Raster Risk Assessment: Determining the Risk of Malaria Transmission in Ethiopia

Written by Carolyn Talmadge and last updated 10/14/2018 for ArcMap 10.6.1

Introduction

Malaria is a serious, acute illness transmitted via the bite of a mosquito. Symptoms include fever, headache, backache, joint pains, and vomiting. If not treated, malaria can quickly become life-threatening. According to the World Health Organization (WHO), about 75% of the land and 60% of the population is exposed to malaria in Ethiopia.

The purpose of this **analysis** is to determine the overall risk of Malaria transmission throughout Ethiopia, when considering the following factors:
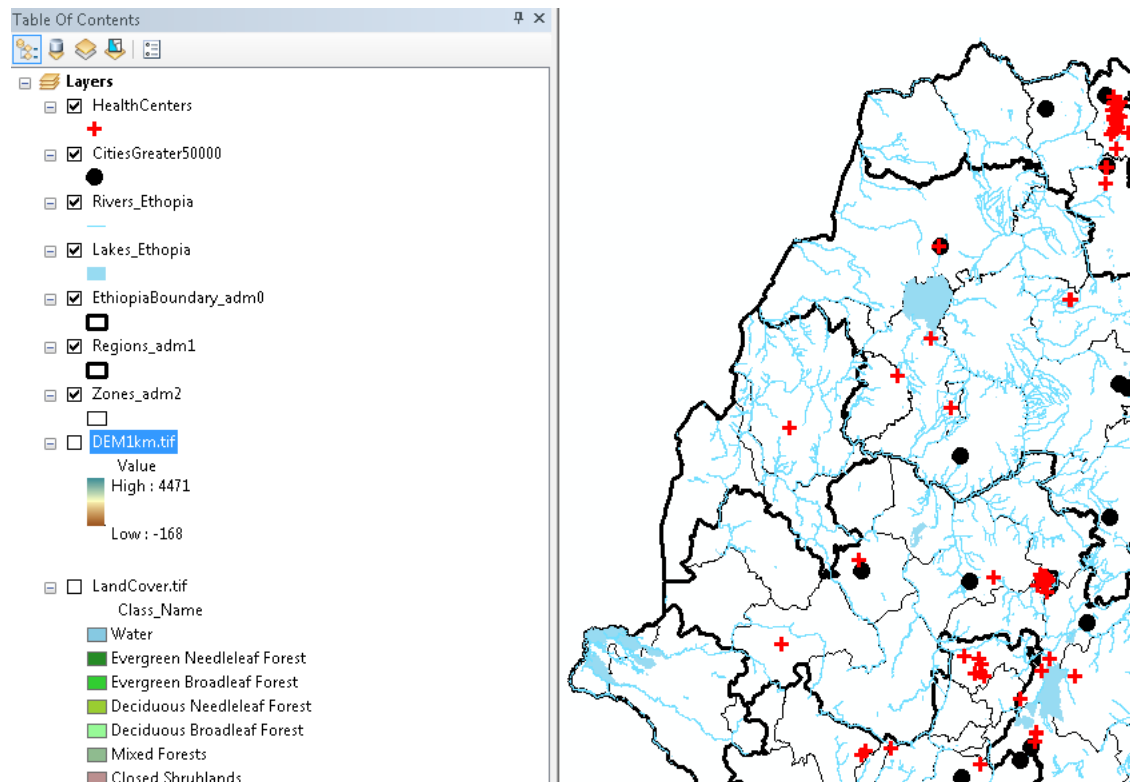
- **Proximity to rivers** – This is a "water related" disease because it is transmitted via mosquitos who like to breed near rivers and stagnant waterbodies. Therefore, proximity to rivers results in higher risk of transmission.

- **Elevation** – Elevation is directly linked to climate, and mosquitos prefer warmer climates and lower elevations. Lower elevations increase mosquito population turnover, which means a higher risk of transmission.

- **Slope** – Gentle slopes equal higher susceptible areas for malaria incidence.

- **Land Cover** – Mosquitos prefer damp areas, so land cover will affect the analysis.

- **Accessibility of Health Clinics** – If not treated, malaria can quickly become life-threatening. Therefore, access to health clinics and hospitals will greatly affect the populations' *vulnerability.*

Accessing the Data and Preparing for the Tutorial

1. The data is stored in *S:\Tutorials & Tip Sheets\Tufts\Tutorial Data\Malaria_Ethopia_RiskAnalysis*. Open a blank ArcMap and in **Catalog**, **copy** the entire folder to your **H: drive.** Alternatively, you can download the data from this link and save it.

2. You need to make the copied folder "writable" – to do this, right-click on the folder in **Catalog**, choose

Properties, and uncheck the **Read-Only box**. When prompted say yes to *apply to subfolders*. If you downloaded the data from the link, you need to unzip it before making it writable.

3. From **Catalog**, pull in the following data files and order the layers logically. Symbologize them as follows:
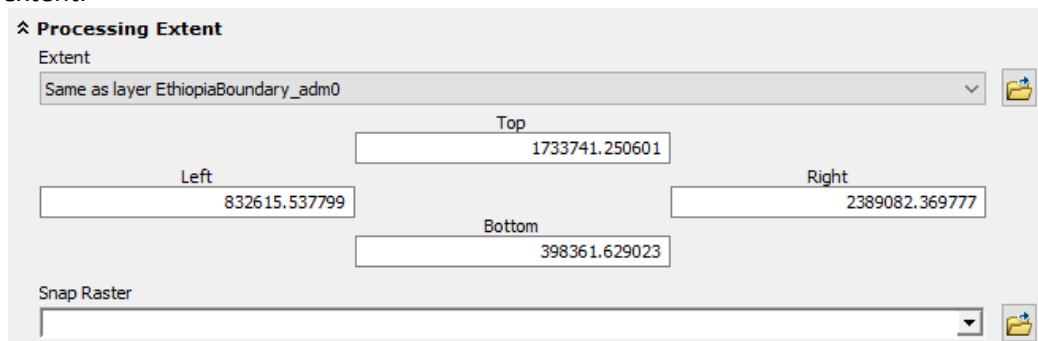

4. Take a couple minutes to explore the different data sets and attribute tables.

5. Check all of the projections. Is the data **projected**? If so, what are the **units** that it is using?

6. Save your **MalariaRiskAnaylsis.mxd.** mapfile in your H Drive.

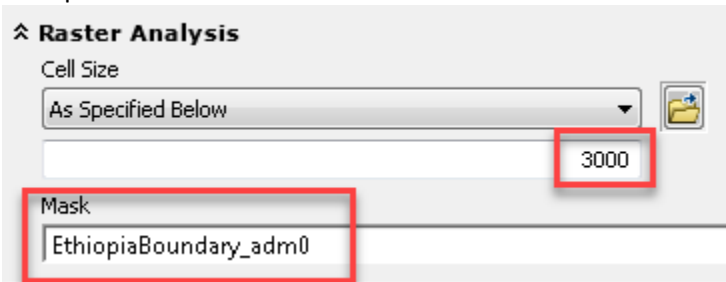## Turning on the Extension & Setting the Extent, Cell Size, and Snap Raster

When performing raster-based overlay analyses, it is very important to set the **extent** of the analysis, the **cell size**, and the **mask**. You can set these options before the analysis by going to **Geoprocessing → Environments**.

Note: You have to set this with each session of ArcMap you start.

1. Go to **Customize → Extensions →** checkmark *Spatial Analyst*. You must do this on any computer whenever you plan to use spatial analyst tools!

2. Click on **Geoprocessing** and select **Environments**…

3. In the Environment Setting box, click on **Processing Extent** to set the Extent to "Same as layer EthiopiaBoundary_adm0". This will make it so all tools (Euc Dist, density) will output to this same square extent.

4. Scroll down further and click on **Raster Analysis.** Set the **Cell Size** to 3000. This means when tools are run, the resulting rasters will have a cell size of 3000 meters by 3000 meters. We know its meters because the projection is in **meters**!

