Lab 1: Importing Data, Rectification, Datums, Projections, and Coordinate Systems

Topics covered in this lab:
i.   Importing spatial data to TAS
ii.   Rectification
iii.   Conversion from latitude/longitude to UTM and resampling procedures
iv.   Datum transformations

Before starting this lab exercise, copy the following files into a working directory on your user space on the hard drive:
Image files:
- DEM_XY.grd
- Batch latlong.rst
- Batch latlong.rdc
- Uplands.dep
- Uplands.tas
- Legend .dep and .tas

Vector Files:
- Contours.asc
- Roads.asc
- Lakes.asc
- Rivers.asc
- Islands.asc

After completing this lab you will be able to:
- Import data into TAS from a variety of sources
- Collect and enter ground control points
- Rectify an image to a new coordinate system within a specified error level
- Converge from latitude/longitude to UTM coordinates
- Perform datum shifts

Introduction:
Very often when you collect spatial data from varying sources you will find that the data are stored using different datum and projections. When you overlay data in differing reference systems they will not align properly. Therefore, it is very important that all of the data in your spatial database use the same datum and coordinate system. This is one of the greatest challenges in building a spatial database and it is likely that you will experience this problem when you are acquiring data for your own project. These next few sections of Lab 1 will help you to resolve some of these problems.

i. Importing Vector and Raster Data

Often you will receive spatial data in different formats (i.e. file structures) and will need to convert the data to one that can be read by a specific Program. In this section of Lab 1, we will import various common types of GIS data into TAS.

Before we can import data into TAS, we need to set up a working directory under your user space where all files will be stored. Create a new folder and call it TAS_Working within your user space. Then, open TAS and set this as the default directory by going to the File Menu/Set working directory. After this is done, set the default palettes by going to File/system settings. The default image palette should be earthtones.plt and the default vector palette should be mono_black.plt. Everything else can remain as is. Now we are ready to import data into TAS.

1. Ensure that you have copied all of the data for this lab into your current working directory.

2. Click the import/export button on the TAS toolbar. The Import/Export Raster sub-program will be launched.
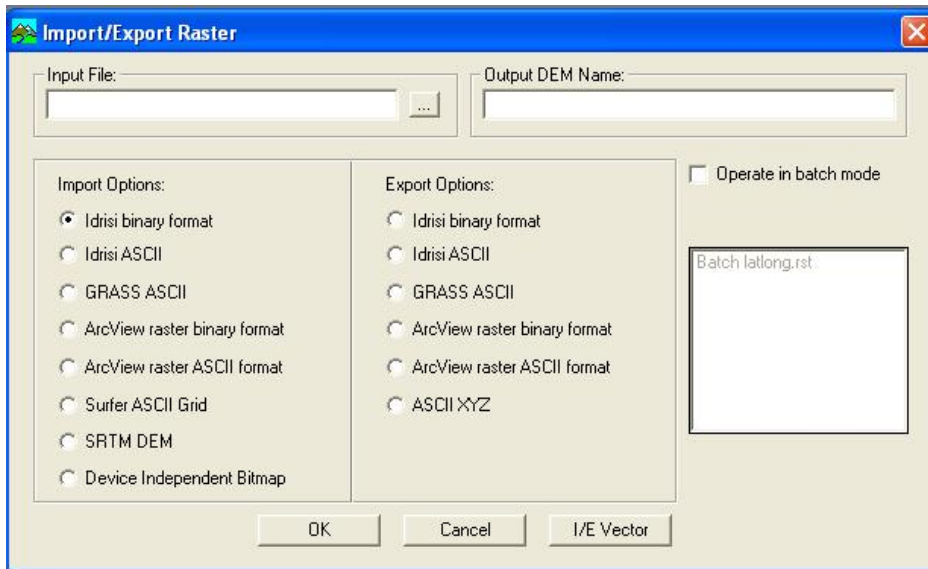
3. Select Idrisi ASCII under Import Options. Select 'Batch latlong' as the input file and call the output file 'Batch latlong'. Press OK. TAS will read the IDRISI ASCII (text) raster image and create a new TAS image with the same name. It will actually create two files 1) a .dep header file, and 2) a .tas data file.

