This is the general guide for landuse mapping using mid-resolution remote sensing data

February 11

2015

This document has been prepared for "Training workshop on REDD+ Research Project in Peninsular Malaysia" based on Multispec & QuantumGIS Open Source Software applications

GENERAL GUIDE FOR LANDUSE MAPPING USING MID-RESOLUTION REMOTE SENSING DATA

1. Classification system and map legend

The general methodology of landuse mapping consists of six steps: (1) Determination of classification system; (2) Preparation of ground truth data based on the predetermined classification system; (3) Preprocessing of satellite data; (4) Classification by satellite data; (5) Correction and modification; (6) Validation using ground truth data. The first one of all steps is about the classification system; it means the classification system is quite important. Land cover classification system is decided considering the purpose of the land cover mapping project. Usually classification systems of different projects vary. Based on this project mapping purpose, we need to define land cover classification system. The detailed definition of each type sees Table 1.

Table1 Landuse classification system

No	Landuse Type	Landuse Definition
1	Inland Forest	Natural forests with > 30% canopy cover, the canopy being > 75% of dipterocarps forest.
2	Cropland	Cultivated, pasture land, rubber and oil palm plantation
3	Mangrove	Natural forests with > 30% canopy cover, composed of species of mangrove tree, generally along coasts in or near brackish or salt water.
4	Urban & Built-up	Includes residential, commercial and industrial, transportation, sport facilities
5	Shrub land	Woody vegetation < 3m in height, with at least 10% ground cover
6	Grassland	Upland herbaceous grasses >10% ground cover
7	Water body	Permanent open water bodies
8	Clouds/Shadows/No data	Areas where land cover interpretation was not possible

2. Data preparation

2.1 Input Remote sensing data

Medium resolution multi-spectral remote sensing data were chosen for landuse mapping in this project, taking into account the price and time limits for the production. The spatial resolution of the remote sensing image ranges from 10 to 30m, while the band number ranges from 3 to 10. In this project we use Landsat multi-spectral data.

Table 2 The key parameters of the multi-spectral remote sensing data to be used in this project

Data sources	Imaging years	Resolution (m)	Bands to be used	Availability
Landsat TM	1982 to present	30 m for multi-spectral bands, 120 m for thermal infrared band; Each scene covers an area of 185 3 185 km. Temporal resolution is 16 days	7	Free, full coverage difficult
Landsat ETM+	1999 to present	15 m for panchromatic band, 60m for thermal infrared and 30m for multi-spectral bands; Temporal resolution is 16 days	7	Free, no full coverage

2.2 Reference data

Collect any archived data useful for defining the location and extent of training samples for classification. For examples: history landuse map, dominate forest types map, forest coverage map etc. High resolution (1-5m) remote sensing data can also be used for classification reference such as by using Google Earth. The user should know the mapping or imaging date of these reference data very well. Don’t use data sources without any metadata information as reference data for landuse mapping.

3. Image Classification

Image classification, in a broad sense, is defined as the process of extracting differentiated classes or themes (e.g. land use categories, vegetation species) from raw remotely sensed satellite data. Obviously this definition includes the preprocessing of images. We here simply refer to the process following the image preprocessing as image classification.

3.1 Supervised versus unsupervised classification

Supervised classification is most appearing if there is much training data available to support training the classifier. However, for classifications in an area without extensive training data, unsupervised classification may be faster and easier.

Unsupervised classification tends to be most useful when the desired map classes are spectrally complex, that is, they are not well characterized by a single multi-variate distribution function. Unsupervised classification allows multiple spectral classes within individual map classes. The best approach when using supervised classification in this situation is to employ a non-parameterc classifier such as neural networks or decision trees or to develop multiple spectral classes for each map class through the training and classification steps.

It should be noticed that the direct results from unsupervised classification can only tell you the distinct spectral classes, however, what we expect from one classification are specific information classes, so the spectral classes should be reclassified or combined into some information classes as in the final landuse classes, this needs the knowledge learned from the ground true data.

3.2 Image classification methods

Image classification method, as one kind of multivariate methods, can be grouped into parametric, non-parametric, spatial-spectral and spectral-temporal classification as shown in table 3. The classification method most suitable to your project is mapping site, project objective, data sources used dependant. It is upon the user to choose one of them according to the specific mapping activities.

Table 2 Summary of classification methods, and their advantages and availability for forest attribute mapping

Classifier	Comments
Parametric	
Maximum likelihood	Works for “well separated” classes (eg. Land cover, forest classes with different physiognomy): requires relatively few training sites but subclasses must be characterized; results improve using iterative “cluster busting”
Minimum distance	Fast; Works for “well separated” classes with similar variability
Box classifier or parallelepiped (mean and standard deviations)	Fast; Works for “well separated”
Generalized Linear Model (GLM)	Useful if “response function” between dependent and independent variables well known, or of interest (eg. Tree species distributions)
Nonparametric	
Box classifier or parallelepiped (minimum, maximum)	Fast; Works for “well separated” classes; useful for “quick look” or to apply before MLC
Artificial Neural Networks	Good results for forest classification; requires extensive training data
Decision (Classification) Trees	Useful for combing remote sensing and other variables; decision rules are explicit; classification accuracies may not be much greater than MLC

4.2 Procedures for supervised landuse classification

The general procedures for supervised classification can be summarized as that shown in Fig. 1. A brief description of the major steps is as follows.

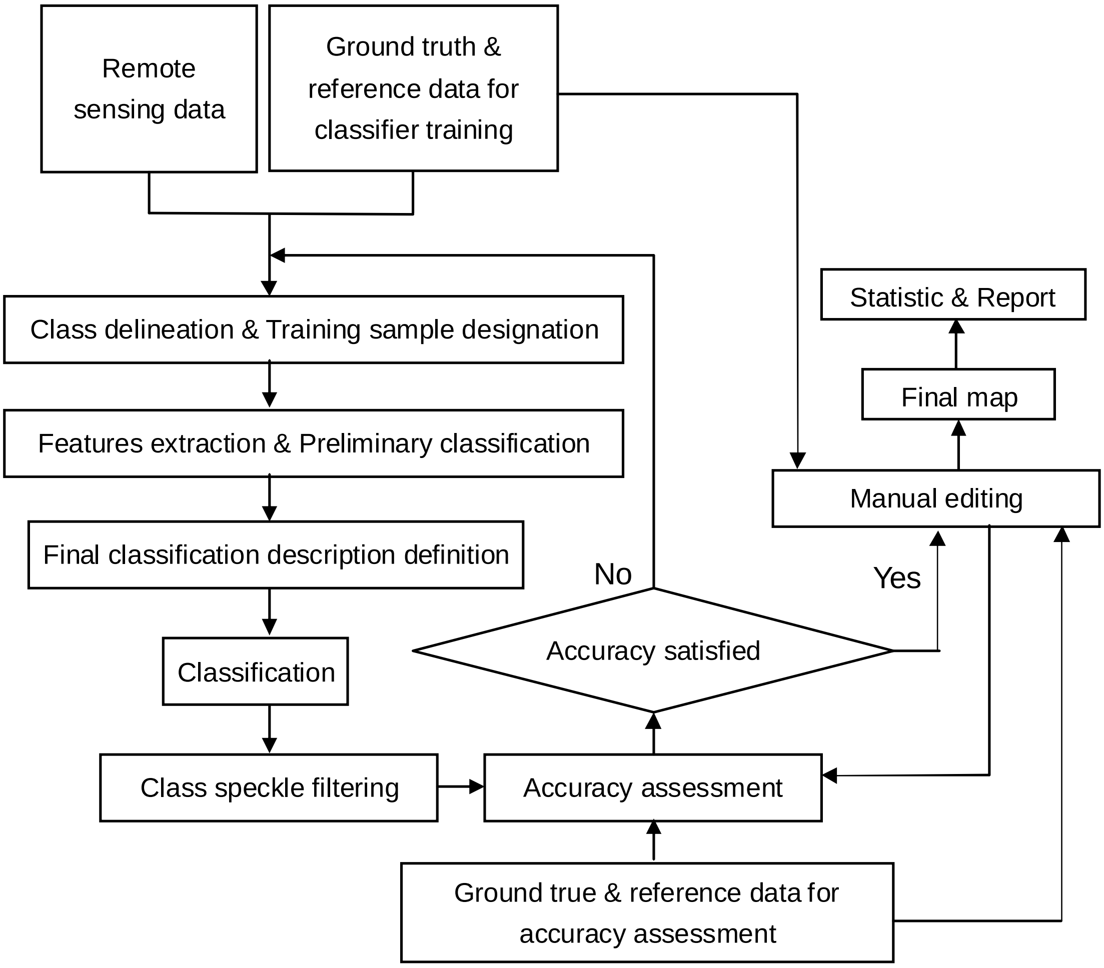


Fig. 1 General procedures for supervised classification

Class Delineation. The analysis process begins with a general overview of the data set to be analyzed, often by viewing a 3-color simulated color infrared presentation of the data in image space. The intent is to create a list of classes which is suitably exhaustive, and which includes the classes of user interest. To the extent possible at this point and from such an image presentation, consideration should be given to picking classes in such a way as to provide for a set that are separable from a spectral standpoint.

Training Sample Designation. Following, or as a part of listing the desired classes, the spectral description of the classes must be designated. How this is done varies widely with the particular data set and the information about the scene that is available to the analyst. Examples of some of the ways this might come about are as follows.

- Observations taken from a portion of the ground scene at the time of the data collection;
- Observations from aerial photographs from which examples of each class can be labeled;
- Conclusions that can be drawn directly from the image space, itself. For example, an urban mapping problem where the spatial resolution was great enough to make objects of human interest recognizable in image space.
- Conclusions that can be drawn about individual pixels by observing a spectral space representation of a pixel. The use of “imaging spectroscopy” characteristics, where specific absorption bands of individual molecules are used to identify specific minerals, are an example of this.

Optimal class definition requires that the classes defined must be

- *Exhaustive*. There must be a logical class to assign each pixel in the scene to.
- *Separable*. The classes must be separable to an adequate degree in terms of the spectral features available.
- *Of informational value*. The classes must be ones that meet the user needs.

These three conditions on the list of classes must be met simultaneously. Note that the exhaustive condition and separability are properties of the data set, while the user imposes the informational value condition. It is the bringing together of these circumstances, those imposed by the data with those imposed by the user’s desires that is the challenge to the analyst. It is further noted that the classes are defined by the training samples selected. That is to say that the definition of classes is a quantitative and objective one, not a semantic one. One has not really defined a class one might wish to call “forest” until one has labeled the training samples to be associated with that class name, thus documenting quantitatively what is meant (and what is not meant) by the word “forest”.

Feature Extraction and Preliminary Classification. At this point one can expect to have training sets defined for each class, but they may be small. There would thus be value in eliminating features that are not effective for the particular set of classes at hand, so as to reduce the dimensionality without loss of information. A feature extraction algorithm would be used for this purpose, followed by a preliminary classification. From the preliminary classification, one can determine if the class list is suitably exhaustive, or if there have been classes of land cover of significant size that have been overlooked. One can also determine if the desired classes are adequately separable. If not, the classification can be used to increase the selection of training samples, so that a more precise and detailed set of quantitative class descriptions are determined.

Final Class Description Determination. With the new augmented training set, in terms of either additional classes having been defined or more samples labeled for the classes or both, any of the following several steps may be taken to achieve the final class descriptions in terms of class statistics.

- It may be appropriate to re-apply a feature extraction algorithm, given the improved class descriptions. In this way, a more optimal subspace may be found.
- The Statistics Enhancement algorithm may be applied. This algorithm is known to be sensitive to outliers, and thus would not be expected to perform well until it is known that the list of classes is indeed exhaustive, as classes not previously identified would function as outliers to the defined classes. The intended result of applying this algorithm at this point is to increase the accuracy performance of the following classification and to improve the generalization capabilities of the classifier from the training areas to the rest of the data set.

Post-Classification Filtering. Resulting classification map may be difficult to interpret, because it has a salt-and-pepper appearance due to inherent variability of the per-pixel classifier. So post-classification filtering is needed to remove pixels and pixel groups not satisfying a minimum requirement. It usually can enhance interpretability and increases classification accuracy. A final classification of adequate quality should be possible at this point.

Classification Accuracy Assessment. If the accuracy is not good enough, depending on the nature of the further improvement needed, a return to any of the above steps may be used.

4.3 Classification Accuracy Assessment

4.3.1 Accuracy assessment stages

Four significant stages have been witnessed in accuracy assessment methods. Accuracy assessment in the first stage was done by visual inspection of derived maps. This method is deemed to be highly subjective and often not accurate.

The second stage used a more objective method by which comparisons of the area extents of the classes in the derived thematic maps (e.g. the percentage of a specific vegetation group in area) were made with the corresponding extents on ground or in other reference dataset. However, there is a major problem with this non-site-specific approach since the correct proportions of vegetation groups do not necessarily mean the correct locations at which they locate.

In the third stage, the accuracy metrics were built on a comparison of the class labels in the thematic map with the ground data for the same locations. Measures such as the percentages of cases correctly (and wrongly) classified were used to evaluate the classification accuracy.

The accuracy assessment at the fourth stage made further refinements on the basis of the third stage. The obvious characteristic of this stage is the wide use of the confusion or error matrix, which describes the fitness between the derived classes and the reference data through using the measures like overall accuracy and kappa coefficient. Additionally, a variety of other measures is also available or can be derived from the error matrix. For example, the accuracy of individual classes can be derived if the user is interested in specific vegetation groups.

4.3.2 Ground truth data collection for accuracy assessment

Generally, there are three major ground truth data collection methods by different reference data sources:

1) Record the landuse type and its location using GPS in the field through ground survey work.
2) Through aerial photo interpretation when the aerial photos is available and timely consistent with the remote sensing data used for classification.
3) Using GIS layer from other sources.

The area standing for by the ground survey point (sample unit) should be consistent with the resolution of the remote sensing image used for classification. It is suggested to use Systematic Non-Aligned Sampling (SNAS) method to determine the sampling location in the field. Figure 2 shows the concept of this sampling method. The whole mapping region is divided into sub-regions, and in each sub-region equal number of points are sampled, whose location can be randomly distributed.