5. Set the *Mask* to **EthiopiaBoundary_adm0**. This then *masks* (aka clips) the output rasters to the shape of Ethiopia.
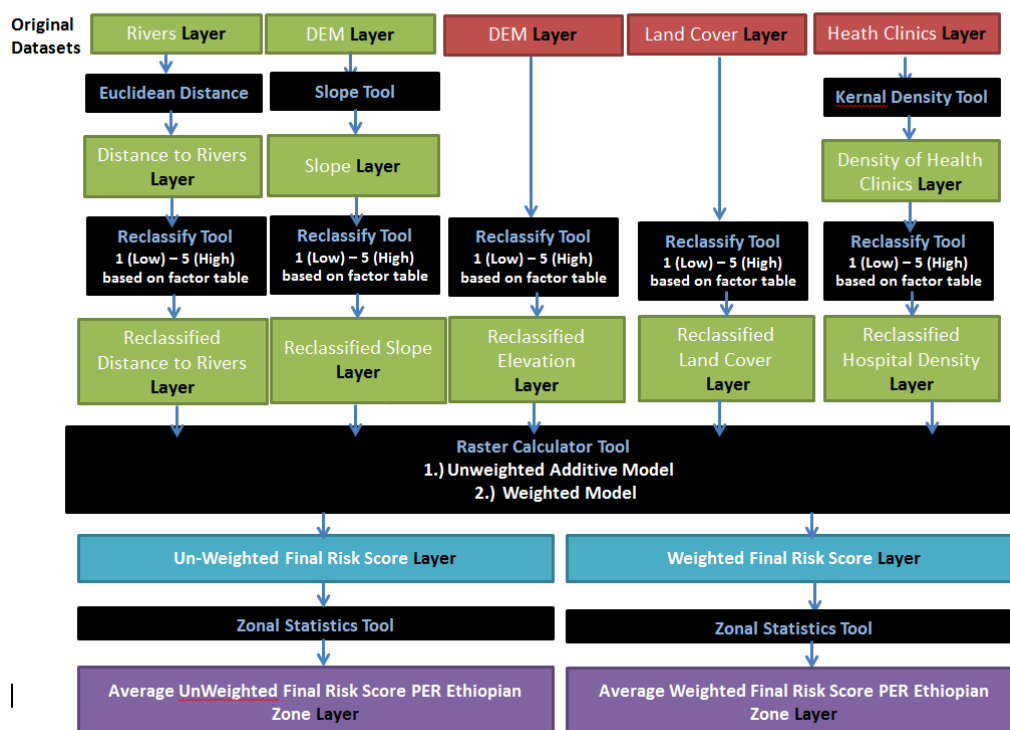

7. Click **OK** to exit back to the map.

Overview of the Analysis

Here are the analysis steps you will follow to determine risk of malaria transmission throughout Ethiopia:

1. **Rivers** → Euclidean Distance → Reclassify 1-5 for distance preference (close = High Risk (5))

2. **Elevation (DEM)** layer → Reclassify for elevation preferences (low elevations= High Risk (5))

3. **Elevation (DEM)** layer → Derive slope → Reclassify for slope preferences (low slope= High Risk (5))

4. **Land Cover** layer → Reclassify 1-5 for land cover preferences

5. **Health Centers** layer → Kernel Density → Reclassify 1-5 for density preferences (low density = High Risk (5))

6. Use Raster calculator to combine the 5 reclassified layers. Do so for un/weighted calculations.

7. Use Zonal Statistics to determine the average un/weighted risk per Malaria Zone (Admin2).

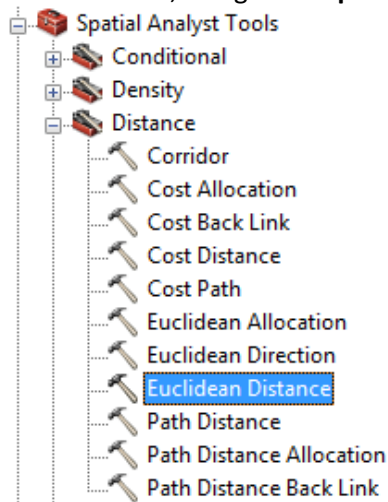8. **Here are the reclassification schemes you will use: NOTE that HIGH RISK (BAD) is a 5**

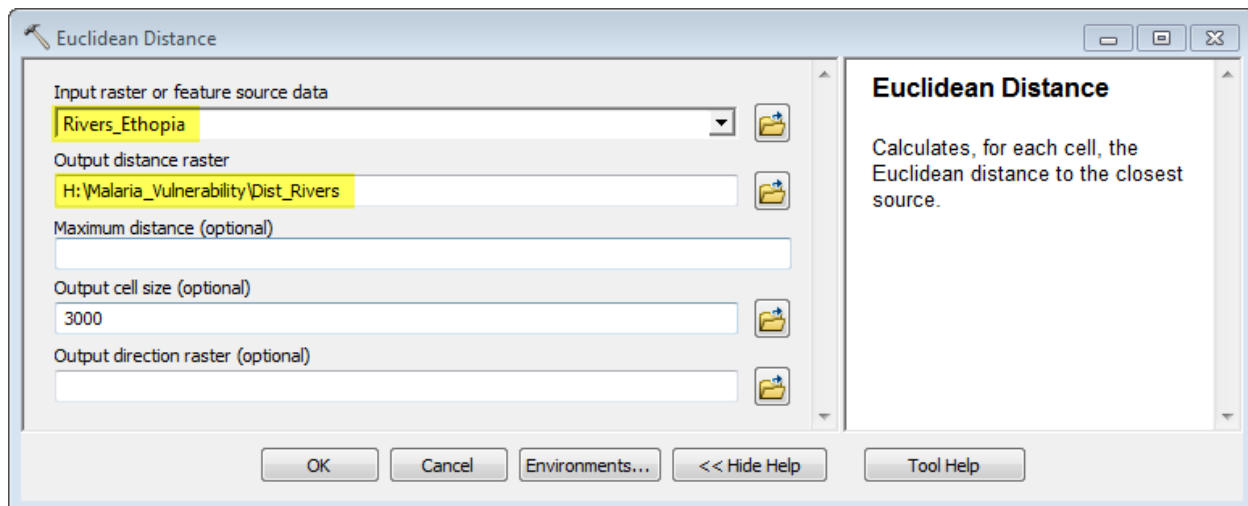# Calculating and Reclassifying Distance to Rivers

Distance to water sources, such as rivers, is a very important part of the analysis since malaria is a "water related" disease. This means it is transmitted via mosquitos who like to breed near waterbodies. Therefore, the closer one is to waterbodies, the higher the risk for contracting malaria.

1. In ArcToolbox, navigate to **Spatial Analyst Tools → Distance** and double click on **Euclidean Distance**.

2. Fill the tool out as follows.

   - Choose *Rivers_Ethiopia* as the *Input Raster.*

   - Set the *Output Raster* to your H: drive with the name **Dist_Rivers.**

   - Leave the rest of the defaults. You will notice that the cell size is properly set to 3000 because we set it in the Geoprocessing Environments already.

   - **Note:** it is VERY important that there are no spaces in the entire file path or name. Raster tools will fail if this is the case.

   After the process is completed, you will see the **Dist_River** add to your Table of Contents. The result is a raster grid, in which each cell is the distance from the *closest* river in **meters** (units determined by projection units).

3. In ArcToolbox, navigate to **Spatial Analyst Tools → Reclass** and double click on **Reclassify.**
   Now, we will use the **reclassify** tool to reclassify, or change the distance cell values to **1 to 5 to indicate low to high risk based of the distance to rivers** based on the following classification break values.

   Remember: These distance values we need to reclassify are in **meters**!

Distance Values	Reclassify Values (Ranking)
< 1 Mile (1609 m)	5 – Very High Risk
1 – 5 Miles (8046 m)	4 – High Risk
5-25 Miles (40233 m)	3 – Moderate Risk
25-75 Miles (120701 m)	2 – Low Risk
>75 Miles (leave the max)	1 – Very Low Risk

4. Set up the **Reclassify** tool as follows:

    - Set *input raster* to **Dist_Rivers** and keep the *Reclass field* as **Value.**

    - Click *Classify* to set the class values.

    - Set the number of *classes* to **5.**

    - Set the **Break Values** (located on the right) to 1609, 8046, 40233, 120701, and leave max (as broken down in table above). Press OK to exit out of the Classification box.

    - Click <u>Reverse New Values</u>. This is **IMPORTANT** because it gives values with a short distance to rivers (0-1609m aka 1 mile) a high risk score of 5, while longer distances (120701m aka > 75 miles) a low risk score of 1.

    - Name the output **Reclas_Rivers**. It is important to change the name because there will be a lot of reclass layers by the end of the tutorial.

The *Reclas_Rivers* will appear automatically in your table of contents, but the colors aren't helpful to interpretation.

