



Lake Tanganyika Project: Zambia National Site Characterisation and Catchment Management Design Workshop

GIS training module – Lesson 6

Thomas Gumbricht, ICRAF



GLOBAL ENVIRONMENT FACILITY
INVESTING IN OUR PLANET

Lake Tanganyika Regional Integrated Management Project

A United Nations Development Programme (UNDP) / Global Environmental Facility (GEF) project covering the Lake Tanganyika riparian countries, Burundi, the Democratic Republic of Congo, Tanzania and Zambia.

Training module created by Thomas Gumbricht,

Last updated: October, 2010

LESSON 6 – Rain Normalized NDVI and land degradation hotspots

Lesson overview

In this lesson you will learn about how to disentangle rainfall driven vegetation changes from other factors affecting vegetation changes using the concept of Rain Normalized NDVI (rNDVI); more advanced grid calculation are introduced; you will also learn about using multiple criteria for evaluating land degradation.

Rain Normalized NDVI (rNDVI)


To overcome the shortcomings of using annual vegetation indexes for tracking land degradation (as in [lesson 5](#)), the concept of Rain Use Efficiency (RUE), which combines rainfall and vegetation information, was developed for arid and semi-arid regions. It is an attractive index in regions where growth is limited by rainfall, and where there is a linear relation between rainfall and vegetation. RUE is reported to be lower in degraded arid lands compared to equivalent non-degraded areas and is hence a useful index for separating the effects of rainfall from human factors on temporal changes of vegetation in rangeland ecosystems. As RUE is calculated using seasonal or annual time-steps it is also hypothesized to bridge short-term fluctuations in vegetation dynamics. It, however also means that RUE is not sensitive to short-term droughts or heavy storms, and thus might be biased by such events. The concept of Rain Use Efficiency was developed for estimating Net Primary Production in dryland ecosystems. The Lake Tanganyika project focuses on identifying areas with potential soil degradation (or soil improvements). We do not use the ratio between vegetation and rainfall for primarily forecasting vegetation growth. As we use the ratio between vegetation growth and rainfall as an index for studying changes and trends in time we prefer the term Rain Normalized (rNDVI) instead of Rain Use Efficiency.


Start a new DIVA-GIS project

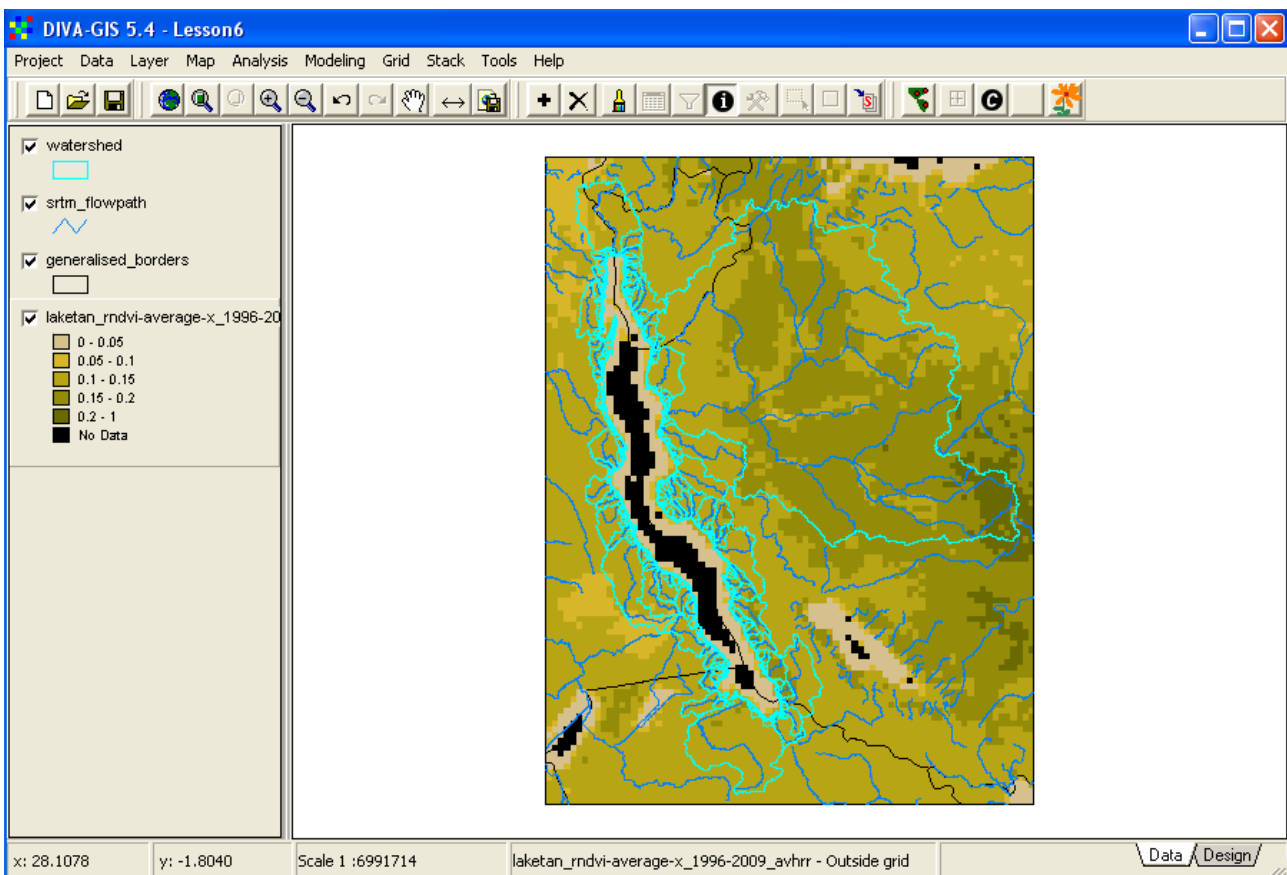
Start DIVA-GIS and create a new empty project, save it in your my_diva folder, in the tutorial the project is saved as lesson6.div. Add some framework data to use as positional reference (e.g. the national borders and rivers that you used in the previous lessons).