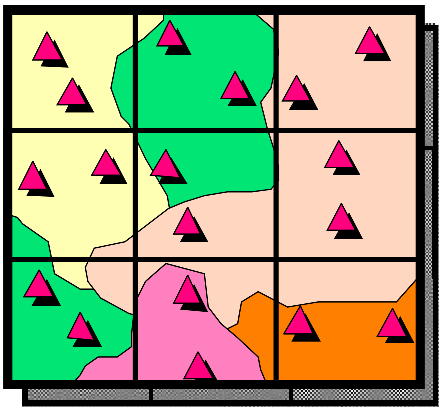


Figure 2 Systematic Non-Aligned Sampling

5. Landuse classification example: Supervised classification using Multispec software.

Exercise	Title
1	Display and Inspection of Image Data
2	Image Enhancement
3	Unsupervised Classification (Cluster Analysis)
4	Supervised Classification - Select Training Fields
5	Overlay Shape Files on Image Window
6	Selecting Areas and the Coordinate View
7	Creating Vegetation Indices Images

6. Landuse change analysis example: Change detection using QuantumGIS

Exercise	Title
1	Vector data preparation
2	Using Analysis Tool for changes analysis
3	Analysis of landuse changes result
4	Producing Map

MultiSpec Exercise 1: Display and Inspection of Image Data

Requirements: MultiSpec application and image titled "training.img".

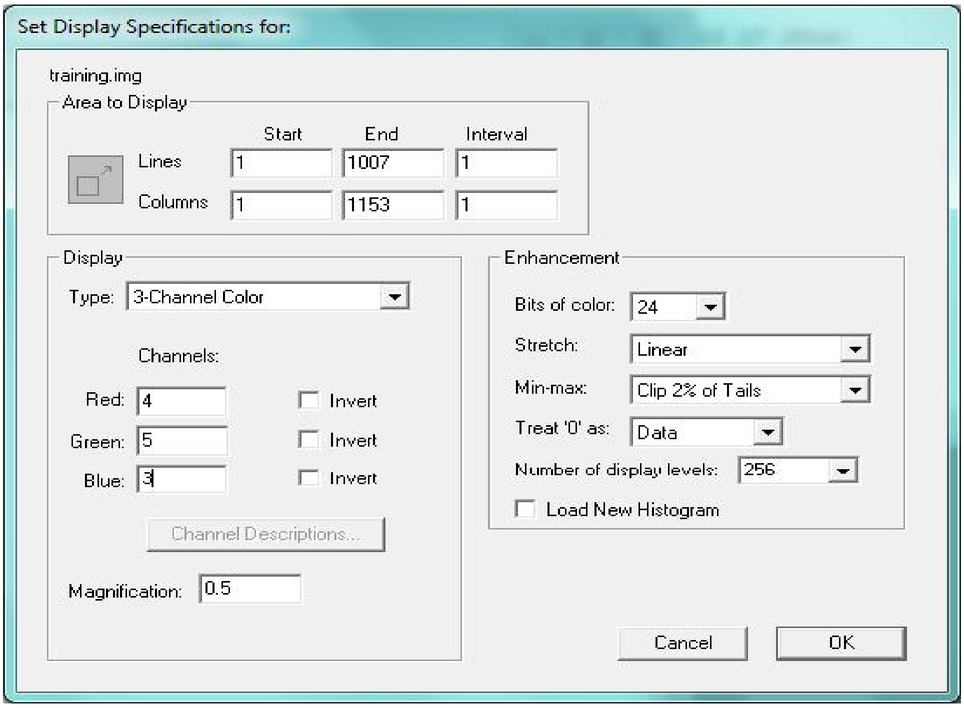
In this exercise, you will display the Landsat satellite image of a portion of the Klang, Selangor and view the data in several ways using MultiSpec.

1.1 Start MultiSpec using the icon on the desktop or from MultiSpec in the Startup Menu.

1.2 From the ***File*** menu choose ***Open Image. . .***. A dialog box will open to allow one to select the data file one wishes to use.

1.3 Select training.img in the Training folder and ***Open,*** or simply double-click on training.img

This is a segment (1007 lines x 1153 columns of pixels) of a 3-channel image of Klang, Selangor collected on 14 January 2015. Next a dialog box will appear to allow one to choose among various options for the image display.



Note that by default, the area designated for display is the whole scene and the ***3-Channel Color*** Display Type is selected. The default settings call for the Red screen color to be derived from band 4, Green from band 5 and Blue from band 3. These particular choices will cause the screen image to be in a 3-color format approximating Color Infrared film.

Select ***OK.***

1.4 This step may not occur for all situations. If the data histogram has not previously been calculated and stored (in a .sta file), another dialog box will be presented allowing the choice of regions to be histogrammed, so that the channel data values can be properly assigned to screen colors. The default options built into this dialog box are satisfactory, so

Select ***OK*** to begin the histogramming.

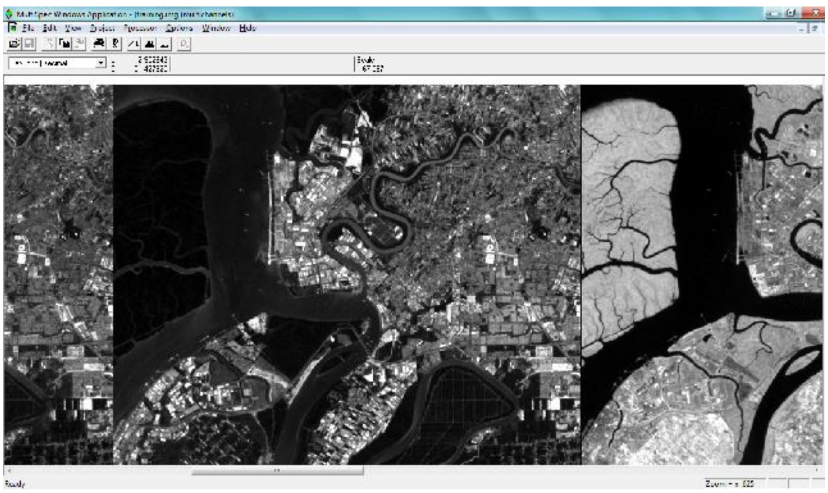
After the histograms of all of the channels have been computed, the information will be saved to a file named “training.sta” so that they will not have to be recomputed when needed again.

1.5 The image of the data will now appear. Notice that just above the image window in the toolbar there are two small boxes with large and small “mountains”. These are image zooming buttons allowing one to zoom in (large mountain) or out (small mountain) from the current image scale. Just to the left of the image zooming buttons is another button which shows ***X 1*** in grayed form. This button allows one to go to X1 magnification directly. The current zoom magnification is displayed along the bottom of the MultiSpec application window in the box labeled “Zoom=”.

Some other options are to hold the ‘Ctrl’ key down while zooming to change the zoom step factor to 0.1 instead of 1. In other words, the zoom factor will change from 1.0 to 1.1 to 1.2 etc. instead of 1, 2, 3, etc.

One can make a selection within the image by left click-hold in the image window, drag to select a rectangle, and then releasing the left mouse button. If a selected area exists in the image, any zooming will be centered on the selected area if possible. Clear selection using the “Delete” key.

1.6 Next one can view a side-by-side channel display for data quality inspection. From the ***Processor*** menu, select ***Display Image…*** to bring up the display dialog box. Then select Display Type “Side-by-Side Channels”, and select OK to display all three channels in the image side by side.

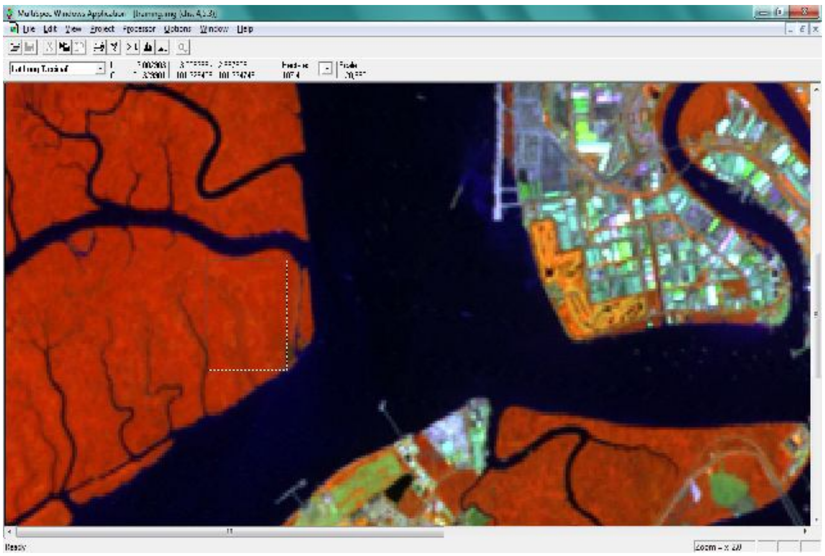


The above image window will be displayed (after zooming out) which shows all eight channels displayed side-by-side. Note that the vegetation areas in channel 4 are brighter than the same areas in channels 1 and 2.

The side-by-side channel display is a good way to verify that the channels are registered correctly. In other words, the same location in the image is at the same pixel location in all channels. To do this, select an area in one channel near a field intersection. This same selected area will be drawn in all of the channels. One can then verify that the selected area is at the same location in each channel.

Redisplay the 3-channel image with channels 7, 4 and 2 as Red, Green and Blue.

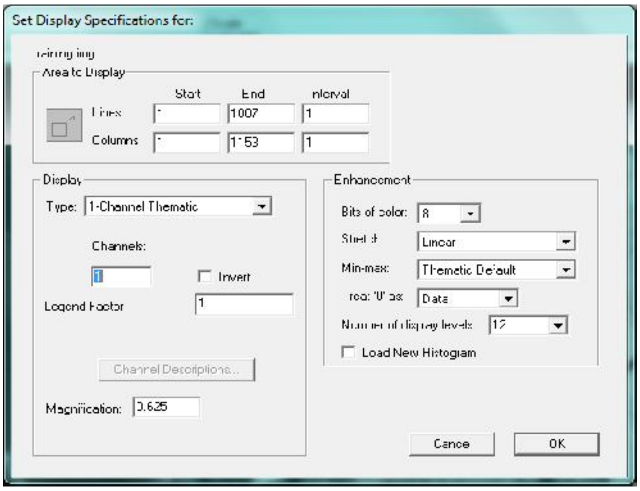
1.7 **Coordinate View.** One can also display a “coordinate view” along the top of the image to present the cursor (mouse) location and selected areas in the image. To do this, select the **View->Coordinate View** menu item in the Windows version of MultiSpec.



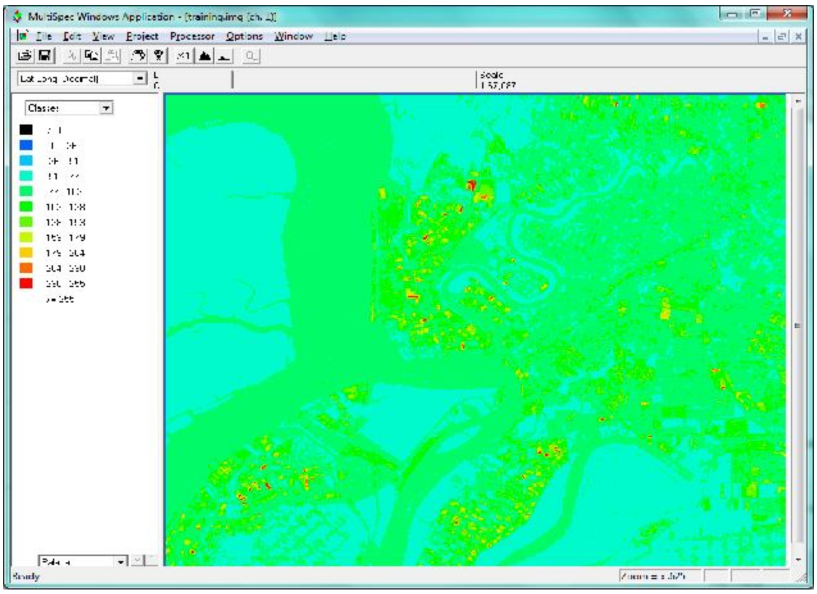
If map coordinate information exists for the image, one can display the coordinates as map units. Use the popup menu on the left side of the coordinate view to select the map units. The area of the selection can be displayed as the number of pixels or in units of acres, hectares, etc using the popup button to the left of “Scale”. The scale of the image will also be displayed.

For this image, if one selects the **Edit–>Image Map Parameters…** menu item and set the UTM Zone to 47 North and the Datum to “WGS 84” and then select OK; one can also display the cursor coordinates as latitude-longitude.

1.8 The 1-channel Thematic display type is useful to display "product" type images such as MODIS NDVI or any of the other many MODIS products.



The data values are grouped into the desired number of levels and a legend is displayed to the left of the image indicating which palette colors are associated with each range of data. (See illustration below.) One can also enter a factor to multiply the data values displayed in the legend by to reflect the actual measurement unit. Sometimes the data value may be the measurement value times 100 or 1000. One can use the Min/Max User Specified dialog box item to set the min and max values for the range of data to be displayed. Black is the default color for data values less than the minimum and white is the default color for values greater than the maximum. (Note: This feature can be considered as a supervised 1-channel levels classifier.)



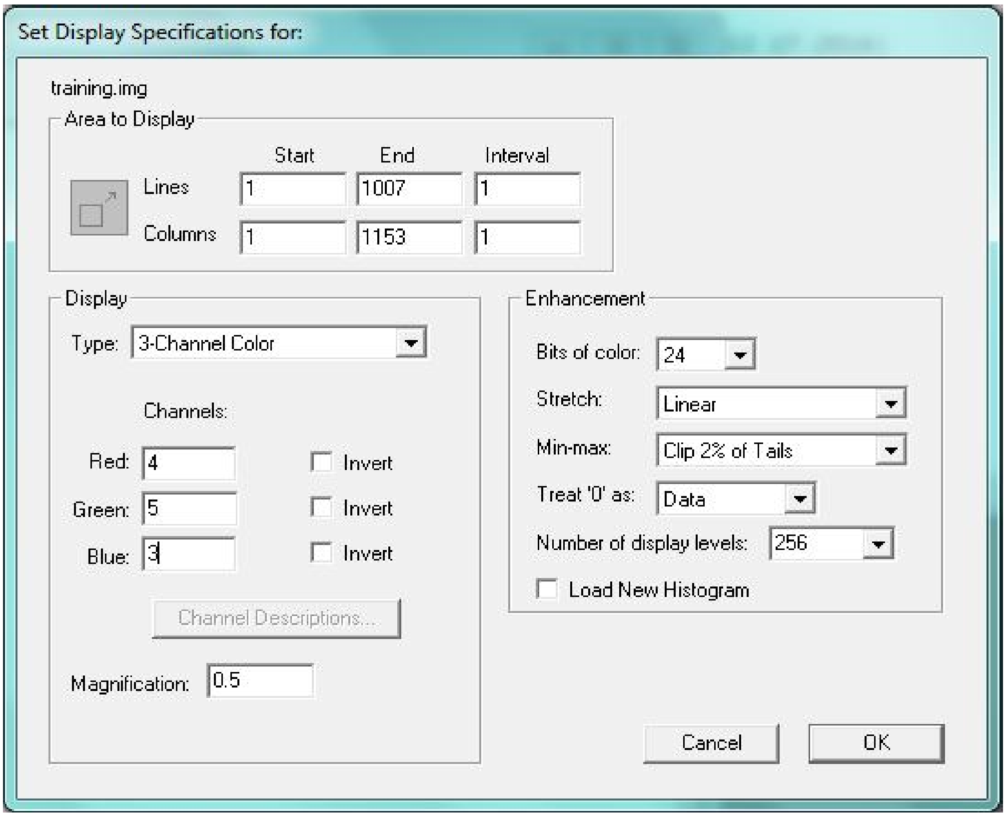
One can use the Reformat - Change Image File Format processor to create thematic images based on what is displayed in the image window.

