World Agroforestry Centre
TRANSFORMING LIVES AND LANDSCAPES

Using Farmer Field Schools Approaches to Overcome Land Degradation in Agro-Pastoral Areas of Kenya

Land degradation assessment – Baseline survey on spatial analysis of land cover / degradation trends and Toolkit Development.

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

Last updated: October 2007

LESSON 5 – VEGETATION AND TIME SERIES DATA

Lesson overview

In this lesson you will learn about satellite derived vegetation data, and how to use it for assessing land degradation; how to visually explore changes in both space and time using GIS; how to analyze time series data using DIVA-GIS; and more advanced methods for grid calculations. The lesson also gives an introduction to different data formats for grid data.

Vegetation changes and land degradation

The lack of adequate ground data has led to most studies utilizing the growing archives of satellite data for mapping vegetation changes and land degradation. In this lesson time series data of vegetation for the period 1981 to 2006 is calculated from 8 km resolution Normalized Difference Vegetation Index (NDVI) data obtained from the Advanced Very High Resolution Radiometer (AVHRR) sensor flown on a series of satellites operated by the National Oceanic Atmospheric Administration (NOAA) of the United States of America (USA). The AVHRR sensor was not designed for vegetation mapping, and NDVI is hampered by several shortcomings. To overcome the influence of soil signals at low vegetation densities affecting NDVI a previously published soil adjustment facto was applied. Information about the data source, and where to download, is in the project web-catalog.

The NDVI data was recalculated to annual indexes of vegetation growth: 1) annual average vegetation, 2) annual maximum vegetation and, 3) annual increments in vegetation. As the main dry season in Kenya occurs in September – October the annual indexes were calculated using the period October to September the following year. This was done in order to capture the annual growing cycle rather than the calendar year annual vegetation.

Earlier studies have used either an annual (or growing season) NDVI average for mapping vegetation growth. Studies restricted to cropland have tended to use the annual maximum NDVI, hypothesized to represent the standing crop before harvest. For this project we adopted a theoretically neutral annual vegetation index, defined as the summed annual increment in NDVI between each of the individual observations (i.e. difference in vegetation between current and previous observation if positive). The starting of each annual vegetation cycle was set to September-October. By

summing the increments we hypothesize that our index better captures the productivity of rangelands, and neutralizes the differences between rangelands, croplands and woodlands. It further has the advantage of neutralizing initial background effects (e.g. soil moisture conditions, influence of woody biomass) at the start of an annual growing cycle. The developed index should hence be more suitable when comparing vegetation in landscapes with transient temporal and spatial changes, compared to either annual average or maximum vegetation used in most other studies. Each of the three indexes has its advantage and drawback: annual average vegetation can be misleading as it is sensitive to cloud contamination, and it overestimates vegetation growth in evergreen forests and shrub lands; maximum vegetation is not suitable for pastoral landscapes with continuous or intermittent grazing, where standing crop is low but growth can still be higher compared to fallow or croplands; the annual vegetation increment is sensitive to cloud contamination, and the index has not been rigorously tested.

About grid data and data formats

In DIVA-GIS a grid layer is built of three separate files; a data file (with the extension *.gri), a header file that contains information on rows, columns, resolution, projection, data format and legend (with the extension *.grd) and an image file that is displayed in the Data View (usually a bitmap - *.bmp). Most GIS only use two files for grids, and do not have a separate image file – the GIS generate the display image when a grid is added. Here is a grid header file from DIVA-GIS, with some explanations:

[General]	# General information
Creator=Thomas Gumbricht	#
Created=Sun Oct 28 17:55:33 2007	#
Title=Afpopd90	#
[GeoReference]	# The Coordinate system of the grid
Projection=GEOGRAPHIC	# Either Geographic or Other (i.e. projected)
Datum=WGS84	# How the Earth roundness is set
Mapunits=0.00000000	# The units in Other projections (see below)
Columns=216	# Nr of columns in grid
Rows=240	# Nr of rows in grid
MinX=33.000000	# Coordinate of left side
MaxX=42.000000	# Coordinate of right side
MinY=-5.000000	# Coordinate of bottom
MaxY=5.000000	# Coordinate of top
ResolutionX=0.041667	# Resolution in Grid X
ResolutionY=0.041667	# Resolution in Grid Y
[Data]	#
DataType=INT2BYTES	# Datatype can be byte, integer or real
MinValue=-0.000000	# Minimum data value in grid
MaxValue=10410.000000	# Maximum data value in grid
NoDataValue=-9999	# Value denoting no data
Transparent=1	# if no data is to transparent or not
Units=nr	# Units in map (here = people/km2)
[Application]	#
Opt0= Imported from ….	# DIVA-GIS operations (entered by DIVA)
Hist0=Population data made ….	# Metadata put the producer/user

Hist1=Produce by CIESIN….	# Any number of Hist entries allowed
[ContLegend]	# Default continuous legend (if no [Legend])
Count=2	# Nr of entries in default legend
Color1=16711680	# Color of entry 1 in default legend
Value1=0	# Value of entry in default legend (= min)
Label1=	# Text for entry 1 in default legend
Color2=255	# Color for label 2 in default legend
Value2=10410	# Value for label 2 in default legend (= max)
Label2=	# Color for label 2 in default legend
	# [ContLegend] in this header is ignored
[Legend]	# If given overrides the [ContLegend]
Count=9	# Nr of entries in the Legend
Color1=13816530	# Color of entry 1 in Legend
Min1=0	# Min value for entry 1 in Legend
Max1=0	# Max value for entry 1 in Legend
Label1=0 - 0	# Label for entry 1 in Legend
.	.
.	.
.	.
.	.
Color9=65634	# Color of entry 9 in Legend
Min9=5001	# Min value for entry 9 in Legend
Max9=1000000	# Max value for entry 9 in Legend
Label9=5001 - 1000000	# Label for entry 9 in Legend
Transparent=1	# If nodata is transparent (1) or not (0)
NoDataColor=16777210	# No data color (shown if Transparent = 0)
NoDataLabel=No Data	# Label for no data
isContinuous=0	# If legend is continuous or has gaps
SpacedByColor=0	# If Legend should be spaced by color

The georeference of a GIS data layer can either be Geographic (unprojected) or Projected (called Other in DIVA-GIS). A Geographic georeference is given in Longitude for X and Latitude for Y. DIVA-GIS can reproject layers from Geographic to many different Projections, but not vice versa. A Geographic georefence is defaulted to have Mapunits in decimal degrees (e.g. 33.000000 and not in degrees-minutes-seconds 33 0 0). For Projected layers Mapunits can be kilometers, meters, millimeters, miles, feet etc. The Datum describes the spherical shape of the Earth, as the Earth is not completely round but is flattened at the poles and thicker around the equator. WGS94 (World Geodetic System 1984) is the default for Geographic georeference. If you reproject data to a Projection, another Datum can be set.

The DataType in GIS grid layers can have many formats, but is normally restricted to three formats:

BYTE (INT1BYTE)   0 .. 255

INTEGER (INT2BYTES)   −32768 ..32767

REAL (FLT4BYTES) 1.5×10^{-45} .. 3.4×10^{38}

 In a Byte image each cell uses one byte (8 bits) for storing each cell value, and the data range can be a whole number between 0 and 255 (called unsigned byte by computer specialists as only positive numbers can be stored). In an Integer image each

cell uses 2 bytes (16 bits) for storing each cell value, and the data range can be a whole number between –32768 and 32767 (signed small integer for computer specialists, as the data can have both positive and negative values). An image with Real values uses 4 bytes (32 bits) for storing each cell value, and the data range can be a ratio number (32 bit real is called Single by computer specialists). The trend images that you generated in lesson 4 contain Real data (FLT4BYTES in the DIVA-GIS header).

The image files for vegetation (and rainfall and rain use efficiency that you will use later), where generated outside DIVA-GIS to make a more full use of color ramping. If you make changes to the legends in these files, DIVA-GIS will regenerate the image file and the pre-made color ramp in the images will disappear, You will not be able to recreate the symbolization using DIVA, so if you change the symbolization, then you can make use of the image files to get the original symbolization back by using an image program (photoshop, gimp etc) to convert the jpg image to a bmp.

