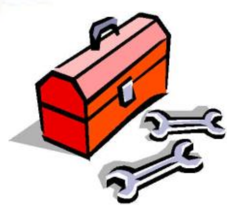
Telecoupling Toolbox v1.7.3b
User Guide



Author: Francesco Tonini
Email: ftonini@anr.msu.edu (ftonini84@gmail.com)

Last Revised: July 25, 2018

CONTENTS

1 INTRODUCTION

In the past decades, the world has undergone dramatic changes and increased interactions between human and natural systems over large distances often led to unexpected outcomes with profound implications for sustainability (Reid et al., 2010). These increased interactions are a direct consequence of globalization and expansion in human population. Trade, transnational land deals, spread of invasive species, and technology transfer occur quicker and are more prevalent than ever before (Liu et al., 2007). Many subsistence needs, e.g. water and food, historically met by local resources are now being met by increased global trade (Kastner et al., 2011; Konar et al., 2011). Understanding distant interactions is a direct response to international calls for transforming the concept of sustainable development into practice (United Nations Secretary-General's High-level Panel on Global Sustainability, 2012). Several disciplines have studied interactions between distant natural or human systems. For example, atmospheric sciences use the concept of teleconnections to describe environmental interactions among climatic systems over long distances (i.e. change in climate at one location affect other places that are hundreds and thousands of miles away). On the other hand, social scientists have studied economic globalization and socioeconomic interactions between distant human systems.

Telecoupling is an interdisciplinary research umbrella concept that enables natural and social scientists to understand and generate information for managing how humans and nature can sustainably coexist worldwide. The telecoupling framework (Fig. 1) treats each place as a coupled human and natural system, in which humans and natural components interact not only locally but also across temporal and spatial scales. It provides an explicit approach to account for and internalize socieconomic and environmental externalities across space. The framework consists of five major interrelated components: coupled human and natural systems; flows of material, information, and energy among systems; agents that facilitate the flows; causes that drive the flows; and effects that result from the flows. Depending on the direction of flows, systems can be classified as three different types. These include sending systems (e.g. exporting countries), receiving systems (e.g. importing countries), and spillover systems (e.g. countries other than the trade partners). Spillover systems are those that affect and are affected by the interactions between sending and receiving systems. For more detailed information see Liu et al. (2015, 2016).

To systematically study telecoupling, it is essential to build a comprehensive set of tools for describing and quantifying multiple reciprocal socioeconomic and environmental interactions between a focal area and other areas. The telecoupling toolbox is (a) spatially-explicit, in that it is developed within a geographic information system (GIS) environment to account for the spatial location of the five major components of the telecoupling framework; (b) multi-scale, in that it can describe socio-ecological systems at a very fine scale as well as at a very coarse global scale; (c) extendible, in that it can be expanded to accommodate as many quantitative/qualitative tools and telecoupling case studies as deemed appropriate; (d) modular, in that it allows the integration of existing tools and software (e.g. InVEST[1]) to assess synergies and tradeoffs associated with policies and other local to global interventions; (e) interactive, in that the user can take full advantage of the interactive functionalities offered by the GIS software environment; and (f) open source, in that its source code and documentation is freely available to users in the public and private sectors.

1.1 Who Should Use the Telecoupling Toolbox?

The telecoupling toolbox is designed for a wide audience of users coming from different research disciplines and both the public and private sectors, interested in applying the telecoupling framework to a specific case study. Case studies of telecoupling can range from agricultural production and trade, payments of subsidies for conservation programs, tourism, spread of invasive species, wildlife migration, and many others. The modular design of the toolbox allows the integration of existing tools and software (e.g. InVEST) to assess synergies and tradeoffs associated with policies and other local to global interventions. The toolbox provides a single, integrated environment to help users map systems, agents, and flows at any spatial scale, while offering descriptive and quantitative tools to determine the causes as well as quantifying how changes in socio-ecological systems are likely to lead to changes in the flows of benefits to people over multiple spatial scales and distances. The telecoupling toolbox can be especially useful for exploring the outcomes of alternative management and climate scenarios or evaluating trade-offs and feedbacks between focal area and any other potentially affected area. For example, changes in agricultural crop production caused by either changes in local/global demand or changes in the natural environment (e.g. climate, land-use change) will likely have repercussions on the market price values and have socio-economic feedbacks on revenues and incomes of all partners involved in the trade chains. At the same time, an increased or decreased crop production may have repercussions on carbon sequestration and soil water retention. Users should utilize the telecoupling toolbox to better describe the entire system and entities involved in a given flow of material/energy while accounting for multiple effects and feedback on both the socioeconomic and natural systems at any affected location. Some of the script tools found inside the toolbox can help organizations and stakeholders decide where to best allocate economic resources to ensure their investments are sustainable and secure.

The telecoupling toolbox can help visualize interacting systems, agents involved, and flow routes on a map. At the same time, the integration with multiple external software tools, e.g. InVEST, helps answer questions like:

- Where do goods, information, and ecosystem services originate and where are they consumed?

[1] http://www.naturalcapitalproject.org/invest/

- How do conservation subsidy programs affect the local population, habitat quality, water quality and recreation?
- How will climate change and population expansion impact the natural environment and biodiversity?
- What are the main factors causing the flow of goods, information, or ecosystem services between focal and receiving areas?
- How will an investment to increase local eco-tourism affect the natural environment and benefit the local population?
- What fishery management policies will be the most sustainable?
- How will the natural and scenic views of marine and coastal seascapes be impacted by future construction plans?

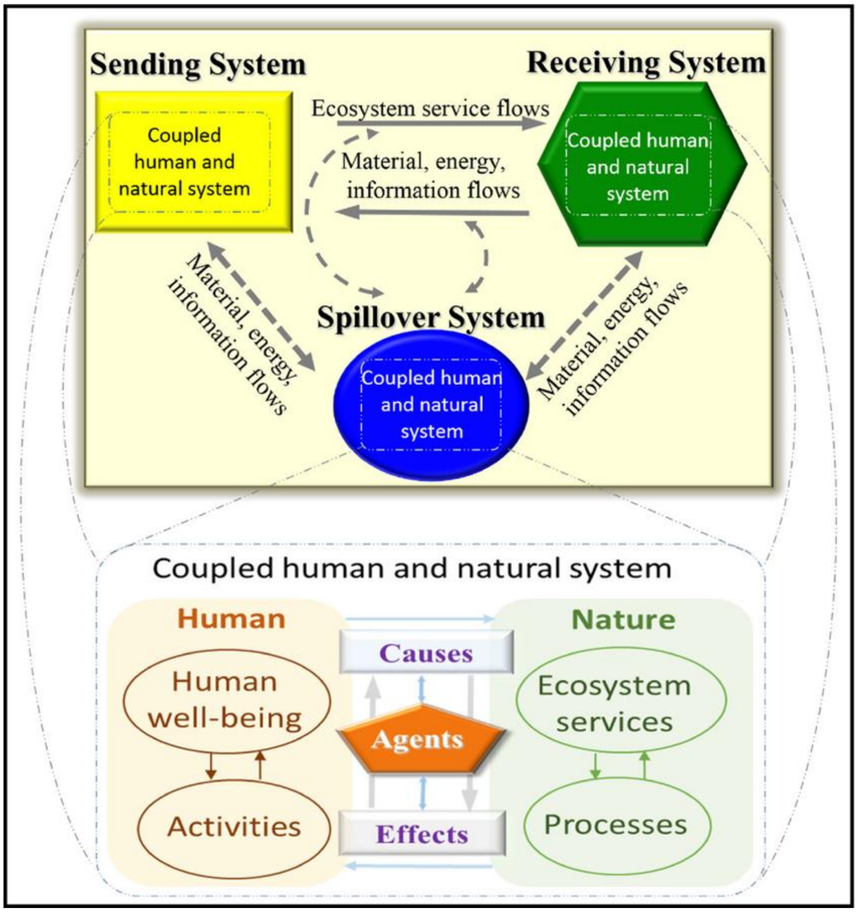


Figure 1. Framing ES in a telecoupling context (adapted from Liu et al. 2013). A "sending" system provides ES to a "receiving" system, and may influence another system (the "spillover" system) in the process. Feedbacks occur among different systems as a result of material energy, and information flows. Each system represents a coupled human and natural system, with two major components: humans and

nature (inset). The human component consists of human well-being (e.g. health), activities (e.g. consumption), and other elements (e.g. population size, not shown). The nature component consists of ES (e.g. clean water), processes (e.g. predation), and other elements (e.g. biodiversity, not shown). The causes, agents, and effects of telecoupling processes, such as flows of ES between systems, may occur within each system. For example, causes may include human activities and ecological processes that affect flows of ES with the facilitation of agents, such as traders and government officials, whereas effects may include impacts of a telecoupling process on ES and human well-being. Arrows indicate the direction of these influences. **NOTE: this figure was taken from Liu et al. (2016)**.

1.2 Major Releases

1.2.1 **Version 1.7** (beta)

This release adds a new tool within the Environmental Analysis toolset, called *Seasonal Water Yield*, customized from InVEST (3.3.3). This tool can be used to estimate the contribution of a pixel in a watershed to baseflow, quickflow, and local recharge.

1.2.2 **Version 1.6** (beta)

This release adds a new tool within the Environmental Analysis toolset, called *Fisheries Harvest*, customized from InVEST (3.3.3). This tool can be used to quantify catch volume and economic value.

1.2.3 **Version 1.5** (beta)

This release adds a new tool within the Systems toolset, called Network Analysis Grouping, written in R and connected to ArcGIS using the R-Bridge from ESRI. This tool uses network information (i.e. nodes, links, and directions) of units (e.g. countries, areas, administrative units) within a telecoupling system and aggregates them into groups (clusters) based on their network connectivity.

1.2.4 **Version 1.4** (beta)

This release adds a new tool within the Environmental Analysis toolset, called *NDR (Nutrient Delivery Ratio)*, customized from InVEST (3.3.3). This tools can be used to map nutrient sources from watersheds and their transport to the stream.

1.2.5 **Version 1.3** (beta)

This release adds a new tool within the Socieconomic Analysis toolset, called *Nutrition Metrics*, a custom tool that estimates the population within an Area of Interest (AOI) by age groups and then calculates the Lower Limit of Energy Requirement (LLER; in kilocalories / day) for age groups within the AOI. NOTE: This tool will only run with Areas of Interest in Africa, Asia, or South America / Central America / Caribbean.

1.2.6 **Version 1.2** (beta)

This release adds a new tool within the Socieconomic Analysis toolset, called *Population Count and Density*, a modified version of the Population Density Metrics tool from the Analytical Tools Interface for Landscape Assessments (ATtILA) by the United States Environmental Protection Agency (EPA).

1.2.7 **Version 1.1** (beta)

This release adds two new tools, the *Habitat Risk Assessment Preprocessor* and the *Habitat Risk Assessment*, both customized from InVEST (3.3.3). Sample data is now shipped separately from the main code repository to reduce its size. The main toolbox file (.tbx) has been updated to reflect the change in version and development stage as well as naming convention (Telecoupling *Toolbox* instead of *Tools*). The old ‘Effect’ toolset has now been partitioned into two new toolsets: Environmental Analysis and Socioeconomic Analysis.

1.2.8 **Version 1.0** (beta)

This version moves the development stage from *alpha* to *beta* after several months of testing. Some tools might still return errors or have bugs in it, but testing was successful on several different machines and settings. This version updates the InVEST tools to version **3.3.3** and its dependencies such as the PyGeoprocessing library (*updated to version 0.3.2*). All .bat files inside the PyLibs folder have been updated to reflect this change. The main toolbox file (.tbx) has been updated to reflect the change in version and development stage.

1.2.9 **Version 1.2** (alpha)

This release eliminates the redundant presence of the “Draw Radial Flows and Nodes” tool (Flows toolset) while incorporating its old functionality into the existing “Draw Radial Flows” tool. The latter can now optionally draw and additional layer on top of the default flow lines, showing nodes at each flow destination, which can then be symbolized according to an attribute (quantity) of interest. The toolbox is still in *alpha* development stage and is being tested for bugs and errors that need to be addressed before releasing a more stable *beta* version.

1.2.10 **Version 1.1** (alpha)

This release includes an upgrade in the CO2 emission script tool accounting for number of wildlife units transferred and transportation capacity of the medium used. The tool now allows testing of future CO2 emission scenarios for wildlife transfer compared to current conditions. The toolbox is still in *alpha* development stage and is being tested for bugs and errors that need to be addressed before releasing a more stable *beta* version.

1.2.11 **Version 1.0** (alpha)

This is the first version officially released for the Telecoupling Toolbox for ArcGIS. The toolbox is still in *alpha* development stage and is being tested for bugs and errors that need to be addressed before releasing a more stable *beta* version.

1.3 Minor Releases

1.3.1 **Version 1.7.3** (beta)

This release adds network analysis metrics (within an output .csv file) to the *Network Analysis Grouping* tool.

1.3.2 **Version 1.7.2** (beta)

This release removed the dependence on mosaic data sets from the *Nutrition Metrics* tool. This tool now relies on rasters grouped in a common folder rather than construction of a mosaic data set.

1.3.3 **Version 1.7.1** (beta)

This release re-organizes the *Crop Production* and the *Fisheries Harvest* InVEST *3.3.3* tools into the Socioeconomic Analysis toolset.

1.3.4 **Version 1.2.1** (alpha)

Fixed bug in the Habitat Quality tool (Effects toolset) potentially producing negative values for habitat quality and degradation output rasters.

1.3.5 **Version 1.0.1** (alpha)

Fixed bug in the script tools linked to InVEST (3.3.1) that caused any output vector and raster files to incorrectly align with the input layers. The issue seems to be caused by the way ArcGIS interprets the spatial reference (projection string) of output files saved by the InVEST software. A workaround has been implemented, automatically re-defining the projection of the output layers before showing in the table of content in ArcGIS.

1.3.6 **Version 1.1.2** (alpha)

Added output layer name fields in the Cost-Benefit Analysis (Wildlife Transfer) and Draw Radial Flows tools.

1.3.7 **Version 1.0.1** (alpha)

Fixed bug in the script tools linked to InVEST (3.3.1) that caused any output vector and raster files to incorrectly align with the input layers. The issue seems to be caused by the way ArcGIS interprets the spatial reference (projection string) of output files saved by the InVEST software. A workaround has been implemented, automatically re-defining the projection of the output layers before showing in the table of content in ArcGIS.

1.4 The User Guide

This guide will help you understand the basics of the Telecoupling Toolbox. The next chapter shows you the installation process and provides general information about the tools and their interface.

2 GETTING STARTED

2.1 Prerequisites

In order to use the telecoupling toolbox, make sure you have the following required software installed:

- R (3.2.0 or later)
- ArcGIS (10.3.1 or later)
- Python 2.7.x **(ArcGIS comes with Python, so no need to install a standalone version of Python!)**

NOTE: versions of ArcGIS prior to 10.3.1 may still work with some of our tools but have NOT been tested. ArcGIS Pro has also NOT been tested yet!

Python 2.7.x ships together with ArcGIS and is automatically installed with it. If possible, avoid installing multiple versions of Python on your system as it may create conflicts and errors. If you are using Anaconda, you may need additional settings to make sure you can run the tools smoothly. For more info, check https://pymorton.wordpress.com/2014/06/17/using-arcpy-with-anaconda/.

2.2 Install Python libraries for 3rd party external software

The Telecoupling Toolbox relies on a number of python libraries that are required to run tools that use external software (e.g. InVEST). If you skip this step, tools using any external software will NOT work.

1. Although your computer may already have a C++ compiler installed, follow this step and **Download** and **install** the Microsoft Visual C++ Compiler for Python 2.7[2]
2. Follow the instructions below depending on the version of ArcGIS installed on your system:

   - **ArcGIS 10.3.1 (standard 32-bit version)**:
     - Download **get-pip.py**[3]

[2] https://www.microsoft.com/en-us/download/details.aspx?id=44266
[3] https://bootstrap.pypa.io/get-pip.py

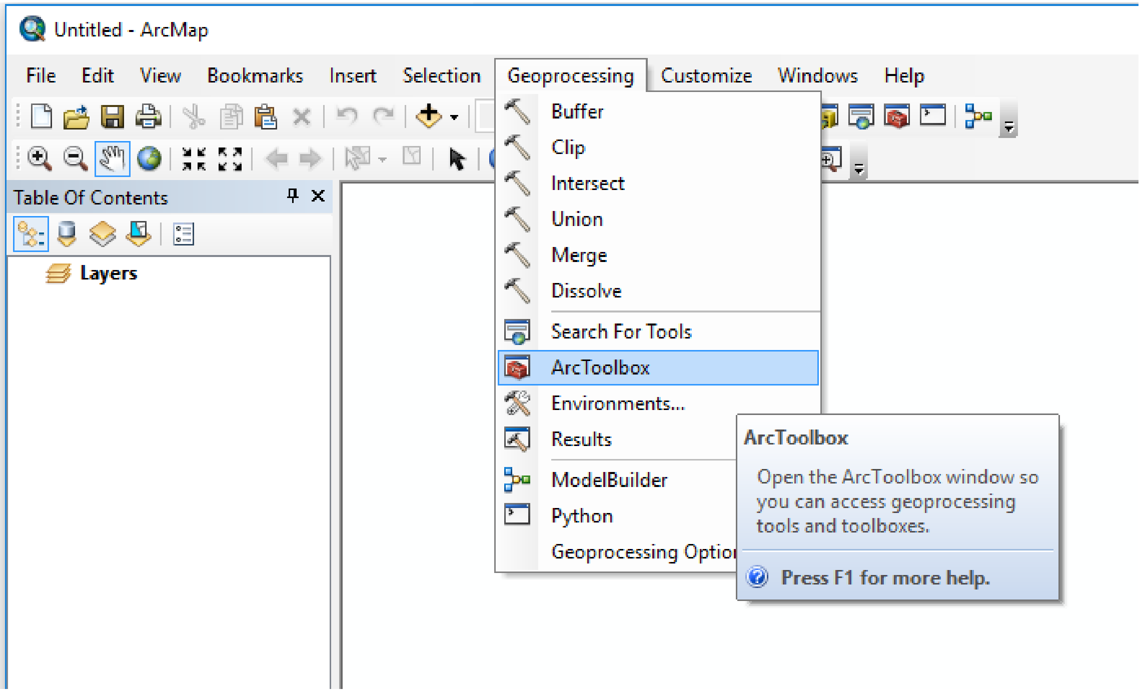
- Open the CMD prompt on Windows and type: `C:\Python27\ArcGIS10.3\python.exe` followed by the full path to `get-pip.py` (*for example if you downloaded and saved the file on your D:\ drive, the full path would be* `D:\get-pip.py`)
    - Hit *Enter* to run the command above
    - Open the folder `PyLibs` found inside the (unzipped) telecoupling project folder
    - Double-click on the `ArcGIS103_Py32_libs.bat` file
- **ArcGIS 10.4.x (standard 32-bit version)**:
    - Open the folder `PyLibs` found inside the (unzipped) telecoupling project folder
    - Double-click on the `ArcGIS104_Py32_libs.bat` file
- **ArcGIS 10.5.x (standard 32-bit version)**:
    - Open the folder `PyLibs` found inside the (unzipped) telecoupling project folder
    - Double-click on the `ArcGIS105_Py32_libs.bat` file

2.3 Install the R-ArcGIS Bridge

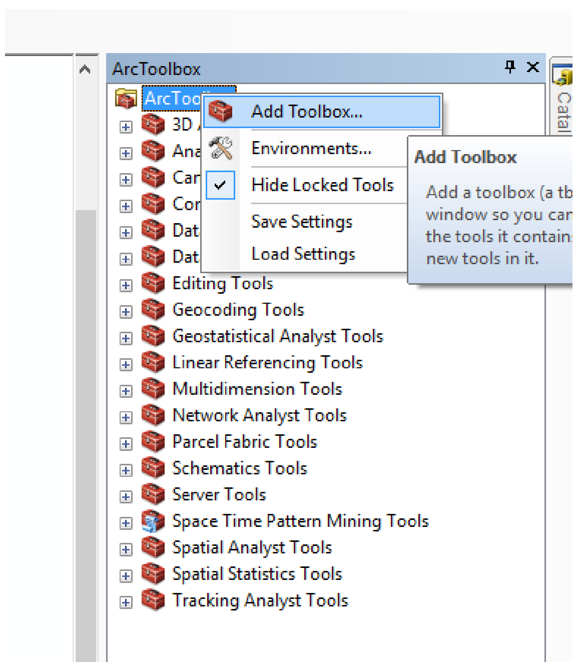
In order to allow interaction between ArcGIS and the R software, you will need to follow the next few steps:

1. Open ArcMap **(NOTE: make sure you have admin rights on your computer or the next steps will not work!)**

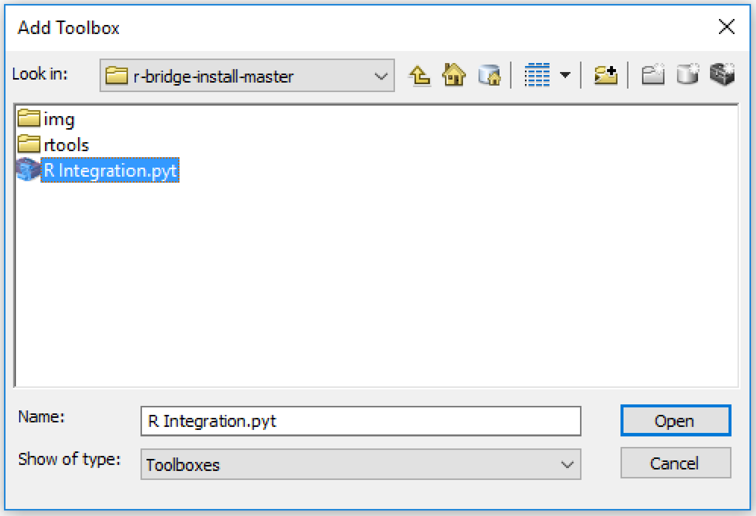
2. Find and open the Geoprocessing ArcToolbox window (Menu > Geoprocessing > ArcToolbox)



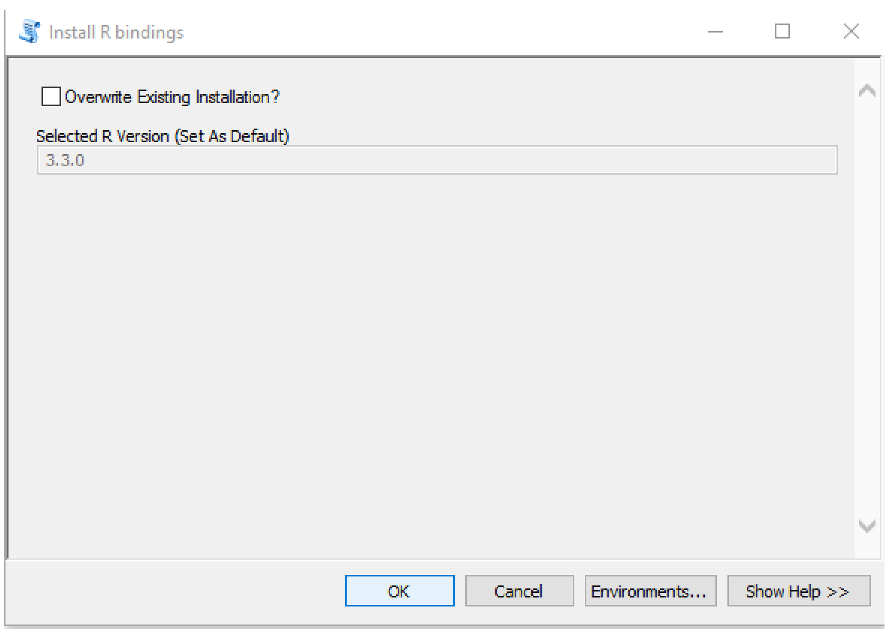
3. Right-click the ArcToolbox folder and select "Add Toolbox..."



4. Browse to the `r-bridge-install-master` folder found inside the unzipped telecoupling project folder and select the `R Integration.pyt` toolbox



5. After the toolbox has been added to the ArcToolbox list, click on it to open it and double-click on the `Install R Bindings` tool to open its interface. Click on OK to run it.



If you need more details and information, ESRI has developed a nice Github webpage[4] with lots of useful documentation on how to install a set of libraries to make sure R and ArcGIS can talk to each other.

2.4 Add the Telecoupling Toolbox to ArcGIS

You are almost done! Now that you installed all Python 3rd party libraries and the R-ArcGIS Bridge, you are ready to use and test the **Telecoupling Toolbox for ArcGIS**.

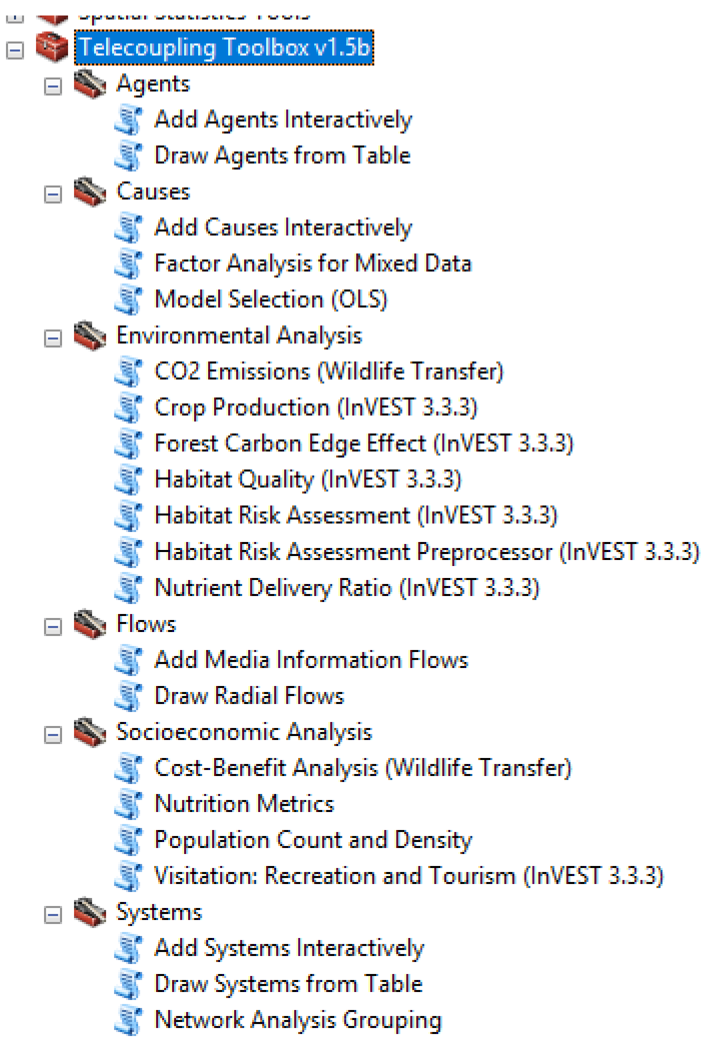
Follow these steps to add the Toolbox to your ArcMap document:

1. Open ArcMap
2. Right-click on the ArcToolbox folder and select "Add Toolbox"
3. Browse to the unzipped folder of the telecoupling project and select `Telecoupling Toolbox v1.2b.tbx`

Inside the Telecoupling Toolbox you should see 6 toolsets (**agents, causes, environmental analysis, socioeconomic analysis, systems**) and a number of python tool scripts inside each one of them.

Each tool will have a help window associated with it explaining what each parameter is and a general description of the tool. To open the help window, click on the 'show help' button found at the bottom of each tool script after opening it (double-click on the tool script to open the user interface).

[4] https://github.com/R-ArcGIS/r-bridge-install



That's it! The Telecoupling Toolbox is now added to the ArcToolbox list and you can start using it with the set of sample data (`SampleData_ArcGIS_v1.7.3b.zip`) that can be downloaded at: https://s3.amazonaws.com/telecoupling-toolbox-sample-data/SampleData_ArcGIS_v1.7.3b.zip After unzipping the sample data folder, you will see a mix of GIS (vector, raster) data and tables (.csv) needed as input parameters by the script tools.

3 TOOLBOX STRUCTURE

3.1 Overview

In ArcGIS, tools and script tools are grouped into toolsets, which are then collected into toolboxes. The telecoupling toolbox was developed as a custom geoprocessing ArcGIS toolbox, made of five nested toolsets and their Python/R script tools. Each of the five toolsets corresponds to one component of the telecoupling framework (see Section 1).

3.2 Systems Toolset

The Systems toolset contains script tools that are meant to map and visualize the geographical location of all areas that are interconnected within the telecoupling of interest. Systems are divided in sending, receiving, and spillover. Sending systems (e.g. exporting countries) are those locations where goods, information, or ecosystem services originate from. Receiving systems (e.g. importing countries) are locations that consume goods, information, or ecosystem services. Spillover systems (e.g. countries other than the trade partners) are locations that directly or indirectly affect or are affected by the interactions between sending and receiving systems.

3.2.1 Add Systems Interactively

Interactively add telecoupling system components (i.e. "sending", "receiving", "spillover") with an explicit spatial location to a map. The tool reads the system spatial coordinates from the map and creates a permanent point feature class with a pre-assigned symbology. Upon successfully completing execution, the tool will automatically export the newly created features and their attributes as a .csv table. Use this tool if you do not have a pre-existing table on file with name, description (optional), and spatial coordinates of all telecoupling system components as part of your case study. Upon successfully completing execution, the tool will automatically export the newly created features and their attributes as a .csv table.

Parameter	Explanation	Data Type
Input_Features	Click on the map to add the spatial location of the system components. ***NOTE: the mouse cursor and its pre-assigned symbology should appear as soon as you move it over the map. The system component will be added to the map as a point feature, so make sure to zoom in to the correct spatial location before clicking on it. It is***	Feature Set

	recommended that you type a name and a description (optional) in the Input_Attributes variable after each time you add a system component so that there is a correct one-to-one correspondence between the spatial location and its attributes.	
Input_Attributes	Type in a name and a description (optional) for each system component added to the map in the previous step. Click on the "+" sign to add a new row or "x" to delete. ***NOTE: The name field needs to be filled out, while the description field is optional. The number of rows needs to equal the number of agents (points) added to the map in the previous step.***	Record Set
Add_XY_Coordinates (Optional)	When checked, XY coordinate (Web Mercator by default) fields are added to the system components attribute table. When unchecked (this is the default), no coordinates are added.	Boolean

3.2.2 Draw Systems from Table

Draw telecoupling system components (i.e. "sending", "receiving", "spillover") on the map by uploading a table stored in a local file on disk. The tool reads XY coordinates from the table and creates a permanent point feature class with a pre-assigned symbology. Upon successfully completing execution, the tool will automatically export the newly created features and their attributes as a .csv table. Use this tool if you already have a pre-existing table on file with name, description (optional), and spatial coordinates of all telecoupling system components as part of your case study. Upon successfully completing execution, the tool will automatically export the newly created features and their attributes as a .csv table.

Parameter	**Explanation**	**Data Type**
Input_Table	The table containing the name, description (optional), and XY coordinates of each system component.	Table View
X_Field	A numerical field in the input table containing the x coordinates (or longitudes) of each system component.	Field

	NOTE: if the XY coordinate fields are not in the Web Mercator coordinate system, the tool will internally reproject to make sure the system components overlap correctly on the basemap. Although the latter step is not necessary, as ArcGIS Desktop can reproject features on-the-fly, the same will not apply when using the tool as a geoprocessing service consumed within a web application.	
Y_Field	A numerical field in the input table containing the y coordinates (or latitudes) of each system component. ***NOTE: if the XY coordinate fields are not in the Web Mercator coordinate system, the tool will internally reproject to make sure the system components overlap correctly on the basemap. Although the latter step is not necessary, as ArcGIS Desktop can reproject features on-the-fly, the same will not apply when using the tool as a geoprocessing service consumed within a web application.***	Field

3.2.3 Network Analysis Grouping

This tool uses network information (i.e. nodes, links, and directions) of units (e.g. countries, areas, administrative units) within a telecoupling system and aggregates them into groups (clusters) based on their network connectivity. Upon successfully completing execution, the tool will automatically export a new shapefile of the telecoupling system along with their grouping outcomes (cluster number) added in attribute table for each of the clustered unit. A network metrics table (in CSV format) indicating some characteristics of the network will also be generated. Optionally, the tool can generate an output plot (in PDF format) showing the clustering outcome of the network analysis. Examples of network data are migration of population (e.g. immigration, tourists, business travelers, students and etc.) or import/export of goods.

The ESRI R Bridge must be installed for this tool to run. Use this tool If you want to separate countries within a system with network analysis algorithm and add an attribute as group information to the system shapefile.

Parameter	Explanation	Data Type
Nodes_Table	Input .csv file containing a list of all the telecoupling systems (e.g. countries, administrative units) that are to be considered as "nodes" in the network grouping analysis. This table must include a unique identifier attribute for each record as well as attributes specifying both the outgoing and incoming flows. Flow can be represented by number of people traveling, export/import of goods, energy, etc. **NOTE: in some case it might be best to use the natural logarithm of total amount of outgoing/incoming flow, especially when flow quantity is very large. Having large numbers will skew the output results so it is recommended to consider natural log as an option.**	File
Nodes_Join_Attribute	Attribute field from the input nodes table that will be used to join with telecoupling systems input layer.	Field
Links_Table	Input .csv file containing a list of all the connections between sending and receiving telecoupling systems (e.g. countries, administrative units) that are to be considered as "links" in the network grouping analysis. This table must include a unique identifier code for both the sending and receiving systems, as well as an attribute specifying the outgoing flow. Flow can be represented by number of people traveling, export/import of goods, energy, etc. **NOTE 1: make sure the codes used for the sending and receiving systems match exactly those used in the nodes input table.** **NOTE 2: make sure the outgoing flow attribute contains the same values used in the nodes input table. Therefore, if you used natural logarithm in the nodes input table, make sure to use the same here.**	File
Telecoupling_Systems_Layer	Input feature layer (e.g. shapefile) representing all the telecoupling sytems (e.g. countries, administrative units) that are to be used in the network grouping analysis.	Shapefile

Layer_Join_Attribute	Input feature layer (e.g. shapefile) representing all the telecoupling sytems (e.g. countries, administrative units) that are to be used in the network grouping analysis.	Field
Clustering_Algorithm	Input algorithm to be used for the grouping analysis. *walktrap*: detecting community strucure via short random walks. The idea is that short random walks tend to stay in the same community because there are only a few edges that lead outside a given community. Walktrap runs short random walks and uses the results of these random walks to merge separate communities in a bottom-up manner. *spin_glass*: this function tries to find communities in graphs via a spin-glass model and simulated annealing. In this model, each node can be in one of c spin states, and the edges specify which pairs of nodes would prefer to stay in the same spin state and which ones prefer to have different spin states. Communities can be defined according to the spin states of nodes after running the model. **NOTE: Refer to R igraph package help page for more information about the algorithms. http://igraph.org/r/doc/communities.html**	String
Weight_Within_Clusters (Optional)	Input number to ajust the distance between nodes within a cluster.	Double
Weight_Between_Clusters (Optional)	Input number to ajust the distance between nodes between different clusters.	Double
Color_Set (Optional)	Select color sets for different clusters. ***Demos of color sets:***  ***NOTE: plotting options. Default value is Set3.***	String
Node_Size (Optional)	Input number to adjust size of nodes.	Double

	NOTE: plotting options. Default value is 0.05.	
Edge_Width (Optional)	Input number to adjust the width of edges. ***NOTE: plotting options. Default value is 0.08.***	Double
Label_Size (Optional)	Input number to adjust the size of labels. ***NOTE: plotting options. Default value is 0.8.***	Double
Network_Graph_Plot	Output PDF file with a plot delineating grouping outcomes in a network graph. **NOTE: Node colors represent different clusters and node size for the amount of links from each node to others (node degrees). Edge width indicates the weight of links. Red links are links to nodes belonging to other groups, while grey links are links to nodes within the same group.**	File
Group_Number	Output shapefile representing all telecoupling systems used for the network grouping analysis. The shapefile will include an attribute that represents the group (cluster) number and can be used to color-code the outcome of the analysis on the map.	Shapefile
Network_Metrics_Table	Output CSV file showing network metrics (degree, closeness, and betweenness) calculated for each node. These metrics are node properties that measure the importance or centrality within the network. *Degree*: Importance score based purely on the number of edges (links) held by each node. A higher number means a more connected node. It is often considered a measure of direct influence. *Closeness*: Scores each node based on their 'closeness' to all other nodes within the network. It calculates the shortest paths between all nodes, then assigns each node a score based on its sum of shortest paths. It tells how many direct, 'one hop' connections each node has to other nodes within	File

	the network. Useful for finding very connected nodes who are likely to hold most information or nodes who can quickly connect with the wider network. Higher values mean less centrality. **NOTE**: *Closeness centrality can help find good 'broadcasters', but in a highly connected network you will often find all nodes have a similar score.* *Betweenness*: measures the number of times a node lies on the shortest path between other nodes. This measure shows which nodes act as 'bridges' between nodes in a network. It does this by identifying all the shortest paths and then counting how many times each node falls on one. Useful for finding the nodes that influence the flow around a system. A high betweenness count could indicate someone holds authority over, or controls collaboration between, disparate clusters in a network.

3.3 Agents Toolset

The Agents toolset contains script tools that are meant to map and visualize the geographical location of all entities (e.g. people, households, organizations, etc.) that facilitate the flow of goods, information, or ecosystem services between sending and receiving systems.

3.3.1 Add Agents Interactively

Interactively add telecoupling agents with an explicit spatial location to a map. The tool reads the agent spatial coordinates from the map and creates a permanent point feature class with a pre-assigned symbology. Upon successfully completing execution, the tool will automatically export the newly created features and their attributes as a .csv table. Use this tool if you do not have a pre-existing table on file with name, description (optional), and spatial coordinates of all telecoupling agents as part of your case study. Upon successfully completing execution, the tool will automatically export the newly created features and their attributes as a .csv table.

Parameter	Explanation	Data Type
Input_Features	Click on the map to add the spatial location of the agents. ***NOTE: the mouse cursor and its pre-assigned symbology should appear as soon as you move it over the map. The system component will be added to the map as a point feature, so make sure to zoom in to the correct spatial location before clicking on it. It is recommended that you type a name and a description (optional) in the Input_Attributes variable after each time you add a system component so that there is a correct one-to-one correspondence between the spatial location and its attributes.***	Feature Set
Input_Attributes	Type in a name and a description (optional) for each agent added to the map in the previous step. Click on the "+" sign to add a new row or "x" to delete. ***NOTE: The name field needs to be filled out, while the description field is optional. The number of rows needs to equal the number of agents (points) added to the map in the previous step.***	Record Set
Add_XY_Coordinates (Optional)	When checked, XY coordinate (Web Mercator by default) fields are added to the agents attribute table. When unchecked (this is the default), no coordinates are added.	Boolean

3.3.2 Draw Agents from Table

Draw telecoupling agents on the map by uploading a table stored in a local file on disk. The tool reads XY coordinates from the table and creates a permanent point feature class with a pre-assigned symbology. Upon successfully completing execution, the tool will automatically export the newly created features and their attributes as a .csv table. Use this tool if you already have a pre-existing table on file with name, description (optional), and spatial coordinates of all telecoupling agents as part of your case study. Upon successfully completing execution, the tool will automatically export the newly created features and their attributes as a .csv table.

Parameter	Explanation	Data Type
Input_Table	The table containing the name, description (optional), and XY coordinates of each agent.	Table View
X_Field	A numerical field in the input table containing the x coordinates (or longitudes) of each agent. ***NOTE: if the XY coordinate fields are not in the Web Mercator coordinate system, the tool will internally reproject to make sure the agents overlap correctly on the basemap. Although the latter step is not necessary, as ArcGIS Desktop can reproject features on-the-fly, the same will not apply when using the tool as a geoprocessing service consumed within a web application.***	Field
Y_Field	A numerical field in the input table containing the y coordinates (or latitudes) of each agent. ***NOTE: if the XY coordinate fields are not in the Web Mercator coordinate system, the tool will internally reproject to make sure the agents overlap correctly on the basemap. Although the latter step is not necessary, as ArcGIS Desktop can reproject features on-the-fly, the same will not apply when using the tool as a geoprocessing service consumed within a web application.***	Field

3.4 Flows Toolset

The Flows toolset contains script tools that are meant to map and visualize the spatial flow of goods, information, or ecosystem services between sending and receiving systems.

3.4.1 Draw Radial Flows

Draw radial flows between two or more telecoupling system components (i.e. "sending", "receiving", "spillover") on the map. Radial flows have a spoke-like pattern where features and places are mapped in nodal form with one place being a common origin or destination. Flows represent the movement of material, energy, or information. Arrows show the direction of the movement. The tool creates a new feature class containing geodetic line features constructed based on the values in a start x-coordinate field, start y-coordinate field, end x-coordinate field, and end y-coordinate field of a table. If desired, the tool will also compute a point vector layer representing nodes at each destination of the radial flows, using the same end x-coordinate field and end y-coordinate field of the attribute table used to produce flow lines. Points (nodes) at flow destinations can be symbolized and rendered as desired by each user to show quantities (attributes) of interest. Use this tool if you already have a pre-existing table on file with (at minimum) spatial coordinates of all the origin and destination points for each flow. If you have the amount of material, energy, or information in the table, use it to show differences in magnitude (e.g. line thickness) between flows.

Parameter	**Explanation**	**Data Type**
Input_Table	The table containing records on each flow between an origin and a destination. The input table that can be a text file, CSV file, Excel file, dBASE table, or geodatabase table ***NOTE: table attributes should at minimum include XY coordinates of both the origin and destination nodes of the flows.***	Table View
Start_X	A numerical field in the input table containing the x coordinates (or longitudes) of the starting points of lines to be positioned in the output coordinate system specified by the Spatial Reference parameter. ***NOTE: if the coordinate fields are not in the Web Mercator coordinate system, the tool will internally reproject to make sure the system components overlap correctly on the basemap. Although the latter step is not necessary, as ArcGIS Desktop can reproject features on-the-fly, the same will not apply when using the tool as a geoprocessing service consumed within a web application.***	Field
Start_Y	A numerical field in the input table containing the y coordinates (or latitudes) of the starting points of lines to be positioned in the output coordinate system specified by the Spatial Reference parameter.	Field

	NOTE: if the coordinate fields are not in the Web Mercator coordinate system, the tool will internally reproject to make sure the system components overlap correctly on the basemap. Although the latter step is not necessary, as ArcGIS Desktop can reproject features on-the-fly, the same will not apply when using the tool as a geoprocessing service consumed within a web application.	
End_X	A numerical field in the input table containing the x coordinates (or longitudes) of the ending points of lines to be positioned in the output coordinate system specified by the Spatial Reference parameter. ***NOTE: if the coordinate fields are not in the Web Mercator coordinate system, the tool will internally reproject to make sure the system components overlap correctly on the basemap. Although the latter step is not necessary, as ArcGIS Desktop can reproject features on-the-fly, the same will not apply when using the tool as a geoprocessing service consumed within a web application.***	Field
End_Y	A numerical field in the input table containing the y coordinates (or latitudes) of the ending points of lines to be positioned in the output coordinate system specified by the Spatial Reference parameter. ***NOTE: if the coordinate fields are not in the Web Mercator coordinate system, the tool will internally reproject to make sure the system components overlap correctly on the basemap. Although the latter step is not necessary, as ArcGIS Desktop can reproject features on-the-fly, the same will not apply when using the tool as a geoprocessing service consumed within a web application.***	Field
ID (Optional)	A field in the input table; this field and the values are included in the output and can be used to join the output features with the records in the input table.	Field

Line_Type (Optional)	The type of geodetic line to construct. GEODESIC— A type of geodetic line which most accurately represents the shortest distance between any two points on the surface of the earth. The mathematical definition of the geodesic line is quite lengthy and complex and therefore omitted here. This line type is the default. GREAT_CIRCLE—A type of geodetic line which represents the path between any two points along the intersection of the surface of the earth and a plane that passes through the center of the earth. Depending on the output coordinate system specified by the Spatial Reference parameter, in a spheroid-based coordinate system, the line is a great elliptic; in a sphere-based coordinate system, the line is uniquely called a great circle—a circle of the largest radius on the spherical surface. RHUMB_LINE—A type of geodetic line, also known as a loxodrome line, which represents a path between any two points on the surface of a spheroid defined by a constant azimuth from a pole. A rhumb line is shown as a straight line in the Mercator projection. NORMAL_SECTION—A type of geodetic line which represents a path between any two points on the surface of a spheroid defined by the intersection of the spheroid surface and a plane that passes through the two points and is normal (perpendicular) to the spheroid surface at the starting point of the two points. Therefore, the normal section line from point A to point B is different from the one from point B to point A.	String
Spatial_Reference (Optional)	The spatial reference of the coordinates found in the input table. On the Spatial Reference Properties dialog box you can Select, Import, or Create a New coordinate system. The default is Web Mercator or the input coordinate system if it is not Unknown.	Spatial Reference
Join_Fields (Optional)	The fields from the join table to be included in the join.	Multiple Value

Create_Destination_Nodes	If checked, a point layer (nodes) is created at each flow destination in addition to a line layer showing flows between locations. These points can be rendered and symbolized according to a specified attribute of interest (e.g. quantities traded). If **unchecked (default)**, only a line layer showing flows between locations will be created.	Boolean
Output_Flows_Layer_Name	Name of the output flows layer	String
Output_Nodes_Layer_Name (Optional)	Name of the output nodes point layer	String

3.4.2 Add Media Information Flows

News media and publication of books and articles heavily contribute to disseminate information on certain topics across the globe. Several online portals let you search through large databases for specific terms or academic publications on a subject of interest. LexisNexis® Academic[5] is an online research database where college faculty and students can find the critical news, legal and business information needed. For more information, visit the LexisNexis online portal, This telecoupling tool uses a search report file from LexisNexis in HTML format and parses the file to quantify the amount of publications, articles, or media outlets by geographic location (e.g. country). For example, a user may be interested in quantifying how much information about a nature reserve has spread out and where globally. Make sure the provided input feature layer corresponds to the geographic locations (e.g. countries) you are interested in searching within the report file.

