West African Drylands Project

An Ecosystem Approach to Restoring West African Drylands and Improving Livelihoods through Agroforestry-based Land Management Interventions.

A United Nations Environment Programme (UNEP) project conducted in partnership with the World Agroforestry Centre (ICRAF), the Centre for Environmental Policy of the University of Florida, and the Governments of Burkina Faso, Mali, Mauritania, Niger and Senegal. The project is funded by the Government of Norway.

Training module created by Thomas Gumbricht, www.mapjourney.com

Last updated: March 2008

LESSON 4 – GRID/RASTER DATA

Lesson overview

In this lesson you will learn about the difference between vector and grid; to symbolize grid data, to work with stacks of multiple grid layers in DIVA-GIS, and the use map algebra to produce new grid layers. The lesson also introduces interpretation of data from multiple layers using both visualization and some built-in analysis tools in DIVA-GIS. The lesson introduces several global datasets on vegetation useful for interpreting land degradation.

Vector and raster data

Open project

Start DIVA-GIS by clicking the icon for DIVA-GIS on the desktop (see lesson 0 to create an icon on the desktop if you do not have one). Open the project from lesson 3 (\data_spatial\sahel\my diva\lesson3.div) by either clicking the Open Project tool, 🖾, or via the menu: Project – Open.

Vector and grid data

Thus far you have worked mostly with vector data, and one example of image (jpg) data (in lesson 1). Now it is time to look at grid (or raster) data. Grid or raster data are built like image data, with rows and columns of cells or picture elements (pixels) that build up a data layer. Each pixel represents a value, like elevation, vegetation density or a land cover class. Each pixel is of a certain size – the resolution of the grid data.

The land cover and vegetation data you will use for the whole of the Sahel in this lesson is built from raster data where each cell represents 500 m * 500 m (in GIS jargon the grid resolution is 500 m). In the raster data the cell is carrying the information about the attribute, not like in vector data where the attributes are stored in a separate database. More advanced GIS software can have a database also associated with a grid file, but not DIVA-GIS. The most common is, however to let grid files represent a single feature.

Natural phenomena like vegetation density, rainfall, elevation and land cover tend to vary continuously over space. Rainfall does not jump at a certain location, but changes gradually. As raster data is portraying a surface with small cells it is often better to use raster data for representing

natural phenomena. Another advantage with raster data is that it is easy to compare two datasets, say for instance rainfall and vegetation growth. A disadvantage with raster data is that it takes more storage space in the computer memory, and that it is less precise than vector data.
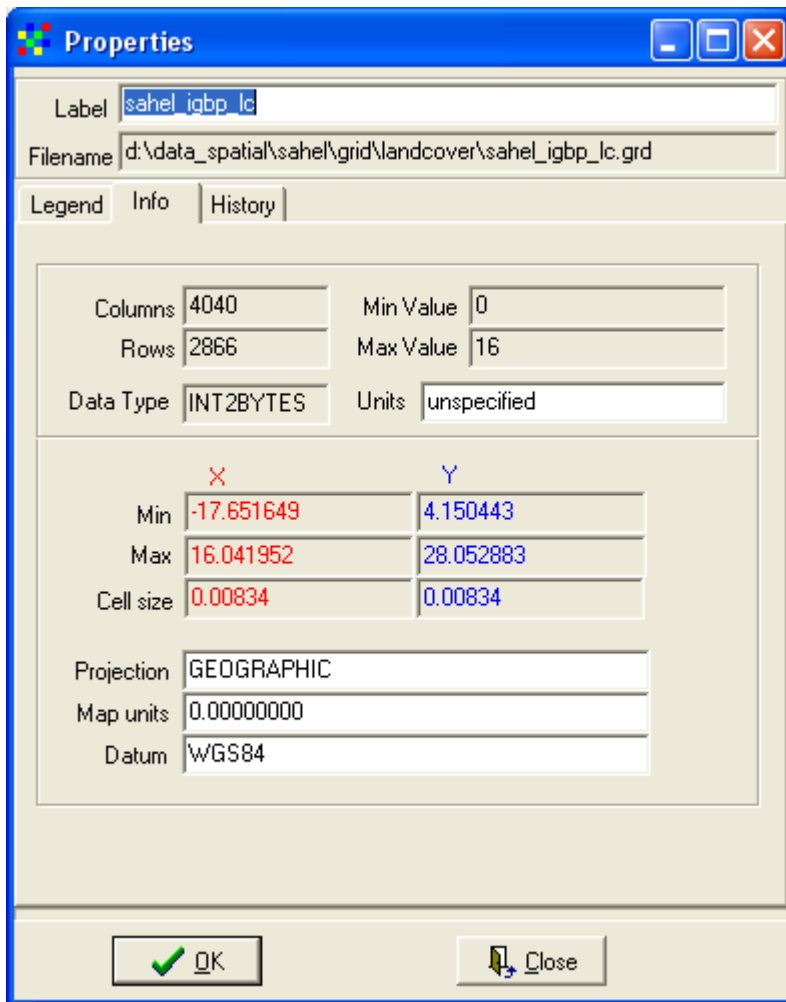
Things of a human origin tend to be discrete (have abrupt boundaries). Hence vector data is more suitable for mapping out objects like roads, administrative boundaries and other phenomena that are of human origin. The advantage with vector data is that it is more precise, takes less storage space, and can be linked to a database containing many attributes. But it is troublesome to use vector data for map calculations and for portraying continuously varying phenomena.

Add grid data

The grid data available for West Africa (Sahel) is under the folder \data_spatial\sahel\grid on the project CD. Navigave to the sub-folder \data_spatial\sahel\grid \landcover and add both the grid (*.grd) and image (*.jpg) file for the layer sahel_igbp_lc. The image file is a ready-made image of the land cover, the grid file contains the data used for creating the image. As the image and the grid have the same name, DIVA_GIS links them and you can have the image as active theme, and use the Identify tool to get information from the grid theme. Note that for both the image theme and the grid theme the Table tool, , remains inactive (appears grey) and you can not use it. That is because there is no attribute data associated with a grid or an image in DIVA-GIS. More advanced GIS packages have solutions for attaching a database also to grid files, but not DIVA-GIS.

Grid data properties

The grid file sahel_igbp_lc is not symbolized, and hence you need to symbolize it. Open the Properties window for the sahel_igbp_lc.grd, make it the active theme and double click it in the Legend. The Properties window for grid layers looks a bit different compared to vector files, but works in a similar way. It has three tabs, Legend, Info and History. Under the Info tab you can see size of the image (Columns and Rows), and the data type (in this case INT2BYTES, which means that the data is made up of integer values), the Min and Max values of the data in the grid (0 to 16 for sahel_igbp_lc.grd), and information about the extent, and which Projection the data is in.

Properties

Label: sahel_igbp_lc

Filename: d:\data_spatial\sahel\grid\landcover\sahel_igbp_lc.grd

Legend | **Info** | **History**

Columns	4040	Min Value	0
Rows	2866	Max Value	16
Data Type	INT2BYTES	Units	unspecified

	X	Y
Min	-17.651649	4.150443
Max	16.041952	28.052883
Cell size	0.00834	0.00834

Projection	GEOGRAPHIC
Map units	0.00000000
Datum	WGS84

✓ OK   Close

Under the History tab, some metadata is usually found about how the layer was created. But for now we are mostly interested in the Legend tab.

Symbolize grid data

Symbolize nominal grid data

In the Properties window click the Legend tab to get to the symbolization options for grid data in DIVA-GIS.
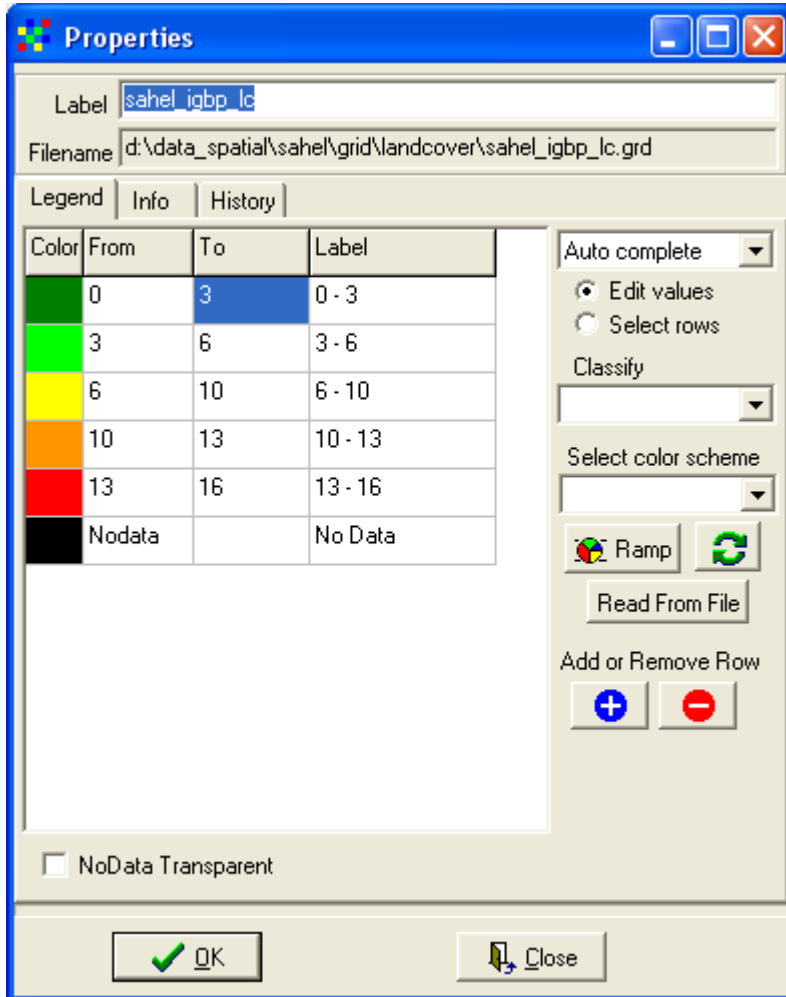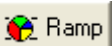

Under the Legend tab you can symbolize the grid theme. The sahel_igbp_lc.grd theme has 17 classes (0-16) as you saw under the Info tab. This is also reflected in the legend. The legend is defaulted to display numerical data of interval or ratio type, but we actually have nominal data (land cover types is a nominal data type, if you need to repeat the different data types in GIS, look at lesson 2). We need to create a symbolization that reflects the 17 classes. First you must create 17 entries (plus Nodata) in the Legend. Click on the Insert Row button, [image], until you get 17 classes.

