**Referee: 1**

General comments:

This manuscript aims to quantify the impact of changes in green spaces on human health in global cities over a ten-year period. The authors use NDVI, aggregated into five-year intervals, to assess city-level changes, which serve as input for an exposure-response function linking NDVI to all-cause mortality. While no major global trends in NDVI or associated health impacts are observed, significant changes are noted in specific cities and years. The paper is well-written and addresses a topic relevant to ERL readers. However, the following issues need to be addressed, as outlined below:

Thank you for your careful read of our paper and for your thoughtful suggestions.

Specific comments:

Health impact assessment

1. It would be helpful to explicitly discuss how increased greenspaces positively impact human health. Which specific diseases see a reduced risk, and through what mechanisms? For example, line 100 mentions that NDVI over a five-year period is used in the health assessment, implying that the exposure-response functions focus on short-term health impacts. Clarifying the pathways through which greenspaces influence health outcomes, along with their respective time scales, would strengthen the analysis. Additionally, how are these effects isolated from other contributing factors? For instance, increased urban greenspaces can lower near-surface temperatures, potentially mitigating extreme events like heat waves. Given that heat-related health impacts can be quantified separately, how do they compare to and differ from the health benefits attributed to NDVI? More importantly, how can these influences be disentangled in the analysis?

Thank you for your comment. In response to the first part of your comment, we’ve added text to the introduction:

1. Since the analysis focuses on the peak season, the contribution of individual years may be highly relevant. Wouldn't it be more appropriate to analyze individual years rather than aggregated periods? Additionally, at what scales have the epidemiological functions been developed?
2. Starting from line 413, the health impact assessment appears to consider different population groups. However, there is no explanation or equation demonstrating how this dependence is incorporated into the exposure-response function. More details should be provided in the methods section.

I don’t really understand this comment. Is this about Brochu et al. looking at 65+?

Methods

1. Line 149 and later: The paper attributes the large interannual variability in NDVI to meteorological factors. If this is the case, to what extent can NDVI reliably represent land-use changes and be used to infer health impacts? Would it be possible to quantify how much NDVI variability is driven by temperature fluctuations versus land-use changes due to urbanization?

Right now we’re just saying this is how health has been impacted by any changes and not inferring (beyond saying it’s a mix of things) how/why those changes have occurred. I’m not sure how to really disentangle climate change, weather, and land-use changes very reliably?

1. Existing literature or simple correlation analyses could help address this question. Additionally, could a land-use dataset that includes both land-use type and urban fraction help identify whether surface property changes have occurred? More evidence is needed to explain and attribute the observed interannual variability in NDVI.

Will look at literature. What about looking at %urban for each year correlated with NDVI? There are so many non-urban categories that I’m not sure how to make the comparison suggested about tracking land changes over time that way

1. Over the 10-year period analyzed, many cities have undergone changes in size. How is this accounted for? If city boundary changes are measured, how are they incorporated into the health impact assessment? It would be helpful to interpret the observed changes for specific cities in relation to their growth, urbanization levels, and other relevant factors. Additionally, is there variability within cities at the sub-city level? If so, how is it considered when computing an aggregate metric for individual cities?

Thank you for your points. We have used the GHS-SMOD shapefiles do define city boundaries, which are based on built-up area and population density from 2019. We use the same urban boundary for all the included years. You are correct that urbanization could be another factor that is driving the changes we observed in city-level NDVI over time. We have added text to the discussion section to reflect your important point:

“Urbanization in the past decade could also contribute to these changes, as we used a consistent urban boundary definition across the ten-year period, however cities may have grown and morphed over this time.”

To your last point, yes there is variability within cities in NDVI and population. We performed the health impact assessment at the 100m x 100m grid cell level to account for this. Population data were only updated every 5-years, so we have used 2015 gridded population data for years 2014-2018 and 2020 data for years 2019-2024.

Technical comments

* 1. Line 33: why is NDVI higher and more stable in European and North American cities? The sentence continues by talking about epidemiological studies performed in those areas, however the two aspects are unrelated. Please clarify.
  2. Line 62: is this true everywhere? I believe cities in the developing world have very different emission regimes than the ones in developed countries. Please revise.
  3. Line 145: what is “i” in Equation 2? Does it refer to the pixels? How is the HR defined? An explicit expression for that should be provided. Also, it is not clear which spatial resolution NDVI and population data have and how they are homogenized in the present analysis.
  4. Figure 3: what do the colored dots represent and the box plots? More details are needed in the caption.
  5. Line 309: is the IQR: 0.13 – 8.5?
  6. Line 319: this section is confusing. What does a median change in mortality of 0.01 fewer means? There is mention of a change in NDVI in that sentence, but how much change? Also, it would be useful to contextualize changes in NDVI due to various factors. E.g. the mentioned 0.19 change in NDVI how can that be interpreted? Is it typical of a greener season due to more favorable weather conditions or is it the change expected when land use changes from urban/concrete to forest? Some references are needed through the text.
  7. Line 326: the ranges expressed as “fewer to more” are quite confusing, so it would be clearer if there was a sentence explaining how this wording will be used and interpreted
  8. Figure 5: remove “Associated”.
  9. Line 385: the impact on individual cities would be very relevant to be discussed and possibly compared with cities in the same region that do not experience such changes.
  10. Line 419: most of the paper mentioned that NDVI had remained stable over the period analyzed. However, the authors mention that NDVI has decreased over time. This seems to contradict what discussed before.
  11. Conclusions are too short. I suggest merging with discussion.