Rain Normalized NDVI

Rain Normalization of NDVI (rNDVI) is a method to disentangle the rainfall driven variation in vegetation growth from other factors governing changes in vegetation ([see above](#)). rNDVI is simply calculated as the ratio between an annual vegetation production index and annual rainfall (some studies use growing season instead of annual indexes, but in this project the annual indexes have been used – see lesson 5). In this study you will find already calculated datasets for the annual Rain Normalized NDVI under the directory `\data_spatial\laketan\RUE`

rNDVI in general has values between 0 and 2.5 with the input units used in the dataset. When calculating rNDVI areas with low rainfall but high vegetation production will generate very high rNDVI values. This can for instance be irrigated areas, or flooding areas. DIVA-GIS sets the default legend after the highest and lowest value resulting from the overlay analysis, so by default your rNDVI map will become a (rather) green map. You can check the actual values by either using the Identify tool, , by making a transect (see lesson 5) or by making a histogram of the calculated rNDVI map (see lesson 5). In this study, the rNDVI layers were calculated for the years from 1996-2009

As an example, you can ADD  the rNDVI for 2009 from the folder `\data_spatial\laketan\RUE\rNDVI\`

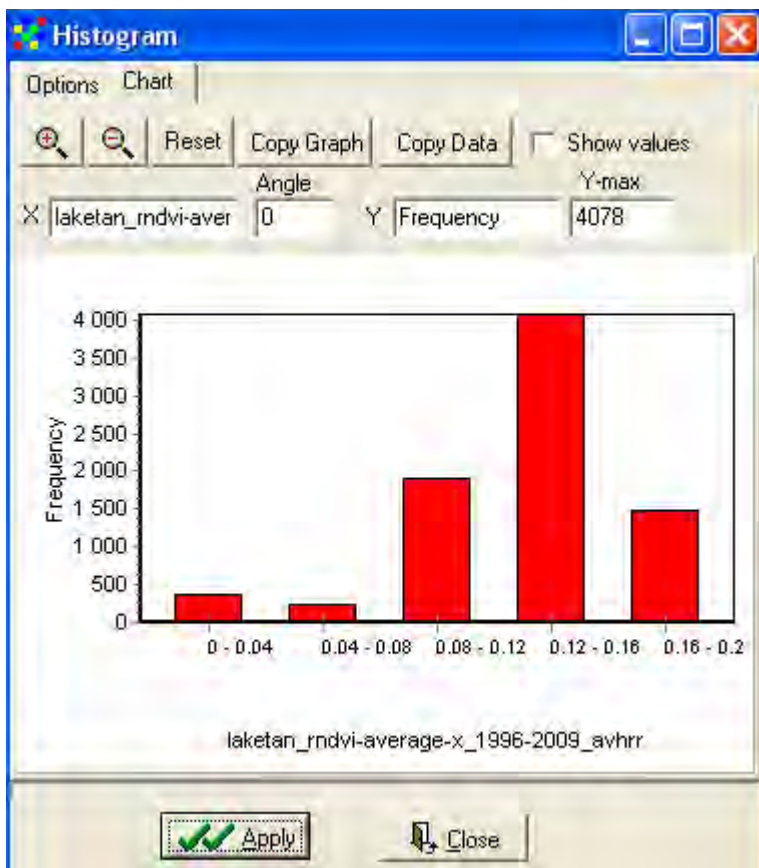


Use some of the visualization tools (for example transect or histogram) that you used in earlier lessons to analyze the distribution of the rNDVI data.