If you added too many rows, use the Delete Rows button, [image], to remove them. As we have nominal data we must fill in the values using the option manual from the drop menu next to the Label column as shown below. Then you must fill in values for each land cover class as shown below.
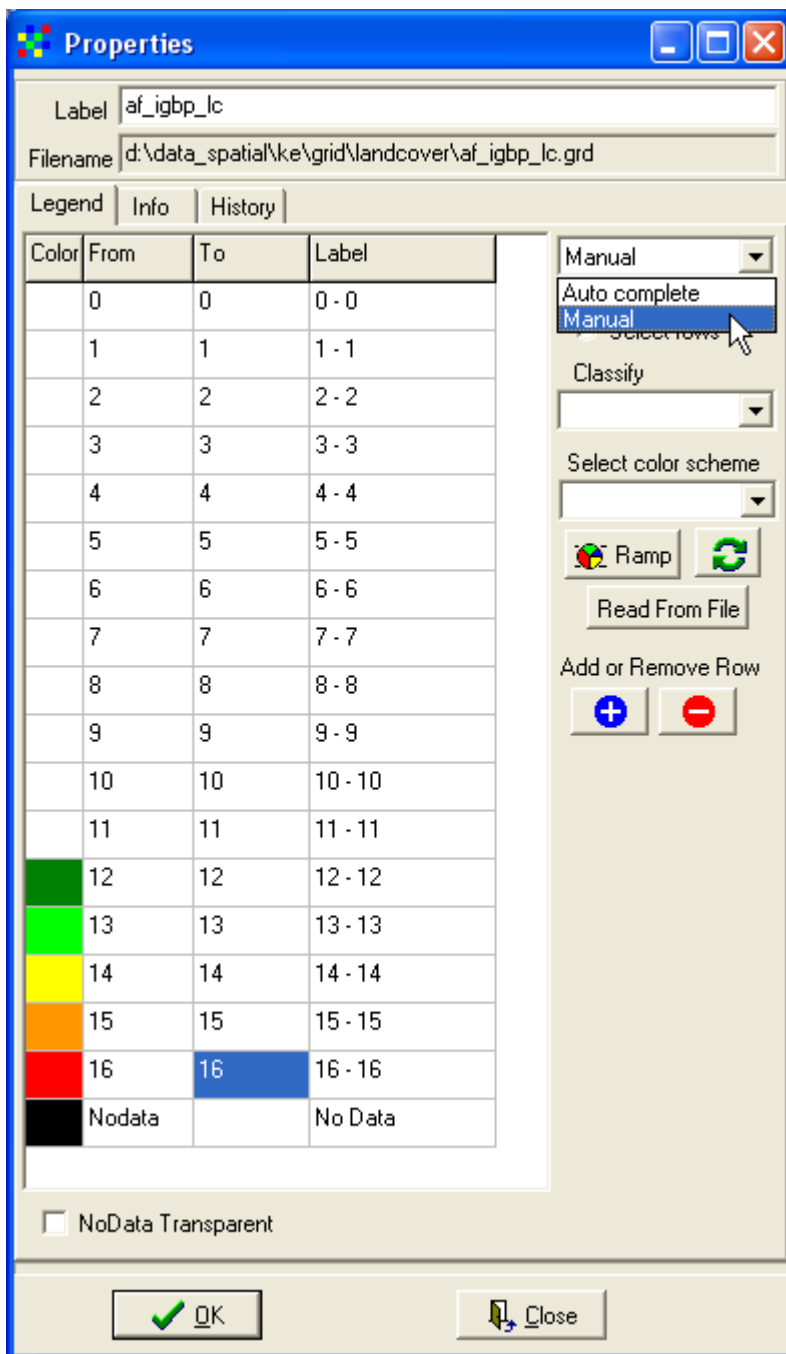
You can put a default symbolization using one of the options from the Select color scheme drop down menu.

If you set the colors of the first and the last rows (excluding Nodata) you can ramp the intermediate colors by pressing the Ramp button, . You can also swap the colors by clicking the Swap

button, . As we, however have nominal data in sahel_igbp_lc, you can not escape setting the colors individually.

The symbolizing of the image file sahel_igbp_lc.jpg is done using standard colors for land cover also used elsewhere. You can get to the metadata and legend through the project data web-catalog, but the symbolization colors are also shown in the figure below.

Ideally you should now try to put the same colors in the Legend for the grid theme with same name, and at the same time enter the correct label in the Label field. To change a color in the Properties window, double click the cell for Color and set the color using the Color window (exactly as with vector data).

The exact color coding for each class can be found via the web-catalog on the project CD. Click OK in the Properties window when you have symbolized the 16 land cover classes.
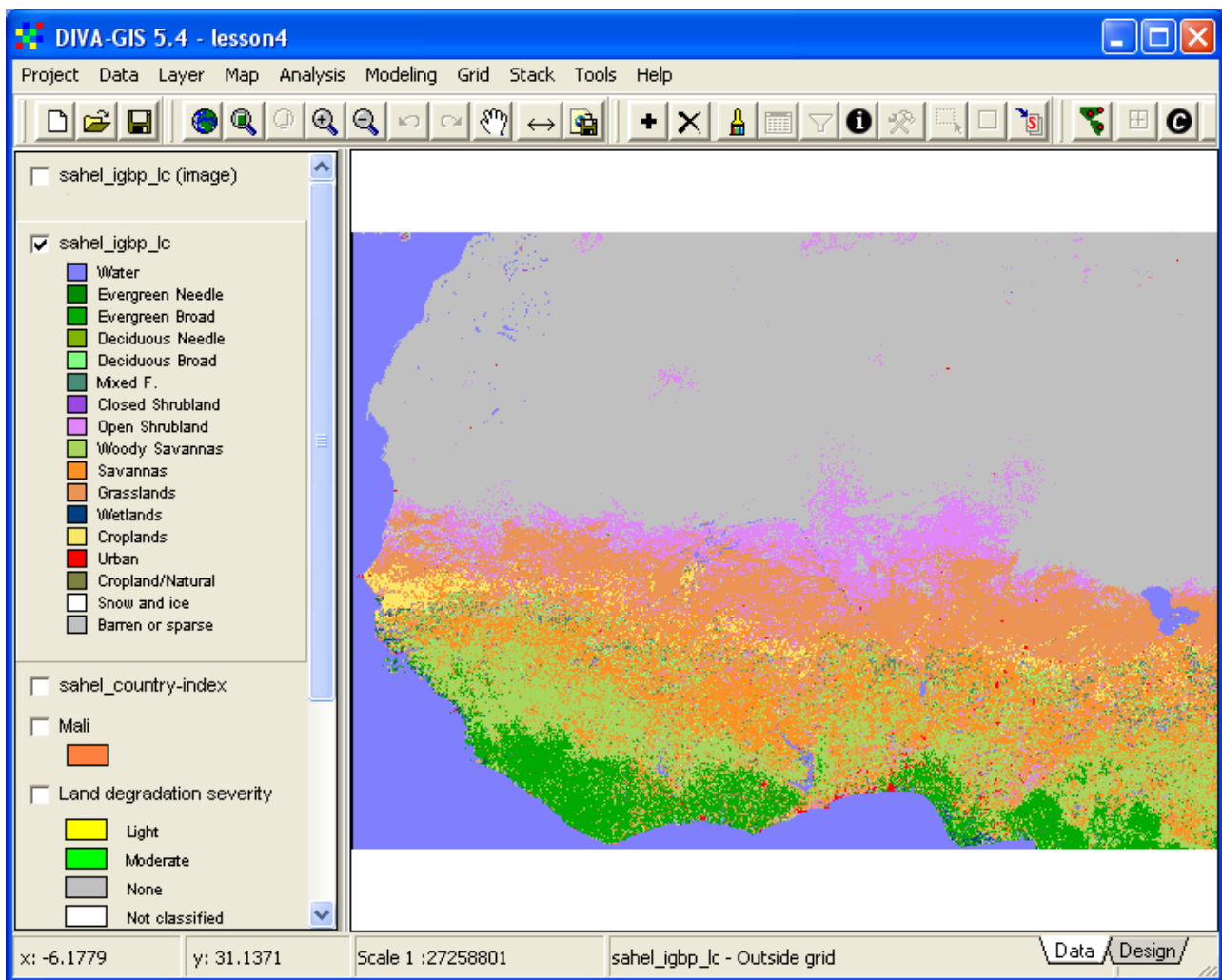
The land cover data that you have used is derived from satellite data. It was done from analyzing the 2001 annual cycle or vegetation growth (the vegetation phenology) using the MODIS (MODerate Image Spectrometer) sensor that takes a picture of the whole Earth every second day, see the project data web-catalog for details.

Symbolize ratio grid data

MODIS data has also been used for creating global maps of woody cover, herbaceous cover and bare soil. These three data layers you can find in the folder \data_spatial\sahel\grid\veg_fields. Use the Add Layer tool, ➕, to add them to the DIVA-GIS project. These vegetation field data layers have values from 0-100, with the value 253 denoting Nodata (water areas). They are hence of ratio type (60 % cover is double as much as 30 % cover). They are much easier to symbolize.

Make the layer for herbaceous cover (herbcov_2001) active and open the Properties window. Use the Insert Row button, ➕, to create 10 classes. Then select Auto Complete from the drop down menu to the right of the Label columns (as shown below). You now only need to fill in the values in the To column and DIVA-GIS will fill in the From column automatically (you only have to put 0 in the topmost From row). Set the color of the first row (0-10) and the last row (90-100) and press the Ramp button, 🌀 Ramp.
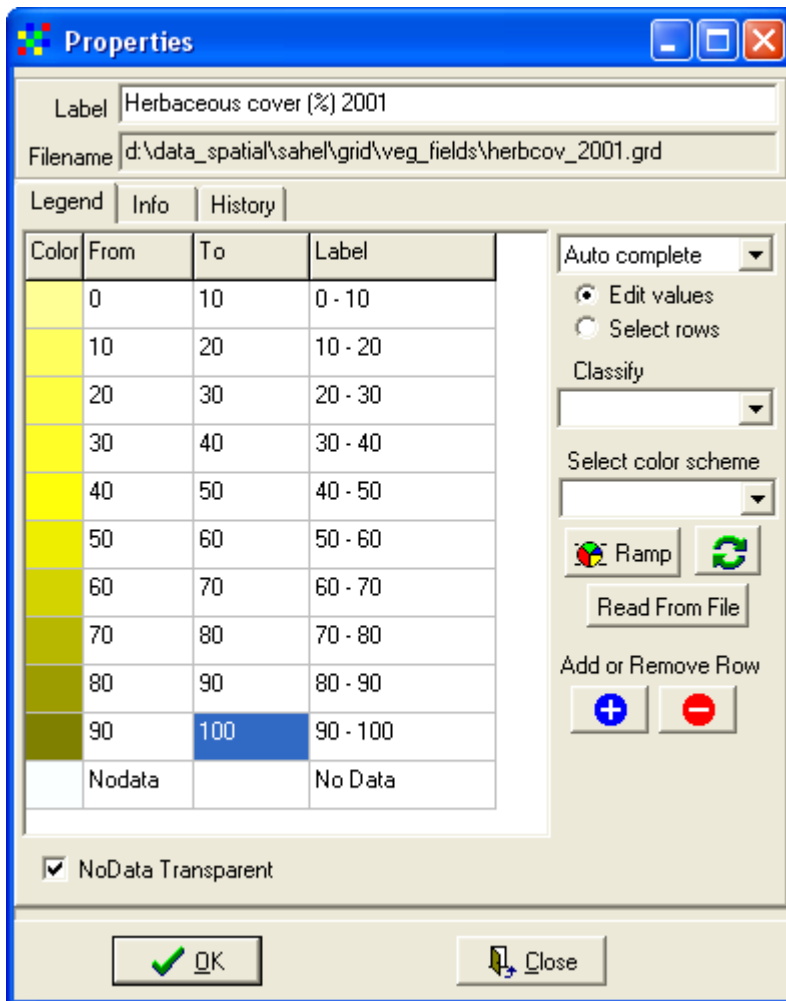
Give the layer a more descriptive label, select the checkbox for NoData Transparent, and click the OK button.

Symbolize by loading legend from file

You can now symbolize the layers for tree cover and bare a bit easier. Make bare cover the active theme and open the Properties window, e.g. by clicking the Properties tool, [icon]. In the properties window select the Legend tab, and click the Read From File button, [Read From File] (just below the middle on the right side under the legend tab – see figure above). Select herbcov_2001.grd and the symbolization you did above will be adopted also for barecov_2001. You now only need to change the colors of the first and last rows and click the Ramp button to generate a new color ramp. Click the OK button to apply the changes to the map in the Data View.

Aggregate grid

As mentioned above grids can be very large, and hence be slow to process. The grid layers for vegetation field cover are rather large, but you can reduce them in size by using the Aggregate function in DIVA-GIS. The grid processing available in DIVA-GIS is found under the grid menu. To reduce a grid size by aggregating go via the menu: Grid – Aggregate. In the example below the resolution of the treecover grid layer was reduced from 500 m to 8 km. The file size is reduced 16*16=256 times, so the 8 km resolution grid is very much faster to work with. If you still want to use the full 500 m resolution data you can cut the grid for the whole of West Africa down to fit your

area of interest (from the menu: Grid – Cut – see lesson 5 for details). In the example below the aggregation was done using mean values and if the rows and columns did not match the reduction factor the new grid was set to be truncated to fit the new resolution. (If you set the fitting function to Truncate the Output grid will get the same number of rows and columns as the vegetation and rainfall grids included in the Sahel dataset, if you set the option Expand the Output grids will be 1 row and 1 column larger, disallowing comparing the grids directly – also see the Note on grid data dimensions at the end of this lesson).