Visual exploration of vegetation

Start a new project

Start DIVA-GIS and open a new project by clicking the New Project button, [image], or via the menu: Project - New. Save the new project, e.g. calling it lesson5.

Add framework layers

In order to get some reference and background data to use for interpreting vegetation trends, start by adding the following framework (basic) data layers

Data layer	Folder	File
Districts	\data_spatial\ke\mapdata\politic	ke_kenya-districts
Rivers	\data_spatial\ke\mapdata\hydro	ke_af_rivers_detailed
Basins	\data_spatial\ke\mapdata\hydro	ke_af_basins_detailed
Climate	\data_spatial\ke\mapdata\climate	ke_climate_monthly_average

Symbolize the layers, and save the project.

Add vegetation grid data

For the analysis of vegetation and rainfall changes covering the whole of Kenya, vegetation data from the AVHRR (Advanced Very High Resolution Radiometer) was used (see the project report on the project CD for details). The AVHRR data is recorded daily, and then composed into 10-day values for vegetation using the maximum recorded values. The maximum value is preferred as clouds and atmospheric pollution tends to decrease the value of NDVI. All the vegetation data for Kenya as whole is under \data_spatial\rs\ndvig. Here we will only use annual indexes, and not the original 10-day datasets. The annual indexes for Kenya were calculated using 12 months starting in October and ending in September. This was done to better depict the annual growing cycle, which has its minimum in September-October in East Africa. As stated above there different annual indexes were calculated, and they are all included on the project CD. Start by adding the indexes for annual average NDVI in the folder \data_spatial\ke\rs\ndvig\year_oct-sep\average. If you are going to add several grid files at the same time, it is better to uncheck the option for Layer is visible when added, in the Option window (found via the menu: Tools – Options; see lesson 3 to set the Options).

Only add the grid (*,grd) files, do not load any image (*.jpg) file. You do not need them, as the symbolization of the grid files is identical to the image files.

Add at least three images showing the annual average vegetation going back to 1981. Use one image from the early 1980's, one from the 1990's and one recent image. In the example below, the images for 1983/84 (very dry), 1994/95 (very wet), and 2004-2005 (normal), were used.

Put the added vegetation images at the bottom of the Legend, and turn on the layer showing the districts of Kenya, then sequentially turn on each vegetation image. The years used in the tutorial show a remarkable difference in average vegetation cover.
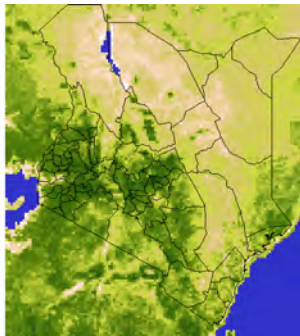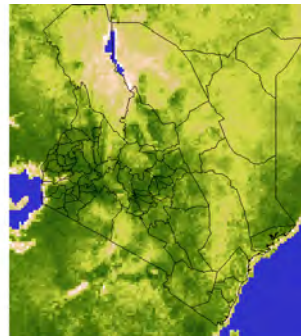
| 1983/84 | 1994/95 | 2004/05 |

The three small map-images above were generated by clicking the Map to Image tool, in DIVA-GIS, using the Clipboard as Target, and then pasted into this document (by clicking Ctrl-v), without being saved to any intermediate image file.
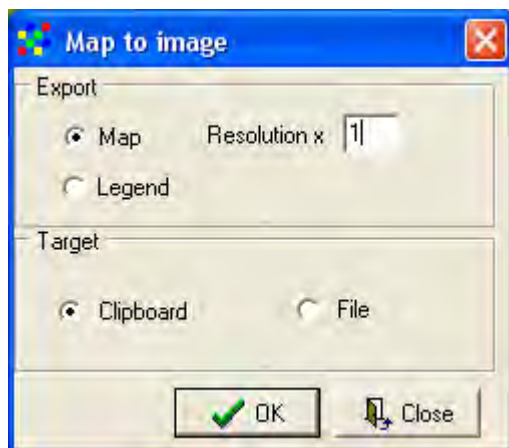
Vegetation and drainage

