Marcel Rodriguez-Riccelli
1/22/2026
Satellite Image Analysis (GEOG 581)
LAB 1 - IMAGE QUERYING, STATISTICS, CORRELATIONS

QUESTION 1: Using the print() command to investigate the image metadata, identify how many bands are in each image (SR and TOA).

Using the Google Earth Engine Python API has a different mechanism for printing image metadata compared to the browser-based JavaScript IDE. By the following code:

```
# Select first image from each filtered collection
L8SR_image = L8SR_filtered.first()
L8TOA_image = L8TOA_filtered.first()

L8SR_band_list = L8SR_image.bandNames().getInfo()
L8TOA_band_list = L8TOA_image.bandNames().getInfo()

L8SR_num_bands = len(L8SR_band_list)
L8TOA_num_bands = len(L8TOA_band_list)

print("SR Bands:", L8SR_band_list)
print("SR Number of Bands:", L8SR_num_bands)
print("TOA Bands:", L8TOA_band_list)
print("TOA Number of Bands:", L8TOA_num_bands)
```

I was able to print:

SR Bands: ['SR_B1', 'SR_B2', 'SR_B3', 'SR_B4', 'SR_B5', 'SR_B6', 'SR_B7', 'SR_QA_AEROSOL', 'ST_B10', 'ST_ATRAN', 'ST_CDIST', 'ST_DRAD', 'ST_EMIS', 'ST_EMSD', 'ST_QA', 'ST_TRAD', 'ST_URAD', 'QA_PIXEL', 'QA_RADSAT']

SR Number of Bands: 19

TOA Bands: ['B1', 'B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'B8', 'B9', 'B10', 'B11', 'QA_PIXEL', 'QA_RADSAT', 'SAA', 'SZA', 'VAA', 'VZA']

TOA Number of Bands: 17

From this printout, we can see that the Surface Reflectance imagery has 19 bands, and the Top of Atmosphere has 17.

QUESTION 2: What band numbers would you use to display true or false color?

For the visualization parameters I used to display a true color image from the Surface Reflectance dataset, I used the following dictionary:

```
# Visualization parameters for SR RGB
rgb_vis = {
  "min": 0,
  "max": 0.4,
  "bands": ['SR_B4', 'SR_B3', 'SR_B2']}
```

that includes Band 4 (Red), Band 3 (Green), and Band 2 (Blue). For the false color image, I used a different dictionary:

```
# Visualization parameters for SR False
false_vis = {
  "min": 0,
  "max": 0.4,
  "bands": ['SR_B5', 'SR_B4', 'SR_B3']}
```

which instead includes Band 5 (Near Infrared), Band 4 (Red), and Band 2 (Blue).

QUESTION 3. How many rows, columns, and total pixels are in the red band?

When running the following command to print the image metadata:

```
L8SR_info = L8SR_image.getInfo()
L8TOA_info = L8TOA_image.getInfo()
```

SR_B4 lists the following dimensionality:

'dimensions': [7791, 7901]

This tells me that the red band contains 7791 rows and 7901 pixels, for a total of 61,556,691 pixels.

QUESTION 4. In two sentences, discuss any visual differences in false color composites between the two different Landsat 8 images. Are changes slight throughout the image or pronounced in some regions? Does changing your stretch make it easier to see differences?

The first obvious difference between the Surface Reflectance and Top of Atmosphere imagery is in the cloud coverage; while the most dense clouds are still present in both images, we can see that the extent of these clouds off of the coast are larger in the Surface Reflectance dataset than in the Top of Atmosphere (*Fig. 1)*, and so can assume that the less dense edges of the clouds have been corrected by the atmospheric correction algorithm used to generate the Surface Reflectance dataset from the raw data. Throughout the image, we can also see

