$$

Figure 13. Figure 12. The figure shows rotation and offset matrix used to calculate row offset, column offset, and vector rotation.

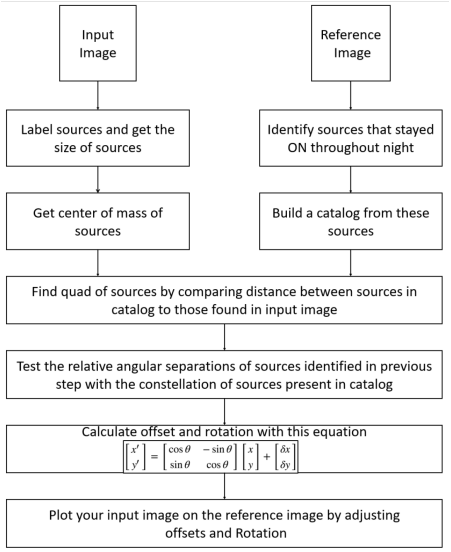

Figure 14. Figure 13. The figure shows graphical illustration of the image registration process

in the catalog and the ones identified in the input image. Subsequently, relative angular separations of sources found in the quad were tested with the constellation of sources in the catalog and row offsets, column offset and the vector rotation was calculated using rotation matrix from linear algebra (Figure 12). We used the resulting parameters to spatially transform the input image in order to bring it into alignment with the reference image (Figure 13).

The diagram below represents the graphical illustration of the process.

As noted, the hyperspectral camera sweeps through the field of view and the frame rate of the camera and pan rate of the pan/tilt mechanism are not constant across one complete scan, so the pixels in the scans are not square or uniformly shaped. In order to compensate for this, another scaling factor was calculated and aspect ratio of the hyperspectral pixels was corrected accordingly. Subsequently, the column pixel locations of sources in broadband were divided by this scaling factor in order to scale them down to the pixel configuration of the hyperspectral image (Figure 14). After applying these scaling factors, the broadband mask was registered onto the hyperspectral mask by calculating the row offset, column offset, and the vector rotation.

Due to the low spatial resolution of the hyperspectral images relative to the broadband images, image registration distorted and shrank the light source groups, so we removed all sources smaller than 20 pixels from the broadband source list. This resulted in 1,265 sources in the final mask

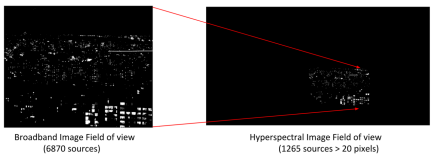

Figure 15. Figure 14. The figure shows the conversion of broadband mask (3072 x 4096 pixels, 6870 sources) to hyperspectral mask (1600 x 3194 pixels, 1265 sources)

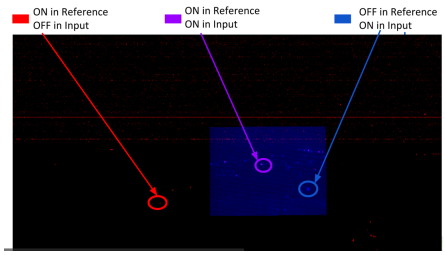

Figure 16. Figure 15. The registered broadband mask superpositioned on the hyperspectral mask. The corresponding sources in both hyperspectral and broadband imaging for each labeled pixel were identified. The red color indicates sources not present in broadband but in hyperspectral, whereas the blue color indicates sources present in broadband but not hyperspectral. Purple indicates sources that were present and aligned in both.

with which to classify lighting types from the hyperspectral images.

5.5 Source type identification

Source type identification was accomplished using the hyperspectral scans, in a 4 stage process as described below. The final lighting types obtained from this process were eventually used for rebound analysis.

Stage 1 - Stacking hyperspectral scans: The hyperspectral scans typically consist of detector noise that obscures the spectral information contained in it leading to low signal to noise ratio (Figure 16). In order to reduce the effect of noise and enhance the signal to noise ratio, multiple hyperspectral scans are stacked together to compute the mean spectral intensity for each pixel in each channel. Since the detector noise is random, this process of computing the mean intensity of multiple scans suppresses the noise. We collected 15 hyperspectral scans on May 26th, 2017 between 10-11PM, which were stacked together for further analysis.