5. Double click on the layer and then the **symbology** tab, assign a sequential color ramp – such as blue to red. In this case, make sure that red shows high risk areas. Note: if the colors are reversed when you change the color scheme, you can highlight the boxes and right click and then select "flip colors".
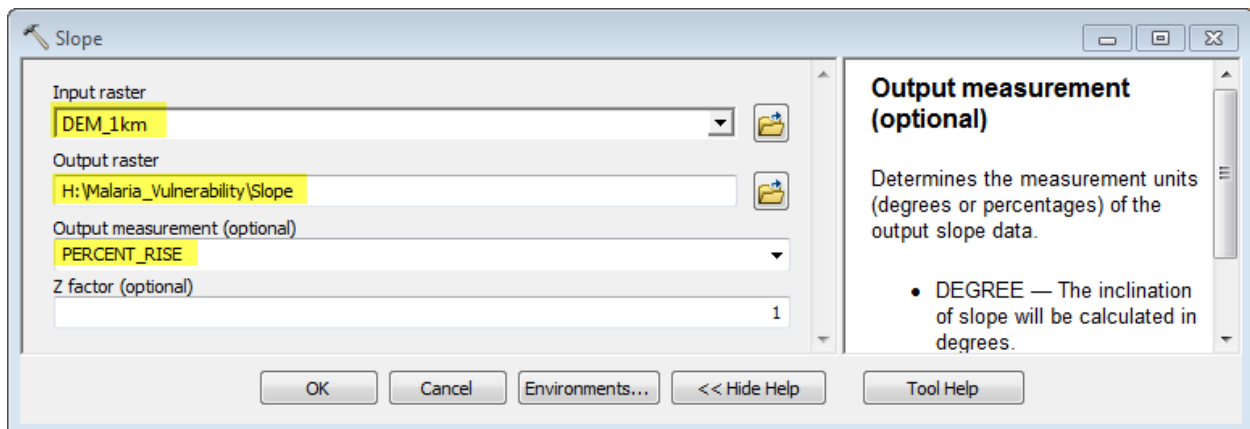

☑ Reclas_rivers
Value
■ 1
■ 2
□ 3
■ 4
■ 5

6. Save your **MalariaRiskAnaylsis.mxd.** mapfile in your H Drive

Calculating and Reclassifying Slope

Since stagnant water is a great breeding ground for mosquitos, low slopes are more likely to have higher risk of malaria because it allows water to pool. You'll make a slope map from the elevation data (DEM_1km), and then reclassify it into a slope preference grid.

1. Turn on **DEM_1km** and turn off all other layers except EthiopiaBoundary_adm0. Notice how the DEM extends outside the country boundary. Since we set the mask in the *Environments*, when we run the tool, the output raster will automatically be clipped to the country!

2. In ArcToolbox, navigate to **Spatial Analyst Tools → Surface** and double-click on **Slope**.

3. Fill out the dialog box as you see here. Be sure to give it the name *Slope*. You want a slope grid that shows **PERCENT_RISE**, not Degrees!
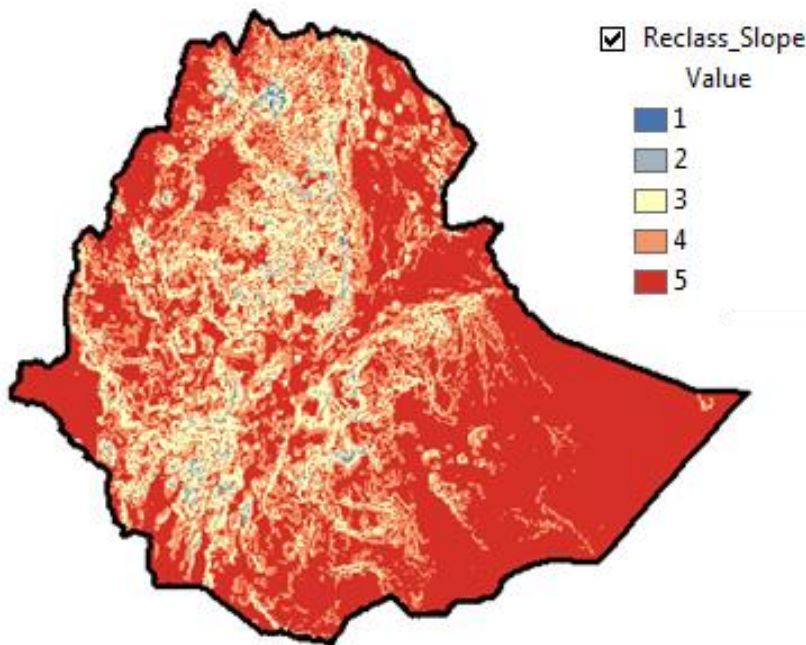

4. Click OK. Let the process run. The result is a slope grid where each cell is the percent slope.

5. Now, using the same steps as we did before when we reclassified distance to rivers, now **reclassify** your

slope using the break values below. Remember to **reverse your break values** so low slope = high risk (5). Save your output raster in your H drive and call it *Reclass_Slope.*

Slope Values	Reclassify Values (Ranking)
< 2%	5 – Very High Risk
2 – 5 %	4 – High Risk
5 – 12%	3 – Moderate Risk
12 – 20%	2 – Low Risk
>20%	1 – Very Low Risk

6. Change the symbology to the same blue to red color scheme for easy comparison. Your reclassified slope layer should look like the following.
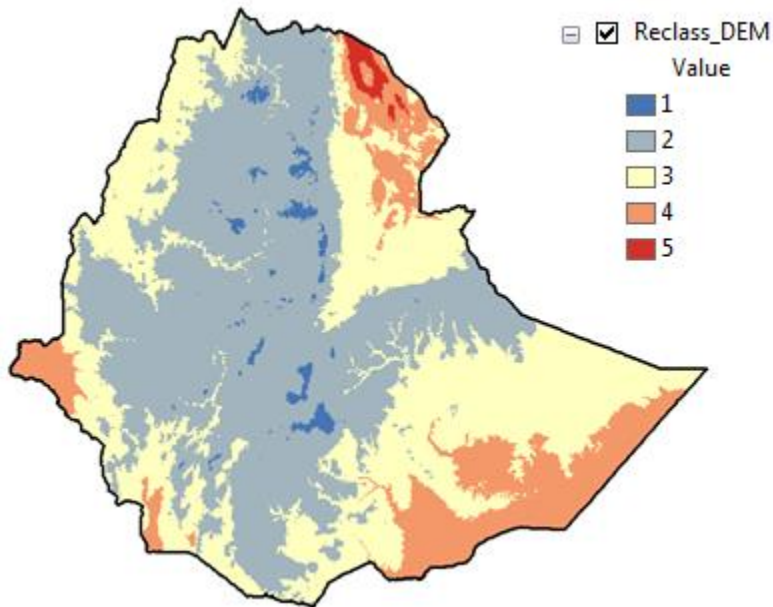

7. Save your map file.

Reclassifying Elevation

Elevation is an important factor when determining the overall risk for malaria transmission. Low elevations pose a higher risk because temperatures are generally warmer at lower elevations, resulting in a quicker life cycle and more breading for mosquitos. High elevations are often colder, resulting in fewer mosquitos.

1. Since we are interested in specific elevations, **the DEM_1km** (Digital Elevation Level) can be directly reclassified to reflect these specific ranges. We do not need to run any tool (such as Euclidean distance or slope) before reclassifying, since this layer already represents elevation.

2. We will **reclassify** the **Elevation** layer using the following break values (which are in meters).   Open the reclassify tool and use the same methods as shown in the previous steps to change the break values. **Remember to reverse your break values** so low elevations = high risk (5) and high elevations = low risk (1). Save your output raster in your H drive and call it *Reclass_DEM.*

Elevation Values	Reclassify Values (ranking)
<0 m	5 – Very high risk
0 - 500m	4
500 - 1200m	3
1200 - 3000m	2
>3000 m	1 – Very low risk