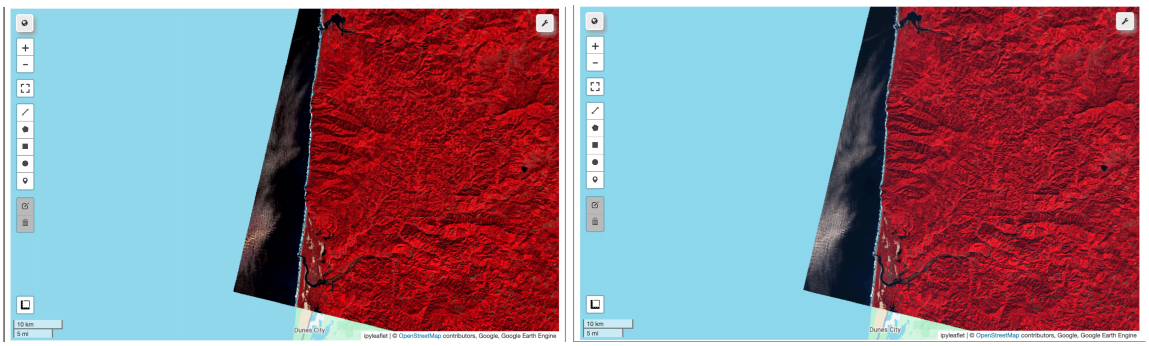


Fig. 1. Example of differences in cloud extent; Left: Landsat 8 Surface Reflectance, Right: Landsat 8 Top of Atmosphere

differences in saturation and brightness throughout *(Fig. 2)*, some of which may be ascribable to the presence of clouds in the Top of Atmosphere image, and others which may be an article of the uniform stretch being applied to each image that is agnostic of the actual range of each band.

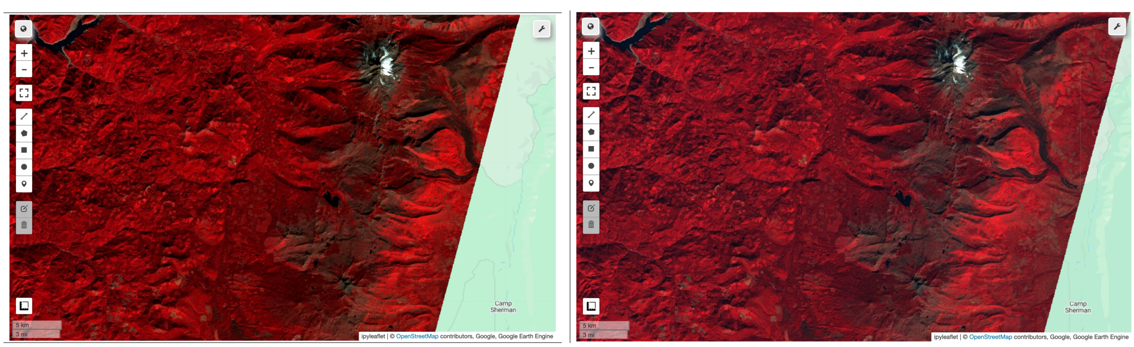


Fig. 2. Example of differences in image saturation; Left: Landsat 8 Surface Reflectance, Right: Landsat 8 Top of Atmosphere

Modifying the minimum and maximum visualization parameters for each image is necessary for identifying actual differences between the two datasets; by reducing the amount of artifacts that are a result of the indiscriminate stretching of each band, we can more accurately detect differences that are instead attributable to the differences between data products.

QUESTION 5. Now, identify two land cover types (see FAO documentation for understanding land cover rather than land use) or specific features that look very *similar* between the two images (SR vs. TOA), and two land cover types or specific features that look very *different*. In what ways do these land cover types or features look different or similar?

A land cover type that appears to be very similar between the Surface Reflectance and Top of Atmosphere can be seen at the aquatic feature Fern Ridge Lake *(Fig. 3)*. Aside from the differences in saturation noted earlier, we can see the same detailed color differences caused by differences in suspended particulate matter distributed non-uniformly by eddies present in the lake fluid. A land cover type where we can see significant differences is in the snow on top

Fig. 3. Similarities in land cover features between Surface Reflectance and Top of Atmosphere datasets exemplified by imagery of Fern Ridge Lake; Left: Landsat 8 Surface Reflectance, Right: Landsat 8 Top of Atmosphere.

of Mt. Jefferson *(Fig. 4)*. Much of the detail from the difference in shadow from the drastic relief in the mountain top covered by snow that is seen in the Top of Atmosphere image is lost in the Surface Reflectance image.