MultiSpec Exercise 2: Image Enhancement

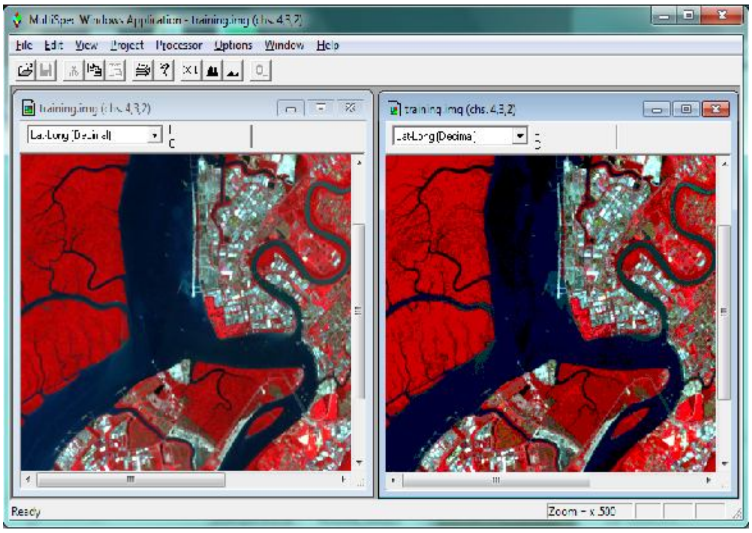
Requirements: MultiSpec application and image titled "training.img".

Open the image if it is not already displayed in a multispectral image window following the guidelines given in exercise 1 (Display and Inspection of Image Data with MultiSpec).

One can control the enhancement of the image in the multispectral image window by setting five different options in the Enhancement portion of the Display Specifications dialog box including Bits of color, Stretch, Min-maxes, Treat '0' as and Display levels per channel.



2.1 The “Bits of color” default is 24 and the “Display levels per channel” default is 256 for all monitors now days for the maximum number of colors possible. One can adjust to lower values if one wishes to see what the affects are.

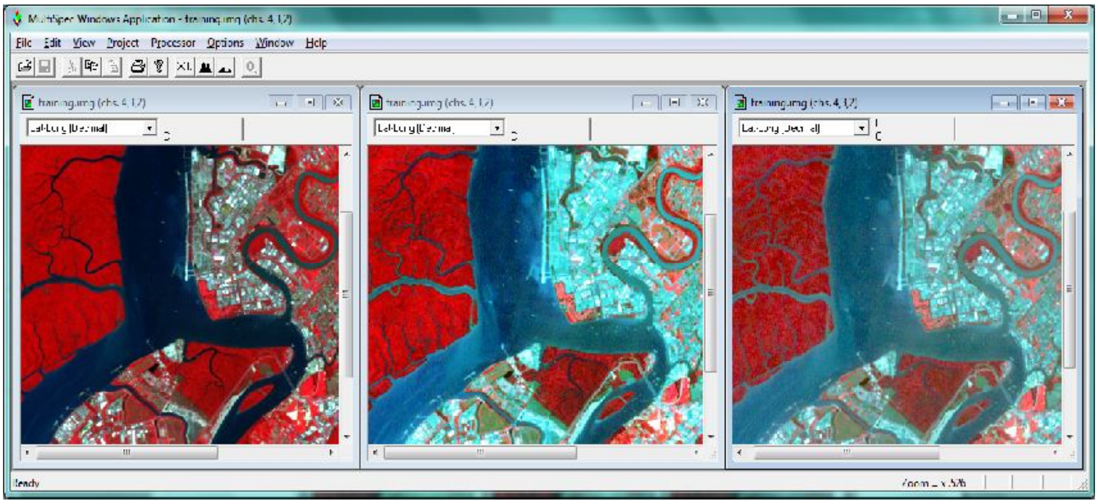


The figure above illustrates 256 display levels per channel (millions of colors) on the left and 2 display levels per channel (6 colors) on the right.

2.2 The Treat ‘0’ as data setting causes 0 values to be displayed as black. However, if 0’s actually represent background or ‘no data’, one may want to select the background option to cause 0’s in all channels to be displayed as white.

2.3 The Stretch and Min-maxes are usually the options used to enhance different parts of the image. They control the process by which each possible data value in the image data is assigned to all possible display levels.

There are three choices for Stretch: Linear, Equal Area and Gaussian. In the case of Linear Stretch, gray scale intervals are equally spaced over the data range, while the Equal Area Stretch choice causes them to be set so that each interval represents about the same number of pixels. Though nonlinear, the equal area stretch will provide maximum contrast. The Gaussian selection causes the distribution of the number of pixels assigned to the gray scale intervals to represent a Gaussian curve. If one holds the 'Option' key down (Mac version) or the 'Right Mouse button down (Windows version) before you select the enhancement popup menu with the left button, you can change the number of standard deviations that the data will be fit to for the Gaussian selection. The default is 2.3.

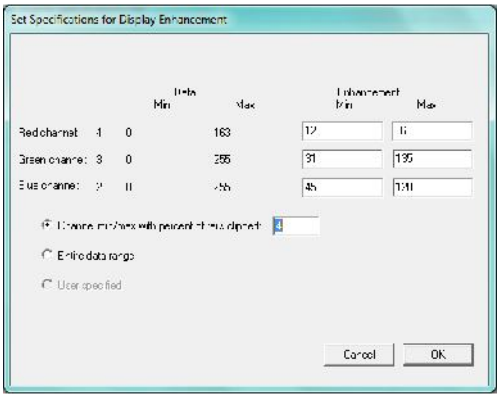


Linear, equal area and Gaussian stretches are illustrated in figure above left to right.

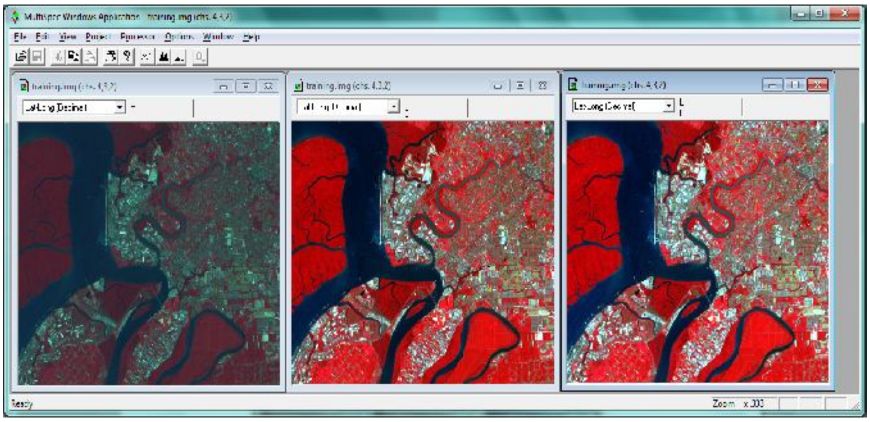
2.4 The Min-maxes option allows one to select the beginning and end data values of the image histogram to be used for assigning the pixels in the gray scale intervals defined by the Stretch option. The “Entire Range” choice for this option would cause lowest data value possible in the image, 0 for 8-bit unsigned data, to be the first data value displayed for lowest display value (black) and 255 to be the last data value displayed for the highest display value (white). However, if the actual range of the data is only 50 to 150, then the data will only be represented by grays not black to white; there will not be much contrast in the image.

The 2 Percent Tails Clipped choice will cause the selected begin and end range of data values for a given channel to represent those data values in which 2 percent of them in the histogram are outside of the selected range. The intent of this choice is to reduce the chance that a small number of extreme outlier data values in the image will unduly influence the display enhancement. This choice usually results in a display of the data that has better contrast over all.

Selection of the Specify Min-Max... choice presents a dialog box (illustrated below) to allow one to set percent of tails clipped to be something other than 2%. You can also set your own min-max values to stretch the gray levels across. The actual data min and max values computed from the histogram are included in the dialog box.



Entire range, 2% clip and 4% clip is illustrated below from left to right.



MultiSpec Exercise 3: Unsupervised Classification (Cluster Analysis)

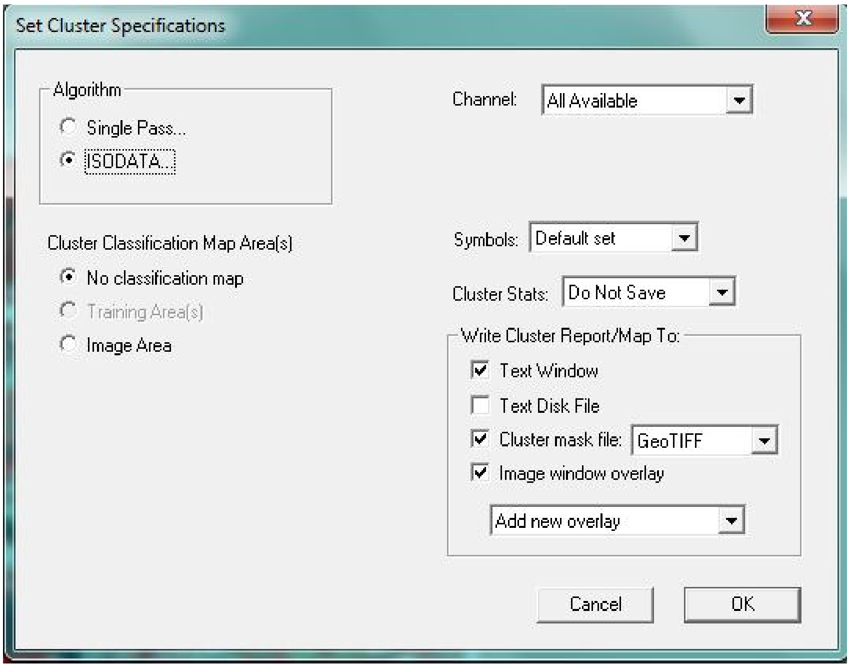
Requirements: MultiSpec application and image titled "training.img".

Two Clustering algorithms are available in MultiSpec. They are useful in grouping similar pixels in the image into clusters or categories. One algorithm implemented is a simple one-pass type. The second is an iterative type called ISODATA. We will use the ISODATA algorithm for this exercise.

To start this exercise, be sure that the "training.img" image that was used in exercise 1 is open. Also clear any selections in the image window by striking the "Delete Key"

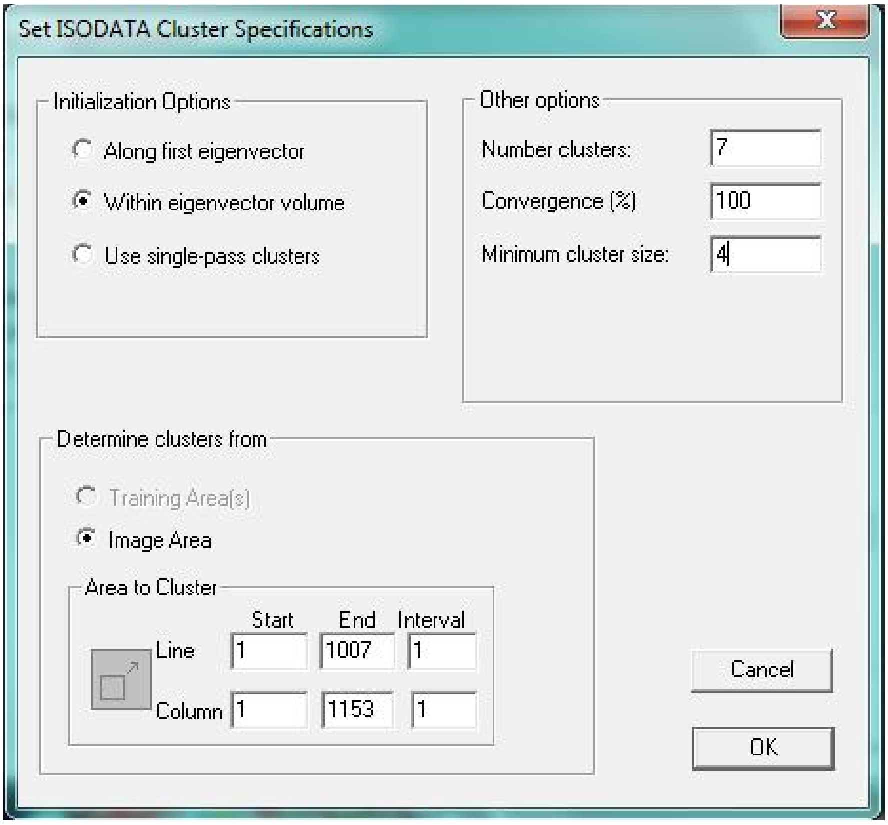
A cluster analysis will be run using the image file represented by the active (top-most) multispectral image window.

3.1 From the ***Processor*** menu, select ***Cluster…*** to bring up the cluster specifications dialog box. Select "Do Not Save" in the Cluster Stats: popup menu. Select "Cluster mask file" and select "Image window overlay" under the "Write Cluster Report/Map To" group.



This will cause a cluster map to be created as a thematic image disk file and displayed as an overlay on the multispectral image window.

3.2 Then select “ISODATA…” This will cause the ISODATA Specifications dialog box to be displayed. Enter 7 for the number of clusters, 100 for the convergence percentage and set the line and column intervals to 1 (if needed) and then select **OK**.



3.3 You are now back to the Cluster Specifications dialog box. Select **OK** to close this dialog box and start the clustering operation.

You will be prompted to enter a name for the cluster map disk file and where to save the file. Just use the defaults by selecting OK in the Save Cluster Map dialog box.