Open the Properties window for the 8 km grid that you used created, and click the Read from file button, Read From File, under the Legend tab. Use the original treecov_2001 grid layer as template by choosing it from the Open File window that opens. Click the OK button to apply the changes in the legend.

In the example below, the layer with your selected countries created in Lesson 3 was put on top in the Legend, with the aggregated tree cover layer below. Then Zoom to Theme, was clicked to zoom into the layer with selected countries.
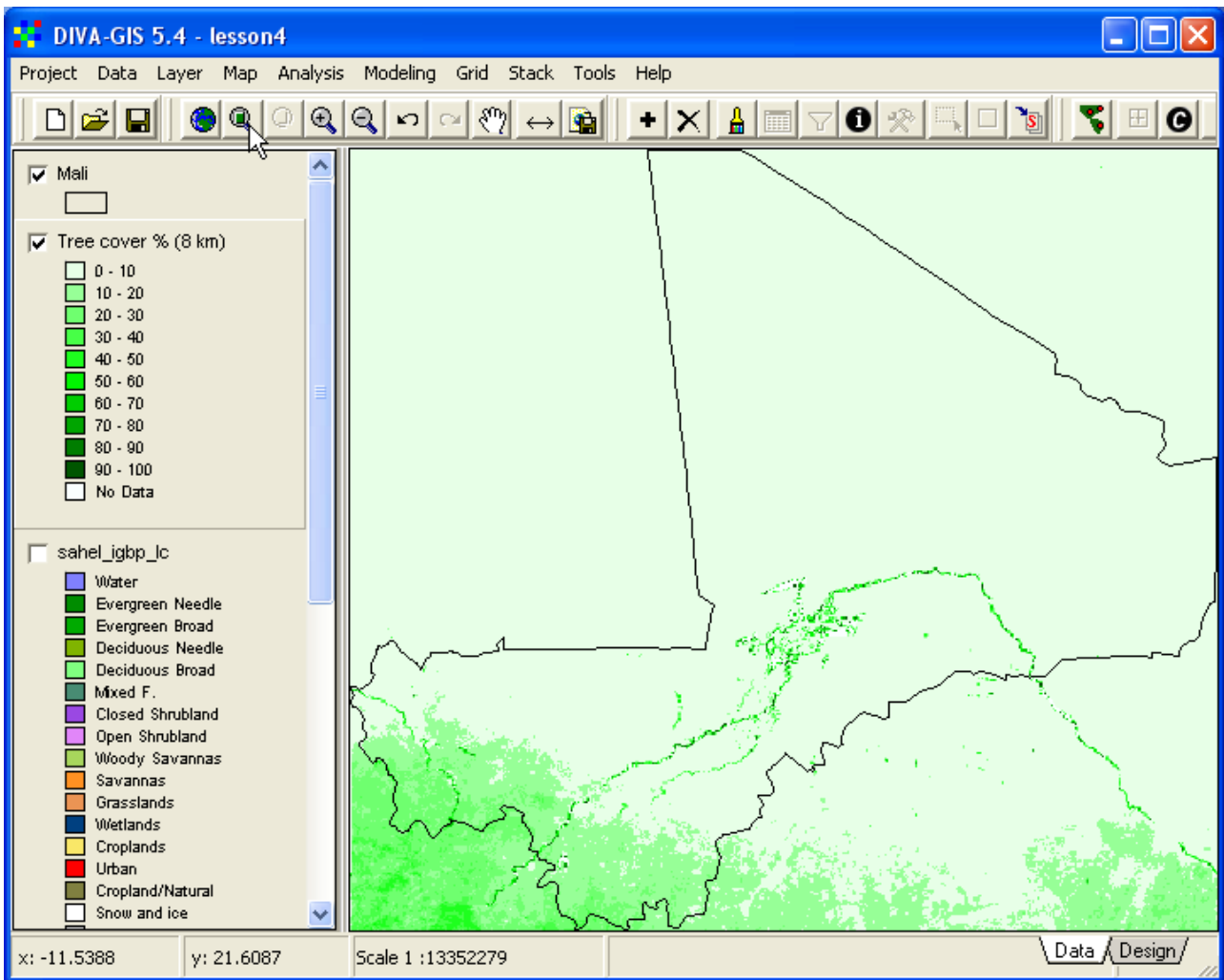
Identify multiple features

You can retrieve information from more than one layer at a time using the Identify Feature tool, . But first you need to have a better overview of your layers, hence hide all legends by right clicking on any Legend entry and select, Hide All Legends, as shown below.
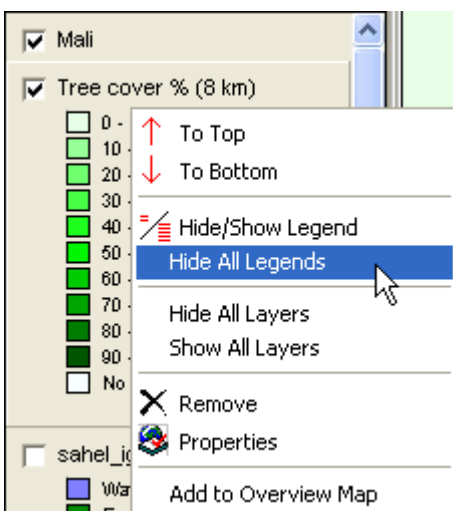
With the legends hidden it is easier to rearrange the display order, and in the example below the vegetation fields are put together in the legend. Make the three aggregated vegetation field layers active, by clicking each of them while holding down both the Ctrl + Shift keys.

Then select the Identify Feature tool, , and click in the map (e.g. in your selected country). The Identify window will popup and have as many tabs as there were selected data layers in the Legend. In the example below there are three tabs (for Tree cover, Herbaceous cover and Bare cover). The location that was clicked is indicated, and as you can see the cover is estimated to be 12 % tree cover. If you look under each tab and calculate the sum of the values, it should sum up to 100 %.
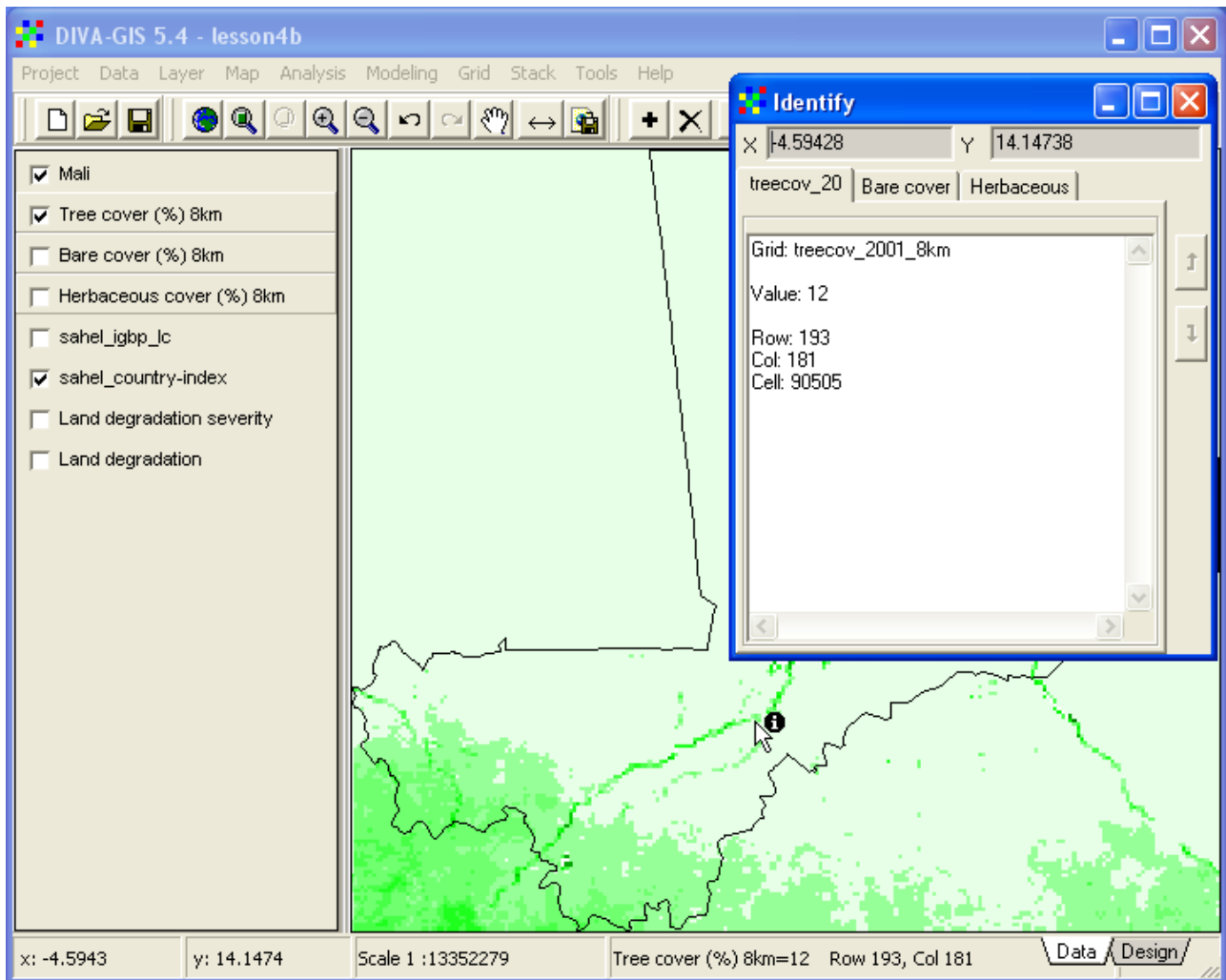

Visual interpretation

Now you should try visually to interpret some of the data on vegetation and land degradation that we have used so far, by relating it to protected areas (national parks, game and forest reserves etc). The project dataset includes international data on protected areas at three different levels. All data is prepared by the International Union for Conservation of Nature (IUCN) and you can find the data layers in the folder \data_spatial\sahel\mapdata\protected. To get to the metadata (the documents describing the data layers), you should open the data web-catalog that is on the project CD.

There are three layers included, each representing a different protection status, listed in the table below.

Data layer	Folder	File
Internationally recognized	\data_spatial\sahel\mapdata\ protected	sahel_international_poly
Stronger nationally protected	\data_spatial\sahel\mapdata\ protected	sahel_national_iucn1to6_poly
Other protected area	\data_spatial\sahel\mapdata\ protected	sahel_national_otheraeras_poly