Stage 2 - Dark current suppression : Dark current is another type of instrument-based noise that was observed, which needed to be cleaned to obtain the true spectra from the light sources. Unlike the detector noise which is evenly distributed in all spectral channels, dark current is unevenly distributed across the spectral dimension as illustrated in the Figure 18 below. In order to compensate for this, we adopted the 3-sigma clipping technique mentioned in Dobler et al. (2016). 3-sigma clipping is a process by which the median of each row/column of the scan across each spectral channel is computed by masking the pixels with intensities

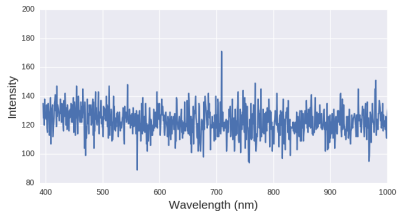

Figure 17. Figure 16. Spectrum of a pixel from raw hyperspectral scan having low signal to noise ratio

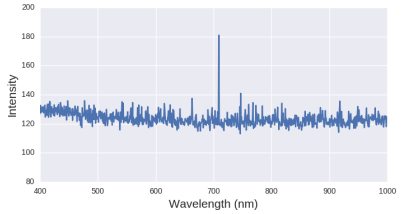

Figure 18. Figure 17. Spectrum of a pixel from stacked hyperspectral scans with improved signal to noise ratio

greater than 3 sigma with in channel multiple times (n=10) and subtracting that median intensity value from the respective row/column of each channel. This process ensures that the unevenly distributed noise gets suppressed and there by true source spectra are obtained.

Stage 3 - Extract mean spectrum for each source based on registered source mask: In order to extract the characteristic spectra for each of the lighting source in the broadband pipeline, we compute the mean spectrum for each source across pixels comprising the light source.

Stage 4 - Correlate source spectra to lab spectra: The final clean source spectra obtained from the previous three stages was correlated with lab spectra of known lighting types. National Oceanic and Atmospheric Administration (NOAA) Elvidge et al. (2007), Elvidge et al. (2010) has published a catalog of high resolution lab spectra between the 100-2500nm spectral range for 43 lighting types falling into 9 lighting categories i.e high pressure sodium, LED, fluorescent, low pressure sodium, oil, metal halide, halogen, incandescent and pressurized gas respectively. From this set of spectral templates we identified 17 sufficiently distinct lighting types. We then correlated our source spectra with these 17 distinct NOAA spectral templates. Each of

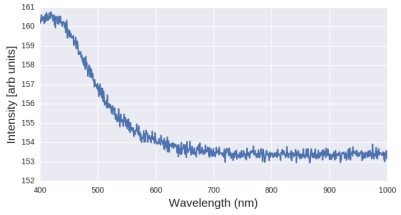

Figure 19. Figure 18. Mean spectrum of complete hyperspectral scan indicating dark current in lower wavelength channels

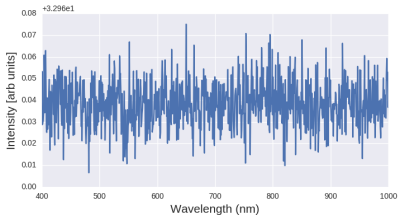

Figure 20. Figure 19. Mean spectrum of complete hyperspectral scan after cleaning using 3 sigma clipping

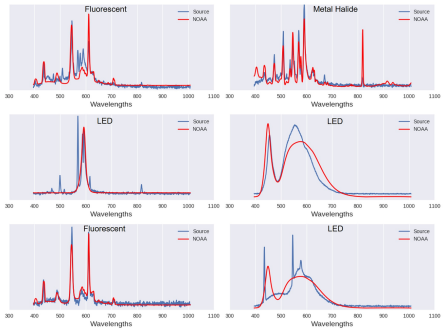

Figure 21. Figure 20. Source types detected for different lighting sources and the corresponding lab spectra

the sources was labeled with the type of the lab spectra template with which they correlated the highest. It should be noted that Dobler et al. (2016) also identified other new lighting types through clustering techniques that were not categorized within the NOAA catalog. Through this process, we were able to correlate and match 196 sources at a correlation threshold of 0.35. Figure 20 illustrates a few source spectra and the corresponding lab spectra they were found to be correlating with. Subsequently, the bar chart in Figure 21 illustrates the counts of different lighting types identified among the 196 sources at 0.35 threshold.

