GEOG357 Fall 2022: Supervised classification

Learn the process; similar assignment to last week: send text and classification TIF layer to wheate@unbc.ca …

0. Start

In Catalyst, Open your copy of the Landsat 5 TM image: **pg14sept2011.pix**
Change display to bands 543 and enhance (root)

Expand the image display so it doesn't need panning

Check your pix file for bands/channels:

Click the + symbol to the left of the .pix filename – it will simply show what is in RGB (543)
The 'maps' tab reflects the display, while 'files' tab indicates the file data

Change to the 'files' tab and click the + next to rasters; this will show your channels – you should have bands 1-7 and then ~4 unsupervised CHANNELS (not bands) and maybe a sieve, plus one or more bitmaps

Switch back to 'maps' tab.

1. Supervised Classification session: Prince George 2011 image

Unsupervised classifications are quick, supervised are not as quick due to the need to identify cover types and create training sites; I'm picturing 10-15 classes for the image: e.g. Water, wetlands, Coniferous, Deciduous, Mixed forest, Fields, Cutblocks, Industrial, residential etc..

A. Steps to start supervised classification:

- Analysis -> Image Classification -> Supervised
- Select your file **.pix ,** click on 'open'
- Click on New Session .. then 'add layer' button (if you need it)
- You will need 3 empty 8-bit channels - unless you already have empty channels
- Display channels should be 5,4,3 (RGB)
- Input channels should be 2,3,4,5,7; tick in Input channel column *
  I chose these 5 for 2 reasons: so that you realise you are not limited to 3, and because the r value between bands 2-3, and 5-7 was .90, so there is some contrast between each pair.
- Set the training channel to an **empty** channel number (tick in that column)
- Set the output channel to another **empty** channel (tick in that column)
- OK

'Training Site Editing' box appears !

2. Training Areas: *In the Training Site Editing box:*

- Add a new class, (class->new) double-click on its name and change it to water,
- In the main focus window (below View), pick the 3rd icon in the 2nd row,– it looks like a zig-zag 'new shapes', and from the dropdown, pick -> raster seeding.

The raster seeding window controls its functionality with two tolerance settings
These are currently set to 10 and 1X, increasing either increases the deviation that is allowed in including pixels adjacent to wherever you might click. For now leave these as is; Zoom into the image with water and click on the water, I suggest the confluence (T'enneh) of the two rivers.

Throughout seeding, the 'raster seeding panel' should be onscreen (or you can't seed) Ensure that the training areas (channel) are highlighted in the table of contents - seeding can be picky

Select 2-3 seed points for water - you will want to cover clear water (Nechako R) and silty water (Fraser); also pick a lake e.g. Tabor ... you could change training site colours, but DON'T PICK BLUE for water or you can't see your sites - the defaults are usually OK
The goal is to find a sample of water pixels, not to 'fill' the river: DO NOT TRY TO FILL THE RIVER WITH SEEDED PIXELS - THIS DEFEATS THE POINT OF THE ALGORITHM.

It's a good idea to hit the 'save' button periodically in the Training Set Editing window ..when done each class, 'save', add the next class, and repeat seeding process by class. You may need to modify the tolerance values for other classes, but I found this was only so for 'residential'

Oooops- what happens when ... you added too much, picked a seed point that spread too much, added something to the wrong class etc.
There are two ways of dealing with this, the easiest is to remove/clear that seed and start over...

- In the training site editing window highlight the class with the mistake
- Select Edit – Clear Selected

You can also use this to remove the entire class selection

- Select Edit – Delete Selected

Usually pick 2-3 seed points for each class training 'set' - and view for different DNs for the same cover type, e.g. due to different illumination (more important in the mountains). For fields and cutblocks, try to pick sites that represent the range of possible values (screen colours). There is overlap between these two classes and variations due to cutblock regeneration - this is where ground knowledge would be handy.

