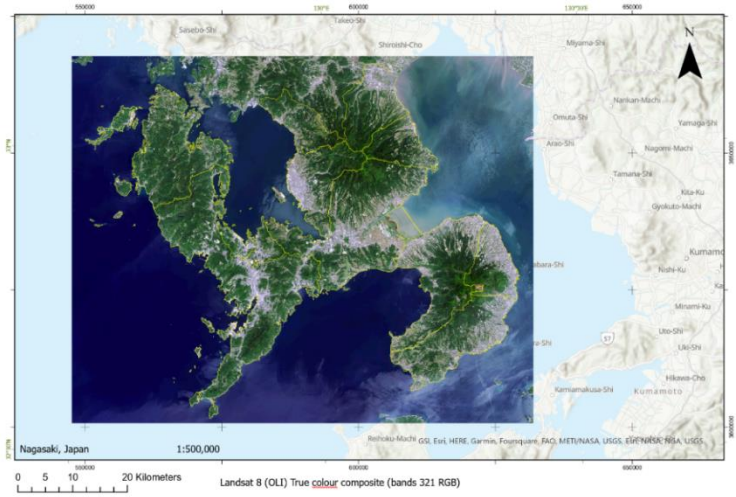
EDSML Environmental Data – Earth Observation week - Group project outline

The task

Imagine you work for a small, start-up company offering data analytics services in environmental science (there are growing numbers of these). Your task is to perform a benchmark exercise to analyse the environmental condition of Nagasaki, Japan, using Earth Observation archive data; the first part of which is to produce a land-cover classification map. This can begin with a single dataset, to develop your workflow and ensure that your processing pipeline works and produces representative classes, and then to build a time-series analysis to examine changes in land-cover. You should try using data acquired every 5 years over the last 25 or perhaps every month for the last 3 years. You can decide this and you will need to consider how much data exists, as well as the likely rate of change for this area.



General strategy

The analysis should begin with generation of colour composites to illustrate the landscape, vegetation and visualisation of the image statistics. You should extract various spectral indices, generating spectral plots etc, and then using classification methods (of your choice) to produce an initial land-cover classification of this region. Next you could segment your extracted land-cover information by administrative region, and then produce metrics of change over time. You could also include analysis of the terrain, e.g. slope angle and aspect, or rainfall data (if you can find any data online), since you may find patterns of landcover type or change that are correlated with particular slopes across the study area. You may also consider animating the classified results to visualise these, or some other novel method of viewing the results, or presenting time-series statistical plots.

We will provide you with a starting Landsat dataset (with spectral bands as individual files, and with all the bands stacked and cropped to the study area!), an Digital Elevation Model (DEM), and administrative boundaries (as vector data, a .geojson format). You will need to develop the basic code to ingest, pre-process and display the images and results. All your plots, images etc must be labelled and displayed in a professional manner.

You should all undertake the exercise but should work as a team (of up to 9 people). Discuss the objective and decide what kind of analysis you want to undertake, decide on tasks that need to be accomplished, and allocated them to members of the team, and thus how to achieve the overall objective. So, within your team/sub-groups, you should also discuss how to approach the tasks, if any other data are available or are needed (?) and how best to display and visualise and present the data and results. You have from until Friday morning and the end goal is for each team to present their approach and outputs (in a professional manner). Each team will have 10-15 minutes to present their work and findings to the class.

