USING SPATIAL INFORMATION TO SUPPORT DECISIONS ON SAFEGUARDS AND MULTIPLE BENEFITS FOR REDD+

STEP-BY-STEP TUTORIAL VERSION 1.1: BUILDING SPATIAL WORKFLOWS TO HELP IDENTIFY POTENTIAL AREAS FOR UNDERTAKING A REDD+ INTERVENTION USING MODEL BUILDER IN ARCGIS 10.X

PLUS ADDITIONAL ANNEX CONTAINING GUIDANCE NOTES ON GEOPROCESSING TOOLS AND TECHNIQUES WITHIN ARCGIS, COMPARING RASTER AND VECTOR DATA ANALYSIS

UN-REDD PROGRAMME

Food and Agriculture Organization of the United Nations

UNDP Empowered lives. Resilient nations.

UN environment

The UN-REDD Programme is the United Nations Collaborative initiative on Reducing Emissions from Deforestation and forest Degradation (REDD) in developing countries. The Programme was launched in September 2008 to assist developing countries prepare and implement national REDD+ strategies, and builds on the convening power and expertise of the Food and Agriculture Organization of the United Nations (FAO), the United Nations Development Programme (UNDP) and UN Environment.

The UN Environment World Conservation Monitoring Centre (UNEP-WCMC) is the specialist biodiversity assessment centre of UN Environment, the world's foremost intergovernmental environmental organisation. The Centre has been in operation for over 35 years, combining scientific research with practical policy advice.

Prepared by Corinna Ravilious, Charlotte Hicks and Simon Blyth

We welcome comments on any errors or issues. Should readers wish to comment on this document, they are encouraged to get in touch via: ccb@unep-wcmc.org.

Citation: Ravilious, C., Hicks, C. and Blyth, S. (2017) Using spatial information to support decisions on safeguards and multiple benefits for REDD+. Step by Step Tutorial Version 1.1: Building spatial workflows to help identify potential areas for undertaking a REDD+ intervention using model builder in ArcGIS 10.x. Prepared on behalf of the UN-REDD Programme. UNEP World Conservation Monitoring Centre, Cambridge, UK.

Acknowledgements: These training materials have been produced from materials developed for joint working with the UN-REDD Viet Nam Phase II Programme to inform spatial analysis, supporting provincial REDD+ action planning in Viet Nam.

Contents

1. Introduction

REDD+ is a voluntary climate change mitigation approach that has been developed by Parties to the UNFCCC. It aims to incentivize developing countries to reduce emissions from deforestation and forest degradation, conserve forest carbon stocks, sustainably manage forests and enhance forest carbon stocks. This will involve changing the ways in which forests are used and managed, and may require many different actions, such as protecting forests from fire or illegal logging, or rehabilitating degraded forest areas.

REDD+ has the potential to deliver multiple benefits beyond carbon. For example, it can promote biodiversity conservation and secure ecosystem services from forests such as water regulation, erosion control and non-timber forest products (NTFPs). Some of the potential benefits from REDD+, such as biodiversity conservation, can be enhanced through identifying areas where REDD+ actions might have the greatest impact using spatial analysis and other approaches.

The purpose of this tutorial series is to help participants in technical working sessions, who are already skilled in GIS, to undertake analyses that are relevant to REDD+. The tutorials have been used to build capacity in a number of countries to produce datasets and maps relevant to their spatial planning for REDD+, and to develop such map products. Maps developed using these approaches appear in a number of publications whose aim is to support planning of strategy options that enhance biodiversity and ecosystem services as well as delivering climate change mitigation (see http://bit.ly/mbs-redd for country materials). There is of course no requirement for countries to use the approaches described in these tutorials.

This tutorial focuses on how to use spatial analysis workflows to help identify potential areas where a REDD+ intervention (also known as REDD+ action, or REDD+ measure) could be undertaken to address a specific driver or barrier. Specific interventions will be suitable for addressing specific drivers or sets of drivers. Where countries have identified biodiversity conservation as a goal for REDD+, and to be consistent with the Cancun safeguards for REDD+ on protecting biodiversity, it is useful to identify areas where specific REDD+ interventions are feasible and can protect threatened species. It may also be useful to identify areas outside forest where threatened species may be vulnerable to the displacement of land-use change pressures or to afforestation.

Spatial analysis can be used to support decision-making in order to identify potential areas for appropriate REDD+ interventions (REDD+ actions), but maps alone are not sufficient for making decisions. This does not mean that all potential interventions can be usefully mapped, but map analysis can be useful alongside methods such as stakeholder consultation and participatory mapping approaches in selecting what REDD+ approaches to apply, and where.

This tutorial uses an imaginary land area to demonstrate how to create and run an analysis using workflows in ArcGIS Model Builder. It covers: mapping the drivers of deforestation and forest degradation, and barriers to the '+' activities; mapping REDD+ interventions; defining spatial logic and creating workflows; sharing the resulting model; and fixing a broken model. The annexes provide more background information on raster data formats and processing for those who are more familiar with vector GIS data, and will guide users in exploring the various raster analysis tools and methods that are available within the ArcGIS Geoprocessing Toolbox.

Annex 1 provides a brief summary of the differences between vector and raster data and an introduction to getting started with raster data analysis. Raster analysis provides users with access to a large number of tools and functions that are not available for vector analysis. Many of these tools are particularly suited to multi-criteria analysis. Raster functions can also be used to generate new data (e.g. hydrology and slope data from digital elevation models).

Annex 2 provides links to additional resources and guidance related to ArcGIS Model Builder.

A similar tutorial that uses QGIS as an open-source alternative platform is available from http://bit.ly/GIStools-redd:
Ravilious, C., Hicks, C. and Blyth, S. (2017) Using spatial information to support decisions on safeguards and multiple benefits for REDD+. Step by Step Tutorial Version 1.0: Building spatial workflows to help identify potential areas for undertaking a REDD+ intervention using the graphical modeler in QGIS 2.14.x. Prepared on behalf of the UN-REDD Programme. UNEP World Conservation Monitoring Centre, Cambridge, UK.

2. Mapping drivers of deforestation, forest degradation and barriers

When thinking about drivers of deforestation and forest degradation, and barriers to the 'plus' activities (conservation, enhancement, sustainable management), we often refer to the current situation (e.g. what are the drivers now and where forest cover change has already happened). But it is also important to consider future pressures and threats on forests, and how they may lead to deforestation or forest degradation, as well as future barriers.

To identify locations for REDD+ interventions, consideration must be given to both location of current drivers and barriers, and the pressures and threats on forests, to help identify where drivers or barriers may be found in the future.

Drivers and barriers may be direct or indirect. Examples of **direct drivers** include expansion of infrastructure, agricultural expansion, fire, mining activities and expansion of plantations (e.g. rubber, oil palm). A direct barrier could be physical features (steep slopes, rocky terrain, distance from roads) that prevent access for reforestation. Examples of **indirect drivers and barriers** include changes in population size and density (perhaps indicating growing demand for land and natural resources), poverty levels (perhaps indicating direct dependence on natural resources), financial incentives (e.g. commodity prices and subsidies that may make certain land-uses more desirable), cultural preferences (perhaps defining how natural resources are used) and political decisions (determining how land-uses are distributed and controlled).

Drivers of change in forest cover/quality may lead to changes in forest functions, levels of biodiversity, provision of ecosystem services and support for forest-based livelihoods.

To identify and map areas affected by or at risk of drivers and barriers we can:

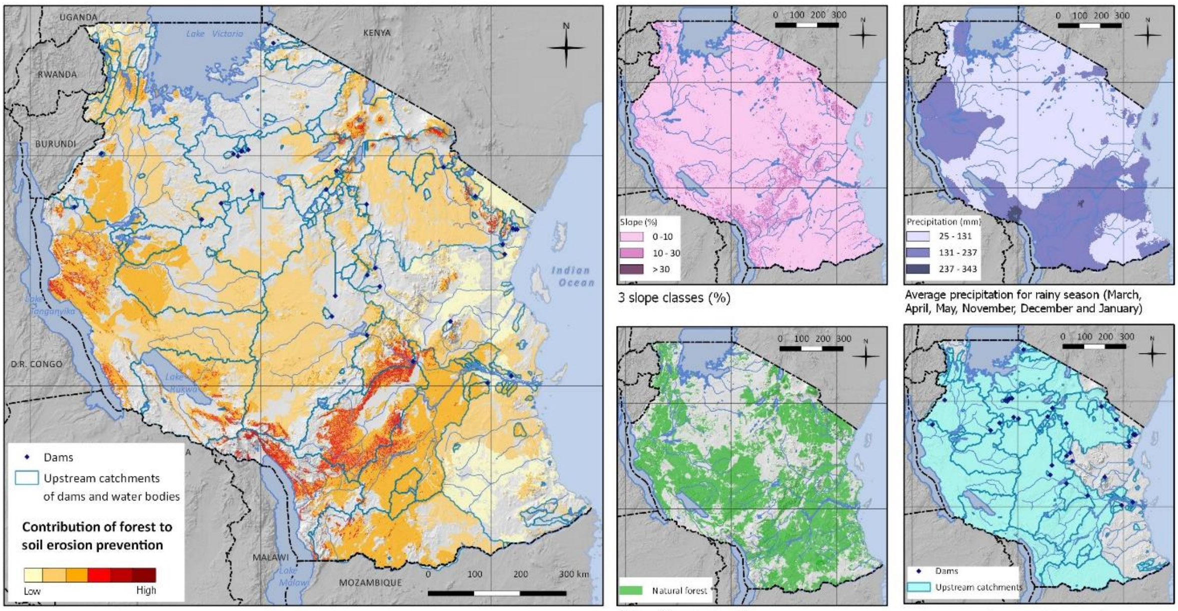
- Use a participatory approach, involving multiple stakeholders from different sectors, organizations and communities with local knowledge to identify drivers, barriers and the areas affected or at risk.
- Map current direct and indirect pressures on forest, examine their relationship with forest cover and forest-cover change, and explore how these may change in the future.
- Map areas that are likely to be affected by particular drivers or barriers in the future (e.g. planned deforestation due to the proposed development of infrastructure or changes to land-use designations).
- Map current areas where conservation, enhancement and sustainable management activities are already being undertaken and examine the relationship between factors that are either hindering or promoting their success.
- Map areas with physical, environmental, and socio-economic potential for the promotion of 'plus' activities.

Maps showing individual pressures/future threats layers and maps using simple overlays can be created, to show where different factors coincide. This does not 'select' particular areas, but these maps can be important inputs into multistakeholder workshops, where participants identify areas suitable for REDD+ interventions based on their expert and local knowledge.

Maps showing areas meeting certain criteria can also be created. For example, a map showing areas at risk of small-scale agricultural expansion may ‘select’ areas based on the presence of existing agriculture, areas designated for future agriculture through the land-use plan, and socio-economic factors such as poverty and population.

This approach can be more subjective in terms of the criteria and thresholds used and needs to be validated by expert and local knowledge. If the knowledge needed to inform the workflow of the map is lacking, or the workflow is not properly validated, it can lead to misinformed use of data. Assumptions should be transparently presented so that planners and stakeholders understand the data that were used to create the map and how they influence the areas that were selected by the spatial analysis process. One way to do this is to show the input layers or assumptions as a series of maps in an annex, so it is clear how the areas were selected, what input layers were used and what criteria were applied. Alternatively, input maps can be presented alongside the final output, as in the example map below, where the input layers are presented on the right-hand side of the final output map.

This map shows areas where forests are particularly important for limiting soil erosion that might cause sedimentation problems for dams in Tanzania. The methodology is based on four parameters: slope, precipitation, locations of dams and water bodies and their catchments, and forest.



Map projection: WGS84 / UTM Zone 36S. Map prepared by Tanzanian Forest Service (TFS), UNEP-WCMC, FAO, Sokoine University of Agriculture (SUA) and Forestry Training Institute (FTI). Date: May 2013.

Natural forest extent*

* Defined according to the national forest definition to include the following land cover classes: montane and lowland forest, open and closed woodland, mangroves, thickets and bushlands.

Upstream catchments of dams and water bodies

Data Sources:
Natural forest and water bodies: NAFORMA. 2013. NAFORMA land-use / land-cover Map 2010.
Dams: Dr. Mark Mulligan, Department of Geography, Kings College, London.
Slope: generated from Lehner, B., Verdin, K., Jarvis, A. 2008: New global hydrography derived from spaceborne elevation data. Eos, Transactions American Geophysical Union, 89 (10) (2008) 93-94. Journal Article.
Precipitation: Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. 2005. Jarvis. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25 (2005): 1965-1978. Journal Article.

3. Mapping REDD+ interventions to address a driver or barrier

Both participatory and spatial analysis approaches may play a role in mapping priority areas for the implementation of REDD+ interventions. During the Provincial REDD+ Action Plan (PRAP) development process in Viet Nam for example, both approaches were used, with workshop participants first indicating areas for particular interventions, and spatial analysis teams conducting further analysis in GIS.

For the spatial analysis approach, two main steps are required:

1. **Defining the spatial logic or a workflow to answer a specified question**. The workflow will outline what datasets and geoprocessing tools needed to produce the desired output map. This workflow should also take into account a) the area affected by/at risk of the driver/barrier, and b) the results of the participatory mapping of interventions.
2. **Putting the workflow into ArcGIS model-builder**. By putting the geoprocessing steps into a single tool, the analysis can be run as a single step, or in fewer steps.

4. Defining spatial logic and creating a workflow (using a REDD+ intervention example)

Defining **workflows** helps you to think about **how you are going to undertake a piece of analysis**: the spatial logic, the technical GIS processes, and the sequence of steps.

This section of the tutorial focuses on the thinking and preparation required before building a workflow in ArcGIS Model Builder. To help identify potential locations for a particular REDD+ intervention, a workflow would combine factors in multi-criteria analysis, so that we **include** areas **suitable** for the intervention and **exclude** areas that are **unsuitable**.

Considering the questions below can help to clarify the input layers, other data and processes you will need in your workflow:

- Where are the areas affected by/at risk from the drivers or barriers? The location of interventions should be informed by the location of drivers/barriers; this may be based on previous maps developed to show drivers/barriers, or on other knowledge about which areas are affected. This could be a large proportion of the land area, for example all forest in a province may be affected by a barrier such as lack of monitoring resources.
- What physical aspects will affect the implementation of the intervention? For example, slope, soil type, forest type, climate.
- What other aspects may affect its feasibility? For example, accessibility of sites, carbon stocks, forest condition, or risks to the success of the intervention (e.g. fire risk)
- What is the potential to enhance benefits from the intervention? For example, can it contribute to poverty reduction, biodiversity conservation, or ecosystem services provision?
- What social and environmental risks are associated with the intervention? What is the potential to reduce risks and support safeguards, for example prevent the conversion of natural forest or reduce the risk of displacement (i.e. reducing the risk of moving deforestation or degradation to another area)?

In the example of Viet Nam, much relevant information should have been gathered during the PRAP process. For instance, previous analysis, and participatory mapping, of drivers and barriers can help identify the driver/barrier area. Meanwhile, the social and environmental risks and benefits assessment should help to identify the risks and benefits that could be associated with each intervention.

The identification of areas suitable for specific interventions can be complementary to, and used alongside other, participatory methods with stakeholder involvement. For example, if a participatory mapping approach has been undertaken to define potential areas, is any further analysis needed? This may vary between interventions. Participatory maps may give a broad picture, for example they may identify areas at the administrative unit level (e.g. the commune) and provide some of the criteria that could feed into a more detailed spatial analysis if required to complement the participatory results.

In defining the spatial logic to be used for the workflow, as a first step it is often useful to think about where **the REDD+ intervention cannot be undertaken,** to **exclude areas** where that REDD+ intervention would **not be possible**. It is often easier to do this before identifying **where the REDD+ intervention can be undertaken** in the remaining area**.**

When thinking about the spatial analysis to be undertaken, you will need to identify what criteria and data you will use to make those exclusions/inclusions, what geoprocessing tools you can use in ArcGIS, and whether the data are available (and available at an appropriate scale).

It is important to document the reasons why you are excluding or including certain areas and what factors have been used to determine these areas. This will help ensure that any analysis can be presented clearly and transparently to policy makers, so they can understand how a map has been created and what the assumptions are that have influenced it.

In this tutorial, an example workflow will be shown for the development of a REDD+ intervention layer for **community-based sustainable forestry to address the driver: small-scale conversion of forest to cassava**, on the basis that forests that are valued by the community are less likely to be converted to agricultural land.

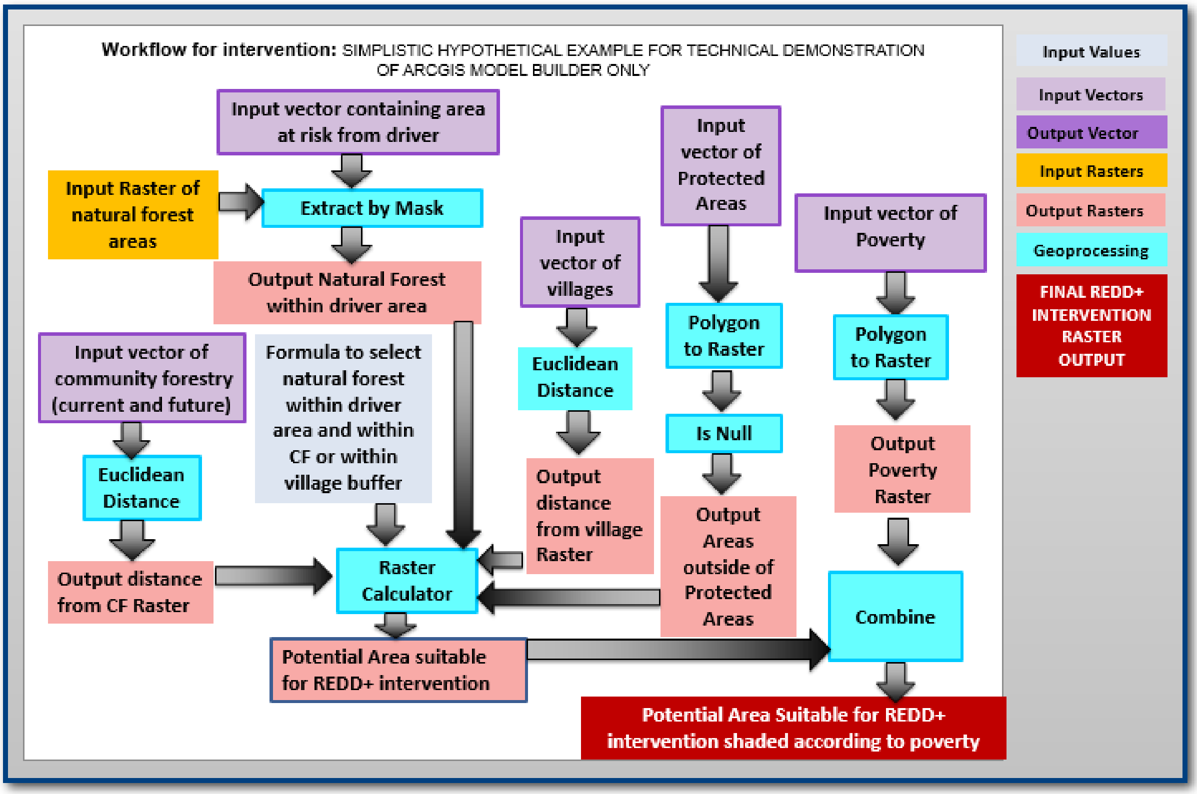
To help define a workflow for this question, we can consider the following questions:

- Where are the areas at risk from small-scale cassava expansion (now and in the future)?
- Where can community-based sustainable forestry feasibly occur? In our examples, areas close to villages are regarded as more feasible.
- Which forest area designations should be included? In our example, strictly protected areas are not available for community forestry.
- Should the intervention occur in natural forest and/or planted forest? In our example, natural forest is preferred.
- Should it be limited to existing community forestry areas or broader? In our example, areas close to existing or planned community forests are preferred.
- What benefits and risks are associated with the intervention? Can they be mapped? In our example, poverty alleviation was the only additional benefit considered.

The GIS analysis would then **exclude** areas not at risk from the driver, and areas where it is not possible to undertake that particular action, and **include** areas that are considered suitable, and where benefits could be enhanced and risks reduced. The table sets out the input layers for an example workflow:

Type of Input layer/data	How to use	Name of dataset in Mythical_Database.gdb
Forest cover	Natural forest area available for action	Natural_forest_mythical
Village locations	Areas near villages with likely demand for cassava production and/or community forestry (CF) activities	Village_mythical
Existing and potential community forests	Category of forest; areas already identified as current or likely CF sites	Community_forestry_mythical
Protected Areas	Exclude strictly protected areas	PA_mythical
Extent of driver	Area affected by / likely to be affected by driver (conversion to cassava) *(in this example a future Land Use Plan has been used to identify broad areas designated for agriculture)*	future_LUP_agriculture_mythical
Potential to alleviate poverty (as a benefit of the intervention)	Areas with medium-high poverty rates	Communes_mythical (using the field pov_hml , values: (1) low (2) medium and (3) high

The workflow can be hand-drawn in a flow diagram style prior to putting it into ArcGIS. The example for the above intervention is presented below.

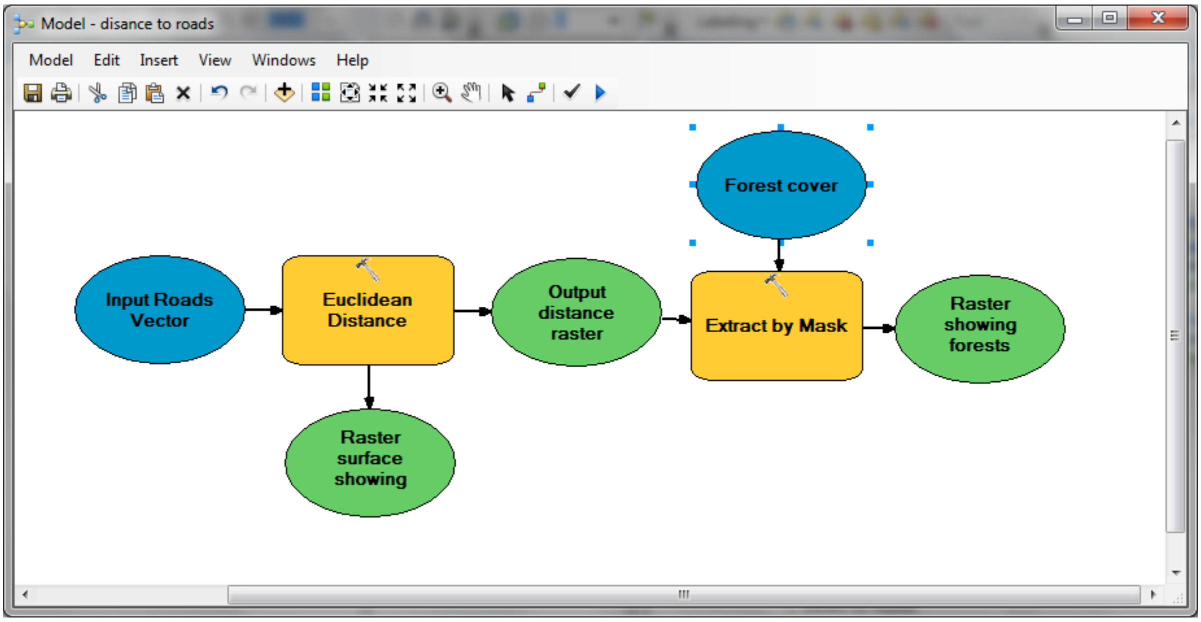


4.1. Using ArcGIS Model builder to implement workflows

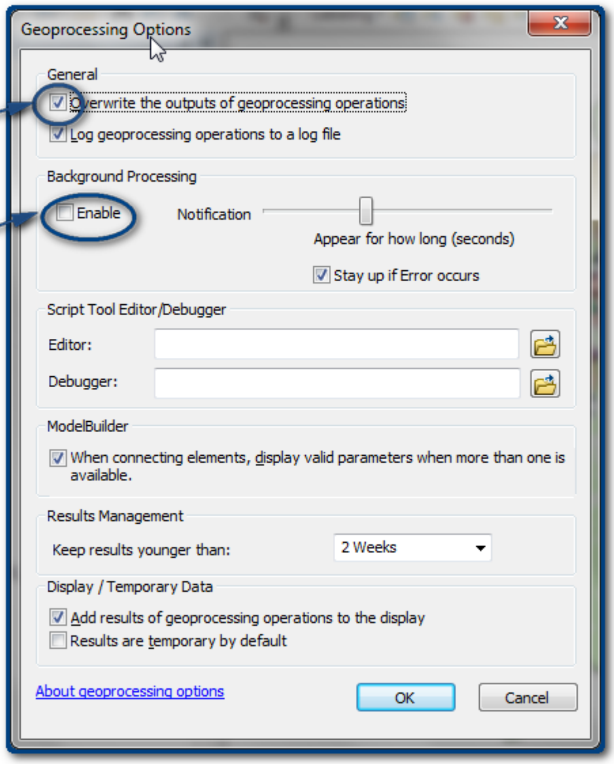
4.1.1. Overview of Model Builder

Model Builder is an application that you can use to **create, edit**, and **manage models**. Models in ArcGIS are workflows that string together sequences of geoprocessing tools, feeding the output of one tool into the next. Benefits of using Model Builder include better organization for improved workflows and faster analysis. By standardizing processes in this way, any future repeat analysis is also made much easier, enabling replication and direct comparison of results.

A simple model may only contain one or two steps and ArcGIS geoprocessing tools. For example, the model below classifies forest according to distance from roads:



4.1.2. Transferring a workflow into ArcGIS (using the REDD+ intervention example)



- **a.** Open ArcMap
- **b.** From the main menu click **Geoprocessing >>Geoprocessing Options**
- **c.** Ensure that the option **Overwrite the outputs of the geoprocessing operations** is **enabled.**

  So that you don't keep having to change the name of an output file when testing the model multiple times
- **d.** Ensure that **Background Processing** is **disabled**.

  ArcGIS Help indicates that custom tools are better run in the foreground, as this will allow you to see each process running and will highlight any errors found.

e. Click **OK** to close the window
f. From the folder **MODEL_BUILDER _DEMO**, Open the project **Mythical_Model_Builder_10_x.mxd.**
g. Familiarize yourself with the data in the Table of Contents in ArcMap and with its location in the **Mythical_Database.gdb** geodatabase in the ArcCatalog window. The table below lists the data layers that are contained in the geodatabase.

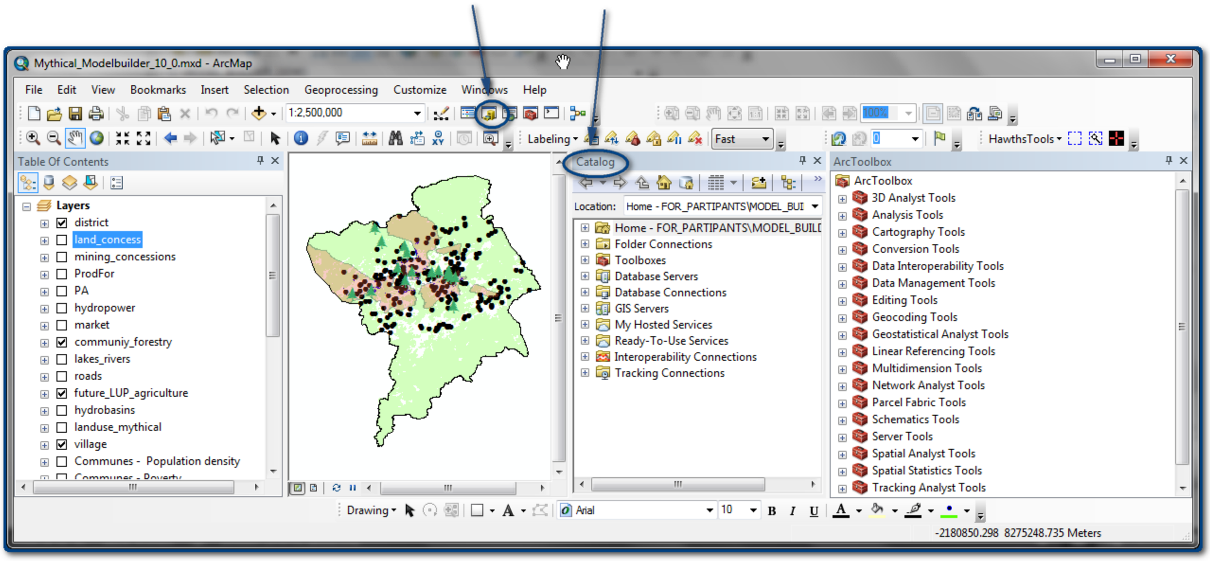
communes_mythical – Administrative boundaries (Communes)	community_forestry_mythical – Current and proposed community forestry sites
DEM_mythical – Digital Elevation model	district_mythical – Administrative boundaries (Districts)
future_LUP_agriculture_mythical – Future land use plan for Agriculture	HILL_mythical – Hillshade generated from Digital Elevation model
hydrobasins_mythical – Hydrological basins	hydropower_mythical – Planned and existing Hydroelectric sites
lakes_rivers_mythical – Water bodies and rivers	land_concess_mythical – Economic land concessions
landuse_mythical – Current landuse	market_mythical – Market locations
natural_forest_mythical – A subset of the current landuse layer that just contains the natural forest classes)	PA_mythical – Protected Areas
mining_concessions_mythical – Mining concessions	roads_mythical – Existing and planned roads
prod_for_mythical – Production forests	Poverty rate_mythical – Poverty rate by commune
village_mythical – Village locations	

Note: not all of these datasets will be used in the demonstration analysis, but they can be used as test data when developing other workflows.

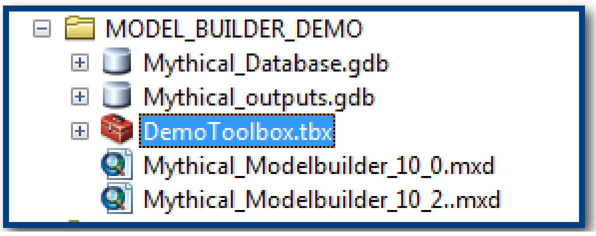
4.1.3. Create a new Toolbox

A toolbox is the area where you store your models, either as individual models or sets of models. A toolbox can be created inside a geodatabase or just in a folder. A toolbox that has been saved in a folder (i.e. outside a geodatabase) is easier to share as it can simply be copied (more detail about sharing models is covered in Section 5).

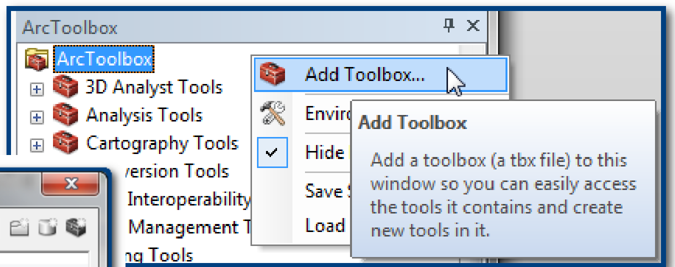
a. To start using **Model Builder**, open the **ArcCatalog** window.