Add / consider these classes for your classification

- Water
- Wetlands (NW portion of the image)
- residential (zoom in to find houses/gardens) – I changed 10 to 5 or it took too much
- industrial / concrete ((likely change tolerance back to 10 ?)
- cutblocks – regenerating
- Recent cutblocks (e.g. from pine beetle)
- agricultural fields
- Parks
- deciduous forest
- coniferous forest (include the shadowed eskers slopes)
- mixed forest – this class can be tricky
- bare ground - chip piles /
- Gravel (river edge)

Any others you can identify ? – remember, you can merge classes afterwards

Once you are ready, in the **training sites window**, check **tools -> signature separabilities** for suitability of classes and training areas - ideally each matrix value is > 1.8
If some pairs are very low (below 1.5) you can either delete one or merge the pair
Class-merge (and I think the method is obvious)

- but this is only a lab exercise, so no need for excessive rigour

Save the signature separability report, as a .txt file (in your labs folder)
- don't worry about the other tools dropdown option except this next one:

Tools-> classification preview
shows how it will look .. go back if it looks doomed and make edits as needed e.g. new classes

You have these options - try them all as they are previews: and you can Click off the classification layer now and then to view the 543 composite to check how they compare with the land cover types

- Maximum likelihood
- Maximum likelihood with null class (null = black) – what is missing, do we need more ?
- Parallelepiped
- Parallelepiped with MLC tiebreaker
- Minimum distance

3: Run the Classification

Pick the best one based on observation of areas you know – I was watching for the eskers area – did one or more avoid confusing the esker shadows with water ?

As per lecture notes, one might expect the best to be the MLC option ?

You should preview how this compares with the RGB bands, by toggling the classification on / off and check areas you know e.g. has it cleanly separated water, which class contains UNBC, how about the residential versus industrial areas in the bowl ?

When you are done, select Save and Close in the Training Site Editing Window.

right-click on the **classification metalayer** in the ‘maps’ list and select **Run Classification**.
Select your preferred classification. Make sure the show report button is ticked. Note the 'use bitmap as mask' option I mentionned in class to restrict a classification to only part of the image.
Save this report as a .txt file

I**n the report window** which classes are most confused (n the confusion matrix) ?
- why are these classes confused and not others? (maybe none!). Note that the accuracy shown is inflated as this is only comparing with training areas, which should mostly be correct.
Did any classes get 0 pixels – if so, why – maybe you need to add another training area

The **confusion matrix** gives a rough guide on the classification based on the training sites
Save the report as a .txt file so you can refer to it if need be after the report window is closed.

Reference data listed in the columns of the matrix represent the number of correctly classified training samples. Errors of omission are represented by the non-diagonal **column** elements, and errors of commission are represented by non diagonal **row** elements.
The Accuracy Statistics report lists statistical measures of overall accuracy and per class.

You can run a second classification if you like and have an empty 8-bit channel.
Right-click on the **classification metalayer** and then select configuration **–** change the output channel to the empty channel, so it doesn’t overwrite the previous one.

Remember that urban classification is challenging due to the mosaic of many small features – and that some features look similar e.g. agricultural and sports fields, some cutblocks.

4. SIEVE

The parameters will be as below: I suggest sieve to 1 acre minimum (4 pixels) for an urban area
input = ## (classification channel) .. note – not the training channel
Polygon size threshold = 4 (for 1 acre minimum), or 12 for a hectare...
Connectedness - can be 4 or 8 (try both and compare) –
exclude values list = ## (where ## is the class number for water - preserve small lakes)
output port should be viewer - PCT
select log tab and run ...

View the result, compare with the unsieved classification (tick the sieved layer on and off)
Q you should know: why does '8' preserve more small pixel clusters than '4'?

If happy with the result, Re-run the sieve putting output to your pix file – use BROWSE button - remember everything goes into the 2011 pix file, don't put it somewhere else … Run

After this is done, view your raster layers (files tab) listing to check you have a sieve layer – I found it's not labelled other than 'contents not specified' .. if so, go to the files tab, double-click on the name and change to something like 'sieve to 1 acre'. [you can also rename under the 'maps' tab but this only changes the name onscreen- the channel remains unnamed if you close it]

Compare with your unsupervised classification:
Layer-add-pseudocolor- and select your 'best' unsupervised channel from last week – finish
It's not easy to compare as the class colours don't match, but you might find more 'order' in today's result as you had more control over the classes.

