I respectfully disagree with most points but ESSD, as a journal, does not promote scientific contention or discussion. It publishes “original research data and data collections” for two purposes: a) make datasets a reliable, fully citable resource upon which other research can build; and b) reward data providers by establishing their priority and recognition. This review slights the data; I find nothing about data availability, organization, etc. Not obvious to this author that reviewer understands ESSD requirements? Not possible nor useful to make changes to manuscript text based on comments form this reviewer.

Review expresses appreciation of inexpensive environmentally-friendly sensors and deployments easily applicable in many regions; good. Review also seems to appreciate suggestions for scientific relevance; also good.

Most comments easily dispensed. Reviewer needs only to read carefully or watch example video. Wheel size, rotation, color etc. had no influence. Look at video. UV used to check visible data; read text following line 280. Supposed near-surface temperature profiles assume a base-point or starting point; this manuscript demonstrates that ‘base-points’ vary as a function of surface type, surface temp, gravel, vegetation, season. shading, etc; check example video. One of these sensors (TMP 117) often serves as primary reference; does one need to validate primary sensor against itself? Swiss and US sensors perfectly validated (see Figs 5, 6). Information about sensor manufacture and performance, rarely available for weather service sensors, included on freely-available data sheets. Sensor outputs compared directly to NOAA Climate Reference station data; only 3 exist in Montana. I use closest (25 km) station (Figure 7). Reviewer needs to assess Figure 11; first (only?) season-long inter-comparison between surface data and CRN data. If reviewer knows better free open-access source this author would like to see it. (Please note: ESSD does not allow CC-BY-NC contributions.)

Of course data only cover specific routes to and from a small city: one person on a bike. I make that point frequently and repeatedly throughout manuscript. I include vegetation and development (e.g. Demuzere 2022) categories, to facilitate comparisons with other studies. Multiple bikes carrying like sensors? Therein lies great promise. None of those cyclists will purchase standard Campbell $15k sensor packages (plus data access fees?) nor try to mount commercial sensors on a bicycle (several failed examples exist). But they might follow guidance provided here to purchase and mount $250 sensor packages of remarkable capability. As I contend in manuscript: bicycles should represent key component of future observing systems

Urban regions undoubtedly contain “roads/paved surfaces, but also include trees, grassland, buildings, and possibly rivers and lakes”. This dataset includes all of the above, geolocated at very high resolution, measured repeatedly over a full year’s worth of insolation, shading, snow cover, etc. So far, no other observations have covered more than a few warm days. What happens in deep shade, in daily- or seasonally-varying shade, when shade coincides with gravel, over grassland or on grassland under shade, or on paved surfaces under snow or ice? One can address these questions, accurately, with these sensors and this bicycle. Do we need more measurements in more regions under more conditions? Yes, as called for explicitly here! Those regions, if judiciously-selected, should allow multiple full-season measurements over pavement, gravel, shade, across all times of day. Until multiple cyclists accept the challenge, this represents a unique open-access single-cyclist survey.

Not sufficient to build “quantitative relationships with satellite remote sensing data”? This manuscript spends several paragraphs on plausible discrepancies between periodic (once or twice per month, only useful under cloud-free conditions) remote sensing data at (rarely) 10 m resolution versus these data at 4 m resolution under all weather, sky and environmental conditions. Review apparently missed those discussions? This author clearly asks a reverse

question: how will remote sensors validate their products against high-res ground truth? Challenges ahead! Based on open-access data sources! What about slight GPS mis-locations?

As clearly mentioned, cooler surface temperatures of gravel surfaces long recognized. But, based on what data? Includes shade? Snow? Real world data, or generated artificially in a laboratory? Year-long or only on warmest days? Reviewer knows free open-access source for temperatures of pervious versus impervious surfaces under all conditions? If researchers want to explore UHI factors using an atlas of building- and vegetation-influenced surface temperatures, this data might make a worthy contribution? Reviewer entirely silent on this aspect. No comments on GPS registration issues, on daily patterns if shading, on snow?