How to start the process and to organise your group task

1. Break down the top-level goal: 'benchmark environmental assessment/analysis of the Nagasaki area'
    a. Hence this is an assessment of the state of the environment now
    b. But it could include a retrospective analysis of the past up to this point, and it could include recommendations for future monitoring methodology
2. Limit the scope of the exercise (i.e. make things a little easier for yourselves)
    a. Spatially - crop the data to a limited geographic area; and/or
    b. Temporally – analyse only one image or analyse once a month for a year, or every season or every year for a number of years. You will need to decide what is appropriate, e.g. change may or may not be rapid enough to detect every month, but detectable over a period of years. So you can search for suitable (cloud-free) Landsat data in the USGS archive acquired over the past 30+ years
3. What factors affect the environment?
    a. Topography – you have a Digital Elevation Model and can derive slope, aspect, hillshade etc
    b. Coast-line – you can extract this from the satellite image but may be able to find vector GIS data online
    c. Landuse/landcover – e.g. forest, urban, grassland, agriculture, upland etc – and you can derive this information through classification of your multi-spectral satellite data (Landsat in this case)
    d. Weather & climate – you may be able to find rainfall, temperature etc records for Japan online somewhere 9this may be a document or a point file or a spreadsheet
4. Processing, integrating and analysing the data (the first 3 parts take a bit of time but this is the interesting part)
    a. You have some image processing skills now – reading the metadata to obtain pixel size and coordinates etc, colour composites, spectral indices, classification, hillshade, slope and aspect – so you can employ these skills and begin to extract information from your data. You also have some polygons marking the local administrative (county) boundaries so you could use these to constrain any statistics you derive, e.g. percentage forest cover per county.