5.6 Point cloud data and image integration

A key to estimating the rebound effect is understanding the interaction between energy efficient lighting technologies and socioeconomic and demographic (SED) characteristics of the underlying population that uses these technologies. Previous, survey-based studies have concluded that, in addition to physical characteristics of buildings, certain household characteristics affect energy consumption Jones et al. (2015). In particular, household income, age composition, and educational level have a significant effect on energy consumption. Therefore, one can expect the size of the rebound effect to vary for different SED groups and building types.

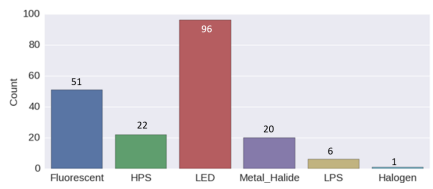

Figure 22. Figure 21. Count of source types detected vs the lighting types

To further explore this relationship, it is necessary to map observed light sources and their corresponding lighting types to exact geographic locations in the physical space. This would enable the extraction of zoning types as well as SED characteristics of the underlying populations that use the lighting technologies and exhibit on and off patterns we observe in the scene. The process of obtaining SED characteristics of the observed cityscape is a multi-step process involving external datasets and such techniques as rasterization, photogrammetry, and spatial joins.

Photogrammetry is the science of obtaining reliable information about the properties of surfaces and objects without physical contact with the objects Schenk (2005) or, put simpler, is the science of obtaining measurements from photographs. The relationship between coordinates in an image plane (in two dimensions) and coordinates in an object space (in three dimensions) can be modelled using mathematical models, namely collinearity equations. In this work, we used collinearity equations to model transformation from each pixel coordinates in the UO broadband image to the x, y, and z coordinates in the physical space.

The projection algorithm takes as inputs the image pixel coordinates as well as seven parameters of the camera: x, y, and z location coordinates, yaw, pitch, roll, and focus. In order to obtain the seven camera parameters, we implement a camera optimization algorithm. This algorithm takes as inputs image pixel coordinates of several unique points of reference, called “fiducial” points, as well as their corresponding coordinates in the LiDAR dataset, and outputs the seven degrees of freedom of the camera (Figure 22). To identify optimal camera parameters, we chose eight fiducial points, which are well-spaced in the image, including buildings close as well as far away from the camera. We then identified the corresponding x, y, z coordinates of these points in the LiDAR dataset (Figure 23). Both sets of fiducial points are represented as red dots in Figures 22 and 23.

We chose the initial parameters of the camera by an informed guess (approximate camera location on the 19th floor of 1 Metro Tech Center pointing south). These values are then used in the collinearity equations along with the coordinates of the eight LiDAR fiducial points to evaluate if the projected values are matching the pixel positions of the fiducial points in the image. The mean squared error between the fiducial points in the image and the projected LiDAR coordinates for each point is minimized by adjusting the seven camera parameters using gradient descent. This process is then iterated several times (with different initial guesses) in order to obtain a minimum value for the error term. The parameters corresponding to the function with the minimum error value are chosen as camera parameters for the photogrammetry. To visualize the goodness of fit of the chosen parameters, we plotted projected fiducial points

Figure 23. Figure 22. Eight fiducial points identified in the image (red dots) and projected LiDAR coordinates (yellow dots) used to optimize camera parameters