Parameter	**Explanation**	**Data Type**
Input_Feature	An OGR-supported shapefile or Feature Layer corresponding to the geographic locations (e.g. countries) you are interested in searching within the media report file.	Feature Layer
Location_Field	A field in the input feature layer that will be used as a geographic location to search within the media report file.	Field
Media_Report_File	An HTML file created from a word search on the online LexisNexis portal.	File

[5] http://www.lexisnexis.com/en-us/products/lexisnexis-academic.page

Source_Location	Select on the map the geographic location from which information has emanated out. This location should correspond to the term used to search through the LexisNexis online database. ***NOTE: Make sure to select ONLY a SINGLE source location on the map.***	Feature Set
Merge_With_Existing_Flow_Layer (Optional)	If checked (default), radial flows generated from the media report file will be merged to an existing OGR-supported shapefile or Feature Layer corresponding to other flows of information.	Boolean
Input_Flow_Layer (Optional)	An optional feature layer representing existing radial flows of information that will be merged with the flows layer generated from the media report file. For example, you may have already generated radial flows of information generated by movement of tourists between sending and receiving systems.	Feature Layer
Line_Type (Optional)	The type of geodetic line to construct. • GEODESIC— A type of geodetic line which most accurately represents the shortest distance between any two points on the surface of the earth. The mathematical definition of the geodesic line is quite lengthy and complex and therefore omitted here. This line type is the default. • GREAT_CIRCLE—A type of geodetic line which represents the path between any two points along the intersection of the surface of the earth and a plane that passes through the center of the earth. Depending on the output coordinate system specified by the Spatial Reference parameter, in a spheroid-based coordinate system, the line is a great elliptic; in a sphere-based coordinate system, the line is uniquely called a great circle—a circle of the largest radius on the spherical surface. • RHUMB_LINE—A type of geodetic line, also known as a loxodrome line, which represents a path between any two points on the surface of a spheroid defined by a constant azimuth from a pole. A rhumb line is shown as a straight line in the Mercator projection.	String

	• NORMAL_SECTION—A type of geodetic line which represents a path between any two points on the surface of a spheroid defined by the intersection of the spheroid surface and a plane that passes through the two points and is normal (perpendicular) to the spheroid surface at the starting point of the two points. Therefore, the normal section line from point A to point B is different from the one from point B to point A.

3.5 Causes Toolset

The Causes toolset contains script tools that are meant to describe or statistically assess the potential leading causes of the flow of goods, information, or ecosystem services between sending and receiving systems.

3.5.1 Add Causes Interactively

Interactively select telecoupling causes from a pre-defined list and add them to the map at a location of interest (preferably linked to one of the telecoupling systems). The tool reads the causes location from the map and creates a permanent point feature class with a pre-assigned symbology. Upon successfully completing execution, the tool will automatically export the newly created features and their attributes as a .csv table. Use this tool if (a) the telecouling causes are merely descriptive or (b) do not have any dataset that can be analyzed with other exploratory statistical tools to help define factors associated with the telecoupling of interest.

Parameter	**Explanation**	**Data Type**
Input_Features	Click on the map to add a specific class of causes associated with the telecoupling of interest. Ideally, click on a spatial location that can be easily associated with one of the telecoupling systems, e.g. if the cause originates from a particular systems (sending, receiving, spillover). ***NOTE: the mouse cursor and its pre-assigned symbology should appear as soon as you move it over the map. Each class of causes is assigned letters that uniquely identify it and this will be added to the map as a point feature. Make sure to zoom in to the desired spatial location before clicking on the map. Type a description in the***	Feature Set

	Describe Causes table for each cause class you added. There must be a one-to-one correspondence between the spatial location of the cause added and its description attribute.	
Describe_Causes	Type in a description for each telecoupling cause added to the map in the previous step. Click on the "+" sign to add a new row or "x" to delete. ***NOTE: The description field is mandatory. The number of rows in this table needs to equal the number of causes (points) added to the map in the previous step.***	Record Set

3.5.2 Model Selection (OLS)

Geoprocessing tool for finding properly specified OLS models by exploring all combinations of candidate explanatory variables. This script tool has been slightly modified from the original ***Exploratory Regression*** GP tool developed by ESRI as part of the Spatial Statistics toolbox. The tool name has been changed but the main functionalities have been maintained. The following documentation is taken “as is” from the original one by ESRI.

- The primary output for this tool is a report file which is written to the Results window. Right-clicking on the Messages entry in the Results window and selecting View will display the Exploratory Regression summary report in a Message dialog box.

- This tool will optionally create a text file report summarizing results. This report file will be added to the table of contents (TOC) and may be viewed in ArcMap by right-clicking on it and selecting Open.

- This tool also produces an optional table of all models meeting your maximum coefficient p-value cutoff and Variance Inflation Factor (VIF) value criteria. A full explanation of the report elements and table is provided in Interpreting Exploratory Regression Results.

- This tool uses Ordinary Least Squares (OLS) and Spatial Autocorrelation (Global Moran's I). The optional spatial weights matrix file is used with the Spatial Autocorrelation (Global Moran's I) tool to assess model residuals; it is not used by the OLS tool at all.

- This tool tries every combination of the Candidate Explanatory Variables entered, looking for a properly specified OLS model. Only when it finds a model that meets your threshold criteria for Minimum Acceptable Adj R Squared, Maximum Coefficient p-value Cutoff, Maximum VIF Value Cutoff and Minimum Acceptable Jarque-Bera p-value will it run the

Spatial Autocorrelation (Global Moran's I) tool on the model residuals to see if the under/over-predictions are clustered or not. In order to provide at least some information about residual clustering in the case where none of the models pass all of these criteria, the Spatial Autocorrelation (Global Moran's I) test is also applied to the residuals for the three models that have the highest Adjusted R2 values and the three models that have the largest Jarque-Bera p-values.

- Especially when there is strong spatial structure in your dependent variable, you will want to try to come up with as many candidate spatial explanatory variables as you can. Some examples of spatial variables would be distance to major highways, accessibility to job opportunities, number of local shopping opportunities, connectivity measurements, or densities. Until you find explanatory variables that capture the spatial structure in your dependent variable, model residuals will likely not pass the spatial autocorrelation test. Significant clustering in regression residuals, as determined by the Spatial Autocorrelation (Global Moran's I) tool, indicates model misspecification. Strategies for dealing with misspecification are outlined in What they don't tell you about regression analysis.

- Because the Spatial Autocorrelation (Global Moran's I) is not run for all of the models tested (see the previous usage tip), the optional Output Results Table will have missing data for the SA (Spatial Autocorrelation) field. Because DBF (.dbf) files do not store null values, these appear as very, very small (negative) numbers (something like -1.797693e+308). For geodatabase tables, these missing values appear as null values. A missing value indicates that the residuals for the associated model were not tested for spatial autocorrelation because the model did not pass all of the other model search criteria.

- The default spatial weights matrix file used to run the Spatial Autocorrelation (Global Moran's I) tool is based on an 8 nearest neighbor conceptualization of spatial relationships. This default was selected primarily because it executes fairly quickly. To define neighbor relationships differently, however, you can simply create your own spatial weights matrix file using the Generate Spatial Weights Matrix File tool, then specify the name of that file for the Input Spatial Weights Matrix File parameter. Inverse Distance, Polygon Contiguity, or K Nearest Neighbors, are all appropriate Conceptualizations of Spatial Relationships for testing regression residuals.

The spatial weights matrix file is only used to test model residuals for spatial structure. When a model is properly specified, the residuals are spatially random (large residuals are intermixed with small residuals; large residuals do not cluster together spatially).

When there are 8 or less features in the Input Features, the default spatial weights matrix file used to run the Spatial Autocorrelation (Global Moran's I) tool is based on K nearest neighbors where K is the number of features minus 2. In general, you will want to have a minimum of 30 features when you use this tool.

Parameter	Explanation	Data Type
Input_Features	The feature class or feature layer containing the dependent and candidate explanatory variables to analyze.	Feature Set
Dependent_Variable	The numeric field containing the observed values you want to model using OLS.	Field
Candidate_Explanatory_Variables	A list of fields to try as OLS model explanatory variables.	Multiple Value
Weights_Matrix_File (Optional)	A file containing spatial weights that define the spatial relationships among your input features. This file is used to assess spatial autocorrelation among regression residuals. You can use the Generate Spatial Weights Matrix File tool to create this. When you do not provide a spatial weights matrix file, residuals are assessed for spatial autocorrelation based on each feature's 8 nearest neighbors. Note: The spatial weights matrix file is only used to analyze spatial structure in model residuals; it is not used to build or to calibrate any of the OLS models.	File
Maximum_Number_of_Explanatory_Variables (Optional)	All models with explanatory variables up to the value entered here will be assessed. If, for example, the Minimum Number of Explanatory Variables is 2 and the Maximum Number of Explanatory Variables is 3, the Exploratory Regression tool will try all models with every combination of two explanatory variables, and all models with every combination of three explanatory variables.	Long
Minimum_Acceptable_Adj_R_Squared (Optional)	This is the lowest Adjusted R-Squared value you consider a passing model. If a model passes all of your other search criteria, but has an Adjusted R-Squared value smaller than the value entered here, it will not show up as a Passing Model in the Output Report File. Valid values for this parameter range from 0.0 to 1.0. The default value is 0.05, indicating that passing models will explain at least 50 percent of the variation in the dependent variable.	Double

Maximum_Coefficient_p_value_Cutoff (Optional)	For each model evaluated, OLS computes explanatory variable coefficient p-values. The cutoff p-value you enter here represents the confidence level you require for all coefficients in the model in order to consider the model passing. Small p-values reflect a stronger confidence level. Valid values for this parameter range from 1.0 down to 0.0, but will most likely be 0.1, 0.05, 0.01, 0.001, and so on. The default value is 0.05, indicating passing models will only contain explanatory variables whose coefficients are statistically at the 95 percent confidence level (p-values smaller than 0.05). To relax this default you would enter a larger p-value cutoff, such as 0.1. If you are getting lots of passing models, you will likely want to make this search criteria more stringent by decreasing the default p-value cutoff from 0.05 to 0.01 or smaller.	Double
Maximum_VIF_Value_Cutoff (Optional)	This value reflects how much redundancy (multicollinearity) among model explanatory variables you will tolerate. When the VIF (Variance Inflation Factor) value is higher than about 7.5, multicollinearity can make a model unstable; consequently, 7.5 is the default value here. If you want your passing models to have less redundancy, you would enter a smaller value, such as 5.0, for this parameter.	Double
Minimum_Acceptable_Jarque_Bera_p_value (Optional)	The p-value returned by the Jarque-Bera diagnostic test indicates whether the model residuals are normally distributed. If the p-value is statistically significant (small), the model residuals are not normal and the model is biased. Passing models should have large Jarque-Bera p-values. The default minimum acceptable p-value is 0.1. Only models returning p-values larger than this minimum will be considered passing. If you are having trouble finding unbiased passing models, and decide to relax this criterion, you might enter a smaller minimum p-value such as 0.05.	Double
Minimum_Acceptable_Spatial_Autocorrelation_p_value (Optional)	For models that pass all of the other search criteria, the Exploratory Regression tool will check model residuals for spatial clustering using Global Moran's I. When the p-value for this diagnostic test is statistically significant (small), it indicates the model is very likely missing key explanatory variables (it isn't telling the whole story). Unfortunately, if you have spatial autocorrelation in your	Double

	regression residuals, your model is misspecified, so you cannot trust your results. Passing models should have large p-values for this diagnostic test. The default minimum p-value is 0.1. Only models returning p-values larger than this minimum will be considered passing. If you are having trouble finding properly specified models because of this diagnostic test, and decide to relax this search criteria, you might enter a smaller minimum such as 0.05.	

3.5.3 Factor Analysis for Mixed Data

Often times, population surveys record a high number of quantitative and qualitative information, whose purpose is to comprehensively describe socioeconomic characteristics associated with the interviewee. In statistical analysis, factor analysis is a useful tool for investigating variable relationships for complex concepts such as socioeconomic status, dietary patterns, or psychological scales. The most common statistical methodologies are principal component analysis/factor analysis for quantitative data, and multiple correspondence analysis (MCA) for qualitative data. Factor analysis of mixed data (FAMD), or factorial analysis of mixed data, is the factorial method devoted to data tables in which a group of individuals is described both by quantitative and qualitative variables. This tool uses the **FactoMineR** and **missMDA** R packages to factor analysis on any combination of continous and categorical variables. The **ESRI R Bridge** must be installed for this tool to run.

Parameter	**Explanation**	**Data Type**
Input_Table	A table with attributes corresponding to qualitative and/or quantitative variables.	Table View
Quantitative_Variables (Optional)	Select one or more quantitative variables from the input table.	Multiple Value
Qualitative_Variables (Optional)	Select one or more qualitative variables from the input table	Multiple Value
Supplementary_Quantitative_Variables (Optional)	Select one or more supplementary quantitative variables from the input table.	Multiple Value
Supplementary_Qualitative_Variables (Optional)	Select one or more supplementary qualitative variables from the input table.	Multiple Value

Compute_Missing_Values	If checked, missing values are replaced using PCA (for quantitative only datasets), MCA (for qualitative only datasets), or FAMD (for mixed qualitative and quantitative datasets) imputation. When unchecked, missing values are replaced by their column average.	Boolean
Number_of_Dimensions	Number of dimensions kept in the results.	Long
Color_Plots_By (Optional)	String corresponding to the color which are used. • Quantitative-only dataset: If "none", no color is used for the individuals; if "ind", a color for each individual ("ind"); if "var", color the individuals based on a categorical variable. • Qualitative-only dataset: . If "none", one color is used for the individual, another one for the categorical variables; if "quali", one color is used for each categorical variables; if "var", colors are used according to the different categories of a categorical variable • For mixed-data: If "none", no color is used for the individuals; if "ind", one color is used for each individual; if "var", colors are used according to the different categories of a categorical variable	String
Categorical_Variable (Optional)	When "var" is selected, this name corresponds to that of a categorical variable selected from the input table. The variable will be used to color final plots according to its different levels/categories.	Field
Individuals_Selection (Optional)	A selection of the individuals that are drawn: • "coord N": select the N elements that have the highest (squared) coordinates on the 2 first dimensions drawn • "contrib N": select the N elements that have the highest contribution on the 2 first dimensions drawn • "cos2 N": select the N elements that have the highest cos2 on the 2 dimensions drawn	String

	• "dist N": select the N elements that have the highest distance to the center of gravity	
Categories_Selection (Optional)	A selection of the categories that are drawn: • "coord N": select the N elements that have the highest (squared) coordinates on the 2 first dimensions drawn • "contrib N": select the N elements that have the highest contribution on the 2 first dimensions drawn • "cos2 N": select the N elements that have the highest cos2 on the 2 dimensions drawn • "dist N": select the N elements that have the highest distance to the center of gravity	String
Add_Label_to_Individuals (Optional)	When checked, individuals drawn on the plot are labelled.	Boolean
Output_PDF	Creates a PDF containing graphs generated from the plot function in the FactoMineR package. These graphs show individual and/or variable plots using the first two dimensions extracted from the analysis. For quantitative-only datasets, a graph of the eigenvalues associated with the extracted principal components is shown in the last page. For more information about these graphs, consult the documentation for the FactoMineR package.	File

3.6 Environmental Analysis Toolset

The Environmental Analysis toolset contains script tools that are meant to quantify actual or potential environmental effects that are directly or indirectly caused by a flow of goods, information, or ecosystem services between sending and receiving systems.

3.6.1 CO_2 Emissions (Wildlife Transfer)

Calculate the amount of CO2 emissions associated with each flow segment of material transported between two spatial locations. The tool reads the total length of each linear feature created by the Flows tool and multiplies it with the number of trips needed to transfer wildlife units (derived from number of units to transfer divided by a per-trip transportation capacity, e.g.

per-flight) and the amount of CO2 (kilograms) emitted per unit of length (e.g. meters) specified by the user. The tool also allows testing of future CO2 emission scenarios for wildlife transfer compared to current conditions. Make sure the unit of length corresponds to the coordinate system used (e.g. meters in the case of Web Mercator). Use this tool if you have a general estimate of how many kilograms of CO_2 per unit of length (e.g. meters) are associated with the type of transportation flow being considered.

Parameter	**Explanation**	**Data Type**
Input_Feature	Input Feature representing the linear flows of material between telecoupling systems across the globe.	Feature Layer
Wildlife_Units	The numeric field containing values for the number of wildlife units to be transferred.	Field
Transportation_Capacity	Integer number representing the maximum capacity of the transportation medium (e.g. airplane) used to transfer all wildlife units. **NOTE: this tool assumes homogeneity of transporation medium (e.g. same type of airplane), hence the estimates do not account for variations in CO2 emissions caused by multiple carriers or weather conditions (e.g. wind) affecting the transfer.**	Long
Units_of_CO2_Emitted	Specify the approximate amount (in kilograms) of CO2 emitted per unit of length. Make sure the unit of length corresponds to the coordinate system used (e.g. meters in the case of Web Mercator).	Double
Input_Feature__Scenario_ (Optional)	Input Feature representing a future scenario of linear flows of material between telecoupling systems across the globe.	Feature Layer
Wildlife_Units__Scenario_ (Optional)	The numeric field containing values for the number of wildlife units to be transferred under a future scenario.	Field
Transportation_Capacity__Scenario_ (Optional)	Integer number representing the maximum capacity of the transportation medium (e.g. airplane) used to transfer all wildlife units under a future scenario. **NOTE: this tool assumes homogeneity of transporation medium (e.g. same type of airplane), hence the estimates do not account for variations in CO2 emissions caused by**	Long

	multiple carriers or weather conditions (e.g. wind) affecting the transfer.	
Units_of_CO2_Emitted__Scenario_ (Optional)	Specify the approximate amount (in kilograms) of CO2 emitted per unit of length under a future scenario. Make sure the unit of length corresponds to the coordinate system used (e.g. meters in the case of Web Mercator).	Double

3.6.2 Forest Carbon Edge Effect (InVEST 3.3.3)

The InVEST carbon edge model extends the approach of the InVEST carbon model to account for forest carbon stock degradation due to the creation of forest edges. It applies known relationships between carbon storage and distance from forest edge to calculate edge effects in carbon storage, and combines these estimates with carbon inventory data to construct the overall carbon map. The model for edge effects pertains to above-ground carbon only, because edge effects have not been documented for the other carbon pools. For all other carbon pools, and for non-tropical forest classes, or if the model is run without edge effects, it follows the IPCC (2006) inventory approach to assigning carbon storage values by land cover class. ***NOTE: This model is recommended over the simple carbon storage and sequestration model***. ***To have a more detailed explanation on this model and its parameters, please read the official online InVEST documentation***.

Parameter	**Explanation**	**Data Type**
Land_Use_Land_Cover_Map	A GDAL-supported raster file, with an integer LULC code for each cell.	Raster Layer
Biophysical_Table	A CSV table containing model information corresponding to each of the land use classes in the LULC raster input. It must contain the fields 'lucode', 'is_tropical_forest', 'c_above'. If the user selects 'all carbon pools' the table must also contain entries for 'c_below', 'c_soil', and 'c_dead'. See the InVEST Forest Carbon User's Guide for more information about these fields.	File
Carbon_Pools_to_Calculate	If 'all carbon pools' is selected then the headers 'c_above', 'c_below', 'c_dead', 'c_soil' are used in the carbon pool calculation. Otherwise only 'c_above' is considered.	String

Compute_Forest_Edge_Effects	If selected, will use the Chaplin-Kramer, et. al method to account for above ground carbon stocks in tropical forest types indicated by a '1' in the 'is_tropical_forest' field in the biophysical table.	Boolean
Global_Forest_Carbon_Edge_Regression_Models (Optional)	A shapefile with fields 'method', 'theta1', 'theta2', 'theta3' describing the global forest carbon edge models. Provided as default data for the model.	Shapefile
Number_of_Nearest_Model_Points_to_Average (Optional)	Used when calculating the biomass in a pixel. This number determines the number of closest regression models that are used when calculating the total biomass. Each local model is linearly weighted by distance such that the biomass in the pixel is a function of each of these points with the closest point having the highest effect.	Long
Forest_Edge_Biomass_to_Carbon_Conversion_Factor (Optional)	Number by which to scale forest edge biomass to convert to carbon. Default value is 0.47 (according to IPCC 2006). This pertains to forest classes only; values in the biophysical table for non-forest classes should already be in terms of carbon, not biomass.	Double
Service_Area_of_Interest (Optional)	This is a set of polygons that will be used to aggregate carbon values at the end of the run if provided.	Shapefile

3.6.3 Crop Production (InVEST 3.3.3)

Expanding agricultural production and closing yield gaps is a key strategy for many governments and development agencies focused on poverty alleviation and achieving food security. However, conversion of natural habitats to agricultural production sites impacts other ecosystem services that are key to sustaining the economic benefits that agriculture provides to local communities. Intensive agricultural practices can add to pollution loads in water sources, often necessitating future costly water purification methods. Overuse of water also threatens the supply available for hydropower or other services. Still, crop production is essential to human well-being and livelihoods. The InVEST crop production model allows detailed examination of the costs and benefits of this vital human enterprise, allowing exploration of questions such as:

- How would different arrangement or selection of cropping systems compare to current systems in terms of total production? Could switching crops yield higher economic returns or nutritional value?
- What are the impacts of crop intensification on ecosystem services? If less land is used to produce equal amounts of food by increasing intensification, is the net result on ecosystem services production positive or negative?

- How can we evaluate different strategies for meeting increasing food demand while minimizing the impact on ecosystem services?

NOTE: to have a more detailed explanation on this model and its parameters, please read the official online InVEST documentation.

Parameter	Explanation	Data Type
Lookup_Table	Filepath to a CSV table used to convert the crop code provided in the Crop Map to the crop name that can be used for searching through inputs and formatting outputs.	File
Crop_Management _Scenario_Map	A GDAL-supported raster representing a crop management scenario.	Raster Layer
Global_Dataset_Fo lder	The provided folder should contain a set of folders and data specified in the 'Running the Model' section of the model's User Guide.	Folder
Yield_Function	The method used to compute crop yield. Can be one of three: 'observed', 'percentile', and 'regression'.	String
Percentile_Column (Optional)	For percentile yield function, the table column name must be provided so that the program can fetch the correct yield values for each climate bin.	String
Fertilizer_Raster_F older (Optional)	Path to folder that contains a set of GDAL-supported rasters representing the amount of Nitrogen (N), Phosphorous (P2O5), and Potash (K2O) applied to each area of land (kg/ha).	String
Irrigation_Map (Optional)	Filepath to a GDAL-supported raster representing whether irrigation occurs or not. A zero value indicates that no irrigation occurs. A one value indicates that irrigation occurs. If any other values are provided, irrigation is assumed to occur within that cell area.	Raster Layer
Compute_Nutrient _Contents	If checked, calculates nutrition from crop production and creates associated outputs.	Boolean
Crop_Nutrient_Inf ormation (Optional)	Filepath to a CSV table containing information about the nutrient contents of each crop.	File

Compute_Financial_Analysis	If checked, calculates economic returns from crop production and creates associated outputs.	Boolean
Crop_Economic_Information (Optional)	Filepath to a CSV table containing information related to market price of a given crop and the costs involved with producing that crop.	File

3.6.4 Habitat Quality (InVEST 3.3.3)

The InVEST Habitat Quality model uses habitat quality and rarity as proxies to represent the biodiversity of a landscape, estimating the extent of habitat and vegetation types across a landscape, and their state of degradation. The model combines maps of land use land cover (LULC) with data on threats to habitats and habitat response. Modeling habitat quality alongside ecosystem services enables users to compare spatial patterns and identify areas where conservation will most benefit natural systems and protect threatened species. This model does not attempt to place a monetary value on biodiversity. ***NOTE: to have a more detailed explanation on this model and its parameters, please read the official online InVEST documentation***.

Parameter	**Explanation**	**Data Type**
Current_Land_Cover	A GDAL-supported raster file. The current LULC must have its' own threat rasters, where each threat raster file path has a suffix of **_c**. Each cell should represent a LULC code as an Integer. The dataset should be in a projection where the units are in meters and the projection used should be defined.	Raster Layer
Future_Land_Cover (Optional)	Optional. A GDAL-supported raster file. Inputting a future LULC will generate degradation, habitat quality, and habitat rarity (If baseline is input) outputs. The future LULC must have it's own threat rasters, where each threat raster file path has a suffix of **_f**. Each cell should represent a LULC code as an Integer. The dataset should be in a projection where the units are in meters and the projection used should be defined. **The LULC codes must match the codes in the Sensitivity table**.	Raster Layer

Baseline_Land_Cover (Optional)	Optional. A GDAL-supported raster file. If the baseline LULC is provided, rarity outputs will be created for the current and future LULC. The baseline LULC can have it's own threat rasters (optional), where each threat raster file path has a suffix of **_b**. If no threat rasters are found, degradation and habitat quality outputs will not be generated for the baseline LULC. Each cell should represent a LULC code as an Integer. The dataset should be in a projection where the units are in meters and the projection used should be defined. The LULC codes must match the codes in the Sensitivity table. If possible the baseline map should refer to a time when intensive management of the landscape was relatively rare.	Raster Layer
Folder_Containing_Threat_Rasters	The selected folder is used as the location to find all threat rasters for the threats listed in the below table.	Folder
Threats_Data	A CSV file of all the threats for the model to consider. Each row in the table is a degradation source and each column contains a different attribute of each degradation source (THREAT, MAX_DIST, WEIGHT). **THREAT:**The name of the threat source and this name must match exactly to the name of the threat raster and to the name of it's corresponding column in the sensitivity table. **NOTE:**The threat raster path should have a suffix indicator (_c, _f, _b) and the sensitivity column should have a prefix indicator (L_). The THREAT name in the threat table should not include either the suffix or prefix. **MAX_DIST:**A number in kilometres (km) for the maximum distance a threat has an affect. **WEIGHT**: A floating point value between 0 and 1 for the the threats weight relative to the other threats. Depending on the type of habitat under review, certain threats may cause greater degradation than other threats. **DECAY:**A string value of either **exponential**or **linear**representing the type of decay over space for the threat.	File

	See the user's guide for valid values for these columns.	
Accessibility_to_Threats (Optional)	An OGR-supported vector file. The input contains data on the relative protection that legal / institutional / social / physical barriers provide against threats. The vector file should contain polygons with a field **ACCESS**. The **ACCESS** values should range from 0 - 1, where 1 is fully accessible. Any cells not covered by a polygon will be set to 1.	Shapefile
Sensitivity_of_Land_Cover_Types_to_Each_Threat	A CSV file of LULC types, whether or not they are considered habitat, and, for LULC types that are habitat, their specific sensitivity to each threat. Each row is a LULC type with the following columns: **LULC, HABITAT, L_THREAT1, L_THREAT2, ...** **LULC:** Integer values that reflect each LULC code found in current, future, and baseline rasters. **HABITAT:** A value of 0 or 1 (presence / absence) or a value between 0 and 1 (continuum) depicting the suitability of habitat. **L_THREATN:** Each L_THREATN should match exactly with the threat names given in the threat CSV file, where the THREATN is the name that matches. This is a floating point value between 0 and 1 that represents the sensitivity of a habitat to a threat. Please see the user guide for more detailed information on proper column values and column names for each threat.	File
Half_Saturation_Constant	A positive floating point value that is defaulted at 0.5. This is the value of the parameter k in equation (4). In general, set k to half of the highest grid cell degradation value on the landscape. To perform this model calibration the model must be run once in order to find the highest degradation value and set k for the provided landscape. Note that the choice of k only determines the spread and central tendency of habitat quality cores and does not affect the rank.	Double

3.6.5 Habitat Risk Assessment (InVEST 3.3.3)

The condition of a habitat is a key determinant of the ecosystem services it can provide. For example, multiple stressors including fishing, climate change, pollution and coastal development threaten the ability of coastal ecosystems to provide the valuable goods and services that people want and need. As human activities continue to intensify, so too does the need for quick, clear and repeatable ways of assessing the risks posed by human activities under various management plans. The InVEST habitat risk assessment (HRA) model allows users to assess the risk posed to coastal and marine habitats by human activities and the potential consequences of exposure for the delivery of ecosystem services and biodiversity. The InVEST HRA model is similar to the InVEST biodiversity model in that both models allow users to identify regions on a landscape or seascape where human impacts are highest. While the biodiversity model is intended to be used to assess how human activities impact biodiversity, the HRA model is better suited to screening the risk of current and future human activities to prioritize management strategies that best mitigate risk. We built and tested the HRA model in marine and coastal systems, and discuss it accordingly, but it easily can be applied to terrestrial systems, or mobile species. Risk of human activities (e.g., salmon aquaculture, coastal development, etc.) to habitats (e.g., seagrasses, kelp forests, mangroves, reefs) is a function of the exposure of each habitat to each activity and the consequences for each habitat. Exposure to stressors can arise through direct overlap in space and time or through indirect effects (i.e. finfish farms in an enclosed bay may degrade water quality and thus impede eelgrass growth throughout the bay, even if the netpens are not situated directly over eelgrass beds). Consequence depends on the effects of activities on habitat area and density, and the ability of habitats to recover from these effects (i.e. through processes such as recruitment and regeneration). Outputs from the model are useful for understanding the relative risk of human activities and climate change to habitats within a study region and among alternative future scenarios. Model outputs can help identify areas on the seascape where human activities may create trade-offs among ecosystem services by posing risk high enough to compromise habitat structure and function. The model can help to prioritize areas for conservation and inform the design and configuration of spatial plans for both marine and terrestrial systems. ***NOTE: to have a more detailed explanation on this model and its parameters, please read the official online InVEST documentation***.

Parameter	**Explanation**	**Data Type**
Use_Species_Risk	Checked: calculate the habitat risk assessment using species layers as input Unchecked (default): calculate the habitat risk assessment using habitat layers as input	Boolean
Species_Habitat_Layers_Folder	This folder should point to a directory containing all of the named habitat or species shapefile layers that you wish to include in this model run. All layers must be projected in the same projection.	Folder

Stressors_Folder	Folder containing stressors to be overlapped with habitats and/or species. This directory should contain ONLY the stressors desired within this model run. All layers must be projected in the same projection.	Folder
Spatially_Explicit_Criteria_Folder (Optional)	If spatially explicit criteria is desired, this input should point to an outer directory for all spatial criteria. A rigid structure **MUST** be followed in order for the model to run. Within the outer spatial criteria folder, there **MUST** be the following 3 folders: Sensitivity, Exposure, and Resilience. Vector criteria may then be placed within the desired folder. Each feature in the shapefiles used **MUST** include a 'Rating' attribute which maps to a float or int value desired for use as the rating value of that spatial criteria area. • Any criteria placed within the Resilience folder will apply only to a given habitat. They should be named with the form: habitatname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore. • Any criteria placed within the Exposure folder will apply to the overlap between a given habitat and a given stressor. They should be named with the form: habitatname_stressorname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore. Any criteria placed within the Sensitivity folder will apply to the overlap between a given habitat and a given stressor. They should be named with the form: habitatname_stressorname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore.	Folder
Criteria_Scores_Folder	Folder containing CSV files providing all criteria information for the run of the Habitat Risk Assessment. There are two types of CSVs- habitat overlap CSVs and the stressor buffer CSV. Habitat CSVs will contain not only habitat-specific criteria information, but also all criteria that impact the overlap between that habitat and all applicable stressors. The stressor buffer CSV will be a single file containing the desired buffer for all stressors included in the assessment. Habitat CSVs should be filled out	Folder

with habitat-specific criteria information as well as any criteria which apply to the overlap of the given habitat and stressors. The Stressor Buffer CSV should be filled out with the desired numerical buffer which can be used to expand a given stressor's influence within the model run. This can be 0 if no buffering is desired for a given stressor, but may NOT be left blank.

Any criteria which use spatially explicit criteria (specified by the user during the HRA Preprocessor) will be noted in the CSV by the word 'SHAPE' in the rating column for that habitat, stressor, or combined criteria. The user should still fill in a Data Quality and Weight for these criteria, but should NOT remove the 'SHAPE' string unless they no longer desire to use a spatial criteria for that attribute.

When preprocessor is run, the CSVs will contain no numerical ratings, only guidance on how each rating might be filled out. The user should use the best available data sources in order to obtain rating information. The column information to be filled out includes the following:

1. "Rating"- This is a measure of a criterion's impact on a particular habitat or stressor, with regards to the overall ecosystem. Data may come from a combination of peer-reviewed sources at the global scale and locally available fine-scale data sources. Model inputs and results can be updated as better information becomes available. We provide guidance for well-known criteria on a scale of 0-3, but it should be noted that if information is available on a different scale, this can also be used. It is important to note, however, that all rating information across all CSVs should be on one consistent scale, regardless of what the upper bound is.
2. "DQ"- This column represents the data quality of the score provided in the 'Rating' column. Here the model gives the user a chance to downweight less-reliable data sources, or upweight particularly well-studied criteria. While we provide guidance for a scoring system of 1-3, the user should feel free to use any upper bound they feel practical, as long as the scale is consistent. The lower

bound, however, should ALWAYS be 1, unless the user wishes to remove the entire criteria score.

3. “Weight”- Here the user is given the opportunity to upweight critiera which they feel are particularly important to the system, independent of the source data quality. While we provide guidance for a scoring system from 1-3, the user should feel free to use any upper bound they feel practical, as long as the scale is consistent. The lower bound, however, should ALWAYS be 1 unless the user wishes to remove the entire criteria score.
4. (Optional) “E/C”- This column indicates whether the given criteria are being applied to the exposure or the consequence portion of the chosen risk equation. These can be manually changed by the user on a single criterion basis, however, we would strongly recommend against it. If the user desires to change that criterion’s allocation, it would be better to change the allocation of the criterion within the Resilience, Exposure, Sensitivity categories using the HRA Preprocessor User Interface. By default, any criteria in the Sensitivity or Resilience categories will be assigned to Consequence (C) within the risk equations, and any criteria within the Exposure category will be assigned to Exposure (E) within the risk equation.

NOTE:

Required ratings data - We recommend users include information about all of the key components of risk (i.e., spatial overlap and other exposure criteria, consequence criteria and the components of consequence, resilience and sensitivity. Nevertheless, the model will produce estimates for risk with only the habitat and stressor spatial layers and no other exposure values (i.e., E = 0 = no score for all other exposure criteria). To produce these estimates, the model does require values for at least one consequence criteria, either sensitivity or resilience. Without this information, the model will return an error message. If the user inputs scores for only sensitivity or resilience, then the consequence score will be based on those data alone.

	NOTE: **Specifying No Interaction Between Habitat and Stressor** - As of InVEST 3.3.0 the HRA model will allow users to indicate that a habitat - stressor pair should have no interaction. This essentially means that the model will consider the habitat and stessor have no spatial overlap. This enhancement is to deal with the issue of having fine resolution vector data where the values may share the same pixel space when converted to a raster grid format. To set a habitat - stressor pair to no overlap, simply fill in each criterias "Rating" column with an "NA" value for the given pair. ALL "Rating" values for that pair must be set to "NA" for the model to consider the pair to have no interaction / overlap. If an "NA" value is found, but not all are set, an error message will be presented.	
Resolution_of_Analysis__meters_	The size in meters that is desired for the analysis of the shapefile layers at a grid cell scale. This will define the width and height of each unique risk grid cell. This must be a whole number. The user should base this size on the resolution of the habitat data and scale at which habitats are distributed in space. For example, small patches of seagrasses and kelp are often about 100-200 square meters, which is about the smallest resolution we recommend running the model. If the input habitat data are coarse, then a minimum of 500 meters is better. If you examine your risk outputs and find that the edges of patches of habitat have regular and distinct variation in risk, such that every high and medium risk cell on the edge of habitat patches are border by low risk cells, consider enlargening your resolution. We recommend running the model for the first time at a low resolution (500 m or 1 km) to verify that the model is running properly. Then use a higher resolution in subsequent runs.	Double
Risk_Equation	This selection chooses the equation that will be used when calculating risk to a given habitat. The user may choose either either a Euclidean risk model, or a Multiplicative risk model. See online documentation of InVEST HRA model for more details.	String
Decay_Equation	This selection influences how the "zone of influence" (i.e., buffer distance) of a stressor will be applied to risk. The stressor buffer distance in the stressor buffer CSV can be degraded to provide a more accurate depiction of the influence of a stressor beyond its footprint. The decay equation decays the overall	String

	exposure rating (e.g., combined spatial overlap, temporal overlap, intensity, management effectiveness) before the value for E goes into the risk equation. For each pixel, the model uses the value of the decayed exposure score. The options for decay are as follows. “None” will apply the full exposure to the full range of the stressor footprint plus buffer, without any decay. “Linear” and “Exponential” will use the stated equation as a model for decay from the edges of the footprint to the extent of the buffer distance.	
Maximum_Criteria_Score	The maximum criteria score is the user-reported highest value assigned to any criteria rating within the assessment. This will be used as the upper bounded value against which all rating scores will be compared. For example, in a model run where the ratings scores vary from 0-3, this would be a 3. If the user chooses to use a different scale for ratings, however, this should be the highest value that could be potentially assigned to a criteria. If the model run is using spatially explicit criteria, this value should be the maximum value assigned to either a criteria feature or to a CSV criteria rating.	Long
Maximum_Overlapping_Stressors	The is the largest number of stressors that overlap withing the analysis zone. This will be used in order to make determinations of low, medium, and high risk for a given habitat. If the number of overlapping stressors provided is too low, results will likely show more medium and high risk areas than are present. Conversely, if the number of overlapping stressors is too high, it will be difficult for areas to break the threshold to show up as medium or high risk. If unsure how many stressors overlap, we recommend running the overlap analysis tool without weighting.	Long
Subregions__shapefile_	The model will use a subregions shapefile to generate an HTML table of averaged exposure, consequence, and risk values within each subregion by habitat and stressor. In addition, if the Risk Equation chosen is Euclidean, the model will also generate a series of figures which clearly display the exposure-consequence ratings and the resulting risk results for each habitat-stressor combination by subregion. It will also create a figure showing cumulative ecosystem risk for all subregions habitats in the study. Each of the subregion shapefile features **MUST contain a ‘Name’ attribute** in order to be properly included in the subregion averaging. If subregion data is not available for the given study region, an AOI for the area could also be used in	Shapefile

order to obtain averaged data per habitat-stressor pair. However, the AOI must also contain a 'Name' attribute.

3.6.6 Habitat Risk Assessment Preprocessor (InVEST 3.3.3)

Before running the HRA model, it is necessary to concatenate and rate all applicable criteria information. This can be accomplished by running the Preprocessor tool, then editing the resulting CVSs. If you have already run the model, or have the 'habitat_stressor_ratings' directory from a previous HRA Preprocessor run, you may skip this step and proceed to running the Habitat Risk Assessment tool.

Parameter	Explanation	Data Type
Use_Species_Risk	Checked: calculate the habitat risk assessment using species layers as input Unchecked (default): calculate the habitat risk assessment using habitat layers as input	Boolean
Species_Habitat_Layers_Folder	This folder should point to a directory containing all of the named habitat or species shapefile layers that you wish to include in this model run. All layers must be projected in the same projection.	Folder
Stressors_Folder	Folder containing stressors to be overlapped with habitats and/or species. This directory should contain ONLY the stressors desired within this model run. All layers must be projected in the same projection.	Folder
Exposure_Criteria	**Choose at least 1 criteria for each category below, and at least 4 total**. Exposure criteria apply to the overlap between provided habitats and stressors, or species and stressors being used within this model run. These criteria will be applied to the exposure portion of habitat risk.	Record Set
Consequence__Sensitivity	Sensitivity criteria apply to the overlap between provided habitats and stressors, or species and stressors being used within this model run. These criteria will be applied to the consequence portion of habitat risk.	Record Set
Consequence__Resilience	Resilience criteria apply ONLY to the habitats or species being used within this model run. These criteria will be applied to the consequence portion of habitat risk.	Record Set

Use_Spatially_Explicit_Risk_Scores_in_Shapefile	Checked: checking this box indicates that model should use criteria from provided shapefiles Unchecked (default): do NOT use spatial criteria from provided shapefiles	Boolean
Spatially_Explicit_Criteria_Folder (Optional)	Each shapefile in the folder directories will need to contain a 'Rating' attribute to be used for calculations in the HRA model. If spatially explicit criteria is desired, this input should point to an outer directory for all spatial criteria. A rigid structure **MUST**be followed in order for the model to run. Within the outer spatial criteria folder, there **MUST**be the following 3 folders: Sensitivity, Exposure, and Resilience. Vector criteria may then be placed within the desired folder. Each feature in the shapefiles used **MUST**include a ‘Rating’ attribute which maps to a float or int value desired for use as the rating value of that spatial criteria area. • Any criteria placed within the Resilience folder will apply only to a given habitat. They should be named with the form: habitatname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore. • Any criteria placed within the Exposure folder will apply to the overlap between a given habitat and a given stressor. They should be named with the form: habitatname_stressorname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore. Any criteria placed within the Sensitivity folder will apply to the overlap between a given habitat and a given stressor. They should be named with the form: habitatname_stressorname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore.	Folder

3.6.7. Nutrient Delivery Ratio (InVEST 3.3.3)

Land use change, and in particular the conversion to agricultural lands, dramatically modifies the natural nutrient cycle. Anthropogenic nutrient sources include point sources, e.g. industrial effluent or water treatment plant discharges, and non-point sources, e.g. fertilizer used in agriculture and residential areas. When it rains or snows, water flows over the landscape carrying pollutants from these surfaces into streams, rivers, lakes, and the ocean. This has consequences for people, directly affecting their health or well-being, and for aquatic ecosystems that have a limited capacity to adapt to these nutrient loads.

One way to reduce non-point source pollution is to reduce the amount of anthropogenic inputs (i.e. fertilizer management). When this option fails, ecosystems can provide a purification service by retaining or degrading pollutants before they enter the stream. For instance, vegetation can remove pollutants by storing them in tissue or releasing them back to the environment in another form. Soils can also store and trap some soluble pollutants. Wetlands can slow flow long enough for pollutants to be taken up by vegetation. Riparian vegetation is particularly important in this regard, often serving as a last barrier before pollutants enter a stream.

Land-use planners from government agencies to environmental groups need information regarding the contribution of ecosystems to mitigating water pollution. Specifically, they require spatial information on nutrient export and areas with highest filtration. The nutrient delivery and retention model provides this information for non-point source pollutants. The model was designed for nutrients (nitrogen and phosphorous), but its structure can be used for other contaminants (persistent organics, pathogens etc.) if data are available on the loading rates and filtration rates of the pollutant of interest.

NOTE: For all raster inputs for this tool, the projection should be defined and the projection's linear units should be in meters.

Parameter	Explanation	Data Type
DEM	A GIS raster layer, with an elevation value for each cell. Make sure the DEM is corrected by filling in sinks. To ensure proper flow routing, the DEM should extend beyond the watersheds of interest, rather than being clipped to the watershed boundaries.	Raster Layer
Land_Use_Land_Cover	A GIS raster layer, with an integer LULC code for each pixel. The LULC code should be an integer.	Raster Layer
Nutrient_Runoff	A GIS raster dataset representing the spatial variability in runoff potential, i.e. the capacity to transport nutrient downstream. This raster can be defined as a quickflow	Raster Layer

	index (e.g. from the InVEST seasonal water yield model) or simply as annual precipitation.	
Watersheds	A shapefile of polygons. This is a layer of watersheds such that each watershed contributes to a point of interest where water quality will be analyzed. The shapefile must include a field titled *ws_id* numbered from 0 to N, where N is the total number of watersheds.	Shapefile
Biophysical_Table	A .csv table of land use/land cover (LULC) classes, containing data on water quality coefficients used in this tool. These data are attributes of each LULC class rather than attributes of individual cells in the raster map. Each row in the table is an LULC class while each column contains a different attribute of each land use/land cover class. The columns must be named as: • *lucode*: Unique integer for each LULC class (e.g., 1 for forest, 3 for grassland, etc.), must match the LULC raster above. • *LULC_desc* (optional): Descriptive name of land use/land cover class. • *eff_n (*and *eff_p)*: The maximum retention efficiency for each LULC class, varying between zero and 1. The nutrient retention capacity for a given vegetation is expressed as a proportion of the amount of nutrient from upstream. For example, high values (0.6 to 0.8) may be assigned to all natural vegetation types (such as forests, natural pastures, wetlands, or prairie), indicating that 60-80% of nutrient is retained. Suffix _n stands for nitrogen, and _p for phosphorous. • *crit_len_n (*and *crit_len_p)* (in meters): The distance after which it is assumed that a patch of LULC retains nutrient at its maximum capacity. If nutrients travel a distance smaller than the retention length, the retention efficiency will be less than the maximum value eff_x, following an exponential decay.	File

	• *proportion_subsurface_n* (or *proportion_subsurface_p)* (optional): The proportion of dissolved nutrients over the total amount of nutrients, expressed as ratio between 0 and 1. By default, this value should be set to 0, indicating that all nutrients are delivered via surface flow.	
Threshold_Flow_Accumulation	Integer value defining the number of upstream pixels that must flow into a pixel before it's considered part of a stream. This is used to generate a stream layer from the DEM. This threshold expresses where hydrologic routing is discontinued, i.e. where retention stops and the remaining pollutant will be exported to the stream. The default is 1 over the pixel area (in km2), i.e. ~1000 for 30m resolution DEM.	Long
Borselli_k_Parameter	Calibration parameter that determines the shape of the relationship between hydrologic connectivity (the degree of connection from patches of land to the stream) and the sediment delivery ratio (percentage of soil loss that actually reaches the stream). The default value is 2.	Long
Subsurface_Critical_Length_Phosphorous (and Nitrogen) (in meters)	The distance (traveled subsurface and downslope) after which it is assumed that soil retains nutrient at its maximum capacity. If dissolved nutrients travel a distance smaller than subsubsurface_crit_length, the retention efficiency is lower than the maximum value defined above. Setting this value to a distance smaller than the pixel size will result in the maximum retention efficiency being reached within one pixel only.	Double
Subsurface_Maximum_Retention_Efficieny_Phosphorous (and Nitrogen)	The maximum nutrient retention efficiency that can be reached through subsurface flow, a value between 0 and 1. This field characterizes the retention due to biochemical degradation in soils.	Double

3.6.8 Fisheries Harvest (InVEST 3.3.3)

Wild capture fisheries provide a significant source of protein for human consumption and directly employ nearly 40 million fishers worldwide. However, poor harvesting practices and habitat loss and degradation can reduce the ability of ecosystems to support healthy, productive fisheries. The InVEST Fisheries Production model produces estimates of harvest volume and economic value of single-species fisheries. The model is an age- or stage-structured population model, and is presented as a generic model that can be adapted to most species and geographies. Inputs to the model include parameters for life history characteristics (e.g., age at maturity, recruitment, migration and natural mortality rates), behavior of the fishery (e.g., fishing pressure), habitat dependencies (e.g., importance and availability of nursery habitat), and, optionally, economic valuation (e.g., price per unit biomass). The model outputs the volume and economic value of harvest within the area(s) designated by the user. It is best to compare outputs from multiple runs of the model, where each run represents different scenarios of habitat extent, environmental conditions and/or fishing pressure.