4. Select ASCII raster format (ArcView) under Import Options. Select 'DEM_XY' as the input file and call the output file 'DEM XY'. Press OK. TAS will import the ArcView ASCII raster file.
5. Now let's import vector data to TAS. Currently, TAS only imports GRASS GIS and IDRISI ASCII vector files. Select GRASS ASCII under the 'Vector' sub-menu, in the 'Import/Export' heading in the 'File' menu. The 'Import/Export Vector File' window opens. Use this sub-program to import each of the .asc files in your current working directory.

ii. Rectification of Images

Sometimes spatial data that you acquire are in an arbitrary X-Y coordinate system. This is often the case with remotely sensed data, scanned aerial photographs, and digitized data where geo-referencing is simply in rows and columns. Image rectification is the process of fitting an image to a map coordinate system using ground-control points (GCPs). GCPs are well-defined and easily recognizable features that can be located accurately in the desired coordinate system and on the corresponding image. GCPs can be collected from a field survey using a global position system (GPS), gathered from a base map, or using another GIS image that has already been registered. Things to keep in mind when collecting GCPs:

- Select points that are easily identifiable on both your map (or registered image) and the image you wish to rectify.
- Use features that are relatively permanent like buildings, road intersections, distinctive land marks, meteorological stations, etc. If the study area is less developed, then inlets and outlets of streams along lake shorelines, islands, and stream tributary junctions are satisfactory GCPs.
- Collect as many GCPs as possible; it is better to have too many than too few.
- Try to get points that are distributed as evenly as possible around the image. Your GCPs shouldn't be bunched up in a corner or in a straight line across the image.

In this part of the lab, we will rectify a digital elevation model (DEM) of the Turkey Lakes area that was created from digitized contours. This DEM is on an arbitrary X-Y coordinate system and we wish to transform the data to the Universal Transverse Mercator (UTM) system. You will use the UTM eastings and northings from the GCPs you collected in class based on a 1:20,000 OBM map of the Turkey Lakes area. You should collect at least 20 GCPs to be used in the rectification.

1. Display the image 'DEM XY'.
2. Use the 'Vector Overlay' button on the 'Image Attributes' toolbar on the left side of your screen to overlay the following vector files on top of 'DEM XY':

Vector File Name	Palette	Fill Polygon	Line Width
contours	mono_brown	No	1
lakes	mono_lightblue	Yes	1
islands	mono_blue	No	1
rivers	mono_blue	No	1
roads	mono_black	No	2

3. Place your cursor over the same GCPs that you collected on your base map, and record the xy coordinates along with their corresponding UTM coordinates in the TAS text editor (File/text editor) using the following structure :

x coordinate	y coordinate	UTM easting	UTM northing
x coordinate	y coordinate	UTM easting	UTM northing
x coordinate	y coordinate	UTM easting	UTM northing
…	…	…	…

NOTICE: numbers must have tabs in between and NOT spaces. When you are finished, save the file as a text file (.txt) called 'Turkey_Lakes_GCPs'.