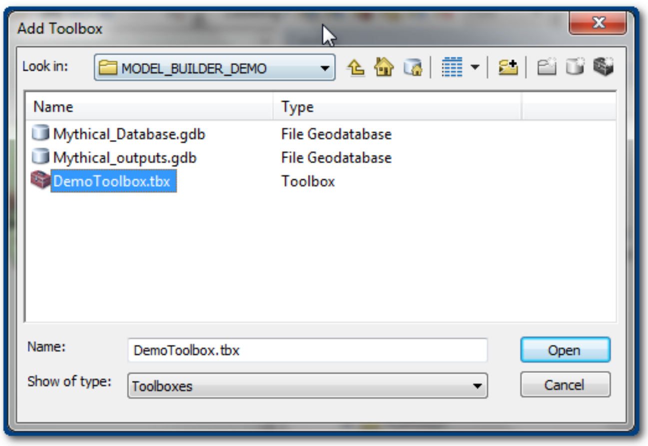


b. Right-click on the model builder folder **MODEL_BUILDER_DEMO>> New >>Toolbox** and give it a name of your choice e.g. in this example **DemoToolbox.tbx**



c. To add your own toolbox to ArcToolbox right-click on **ArcToolbox>>Add Toobox**

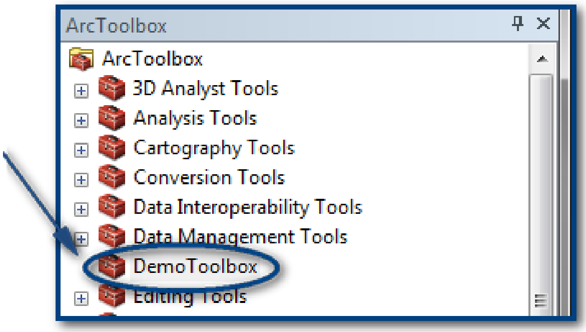




d. Navigate to your toolbox e.g., in this example it is **DemoToolbox.tbx**

e. Click **Open**

The toolbox will be added in alphabetical order to the ArcToolbox window



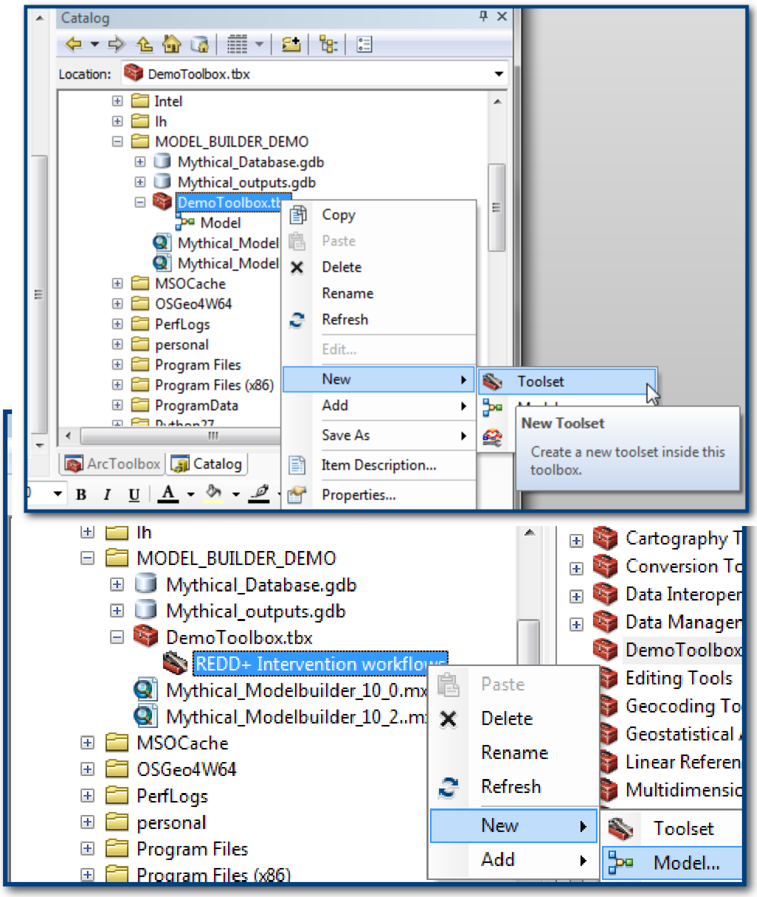
4.1.4. Adding a new Toolset and models to a Toolbox

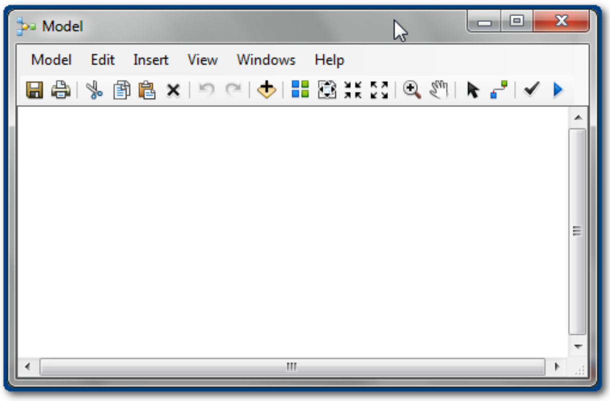
Tools can be added directly to a toolbox, or a toolset can be created to group related Models. The toolset could contain models relating to a particular theme, or developed for a particular project.

If you are creating lots of models it is useful to put them in a toolset.

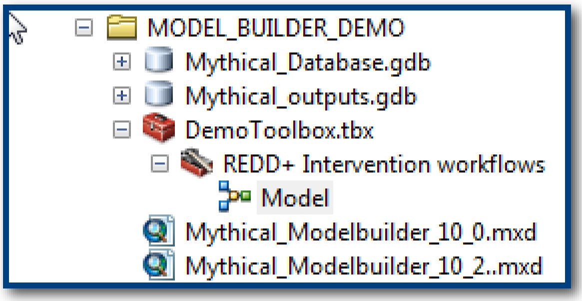
- **a.** Either in the **ArcCatalog** window or the **ArcToolbox** window right-click on the new toolbox **DemoToolbox.tbx>> New>>Toolset**
- **b.** Give the toolset a name, e.g. **REDD+Intervention workflows**
- **c.** Next add a Model to the toolset:

  Right-click on the new toolset **REDD+Intervention workflows>>New>> Model.**





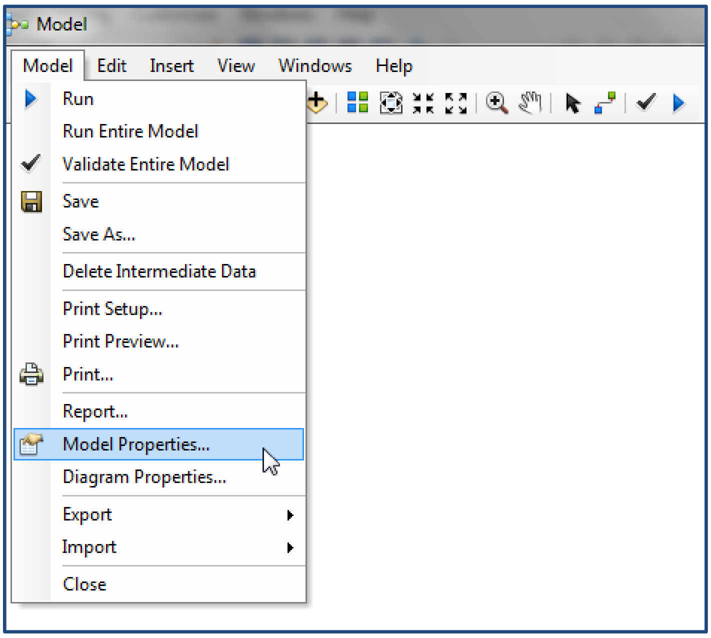
This creates a new empty model and opens it in the model builder window (left). This is the starting point for adding the tools that make up your workflows.



Note that the model is automatically named "Model" at this stage.

Before you start using the model, it is wise to rename it, so it can be distinguished from other models and make it easier for other people to use.

d. From the Model builder window click on **Model>>Model Properties**



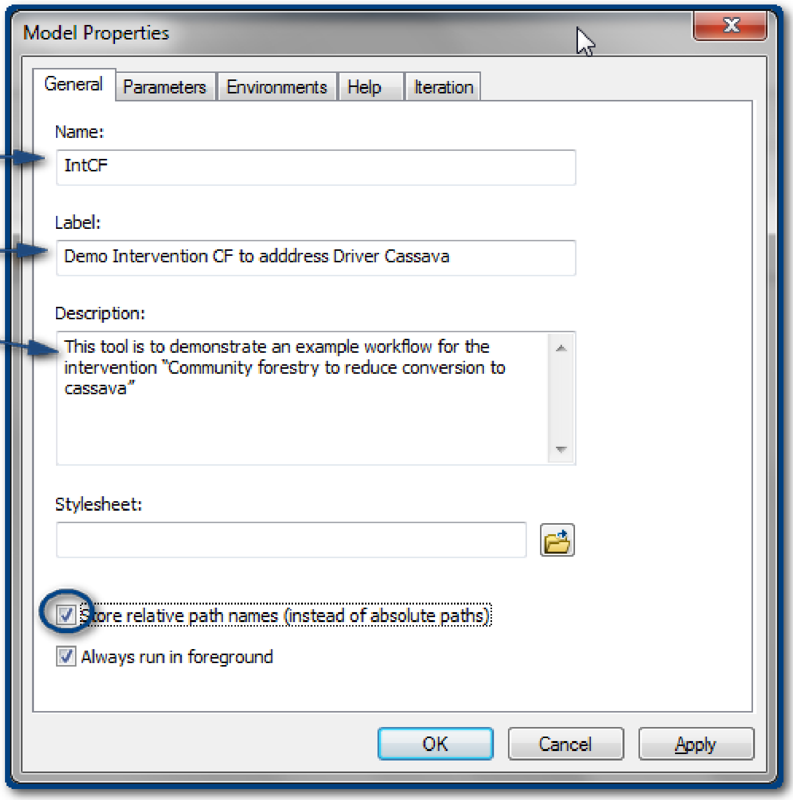
Next use the following steps to fill in the **Name** (with NO spaces), **Label** and **Description.**

*The **label** is the name that appears in ArcToolbox and the **name** is the name that ArcGIS uses to distinguish the tool from other tools in ArcToolbox*

e. Give the tool a **name e.g. IntCF**

f. Give the tool a **label** e.g. **Demo Intervention CF to address Driver Cassava**

g. Give the tool a **Description** e.g. **This tool is to demonstrate an example workflow for the intervention "Community forestry to reduce conversion to cassava"**

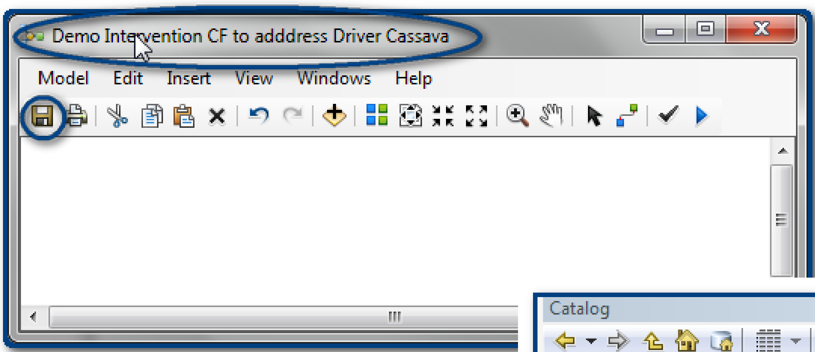


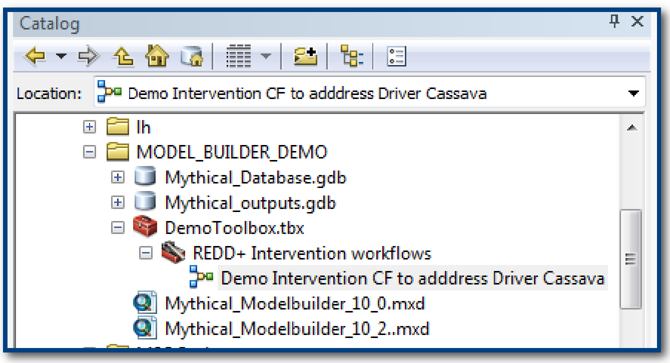
h. Tick to **Store relative path names**

If you check this option, then all path names within the tool will be stored relative to the toolbox containing the tool. This option is really useful if you plan to share your model tool, or move your model data and model to a different location.

i. Click **OK** to close the window

See that the Model name has changed name in the Model Builder window:





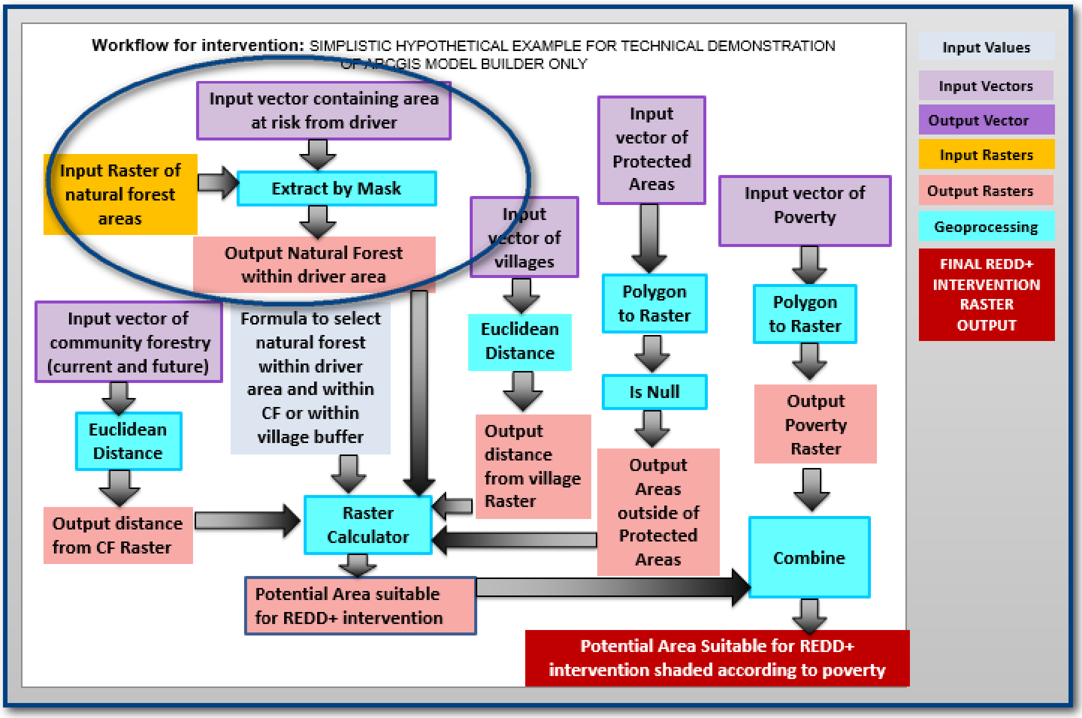
j. Click the **save** button to save your changes to the model, and see the Model label change to your new name in ArcCatalog and ArcToolbox.

4.1.5. Adding the workflow to the model

We will use the example workflow developed and described in **Section 4** to demonstrate this. The first step is select the areas at risk from the driver* "conversion to cassava" and within those areas ensure that the intervention "Community Forestry" can only occur in natural forest areas.

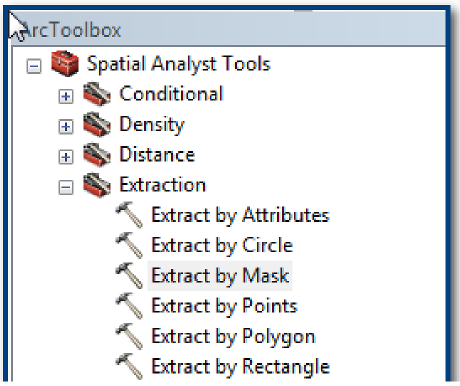
*(*Remember in this example a shapefile already exists for the driver area).*

For Step 1, we had planned to use the "**Extract by Mask**" to clip out the **natural forest areas** within the **driver extent**.



The easiest way to add tools to the model is to drag the tool from ArcToolbox and drop it into the Model Builder window.

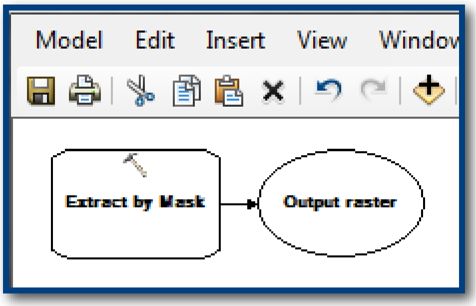
.



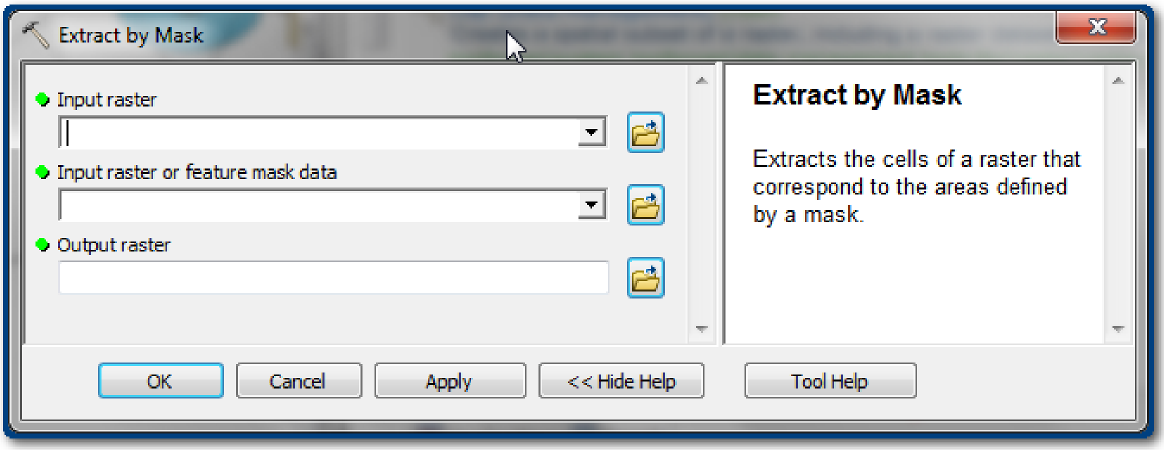
a. Search for the **Extract by Mask tool** (or find it located in the **Spatial Analyst>> Extraction toolbox**)

b. **Drag the tool** into the model builder window

c. Double-click on the **Extract by Mask tool** box in the model builder window to see the tool inputs.



The **Extract by Mask tool** opens:



See that the tool requires:

- An input raster layer (from which the data are to be extracted): the forest layer
- The Mask dataset (which will be used to clip out the areas within the input raster): the driver extent
- A name and location for a new output raster

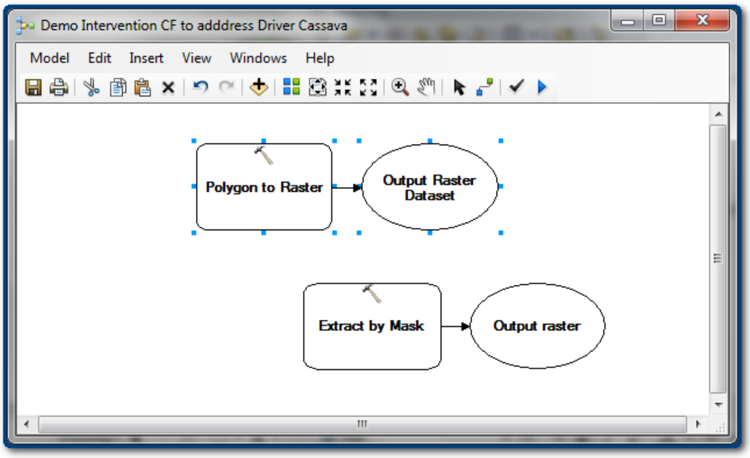
In our example, the input dataset will be a forest layer, the mask will be the driver extent and the output file will be the forest area within the driver extent.

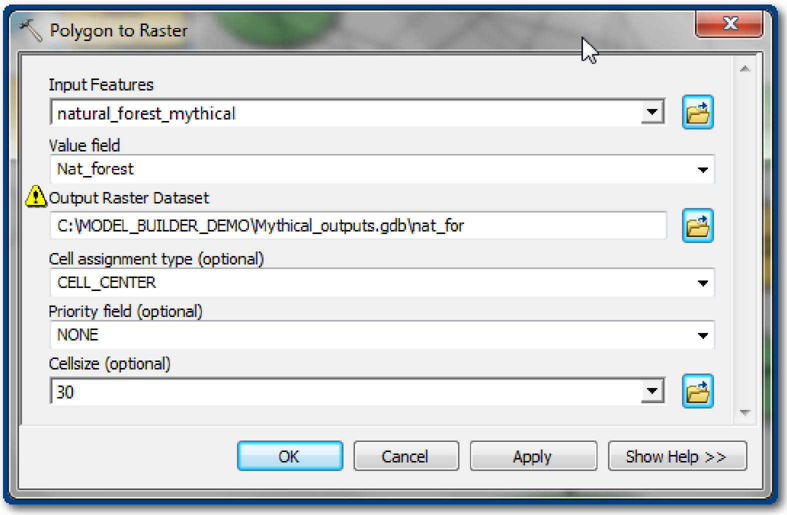
We immediately find a problem for our planned workflow as our natural forest layer is a vector dataset. Therefore we have to add a new step to our model, prior to the extract by mask.

d. Click **Cancel** to close the **Extract by Mask tool window**

e. Find the **Polygon to Raster tool** in ArcToolbox **Conversion>>To Raster** toolbox and drag it into the Model Builder window above the **Extract by Mask tool**

f. Double-click on the **Polygon to Raster tool** in the Model Builder window



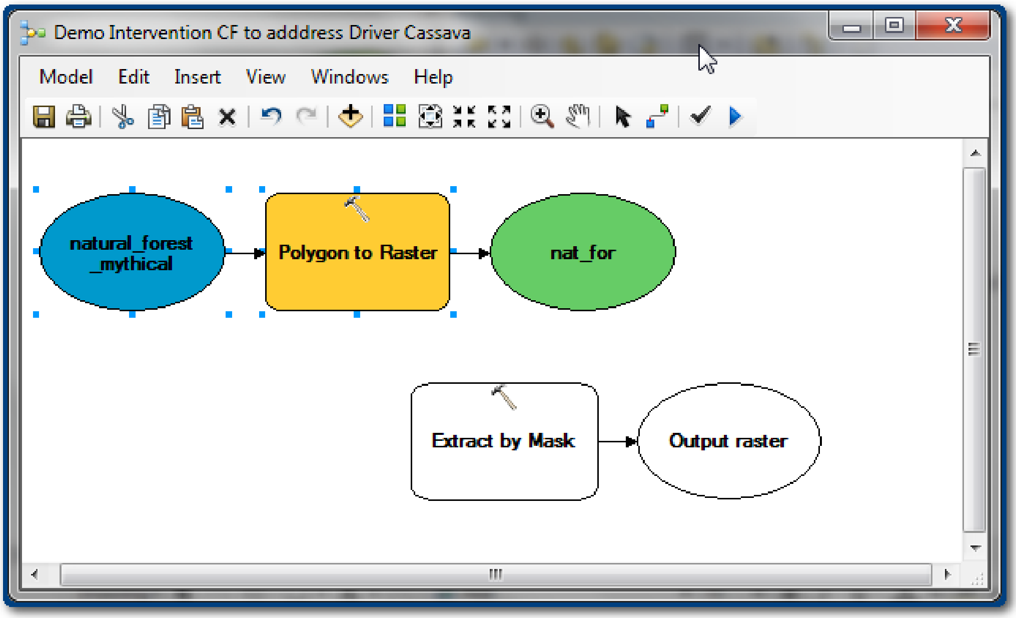


- Select the **natural_forest_ mythical** vector layer from the **Mythical_Database.gdb**
- Choose **Nat_Forest** for the value field (The Nat_forest field in this dataset contains the Value 1 for all the polygons that are Natural forest)
- Put the output raster in **Mythical_output.gdb** and call it nat_for. i.e. **C:\MODEL_BUILDER_DEMO\Mythical_outputs.gdb\nat_for**
- Leave the cell **alignment type** as **CELL_CENTER**
- Leave the **priority field** as **NONE.**
- The **cell size** should be set to the **scale at which you want to undertake all the analysis** in your workflow**.** In this example we chose **30m**

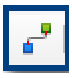
g. Click **OK**

The model now appears with an input and output for the **Polygon to Raster tool**
Note how the tool changes so that the **tool** appears in **Yellow,** the **inputs** appear in **Blue** and **outputs** appear in **Green** in the Model Builder window.

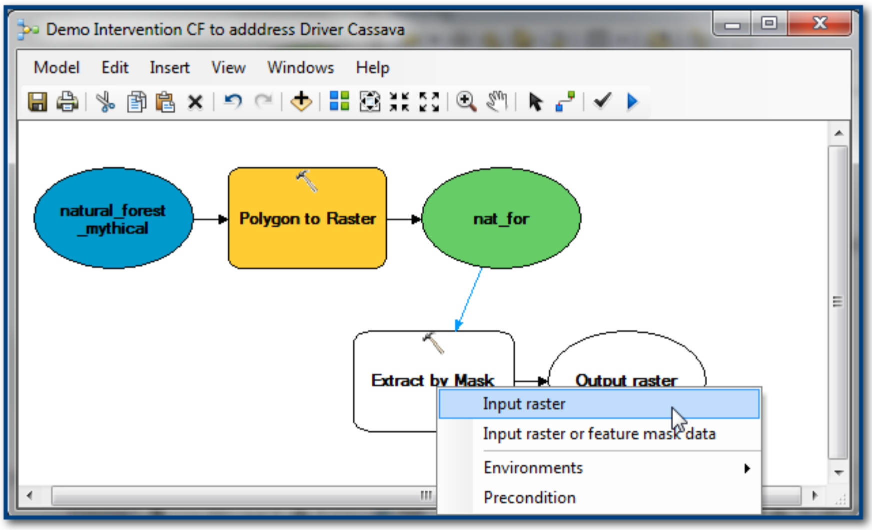
Note: The tool only changes to a colour when all the required parameters (inputs and outputs) are correctly defined.



The input layer in the **Extract by Mask tool** is the output from the **Polygon to Raster tool**, i.e. in this example **nat_for**

h. Click on the  **connect** button and drag a line **from nat_for** to the **Extract by Mask tool**.

i. The pop up menu ask what to connect **nat_for** to. Chose **Input raster**

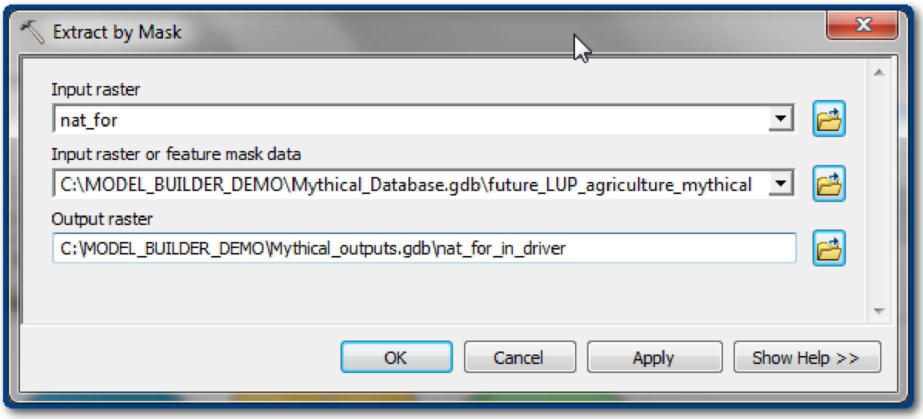


Notice that the model still does not change colour. The tool still requires some inputs.

j. Double-click on the **Extract by Mask tool**

See that the input raster is already set as nat_for (from the previous step)

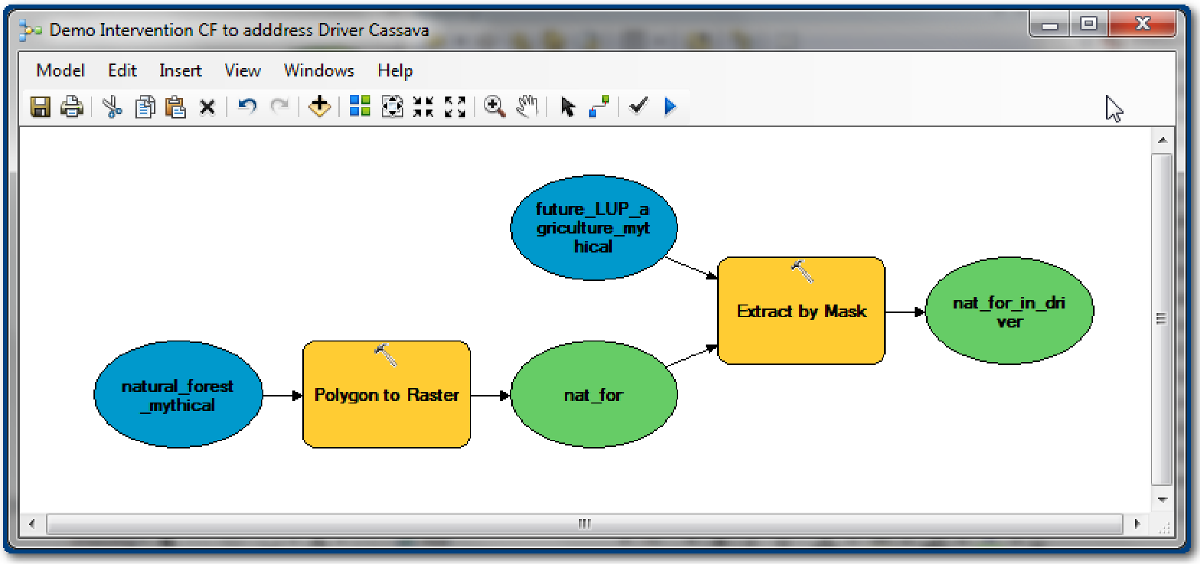
- Choose the **driver extent** as the mask e.g. **future_LUP_agriculture_mythical**
- Chose a **location and name** for the output raster e.g. **C:\MODEL_BUILDER_DEMO\Mythical_outputs.gdb\nat_for_in_driver**



k. Click **OK** to close the tool

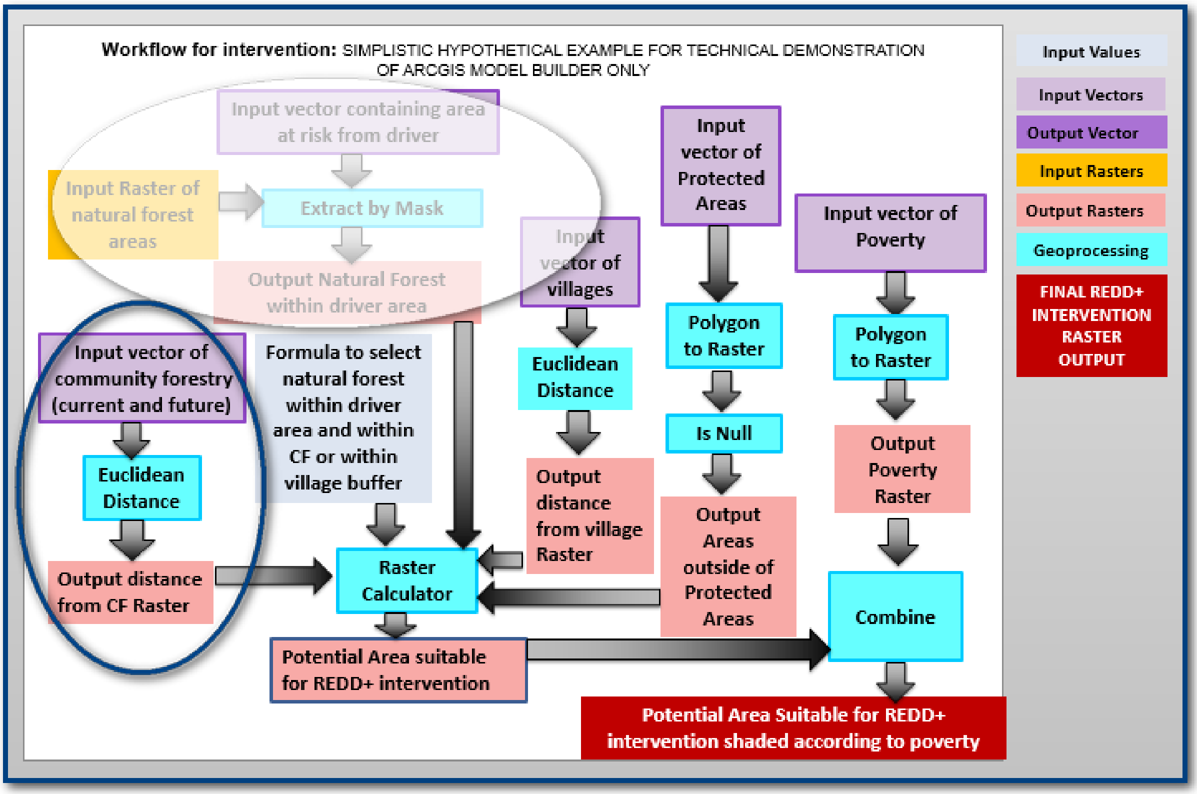
The **Extract by Mask tool** is now coloured and the first two steps of the tool are now complete.

l. If you'd like to tidy the tool and rearrange them, you can move them using your mouse, or use the Auto Layout  and Full Extent buttons on the toolbar

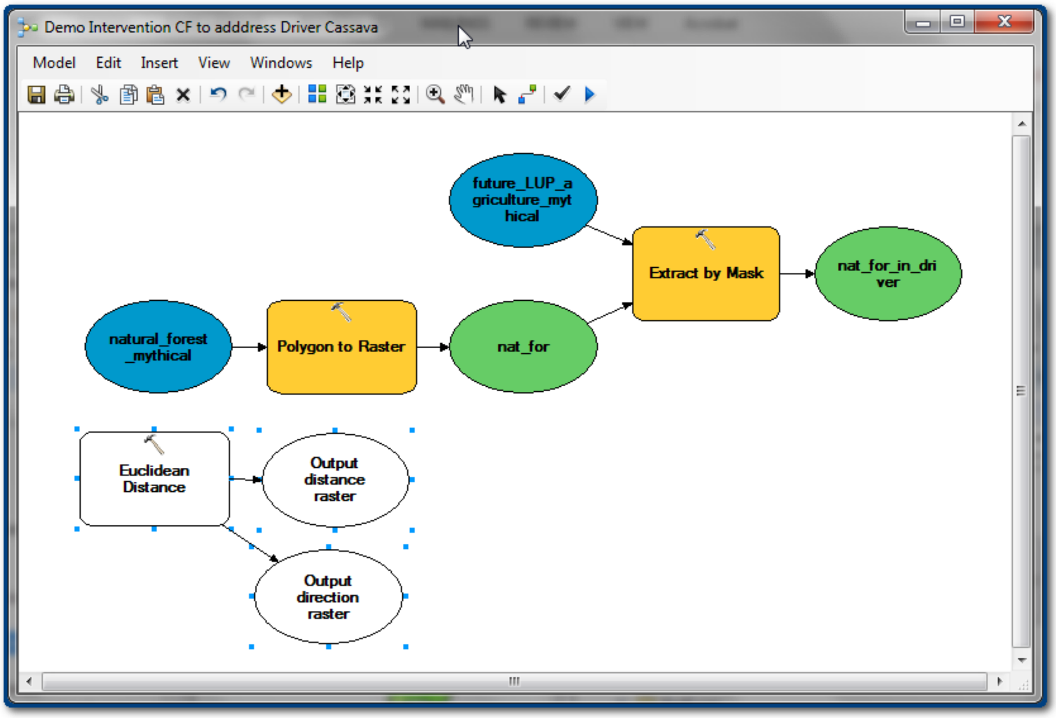


m. Click the **save** button to save the model.