**Referee: 2**  
  
COMMENTS TO THE AUTHOR(S)  
In this study, the authors attempt to quantify the contribution of NDVI to mortality in global cities. The attempt is ambitious and interesting, and I appreciate the authors' efforts. However, the validity of the methods and assumptions used in this study are questionable, and thus, the results and conclusions are not convincing enough. I do not recommend that this article be published in ERL.

Thank you for taking the time to read and review our study.  
  
In this study, temporal changes in NDVI (i.e., the difference between 2014-2018 and 2019-2023) are translated to the differences in mortality. Temporal changes should be discriminated against spatial variations. In the original cohort studies on which the meta-analysis (Rojas-Rueda et al., 2019) is based, in principle, the risks are calculated from the spatial variation of the NDVI. It is also indicated that the temporal change in NDVI is not associated with mortality by a cited study (Ji et al., 2019). There is no justification for attributing the mortality to temporal changes in NDVI, as assumed in this study.

Thank you for your comment. The exposure-response function provided by the Rojas-Rueda et al. meta-analysis does not specify an exposure time scale. The studies included measure NDVI in various ways including contemporaneous and cumulative NDVI exposure. The (Ji et al, 2019) study is the only study to assess changes over time and did not find a significant association with mortality using this exposure definition. However, they defined changes in NDVI by grouping areas into significant increases, no significant change, and significant decreases based on the slope coefficient from a linear regression of annual average NDVI from 2000-2014. We observed substantial inter-annual changes in NDVI that jumped up and down across the years. We averaged two 5-year periods to minimize some of this meteorological noise and attempt to capture changes due to urbanization, climate change, and greenspace interventions that will impact residents’ current and future greenspace exposure.

To address your comment, we’ve added text discussing the temporal aspect of the exposure measurement to the introduction section:

We’ve also added text to the discussion section to be explicit about the limitations of our approach with respect to the temporal nature of the exposure measurment:

The authors apply one single risk function to all the cities all over the world. This risk function is based on the nine studies, most of which are conducted in developed countries in temperate climate zones. The derived relationship might be applicable to similar cities. For example, however, in developing countries in tropical climate zones, higher NDVI could be associated with higher exposure to vector-borne diseases transmitted by mosquitoes (e.g., malaria). Therefore, applying the relationship to global cities is questionable and probably inadequate.

We agree that the use of one global exposure-response function is a limitation of this study. However, the current evidence base linking greenspace and all-cause mortality does not support a city-specific approach. To address this, we chose a large-scale meta-analysis to be as generalizable as possible. While ideally, we would have city-, age-, gender-, and socioeconomic-specific risk curves, the approach we use here reflects the current state of the science and is commonly used in multi-city and multi-country health impact assessments. Many studies quantifying the health impacts of global environmental exposures use a generalized exposure-response function. For example, the Global Burden of Disease study uses a global relative risk when quantifying the associated health burden of ambient air pollution including particulate matter1 and nitrogen dioxide.2 Additionally, many health impact assessments from the past five years of greenspace3-10, air pollution8-11, noise8-9, and heat9-10 have relied on a single exposure-response functions.

References:

1. Ambient particulate matter pollution—Level 4 risk. <https://www.thelancet.com/pb-assets/Lancet/gbd/summaries/risks/ambient-particulate-matter-pollution.pdf>
2. https://www.healthdata.org/sites/default/files/methods\_appendices/2021/wozniak\_air\_no2\_writeup\_GBD2021\_030222\_AC\_updated030624.pdf
3. Brochu P, Jimenez MP, James P, Kinney PL, Lane K. Benefits of Increasing Greenness on All-Cause Mortality in the Largest Metropolitan Areas of the United States Within the Past Two Decades. Front Public Health. 2022 May 10;10:841936. doi: 10.3389/fpubh.2022.841936. PMID: 35619828; PMCID: PMC9127575.
4. Barboza EP, Cirach M, Khomenko S, Iungman T, Mueller N, Barrera-Gómez J, et al. Green space and mortality in European cities: a health impact assessment study. The Lancet Planetary Health. 2021 Oct;5(10):e718–30.
5. Garber MD, Guidi M, Bousselot J, Benmarhnia T, Dean D, Rojas-Rueda D. Impact of native-plants policy scenarios on premature mortality in Denver: A quantitative health impact assessment. Environment International. 2023 Aug;178:108050.
6. Dean D, Garber MD, Anderson GB, Rojas-Rueda D. Health implications of urban tree canopy policy scenarios in Denver and Phoenix: A quantitative health impact assessment. Environmental Research. 2024 Jan 15;241:117610.
7. Kondo MC, Mueller N, Locke DH, Roman LA, Rojas-Rueda D, Schinasi LH, et al. Health impact assessment of Philadelphia’s 2025 tree canopy cover goals. The Lancet Planetary Health. 2020 Apr 1;4(4):e149–57.
8. Pereira Barboza E, Montana F, Cirach M, Iungman T, Khomenko S, Gallagher J, et al. Environmental health impacts and inequalities in green space and air pollution in six medium-sized European cities. Environmental Research. 2023 Nov;237:116891.
9. Mueller N, Rojas-Rueda D, Khreis H, Cirach M, Andrés D, Ballester J, et al. Changing the urban design of cities for health: The superblock model. Environment International. 2020 Jan;134:105132.
10. Pereira Barboza E, Nieuwenhuijsen M, Ambròs A, Sá THD, Mueller N. The impact of urban environmental exposures on health: An assessment of the attributable mortality burden in Sao Paulo city, Brazil. Science of The Total Environment. 2022 Jul;831:154836.
11. Southerland VA, Brauer M, Mohegh A, Hammer MS, Van Donkelaar A, Martin RV, et al. Global urban temporal trends in fine particulate matter (PM2·5) and attributable health burdens: estimates from global datasets. The Lancet Planetary Health. 2022 Feb;6(2):e139–46.
12. Malashock DA, Delang MN, Becker JS, Serre ML, West JJ, Chang KL, et al. Global trends in ozone concentration and attributable mortality for urban, peri-urban, and rural areas between 2000 and 2019: a modelling study. The Lancet Planetary Health. 2022 Dec;6(12):e958–67.

