

# THE URBAN OBSERVATORY: A MULTI-MODAL IMAGING PLATFORM FOR THE STUDY OF DYNAMICS IN COMPLEX URBAN SYSTEMS

Gregory Dobler<sup>†,1234</sup>, Federica B. Bianco<sup>1234</sup>, Mohit S. Sharma<sup>4</sup>, Andreas Karpf<sup>4</sup>, Julien Baur<sup>4</sup>,  
Masoud Ghandehari<sup>45</sup>, Jonathan Wurtele<sup>6</sup>, Steven E. Koonin<sup>45</sup>

April 15, 2019

**Abstract:** We describe an “Urban Observatory” facility designed for the study of complex urban systems via persistent, synoptic, and granular imaging of dynamical processes in cities. An initial deployment of the facility has been demonstrated in New York City and consists of a suite of imaging systems – both broadband and hyperspectral – sensitive to wavelengths from the visible (~400 nm) to the infrared (~13 micron) operating at cadences of ~0.01 – 30 Hz (characteristically ~0.1 Hz). Much like an astronomical survey, the facility generates a large imaging catalog from which we have extracted observables (e.g., time-dependent brightnesses, spectra, temperatures, chemical species, etc.) to a parallel source catalog. We have demonstrated that, in addition to the urban science of cities as systems, these data are applicable to a myriad of domain-specific scientific inquiries related to urban functioning including energy consumption and end use, environmental impacts of cities, and patterns of life and public health, and that an Urban Observatory facility of this type has the potential to improve both a city’s operations and the quality of life of its inhabitants.

## 1. Introduction

With millions of interacting people and hundreds of governing agencies, urban environments are the largest, most dynamic, and most complex human systems on earth. Some 80% of the US population (and 50% of the global population) now live in cities, and within those urban boundaries are a multitude of interactions between the three fundamental components of urban systems: the inhabitants, the natural environment, and the built environment. The study of cities as complex systems is not new. Research in urban planning, engineering, transportation, etc., all have a rich history of quantifying city life at

multiple scales and with networked interactions. The application of that work addresses everything from quality of life to public health to sustainability and resilience. However, there are two recent revolutions that are leading to a dramatic change in the way we understand complexity in urban environments: the systematic collection, digitization, and curation of vast quantities of urban records data and the advent of computational techniques that make it possible to jointly analyze large, disparate data sources.

An important field of study for complex urban environments is remote sensing and the associated data collection from overhead

---

<sup>†</sup> gdobler@udel.edu

<sup>1</sup> Biden School of Public Policy and Administration, University of Delaware

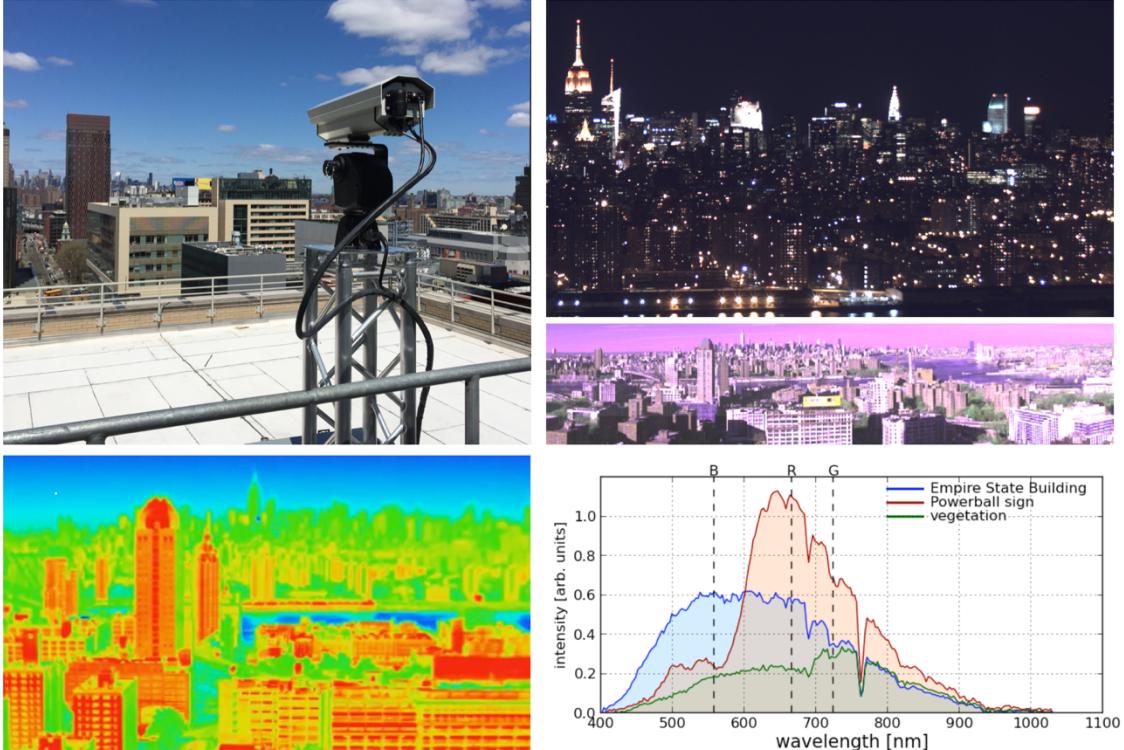
<sup>2</sup> Department of Physics and Astronomy, University of Delaware