Parameter	**Explanation**	**Data Type**
Area_of_Interest	The provided shapefile is used to display outputs within the area of interest. The shapefile must contain a column titled "Name", which corresponds to the subregion specified in the Population Parameters CSV file.	Shapefile
Number_of_Steps_for_Model_Run	The number of time steps the simulation shall execute before completion. Must be a positive integer.	Double
Population_Model_Type	Specifies whether the classes provided in the Population Parameters CSV file represent ages or stages. If the stage-based model is selected, the Population Parameters CSV File must include a ‘Duration’ vector alongside the survival matrix that contains the number of time steps that each stage lasts.	String
Population_Classes_are_Sex_Specific	Specifies whether or not the population classes provided in the Population Parameters File are distinguished by sex.	String
Harvest_by_Individuals_or_Weight	Specifies whether the harvest output values are calculated in terms of number of individuals or in terms of biomass (weight). If ‘Weight’ is selected, the Population Parameters CSV File must include a ‘Weight’ vector alongside the survival matrix that contains the weight of each age/stage, as well as sex if the model is sex-specific.	String

Population_Parameters_File_CSV	The provided CSV file should contain all necessary parameters for population groups based on age/stage, sex, and subregion: • **Classes (required)-** The leftmost column should contain the age/stage names of the given species listed in chronological order. Each name can be an alphanumeric string. If the population classes are sex-specific, all age/stage names of one sex must be listed first, followed below by the age/stage names of the other sex. • **Subregions (required)-** The top-most row should contain the subregion names considered by the model. Each name can be an alphanumeric string. If the AOI shapefile is to be provided, the subregion entries must each match a corresponding 'Name' attribute value in a feature of the AOI shapefile. An entry must be provided even if the model is considering only one subregion. • **Survival Rates from Natural Mortality Matrix (required)-** Each unique pair of age/stage and subregion should contain a survival rate from natural mortality, expressed as a decimal fraction. • **ExploitationFraction (required)-** A row starting in the first column with the label 'ExploitationFraction'. The exploitation fraction is the proportion of the vulnerable population in each subregion that is harvested (0=0% harvested, 1=100% harvested). Each subregion is treated independently (i.e. up to 100% of the vulnerable	File

population in each subregion may be harvested).

- **LarvalDispersal-** A row starting in the first column labeled 'LarvalDispersal'. The larval dispersal is the proportion of the cumulative larvae pool that disperses into each subregion. Each subregion column should have a decimal to represent this. Dispersal across all subregions should add up to 1. If larval dispersal isn't provided, larvae will be dispersed equally across all subregions.
- **VulnFishing (required)-** A column labeled 'VulnFishing', which is the relative vulnerability to harvest for each class. A decimal value for each class listed in this column is required. The most vulnerable age(s)/stage(s) should have a value of 1.0, indicating full vulnerability.
- **Maturity-** A column labeled 'Maturity'. It represents the fraction of that age or stage which is mature and contributes to the spawning stock. A decimal value for each age/stage is required if maturity is included. For classes which do not reproduce, this should be 0.
- **Duration-** A column labeled 'Duration'. This column is required for stage-based models. It represents the number of time steps for which an average individual will be in that stage before moving to the next one.
- **Weight-** A column which is required if 'Spawners by Weight' or 'Harvest by Weight' is selected. This is the average biomass of an individual of the population at each age/stage

	expressed in model-agnostic units and is required for each of the ages/stages listed in the classes column.	
Total_Initial_Recruits	The initial number of recruits in the population model at time equal to zero. If the model contains multiple subregions of interest or is distinguished by sex, this value will be first divided into subregions using the LarvalDispersal vector and then further divided evenly by sex of each subregion.	Double
Recruitment_Function_Type	This equation will be used to calculate recruitment into each subregion in the area of interest (options: Beverton-Holt or Ricker).	String
Spawners_by_Individuals_or_Weight	Specifies whether the spawner abundance used in the recruitment function should be calculated in terms of number of individuals or in terms of biomass (weight). If ‘Weight’ is selected, the user must provide a ‘Weight’ vector alongside the survival matrix in the Population Parameters CSV File. The ‘Alpha’ and ‘Beta’ parameters provided by the user should correspond to the selected choice.	String
Alpha	Specifies the shape of the stock-recruit curve.	Double
Beta	Specifies the shape of the stock-recruit curve.	Double
Estimate_Migratory_Patterns	Select this option if you wish to estimate migratory patterns.	Boolean
Migration_Matrix_CSV_Folder (optional)	Specify the path to a folder containing a migratory matrix CSV file. Consider the following when formating the migratory CSV file: For each age/stage where migration occurs, there should be a single CSV within the migration directory. The name of the CSV can be anything, but MUST end with an underscore followed by the name of the age or stage. This MUST correspond to an age or stage within the Population Parameters CSV File. For migration from the ‘adult’ class for example, a migration file might be named ‘migration_adult.csv’. The CSV should contain nothing besides subregion names and migration values.	Folder (containing Migration Matrix CSV file)

	The first row and first column should be the names of the subregions in the Population Parameters CSV File, listed in the same order. The columns represent the sources — the subregions FROM which the migration occurs; each column should therefore sum to 1. The rows represent the sinks — the subregions TO which the migration occurs. The cells within the matrix should be a DECIMAL REPRESENTATION of percentage of the source's population which will migrate to the sink.	
Estimate_Harvest_ Value	Select this option if you wish to estimate harvest value.	Boolean
Fraction_of_Harvest_ Kept_After_Processing (optional)	This is the decimal representation of the percentage of harvested catch remaining after post-harvest processing is complete.	Double
Unit_Price (optional)	Specifies the price per harvest unit.	Double

3.6.9 Seasonal Water Yield (InVEST 3.3.3)

There is a high demand for tools estimating the effect of landscape management on water supply service (e.g. for irrigation, domestic use, hydropower production). The InVEST seasonal water yield model seeks to provide guidance regarding the contribution of land parcels to the generation of both baseflow and quick flow. The model computes spatial indices that quantifies the relative contribution of a parcel of land to the generation of both baseflow and quick flow.

Parameter	**Explanation**	**Data Type**
Evapotranspiration_Directory	Maps of monthly reference evapotranspiration (mm). NOTE: Provide path to the folder containing the evapotranspiration rasters. There must be a raster for each month, and rasters must be saved with the month number (e.g., ET_1.tif). Rasters should be saved as .tif files.	Folder
Precipitation_Directory	Maps of monthly precipitation (mm). NOTE: Provide path to the folder containing the precipitation raster. There must be a raster for each month, and rasters must be saved with the month number (e.g., Precip_1.tif). Rasters should be saved as .tif files.	Folder

Digital_Elevation_Model_DEM	Digital elevation model.	Raster Layer
Land_Use_Land_Cover_LULC_Raster	Map of LULC. Each LULC must have a corresponding code in the biophysical table.	Raster Layer
Soil_Group_Raster	Map of SCS soil hydrologic groups (A, B, C, or D), used in combination of the LULC map to compute the CN map. Values are entered as integers, with 1, 2, 3, and 4, corresponding to groups A, B, C, and D, respectively. Soil groups are determined from hydraulic conductivity soil depths. FutureWater has created a global map of hydraulic conductivity available at: http://www.futurewater.eu/2015/07/soil-hydraulic-properties/ One can also find guidance online, e.g.: www.bwsr.state.mn.us/outreach/eLINK/Guidance/HSG_guidance.pdf.	Raster Layer
Watershed_AOI_Shapefile	Shapefile delineating the boundary of the area(s) of interest, or watershed to be modeled.	Shapefile
Biophysical_Table	Table comprising, for each LULC type: CN for each soil type Monthly Kc values CN can be obtained from the USDA handbook: (NRCS-USDA, 2007 Chapter 9). Monthly Kc values can be obtained from the FAO guidelines: Allen et al., 1998. For water bodies and wetlands that are connected to the stream, CN can be set to 99 (i.e., assuming that those pixels rapidly convey quickflow). NOTE: when the focus is on potential flood effects, CN may be selected to reflect wet antecedent runoff conditions: CN values should then be converted to	File

	ARC-III conditions, as per Chapter 10 in NRCA-USDA guidelines (2007). Column names within the csv file must be formatted as follows: CN_A, CN_B, CN_C, CN_D, Kc_1, …, Kc_12. Refer to sample data for assistance formatting biophysical table.	
Rain_Events_Table	Table with 12 values of rain events per month. A rain event is defined as >0.1mm. CSV must be formatted with two columns, the first titled *month* and the second titled *events*. The month column should be numbered 1 to 12, and the events column should indicate the number of rain events in each corresponding month.	File
alpha_m_Parameter	Model parameter. Default value: 0.083333	Double
beta_i_Parameter	Model parameter. Default value: 1.0	Double
gamma_Parameter	Model parameter. Default value: 10	Double
Threshold_Flow_ Accumulation	The number of upstream cells that must flow into a cell before it is considered part of a stream, which is used to classify streams in the DEM. Default value: 1000.	Long

References:

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration - Guidelines for computing crop water requirements, FAO Irrigation and drainage paper 56. Rome, Italy.

NRCS-USDA, 2007. National Engineering Handbook. United States Department of Agriculture,http://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/national/water/?cid=stelprdb1043063.

3.7 Socioeconomic Analysis Toolset

The Socioeconomic Analysis toolset contains script tools that are meant to quantify actual or potential socioeconomic effects that are directly or indirectly caused by a flow of goods, information, or ecosystem services between sending and receiving systems.

3.7.1 Cost-Benefit Analysis (Wildlife Transfer)

Transfer of wildlife between tourism hot-spot locations (e.g. zoos) has become increasingly common worldwide because of the increased demand for exotic species exhibits. As a natural consequence, this type of business comes with a set of associated costs (e.g. airline transportation, food, maintenance labor) and benefits (e.g. economic revenues from the sale of the wildlife species, food production for feeding purposes, and revenues from tourism). This tool uses estimated costs and revenues associated with a given wildlife transfer (e.g. Panda loan) and telecoupling system involved, and calculates the total returns of investment (returns = total revenues - total costs). New fields showing total costs, revenues, and returns will be added to the feature layer provided as input and can be symbolized like any other quantitative feature attributes. ***If you do not have information on certain or all approximate costs/revenues for a given telecoupling system (receiving, sending, spillover), leave the value blank or n/a in the table required by tool as one of the inputs***.

Parameter	Explanation	Data Type
Input_Feature	An OGR-supported shapefile or Feature Layer containing information about all telecoupling systems. The layer should have at minimum a field indicating the name of the system (e.g. country).	Feature Layer
Input_Join_Field	A field in the input feature layer used by the tool to join with the economics table provided by the user.	Field
Economics_Table	A CSV table containing costs and revenues for each telecoupling system. The table should have a field that can be joined to the input feature layer representing all telecoupling systems.	Table View
Economics_Table_Join_Field	A field in the economics CSV table that will be joined to the input feature layer representing all telecoupling systems.	Field
Output_Layer_Name	Name of the output cost-benefit analysis layer	String

3.7.2 Visitation: Recreation and Tourism Storage and Sequestration (InVEST 3.3.3)

Recreation and tourism are important components of many national and local economies and they contribute in innumerable ways to physical wellbeing, learning, and quality of life. To quantify the value of natural environments, the InVEST recreation model predicts the spread of person-days of recreation, based on the locations of natural habitats and other features that factor into people's decisions about where to recreate. In the absence of empirical data on visitation, we parameterize the model using a proxy for visitation: geotagged photographs posted to the website flickr. Using photographs, the model predicts how future changes to natural features will alter visitation rates and outputs maps showing current and future patterns of recreational use. ***NOTE: this computer must have an Internet connection in order to run this model***. ***To have a more detailed explanation on this model and its parameters, please read the official online InVEST documentation***.

Parameter	**Explanation**	**Data Type**
Area_of_Interest	An OGR-supported vector file representing the area of interest where the model will run the analysis.	Feature Layer
Start_Year	Year to start PUD calculations, date starts on Jan 1st. Inclusive, must be >= 2005.	String
End_Year	Year to end PUD calculations, date ends and includes Dec 31st. Inclusive, must be <= 2014.	String
Compute_Regression	When checked (default), run a simple regression model using the parameters specified below.	Boolean
Predictor_Table (Optional)	A table that maps predictor IDs to files and their types with required headers of 'id', 'path', and 'type'. The file paths can be absolute, or relative to the table.	File
Scenario_Predictor_Table (Optional)	(*NOT REQUIRED*) A table that maps predictor IDs to files and their types with required headers of 'id', 'path', and 'type'. The file paths can be absolute, or relative to the table.	File
Grid_the_AOI	When checked (default), overlay grid cells of a given shape type on top of the area of interest. Grid cells will be used to aggregate predictors and counts over the study area.	Boolean
Grid_Type (Optional)	Select the shape of the grid used by the model.	String
Cell_Size (Optional)	The size of the grid units measured in the projection units of the AOI. For example, UTM projections use meters.	Double

3.7.3 Population Count and Density

Human population density varies significantly by location. Areas with high population densities may have more available services, but might also exhibit greater inequalities. On the other hand, low population density may signal availability of resources such as land for agriculture and livestock, but may also suggest limited opportunities for wage employment. The Population Count and Density tool produces estimates of the number of individuals and density within a specified geographic area. Population density is reported as the number of people per km2. Population change can also be computed if population estimates for two time points are provided. This tool is quite versatile and can be used beyond its stated purpose. For example, instead of selecting population counts from different time periods, a user could select fields containing counts of male and female members of a population at a given time. The tool could then be used to show the percentage difference between the number of males and females for the given reporting unit. This tool is a modified version of the Population Density Metrics tool from the Analytical Tools Interface for Landscape Assessments (ATtILA). The ATtILA toolbox is produced by the United States Environmental Protection Agency. ***NOTE: This tool assumes that population is distributed evenly throughout each census feature polygon.***

Parameter	**Explanation**	**Data Type**
Reporting_Unit_Feature	The vector polygon dataset that defines the reporting units. **Note: The polygon must be in a projection where the units are in meters.**	Feature Layer
Reporting_Unit_ID_Field	The field in the Reporting Unit Feature layer that contains the unique ID for each reporting unit. It may be an integer or a string data type.	Field
Census_Feature	The vector polygon dataset that contains population data. **Note: The polygon must be in a projection where the units are in meters.**	Feature Layer
Population_Field	The field in the Census Feature layer that contains population data.	Field
Output_Feature_Class	The output reporting unit metrics feature class to be created.	Feature Layer
POPCHG (Optional)	Specifies whether population change over time (POPCHG) metrics will be included in the output feature class. False – No POPCHG metrics will be included. This is the default. True – POPCHG metrics will be included.	Boolean

Census_T2_Feature (Optional)	The optional vector polygon dataset (Census time 2) that contains population data for the second time point. It may be the same feature layer as Census Feature. **Note: The polygon must be in a projection where the units are in meters.**	Feature Layer
Census_T2_Field (Optional)	The field in the Census Feature layer that contains population data for the second time point.	Field

3.7.4 Nutrition Metrics

Feeding humanity's growing population relies on knowledge of nutritional needs within the world's most vulnerable locations. To provide information on minimum dietary requirements for vulnerable populations, the Nutrition Metrics tool relies on population estimates for Latin America, Africa, and Asia coupled with algorithms for estimating the minimum dietary energy requirement (MDERs) by age group within an area of interest (AOI). Age groups are separated by sex and provided in four year intervals (e.g., female 0-4, male 0-4, female 5-9, male 5-9, etc). Population estimates are provided by WorldPop (http://www.worldpop.org.uk/) for the years 2000, 2005, 2010, 2015, and 2020, and algorithms published by the Food and Agriculture Organization of the United Nations (FAO) are used to estimate calorie requirements. The Nutrition Metrics tool produces a report of calorie requirements by age group for an AOI, which the user defines by uploading a shapefile specifying the relevant location. The report includes the number of people by age group within the AOI. It also includes the total daily lower limit of energy requirement (LLER) for each age group in kilocalories per day. From this report, more informed decisions can be made to address questions such as:

- How have population and calorie requirements changed within this location during the past twenty years?
- Is the agricultural production potential within this area sufficient to meet the nutritional needs of the population?
- Do nutritional requirements differ between nearby areas of interest?

For an explanation of the LLER calculation process, refer to the tool's usage information in ArcCatalog. To do this, choose the Nutrition Metrics tool in ArcCatalog by locating it in the catalog tree, then select the Description tab and scroll to the Usage section).

NOTE: This tool relies on population rasters developed by WorldPop (www.worldpop.org). These rasters are downloaded with the Telecoupling Toolbox sample data and are separated by continent.

Parameter	Explanation	Data Type
Area_of_Interest	Shapefile indicating the Area of Interest (AOI). Smaller AOIs will have shorter processing times. If the uploaded shapefile contains more than one feature, all features will be aggregated together. It is therefore recommended that the uploaded shapefile contain only one feature. NOTE: This tool will only run with AOIs in Africa, Asia, or South America / Central America / Caribbean.	Shapefile
Continent_(continent_of_area_or_interest)	Folder of rasters (separated by continent) representing information on population and age structures. Select the folder that is consistent with the continent on which the AOI is located. NOTE: Africa – includes the entire continent from North Africa to the southern extent of sub-Saharan Africa Asia – includes mainland countries between Turkey and China, but excluding all of Russia. Countries off the coast of mainland Asia, including Japan, Taiwan, and those constituting South-East Asia are also included. South America / Central America / Caribbean – includes the entire continent of South America, countries extending southward from Mexico to Panama, as well as the countries of the Caribbean.	Folder
Year_of_Interest	The year for which population and LLER estimates will be produced. Options include 2000, 2005, 2010, 2015, and 2020. NOTE: The Africa mosaic dataset does not include a record for the Male 65+ age group in the Year 2020. Thus, ‘999999999’ is returned for both the population and LLER in this age group.	String
Average_Male_Stature_in_cm_(height)	Average adult male height in the AOI. This must be provided in centimeters (cm). If the average adult male height within the AOI is not known, it is recommended	Double

	that the average male height within the country where the AOI is located is used.	
Average_Female_Stature_in_cm_(height)	Average adult female height in the AOI. This must be provided in centimeters (cm). If the average adult female height within the AOI is not known, it is recommended that the average female height within the country where the AOI is located is used.	Double

3.8 Future Development

Updated and new versions of the telecoupling toolbox are released periodically where new script tools are added or modifications are made to existing tools to fix errors or improve their functionalities. The source code, sample data, and tool documentation are publicly available on Github[6], a common web-based hosting service for projects that use revision control systems. Although the present toolbox was developed to work within ESRI's ArcGIS software environment, thus limited to the Microsoft Windows platform, we also offer a web-based application, the **GeoApp**[7], with most of the tools available for the desktop version for ArcGIS. The major advantage of our web-based application is to free up users from the hassle of installing several required software and libraries, while enhancing the responsive and interactive components that are typical of modern web applications. A note on the telecoupling toolbox versioning: integer changes will reflect major changes, e.g. from 1.2 to 2.0. Increments in the digit after the primary decimal indicates major new features (e.g, the addition of a new tool) or major revisions. The third decimal reflects minor feature revisions or bug fixes with no new functionality.

3.9 References

Kastner, T., Erb, K.-H., and Nonhebel, S. 2011. International wood trade and forest change: A global analysis. *Global Environmental Change* 21(3), pp. 947-956. http://dx.doi.org/10.1016/j.gloenvcha.2011.05.003

Konar, M., C. Dalin, S. Suweis, N. Hanasaki, A. Rinaldo, and I. Rodriguez-Iturbe. 2011. Water for food: the global virtual water trade network. *Water Resources Research* 47:W05520. http://dx.doi.org/10.1029/2010WR010307

[6] https://msu-csis.github.io/telecoupling-toolbox/
[7] https://telecoupling.msu.edu/geo-app/

Liu, J., Dietz, T., Carpenter, S.R., Folke, C., Alberti, M., Redman, C.L., Schneider, S.H., Ostrom, E., Pell, A.N., Lubchenco, J., Taylor, W.W., Ouyang, Z., Deadman, P., Kratz, T., and William Provencher. 2007. Coupled Human and Natural Systems. *AMBIO: A Journal of the Human Environment* 36(8), pp. 639-649. http://dx.doi.org/10.1579/0044-7447(2007)36[639:CHANS]2.0.CO;2

Liu, J., Hull, V., Batistella, M., DeFries, R., Dietz, T., Fu, F., Hertel, T.W., Izaurralde, R.C., Lambin, E.F., Li, S., Martinelli, L.A., McConnell, W.J., Moran, E.F., Naylor, R., Ouyang, Z., Polenske, K.R., Reenberg, A., de Miranda, R.G., Simmons, C.S., Verburg, P.H., Vitousek, P.M., Zhang, F., and Zhu, C. 2013. Framing sustainability in a telecoupled world. *Ecology and Society* 18, art26. 10.5751/ES-05873-180226

Liu, J., Mooney, H., Hull, V., Davis, S.J., Gaskell, J., Hertel, T., Lubchenco, J., Seto, K.C., Gleick, P., Kremen, C., and Li, S. 2015. Systems integration for global sustainability. *Science* 347(6225). 10.1126/science.1258832

Liu, J., Yang, W., and Li, S. 2016. Framing ecosystem services in the telecoupled Anthropocene. *Frontiers in Ecology and the Environment* 14(1), pp. 27-36. 10.1002/16-0188.1.

Reid, W.V., Chen, D., Goldfarb, L., Hackmann, H., Lee, Y.T., Mokhele, K., Ostrom, E., Raivio, K., Rockström, J., Schellnhuber, H. J., and Whyte, A. 2010. Earth System Science for Global Sustainability: Grand Challenges. *Science* 330(6006), pp. 916-917. 10.1126/science.1196263

United Nations Secretary-General's High-level Panel on Global Sustainability. 2012. *Resilient people, resilient planet: a future worth choosing*. United Nations, New York, New York, USA.

3.10 Acknowledgements

The development of the Telecoupling Toolbox has been supported by grants from the National Science Foundation, Michigan State University, and Michigan AgBioResearch.

3.11 License & Copyright

4 TUTORIALS

The following tutorials will guide you through examples of potential application of each script tool using the sample data provided with the toolbox (*SampleData_ArcGIS_v1.7.3b.zip*). All ESRI basemaps are by default using a Web Mercator coordinate system, typically found in most online web applications (e.g. Google Maps, Bing Maps). If you decide to use your own background administrative layer that is in a different coordinate system, ArcGIS will automatically re-project any additional layers (including output from geoprocessing tools) on-the-fly to the coordinate system of the first layer in your map.

NOTE: Although ArcGIS projects on-the-fly to avoid projection mismatch, we always encourage the user to have all layers in the same coordinate system appropriate for the scale of analysis and case study.

NOTE: some basic level of proficiency in ArcGIS Desktop is necessary to better follow the tutorial examples shown in Chapter 4.

4.1 SAMPLE DATASET

The sample dataset provided with the telecoupling toolbox has data on six different types of telecoupling: wildlife transfer, tourism (eco-tourism), agricultural trade, industrial trade, conservation subsidies, and information dissemination. Inside the *SampleData_ArcGIS_v1.7.3b* folder, you will also find several other subfolders divided by topic which will be used in some of the following tutorials (Fig. 3).

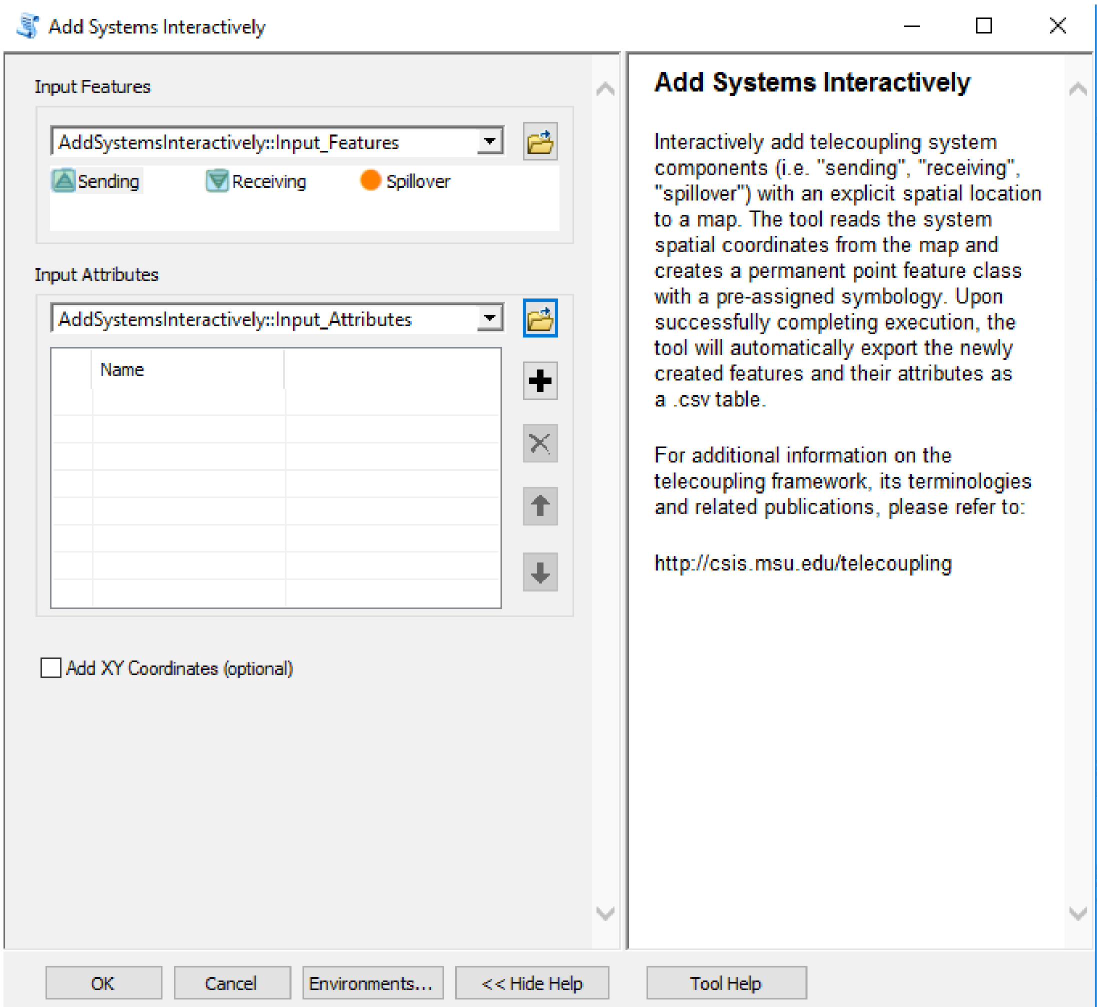
DISCLOSURE ON DATA USAGE: the sample datasets were either partially modified, or their format changed in some case from the original datasets to protect privacy and sensitive data when deemed appropriate. Use the sample datasets for learning purposes only and under no circumstances data should be used to inform any policy. Please refer to the Telecoupling Toolbox license agreement for more details.

Name	Date modified	Type
Agents	10/18/2017 9:34 AM	File folder
Causes	10/2/2017 9:34 AM	File folder
Environmental Analysis	10/2/2017 9:36 AM	File folder
Flows	10/2/2017 9:34 AM	File folder
Socioeconomic Analysis	10/2/2017 1:49 PM	File folder
Systems	10/18/2017 9:32 AM	File folder

Figure 3. File structure of the SampleData_ArcGIS_v1.7.3b folder provided with the telecoupling toolbox.

4.2 SYSTEMS TOOLSET

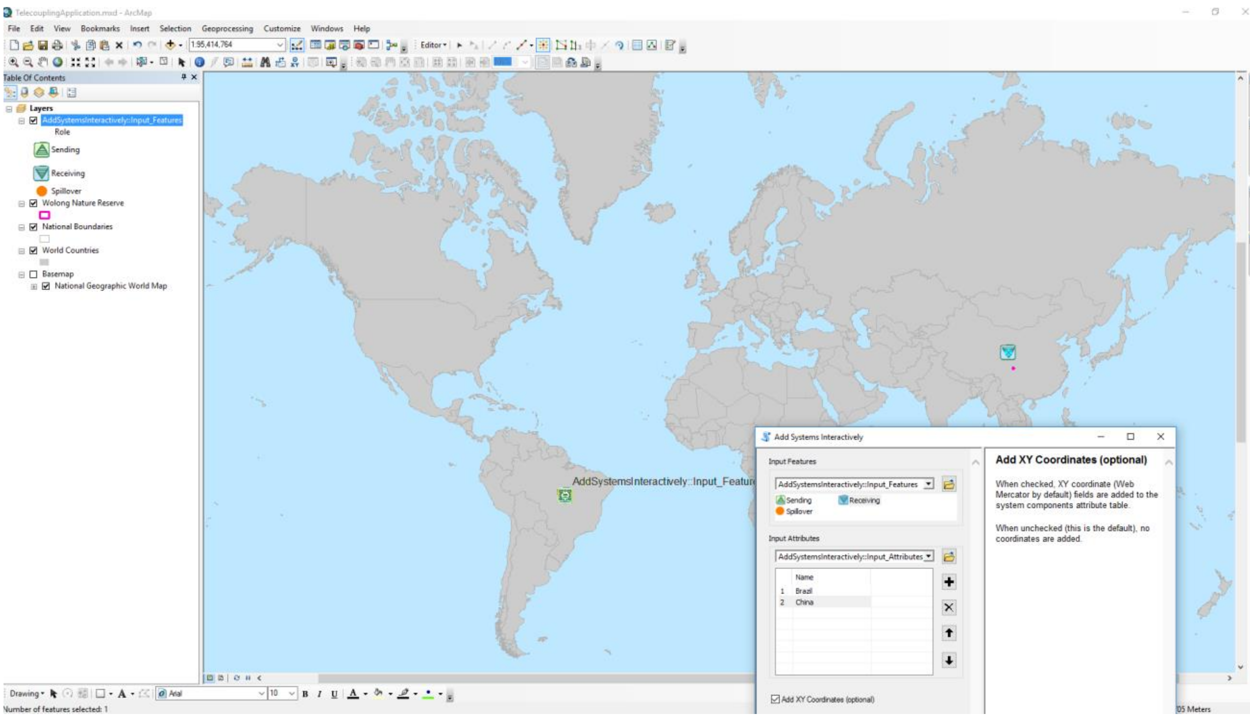
4.2.1 Add Systems Interactively



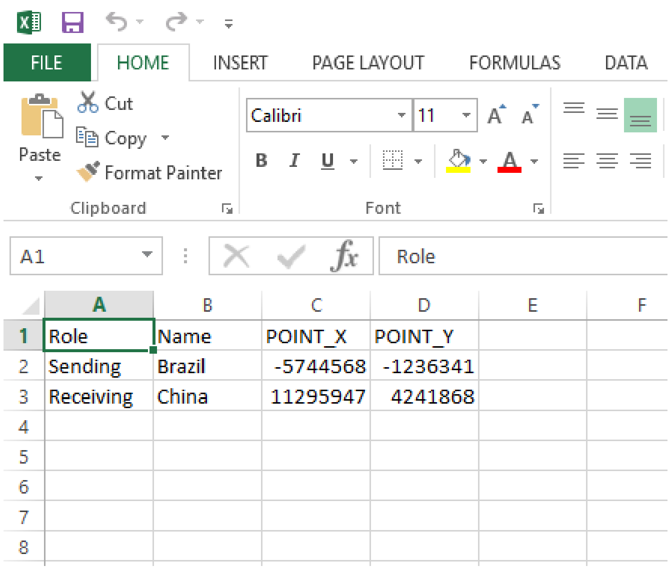
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

The first parameter asks you to select a category for the system you would like to add to the map (sending, receiving, spillover) with its own symbology. Once you select the desired category, move the mouse cursor over the map and you should see the chosen symbol ready to be placed. Make sure to zoom in/out to the extent that best captures your system (e.g. census block, city, province, region, country). Click on the map to mark your system down and type a name in the second parameter of the user interface that best describes your system (e.g. country name, city name, etc). In order to add a name to the second parameter, you need to click on the "+" sign and then start typing in the corresponding line. Repeat the above procedure for all other systems in your case study. Typically, you should have a *minimum of two points* (systems) on the map, one sending and one receiving, or it will not make sense to have a flow between the two. Once you click OK to run the tool, ArcGIS will create a permanent output point feature class and save an output table to your workspace directory with a list of all the systems and names you added to the map. If you wish to add spatial coordinates (in Web Mercator by default) to the attribute table, make sure to check the "Add XY Coordinates" box found in the tool interface.

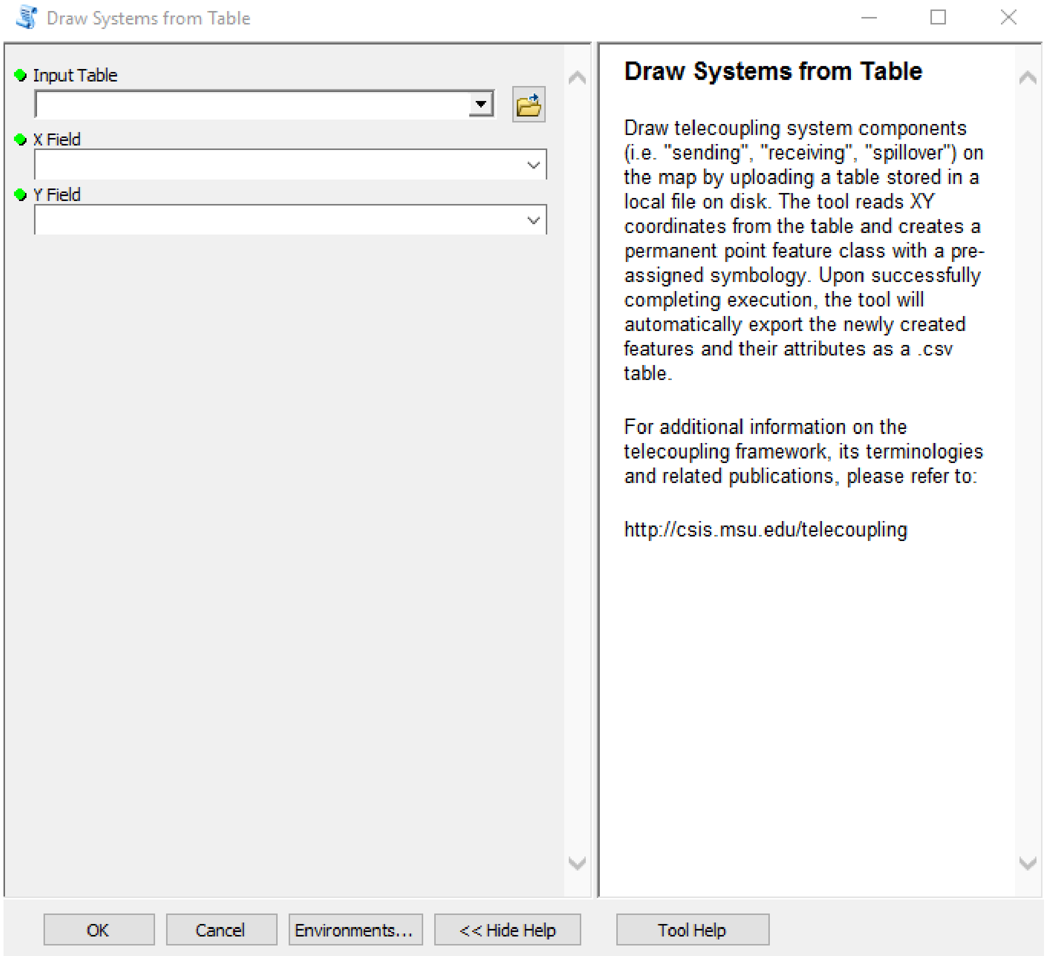
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



The output .csv table created by the tool in the scratch workspace directory should look something similar to the following figure:


FILE
HOME
INSERT
PAGE LAYOUT
FORMULAS
DATA
Cut
Copy
Format Painter
Paste
Clipboard
Calibri
11
Font
A1
Role
A B C D E F
Role Name POINT_X POINT_Y
Sending Brazil -5744568 -1236341
Receiving China 11295947 4241868

4.2.2 Draw Systems from Table

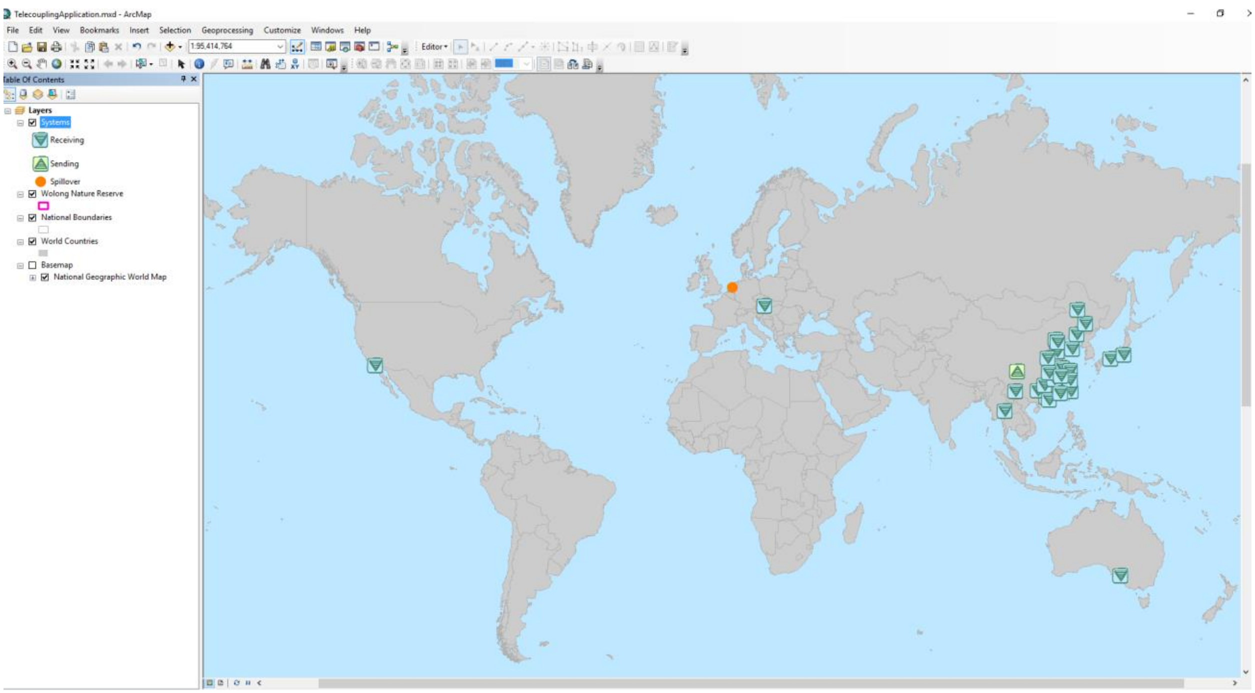


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

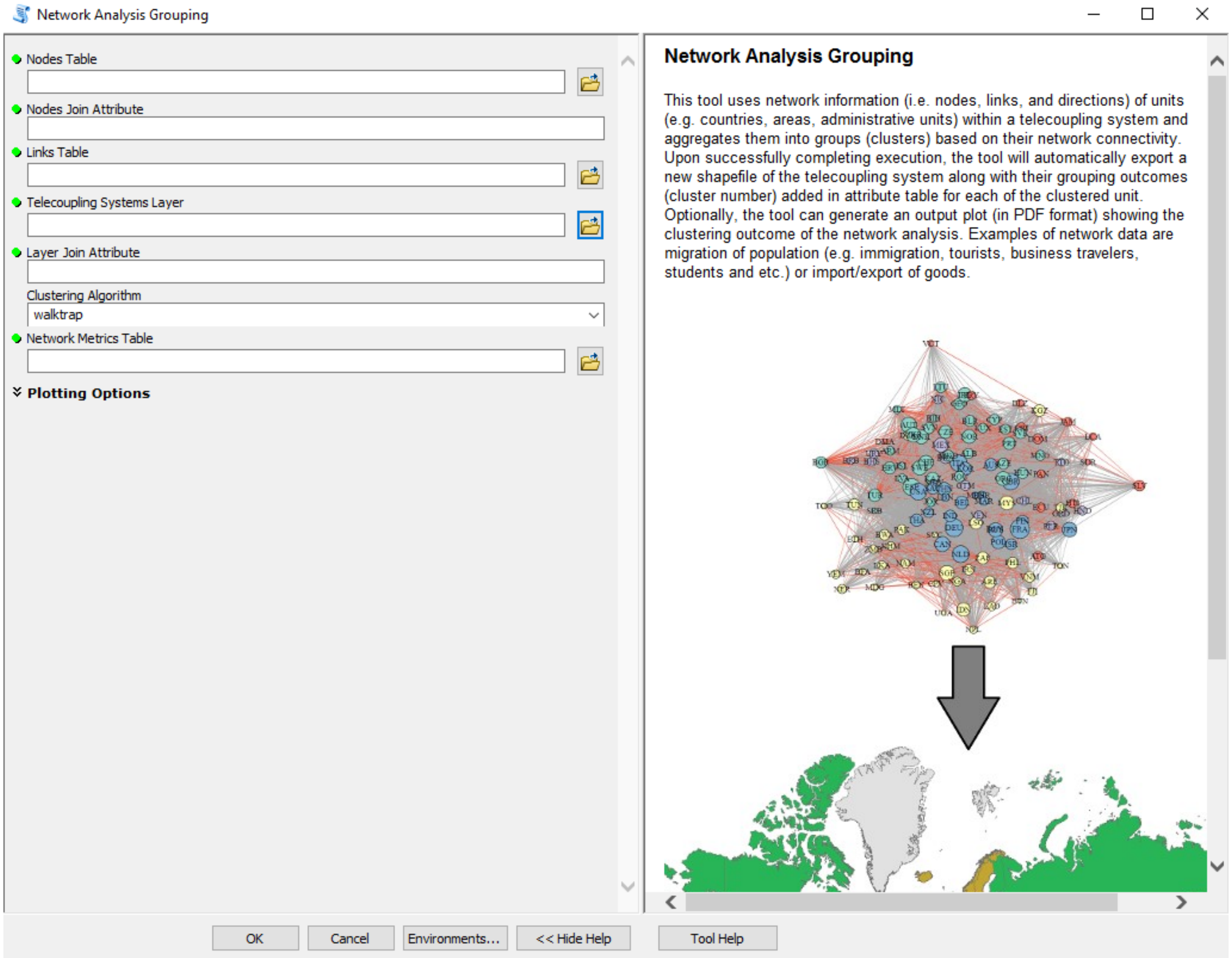
The first parameter asks you to select a table on disk containing records on your telecoupling systems. The table must at least have a column specifying the name of the system (e.g. USA, Chicago, Michigan), the Role of the system (i.e. Sending, Receiving, Spillover), and the spatial coordinates of it. Coordinates can be in any coordinate system, since the tool will automatically re-project them to Web Mercator (by default). The second and third parameter of the tool ask the user to specify which fields (columns) in the chosen table on disk correspond to the X and Y coordinates. Typically, you should have a *minimum of two points* (systems) inside the table, one sending and one receiving, or it will not make sense to have a flow between the two. Once you click OK to run the tool, ArcGIS will create a permanent output point feature class and add it to the map.

If you are using the sample data provided with the toolbox, you can use any of the .csv tables ending with the word "_Systems" (e.g. */SampleData_ArcGIS_v1.7.3b/Systems/Draw Systems from Table/wildlife_Systems.csv*).

<u>*NOTE*</u>*: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.*



4.2.3 Network Analysis Grouping



Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

This tool can be used for network grouping of the nodes under analysis in a telecoupling system based on given network information including nodes, their connected components and in- and out- degrees of the nodes. For example, use the sample dataset within this toolbox for Network Analysis Grouping tool, including nodes.csv, links.csv, and World_countries_2002.shp.

If you are using the sample data provided with the toolbox, browse to the nodes.csv table (*SampleData_ArcGIS_v1.7.3b/Systems/Network Analysis Grouping/nodes.csv*).

The input nodes table should contain telecoupling systems in each record (nodes in the network analysis terminology) and have a unique identifier code tied to it. Make sure your table also has two other columns with the number of outgoing tourists and incoming tourists (in our sample data we took the natural log of this number to avoid issues with node size in the output results).

Specify that you want the column attribute called “CODE” as the join attribute for the nodes table.

For the input links table, browse to the links.csv table in our sample data (*SampleData_ArcGIS_v1.7.3b/Systems/Network Analysis Grouping/links.csv*). The table contains a list of all the connections between sending and receiving telecoupling systems (countries in this case) that are to be considered as "links" in the network grouping analysis. This table must include a unique identifier code for both the sending and receiving systems, as well as an attribute specifying the outgoing flow.

NOTE 1: make sure the codes used for the sending and receiving systems match exactly those used in the nodes input table.