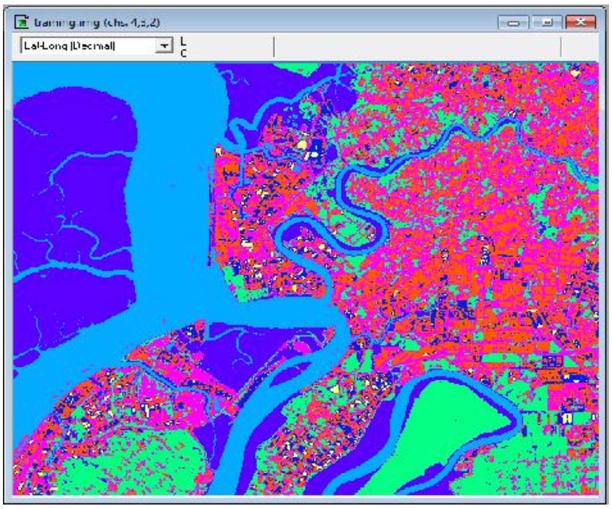
A cluster map will now be created with around 7 classes in an unsupervised manner. You will notice the colors change in the image window as the pixels are sorted into cluster classes during each iteration. After the final iteration, a thematic image file with a map of the cluster classes will be saved to disk.

The text output for the cluster operation will be at the end of text output window. The information includes the mean values for each of the channels for each cluster for both the intial condition and the ending condition. If the map information is available for the image, the final area for each cluster is listed in the units specified in the coordinate view for the image window.

Usually the convergence is set for a little less than 100 so that the process does not take too long to complete. We used 100 in this example so that you have a chance to watch the pixels change cluster classes which illustrates the nature of the ISODATA algorithm.

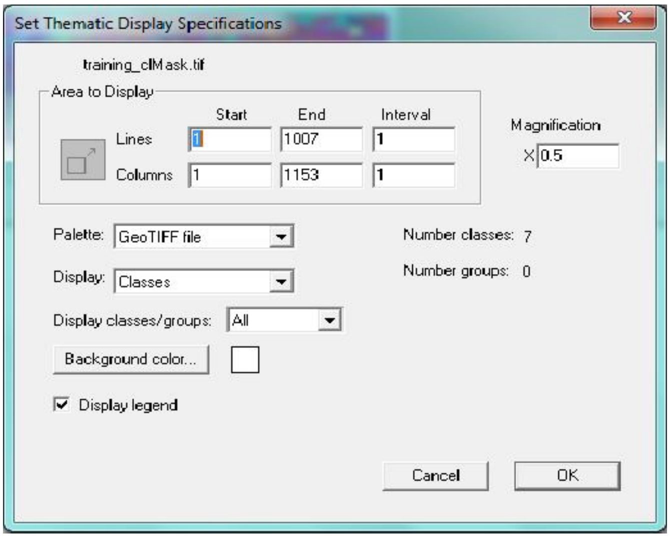
The cluster map overlay on the multispectral image window will look similar to the following.

You can turn the overlay on and off by using the “Red O” popup menu button in the toolbar to the right of the “small mountain” zoom button.



3.4 Now open the cluster map thematic image file. This will be the same image as is shown in the overlay above but you will have more control over the cluster classes.

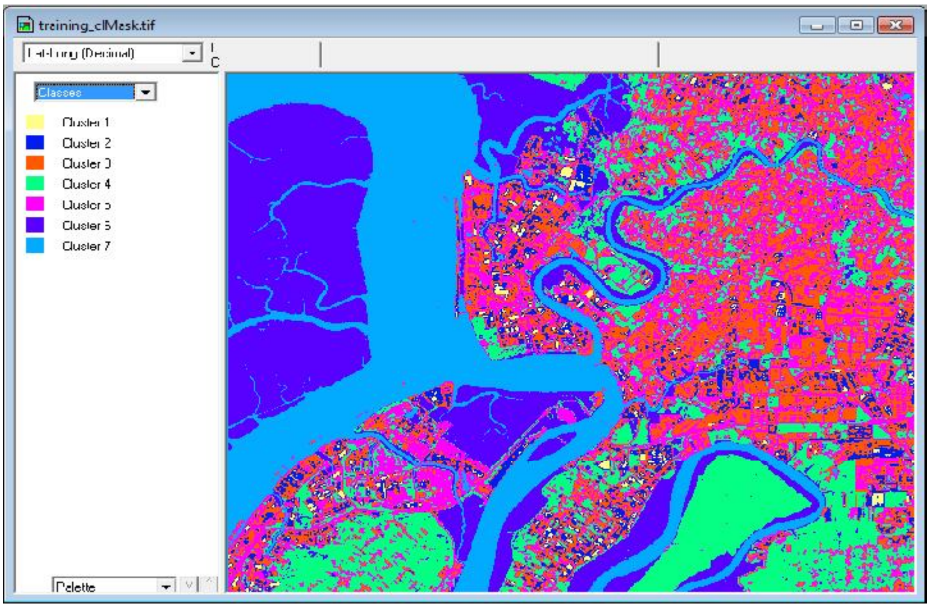
From the ***File*** menu, select ***Open Image…*** to bring up the open image dialog box. You may have to change the “Files of type” popup to All Files or Thematic Files. Then select “training_clMask.tif” and then select **OK**. One may need to select “Thematic” for the *Files of Type* popup menu.



The Thematic Display Specifications dialog box to the right will be displayed.

The default settings are fine; select “OK” in the Display Thematic Image dialog box. This opens a thematic image window.

3.5 The cluster class legend is on the left in the thematic image window below.



Compare the Cluster map with the original image window. Note clusters that represent trees, sparse vegetation, thicker vegetation, different soil colors and waterbody.

There are several things one can do to evaluate the image. One can move the cursor over a color chip, hold the shift key down (cursor will change to an open eye) and click the left mouse button down and up to cause the colors for that class to blink off and on (alternate between white and the color). If one holds down both the shift and ctrl keys and then click the left mouse button down and up, then all of the other classes will blink off and on. These procedures are helpful in understanding the extent of the classes in the image.

One can also change the class color by double clicking on the color chip.

One can change the cluster class names by double clicking on the name to the right of the color chip.

3.6 One can also group the cluster classes together in information groups by selecting Classes/Groups in the popup menu above the legend. Then drag the cluster classes into similar information group categories. Again one can double click on the group name to change the name. For example one could change the appearance of the thematic image to represent Bare Soil, Vegetation and Non-image informational classes. The popup in the legend allows one to display the original cluster (spectral) classes.

MultiSpec Exercise 4: Supervised Classification - Select Training Fields

Requirements: MultiSpec application and image titled "training.img".

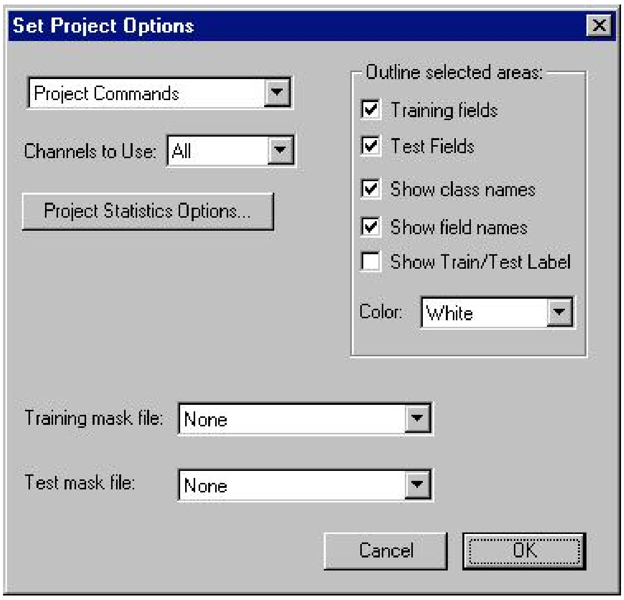
One can also do a supervised classification by selecting training areas for specified classes from known areas. Open the image if it is not already displayed in a multispectral image window following the guidelines given in exercise 1.

4.1 Select Training Fields

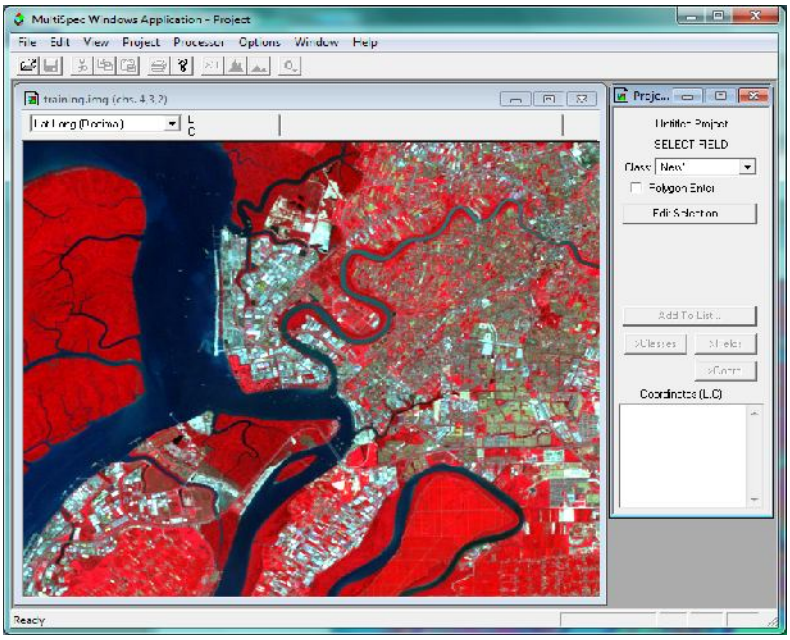
4.1.1 If a project window is open, from the ***Window*** menu, select ***Project***, and then double click in the upper left to close the current project and project window. One may have been created during the cluster analysis.

4.1.2 Now one will select training fields.

From the ***Processor*** menu select ***Statistics*** and select ***OK*** in the Set Project Options dialog box. (The default settings for this exercise are satisfactory.)



A new window labeled ***Project*** will appear to the right of the screen that will be used in a moment. To select training fields for each class, one must simply "drag" a rectangular area on the image (or, with polygon option selected, click on the corners of the desired polygon), and then "Add that field to the list." Thus,

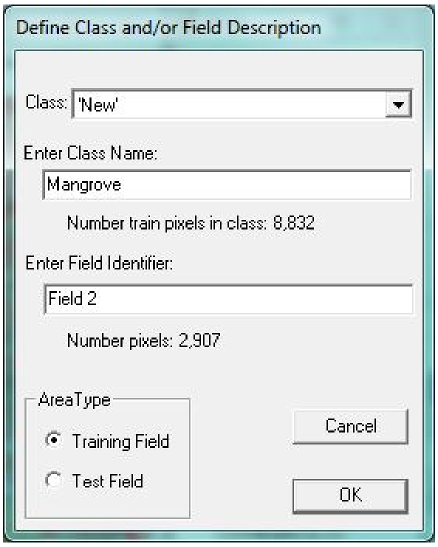


Drag from the upper left corner to the lower right corner of the Tree training field in the ***image window***. If upon inspection, one does not like the exact boundaries resulting, one may immediately repeat the process.

Caution: A mistake that is made many times is to select training areas very near edges of a field. One should stay away from the edges by a couple of pixels to reduce the chance of edge affects.

Note in the ***Project dialog box***, that the coordinates (row and column numbers) of the upper left corner and the lower right corner of the selected area appear in box near bottom. Now,

4.1.3 Select the ***Add to list*** button. A dialog box will appear to allow one to name the class and give the field a special designation, as desired. Thus, Type ***Trees*** into the Class Name box and then select ***OK***.



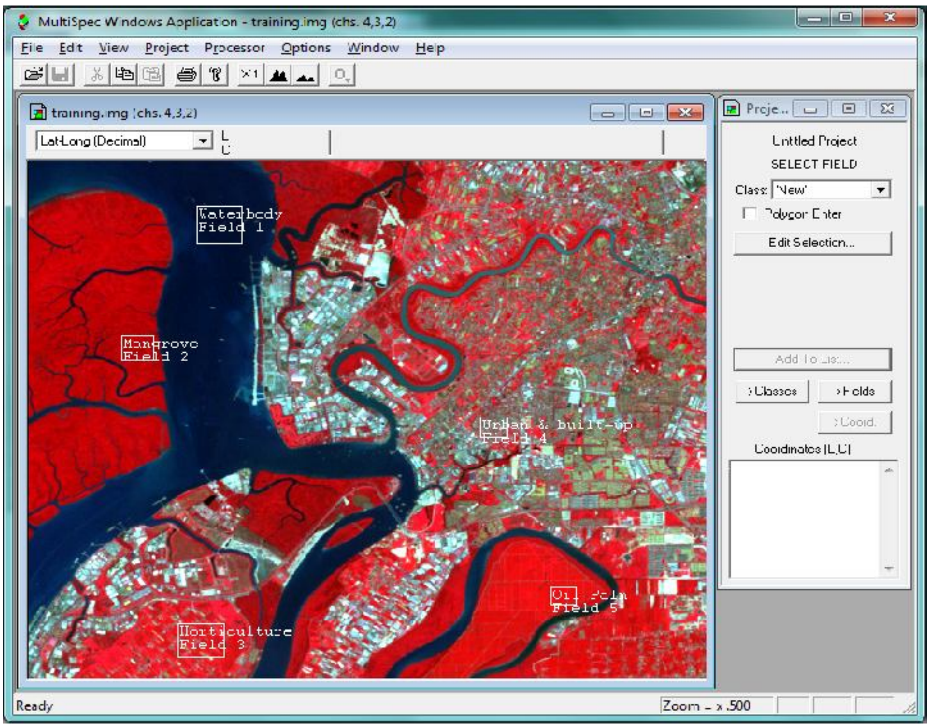
Note that one can designate the selected area as a training or test area.

4.1.4 Since there is to be only one training field for this class, we are ready to select the training for the second training class. Thus next,

Drag across the second training field in the ***Image Window*** shown below for Waterbody.

Select the ***Add to list*** button in the ***Project*** window.

Select the training areas for the rest of the four classes – Urban & built-up, Horticulture, Oil Palm, Bareland



The Project Window can have four different modes – the Select training field mode, the Class list mode, the Field list mode and the Coordinate list mode. The modes are controlled by four buttons just above the list box at the bottom of the Project Window. The “>Select” button causes the Project window to be in the select mode. The “>Classes” button causes the Project classes to be displayed. The “>Fields” button causes the fields for the selected class to be displayed. The “>Coord.” button causes the coordinates for the selected field to be displayed.