3. Change the symbology to the same blue to red color scheme for easy comparison. Your reclassified DEM layer should be clipped and look like the following:
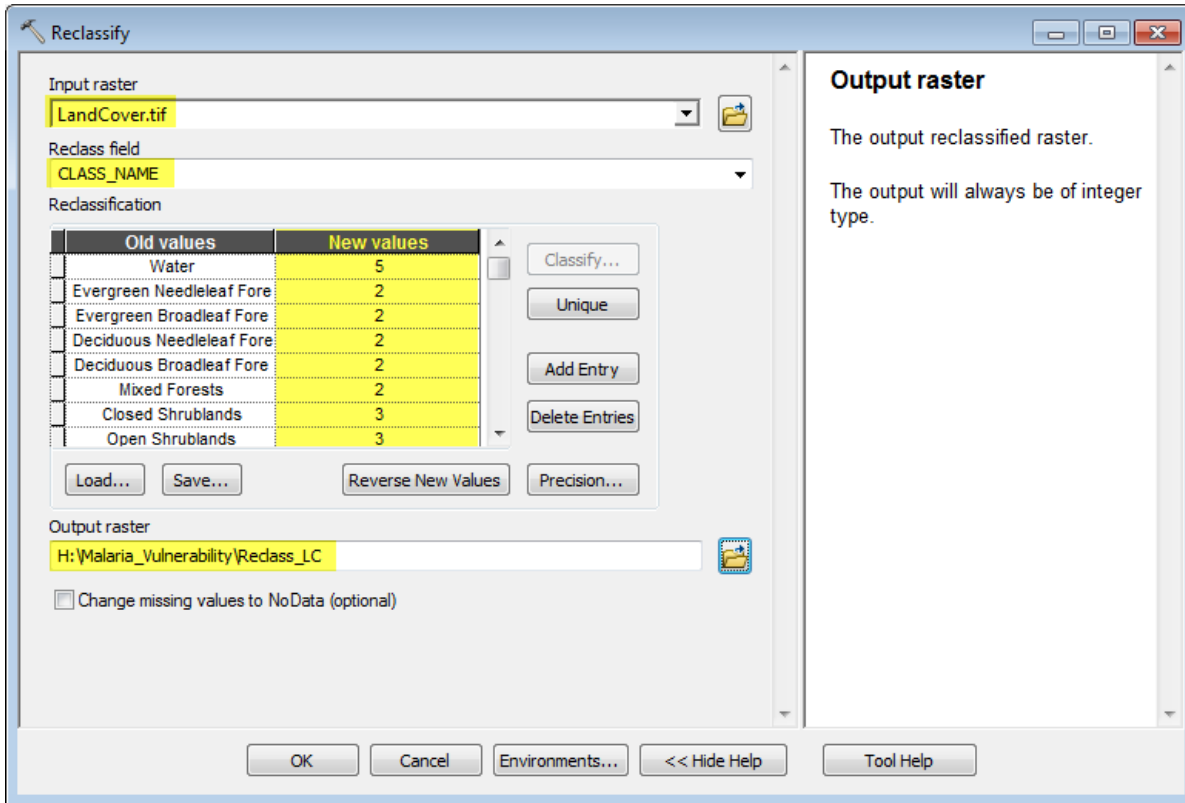

Reclassifying Land Cover

Mosquitos also thrive in specific land covers (moist areas, such as wetlands), but die off quickly in others (deserts or barren land). Therefore, the type of land cover will certainly affect the risk for malaria transmission.
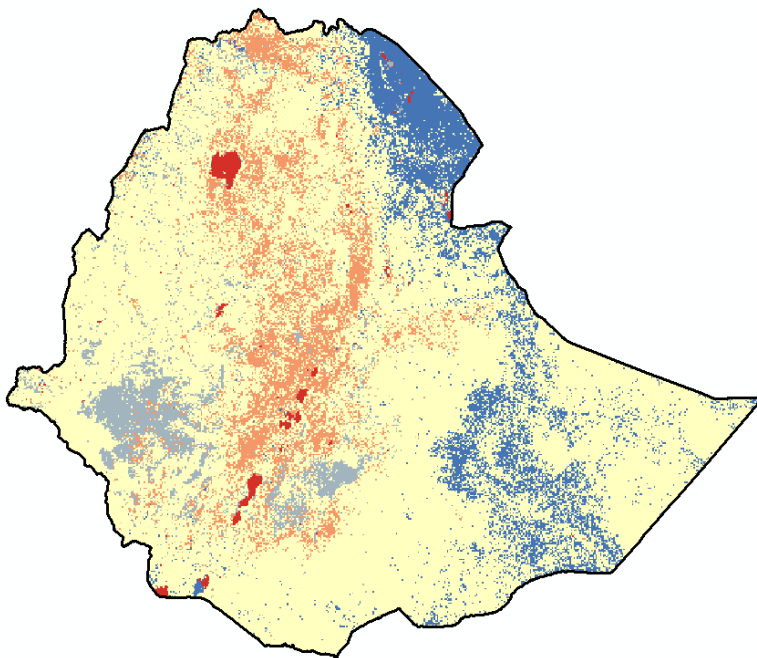
1. Similarly to elevation, **LandCover** can be directly **reclassified** without the need to run any other tool beforehand. However, since land cover is *categorical* data, and not *numerical*, the process of reclassifying is a bit different.

2. Open the reclassify tool. We will use the following land cover categories and reclassify them into their appropriate ranking.

Land Cover Types	Reclassify Values (ranking)
Water and All Wetlands	5 – Very high risk
Croplands and Natural Vegetation	4
All Shrublands, All Grasslands, and all Savannas	3
All Forests, Urban/Built up	2
Snow/Ice, Barren	1 – Very low risk

3. In the input field, select **Land Cover** and set the Reclass field as **CLASS_NAME.**

4. Since this is "categorical" data, you will notice the classify button is grayed out. We will directly type reclassified values for each land cover right into the **Reclassification Table.** Use the above Land Cover Types and associated values and keep no data as nodata.

5. Name the new layer **Reclass_LC.** The tool set up should look like the following. Look at the table above for the values. :
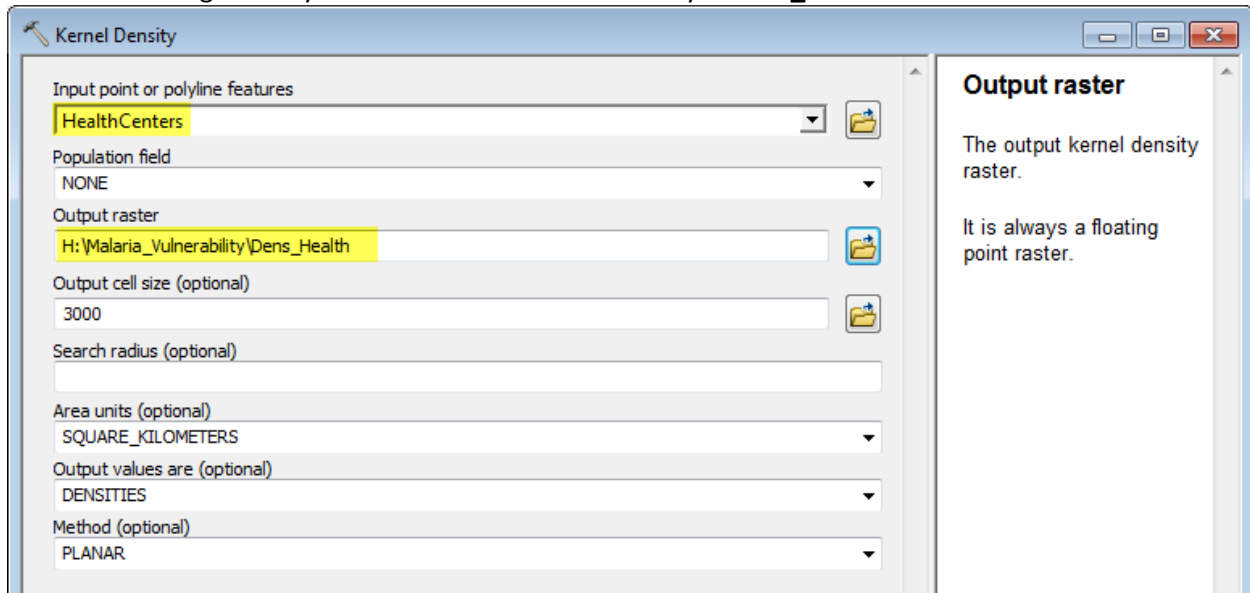

