Session A

Part 2 Question & Answer Session A

Please type your questions in the Question Box. We will try our best to answer all your questions. If we don't, feel free to email Savannah Cooley (savannah.cooley@nasa.gov) and/or Glynn Hulley (glynn.hulley@jpl.nasa.gov). This document will be shared to the training webpage within one week.

1. **Given the possibility of a moderate to strong El Niño this year, how could El Niño interact with long-term climate warming to increase urban heat island risks and affect glacier mass balance? In particular, how might these impacts differ across regions, such as tropical cities, mid-latitude urban areas, mountain glaciers, and polar ice environments?**
    a. El Niño results in positive temperature anomalies on top of climate warming, and so UHI effects would be amplified through factors such as solar loading, reduced cloud cover, and drought-stressed vegetation (reduced evaporative cooling). These effects are likely worse in tropical cities where humidity nonlinearly accelerates heat stress.
    b. Models currently show that confidence is high that a strong El Niño phase will develop heading into summer 2026. If correct, a strong El Niño could lead to 2026 or 2027 becoming Earth's warmest year on record, elevating the UHI even more.

2. **How does this urban cooling effect during the day differ depending on surrounding land type (e.g. arid / agricultural vs forested)**
    a. Arid, bare soil, senesced crops or vegetation surrounding cities usually result in an urban cooling effect (e.g. Delhi, Athens as demonstrated in the presentation). The reverse is true for forested rural regions because dense vegetation maintains strong evapotranspirative cooling that keeps rural LST 10–20°C below urban surfaces, i.e. urban heat island effect. These effects may reverse at nighttime depending on thermal capacity of urban vs rural materials/type.

3. **Why does near surface ozone concentration increase in the summer?**

   a. It is in fact a photochemical reaction from car exhaust and other pollutants (NOx) that reacts with sunlight. So the hotter and sunnier the day, the higher O3 levels builds up near surface.

4. **Can ECOSTRESS land surface temperature data be integrated with hydrological or watershed models to jointly analyze urban heat, evapotranspiration, runoff, and the benefits of green infrastructure? Also, are there any NASA-related application case studies that demonstrate this type of integrated analysis?**
   a. Yes, ECOSTRESS Evapotranspiration (ET) products (ECO3ETPTJPL, ECO4ESIPTJPL) feed directly into watershed water balance models as the latent heat term.
   b. For the LA28 Olympics, we are working on a project that uses ECOSTRESS LST and ET in ML models to jointly quantify the heat mitigation and hydrological co-benefits of cool pavement, tree canopy, and green infrastructure investments that would strengthen the case for nature-based solutions considerably.
   c. Two good papers on biosphere influence on urban heat and carbon uptake in urban areas:
      Coleman, R.W., Stavros, N., Hulley, G., & Parazoo, N. (2020). "Comparison of thermal infrared-derived maps of irrigated and non-irrigated vegetation in urban and non-urban areas of southern California." *Remote Sensing*, 12(24), 1–19.

      Parazoo, N.C., Coleman, R.W., Yadav, V., Stavros, E.N., Hulley, G., & Hutyra, L. (2022). "Diverse biosphere influence on carbon and heat in mixed urban Mediterranean landscape revealed by high resolution thermal and optical remote sensing." *Science of the Total Environment*, 806, 151335.

5. **What is the way to understand the evapotranspiration capability of a place so that we can understand and demonstrate if, when water is scarce, planting trees is not a solution and we have to deal with soil moisture or underground water? Is there any connection?**
   a. The connection between water availability and green infrastructure heat mitigation lies in a plant's physiological response to water stress. Plants

rely on evapotranspiration (ET) - specifically transpiration- to dissipate absorbed solar energy as latent heat. When soil moisture or groundwater supplies are depleted, plants experience hydraulic stress, forcing them to close their stomata to preserve water. When stomata close, transpiration drops to near-zero, transitioning the surface energy balance so that incoming energy is shed almost entirely as sensible heat. This rapidly spikes the canopy temperature, effectively turning a cooling asset into an active thermal emitter.

b. Therefore, in arid environments or under prolonged drought, simply planting trees without addressing deep soil moisture or sub-surface aquifers will fail to resolve urban heat. This distinction is explicitly shown in the comparison between irrigated urban green spaces (like the managed nature preserves) and adjacent unbuilt, unmanaged hillsides. While both have vegetative cover, unmanaged spaces lack the water supply to sustain transpiration under high thermal loads, causing their LST to remain almost as elevated as impervious urban zones.

6. **I have a question regarding downloading. Can we specify the cloud cover, such as <10%?**
    a. This is possible in the eco-sharp GEE code (but not AppEEARS).