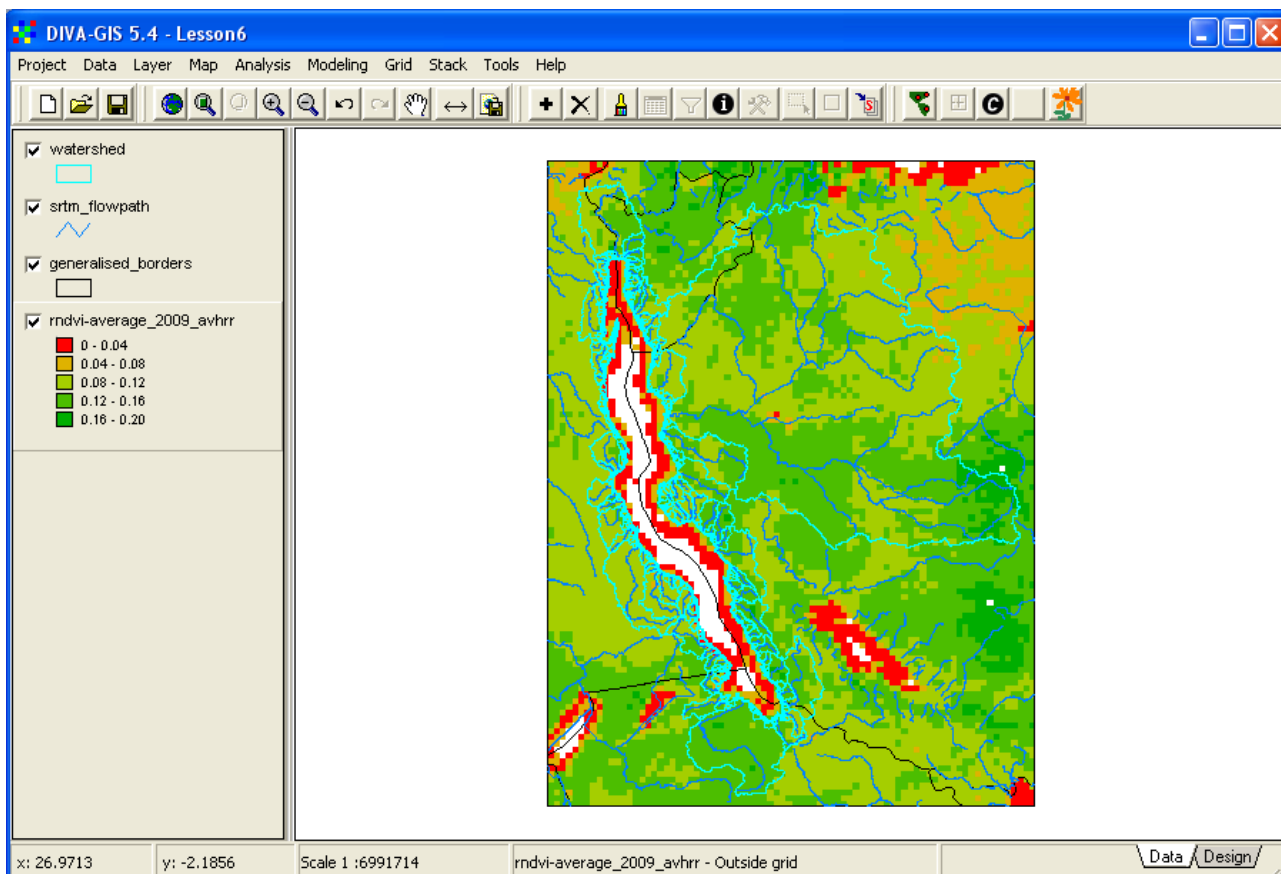
To make a histogram of the rNDVI grid you must set the histogram options manually, for example as shown below. If you use the default setting of displaying the histogram using the Legend entries, you will get almost all pixels falling in one class.



With the manual set Histogram options you see that most rNDVI values are lower than 1.



If you look at the histogram above you can get an idea of how to rearrange the legend in order to symbolize the calculated rNDVI layer. Use this information to set the legend for the rNDVI layer (the rNDVI data is of the type ratio, see lesson 4 to repeat symbolization of grid ratio data). From the symbolized rNDVI example below you can clearly see how some patches get very high values. The rainfall in the north-eastern part of the Lake Tanganyika Basin is high, and not limiting for vegetation production. The rNDVI values in Burundi and northern DRC are hence not reliable, for the Tanzanian and Zambian parts of the Basin rNDVI should be a more appropriate index to use for evaluating land health.