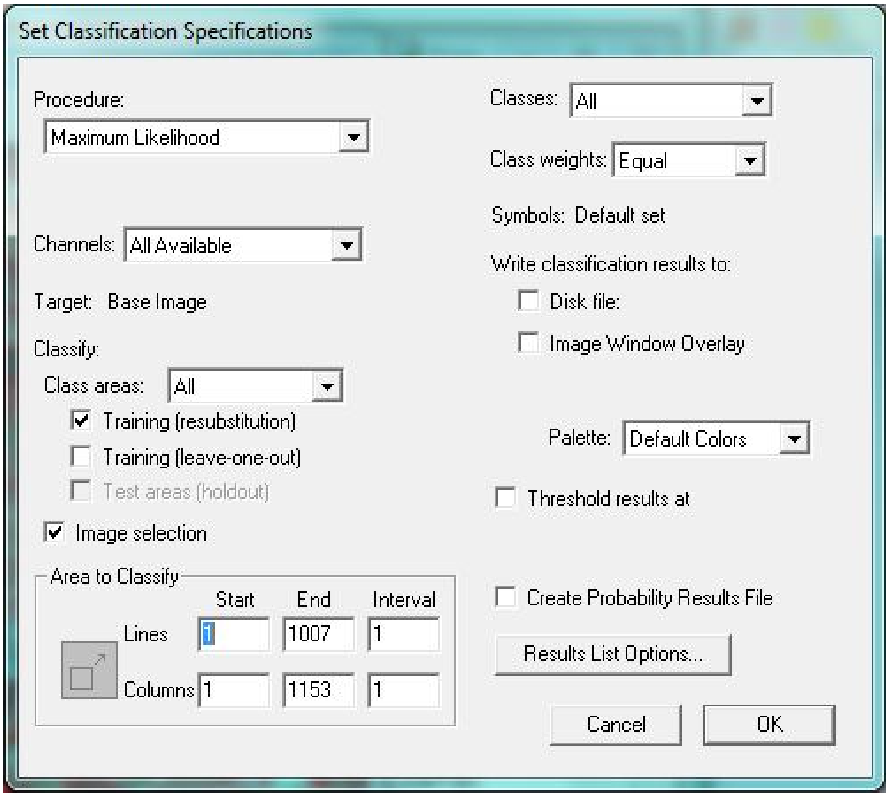
One can delete a class by selecting the class in the class list and then selecting “Cut Class” in the Edit menu. One can also do the same for deleting a specific field.

One can also use polygonal type fields to define training classes. To do this, select the “Polygon Enter” checkbox in the Project Window when in Select mode. Click in the image window to define each corner of the polygon. Double click on the last point. To turn the polygon type selection off, just select the “Polygon Enter” checkbox to deselect it.

Note that the clustering step described in Exercise 3 can be useful in the classifier training step in determining how many classes might be separable in a given data set and where to define training areas such that the spectral characteristics of the pixels are similar.

4.2 Classification

4.2.1 From the ***Processor*** menu select ***Classify…***. In the Set Classification Specifications dialog box which appears, select the √ near ***Image Selection*** under Classify to de-select it since, during this pass, it is desired to classify only the training fields in order to obtain an initial estimate of the quality of the class definition and training.

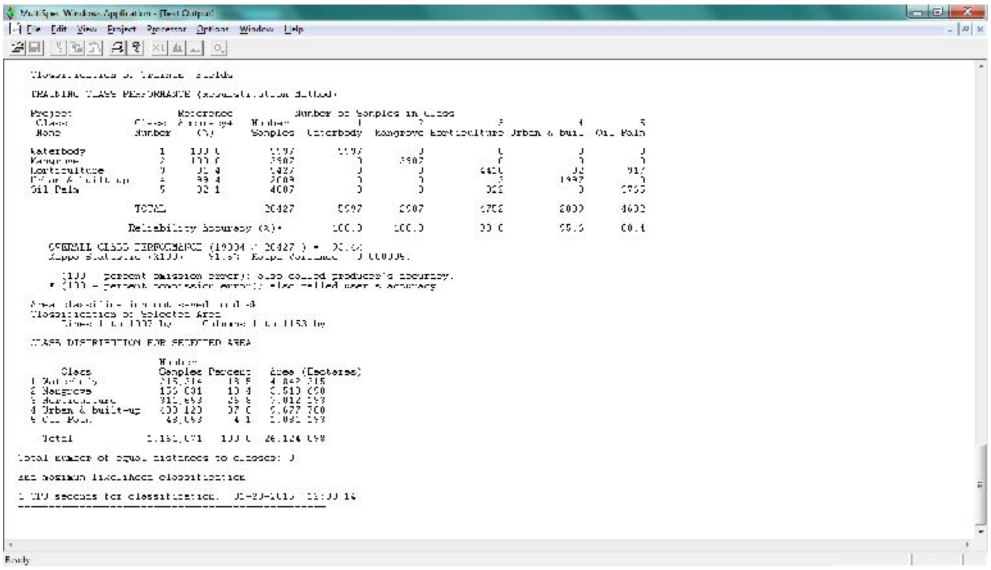


Note that under ***Write classification results to:*** One can also select the ***Disk File*** button causing a disk file version of the results to be written. Since we have no need for this file in this case, leave this button unselected.

Since the other default options are satisfactory, select ***OK*** and then ***Update*** to the "Update Project Statistics" dialog box to begin the classification.

The classification will be complete momentarily.

4.2.2 From the ***Window*** menu select ***Text Output***, to bring the text window forward and make it active, since it contains the classification results. The "TRAINING CLASS PERFORMANCE (Resubstitution Method)" table tabulates how the pixels of each field and class were classified. See example table below. There should be nearly 100% accuracy on the training fields. If the Reference Accuracy is particularly low (say less than 50%) for a class, then the training pixels for that class should be reexamined and new training pixels selected.



4.2.3 Assuming satisfactory results, we are ready to classify the whole area. From the ***Processor*** menu choose ***Classify***…

- Under ***Class Area*** de-select ***Training (resubstitution)*** by selecting the √ by it, and,

- Select ***Image selection***. Make sure that the entire area of the image is to be classified (lines 1-1007 and columns 1-1153). Select the square button, if activated, to the left of lines and columns, to force all lines and columns in the image to be used.

- Also select ***Disk File*** under ***Write classification results to:*** so that a disk file for later use will be created.

- One can also select ***Image Window Overlay*** to cause the classification to be displayed as an overlay on the multispectral image window if you wish to.

- Also select the ***Create Probability Results File*** checkbox so that a classification probability map will be saved to a disk file.

- Then select ***OK***.

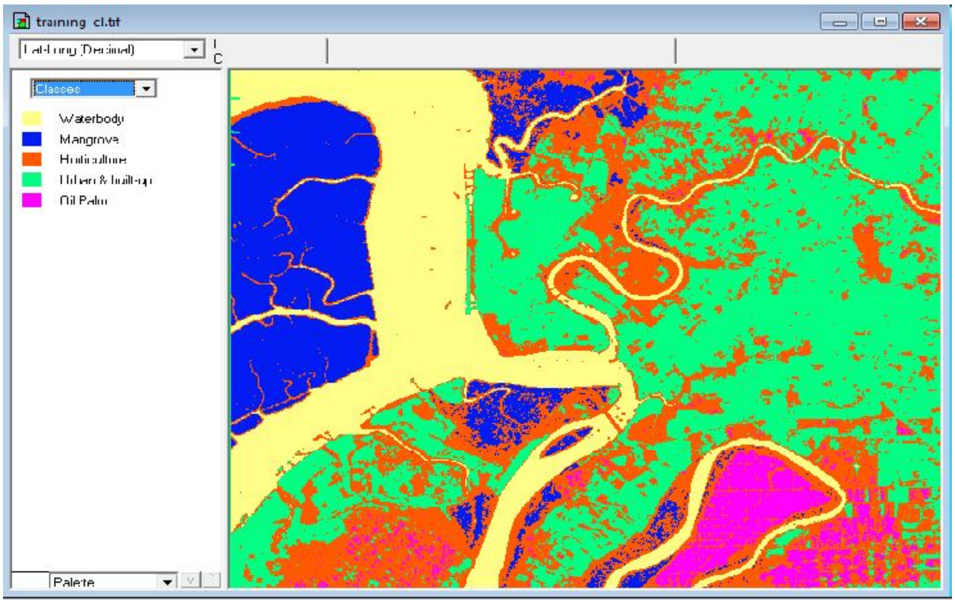
- Select ***Save*** in the dialog box that follows regarding a file name for the results. We will use the default name and location for the output classification file and for the probability map file.

As soon as the classification is complete, one will see a summary of the results displayed in the text window.

4.2.4 You save the project using the **File–>Save Project** menu item. You will be presented with a dialog box to enter the name (or use the default name). The training and test areas that you selected will be saved. You can open this file up at a later time to continue your analyses.

4.3 View Classification Map.

4.3.1 Now open the classified image named “ag020522_DPAC_cd.gis”. It should appear similar to the image below. This is a Thematic type image.



4.3.2 After displaying the classified image “training_cl.tif” from the ***Project*** menu, select ***Add as Associated Image*** to cause the training field outlines to be drawn on the image. You can change the field outline color to black using the **Processor->Statistics…** menu item and selecting “Black” in the Color popup under the “Outline selected areas:” group.

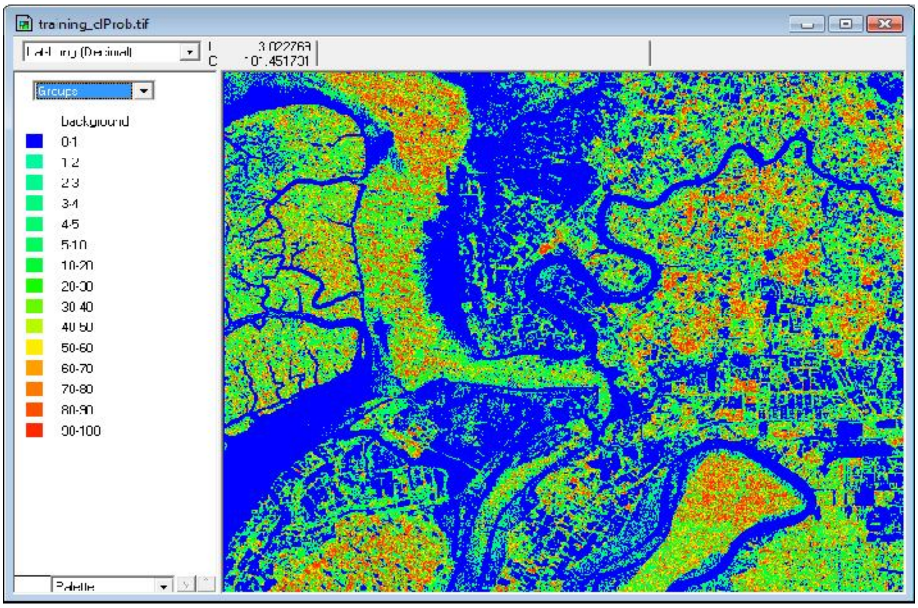
There are some things that one can do to evaluate the results. One is to move the cursor over a color chip, hold the shift key down (cursor will change to an open eye) and click the left mouse button down and up to cause the colors for that class to blink off and on (alternate between white and the color). If one holds down both the shift and ctrl keys and then clicks the left mouse button down and up, then all of the other classes will blink off and on. These procedures are helpful in understanding the extent of the classes in the image and to determine where classification errors may be. One may need to change training fields or add other classes if there appears to be confusion between the categories that you wish to identify.

One can also change the class color by double clicking on the color chip as was done in the cluster section.

4.4 Classification Probability Map.

One can view the classification probability map to evaluate which portions of the image have lower and higher probabilities of being classified correctly. There may be other classes in the image that our training fields do not adequately represent.

4.4.1 Open the probability map named “training_clProb.tif”. It should appear similar to the image below after associating it to the Project. This is a Thematic type image.



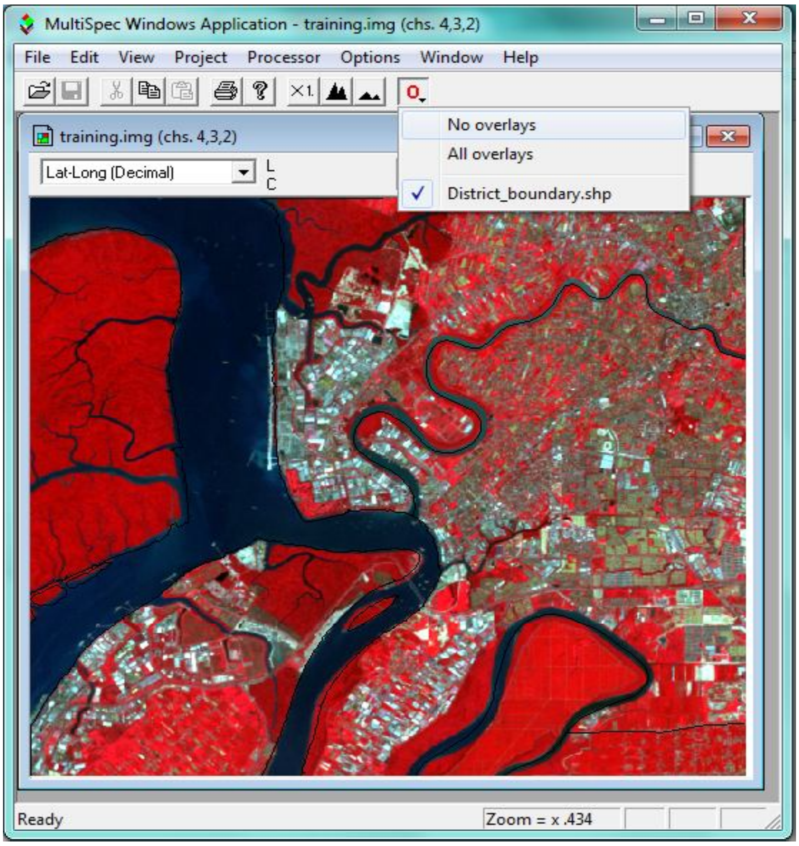
Yellow to red colors represent a high probability of being correct. These pixels are very close to our training pixels for the classified class. Dark green to blues represent a low probability of being correct. These pixels are very far from the training pixels for all of the classes. The classified image though may still represent the area well enough for our purposes.

MultiSpec Exercise 5: Overlay Shape Files on Image Window

Requirements: MultiSpec application, image titled "training.img" & district_boundary.shp.