7. **Is NASA EarthData also a valid source for the most recent ECOSTRESS LST, QA, and watermask data across the globe?**
    a. Yes. NASA EarthData search accesses the same ECOSTRESS LST, QA and cloud mask layers that we accessed in AppEEARS for Demo1.
    b. ECOSTRESS coverage is +/- 55 degrees latitude.

8. **What is the geo location accuracy of ECOSTRESS LST in AppEEARS and GEE?**
    a. It varies and depends on the amount of cloud and/or open ocean on the orbital track but on average is ~50m (sub-pixel). Collection 3 data will include geolocation flags in all products (best, good, bad etc). Geolocation corrections are applied using the ISS orbital and attitude data combined with ground control points.
    b. The next version of eco-sharp in GEE (v1-2) has a geocorrection option to snap the LST on the tile to the build-in land water mask from Sentinel-2.

9. **When I download files from APPEARS, I get asked if I want to save each file individually. How do I avoid this?**
    a. You can select the checkbox on the upper left corner by “Name” to select all of the files in the list. Then click “Download files” from the Download list on the upper right corner of the Download Sample page.

10. **Could you share written steps for the demonstration?**
    a. It is not standard ARSET practice to capture or share a training transcript. That said, the slides are posted to the training website and the demo scripts on GEE and Github have a readme file and comments throughout the code so it is documented in that way.

11. **Can we have a python equivalent of this R code?**
    a. That is not provided in this training but is certainly possible to do. In Python, the data loading, cloud masking, and QC bit-manipulation shown in the training can be easily implemented using Rasterio or RioXarray for raster operations, Geopandas for isolating vector areas of interest, and Xarray or NumPy to perform the Kelvin-to-Celsius conversions and mathematical masking. Furthermore, if you wish to run the machine learning downscaling pipeline in Python, you can utilize the ee (Google Earth Engine) Python API package to execute the exact same Random Forest regression script in a Python environment, or handle it locally using Scikit-learn.

12. **Are there data available for thermal IR of water bodies?**
    a. Yes. ECOSTRESS and other thermal instruments provide LST retrievals over both land and water bodies.
    b. Collection 3 data (starting to become available at LPDAAC) have a dedicated water surface temperature product that is optimized for water surfaces. The LST standard product algorithm is not optimized for water surfaces.

13. **Do you groundtruth the LST data? If so, how do you do it?**
    a. Yes, using ground-based radiometers (engineered at JPL and calibrated to within <0.1 degree accuracy). The measurements on ground at validation sites (e.g. Lake Tahoe, Russel Ranch etc) are matched contemporaneously (in space and time) with the satellite overpass

estimation of LST. Over time, you can start analyzing trends in differences, biases, etc.

b. Satellite-based Land Surface Temperature (LST) products are validated using **ground-based thermal radiometers** that measure surface-leaving radiance at the exact moment of a satellite overpass. This recorded radiance is mathematically converted into a kinetic surface temperature and directly compared against the satellite's retrieved pixel value to calculate accuracy and bias
    i. Ex: Lake Tahoe site. Water bodies are highly reliable because their emissivity is quite stable and known a priori, isolating atmospheric correction accuracy
    ii. https://calval.jpl.nasa.gov/
c. Dedicated flight missions utilizing airborne sensors, like the **Hyperspectral Thermal Emission Spectrometer (HyTES)**, capture ultra-fine spatial resolution thermal pixels. When a HyTES airborne campaign directly coincides with a satellite overpass (such as an ECOSTRESS matchup), these high-resolution pixels are used to evaluate sub-pixel heterogeneity and validate sensor performance over complex land structures.

14. Do cloud-affected pixels in thermal satellite images have lower temperatures than clear-sky pixels?

a. Yes. Clouds block (mostly through absorption) the thermal emission from the surface, and so the sensor in space is mostly measuring the radiance emitted from the cloud top surface itself, which is much cooler than the land surface. Sometimes low level warm clouds during wintertime have very similar temperatures to land, making cloud masking difficult with thermal only data.

15. Will you show the methodology of estimating second degree burn exposure?

a. It was estimated from a generalized model. I can find the journal article.

16. During a heatwave event, cloud cover might ruin our optical data. How resilient is the Random Forest downscaling model if we have to use Sentinel-2 optical predictors from a few days before or after the clear ECOSTRESS thermal acquisition?

a. The GEE code has an option to select your cloud tolerance which is set at <30% as standard. You could enforce less than that if you wish. The L2T data in GEE already come cloud screened, so the negative effects would only be see if in fact there was undetected cloud on the scene. Sentinel-2 predictors will be found for a max of +/- 1 month of the ECOSTRESS observation, if there is significant phenology change (e.g. sudden green-up or fire etc) then those temporal discrepancies would likely result in confusion in the ML training. Usually over summertime using LA as an example, summertime reflectances are usually stable and don't change much.

