GEOG432 / 632 Fall 2015: Supervised classification

Start Geomatica2015 (Duh!)

0. Good habits - labelling new channels

Once you start creating new channels, its best to label them so you don't forget what they are

Open the PG file from last week's lab: **pg14sept2011.pix**
switch the maps tab to files
expand the rasters list
double-click on the channel to edit (or right-click -> properties
edit the classification channel descriptions: kmeans, fuzzy, isodata etc.. if you remember them ...

1. Creating a ratio channel

A ratio can assist classification as by reducing the effect of shadow (especially in the mountains) and increase the albedo effect by enhancing contrasts. So we will first create a 4/3 (NIR/Red) ratio to include in the classification; first for PG, then McBride

using **Tools -> Raster Calculator** in Focus
create this model, **double-clicking** on bands, single click on operators
- note that selection of a band returns %n for example %4 for band 4
%4 / %3 ... RUN (= hit the 'run' button... the wee man running)
... this creates a ratio onscreen (view) .. it hasn't yet been created as a new layer

Now you have an image related to vegetation biomass or greenness, but it only has a few different DN values - what is the range? (you will need to close the raster calculator to do this) ...

Now switch the output parameter from '8 bit unsigned channel' to '32 bit real channel' and RUN again - close the raster calculator window, and HIGHLIGHT the new layer in the maps menu - this time the output should show a full display range as the output values are all different - but this creates decimal DNs ... what is the range of DN values ?
Generally vegetated areas have a ratio DN > 1 while non-vegetated are < 1
i.e. near-IR < Visible reflection

Estimate an approximate scalar multiplier to raise the Ratio values to be similar to the Band 4 DN range - you don't have to be exact. Then run Raster Calculator one more time. This time add the multiplier in the formula: n*%4 /%3 where 'n' is the scalar value. Write to an 8-bit channel, and give the file as your copy of the pg file - NEVER accept 'Untitled.pix' !!!
Remove the temporary previous ratio images from the display and review the new channel.
Good habit: label it as **NIR/Red 4/3 ratio**

Now run an unsupervised classification, using bands 3,4,5 PLUS the new ratio as input. You will need to create a new 8 bit raster layer to the file but you can do this by clicking on 'new layer' when the unsupervised panel opens (I forgot this last week!)
You can use one of the sessions from last week, just change the input and output channels, and select your best classification algorithm from last week. Review the new classification to see if the addition of the ratio has clarified the shadows on the eskers ..it may have added other issues !

McBride
Now we'll do the same for the McBride image and start a new Focus session to avoid confusion. Leave the PG session open and click the focus button on the start panel – yes, you do want to start another application.
Load the McBride image and run raster calculator to create a NIR/Red image (5/4) – first using a 32 bit output to screen to see what the range of values are. Then decide on a multiplier (10,000?) and output to a 16 bit unsigned channel in your mcbride file (NOT untitled.pix !!!!).
Label it as NIR/Red ratio

Now run an unsupervised classification as you did for PG, using bands 4,5,6 PLUS the new ratio as input. You can use one of the sessions from last week, just change the input and output channels, add a new layer to hold the classification and select your best classification algorithm from last week. Review the new classification to see if the addition of the ratio has clarified the shadows – it may only be partial

2. Supervised Classification session

Unsupervised classifications are quick, supervised are not as quick due to the need to create training sites; This part of the lab may progress slower ....

Your choice – pick the PG or the McBride image ..

I'm picturing maybe 8 classes for either image:

PG: Water, Coniferous, Deciduous, Mixed forest, Fields, Cutblocks, Industrial, residential

McBride: Water, Coniferous, deciduous (valleys); alpine meadows (above the forest), avalanche slopes, Bare ice, snow cover, bare rock

- Analysis -> Image Classification -> Supervised
- Select your file **.pix** and then select new session
- Click on Supervised .. then click 'add layer' button
- You will need 2 empty 8-bit channels - unless you already have empty channels

- Display channels are 5,4,3 (RGB)
- Input channels should be 3,4,5; and the new ratio; tick in Input channel column
- Set the training channel to an **empty** channel *
- Set the output channel to another **empty** channel *
- Accept

*Assumes you have empty or available channels !