<sup>3</sup> Data Science Institute, University of Delaware

<sup>4</sup> Center for Urban Science and Progress, New York University

<sup>5</sup> Civil and Urban Engineering, Tandon School of Engineering, New York University

<sup>6</sup> Department of Physics, University of California, Berkeley



**Figure 1.** *Top left:* an example of one of our Urban Observatory (UO) deployments of a Visible Near-Infrared (VNIR) Hyperspectral camera. The camera itself is encased in a weatherized housing and sits atop a pan/tilt, both of which are connected to (and controlled by) our backed computing infrastructure. *Upper right:* An example broadband nighttime image of Midtown Manhattan and the Lower East Side captured by that vantage point [1]. *Lower left:* An example infrared image captured from the same vantage point. *Middle and lower right:* An example VNIR hyperspectral scan of the same scene shown in false color. The *middle* image is integrated over all wavelengths while the *lower* panel shows three spectra from the scene [2].

satellite platforms at a variety of wavelengths. This satellite data is both large in volume and rich in information content, and its analysis has been used to correlate remote observables with land use characteristics, energy consumption, light pollution, poverty, and urban growth (among many others). Characteristically, these overhead platforms provide spatial resolution of  $\sim 10$ s of meters and temporal resolution  $\sim 1$  day or more.<sup>1</sup> However, at present, this resolution is insufficient to study dynamical interactions in cities on the time scale of minutes, and the overhead vantage point –

while ideal for total spatial coverage – only captures photons that are beamed upward providing an incomplete picture for several science cases (e.g., the ecological and public health impacts of light pollution).

In this paper, we describe an “Urban Observatory” (UO) for the study of dynamical processes in complex urban systems. As we outline in §2, our initial deployment of this UO in New York City (NYC) consists of side-facing imaging systems mounted atop tall buildings, providing a synoptic view of city

<sup>1</sup> though recent advances in satellite technology are pushing towards increased granularity on both fronts.

skylines (see Figure 1) [1]. The instrumentation of the UO is sufficiently granular to segment the imaging at the sub-building level, and – inspired by astronomical observatories operating in survey mode such as the *Sloan Digital Sky Survey* or PanSTARRS – the system operates persistently in order to observe multiple temporal scales: minutes, hourly, daily, weekly, monthly, annual, etc. In §3, we give an in depth description of the range of urban phenomenology that is accessible via this system, including work done to date on three key components of complex urban systems and urban metabolism more generally: energy use in cities, environmental impacts, and human factors. In §4 we conclude with what we envision to be important future deployments of UO facilities in other cities.

## 2. Instrumentation

The core UO platform consists of a suite of imaging systems mounted atop tall buildings at distance of 1-2 miles from the area of study, along with associated computing hardware located both at the device (where minimal edge computations are performed) and at a central server location. The former consists of small mini-computers (Raspberry PIs or similar), while the latter is a multi-purpose backend infrastructure that

1. drives the imaging devices remotely;
2. pulls and queues data transfer from the deployed devices;
3. performs signal processing, computer vision, and machine learning tasks on the images and associated products;
4. holds and serves imaging and source databases associated with the data collection and analysis.

This backend infrastructure/architecture is described in detail in Appendix A.

While we envision that the UO will ultimately carry out observations across the full (available) electro-magnetic spectrum at

multiple time scales, the current deployments – that yield the science content described in §3 – consist of the following modalities.

### 2.1 Broadband Visible

The first deployed camera system was a three color (RGB), 8 Megapixel USB camera with a 35mm lens mounted atop a tall building in Brooklyn, NY with a panoramic view of the Manhattan skyline (Figure 1). The instrument was set to a fixed pointing, was enclosed in a weatherized housing, and was triggered to acquire images via a co-located laptop in a weatherized casing. The initial image capture rate was  $f = 0.1$  Hz [1]. Current deployments have incorporated a pan/tilt mechanism for increased field-of-view (FOV) via multiple pointings, mini-PCs for triggering and data transfer, and both ethernet controlled, 20 Megapixel cameras and DSLRs operating in video mode at a sampling frequency of  $f \approx 30$  Hz.

### 2.2 Broadband Infrared

Our current broadband infrared (IR) devices are FLIR A310/320 cameras with a pixel resolution of 320x240, wavelength range of 7.5–13 micron, and temperature sensitivity of  $\pm 2^\circ$  C. As with our visible wavelength imaging, our initial IR deployment was encased in a weatherized housing, had a fixed pointing, and operated at  $f = 0.1$  Hz (Figure 1 shows an example image), while subsequent deployments incorporate a pan/tilt mechanism for increased FOV.

### 2.3 Visible and Near-IR Hyperspectral

In addition to our broadband imaging devices, we have deployed hyperspectral imagers operating at visible and near-infrared (VNIR) wavelengths. These instruments are single-slit spectrographs: the aperture is a vertical slit while a diffraction grating behind the slit

generates the wavelength information. They are mounted atop pan/tilt mechanisms that provide horizontal information as the detector is exposed at  $\approx$ 30 frames per second during the pan. The wavelength range is 0.4–1.0 micron with an average bandwidth of 0.75 nm resulting in  $\sim$ 850 spectral channels [2]. Scans are captured at cadences of  $\sim$ 10<sup>-3</sup> Hz. An example image and associated spectra are shown in Figure 1.