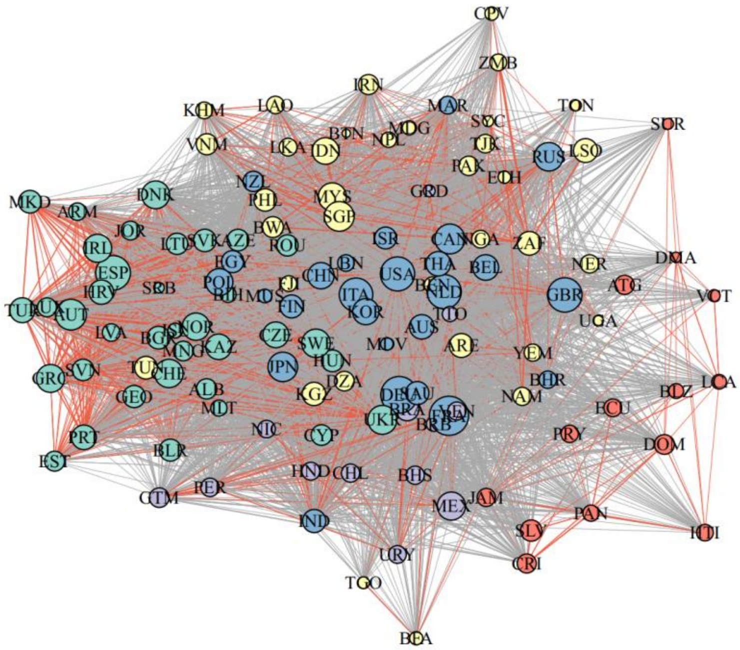
NOTE 2: make sure the outgoing flow attribute contains the same values used in the nodes input table. Therefore, if you used natural logarithm in the nodes input table, make sure to use the same here.

The sample data links.csv shows countries in the tourism network and the amount of tourists heading out from a given country (sender) into a different one (receiver).

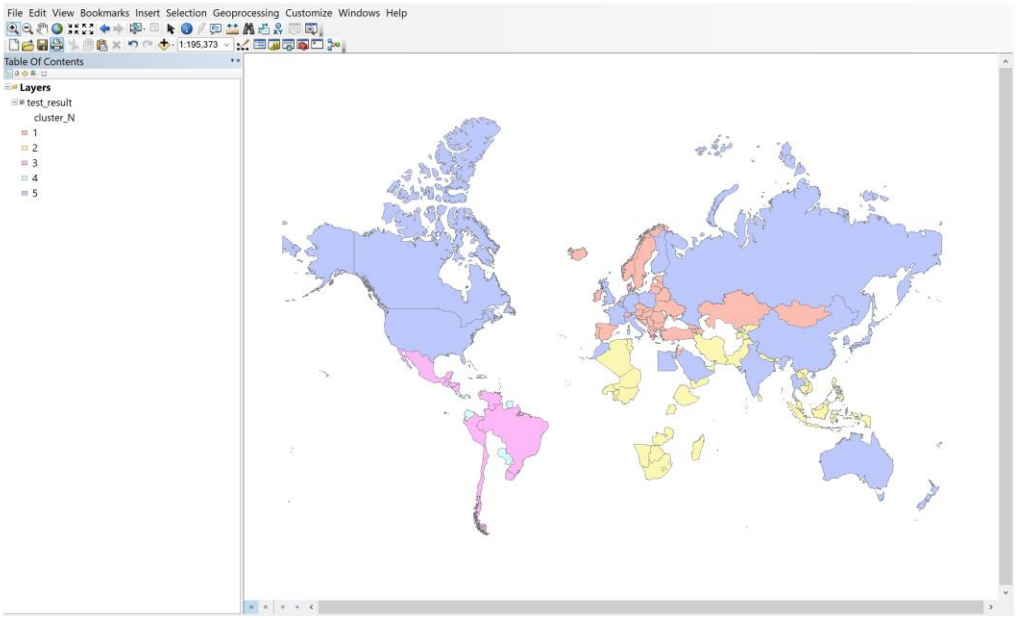
The sample shapefile is used as the input telecoupling systems feature layer (*SampleData_ArcGIS_v1.7.3b/Systems/Network Analysis Grouping/World_countries_2002.shp*). The shapefile contains 211 records of countries with “ISO_3_CODE” as one of its attribute. Select this attribute as the Layer Join Attribute that will be matched against the input nodes join attribute to attach a cluster number when the tool is done running.

As a clustering algorithm, you can choose *walktrap* or *spin_glass* (you can choose the default algorithm walktrap here). A pdf file with a plot of the outcome network grouping will be exported if a local directory and file name is offered for the output report file. The tool does not export this report if you leave this output option blank. As output shapefile, browse to a local directory and give it whatever name you prefer. Otherwise, accept the default that ArcGIS fills out for you. We will leave all the plotting options as default, but feel free to modify any or all of them to see how it affects the output network grouping. Simply expand the plotting options and change the values if you wish to.

Click OK to run the tool. Once completed, a pdf file and a .csv file and a shapefile will be exported. The image in the pdf file delineates the clusters in this tourism network. Countries in different clusters are shown in different colors. Countries without information in this system are not shown in the plot, and they are also removed for the output feature class. Degree, closeness and betweenness are calculated for each node in the network and put in the table. A new attribute called cluster_N can be found in the attribute table of output feature class shows which clusters are the countries belong to according to the network analysis. You can easily color the map for different clusters or pick specific clusters to show in this map within ArcGIS.

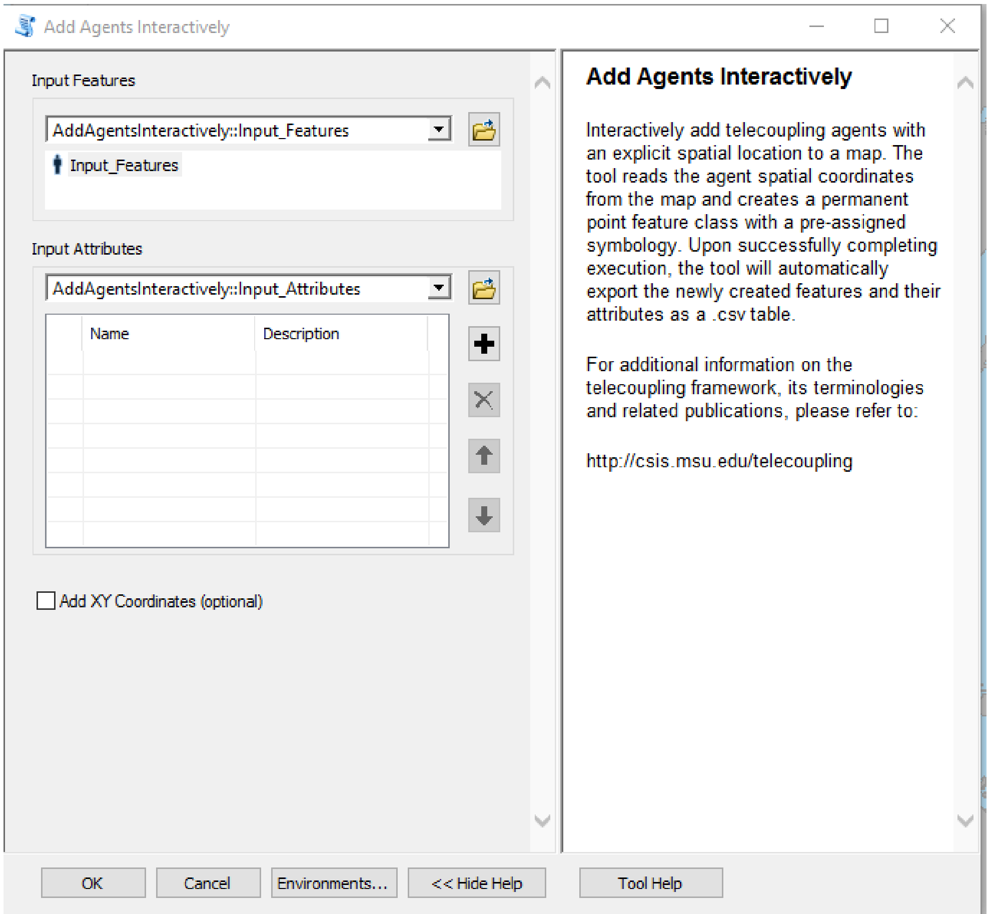


	degree	closeness	betweenness
ALB	109	0.154135338	20.63002545
ARE	65	0.159326425	9.322103636
ARM	94	0.151851852	14.7098935
ATG	30	0.151477833	0.56159059
AUS	215	0.168032787	591.1386095
AUT	138	0.171309192	55.75180001
AZE	107	0.152605459	20.0273266
BEL	222	0.173239437	707.1005371
BEN	90	0.152605459	16.84494106
BFA	37	0.148910412	2.790395562
BGR	120	0.156091371	31.51261066
BHR	153	0.136971047	67.38083755
BHS	145	0.151851852	76.0193102
BIH	78	0.151851852	5.978454459
BLR	108	0.150366748	13.38024193
BLZ	49	0.154522613	2.871817264
BRA	106	0.168493151	50.95204933
BRB	137	0.152985075	82.70895544


File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
1:195,373
Table Of Contents
Layers
test_result
cluster_N
1
2
3
4
5

4.3 AGENTS TOOLSET

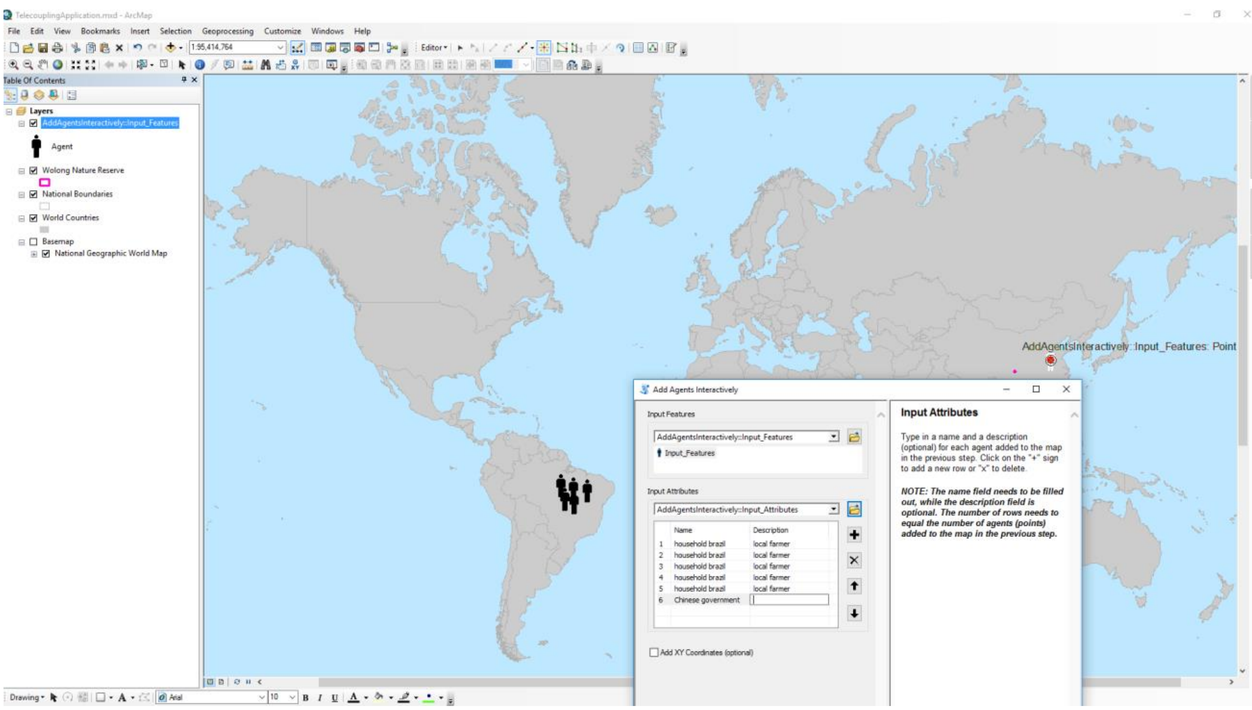
4.3.1 Add Agents Interactively



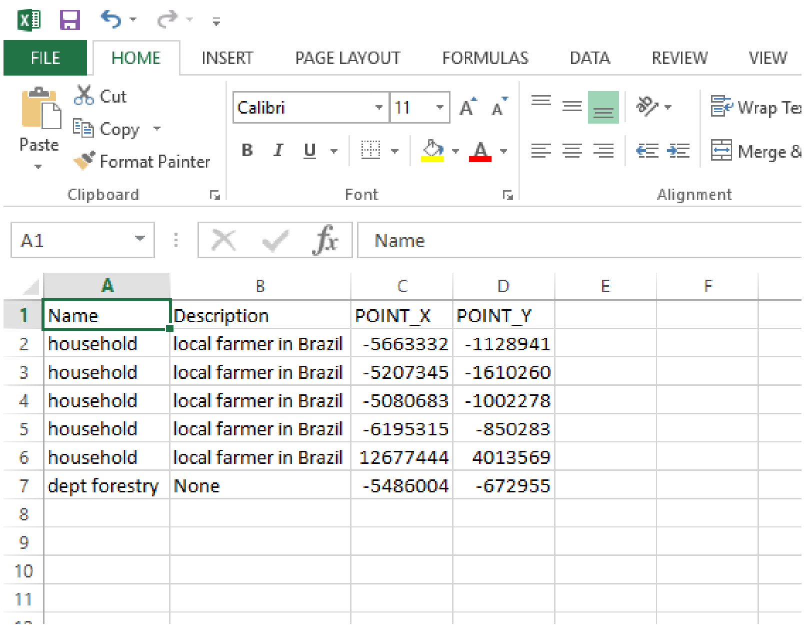
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

The first parameter asks you to select agents (e.g. household, organization, governmental departments) you would like to add to the map. Move the mouse cursor over the map and you should see an agent symbol ready to be placed. Make sure to zoom in/out to the extent that best captures your agent size or location. Click on the map to mark your agents down and type a name and (optionally) a description in the second parameter of the tool interface. You can pick a name that best characterizes your agents and add a description of it if deemed necessary. In order to add a new record to the second tool parameter and start typing name/description, you need to click on the "+" sign. Repeat the above procedure for all other agents in your case study and make sure the number of agents on the map corresponds to the number of rows in the table of the tool interface. Once you click OK to run the tool, ArcGIS will create a permanent output point feature class and save an output table to your workspace directory with a list of all the agents names and description you added to the map. If you wish to add spatial coordinates (in Web Mercator by default) to the attribute table, make sure to check the “Add XY Coordinates” box found in the tool interface.

<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the “Environments…” button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

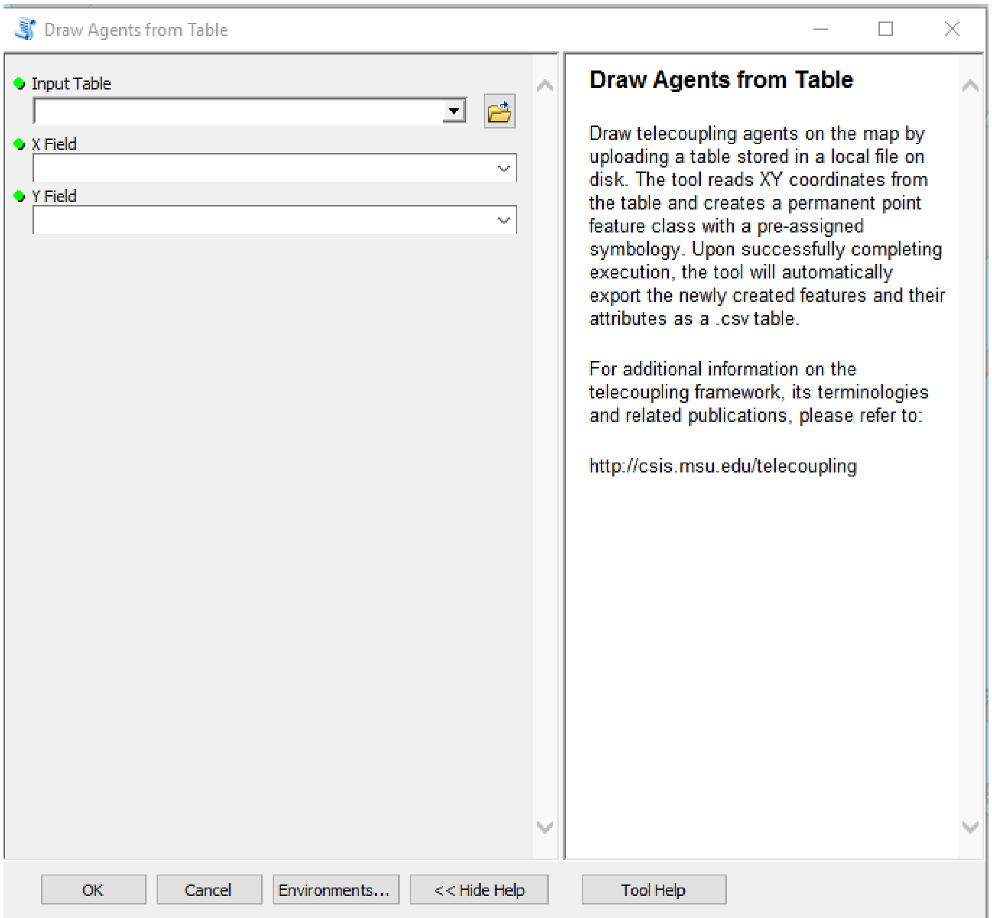


The output .csv table created by the tool in the scratch workspace directory should look something similar to the following figure:



	A	B	C	D	E	F
1	Name	Description	POINT_X	POINT_Y		
2	household	local farmer in Brazil	-5663332	-1128941		
3	household	local farmer in Brazil	-5207345	-1610260		
4	household	local farmer in Brazil	-5080683	-1002278		
5	household	local farmer in Brazil	-6195315	-850283		
6	household	local farmer in Brazil	12677444	4013569		
7	dept forestry	None	-5486004	-672955		
8						
9						
10						
11						

4.3.2 Draw Agents from Table

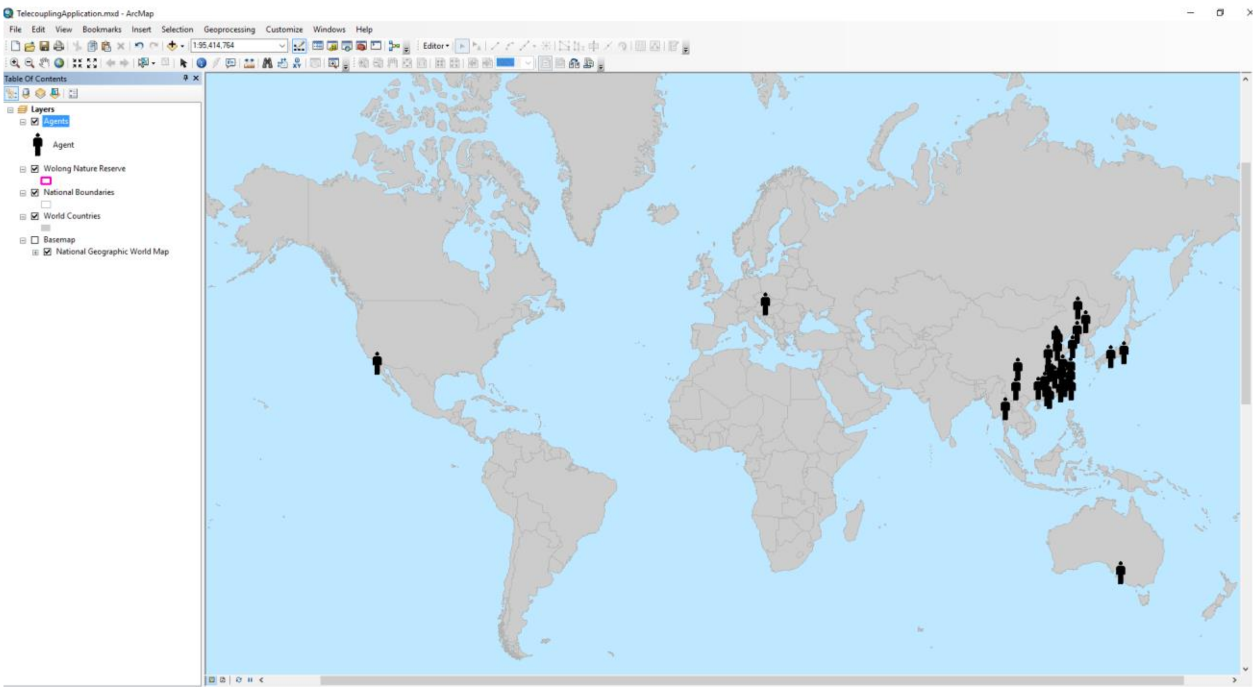


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

The first parameter asks you to select a table on disk containing records on your telecoupling agents. The table must at least have a column specifying the name of the agent and a description (this field can be left empty), and the spatial coordinates of it. Coordinates can be in any coordinate system, since the tool will automatically re-project them to Web Mercator (by default). The second and third parameter of the tool ask the user to specify which fields (columns) in the chosen table on disk correspond to the X and Y coordinates. Once you click OK to run the tool, ArcGIS will create a permanent output point feature class and add it to the map.

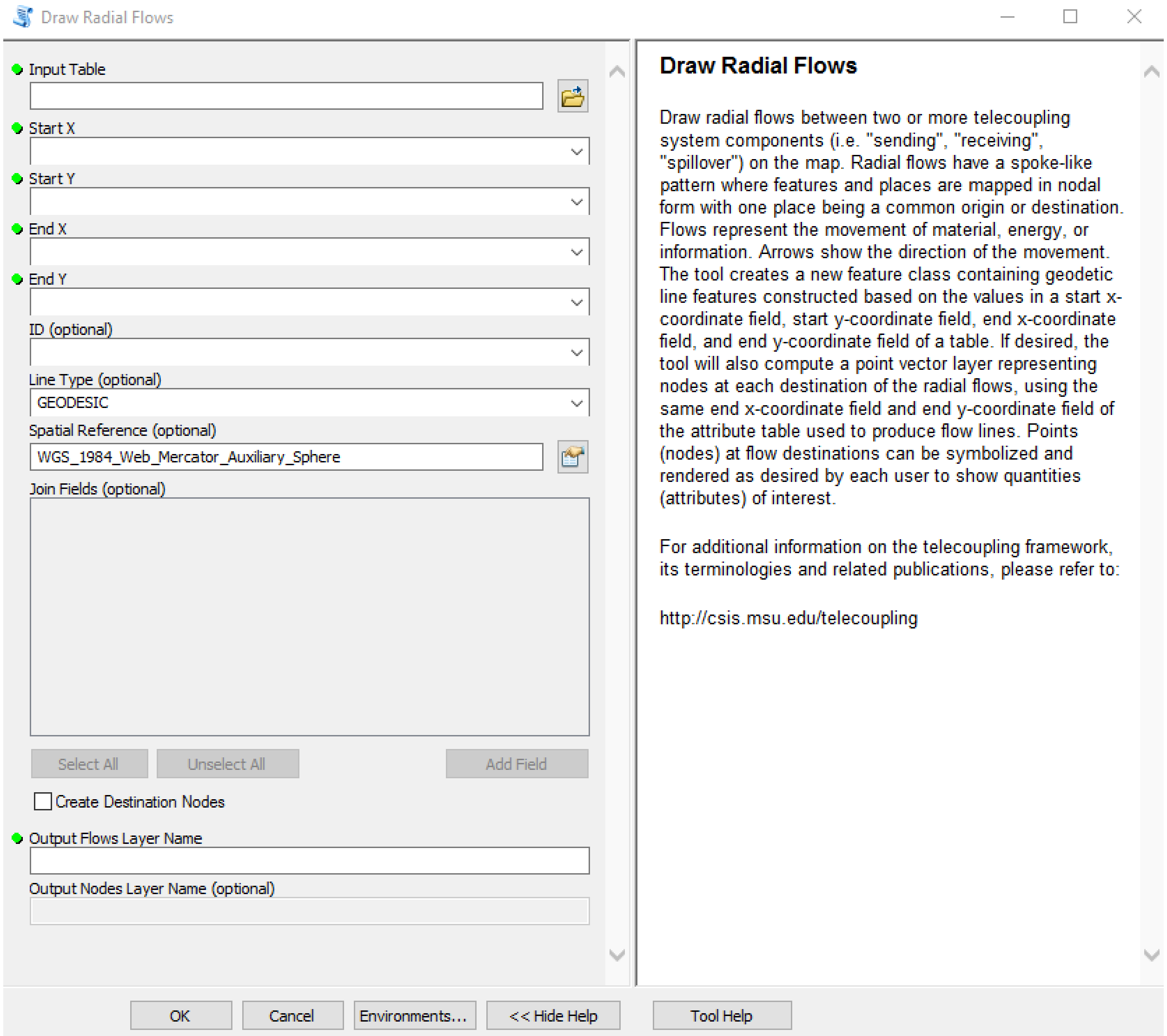
If you are using the sample data provided with the toolbox, you can use any of the .csv tables ending with the word "_Agents" (e.g. */SampleData_ArcGIS_v1.7.3b/Agents/Draw Agents from Table/wildlife_Agents.csv*).

<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



4.4 FLOWS TOOLSET

4.4.1 Draw Radial Flows



Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

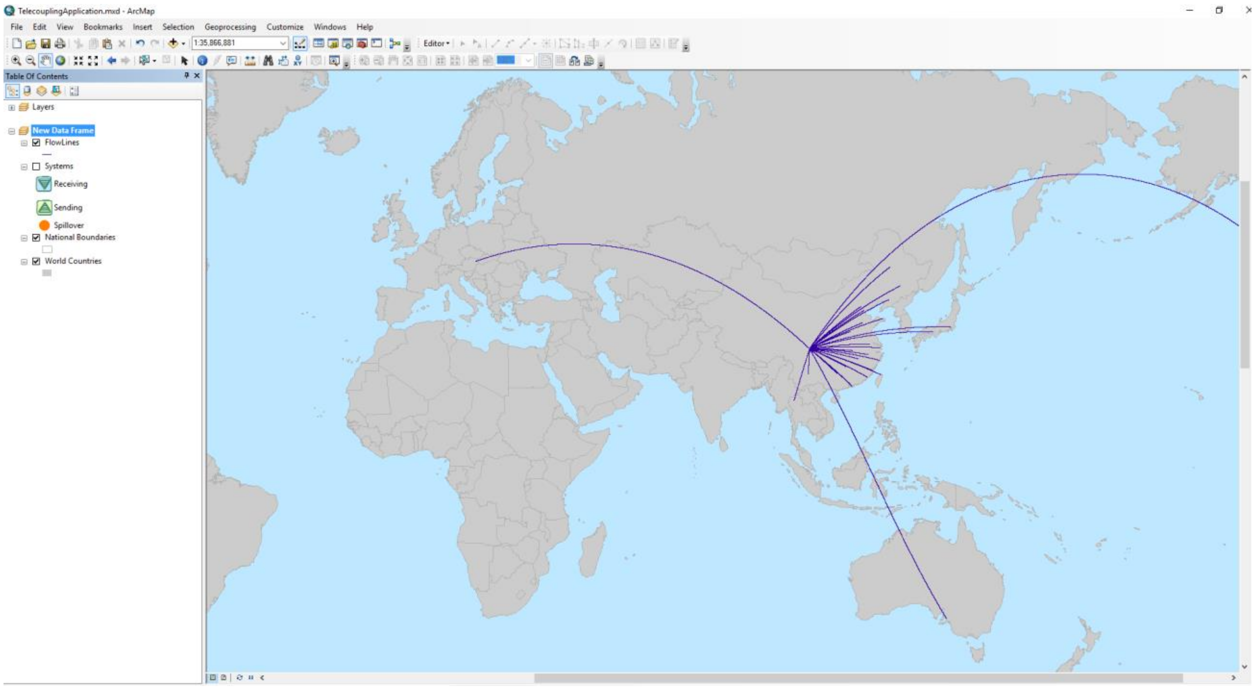
The first parameter asks you to select a table on disk containing records on start and end points of the flow. The next four parameters need the field name (column) of the input table containing

X and Y coordinates of start and end points, respectively. The ID parameter is optional and should only be used in case you have additional columns in the input table with data you would like to use to symbolize the flow lines (e.g. quantitative variables such as the amount of traded goods, or fees paid for the transaction, or a simple categorical label describing the flow). Ideally, you should always have some data on the amount of flow between two points, whether it is material or any other transferred quantity. If you have these additional columns in the input table, as ID parameter select a numerical sequential indicator found in the table that will be used to join the additional columns to the output flow table created by the tool. The Line Type parameter ("GEODESIC" by default) lets you choose what type of flow lines to draw between points. Please refer to the tool description in Section 3 for more details on the options. The Spatial Reference parameter (Web Mercator by default) should correspond to whatever coordinate systems your start and end points in the input table are. For example, if your coordinates in the table are in Geographic Coordinate System WGS 1984, then make sure to change this parameter in the tool interface to match them. ***NOTE*: If you forget this step and run the tool, you may not see any flow lines in output because of a projection mismatch**. The last optional parameter, called "Join Field(s)", lets you choose which additional columns from the input table you want be joined to the output flow attribute table.

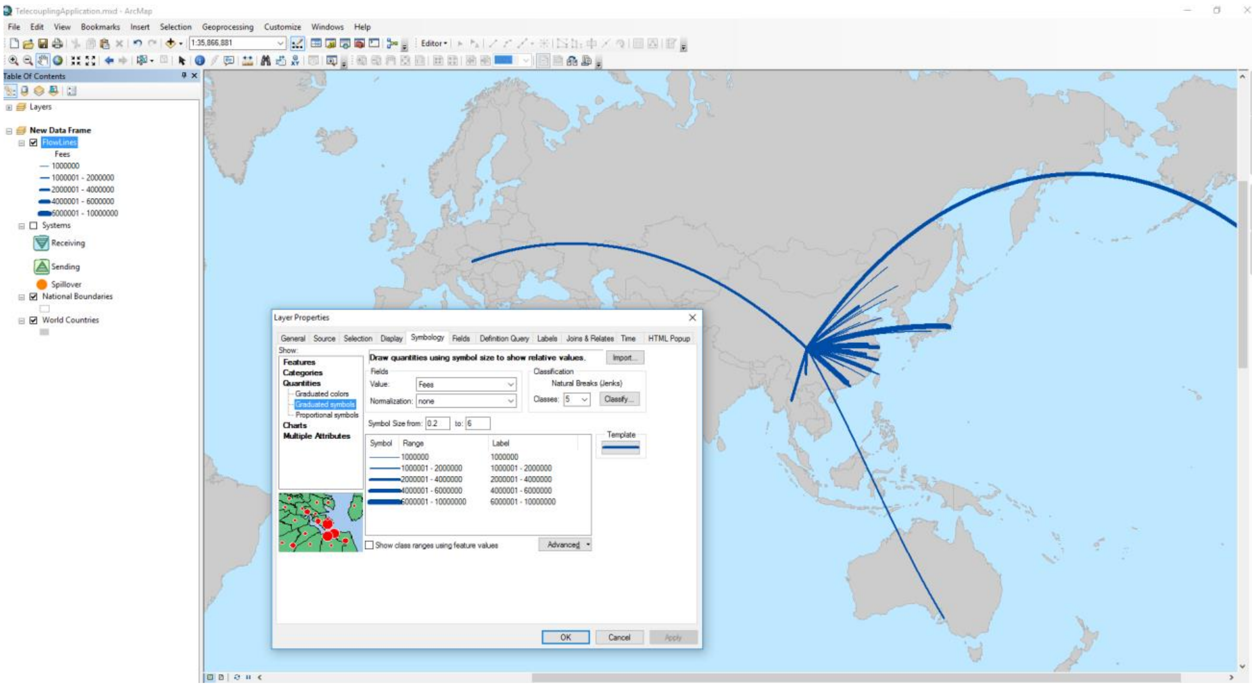
Once you click OK to run the tool, ArcGIS will create a permanent output line feature class and add it to the map.

If you are using the sample data provided with the toolbox, you can use any of the .csv tables ending with the word "_Flows" (e.g. */SampleData_ArcGIS_v1.7.3b/Flows/Draw Radial Flows/wildlife_Flows.csv*). Use the column called "FID" as ID parameter, select WGS 1984 as a spatial reference for the spatial coordinates found in the sample table, and choose one or more additional columns that show the quantity of the flow.

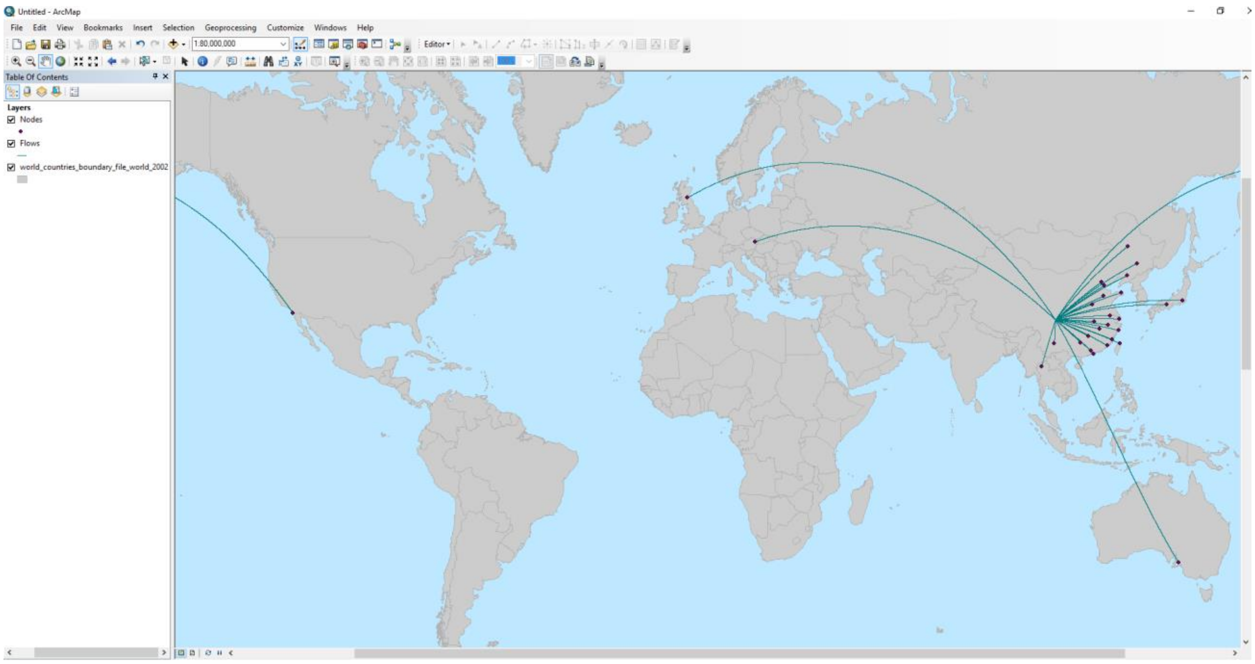
NOTE: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



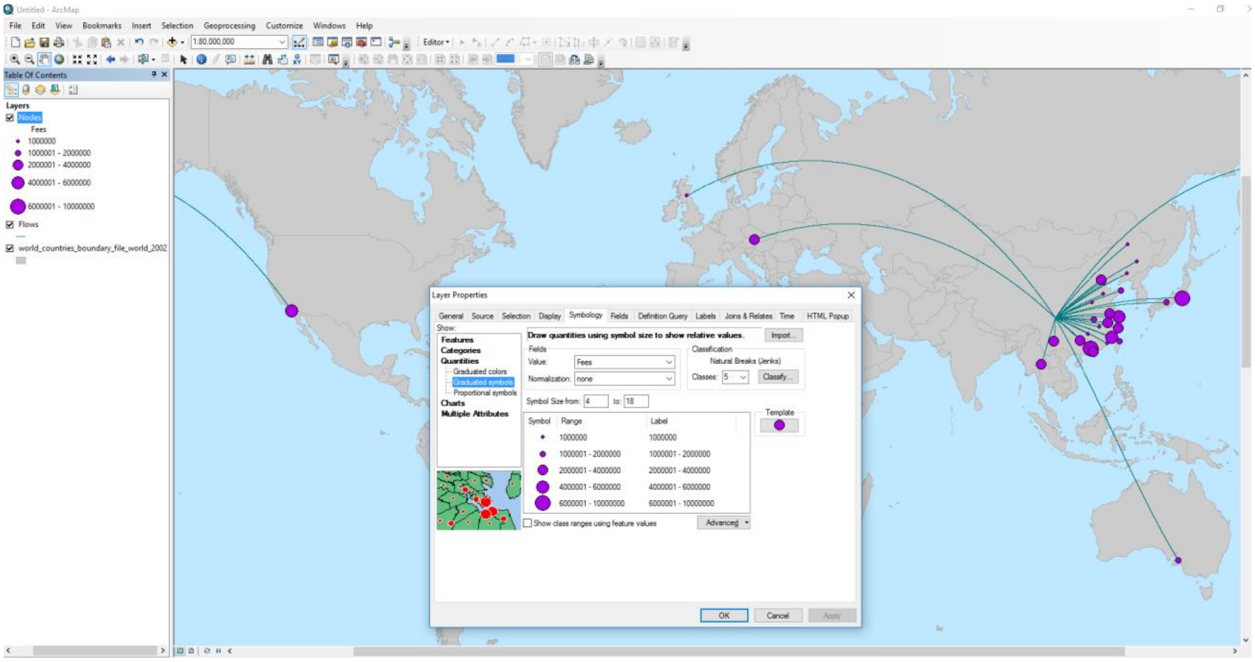
Once the output flow feature class has been created, the lines will all have the same color and thickness. If you have selected additional variables to represent the flow amount, you can use one of them to improve the visualization of the lines. To do so, right-click on the output flow layer found in the table of content (left-side of the map typically) and select the tab called "symbology". Then, you can select "quantities", and something like "graduated symbols" to show the line thickness proportional to the chosen variable of interest (e.g. flow amount, fees). Choose how many bins you want to group your values and a desired color. The next figure shows you an example using the "*wildlife_Flows.csv*" sample table, where flows have been symbolized proportionally to the amount of fees paid for the wildlife transfer transaction.



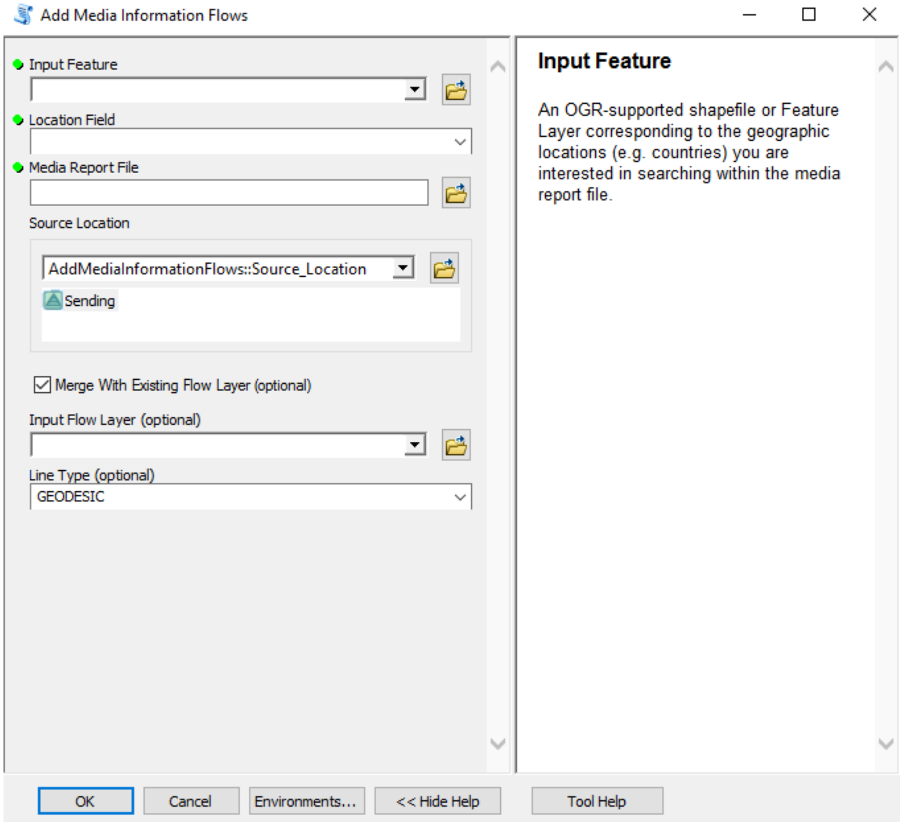
If in addition you desire to create a point layer with nodes at flow destinations, repeat all the steps above, making sure you check the box next to “Create Destination Nodes”. Give a name to the layers and click OK. ArcGIS will create a permanent output point feature class representing destination points of the flows between systems, and add both to the map.



Once the output nodes feature class has been created, the points will all have the same color and size. If you have selected additional variables to represent the flow amount, you can use one of them to improve the visualization of the points. To do so, right-click on the output nodes layer found in the table of content (left-side of the map typically) and select the tab called "symbology". Then, you can select "quantities", and something like "graduated symbols" to show the point size proportional to the chosen variable of interest (e.g. flow amount, fees). Choose how many bins you want to group your values and a desired color. The next figure shows you an example using the "*wildlife_Flows.csv*" sample table, where nodes have been symbolized proportionally to the amount of fees paid for the wildlife transfer transaction.



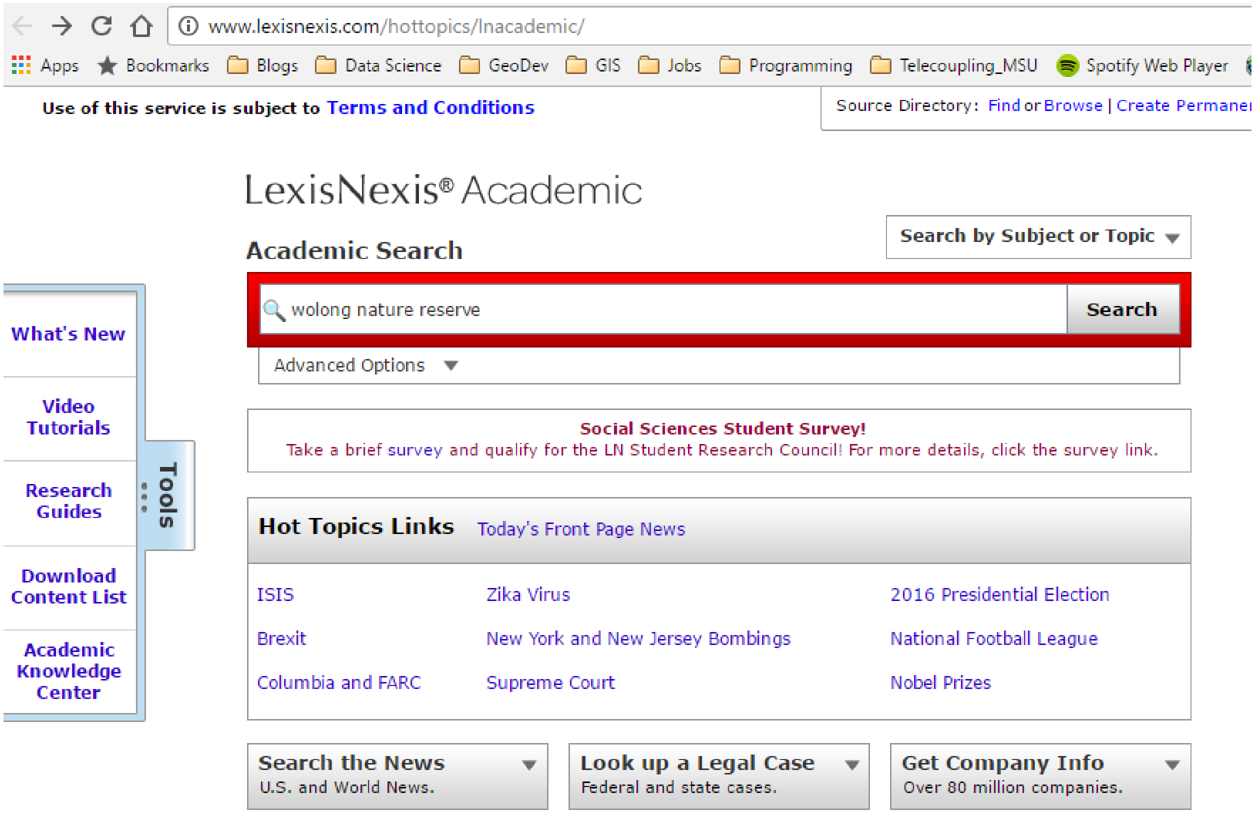
4.4.2 Add Media Information Flows



Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

This tool should be used to map flows of information that disseminate from a source location to other locations globally. For example, the Academic LexisNexis website is often used to search word terms, e.g. "Wolong Nature Reserve", in a vast online database made of academic and scientific articles, books, news media, and magazines just to name some. When using the media information flows tools, the user decides whether to map flows using only the online report generated from LexisNexis, or merge the media flow to an existing flow layer previously generated using the radial flows tools (i.e. Draw Radial Flows, Draw Radial Flows and Nodes). In this tutorial, you will first use LexisNexis to search word terms and generate an HTML report file out of it. If you wish to skip this step and use a pre-generated HTML report, please use the "*./SampleData_ArcGIS_v1.7.3b/Flows/Add Media Flows/LexisNexis.html*" file provided with the sample dataset.