Add the layers with protected areas and then symbolize them so that you can differentiate them in the Data View. Open the attribute table for each of the three layers and look at the attributes, find out which attribute to use for labeling, and label each of the three themes. In the example below different fonts are used for the different protection levels. You can now zoom and pan in the Data View and visually explore the difference in vegetation coverage inside and outside protected areas. Note that the dataset over protected area is from global datasets and the geometry is not completely accurate.
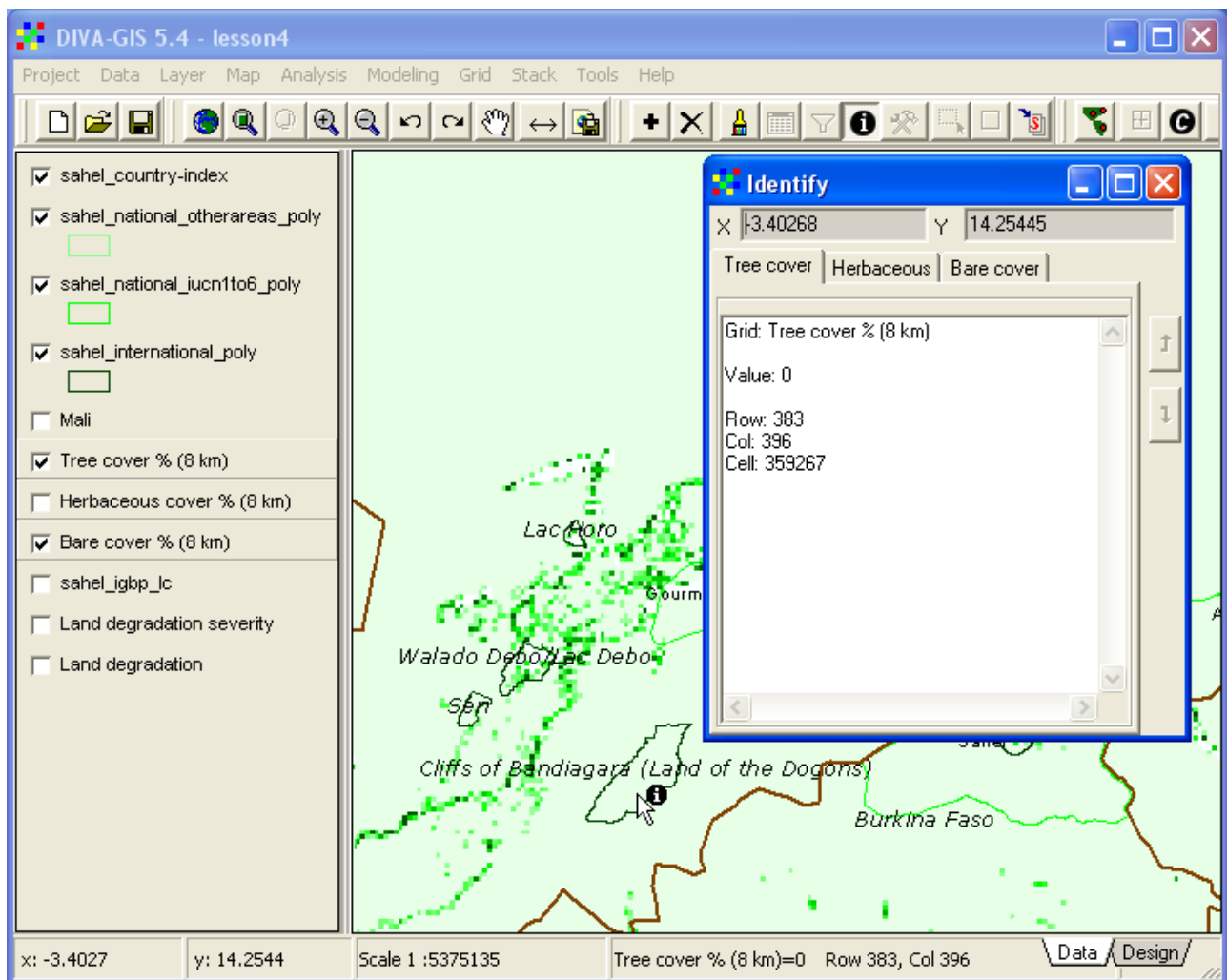
Visualization tools

DIVA-GIS comes with some visualization tools for grid data. You can use them to further explore both spatial and temporal relations in the dataset.

Transect tool

The first tool is a simple transect tool where you can visualize East-West or North-South transects. As these tools are a bit unstable, they can cause the DIVA-GIS program to stop responding, hence it is good to save your project before trying them out. Make one of the vegetation field grid themes the active theme, then use the menu: Grid – Transect, to open the transect window.
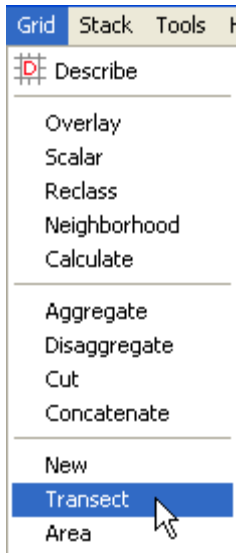
In the Transect window click on the Click on Map button, , and then click in the map. The transect will show the values in the input file (treecov_2001.grd in the example below) from one edge of the grid to the other. The default is to visualize Columns (West to East), but you can change that to Rows (South to North) in the drop down menu to the left in the Transect window. In the Sahel it is more interesting to see the South-North transect as vegetation changes from the Guinea wet savannah (or rain forest) in the South to the Sahara desert in the North. In the example below the check box for Show moving average was checked, and the red line shows the moving average and the green dots the individual cell values. Remember that Nodata values in the vegetation fields are set to 253, and will show up as high peaks in the transects. The transect shown below is for South-North; the first high flat Nodata area is the Atlantic Ocean at the coast in the South, then tree cover gradually decreases, but raises again over the Niger Inland Delta in Mali and then goes to zero in the Sahara Desert.
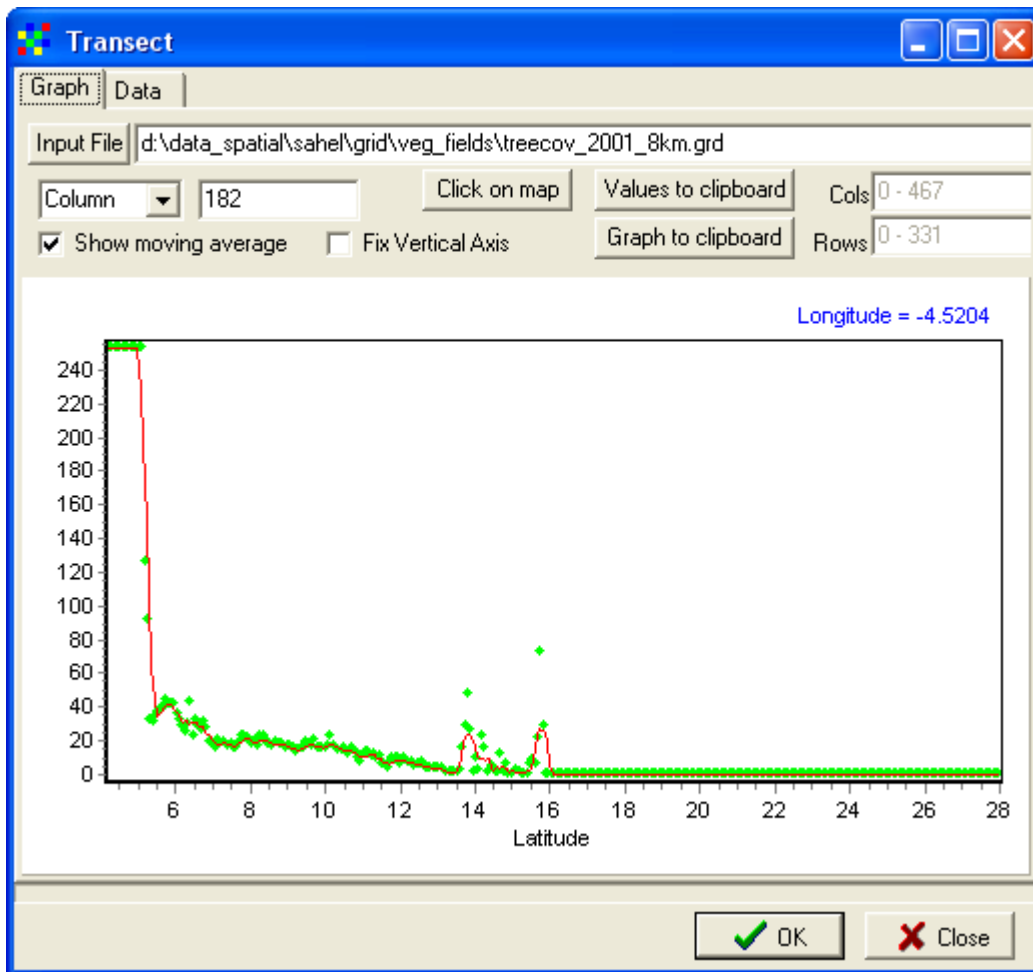
You can copy the generated transect either as values (and paste it into e.g. excel) or as a graph (and paste it into e.g. a word document or a presentation). Just click the Values to clipboard button, or the Graph to clipboard button, then open an excel sheet or word, and use the paste function (or simply click ctrl-v inside excel/word) and the data will be pasted and ready to use. Close the Transect window.

Grid stacks and map algebra

<u>Make a stack</u>

A more advanced tool is to analyze the data value of a point for two or more grid files at the same time. DIVA-GIS has built in function to create stacks of grid files that can then be analyzed together, both graphically and using advanced mathematical and statistical functions. You must first build the stack, and then you can use it for various analysis purposes. You build a stack via the menu: Stack – Make Stack.
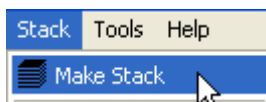

In the Make Stack window that pops up, click the Add Grid button, , and add the three vegetation fields to the stack.
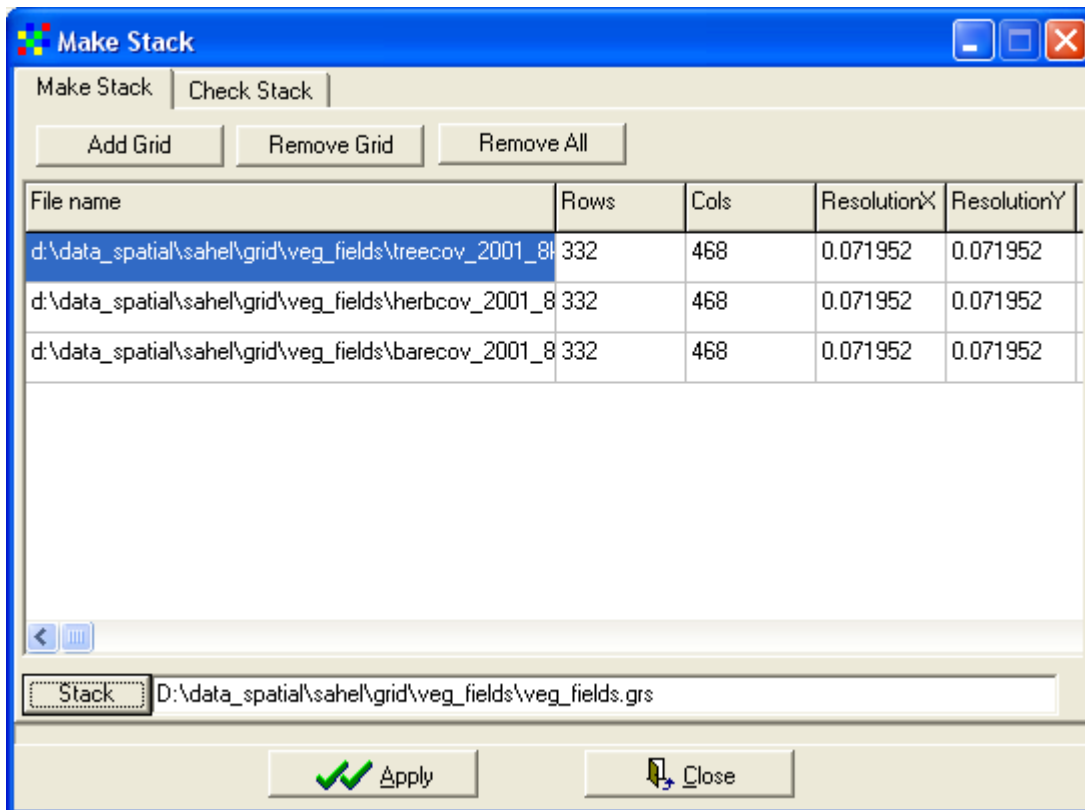
Click the Apply button and then Close the Make Stack window.