#### 2.4 Longwave IR Hyperspectral

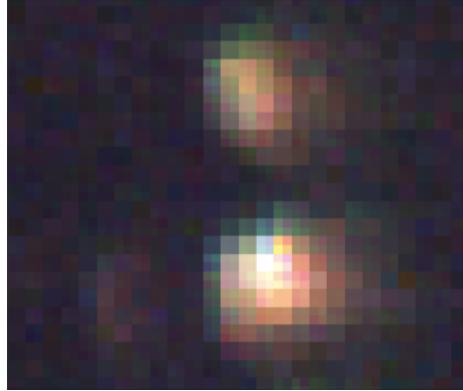
In April of 2015, the UO carried out a test deployment of a Long Wave IR (LWIR) hyperspectral camera. This actively cooled instrument was sensitive to 7.5–13.5 micron in 128 spectral channels and was operated in survey mode at roughly  $f \sim 0.01$  Hz, using the same panning mechanism described for the VNIR camera above. The deployment was done in collaboration with the *Aerospace Corporation* from whom the UO rented the equipment for an 8-day observational campaign [3].

#### 2.5 Privacy Protections

Given the nature of our observational platform, it is of highest importance that the UO ensure privacy protections for inhabitants of cities. This imperative is not unique to the UO nor is it restricted to imaging modalities; all observational platforms (be they acoustic sensors, air quality sensors, traffic sensors, etc.) *must* enact strict limits on the type and content of the data they collect – and the analysis procedures applied to that data – to protect the privacy of individuals.

The UO’s privacy protection policies consist of four core components

1. No personally identifiable information (PIII) is collected.
2. All imaging is strictly limited in pixel resolution so that building interiors cannot be seen (see Figure 2).



**Figure 2.** A zoomed-in example of two of the closest sources in the upper right panel of Figure 1 (figure reproduced from [1]). The resolution of all UO imaging is sufficiently low that no interior information is visible or detectable.

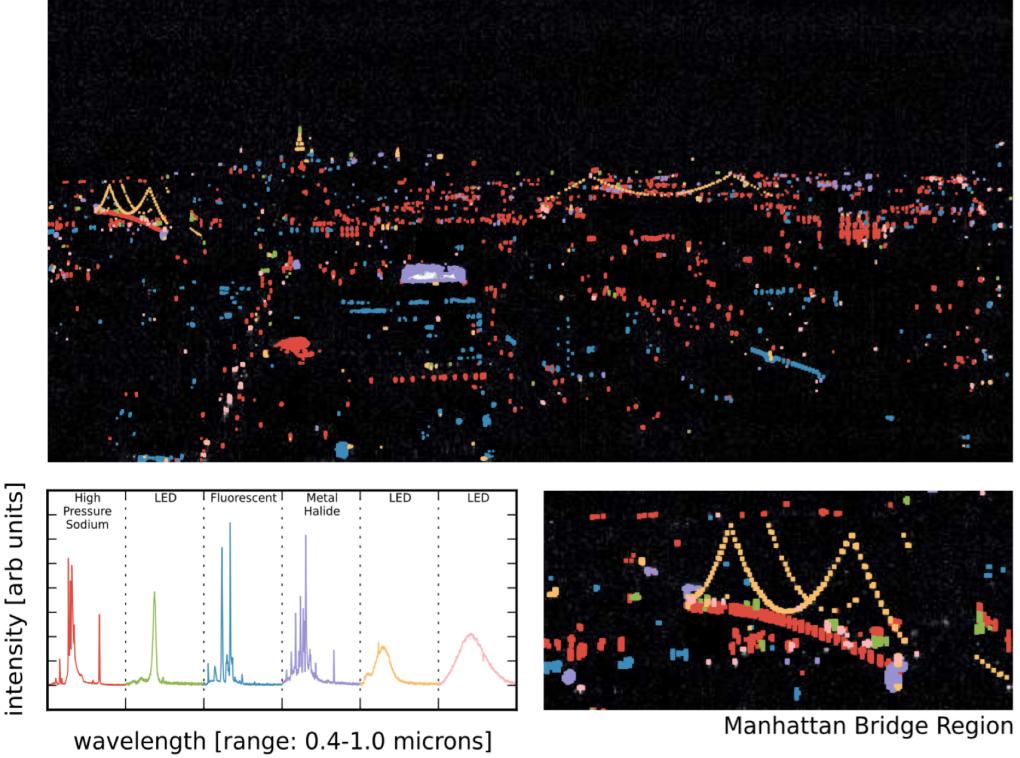
3. All analysis is aggregated and de-identified.
4. An individual source cannot be tracked across multiple modalities to infer PII.

In addition, it is UO policy that data and analyses are not shared with enforcement agencies (e.g., law enforcement, environmental regulatory enforcement, etc.).

### 3. Urban Science and Domains

The instrumentation and the associated operational modes described in §2 enable a wide variety of domain science inquiries related to the dynamics of urban systems from sub-second to yearly time-scales. Below, we describe several core aspects of city functioning that we are exploring through the application of computer vision, image processing, machine learning, and astronomical analysis techniques to the UO data streams. These lines of study can be largely grouped into three categories: Energy, Environment, and Human Factors, topical areas that are closely associated with the fundamental components of urban systems. The science focus of the UO is to develop a deeper understanding of cities through

### New York City Lighting Technologies



**Figure 3.** A night time hyperspectral scan of the scene in Figure 1 (reproduced from [2]). The *top* panel shows each source in the scene color coded by the lighting type determined from the spectra (*lower left*). The *lower right* panel shows a zoomed in view of the Manhattan Bridge region with LED necklace lights, High Pressure Sodium street lights, LED car headlights, and Metal Halide spotlights.

observational determination of the dynamical interplay between these.