Look at your workflow design again to see what the next step is to add to Model Builder. There is a **Euclidean Distance tool** on the left. This will generate an output distance raster to the community forest sites Raster.

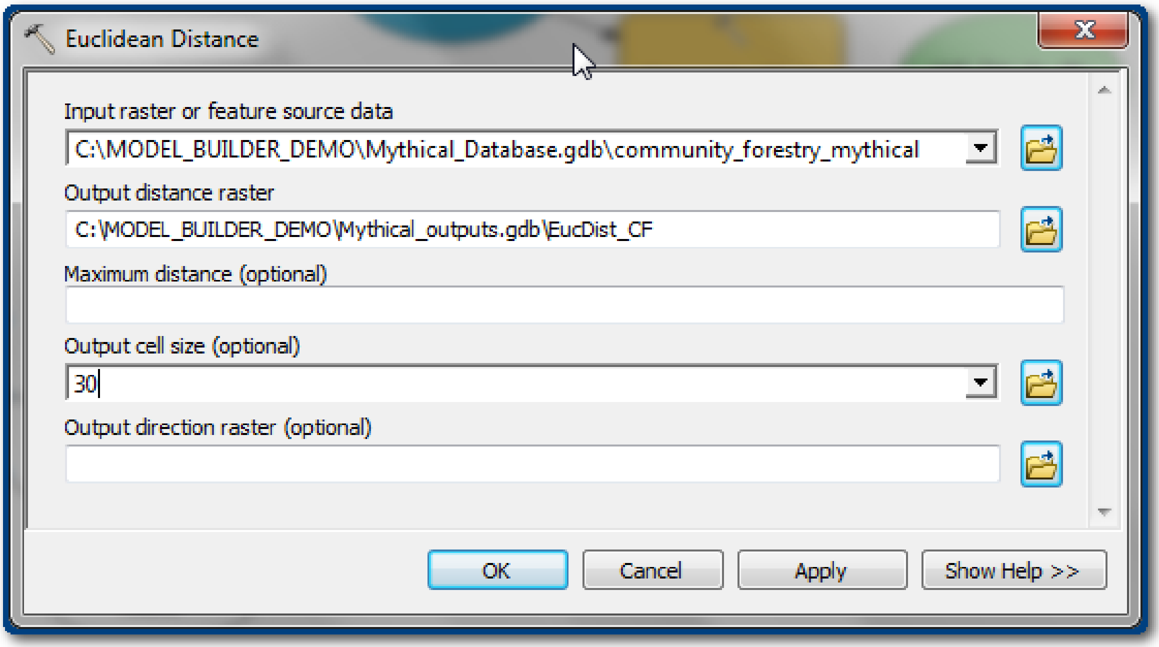


n. Search for the **Euclidean Distance tool** or find it in in the **Spatial Analyst>>Distance** toolbox and drag it into the Model Builder window.

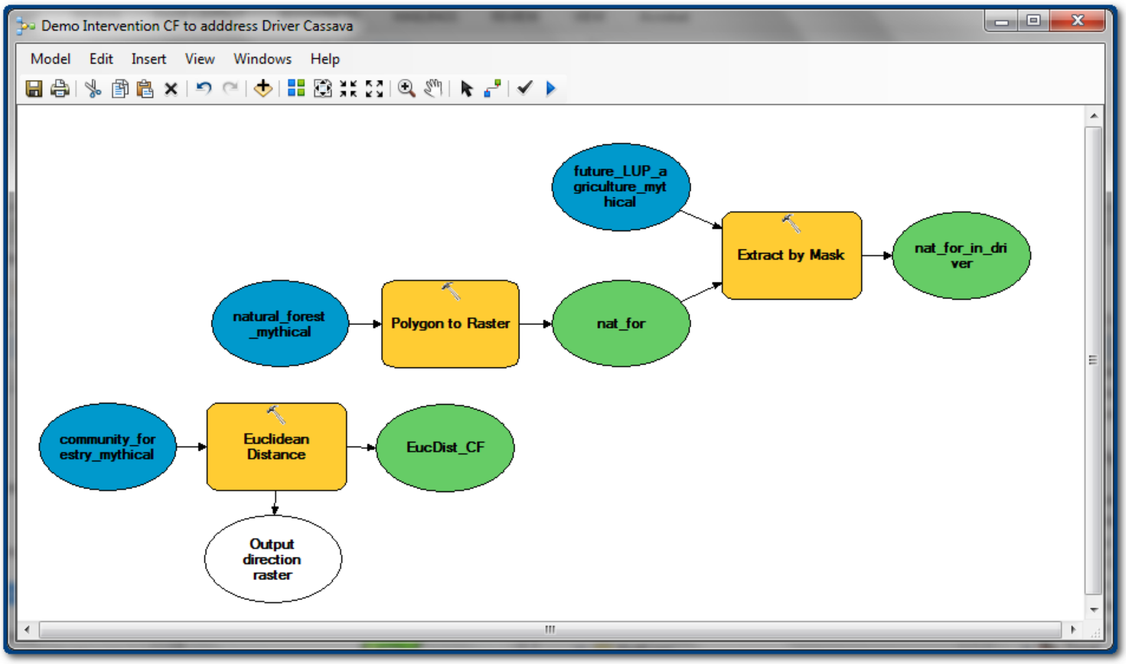


o. Double-click on the **Euclidean Distance tool**:

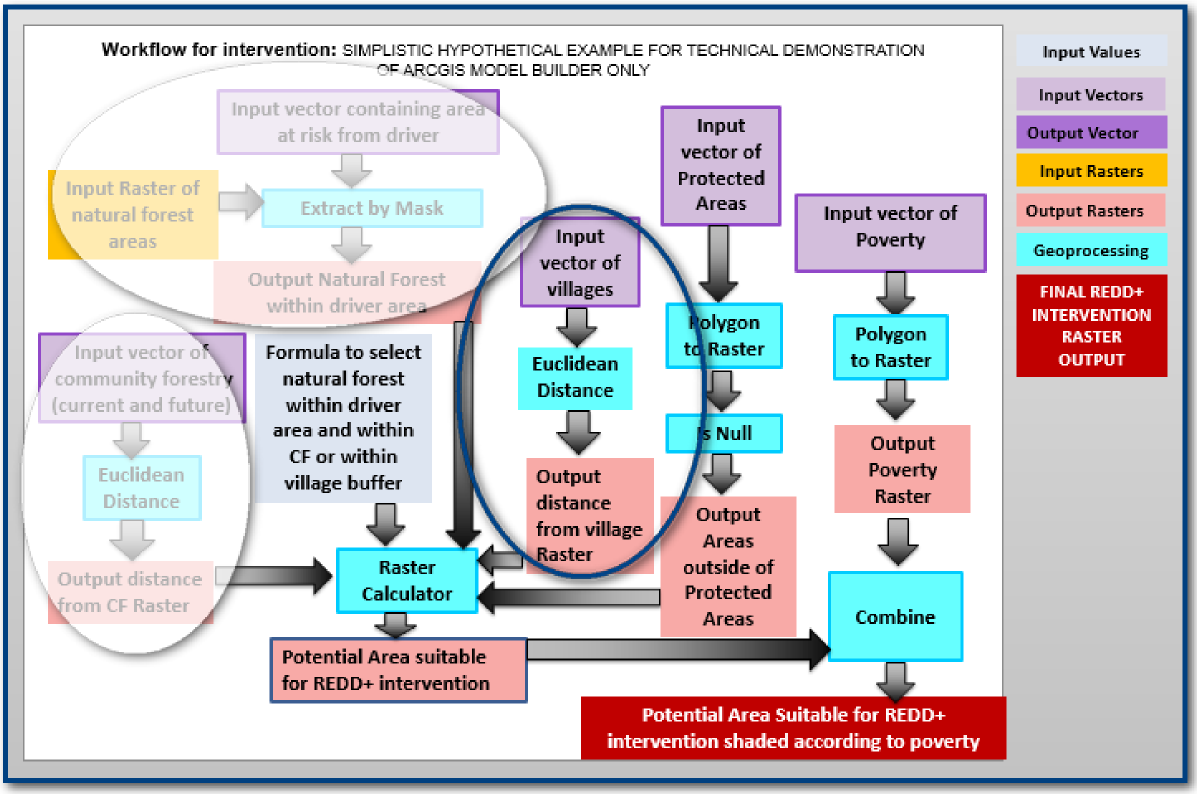
- Choose the **Input raster or feature source data** to generate distance from (i.e. choose the vector layer of current and potential community forest sites - **C:\MODEL_BUILDER_DEMO\Mythical_Database.gdb\community_forestry_mythical** in this example).
- Set the **name and location** of the output distance raster e.g. **C:\MODEL_BUILDER_DEMO\Mythical_outputs.gdb\EucDist_CF**
- Leave the **Maximum distance** blank
- Set the **output cell size** to be the same as the cell size you chose for the whole workflow e.g. **30** in this example
- Leave the output direct raster blank



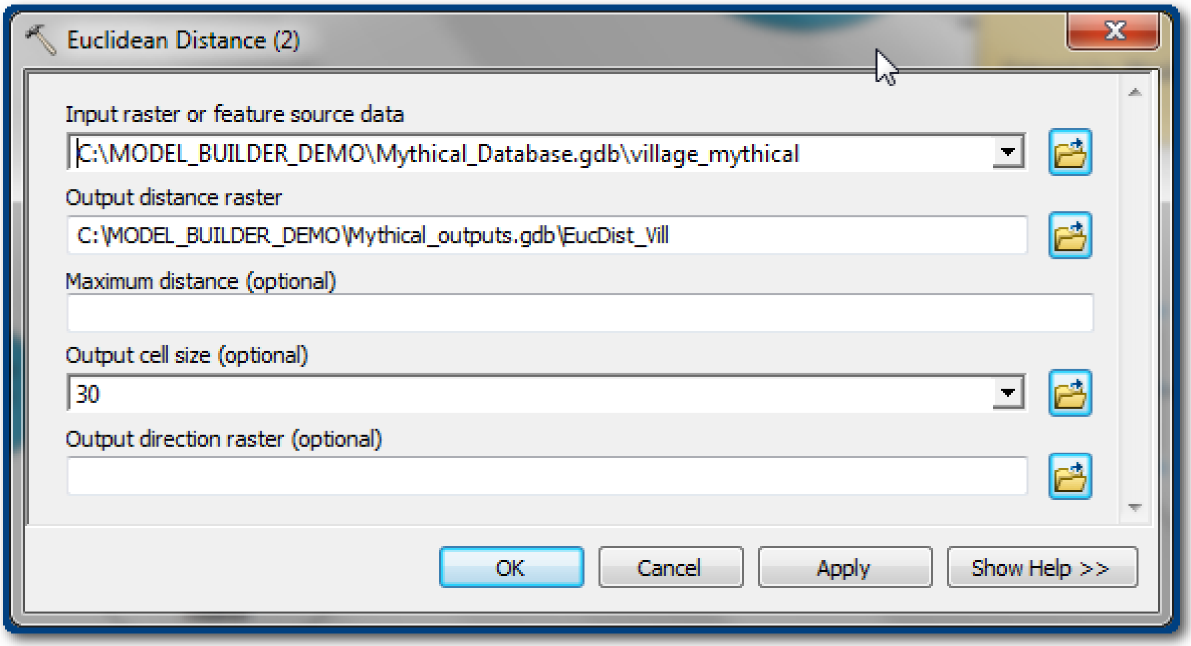
The tool becomes coloured except for the direction raster, which we left blank because it was optional.



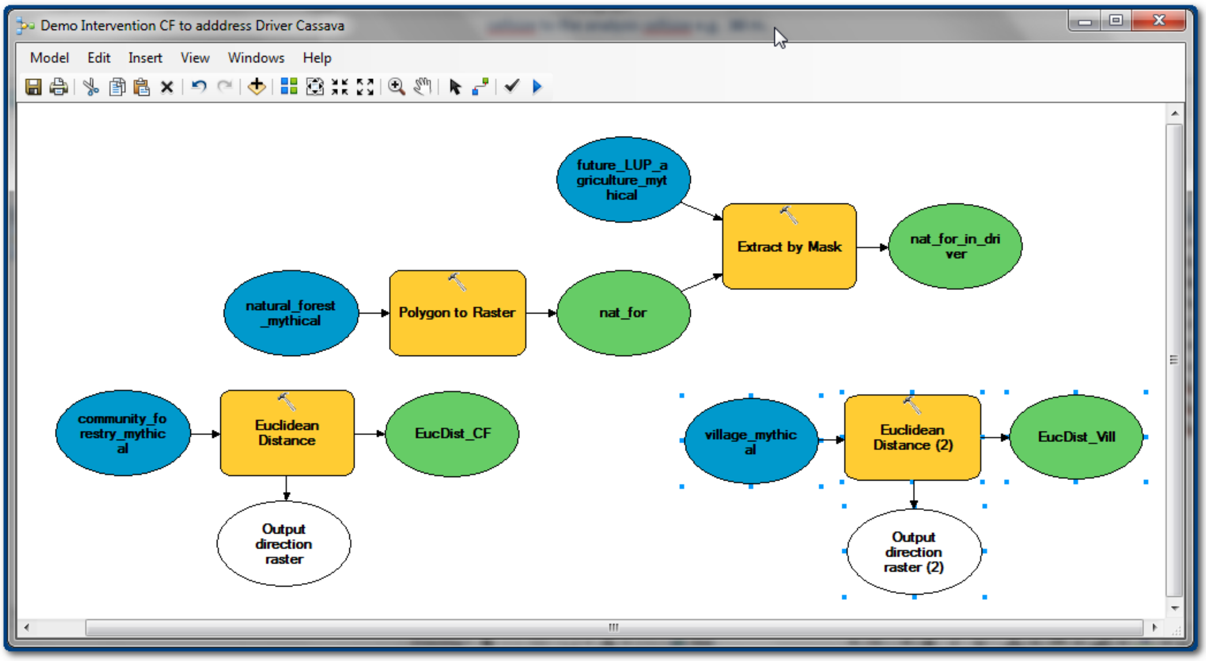
Look at your workflow design again to check what the next step is to add to Model Builder. There is **a Euclidean Distance tool** on the right. This will generate an output distance to villages raster.



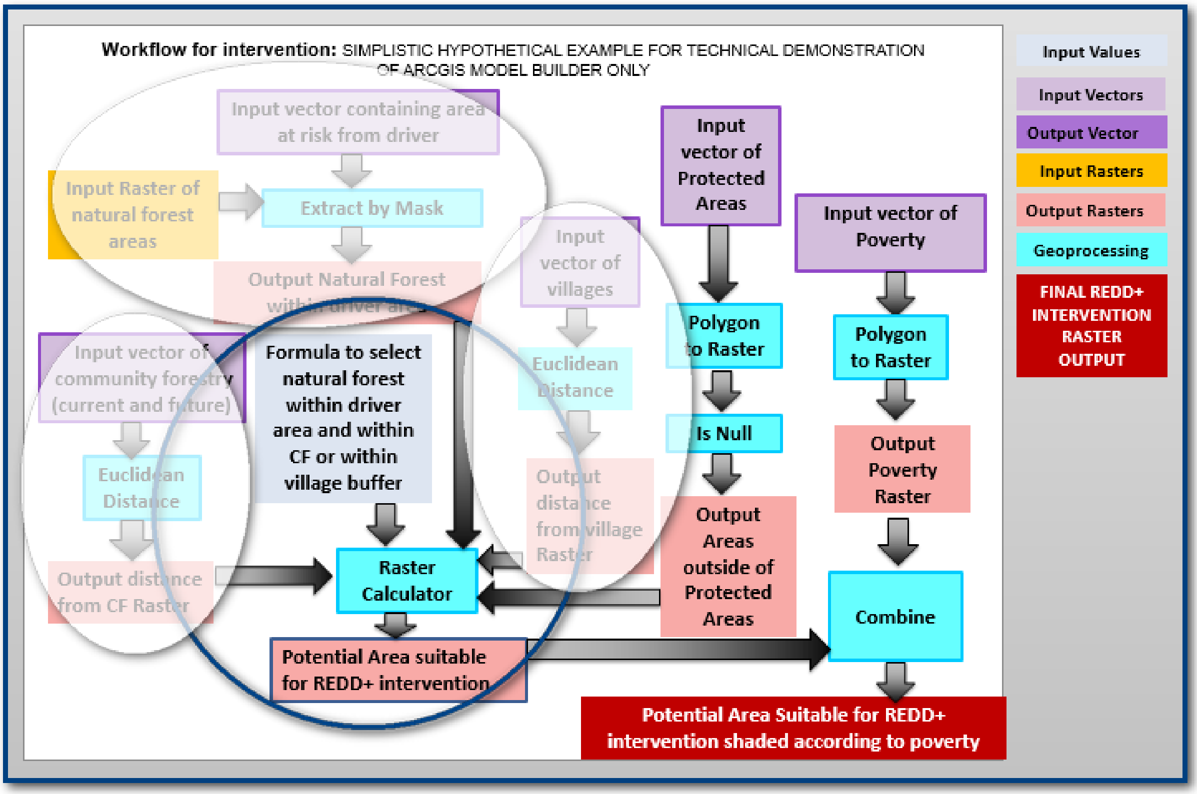
p. Drag in another **Euclidean Distance tool** and, in the same way as the previous steps, set the input raster or feature source data as the villages vector layer e.g. **C:\MODEL_BUILDER_DEMO\Mythical_Database.gdb\village_mythical,** the output distance raster to e.g. **C:\MODEL_BUILDER_DEMO\Mythical_outputs.gdb\EucDist_Vill** and output cell size to the analysis cell size e.g. **30 m**.



q. Click **OK** and the second Euclidean step is now in the model



Look at your workflow design and see what the next step is to add to Model Builder. In the workflow diagram there is a **Raster Calculator process tool**, which will use all the outputs from the **Extract by Mask tool** and **Euclidean Distance tools** to select out natural forest areas within the driver extent within a specified distance from community forestry sites and villages.



In the **Raster Calculator tool** you can type an expression to select out the required areas (See Annex 1, section A.3.5 for help on the syntax for **Raster Calculator tool**). Writing a **Raster Calculator tool** expression as planned in the workflow is complex, so we will break down this step into multiple steps.

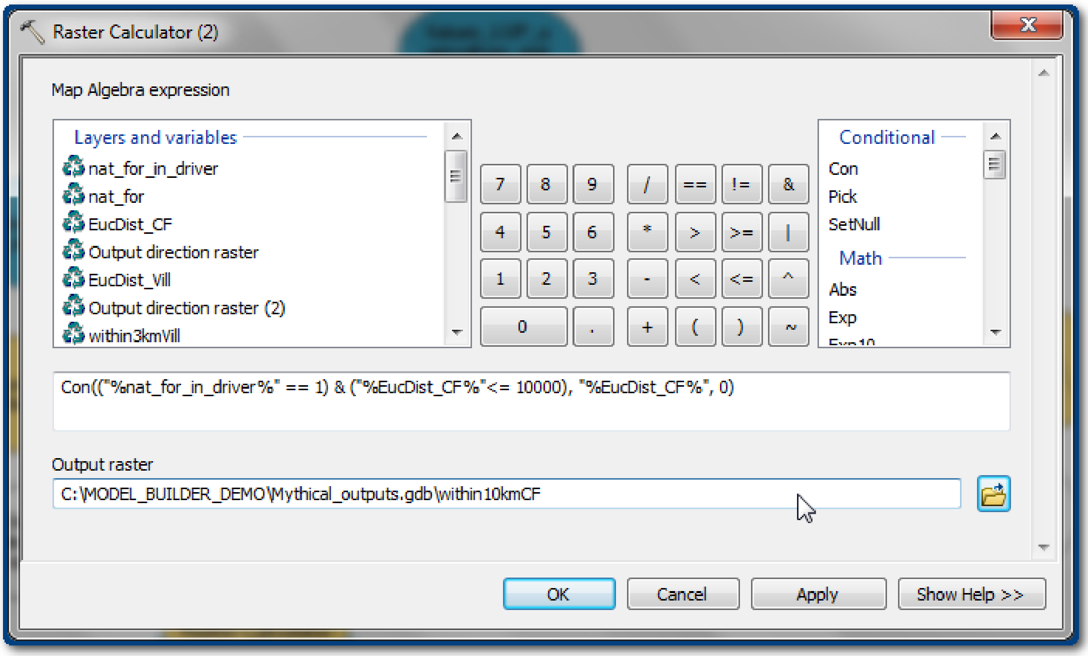
Whilst the workflow is needed to guide us through the analysis, it is common to find that the tool you planned to use is either not appropriate or the step can be done in a better or simpler way.

So in this case, we will add two raster calculator steps: 1) to select out forests within the driver extent within 10km of existing and proposed community forestry sites; and 2) to select out forests within the driver extent within 3 km of villages.

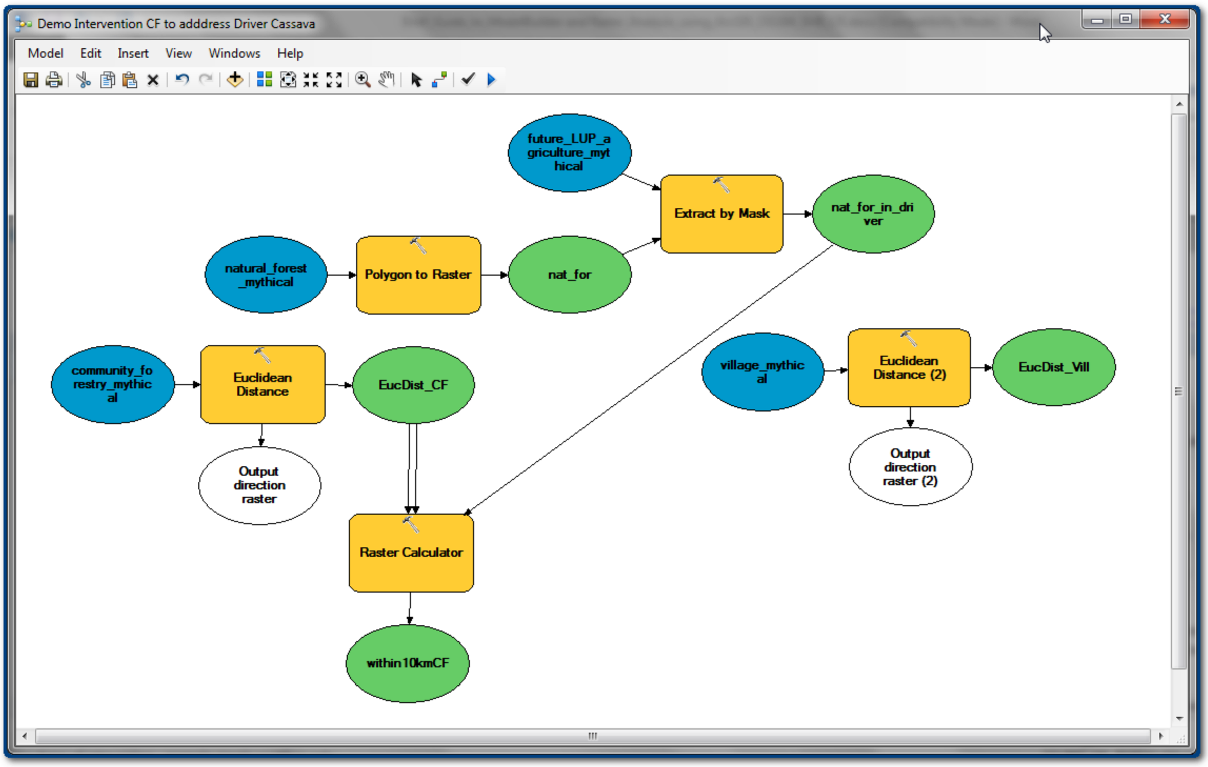
r. Search for the **Raster Calculator tool** or locate it in the **Spatial Analyst>>Map Algebra Toolbox** and drag the **Raster Calculator tool** into the Model Builder window.
s. Double-click on the **Raster Calculator tool** and in the white panel **(expression box)** type:

Con(("%nat_for_in_driver%" == 1) & ("%EucDist_CF%" <= 10000), "%EucDist_CF%", 0)

This means **IF** cells have a value of **1** in the **nat_for_in_driver** raster (i.e. forest within the driver extent) then make the value in the output dataset **1** otherwise make the values in the output dataset **0**, **AND IF** cells have a value of **less than or equal to 10km** in the **EucDist_CF** (i.e. distance to proposed and existing community forestry sites) make the values the **same as the distance value in the EucDist_CF** dataset, otherwise make the values in the output dataset **0**. The output is the **forest areas within 10km of proposed and existing community forestry sites**.



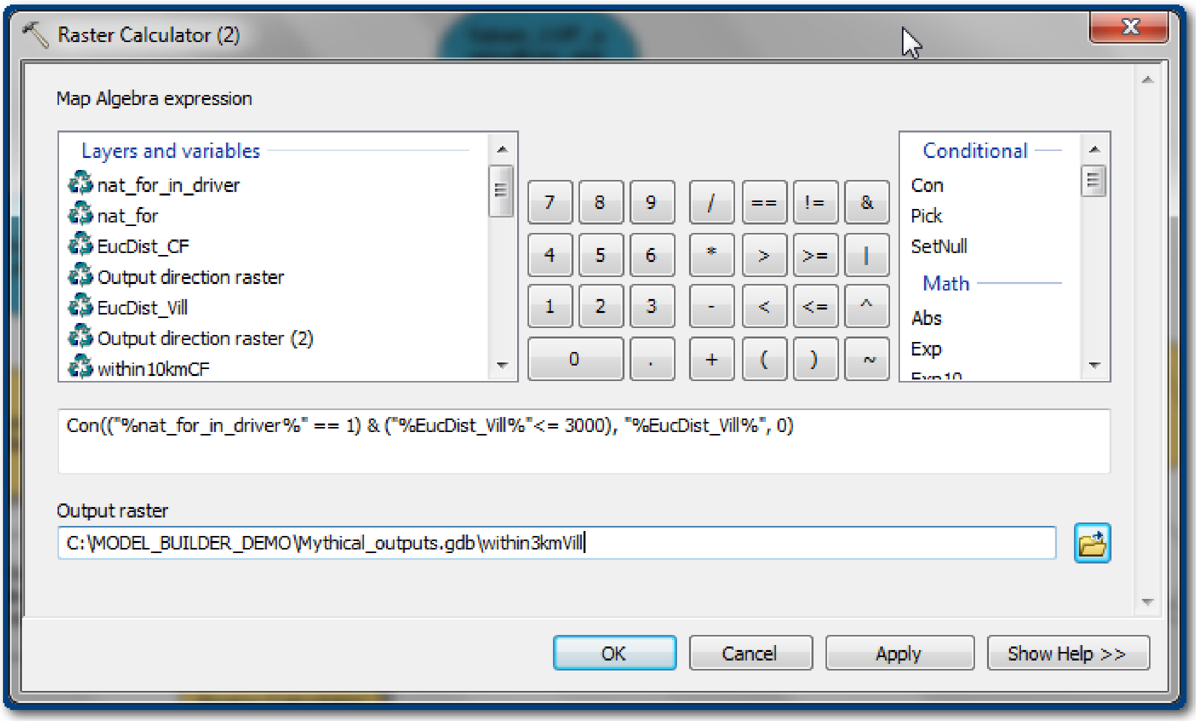
The model should now look like this:



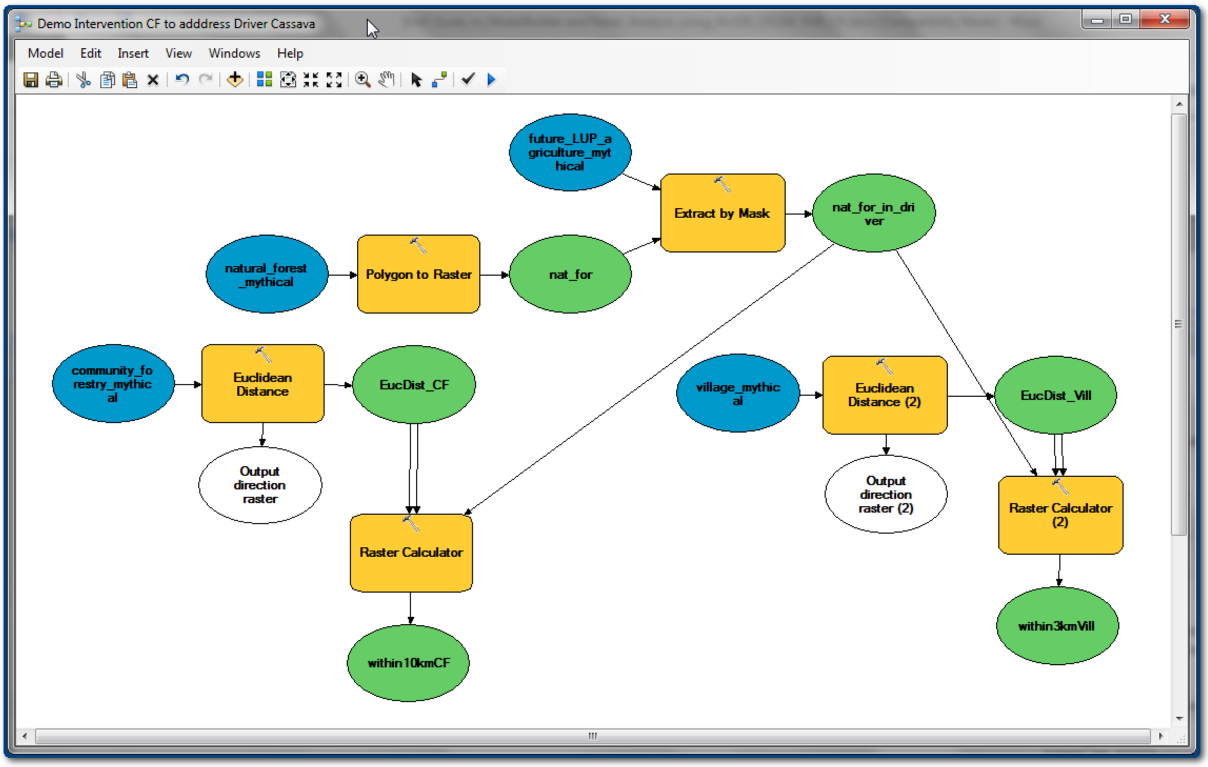
t. Add another **Raster Calculator tool** to the model and this time type in the following expression the expression window type:

Con(("%nat_for_in_driver%" == 1) & ("%EucDist_Vill%" <= 3000), "%EucDist_Vill%", 0)

This means **IF** cells have a value of **1** in the **nat_for_in_driver** raster (i.e. forest within the driver extent) then make the value in the output dataset **1**, otherwise make the values in the output dataset **0**, **AND IF** cells have a value of **less than or equal to 3km** in the **EucDist_Vill** (i.e. distance to villages) make the values the **same as the distance value in the EucDist_Vill** dataset, otherwise make the values in the output dataset **0**. The output is the **forest areas within 3 km of villages**.