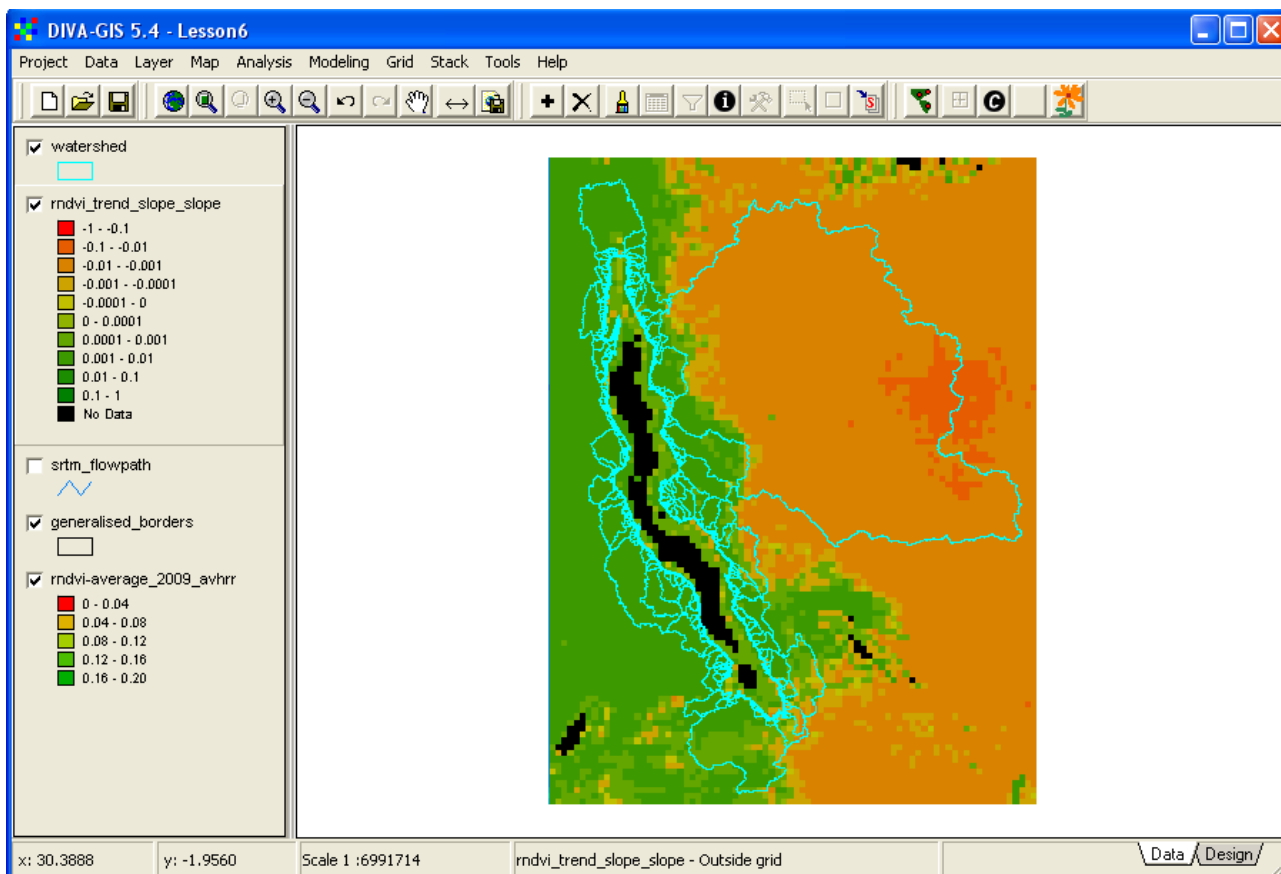
The rNDVI layer shows the variation in vegetation growth with the influence of rainfall taken away. If the vegetation data had been in units of biomass production per area (e.g. g/m^2), rNDVI would have shown the absolute growth per mm of rainfall (and hence a proper Rain Use Efficiency - RUE). There are documented methods to convert NDVI to absolute biomass growth, but detailed ground surveys are needed to adopt those methods. And in this study we are mostly interested in finding areas that have relatively higher and lower rNDVI and areas where rNDVI has changed over time, hence the indexed rNDVI that you just calculated will be sufficient.

The rNDVI for Lake Tanganyika is already calculated for each year 1996 to 2009. All the datasets are in one folder, `\data_spatial\laketan\RUE\rNDVI`

For each index there are layers representing the mean situation for the period 1996-2009, the normalized trend slope 1996-2009 based on z-scores (see [lesson 5](#) to repeat normalization and z-scores), the normalized trend slope coefficient of determination (regression r^2) 1996-2009, and layers showing the significance (p) of positive and negative trends. The significance was calculated by comparing the actual trend with 999 random trends (permutations).