4. Close any open files. Select 'Rectification' under the 'Pre-processing' menu. TAS's rectification sub-program will be launched. This sub-program can be used to rectify vector data as well as images, although we will only be rectifying images in this lab.
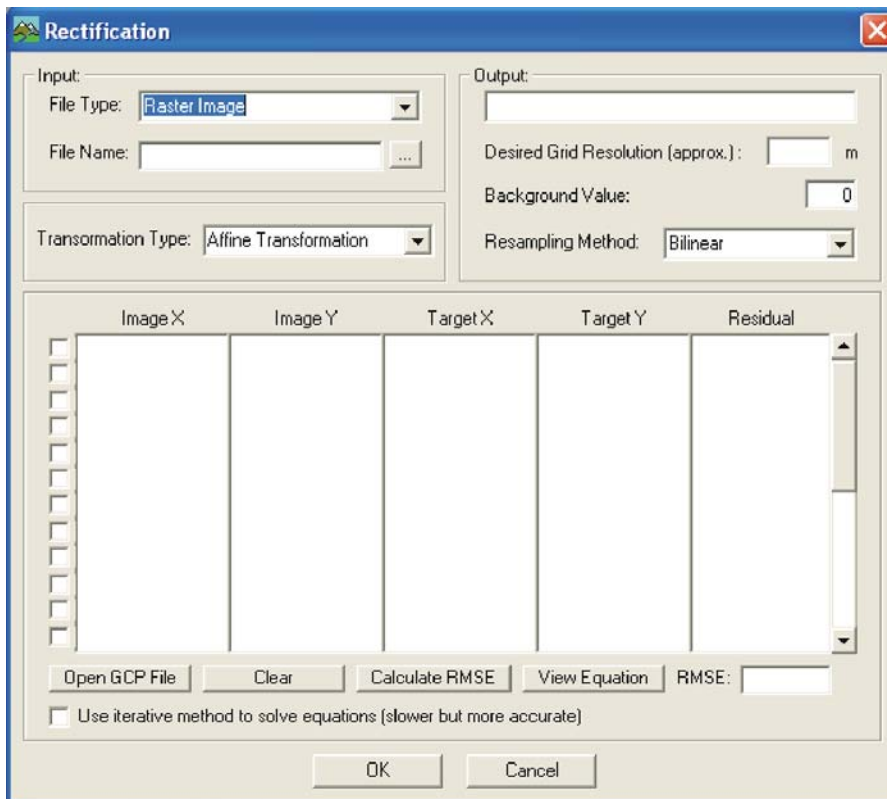
5. Choose 'DEM XY' as the input file name and call the output file 'DEM UTM'. We would like the rectified image to have the same grid resolution as the XY image, which is 5 m.
6. Open your GCP text file using the 'Open GCP File' button near the bottom left of the window. You should find that your points appear in the Image X, Image Y, Target X, and Target Y columns.
7. Press the 'Calculate RMSE' button. This will calculate the transformation equation, allow you to view the residual error for each GCP, and calculate the root mean square error (RMSE), which is a measure of the average error in the transformation.

8. Each GCP should have a check mark at the left-hand side, indicating that it was used to calculate the transformation equation. If there are any points that have residual errors that are much higher than most of the other GCPs, uncheck its box and re-calculate the RMSE. How does this affect your RMSE?

You have just performed a regression analysis to find the plane that fits your GCPs the best (i.e. with the least residuals). Notice that under the transformation type combo box we have selected the default of 'Affine Transformation'. An affine transformation fits a first-order polynomial (i.e., a flat plane) to your GCPs. This transformation type can accomplish scaling, rotation, shearing, and reflection but not warping. The other options in the transformation type combo box include second-order polynomial and third-order polynomial.

9. Check the 'Use iterative method to solve equations' box at the lower left-hand corner of the window. Re-check any GCPs that you have previously unselected. Choose a second-order polynomial transformation type and re-run your analysis. Now try a third-order transformation type.

Q1: Comment on the differences between the residual error and RMSE derived using each of the transformation types. Explain the effect of higher-order transformations on rectification of the Turkey Lakes DEM. (3 marks)

10. Select a third-order polynomial transformation type and remove any GCP points that have anomalously high residual errors. When you are satisfied that the remaining error is acceptable press the 'View Equation' button. A text box will open displaying the transformation equation and the residual for each point.

Q2: Print the equation document and hand it in with the written component of your assignment (1 mark).

11. Press the OK button. The program will run and 'DEM XY' will be transformed according to the calculated transformation equation. When it has finished, display 'DEM UTM'.

Q3: How has rectification changed 'DEM XY'? Examine the rectified image's file attributes (file menu ==> 'File Attributes'). What are the minimum and maximum X and Y values? (3 marks)

Q4: Define Root Mean Square Error. In the process of map rectification, what is the threshold in RMSE that they GIS analyst attempts to stay below? How did your rectification compare to this threshold? Suggest some ways that you could have improved the accuracy of the rectification (5 marks)

Q5: Comment on the distribution of your GCP points (only the ones that were used in the final equation). Have you used a higher density of points in some regions of the Turkey Lakes Watershed than others? (2 marks)

iii. Conversion from Lat/Long to UTM

Spatial data are frequently stored in geographic coordinates (i.e., latitude and longitude) while analysis of topographic data often requires data in the Universal Transverse Mercator (UTM) coordinate system. Therefore, you must convert any data you have in geographic coordinates to the UTM system before performing any *TAS* analysis. The *TAS* module, 'Coordinate System Transformations' can be used to convert geographic coordinates to UTM eastings and northings and vice versa. You will find this module under the 'Pre-processing' menu.