Plot stack

Go back to the Stack menu, but this time select the Plot option. The Stack Plot window opens. Click the Stack button, [Stack], and select the stack that you just created. Then click on the Click on Map button, [Click on map], and click in the map. The values for the grid layers in the stack for the click-point will be displayed. You can zoom and pan with the Stack Plot window open, you must just click the Return button, [Return], that replaced the Click on Map button. Then zoom or pan, and click the Click on Map, [Click on map], button again. Also from the Stack Plot you can copy either the data or the graph into the windows built-in copy-paste system and hence take the data directly to other applications.
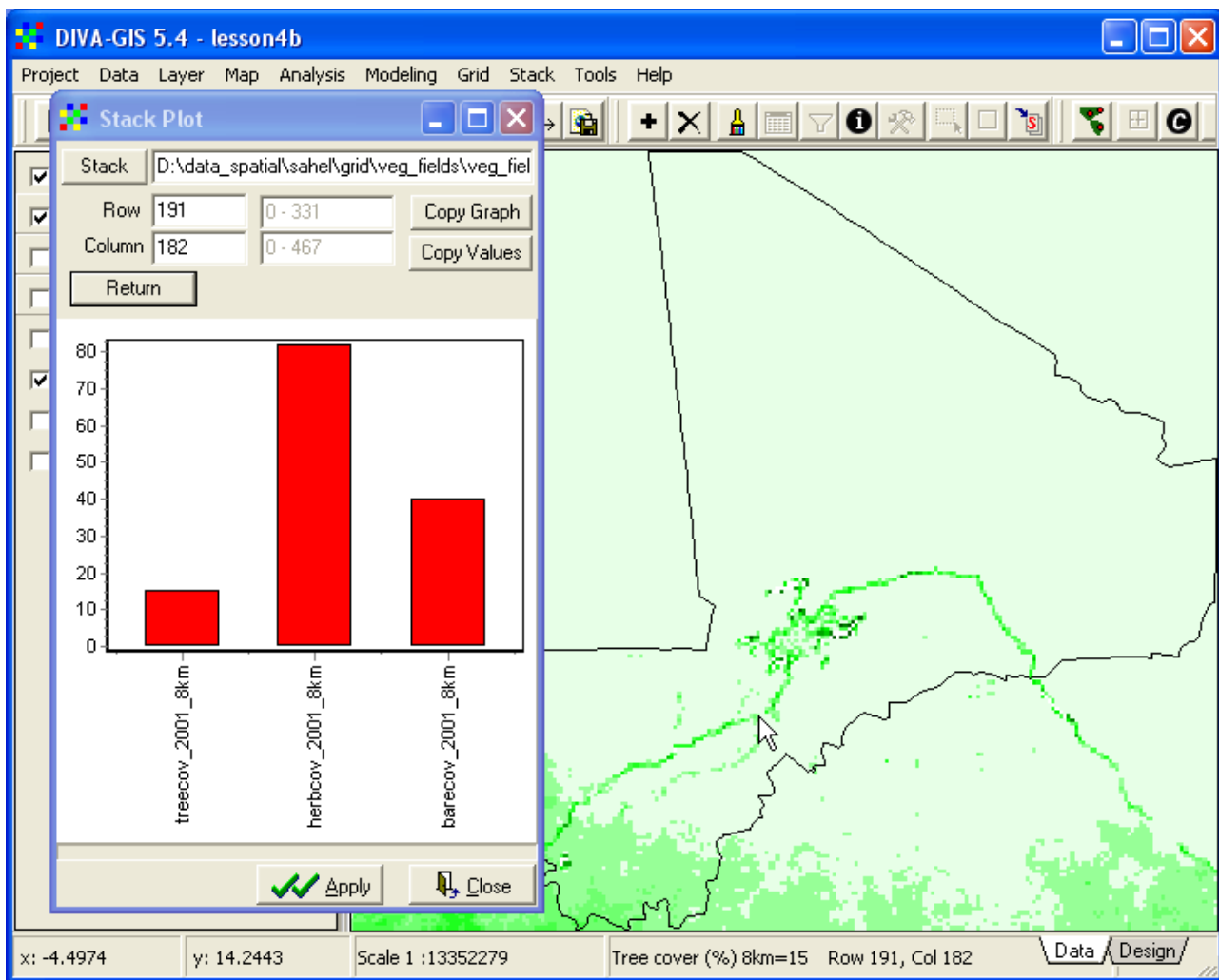
The mathematical and statistical tools you can apply to stacks in DIVA-GIS are rather advanced. But here we will just use one simple option, namely to find out which vegetation cover is dominating (woody, herbaceous or bare) each cell in the grid. But first you shall create a mask that identifies the cells with valid data for vegetation fields.

Reclass

DIVA-GIS contains many useful functions for doing calculations on grid maps (map algebra). You can reach them via the Grid menu. You should use the Reclass function to create a Boolean (0/1) mask for the vegetation field datasets. You want a mask that identifies cells with valid data for the vegetation fields (1 in the mask), and excludes the Nodata cells (0 in the mask). Go to the Reclass window via the menu: Grid – Reclass, and select any of the vegetation fields as Input (all three have Nodata in exactly the same cells). Click the Output button and create a new logical file name (veg_mask.grd in the example below). Then enter the reclass values as shown below. Remember that we are only interested in values from 0 to 100 (the valid data) and hence we reclass values in the range 0-100 to 1, other values will be set to 0 in the second row. Click OK to perform the reclassification. You can save your reclass definition and use it again at later time by using the Save RCL and Read RCL buttons to right in the Reclass window.
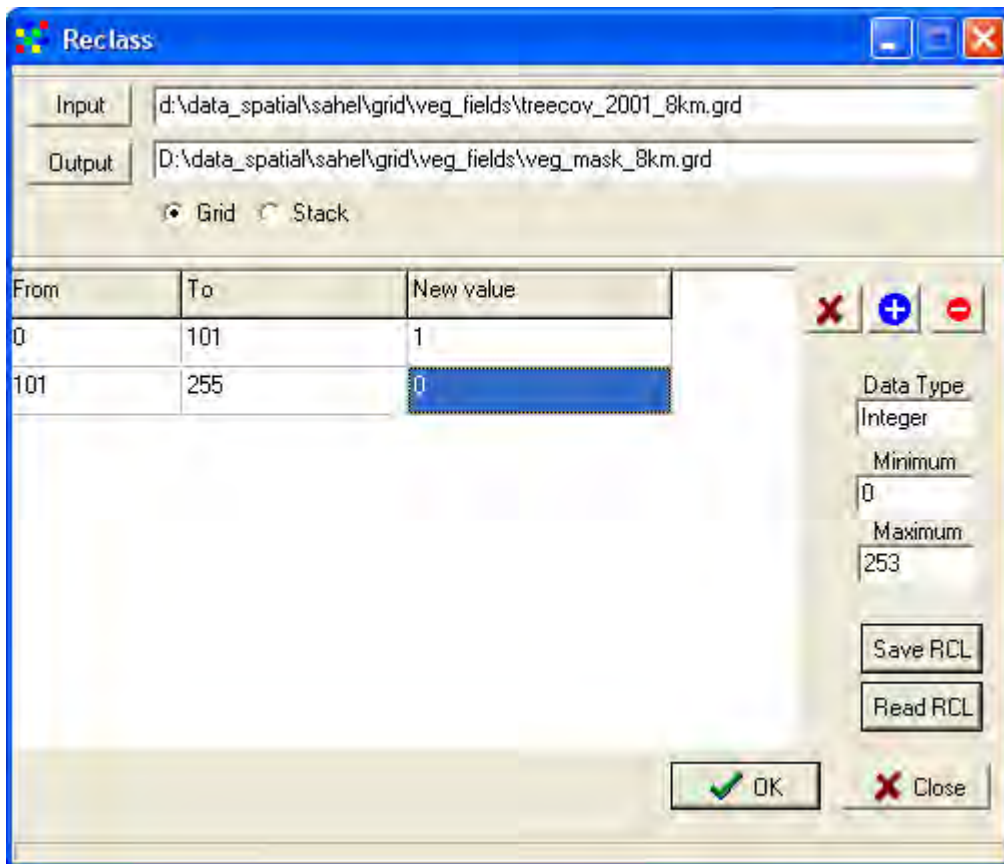
The mask that you created should look like the one below.
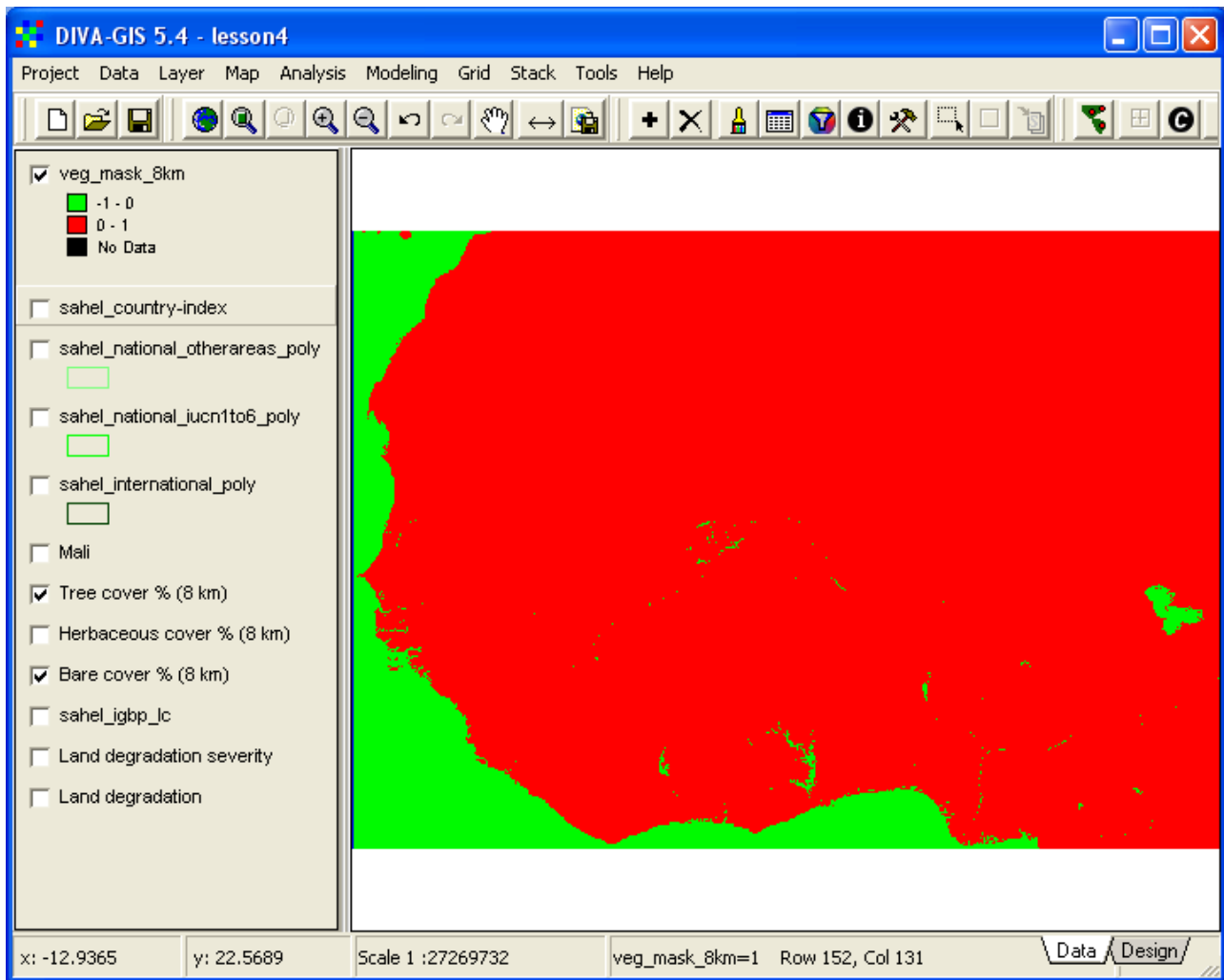
Stack calculations