Save your sieved classification same as you did last week
(right-click layer in maps list-> save as -> TIF)

Assignment #2 – due Wednesday 12 October

Email me:

1. Classification report (.txt)
2. Signature separability report (.txt)
3. Your classification as a TIF

5. Accuracy Assessment (optional extra) – no lab next week !

How could you more fully assess your classification - you'd need some independently derived 'ground truth' plots to test against: (these would usually be generated from ground work or other knowledge). The purpose of the accuracy assessment is to create a report which indicates the accuracy of classification results compared to the raw image data. It compares what is assumed to be correct with an image classification based on pixel groupings.

Select Analysis dropdown -> Image classification -> Supervised-> Post Classification Analysis-> select "Accuracy Assessment".
Select your file (yet again)
In the new accuracy assessment window, click on ‘select classified image’
And choose the **Sieved** results of the Supervised Classification
Select "Load Reference Image" and load channels 5, 4, 3.
This image will be used during the random sample selection.

Generating Random Samples

Select "Generate Random Sample" and increase the sample number to 30.
Leave the "Sample Control Options" section to the default answer, "yes". Press "Accept".
In the new Accuracy Assessment Panel, the idea is to highlight the first point in the sample list and observe where it is located on the image (the cursor will automatically be placed at this location). You should likely zoom in to both find the cursor and identify the class type (as best you can). Compare this pixel to the list and select the class you believe it should belong.

Ideally the pixel is in a relatively homogenous (not mixed) area .. but if this is not the case, you can ignore/reject that pixel. Once the random sample you are working with has been highlighted, you can simply select the class it belongs to (click on that class) and the information automatically transfers to the random sample list. This works the same as the "Transfer" button.
Samples could also be taken from vector segments e.g. verified vegetation polygons.
The user is given the option of selecting the channel and the class attribute associated with it.

Produce Accuracy Report
Once all 30 samples have been assigned, select "Accuracy Report" then 'generate report'

The results of the accuracy assessment are shown in the form of a Sample report Listing, Error (confusion) matrix and as accuracy statistics. Don’t be discouraged by low values – in a project/thesis you would take many more samples to get higher stats – at least 30 per class.
The Sample Report Listing can be used to determine which samples were classified correctly.
Save this report as a .txt file in your folder for later reference.

---

6. Landsat 8 OLI image data: pg3sept2013.pix

- more likely we will do this in the next lab after Thanksgiving

We should see how the newer 16-bit data from Landsat 8 OLI behaves in comparison.

Repeat the classification (step 1) for the 2013 image – Landsat 8 OLI (use bands 3-7 ... for input), and you may find the same algorithm seems to work best. Note that the **two tolerance settings in the raster seeding step will need to be <u>much</u> bigger**. Sadly, the help does not help – experiment – I found tolerance at 50 and 4x for water, 50 and 16x for the other classes.

Do one classification and sieve if you wish - accuracy assessment is optional
Comparing 2011 and 2013 might be interesting but you'd need to match the class colours

Footnote: selecting an area of interest and running a detailed unsupervised and supervised classification could the basis for a suitable project in the last 3 weeks of term, though not the only type – see future labs as well ..

FOR REFERENCE ONLY: Training site steps: the non-seeding (polygon) method
Click in your chosen area

- Zoom in maybe to +4 or even +8
- Select 'Trace and Close '
- Draw your training area (holding the left mouse button down- don't 'click') in a polygon
- make sure you draw a closed polygon; make it fairly large within the feature
- Select 'Fill'
- Click inside your polygon: it should fill
- You could add another training site, if you feel this one has not captured a good cross-section of the class (keep it simple at this point)
- Type in the class name (water) in the description column
- Save
- New