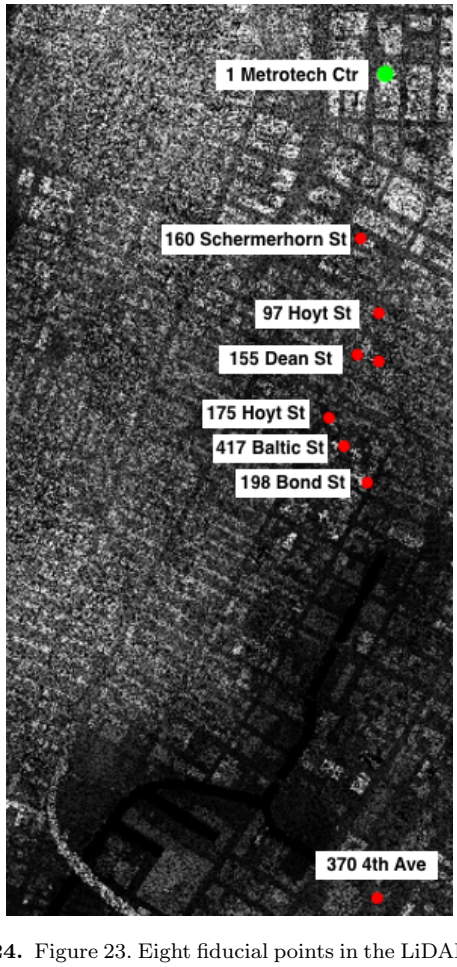

Figure 24. Figure 23. Eight fiducial points in the LiDAR dataset (in red). 1 Metro Tech Center (in green) is included for perspective.

on the image as yellow dots (Figure 22).

Once we optimized the camera parameters, we passed them along with the image dimensions into the collinearity equations, which converted the two-dimensional pixel coordinates into three-dimensional coordinates for a given distance from the camera. This is done using a transformation matrix containing the seven degrees of freedom of the camera and the collinearity equations. The photogrammetry algorithm projects x, y and z coordinates at one foot distance from the camera and checks whether at a given location, there exists a point in the LiDAR dataset that has higher elevation than the projected one. Of those points in space, the algorithm then stores the points closest to the camera. The algorithm returns three n-dimensional arrays as outputs: x and y coordinates of projected points as well as the distance in feet between projected points and the camera (Figure 24). The projected x and y coordinates (latitude and longitude) then allow us to map every pixel to a corresponding BBL using MapPLUTO shapefile for Brooklyn.

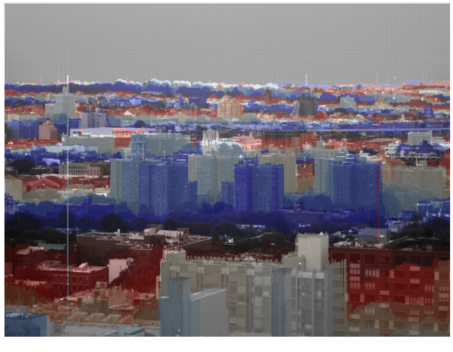

Figure 25. Figure 24. Distance between projected points and the camera. The color coding goes from red to blue in a continuous cycle. Notable that buildings that were built after 2010 are not included in the projection (e.g. the closest building in the bottom right corner of the image).

6 RESULTS AND IMPLICATIONS

6.1 Results

Taken together, the separate work steams produced a dataset of light sources classified by the lighting technology type and assigned features of on/off transitions, total on duration, and geospatial location coordinates that can be further integrated with SED data. The original source extraction identified 6,870 sources, further reduced to 1,265 after eliminating very small sources, of which 196 were classified with known light bulb types–189 with clean enough on and off transitions with which to calculate total duration of on time. Summary statistics for these sources are as follows:

Source Type	Number Classified	Number with clean "on ti
LED	96	91
Fluorescent	51	50
High Pressure Sodium	22	21
Metal Halide	20	20
Low Pressure Sodium	6	6
Quartz Halogen	1	1
All Types	196	189

While only several light sources in the Gowanus Houses region were classified by light type, there were 165 sources with clean on time duration and summary statistics as follows:

Total Sources (not just those classified)	Number with clean "on time" duration	M p a

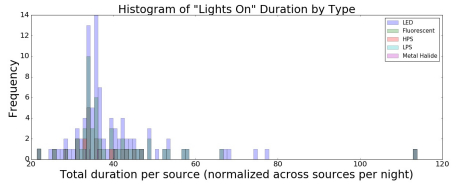

Figure 26. Figure 25. Histogram of total duration per source for each light bulb type (bin size=100) Duration was normalized across all sources for each night to produce independent observations. While the sample size was small overall (n=196), we can observe that the 5 types follow similar distributions. Quartz halogen was omitted with just 1 classified source.