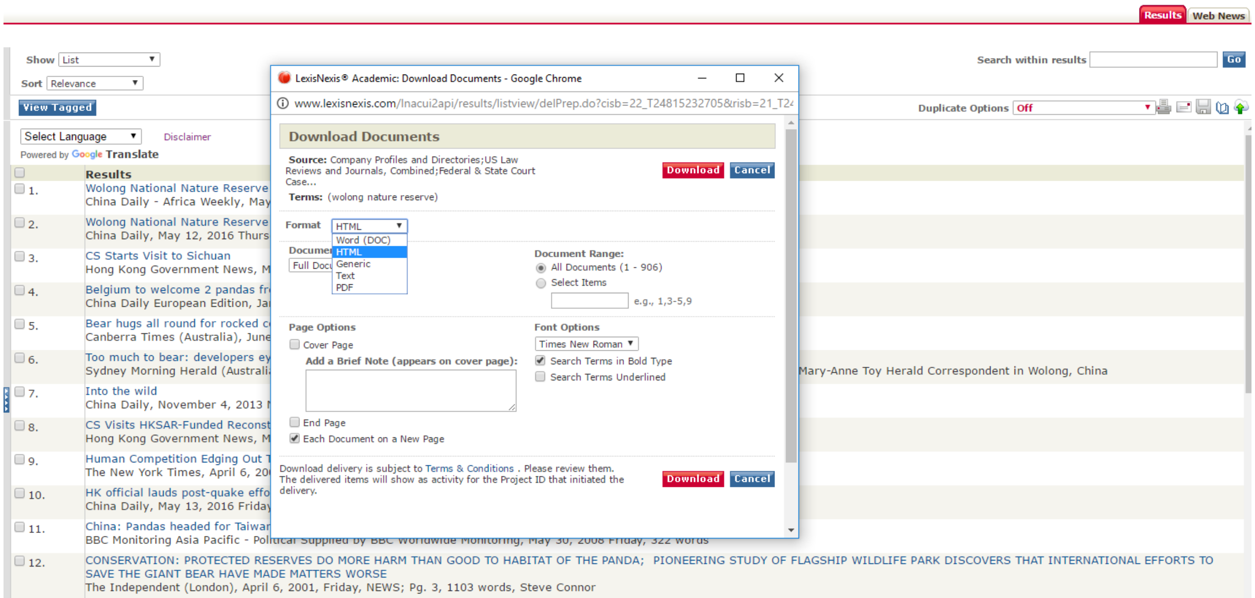
Go to the LexisNexis website[8], and you should see a window similar to the next figure:



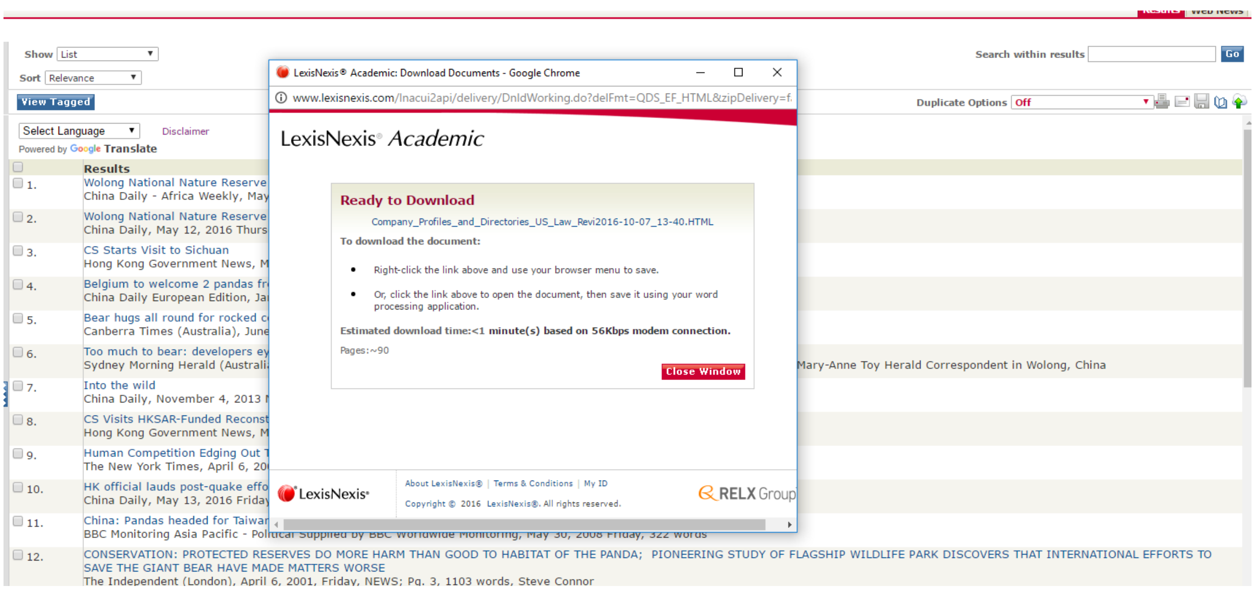
Inside the search box, type something like “wolong nature reserve”, and click the search button. The result page should list a number of records that were found to contain the searched term:

[8] http://www.lexisnexis.com/hottopics/lnacademic/

In the far right corner of the result page, click on the "Save" icon to open the following window:



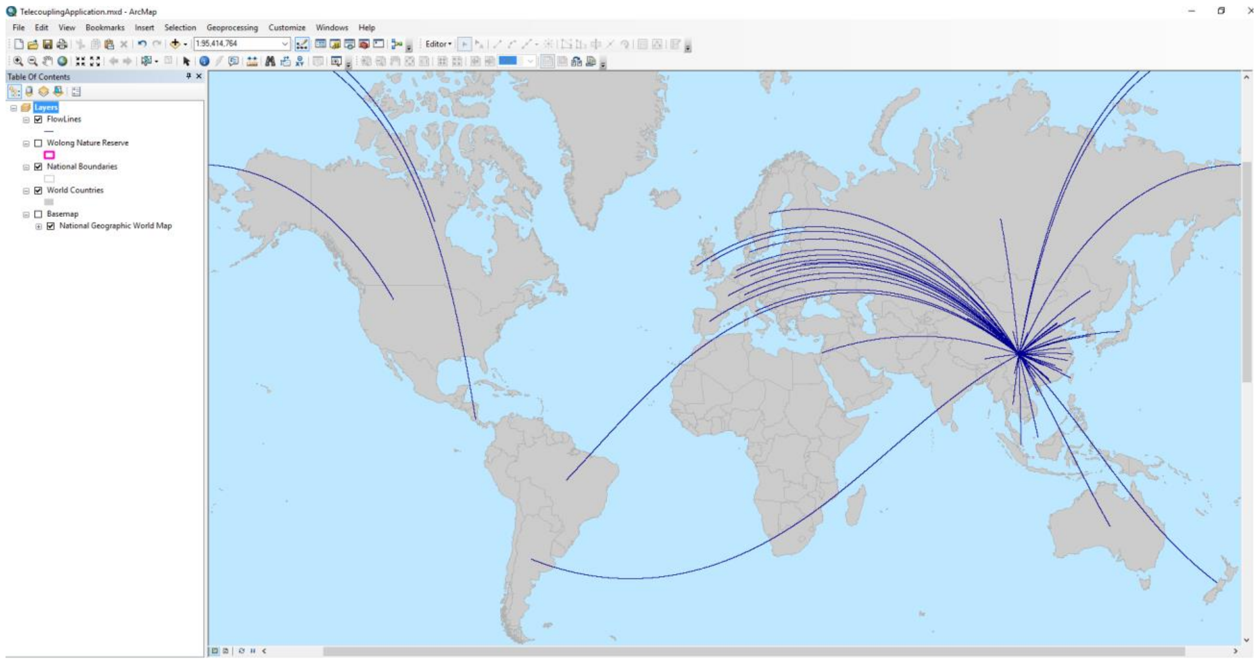
In the "Format" drop-down menu, select HTML. The document range tells you how many records have been found and whether you want to save all of them (default) to the HTML file. If you have too many records, an error message should come up asking to reduce the number of selected documents. In this example, select the second option "Select Items" and type "1-50" as range. Click on the download button and wait to receive a screen that looks like this:



You are now ready to download the HTML report file and save it to your local disk.

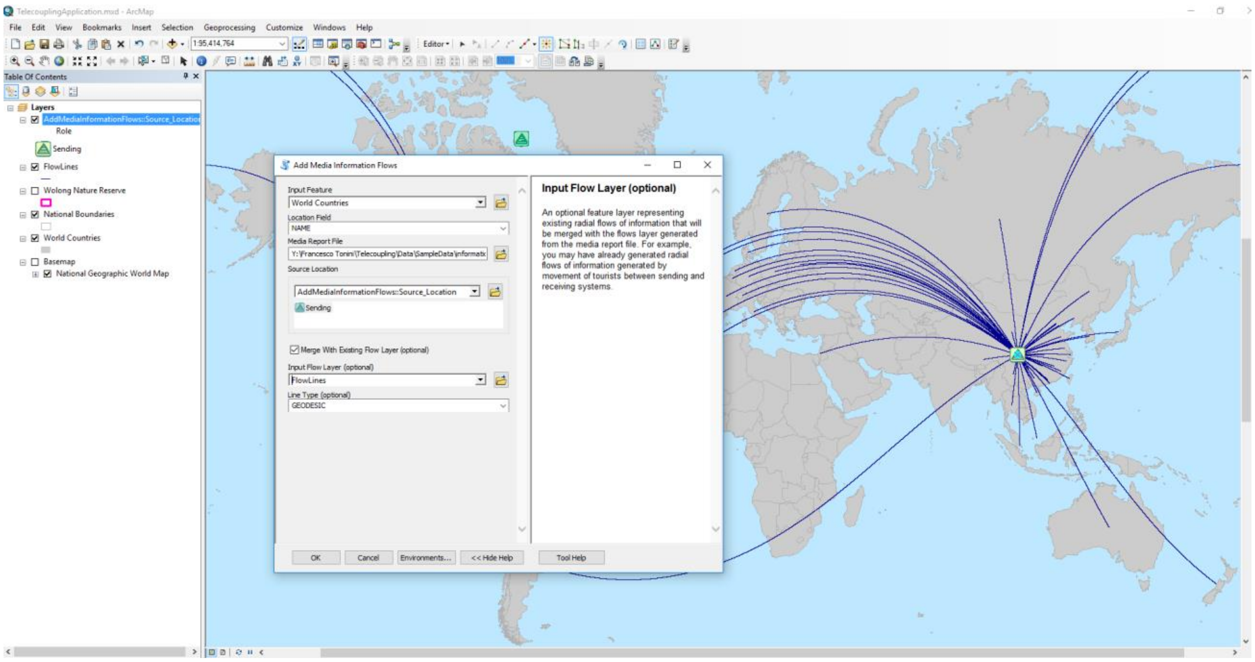
If you decided to skip the LexisNexis online word search, continue from this point on. For this tutorial, we are going to merge the media flows obtained from the HTML report file with an existing flow layer. To do this, use the "Draw Radial Flows" tool, and use the "*./SampleData_ArcGIS_v1.7.3b/Flows/Draw Radial Flows/info_Flows.csv*" table as input. This table contains records on start and end points for flows of tourists between a sending system (the Wolong Nature Reserve) and all receiving systems (i.e. countries of origin for the tourists). This is just one way information about the nature reserve can be spread to other countries by word of mouth. You can leave all other parameters similar to the example shown, and select "Frequency" as Join Field(s) parameter. Run the tool and you should get something like this:

***NOTE*: there is no way to quantify how much information (what units?) has been spread by each visiting tourist to their home countries. We acknowledge this may be a simplification but it helps understanding where the flow of information is likely going to. Finally, we chose to assign a "Frequency" equal to one to each single flow of tourist, given the unitless nature of the information amount in this case.**

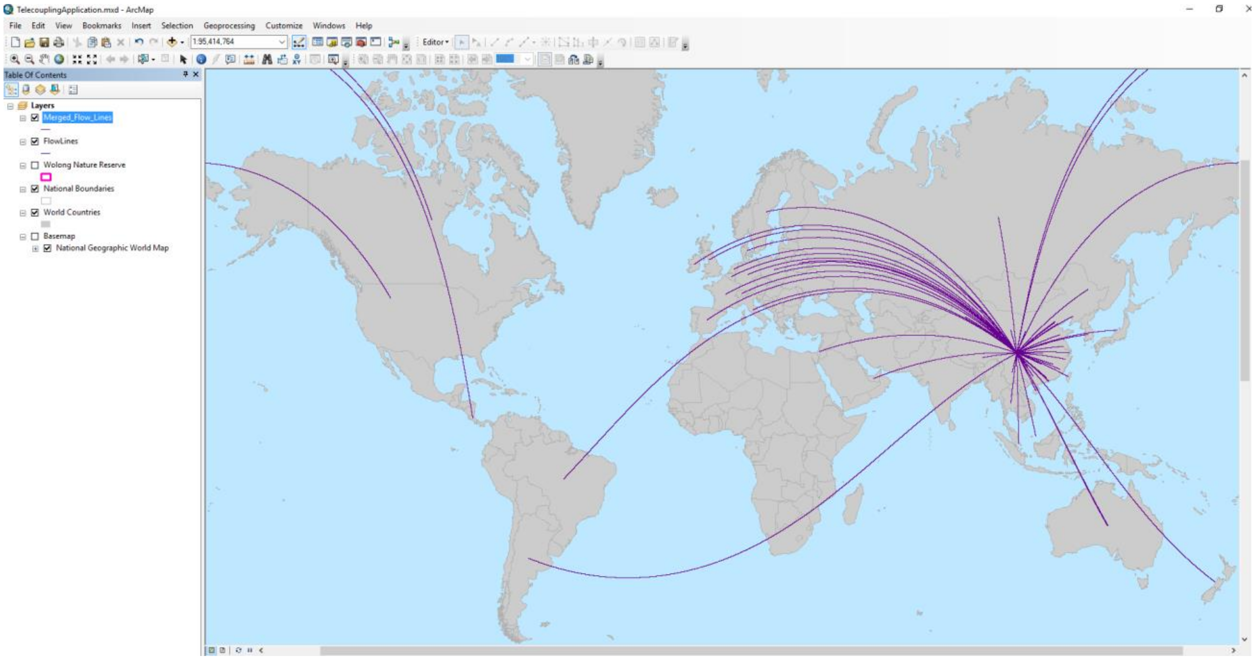


Now we are ready to use the “Add Media Information” tool. If you are using the *TelecouplingApplication.mxd* file, select “World Countries” as the input feature layer for the first parameter of the tool. If you are using a brand new map, then select the “*./SampleData_ArcGIS_v1.7.3b/Flows/Add Media Flows/World_countries_2002.shp*” layer. As location field parameter, select an attribute from the input layer that contains the name of a geographical location (e.g. country) that you would like to match against the geographic locations found inside the HTML report file. If you select an attribute with spatial location names that are not found in the report, no media flow layer will be generated and merged to the existing flow layer. Select “NAME” as location field parameter, and browse to the HTML report file saved on disk (if you followed this example step-by-step) or the “*./SampleData_ArcGIS_v1.7.3b/Flows/Add Media Flows/LexisNexis.html*” file provided with the sample dataset. Then, the user is asked to select a source location for the media flow by clicking on a location of interest on the map. This location should correspond as close as possible to the one where the information has disseminated from and that you searched on the LexisNexis web portal. In this example, please click on a location corresponding to the wolong nature reserve in China. Make sure the “Merge with Existing Flow Layer” box is checked since we chose to do so in this tutorial. As input flow layer, select the layer previously generated using the Draw Radial Flows tool. As Line type, you can leave the default (GEODESIC) or pick one that you like. Note that you should ideally pick the same line type as the existing flow layer to avoid strange visualization effects.

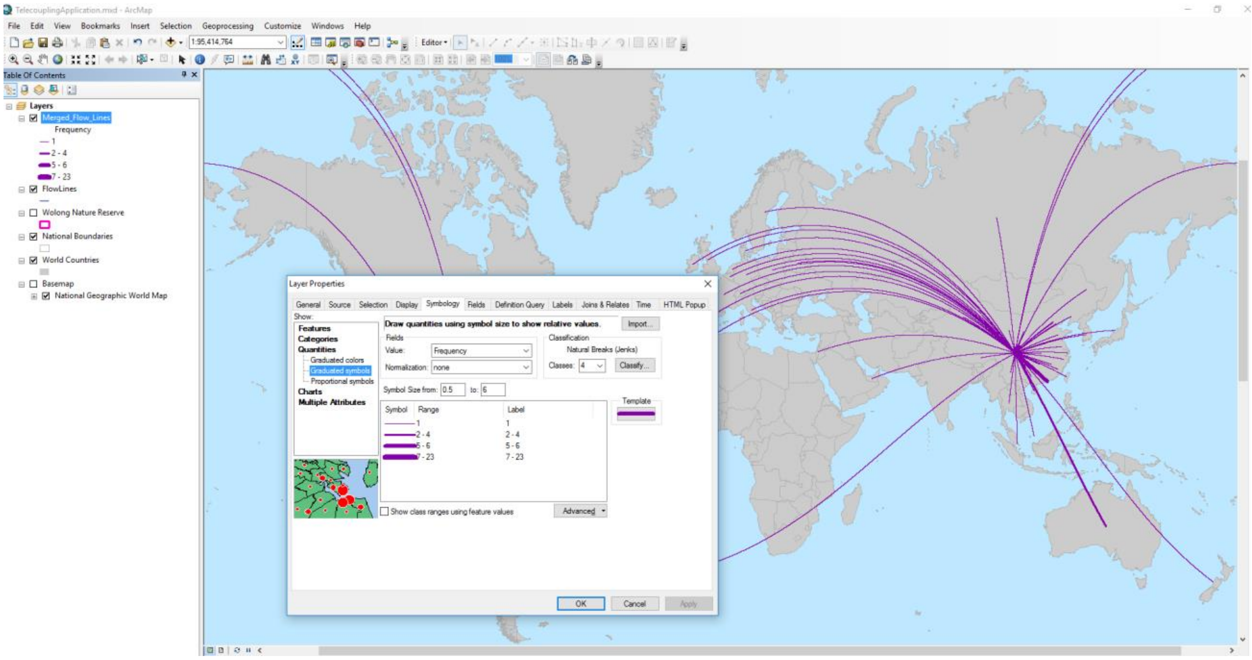
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



Once you click OK to run the tool, the program will start scraping (searching and extracting) the report file for geographical locations and count their frequencies matching them against the values found in the "Location Field" parameter of the tool. ArcGIS will create a permanent output line feature class representing merged information flows between systems and add it to the map.

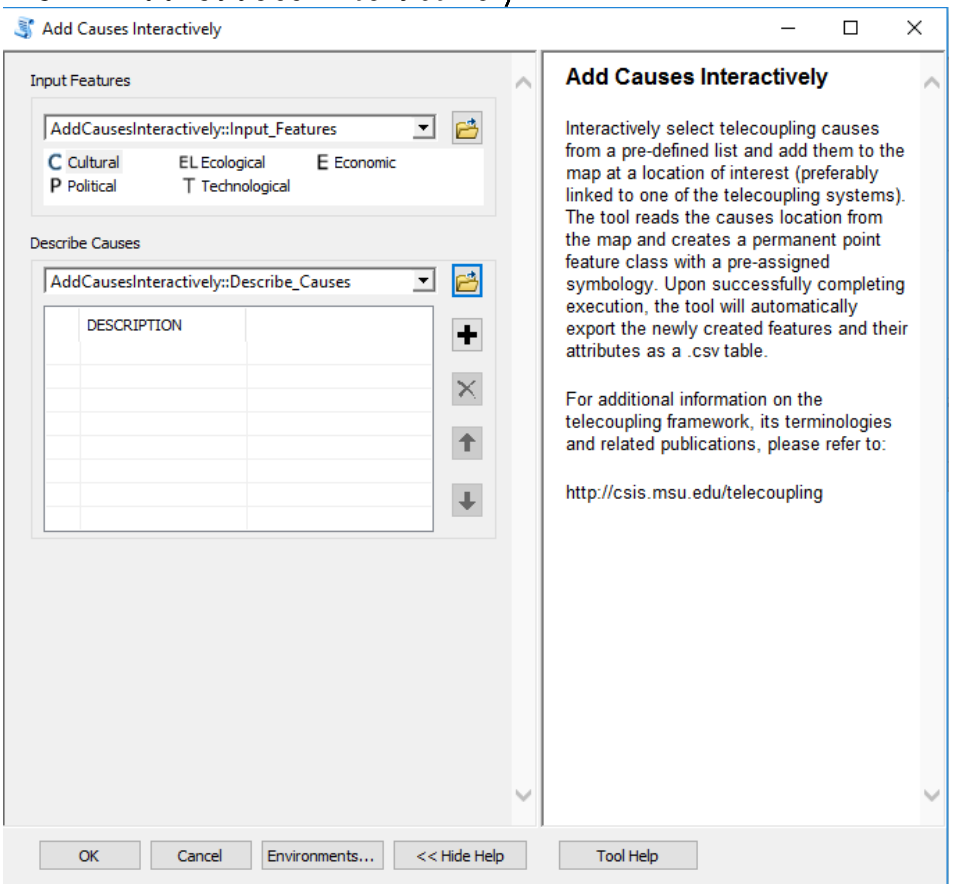


Once the output flow feature class has been created, the lines will all have the same color and thickness. If you have selected additional variables to represent the flow amount, you can use one of them to improve the visualization of the lines. To do so, right-click on the output flow layer found in the table of content (left-side of the map typically) and select the tab called "symbology". Then, you can select "quantities", and something like "graduated symbols" to show the line thickness proportional to the chosen variable of interest (e.g. frequency). Choose how many bins you want to group your values and a desired color.



4.5 CAUSES TOOLSET

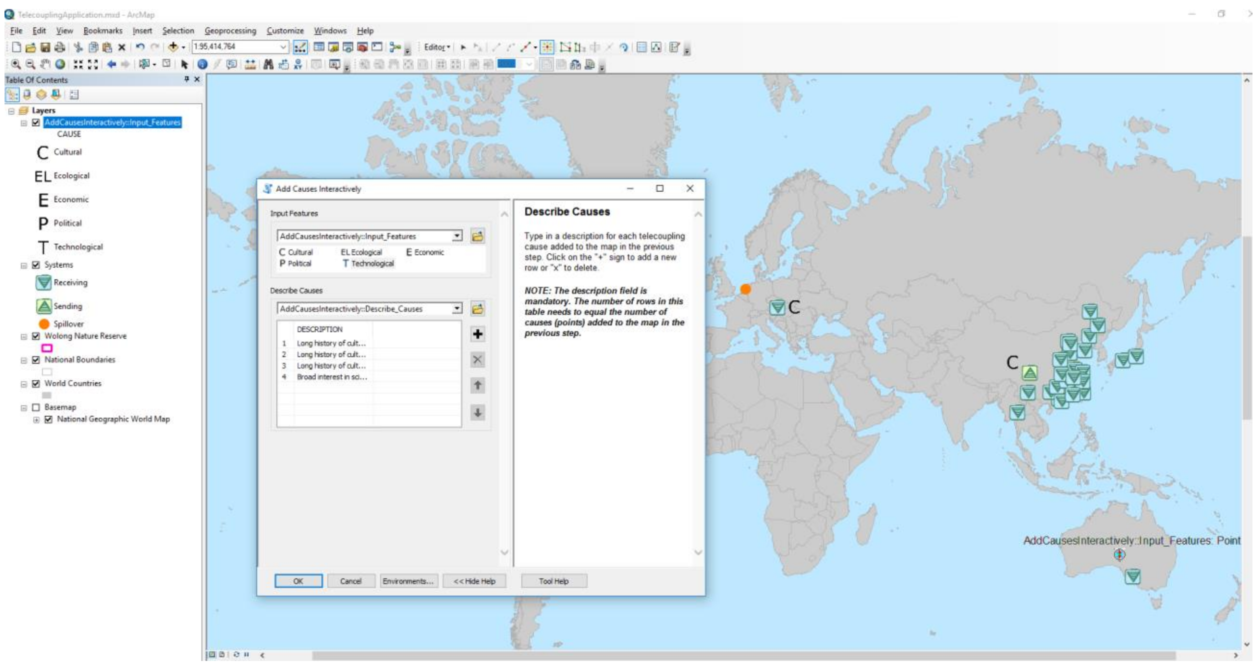
4.5.1 Add Causes Interactively



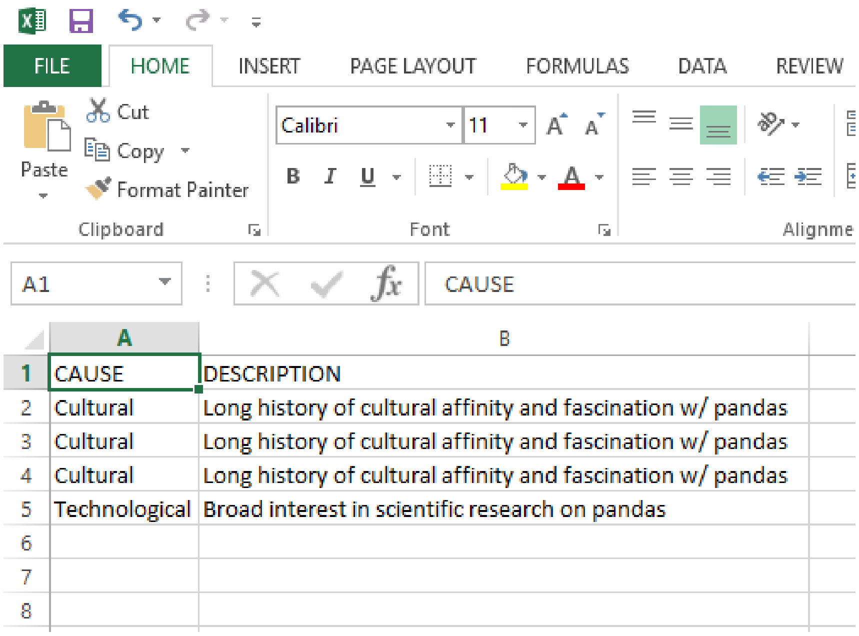
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

This tool should be used to **qualitatively** describe potential causes of an observed telecoupling flow by choosing within a pre-compiled list of broad causal categories (e.g. cultural, ecological, economic, political, technological). Once you select the category of interest, click on the map near a spatial location that corresponds to the telecoupling systems that is related to chosen cause. For example, a wildlife transfer flow may be instigated by a long history of cultural affinity and fascination with the species within one or more receiving systems (e.g. giant pandas are transferred to zoos in provinces and countries that have a cultural affinity for them). Alternatively, the factor instigating the flow may be solely economic given the anticipated return of investment from large amounts of tourists visiting the receiving system. After clicking on the map, add an informative description by clicking on the "+" sign of the second interface parameter and typing in the appropriate box. Repeat the above procedure for all other causes you wish to add to describe your telecoupling system and make sure the number of records in the description table corresponds to the number of points previously added to the map. Once you click OK to run the tool, ArcGIS will create a permanent output point feature class and save an output table to your workspace directory with a list of all the casual categories and their description you added to the map.

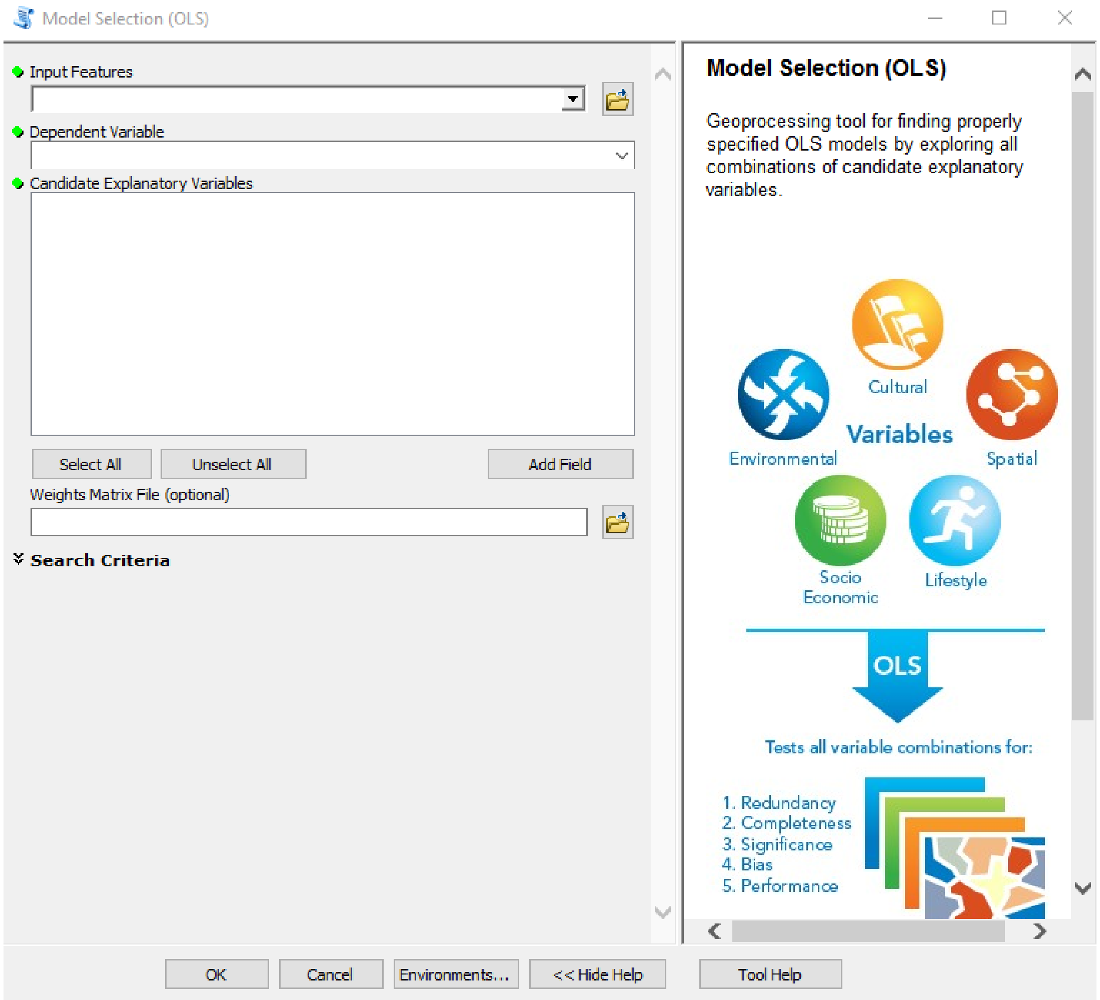
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



The output .csv table created by the tool in the scratch workspace directory should look something similar to the following figure:

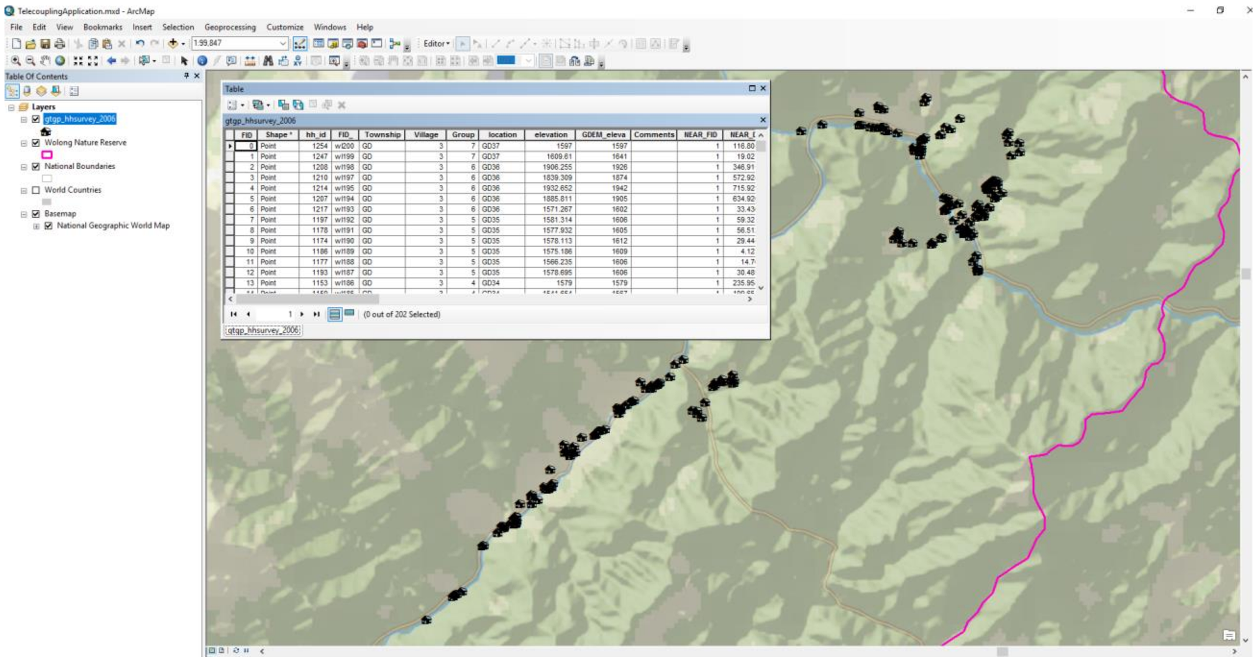

FILE
HOME
INSERT
PAGE LAYOUT
FORMULAS
DATA
REVIEW
Cut
Copy
Format Painter
Paste
Clipboard
Calibri
11
Font
Alignme
A1
CAUSE
A
B
1 CAUSE DESCRIPTION
2 Cultural Long history of cultural affinity and fascination w/ pandas
3 Cultural Long history of cultural affinity and fascination w/ pandas
4 Cultural Long history of cultural affinity and fascination w/ pandas
5 Technological Broad interest in scientific research on pandas
6
7
8

4.5.2 Model Selection (OLS)



Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

This tools, slightly modified from the original ESRI Exploratory Regression script tool, should be used to identify a set of factors that are most associated with a quantitative variable of interest, e.g. amount of cropland owned by local households in China. Factors that are strongly associated with the dependent variables, will be found statistically significant. The relative importance of significant factors can be determined by looking at their estimated OLS regression coefficients. Contrarily to the tool described in the previous section, the user must have a collection of spatial locations at which the dependent variable and a set of potential factors (explanatory variables) are recorded. For example, use the shapefile found with the provided sample dataset ("*./SampleData_ArcGIS_v1.7.3b/Causes/Model Selection OLS /gtgp_hhsurvey_2006.shp*").



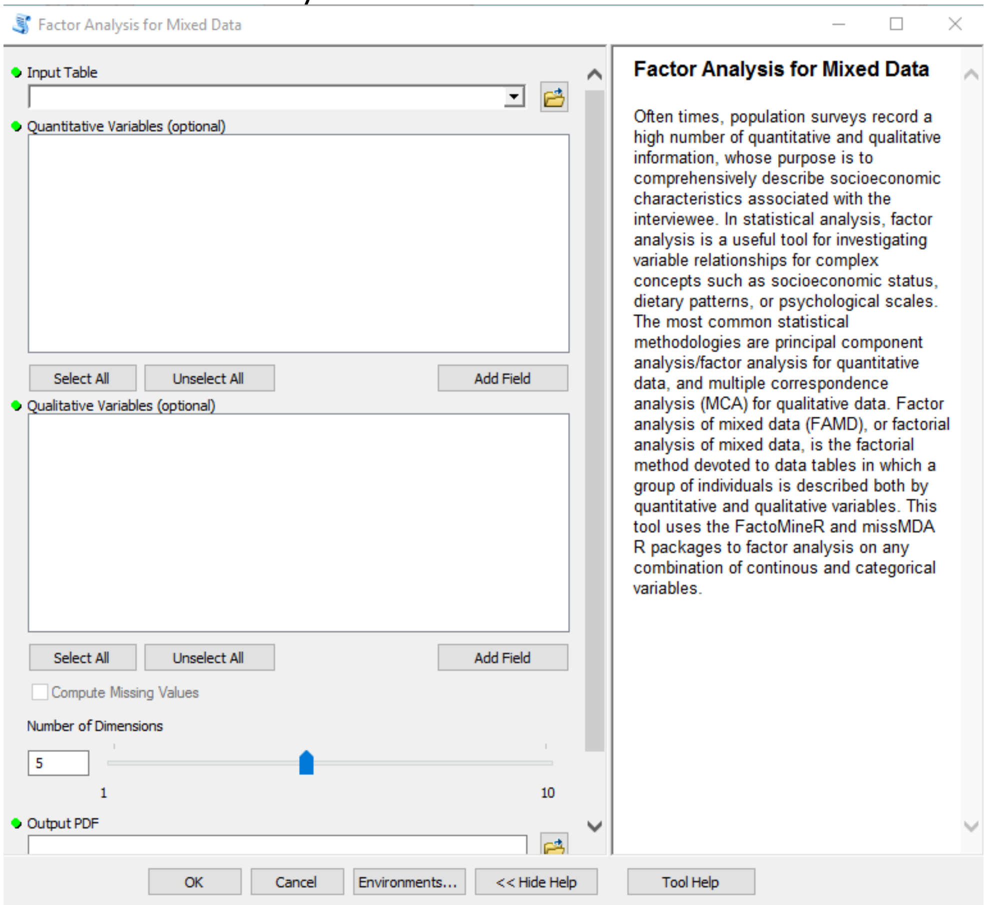
This shapefile describes spatial location of local households in the Wolong Nature Reserve, China, and a number of socio-economic variables recorded in 2006 with a targeted household survey including cropland size, household age, median income, educational level, and others. You may be interested in identifying the factors (variables) that are most associated with the amount of crop produced by each household. Select the shapefile as input feature. As a dependent variable, select the "Crop_produ" attribute. For the candidates explanatory variables, select "*GTGP*" (the percentage of cropland enrolled in the Grain-to-Green-Program, subsidies paid by the Chinese government to stimulate conversion of cropland to forest land), "*GTBP*" (the percentage of cropland enrolled in the Grain-to-Bamboo-Program, subsidies paid by the Chinese government to convert cropland to bamboo land), "*Tourism_pa*" (does the household have a member who directly participated in tourism activities in 2005?), "*Log__dista*" (log distance to the main road), and "*Number_of*" (total number of laborers in the household). Leave all other tool options unaltered and click OK to run the tool. If all goes smoothly, the tool identifies which factors best associate with cropland owned for agricultural production among the variables selected. If you open the result window to inspect the results (Menu > Geoprocessing > Results), under "Current Session", right-click the "Message" icon and select "View". A separate window should open showing you the entire process executed by the tool and all the OLS models that have been tested.

```
Messages
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0.41 935.05 0.00  0.00 1.31 0.04  -GTGP***  -GTBP***  -TOURISM_PA**  +NUMBER_OF***  +LOG__DISTA
      Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model

******************************************************************************
************* Exploratory Regression Global Summary (CROP_PRODU) *************

              Percentage of Search Criteria Passed
                   Search Criterion Cutoff Trials # Passed % Passed
             Min Adjusted R-Squared > 0.50     31        0     0.00
            Max Coefficient p-value < 0.05     31       22    70.97
                      Max VIF Value < 7.50     31       31   100.00
            Min Jarque-Bera p-value > 0.10     31        0     0.00
Min Spatial Autocorrelation p-value > 0.10     13        7    53.85

------------------------------------------------------------------------------

      Summary of Variable Significance
Variable   % Significant % Negative % Positive
GTGP              100.00     100.00       0.00
GTBP              100.00     100.00       0.00
TOURISM_PA        100.00     100.00       0.00
NUMBER_OF         100.00       0.00     100.00
LOG__DISTA         43.75       0.00     100.00

------------------------------------------------------------------------------

   Summary of Multicollinearity
Variable    VIF Violations Covariates
GTGP       1.16      0      --------
GTBP       1.31      0      --------
TOURISM_PA 1.09      0      --------
NUMBER_OF  1.02      0      --------
LOG__DISTA 1.12      0      --------

------------------------------------------------------------------------------
```

The model summary shows that the most significant factors are all but the distance to the main road. Some factors such as GTGP, GTBP, TOURISM_PA have a negative association, meaning the higher their value, the lower the cropland amount is. In the case of a dummy variable (0-1, yes-no) like TOURISM_PA, this means that a participation of household members in tourism activities in 2005 had a negative effect on the total cropland amount. On the other hand, NUMBER_OF has a positive association with cropland amount, i.e. the more laborers in the household the higher its cropland owned for agricultural production was.

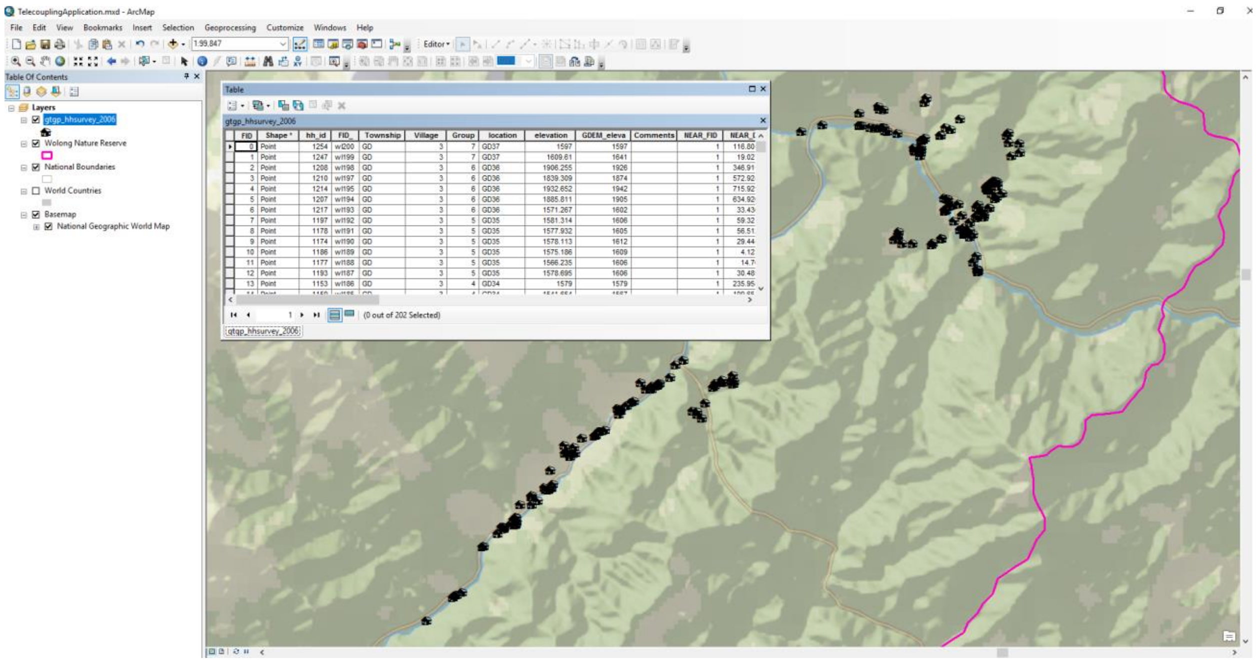
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

4.5.3 Factor Analysis for Mixed Data

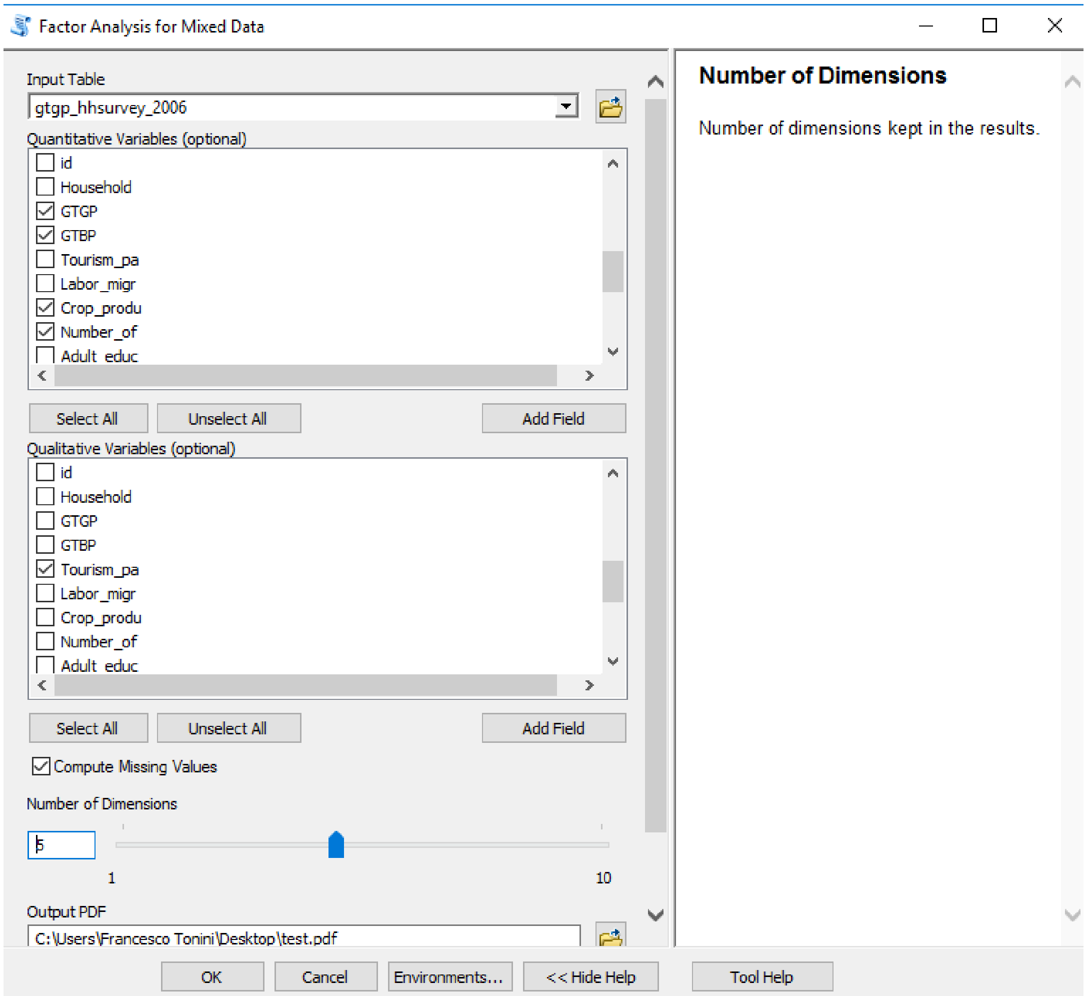


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

This tools can be used to explore qualitative and quantitative variable relationships in order to identify distinct groups of individuals or variables that may be grouped into a smaller number hidden factors. For example, use the shapefile found with the provided sample dataset ("*./SampleData_ArcGIS_v1.7.3b/Causes/Model Selection OLS/gtgp_hhsurvey_2006.shp*").



