Questions & Answers Part 3

Please type your questions in the Question Box. We will try our best to get to all your questions. If we don't, feel free to email Sean McCartney (sean.mccartney@nasa.gov), Kathryn Conlon (kcconlon@ucdavis.edu) or Evan Mallen (emallen3@gatech.edu).

Question 1: How is surface reflectance different from the top-of-atmosphere data, and which is more applicable to use in representing exposure? In terms of LST, SR temperature values are relatively high compared to TOA.
Answer 1: Any time we conduct urban-scale heat risk analysis or heat vulnerability indices, I recommend using data that is as close as possible to the population of interest. In this case, we can use a product like the Landsat Analysis Ready Dataset (ARD) land surface temperature which is already processed for you. Ideally, we would use a measure of air temperature, but this is rarely available on a scope and scale necessary for such analyses. Most often these are derived either through field measurements like a distributed sensor network, or a regional climate model. Both of these options are more city or project-specific and are created on-demand, which substantially limits the availability of this data.

Question 2: There are two Bands For LST calculation, Band 10 & Band11 . What is the difference in the final results of the two bands?
Answer 2: Band 10 and Band 11 are both within the thermal infrared range of wavelengths, just with slightly differing ranges of wavelengths. I more often see B10 used to derive the land surface temperature in the Landsat ARD dataset. Both are viable sources for LST given that they are in the thermal infrared range. Refer to the presentation.

Question 3: Does the Census's Bureau differential privacy process impact the American Community Survey (ACS) datasets or just the 2020 decennial data below the state level?
Answer 3: The differential privacy process existed in the 2010 dataset as well, but has been modified for 2020. The documentation I have seen states that these new differential privacy measures will apply to the ACS starting in 2025. In general, I recommend using the ACS over the decennial census, which comes in 1, 3, and 5-year increments, each containing more data collected over a greater time period, which can

reduce your uncertainty. However, for HVI purposes, the uncertainty is rarely taken into account, using only the values as reported for comparative purposes across the geography of interest.

Question 4: Given that the Landsat Collection 2 ARD dataset is available for the US only on the USGS Earth Explorer site, is there a published workflow for generating this data for other territories?
Answer 4:
"Global Landsat ARD specifications are in development. Information will be provided when details are available."
[https://www.usgs.gov/landsat-missions/landsat-us-analysis-ready-data]
Dwyer, J.L., Roy, D.P., Sauer, B., Jenkerson, C.B., Zhang, H., and Lymburner, L., 2018, Analysis ready data—Enabling analysis of the Landsat archive: Remote Sensing, v. 10, no. 9, art. no. 1363. doi.org/10.3390/rs10091363.

Question 5: Why would we drop correlated variables when constructing the index?
Answer 5: You don't need to drop correlated variables, this choice is up to you. I raise the issue simply to show that you may inherently be weighting your HVI toward particular indicators by including them as independent, when in fact, they may be closely related. For example, including "Over 65," "Living Alone," and "Over 65 and Living Alone" as separate indicators that all contribute equally to the final HVI score will inherently weight the HVI toward these concepts unless you account for the potential overlap by applying an additional weighting scheme.

Question 6: We have crews in the field, in urban and rural areas. Summers in Louisiana have high temps and humidity. Can I use this approach to warn BOTH rural and urban crews?
Answer 6: Great question, you can use this approach for both urban and rural areas. However, HVIs are often used on more of a long-term basis reflecting average conditions rather than extremes. If you're alerting crews, you may use another method such as the National Weather Service declaring heat warnings on a more acute or short-term basis to help determine when heat conditions are unsafe for outdoor work. Speaking of rural areas, there is emerging research finding that rural areas may be at higher risk than urban areas for heat-related health outcomes due to a lower density of healthcare resources. So while urban areas may be hotter due to the urban heat island,

rural areas may have fewer resources to cope with their heat exposure when it is hot throughout the region, such as during a heat wave.

Question 7: When using a dataset that has a few census tracts with 'no data' values, when conducting the HVI, how are these tracts given index values?
Answer 7: Great question, these tracts usually are not included in the HVI analysis because they are not comparable. Most of the "no data" tracts will be tracts without any population, such as large parks, airports, or other non-residential land. I would not recommend assigning a score to these tracts based on the data you do have available, as the scoring will not be on the same scale. That is, it may be inherently lower than other tracts because it will have fewer contributing indicators in the dataset. But this may also depend on your intervention. For example, a tree-planting plan may still be viable in non-residential areas to help cool the city, so you can still use a more exposure-based HVI to help prioritize areas for tree planting. On the other hand, a cooling center may not be appropriate to locate in an area with zero population.

Question 8: How do I assign weights to the indicators?
Answer 8: The weighting scheme design is ultimately up to you. You can choose to not weight indicators at all and simply add the individual HVI scores, or you can add weights depending on your intervention as appropriate. For example, a tree planting plan may be weighted toward exposure indicators while a cooling center plan may be weighted toward adaptive capacity or sensitivity. Applying the weights can be as simple as multiplication factors or proportions. A simple example is below:

[Weighted HVI Score] = 0.8 * [Indicator HVI Score 1] + 0.2 * [Indicator HVI Score 2]

The above example is a simple HVI with two indicators, weighting the first indicator as much more important, or 80% of the total, and the second as much less important, at only 20% of the total.

Question 9: I just noticed that the HVI_U you calculated seems to be a normal distribution. Is that a regular property of HVI?
Answer 9: This will depend on your underlying data and the distribution across your selected geography. But it is common to see a normal distribution when using z-scores, as these are designed to distribute relatively normally across the dataset relative to the mean and standard deviation of a given indicator.