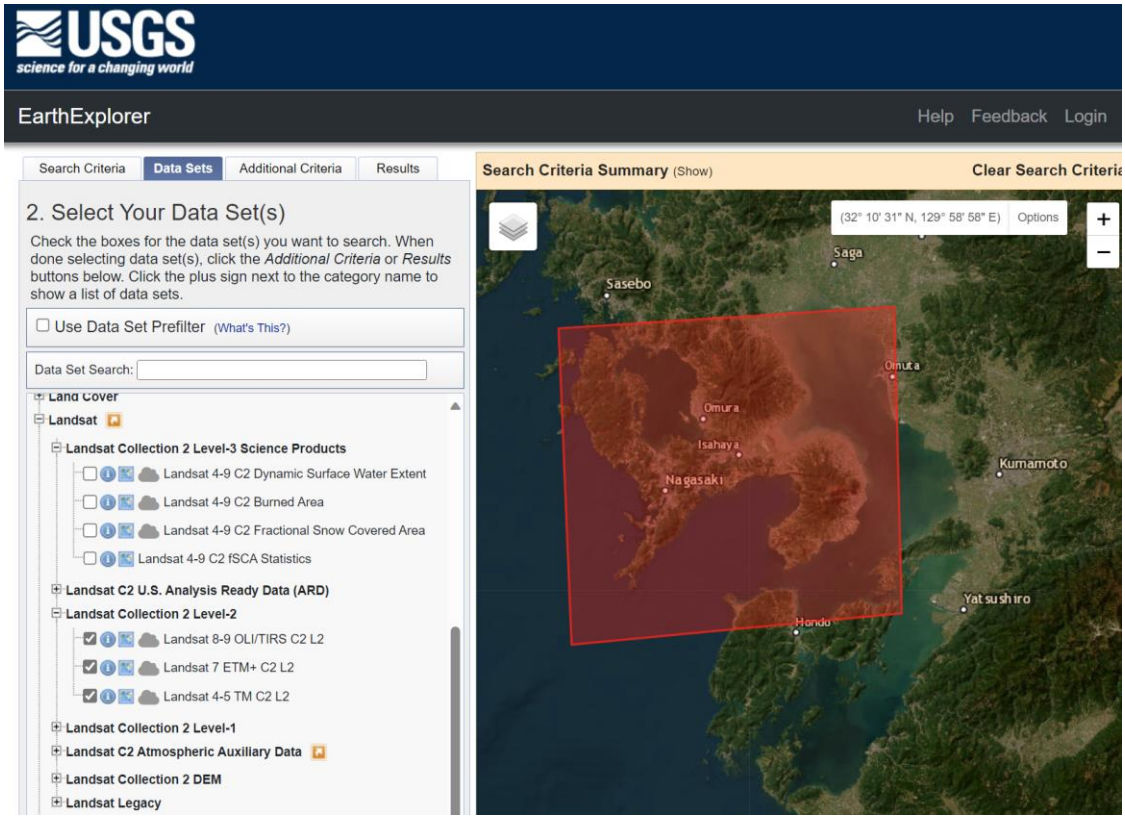
Starting materials provided

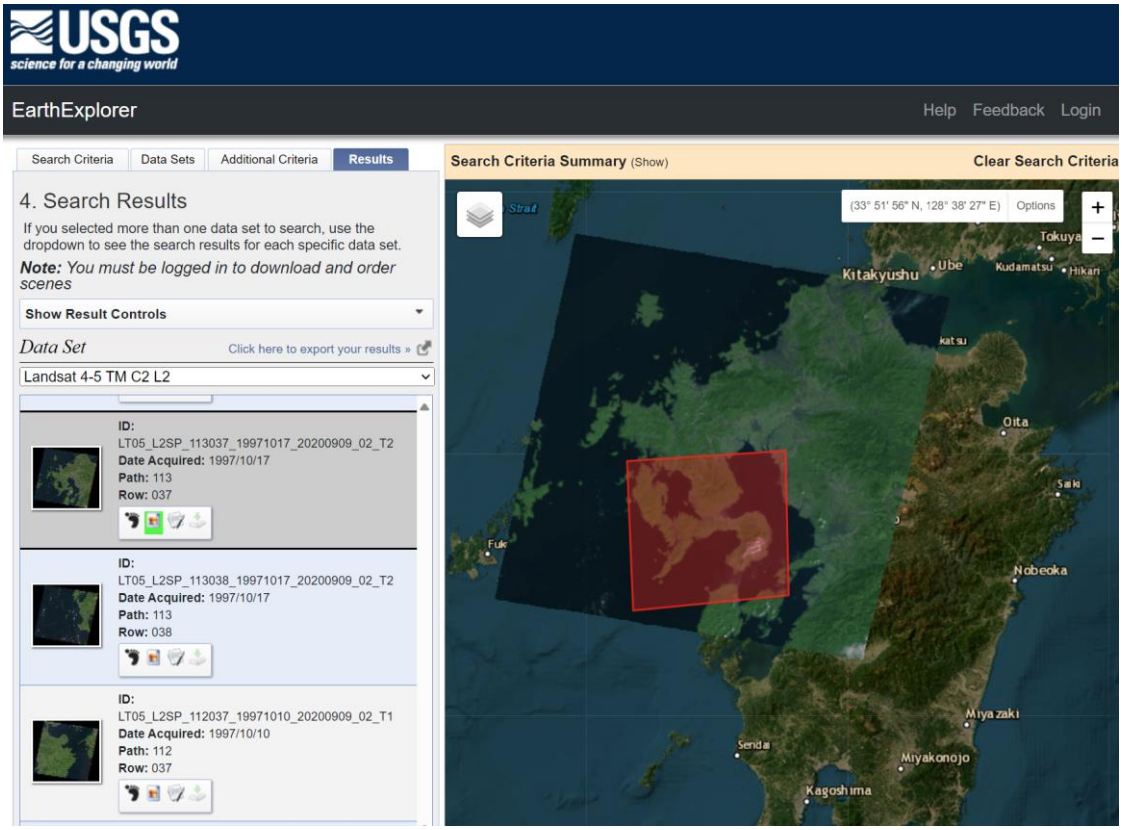
- Landsat 8 image bands (Geo tiff, path 112, row 037, acquired May 9th 2023)
- Landsat 9 image bands (Geo tiff, path 112, row 037, acquired May 12th 2023)
- SRTM DEM (4 Geo tiff tiles)
- Administrative boundaries (shapefile and .geojson)
- An ArcGIS Pro Project has been created which shows the area and some of the data. ArcGIS Pro is available under College license for Windows and is available via the College's Apps Anywhere Software Hub. It can be accessed by Mac users via the Azure Virtual Deskop. You could use QGIS instead if you have this, this can also be used to load up the data)

All available for download (read only) here.

For more Landsat data (for time-series analysis, for instance):

USGS EarthExplorer https://earthexplorer.usgs.gov





Bear in mind that Landsat 7 developed a line-scan correction fault in 2004 so only L7 data before that time are usable.

Also that you are likely to encounter significant cloud in this part of the world!! Which will limit the usable data available to you.