The framework dataset includes rivers and river basins, and you need to symbolize these to be able to visually interpret if they have any effect on the vegetation pattern. The river basins really only need to be symbolized in order to separate them from the district boundaries. In the example below the width of the basin boundaries was set to 2, and the color to magenta. From a visual interpretation of the image it seems that some basins act as dividers for vegetation growth, but that needs to be further investigated, including using topography. Topographic data of different origins are included on the project CD (see the web-catalog), but will not be further analyzed in this lesson.
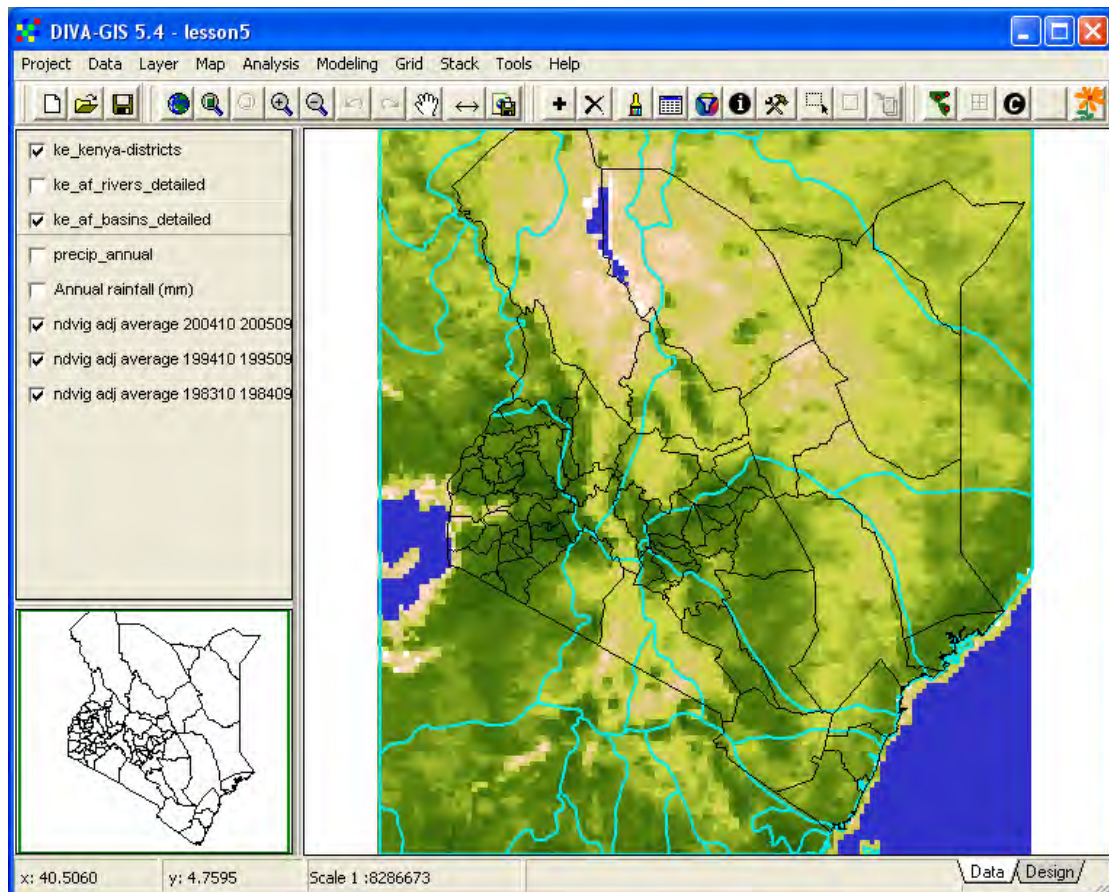
Symbolize the river theme, and use visual interpretation to see if there is any relation between the stream network and the vegetation pattern. The river layer contains data on stream order, where a higher value denotes that more streams have fed into that particular stream segment. To generate stream order, the river data must have a correct topology (see lesson 2), that is to say that the streams must be connected to each other with no gaps, and a new stream segment at each flow join or bifurcation (splitting of one stream to two). Stream order can be seen as ordinal data (click here to get to lesson 2 explaining ratio), and as this ordinal data is numerical you can actually use the classification option to symbolize the streams into small and large (as shown below).
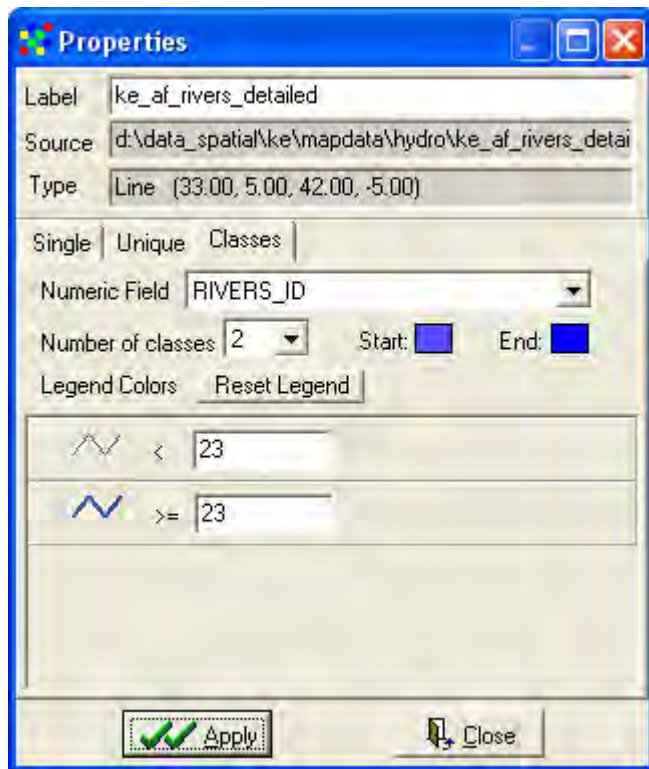
If you look carefully at the resulting map you can see that it is more green around some of the larger rivers. And that at the origin of some streams it is less vegetation, but the latter phenomena is restricted to Mt Kenya and Mt Kilimanjaro, where rainfall is both higher (generating many streams) but at the same time devoid of vegetation (due to glaciers and high elevation). The dense vegetation around these mountains stand out clearly around the peaks with no or little vegetation even in the low resolution data derived from the AVHRR sensor.
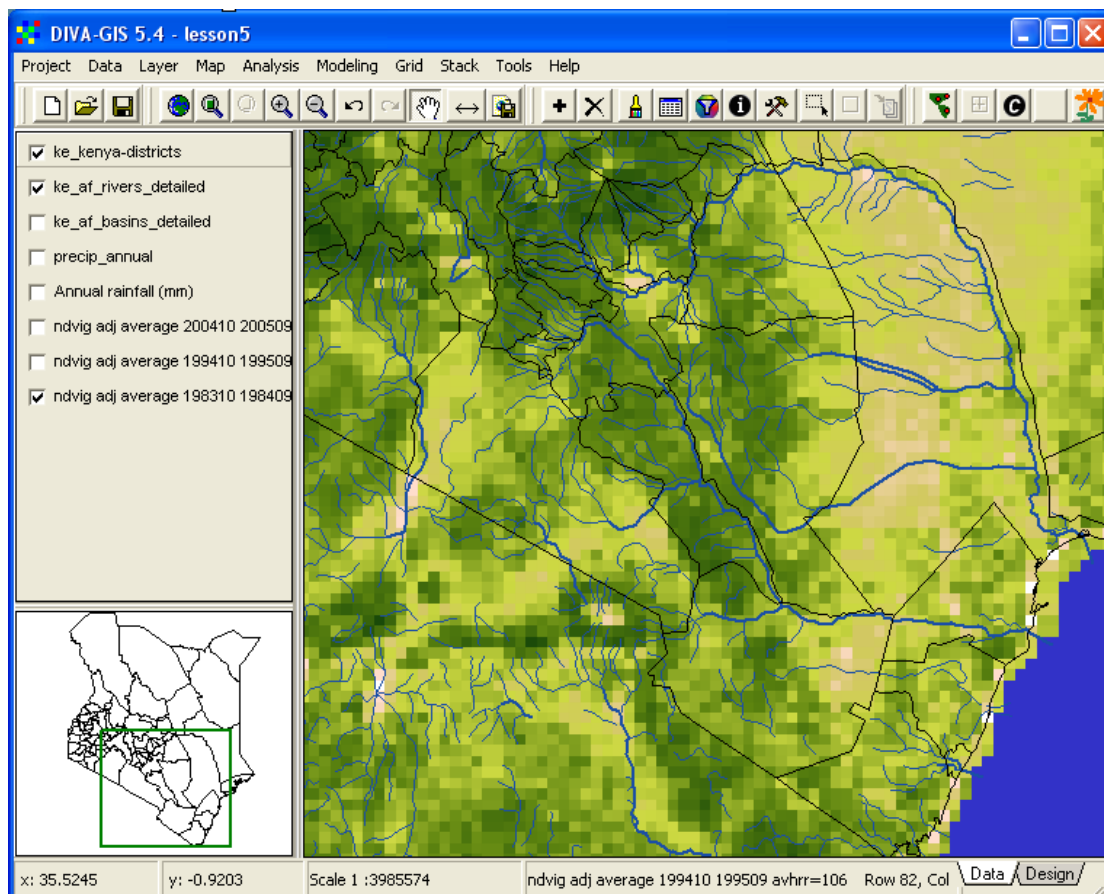
Vegetation and rainfall