'Training Site Editing' box appears

 Add a new class, and change its name to water, Zoom into the image area with water
 Pick the 3rd icon in the main focus window directly above the 'files' tab and select raster seeding. (one can also choose to edit create training polygons ...)
 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 and click on an area of water, if the resulting seed doesn't cover enough area increase the 10 to 20

Ensure that the training areas are highlighted in the table of contents - seeding can be picky (otherwise seeding can't be selected)
Select 2 seed points for water around the scene - note that you will want at least one each to cover clear water (Nechako) and silty water (Fraser) ... you can change training colours, but DON"T PICK BLUE for water or you can't see your sites...
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 EXERCISE.

its 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 for other classes. Usually pick two seed points for each class training 'set' - and view for different DNs for the same cover type, e.g. due to illumination.

Oooops- what happens when ... you added too much, added something to the wrong class etc. There are two ways of dealing with this, the easiest is to turf the class and start over...
 In the training site editing window highlight the class with the mistake
 Select Edit - Delete Selected

Alternatively, you could delete the infected area by drawing a polygon around it, but this will delete everything within the polygon, possibly including other seeded points.
 Close the Raster Seeding window, the fifth button on the second toolbar has an erase tool

- It has a drop down list of erase tools, select erase polygon
- Draw a polygon around the infected area by single clicking around it, it will automatically link the last clicked point to the first, when you are finished, double-click.

Add the classes (PG) to your classification … or their equivalents for McBride

- Water
- residential
- industrial / concrete
- regrowth (cutblocks)
- agriculture
- deciduous forest
- coniferous forest
- mixed forest

If you think you might be ready, in the training sites window, check **tools -> signature separabilities** for suitability of classes and training areas - ideally each matrix value is > 1.8
(but this is only a lab exercise, so no need for excess rigour)

Also check classification preview - MLC or MLC with null - MLC will classify the whole image, while with null indicates if you may be missing a class type. You can preview how this compares with the RGB bands, by toggling the classification on/ off
When you are done, select Save and Close in the Training Site Editing Window.

right-click on the classification metalayer in the table of contents and select **Run Classification**. Perform a maximum likelihood classification. Make sure the show report button is depressed.

I**n the report window** which classes are most confused, why are these classes confused and not others? (maybe none!)

If you perform a classification for real, what is generally done is to create a set of seeded (or known) locations, and to separate these into two sets of locations, one for training the classification, and one to use afterwards for accuracy assessment. This makes sure that the accuracy assessment is independent of the training performance.

3. SIEVE

The classification can be cleaned using the SIEVE filter, as with the unsupervised classification
The parameters will be similar to these:
input = ## (classification channel)
Polygon size threshold= 11 (11 pixels = 1 hectare),
Connectedness - can be 4 or 8

exclude values list = ## (where ## is the class number for water)
output port should be viewer -PCT
select log tab and run ...
View the result, compare with the unsieved classification.
---

4. Accuracy Assessment

How could you assess which classification was better - you'd need some independently derived 'ground truth' plots to test against: (these would need to be generated from ground work or other background 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.
Under the CLASSIFY menu, select "Accuracy Assessment".
Select "Select Classified Image" and choose the results of the Advanced
Supervised Classification (the Maximum Likelihood Classification).
Select "Load Reference Image" and load channels 3,4 and 5. You will notice that a full resolution window is also created. This will be used during the random sample selection.

Generating Random Samples
Select "Generate Random Sample" and increase the sample number to 5. Leave the "Sample Control Options" section to the default answer, "yes". Press "Accept". In the Accuracy Assessment Panel, you will notice that 5 sample points have been added to the "Random Sample List". The idea here 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).
Compare this location to the class list and select the class to which you believe it should belong. Ideally the pixel is in a relatively homogenous (not mixed) area .. if this is not the case, ignore/reject that pixel if you can. Once the random sample you are working with has been highlighted, you can simply select the class it belongs to and the information automatically transfers to the random sample list. This works the same as the "Transfer" button. Samples can also be taken from vector segments. The user is given the option of selecting the channel and the class attribute associated with it.

Produce Accuracy Report
Once all 5 samples have been assigned to a class, select "Produce Accuracy Report".
NOTE: This option will not be available unless a classified image has been selected.
The results of the accuracy assessment are shown in the form of a Random Sample Listing, Error (confusion) matrix and as an accuracy statistic.
The Sample Report Listing can be used to determine which samples were classified correctly.

The Error (Confusion) Matrix is a method for displaying the results of the accuracy assessment process. Reference data are listed in the columns of the matrix represents the number of correctly classified 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 different statistical measures of overall accuracy and accuracy for each class.

FOR REFERENCE: Training site steps: the non seeding method

Click in your chosen area

- Zoom in maybe to +4 or even +8
- Select 'Trace and Close '
- Draw your chosen 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