17. What should be the probable research scope using ECOSTRESS data rather than [in addition to] urban heat island detection?

a. Beyond urban heat island mapping, the unique features of ECOSTRESS data - specifically its diurnal sampling capability from the International Space Station and its high spatial resolution- open an expansive research scope across agriculture, hydrology, ecology, and climate modeling. A major research area is hydrological modeling and agricultural water management, where ECOSTRESS LST is utilized to derive accurate Evapotranspiration (ET) products. This allows monitoring of drought conditions, track soil moisture depletion, and management of agricultural water use by calculating how net surface energy partitions into latent and sensible heat fluxes.
b. Another research avenue is ecological conservation and wildfire risk assessment. We can detect ecosystem thermal stress and vegetation canopy warming before physical signs of wilting become visible in optical data, as well as study long-term thermal impacts over disrupted ecosystems like burned forests.
c. Additionally, ECOSTRESS enables climate change trend analysis and surface energy balance modeling. This supports the characterization of subtle warming indicators in topographically complex areas or aquatic environments, such as long-term lake water temperature tracking and micro-climate buffering.

18. If I understand correctly, we first downloaded the LST data for the region of interest through AppEEARS and opened it in QGIS for visual inspection.

Then, in a separate process, we are performing a similar analysis using an R script, plus doing further analysis. Is that correct?

a. That is the general training workflow for Demo 1. The process splits into a visual inspection phase (QGIS) and a quantitative computing phase as complementary analyses. In the first stage, the analysis-ready ECOSTRESS Level 2 Tiled (L2T) GeoTIFF data (originally requested and packaged via the AppEEARS tool) is loaded directly into QGIS alongside vector bounding boxes of the study region. Toggling these multi-temporal rasters allows for a rapid qualitative evaluation of thermal patterns, such as visually identifying relative warmer versus cooler surface tones across different land covers.
b. The R script then takes this process further by shifting from visual interpretation to automated, repeatable data refinement and precise statistical quantification. The R code automatically handle the necessary data cleaning by converting raw Kelvin values into Celsius, evaluating bit-packed mandatory quality control (QC) layers to isolate reliable data, and pairing them with independent cloud masks to filter out atmospheric noise.

19. **The downscaling relies on daytime optical data. Since it doesn't work at night, what kind of data do we need to successfully downscale a nighttime ECOSTRESS image to 10 meters?**
    a. Daytime optical reflectance properties (NDVI, albedo, impervious fraction) are still physically linked to surface material composition, which governs thermal inertia and heat capacity regardless of time of day. So a Random Forest trained on daytime Sentinel-2 predictors would likely still capture much of the nighttime LST spatial structure, just through the material-composition pathway rather than the direct solar-loading pathway.

20. **My studies on LULC, LST (daytime and nighttime), and UHI hotspot span from 1984 to 2024. For LST determination, between MODIS and ESCOSTRESS, which fits very well?**
    a. Depends if you need high spatial (ECOSTRESS, 70m every 3-5 days) or temporal resolution (MODIS is twice-daily but 1km)?

21. **To what extent can ECOSTRESS land surface temperature data be integrated with hydrological and watershed models to jointly analyze urban heat dynamics, evapotranspiration, runoff behaviour, and the performance of green infrastructure and are there NASA-supported case studies demonstrating such multi-layer integration (e.g., ECOSTRESS + ALEXI/ESI + land–water thermal interactions)?**
    a. See answer to question 4

22. **Is thermal infrared data over water bodies available or derivable for such analyses, and how can cloud-affected pixels be reliably handled in workflows that require temporal consistency (e.g., downscaling with Sentinel-2 predictors)?**
    a. The ECOSTRESS LST product is produced everywhere, so over land and water. In the L2 swath products a cloud mask is included for removing clouds, in the L2T tiled products the mask is already applied.

23. **Do you have the ECOSTRESS data processing package in python?**
    a. While the hands-on materials for this specific ARSET training series utilize an R-based script (ECOSTRESS_LST_Demo1.R) and the Google Earth Engine JavaScript API, these workflows can be fully replicated in Python using open-source geospatial libraries. Because NASA distributes its Level 2 Tiled (L2T) products as standard analysis-ready GeoTIFFs via AppEEARS, you do not need a specialized proprietary package to handle the data.
    b. In Python, the data loading, cloud masking, and QC bit-manipulation shown in the training can be easily implemented using Rasterio or RioXarray for raster operations, Geopandas for isolating vector areas of interest, and Xarray or NumPy to perform the Kelvin-to-Celsius conversions and mathematical masking. Furthermore, if you wish to run the machine learning downscaling pipeline in Python, you can utilize the ee (Google Earth Engine) Python API package to execute the exact same Random Forest regression script in a Python environment, or handle it locally using Scikit-learn.