The framework dataset includes a vector layer with climate long term average data (ke_climate_monthly_average). You can use that layer to see the general rainfall pattern over Kenya. Open the Properties window for this layer, and symbolize it using the attribute field PANN (precipitation annual), as shown below (precipitation is an attribute of the type ratio, and symbolization should be done using classes – click here to get to lesson 2 explaining ratio data and other attribute data types). Change the label of the theme to Annual rainfall (mm) (or other relevant label), the label will then be changed in the Legend of the Data View.
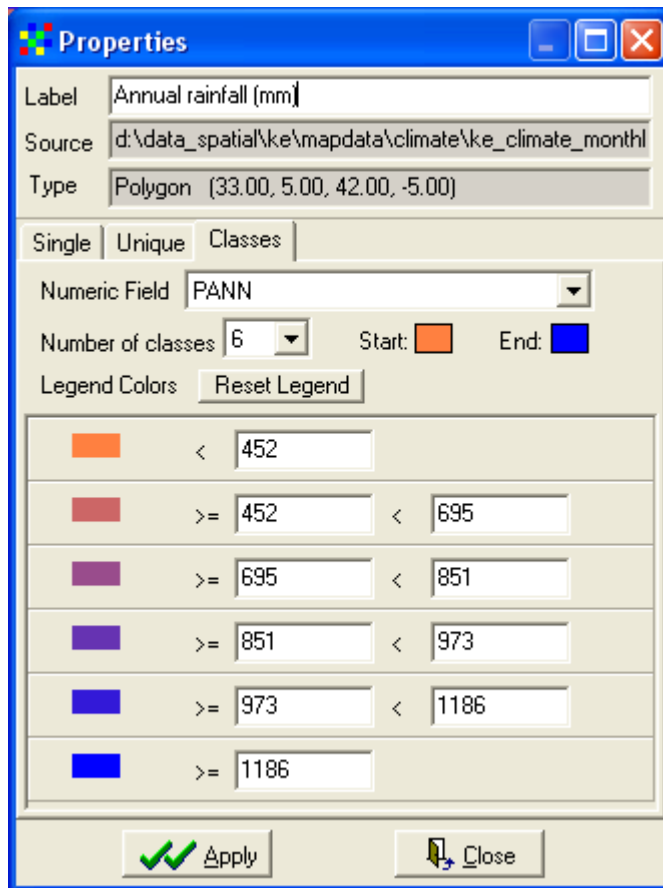
The overall vegetation pattern of Kenya is clearly related to rainfall pattern.
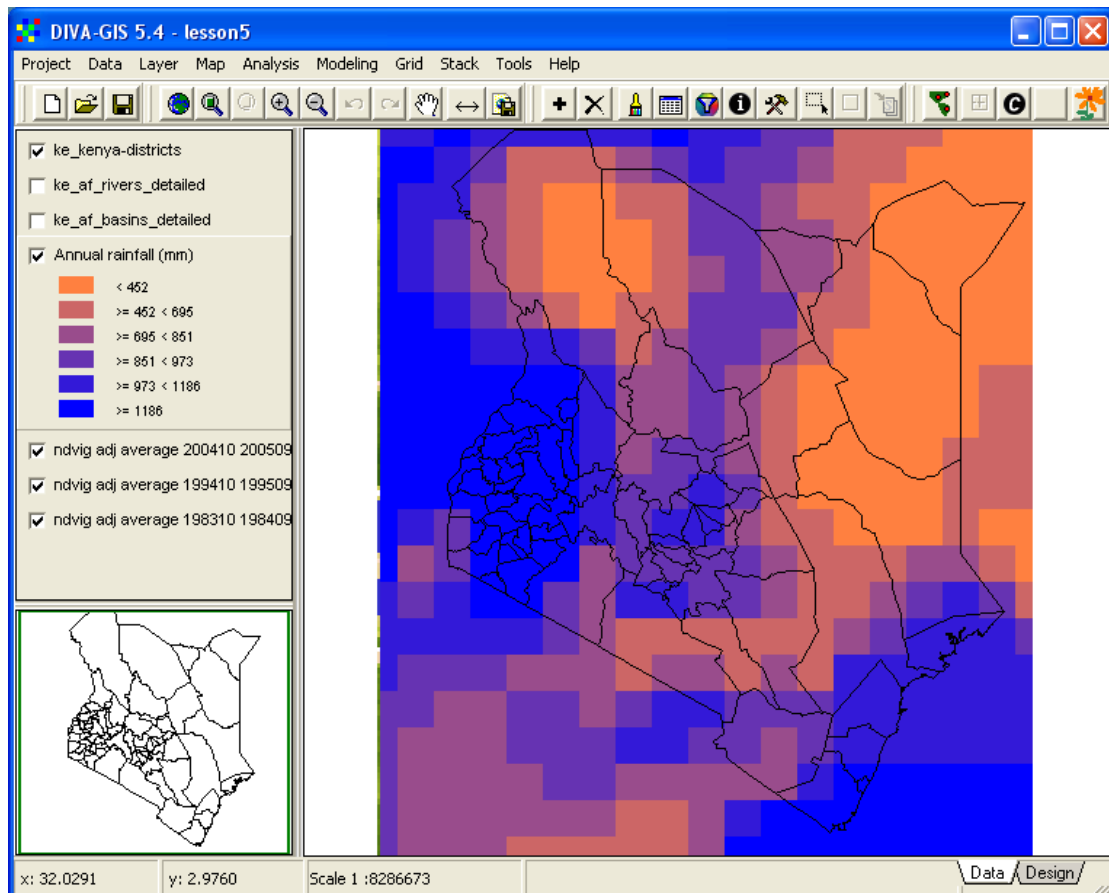
Vector to grid conversion

The climate statistical layer contains monthly data not only for precipitation, but also for temperature and soil water, and can be used for making further analysis. If you want to use the climate statistical data layer for calculations, DIVA-GIS can do that if you first convert the polygons to a grid, via the menu: Data – Polygon to Grid. In the Polygon to Grid window you can set the parameters of the output grid to exactly fit the vegetation data. As the grid layers only contain one attribute, you need to select each attribute of the vector to generate a new grid file (that is why vector files consume less storage space in the memory). For now it is enough if you generate one grid file for the annual average rainfall pattern.
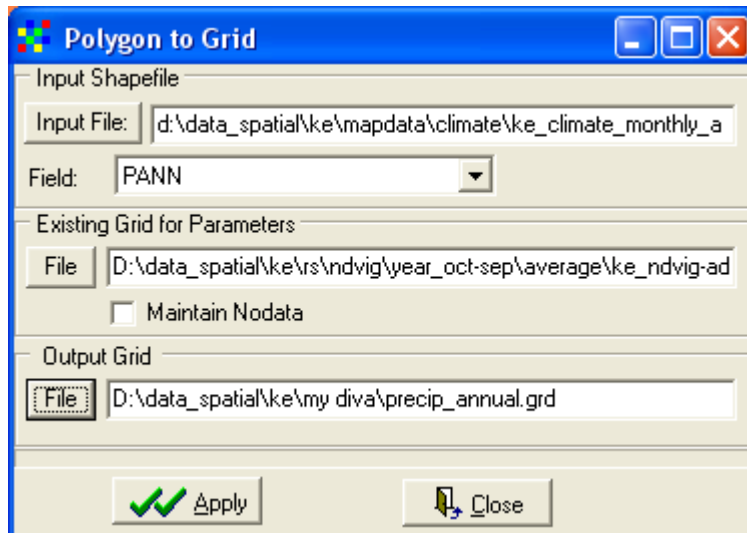

The image below shows the grid layer for annual rainfall generated by DIVA-GIS.
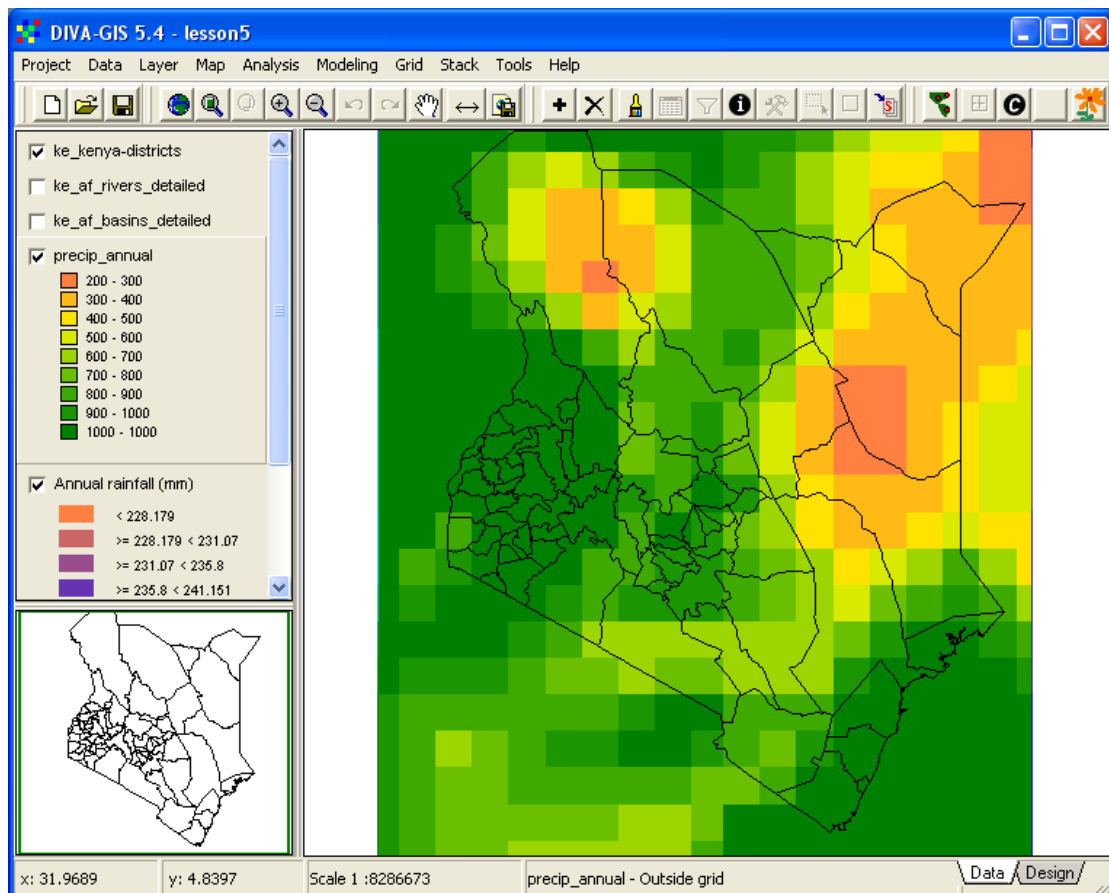
Time series analysis