We note that segmentation of each image into individual buildings (i.e., geolocating each pixel in the image) is an essential first step in all of the topics described below. Furthermore, this “building ID” step allows us to fuse imaging data from the UO with publicly available records data which can enable lines of inquiry that are otherwise impossible. The details of our methods for image segmentation and building ID are given in Appendix B.

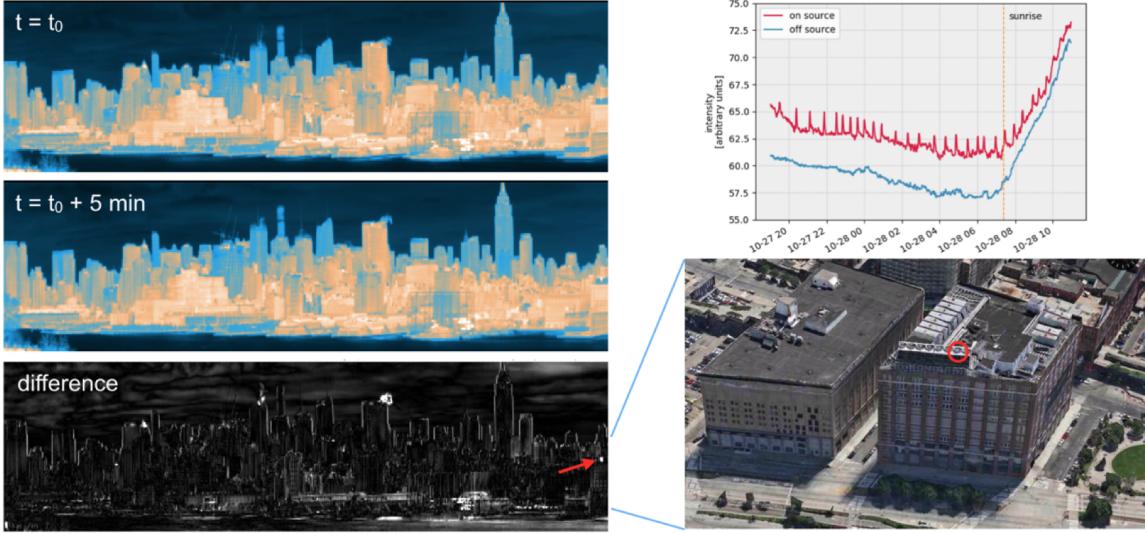
#### 3.1 Energy

Urban energy use – including electricity, natural gas, heating oil, etc. – drives city functioning and is at the heart of urban science questions related to resource consumption and

distribution, infrastructure performance and maintenance, and future urban planning. Furthermore, it serves as the source of much of the environmental impacts of cities. The UO’s imaging capabilities with their multiple spatial and temporal scales allow for numerous unique ways of quantifying urban energy use characteristics from remote vantage points.

##### 3.1.1 Remote Energy Monitoring

Lighting variability in the cities can be observed via broadband visible imaging at 0.1 Hz (10 seconds between images). In [1] we showed that aggregated lighting variability in NYC displays a diurnal pattern of use (on/off behavior) that is roughly repeating night-to-night. This variability encodes occupancy characteristics that serve as strong correlates to



**Figure 4.** *Left:* two IR images separated by 5 minutes and the difference between the two. The difference image shows both steam plumes as well HVAC vents (*lower right*) that heat up and cool down with building energy use. *Top right:* the time-dependent signature of a similar source (in red) compared to the building facade (in blue) over one night shows the cadence of the building’s heating/cooling system.

total energy consumption. We find that using these imaging data as an input to a convolutional neural network (CNN) trained on ground truth consumption data results in a 10-15% improvement on network-level electricity consumption prediction models trained on temperature and humidity data alone.

### 3.1.2 Lighting Technologies and End-Use

The VNIR hyperspectral instrumentation described in §2.3 has sufficient spectral resolution and photometric sensitivity to identify emission lines in nighttime lighting at a distance of several miles. In [2,4] we showed that this information can be used to determine lighting type of individual sources using spectra measured in the lab and developed Template Activated Partition (TAP) clustering to identify lighting types that were not previously cataloged. An example of the technique is shown in Figure 3.

Lighting technology identification has several important energy use cases including identification of targets for energy efficient

lighting upgrades by practitioners, empirical determination of end-use characteristics by consumers, and rates of adoption of modern lighting technologies by developing cities.

#### 3.1.2 Grid Stability and Phase

The 60hz frequency of the AC mains is reflected in a sinusoidal oscillation of flux in the lights connected to the mains for a subset of lighting technologies (e.g. incandescent, traditionally ballasted halogen, fluorescent). In [5] we showed that observing city lights at high temporal frequencies allows us to monitor the grid frequency at city scale. However, since persistent observations at frequencies of 100s of Hz is not feasible (e.g., due to data transfer and storage limitations), in [5] we used a chopper wheel operating at  $\sim$ 119 Hz to generate an  $\sim$ 1 Hz beat frequency that is observable with cadences of  $\sim$ several Hz. As we described, this capability allows for monitoring of the stability of the grid across an urban distribution network, detection of relative phase of sources, and (with sufficiently accurate observations) potentially also

transient phase shifts due to load at the unit level.