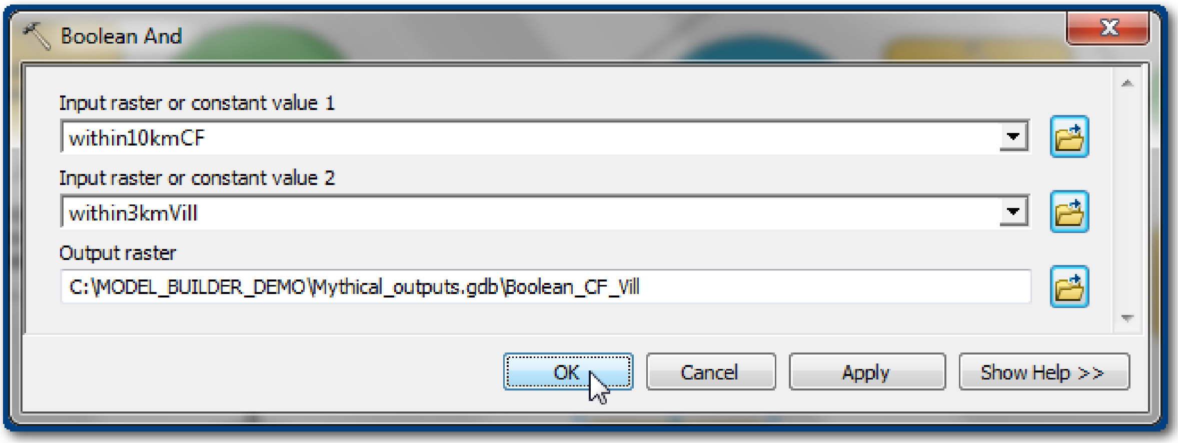


The model should now look like this:

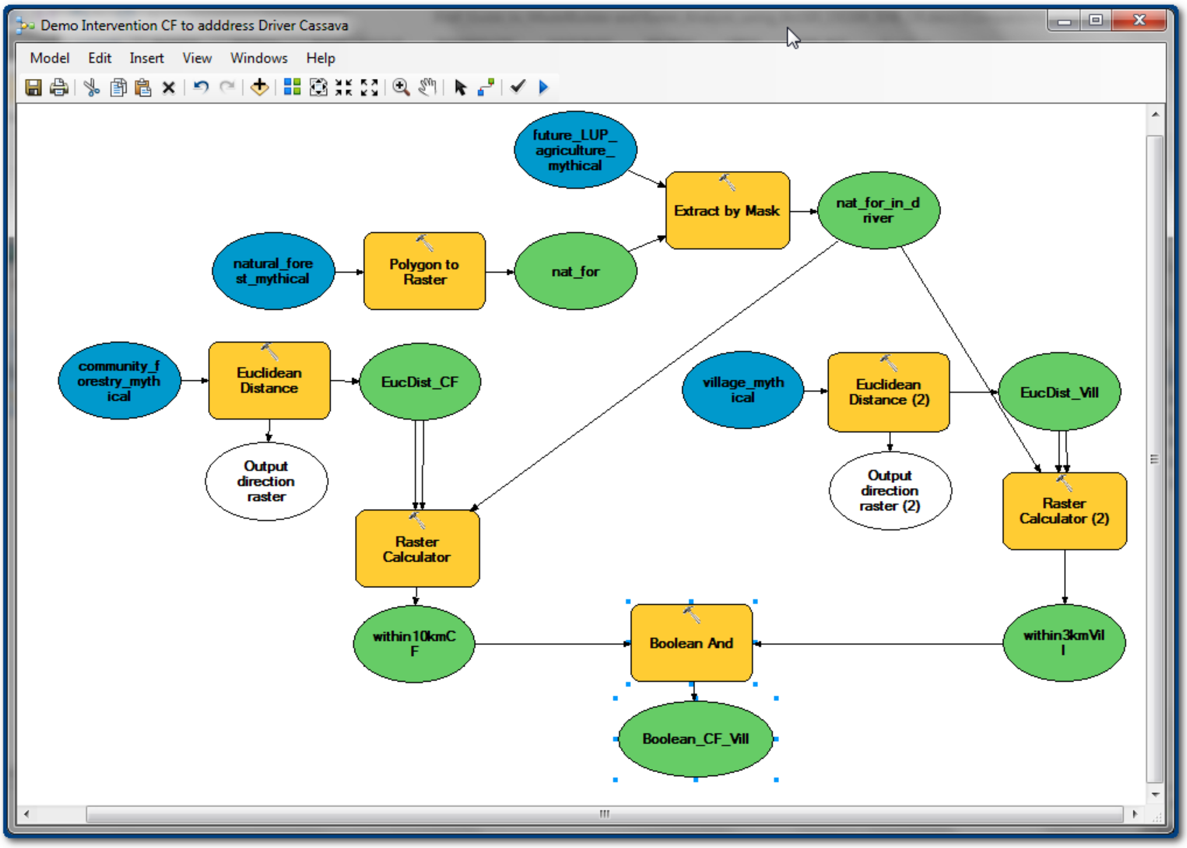


u. We need another step to combine the results from the **Raster Calculator tool** outputs together. So the next step is to add a **Boolean And tool** (located in the **Spatial Analyst>>Math>>Logical toolbox**) to select raster cells that both within 10km of Community Forest or within 3 km of Villages.

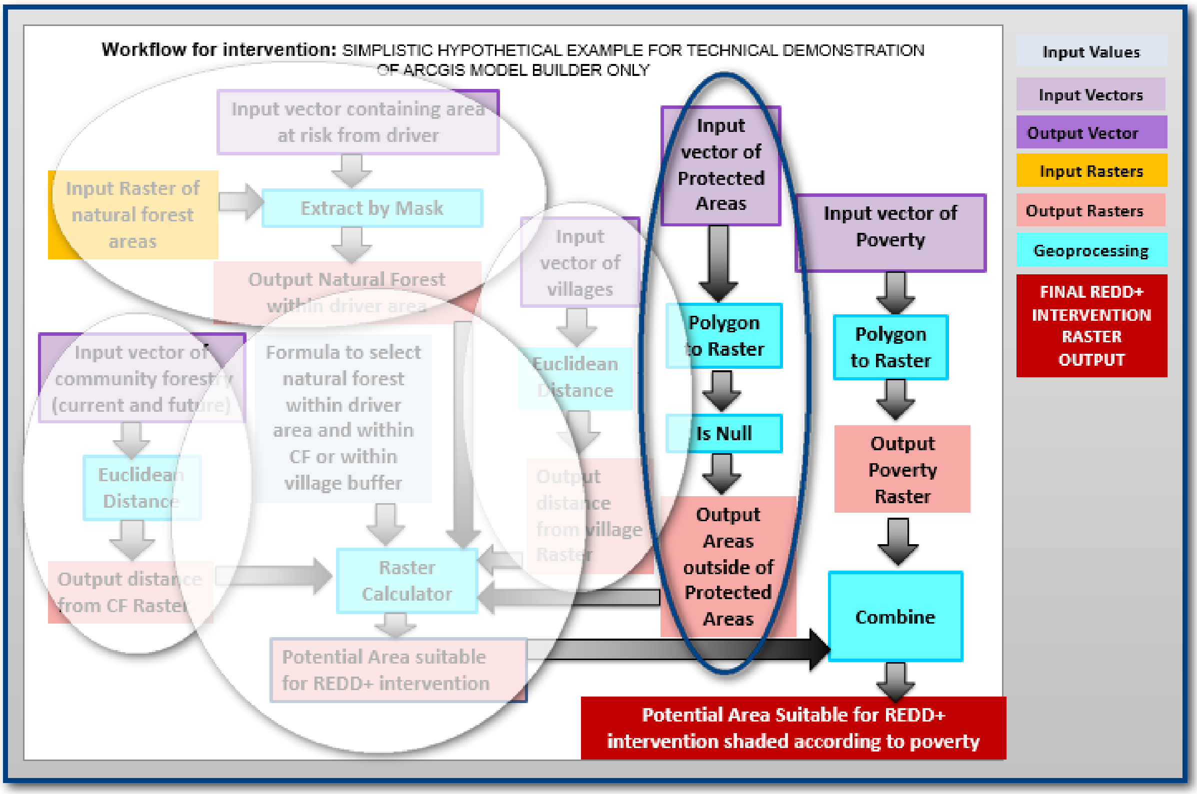
If both input values are true (non-zero), the output value is 1. If one or both inputs are false (zero), the output is 0.



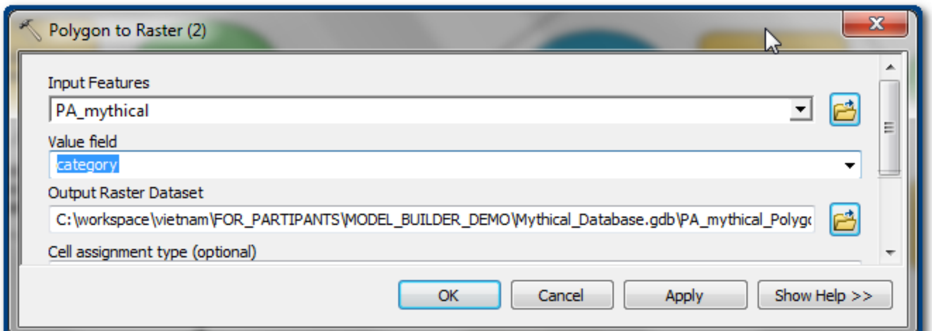
The model should now look like this:



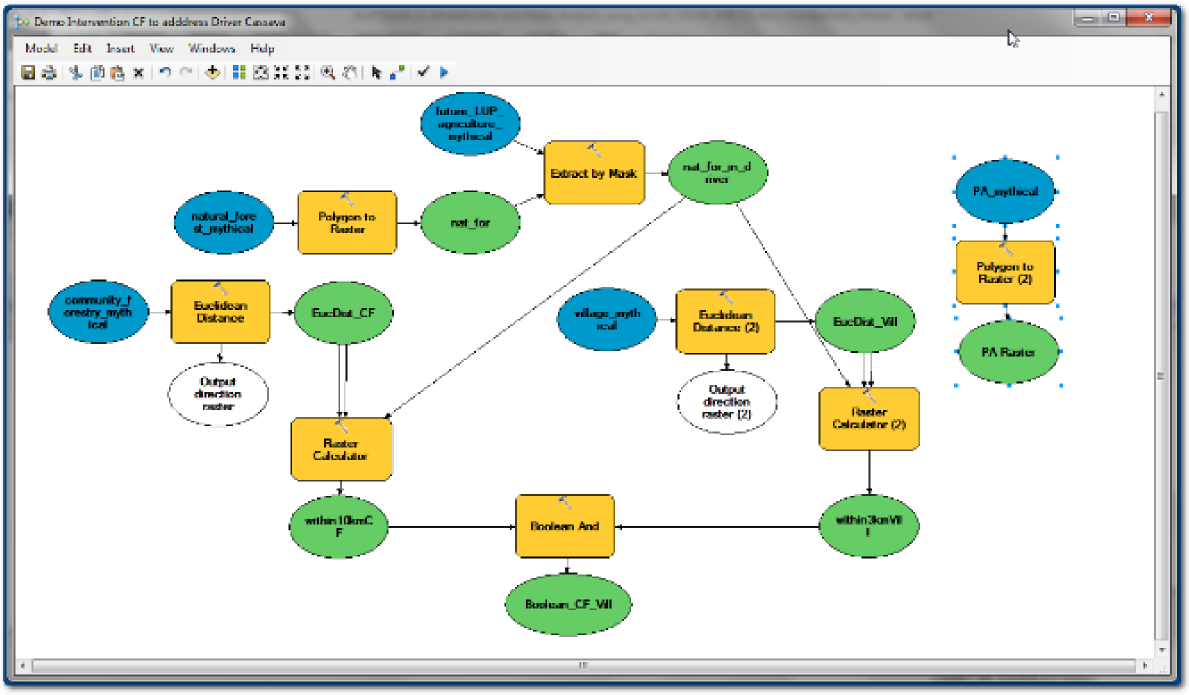
Check the model against the list of datasets we identified at the beginning and the workflow. We can see that the model matches the workflow but we still need to '**exclude strictly protected areas'**



The protected areas are a vector layer so add a **Polygon to Raster tool** to Model Builder.

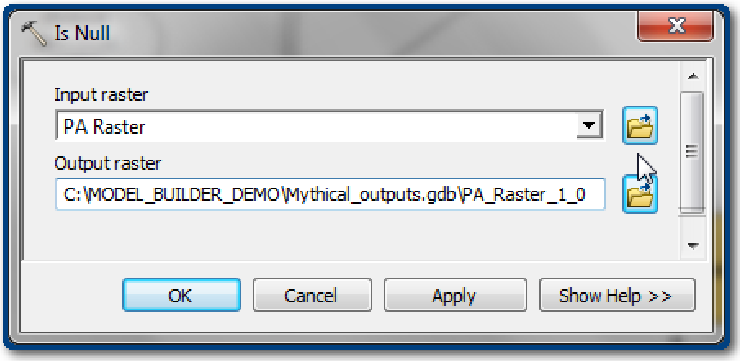


The model should now look like this:

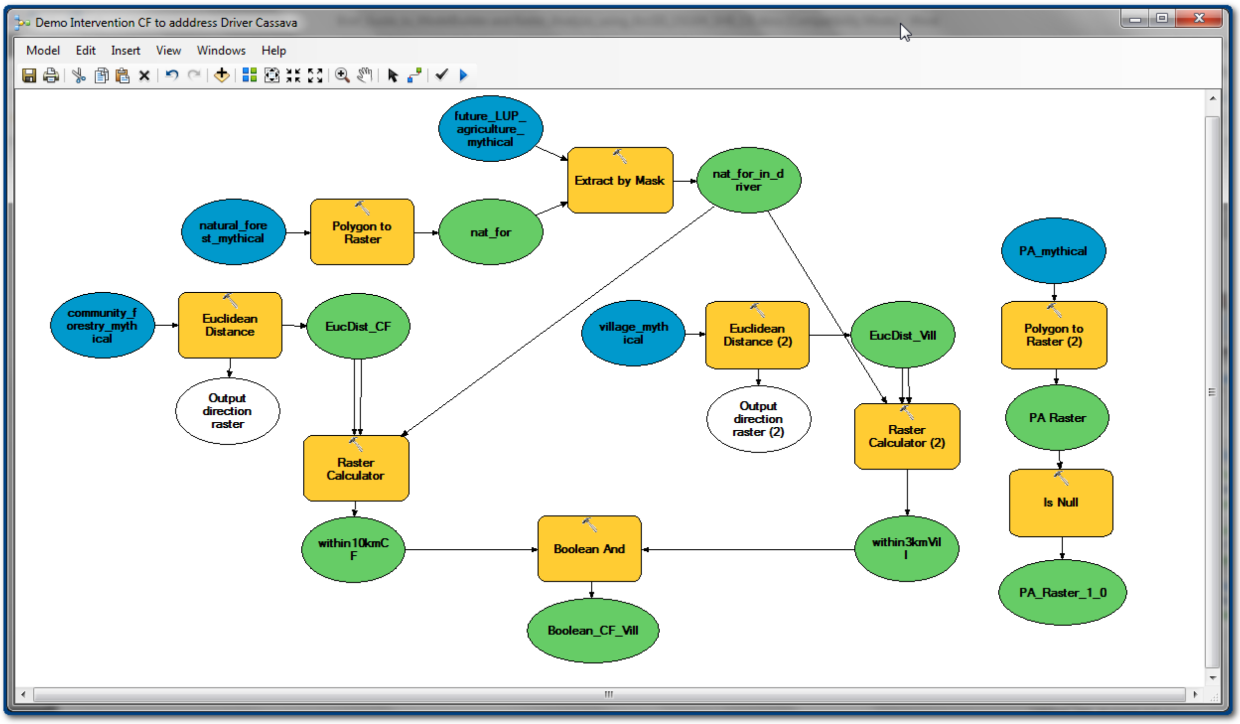


Now we want to select areas that are not under strict protection, so we will use the **Is Null tool** to create a **Not Protected Layer**. **The Is Null tool** returns a value of 1 if the input Value is **No Data** and a value of 0 for cells that are not No Data.

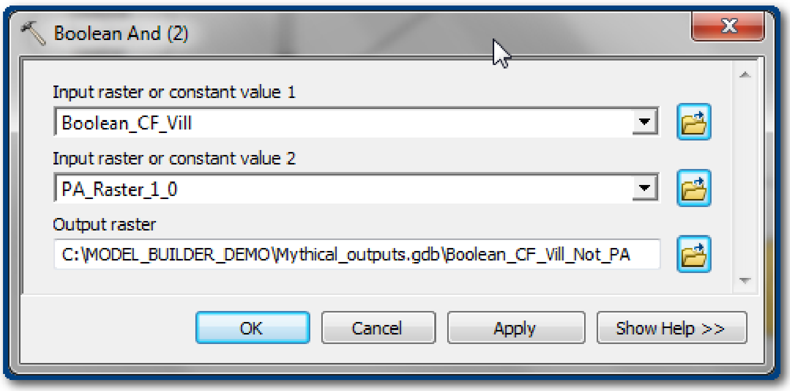
v. Search for the **Is Null tool** (or find it located in the **Spatial Analyst>>Math>> Logical** toolbox) to the model builder window.



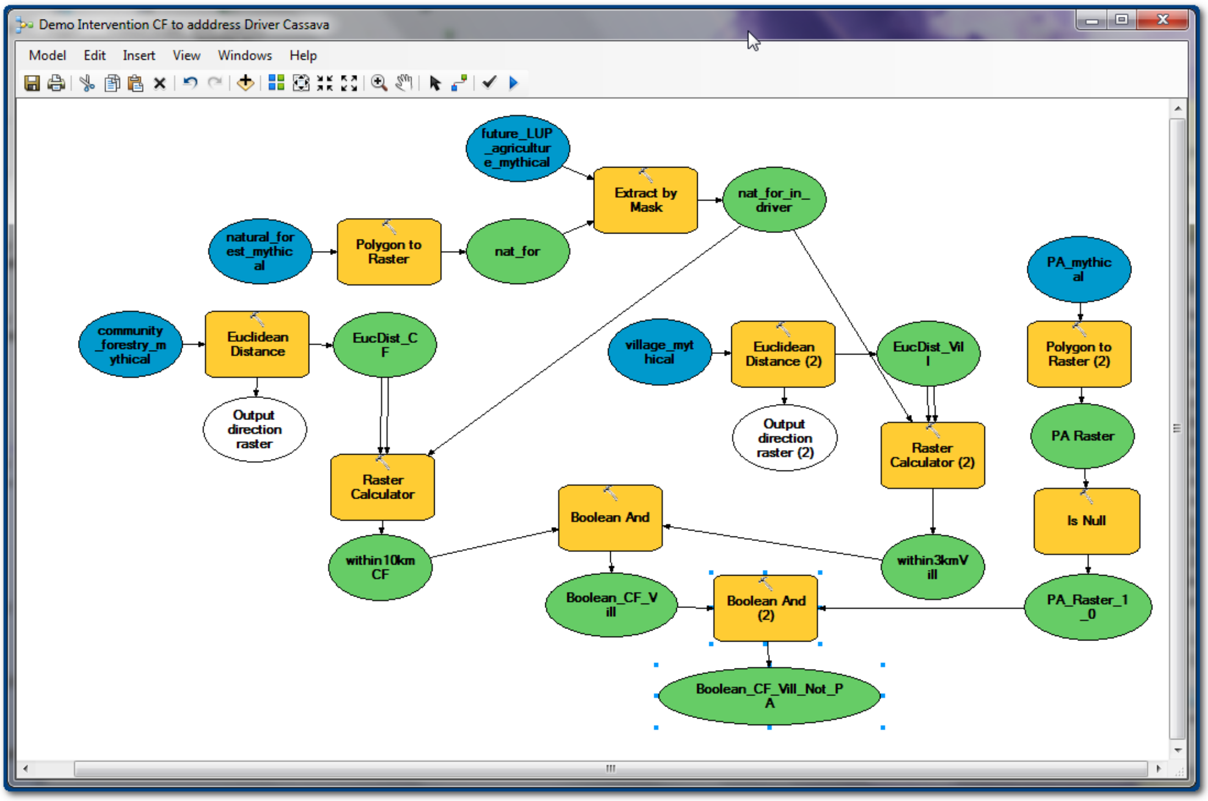
The model should now look something like this:



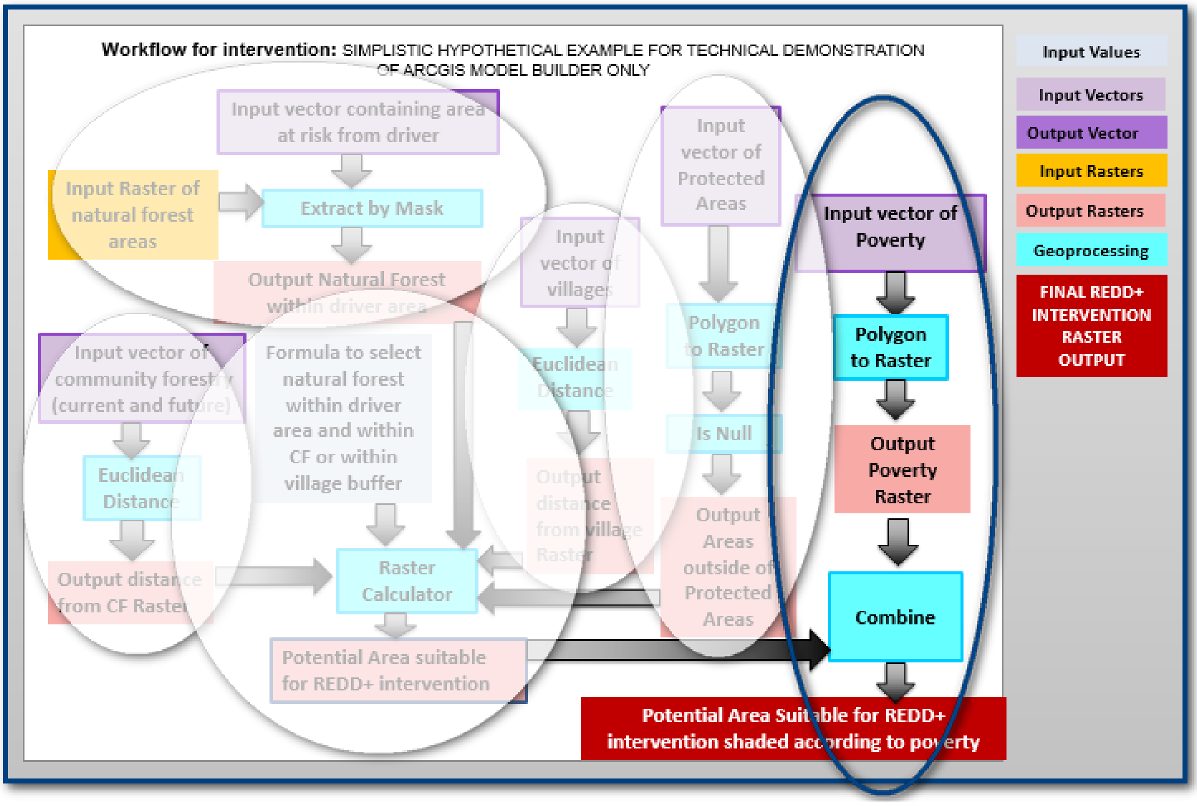
w. We use another **Boolean And tool** to select only those areas that are within 10 km of potential or current community forest sites and within 3km of villages that not in protected areas.



The model should now look like this:

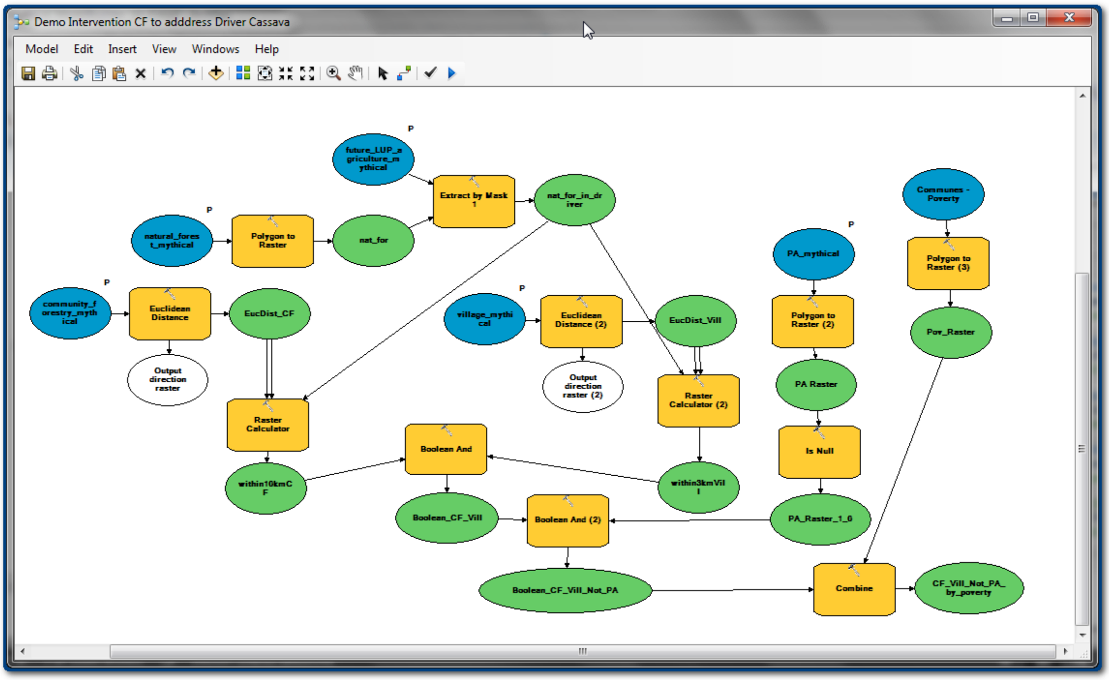


There is one last step in the original workflow, which is to classify the areas selected so that areas can be prioritized based on poverty. This will address one of the possible additional benefits of REDD+:



x. Add a **Polygon to Raster tool** to the Model and a **Combine tool** to overlay the Potential Area suitable for REDD+ Intervention with the Poverty Layer

The model should now look like this:



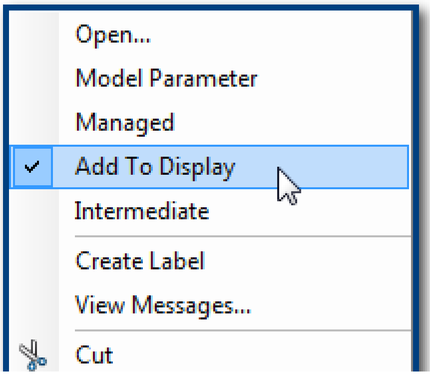
By default all the layers except for the last output layer are set as Intermediate, so output dataset are not automatically added to ArcMap

y. Right-click **on each of the green output layers**

- Click on **Intermediate** to **untick** it
- Click on **Add To Display** (so that when the tool is run the output is added to ArcMap)

z. Click **Save.**

The model is now ready to test.



4.1.6. Testing and modifying the model

a. From the Model Builder window Click on **Model>>Validate entire model** or the validate  button

b. Click **Model>>Run entire model** or click on run  button.

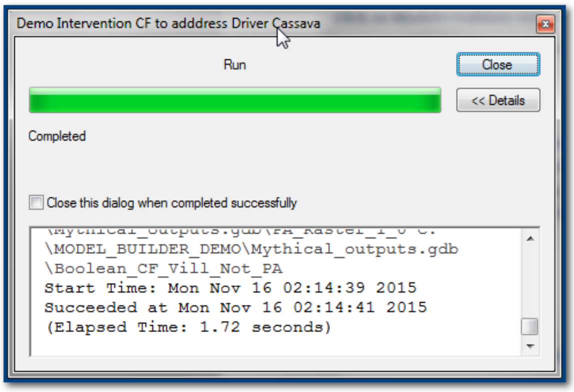
c. Make sure the **Close this dialog when completed successfully** box is **unticked** so that you can check the window has not reported any errors and that you can see how long the tool took to run.

As it is running through the model, Model Builder highlights each tool in red as it is processing it. Once a step in the model has successfully run, it highlights the step by putting a grey shadow around the tool. That way, if the model fails at any point you can easily see which tool it stopped on, as well as looking at the error messages in the tool window.

d. If the model runs it will display 'Completed' and will say 'Succeeded at ...' giving the current date and time (see right).

e. Click **Close** to close the model window once the workflow has completed**.**

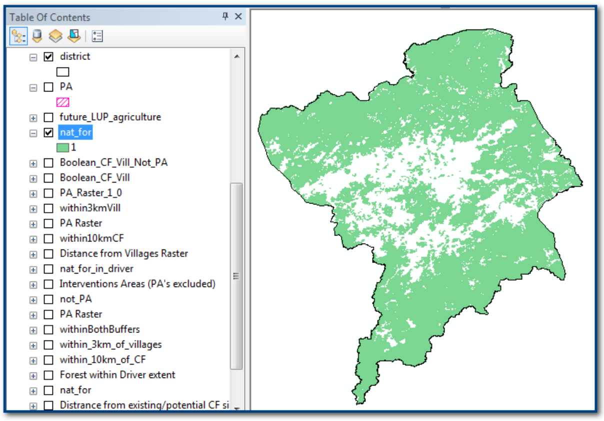
If the model fails you can choose to run the model from the point at which the model stopped by right-clicking on each individual tool step within the Model Builder window and clicking run.