Trends in rNDVI 1996 to 2009

The trend in rNDVI is a strong candidate for identifying potential areas with vegetation loss and land degradation. To calculate the trend in rNDVI over the period 1996 to 2009 a pre-made stack is available in the rNDVI sub-folder. To calculate the trend in rNDVI go the menu: Stack – regression. Navigate the rNDVI folder and use the stack `rndvi.grs`. Change the legend to show the result more clearly, in the example below the legend has been set to a logarithmic scale.



The rNDVI slope is a poor estimate of vegetation performance in the western part of the basin, as rainfall here is not limiting for vegetation production.



Multi Criteria Evaluation (MCE) for identifying hotspots and coolspots

Another alternative for identifying hotspots (and coolspots with improved vegetation growth) is to combine several indicators or vegetation changes. The method of combining two or more criteria (maps) for identifying a single objective (land degradation in this case) is called Multi Criteria Evaluation, and is a standard decision tool also in other fields than GIS. Criteria can be of two kinds – constraints and factors. A constraint is similar to a mask in that it allows the identification of the objective or not. A factor is assigned a weight that is used as a criterion in the decision model for identifying the objective.

Here you shall use one constraint, and four factors for identifying hotspots. You should constrain the evaluation to the land part of the Lake Tanganyika Basin. The factors are four images derived from statistical analysis of the NDVI and rNDVI data.

Create a constraint map

To create a constraint (0/1) map identifying the Lake Tanganyika watershed area, you must first convert the `country_index.shp` to a grid map, then reclassify this grid map to have the value 1 for included countries and 0 (or No Data) for other areas.

There are several options for selecting the countries included in the project. Either make the layer with countries as polygons the active layer, and open the Select window (e.g. by clicking the Query tool, , or via the menu: Layer – Select Records), and select the countries. Or use the interactive Select Feature tool, .

When you have selected the countries, use the Convert to Shapefile tool or go via the menu: Data – Selection to New Shapefile, and save the included countries as a separate shapefile.

Then convert the new vector layer to a grid by using the menu: Data – Polygon to Grid. Use any field you can refer back to as the value field to use in the grid to be produced, and one of the existing grid files of rNDVI as template file.

DIVA-GIS automatically sets areas outside the polygons in the vector layer to No Data in the grid output file. You must now use reclass (from the menu: Grid – Reclass) to set the grid value of all included countries to 1.

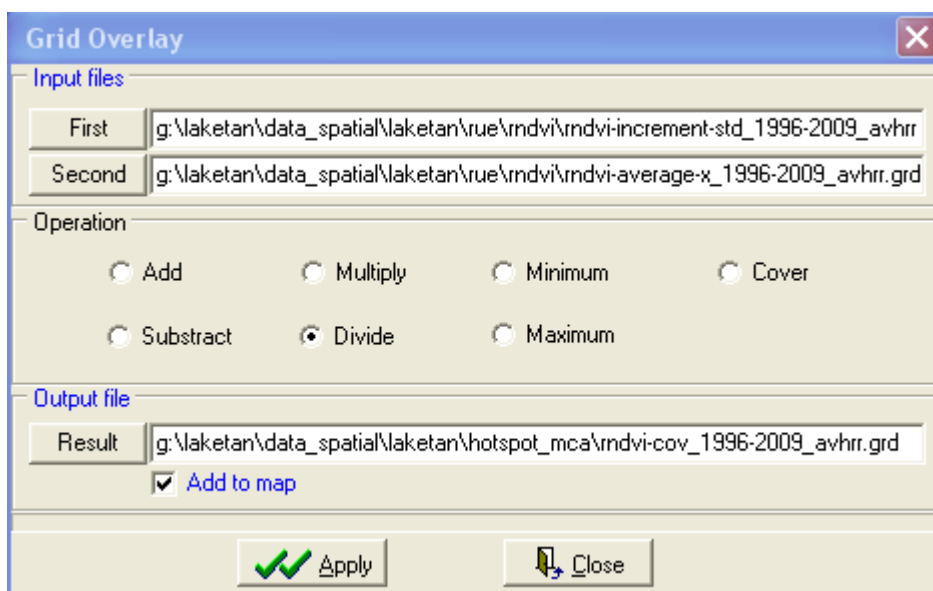
If you want to add more constraints to you MCE you can do that, and then use overly to multiply the constraints (0/1) layers to create an overall constraint. As shown above the lake shores causes anomalous data, and need to be removed. Under /data_spatial/hydro/grid there is a prepared lake mask that you can use.