This shapefile describes spatial location of local households in the Wolong Nature Reserve, China, and a number of socio-economic variables recorded in 2006 with a targeted household survey including crop production, mean household age, median income, educational level, and others. You may be interested in investigating the relationships among a number of quantitative and qualitative variables recorded in this survey, with the goal of comprehensively describe socioeconomic characteristics associated with the interviewees. Select the shapefile as input feature. As quantitative variables, select "*GTGP*" (the percentage of cropland enrolled in the Grain-to-Green-Program, subsidies paid by the Chinese government to stimulate conversion of cropland to forest land), "*GTBP*" (the percentage of cropland enrolled in the Grain-to-Bamboo-Program, subsidies paid by the Chinese government to convert cropland to bamboo land), "*Crop_produ*" (cropland owned by the household for agricultural production in 2005), and "*Number_of*" (total number of laborers in the household). As qualitative variables, select "*Tourism_pa*" (does the household have a member who directly participated in tourism activities in 2005?). Select a name and location on disk where to save an output pdf report file with plots of the factorial analysis. Leave all other tool options unaltered and click OK to run the tool.



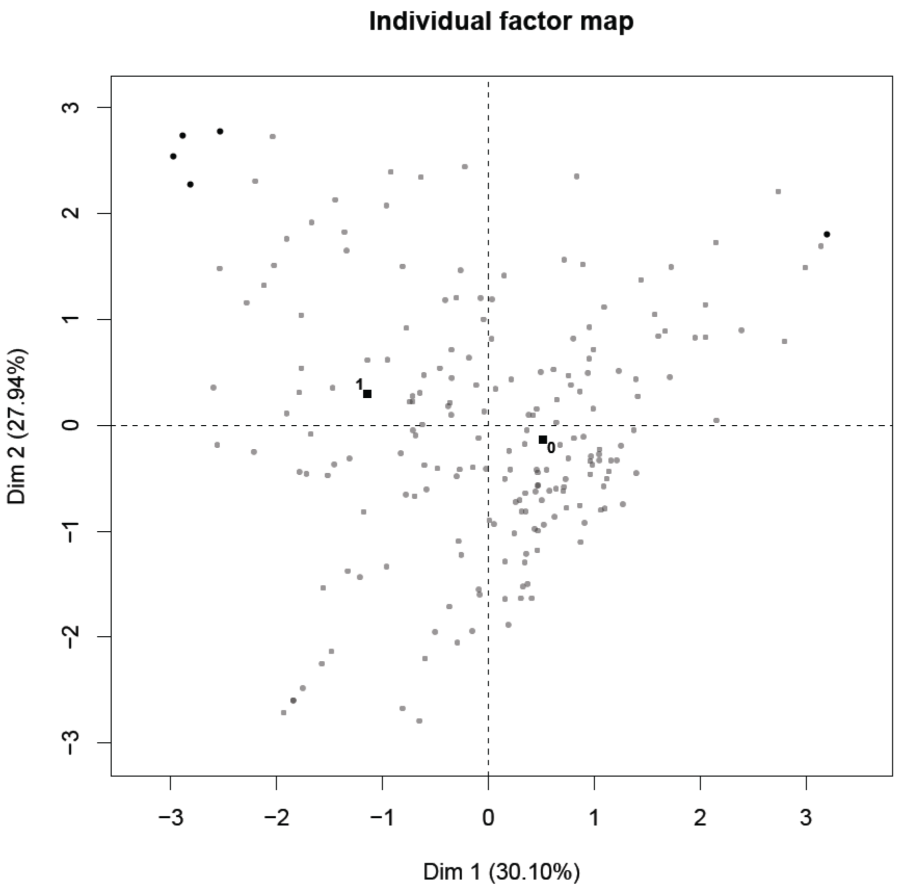
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

If all goes smoothly, the tool identifies whether distinct groups of household units exist and if any of the variables selected are highly correlated with each other and the first two extracted dimension. If you open the result window to inspect the results (Menu > Geoprocessing > Results), under "Current Session", right-click the "Message" icon and select "View". A separate window should open showing you the entire process executed by the tool and all the estimated scores on the first N dimensions for all the selected variables:

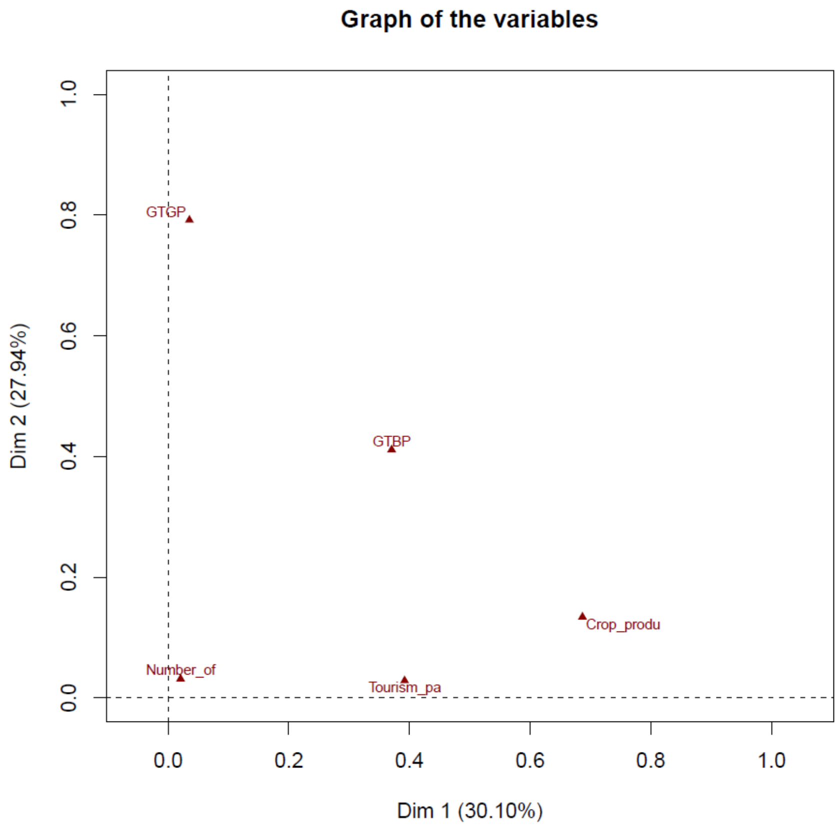
Messages

```
Creating Dataframe from Input Table...
Checking for missing data in the dataset...
...No missing values found in the dataset!
Running Factor Analysis...
Call:
FAMD(base = df, ncp = as.integer(num_fact), graph = FALSE)
Eigenvalues
                       Dim.1   Dim.2   Dim.3   Dim.4   Dim.5
Variance               1.505   1.397   1.112   0.699   0.286
% of var.             30.104  27.945  22.236  13.990   5.726
Cumulative % of var.  30.104  58.048  80.284  94.274 100.000
Individuals (the 10 first)
               Dist    Dim.1    ctr   cos2    Dim.2    ctr   cos2    Dim.3
1           |  2.651 | -2.284  1.716  0.742 |  1.157  0.474  0.190 | -0.172
2           |  2.740 | -1.767  1.027  0.416 |  0.538  0.103  0.039 | -1.161
3           |  2.711 |  1.726  0.980  0.405 |  1.493  0.790  0.303 | -0.992
4           |  2.076 |  1.605  0.848  0.598 |  0.841  0.251  0.164 |  0.756
5           |  1.381 |  0.463  0.070  0.112 | -1.180  0.493  0.731 |  0.216
6           |  1.082 |  0.729  0.175  0.454 | -0.508  0.091  0.220 | -0.603
7           |  0.959 |  0.898  0.266  0.878 | -0.109  0.004  0.013 |  0.058
8           |  2.935 | -0.964  0.305  0.108 |  2.074  1.524  0.499 | -1.392
9           |  1.160 | -0.146  0.007  0.016 | -0.398  0.056  0.118 | -0.850
10          |  0.893 |  0.195  0.013  0.048 | -0.242  0.021  0.074 | -0.135
              ctr   cos2
1            0.013  0.004 |
2            0.601  0.180 |
3            0.438  0.134 |
4            0.254  0.133 |
5            0.021  0.024 |
6            0.162  0.310 |
7            0.002  0.004 |
8            0.863  0.225 |
9            0.322  0.537 |
10           0.008  0.023 |
Continuous variables
              Dim.1    ctr   cos2    Dim.2    ctr   cos2    Dim.3    ctr   cos2
GTGP        | -0.188  2.343  0.035 | -0.890 56.656  0.792 |  0.251  5.658  0.063
GTBP        | -0.609 24.614  0.370 |  0.641 29.429  0.411 | -0.114  1.177  0.013
```

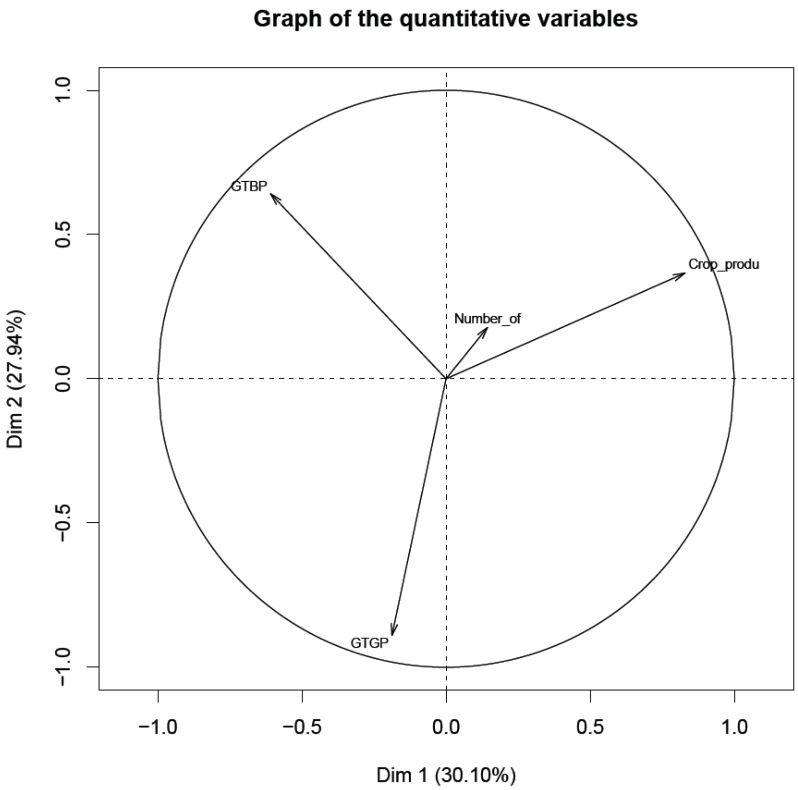
Within each extracted dimension, higher values indicate a higher contribution of an individual or variable to the composition of that dimension. If distinct groups of individual units (e.g. households) or variables exist, higher scores will show for different units/variables in separate dimensions. In most cases with a lot of individuals (e.g. > 200 households) it becomes hard to inspect these particular scores, thus it better to have a visual aid in the output plots from the analysis. Individual factor map output, using default options for the graphical parameters in the tool interface, shows a dispersed cloud of points. Therefore, there does not seem to be evidence of a distinct cluster of households based on the chosen socioeconomic variables. The 0-1 labelled squares in the plot show association between the categorical variable “tourism_pa” and the first dimension (positive and negative, respectively). A small group of household units seems to be associated with the 0 value of this variable.



The graph of the variables shows that “GTGP” is strongly associated with the second dimension, while “crop_produ” and “tourism_pa” have a medium-to-high association with the first dimension.

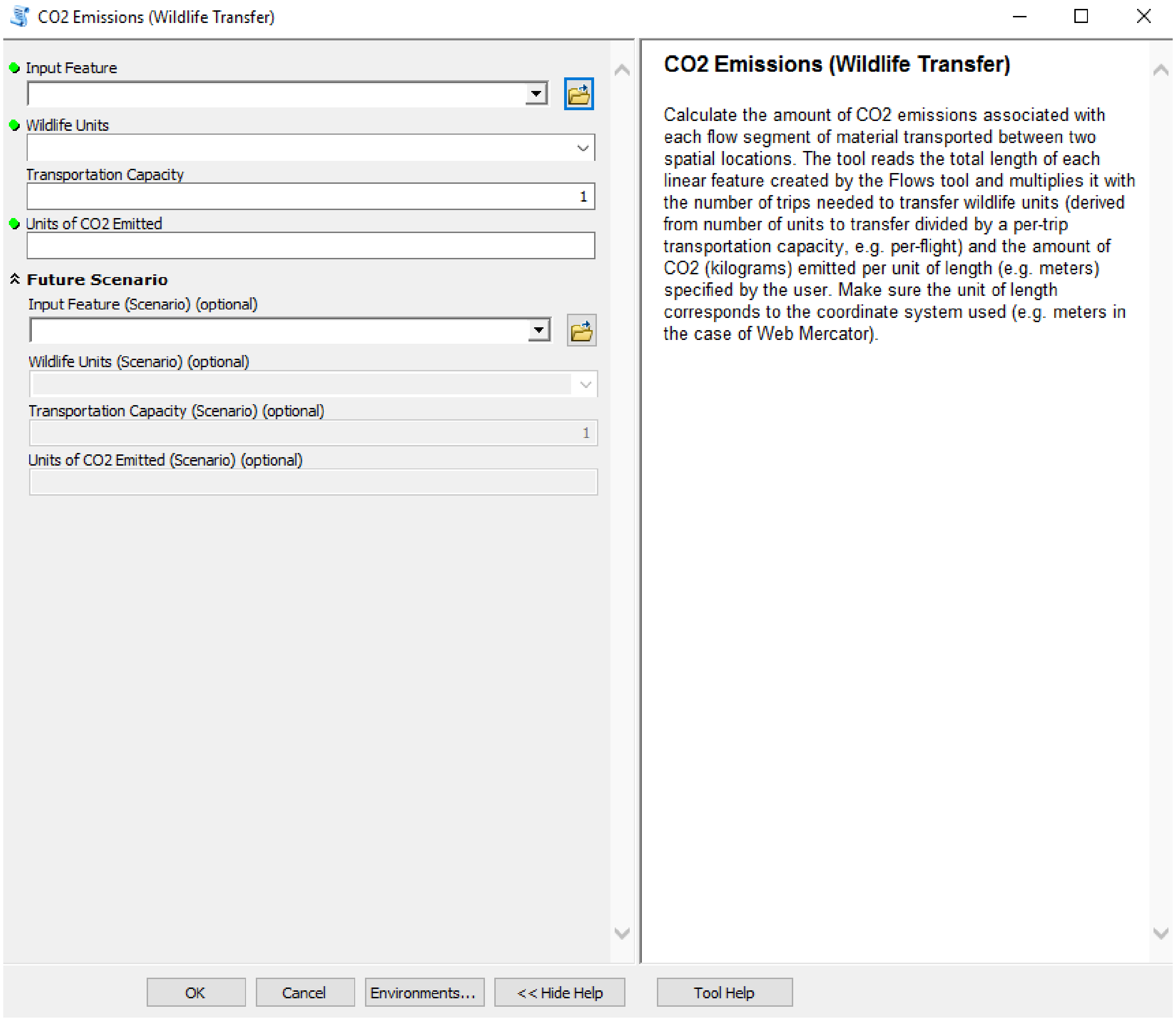


The graph of the quantitative variables on the unit circle, tells which quantitative variables are mostly correlated with each other as well as with the first two dimensions. The "GTBP" has a negative correlation with the first dimension and a positive one with the second one, while the angle almost diagonal to the second quadrant on the top-left, indicates the variable is not well defined in either of those dimensions. The "GTGP" variable shows high negative association with the second dimension, while the "Crop_produ" shows a high correlation with the first dimension. Therefore, the first two dimensions are mostly dominated and defined by "GTGP" and "Crop_produ". The plot indicates that the lower the percentage of cropland enrolled in the GTGP subsidy program (i.e. household receives less subsidies) the higher the amount of cropland owned by the household for agricultural production. Moreover, "tourism_pa" helps defining the first dimension as well, with values of 1 (household members participated in tourism-related activities) negatively associated with it, i.e. higher participation in tourism corresponds to lower cropland amounts dedicated to agriculture. These conclusions are what you would expect, given that households enrolling less cropland in the GTGP program, thus receiving less subsidies, are likely to retain a higher amount of cropland for agriculture. At the same time, if the household income is partially made of tourism-related activities, it is more likely that they will have less land dedicated to agriculture and more willing to enroll it into the GTGP.



4.6 ENVIRONMENTAL ANALYSIS TOOLSET

4.6.1 CO_2 Emissions (Wildlife Transfer)

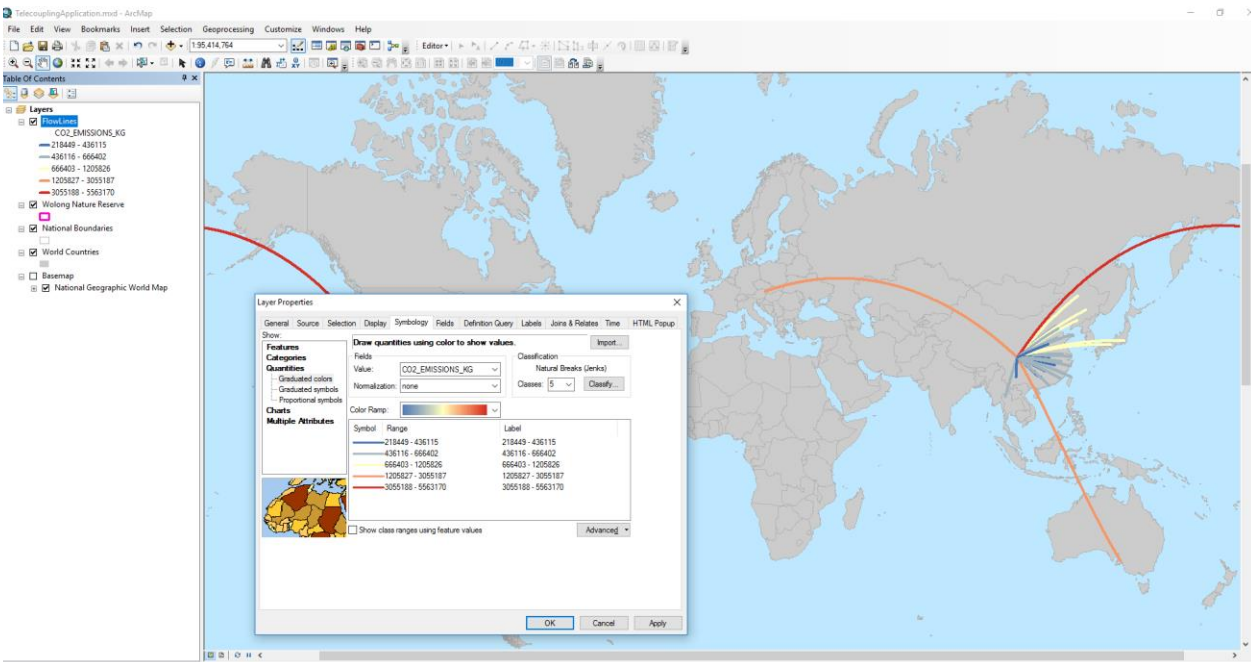


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

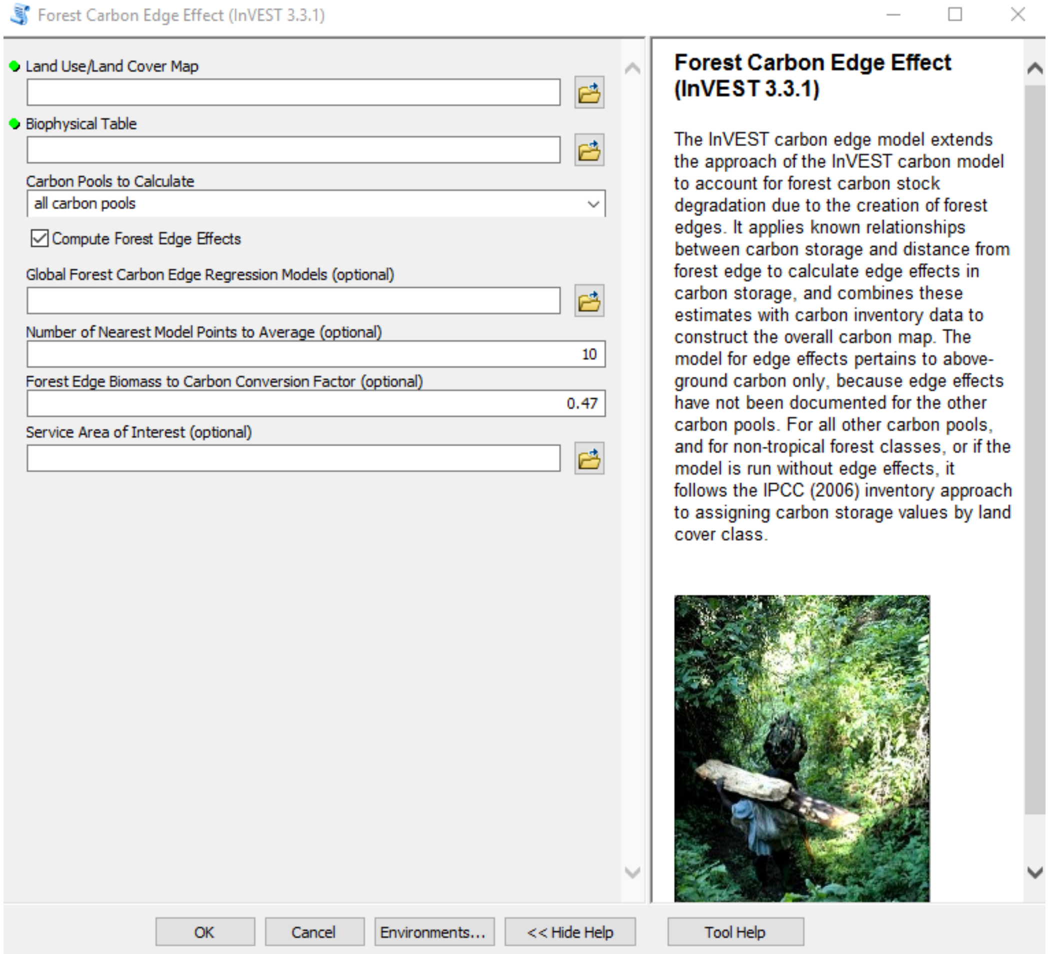
This tool uses the flow layer computed with the Draw Radial Flows tool and calculates the amount of CO_2 emissions based on the length of the flow lines, number of units to transfer,

maximum transportation capacity of the medium used, and an estimated amount of CO_2 (kilograms) emitted per unit-length (e.g. meters). For this example, make sure you first computed the radial flows and select that layer as input for the first parameter here. Then, choose a field from the input layer that represents the number of wildlife units that are to be transferred. In the capacity parameter (default = 1) specify the maximum capacity of the transportation medium in terms of units of wildlife. Finally, type in an estimated amount of CO_2 emitted per unit of length and click OK to run the tool. Once completed, open your flow layer attribute table (right-click on the layer) and make sure there is an extra field with values generated for CO_2 emissions.
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

You can now symbolize the flows by the amount of CO_2 emitted in the atmosphere, as shown in the next figure:



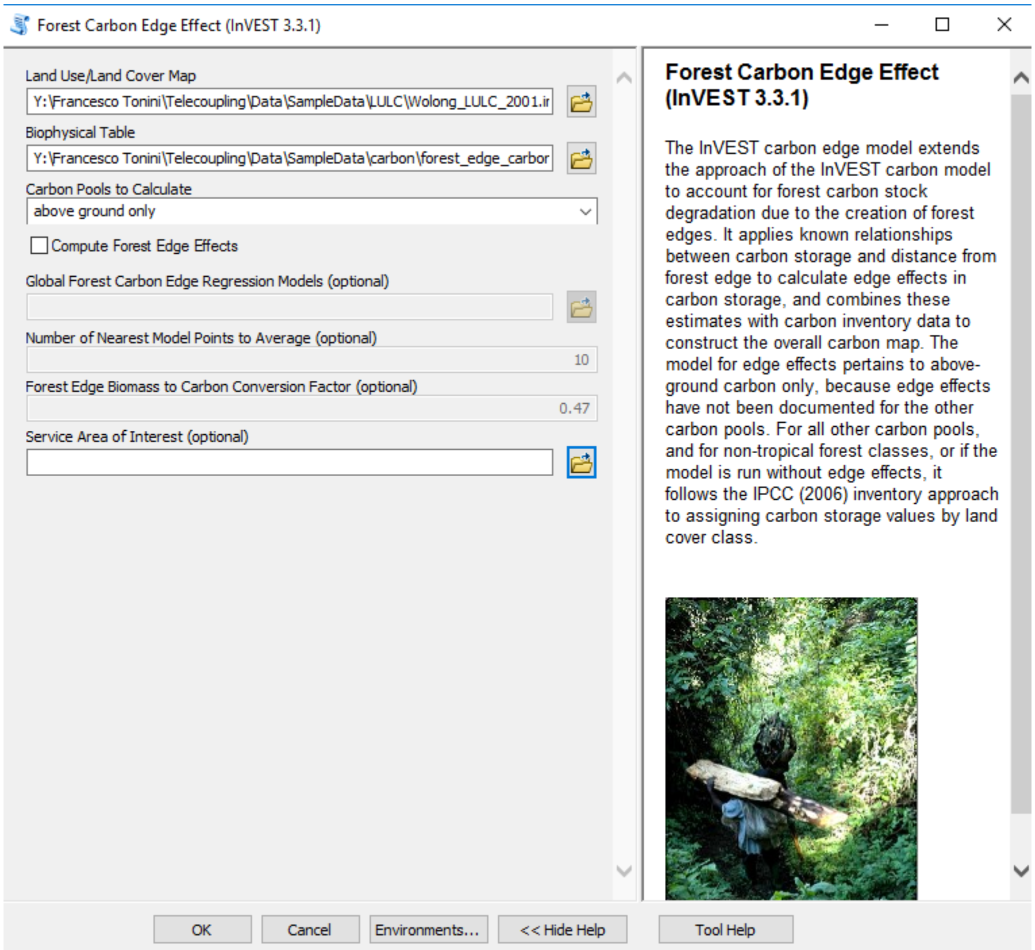
4.6.2 Forest Carbon Edge Effect (InVEST 3.3.3)



Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

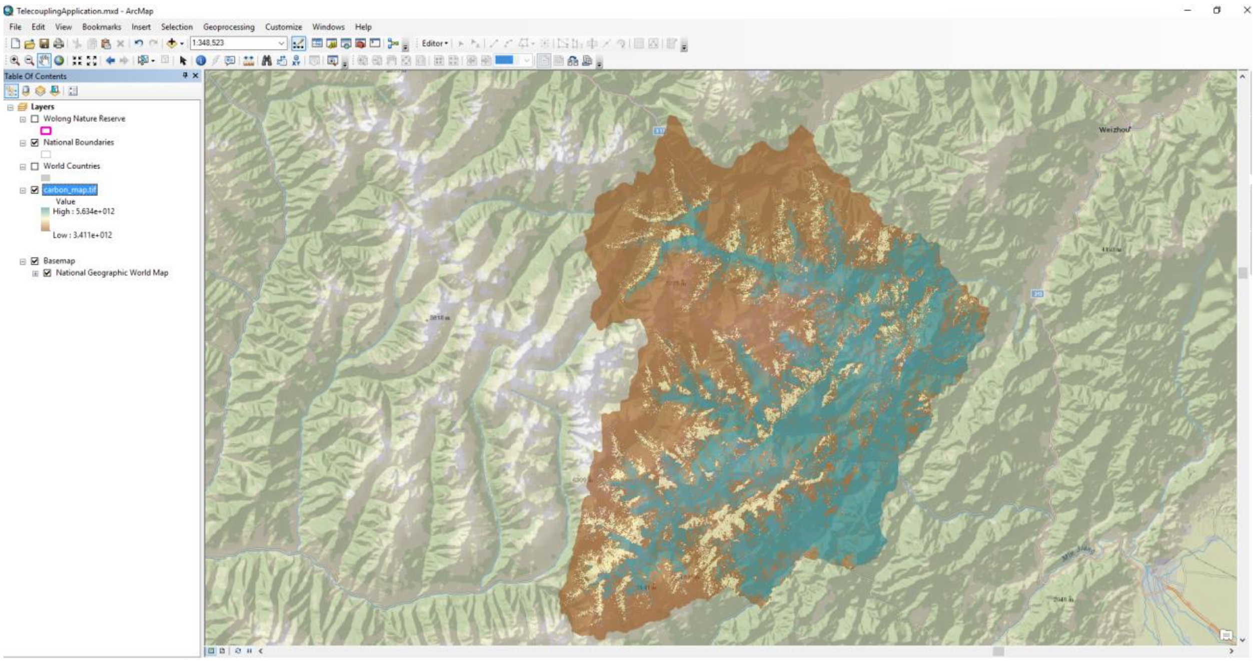
NOTE: This model is recommended over the simple carbon storage and sequestration model.

This tool is linked to the InVEST 3.3.3 forest carbon edge effect model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this example, let's use the land use/land-cover raster files from 2001 for the Wolong Nature Reserve area ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Forest Carbon Edge Effect/Wolong_LULC_2001.img*") to calculate above-ground carbon amounts in 2001. The tool also needs a carbon pool table input ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Forest Carbon Edge Effect/forest_edge_carbon_lu_table_2000.csv*") where each land cover type found in the input raster is associated with values of carbon (above, below, dead). This table is very important as it used by the biophysical model to determined total carbon values per pixel. There are multiple carbon grid products on the web (e.g. MODIS NPP). In our case, the table was compiled using MODIS NPP data for the closest year (2000) to our baseline land cover raster (2001). Because the NPP grid has, like in most situations, different spatial resolution and units of measurements from the one needed by the tool, we had to do some pre-processing work to calculate the average amount of above ground carbon found within each class of land cover used. If you are able to obtain data on below-ground or dead carbon, please add it in your table, as this will make your analysis more accurate and account for all carbon pools. Select above-ground only to be calculated (we do not have data on other carbon pools), and uncheck the "Compute Forest Edge Effects" box. This option becomes useful in tropical areas where the model uses a carbon edge regression model to better account for carbon pool near the edge of forest patches. Leave the Service Area of Interest parameter blank, but keep in mind that you can specify a shapefile of an area of interest for which you would like an overall aggregate estimate of carbon. For example, if the shapefile has three different zoning areas, this tool will return an overall aggregate carbon estimate for each of the three zones.

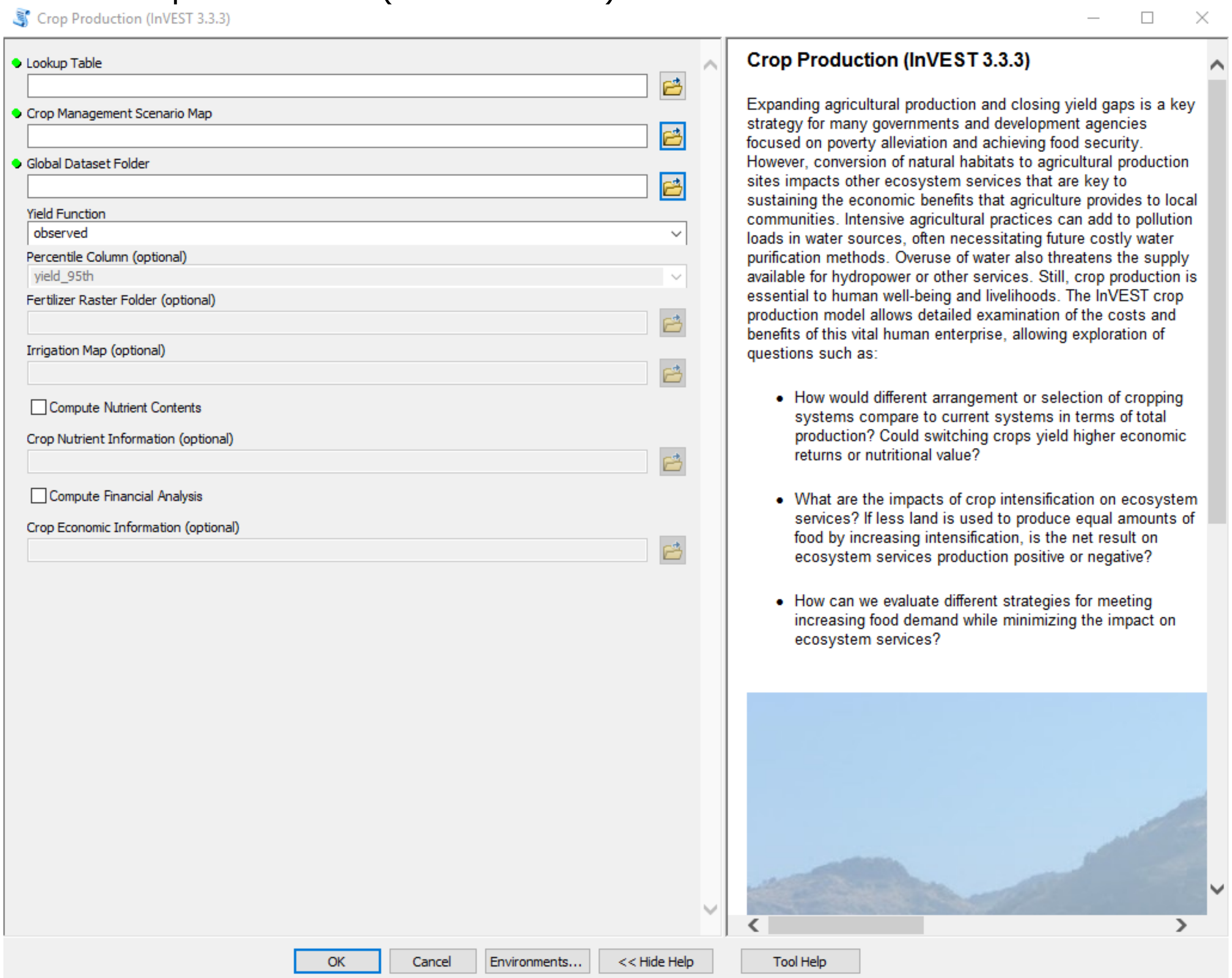


<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

Click OK to run the tool. Once completed, you should have an output raster with estimated above-ground carbon. To improve the visualization of the raster, let's choose a better color ramp using the symbology tab (right-click the raster layer and select Properties). Select a continuous color ramp of your choice. Click on OK and you should see something like the following image:


TelecouplingApplication.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Table Of Contents
Layers
Wolong Nature Reserve
National Boundaries
World Countries
carbon_map.tif
Value
High : 5.634e+012
Low : 3.411e+012
Basemap
National Geographic World Map
Weizhou

4.6.3 Crop Production (InVEST 3.3.3)

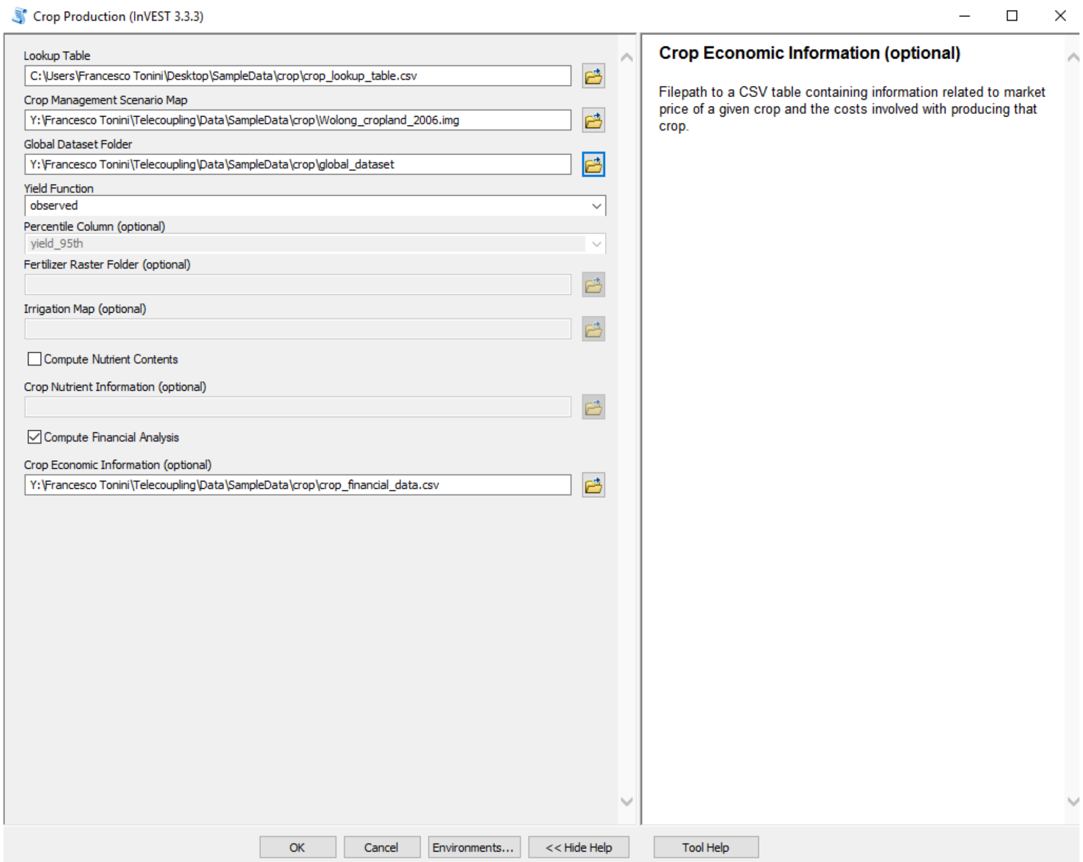


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

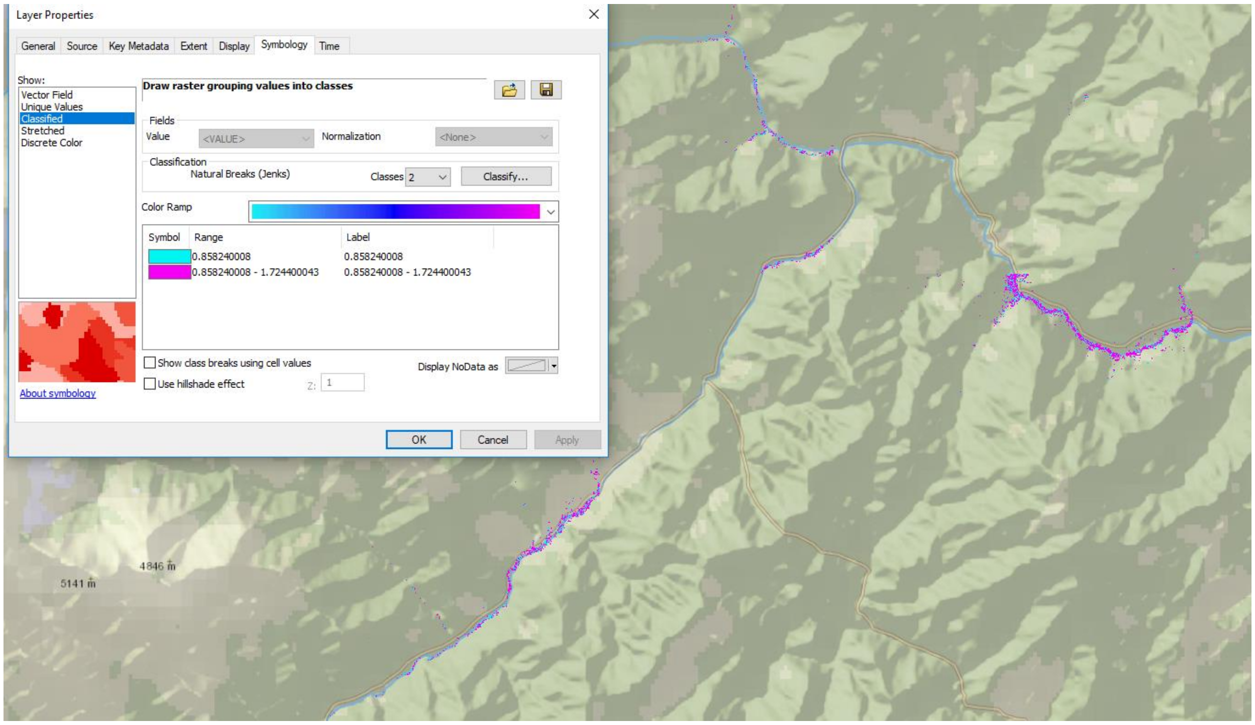
NOTE: At this very moment, this InVEST model is NOT recommended for real decision-making as it needs more validation. However, feel free to use it for exploratory and testing purposes.

This tools is linked to the InVEST 3.3.1 crop production model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this example, let’s use the crop management scenario raster file from 2006, for the Wolong Nature Reserve area ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Crop Production/Wolong_cropland_2006.img*"). The main crop produced in this area is cabbage, followed by potatoes and corn (usually in rotation within the same crop field). The crop management map shows areas where cabbage and potatoes are cultivated by local farmers, and areas where no cropland is present. The crop lookup table ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Crop Production/crop_lookup_table.csv*") needs to have a one-to-one correspondence with the codes found in the crop management raster file to work properly. Specify the location of the global raster dataset folder, which contains data for observed yield as well as climate percentile and regression yields used by the InVEST model ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Crop Production/global_dataset*"). Check the box to compute financial analysis and specify the crop economic information table ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Crop Production/crop_financial_data.csv*"). If you wish to calculate the total nutritional content of the chosen crop, check the appropriate box and use the nutrient table (*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Crop Production/crop_nutrient_data.csv*).