One can open ArcView Shape files (as long as an image window is opened first). The popup menu button for the Windows version is next to the zoom buttons in the toolbar. The shape file will only be added to the active image window.

The training.img image has been provided with several shape files in the Training folder to allow one to practice working with shape files.



5.1 Display the training.img image file in a multispectral image window using the **File–>Open Image…** menu command. You can display the image using any band combination that you wish.

5.2 Use the **File–>Open Image…** menu command to open a shape file such as district_boundary.shp in the Training folder. You may need to change the filter (Files of types) in the open image dialog box to Shape or All. The shape file will be over-laid onto the image window as illustrated above with black lines.

MultiSpec will automatically convert lat-long shape files to respective map coordinates for images in Albers Equal Area, Cylindrical Equal Area, Equirectangular, Lambert Azimuthal Equal Area, Transverse Mercator, Orthographic, Sinusoidal and UTM map projections. MultiSpec first assumes the shape file is in the same units as the map projection. If there is no overlap, MultiSpec checks if the input shape file units are within the range possible for decimal latitude-longitude degrees. If so, MultiSpec assumes the shape file is in lat-long units and converts them to map projection units. If the converted shape file values overlap with the image, then the shape file is overlaid onto the image. If a shape file has been converted, _ltom is appended to the shape file name in the overlay list for the window. (Comment: ltom stands for lat-long to meters.)

5.3 One can obtain a dialog box for editing the vector line width and color in the Windows version by holding the shift key down *before* selecting the Overlay menu button with the left mouse button and then select the shape file overlay to be edited. Note that … now follows the overlay name indicating that a dialog box will be displayed.

Change the color of the shape files lines that have been drawn to white with a thickness of 2.

5.4 Notes:

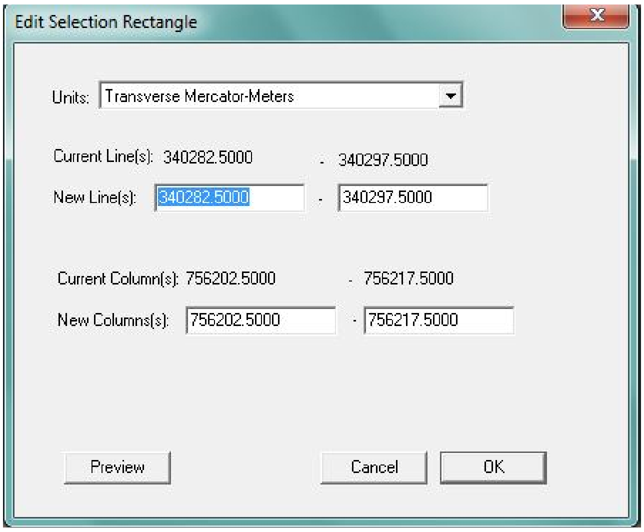
- One can use the **Edit->Clear Overlay** menu item to remove the selected shape file from memory. The list of shape files in this list will include all shape files drawn in all open image windows. If one also has an image in geometric (lat-long) projection, shape files on these images will be treated as a separate shape files in the **Edit–>Clear Overlay** list

MultiSpec Exercise 6: Selecting Areas and the Coordinate View

Requirements: MultiSpec application and training.img image files.

One can make selections of areas within an image using line-column, map or latitude-longitude units. The map and latitude-longitude units are only available for images where the required map projection information is available. Currently, MultiSpec can handle selections in latitude-longitude for image in geometric (latitude-longitude) and UTM, Transverse Mercator, Albers Conical Equal Area, Cylindrical Equal Area, Lambert Azimuthal Equal Area, Sinusoidal, Orthographic, Krovak and a few State Plane map projections.

6.1 One can use the **Edit–>Selection Rectangle** menu item to do this.



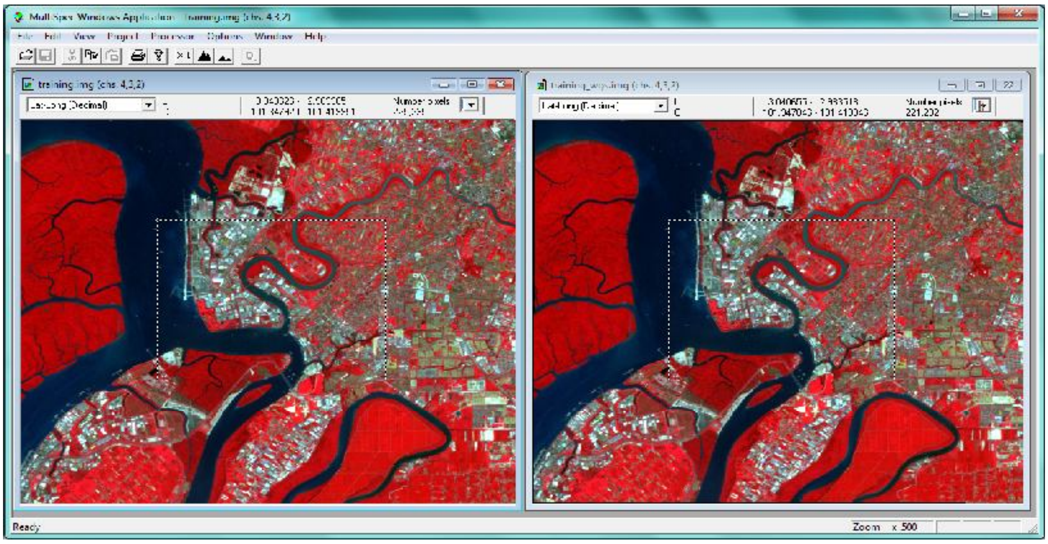
- A checkbox is available to allow one to apply the selection to all open image windows. If the selections are being made in line-column units one can turn off the option to take into account the start line and column of the images. If the units being used are map or latitude-longitude units then the same area will be defined as well as possible across all image windows even if the pixel sizes are different.

- A Preview button is available to apply the change for viewing before closing the dialog box.

6.2 Similar capability to the **Edit->Selection Rectangle** menu item above, is also available using the coordinate views of the images and the mouse cursor. If you make a selection in an image window with the coordinate view in lines/column units and hold the control key down when making selections, the same line and column selection will be made if possible in all open image windows. If one also holds the shift key down when doing this, any difference between the start line and column values will not be used.

If the coordinate view is in latitude-longitude units, the same latitude-longitude will be selected in all open image windows that contain the map projection information (which makes this possible). The same is true when selecting in map units of meters, etc.

The training.img (Transverse Mercator), and training_wgs.img (Latitude-Longitude) files can be used to illustrate these features. The figure below illustrates the result of a latitude-longitude selection in one image that was copied to the other open image windows when the control key was used in making the selection.



Note that the rectangle in the right image has a different shape (wider) than the selections in the other two images because the pixels in this image file were mapped to latitude-longitude units and the pixels in the other two images were mapped to other projections.

MultiSpec Exercise 7: Creating Vegetation Indices Images

Requirements: MultiSpec application and training.img

One can create images that represent algebraic combinations of the original channels of an image to try to enhance the image. This technique is used to enhance the vegetation or mineral variations in the image. One example is the Normalized Vegetation Index (NDVI) image. These images represent an algebraic combination of the red and near infrared bands to represent the amount of green vegetation in the image.

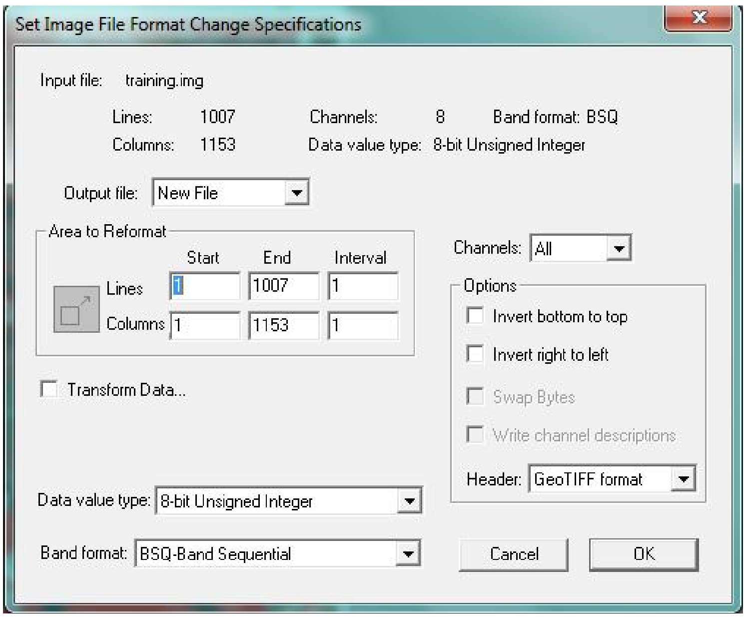
The formula is:

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

Where NIR represents the Near Infrared channel or band and Red represents the red channel or band. This formula results in a value that nominally varies between -1 (usually water) to +1 (strongest vegetation growth).

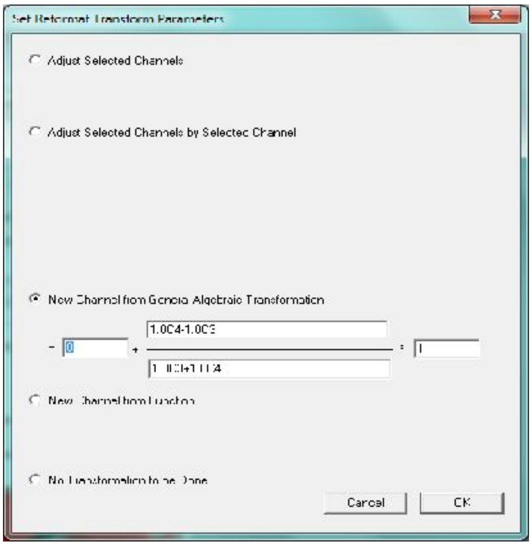
To start this exercise, be sure that the "training.img" image that was used in exercise 1 is open and represents the top-most image window in MultiSpec. Also clear any selections in the image window by striking the "Delete Key"

7.1 From the ***Processor*** menu, select ***Reformat->Change Image File Format…*** to bring up the Image File Format Change Options dialog box.



7.2 Select the “Transform Data…” checkbox. This will cause the “Set Reformat Transform Parameters” dialog box to be displayed.

7.3 Then select “New Channel from General Algebraic Transformation” This will cause the window below to be displayed.

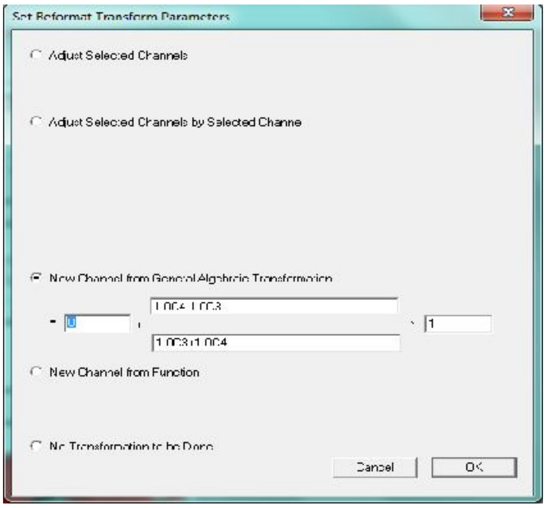


The editable boxes in the above window allow one to define an algebraic combination of the original input channels to create a new channel such as NDVI. The transformation is in the form of = offset + (a0 + a1C1 + a2C2 + ...)/(b0 + b1C1 + b2C2 + ...) ∗ factor; where offset, factor, a0, a1, b0, b1, ... are real constants which can be positive or negative and C1, C2, ...Cn represent the channel number to be used. Both 'c' or 'C' can be used. The number of constant∗Channel value (e.g. a1C1) combinations in either the numerator or the denominator is limited to the number of channels in the active image file plus 1. Also the number of characters in the numerator and denominator is limited to 255. Note that one can create only one new channel at a time.

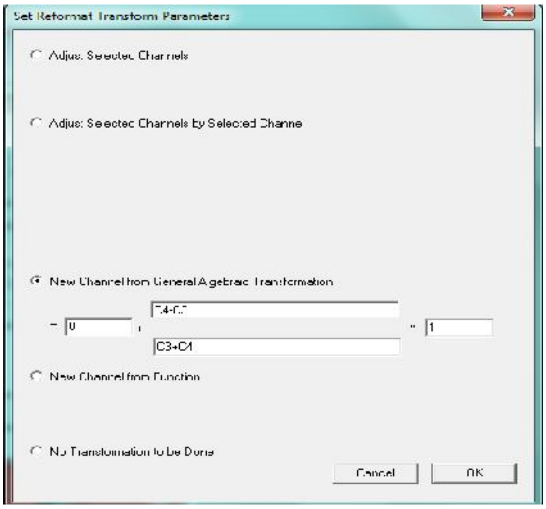
7.4 This tutorial is based on the **training.img** image file in which the red band is channel 3 and the near-infrared band is channel 4. Therefore to create a NDVI image for this image as defined in the introduction for this tutorial the equation (algebraic combination) will look like:

$$NDVI = 0 + \frac{1.0C4 - 1.0C3}{1.0C3 + 1.0C4} * 1$$

After entering the values needed to compute a new channel representing NDVI, the Transform Parameters dialog box will look like:

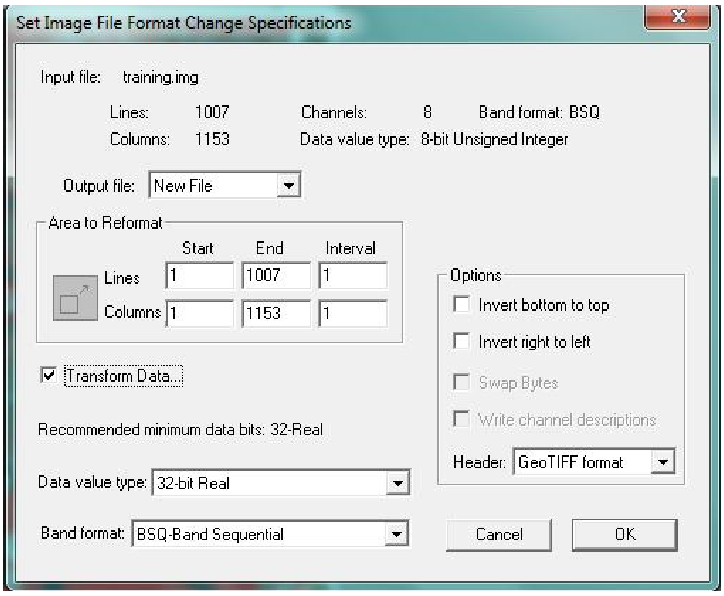


Since the real constants a1, a2, b1, b2 are 1.0, one can also enter the transformation as illustrated by the dialog box below.



