Part 3 Questions & Answers Session

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 Sativa Cruz (sativa.cruz@nasa.gov) or Justin Fain (justin.j.fain@nasa.gov)

Question 1: How does hyperspectral remote sensing differentiate between native and invasive species in grasslands ecosystems?
Answer 1: Effective mapping of invasive plants depends on several factors. For example, contrasting seasonal phenology of invasive plants compared to native plants. Or, distinct biochemical, physiological, and structural characteristics (or functional traits) of invasive plants vs. native plants. The strength of hyperspectral data is their really fine spectral resolution – meaning that these data can distinguish minute spectral differences between native and invasive species. So, if we have invasive plants with different functional traits compared to those of natives, we can detect them using suitable hyperspectral data (e.g., fine spatial resolution). Eventually, successful application of remote sensing data regardless of data type, depends on the spatial resolution of remotely sensed data compared to size of the phenomenon we are observing. This issue was discussed in Part 2 of the webinar. Please also see the answer to question 7.

Question 2: Can hyperspectral remote sensing detect early-stage invasions before they become widespread? If so, how?
Answer 2: It primarily depends on whether the spatial resolution (i.e., grain size, pixel size) matches the phenomenon we are observing and whether the invasive species are spectrally different from native plants. If we are dealing with large invasive plants compared to the scale of remote sensing observations, then the likelihood of successful mapping will be higher. If we are dealing with small individuals, similar to those in grasslands, then we will need fine resolution hyperspectral data for early detection.
To do detection, we have many options. One example is classification (e.g., hard classifier or subpixel classifier) which was covered in Part 2 or you can use trait-based approach similar to what is discussed in this webinar.

Question 3: I have a question referring to the previous part (part 2), specifically about the spectral angle mapper technique. I wonder how to assess that the angle between the curve in a pixel and the reference curve defines that the mapped material/substance is present in a given pixel. Is there an optimal way to determine the cutoff threshold for the angle value?
Answer 3: The threshold is user defined. This paper has a nice overview of the error metrics and constraints for Spectral Angle Mapper and Spectral Mixture Analysis techniques introduced in Part 2. Optimizing these error metrics is how most people identify the optimal threshold.

Question 4: Can we detect invasive plants using Sentinel 2 images?
Answer 4: Yes, potentially. It depends on the invasive plant, primarily its prevalence, size, phenology, and whether the invasive plants are spectrally different from native plants. Please see responses to questions 1 & 2.

Question 5: How does one create a spectral graph per pixel from a multispectral drone image?
Answer 5: A drone-based multispectral dataset is similar to those we get from Landsat, Sentinel 2 constellation, etc. After preprocessing the data (e.g., geometric, radiometric, atmospheric correction) , we can extract spectral graphs per pixel from drone data using programming languages, such as Python, R, MATLAB or software packages like ENVI or ERDAS.

Question 6: Is there a global equivalent for eddmaps?
Answer 6: Yes, there are. One example is Global Biodiversity Information Facility (GBIF): https://www.gbif.org/.

Question 7: Cogongrass (Imperata cylindrica) is an invasive that grows along highways, radiating away from port facilities. It grows in right-of-ways, then adjacent fields. How big a "patch" is necessary to appear on an RS image? By the time it is large enough to be found on an image, it is difficult to eradicate. Can remote sensing be used here?
Answer 7: The Minimum Mapping Unit (MMU) should ideally dictate the appropriate spatial resolution for a given application. The MMU represents the smallest feature size that can be depicted on a map. Features smaller than the MMU will not be represented. The specific MMU will vary depending on the application or management objectives.

Once the MMU is established, the spatial resolution of remote sensing imagery should ideally be finer than the MMU, although explicit guidelines are often lacking in the literature.

In practice, however, the choice of spatial resolution is frequently influenced by the availability and cost of data, as well as trade-offs with temporal and spectral resolution. With coarser resolution data, the mixed-pixel problem is an issue. Techniques like soft or fuzzy classification methods are commonly employed to handle this issue, where each pixel is assigned varying degrees of membership across multiple classes. A prominent approach is spectral mixture analysis, which models the spectra of each pixel as a linear combination of pure spectra from different landscape components (endmembers like bare soil, water, or vegetation). This allows for the mapping of each pixel as a fraction of different endmembers. This topic was explored in detail in Part 2.