From the vegetation field data you shall now create a new grid that shows the dominating vegetation field (tree, herbaceous or bare). The stack calculation functions are reached from the menu: Stack – Calculations. Select the stack you created above as the Input Stack. Click the top radio button and select the calculation to be Layer with highest value. The NULL as zero does not work, as we have not defined any NULL value (albeit we know that 253 is NULL, DIVA-GIS does not). Click the Output button and give the new layer to be created a logical name.
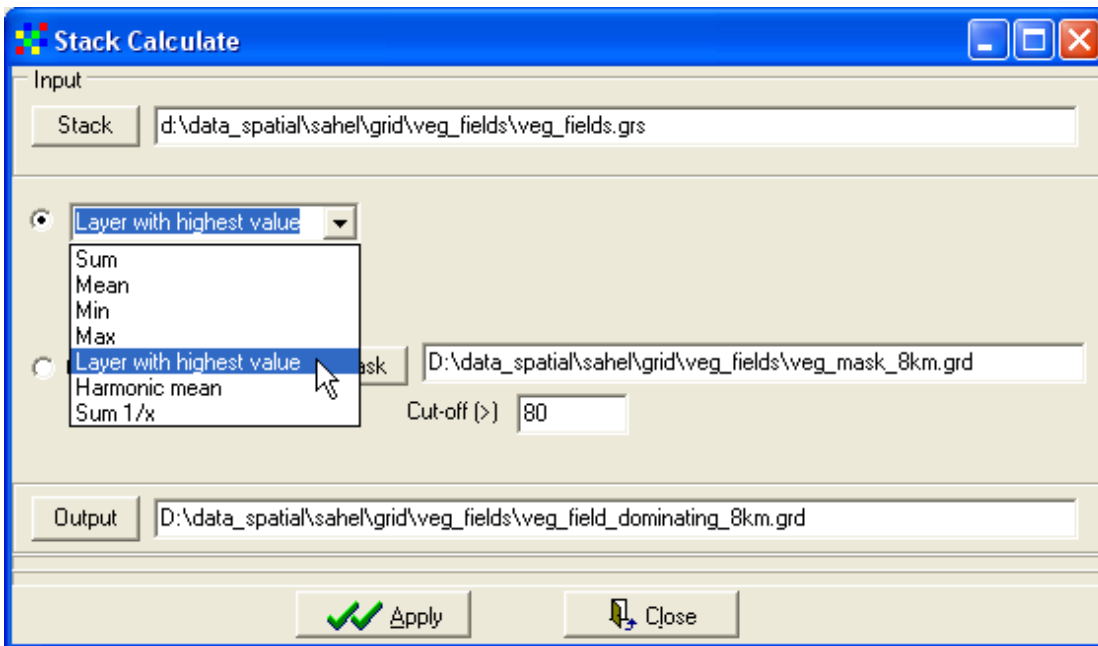
Click Apply to perform the calculation

The output grid is automatically added to your project, but it does not look too good. DIVA-GIS finds the dominating class where there are valid data, but somehow misses the cells were the stack images have equal values (the Nodata areas that have the 253 in all three grids included in the stack). But now we can use the mask we created before and get rid of those areas where there are no valid data.

Overlay

In GIS-jargon it is called overlay when two layers are used in an algebraic calculation. You shall use the overlay function of multiplication to get rid of the Nodata areas in your grid showing the dominating vegetation field. The overlay function is in the menu: Grid – Overlay. Select the First input file to be your dominating vegetation field that you just created, and the second to be your mask, then click the radio button for Multiply and give a logical name for the Output file, Click Apply to start the Overlay calculation.
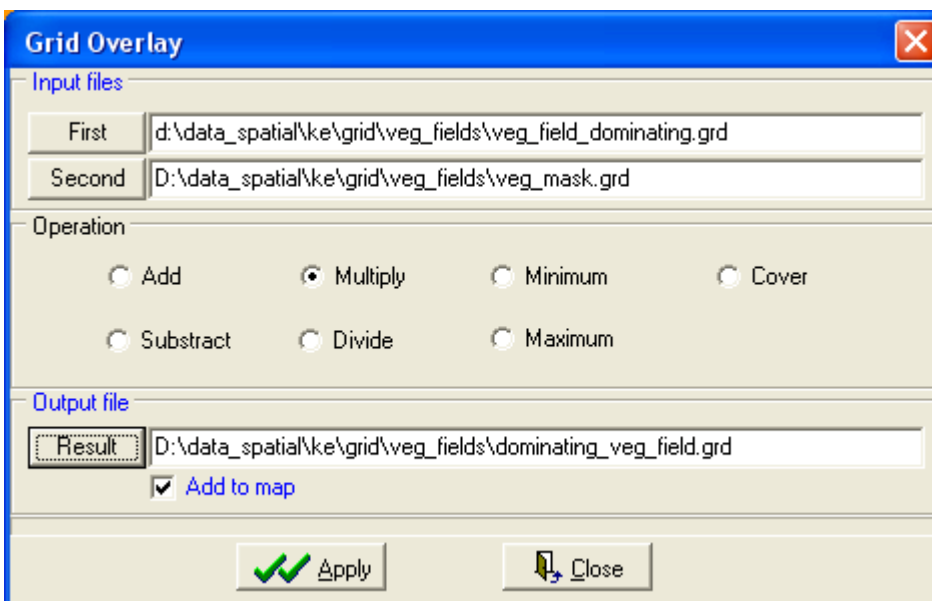
The grid resulting from the Overlay calculation is automatically added to the project, and now the map looks better. In order to check out which of the colors (classes) represent the different vegetation fields, make all the original vegetation field layers active (see above). With the resulting grid file on top in the Legend use the Identify tool to understand the classes in the result file. When you know which class represent which dominating vegetation field, you must symbolize the grid file showing dominating vegetation fields, e.g. as in the example below (the grid to symbolize has nominal values).
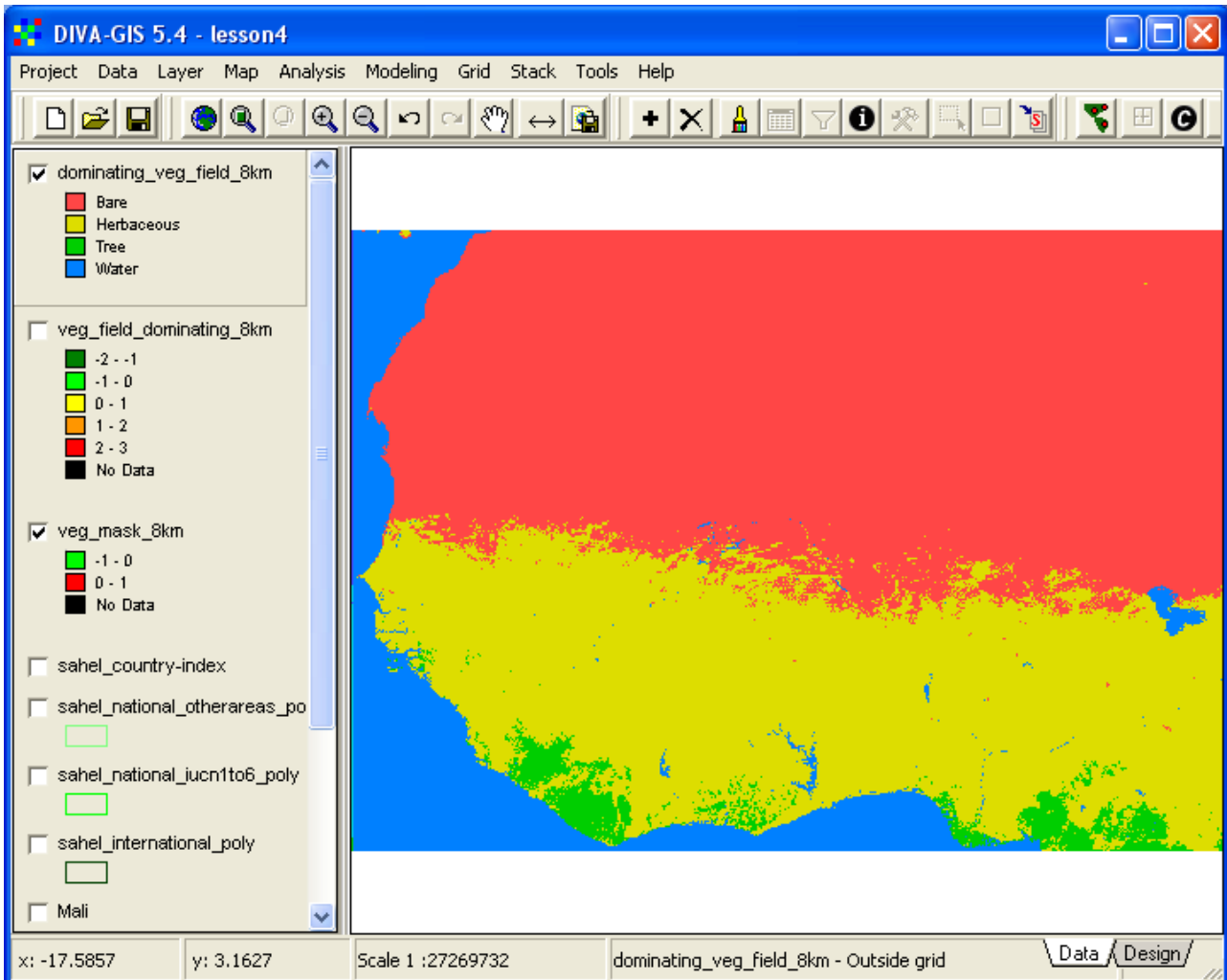

Land degradation and land cover

With the datasets we now have created and symbolized, we should be able to find out if there is any relation between the estimated land degradation from GLASOD, and land cover. Unless the GLASOD map is already in your DIVA-GIS project, use the Add Layer to add the original GLASOD map (\data_spatial\sahel\mapdata\landstatussahel_glasod_geo.shp).

To interpret the attribute data for sahel_glasod_geo you can go back to lesson 1, or read the document describing GLASOD, included in the web-catalog on the project CD, with a direct link here.

Below you see the GLASOD map symbolized using a diagonal cross fill style without borders in order for the underlying maps to be seen through. Can you find any relation between vegetation

cover and land degradation, or between the dominating vegetation field and land degradation? It is perhaps easier to see if you toggle the GLASOD layer on and off.
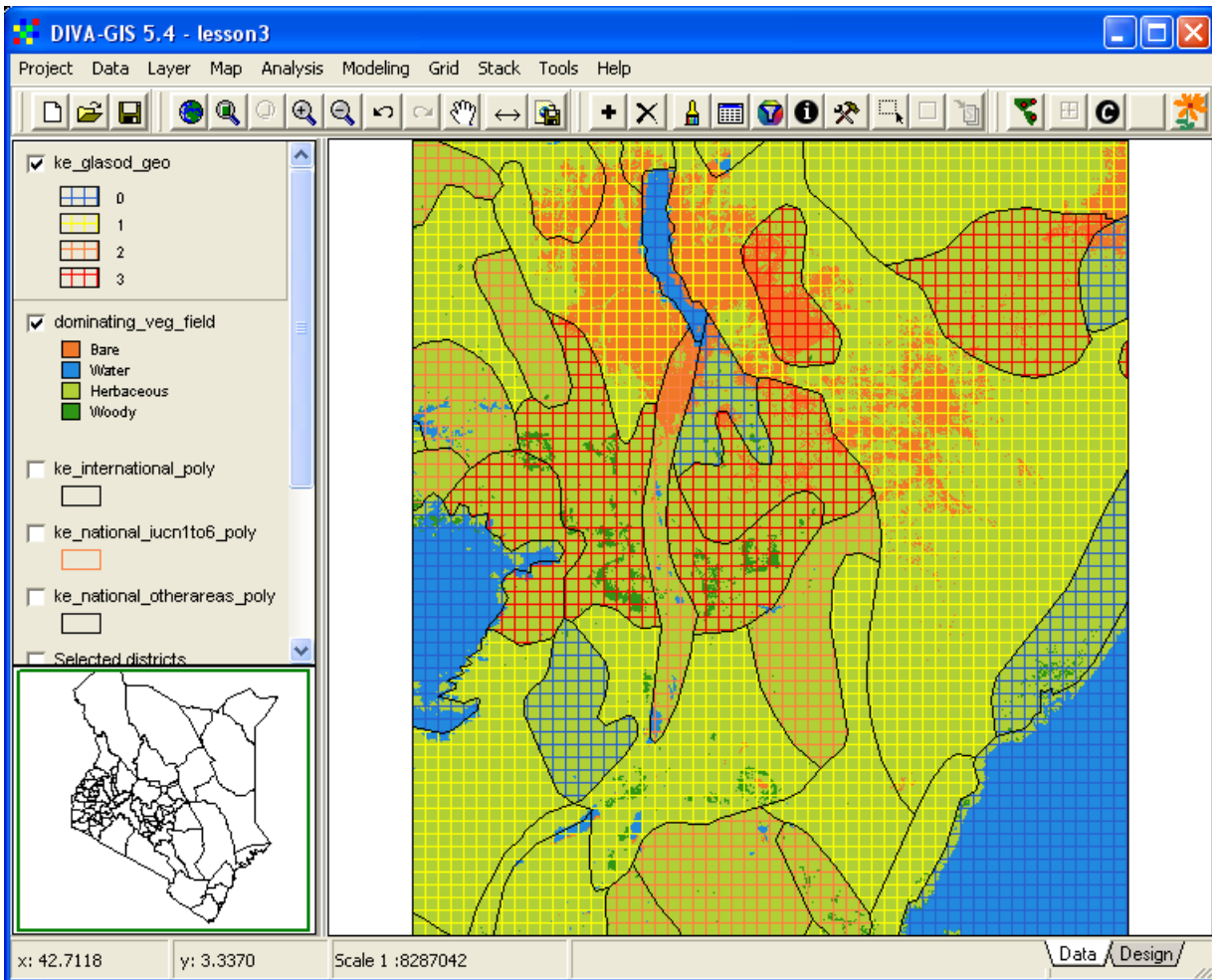

It seems that you have to do more work to track down the causes of land degradation.