### 3.1.3 Building Thermography at Scale

In Figure 4 we demonstrate the use of time-dependent broadband infrared imaging to study thermographic envelopes of buildings in the city skyline. In particular, not only are efficiency characteristics such as heat leaks and thermal couplings detectable [6], but the figure shows that individual HVAC vent duty cycles can also be seen at a distance.

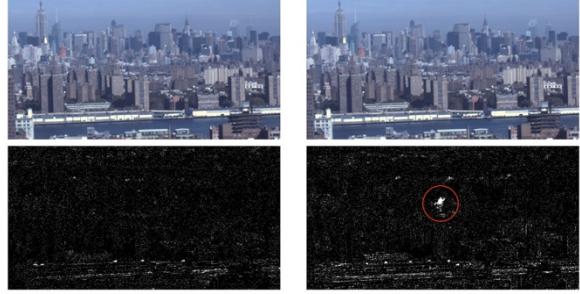
As with the broadband visible wavelength imaging, this type of source variability can serve as input to energy consumption models trained on consumption data. In addition, BMS operations and heating/cooling efficiency can be measured at scale across thousands of buildings from a single IR sensor.

## 3.2 Environment

Energy use in cities, and urban metabolism more broadly, results in byproducts of that use that have significant environmental impact. These impacts have local effects (for example, correlation of air quality with breathing-related health outcomes), regional effects, and – due to the physical and population size of large cities – global effects. The UO instrumentation extends traditional remote sensing of environment by satellites to increased spatial and temporal resolution (albeit with decreased geo-spatial coverage) to allow for the study of *dynamical* detection of environmental impacts of cities on sub-minute timescales.

### 3.2.1 Soot Plumes and Steam Venting

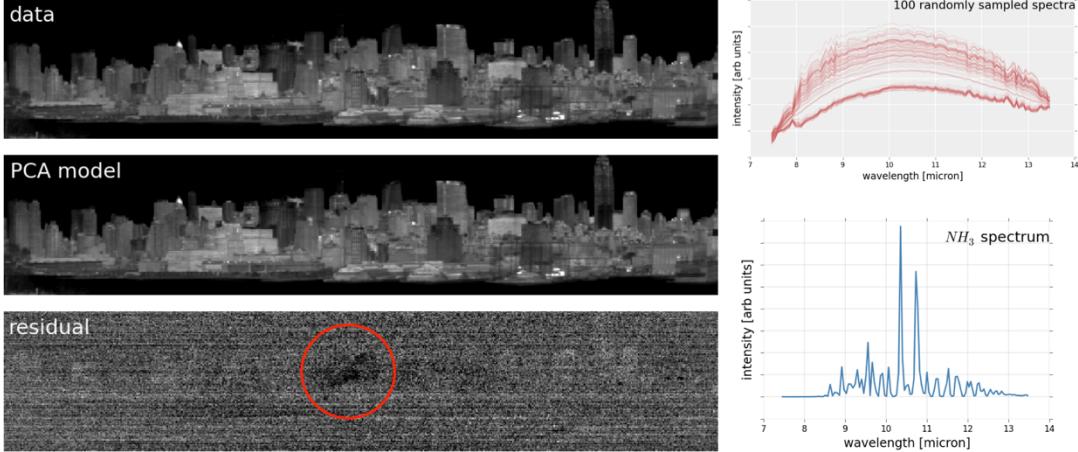
As buildings in urban environments burn oil for heat, they produce soot plumes that are dispersed through the ambient air as they are advected away from the source by local winds. Such plumes are responsible for ~75% of greenhouse gas production in NYC, and are detectable in UO visible wavelength observations at cadences of 0.1 Hz. Figure 5



**Figure 5.** Top panels: Two broadband visible daytime images separated by one minute. Bottom panels: the same two images, but with the time-independent background removed. The application of foreground/background separation techniques clearly reveals a soot plume (circled in red) that has been ejected from one of the buildings in the scene.

shows an example of one such detection in which a plume appears as a transient, dark feature in the raw imaging, although the very low surface brightness of the feature makes it extremely difficult to detect directly from these raw data. However, foreground/background separation techniques reveal the plume clearly towards the center of the image. In subsequent images, the plume is blown to the right as it disperses. In addition to soot plumes, the venting of steam from building heating and cooling systems is also visible.

These plumes and smoke stacks have significant potential for not only detecting and tracking the total number of plumes produced and the effects on local air quality, but their motion can also be used as tracers of urban winds, informing studies of flows through complex urban terrains (including simulations of such effects). The plume in Figure 5 is a particularly striking example, however most plumes are quite difficult to detect through classical computer vision techniques due to the complex urban background, time-dependent shadowing, low source brightness, and amorphous nature of the plumes. We have recently found that applications of regional convolutional neural networks (R-CNN) can be tuned to detect such plumes and we have



**Figure 6.** Left: A Principle Component Analysis (PCA) model fit to an Long Wave Infrared LWIR hyperspectral image. The PCA model (middle) was reconstructed from the first 10 principle spectral components of spectra in the scene (*top right*) and the model at 10.3 micron was subtracted from the raw data (top) to produce the 10.3 micron residual in the *bottom left* panel. The result is a clear, extended Ammonium plume, which has a 10.3 micron absorption feature as shown in the *lower right*.

developed end-to-end tracking systems for application to these type of data [7].