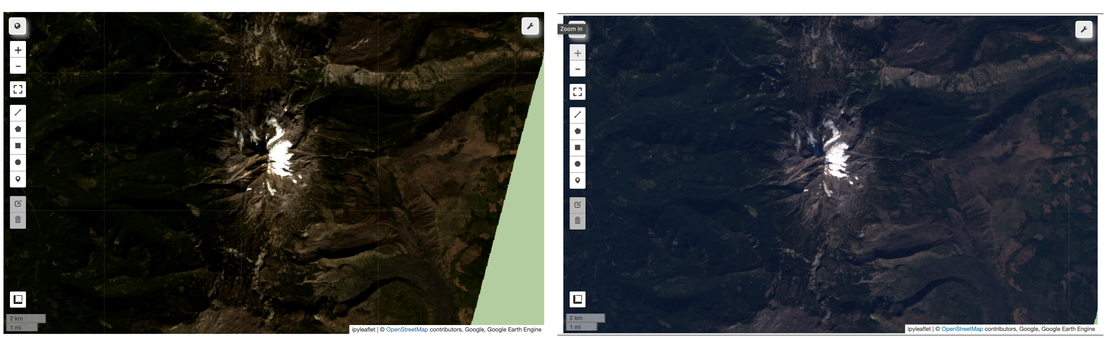


Fig. 4. Difference in land cover features between Surface Reflectance and Top of Atmosphere datasets exemplified by snow atop Mt. Jefferson; Left: Landsat 8 Surface Reflectance, Right: Landsat 8 Top of Atmosphere.

QUESTION 6. Should we use the same range in visualizing the false color and true color images? Why or why not?

I think it depends on the analysis use-case, but one should generally use different ranges when visualizing false color and true color images. Each band has a distinct bandwidth, and the sun does not emit light evenly across all bands, so the range of values between bands will vary greatly. Since the general goal is to increase contrast throughout the image, each band should be stretched to increase contrast according to its bespoke range of values.

QUESTION 7. Please identify one pixel location at each of the four feature types, and list the latitude-longitude coordinates for each site to four decimal places.

1. Forested - Outside of Harlan, Oregon, Point (-123.6743, 44.5437)
2. Urban - Salem, Oregon, Point (-123.0361, 44.9436)
3. Snow - Mt. Jefferson, Point (-121.7916, 44.6769)
4. Cloud - Just west of Gales Creek, Oregon, Point (-121.7916, 44.6769)

QUESTION 8. Using the Landsat 8 SR product, which land covers are generally more distinct in the false color composite compared to the true color image?

Details in certain land types such as forested areas and grass areas are much more distinct in the false color image when compared to the true color image. This is to be expected; we know from class that the Near Infrared Band is commonly used to differentiate between healthy and unhealthy vegetation.

QUESTION 9. Comparing the Landsat 8 SR and TOA images, how does the multispectral reflectance of dark water pixels differ in the three visible light bands? Make sure to specify differences in each of the three bands visualized.

I used the inspector tool to get the following visible light band reflectance values for a dark pixel in Fern Ridge Lake at (-123.3282, 44.0828):

Surface Reflectance

SR_B2: 0.052092499999999986
SR_B3: 0.0814625
SR_B4: 0.078575

Top of Atmosphere

B2: 0.11750970035791397
B3: 0.10682699829339981
B4: 0.09258340299129486

From this data, we can see that the Top of Atmosphere shows higher reflectance values for all three bands. We also see a lower range of values between the three bands in the Top of Atmosphere values for this pixel. This tracks with the previous note made about the difference in saturation and brightness between the bands *(Fig. 3)*; higher values will result in brighter pixels, while a lesser range will result in desaturation.

QUESTION 10. Compare and contrast the brightness characteristics of snow features in the NIR bands of the Landsat 8 SR and TOA images.