Then select **OK** to accept the defined algebraic transformation. This will close the “Set Reformat Transform Parameters dialog box and bring the “Image File Format Change Options” dialog box back to the top.

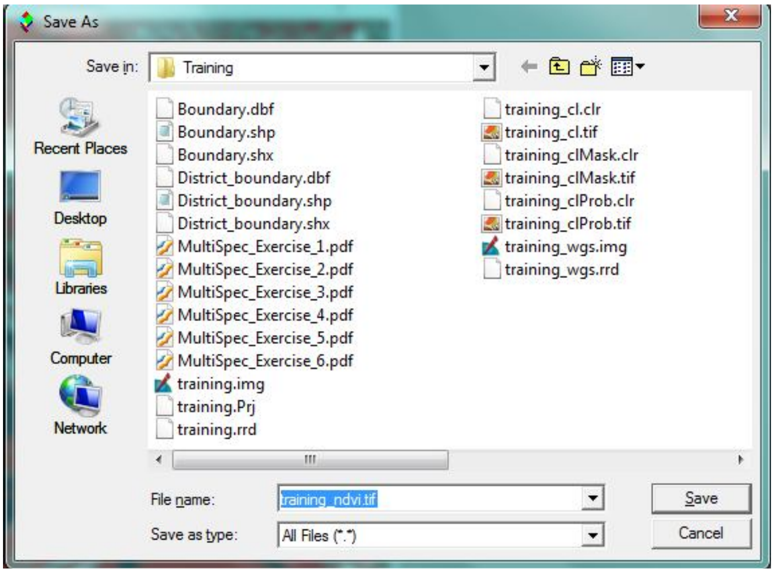
7.5 In the “Image File Format Change Options” dialog box, select “32-bit Real” in the Data value type popup menu.



Note that requesting the output data values to be real numbers is important since the resulting data values nominally range between -1 and +1.

If the output data values are converted to signed integer values, one will end up with only -1’s, 0’s, or 1’s or if to unsigned integer values, one will end up with only 0’s and 1’s.

7.6 Select **OK** to close this dialog box and start the processing to create the NDVI image file. The “Save New Image File As” dialog box will now appear, as given below, to allow one to enter the desired file name. Then select “**Save**”. The new NDVI image file will be written to the disk at the requested location.



Check the log output in the text window for this operation to verify that the creation of the NDVI image went well. The log output should be similar to the following:

```
Reformat: Change Image File Format  01-20-2015  16:04:23
(MultiSpecWin32_12.17.2014)

  Input Parameters:

    Image file = 'training.img'
    Create transformed channel image
    = 0.000000 + (C4-C3)/(C3+C4) * 1.000000
     Lines 1 to 1007 by 1.  Columns 1 to 1153 by 1

    Channels used:
        3-4

  Output Information:

    New output image file name: 'training_ndvi.tif'
    File format:               GeoTIFF
    Image type:                Multispectral
    Band interleave format:    BSQ
    Data type:                 Real
    Swap bytes:                No
    Signed data:               Yes
    Number of lines:                     1007
    Number of columns:                   1153
    Number of channels:                     1
    Number of bytes:                        4
    Number of bits:                        32
    Number of header bytes:               522
    Number of pre-line bytes:               0
    Number of post-line bytes:              0
    Number of pre-channel bytes:            0
    Number of post-channel bytes:           0
    Line start:                             1
    Column start:                           1

    -0.647059 is lowest calculated value
    3.40282e+038 is highest calculated value
    0 data values saturated at low end: -3.40282e+038 2,159
    data values saturated at high end: 3.40282e+038

  0 CPU seconds for reformatting.  01-20-2015  16:07:19
```

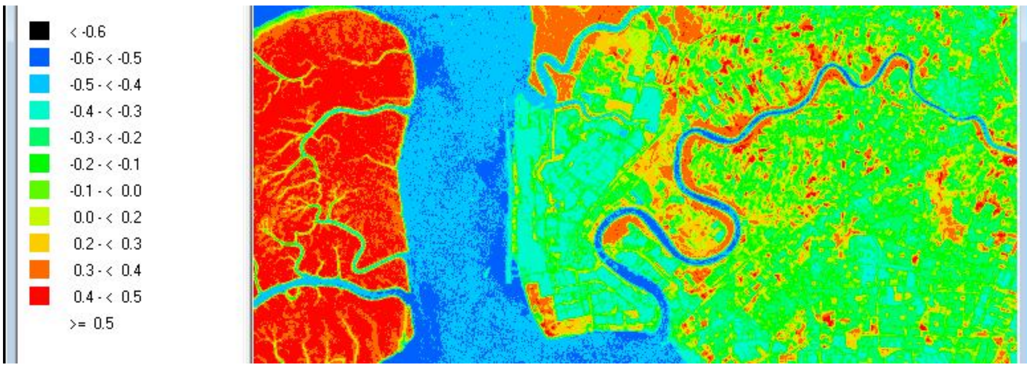
One should verify that the transformation equation used is as you expected. One should also check the lowest and highest calculated value to verify that these values make sense. One should also verify that the number of pixels saturated at the low and high ends are 0 or at least a very small portion of the total number of data values in the image. For this image, the highest calculated value is very high … the maximum possible 32-bit real number which seems a little strange. On checking though one will find that these 2,159 high values represent the calculated NDVI for the portion of the scene (lower left) which contain missing data (all 0’s); a limit value for these pixels make sense.

NOTE:

A large number of saturated data values may indicate that the data output type needs to be changed from integer to real. A very small range in data from the lowest to highest data value may also indicate that one needs to change the data type from integer to real or one needs to provide scaling and offset values to adjust the range of the output data values for better precision.

7.7 One can now display the resulting image as either a 1-channel gray scale type image or as a 1-channel thematic type as show below. See Exercise 1 for steps to do this if needed.

1-channel type display of the resulting NDVI image. High NDVI values which represent significant amounts of green vegetation are illustrated by light tones. Bare soil is represented by darker tones.



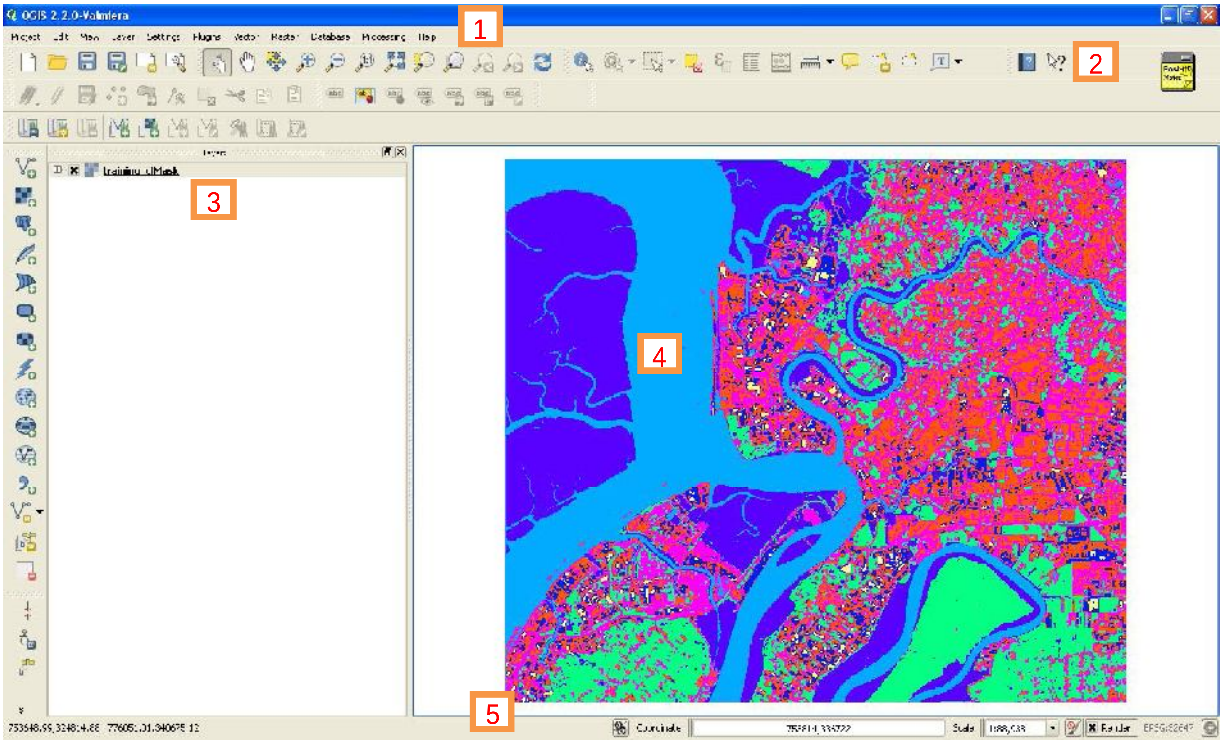
1-channel thematic display of the resulting NDVI image using the MODIS NDVI palette. The legend on the left relate the color tones with the NDVI values.

Quantum GIS (QGIS) Exercise : Working with vector and change detection

What is QGIS

QGIS is a user friendly Open Source Geographic Information System (GIS) licensed under the GNU General Public License that you helps you visualize, manage, edit, analyze, and compose maps with geographic data. QGIS is an official project of the Open Source Geospatial Foundation (OSGeo). It runs on Linux, Unix, Mac OSX, Windows and Android and supports numerous vector, raster, and database formats and functionalities. QGIS provides a continously growing number of capabilities provided by core functions and plugins

QGIS GUI



QGIS GUI can be divided into 6 main components:-

1. Menu Bar
2. Tool Bar
3. Map Legend
4. Map View
5. Status Bar

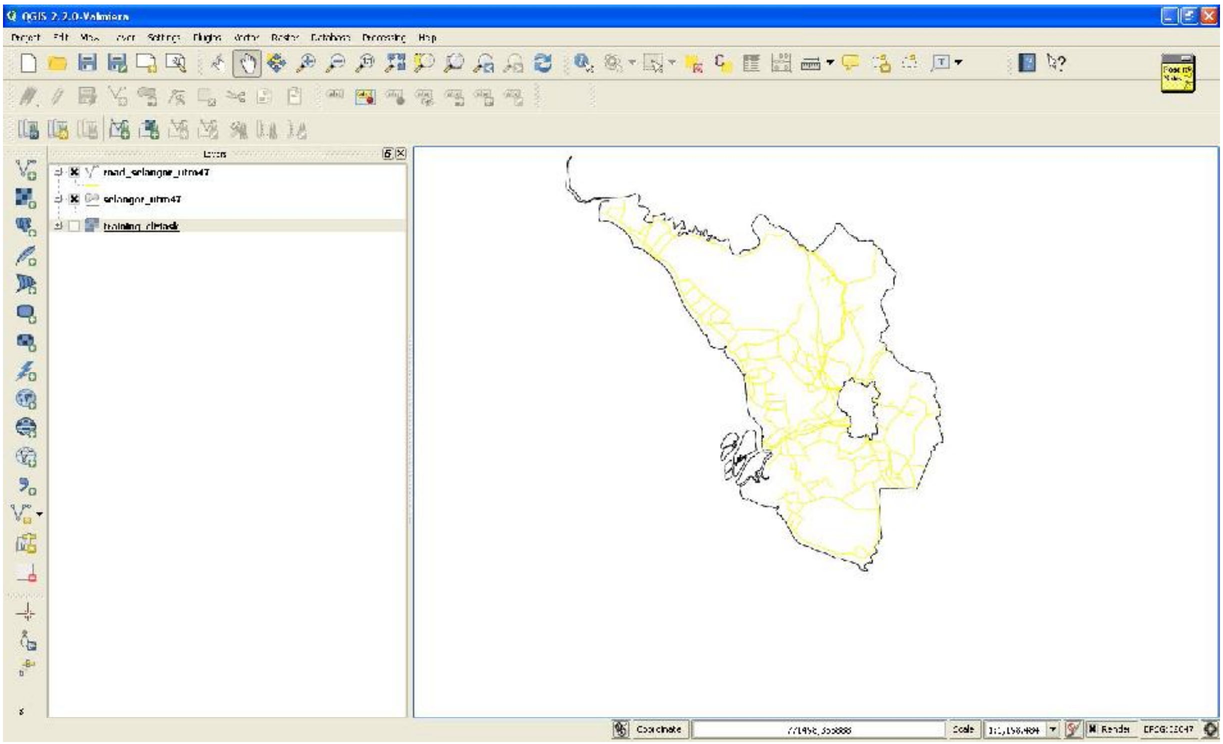
1. Working with vector data

Vector data is arguably the most common kind of data you will find in the daily use of GIS. It describes geographic data in terms of points, that may be connected into lines and polygons. Every object in a vector dataset is called a **feature**, and is associated with data that describes that feature. Vector data can be change in term of size, position and colors.