<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

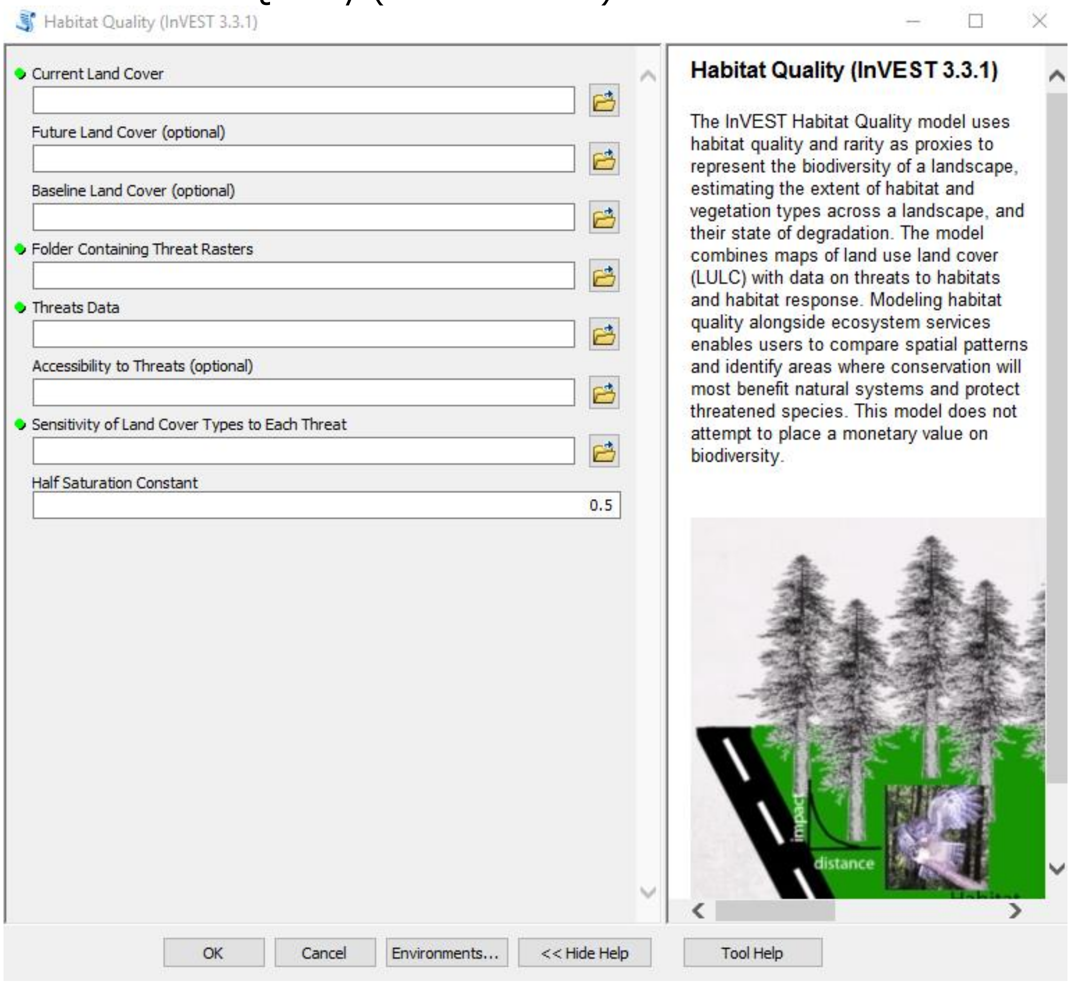


Click OK to run the tool. Once completed, you should have an output raster with estimated yield and a financial analysis table with total yield, costs, revenues, and returns for the chosen crops. To improve the visualization of the yield raster, let’s choose a better color ramp using the symbology tab (right-click the raster layer and select Properties). Select a Classified color ramp of your choice. Click on OK and you should see something like the following image:



Layer Properties
General
Source
Key Metadata
Extent
Display
Symbology
Time
Show:
Vector Field
Unique Values
Classified
Stretched
Discrete Color
Draw raster grouping values into classes
Fields
Value
<VALUE>
Normalization
<None>
Classification
Natural Breaks (Jenks)
Classes 2
Classify...
Color Ramp
Symbol
Range
Label
0.858240008
0.858240008
0.858240008 - 1.724400043
0.858240008 - 1.724400043
Show class breaks using cell values
Use hillshade effect
Z: 1
Display NoData as
About symbology
OK
Cancel
Apply
4846 m
5141 m

4.6.4 Habitat Quality (InVEST 3.3.3)

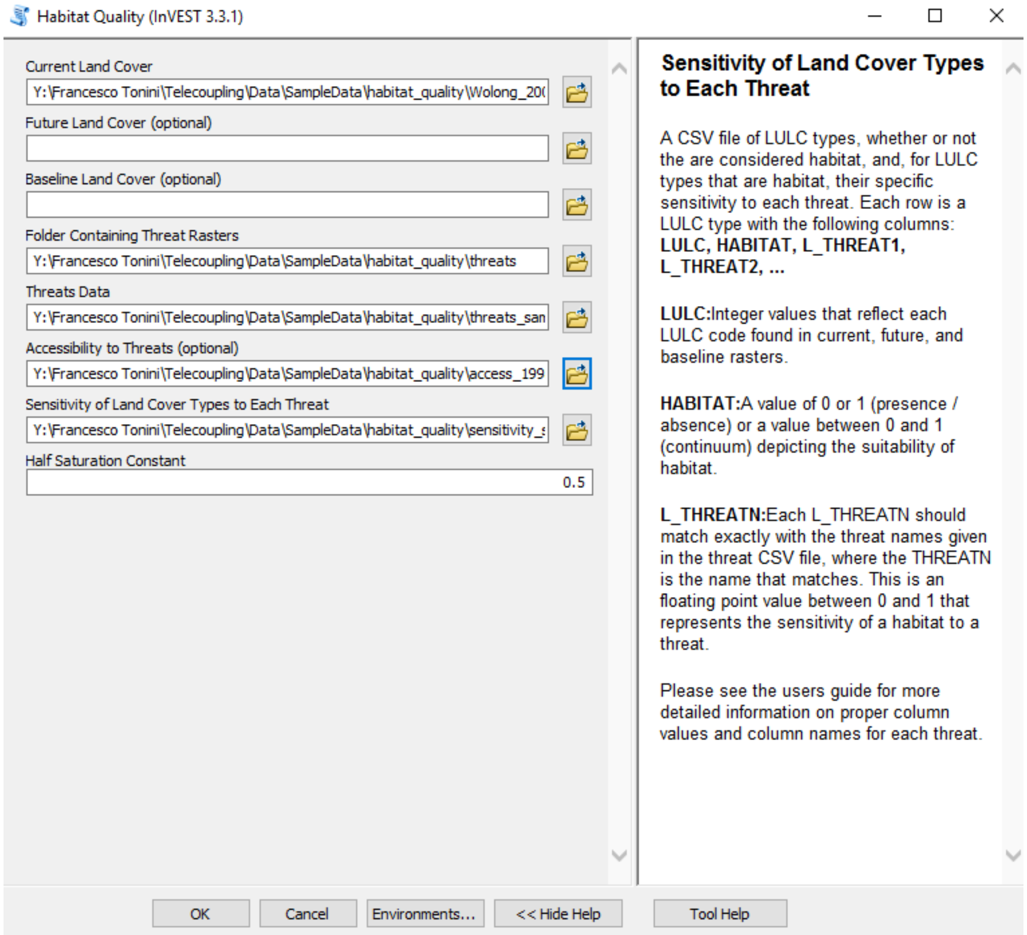


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

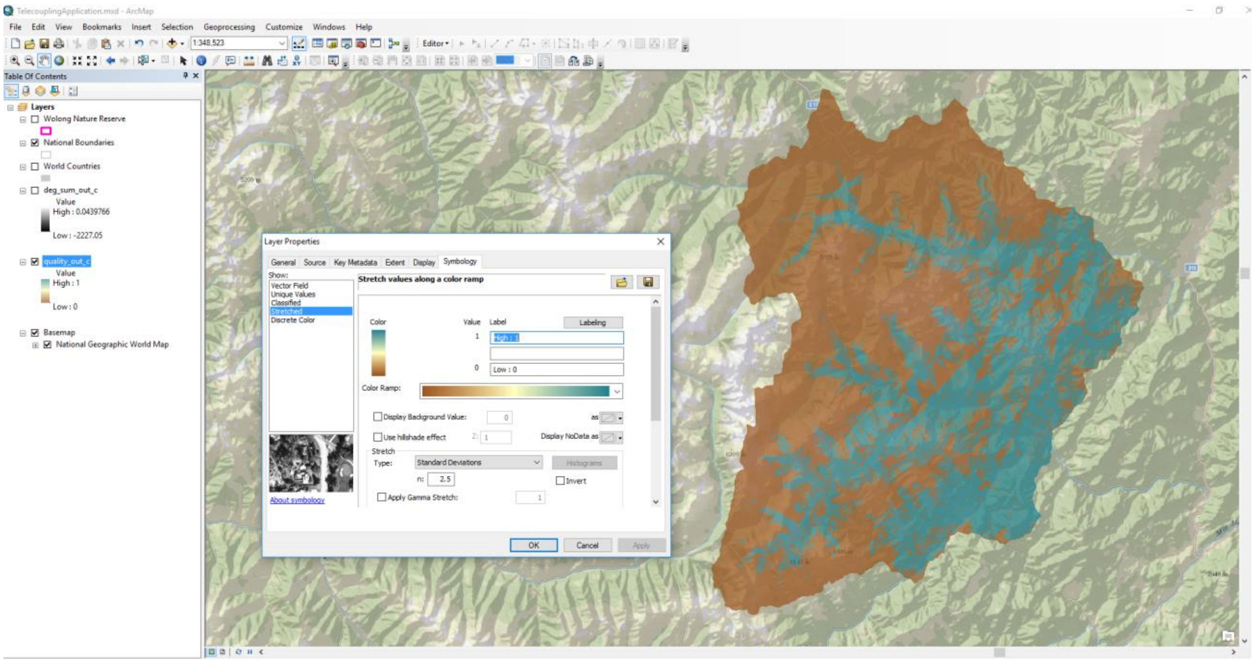
This tools is linked to the InVEST 3.3.3 habitat quality model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this tutorial, we will use the Wolong Nature reserve as a study area to investigate habitat quality for giant pandas. The current land cover parameter needs a raster layer. Because giant panda's habitat is mostly forested areas, we will use a simple land cover layer with code equal to 1 for forest and 0 for any other land cover class. Adjust this according to your species of interest. Select

"*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Habitat Quality/Wolong_2001_forest.img*" as input land cover. Leave future and baseline land cover parameters empty in this tutorial but we recommend to explore these options as well. Specify the folder containing raster layers of each threat to panda's habitat ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Habitat Quality/threats*"). As threats data, select a table that contains all threats you want the model to consider. The table contains information on the each threat's relative importance or weight and its impact across space ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Habitat Quality/threats_samp.csv*"). In our case, we are going to use cropland, roads, and houses as a threat to panda habitat. As accessibility to threats, we are going to select a shapefile that defines zoning in Wolong Nature reserve in 1998, defining core, buffer, and experimental areas for development ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Habitat Quality/access_1998.shp*"). To define the relative sensitivity of each habitat type to each threat, select the table "*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/Habitat Quality/sensitivity_samp.csv*". Leave the half saturation constant with the default value.

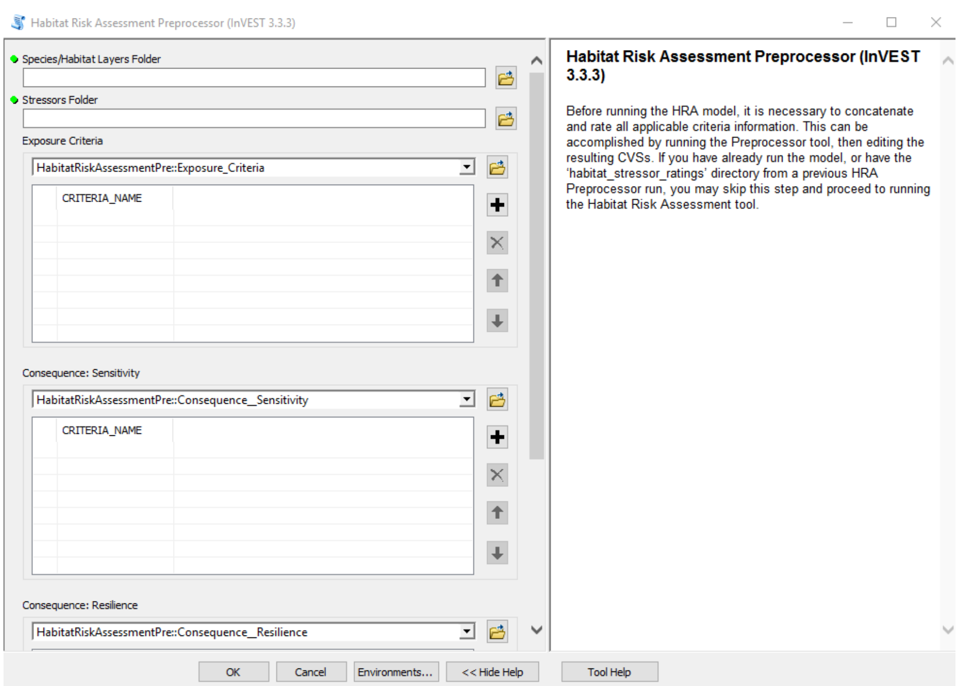
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



Click OK to run the tool. Once completed, you should have an output raster with estimated habitat quality on the current landscape, as well as a raster of relative levels of habitat degradation. Higher numbers for habitat quality indicate better habitat quality vis-a-vis the distribution of habitat quality across the rest of the landscape. Areas on the landscape that are not habitat get a quality score of 0. On the other hand, relative level of habitat degradation get a high score when habitat degradation in a cell is high relative to other cells. The following image shows the habitat quality output, with a different color ramp (right-click raster layer, go to Properties and then change the symbology color ramp if you wish):

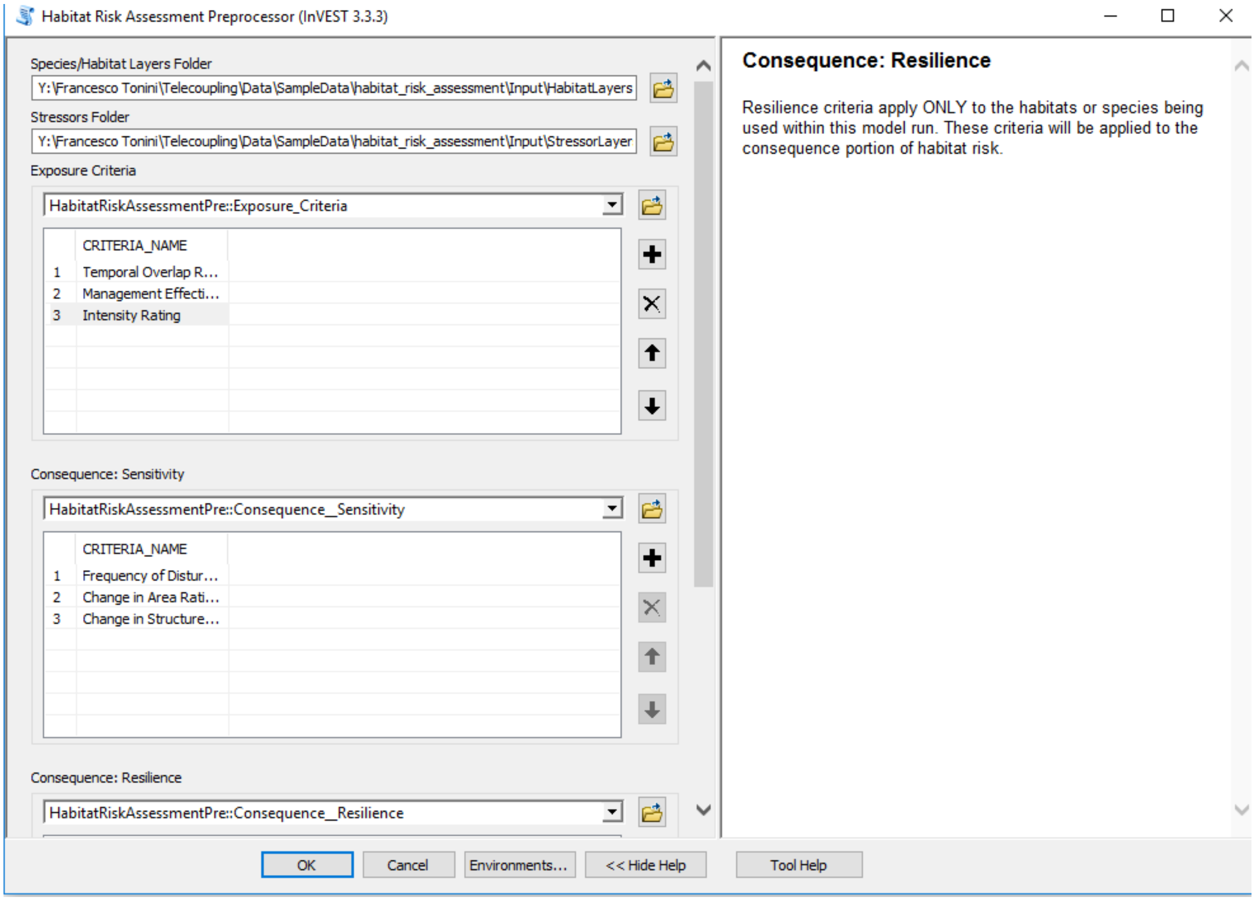


4.6.5 Habitat Risk Assessment Preprocessor (InVEST 3.3.3)

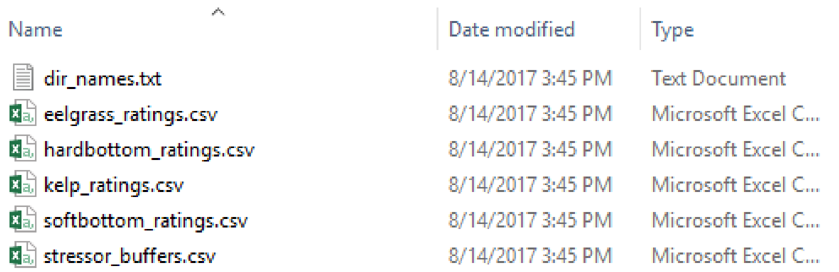


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

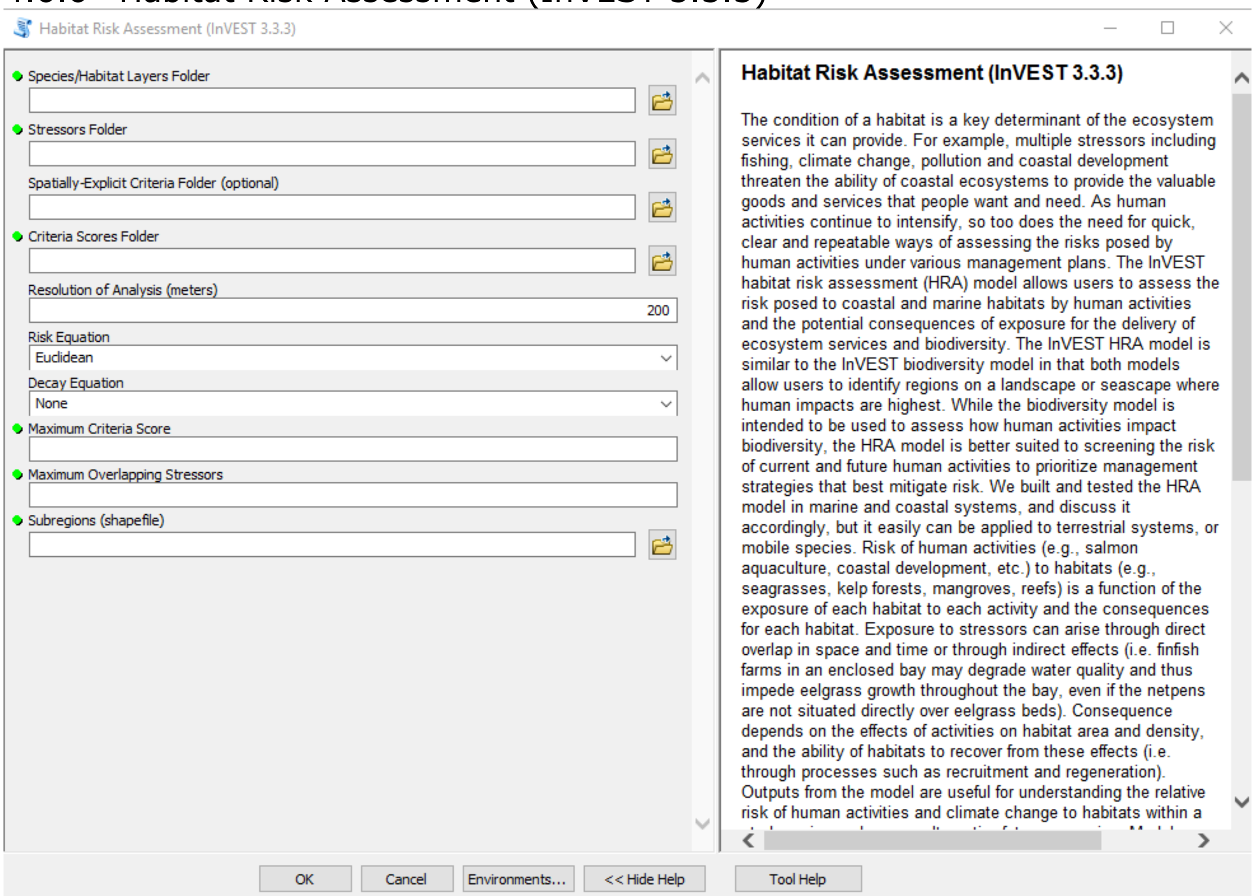
This tools is linked to the InVEST 3.3.3 habitat risk assessment preprocessor model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this tutorial, we will use the sample data provided by the InVEST development team. This tools/model helps you generate a spreadsheet-like file with a space to fill out all the rating scores that you want to assign to each human-related stressor on each habitat. If you feel comfortable using the sample ratings file provided with the sample data and customize it, you can skip this preprocessor tool tutorial and simply follow the habitat risk assessment tutorial in the next section. The species/habitat layer folder parameter needs you to specify a folder on disk containing a shapefile for each habitat/species of interest. Select the folder inside the sample data folder (*./habitat_risk_assessmen/Input/HabitatLayers*) and do the same for the stressors folder parameter, which contains a shapefile for each human stressor affecting one or more of the habitat/species layers specified above (*./habitat_risk_assessmen/Input/StressorLayers*). The following three parameters are a list of textual information on what exposure, consequence (sensitivity, resilience) criteria you would like to account for in your study. For exposure, click on the + sign and type in “Temporal Overlap Rating”, “Management Effectiveness”, “Intensity Rating”. For consequence (sensitivity), select “Frequency of Disturbance”, “Change in Area Rating”, “Change in Structure Rating”. For consequence (resilience), select “Recruitment Rate”, “Natural Mortality Rate”, “Connectivity Rate”, “Recovery Time”. Leave the spatially-explicit parameter section unchecked. You can use this (data is also provided in the sample data folder) on your own if you want to generate a scoring table that accounts for spatially-varying stressors on some (or all) of the habitat/species layers provided.



Click OK to run the tool. Once completed, you should have an output folder containing a txt file ("dir_names.txt") and a csv file for each habitat/species provided as input. Double-click one of the files to explore its content and notice how the criteria you chose have automatically been placed in the appropriate slots inside the spreadsheet. At this stage, you have to fill out each cell combination with a scoring value that relates to the exposure level, or consequence for each habitat/species-stressor combination. Please refer to the official InVEST model online guide (http://data.naturalcapitalproject.org/nightly-build/invest-users-guide/html/habitat_risk_assessment.html) for examples of criteria scoring.

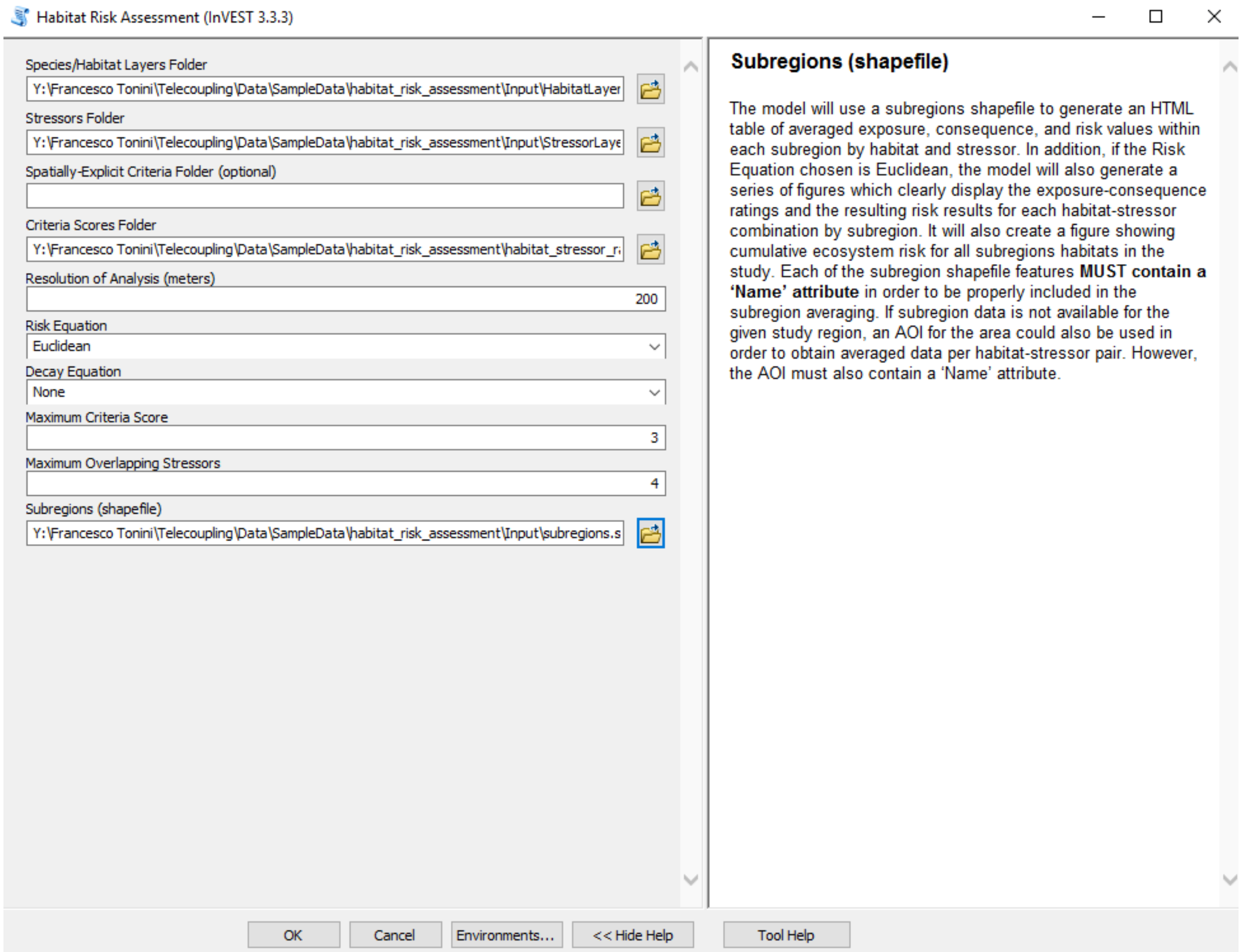


4.6.6 Habitat Risk Assessment (InVEST 3.3.3)

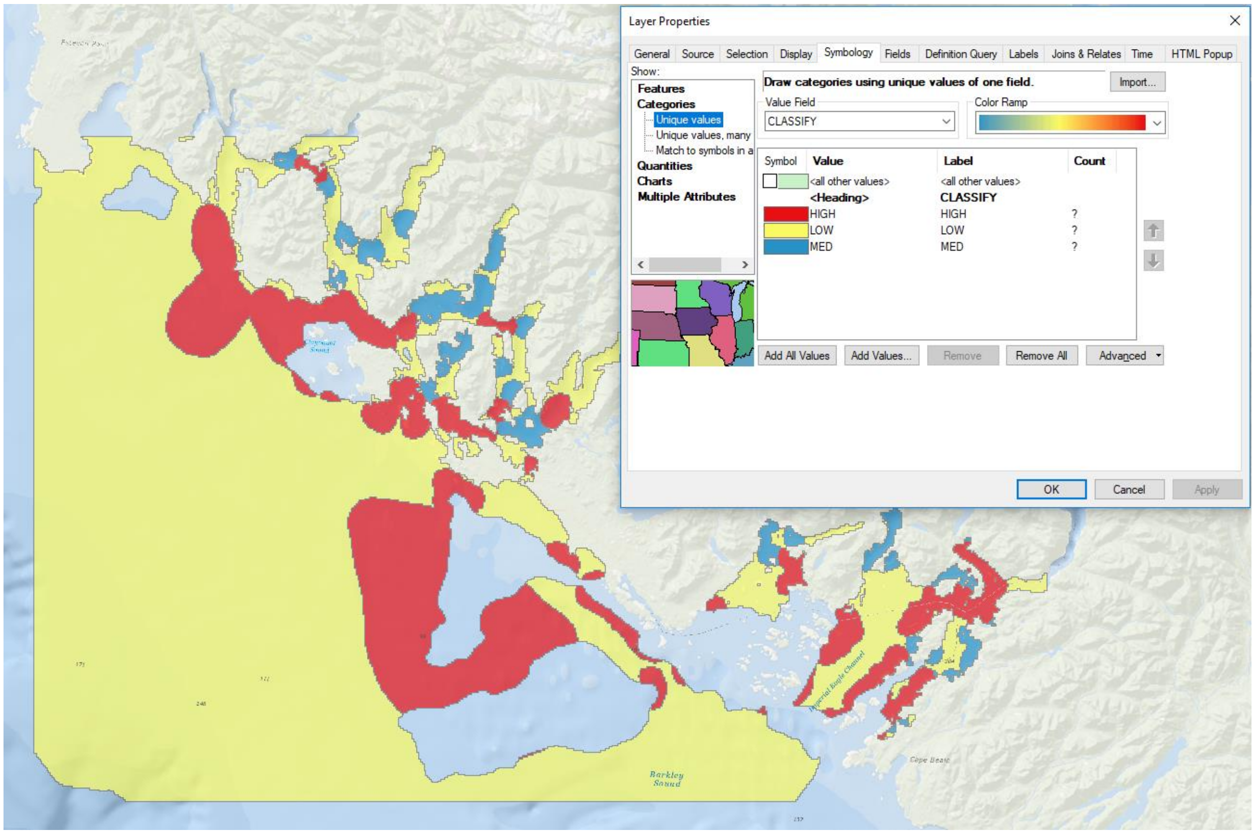


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

This tools is linked to the InVEST 3.3.3 habitat risk assessment model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this tutorial, we will use the sample data provided by the InVEST development team. This tools/model will generate a series of risk shapefiles for each habitat/species provided as input, as well as a zipfile with a number of HTML plots and tables. This tutorial assumes you either ran the preprocessor tool to produce scoring criteria for each habitat-stressor pair or that you took the sample data provided and customized the criteria and scores on your own to be used in this tool. The species/habitat layer folder parameter needs you to specify a folder on disk containing a shapefile for each habitat/species of interest. Select the folder inside the sample data folder (*./habitat_risk_assessmen/Input/HabitatLayers*) and do the same for the stressors folder parameter, which contains a shapefile for each human stressor affecting one or more of the habitat/species layers specified above (*./habitat_risk_assessmen/Input/StressorLayers*). As a criteria scores folder parameter, select the folder that was generated as output in the preprocessor tool. If you have skipped the preprocessor tutorial, then select the folder provided inside the sample data (*./habitat_risk_assessment/habitat_stressor_ratings_sample).* Choose the resolution of the analysis that will be used internally by the model to generate raster files for habitat-stressor risks. Pick a resolution that makes sense for your analysis. In this case, we will choose 200 meters. Leave default values for both risk and decay equation parameters (see official InVEST model guide for more details on these parameters and how they affect the final results, http://data.naturalcapitalproject.org/nightly-build/invest-users-guide/html/habitat_risk_assessment.html). For maximum criteria score type the number 3, and for maximum overlapping stressor type the number 4. The maximum criteria score MUST reflect the scoring system used in your criteria scoring input files. Similarly, the maximum overlapping stressor number MUST equal the number of stressor shapefiles chosen for the analysis. Finally, select a shapefile that represent one or more management areas (subregions) for the case study (*./habitat_risk_assessment/Input/subregions.shp*).

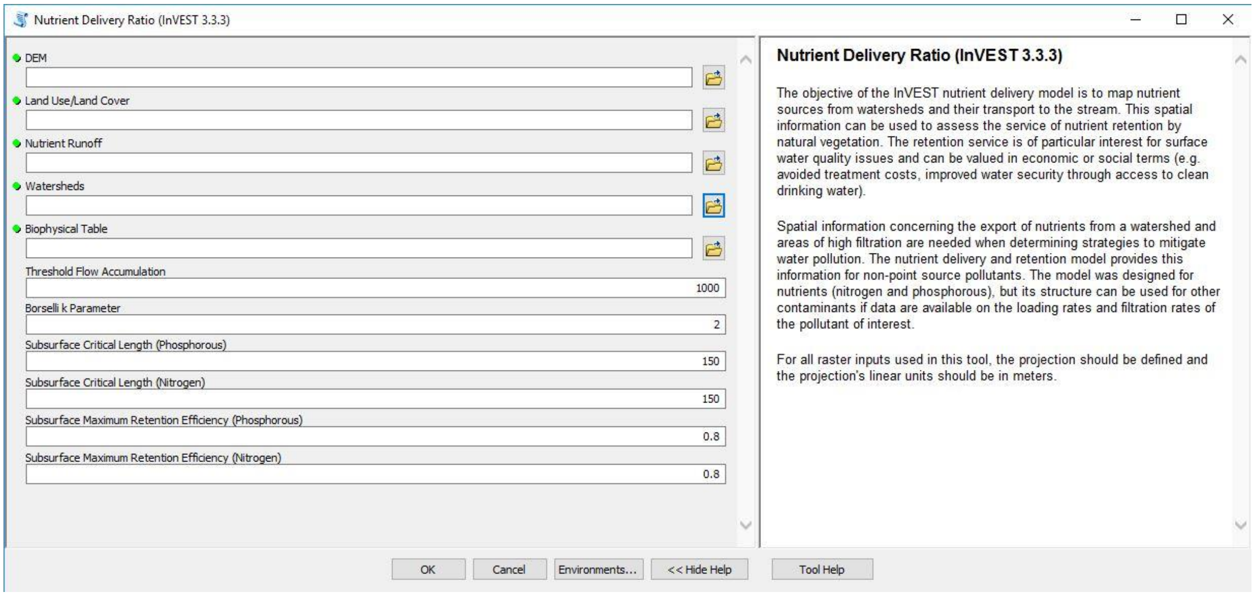


Click OK to run the tool. Once completed, you should have a list of shapefiles with estimated risk for each habitat/species layer provided as input, plus a zipped file with HTML plots and tables. The risk in each shapefile has been classified into LOW, MEDIUM, HIGH internally by the model. The following image shows the risk output, with a chosen qualitative color ramp (right-click shapefile layer, go to Properties and then change the symbology color ramp if you wish):



Layer Properties
General
Source
Selection
Display
Symbology
Fields
Definition Query
Labels
Joins & Relates
Time
HTML Popup
Show:
Features
Categories
Unique values
Unique values, many
Match to symbols in a
Quantities
Charts
Multiple Attributes
Draw categories using unique values of one field.
Import...
Value Field
CLASSIFY
Color Ramp
Symbol
Value
Label
Count
<all other values>
<all other values>
<Heading>
CLASSIFY
HIGH
HIGH
?
LOW
LOW
?
MED
MED
?
Add All Values
Add Values...
Remove
Remove All
Advanced
OK
Cancel
Apply
Barkley Sound

4.6.7. Nutrient Delivery Ratio (InVEST 3.3.3)



The above window will open when you double-click the Nutrient Delivery Ratio tool. Helpful instructions in terms of using the tool are provided by clicking on the "Tool Help" button.

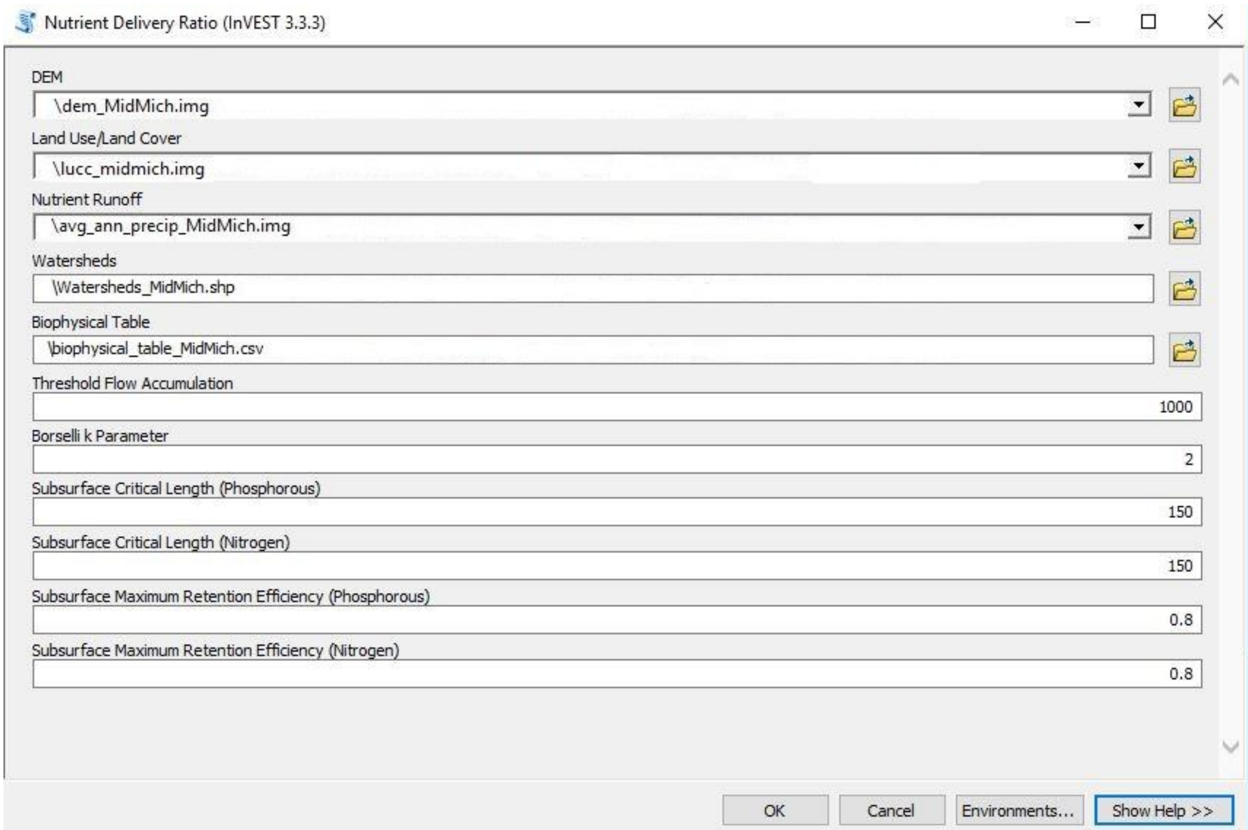
This tools is linked to the InVEST 3.3.3 Nutrient Delivery Ratio model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then returns the results produced back to the main application. In this tutorial, we will use sample data specific to the middle of Michigan in the United States. This tool/model will generate two rasters, one showing how much nitrogen load from each pixel eventually reaches the stream (in kg/pixel), the other showing the same information for phosphorous. The tool also produces a shapefile that aggregates the nutrient model results per watershed for both nitrogen and phosphorous. Opening the attribute table of this shapefile returns information on:

- *n(or p)_load_tot:* kg.yr^{-1}: total nutrient loads (sources) in the watershed, i.e., the sum of the nutrient contribution from all LULC without filtering from the landscape.
- *n(or p)_exp_tot:* kg.yr^{-1}: total nutrient export from the watershed.

For the tool's DEM, select the DEM for Mid-Michigan *("./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/NDR/dem_MidMich.img").* The Land Use/Land Cover should be a land cover image for this same location in Mid-Michigan *("./SampleData_ArcGIS_v1.7.3b/Environmental*

Analysis/NDR/lucc_midmich.img”). Nutrient Runoff is a GIS raster dataset that represents spatial variability of runoff potential. We use a raster of annual precipitation for this model parameter *(“./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/NDR/avg_ann_precip_MidMich.img”).* For the watershed polygons, select the following sample data: *“./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/NDR/Watersheds_MidMich.shp”*. The biophysical table includes coefficients for each LULC class for the Mid-Michigan region. Use the following .csv file: *“./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/NDR/biophysical_table_MidMich.csv”.* The remaining parameters should be kept at their default values.

Your NDR tool should ultimately be populated as follows:



As mentioned above, the tool’s output will include three files. Here we show the aggregated shapefile with its attribute table opened to display the final results.

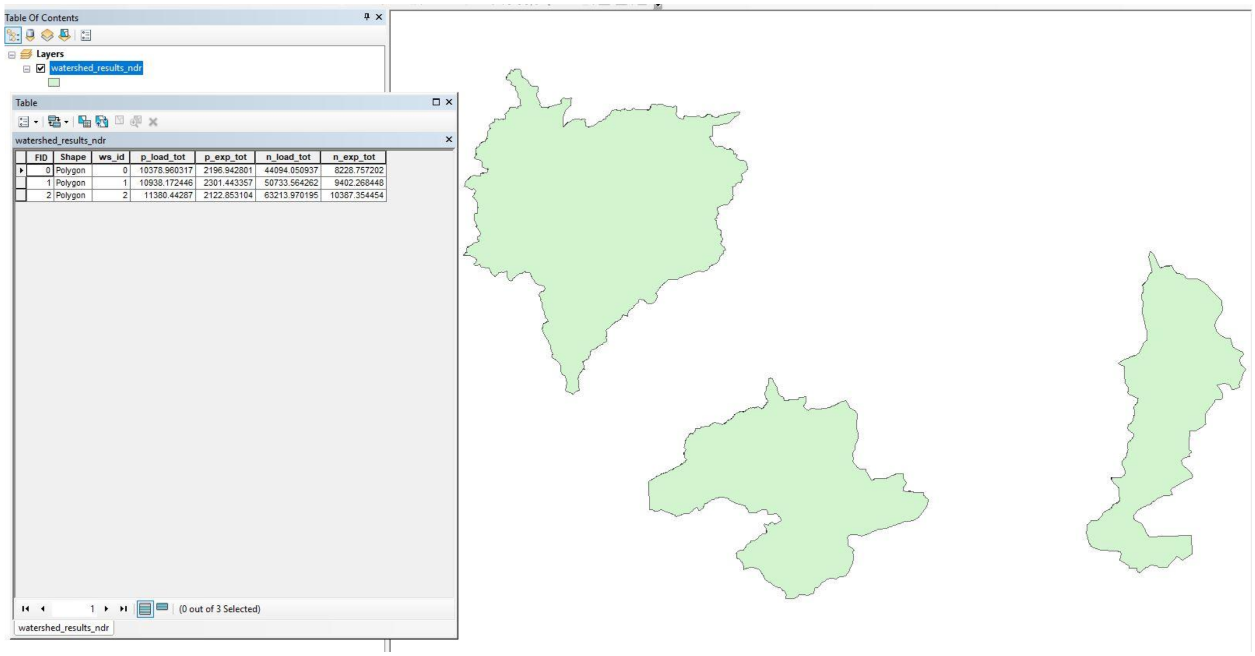
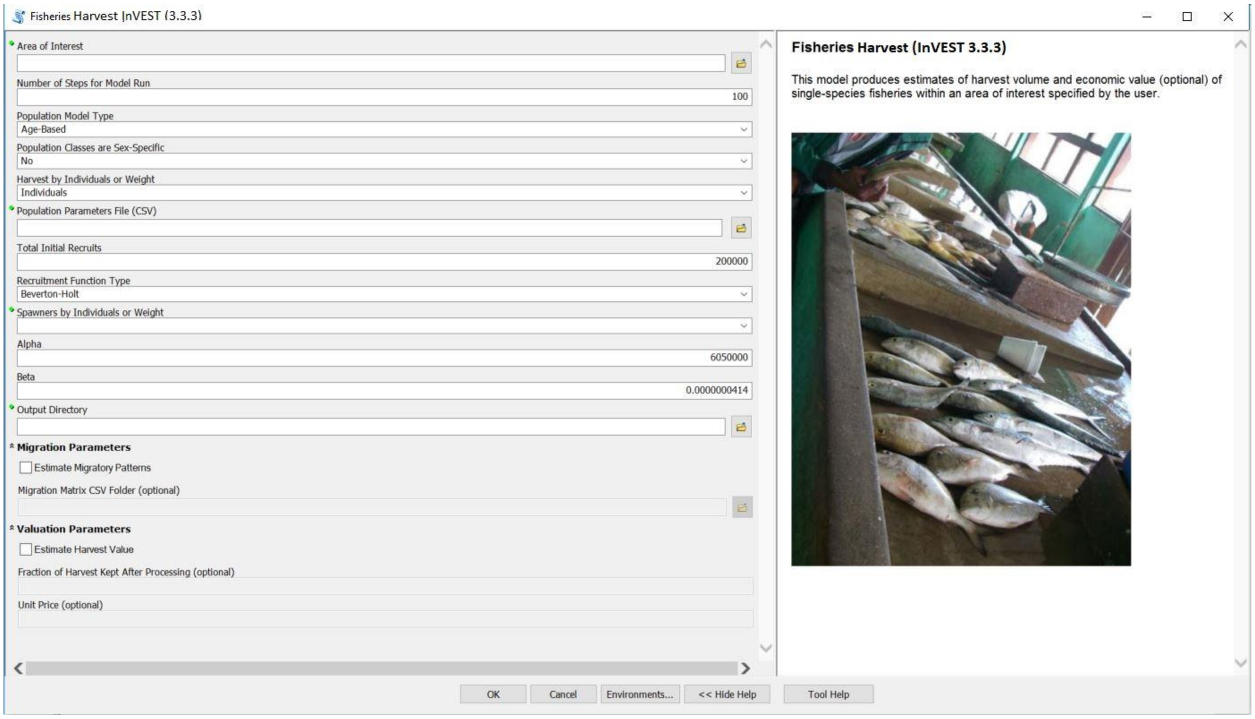


Table Of Contents
Layers
watershed_results_ndr
Table
watershed_results_ndr
FID Shape ws_id p_load_tot p_exp_tot n_load_tot n_exp_tot
0 Polygon 0 10378.960317 2196.942801 44094.050937 8228.757202
1 Polygon 1 10938.172446 2301.443357 50733.564262 9402.268448
2 Polygon 2 11380.44287 2122.853104 63213.970195 10387.354454
(0 out of 3 Selected)
watershed_results_ndr

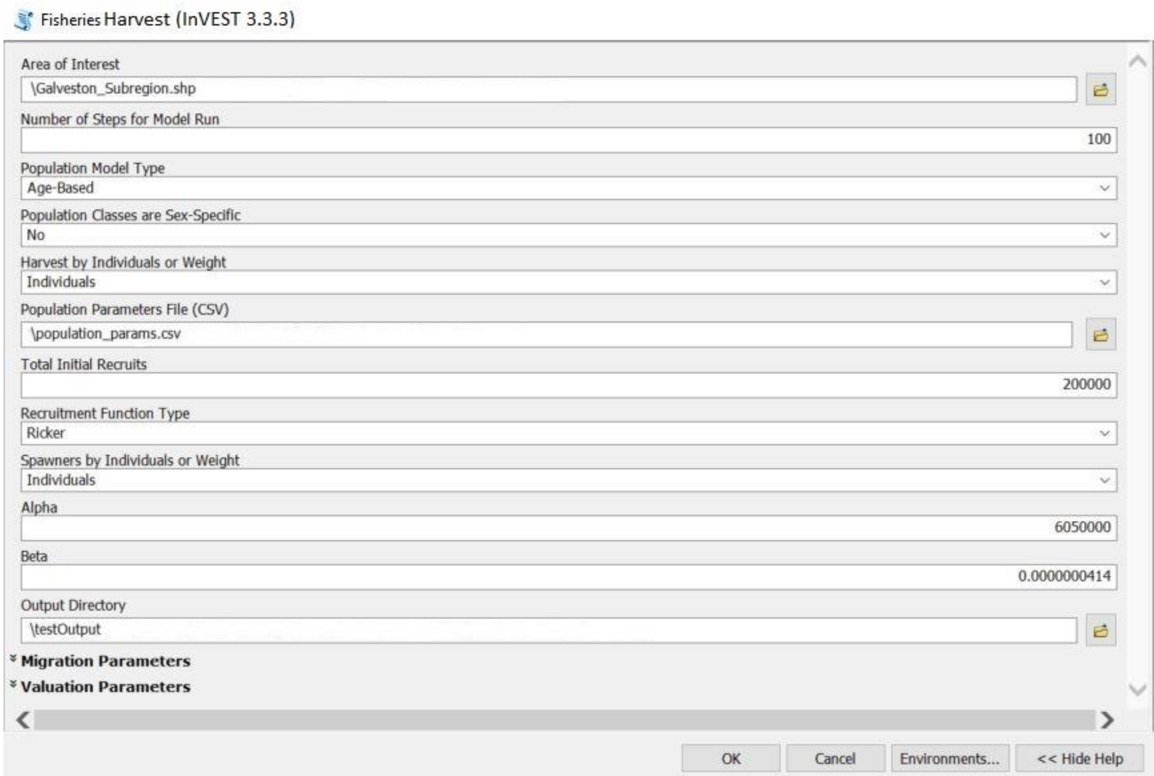
4.6.8 Fisheries Harvest (InVEST 3.3.3)



Once you open the tool, be sure to refer to the help window that appears on the right of the tool interface. The help is there to provide general information about the tool as a whole, as well as to describe in detail each single parameter after clicking on it.

This tool is linked to the InVEST 3.3.3 Fisheries model. All input parameters that are specified in the ArcGIS tool interface will be sent to the appropriate InVEST model which then returns the results back to the main application. In this tutorial, we use sample data that were originally provided by InVEST. The data relates to blue crab in Galveston Bay, Texas. The tool will generate a shapefile of the area of interest with harvest volume, as well as a CSV file summarizing the results.