Notice how the new reclassified land cover raster is clipped to Ethiopia. That's because we set the environments! Your map should now look like this:
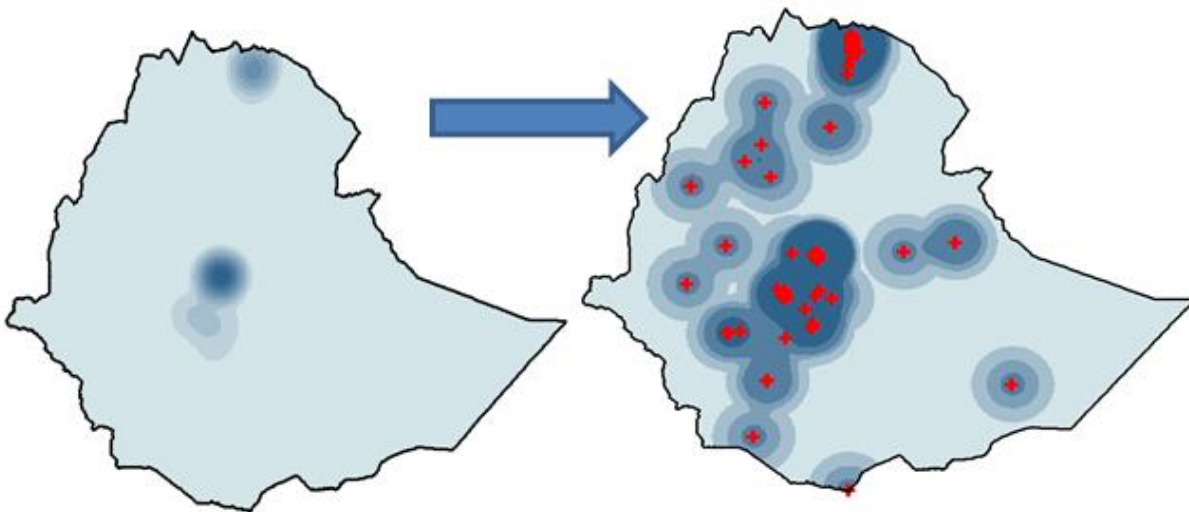
Creating Density Layer of Health Clinics

If not treated, malaria can quickly become life-threatening. Therefore access to health clinics and hospitals in the area would greatly affect the populations' *vulnerability.* We will use the **kernel density tool** to create a raster layer that represents access to health care. We will use the kernel density tool instead of Euclidean distance because areas with many hospitals/health clinics will likely be less vulnerable than areas with only 1 health clinic. In this case, the *number* or *density* of health clinics will model access to health care.

1. Turn on **HealthCenters** and turn off the other layers so you can see it.

2. In ArcToolbox, navigate to **Spatial Analyst Tools → Density → Kernel Density.**

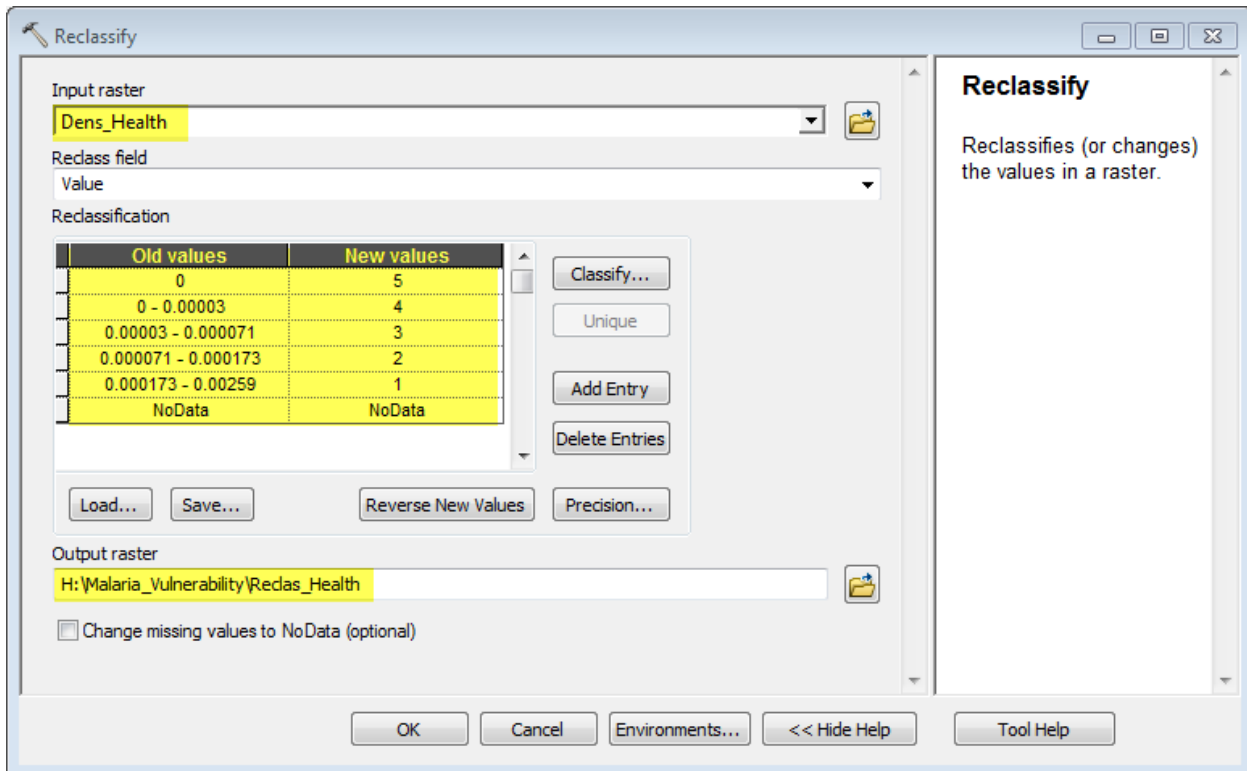3. Fill out the dialog box as you see here and call the new layer **Dens_Health**

4. The new density layer will be automatically added to your table of contents and map. Open the symbology of this new raster.

5. Notice how the default output has just two main clusters. This is because the classification scheme is set to **Equal Interval.** In symbology, click on the Classify button and change the classification scheme to **Quantile and the number of classes to 5.** This will visually represent the density of our health clinics more appropriately. This will also be the classification scheme we use to reclassify the data.
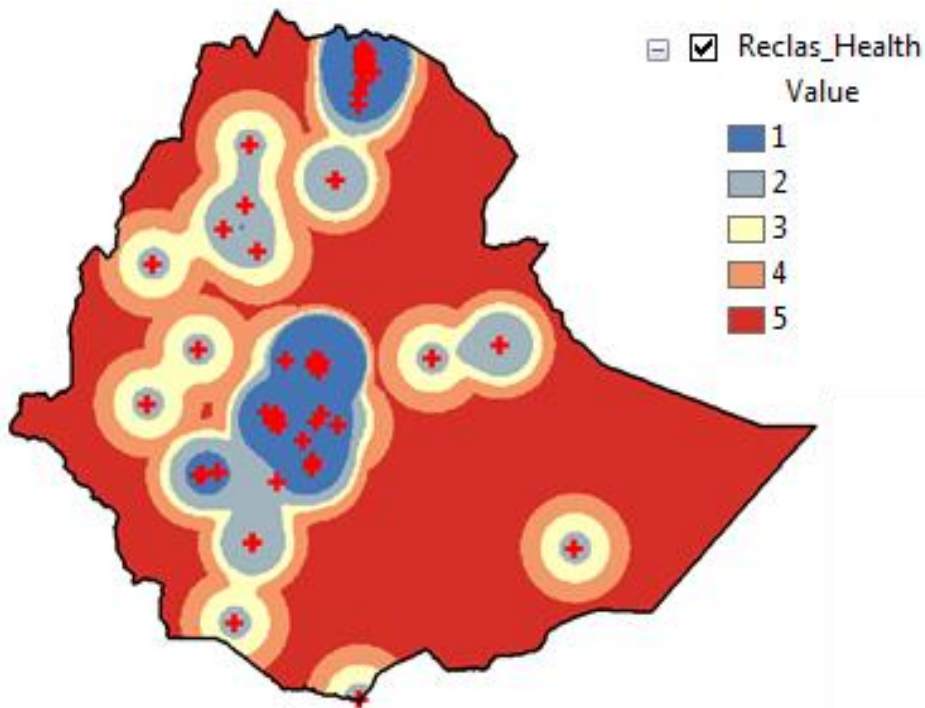

6. Now, use the **reclassify** tool to rank the density of the health clinics using the **Quantile** break values. Remember, low density of health clinics = High risk and high density of health clinics = low risk. **So remember to reverse your values.**

10

7. Before running the tool, make sure your values match the following tool set up. Name your new layer **Reclas_Health.**

8. Change the symbology to the same blue to red color scheme for easy comparison. Your reclassified Health Center Density layer should look like the following:

9. Save your mxd again.

10. Now we should have the following **reclassified** rasters: Distance to rivers, slope, elevation, land cover, and density of health clinics.