<u>Vegetation trends</u>

For mapping out land degradation it is more interesting to analyze the trends in vegetation than the vegetation cover a particular year. In <u>lesson 4</u> you learnt to build a stack and plot data from a selected position including values from all layers in the Stack. To create a stack with many layers is a bit tedious, and the project CD hence contains ready-made stacks for all time series vegetation data. You just need to open the stack for plotting via the menu: Stack – Plot. The stack for annual average NDVI is in the same folder as the grid data, click the Stack button and navigate to open it (the folder only contains one stack). You can make the Stack Plot window larger by grabbing the blue border with the left mouse button, and dragging the mouse. Click the Click on Map button, Click on map, and then just click in the map and you will get the vegetation development of the click-point over the last 25 years. The plot below shows vegetation development in the central (drier) parts of Narok District 1981-2006.
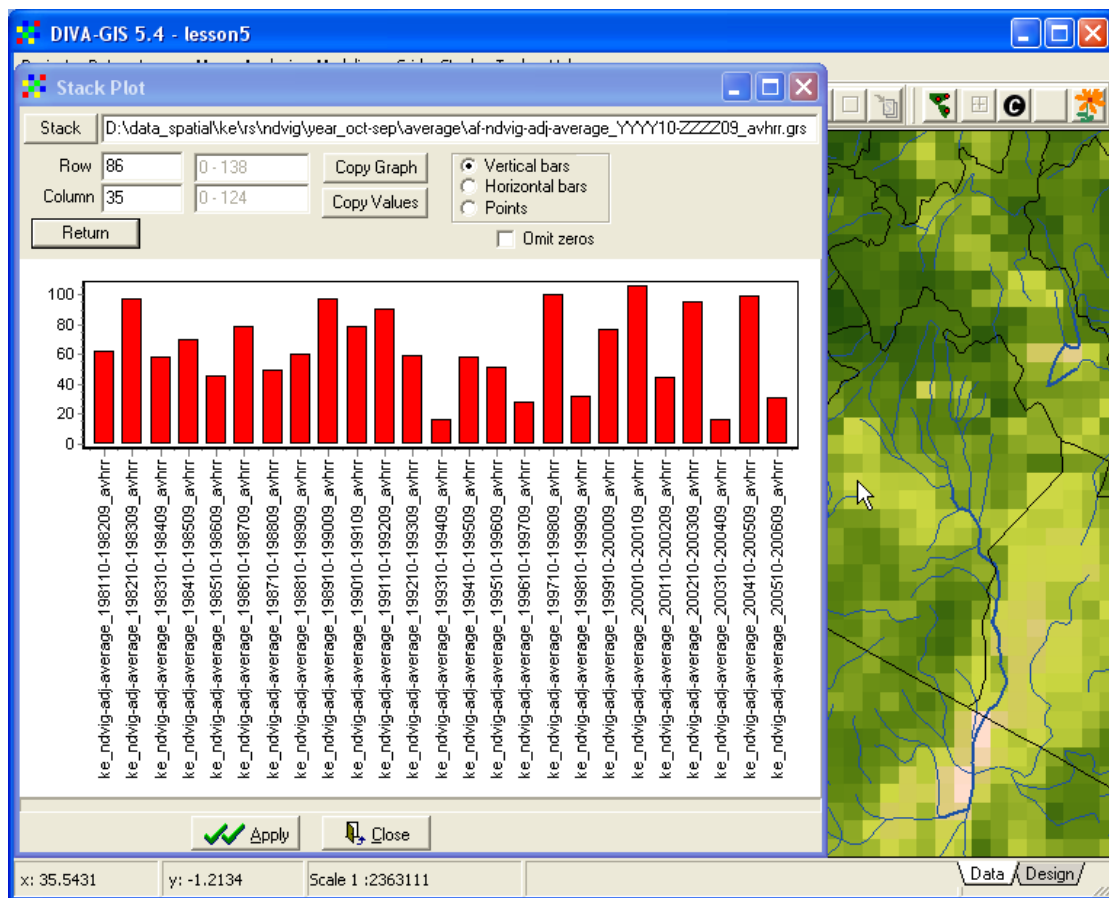

Instead of clicking in the map, you can also give values for which row and column to plot. After entering values in the text fields, just click Apply for DIVA-GIS to analyze you selected point. You can again copy the stack plot graph or the values behind it to the clipboard and just paste the graphics or numerical values into any other windows application. An efficient way to add data to excel, make an image for a presentation, or for including in a report. Close the Stack Plot window.

<u>Stack regression</u>

A stack can be used as input for map calculations in DIVA-GIS. One such option is to calculate a regression trend using a stack. DIVA-GIS supports calculating an ordinary

least square (OLS) regression using a stack as input. We can hence use DIVA-GIS to find areas where vegetation has increased or decreased over the last 25 years. The regression option is reached via the menu: Stack – Regression. Select the same stack as you used for the stack plot above, click the Output button and give a logical output name, and click Apply to perform the regression.
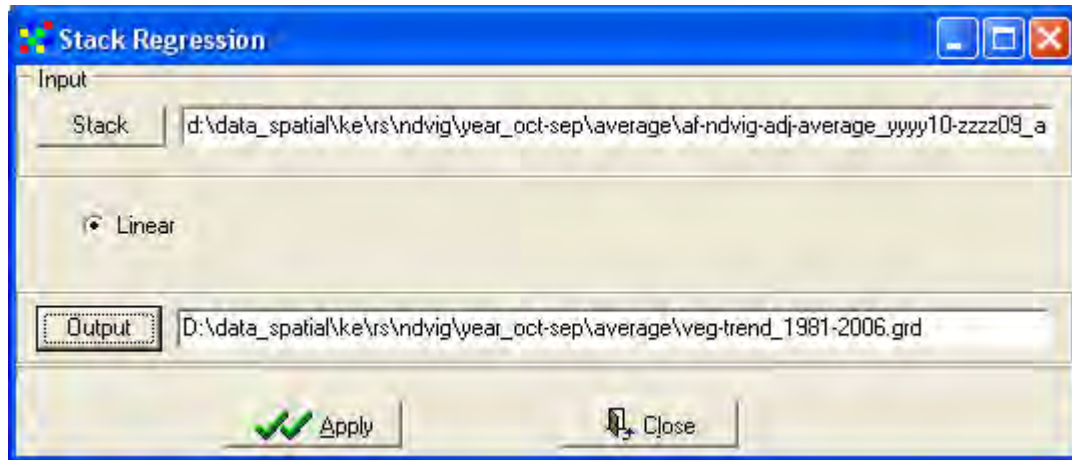

