



Questions & Answers, Session 1

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 Cindy Schmidt (cynthia.l.schmidt@nasa.gov) or Amber Jean McCullum (amberjean.mccullum@nasa.gov) your question(s).

Question 1: I've noticed that changing projection of point samples from UTM to GCS to be used with the GPS device or google earth result in a shift in the point locations, which in turn affects the accuracy of the collected reference data. How can we manage this?

Answer 1: Really good question - sometimes what people do is rather than collecting a single individual point, they collect points, say, in a particular area - if you know your pixel size is 30m x 30m, you might collect info in a larger area than that so you can compensate for any difference in those pixel locations.

Question 2: Typically scientific method dictates 20 +1 samples. why is it so much higher for error estimates? (you recommend 50?)

Answer 2: Don't quite understand where the scientific method dictates that - getting samples for doing accuracy assessment is dependent on the number of classes you have and understanding the statistics that go into creating a stratified random sample. Generally, the recommendation is have at least 50 per class.

Question 3: Can we use QGIS for our assignment?

Answer 3: We have great advice coming in, and we highly encouraging using QGIS - we chose Arc because it was a little easier to put together the webinar. If repeating, we'd like to use QGIS, so any advice people are giving online - if you're interested in using QGIS you should go for it! And maybe send us the process you used to get there and some of the issues you've run into.

Question 4: What is the difference between Kappa index and the error index

Answer 4: We chose not to include the Kappa index in this webinar, primarily because there's a lot of evidence in the literature that the Kappa index isn't very useful to use. So we chose not to include the Kappa index in our webinar. Traditionally it was used in the past, but after the other accuracy was done, a lot of times people would use Kappa to give you another good measurement of accuracy. I think next week when we do unbiased area estimation, that will be much more useful to you.



Question 5: After completing the extract values to points, the attribute table ends up having many RASTERVALU stating -9999 instead of a raster class (using ArcGIS 10.5) But the points are being provided by the shape file in the exercise, and they appear correctly?

Answer 5: If this occurs, I would try completing that step again. Having a value of -9999 indicates there are “no data.” But if you are using the files we provided, this should not be an issue. Just make sure you download that Landsat .tif as well as the reference data we provided on the website. If you’re creating a table where you don’t see any values or the file can be opened in ArcMap - recommend doing those steps again very slowly in the exercise doc. If you’re using Arc 10.0 or higher, you should be able to follow along without any issue. If you have the problem repeatedly, I also suggest saving arcmap document and restarting Arc. Another tip is make sure you’re saving the files to your computer - sometimes if you’re working off a thumb drive, there might be issues with opening the data or saving it in the right place.

Question 6: What was your reasoning to not simply rename the fields with the alter field tool?

Answer 6: Initially, we were not aware of that tool. However, it seems that if you use that function, the correct names will not be stored in files that you export in the subsequent steps of the exercises. So we recommend following the exercise as shown.

Question 7: Please the classification depend of the quality of spatial resolutions or others aspects ?

Answer 7: We don’t really have time to answer classification questions in this webinar. If you do have questions, look at our webinar from last year on [land cover classification](#) and if you still have questions, you can email us and we’ll be able to help out (hopefully) or give you info that will help.

Question 8: How do you minimize the error of commission?

Answer 8: The use of the accuracy assessment allows you to investigate the errors to see where they are occurring, and if you need to go back and redo your classification. For example: you might have errors of commission in vegetation - in areas that have vegetation but they occur on different elevations. So you might be able to bring in an elevation dataset that helps you separate different vegetation types depending on where it is on the elevation spectrum. So there’s instances where you can go back and affect your classification based on the results of your accuracy assessment.



Question 9: Can you use Google Earth to generate reference points or to evaluate the land cover classes in your Landsat imagery?

Answer 9: There are good resources out there on the best way to do that - any kind of imagery is a higher spatial resolution than the land cover classification will give you better info than having no info at all. We'll get you info on best way to use Google Earth imagery.

Question 10: If you are working with 2015 images, is it okay to use 2016 reference data? Assuming that is the only close reference data available.

Answer 10: Ideally, you want to use reference data that's exactly the same time as your classification because if any changes have occurred in that time (in this case between 2015 & 2016) then your reference data will be different than your classification data. But if you know where the change has occurred - if that's the only reference data you have, you can go ahead and use it.