11

Using Raster Calculator to Create a Risk/Vulnerability Assessment
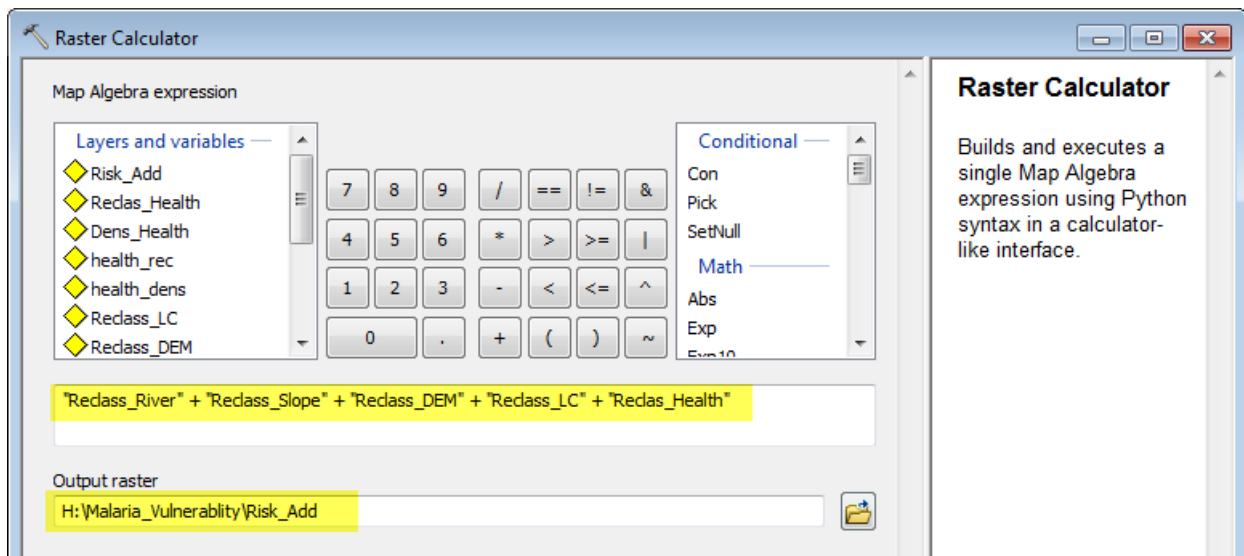
Now that we have all the individual **factors** reclassified using the same ranking scale, 1(low risk) to 5 (high risk), we can use the **raster calculator** to combine the layers to create a final risk assessment score.

We will create both an **unweighted** Risk Assessment score and a **weighted** Risk Assessment score.
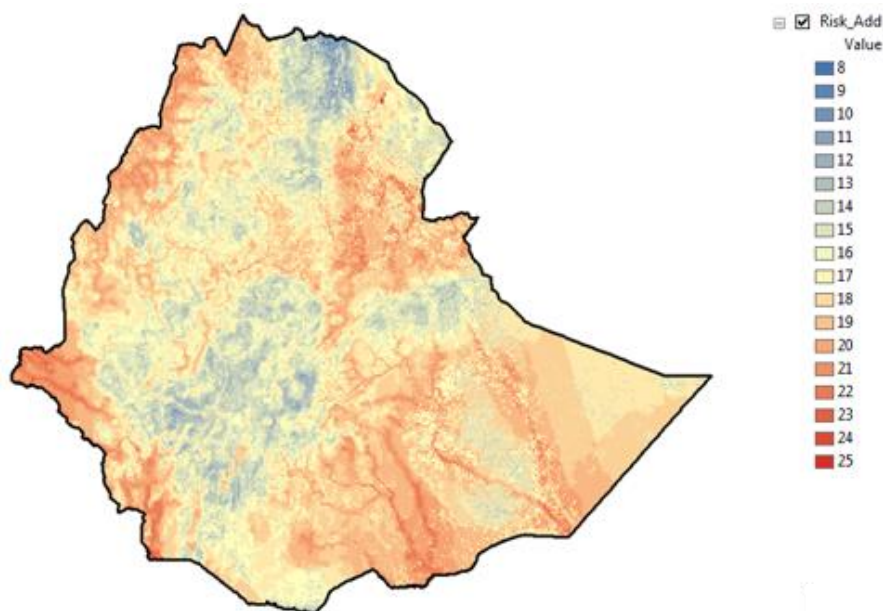
Creating a Unweighted Vulnerability Score

We will use a tool called **Raster Calculator** to combine our reclassified grids.

1. In ArcToolbox, navigate to **Spatial Analyst Tools → Map Algebra** and open **Raster Calculator**.

2. Follow the diagram below to create an "**Additive Model**". In this type of unweighted model, all values are ranked equally by simply adding them together. Therefore, the final raster will have values that range from 5 (lowest possible risk) to 25 (highest possible risk).

3. Fill out the **raster calculator** as follows and call the unweighted raster **Risk_Add.**

4. Change the symbology to match the same blue to red color scheme. Your final unweighted Vulnerability Assessment should look like the following map, with values ranging from 8 to 25.
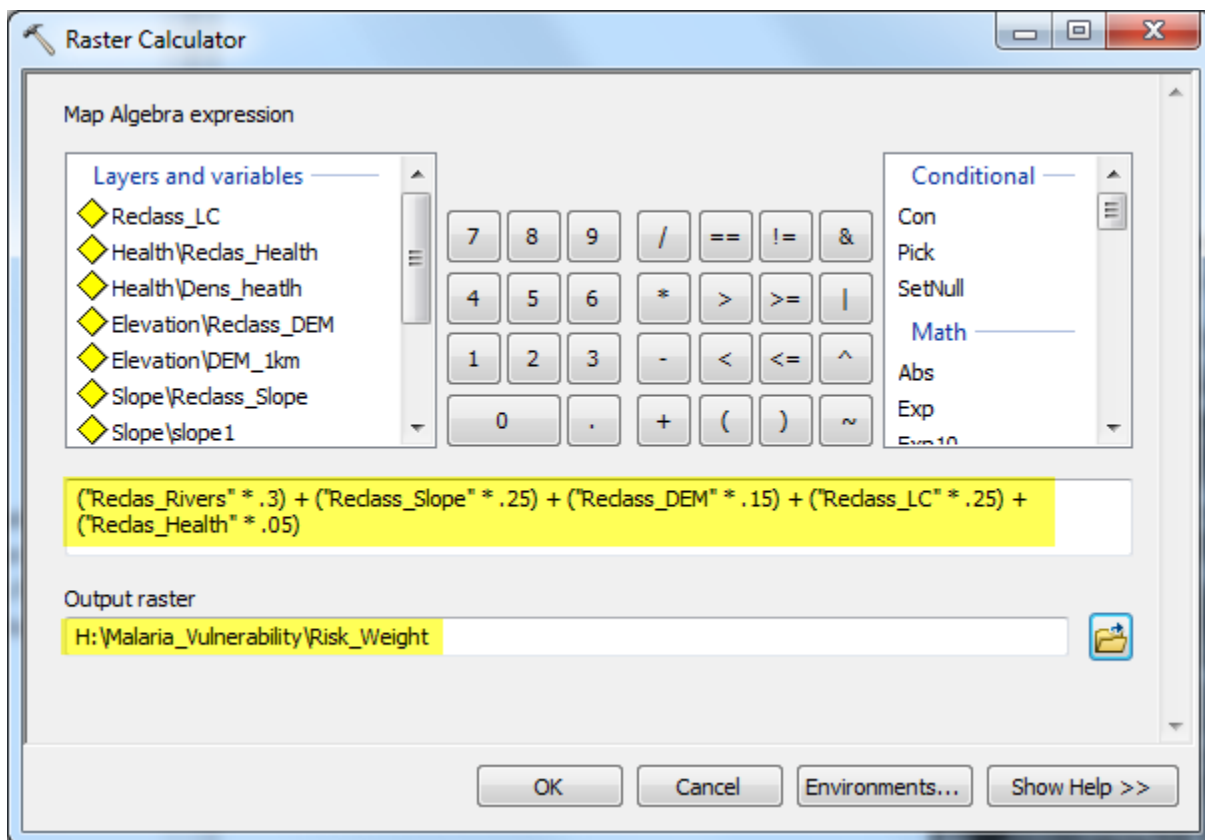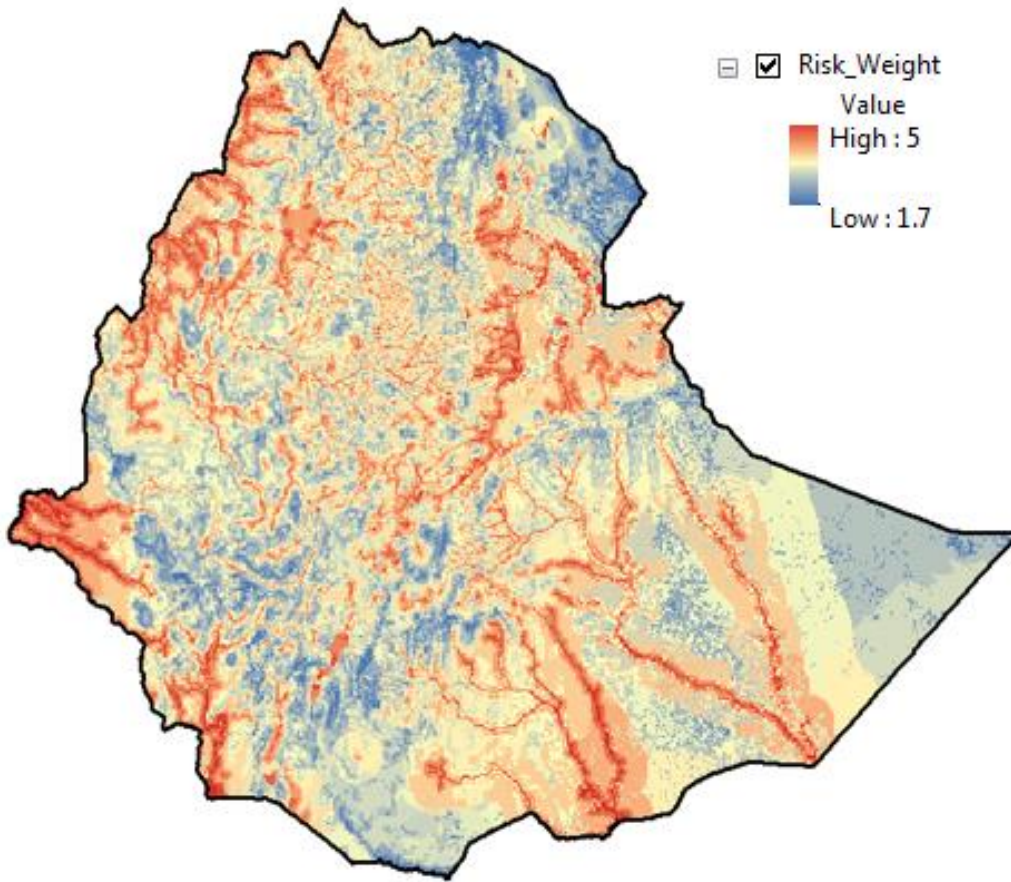
Creating a Weighted Vulnerability Score

