Dear Anonymous Reviewer,

Thank you for taking the time to review our manuscript and for your constructive comments and suggestions. You raise some good points that needed to be addressed. We appreciate the time you took to provide this thorough evaluation of our methods and their weaknesses. Please see our responses to your three major areas for improvement below.

1. Throughout the paper, it is unclear what is the basic requirement for a high-quality nighttime aerosol retrieval using the available sample of stable point source data.

    a. While this work is described by co-authors (J. Zhang), it is still unclear what is the minimum required sample to establish an aerosol retrieval for each specific study domain. Would a single pixel-based source be sufficient to estimate nighttime AOD over the regions of interest used (256 x 256 pixels)? Are more stable points needed? If so, what would be needed, sampling-wise, in order to ensure a routine retrieval of aerosols using this method?

        Response:
        We have added a new paragraph on Page 3 describing the requirements of aerosol retrieval algorithms that rely on city lights as their source of visible light. The paragraph also clarifies that the results of this study are not intended to be specific to one algorithm and should be useful for any algorithm that uses city lights as their source of visible light. The new paragraph reads:

        "This analysis provides a first step towards characterization of anthropogenic light sources for use in retrievals and is intended to be agnostic of the retrieval algorithm. Some algorithms may be capable of performing retrievals using attenuation of light emitted by isolated point sources. Others may rely on the amount of "blooming" observed around a light source to retrieve optical depth. It may also be possible to retrieve optical depth by observing changes in the brightness or spatial structure (e.g. spatial variance) of groups of well-characterized light sources. Regardless of the algorithm employed, understanding the variability in anthropogenic light emissions and its causes is important to the problem of retrieving optical depth at night."

    b. Then, can you comment on how the methodology be scaled up at the region-level and over a sufficiently large global sample (urban-lit areas, only comprise less than 3% of the Earth's land mass). The authors should provide some commentary (and background/statistics/sampling estimates), which would then put the results of this paper into perspective.

        Response:
        Estimates of urban extent vary widely (Schneider et al., 2009). Given the estimate range, it might be reasonable to estimate that total urban extent is approximately 0.5% of Earth's land area (about 700,000 km^2). If all of the urban area were useable in optical

depth retrievals, a single VIIRS instrument, whose spatial resolution is 750 m, would be able to make about 1.2 million observations per night. If only 1% of the urban-lit area is usable the number would drop to about 12,000 observations per night. How much of the global urban area is useful for optical depth retrievals will be algorithm and application dependent as well as dependent on how well the brightness of each light source can be constrained.

2. The authors seem often confound basic terminology of reflectance/and nighttime remote sensing nomenclature, particularly when it comes to surface-related phenomena. In most cases, what they deem "Lunar Geometry" or "Surface Scattering" is formally defined as "Lunar-reflectance", Lunar BRDF, as well as surface albedo and reflectance anisotropy. That is, the authors described the surface phenomena as if it is only influenced by lunar phase and satellite/source view illumination angle of capture, when in reality there are additional factors that are intrinsic to surface conditions, which are virtually ignored.

   This lack of understanding stems from the author's push towards characterizing stable point sources of high-intensity radiances >20nW, which are well-above the magnitude of lunar variation (< 10nW). However, previous studies have documented how surface phenomena is in fact an influencing factor of stability of stable point sources. For instance:

   N. Levin, Q. Zhang, A global analysis of factors controlling VIIRS nighttime light levels from densely populated areas Remote Sens. Environ., 190 (2017), pp. 366-382, 10.1016/j.rse.2017.01.006

   M.O. Román, E.C. Stokes Holidays in lights: tracking cultural patterns in demand for energy services Earth's Futur., 3 (2015), pp. 182-205, 10.1002/2014EF000285

   Román, M.O.; Wang, Z.; Sun, Q.; Kalb, V.; Miller, S.D.; Molthan, A.; Schultz, L.; Belle, J.; Stokes, E.C.; Pandey, B.; et al. NASA's Black Marble nighttime lights product suite. Remote Sens. Environ. 2018, 210, 113–143.

   M.M. Bennett, L.C. Smith: Advances in using multitemporal night-time lights satellite imagery to detect, estimate, and monitor socioeconomic dynamics, Remote Sens. Environ., 192 (2017), pp. 176-197, 10.1016/j.rse.2017.01.005

   The most obvious effect not being considered is the influence of snow cover, which will most likely affect the RSD of Study Domains in temperate regions, given the 18-month sample size (Particularly St George UT). See for example Figure B1 in (Roman and Stokes, 2015), which shows how the presence of snow cover influences the stability of point sources for large US cities.

   Response:
   This is a good point that will need to be explored as this type of analysis is employed over larger regions. At least with respect to snow and NDVI, however, this is a relatively small

concern for the domains chosen. None of the study domains have significant snowfall in an average year. The study domain with the highest average annual snowfall is St George, UT which receives 1.4 inches per year and an average of 0.4 days per year with snowfall greater than 0.1 inches. Additionally, most of the study domains are desert regions and have little annual cycle in NDVI. The lone exception, San Francisco, does have an annual cycle in NDVI that may need to be considered in future studies.

We have added Section 7 which discusses some of the limitations of this study including this limitation.

3. The authors seem to make conclusions about RSD and its dependency to urban built environments based on simple inferences collected from manual examination of specific areas within cities (e.g., large/tall building located in the San Francisco region, residential sectors, large roads, etc., as noted in page 13, line 32.) The analysis needs to be strengthened by a more quantitative assessment based on a control variable that characterizes the urban built area (e.g., % Urban cover). A dependency between urban density and RSD should be explored in this paper to make a more definitive conclusion.

   The datasets are already available to come up with such an assessment. See for example, work by:

   T. Esch, W. Heldens, A. Hirne, M. Keil, M. Marconcini, A. Roth, J. Zeidler, S. Dech, E. Strano Breaking New Ground in Mapping Human Settlements From Space The Global Urban Footprint (2017)

   T. Esch, M. Marconcini, A. Felbier, A. Roth, W. Heldens, M.Huber, M. Schwinger, H. Taubenbock, A. Muller, S. DechUrban footprint processor-fully automated processing chain generating settlement masks from global data of the TanDEM-X mission, IEEE Geosci. Remote Sens. Lett., 10 (2013), pp. 1617-1621, 10.1109/LGRS.2013.2272953

   Also see Figure 12a in, Roman et al., 2018: https://www.sciencedirect.com/science/article/pii/S003442571830110X#bb0125

   Response:
   This is a good point that should be considered. We intend to explore this issue in future work. We have added language to the conclusions to discuss future work on this and other issues.