DIVA-GIS generates two files from the regression, one with the suffix _slope (the strength of the slope) and a second with the suffix _r2 (the correlation in the slope). Only the first file is automatically added to the project, the second must be added using the Add layer tool. Below you see an example of how the regression slope result has been symbolized using color ramping in 8 classes.
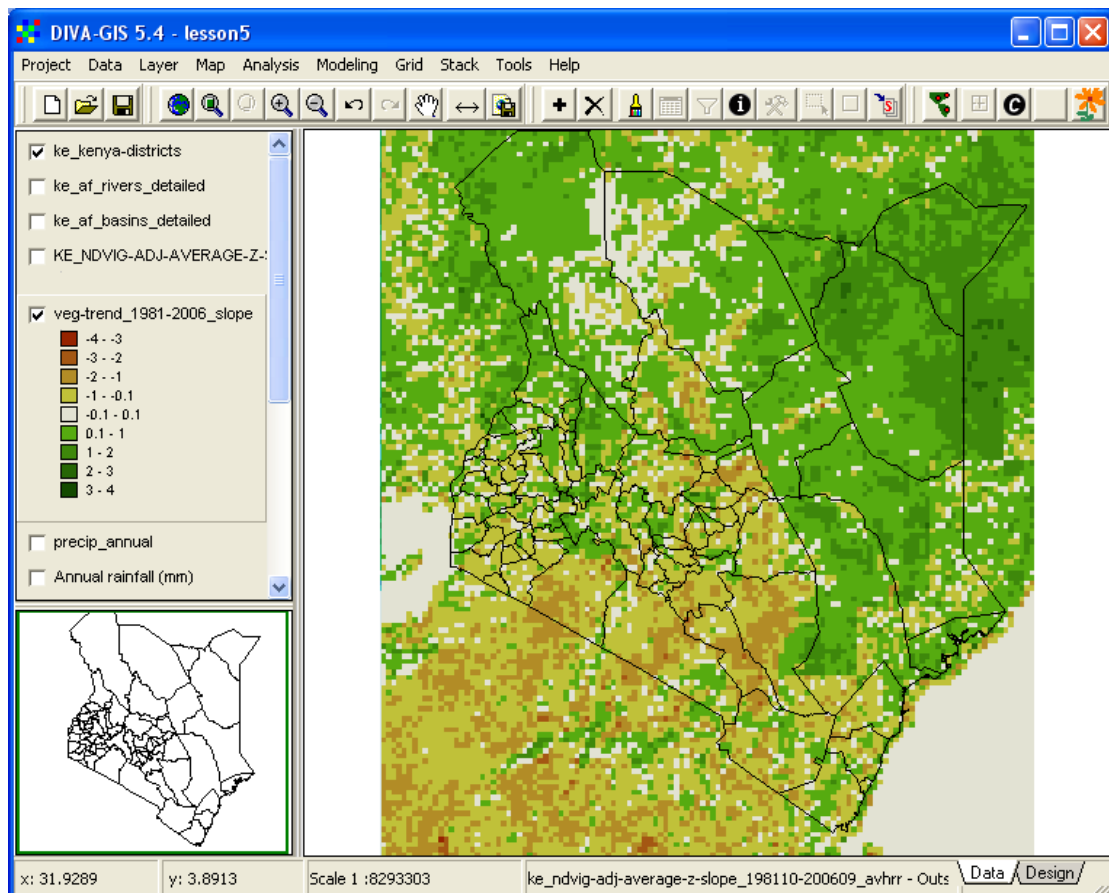
Normalized vegetation trends

The use of Ordinary Least Square Regression for identifying trends in time-series data with large variation is a bit questionable. For a first test it works but more advanced

statistical methods should ideally be applied before any management or policy is chosen. On the project CD there are ready-made trends using normalized vegetation data for each annual vegetation index included (the section on normalization at the end of this lesson to get more information on the normalization method). The normalized trends are also derived from OLS (and hence a bit questionable). The normalization (standardization) was done by calculating the average and variation (standard deviation) for each cell for the period 1981-2006. A value (called z-score) was the derived that measures the relation between the original NDVI value of each cell and year, and the time series average and standard deviation for that cell. In short this method is called is applied to remove the influence of extreme (high or low) values that otherwise have an un-proportional large influence on the regression outcome. You can now compare your regression using the original data, with the regression done using normalized data. The z-score regression is in the same folder as the original data, but with "z-slope" inserted – it is usually the last file in the list when you open the Add layer window. Below you see the normalized trend slope for average NDVI 1981-2006, apart from the colors in the symbolization and the numerical values (that can not be compared anyway) the trends looks similar to the one above.
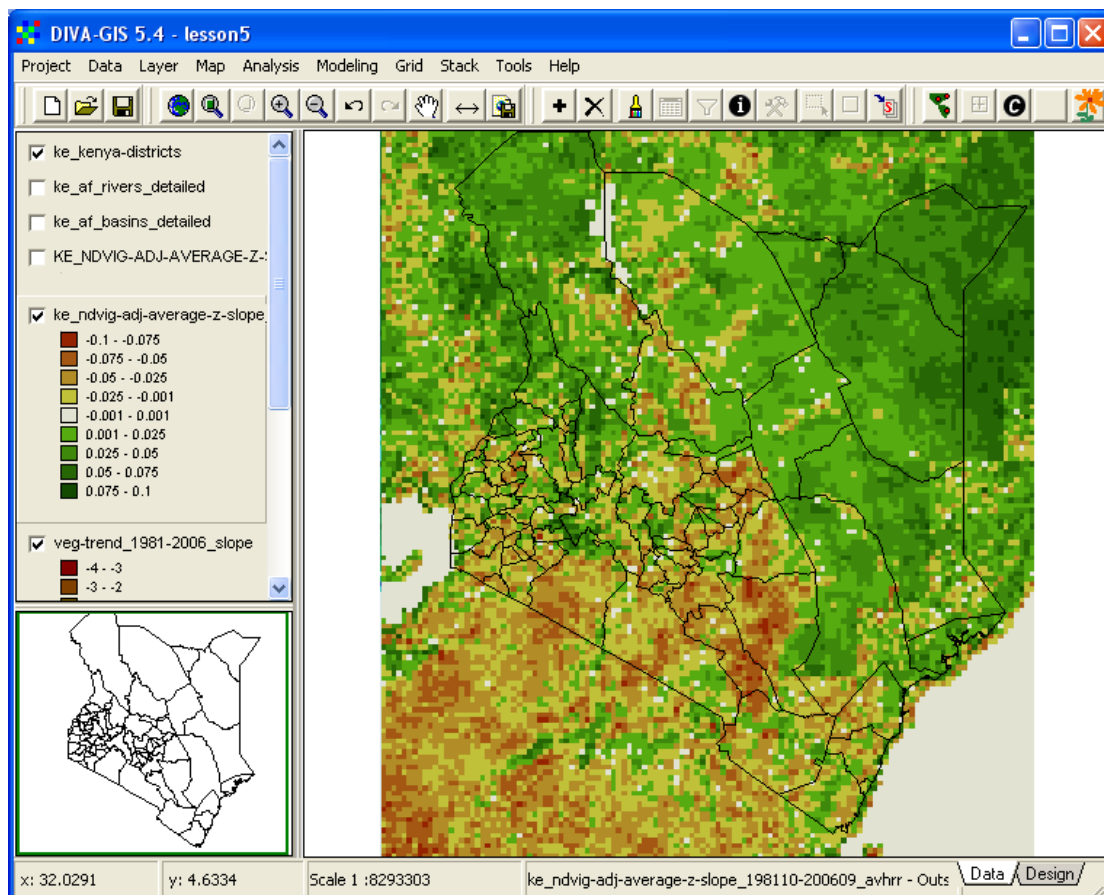

Annual Maximum and increment NDVI

Now you can repeat the steps done using the annual average NDVI, with the annual maximum and annual increment NDVI indexes, The difference between these indexes are discussed at the beginning of the this lesson. There are already stacks available for both indexes, and also the normalized ("z-score') trends are included.