If one of your "factors" is more important in a given analysis, we can create what's called a weighted **grid**. We can use the **Raster Calculator** again to give different weights to each layer and create a weighted grid –as follows:

- Distance to Rivers is the most important factor (weighted 30%)
- Slope is the next most important factor (weighted 25%)
- Land cover is an equally important factor (weighted 25%)
- Elevation is somewhat less important (weighted 15%)
- Hospitals are least important (5%)

1. In ArcToolbox, navigate to **Spatial Analyst Tools → Map Algebra** and open **Raster Calculator**.

2. Follow the diagram below to create a "**Weighted Model**". In this type of weighted model, all values multiplied by their weight percentages and then added together. With this type of raster, the highest possible value will be a 5 and the lowest possible value will be a 1.

3. Fill out the **raster calculator** as follows and call it Risk_Weight.
   Note: Always use the "buttons" to add to the expression box!

4. The result is a grid where the values have a *continuous* value from 1.7 to 5. Change the color ramp to the blue to red one we have been using for consistency. It should look something like this:
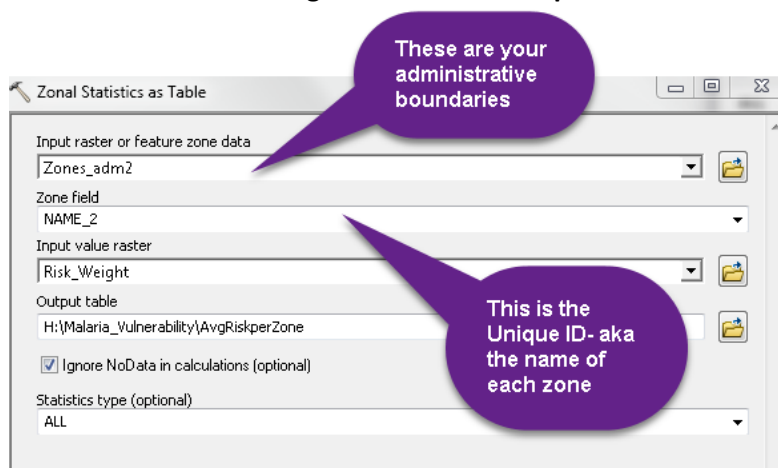
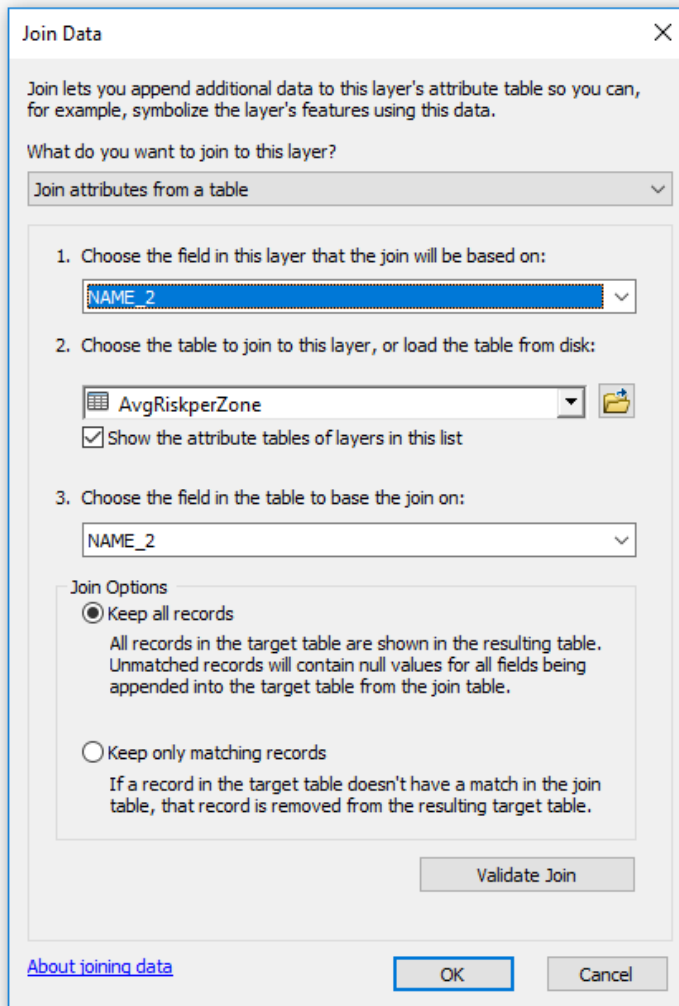## Using Zonal Statistics to Determine Average Risk per Zone.

While the risk scores might be helpful for determining specific locations with high risk, public health officials might like to know which zones (similar to provinces) within Ethiopia are most at risk. Therefore, these administrative boundaries called **zones** can allocate funds towards assisting with malaria prevention.

The purpose of **Zonal Statistics** is to get an *average score per zone.* The tool looks for "zones" (aka some sort of administrative boundary typically), but it is just ironic that the administrative 2 boundary in Ethiopia just happens to be called "zones".
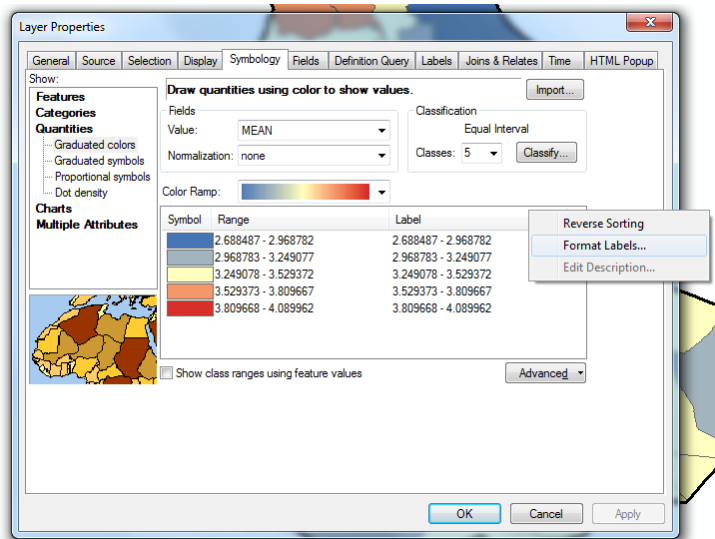
1. In ArcToolbox, navigate to **Spatial Analyst Tools → Zonal** and open **Zonal Statistics as a Table**. You can do this as a table and then join it to your zones, or just as Zonal Statistics (but then you only have 1 Statistic option).