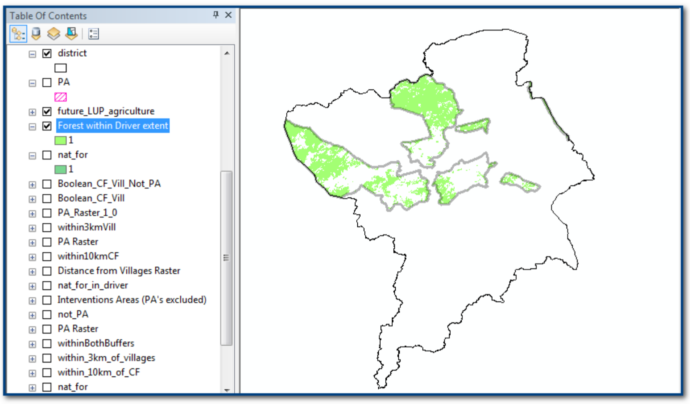
*****It is VERY IMPORTANT to check the outputs of the Model, to make sure that there were no errors in the set-up. Just because it has run, does not mean that it has run correctly to your design*****

For example, in our demonstration model the outputs are below:

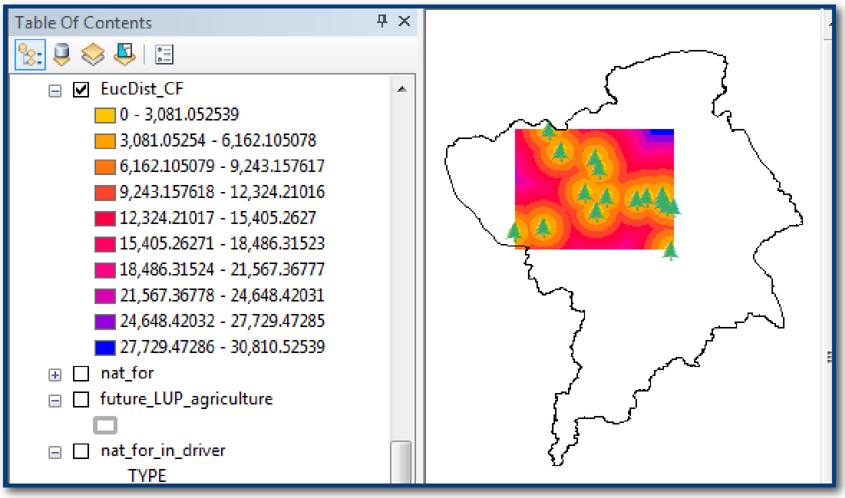
- The output of Step 1 is a natural forest raster layer (green). We can see it correctly covers the full extent of the district and there are white areas of **No Data** where there is no natural forest.



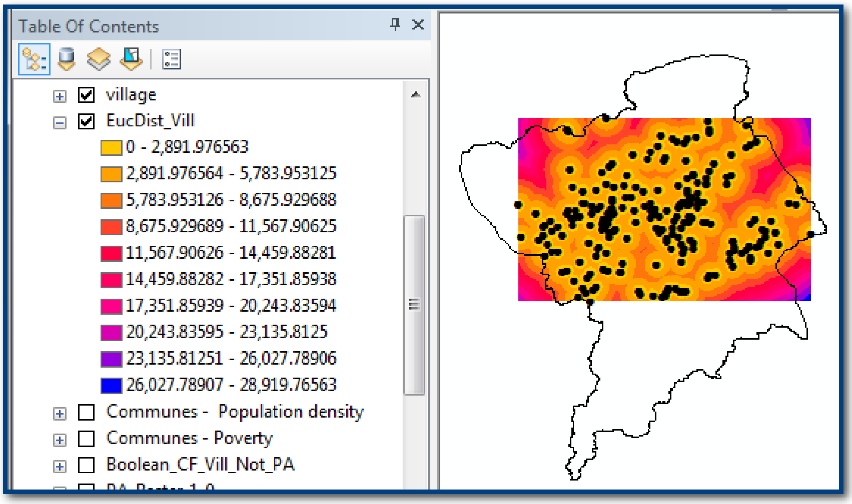
- The output of Step 2 is the natural forest within the driver extent. We can see it correctly covers the full extent of the driver areas and there are white areas of **No Data** where there is no natural forest.



- The output of Step 3 is the Euclidean distance from community forestry areas. We can see that it has <u>not done it</u> correctly. It only covers the extent of the community forestry shapefile. We wanted it to cover the full extent of the district.



- The output of Step 3 is the Euclidean distance from villages. We can again see that it has <u>not done it</u> correctly. It only covers the extent of the villages shapefile. We wanted it to cover the full extent of the district.



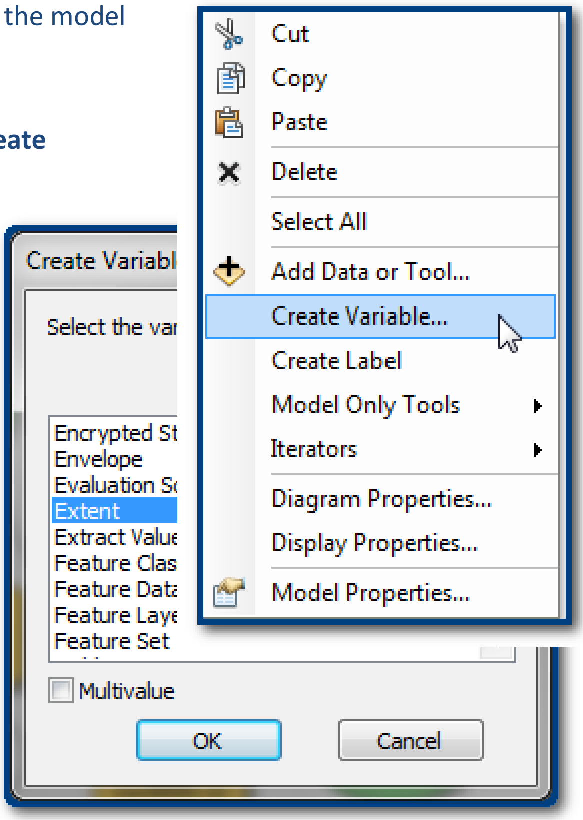
So we need to edit the model to make sure that during each step ArcGIS is processing data for the right extent (see next step).

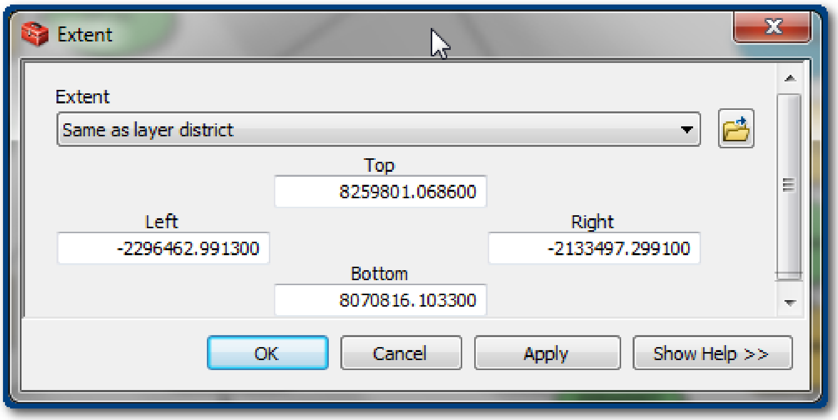
f. If you have closed your model, right-click on the model and click **Edit** to open it again

g. In a blank area of the model, right-click>>**Create Variable**

h. Scroll down and click on **Extent**

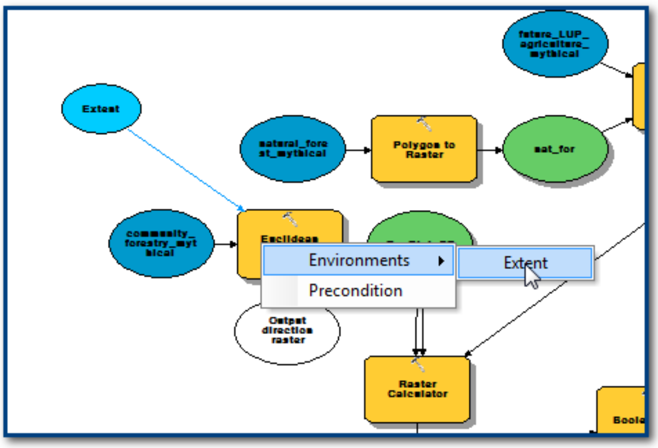
i. Double-click on the **Extent circle** and pick the **Same as district layer** to define the extent





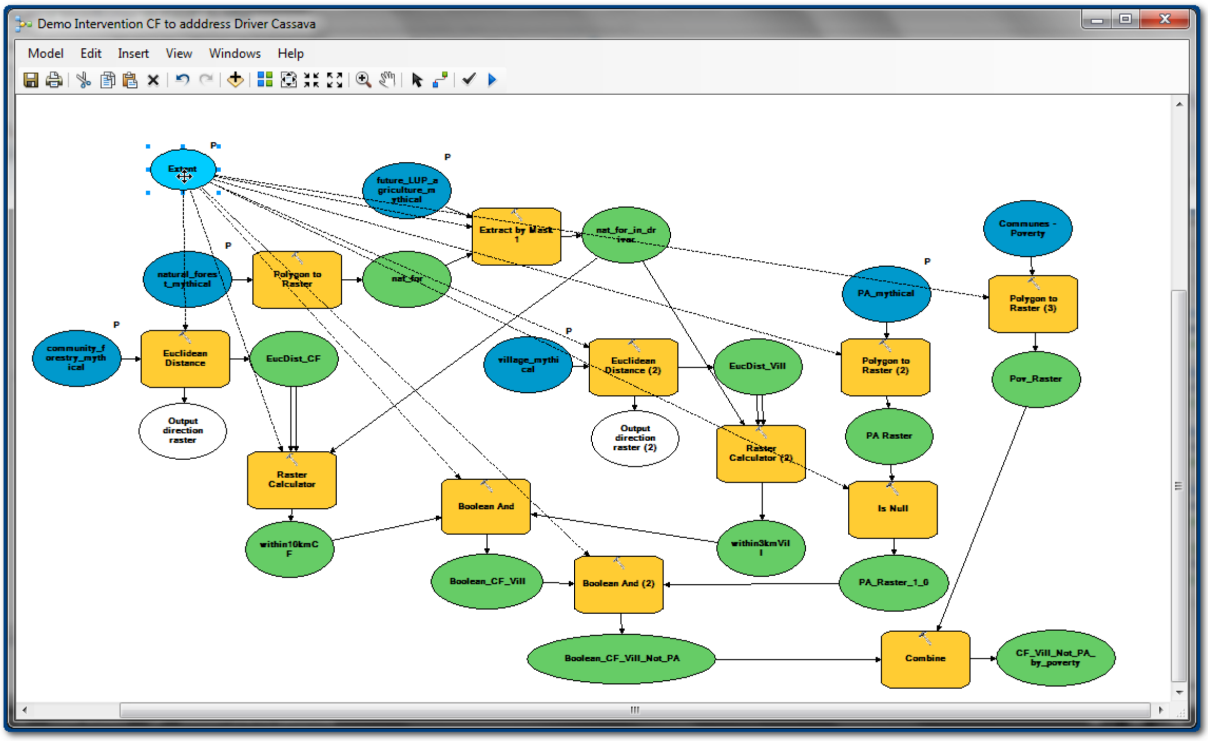
j. Click OK and the **Extent** variable turns light blue

k. Use the connector tool to connect the variable to the **Euclidean distance tool**, and choose the **Environments>>Extent** setting



l. To ensure the extent is correct for all the steps in the analysis, connect the extent variable to all the steps in the tool as done in step k.

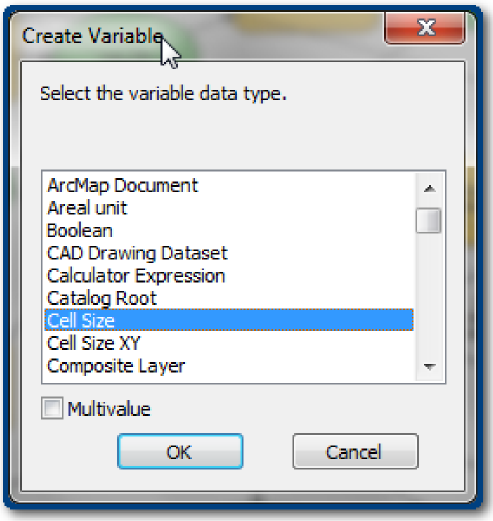
The model now looks messy but at least we can be sure that the full extent is being considered in each of the steps:

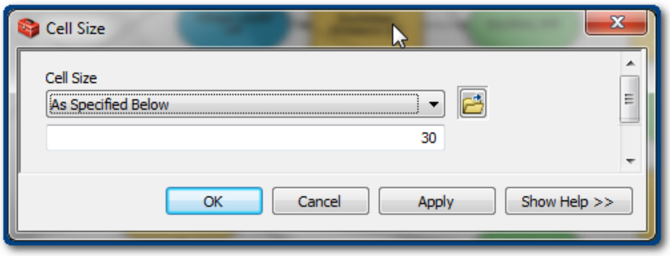


There are two other variables that can be added:

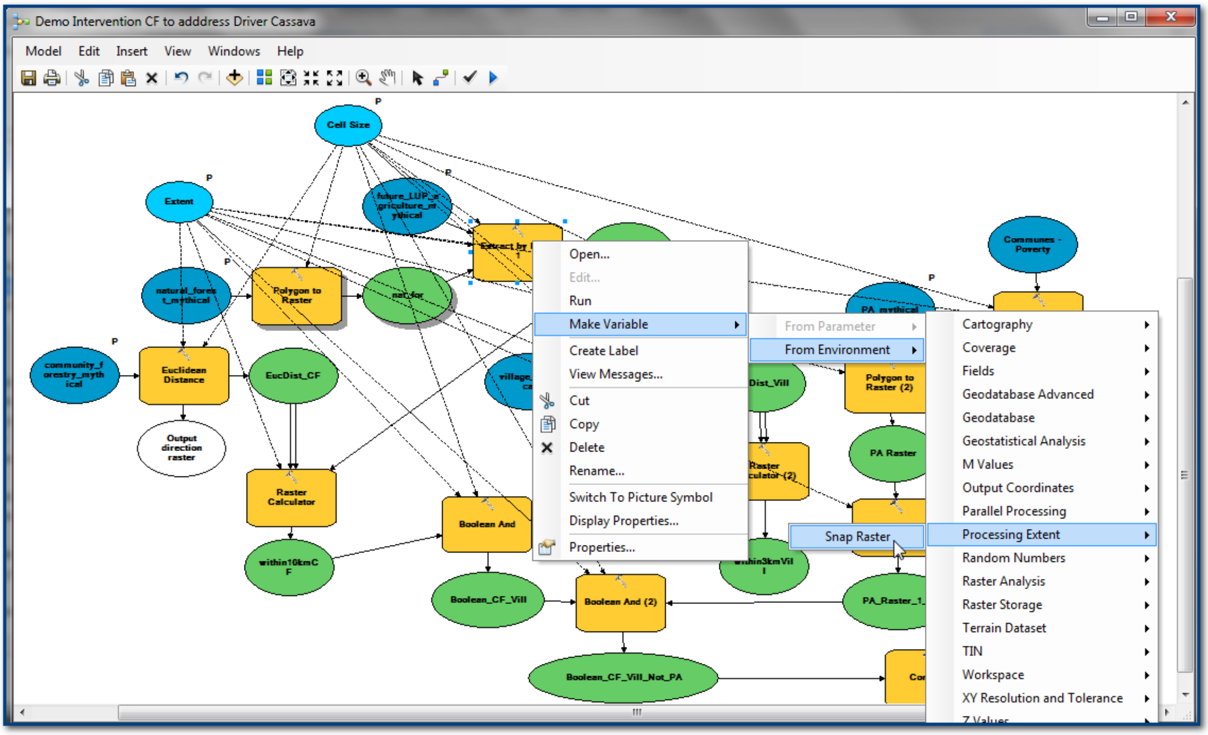
m. **Create Variable** as above, and this time select the **cell size** (see right)

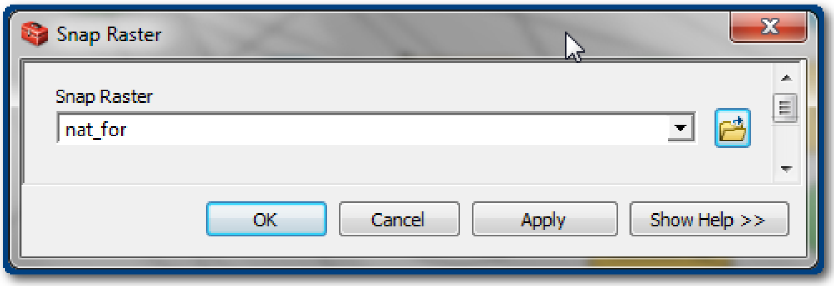
n. Use the connector tool to connect the variable to the all the tools in the model builder and connect it to the **Environments>>cell size** setting





o. Another way to add the Environment variable is to right-click on the **Extract by Mask tool** and click **>>Make variable >> From Environment>>Processing Extent >>Snap Raster**



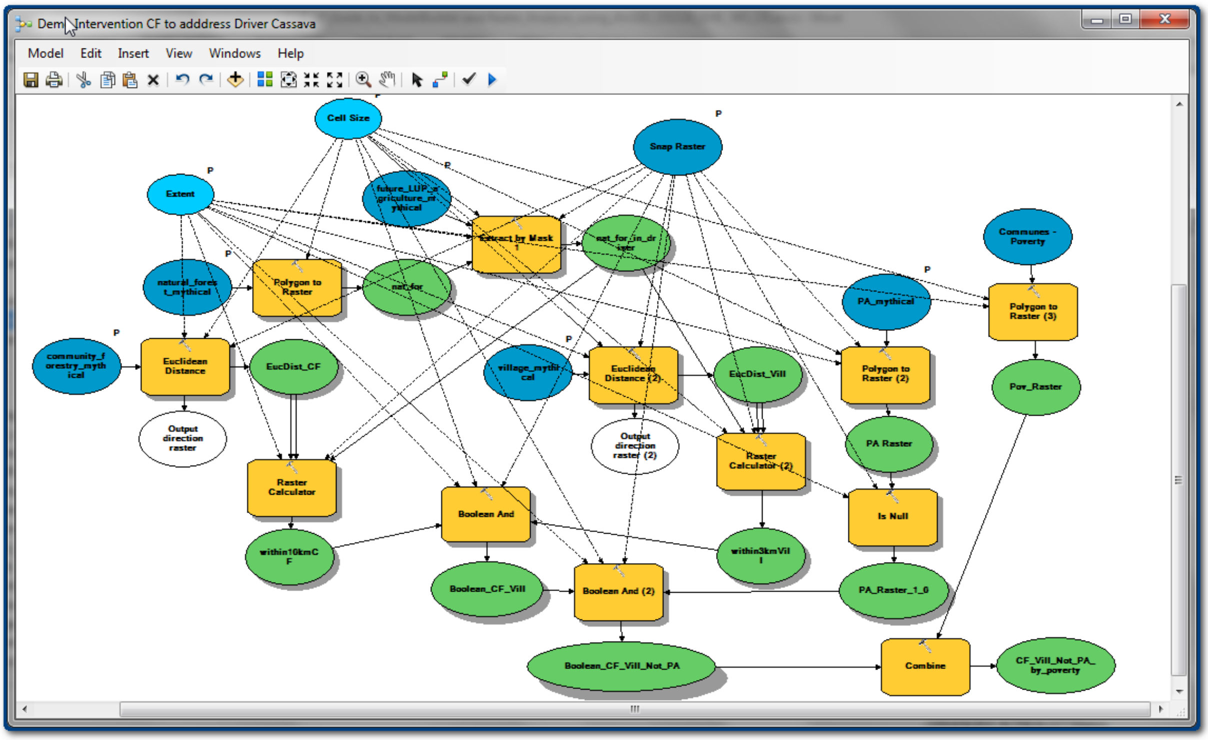


p. Set the **Snap Raster** to be the **nat_for** layer.

This is to ensure all the cells in the output layer align with the nat_for layer.

q. Use the **connect tool** to connect this **Snap Raster** to all the other **raster tools environment>>Snap Raster**

The model should now look similar to the one below:

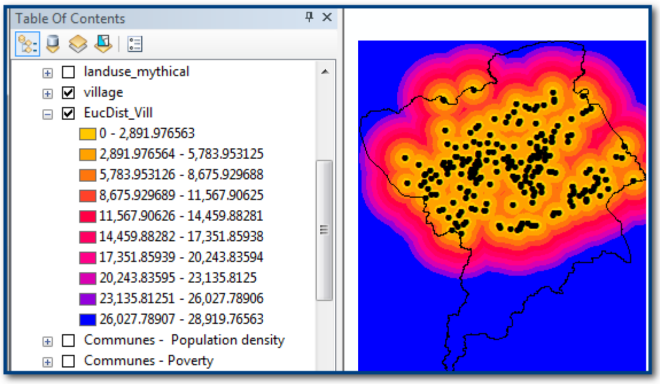


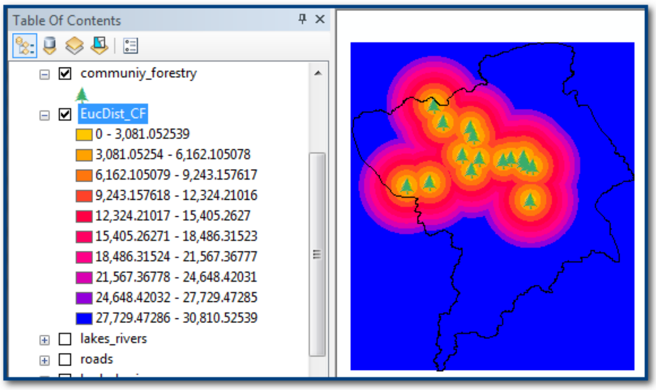
r. Click on **Model>>Validate Model**

s. Click on **Model >>Run Entire Model**

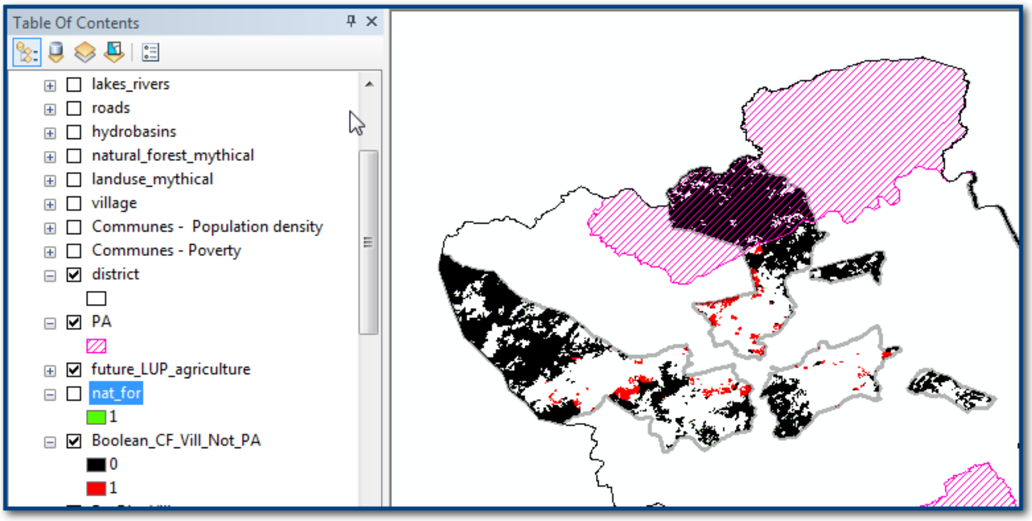
Now check the results again. You should see that the output at each of the steps covers the correct extent.

See that both distance to villages and to community forestry layers are now covering the full extent.



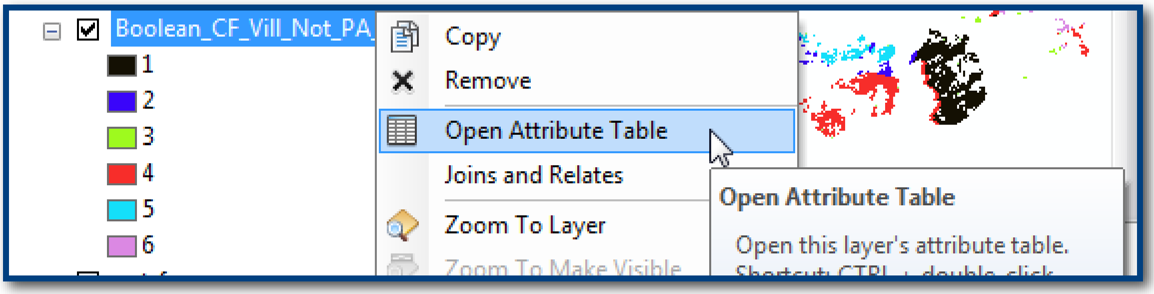


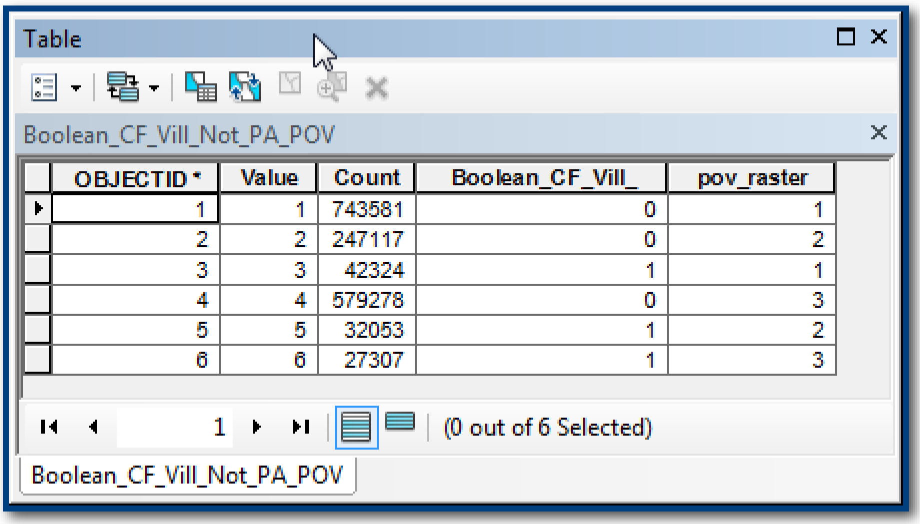
If we look at the final layer, we can see that the **red areas** are the areas that have been selected as areas within natural forest, that are within 10km of existing or proposed community forestry sites. The **black areas** are the natural forest areas that have been excluded. Note also that there are **no sites selected within protected areas**.



And if we look at the final combined layer we can shade the areas by poverty class. You will see that the poverty class has values of 1-6.

t. **Right-click on the combined raster e.g. Boolean_CF_Vill_Not_PA_POV** in this example and click **Open the attribute table**





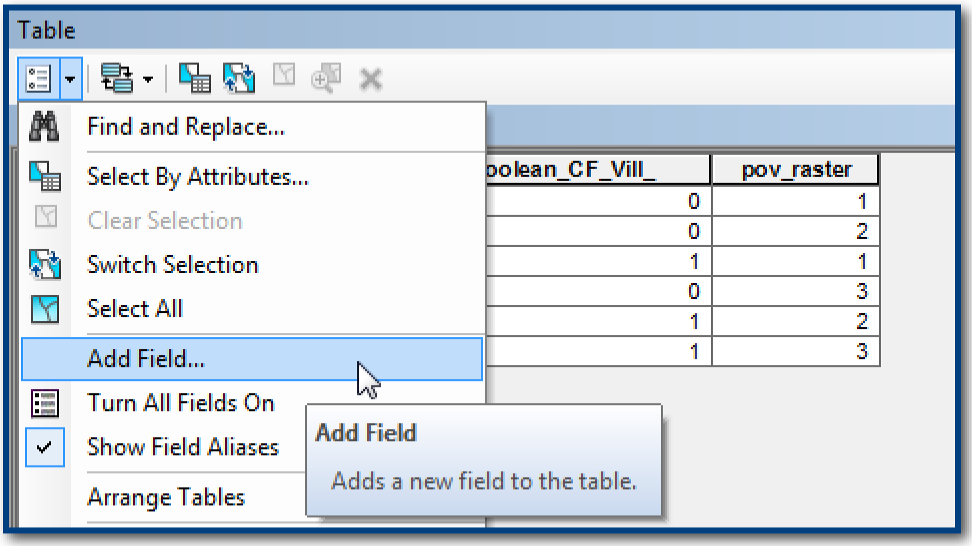
OBJECTID *	Value	Count	Boolean_CF_Vill_	pov_raster
1	1	743581	0	1
2	2	247117	0	2
3	3	42324	1	1
4	4	579278	0	3
5	5	32053	1	2
6	6	27307	1	3

See that the VALUE is a new unique ID that has been given each unique combination of input layers. See that the two additional fields are the names of the two input layers in the Combine. **Boolean_CF_Vill_** and **pov_raster** values.