Question 10: How can we connect air pollution with UHI?
Answer 10: https://www.mdpi.com/2071-1050/14/15/9234
https://doi.org/10.1016/j.ecolind.2021.107976
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7434321/
https://www.jstor.org/stable/26328061

Question 11: Is the factor/component loading the same as factor scores?. I read from a paper that "the calculated factor scores were normalized to have a mean of 0 and a standard deviation of 1". Can I get little more explanation about the main in-put parameters from the PCA analysis that were used for the following z-score equation; z = x – μ/ σ
Answer 11: In general, we do not recommend PCA for HVIs because of the complexity of the method. It is often difficult to interpret the output of the HVI because they are so heavily processed. But yes, you are correct, that in the PCA method, we can take the factor scores output and convert those to z scores before applying the HVI scoring scheme such as the Reid et al. 2009 scheme we use in this training. You can then add those factor-based HVI scores to obtain your final HVI score.

Question 12: Sometimes LST data is on different tiles to a given boundingbox. Which tools can we use to join those tiles in a single one?
Answer 12: You're right, it's quite common to find your area of interest bisected into multiple Landsat tiles. In QGIS, you can use a tool called Merge to easily combine multiple rasters into a single raster for your analysis. If you are using ArcMap, you can use the Mosaic (Data Management) tool to accomplish this.

Question 13: Hello. I have one question to address. Do you think that it would be possible to integrate in the HVI calculation data such as disease statistics (Covid-19 positive confirmed cases, for example)?
Answer 13: Certainly, you can include any indicators you find to be relevant to the intervention you have in mind. Disease statistics can often be included in HVIs as sensitivity indicators, such as diabetes used in our example HVI. But it is also common to use other conditions such as cardiovascular or renal diseases or obesity, as they have been associated with heat-related illness and death in the literature. COVID-19 has been a big concern when considering interventions like cooling centers since users may be cautious about convening in an enclosed space during a pandemic.

Question 14: Is it possible to have adaptive and sensitivity data for the other parts of the world? Africa for example

Answer 14: Yes, though I am most familiar with US-based datasets. I would recommend that anyone engaging in HVI analysis to begin with the local government of interest. For example, many cities now offer GIS data portal services online containing local data they've used in their own analyses. National-level datasets may also be useful when considering demographic data such as in a national census, or disease prevalence data from local health organizations. However, the scale of this data may vary, so you may need to adapt your scope or scale based on data availability. You may find some useful resources from this earlier ARSET training on gridded population datasets:

https://appliedsciences.nasa.gov/join-mission/training/english/arset-introduction-population-grids-and-their-integration-remote

Question 15: Do you have tips for selecting imagery for areas with high cloud cover (i.e. humid tropics) ?

Answer 15: While the Earth Explorer tool is not yet global, it will become available in the near future. This data portal has a tool to select only images with low cloud cover. I usually set this to no more than 5 or 10% to ensure my images are clear. Alternatively, you can also consider the MODIS land surface temperature dataset, which is global. It does have lower resolution at 1km, but it has a great data product called the 8-day LST, which shows an average LST over 8-day periods for a more broad-scale view of surface temperature conditions than just a single snapshot. You can also get both day and night-time images from the MODIS datasets.

"Global Landsat ARD specifications are in development. Information will be provided when details are available."

[https://www.usgs.gov/landsat-missions/landsat-us-analysis-ready-data]

Dwyer, J.L., Roy, D.P., Sauer, B., Jenkerson, C.B., Zhang, H., and Lymburner, L., 2018, Analysis ready data—Enabling analysis of the Landsat archive: Remote Sensing, v. 10, no. 9, art. no. 1363. doi.org/10.3390/rs10091363.

Question 16: If we want to calculate an average HVI for a given city, it would be better to use a multi-year average LST for that purpose. I was wondering if it is possible to cover how to download the calculated multi-year average LST based on Landsat from GEE?

Answer 16: Good question, I have not seen an analysis use multi-year LST for this purpose. In general, I recommend that you try to maintain consistency between the

various data sources, so if you have socioeconomic data for only one year, you can try to use LST from that year. I would also be concerned that in a multi-year analysis, you may end up with many images with heavy cloud cover, which would interfere with the accuracy of your analysis. It is standard practice to use one or a few LST images for HVIs, but it will help to carefully select these images, such as by selecting days with very low cloud cover or during hot time periods.

Question 17: Do the LSTs that we created in part 1 have to be scaled after downloading? Or is scaling included in the GEE script?
Answer 17: In the GEE script also (Part 1) LST are scaled and converted to Celcius (Please see the script Landsat_LST_SUHI lines 89-96).

Question 18: Is it possible to do Vulnerability Indices beyond temperature? i.e. Pollution, Housing, Endangered Species, etc. Would that require a major redesign of the VI calculation to implement, beyond the data choices themselves?
Answer 18: It is possible. We covered this in depth during the previous session. This is part of a Social Vulnerability Index, which can include other metrics such as housing, pollution, flooding, etc.

Question 19: During the conversion of the unprocessed data, I got an error because I wasn't sure which CRS to define. What is the appropriate CRS for Detroit?
Answer 19: At the urban scale, it does not matter what Coordinate Reference System (CRS) is used, but it does matter that all of your data use the same CRS. Loading in the shapefile first (.shp) will help to set the CRS for other data.

Question 20: Is there a method to account for differences in surfaces which are very high from the ground (i.e. tops of tall buildings in cities) vs. at ground level?
Answer 20: If you don't want surfaces that are high from the ground, you could go into a city GIS portal and mask out a building layer to get just ground level measurements. This is mostly applicable to a city with many high rise buildings. Most people use LST data as is.