I used the inspector tool to get the following visible light band reflectance values for a snow pixel atop Mt. Jefferson at (-121.7908, 44.6748):

Surface Reflectance

SR_B2: 0.41099500000000005
SR_B3: 0.43656999999999996
SR_B4: 0.44955000000000006

Top of Atmosphere

B2: 0.4137834310531616
B3: 0.4074491262435913
B4: 0.4259041845798492

The values are somewhat close, but we can see that the Top of Atmosphere has lower values across all bands and is therefore darker. This can be linked directly to the lack of relief detail noted as a response to an earlier question. *(Fig. 4)*

QUESTION 11. What does the '38m/px' text refer to in the Inspector window? Hint: see the "m / pixel (on Equator)" table field in the "Zoom levels" page here.

The 'Xm/px' text refers to the spatial extent of one pixel currently being displayed in the map window. According to the OpenStreetMap Zoom Levels table, the spatial extent of a pixel is 38m/px when the Zoom level is 12. Google Earth Engine uses the assumption that all pixels were on the equator (0° latitude), and to get the true spatial extent per pixel, we should multiply this value by the cosine of the latitude. For the point (-122.9058, 42.3539), at Zoom level 12 at 38m/px, the actual spatial extent is given by:

$$38 \times \cos(42.3539) \approx 28.08 \ m/px$$

QUESTION 12. In your own words, explain what each line of this script does.

Here's the same code, syntactically modified for use with the Python Google Earth API:

mean_dictionary = L8SR_image.reduceRegion(reducer=ee.Reducer.mean(), scale=30, maxPixels=1e9)
print(mean_dictionary.getInfo())

The explanation for each line is as follows:

mean_dictionary =

Instantiate a variable named mean_dictionary to store the function output.

L8SR_image.reduceRegion()

Apply the reduceRegion function to the L8SR_image Landsat 8 Surface Reflectance image for our region of interest, which converts the image data into a statistical representation of the image data based on the arguments.

reducer=ee.Reducer.mean()

Set the Reducer type to retrieve a mean value for each band of the image.

scale=30

Specify a grid of 30 meters per pixel to be sampled during the reduction.

maxPixels=1e9

Specify the maximum number of pixels considered in the calculation to prevent computation failure for large regions.

print(mean_dictionary.getInfo())

Print the dictionary with the mean pixel value of each band stored in the mean_dictionary variable.

The console line from the print yields:

{'QA_PIXEL': 21877.403388262373, 'QA_RADSAT': 0.0001275457097965396, 'SR_B1': 0.011170529570103734, 'SR_B2': 0.01679632509854225, 'SR_B3': 0.03433548429530112, 'SR_B4': 0.03245246272256908, 'SR_B5': 0.2080610830929151, 'SR_B6': 0.11118741027800096, 'SR_B7': 0.06128205222864313, 'SR_QA_AEROSOL': 95.90322442166497, 'ST_ATRAN': 7715.283047338548, 'ST_B10': 293.7617752209188,

'ST_CDIST': 781.1459727962192, 'ST_DRAD': 803.8656454086474, 'ST_EMIS': 9810.568480662956, 'ST_EMSD': 92.05128440578929, 'ST_QA': 248.59721907791894, 'ST_TRAD': 8290.729351129188, 'ST_URAD': 1666.2420424347417}

From our arguments, we can assume that each value represents the mean values for each band.

QUESTION 13. You'll see that a reduceRegion() function returns a Dictionary. Using the information under the Docs tab—in your own words, what is a 'Dictionary' in Earth Engine and what is it used for?

A dictionary is a data structure that stores data in key/value pairs. The value for a key (X) can be retrieved by using the .get("X") function. Dictionaries can also be passed in as arguments to functions or objects which expect dictionaries. In Google Earth Engine, dictionaries are stored server-side, and are commonly used to hold metadata.

QUESTION 14. What are the mean values for the near infrared band in each image?