24. **Is ECOSTRESS data useful for small waterbody for wetland nighttime radiation monitoring?**

a. Yes. ECOSTRESS data is valuable for monitoring nighttime radiation, surface energy balances, and thermal patterns across wetlands and small waterbodies. Traditional high-resolution optical and thermal sensors onboard sun-synchronous satellites are limited by morning-only overpass schedules, meaning they miss nocturnal thermal dynamics entirely. Because ECOSTRESS collects data from the International Space Station, its precessing orbit enables it to capture purely emitted thermal radiation through the 8–14 μm atmospheric window at various times of night when waterbodies are slowly shedding heat.

25. **Is there ongoing research or planned development within NASA or partner institutions to integrate ECOSTRESS LST dynamics with human-centred resilience models — for example, linking thermal stress patterns, nighttime heat retention, and micro-scale surface heterogeneity with indicators of physiological vulnerability, emergency response readiness, or exposure-risk forecasting? If so, which data fusion pathways (e.g., ECOSTRESS + VIIRS + socio-environmental layers) are considered most promising for operational use?**
    a. See Hulley et al. (2019): “New ECOSTRESS and MODIS Land Surface Temperature Data Reveal Fine-Scale Heat Vulnerability in Cities: A Case Study for Los Angeles County, California”. This study combines ECOSTRESS, MODIS and SEDAC socioeconomic data to assess Heat Vulnerability https://www.mdpi.com/2072-4292/11/18/2136
    b. Most heat risk models (e.g. NOAA NWS, CalEPA) use model derived surface air temperatures, and not LST directly. The paper above merges LST with sociodemographic variables to estimate heat vulnerability, but we really need to get at the physiological effects of temperature, e.g. heat index, UTCI, or mean radiant temperature. There are digital twin models out there that estimate these variables but not directly from NASA derived LST, which provides the outgoing longwave radiation component. There is a project upcoming called HEAT28 that will directly integrate all high res NASA remote sensing datasets into a Digital Twin model to estimate fine scale urban heat for the upcoming LA28 Olympics.

26. **Can you provide an example of the difference in night time heat and daytime heat (as in when it was an indicator that something that is**

supposed to cool naturally hasn't cooled and led to detecting some other cause)?

a. An example we can consider is the Chatsworth Nature Preserve, which we monitored in Demo 1 alongside an adjacent commercial zone. Normally well-irrigated vegetated areas transpire to allow for a cool nighttime temperature relative to urban infrastructure. However, if a multi-day heatwave intensifies and water supplies are depleted, plants close their stomata to conserve moisture, which halts evaporative cooling. When this happens, nighttime satellite data reveal that the expected thermal gap between green spaces and the asphalt substantially narrows. This narrowing of the temperature gap is an indicator that the vegetation is experiencing water scarcity or groundwater exhaustion, identifying an instance of localized heat-induced ecological stress rather than standard urban heat retention.

27. **When GEE is used in everyday analysis, Landsat and Sentinel are used. Is ECOSTRESS openly available like Landsat? Do we need specific access?**
    a. Yes ECOSTRESS L2T Collection 2 product is available over the continental US.

28. **Is it possible to share some geojson files etc and other samples on part 2's walk through for areas already detected to have heat waves, especially ones that do not subside?**
    a. We encourage participants to find their own areas of interest to test the workflows we provided in areas of interest.

29. **Is there any quicker, one-shot way before we get into coded scripts to quickly know the red flag where heatwave is to be picked for analysis?**
    a. Check for historical heatwaves that made the news, or keep an eye out on your local weather forecast.

30. **Is there a definition of "heat wave"? Is there a canonical source for data on instances of heat waves in various locations?**
    a. There are many definitions, but most use either a daily average temperature that exceeds a certain threshold lasting 3+ days.

31. **Is it possible to use this downscaling method in thermal data obtained by other thermal sensors? Is there any downside to downscaling? Why not just use Sentinel at 10 m instead of downscaling ECOSTRESS? PlanetScope also capture RGB and NIR but at a 3m resolution. Is it possible to downscale ECOSTRESS to 3m? Or is it just a waste of resource since the other datasets used have 10-30m resolution?**
    Yes you can theoretically downscale any LST product in a similar manner, e.g. MODIS, VIIRS, ASTER but there are limitations, e.g downscaling MODIS from 1km to 10m is 2 orders of magnitude factor and likely would not produce reliable or accurate results. Sentinel-2 does not have thermal bands that allow for LST estimation. If we directly modeled the LST from S2 you would introduce large errors (in absolute magnitude) anywhere the LST and visible/vswir spectral bands do not correlate well, e.g. due to surface composition, thickness (substrate), soil moisture, wind etc.