### 3.2.2 Remote Speciation of Pollution Plumes

A variety of molecular compounds have strong absorption and emission features in the 7.5–13.5 micron range of the LWIR instrument described in §2.4. Over the course of that observational campaign we showed in [3] that numerous molecular compounds could be identified in plumes emitted from buildings along the NYC skyline including Ammonia, Methane, Acetone, Freon-22, CO<sub>2</sub>, etc. In Figure 6, we show a simple detection of an Ammonia plume using a Principle Component Analysis (PCA) decomposition of a single data cube produced by the LWIR instrument. The various PCA components capture blackbody spectra, atmospheric (including water vapor) effects, and instrumental artifacts. The PCA model for each pixel, when subtracted from the raw data reveals a deficit in the 10.35 micron residual that is due to the absorption line of Ammonia at that wavelength. This technology has significant applications for both environmental studies of cities as well as

emergency management and tracking of toxic materials release.

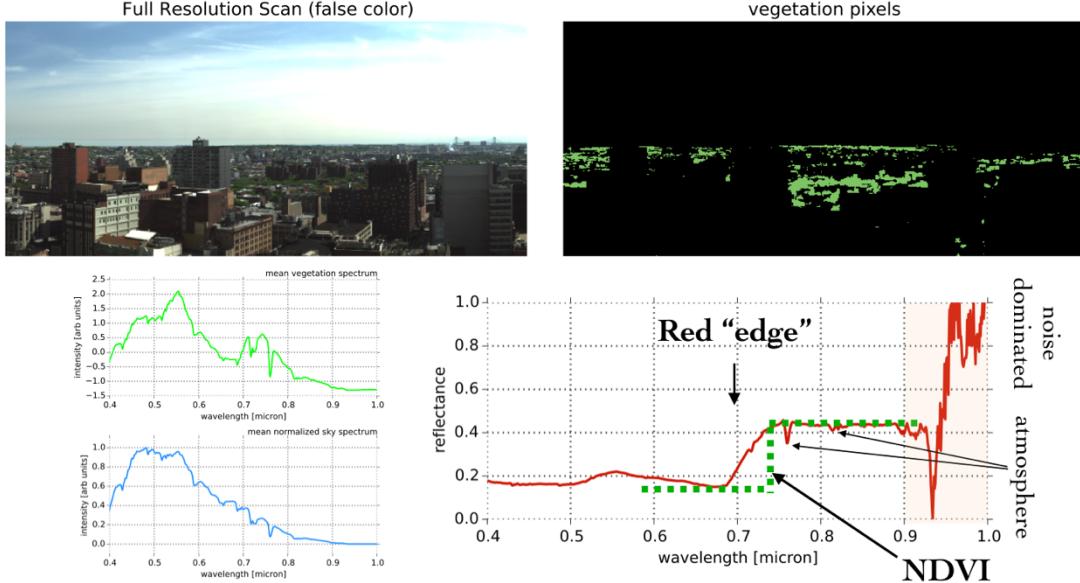
### 3.2.3 Urban Vegetative Health

The UO’s daytime VNIR observations have sufficient sensitivity to directly measure chlorophyll features in the reflectance spectra of urban vegetation. In particular, Figure 7 shows the characteristic “red edge” at ~700 nm that can be used to determine vegetative health indices such as the Normalized Difference Vegetation Index (NDVI).

Furthermore, because of the extremely high spectral resolution and sensitivity, as well as the persistent nature of the UO’s observational strategy, new metrics on vegetative health can be developed, low signal-to-noise effects such as solar-induced fluorescence can be measured, and short timescales response to local air quality can be determined to high precision.

### 3.2.3 Ecological Impacts of Light Pollution

It is well known that city lighting has detrimental impacts on local ecologies including affecting migratory avian behavior. These effects are both regional and highly local



**Figure 7.** *Top left:* a false color representation of a VNIR scan of Brooklyn, NY during the daytime. *Top right:* the vegetation pixels in the scene identified by clustering-based image segmentation. *Bottom left:* the mean vegetation spectrum and mean sky spectrum in the scene. *Bottom right:* the ratio of those two giving plant reflectance as a function of wavelength. The Normalized Difference Vegetation Index (and other vegetative health indices) is easily determined given the spectral resolution and sensitivity of our VNIR instrument.

(i.e., individual light installations). In collaboration with the New York City Audubon bird conservancy, we have deployed visible wavelength cameras acquiring images at 0.1 Hz to detect time-dependent lighting in lower Manhattan. By combining this data with regional NEXRAD radar scans we are able to measure the ecological impacts from urban lighting on migratory bird densities at scales of  $\sim$ 100s of meters and at time scales  $\sim$ minutes.

### 3.3 Human Factors

Urban functioning is fundamentally driven by human decision making. Infrastructure use, transportation choice and their management, economic transactions, etc., all have, at their core, a human element that drives observable patterns of activity. This micro-behavior aggregates to macro-scale patterns of life with diurnal, weekly, monthly, and annual rhythms that can be detected by deployed sensing platforms like the UO.