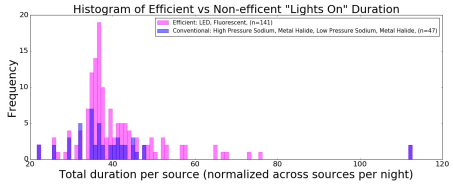

Figure 27. Figure 26. Histogram of total duration per source for the efficient and non-efficient lighting groups (bin size = 100) The two groups demonstrated similar distributions of total on time duration across all 29 observed nights.

group of "efficient" light types and a comparison group of "non-efficient" light types. These assignments were based on a basic comparison of light output in lumens per watt, a metric for light efficiency known as efficacy He et al. (1997), Thimijan and Heins (1983). An arbitrary threshold was set that bulbs with 40 lumens / watt efficacy or more were efficient Koekemoer (2014). Metal halide and high pressure sodium lamps may have relatively high source efficiency at the filament, but losses from trapped light, covers and lenses, inefficient ballasts and high operating temperatures mean these lamps typically operate at around 40 lumens / watt Light (2017). As such, LED and fluorescent lights were deemed efficient for the test group (n=141) and the remaining types deemed non-efficient for the comparison group (n=47). The on time duration values were scaled across sources for each given night and then aggregated across all 29 nights to produce independent observation samples of total on time duration per source.

The null hypothesis states that: there is no significant difference between total light use — to be measured as the duration that a given light is "on" — for residents who use energy efficient lighting compared to those who do not. Since the sample sizes vary (141 and 47) and no probability distribution was specified, an Anderson-Darling k-sample test Anderson and Darling (1952), Scholz and Stephens (1987), was conducted to determine if the two samples were drawn from a single population. The test failed to reject the null that the two distributions were significantly different (test statistic of 0.7164, with $p>0.05$).

It is noted that a larger sample size is necessary to robustly test the null hypothesis. A larger sample size would open up the opportunity to use other nonparametric tests, such as the Kolmogorov-Smirnov, and facilitate parameter estimation to support other tests. A histogram reveals a similar distribution for these two groups as seen in Figure 27.

6.2 Implications

One of the goals of this work is to contribute a body of empirical observations in an attempt to address some well-known challenges with measuring the rebound effect. Quasi-experimental approaches are vulnerable to selection bias, which can exacerbate conditions of fairness Sorrell et al. (2009). Subjects that are approached and consent to participate in surveys or monitoring likely do not represent a random sample of the population and can contribute to an "interpretation-after-the-fact" problem O'Neil (2013) that equates the results of a surveyed population with a larger population and can misguide conclusions. Econometric approaches that use price elasticity of demand, which were discussed earlier, make certain assumptions about human behavior which may not be a true representation of the real world.

The empirical approach to observing lighting use seeks to reduce the selection bias or econometric methods by directly observing behavior remotely. While this provides several advantages, caution should be taken not to underestimate the level of bias in this approach, either, particularly if data used in later studies. Even data that are objectively and remotely collected may contain biases if observations fail "to represent different groups in accurate proportions." Barocas and Selbst (2016)

This work seeks to map observed light sources to particular geographic locations and integrate this new data with existing records, such as SED information aggregated at the census tract level. Along with addressing privacy concerns referenced later in this paper, aggregating light sources to census tracts has several advantages. For example, it allows for the conditioning of covariates, such as building type and socioeconomic and demographic data, that can more fairly and accurately define groups represented in the data. For example, one popular statistical model, the potential outcome model created by Neyman and further developed by Rubin Rubin (2005), provides a framework with which to make assumptions of unconfoundedness. This includes methods to condition the treatment and control groups with covariates in order to determine a propensity score that can more accurately weight these covariates and isolate a treatment effect. For example, conditioning the observed lighting behavior for a specific census tract by weighting the proportion of income levels in the tract could address some of these issues. Future work should leverage the techniques developed here to assign covariates to observed light sources for such a purpose.

7 LIMITATIONS

Similar limitations as discussed in previous UO research apply to this project, among them that "curtains, interior reflections and mixed lighting type use can complicate the interpretation" Dobler et al. (2016). In addition, certain

limitations that are particular to this capstone project should be kept in mind when interpreting results. These include:

• Only North facing sources are captured. The field-of-view (FOV) of the images taken from a vantage-point in Downtown Brooklyn facing South exposes mostly the North-facing light sources of buildings. If there are any correlated aspects of facing North or downtown, those biases would be present in the data.