**Referee: 3**  
  
COMMENTS TO THE AUTHOR(S)  
This paper performs a health impact assessment of changes in urban NDVI in more than 1000 cities across the globe, and finds on average mild changes although with important variability between cities.  
  
It is clearly of interest to provide an idea of to which extent green space impact public health and the global scope is an additional strength of this study. The method is also appropriate overall although I think it comes with important uncertainty that is not really acknowledged (more on that in the below comments). It is not very clear to me however why these specific counterfactuals have been used (2014-18 vs 2019-23) instead of more usual counterfactual such as a fixed increase or some policy objective.

Thank you for taking the time to review our study and for your thoughtful feedback.

## Comments  
1. It seems that the exposure used here differs slightly from the studies estimating the HR, and I wonder to which extent this could impact the mortality estimates.  
- Has the exclusion of water pixels done in the epidemiological studies? It is mentioned in the introduction that blue spaces have alleged benefits on health, and that negative NDVIs tend to represent water. One could then imagine a kind of V-shaped relationship between NDVI and mortality, that could perhaps impact the estimation of HR. Maybe removing water pixels would change the potential range of NDVI here and therefore represent a slightly different exposure?

Thank you for your comment. Typically, water pixels have been removed from analyses using NDVI.

- I cannot find any mention of "greenest season" NDVI in the meta-analysis of Rojas-Rueda et al. (2019) and was wondering to which extent it could change the picture here.

The nine studies included in the meta-analysis use a range of different NDVI-based exposure metrics.   
2. I would say the main shortcoming of this analysis is the lack of any uncertainty estimate (save for the grand mean). There seems to be a lot heterogeneity in the reported HRs in Rojas-Rueda et al. (2019) which means high variability of the overall one. Additionally, I can imagine there is some uncertainty related to exposure assessment, notably because of cloud cover. I think confidence intervals (or any measure of uncertainty) should be included in the analysis for the reader to get a sense of how uncertain the whole exercise here (and high uncertainty is not a flaw of any analysis). Note that it is quite easy to sample from the distribution of the HR and compute CIs from there.

Do you think we should switch to this approach? The way I have it, the CIs provided are just using the lower and upper bounds of the HR estimate.

3. Why restrict the analysis to cities larger than 500,000 inhabitants (or the biggest one of the country)? The GHS-UCDB is much more comprehensive than that and the NDVI is a global product so, unless I am missing something, nothing prevents from expanding the number of cities covered.

You are correct that the GHS-UCDB contains many more cities (~10,000 cities). We restricted to cities in this way to reflect the universe of cities included in The Lancet Countdown. We have added language to be clearer about how and why the study population was chosen in the methods section:

4. It would be a nice addition to provide absolute figures in the results, i.e. in this set of cities, how many deaths have been lost/averted due to NDVI changes? At the moment only death rates are reported. Note that Eq. 3 promises attributable numbers to the reader.

Thank you for your comment. We do provide absolute numbers in the appendix,

Make fig 5 a 4 panel figure with absolute results?  
  
## Minor  
- 112: "Joint Research Commission" is actually the "Joint Research Centre"  
- I think the title of the paper is slightly misleading and should indicate that the paper looks at changes in NDVI. "A quantitative health impact assessment" promises a more expanded assessment than what is done in the paper.  
- The figures are a bit difficult to read and I suspect they would be difficult to analyse for someone with colour-blindness. I would suggest thinner and fainter lines for border in Figure 2 and boxplots in figure 3, with perhaps slightly enhanced points.  
- Similarly, I would suggest reordering the regions in Figure 3, so that regions from the same continent are next to each other. Another ordering could be by "average latitude" or something like that.