Coordinate system transformations can be applied to both vector and raster data by specifying the appropriate 'File Type'.
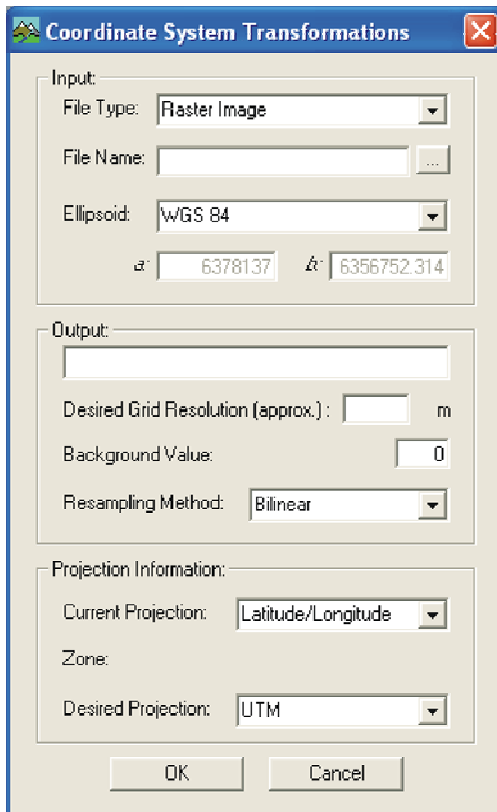

1. Choose 'Batch latlong' as the input file name. This image is a Shuttle Radar Topography Mission (SRTM) DEM of a portion of the Batchawana River, just north of Sault Ste. Marie.
2. SRTM DEMs use the WGS 84 ellipsoid, so keep the default settings for the 'Ellipsoid' combo box.
3. Call the output image 'Batch UTM'.
4. Each grid cell in the SRTM DEM has a length of 0.0008333^o (3 arc seconds), which is approximately 90 m on the ground. Specify that you would like the output grid resolution to be 90 m.
5. Maintain the default settings for everything else and press 'OK'.
6. Display 'Batch latlong' and 'Batch UTM' using the 'earthtones' palette. Set the min Z value of 'Batch UTM' so that it is the same as 'Batch latlong' and compare the two images.

Q6: What coordinate system is used for the Earth's spheroid? In the transformation of the Earth from a spheroid to a flat map, what geometric properties of the map may be altered? What projection and coordinate system are commonly used for maps of the Earth's topography? (3 marks)

A Note on the Resampling Method:
When we apply a coordinate transformation to a raster image *TAS* essentially maps the existing grid in Lat/Long coordinates onto a new grid in UTM coordinates. It is likely that the centre of each cell in the first grid will not map onto the centre of a corresponding cell in the second grid. So how do we know what value to assign to the cells in our UTM grid?
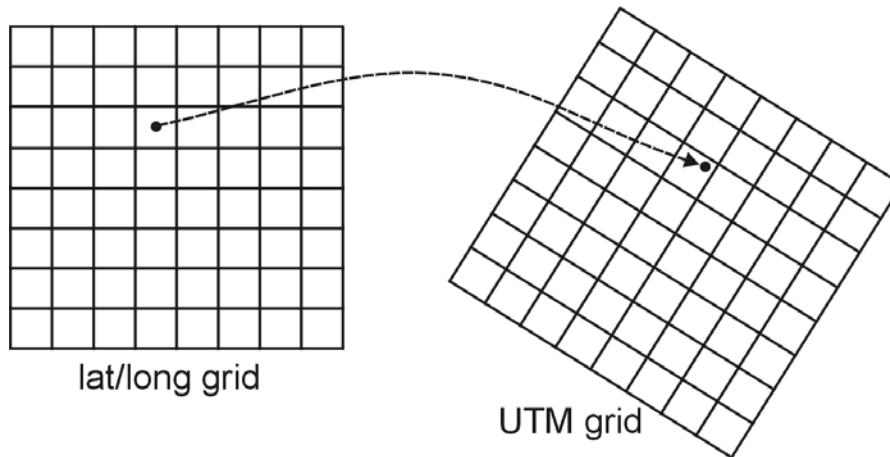
Resampling procedures are used to solve this problem. The simplest resampling procedure is called the *nearest neighbour* method and works by assigning each cell in the UTM map the value of the closest cell centre in the Lat/Long image. The second (and default) resampling procedure in *TAS* is called *bilinear* interpolation. This procedure uses a distance-weighted combination of the four closest grid cells in the Lat/Long grid to assign values to the UTM grid. Nearest neighbour resampling is generally faster to run and ensures that values in the output grid are identical to those of the input grid. Bilinear interpolation, however, provides a more accurate output image for transforming continuous data like DEMs.