• Far-field census tracts might be under-represented. Another aspect of the data collection methodology is the possibility of under representation of the census tracts that are present in the far-field and over representation of census tracts in the near field, since more observable and distinct data sources may be visible. However, as noted earlier, it's also possible that the perspective of far-field light sources is more dense and causes them to be over-represented. Integrating sources with the built environment and aggregating to the census tract level should normalize this somewhat.

• Shorter night duration in summer. The northern hemisphere experiences summer solstice in June, yielding longer daylight hours. This project's observations take place in and near June, so changes in lighting use due to longer days will clearly influence this data. Hence, the same analysis could lead to very different results in terms of rebound characteristics when performed in a different time of the year. Clearly describing the data collection methods and dates and normalizing for varying nighttime hours will address this.

• Missing single-pixel light sources. The light sources were identified and extracted using an adjacent pixel correlation algorithm, so light sources smaller than the resolution of one pixel within the image would not get captured. While the broadband images have high resolution, this may become non-trivial with respect to the far-field sources, introducing another bias based on distance from the camera.

7.1 Ethics and privacy concerns

One of the broader objectives of the UO capstone project is to contribute to a better understanding of the relationship between humans and technology. The project aims to investigate whether the use of energy efficient lighting produces a behavioral response which offsets the beneficial effects of using that technology by collecting and processing of thousands of images of the urban landscape in order to identify light sources.

Since the camera view covers a large swath of the city, the project cannot specify a well-defined group of participants or obtain consent from all persons whose windows could potentially be captured in a UO image; however, it is important to note that no Personally Identifiable Information (PII) is obtained and the light exterior to buildings is in the public domain, which does not require consent C.F.R. (2009).

After careful consideration, we concluded that identities cannot be readily ascertained from the information collected by the UO. Furthermore, individual identities cannot be ascertained after combining observational data with publicly available socioeconomic and demographic data because U.S. Census data are already de-identified and aggregated to a census tract level.

To better protect the privacy and restrict the data from public access, the raw data are restricted (permission-based) within the CUSP data facility.

8 CONCLUSION AND FUTURE RESEARCH

This work put forth a novel approach to quantifying the rebound effect and contributed in three general ways: with urban image detection, urban informatics more broadly, and with respect to research into the rebound effect.

8.1 Urban image detection

Perhaps the most important contribution is establishment of an empirical constraint on the amplitude of the rebound effect upon which to base future scientific inquiry. Additionally, we developed an analytical pipeline for observing and analyzing light use. The effectiveness of previous techniques Dobler et al. (2015), Dobler et al. (2016) was further validated and enhanced, and critically, the two previously separate work streams, broadband light curve extraction and hyperspectral lighting type classification, were brought together for the first time and integrated with the 3-D LiDAR data projection to produce a completely novel methodology for analyzing light use.

The work was technically challenging, particularly when seeking such a heightened level of precision (exact light use for individual light sources) while dealing with a low signal / noise ratio; however, the benefit was a demonstrative success in quantifying light usage for specific light types. Additional work to improve the methods used as well as collect more data will surely increase the robustness of statistical analyses. Furthermore, the ability to integrate light sources with SED data vis-a-vis the 3-D projection technique opens an entirely new avenue of inquiry.

8.2 Rebound effect

The rebound effect exemplifies some of the most challenging aspects of applied science in that it seeks practical insights that can potentially influence policy, business, and human behavior, yet must maintain the integrity of scientific pursuit. This is particularly evident when considering the issues of bias examined earlier. Bringing a new empirical

method to bear on the complex engineering, social, economic dynamics of rebound effect is an important aspect of this work. For example, characterizing sample populations for SED attributes, such as income level, is crucial in developing robust statistical assumptions and avoiding issues of bias our confoundedness, yet most research utilizing such geo-specific data on behavior must rely upon sources like the American Community Survey that produces data aggregated at the census tract level and usually only every five years. Compared to this, gaining insight into a specific light source as observed in arbitrarily precise time units is a significant step forward.

This work also uses a methodology that should be easily replicated and generalized for other geographical locations and, certainly most tangibly, contributed a body of raw observed data that can be measured in Terabytes and will hopefully encourage wider adoption of these techniques by researchers investigating the rebound effect.