Try it on your own - population data

One of the discussed causes of land degradation, especially during the droughts in the 1970s and 1980s is population growth and grazing pressure. The dataset contains gridded data on population covering the last 50 years (1960, 1970, 1980, 1990 and 2000). All the population data is in the folder \data_spatial\sahel\grid\population. The data in the population layers represent population density given as number of people / square kilometer.

Add the population layers to your project (afpopdxx, where xx is the decade – the layer af_pop.grd contains more recent data with higher accuracy and resolution and can not be used together with the other layers).
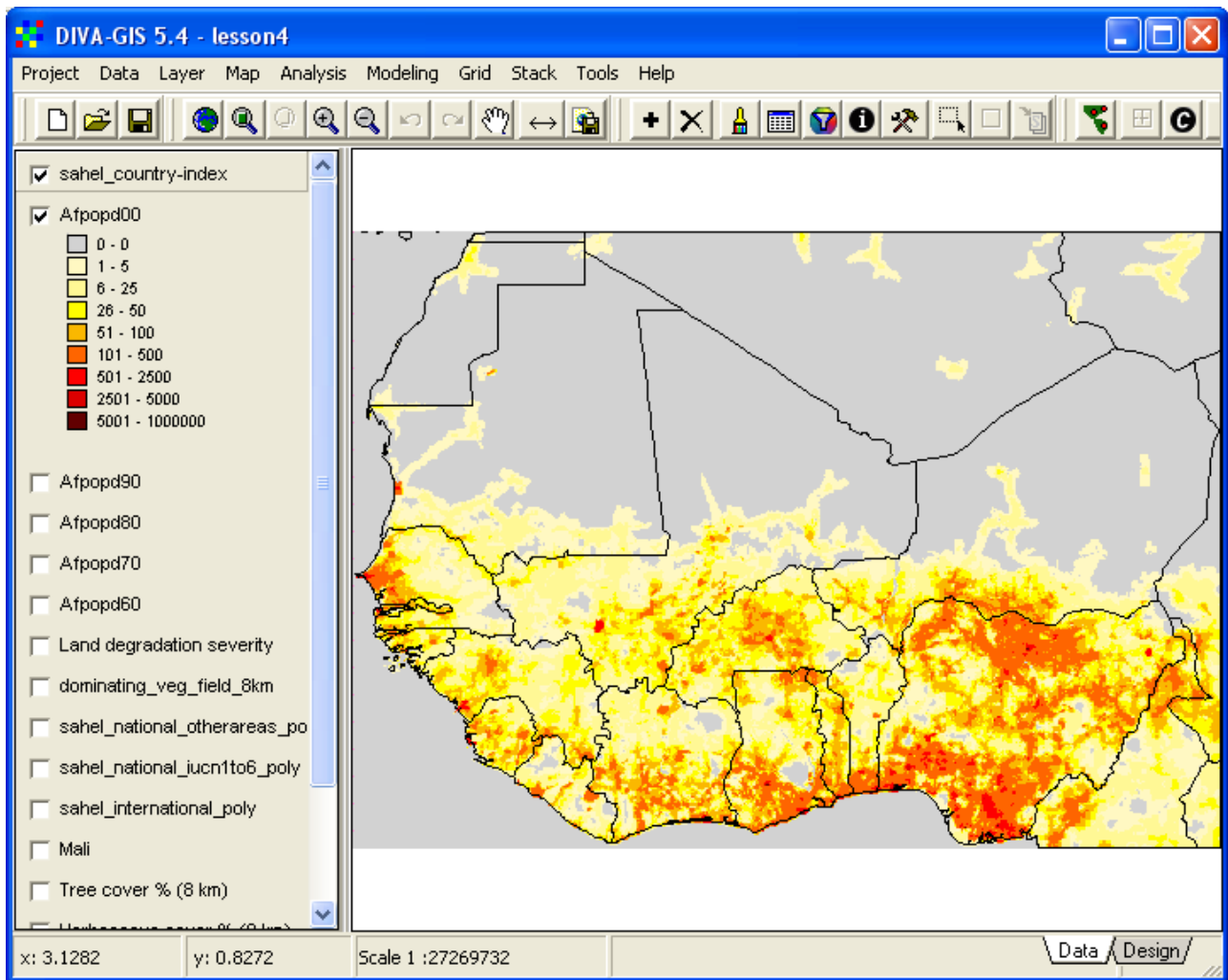
Build a stack including the five population layers. Make sure the data is entered in the right sequence with the oldest population data coming first and the most recent last in the stack, as shown below.
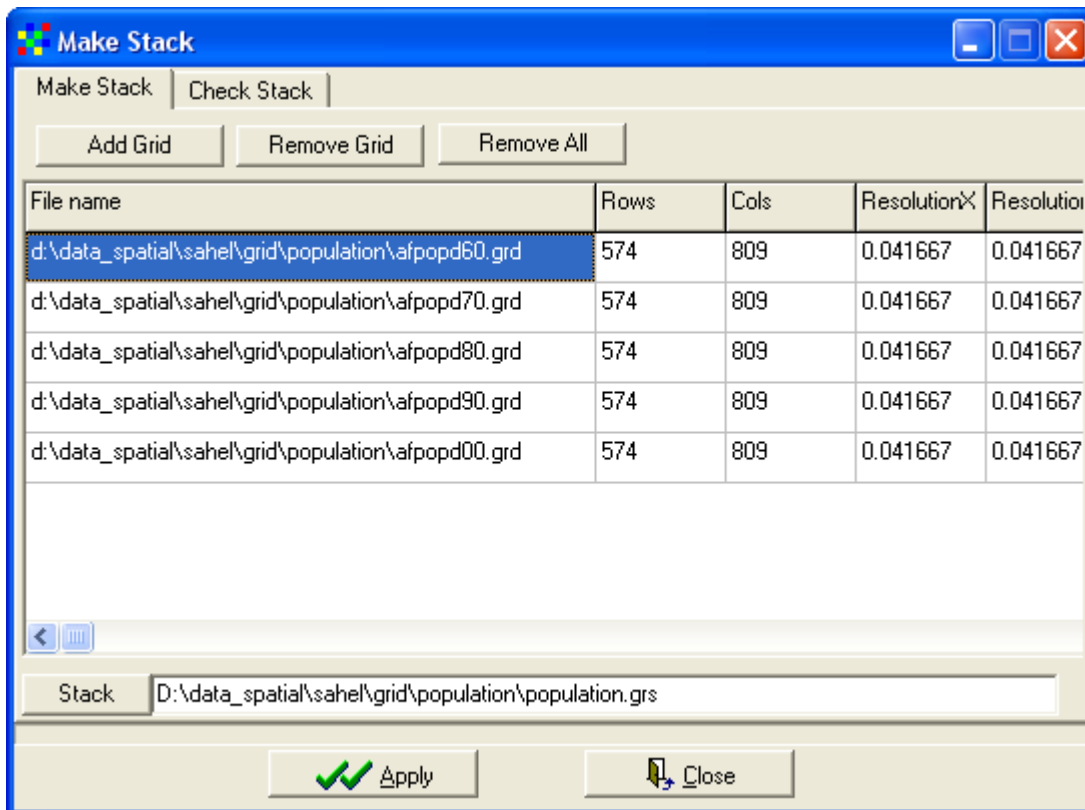
Use the stack to probe the population change at any point in the map. You can also do a transect to study the spatial changes in population density from South to North or East to West. As you will discover, population has grown very steeply over the last 50 years in most of West Africa south of the Sahara.
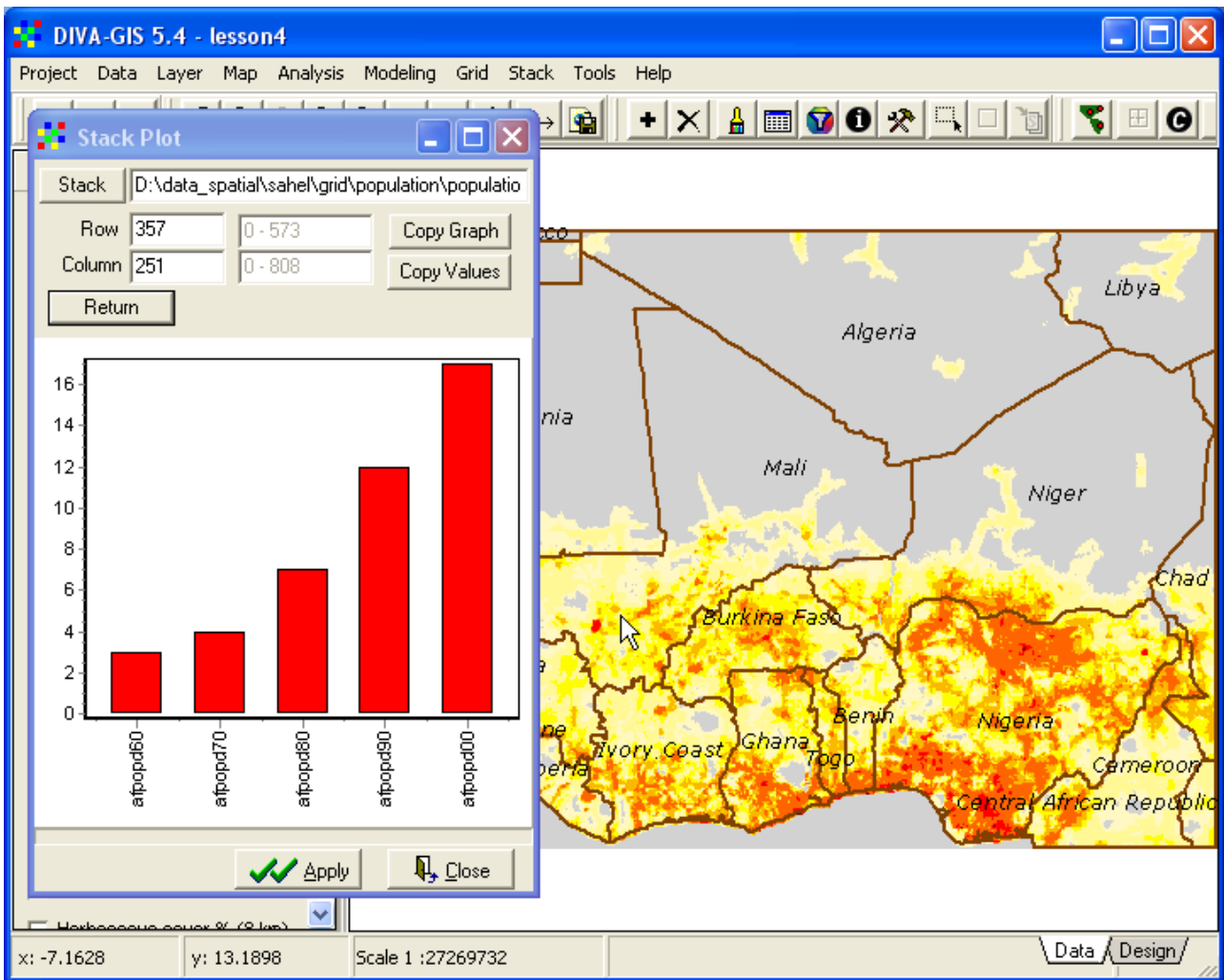
You can also do a more advanced testing of population change by using DIVA-GIS to do a time-series regression. Go via the menu: Stack – regression and use the population stack you just created as Input Stack and give an Output name (note that throughout these lessons we have used rather long and descriptive file names, which is a good custom as you otherwise easily loose track of your data). Click apply to perform the regression analysis.
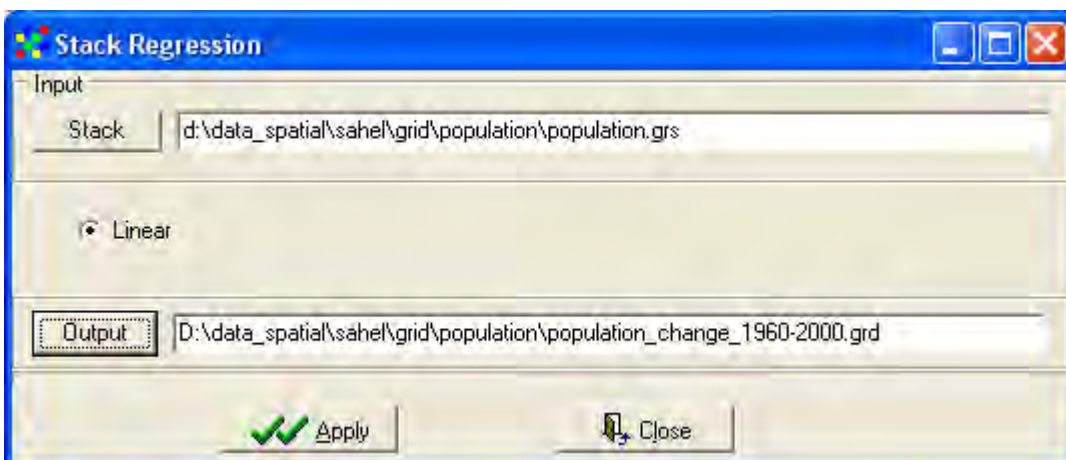