32. **Can I downscale using ArcGIS/QGIS?**
    a. Currently the downscaling workflow is only available on Google Earth Engine. That said, it could be possible to reproduce in Python or R. if you wish to run the machine learning downscaling pipeline in Python, you can utilize the ee (Google Earth Engine) Python API package to execute the exact same Random Forest regression script in a Python environment, or handle it locally using Scikit-learn.

33. **Do you consider spectral fidelity during downscaling from 70m to 10m using sharpening technique with sentinel?**
    a. Yes, spectral and radiometric fidelity is directly preserved during the downscaling process through an automated residual correction step implemented within the computing pipeline. While the Random Forest model captures fine-scale spatial variations at 10 meters based on optical predictors, machine learning regressions can introduce local structural biases. To correct for this, the script automatically takes the initial 10-meter sharpened LST estimation, aggregates it back up to the native 70-meter scale using a spatial mean reducer, and subtracts it from the true ECOSTRESS satellite observation to isolate the exact modeling error. This 70-meter residual error is then bilinearly resampled back down to the fine 10-meter grid and added directly to the initial downscaled model

prediction. By executing this radiometric normalization loop, the pipeline guarantees that the total thermal energy across the sharpened 10-meter sub-pixels remains strictly consistent with the absolute physical measurements recorded by the spaceborne instrument.

34. Can we do all this processing of ECOSTRESS data in Earth Engine?

a. The entire thermal sharpening and downscaling pipeline can be executed directly within the cloud-based Google Earth Engine (GEE) platform. The fully interactive user interface script (Eco-sharp-v1-1) handles image querying, time-of-day filtering, cloud screening, Random Forest model training at 70 meters, high-resolution prediction at 10 meters, and residual correction without requiring any local software installation. The final 70-meter and 10-meter GeoTIFF products are then exported directly to your Google Drive folder.
b. Note that certain specialized statistical workflows, such as the pixel-level longwave distribution analysis and historical "thermal gap" evaluations demonstrated in Demo 1, utilize a local standalone script written for R and RStudio. Both GEE and R workflows complement each other to bridge the gap between cloud-scale data processing and detailed regional statistics.

35. Does the fusion still work well when the high resolution images are acquired at different times compared to ECOSTRESS?

a. Yes, but will depend on location and time of year, e.g. phenological changes in the high res images may be different to when the thermal was acquired.

36. How do you actually validate that the 10m results are real and not artifacts? How was downsacled ECOSTRESS @10m resolution validated? What validation approach did you use? Cross-validation?

a. We are attempting validation with higher resolution airborne thermal data, e.g. HyTES at <5m resolution (see hytes.jpl.nasa.gov).

37. In what manner does a water body in an urban area affect the LST and air temperature at night and during the day?

a. Water bodies have high heat capacity so are cooler than surrounding areas during the day, but warmer than surrounding areas at night, particularly during high temperatures, heatwaves etc.

38. You train your model at 70m and then predict at 10m, so you are assuming that the same relationship holds at 70m and at 10m? But that is not necessarily true or known, correct?

a. Yes, strictly we should be training on purely homogeneous pixels at 70m scale and assume that relationship holds at 10m, but often in urban areas we have insufficient homogenous pixels at 70m to provide enough data for the RF model. In the next version, we are exploring the effects of these discrepancies, but good question.

39. Given that urban surfaces are highly heterogeneous, how does the TES algorithm on ECOSTRESS handle the spectral mixing of materials with very different emissivities within a single 70m pixel before the sharpening process?

a. TES is trained on spectra from mostly natural surfaces (rocks, soils, sands, veg, and some urban). In a mixed urban environment these assumptions on spectral variability are usually okay, but there are certain materials (e.g. some types of metals like aluminum that have unusually low emissivity and low spectral contrast) that don't follow the typical emissivity calibration curves we use in TES. In these instances we may underestimate the temperatures due to overestimating the emissivity. Urban-calibrated TES curves have been explored but never implemented.

40. Can ECOSTRESS land surface temperature data be integrated with hydrological or watershed models to jointly analyze urban heat, evapotranspiration, runoff, and the benefits of green infrastructure? Also, are there any NASA-related application case studies that demonstrate this type of integrated analysis?

a. Yes, ECOSTRESS land surface temperature (LST) and evapotranspiration (ET) products can be tightly integrated with hydrological or watershed models to evaluate urban energy balances, runoff dynamics, and green infrastructure efficacy. Because LST governs how net radiation splits into

latent heat flux (evaporation and transpiration) and sensible heat flux, ingesting thermal data allows land surface and hydrological models to better calibrate soil moisture availability, estimate urban moisture stress, and simulate how green infrastructure modifies heat retention and local runoff properties.