8.3 Urban informatics

The pipeline described also reinforces the interdisciplinary nature and creativity of the field of urban informatics. Previous UO research leveraged astronomical techniques in a novel way Dobler et al. (2015), Dobler et al. (2016) as well as edge detection techniques used for computer vision Canny (1986), which were key components in our methodology. We also introduced photogrammetry as a tool in the pipeline useful for 3-D projection. This research also demonstrated how urban informatics can apply empirical methods to ever broadening lines of inquiry, such as socio-economic issues. It speaks to the power of urban instrumentation as a means to better understand latent characteristics of the urban ecosystem, perhaps furthering the claim that urban science holds a unique place in the scientific realm Bettencourt and West (2010).

9 AUTHOR CONTRIBUTIONS

Acknowledgements: Professor Gregory Dobler provided instrumental guidance to the team and was actively involved in every step of the process. The team would also like to acknowledge the wider UO team, in particular Mohit Sharma and Andreas Karpf.

Daynan Crull: contributed much of the literature review on the rebound effect as well as conducted the data processing and analysis within the "broadband track" of the analytical pipeline, including source identification and extraction, light curve extraction, on/off edge detection and (forthcoming) total light usage duration measurement.

Akshay Penmatcha: contributed to the source type identification tasks in the pipeline. Worked on the Hyperspectral imaging data to identify the lighting source type. Used vectorized computations and correlation techniques in order to accomplish the tasks.

Anastasia Shegay: worked on integrating image data with point cloud data, which included applying photogrammetry techniques to convert rasterized 3-dimensional x, y and z data to a 2-dimensional image plane.

Priyanshi Singh: Contributed in analyzing broadband images by performing image registration, finding pixels deviations, and conversion of broadband mask to HSI mask which makes use of registration techniques by using Linear Algebra and rotation matrix computations.

References

World Energy Investment 2017. IEA, jul 2017. doi: 10.1787/9789264277854-en. URL https://doi.org/10.1787%2F9789264277854-en.

U.S. Energy Information Administration. State Energy Data System. 2015.

Theodore W Anderson and Donald A Darling. Asymptotic theory of certain goodness of fit criteria based on stochastic processes. *The annals of mathematical statistics*, pages 193–212, 1952.

Solon Barocas and Andrew D Selbst. Big data's disparate impact. 2016.

Peter H.G. Berkhout, Jos C. Muskens, and Jan W. Velthuijsen. Defining the rebound effect. *Energy Policy*, 28(6-7):425–432, jun 2000. doi: 10.1016/s0301-4215(00)00022-7. URL https://doi.org/10.1016%2Fs0301-4215%2800%2900022-7.

Luis Bettencourt and Geoffrey West. A unified theory of urban living. *Nature*, 467(7318):912–913, oct 2010. doi: 10.1038/467912a. URL https://doi.org/10.1038%2F467912a.

John Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, nov 1986. doi: 10.1109/tpami.1986.4767851. URL https://doi.org/10.1109%2Ftpami.1986.4767851.

C.F.R. 45 C.F.R. § 46.102 (Protection of Human Subjects 2009). 2009.

Urban Green Council, City of New York, NYU Center for Urban Science, and Progress. NYC Energy and Water Use 2013 Final Report. *http://www.nyc.gov/html/gbee/downloads/pdf/nyc$_e$nergy$_w$ater$_u$se$_2$013*

Gregory Dobler, Masoud Ghandehari, Steven E. Koonin, Rouzbeh Nazari, Aristides Patrinos, Mohit S. Sharma, Arya Tafvizi, Huy T. Vo, and Jonathan S. Wurtele. Dynamics of the urban lightscape. *Information Systems*, 54: 115–126, dec 2015. doi: 10.1016/j.is.2015.06.002. URL https://doi.org/10.1016%2Fj.is.2015.06.002.

Gregory Dobler, Masoud Ghandehari, Steven Koonin, and Mohit Sharma. A Hyperspectral Survey of New York City Lighting Technology. *Sensors*, 16(12):2047, dec 2016. doi: 10.3390/s16122047. URL https://doi.org/10.3390%2Fs16122047.