Because of urbanization the regression will yield some extreme values where cities have grown. You hence have to rearrange the results to get a better symbolization. You can either do it by

defining the entries in the legend manually or by creating a mask and putting all extreme values to the same value. In the example below the change was done using the legend. First a legend with 12 entries (rows) was created, with break points manually defined. The legend was then ramped, and the colors for the negative, stable and extremely high (urbanization) growth classes were symbolized manually to stand out better. The final map was then designed in the Design view.
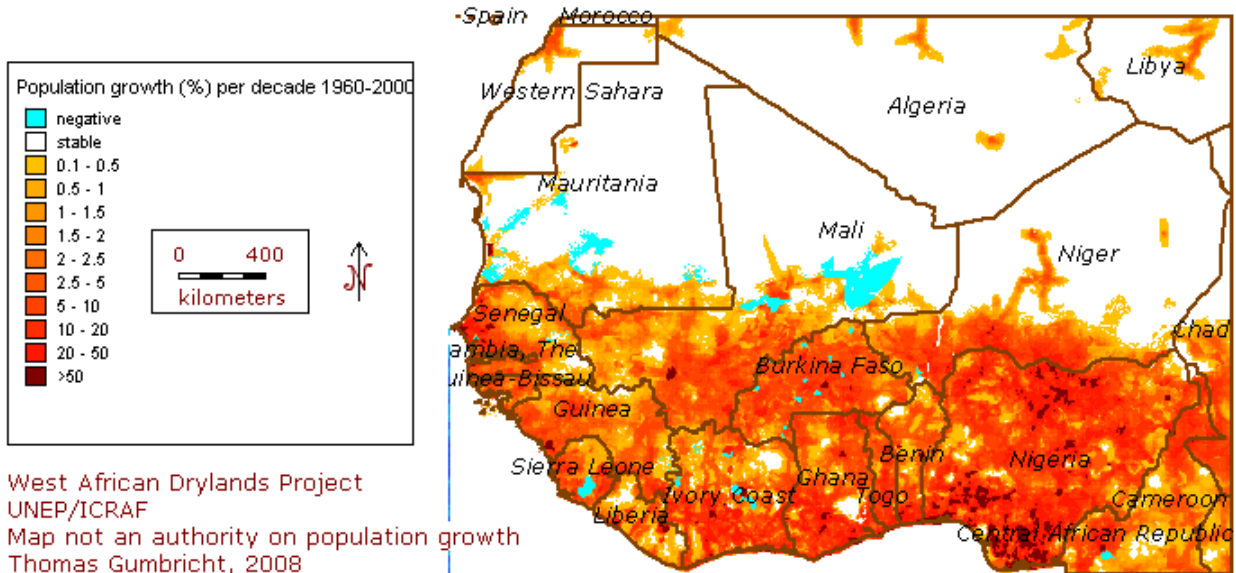

Remember to save the project before finishing the lesson.

A note on grid data dimensions

If you have grid data of different origin and want to use them together in a stack or in map algebra (overlay) they must have exactly the same dimensions. In this lesson you aggregated the vegetation field datasets to have the same number of columns (468) and rows (332) as the satellite derived vegetation and rainfall data that will be used in later lessons. When you aggregated the data, the Output layer that was produced automatically also got new coordinates and resolution – it you open the attribute table for one of the aggregated vegetation field maps you will see this information under the Info tab, as shown below (a small bug in DIVA-GIS causes the cut grid to loose Projection, Map Units and Datum – but below you will learn how to put them back).
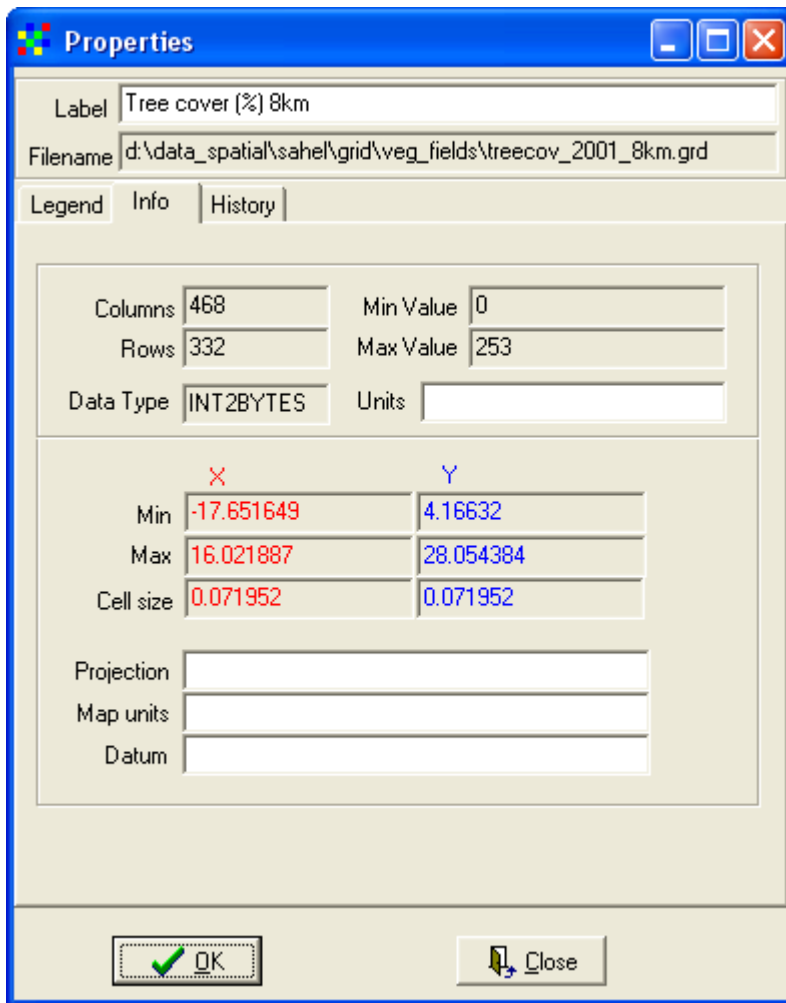
If you add one of the grid layers with satellite derived estimates of rainfall (under the folder \data_spatial\sahel\grid\precip\annual\sum) or vegetation (e.g. \data_spatial\sahel\rs\ndvig\year\average) and open the attribute table, you will see that the columns and rows are equal, but that the coordinates and resolution are slightly different.
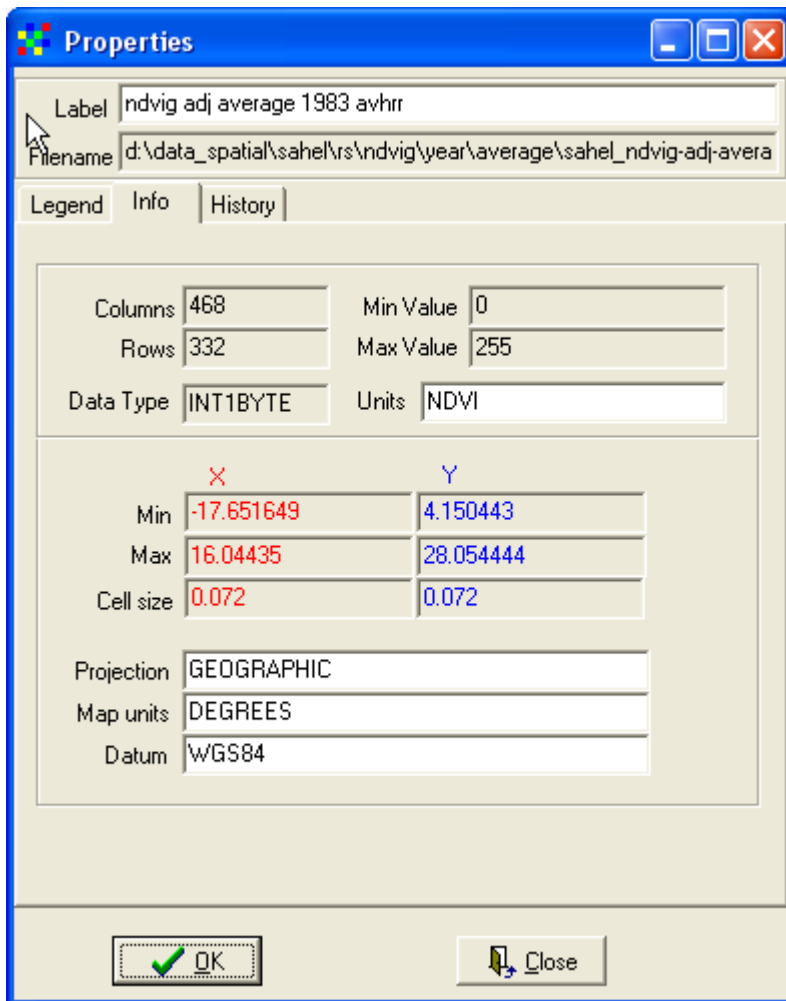
This slight difference in the data layers prevents DIVA-GIS from understanding that the layers represent the same area. To overcome this problem you must use a text editor and change the information manually in the header file (*.grd) of the file you generated from Aggregation. The easiest is to open both the header (*.grd) file for the aggregated layer, and a "correct" header for one of the original files in the dataset, and copy and paste the coordinate and resolution information from the original header to the aggregated header, and then save the aggregated header. The aggregated layer will now work together with the other datasets on the project CD. In the example below the header information from one of the original files have been cut and pasted using Notepad (comes with windows – to find it click Start – Programs – Accessories – Notepad on the Desktop of your computer). Open an original header and mark the section called [Georeference] in the header, press ctrl-c or select Edit-Copy. Then open the aggregated header, mark the same section and press ctrl-v (or Edit – Paste), as shown below. Save the updated header file by pressing ctrl-s (or File – Save).

This manual editing only works if the grid columns and rows are identical.