Identify degradation hotspots

Each of the annual vegetation indexes; annual average, annual maximum and annual increments, have merits and drawbacks. Here you are going to create a degradation hotspot map by combining all three indexes into one. Hotspots for potential land degradation will be defined as cells (pixels) that in all indexes have shown a clear negative trend over the last 25 years. To accomplish that, first build a stack that includes the three z-score slopes, as shown below. The stack contains grid files from different folders, but that is no problem. In the example the stack is saved under the folder my diva.
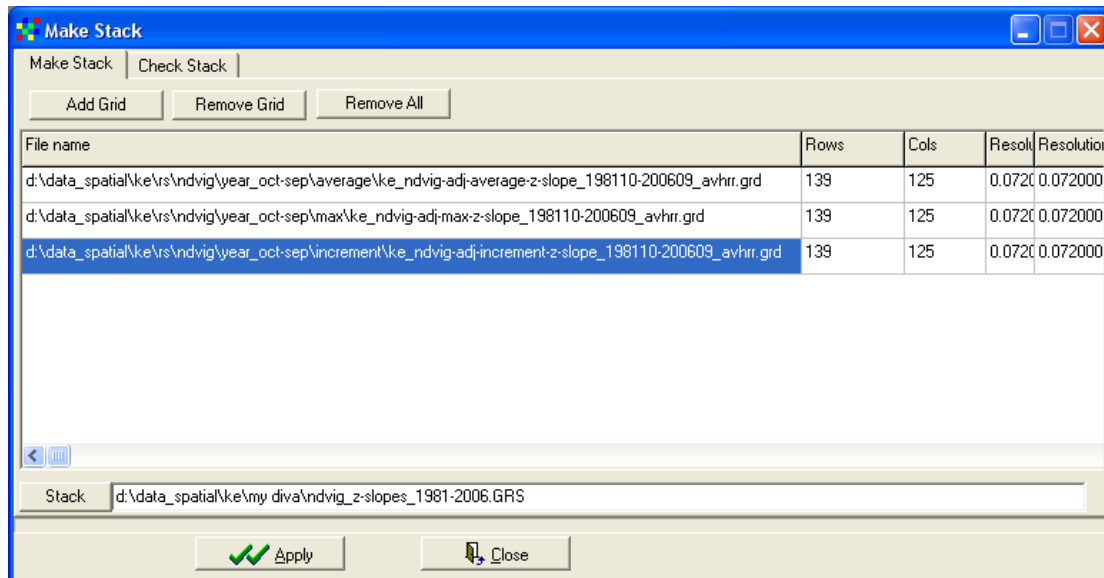

You can use the Stack Plot function to probe any cell in the map for changes in all three vegetation indexes. But it is visually more efficient to create a layer showing the hotspots.

To make the combined hotspot map, you should use the reclass function, open it via the menu: Grid – Reclass. Reclass in DIVA-GIS has the option of operating on all grid layers in a stack, which is what you are going to do, so click the radio button for Stack. Then Click the Input button and open the stack you just created (ndvig_z-slope_1981-2006.grs in the example). Then click the Output button, navigate to your my diva folder and give the stack to be generated a logical name. The Reclass function will generate one new output file for each layer included in the stack, keeping the name of the input layer, but adding a suffix. Give the suffix you want the new layers to get in the text box Suffix for output filenames. The trend slope of normalized time-series data is always in the range –1 to +1, and hence your reclass values should be set so that negative values get a new value of 1 (= areas identified as having potential land degradation) and positive values get a new value of 0 (= no land degradation). You must use 0 and 1, as in a later step you shall use the multiply function to combine the three maps into one. In the example below the reclass threshold was set using a negative value (-0.05) rather than zero. By using a negative value all areas with no trend (e.g. water areas) will get a new value of zero, and the hotspot map will identify a smaller number of cells, but with more certain negative trends. You have to try iteratively which threshold to use.
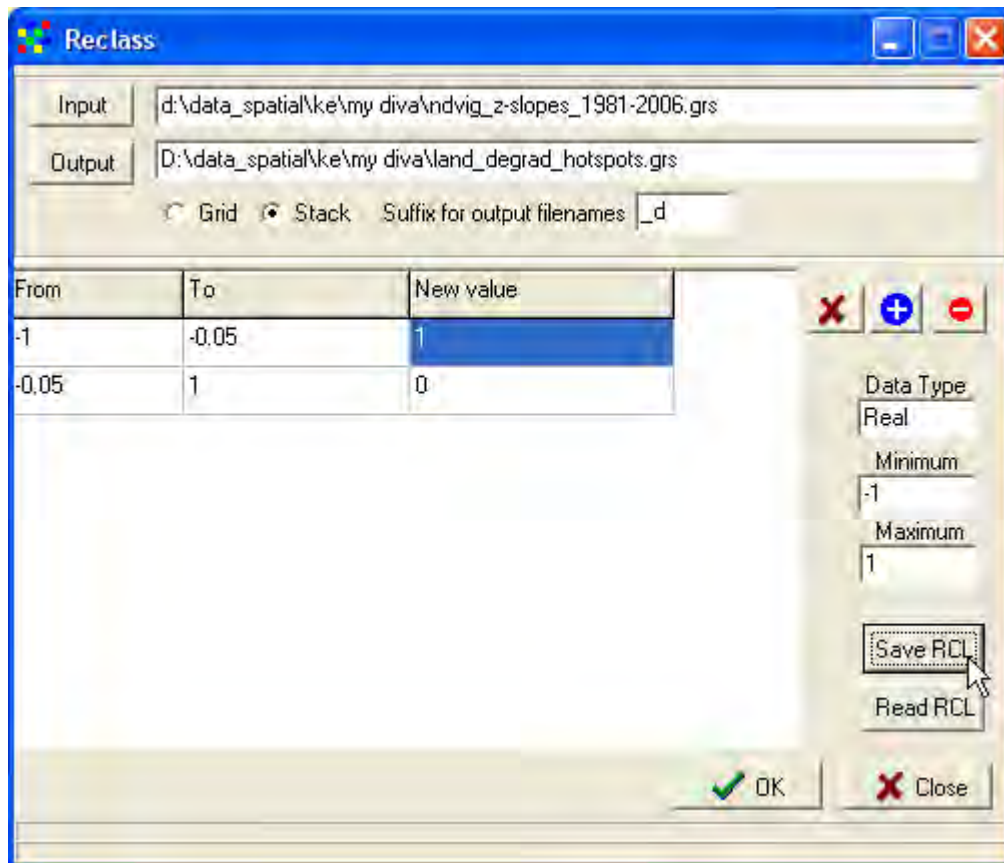
Click OK to perform the reclassification. The new stack, and the three reclass grid layers are not added to the project, but you can find them in the my diva folder. Below you see the potential degradation sites for each index.
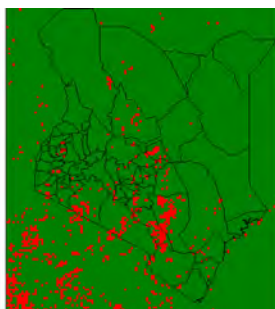

| Average | Max | Increment |

To derive a single hotspot map you can multiply or add the three Boolean (0/1) hotspot maps. This is best done using the Grid calculator that you find from the menu: Grid – Calculate. In the Grid Calculator, start by clicking the Add Grid button,

, and add the first Boolean (0/1) hotspot map, then click the

multiply button, . In this way you can multiply all three Boolean (0/1) maps together to identify areas were all vegetation trends have been negative. If you want