Creating factor maps

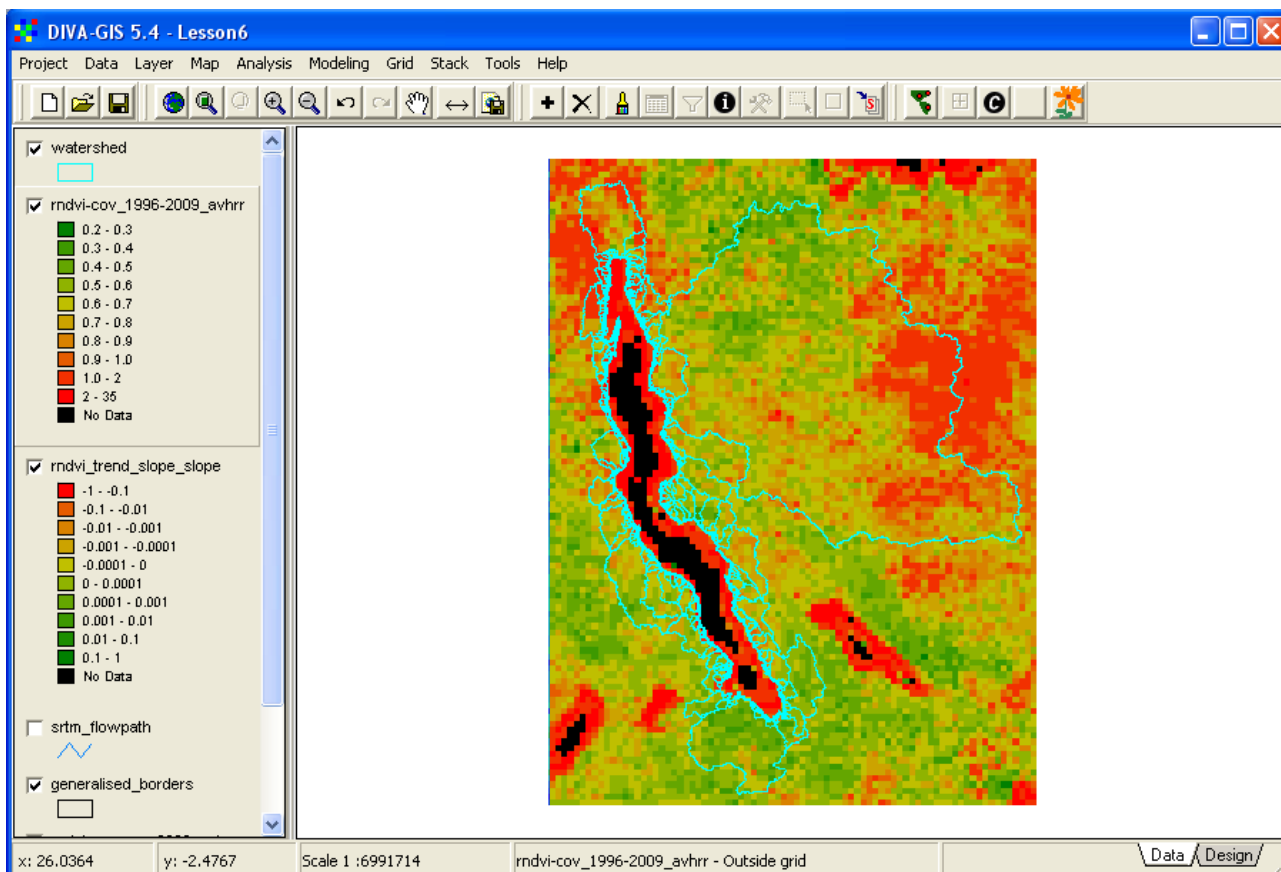
Above you created the trend map for rNDVI, which will be used as the first factor map in the MCE of land degradation. Another way to analyze vegetation performance is to analyze the variation (not the trend) over time. Large inter-annual variations signifies an unstable (or unsustainable) vegetation situation. A common statistical method to use for analyzing stability is the Coefficient of Variation (CoV):

$$\text{CoV} = \text{standard deviation} / \text{average}$$

You can calculate the CoV for rNDVI by using the layer rndvi-average-std_1996-2009_avhrr.grd for the standard deviation and the rndvi-average-x_1996-2009_avhrr.grd layer as the average.



The CoV map will have the highest values around the lake shores, this is an artificial variation stemming from the fact that these pixels are a mix of land and water, When visually inspecting the results of the CoV analysis you need to arrange the legend to accommodate for this artifact.



You should now have 2 statistical indexes indicating land degradation hotspots – the rNDVI trend and the rNDVI CoV. As rNDVI is only really useful when rainfall is below about 1000 mm per year, we will also use the CoV and trend from the original NDVI data. Calculate the NDVI trend (for 1982 to 2009) using the annual average NDVI in the folder /data_spatial/rs/NDVI/GIMMS_NDVI-annual-average. Use the stack available in that folder to calculate the trend. Then use the NDVI standard deviation and average layers in the sub-folder GIMMS_NDVIg-annual-increment-stat, to calculate the CoV for NDVI.