2. We will look at the **average WEIGHTED score per zone**. Fill out the tools as follows:
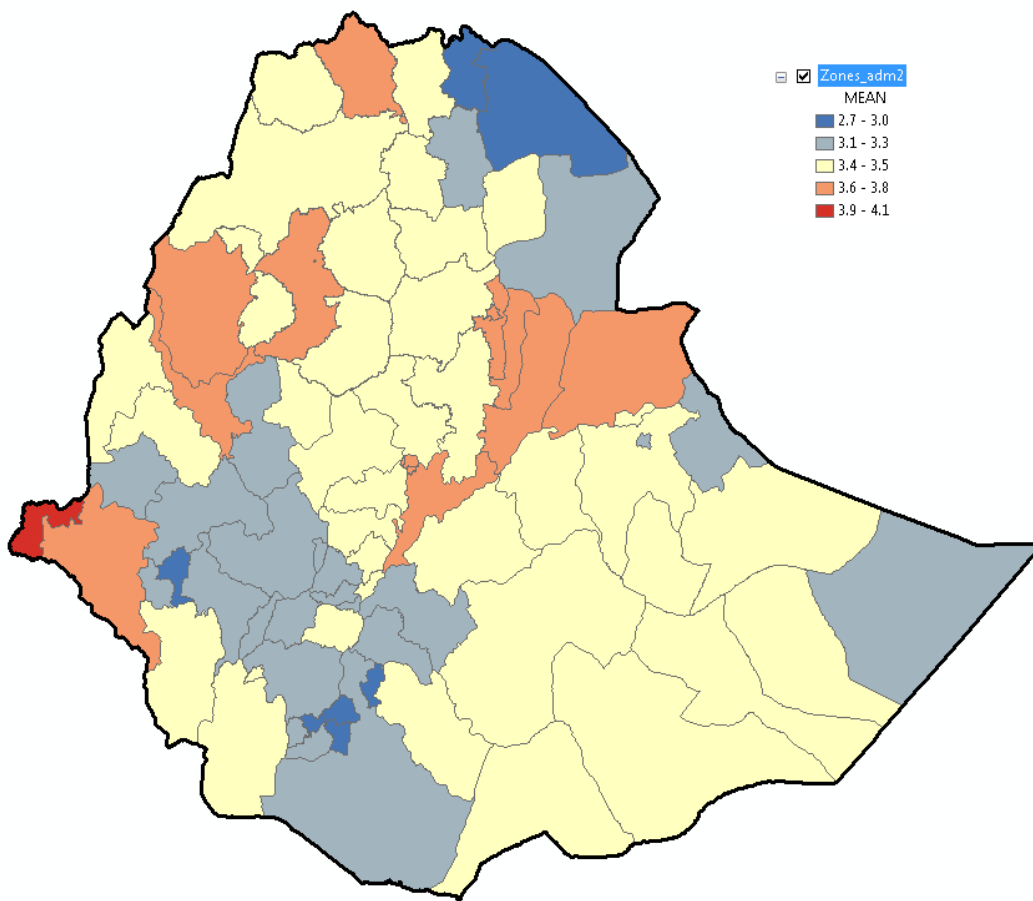
3. Your output will be a table showing various statistics about the risk scores **per** Ethiopia Zone. Now we must **join** this table to the **Zones_adm2** Shapefile.

4. Right click on **Zones_adm2** and select **Joins and Relates…** and then **Join.** Fill out the tool as follows:

5. Say yes when asked if you want to index the data.

6. Now, when you examine the Zones attribute table, all of the statistic information is added to the end of the table. Let's **symbologize** the Mean Score per Zone. Open the **symbology** properties for the **Zones_adm2** layer.

7. Select Quantities and choose **Mean** as the value field. In **Classify**, choose 5 classes and **equal interval.** Then press ok.

8. Since this is an average, there are a lot of decimal places linked to these average risk scores. In symbology, click on **Label → Format Labels…**. Then, change the number of decimal places to 1 and press ok.

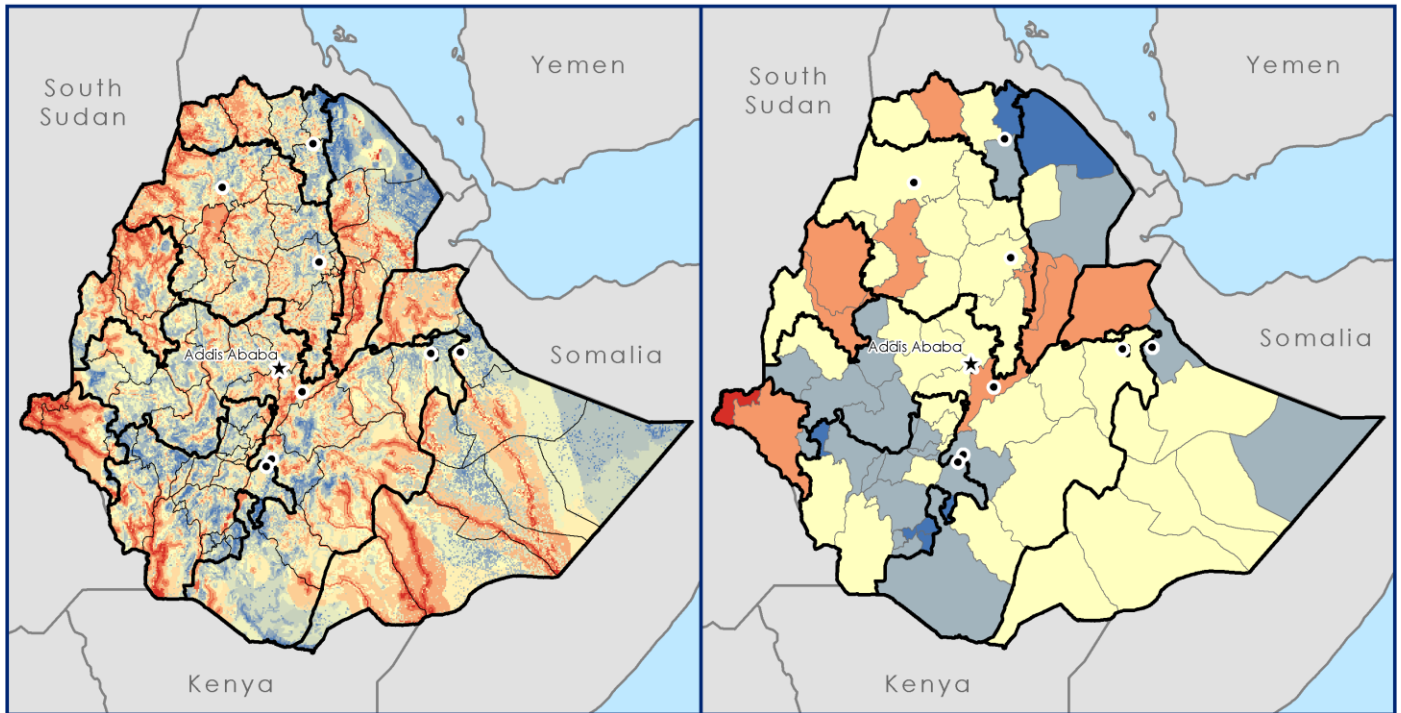9. Your final Risk Score per Zone map should look like this:

10. Save your map.

11. That's it! Now we know where the high and low risk for malaria is in Ethiopia with/out weighting and which zones on a whole are at the highest risk. Think about what other factors you could add in to this analysis to make it more robust.

Check below to see a final map of both the Weighted Risk Assessment and the Average Risk per Zone!

Risk Assessment for Malaria Transmission in Ethiopia | 2015
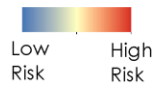
Weighted Risk Assessment

Average Risk per Zone

Risk Assessment Criteria
1.) Proximity to rivers
2.) Elevation
3.) Land Cover
4.) Slope
5.) Density of Health Centers

Data Sources: GADM, SRTM, ESRI

Carolyn Talmadge | GIS for Conservation Medicine | October 12, 2018

Risk Assessment Score

Low Risk — High Risk

• Large Cities
★ Capital
☐ Zones
☐ Regions

Average Risk per Zone

2.7 - 3.0   3.1 - 3.3   3.4 - 3.5   3.6 - 3.8   3.9 - 4.1

Kilometers
0   250   500

NORTH