41. **Why are you using NDVI as opposed to EVI (Enhanced Vegetation Index)? If I wanted to establish a relationship between LTS or UHI and vegetation infrastructure, which index would be the most appropriate to extract vegetation cover?**
    a. The downscaling workflow uses the Normalized Difference Vegetation Index (NDVI) because it effectively captures vegetation greenness by measuring the distinct spectral contrast between red chlorophyll absorption (Band 4) and near-infrared canopy reflectance (Band 8). Within the Random Forest regression model, raw reflectances from Sentinel-2's Blue, Green, Red, and Near-Infrared channels are fed directly into the algorithm alongside auxiliary grids. Because the decision trees map highly complex, non-linear relationships directly from these fundamental spectral bands, the marginal benefit of using an advanced index like EVI vs. NDVI is typically very small.

42. **The GEE script/materials are missing from the Github/the code is different (asset is not provided). Can you add them?**
    a. Link to GEE script: https://code.earthengine.google.com/f399ecd63e7eeac7d705634d6d64470d?hl=it

43. **Does keeping smileRandomForest output mode as classification improve performance or is it used only for feature importance analysis?**
    a. I've used it for both.

44. **For UHI analysis, is it advisable to combine optical remote sensing data (e.g., Landsat 8/9 and MODIS Aqua) with thermal remote sensing data (ECOSTRESS), or is thermal remote sensing alone sufficient, particularly after downscaling ECOSTRESS land surface temperature data from 70 m to 10 m resolution?**

a. If you need higher temporal coverage I'd suggest augmenting with Landsat and MODIS/VIIRS.

45. Can I use Landsat 8 for UHI applications instead of ECOSTRESS?

a. Yes but be aware that Landsat 8 observations are at 10am and will miss the hotter times of day.

46. LA has relatively low humidity compared to New Orleans. If the same analysis was performed in a "High humidity" location, how would it change the observed results? Air temperatures don't change much, diurnally, does remotely sensed LST?

a. Performing this analysis in a high-humidity environment increases atmospheric attenuation and path radiance due to intense infrared absorption by column water vapor. In the thermal infrared window (8–14 microns), water vapor carves out structural noise that obscures surface emission features before they reach the satellite sensor. To resolve accurate land surface temperatures in a humid region, executing a rigorous atmospheric correction workflow (such as the Temperature-Emissivity Separation algorithm used by ECOSTRESS) is mandatory to mathematically untangle surface radiance from humid atmospheric background noise.

b. Regarding diurnal patterns, unlike ambient air temperatures which can stay relatively stable or buffer slowly over a 24-hour cycle, remotely sensed LST can exhibit strong diurnal fluctuations. Because land surface skin temperatures respond directly to solar heating and localized material heat capacities, thermal emissions spike rapidly during the day and drop sharply at night based on the surface type. Consequently, even in humid environments, ECOSTRESS's unique ability to track these dramatic surface temperature shifts across changing times of day remains fully operational and distinct from stable diurnal air trends.

47. Could you please explain how to setup and access GEE? I already registered for an account but I cannot access the links

a. To access the cloud-computing workspace, open a web browser and navigate directly to the primary developer portal at code.earthengine.google.com. Log in using your registered credentials to launch the interface. If you encounter access errors, your profile may not

be fully approved or linked yet; you can verify or re-submit your registration details by visiting the signup page at earthengine.google.com/signup.

Session B

Part 2 Question & Answer Session B

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 Savannah Cooley (savannah.cooley@nasa.gov) and/or Glynn Hulley (glynn.hulley@jpl.nasa.gov). This document will be shared to the training webpage within one week.

1. **Can we use these temperature datasets to detect or predict earthquakes (over land or water)? Can the temperature anomalies be correlated with tremors that are seismic in nature?**
    a. This is outside the area of expertise of the training instructors. We encourage you to search through peer-reviewed studies on this topic.

2. **Since ECOSTRESS was launched in 2018, does that mean that is the year it has data from? How best can we use it to analyse historical heatwaves?**
    a. Correct, ECOSTRESS data stream strictly begins from that year forward. It cannot be used directly to view thermal observations prior to 2018.
    b. To evaluate long-term trends, scientists use historical baselines from MODIS or Landsat to establish what a "normal" multi-decade seasonal profile looks like and determine localized warming rates over time. For example, decades of consistent spaceborne records are what allow researchers to extract precise, long-term climate signals like lake surface warming or a rising frequency of hot summer nights.
    c. Once the pre-2018 baseline is established, you can layer in ECOSTRESS to solve a critical blind spot of traditional historical data: overpass time. Satellites like Landsat operate in sun-synchronous orbits that pass overhead at roughly 10:30 AM local time, missing the peak afternoon heat. Because ECOSTRESS features a precessing orbit, it samples across the entire diurnal cycle. Fusing historical datasets with modern ECOSTRESS imagery lets you leverage the long-term context of the past to evaluate current peak afternoon heat stress and nighttime heat-retention failures at a fine neighborhood scale.