Creating a compound land degradation hotspot index

You should now have 4 factors that we can regard as factors indicative of land degradation:

- NDVI trend 1982 to 2009
- NDVI Coefficient of Variation 1982 to 2009
- rNDVI trend 1996 to 2009
- rNDVI Coefficient of Variation 1996 to 2009

Normalizing the factors

The factors represent different statistical properties, and have very different values. The original NDVI are in the range 0 to 250, whereas rNDVI is in the range 0 to 2. You must account for this by normalizing the factors – given them a similar span (say 0 to 100). You can either use various grid functions to achieve this, or you can account for it by combining normalization and weighting in one step. Also note that a positive CoV is indicative of land degradation, whereas a negative trend signifies a loss in vegetation production. Also this you must account for when compounding the factors.

Weighting the factors

The weighting of the factors is a major part of MCE. It is normally done using expert panels, where experts from different disciplines discuss and decide on the relative importance of factors (this was done for the MCE in lesson 8). In this case you simply assign the weights you want, or use the equal weights for the 4 factors.

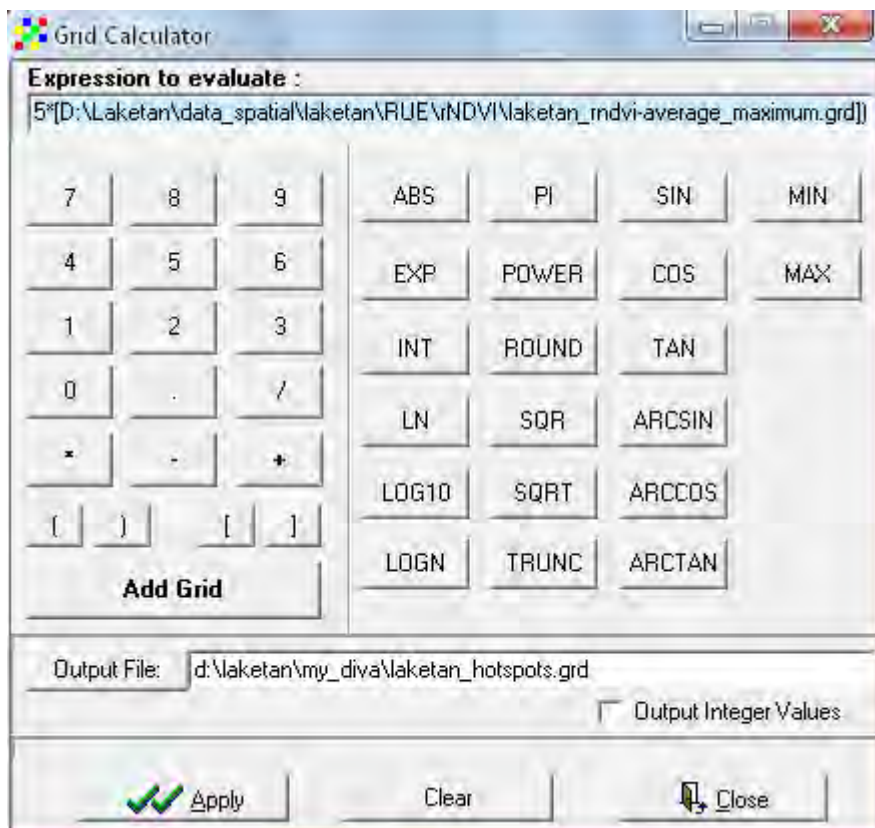
Grid calculation

As you have the constraint map and the four factor maps, you are ready to execute the Multi Criteria Evaluation. The MCE is done as a map calculation:

$$\text{MCE} = \text{constraint} * (\text{factor1} * \text{weight1} + \text{factor2} * \text{weight2} + \text{factor3} * \text{weight3} + \text{factor4} * \text{weight4})$$

Where the weights should add up to 1 or 100.

You can easily implement the above formula in DIVA-GIS by using the Grid calculator from the menu: Grid – Calculator. The Grid Calculator does not work in version 7.2, instead you have to use the scalar and overlay functions under menu – grid, to implement the formula above.