1.1 Start QGIS using the icon on the desktop or from QGIS Desktop from Startup Menu.

- 1.2 Select menu **Layer** and choose **Add Vector Layer**

- 1.3 At the dialog box vector layer, browse to **c:\latihan\selangor_utm47.shp** and **road_selangor_utm47.shp**
- 1.4 Try to change the color of the polygon and line



2. Working with raster data

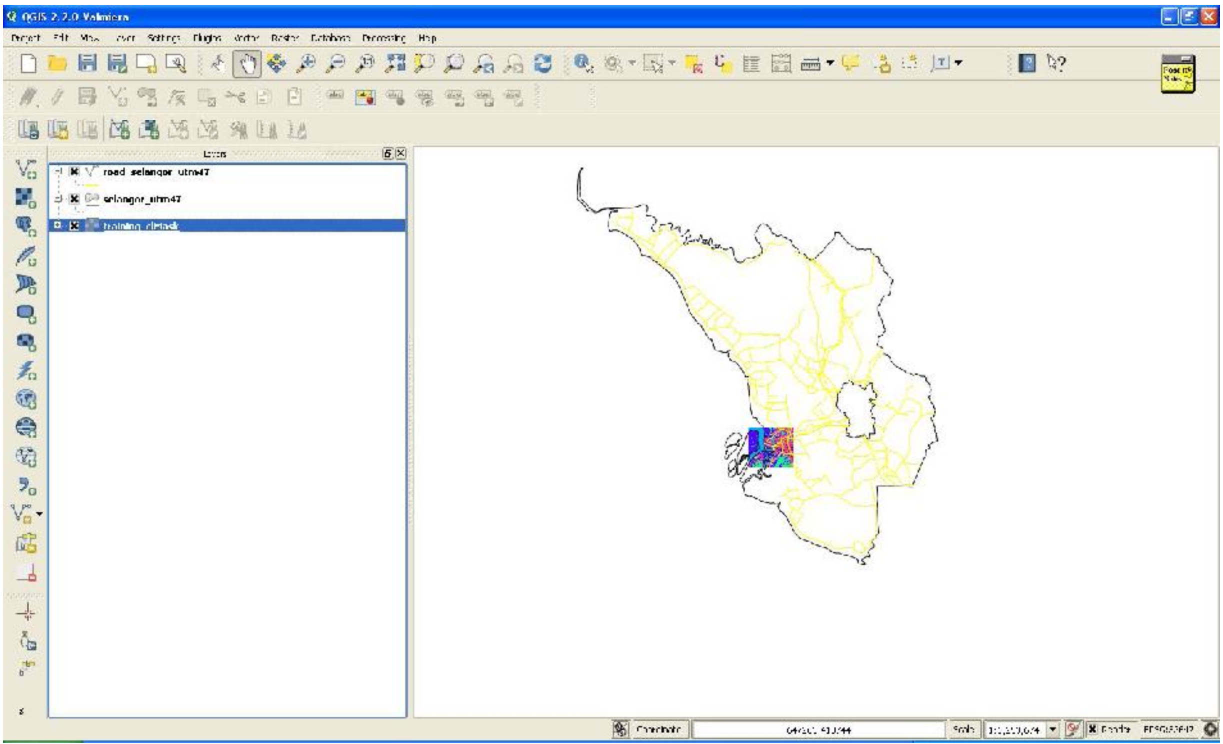
Raster data in GIS are matrices of discrete cells that represent features on, above or below the earth's surface. Each cell in the raster grid is the same size, and cells are usually rectangular (in QGIS they will always be rectangular). Typical raster datasets include remote sensing data, such as aerial photography, or satellite imagery and modelled data, such as an elevation matrix.

Unlike vector data, raster data typically do not have an associated database record for each cell. They are geocoded by pixel resolution and the x/y coordinate of a corner pixel of the raster layer. This allows QGIS to position the data correctly in the map canvas.

QGIS makes use of georeference information inside the raster layer (e.g., GeoTiff) or in an appropriate world file to properly display the data

2.1 Select menu **Layer** and choose **Add Raster Layer**

2.2 At the dialog box raster layer browse to **c:\latihan\training_clmask.tif**



3. Raster to vector

Converting between raster and vector formats allows you to make use of both raster and vector data when solving a GIS problem, as well as using the various analysis methods unique to these two forms of geographic data. This increases the flexibility you have when considering data sources and processing methods for solving a GIS problem.

To combine a raster and vector analysis, you need to convert the one type of data to the other. Let’s convert the raster result of the previous lesson to a vector.

The goal for this lesson: To get the raster result into a vector that can be used to complete the analysis.

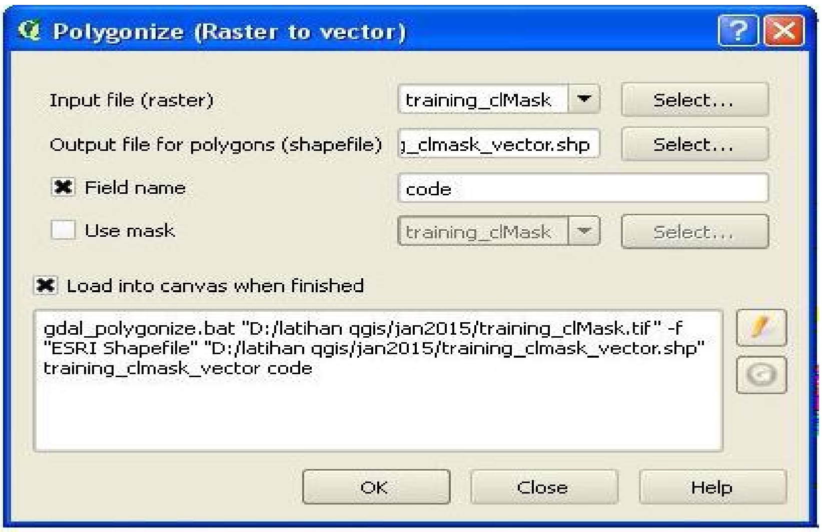
3.1 Select menu **Raster > Conversion** and choose **Polygonize (Raster to Vector)**

3.2 At the raster to vector dialog box, browse the input file to folder **c:\latihan\training_clmask.tif**

3.3 At the output file, save the output to **c:\latihan\training_clmask_vector.shp**

3.4 Type **Code** to the field name

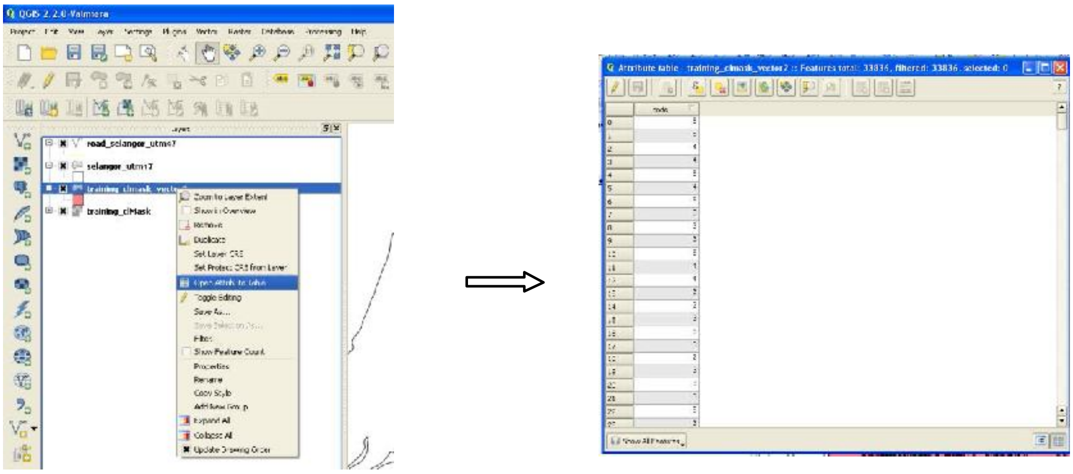
3.5 Click **close** when finish



4. Working with Attributes (Open Attribute Table)

GIS data has two parts - features and attributes. Attributes are structured data about each feature. Here you can make query, add another column, calculate area etc..

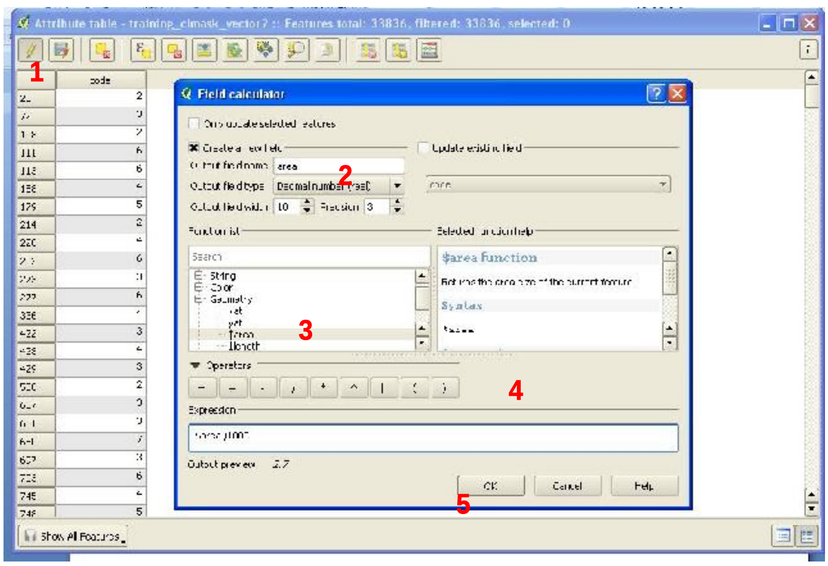
4.1 Right click on the layer and choose **Open Attribute Table**



4.2 Area Calculation in hectare

4.1 Click Start editing 

4.2 Click Open Field Calculator 



4.3 Save the changes 

5. Layer Symbology

The symbology of a layer is its visual appearance on the map. The basic strength of GIS over other ways of representing data with spatial aspects is that with GIS, you have a dynamic visual representation of the data you're working with.

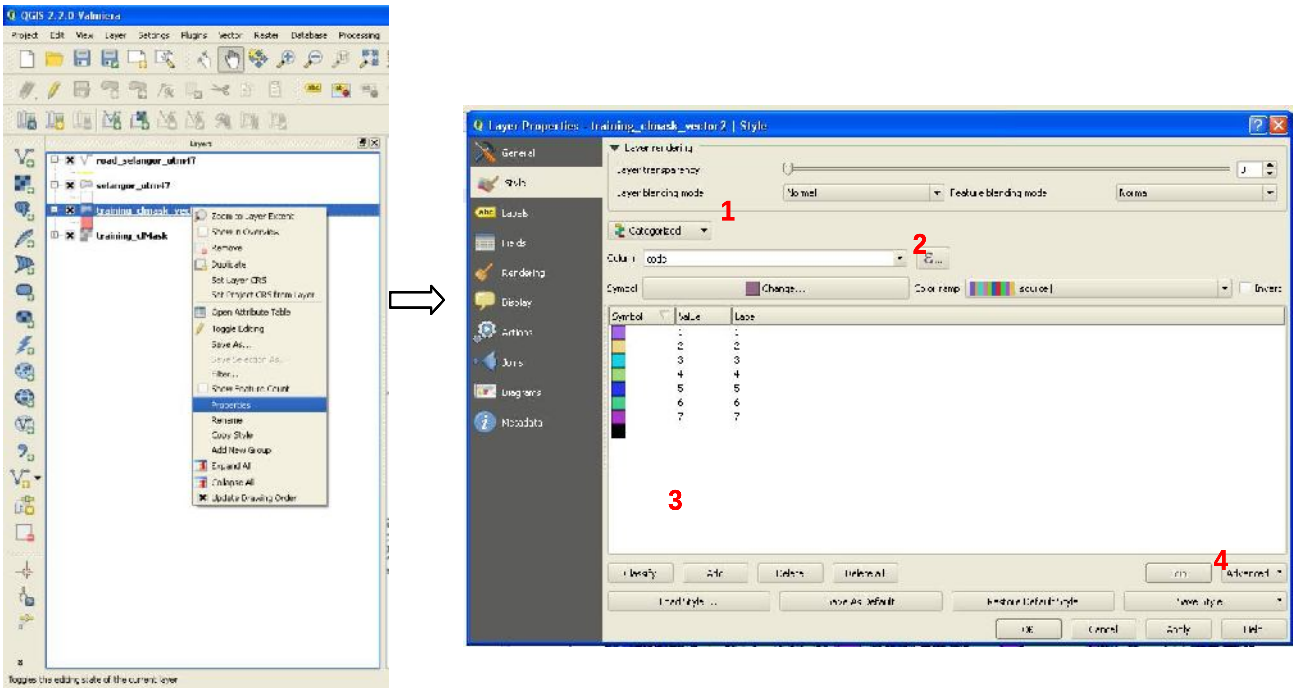
Therefore, the visual appearance of the map (which depends on the symbology of the individual layers) is very important. The end user of the maps you produce will need to be able to easily see what the map represents. Equally as important, you need to be able to explore the data as you're working with it, and good symbology helps a lot.

In other words, having proper symbology is not a luxury or just nice to have. In fact, it's essential for you to use a GIS properly and produce maps and information that people will be able to use.

The goal for this lesson: To be able to create any symbology you want for any vector layer.

5.1 Right click on the layer and choose Properties

52 Layer Properties dialog box will appear and follow the step below.

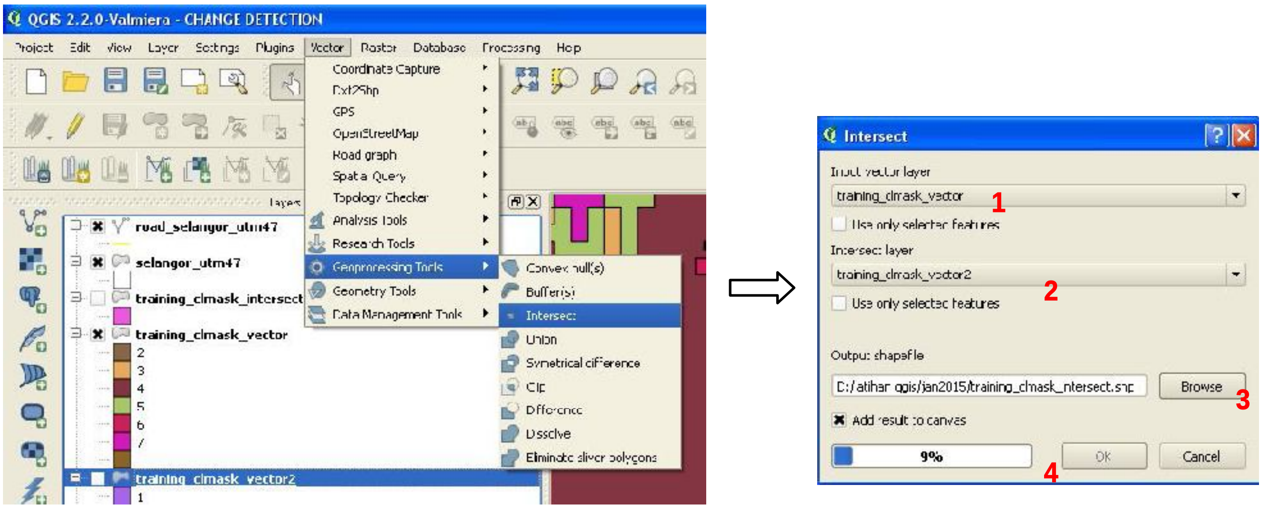


6. Vector Analysis (Change Detection using Intersect processing)

Vector data can also be analyzed to reveal how different features interact with each other in space. There are many different analysis-related functions in GIS, such as union, intersect, merge etc..

6.1 Select Menu Vector and choose Geoprocessing Tools and click Intersect

6.2 The Intersect dialog box will appear and follow the step below



Reference:

1. http://www.qgisworkshop.org/html/workshop/qgis.html
2. http://docs.qgis.org/2.2/en/docs/user_manual/
3. http://manual.linfiniti.com/