3. **What's the margin of error when measuring the temp? For example, would a 0.2°C decrease be less significant?**

a. Over uniform calibration targets like lakes or oceans, the absolute precision (RMSE) of spaceborne thermal data processed via the Temperature-Emissivity Separation (TES) algorithm is typically falling between 0.5 K and 1.5 K
    i. See Part 1 slides 32, 33, 34 Mauna Loa Caldera example
b. Over more complicated land patterns (like deserts or rocky lava flows), standard Split-Window methods can accumulate systemic cold biases of 3 K to 5 K if their lookup tables assume the wrong emissivity parameters. Because the TES workflow physically calculates emissivity band-by-band, it keeps errors much lower, maintaining an accuracy profile within approximately 1.5 K across diverse geographic landscapes
c. In the context of a heatwave analysis, a 0.2°C change is relatively small on its own in the context of urban heat analyses and generally indicates that the baseline thermal landscape or cooling dynamic has remained essentially stable between observations.
d. Measurement noise from instrument itself is at the ~0.2 C level.

4. **The sharpening looks amazing. Can this downscaling technique work over shallow coastal environments, such as coral reefs?**
    a. No, this specific downscaling technique cannot be applied directly over shallow coastal environments or coral reefs without completely restructuring the model. The underlying RF machine learning framework relies strictly on terrestrial land surface predictors (such as vegetation vitality (NDVI), land cover types, and elevation) to model localized thermal variations. the thermal behavior, high heat capacity, fluid mixing, and heat dissipation of open water bodies are governed by different physical processes than terrestrial land surfaces and for this reason the downscaling technique will not work over coastal environments. The script screens out water bodies.

5. **Why did you select these co-variates (talking about Glynn's demo)?**
    a. The co-variates for the Random Forest model were selected because they represent the fundamental, high-resolution physical structures of the landscape that directly dictate surface energy balance and thermal retention. We seek covariates that correlate well with LST:
        i. Sentinel-2 Bands (B2, B3, B4, B8) reflectances capture fine-scale spatial patterns and surface albedo indicators. The NIR band (B8)

is highly sensitive to vegetation density, water content, and biomass, making it a strong predictor of spatial thermal variability across urban-rural gradients. B2 is more sensitive to solar insulation / UV. Also good correlations with SWIR data (e.g., 1.6 micron band).

ii. Land use classification data characterizes surface type, allowing the model to separate distinct heat-emission signatures belonging to classes like buildings, soil, vegetation, or open areas.

iii. Elevation accounts for orographic temperature gradients driven by the atmospheric lapse rate, which naturally explains substantial LST variations across topographically complex terrains.

6. **Is the ECOSTRESS instrument the small box with the logo on the ISS?**
    a. The logo itself is a graphical element added to the slide formatting, but the actual physical instrument is indeed the rectangular payload highlighted inside the black circle in that image.

7. **Does it really matter if the ECOSTRESS and Sentinel images aren't from the same day?**
    a. In this workflow, it does not substantially matter if the images aren't collected on the exact same day, which is why the Google Earth Engine pipeline safely uses a Sentinel-2 search window spanning several weeks time period around the ECOSTRESS acquisition date. This timeline is valid because the chosen co-variates (e.g., topography, urban building fractions, and broad land cover classes) represent structural characteristics of the landscape that change very slowly over time.
    b. However, if there is rapid LC change (ex. Rapid agricultural changes, fire, etc), then participants may consider reducing the time window from +/- 1 month to ~1 week. The only issue with this is that users might not actually find a clear Sentinel scene.
    c. Nice feedback that we will consider adding to the next version of the downscaling code. Next version will also have improved ECOSTRESS geolocation.

8. **What is the source of local temperature introduced on the model?**
    a. the source of temperature used in the machine learning script is the ECOSTRESS Land Surface Temperature (LST) product itself (derived from

the Level 2 Tiled LSTE collection). The model does NOT introduce local ambient air temperatures from ground-based meteorological weather stations.

9. **When applying residual correction to downscaled 10 m LST data, how does the model ensure radiometric consistency with the original 70 m ECOSTRESS native product?**
    a. There is a residual correction pipeline applied to adjust the model's outputs back to the original satellite observation. The original, true 70 m ECOSTRESS LST pixel value is subtracted from this upscaled model prediction to create a localized "residual difference" image representing the exact modeling error. This 70 m residual map is smoothed back down to a fine 10 m grid using a bilinear resampling method and added directly back to the initial 10 m downscaled model prediction.

10. **Does the script auto-select the corresponding Sentinel2 / SRTM / Built Surface Data if we select a different study area or do we have to find and select them manually beforehand?**
    a. Yes, the script auto-selects all corresponding covariates (including Sentinel-2 imagery, SRTM elevation models, JRC Global Human Settlement Data, and ESRI land cover maps) whenever you select a new study area. There is no need to manually search for, tile-match, or filter the ancillary baseline sheets beforehand.