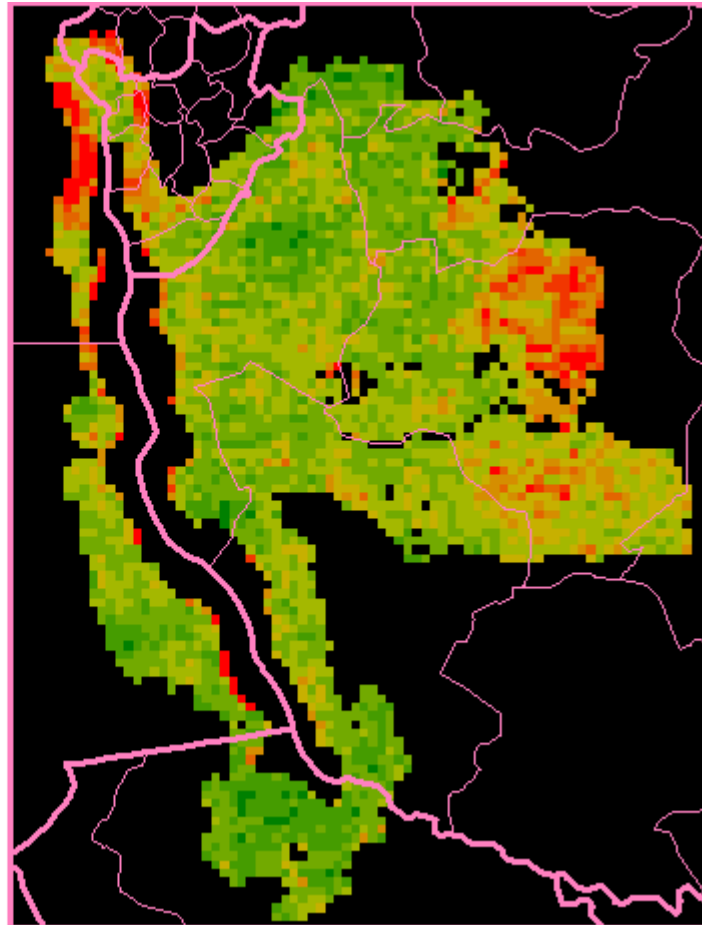
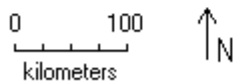


Use the Add Grid button, **Add Grid**, to add the grid layers and the other buttons to put the operators and parenthesis needed.

Give the MCE for identifying hotspots a logical name, and press Apply to run the MCE calculation.

You can analyze the MCE for identifying hotspots by looking at the histogram (see above) and then use the Legend to threshold the MCE to show only the most affected cells, as in the example below.

Lake Tanganyika
Land degradation hotspots



Try it your own 1 - Correlation between temporal trends in rainfall and vegetation production

In lesson 5 you did a statistical regression using an independent variable (rainfall) to explain the variation in a dependent variable (vegetation production) in a spatial context. That analysis clearly showed vegetation production to be strongly dependent on rainfall up to around 800 to 1000 mm of annual rainfall.

With the trend images of rainfall and vegetation production you can do a similar regression to test the influence of the rainfall trend on the vegetation production trend. It is better to use trends derived from the original data and not the normalized trend. Using a trend derived from the original data will be interpretable in terms of how much changes in rainfall affects vegetation production. If you use normalized data the regression formula can not be interpreted explicitly and is only valid for the period you analyzed (i.e. it can not be used for extrapolation) The principal steps you need to take include:

1. Make a stack with annual rainfall grids for the period you want to analyze
2. Make a stack for annual vegetation production for the same period (you can use any, or compare all, of the vegetation production indexes included in the dataset),
3. Calculate the trends in rainfall and vegetation production,
4. Cut the trends to fit a country, river basin or district,
5. Convert the selected country, river basin or district to a grid, using the cut trend grids as template,
6. Create a Boolean (0/1) mask of your selected country, river basin or district,

7. Apply the mask to your trend grids to remove pixels outside your study area by multiplication,
8. Make a regression using the cut and masked rainfall trend as independent and the cut and masked vegetation production trend as dependent variable,
9. Analyze the results from the regression, by using the transect tool and a histogram,
10. Symbolize two maps showing the cut and masked trends in rainfall and vegetation growth,
11. Compile the results into a short report on changes in rainfall and vegetation production in your selected area, including the two designed maps, histograms of the derived trends, one or two transects for both rainfall and vegetation production trends (using the same position for the transects in both maps), and both graphical and numerical results of the regression.

Try it your own 2 – Identify hotspots in a sub-region

You should finally also identify potential land degradation sites in the region you used above (Try it on your own 1). Follow the steps taken above in this lesson to create a hotspot map from either the ranking maps, or from the normalized trends in rain normalized *NDVI. To make the analysis even better, you can include the layers containing information on the significance of the trends (use them as constraints in the MCE). If you restrict your analysis to your home district, the principal steps you need to take include:

1. Cut the grid layers you want to use as criteria to fit the grid data you used above (Try it on your own 1).
2. Normalize the input data to a common scale (either use the grid calculator, or the scalar function under the grid menu)
3. Assign weights to the factors you want to use in the MCE
4. Execute the MCE
5. Analyze the histogram of the MCE
6. Symbolize and design a map showing the identified hotspots
7. Compile the results into a short report on hotspot identification using MCE, including the designed maps and the histogram of the MCE result.