#### 3.3.1 Patterns of Lighting Activity and Circadian Phase

In [1] we showed that aggregate lighting activity in NYC displays clear diurnal patterns that repeat day-after-day and week-after-week. These patterns (on/off transitions of individual light sources) differ for residential versus commercial buildings and – as noted above – can serve as proxies for occupancy characteristics of buildings. In addition, these patterns for residential buildings correlate with circadian behavior of the population as shown in Figure 8. Variations of these patterns with ambient lighting intensity can quantify the effects of light pollution on public health.

Interestingly, in [1] we also showed that, while the aggregate patterns were strictly repeating – albeit with different behavior on weekends versus weekdays – a given source does *not* strictly repeat from one day to the next (nor from one day to one week later). This type of

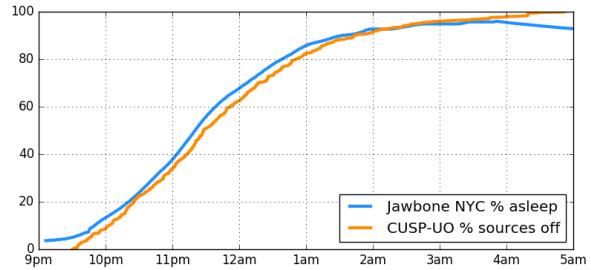
UO data directly address micro/macro behavioral properties of population and the scale of the transition between the two.

### 3.3.2 Technology Adoption and Rebound

Technological choice, and in particular the correlation of choice with socio-economic and demographic characteristics of the population, is an important indicator of population dynamics. In the energy sector, choice is an end-use characteristics studied most commonly by surveys of users. By combining the lighting patterns described in §3.3.1 with the lighting technology identification described in §3.1.2, our UO platform is ideally suited to not only quantify choice as a function of population characteristics via fusion (as described in Appendix B) of imaging data with census data, but also allows for an empirical measurement of the amplitude of the rebound effect in which the benefits of energy efficiency are offset by increased use stemming from a decreased economic incentive of curtailing use.

## 4. Conclusions

Modern cities are systems with tremendous complexity, potential for resource and environmental impact, and opportunities for growth. The dynamical interactions between the fundamental components of urban systems require measurement at high spatial and temporal resolution, with sufficient coverage to generate a representative sample of that system as a whole. We have presented an observational platform for the collection of data that can serve as inputs to machine learning, computer vision, image processing, and astronomical analysis techniques that extract information relevant to the functioning of cities. Our realization of that platform in New York City is the Urban Observatory facility consisting of imaging systems sensitive



**Figure 8.** An overlay of the percentage of sources “off” after 9:30pm on a single night (from UO broadband visible night time imaging of the scene in Figure 1) with the fraction of Jawbone™ users asleep in NYC on Monday March 31st 2014. There is a clear correspondence between our observations of lighting variability and the aggregate circadian patterns of NYC.

to visible and infrared wavelengths (both broadband and hyperspectral) with an operational mode that is persistent, synoptic, and granular. The urban science and domain topics that this data can address is broad ranging from energy use and its environmental impacts to patterns of life and public health. As the technology develops, deployment of similar Urban Observatories to cities of various sizes, local environments, and localities will enable a comparative study of diverse cities, deepening our core understanding of complex urban systems.

**Acknowledgements** GD’s and MSS’s work has been supported by a Complex Systems Scholar Award from the *James S. McDonnell Foundation* (JSMF). FBB, GD, JB, and MSS have been partially supported by Department of Energy (DOE) ARPA-E IDEAS grant. Imaging deployments have been supported by the JSMF, DOE ARPA-E, and Leon Levy Foundation. We thank the New York City Audubon for their partnership and collaboration.

## References

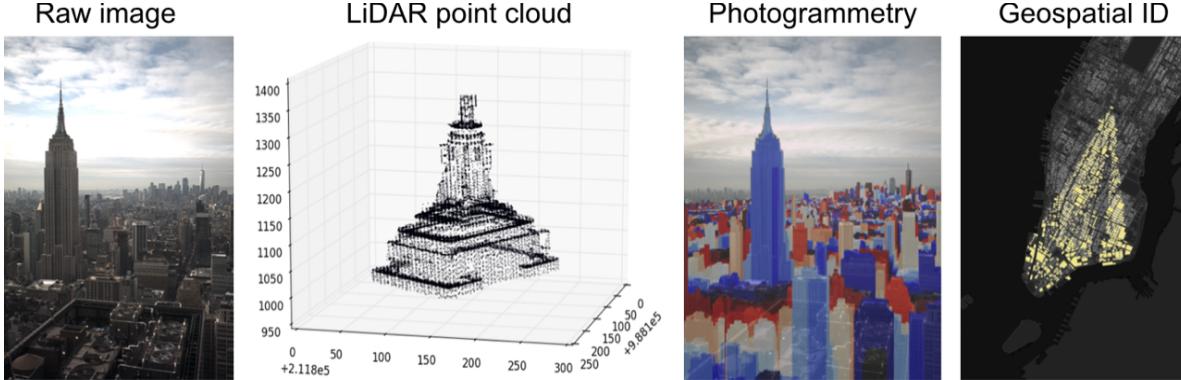
1. Dobler, G., Ghandehari, M., Koonin, S.E., Nazari, R., Patrinos, A., Sharma, M.S., Tafvizi, A., Vo, H.T. and Wurtele, J.S., 2015. Dynamics of the urban lightscape. *Information Systems*, 54, pp.115-126.
2. Dobler, G., Ghandehari, M., Koonin, S. and Sharma, M., 2016. A hyperspectral survey of New York City lighting technology. *Sensors*, 16(12), p.2047.
3. Ghandehari, M., Aghamohamadnia, M., Dobler, G., Karpf, A., Buckland, K., Qian, J. and Koonin, S., 2017. Mapping refrigerant gases in the new york city skyline. *Scientific reports*, 7(1), p.2735.
4. Baur, J., Dobler, G., Bianco, F., Sharma, M. and Karpf, A., 2018, November. Persistent Hyperspectral Observations of the Urban Lightscape. In 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP) (pp. 983-987). IEEE.
5. Bianco, F.B., Koonin, S.E., Mydlarz, C. and Sharma, M.S., 2016, November. Hypertemporal imaging of NYC grid dynamics: Short paper. In Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments (pp. 61-64). ACM.
6. Ghandehari, M., Emig, T. and Aghamohamadnia, M., 2018. Surface temperatures in New York City: Geospatial data enables the accurate prediction of radiative heat transfer. *Scientific reports*, 8(1), p.2224.
7. Li, C., Dobler, G., Feng, X. and Wang, Y., 2019. TrackNet: Simultaneous Object Detection and Tracking and Its Application in Traffic Video Analysis. *arXiv preprint arXiv:1902.01466*.