Question 8: How would one create a spectral signature for a specific plant from drone multispectral imagery?
Answer 8: As discussed in question 5, after preprocessing the data (e.g., geometric, radiometric, atmospheric correction), we can extract spectral graphs per pixel from drone data using programming languages, such as Python, R, MATLAB or software packages like ENVI or ERDAS. To achieve this goal, accurate geometric correction of the data is necessary so that each pixel is in the correct location.

Question 9: How effectively can the maximum entropy (MaxEnt) algorithm be applied to this use case?
Answer 9: MaxEnt is a species distribution model (SDM). The focus of this talk was direct mapping of invasive plants but we can use MaxEnt and other SDMs to model the spatial distribution of invasive plants. To develop an SDM, we will need additional environmental variables, such as temperature and topography.

Question 10: What hyperspectral imager did you use?
Answer 10: We collected airborne hyperspectral data (AISA Fenix 1 k, Specim, Oulu, Finland). The data covered the 400–2450 nm range in 323 spectral bands with a spectral resolution of approximately 4.5 nm in the 400–970 nm range and 14 nm in the 970–2450 nm range.

Question 11: How do you assess accuracy?
Answer 11: We can use metrics called classification accuracy assessment metrics as discussed in slide 35. To conduct accuracy assessment, we used field-based data and compared them to remote sensing classification results.

Question 12: Which model is used to determine nitrogen content in *L. cuneata*? What were the model's inputs and function for determining nitrogen content in the *L. cuneata* field? An example of such a case would be helpful.
Answer 12: We used a data-driven model called partial least squares (PLSR). Independent variables were spectral reflectance data and dependent variables were functional traits (e.g., foliar nitrogen content) that we quantified from foliage samples collected in the field. More details can be found in the paper:
URL: https://doi.org/10.1016/j.rse.2022.112887

Question 13: Is EMIT open source? How can I access its data? Does it require Python code?
Answer 13: Yes, EMIT is publicly available. Please check these websites for info on EMIT and how to access the data: Data Products | Data – EMIT (nasa.gov).
Also: LP DAAC - EMIT Overview (usgs.gov)
https://earth.jpl.nasa.gov/emit/data/data-portal/Greenhouse-Gases/

Question 14: I guess the functional traits vary temporally over the growing season. Did you take that into account when mapping these traits? How does this affect the classification?
Answer 14: Great point. Yes. traits vary in time. To minimize the impact of temporal variability of traits, we collected our remotely sensed data and field data as close as possible and within a few days. But the temporal mismatch between remotely sensed data and field data is a source of uncertainty as traits vary significantly over time.

Question 15: Could you explain a little bit about the technique of image fusion to improve resolution? Is it pan-sharpening in another language?
Answer 15: There are many methods for image fusion. In our case, we used a method called Hyperspectral Image Superresolution (HySURE). HySURE is a subspace method that extracts endmember information from the coarse spatial resolution hyperspectral data and estimates the abundance fractions of the endmembers from the fine spatial

resolution multispectral data. We conducted this analysis in MATLAB. More details on fusion can be found here: https://ieeexplore.ieee.org/document/7946218
Also, you can find a great collection of studies and freely available codes in this GitHub repository:
https://github.com/junjun-jiang/Hyperspectral-Image-Super-Resolution-Benchmark?tab=readme-ov-file

Question 16: What are the main challenges in applying hyperspectral remote sensing for invasive species monitoring in grasslands?
Answer 16: Data availability (e.g., access to data with fine resolution), pre-processing particularly atmospheric correction, uncertainty in our analysis, and large data volume. It is also preferred to collect field data to develop models and validate them, so another challenge is conducting field campaigns, such as harvesting and storing foliage samples.

Question 17: I would like to ask what the process of resampling data up to 3m looked like (DESIS to PLANET), was it "traditional" like bi-linear? Was another algorithm or machine learning method used?
Answer 17: Great question. We fused the data from DESIS with those from Planet using image fusion. Please see answer to question 15.

Question 18: Is there a repository where we are able to access the indices for common invasive species or does one have to craft them individually/refer to available journals?
Answer 18: In a separate project that we did not discuss in the webinar, we used vegetation indices obtained from multispectral data to map invasive species. Something important to consider when using vegetation indices for mapping invasive species is that there should be a justification behind the choice of vegetation indices. In other words, different invasive species might require different vegetation indices for satisfactory mapping. Please also see our response to question 23 and this book:
URL: https://doi.org/10.1201/9781315159331

Question 19: How would you suggest we best navigate with remote sensing of invasive plants in areas with more tropical climates where phenology and functional trait diversity has much greater variation?