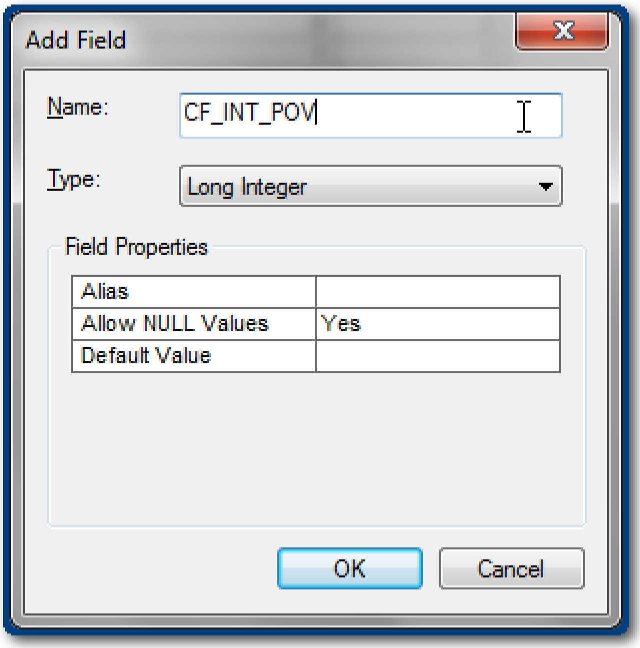
*Note also that the **Boolean_CF_Vill_** should logically have been **Boolean_CF_Vill_Not_PA_POV** but ArcGIS had truncated the name as it was too long for a field name.*

If we want to shade the areas that were selected as suitable for the intervention, by poverty class, we need to add a new field.

u. **Click on Table>>Add Field**



v. Add a field which will contain the Intervention Area classified by poverty class

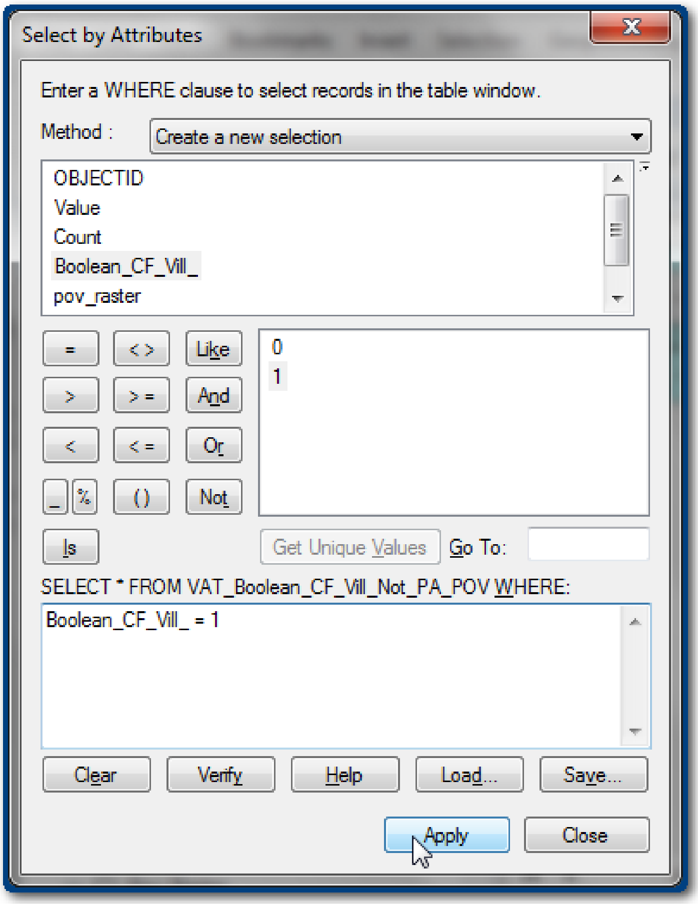


w. Click on the Table menu >>**Select by Attribute**

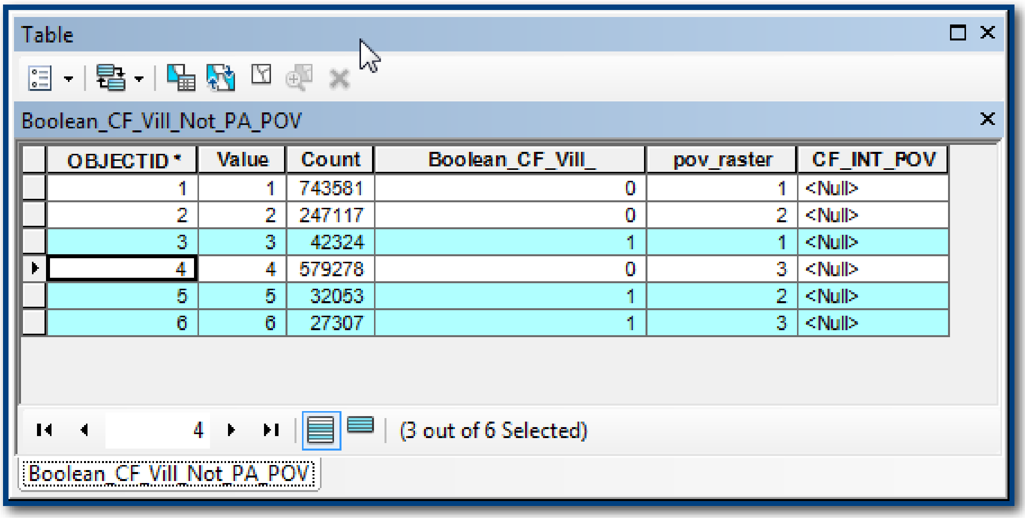
- Double-click on **Boolean_CF_Vill_** to bring it down into the white SELECT panel
- Click on **=**
- Click on **Get Unique Values** and double-click on **1**

The expression in the select window should read **Boolean_CF_Vill_ = 1**

- Click **Apply**



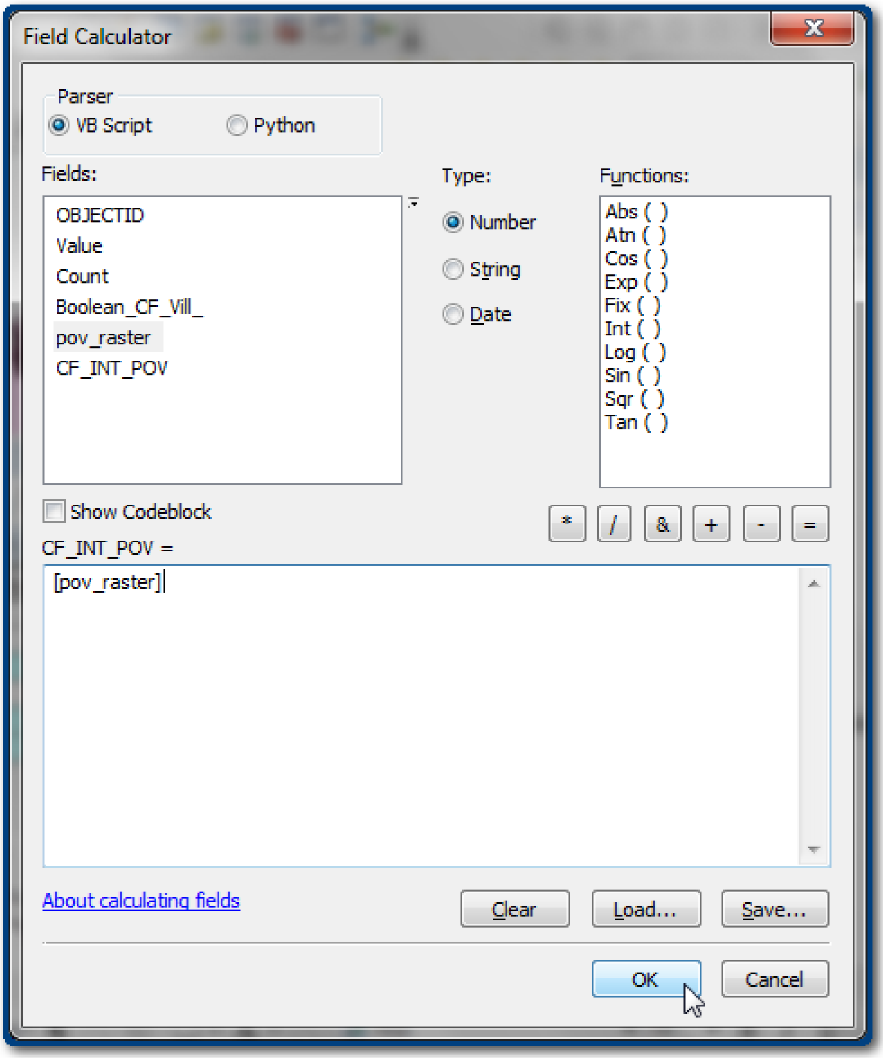
The selected records will be highlighted in blue:



OBJECTID *	Value	Count	Boolean_CF_Vill_	pov_raster	CF_INT_POV
1	1	743581	0	1	<Null>
2	2	247117	0	2	<Null>
3	3	42324	1	1	<Null>
4	4	579278	0	3	<Null>
5	5	32053	1	2	<Null>
6	6	27307	1	3	<Null>

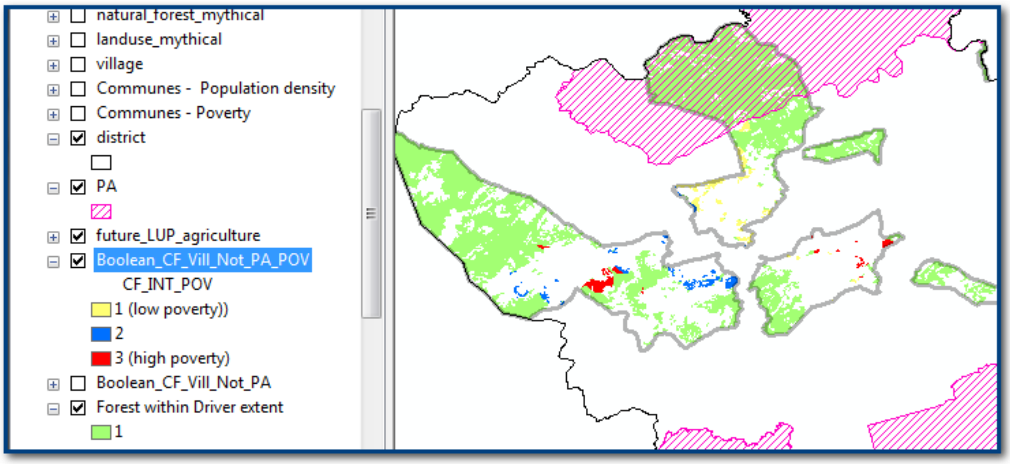
x. Right-click on the new field **CF_INT_POV** and click **>>Field Calculator**

y. Double-click on **pov_raster** to transfer the **pov_raster** values into the new field for the selected records.



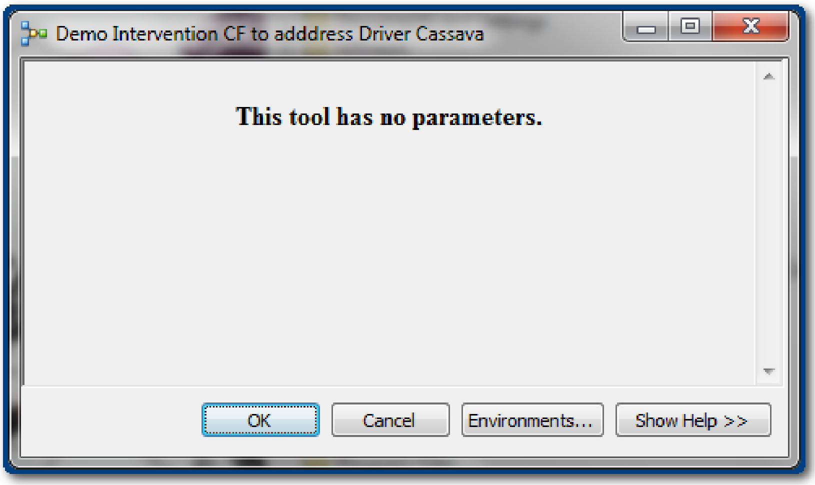
z. From the **Table Menu>>Clear Selection**, and close the table

Now you can shade the map on the **CF_INT_POV** field.



4.1.7. Enabling User Selection of Parameters

a. The model that you create using Model Builder can be run like any other tool from ArcToolbox. If you double-click on the tool it will bring up the following window.

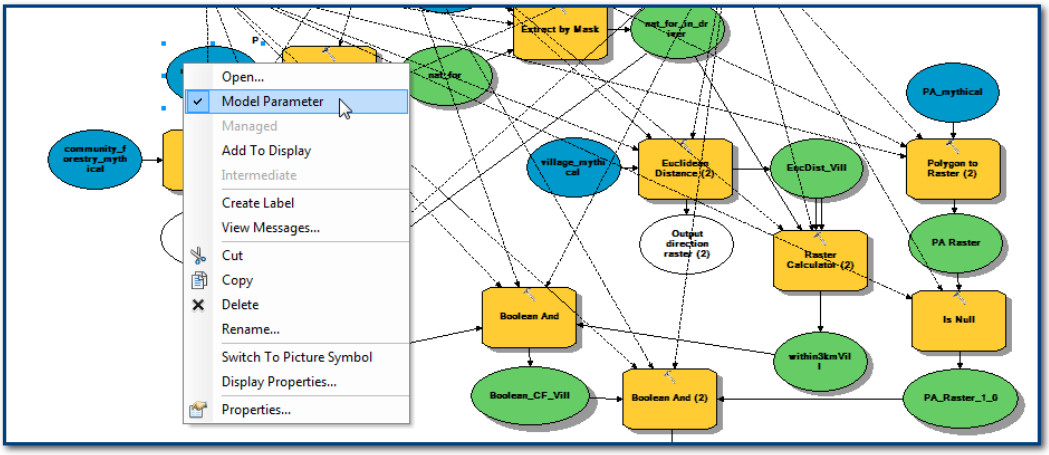


b. Click OK and the model will run.

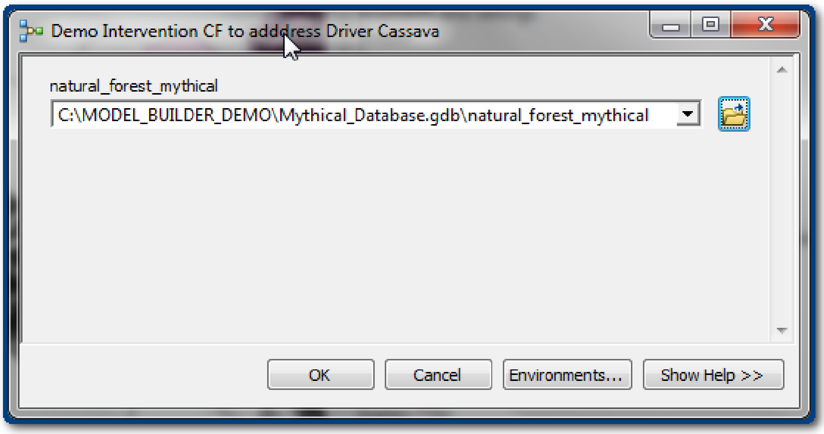
To make the tool more useful, you can allow a user to enter their own parameters so they can use the tool with their own data.

c. Right-click**>>Edit** to edit your model, then right-click on one of your input datasets **e.g. natural_forest_mythical** and **click Model Parameter**

d. Click **Save** and close your model.

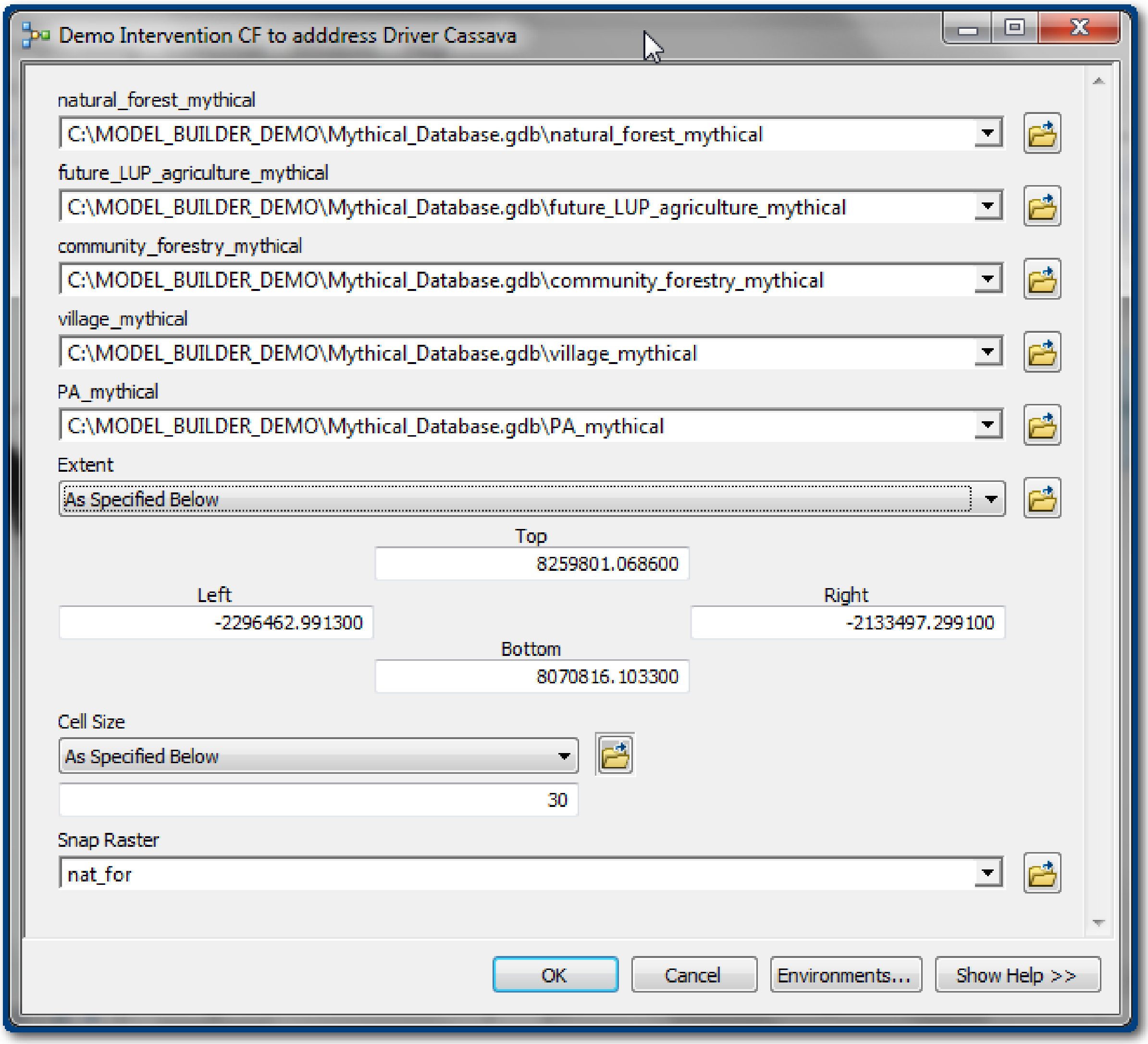


e. Double-click on your model and see that **the user is now able to select their own data for the natural forest layer.**



f. The inputs appear in the order you made them a parameter.

If you set all your inputs as parameters (including the **Extent**, **Cell Size** and **Snap Raster** variables) your tool will look something like this:



4.1.8. Symbology

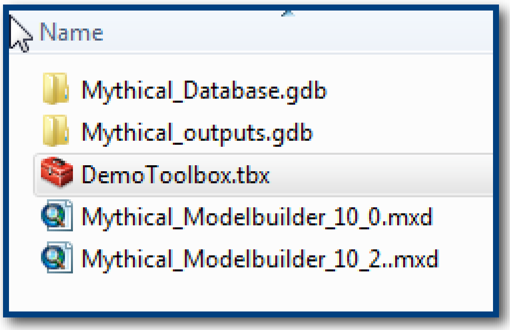
When the output of the tool is added to the Table of Contents, the Symbology of the layer changes each time. You can pre-set the Symbology for your results. In your model, right-click on the **Output Feature Class** and select **Properties**. In the **Layer Symbology tab**, you can assign a Layer File to determine the Symbology.

4.1.9. Documenting the model

It is important to have metadata for your tool, both for future reference and to facilitate sharing with your colleagues and others. To fill this in, right-click on your model, select **Item Description >>Edit** and fill in as much information as you want.

5. Sharing an ArcGIS Model with other users

The easiest way to share a tool you have created is to **simply copy the toolbox using Windows Explorer,** in the same way as you would any other file.

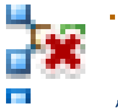


If you want to share the data as well as the tool, you can zip up the entire folder, or In ArcMap create a Geoprocessing Package. To do this browse to **Geoprocessing >>Results**. Expand the current session, **right-click on one of your results >>Share As>>Geoprocessing Package**.

Browse through the different sections and fill in where required. When this is complete, click **Analyze** on the toolbar to analyse your tool for any errors. If there are no errors, click **Share,** and your Geoprocessing Package will be ready to be distributed to your colleagues!

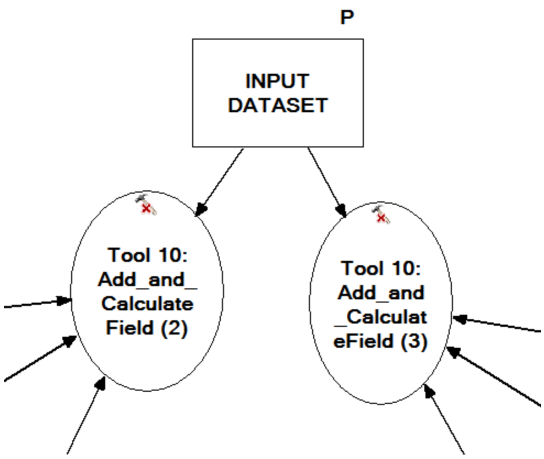
More information can be found here:
http://resources.arcgis.com/en/help/main/10.2/index.html#/What_is_a_geoprocessing_package/00570000004p000000/

6. Fixing a broken model

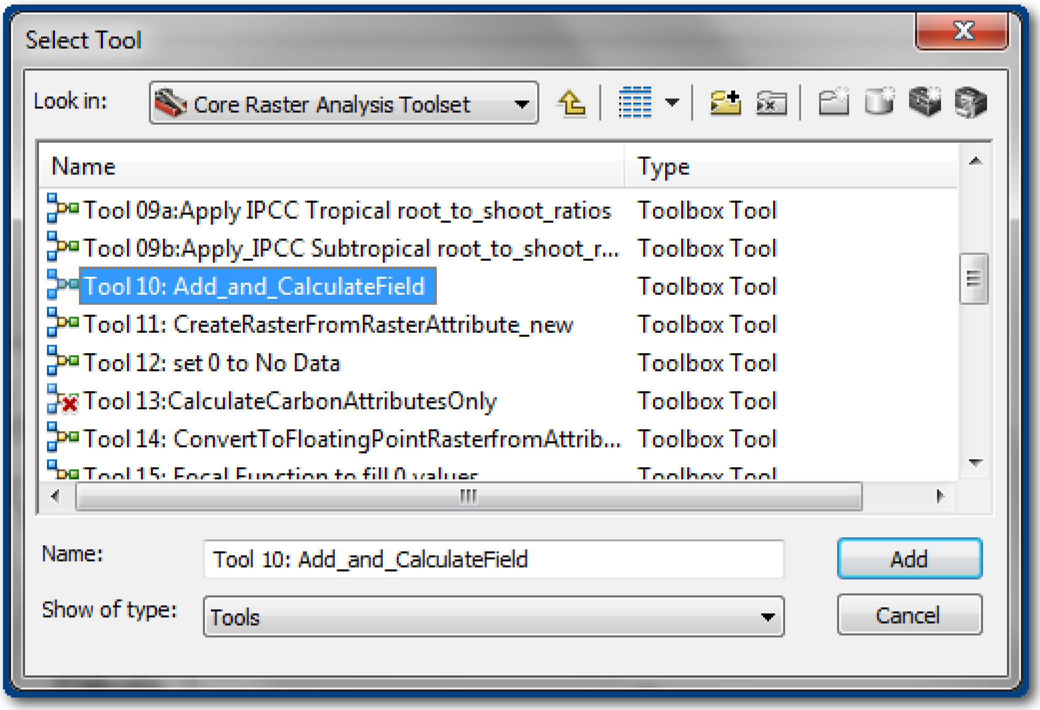
If you see a broken link in a model  , it means that ArcGIS cannot find at least one of the tools that the model uses. This can be fixed by pointing the model to the right files.

a. **Right-click>>edit the model**

See that the broken parts of the tool are flagged with a broken tool  symbol



b. To fix the broken link, double-click on one of the broken tools within the model builder window and navigate to the location of the tool.



c. Click **Add** to add the tool.

d. Click **Save** to save the tool. Close Model Builder and see that the  has disappeared.

Some custom tools break when ArcGIS versions change, usually because a geoprocessing tool has changed or works slightly differently to the old version. In these instances it may be necessary to select and delete that step in the Model Builder window and drag in a new version of the tool.

A Annex 1: Guidance notes on raster analysis and geoprocessing tools in ArcGIS

A.1. Vector vs raster

A.1.1. Differences between vector and raster datasets

Data can be divided into two broad types: **discrete data**, which represents objects and can be stored either as **vector features or as raster**, and **continuous data**, which are usually (and best) stored in **raster data formats.**

With **discrete data**, there is a clear starting point and ending point. It is easy to define the boundary of the object - a lake, for example, is a discrete object as its boundary can be definitively established. Other examples of discrete objects include administrative units (i.e. boundaries defining a region or country) or boundaries between different land-cover classes. Cells belonging to each discrete object are given the same values and are stored as Integers. **Integer** rasters cannot store decimal places. Discrete data are sometimes called categorical data.

With **Continuous data,** values change smoothly across a landscape or surface and there are no discrete class breaks. Cells have individual values, with neighbouring cells often being different from each other, but having a relationship between each other. Examples of continuous data include elevation, slope and aspect, or data such as Normalized Difference Vegetation Index (NDVI) from a remotely sensed image. The numeric cell values can be either be stored as **integer** or **floating-point** rasters. **Floating-point** rasters are necessary for storing **continuous** data which have value attributes containing decimal places.

The table below summarises some of the differences between vector and raster data

Vector data	**Raster data**
Points, lines, polygons	Surface of regular sized **grid cells or pixels**
Storing discrete data e.g. data with discrete boundaries such as country boundaries, roads, streets, location of a place.	Raster cells storing either **discrete** data e.g. classified landcover data or **continuous** data e.g. elevation or satellite imagery. Raster data are especially suited to continuous data.
Complex spatial relationships can exist within and between vector layers. Topology rules can be set up (i.e. relationships between adjacent or neighbouring features) such that they do not overlap.	Spatial relationships are based on the location, size and alignment of the cells/pixels only.
Vector data have attribute tables. For each feature (in singlepart datasets) or group of features (in multipart datasets), fields can be added to store additional information about the polygon(s), point(s) and line(s)	For integer rasters, cell values may be stored in summary tables known as Value Attribute Tables (VATs). Additional information can be stored in fields for groups of cells with the same value. Attributes cannot be added to individual cells/pixels. *Note: ArcGIS software can create and see raster attribute tables but most other GIS software (e.g. QGIS) do not recognize tables for raster data.* *Important note: Floating point rasters <u>do not</u> have attribute tables. The workaround to create an Integer raster from a floating point raster, without the data*

	values being generalized, is described in the Exploring Multiple Benefits toolbox manual, which can be accessed at: http://bit.ly/GIStools-redd ***Extreme care must be taken with this workaround*** *as it involves multiplying values by 100 to allow conversion to an integer grid (thus preventing loss of information after the decimal point). A field showing these values will need to be subsequently divided by 100 to restore the true values.*
Query, select and overlay analysis is possible. Overlay of too many datasets can result in small sliver polygons where boundaries do not exactly match. This tends to slow down analyses or when severe, can cause analyses to fail when combined data become too large for processing.	As well as query, select and overlay analysis, there are many analyses better suited or only possible with raster data. Data are transformed to a consistent cell resolution to perform analyses, allowing quick and efficient analysis of multiple datasets.

A.1.2. Vector vs raster analysis

Many analyses are better suited or only possible with raster data. You can create, query, map, and analyse data with both vector and raster data. However, vector analysis only allows you to undertake some simple overlay operations compared to some of the more complex multi-criteria analyses options that are available with rasters. Raster analysis functions also allow you to do integrated raster and vector analysis, derive new information from existing data and run multi-criteria analysis queries across numerous data layers.

Table 2 provides a brief overview of some of the common vector and raster functions. This will help identify and locate raster functions that are equivalent/similar to vector functions. There may also be other functions that perform similar tasks which are not listed. **It is important to read the help for the individual tools as the table gives only a general indication of the function of the different tools listed. Individual tools may perform slightly differently**. The descriptions and tool names are mainly based on information in the ArcGIS and QGIS Processing toolboxes. In ArcGIS from the main menu click on **Help >>ArcGIS Desktop Help** or access Help Information held within the individual tools in ArcGIS toolbox. In QGIS see the QGIS user guide accessible from the main menu **Help>>Tool** or by clicking on the **Help** button on an individual tool dialogue window. Note that in QGIS help within the individual tools is not always present and you will see the message “Sorry, no help is available for this algorithm”.

Table 2: Brief overview of some of the common vector and raster functions

Extraction tools	Example Vector Tools		Example Raster Tools	
	ArcGIS	**QGIS**	**ArcGIS**	**QGIS**
Vector: Extracts input features that fall within the boundaries of the clip features. Features that fall partially within clip layer feature(s) are split along the boundary of the clip layer feature(s). **Raster:** Extracts the cells of a raster that correspond to the areas defined by a mask, polygon, within a circular radius or bounding box (extent).	**Clip**	**Clip (QGIS)** **Clip vectors by extent (OGR)** **Clip vectors by polygon (OGR)** **Clip points with polygons (SAGA)** **Cut vector layer (SAGA)** **v.overlay (GRASS)** **v.select (GRASS)**	**Extract by Mask** **Extract by polygon** **Extract by Rectangle** **Extract by circle**	**Clip raster by extent (GDAL)** **Clip raster by mask layer (GDAL)** **Clip raster with polygon (SAGA)**
Vector: Extracts features / table records using expressions. ArcGIS uses Structured Query Language (SQL). QGIS uses it's own expressions which are documented in the user guide https://docs.qgis.org/2.14/en/docs/user_manual/working_with_vector/expression.html **Raster:** Extracts the cells of a raster based on a logical query. Again ArcGIS and QGIS differ in their syntax. In QGIS syntax may also differ dependant upon whether a QGIS, GDAL, SAGA or GRASS.	**Select** **Table Select**	**Select by attribute (QGIS)** **Select by attribute sum (QGIS)** **Select by expression (QGIS)** **v. extract (GRASS)**	**Extract by Attribute**	**Add raster values to points (SAGA)** **Add raster values to features (SAGA)** **Raster values to points (SAGA)** **Raster values to features (SAGA)**

Extraction tools	Example Vector Tools		Example Raster Tools	
	ArcGIS	**QGIS**	**ArcGIS**	**QGIS**
Vector: Splitting the Input Features to create subset of multiple output layers from a split layer. **Raster:** Creates a tiled output from an input raster dataset.	**Split**	**Split vector layer (QGIS)**	**Split Raster tool**	**Gdal2tiles (GDAL)**
Vector: Selects features in a layer based on their relationship to features in another layer	**Select layer by Location**	**Extract by location(QGIS)**	**N/A**	**N/A**

Proximity tools	Example Vector Tools		Example Raster Tools	
	ArcGIS	**QGIS**	**ArcGIS**	**QGIS**
See summary here:- http://resources.arcgis.com/EN/HELP/MAIN/10.1/index.html#//018p00000007000000 **Vector:** Straight line distance tools to create buffers at specified distances **Raster:** Various distance tools that do more than straight line distance e.g. to take into account of elevation etc.	**Buffer** **Multiple Ring Buffer**	**Fixed distance buffer (QGIS)** **Fixed distance buffer (SAGA)** **Variable distance buffer (QGIS)** **Variable distance buffer (SAGA)** **v.buffer.distance (GRASS)** **v.distance (GRASS)** **v.distance.toattr (GRASS)**	**Euclidean Distance** **Cost distance** **Path distance**	**Proximity (raster distance) (GDAL)** **r.grow.distance (GRASS)** **r.grow (GRASS)** **r.cost (GRASS)** **r.cost.full (GRASS)** **r.cost.full.raster (GRASS)** **r.spreadpath (GRASS)** **Least cost paths (SAGA)** **Accumulated cost (anisotrophic) (SAGA)** **Accumulated cost (isotrophic) (SAGA)**

Proximity tools	**Example Vector Tools**		**Example Raster Tools**	
	ArcGIS	**QGIS**	**ArcGIS**	**QGIS**
Vector: Each Thiessen polygon contains only a single point input feature. Any location within a Thiessen polygon is closer to its associated point than to any other point input feature. **Raster:** divides an area up and allocates each cell to the nearest input feature that has a value.	**CreateThiessen polygons**	**Voronoi polygons (QGIS)** **v.voronoi (GRASS)**	**Euclidean allocation**	**r.grow.distance (GRASS)**

Statistics tools	**Example Vector Tools**		**Example Raster Tools**	
	ArcGIS	**QGIS**	**ArcGIS**	**QGIS**
Vector: Creates a table with statistics based on polygon contiguity (overlaps, coincident edges, or nodes). **Raster:** Calculates for each input cell location, a statistic of the values within a specified neighbourhood around it.	**Polygon Neighbours**	**NNjoin (QGIS plugin)**	**Neighbourhood toolbox: Focal Statistics tool**	**Raster layer statistics (QGIS)**
Reads a table and a set of fields and creates a new table containing unique field values and the number of occurrences of each unique field value.	**Frequency**	**Frequency Analysis (QGIS)**	For Integer rasters the value attribute table is similar to a frequency table in that it contains the unique field values and the count of the number of cells	**Unique values count (SCRIPT)** **Extract raster values to CSV (SCRIPT)**