The following code:

print("L8SR Band 5 Mean:", mean_dictionary_SR.getInfo().get('SR_B5'))
print("L8TOA Band 5 Mean:", mean_dictionary_TOA.getInfo().get('B5'))

Output the following:

L8SR Band 5 Mean: 0.2080610830929151
L8TOA Band 5 Mean: 0.2098948093141057

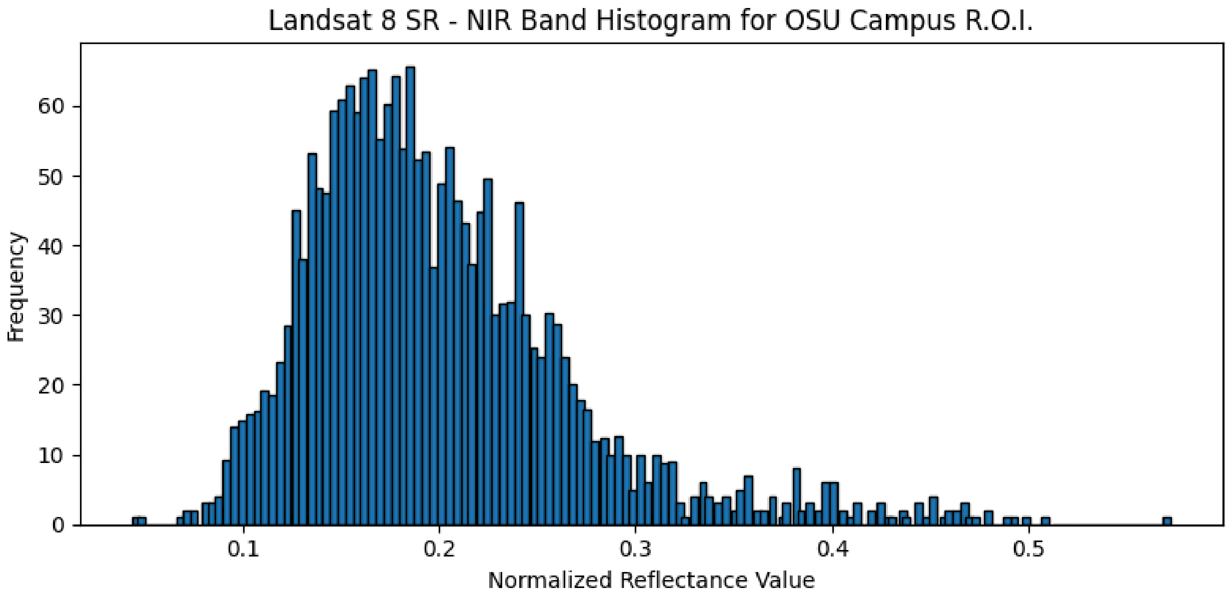
From this print out, we can see that the Near Infrared Band (5) from the Surface Reflectance dataset has a mean normalized reflectance value of ~0.208, while the same band from the Top of Atmosphere dataset has a mean of ~0.209.

QUESTION 15. What is the range of values for the near infrared band in each image? Hint: Look in the Docs tab to find Reducers other than ee.Reducer.mean() that might be helpful to answer this question.

I used the ee.Reducer.minMax() function in the reducer argument field, and referenced the resulting dictionary to get the following output:

L8SR Band 5: Min = -0.013165, Max = 1.5859, Range = 1.5991
L8TOA Band 5: Min = -0.0044511, Max = 1.9424, Range = 1.9469

QUESTION 16. Include an output of both histograms in your submitted answers.