Answer 19: In cases where hyperspectral data are not available, mapping invasive species in ecosystems with high spatio-temporal variability will be more successful if we have access to data with fine temporal and spatial resolution. Some options include collecting data using commercial satellites (costly option), using readily available multispectral data such as PlanetScope (cheaper and free under educational license), or using drone-based data which cover smaller areas.

Question 20: How does the team see the future of invasive species mapping with the launch of the Landsat Next?
Answer 20: Landsat Next will be an improved version of Landsat 8/9 with finer spatial, spectral, and temporal resolutions; therefore, we expect that it will improve our capability to map invasive plants compared to Landsat 8/9 and Sentinel 2 constellation.

Question 21: Can you please explain more on how to find traits that can distinguish my particular weed from the rest of the plants? Is it possible to have a singular trait with which can distinguish multiple weeds (10 or more), if yes, how can I find those traits?
Answer 21: In our case, we started with looking at previous literature to identify key functional traits that distinguish *Lespedeza cuneata* from native plants and lead to its success. Examples of the traits that we selected included nitrogen content, chlorophyll content, and height. We focused on those traits that are remotely-observable as not all traits can be estimated from remotely-sensed data.

Yes, it is possible. If the weeds of choice have common traits that distinguish them from co-occurring native plants, those traits can help with mapping invasive plants. Based on what we have learned from our previous experiments, including more than one trait can improve mapping results.

Question 22: Can we combine multiple traits for distinguishing multiple weeds?
Answer 22: Yes, in the case study presented in the webinar, we used 12 traits together to distinguish the invasive plant from native species. Based on what we have learned from our previous experiments, including more than one trait can improve mapping results. Please also see our response to question 21.

Question 23: Would machine learning with accuracy tuning using proxy data sets and those from indices (such as NDVI, WQI, Built-index, etc.) enhance the effectiveness of invasive plants detection?
Answer 23: This is a great question. Yes, in a separate project that we did not discuss in the webinar, we used vegetation indices obtained from multispectral data to map invasive species. While imaging spectroscopy or hyperspectral remote sensing imagery has shown its effectiveness in detecting invasive plants, challenges such as limited access to hyperspectral imagery, coarse temporal resolution, high costs, and lack of standardized analysis approaches have limited their application in monitoring biological invasions. So, instead of estimating functional traits from hyperspectral data (like what we discussed in the case study presented in the webinar), we used vegetation indices as proxies of plant functional traits. This approach showed lower overall classification accuracy for detecting *Lespedeza cuneata* compared to the case study presented in the webinar. The lower detection accuracy of this alternative approach is likely due to the use of vegetation indices estimated using only two to three spectral bands instead of using more than 200 spectral bands provided by hyperspectral data. Moreover, vegetation indices do not estimate traits directly but are proxies of these traits. Nevertheless, our results showed that mapping invasive plants using vegetation indices from multispectral data is feasible. Something important to consider when using vegetation indices for mapping invasive species is that there should be a justification behind the choice of vegetation indices. In other words, different invasive species might require different vegetation indices for satisfactory mapping. More details can be found in paper below:
URL: https://doi.org/10.1080/01431161.2023.2275321

Question 24: Can you explain more about the data fusion process (e.g. Planet + DESIS that was mentioned), or link to resources where we can learn more?
Answer 24: There are many image fusion techniques. You can find a great collection of studies and freely available codes in this GitHub repository:
https://github.com/junjun-jiang/Hyperspectral-Image-Super-Resolution-Benchmark?tab=readme-ov-file
Please also see our response to question 15.

Question 25: If using drone imagery flown at 30m, is it still necessary to do preprocessing such as atmospheric correction?

Answer 25: Yes, it is strongly recommended to atmospherically correct spectral data collected from drones even if the data are collected at low altitudes.

Question 26: What part of the presentation was the phenological events that help distinguish between species covered?
Answer 26: Phenology was briefly discussed on slide 13. But in the case study presented in the webinar, we focused on trait-based mapping at one point in time, and phenology was not discussed in the case study, primarily because we did not have multitemporal airborne hyperspectral data as collecting multitemporal airborne hyperspectral data is costly. But in a separate study (not discussed in the webinar), we used multitemporal multispectral data to detect invasive plants. URL: https://doi.org/10.1080/01431161.2023.2275321
Please also see our response to question 23.