## Appendix A: Backend Infrastructure

The deployed instruments and their associated operational modes described in §2 require a flexible and robust backend computing infrastructure to collect, store, and analyze the large volume of data generated by the UO. This infrastructure consists of both hardware and software components that, taken together, operate continuously to yield persistent observations. Our backend infrastructure consists of the following core components with associated functionality.

*Camera control devices* – Each imaging device is equipped with a mini-computer that opens a direct link with the camera itself. This machine is tasked with communicating directly with the camera and issues image acquisition commands. In certain instances, this computer can also be used to perform edge computations including compression or sub-sampling of the data. Acquired data may be saved temporarily on disk on this machine for buffered transfer back through the gateway server, or be written directly to bulk data storage.

*Gateway server* – The main communications hub between our computing and data storage servers and the deployed instrumentation is a gateway server that issues scheduled commands to the edge mini computers. This hub is also responsible for the pull (from the deployment) and push (to the bulk data storage) functionality for the data acquisition as well as the gateway for remote connections of UO users to interact with our databases.



**Figure B1.** *Left:* a broadband visible image from a UO deployment in midtown Manhattan. *Center-left:* a LiDAR point cloud of the top of Empire State Building. Photogrammetric techniques allow us to use LiDAR to identify the buildings observed by each pixel (*center-right*) given the building tax-lot footprints from a publicly available data base (*right*). With these building IDs, we can integrate geospatial information with image data.

*Bulk data storage* – At full operational capacity, a UO site (consisting of a broadband visible camera operating at 0.1 Hz, broadband infrared camera operating at 0.1 Hz, a DSLR operating in video mode, and a VNIR hyperspectral camera operating at  $10^{-3}$  Hz) acquires roughly 2–3 TB per day. This data rate necessitates not only careful data buffering and transfer protocols to minimize packet loss from the remote devices, but also a large bulk data storage with an appropriate catalog for the imaging data. This ~PB-scale storage server is directly connected to our computing servers for computational speed. This storage server also hosts parallel source catalogs that store information extracted from the data.

*Computing server* – Our main computing server that is used to process UO data consists of a dedicated >100 core machine that is primarily tasked with processing pipelines including: registration, image correction, source extraction, etc. We have designed our own custom containerized interface that seamlessly allows UO users to interact with the data while background data processing and cataloging tasks operate continuously.

*GPU mini-cluster* – Several of the data processing tasks described in §3 require the building and training of machine learning models with large numbers of parameters including convolutional neural networks. For these tasks, we use a GPU mini-cluster that is directly connected to our main computing server and which is continuously fed streaming input data from which objects and object features are extracted.

## Appendix B: Data Fusion

To maximize the utility of our UO facility, it is important that we be able to integrate the massive imaging data sets that we generate with available geo-spatial data including publicly available data such as census data, building-level energy consumption, fuel types, external sensor data (e.g., air quality sensors or radar data), etc.

Our data fusion utilizes publicly available LiDAR data from NYC to locate the 3-space coordinate that is covered by each pixel. Namely, using the collinearity equations, we project the topographic LiDAR point cloud into the 2D image plane (given the position, pitch, yaw, roll, and focus of the

camera) and for each pixel, choose the closest LiDAR point. Since the LiDAR resolution is  $\sim$ 1 foot, there are pixel lines-of-sight for which the nearest LiDAR point is behind the nearest surface in that direction, and so we make a “no overhang” approximation and assign a given pixel the same 3-space coordinate as the pixel above it if the pixel above it is found to be closer to the camera. Finally, we use the publicly available MapPLUTO (Primary Land Use Tax-lot Output) data that contains the geospatial footprints for each building to associate the x,y components of the 3-space coordinate of a pixel with a given building footprint. Thus we are able to tag each pixel as “observing” a given building (see Figure B1). Additional geo-spatial data at coarser resolution (e.g., census tracts, utility distribution zones, etc.) can be associated as well.