to, you can use addition, , instead of multiply, and then you will get a map with values from 0 to 3, with three denoting cells where all trends have been negative.
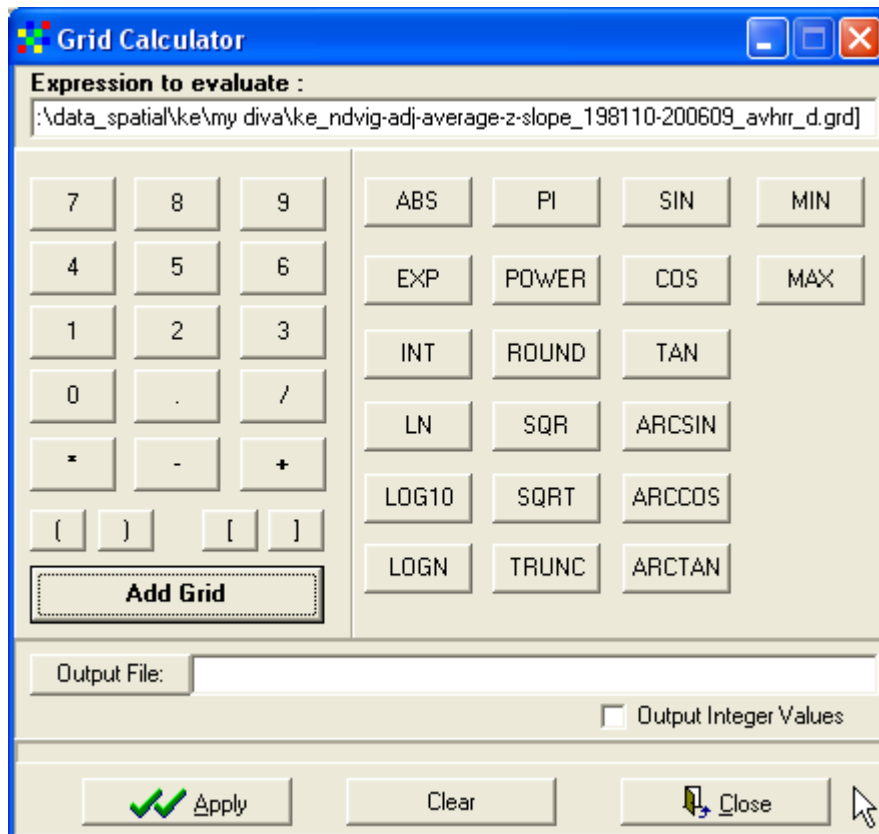
The example below shows the result of summing the three Boolean hotspot maps, and then symbolizing the resulting layer.
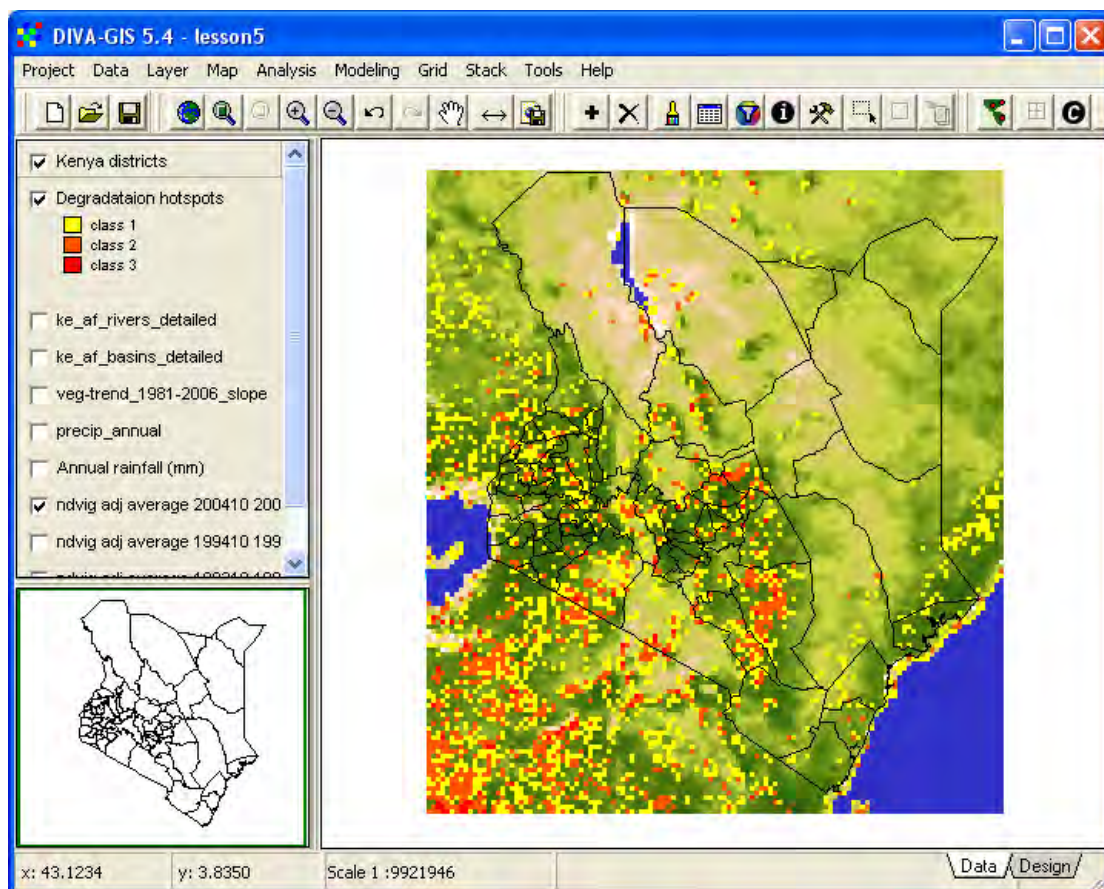
In a similar way you can also identify cells that have had a positive vegetation trend ("coolspots") in the period 1981-2006. As the resulting map(s) could be of a wider interest, you should symbolize and label the map in the Data View and then design a map for publication. See lesson 3 to repeat designing a map in DIVA-GIS. The designed map below was created in several steps:

1. Converting the district polygon vector layer to a grid, using one of the vegetation images as template for extension and resolution (link),

2. Reclassification of the derived grid map for districts to a Boolean (0/1) mask showing Kenya only (link),

3. Overlay (multiplication) of the Boolean mask with the hotspot map to derive a hotspot grid map including values only for Kenya (link),

4. Symbolizing the of the Kenya hotspot map by loading the legend from the previously symbolized hotspot map (link),

5. Arranging the layers in the Legend of the Data View (link),

6. Setting the theme labeling in the Properties window (link),

7. Using the Design View to compose a map design (link).

Land degradation hotspots in Kenya
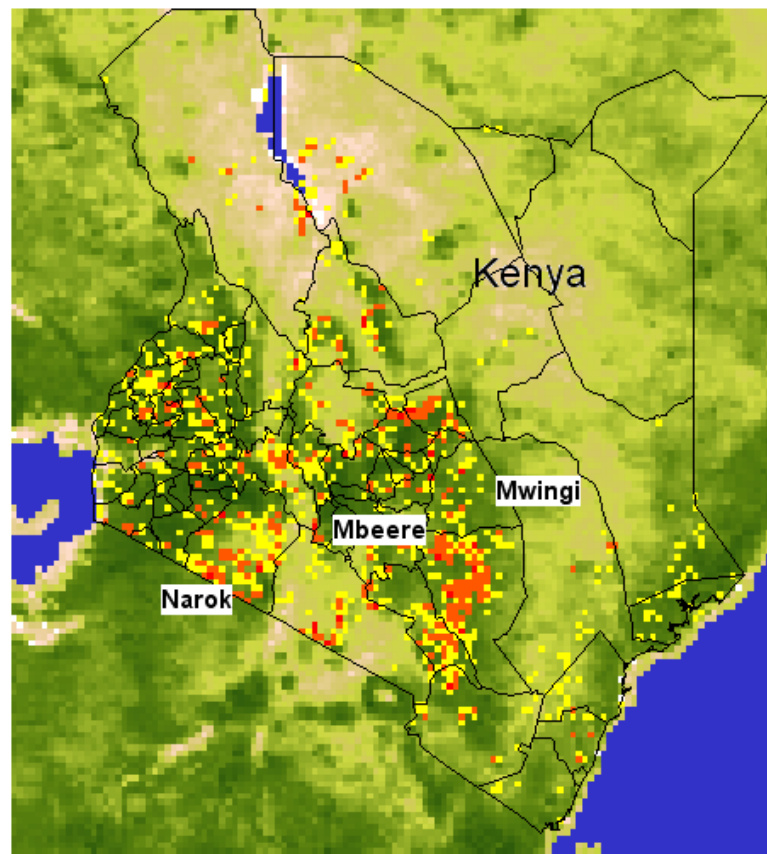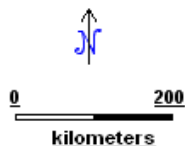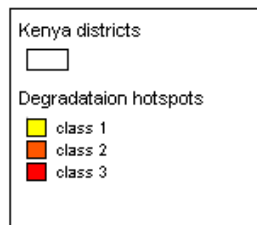
Normalization (from wikipedia.org)

In statistics, the standard score, also called the *z*-score or normal score, is a dimensionless quantity derived by subtracting the population mean from an individual raw score and then dividing the difference by the population standard deviation. This conversion process is called standardizing or normalizing.

The standard score indicates how many standard deviations an observation is above or below the mean. It allows comparison of observations from different normal distributions, which is done frequently in research.

The standard score is: $z = \dfrac{x - \mu}{\sigma}$

where:
    x is a raw score to be standardized
    σ is the standard deviation of the population
    μ is the mean of the population.

The quantity *z* represents the distance between the raw score and the population mean in units of the standard deviation. *z* is negative when the raw score is below the mean, positive when above.