Question 27: Are there any free hyperspectral data sources?
Answer 27: All NASA hyperspectral data are free and open access. See response to question 13 about how to access EMIT data (spaceborne). There are also a number of airborne hyperspectral instruments that collect data and are available for free and open access, for example the AVIRIS and AVIRIS-NG sensors (airborne) discussed in Part 2.

ARSET has some useful trainings on these topics:

- Hyperspectral Data for Land and Coastal Systems
- Biodiversity Applications for Airborne Imaging Systems

Question 28: Are there any computational imaging techniques that allow us to get data from multispectral drones that would be competitive enough with hyperspectral imagery to detect invasive species?
Answer 28: The advantage of multispectral data collected from drones is their fine spatial resolution (if flown at low altitudes) which can significantly improve detection results. Other advantages include lower data volume compared to hyperspectral data and lower computational cost for data processing. Additionally, drones allow us to gather data over time, improving our ability to detect phenological changes in invasive plants. The main limitation of drone-based data is the limited area coverage due to battery life. In addition, flying drones out of line of sight is not allowed in some countries.

Question 29: With, for example, 320 bands, how do we go about choosing the correct ones? What composites are useful?
Answer 29: We used all the spectral bands in our airborne hyperspectral data. The only bands we did not use were noisy and water vapor absorption bands. That being said, some bands have higher “importance” at estimating specific traits and mapping invasive plants. Please see figures 7 and 10 in this paper:
URL: https://doi.org/10.1016/j.rse.2022.112887

Question 30: Do we require image fusion if we have multispectral sensor drone imagery and we upscale it to medium resolution Sentinel data?
Answer 30: If you have fine-resolution drone imagery, you can use them for mapping without fusion with Sentinel data. In our case study, we used image fusion to fuse spaceborne hyperspectral data with coarse spatial resolution and multispectral data with fine spatial resolution to create a hyperspectral data with fine spectral and spatial resolution. Because multispectral drone imagery already has very fine spatial resolution, you do not need to fuse it with another multispectral data with coarser spatial resolution.

Question 31: Do hyperspectral images help with Lantena identification? Also, what are the best methods and instruments for biochemical sampling? How can I get the methodology for classifying L. cuneata species?
Answer 31: Hyperspectral data, or in general remotely sensed data, can potentially help us map different species. In the case of *Lantana camara*, if we focus our data collection efforts during the flowering period of this plant, we might have a better chance of mapping it. But as discussed earlier in this Q&A document, while remote sensing has great potential at mapping invasive species, there are several factors that can affect the success of detection.

Details on field sampling, methodology, and classification can be found in the paper below:
URL: https://doi.org/10.1016/j.rse.2022.112887

Question 32: What are the most effective spectral bands or indices for differentiating ISPPs from native grassland species?

Answer 32: As discussed in question 29, we used all the spectral bands in our airborne hyperspectral data. The only bands we did not use were noisy and water vapor absorption bands. That being said, some bands have higher “importance” at estimating specific traits and mapping invasive plants. Please see figures 7 and 10 in this paper:
URL: https://doi.org/10.1016/j.rse.2022.112887

In a separate project that we did not discuss in the webinar, we used vegetation indices obtained from multispectral data to map invasive species. More details can be found in our response to question 23 and in paper below:
URL: https://doi.org/10.1080/01431161.2023.2275321

Question 33: Did you use points or polygons for accuracy assessment? What are your thoughts on the controversy of using overall accuracy metric?
Answer 33: We used field-based observations for validating our results. Specifically, we considered two scenarios for our validation. In scenario one, we identified large and homogeneous patches of *Lespedeza cuneata* and recorded the area of these patches using a GPS unit. In the second scenario, we documented the abundance of *Lespedeza cuneata* at 133 equal-sized 60 m by 60 m plots. The reason for including this second scenario is that invasive species do not necessarily always grow in large patches; they can also occur in low abundance mixed with co-occurring native species.

No classification accuracy assessment metric is perfect. Some of these metrics, such as overall classification accuracy, are being widely used in the literature but this does not mean that overall classification accuracy is perfect. As a result, researchers often use more than one metric to report classification accuracy.

Question 34: Can you explain the type of challenges that arise when detecting ISPPs in heterogeneous grassland environments using hyperspectral data?
Answer 34: As discussed in question 16, some challenges include data availability (e.g., access to data with fine resolution), pre-processing particularly atmospheric correction, uncertainty in our analysis, and large data volume. It is also preferred to collect field data to develop models and validate them, so another challenge is conducting field campaigns, such as harvesting and storing foliage samples.