7. Close all open images.
8. The image 'uplands' is a Boolean image (i.e., 1's and 0's) identifying all the areas in 'Batch latlong' with an elevation greater than 450 m. Convert this image from geographic coordinates to UTM coordinates using the same settings described in steps 1 through 6. Call the output image 'uplands UTM bilinear'
9. Convert 'uplands' to UTM coordinates again, however, this time use the nearest neighbour resampling method. Call the output image 'uplands UTM nearest neighbour'.
10. Display 'upland', 'uplands UTM bilinear' and 'uplands UTM nearest neighbour' using the 'quantitative' palette.

Q7: Which resampling methods is more appropriate for transforming discrete (i.e. non-continuous) entity data and why? (2 marks)

iv. Changing Datums

There are several reasons that you may want to perform a datum conversion on either a vector file or raster image. If you acquire data from varying sources that are referenced to different datum you will need to transform the data to a single datum. Also, international standards for ellipsoids and datums can change so you may need to update your spatial database occasionally. 'Batch UTM' is currently referenced to the WGS 84 ellipsoid and datum. In this section of the lab we will perform a datum conversion of this image to the North American Datum 27 (NAD 27), which uses the Clarke 1866 ellipsoid.

1. Select 'Datum Transformation' under the 'Pre-processing' menu.
2. Choose 'Batch UTM' as the input file name. This file has a UTM coordinate system. Change the 'Zone' combo box to read 16N (i.e., UTM zone 16 northern hemisphere), the UTM zone for this particular part of the world. Keep the default elevation units; this DEM is in meters.
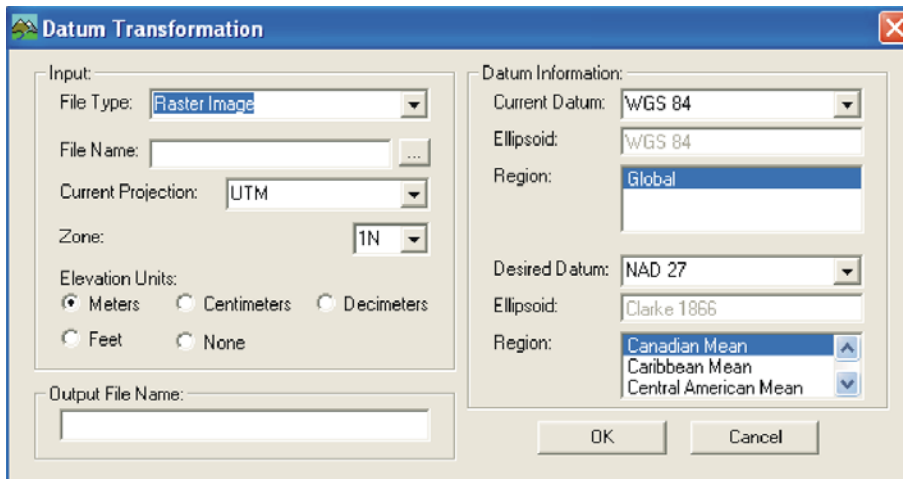
3. Call the output file 'Batch UTM NAD27'.
4. All SRTM DEMs are referenced to the WGS 84 datum and ellipsoid, so keep the default for the current datum.
5. We wish to transform this DEM to the NAD 27 datum, the default. However, to familiarize yourself with this program, try selecting a few other datums in the 'Desired Datum' combo box. Notice that when you change the datum, the reference Ellipsoid and Region boxes are updated. The Region box indicates specific locations for which the selected datum is commonly used. Using the closest possible region to match your study area will improve the accuracy of the datum transformation. Re-select the NAD 27 datum and select the Manitoba & Ontario region. Press OK to perform the datum transformation.

6. Display 'Batch UTM' and 'Batch UTM NAD27' using the earthtones palette.

Q8: How has the datum transformation affected the elevations of the DEM? What about the minimum and maximum X and Y coordinates? Explain why these changes have occurred. (3 marks)

Note: File attributes for the files can be found under the File menu/File Attributes.