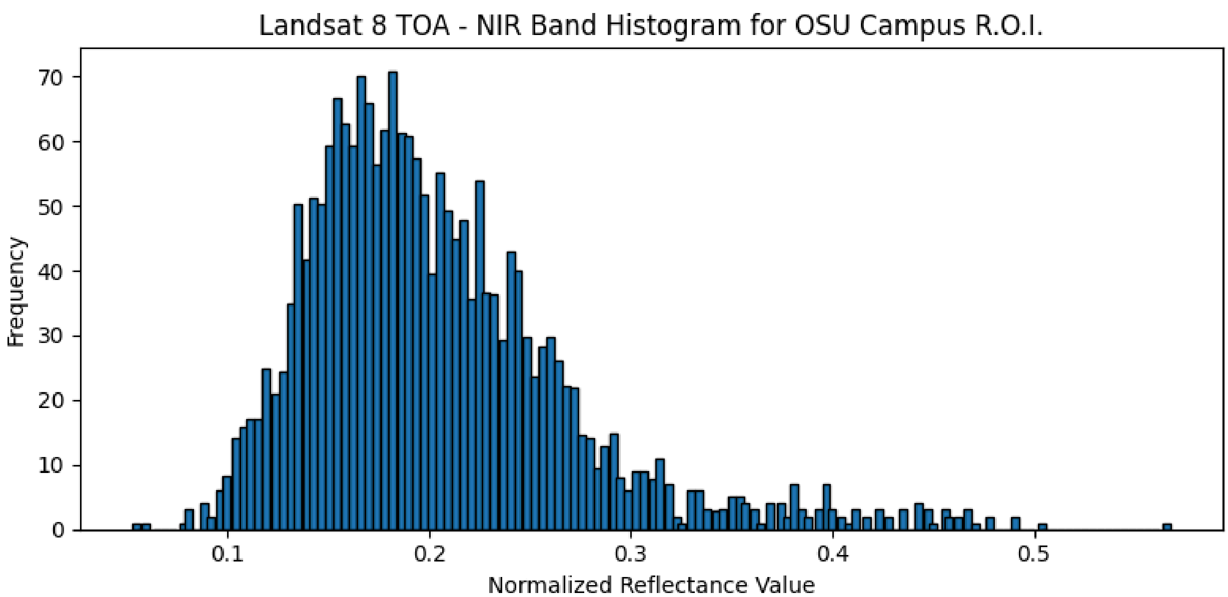


Fig. 5. Comparison of histograms charting frequency of normalized reflectance values for Near Infrared Band for Oregon State University Campus area of interest ; Top: Landsat 8 Surface Reflectance, Bottom: Landsat 8 Top of Atmosphere.

QUESTION 17. What is the apparent (visual) range of values for SR and TOA image bands? How does this band range compare to the image statistics values that you generated above?

Approximating visually from the histograms *(Fig. 5)*, range of normalized reflectance values for the Surface Reflectance and Top of Atmosphere NIR Bands in the histogram appear to be from ~0.05 - ~0.6 for both datasets, as compared to the full image range provided by the ee.Reducer.minMax() function that were from ~0.01 - ~1.5 for SR and ~0.004 - ~ 1.9 for TOA, meaning that our area of interest has a significantly smaller range of values than that of the full image.

QUESTION 18. Is there any apparent difference in the number of peaks (unimodal, bimodal, etc.), skewness, or kurtosis between the two image band histograms?

The peaks are very similar: they are both largely unimodal with a positive skew and a mode of ~1.8. The histogram for the Top of Atmosphere histogram has slightly higher kurtosis than the Surface Reflectance histogram.

QUESTION 19. For each image, are the range of values within your OSU geometry (calculated in your histogram) representative of the values across the entire band (calculated in Q15)?

Neither the range of normalized reflectance values in the Surface Reflectance or Top of Atmosphere histograms reflect the full range of values across the entire image, calculated in the previous problem. I applied the same ee.Reducer.minMax() to our subset bounded by the OSU geometry and printed the resulting metadata to get:

> *{'SR_B5_max': 0.5704125, 'SR_B5_min': 0.045657499999999999}*
> *{'B5_max': 0.5659776926040649, 'B5_min': 0.054337963461875916}*

We can see that the range of the values in each of the subsetted datasets are significantly smaller than the values across the entire band.

QUESTION 20. Based on your visual assessment, which image type (SR or TOA) shows higher correlation between Red and NIR bands?

Although the scatterplots of Red and NIR normalized reflectance values for both the Surface Reflectance and Top of Atmosphere datasets are similar *(Fig. 6)*, the Top of Atmosphere dataset shows a tighter spread of values and therefore higher correlation.