Question 11: What spatial resolution (raster pixel size) do you suggest to use for a national scale in a land cover generation of 7 classes based on remote sensing?

Answer 11: We aren't really focusing on land cover classification in this particular webinar. We did a land cover classification webinar about a year ago, and we suggest you take a look at that. I am not an expert at national land scale generation, but I know there are some global efforts going on to map land cover on a national scale using landsat spatial resolution imagery - around 30m.

Question 12: How we can explain the reason for low and high percentages for user's accuracy and producer's accuracy?

Answer 12: Those are errors - the low percentages are errors in your classification. That's letting you know that some of your classes are being confused with each other. At that point, you need to go back to your classification and see where the confusions are occurring and see where you can improve the classification with ancillary data, divide your region into bio regions, or something like that.

Question 13: Are there any recommendations (or rules of thumb) on defining the optimal sample size with respect to the area of analysis (pixels) and the number of classes?

Answer 13: There are fairly rigorous statistical processes you can use to determine your sample size. We aren't going through that in our webinar this week, however, I know that in practice a lot of people will use a minimum of 50 points per class - that's



just a ballpark - it's also dependent on what you're using for your reference data. If you're doing fieldwork to collect reference data, it may be really difficult to get 50 points per class because of limitations of field work in addition that it's very expensive. If you're using aerial photography or high res spatial imagery, then you might be able to get at least 50 points per class.

Question 14: When you are doing a Oriented Object Classification, how can i get theses accuracy assessment?

Answer: Using segmentation, it's the same process. You get reference points and relate reference points to the segments, essentially.

Question 15: In these exercise of land cover classification was supervised or unsupervised, the error variated in those cases?

Answer: Is the error different depending on if you're using supervised or unsupervised - every classification will be different whether you're doing supervised or unsupervised, so there's no good way to answer that question. What we provided you with was a map generated by using supervised classification.

Question 16: Can we use this method for accuracy assessment with polygons as referenced/classified data, instead of points?

Answer: Yes, you can. You can use polygons instead of points and collect sort of average info about the land cover class within a polygon. There's various ways to do it - it's easier with the points, but then you run into issues with the accuracy of the points.

Question 17: In the case of this exercise, how did you choose the 30 m cell size of your raster dataset? Is it based on the scale, number of classes, spatial resolution of your imagery?

Answer: It's based on Landsat imagery for this exercise, as we do in most exercises, and the cell size of Landsat imagery is 30 m - we didn't choose it, that's the characteristics of this dataset.

Question 18: What about working with pansharpening of images?

Answer: Not sure what the question is exactly - if you're asking about using pan sharpened images for your reference data - that will often work. It's similar to using any high resolution spatial imagery where you're getting slightly better data than a non-pan sharpened image. You could use it for your reference data if you didn't have anything else.



Question 19: How to handle pixel shifting in time series data (for example with Landsat 5, 7 and 8) for classification?

Answer: The first thing you should always do with imagery, if you're doing time series and looking at change detection or something like that. Is overlay imagery on top of each other to determine if there's pixel shifting, and you have to fix that **first** before you do any other processing. This is more addressing time series issues rather than accuracy assessment.

Question 20: Can we produce reference data from NDVI calculation?

Answer: That's probably not a preferable choice for reference data. If you have an NDVI calculation and you're trying to identify a land cover classification, then that doesn't make much sense.

Question 21: Any suggestions for improving accuracy when you have spectral mixing within a pixel?

Answer: Not exactly sure if I understand the question - all pixels, if you're using landsat or whatever, will have a combination of land covers within that pixel (most of the time). Within a pixel, you might have more than one land cover class - but you might have a dominant class in that pixel. Depending on how you're collecting the accuracy in that info you have to understand if there's more than one land cover type in that pixel, you choose the dominant type. That's one way of addressing that question.

Question 22: Can error be explained by the pixel size as well and proximity with the boundary between two different classes?

Answer: Absolutely - error can be explained by a lot of different things. The pixel size is one factor, depending on the detail of your reference data. If you have a resolution of data with a pixel size of 30 m or more, there can be some error due to those discrepancies