Statistics tools	Example Vector Tools		Example Raster Tools	
	ArcGIS	QGIS	ArcGIS	QGIS
Calculates summary statistics for field(s) in a table	**Summary Statistics**	**Basic Statistics for numeric fields (QGIS)** **Basic statistics for text fields (QGIS)** **Statistcs by Categories (QGIS)**	N/A	N/A
Calculates statistics on values of a raster within the zones of another dataset. Summarizes the values of a raster within the zones of another dataset and reports the results to a table	N/A	**N/A**	**Zonal statistics tool, Zonal statistics as table tool**	**r.stats (GRASS)** **r.univar (GRASS)** **r.statistics (GRASS)** **Zonal raster statistics (SAGA)** **Zonal Statistics (QGIS)** **Raster statistics for polygons (SAGA)**
Calculates cross-tabulated areas between two datasets and outputs a table.	N/A	N/A	**Tabulate Areas**	**Cross-Classification and Tabulation (SAGA)**
Calculates a per-cell statistic from multiple rasters. The available statistics are Majority, Maximum, Mean, Median, Minimum, Minority, Range, Standard Deviation, Sum, and Variety.	N/A	N/A	**Cell statistics**	**Statistics for rasters (SAGA)** **r.report (GRASS)**

Overlay tools	Example Vector Tools		Example Raster Tools	
	ArcGIS	QGIS	ArcGIS	QGIS
Vector: Union tool - All features and their attributes will be written to the output layer. **Identity tool** - The input features or portions thereof that overlap identity features will get the attributes of those identity features. **Intersect tool** - Features or portions of features which overlap in all layers and/or feature classes will be written to the output layer. **Raster:** Combines multiple rasters so that a unique output value is assigned to each unique combination of input values. Note: Also see Multicriteria analysis tools which are listed separately in this table under the Raster only tools.	**Union** **Identity** **Intersect**	**Union (QGIS)** **Polygon identity (SAGA)** **Intersect (SAGA)** **Intersection (QGIS)** **v.overlay (GRASS)** **v.patch (GRASS)**	**Combine** (Note: any cells in the input rasters that have No Data will be excluded from the output raster even if those cells contain data in some of the input rasters)	**r.cross (GRASS)** **Cross-classification and tabulation (SAGA)**
Vector: Computes a geometric intersection of the Input Features and Update Features. Attributes and geometry are updated by the update features in the output layer. **Raster:** Mosaics multiple raster datasets into a new raster dataset.	**Update**	**Polygon update (QGIS)**	**Mosaic to New Raster**	**Mosaic raster layers (SAGA)** **r.patch (GRASS)**

Raster only tools	Example Raster Tools	
	ArcGIS	**QGIS**
Multicriteria Analysis tools: Overlay analysis tools allow you to apply weights to several inputs and combine them into a single output. The most common application is for suitability modelling.	**Fuzzy Membership** Transforms the input raster into a 0 to 1 scale, indicating the strength of a membership in a set, based on a specified fuzzification algorithm. **Fuzzy Overlay** Combine fuzzy membership rasters data together, based on selected overlay type. **Weighted Overlay** Overlays several rasters using a common measurement scale and weights each according to its importance. **Weighted Sum** Overlays several rasters, multiplying each by their given weight and summing them together. For more detail see the Help pages:- http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/An_overview_of_the_Overlay_tools/009z000000rm000000/ and this article on Fuzzy logic: http://www.esri.com/news/arcuser/0410/fuzzylogic.html	**Fuzzify (SAGA)** Translates grid values into fuzzy set membership as preparation for fuzzy logic analysis. **Fuzzy union (or) (SAGA)** Calculates the union (max operator) for each grid cell of the selected grids. **Fuzzy Intersection (and)** Calculates the intersection (min operator) for each grid cell of the selected grids. **Ordered weighted averaging (SAGA)** Overlays several rasters using a common measurement scale and weights each according to its importance. See the help for the Grid Calculus tools in SAGA: http://www.saga-gis.org/saga_tool_doc/3.0.0/grid_calculus.html and the grid analysis tools: http://www.saga-gis.org/saga_tool_doc/3.0.0/grid_analysis.html
Conditional evaluation: Performs a conditional if/else evaluation on each of the input cells of an input raster. For example if a value is > x then do this or if it isn't, do something else	**Con tool** **(function can also be performed in the raster calculator)**	**Function can only be performed in the raster calculator** **See http://gis.stackexchange.com/questions/69734/what-is-the-equivalent-of-arcpy-con-in-qgis-and-or-r-raster-package**

Raster only tools	Example Raster Tools	
	ArcGIS	**QGIS**
Set Null tool sets identified cell locations to No Data based on a specified criteria. It returns No Data if a conditional evaluation is true, and returns the value specified by another raster if it is false.	**Set Null** **(function can also be performed in the raster calculator)**	**Function can only be performed in the raster calculator**
Math tools that perform mathematical operations on rasters. For Multicriteria analysis **Logical Math tools** are useful. These evaluate the values of the inputs and determine the output values based on Boolean logic.	The **Math toolset** in ArcGIS contains tools that perform mathematical operations on rasters. **Logical Math toolset** contains tools for performing logical evaluations on rasters which look at whether statements are true or false. The tools are in the following categories: **Boolean, Combinatorial, Relational**, and **Conditional.** ➢ The **Boolean tools** evaluate the inputs only as True or False conditions and return the result of the particular tool as a 1 or 0 (True or False) Boolean value. ➢ The **Combinatorial tools** identify unique combinations of input values based on the logic of the particular tool and return a different value for each unique combination. ➢ The **Relational tools** compare the values of one input relative to another and return the result of the particular tool as 1 or 0 Boolean value. **(functions can also be performed in the raster calculator)**	The **Raster Calculus toolset (SAGA)** in QGIS contains tools that perform mathematical operations on rasters. **No Logical math tools in QGIS. Functions can only be performed in the raster calculator.**

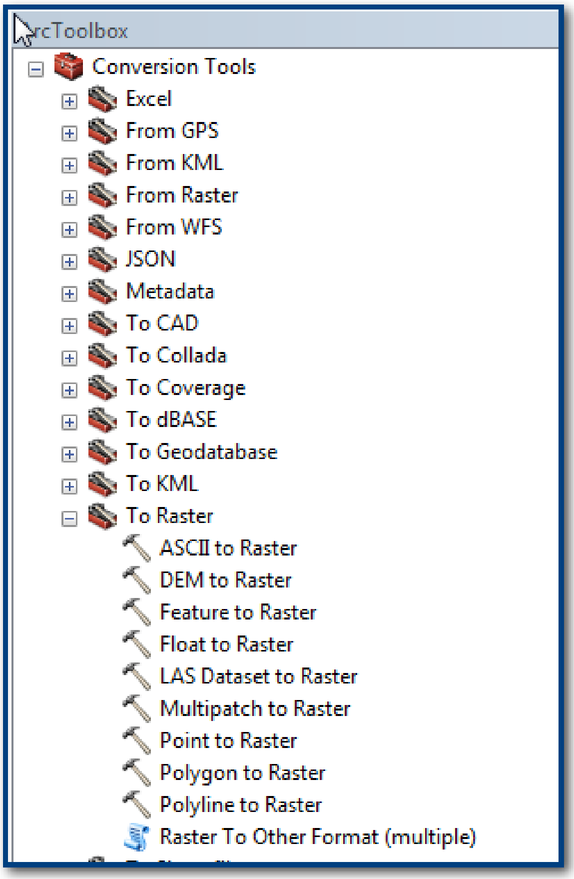
Raster only tools	Example Raster Tools	
	ArcGIS	**QGIS**
Raster Calculator Allows you to perform custom calculations based on values within raster datasets.	**Raster calculator** Builds and executes a single map algebra expression using python syntax in a calculator type interface	Syntax varies depending on which calculator you use:- **Raster calculator (QGIS)** https://docs.qgis.org/2.2/en/docs/user_manual/working_with_raster/raster_calculator.html **r.mapcalculator (GRASS)** https://grass.osgeo.org/grass73/manuals/r.mapcalc.html **Raster calculator (SAGA)** http://www.saga-gis.org/saga_tool_doc/3.0.0/grid_calculus_1.html

A.2. Raster analysis and the Spatial Analyst extension

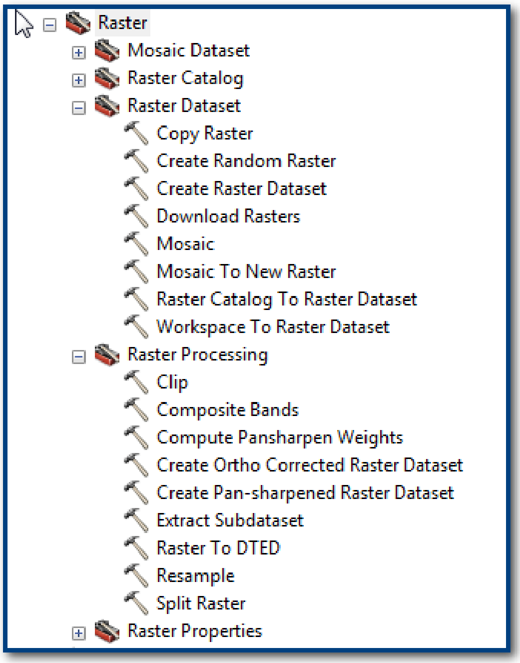
The **Spatial Analyst extension** provides desktop ArcGIS users with access to **raster analysis tools** within ArcToolbox. This extension is a necessity if analysis is to be undertaken.

To convert from vector to raster, find tools in the **Conversion Toolbox**:

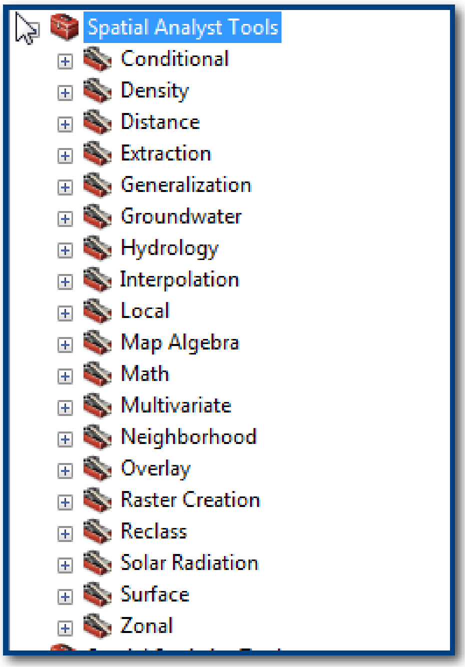
Note: Feature to raster uses cell centre to determine pixel value. ***Polygon to Raster tool*** *or* ***Point to Raster tool*** *offers more options to enable greater control over how input geometries are converted.*



Various raster tools are located under the **Data Management Toolbox** for creating, mosaicking, clipping, resampling, etc., of raster data:



Most of the **Raster analysis tools** are within the **Spatial Analysis toolbox**:



Some of the advantages of Raster analysis include the ability to:

- Integrate raster and vector analysis
- Derive new information from existing data
- Query information across multiple layers
- Undertake complex analyses in a more timely manner

A.2.1. Raster analysis and the environment settings

You may be familiar with setting **Environment Settings** for Geoprocessing in ArcMap. The settings are additional parameters that can be implemented to affect a tool's results.

In raster analysis this is very important, to make sure that cells are not misaligned (causing data to shift) in output layers.

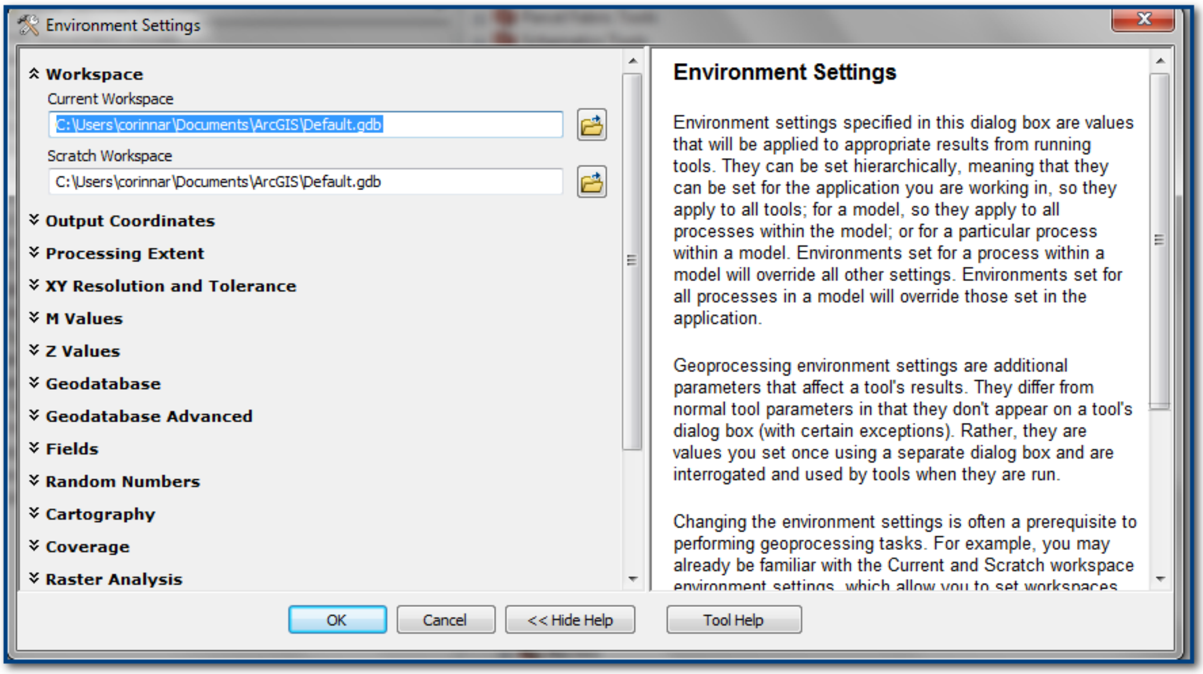
The environment settings can be set at four different levels:

- For the application (e.g. ArcMap) you are working in so they apply to all tools
- For an individual tool so it only applies to this tool and overwrites the application (e.g. ArcMap) environment settings
- For a model so the settings apply to all processes within the model
- Or for a particular process within a model

A.2.2. Workspaces

When performing any geoprocessing tasks in ArcGIS, you usually have to indicate where to save the output dataset. There are two environment settings designed to make specifying input and output datasets much easier: the **Current Workspace** (the default location for geoprocessing inputs and outputs) and **Scratch Workspace** (the default location for temporary outputs, primarily used by model builder intermediate outputs).

a. From the main menu click on **Geoprocessing>>Environments** or when running an individual tool click on the **environments tab at the bottom of the tool interface**. The Environment Settings window appears and you can expand the **Workspace** tab and set the **Current** and **Scratch** workspaces. Here you can change these locations.



A.2.3. Output Coordinates

The Output Coordinates section allows you to pre-set the Output Coordinate System for the output datasets. The also means that processing (calculation of geometric relationships and modification of geometries) occurs in the same coordinate system as the output dataset.

a. In the Environment Settings dialog, expand **Output Coordinates**. For Output Coordinate System, choose **Same as Input** as the default. However, by choosing **As specified Below,** a consistent output coordinate system can be selected for all outputs.

You also have the option to add one or multiple Geographic Transformations so that outputs are projected in the most accurate way.

A.2.4. Processing Extent and Snap Raster

Setting up a **Processing Extent** means that tools will only process features or rasters that **fall within the extent specified in this setting**. This can be helpful for constraining the area of the results and can also quicken geoprocessing time.

a. In the Environment Settings dialog, expand **Processing Extent**. From the Extent drop-down list, choose from **Union of Inputs, Intersection of Inputs, As Specified Below or Same As Display**.
b. Choose a **Snap Raster** to ensure cells are aligned correctly in the output. Cells should be adjusted to ensure they align*. Cell alignment is particularly important to set when input datasets have different cell alignments, cell resolutions or coordinate systems.

**It may be necessary to set both the snap raster and the processing extent to be the same input raster to make sure cells align to that input dataset. Making a copy of an existing raster that you want to use as a snap raster and calling it 'snapgrid', is a good way to remember which file you are using to snap to.*

A.2.5. Cell size

Setting the cell size in raster analysis means that tools use this cell size for the output raster cell size, or resolution, for the operation. By default, the output resolution is determined by the coarsest of the input raster datasets **

a. In the Environment Settings dialog, scroll down and expand **Raster Analysis**. From the **cell size** drop-down list, choose **As Specified Below** and type in the value to use for the analysis.

*** It is important to read the help associated with the tool you are running to ensure that it is transforming the data to a different resolution in an appropriate way. In some cases it may be better to use the **resample** or **aggregate tools** prior to an analysis, as cells are resampled using nearest neighbour resampling in the environment settings and may not always be the most appropriate.*

A raster dataset can always be resampled to have a larger or smaller cell size, but this will not increase the level of detail of the dataset. Important factors for choosing cell size include:

- Spatial resolution of input datasets
- Type of analysis
- Size of output dataset
- Speed of running the analysis

Generally a **smaller cell size** should be used for **large scale data** with **high feature spatial accuracy** (so fine detail is retained), but viewing this over a large area can lead to slow drawing times. A l**arger cell size** usually is selected **for small scale data,** as this means **faster drawing times**.

A.2.6. Mask

Setting up a **Mask** means that tools will only process features or cells that **fall within the analysis mask**. This can be helpful for constraining the area of the results and can also reduce geoprocessing time.

a. In the Environment Settings dialog, scroll down and expand **Raster Analysis**. Choose a **Mask** raster or feature class.

A.3. An Introduction to Some Raster Analysis Techniques

The following guidance provide a brief introduction to some of the common raster analysis functions and finally how the **Raster Calculator tool** can be used to build more complicated expressions.

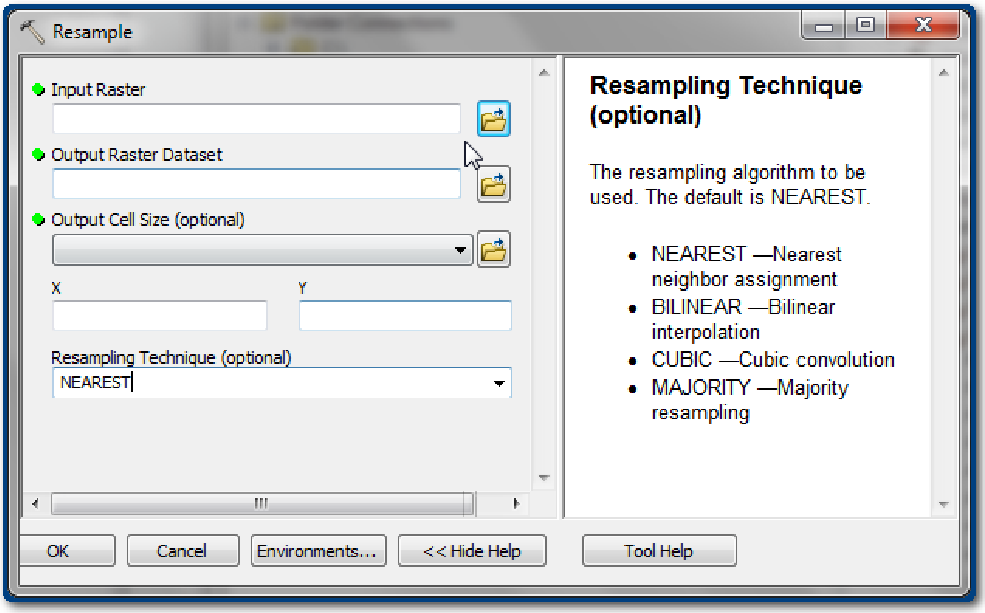
A.3.1. Standardizing Raster Data

There are two main methods: 1) Resampling and; 2) Aggregation

The **Resample tool** allows the user to change the cell size and choose the resampling method.

The default is NEAREST. But users have the choice of four options:

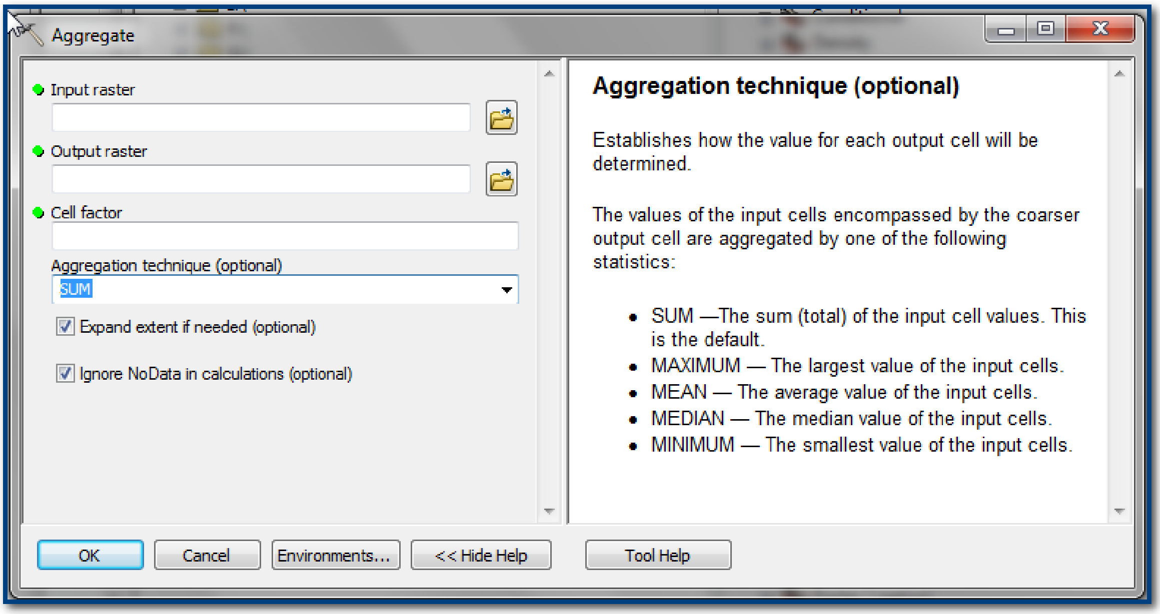
- NEAREST —Nearest neighbour assignment
- BILINEAR —Bilinear interpolation
- CUBIC —Cubic convolution
- MAJORITY —Majority resampling



The **Aggregate tool** allows the user to choose a cell factor to create a reduced resolution version of a raster dataset. Each output cell contains either the Sum, Minimum, Maximum, Mean, or Median of the input cells that are encompassed by the extent of that cell. For example, a user wants to up-scale data from 30m cell size to ~1km cell size. A cell factor of 33 would result in a cell size 33 time larger (i.e. 990m). The cell factor has to be a whole number greater than 1.

To determine how the output cell value is calculated the user has a choice of 5 options (Sum is the default):-

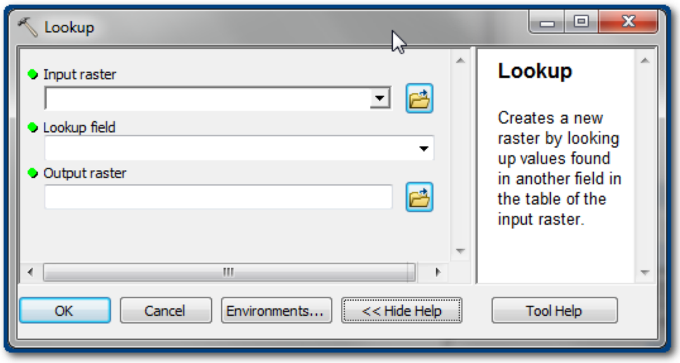
- SUM —The sum (total) of the input cell values.
- MAXIMUM — The largest value of the input cells.
- MEAN — The average value of the input cells.
- MEDIAN — The median value of the input cells.
- MINIMUM — The smallest value of the input cells.



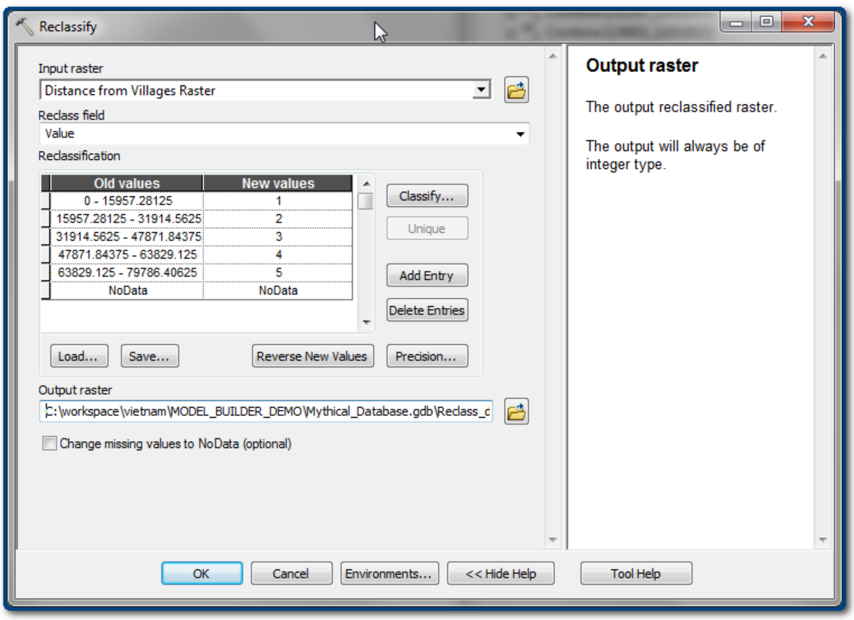
Note that resampling and aggregate provide different methods for determining the output values of the cells. Resampling allows users to change to a smaller and larger cell size, but aggregation is for generalization to a larger cell size only.

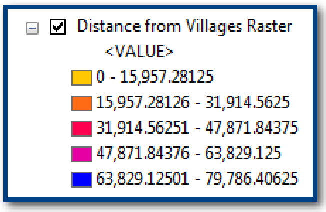
A.3.2. Reclass functions

Many raster processes use the value field when processing data, so if the information you want to use is stored in an additional field it may be necessary to make a new raster from this field so that it becomes a value in the new datasets. The **Lookup tool** can be used to do this. Located in the **Spatial Analyst>>Reclass toolbox**

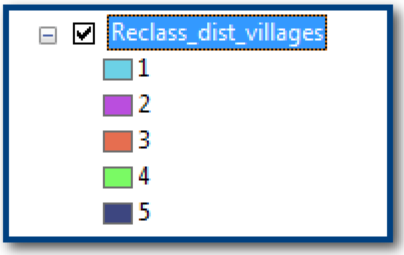


The **Reclassify tool** can be used with or without adding an additional field. Before clicking on the **Reclassify tool**, it is helpful to symbolize the data into the required class breaks (as illustrated below).





Then when the **reclassify tool** is opened it defaults to the class breaks that have been chosen in the symbolization.



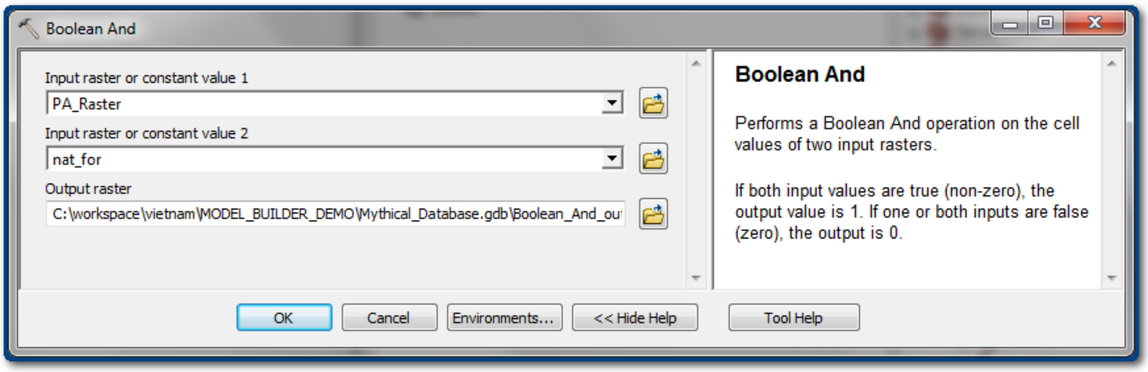
In this example, the new output raster will have values of 1,2,3,4,5.

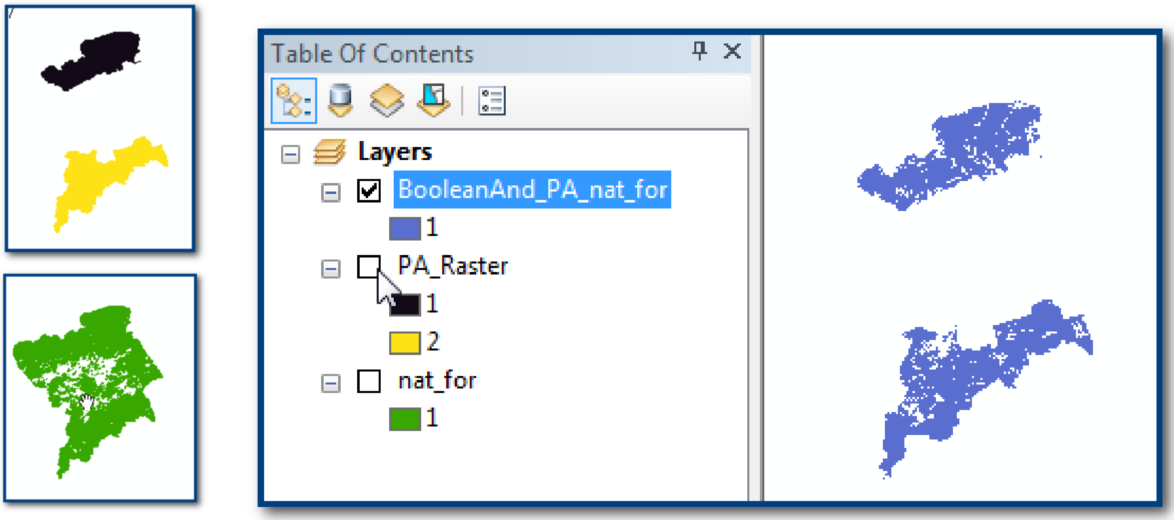
If you only need to convert No Data values to another value, a simpler method to do this is to use the 'Con' expression in **Raster Calculator tool** (or Tool 4 in the Exploring Multiple benefits toolbox) – see section A.3.7 below.

A.3.3. Boolean Intersection

Boolean intersection is the simplest variant of criteria processing and is often referred to as constraint mapping. Prior to the combination, each input criteria is standardized to a certain scale of suitability (or reclassified into classes of 1 and 0), i.e. reducing all the factors to Boolean raster datasets of suitable and unsuitable areas. The separate 1-0 datasets can be combined together using **Boolean tools**; in ArcGIS found **the Spatial Analyst>>Math>>Logical toolset**. These tools are described below.

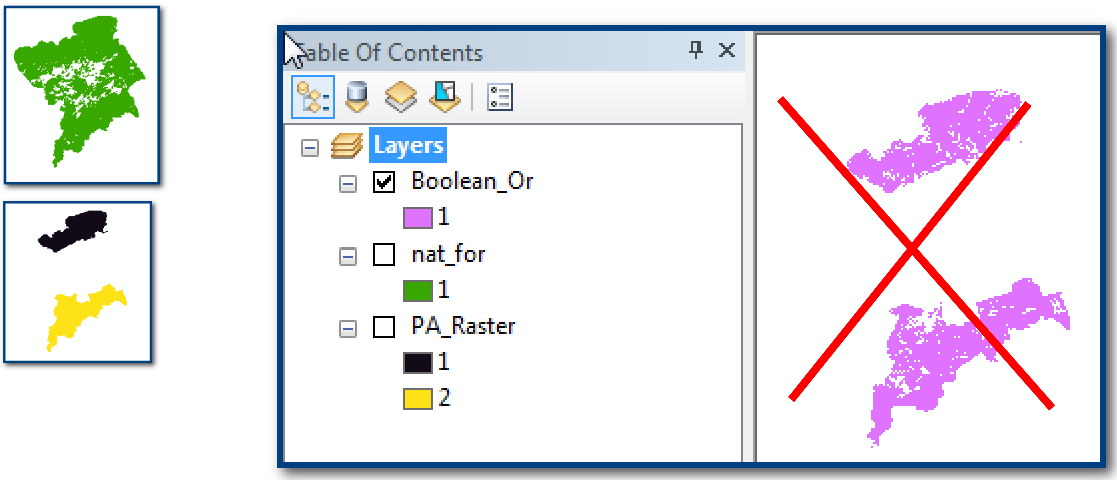
In an example **Boolean And tool** analysis, where data for forests contain the value of 1, and data for Protected Areas contain the values of 1 and 2, and outside of Protected Areas and outside of forest area has **No Data** value, the output raster is 1 where forests and Protected Areas coincide. Remaining areas are classified as **No Data**.