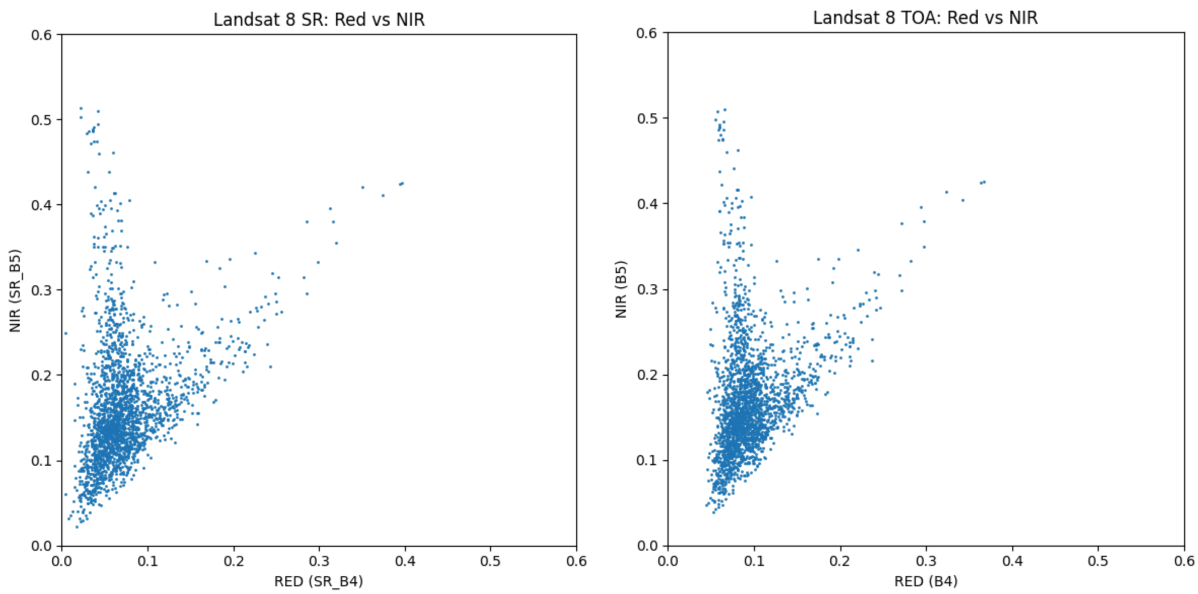


Fig. 6. Comparison of Red Visual and NIR band reflectance values for Oregon State University Campus area of interest; Left: Landsat 8 Surface Reflectance, Right: Landsat 8 Top of Atmosphere.

QUESTION 22. Looking at the code on line 3 of the reference script for this part,

```
var image = ee.Image(L8_SR.filterBounds(region).filterDate("2020-12-01","2020-12-16").first())
```

Please describe the roles of each component of this line of code.

Here's the same code, syntactically modified for use with the Python Google Earth API:

```
image = ee.Image(
  L8_SR.filterBounds(region_geom)
     .filterDate("2020-12-01", "2020-12-16")
     .first()
)
```

Line by line, each component plays the following role:

Image =

Instantiate a variable to hold the output of the ee.Image function, which will be an image object.

ee.Image()

The Image variable will be a type that is the Earth Engine image class.

L8_SR

The L8_SR image collection object previously defined with the full imported Landsat 8 Surface Reflectance dataset.

filterBounds(region_geom)

A function which filters the images available in the L8_SR collection according to the bounds of polygon drawn around OSU campus in the previous part of this lab.

.filterDate("2020-12-01", "2020-12-16")

Another filter function, that after the filterBounds() function, filters the resulting filtered image collection for only images available between December 1st and December 16th 2020.

.first()

A third function in the chain that selects the first image in the collection after it has been filtered by .filterDate() and .filterBounds(), which is the image with the earliest timestamp.

QUESTION 23. The code will not run successfully without using *first()*. Why not? What do the error messages tell you?

Without selecting the first image, the argument for the Image() class would remain a collection of images, as opposed to a single image, as is required. We can confirm this by removing the .first() function from the chain, and reading the error print out:

EEException: Image.visualize, argument 'image': Invalid type.
Expected type: Image<unknown bands>.
Actual type: ImageCollection.

This confirms a single image is required as an input to the Image() class, as opposed to an ImageCollection.

QUESTION 24: Include shareable links from both scripts (at the end of Parts II and III) in your submission.

Parts I & II:
https://colab.research.google.com/drive/1DnmLw-aDPCvFgEShjj-f0rx_7Ldn4LPv?usp=sharing

Part III:
https://colab.research.google.com/drive/11UOABQDwzEspCXfLh5cvNKtsUMLT4RPa?usp=sharing