C. D. Elvidge, P. Cinzano, D. R. Pettit, J. Arvesen, P. Sutton, C. Small, R. Nemani, T. Longcore, C. Rich, J. Safran, J. Weeks, and S. Ebener. The Nightsat mission concept. *International Journal of Remote Sensing*, 28(12):2645–

2670, jun 2007. doi: 10.1080/01431160600981525. URL https://doi.org/10.1080%2F01431160600981525.

Christopher D. Elvidge, David M. Keith, Benjamin T. Tuttle, and Kimberly E. Baugh. Spectral Identification of Lighting Type and Character. *Sensors*, 10(4): 3961–3988, apr 2010. doi: 10.3390/s100403961. URL https://doi.org/10.3390%2Fs100403961.

Manuel Frondel and Christoph M Schmidt. Evaluating environmental programs: The perspective of modern evaluation research. *Ecological Economics*, 55(4):515–526, 2005.

Kenneth Gillingham, Matthew J Kotchen, David S Rapson, and Gernot Wagner. Energy policy: The rebound effect is overplayed. *Nature*, 493(7433):475–476, 2013.

Lorna A. Greening, David L. Greene, and Carmen Difiglio. Energy efficiency and consumption — the rebound effect — a survey. *Energy Policy*, 28(6-7):389–401, jun 2000. doi: 10.1016/s0301-4215(00)00021-5. URL https://doi.org/10.1016%2Fs0301-4215%2800%2900021-5.

James D Hale, Gemma Davies, Alison J Fairbrass, Thomas J Matthews, Christopher DF Rogers, and Jon P Sadler. Mapping lightscapes: spatial patterning of artificial lighting in an urban landscape. *PloS one*, 8(5):e61460, 2013.

Raymon S Hartman. Self-selection bias in the evolution of voluntary energy conservation programs. *The Review of Economics and Statistics*, pages 448–458, 1988.

Y. He, M. Rea, A. Bierman, and J. Bullough. Evaluating Light Source Efficacy under Mesopic Conditions Using Reaction Times. *Journal of the Illuminating Engineering Society*, 26(1):125–138, jan 1997. doi: 10.1080/00994480.1997.10748173. URL https://doi.org/10.1080%2F00994480.1997.10748173.

Pamela Hinds, John Tang, Jakob Eyvind Bardram, Jian Wang, and Nicolas Ducheneaut. Proceedings of the ACM 2011 conference on Computer supported cooperative work. 2011.

Rory V Jones, Alba Fuertes, and Kevin J Lomas. The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings. *Renewable and Sustainable Energy Reviews*, 43:901–917, 2015.

J Daniel Khazzoom. Economic implications of mandated efficiency in standards for household appliances. *The energy journal*, 1(4):21–40, 1980.

Charles Koekemoer. https://insights.regencylighting.com/author/charles-koekemoer. 2014.

Innovative Light. http://www.innovativelight.com/hid-vs-led-lighting/. 2017.

Breed D Meyer. Natural and quasi-experiments in economics. *Journal of business & economic statistics*, 13(2): 151–161, 1995.

Cathy O'Neil. *On being a data skeptic*. O'Reilly Media, Inc., 2013.

John M Polimeni. *The Jevons paradox and the myth of resource efficiency improvements*. Earthscan, 2012.

Donald B Rubin. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322–331, 2005.

Toni Schenk. Introduction to photogrammetry. *The Ohio State University, Columbus*, 2005.

F. W. Scholz and M. A. Stephens. K-Sample Anderson–Darling Tests. *Journal of the American Statistical Association*, 82(399):918–924, sep 1987. doi: 10.1080/01621459.1987.10478517. URL http://www.johnloomis.org/ece564/notes/tform/image_reg/reg1/image_reg1.html.

L.E. Singer and D. Peterson. International Energy Outlook 2016. 484, 2016.

Steve Sorrell and John Dimitropoulos. UKERC Review of evidence for the rebound effect: Technical Report 2: Econometric studies. 2007.

Steve Sorrell, John Dimitropoulos, and Matt Sommerville. Empirical estimates of the direct rebound effect: A review. *Energy policy*, 37(4):1356–1371, 2009.

Richard W Thimijan and Royal D Heins. Photometric, radiometric, and quantum light units of measure: a review of procedures for interconversion. *HortScience*, 18(6):818–822, 1983.