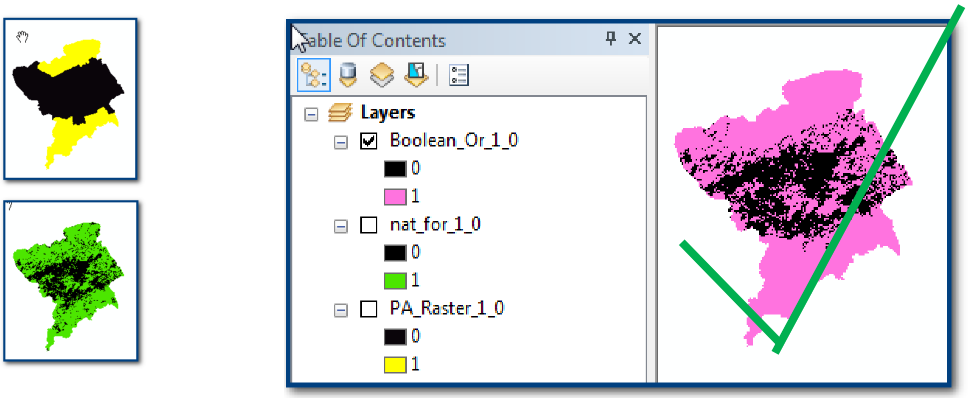


The No Data is not a problem with the **Boolean And tool** analysis, as the analysis is selecting values present in BOTH datasets only. When it comes to **MOST OTHER tools in the logical toolbox** you will see that **No Data values can cause problems** and it is necessary to convert them to 0 before using the data (see section 1.4.1 on how to convert No Data values into data).

The **Boolean Or tool** function selects where data values are greater than 0 in either of the two datasets. However, if No Data values are present in either of the datasets, this will take precedence over any other value and acts as an additional constraint. For example, if a cell has a value of 1 in the forest dataset and a value of No Data in the Protected Area dataset, the output cell would receive a value of No Data. *** So, with **No Data values** present **in the input layers, the output from the Boolean OR tool** will be **WRONG**!***

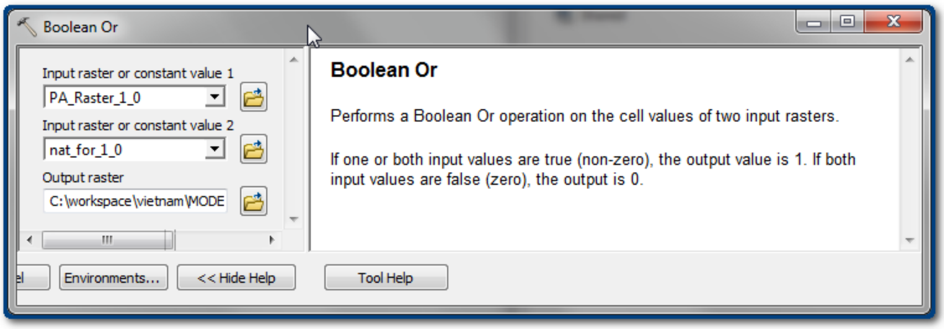


The result should display values of 1 where protected area OR natural forest is present. No Data values therefore need to be changed to 0. The illustration below shows the correct result using inputs for protected areas, where 1 is protected and 0 for not protected, and 1 is natural forest and 0 for not natural forest.



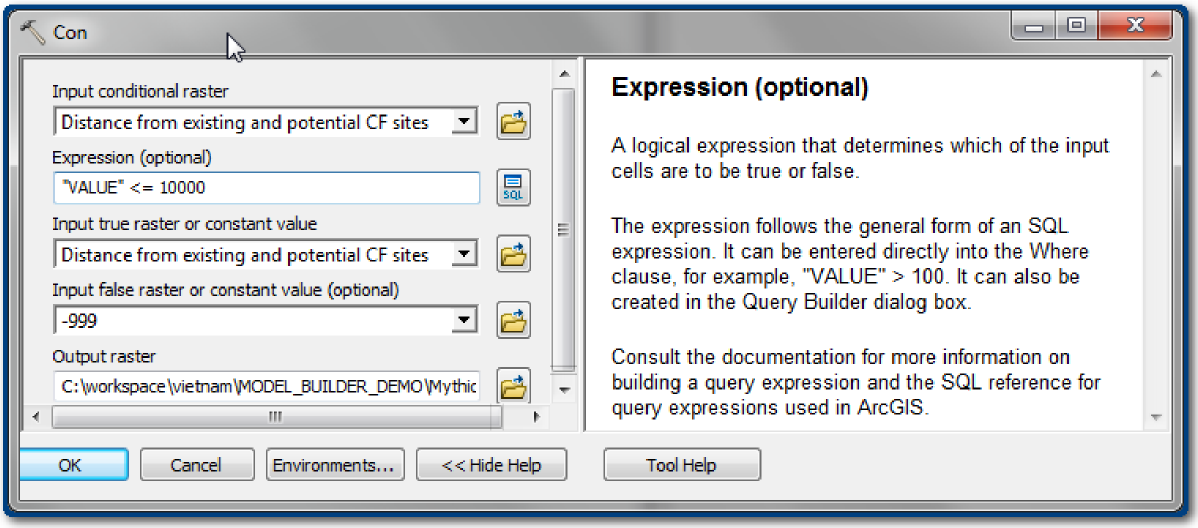
To convert the input datasets to 1-0 rasters where No Data values are set to 0 and all other values are set to 1, you could use the **Con** expression in the **Raster calculator tool**. Alternatively, we have created a tool (Tool 4) in the **Exploring Multiple benefits toolbox** that is easier to use, as it contains the syntax of the **Con** expression (see section A.3.7 for more details on converting No Data values).

Once both the inputs have 1's and 0's, the **Boolean Or tool** can be run.



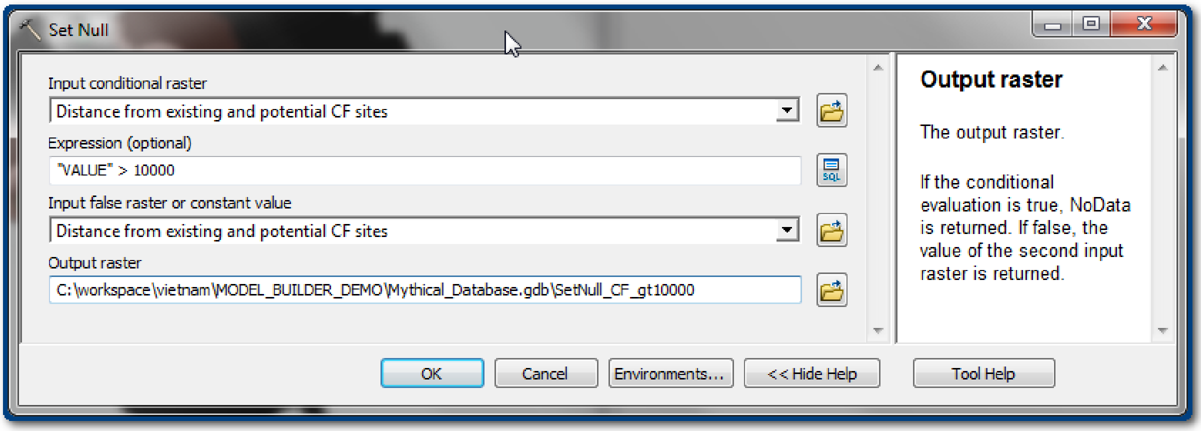
A.3.4. Conditional Tools

Conditional expressions for a single raster dataset can be run using the **Con tool** in the Conditional toolbox. For example, in a dataset showing distance from potential and existing community forestry sites we want to select out areas smaller than or equal to 10000.



The output raster will contain a distance value where the expression is True (i.e. where distance is <= 10000, and -999 where the expression is false).

Alternatively the **Set Null tool** can be used. Using the same example, this time you are using a conditional statement to select cells in which to set the values to be Null (No Data). Therefore, the opposite expression is used, all areas > 10,000.

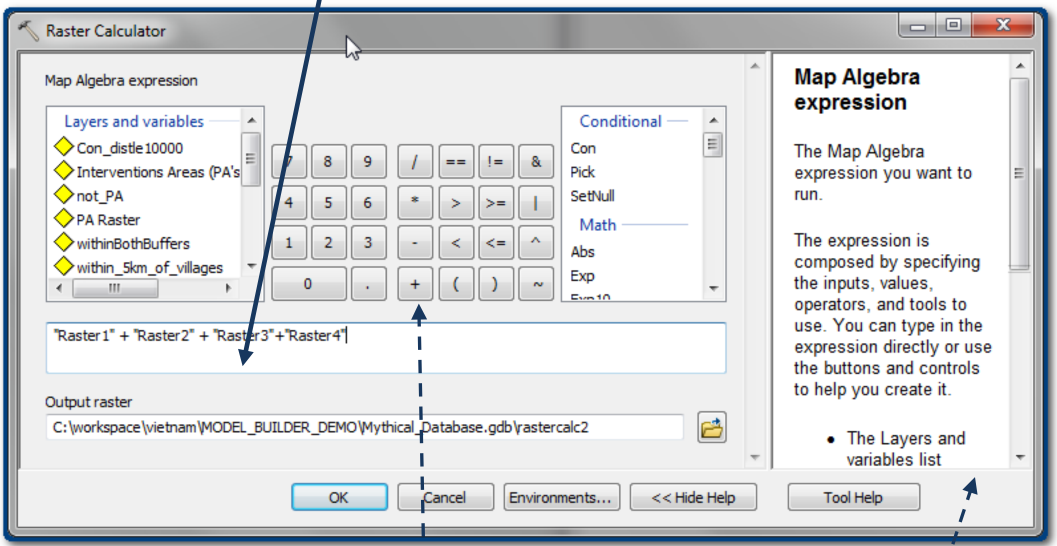


The output raster will be set to No Data where the expression is True (i.e. where distance is > 10000) and will maintain the distance value where the expression is false.
For a more complex conditional statement, involving more than one raster dataset, the **Raster Calculator tool** can be used (Found in **Spatial Analyst Tools>>Map Algebra** toolbox).

A.3.5. Raster calculator

You can also run raster tools using **Map Algebra** expressions typed into the **Raster Calculator tool**. Map Algebra expressions can be used to run a single tool or operators, or to create a more complex string of tools or expressions.

The expression is typed in here. In this example four raster datasets are being summed together.



Raster1		
2	5	0
8	6	1
5	1	1

Raster2		
1	0	1
0	1	0
0	1	1

Raster3		
2	1	1
1	1	1
1	3	1

Raster4		
2	1	5
0	0	3
3	3	3

=

Output Raster		
7	7	7
9	8	5
9	8	6

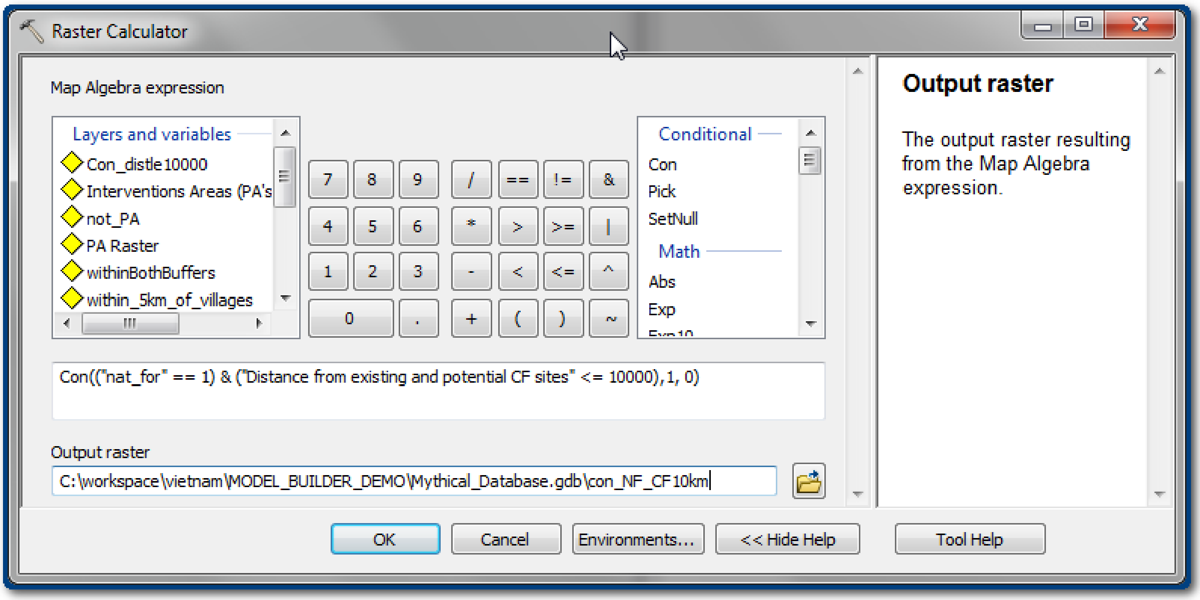
e. The **left hand panel** shows the **list of layers and variable names** that can be used in the expression. These are the raster layers that you have added to your ArcMap session. By double-clicking on a layer it will bring it into the expression box in quotation marks.

f. **The middle panel contains a list of operators. Click once on these to add to the expression box.** There are simple operators to **Add (+), Subtract (-), Divide (/) or multiply (*)** as well as those that are equivalent to the **Boolean logic tools, and operators** to enable the user to select out certain combinations values from a dataset(s):

& Boolean And	== Equal to	>= Greater than or equal to
| Boolean Or	> Greater than	<= Less than or equal to
^ Boolean XOr	< Less than	
~ Boolean Not	!= Not Equal to	

In the right hand panel are a list of **tools** that are comonly used in the **Raster Calculator tool**, however you can also run any of the other tools from the geoprocessing toolbox (see the Map Algebra syntax in the ArcGIS help).

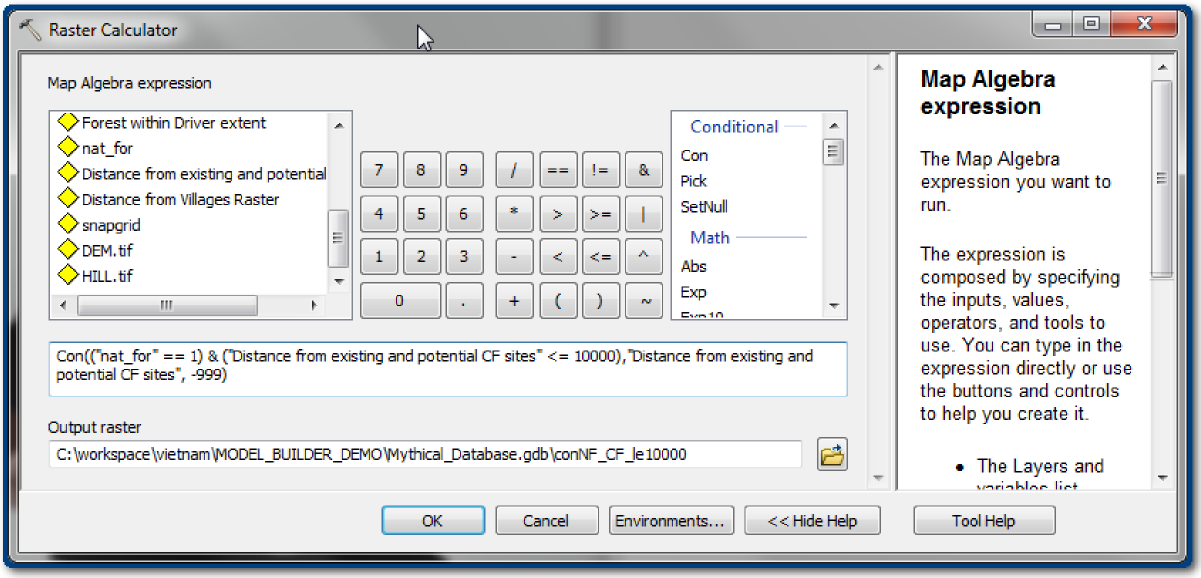
For example, if we want to select out **only Natural Forest Areas** at a distance **<= 10000 m from potential or existing Community forestry sites,** we can use the raster calculator with the following expression:



This means that if **nat_for** has a value equal to **1** and is **less than or equal to 10000m** from potential or existing community forestry sites, set the value in the output dataset to **1**. If this expression is **false**, set the value in the output dataset to **0**.

*Note that the '&' is the same as the **Boolean And tool**.*

Alternatively, you can keep the distance value if the expression is true and make all values -999 if the expression is false:



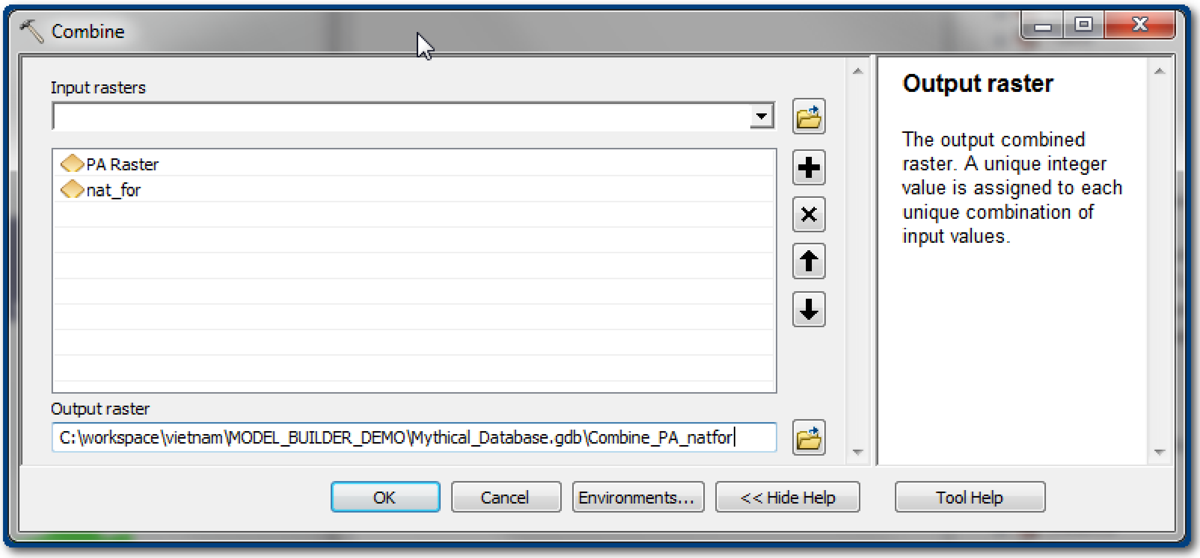
Con(("nat_for" == 1) & ("Distance from existing and potential CF sites" <= 10000), "Distance from existing and potential CF sites", -999)

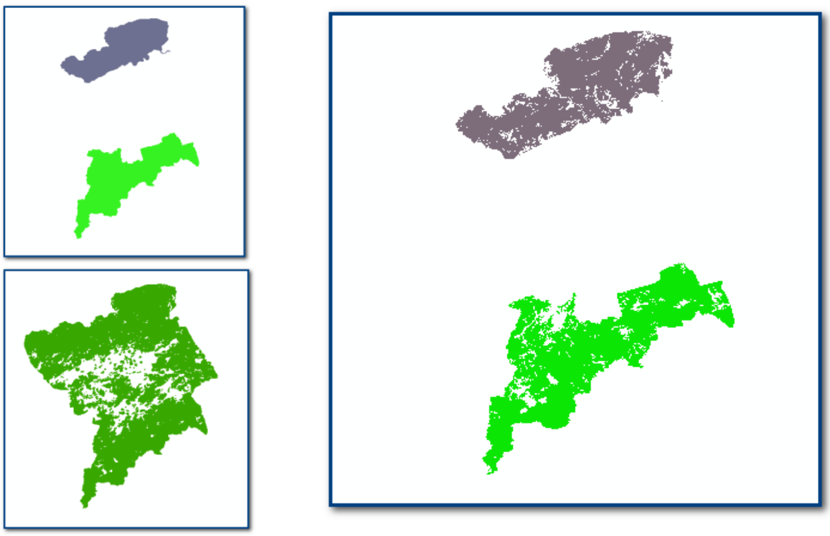
The **ArcGIS Help** is very good at showing the syntax of the map algebra for the various tools.

A.3.6. Combine function

The **Combine function tool** allows users to overlay multiple rasters together so that a new output value is assigned to each unique combination of input values.

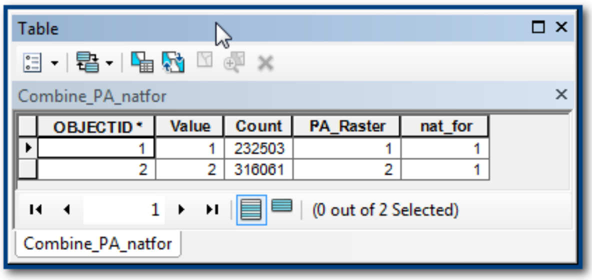
The attribute table contains the new unique output value and the fields from the datasets being combined. In this example there are **548564 cells** which have a **value of 0 in the protected areas data and a value of 1 in the natural forest dataset,** and **551031 cells** that have a value of 1 in both datasets. Note that there are no cells in the dataset that have a value of 0 in the natural forest layer.





IMPORTANT: Note that the resulting output file contains only forest within the Protected Area layer. This is because the area outside the Protected Areas has a value of No Data, and the Area outside the forest has a value of No Data. If a cell has a No Data value in <u>**either**</u> one of the files, it is excluded from the analysis.

This means that if you are using the combine function and want to keep all cells in the analysis, you need to first change No Data values to another Value (see next page).



The output table shows the attributes from the input layers. So in this example, there were 23503 cells that had a value of 1 in the Protected Areas layer and a value of 1 in the forest layer.

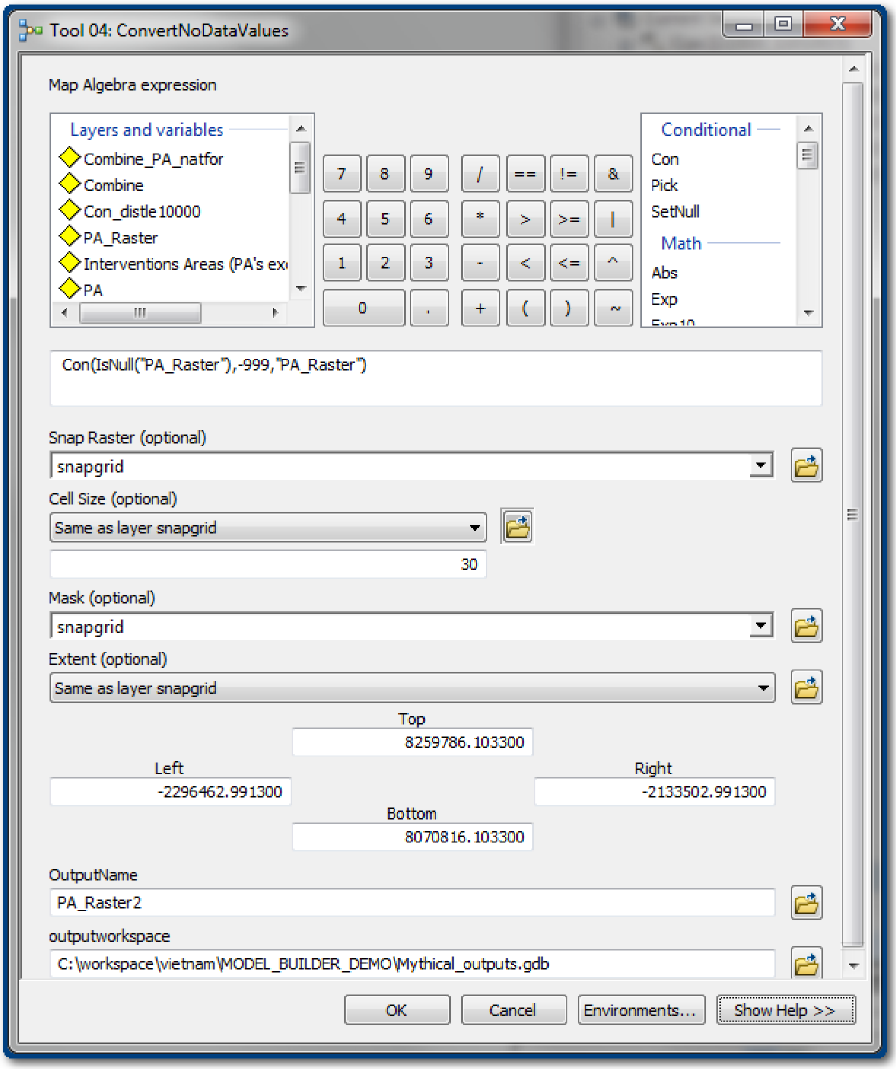
A.3.7. Changing No Data values to a Data value

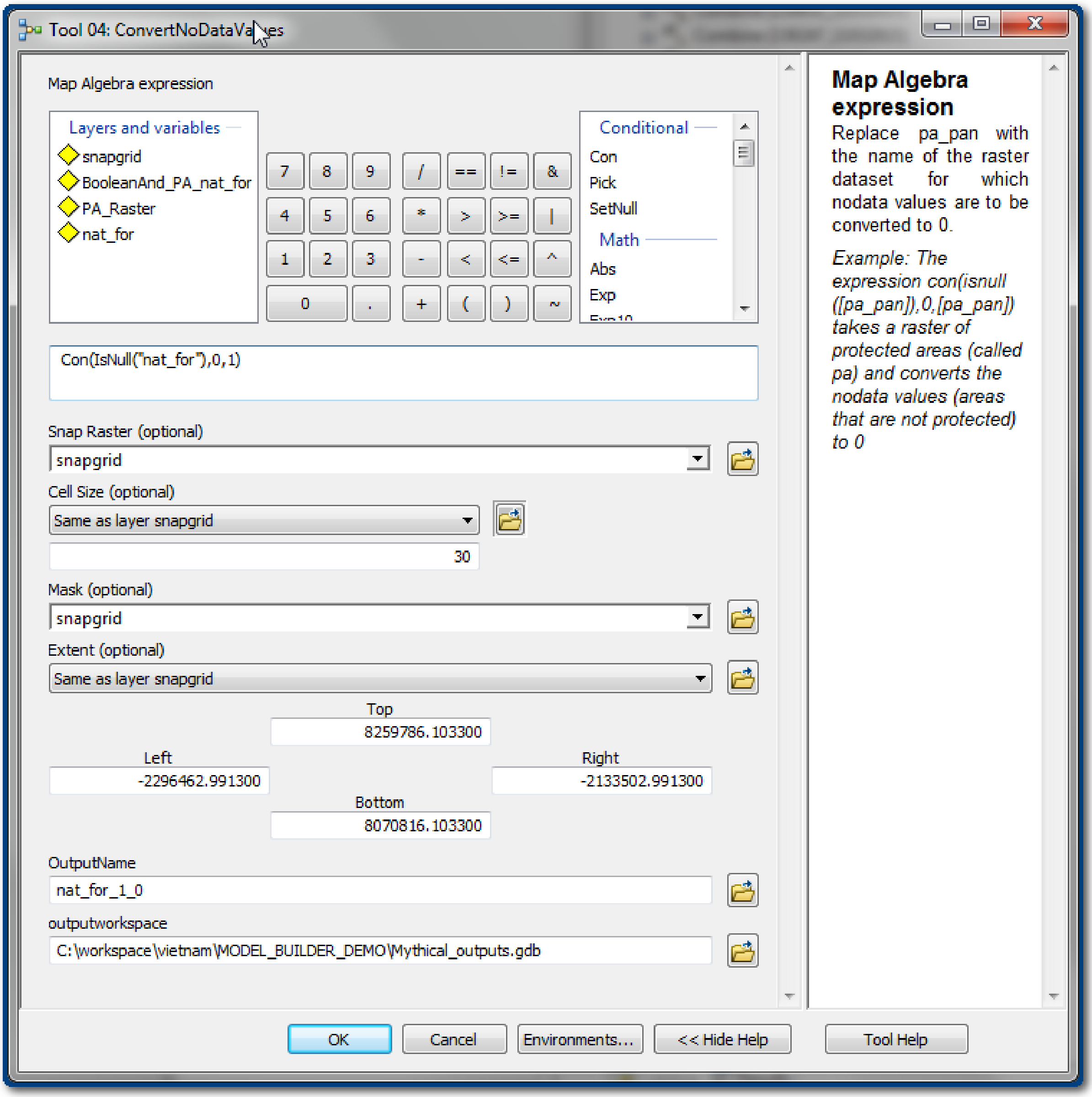
You change No Data values to another value using the **Con** function in the **Raster calculator tool** e.g. **Con(IsNull("PA_Raster"),-999,"PA_Raster")** to convert No Data values to -999 in a protected areas raster, where -999 represents cells that are not protected and all other values remain as they were in the protected areas raster. Alternatively use **Con(IsNull("nat_for"),0,1)** to convert No Data values to 0 in a natural forest layer, where 0 represents not natural forest and 1 represents natural forest.

In order for the full extent of your area of interest to be processed, you will need to make sure you set the **relevant environment settings**. In particular, the **processing extent** to cover the full area of interest and the **snap raster** set to the raster to which you want the cells to align.

Alternatively you can use **Tool 4** in **the Exploring Multiple benefits toolbox,** which asks you to set the relevant settings on the tool interface without having to find and set them in the environment settings tab.

If these environment settings are not set correctly, ArcGIS will use default settings which are not always correct. For example, if you want to use the combine function to combine a country protected areas layer with a country forest layer, the bounding box of the protected areas layer (left) may differ from that of the forest layer (next page).





Helpful hint: If you are undertaking a raster analsyis it is useful to create a 'snapgrid' e.g. from a country or provincial boundary, or an existing raster that covers the full extent of interest. By doing this you can use this 'snapgrid' dataset for environment settings. That way you don't have to remember which dataset you were snapping to and will ensure consiteny thoughout your analysis. To ensure the extent is correct and that there is no mis-alignment of cells, the ***mask****,* ***processing extent*** *and* ***snap raster*** *are the three settings that are usually necessary to set These can all be set the same as the'snapgrid'. The mask is included in this list because ArcGIS has been known to mis-align cells, which masking to the full extent helps to prevent! There is further explanation in the Exploring Multiple Benefits Manaual and Tools 1a and 1b in the Exploring Multiple Benefits toolbox. These can be used to create the snapgrid from an existing vector or raster dataset.*

A.3.8. Zonal statistics

Zonal statistics tools are very useful and are found in the **Spatial Analyst>>Zonal toolbox**. They can be used to summarize an input raster data layer by a zone layer. The **Zone** layer can be either vector or raster and can, for example, be used to summarize the amount of different forest types within individual protected areas.

This guidance has presented only a few tools in detail, to help to get users started on raster analysis and should not be seen as a comprehensive guide. Users can explore the toolboxes in ArcGIS for themselves to find and identify tools that may be relevant to the question they are trying to answer.

Annex 2: Additional ArcGIS Model Builder Resources

- **Tutorials**
- **Resource Center**
- **Online Examples**
- Model Builder **Terms**
- Blogs:
  - **Troubleshooting Tips**
  - **Improving Performance**