11. **What is the 'truth' that the RF is predicting? I get the input covariates but what is it predicting against?**
    a. The underlying "truth" that the RF regression model uses as its target training variable is the original, native 70-meter ECOSTRESS Land Surface Temperature (LST) observation layer. The model evaluates the actual thermal energy emitted from the ground surface at that precise day and hour as its baseline response variable.

12. **For the data downscaling, is it possible to pick a specific city? If not, how do you extract a specific boundary?**
    a. The interactive tool panel does not feature a simple city-name text lookup bar. However, you can easily target any specific city or neighborhood by utilizing the graphical custom boundary selection modes. The utility is

designed to be fully interactive within the main Google Earth Engine mapping visualization panel.

13. **Are these predictors fixed at a single time point, or are they matched to the closest available time?**
    a. A combination of both. The predictor variables in this data workflow combine matched time windows for rapidly changing surface environments with fixed historical intervals for static structural land features:
        i. The Sentinel-2 spectral surface reflectance bands are actively searched and matched. The script scans the image repository dynamically to extract the single clearest, cloud-masked optical scene captured within a several week- to-moth window centered directly on your active ECOSTRESS overpass date.
        ii. The ancillary environmental co-variates are from static baseline collections. This includes the ESRI Global Land Cover map from 2020, the JRC Global Human Settlement Layer from 2018, and the USGS SRTM digital elevation dataset. These layers operate at single fixed time points because regional topographic elevation slopes and coarse and block-level urban concrete arrangements change very slowly compared to daily atmospheric cycles and LST.

14. **It looks like only LA area is available on GEE. When are additional areas going to be added, and are there any plans to process and upload L2 from other regions (e.g. Pacific Islands)?**
    a. The complete historical archive of ECOSTRESS LST observations has been ingested into Google Earth Engine over the entire continental United States. You can leverage this machine learning pipeline to generate fine-scale thermal maps across any American city or region observed since the sensor went operational in 2018.
    b. We currently do not have ECOSTRESS data on GEE for areas located outside the US mainland. However, it would be possible to reproduce the GEE script in Python or R. After acquiring ECOSTRESS LST and other ancillary data sets over your area of interest (e.g., with ApPEEARS, NASA EarthData search and other ancillary sources), you could process these data with an adapted downscaling R/ Python script. This would require substantial time investment.

15. **What OS or platform do the ARSET staff use or recommend - Windows, Linux, or Mac?**
    a. We do not require or recommend any specific operating system. We build courses for a diverse international audience, so our structural guideline is to design workflows exclusively around free, platform-agnostic, and open-source software tools.

16. **Is it possible to use colab for the exercise?**
    a. Yes. For the Demo2 GEE workflow: Colab supports the Google Earth Engine Python API (ee). You can authenticate your free Earth Engine account in a standard Python notebook, import your processing coordinates, and execute the exact same Random Forest downscaling logic locally or via cloud export tasks.
    b. For the R Demo1 workflow: while Colab defaults to Python, you can launch a native R notebook by navigating to https://colab.to/r. This environment supports the installation and execution of all geospatial packages used in the script, including terra, tidyverse, and tidyterra.

17. **Is there a tutorial to add the legend/colorbar on python or R?**
    a. We do not have a dedicated training for this. However, both languages offer ways to implement synchronized legends and dynamic colorbars for geospatial raster datasets.
    b. In R (ggplot2 + tidyterra), as shown in the Part 2 R script, the tidyterra package provides dedicated scales that handle spatial raster plotting alongside standard ggplot2 syntax.
    c. In Python the libraries to use would be: matplotlib + Xarray/rasterio.

18. **There is a lot of significance put on the vegetation. But does vegetation really correspond to cool islands in all cases? If a roof of a house is lined with grass, it wouldn't really make the house a cool island, at least not to the degree the model would predict. When proposing measures for heat adaptation in cities isn't it easy to end up making false assumptions about effects of vegetation?**
    a. Vegetation does not automatically correspond to a cool island in all cases. The cooling capacity of vegetation is not driven by its color or presence alone. Rather, it is almost entirely governed by

evapotranspiration (ET), which requires a sustained water supply. In topographically complex or arid environments, a clear thermal split is observed between irrigated green spaces (like managed city parks or preserves) and adjacent unmanaged, unbuilt hillsides. Even though both areas feature vegetation, the unmanaged hillsides lack irrigation. During a severe heatwave, unwatered plants cross a strict physiological threshold: they close their stomata to conserve moisture, effectively shutting off their natural air-conditioning and shedding energy as sensible heat instead. This causes their surface temperatures to spike, making them look nearly as hot as impervious urban surfaces.