For the shapefile specifying the area of interest, locate the following file in the Sample Data folder: *“./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/FisheriesHarvest/Galveston_Subregion.shp”.* The following file should be used for the Population Parameters CSV File: *“./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/FisheriesHarvest/population_params.csv”*. The output directory should be the location where you would like the output saved. All other parameters should be entered as follows:



The final output will include a CSV file summarizing the results. This file will be saved to the Output Directory specified by the user. The final output will also include a shapefile of the AOI with information related to total harvest volume. This shapefile will also be saved to the Output Directory and will automatically be loaded to the map viewer. The shapefile for the tutorial is shown here:

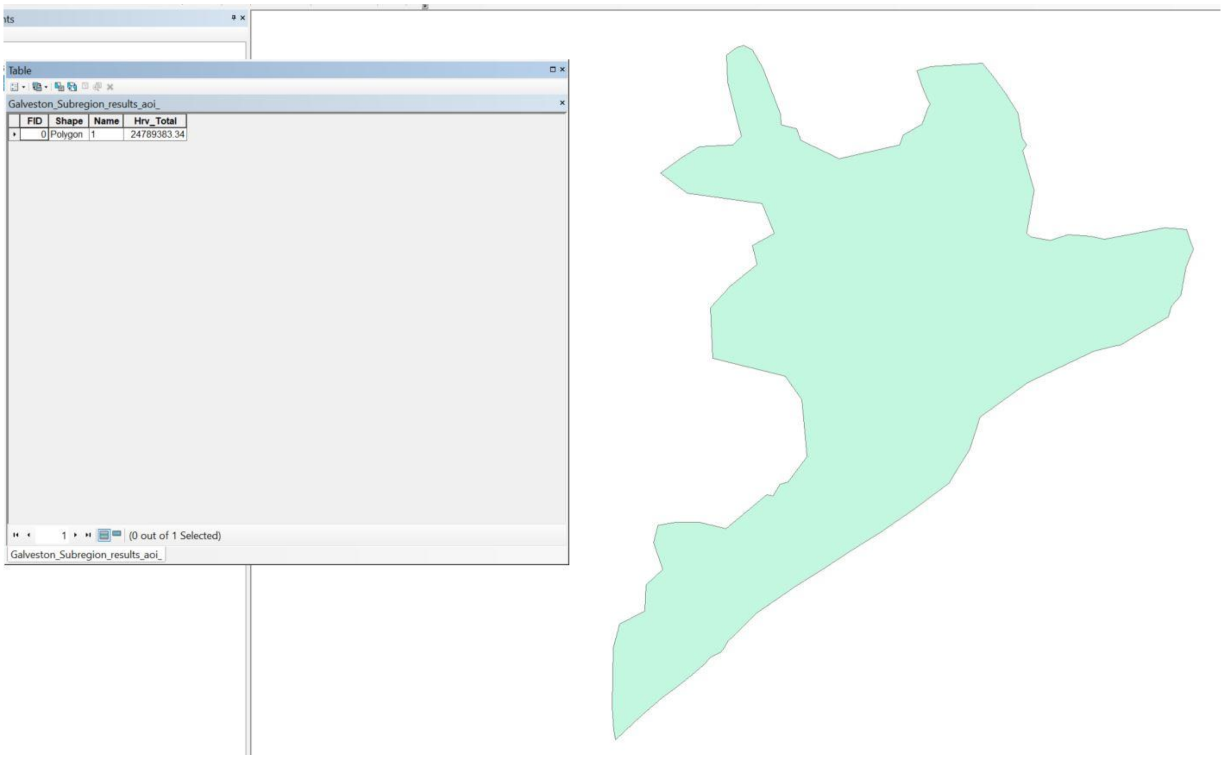
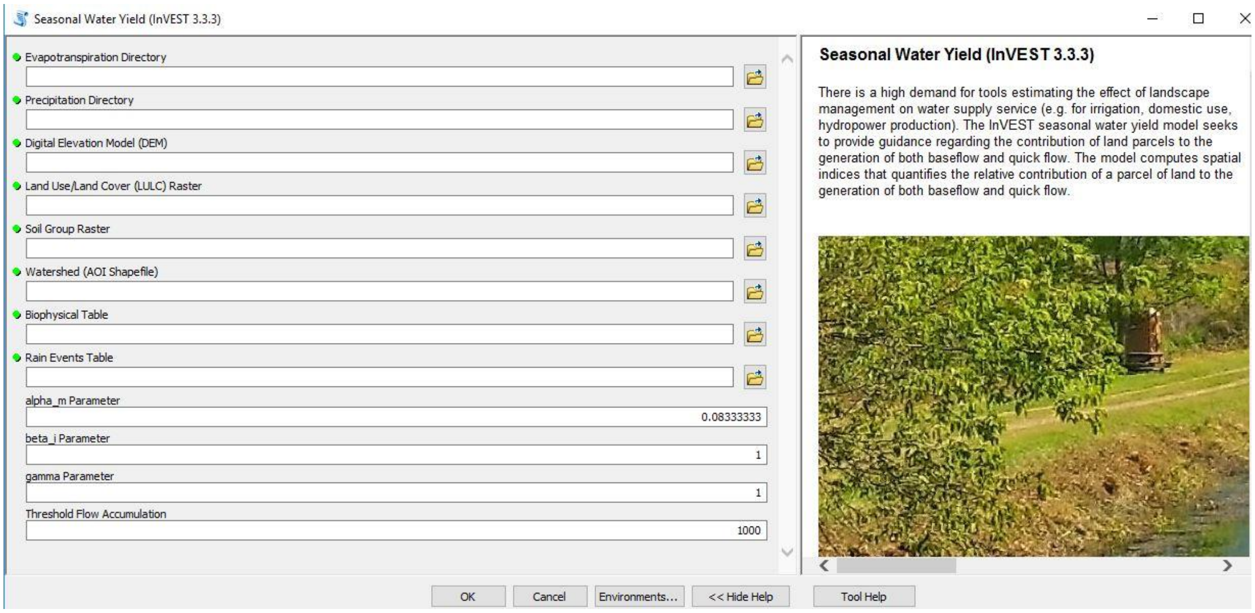


Table
Galveston_Subregion_results_aoi_
FID
Shape
Name
Hrv_Total
0
Polygon
1
24789383.34
(0 out of 1 Selected)
Galveston_Subregion_results_aoi_

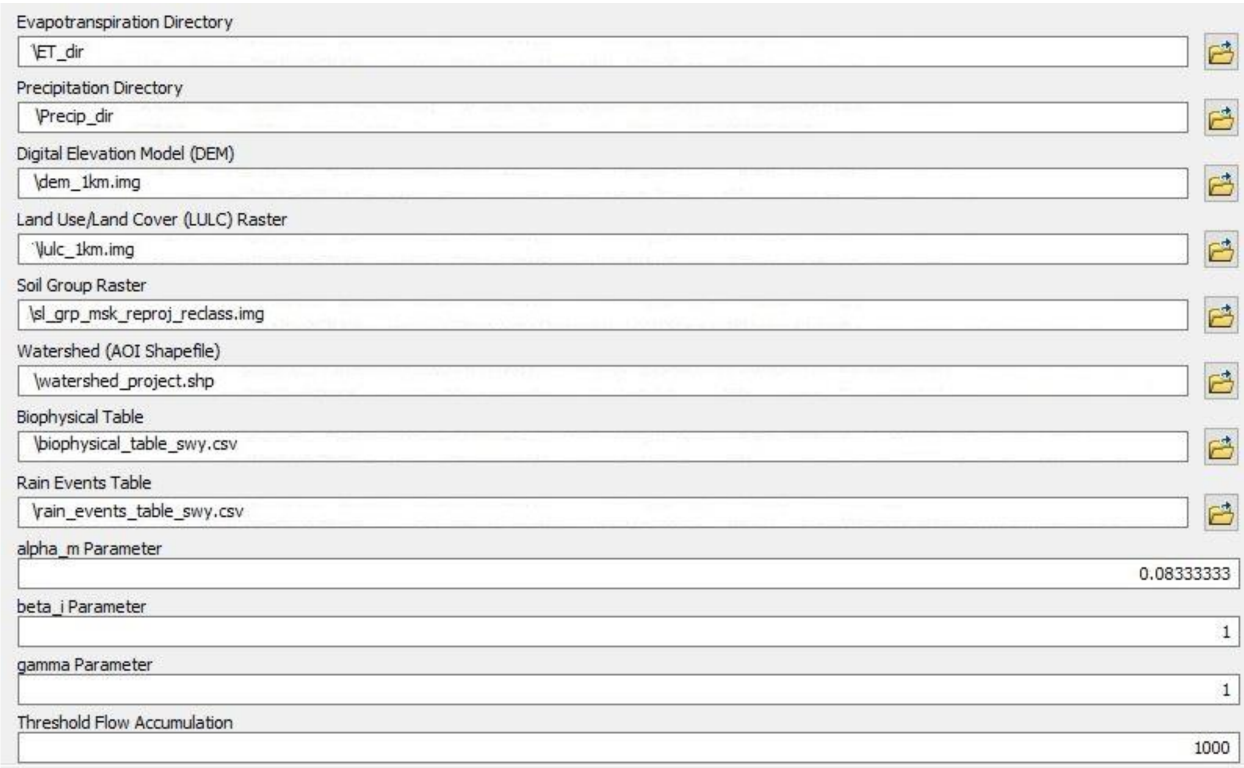
4.6.9 Seasonal Water Yield (InVEST 3.3.3)



This tools is linked to the InVEST 3.3.3 Seasonal Water Yield model. All input parameters that are specified in the ArcGIS tool interface will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this tutorial, we will focus on a watershed that spans the Brazilian states of Minas Gerais and Espirito Santo before ultimately emptying into the Atlantic Ocean. The output will return spatially-explicit information of quickflow, baseflow, and local recharge, among others, as they relate to the watershed of interest.

When you open the Seasonal Water Yield tool, the interface should appear as the image above. For the Evapotranspiration Directory select the following folder: *"./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/SeasonalWaterYield/ET_dir"*. Similarly, the Precipitation Directory should be specified as follows: *"./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/SeasonalWaterYield/Precip_dir"*. Next, specify the following rasters: a Digital Elevation Model (DEM) raster ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/SeasonalWaterYield/dem_1km.img"*), a LULC raster ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/SeasonalWaterYield/lulc_1km.img"*), and a Soil Group raster ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/SeasonalWaterYield/sl_grp_msk_reproj_reclass.img"*). The Watershed AOI shapefile should be specified next ("*./SampleData_ArcGIS_v1.7.3b/Environmental*

Analysis/SeasonalWaterYield/watershed_project.shp"). Finally, specify both the Biophysical Table ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/SeasonalWaterYield/biophysical_table_swy.csv"*) as well as the Rain Events Table ("*./SampleData_ArcGIS_v1.7.3b/Environmental Analysis/SeasonalWaterYield/rain_events_table_swy.csv"*). Keep the default values for the alpha, beta, gamma, and threshold flow accumulation parameters. Your populated interface should be as follows:



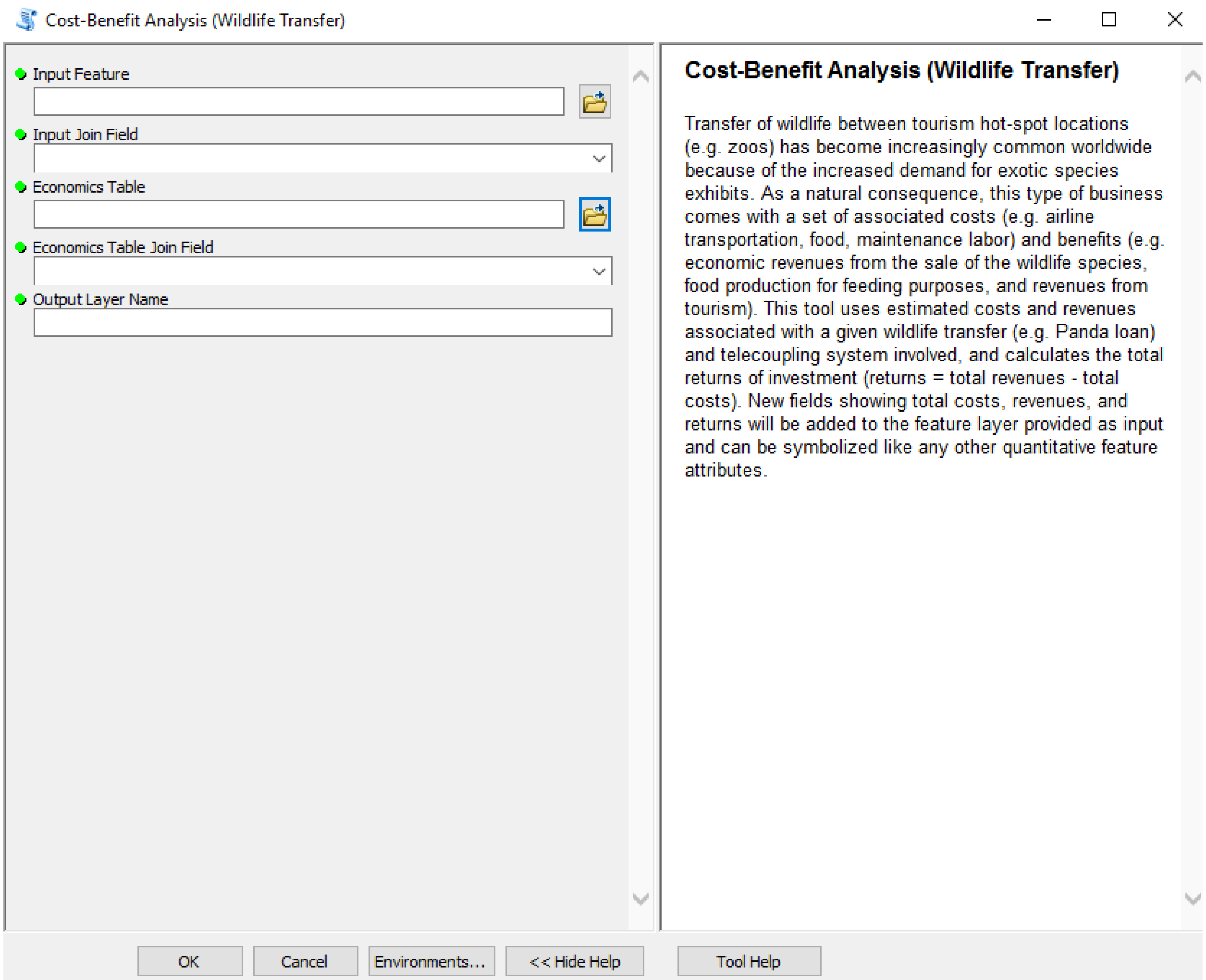
Click OK to run the tool, and ArcGIS will return nine rasters and one shapefile. The output are as follows:

CN.tif	Raster of curve number values
QF.tif	Raster of quickflow values (mm)
L.tif	Raster of local recharge values (mm)
L_avail.tif	Raster of available local recharge (mm)
B.tif	Raster of baseflow values (mm); the contribution of a pixel to slow release flow
B_sum.tif	Raster of the sum of baseflow values (mm) contributed by all upslope pixels

L_sum.tif	Raster of the sum of local recharge values (mm) contributed by all upslope pixels
L_sum_avail.tif	Raster of *Lsum* values (mm) – the available water to a pixel – contributed by all upslope pixels
Vri.tif	Raster of the values of recharge contribution to the total recharge
aggregated_results.shp	Shapefile specifying the Q_b value (ie, the annual average baseflow in millimeters)

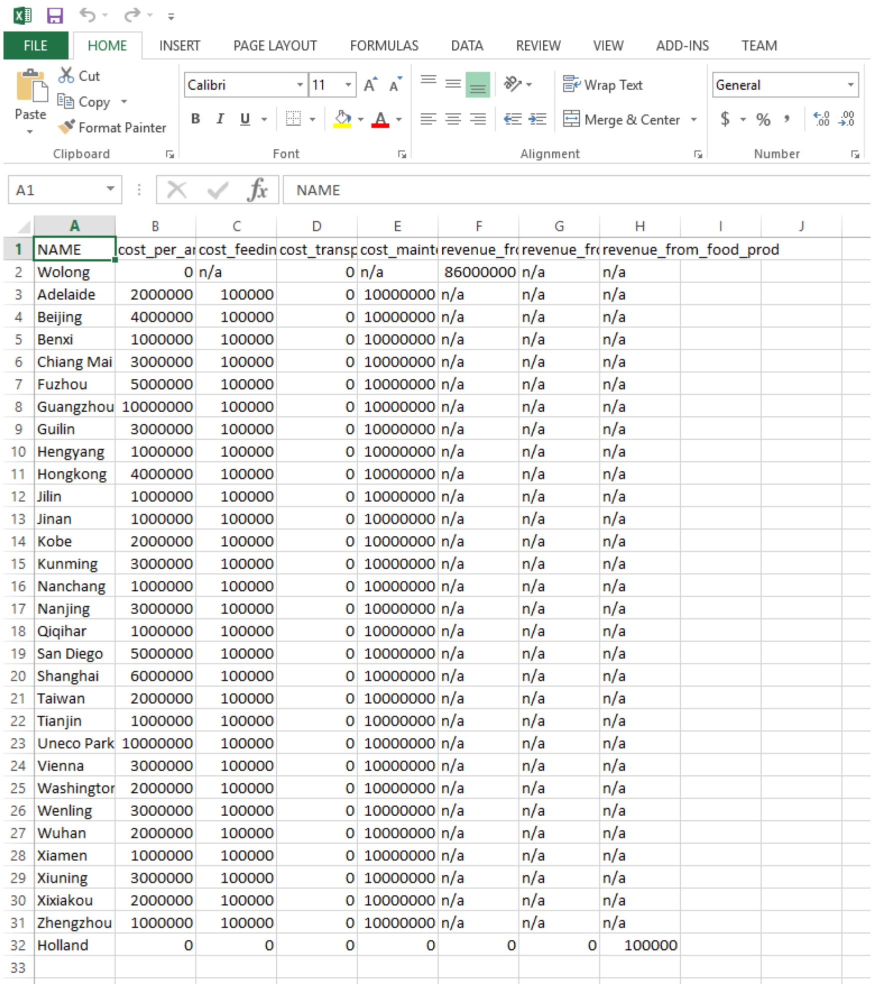
4.7 SOCIOECONOMIC ANALYSIS TOOLSET

4.7.1 Cost-Benefit Analysis (Wildlife Transfer)



Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some cases, the tool help windows will have a nice graphic with it.

This tool can be used whenever you have some data on costs and revenues involved with different aspects of a wildlife transfer. For example, the transfer of giant pandas from the Wolong Nature reserve to other location worldwide involves cost of transportation, maintenance, feeding material, laborers, as well as revenues from the transaction or from tourism-related activities increased (ideally) after such a transfer just to name some. If you do not have data with respect to one or more of the costs/revenues factors, you might want to collect them or simply use what you have and type "n/a" or leave blank in the appropriate input table fields. In this tutorial, we will use the financial table found at "*./SampleData_ArcGIS_v1.7.3b/Socioeconomic Analysis/Cost-Benefit Analysis/wildlife_financial_data.csv*":



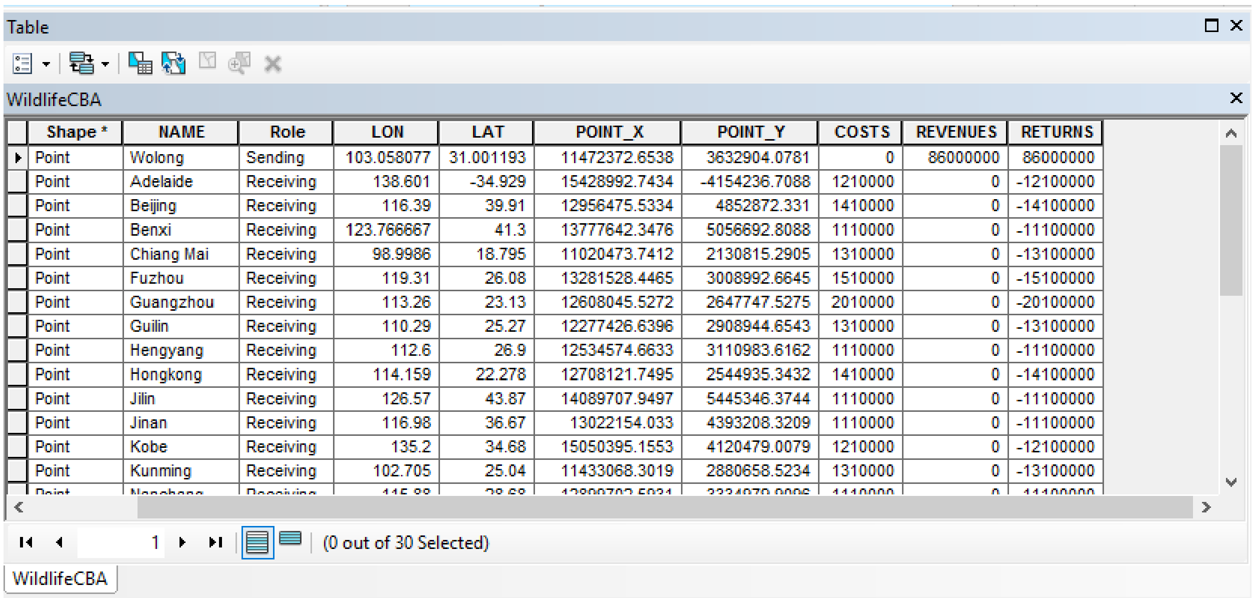
NAME	cost_per_ar	cost_feedin	cost_transp	cost_maint	revenue_fr	revenue_fr	revenue_from_food_prod
Wolong	0	n/a	0	n/a	86000000	n/a	n/a
Adelaide	2000000	100000	0	10000000	n/a	n/a	n/a
Beijing	4000000	100000	0	10000000	n/a	n/a	n/a
Benxi	1000000	100000	0	10000000	n/a	n/a	n/a
Chiang Mai	3000000	100000	0	10000000	n/a	n/a	n/a
Fuzhou	5000000	100000	0	10000000	n/a	n/a	n/a
Guangzhou	10000000	100000	0	10000000	n/a	n/a	n/a
Guilin	3000000	100000	0	10000000	n/a	n/a	n/a
Hengyang	1000000	100000	0	10000000	n/a	n/a	n/a
Hongkong	4000000	100000	0	10000000	n/a	n/a	n/a
Jilin	1000000	100000	0	10000000	n/a	n/a	n/a
Jinan	1000000	100000	0	10000000	n/a	n/a	n/a
Kobe	2000000	100000	0	10000000	n/a	n/a	n/a
Kunming	3000000	100000	0	10000000	n/a	n/a	n/a
Nanchang	1000000	100000	0	10000000	n/a	n/a	n/a
Nanjing	3000000	100000	0	10000000	n/a	n/a	n/a
Qiqihar	1000000	100000	0	10000000	n/a	n/a	n/a
San Diego	5000000	100000	0	10000000	n/a	n/a	n/a
Shanghai	6000000	100000	0	10000000	n/a	n/a	n/a
Taiwan	2000000	100000	0	10000000	n/a	n/a	n/a
Tianjin	1000000	100000	0	10000000	n/a	n/a	n/a
Uneco Park	10000000	100000	0	10000000	n/a	n/a	n/a
Vienna	3000000	100000	0	10000000	n/a	n/a	n/a
Washingtor	2000000	100000	0	10000000	n/a	n/a	n/a
Wenling	3000000	100000	0	10000000	n/a	n/a	n/a
Wuhan	2000000	100000	0	10000000	n/a	n/a	n/a
Xiamen	1000000	100000	0	10000000	n/a	n/a	n/a
Xiuning	3000000	100000	0	10000000	n/a	n/a	n/a
Xixiakou	2000000	100000	0	10000000	n/a	n/a	n/a
Zhengzhou	1000000	100000	0	10000000	n/a	n/a	n/a
Holland	0	0	0	0	0	0	100000

You can use this table and modify it with your own case study if you prefer. As first tool parameter, specify a feature class that contains all your telecoupling systems. If you have not run any of the Systems tools, make sure to do so before running this tool, as it is a required input. In the second parameter, select an attribute from the input systems feature class whose values will be used to join the financial table with. Select NAME as this attribute contains names of all geographic locations that are found within the financial table name attribute. ***NOTE: if you do NOT have the exact names in both the systems table and the financial table, or if you have a***

different number of records between the two tables, the join operation should still complete successfully with a warning message. Specify the economics table provided with the sample dataset above and use NAME as a join field. You do NOT need to have the same column labels (NAME) like in this case, as long as you specify which attribute in both tables to use the join on.

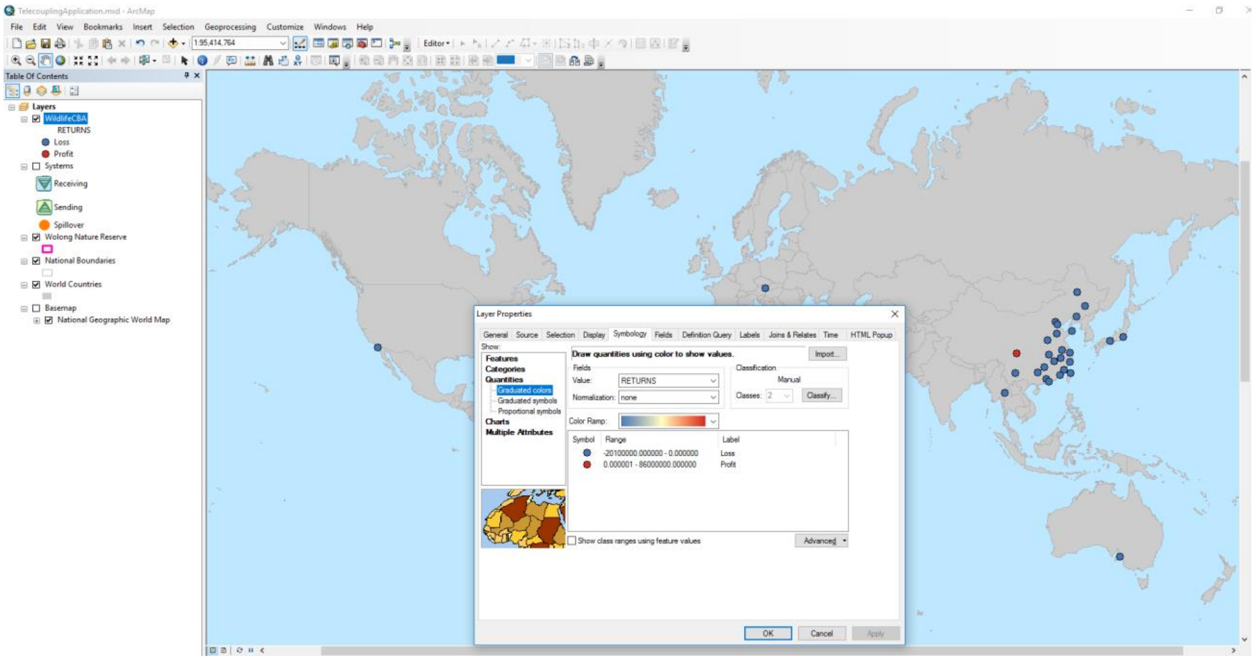
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

Click OK to run the tool and ArcGIS will return a new feature class in output with three new attributes specifying COSTS, REVENUES, and RETURNS.

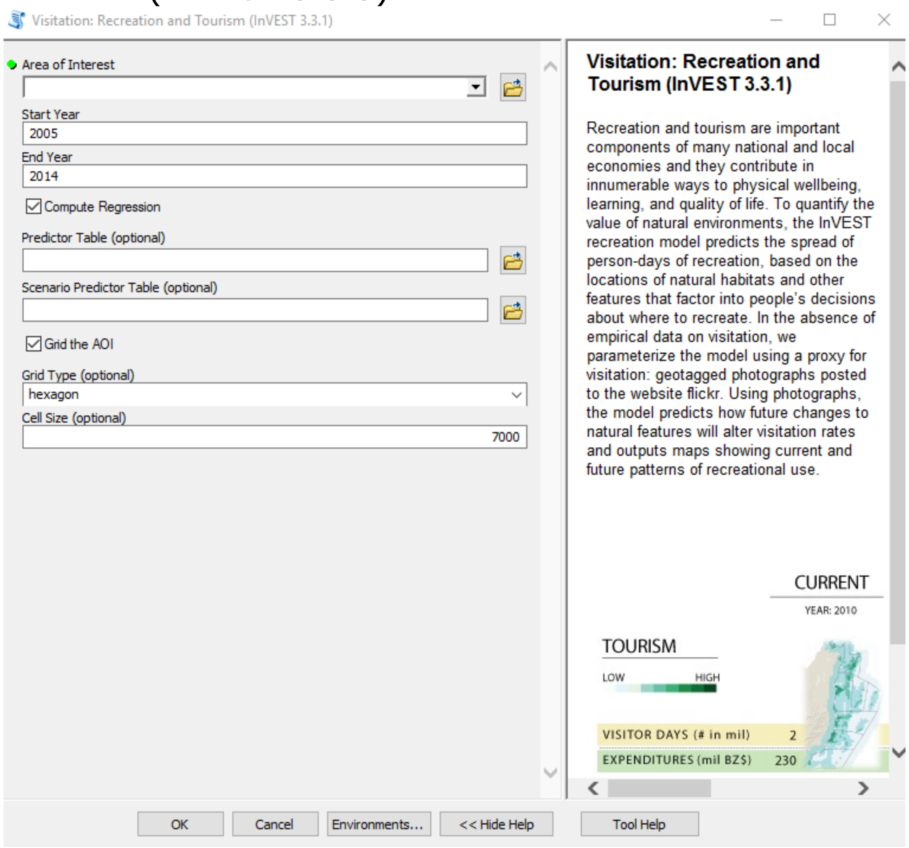


Shape *	NAME	Role	LON	LAT	POINT_X	POINT_Y	COSTS	REVENUES	RETURNS
Point	Wolong	Sending	103.058077	31.001193	11472372.6538	3632904.0781	0	86000000	86000000
Point	Adelaide	Receiving	138.601	-34.929	15428992.7434	-4154236.7088	1210000	0	-12100000
Point	Beijing	Receiving	116.39	39.91	12956475.5334	4852872.331	1410000	0	-14100000
Point	Benxi	Receiving	123.766667	41.3	13777642.3476	5056692.8088	1110000	0	-11100000
Point	Chiang Mai	Receiving	98.9986	18.795	11020473.7412	2130815.2905	1310000	0	-13100000
Point	Fuzhou	Receiving	119.31	26.08	13281528.4465	3008992.6645	1510000	0	-15100000
Point	Guangzhou	Receiving	113.26	23.13	12608045.5272	2647747.5275	2010000	0	-20100000
Point	Guilin	Receiving	110.29	25.27	12277426.6396	2908944.6543	1310000	0	-13100000
Point	Hengyang	Receiving	112.6	26.9	12534574.6633	3110983.6162	1110000	0	-11100000
Point	Hongkong	Receiving	114.159	22.278	12708121.7495	2544935.3432	1410000	0	-14100000
Point	Jilin	Receiving	126.57	43.87	14089707.9497	5445346.3744	1110000	0	-11100000
Point	Jinan	Receiving	116.98	36.67	13022154.033	4393208.3209	1110000	0	-11100000
Point	Kobe	Receiving	135.2	34.68	15050395.1553	4120479.0079	1210000	0	-12100000
Point	Kunming	Receiving	102.705	25.04	11433068.3019	2880658.5234	1310000	0	-13100000

At this stage, let's symbolize the point feature class with colors based on the "RETURNS" attribute, to better see the spatial locations of telecoupling systems that have positive or negative returns as a result of the wildlife transfer. ***NOTE: values that are found as "n/a" or blank in the economics table, will be ignored in the calculation of the returns and you will get a warning message about it***. Right-click the output layer, go under the Properties and select Quantities > graduated colors. Select the RETURNS field and use as many classification bins as you deem appropriate. In this case, we will select only 2 bins and manually modify the range to have negative values (losses) and positive values (gains) in two different colors like the following figure shows:



4.7.2 Visitation: Recreation and Tourism Storage and Sequestration (InVEST 3.3.3)

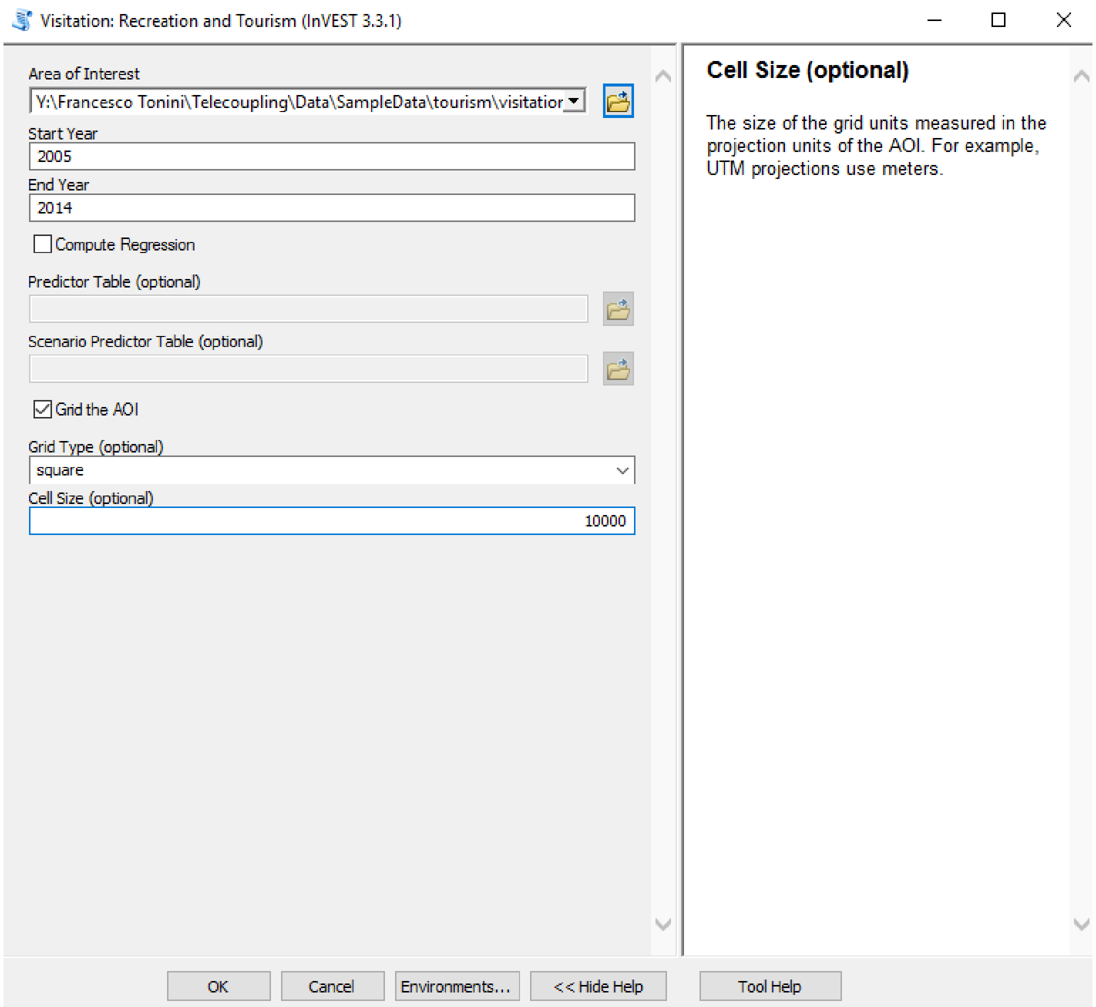


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case, the tool help windows will have a nice graphic with it.

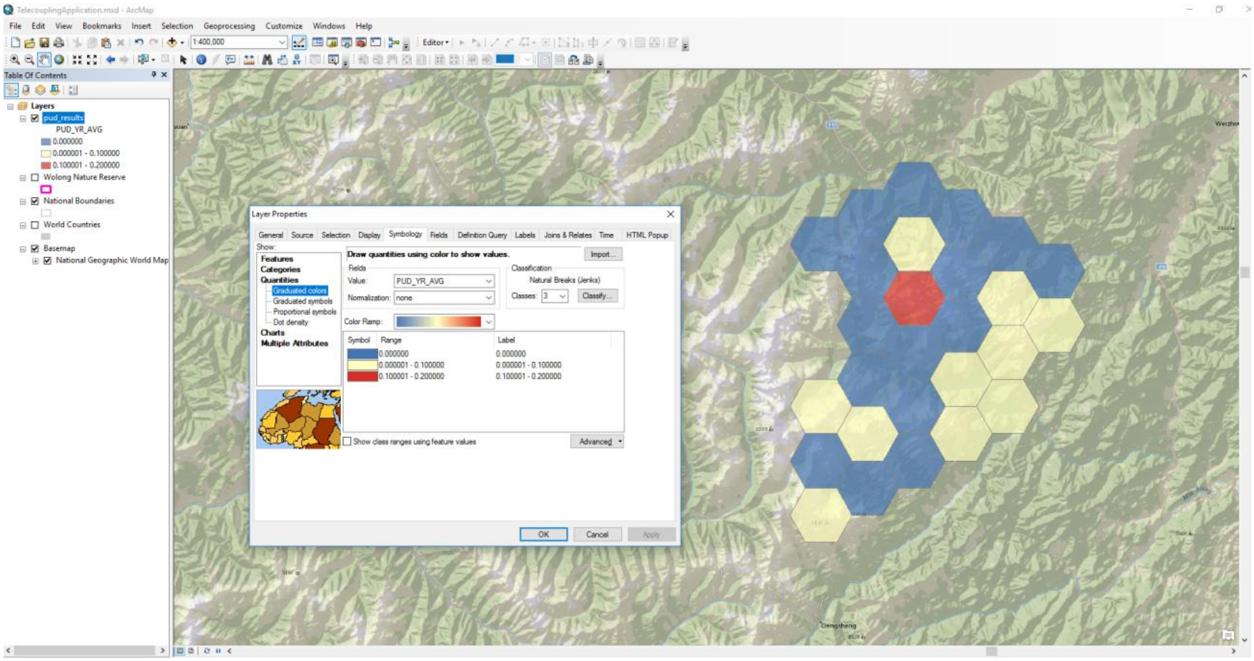
This tools is linked to the InVEST 3.3.3 visitation: recreation and tourism model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this tutorial, we will use the Wolong Nature reserve as a study area to investigate tourist visitation rates from geotagged photographs posted to the website Flickr. ***NOTE: this tool should be used keeping in mind that certain areas may have very few to none geotagged photographs on Flickr. If you have a similar situation, please consider using a different tool to study tourism, e.g. linear regression***.

As area of interest select the shapefile of the Wolong Nature reserve boundaries ("*./SampleData_ArcGIS_v1.7.3b/Socioeconomic Analysis/Visitation Rate/Wolong_NatReserve_bnd.shp*"). The start year is constrained to be 2005 or later while the end year is constrained to be 2014 or earlier. Uncheck the box that computes regression, since in this example we are not interested in estimating the contribution of socio-ecological factors to the visitation rates, but rather calculate the average visitation rates on the study area between 2005 and 2014. Check the box to grid the AOI, and use hexagon as a grid type to divide our study area and calculate visitation rates within each cell. As cell size you can leave the default or specify a different one (in the units of the coordinate systems used, e.g. meters). In this example, we are choosing 10 km (10,000m) as a cell size.

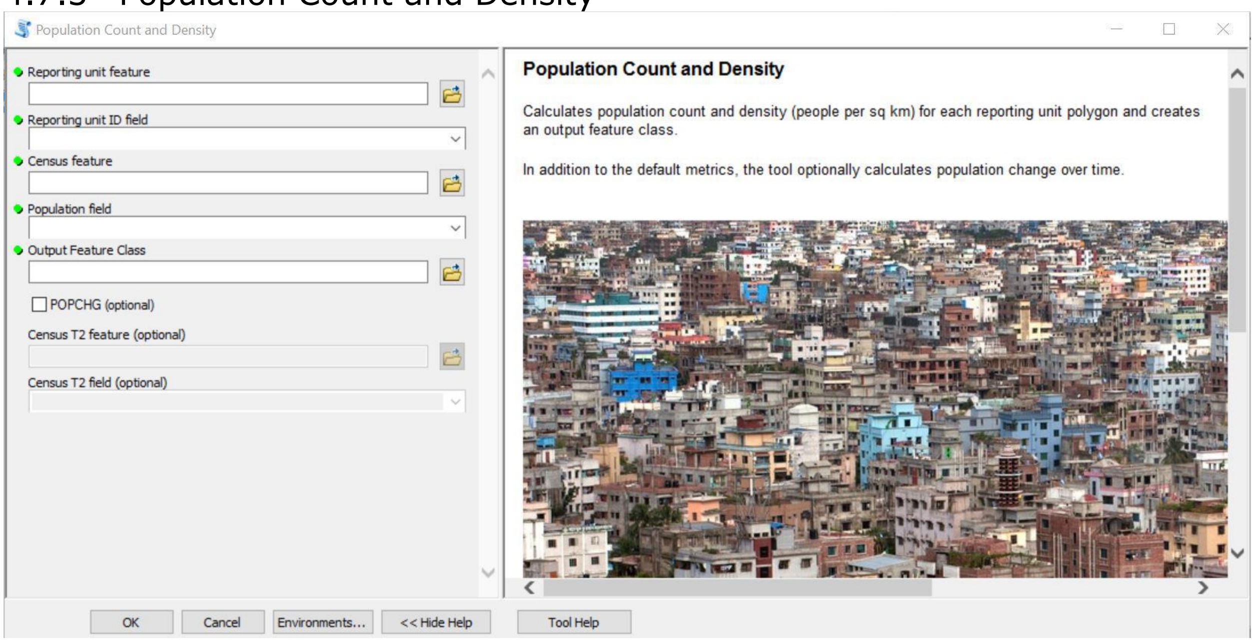
NOTE: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



Click OK and run the tool. The size of your study area, the chosen cell size, and whether or not to run a regression analysis will greatly affect the computation time for the tool to complete. After the tool is done running, ArcGIS returns a shapefile made of hexagonal grid cells and a csv table with extracted monthly counts of geotagged photograph in the selected time frame. The shapefile can be symbolized with colors representing classes of average visitation rates, e.g. yearly average visitation rate (2005-2014). Right-click the shapefile in the table of content, select Properties, go to the symbology tab to choose Quantities > Graduated colors and use the PUD_YR_AVG field as values:

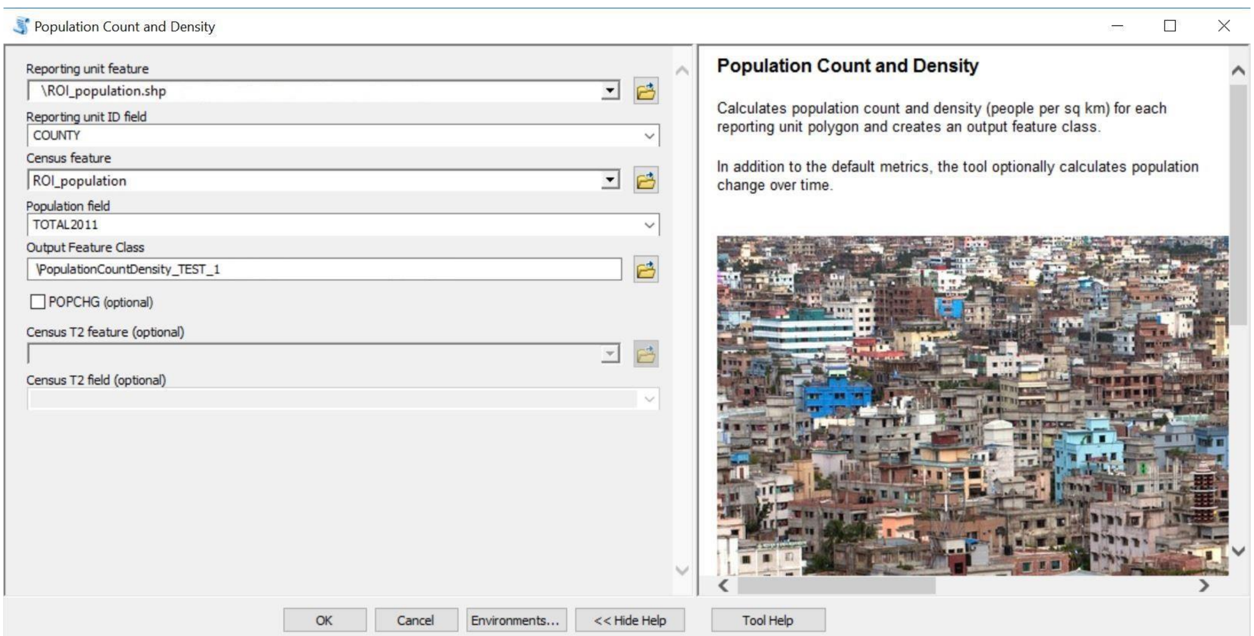


4.7.3 Population Count and Density



Once you've opened the tool, click the folder icon next to the Reporting unit feature field and navigate to the dataset "ROI_population" within the sample data folder ("*./SampleData_ArcGIS_v1.7.3b/Socioeconomic Analysis/Population Count and Density/ROI_population.shp*"). This represents the Republic of Ireland's population from the 2011 Census. ***Note that this shapefile has linear units of meters. When using this tool, your shapefile must be in a projection where the units are in meters.*** You've now indicated that the reporting units will be derived from Ireland's administrative boundaries. Next, click the drop-down arrow for the Reporting Unit ID Field and select "County." This will be the level at which population counts and densities will be calculated. For the Census Feature, select the "ROI_population" shapefile again. This identifies the shapefile that provides the population information. For Population Field, select "TOTAL2011" as this column contains the population from the 2011 Census. Finally, for Output Feature Class, select a location where you'd like to save the tool's results.

Your populated tool should look like this:



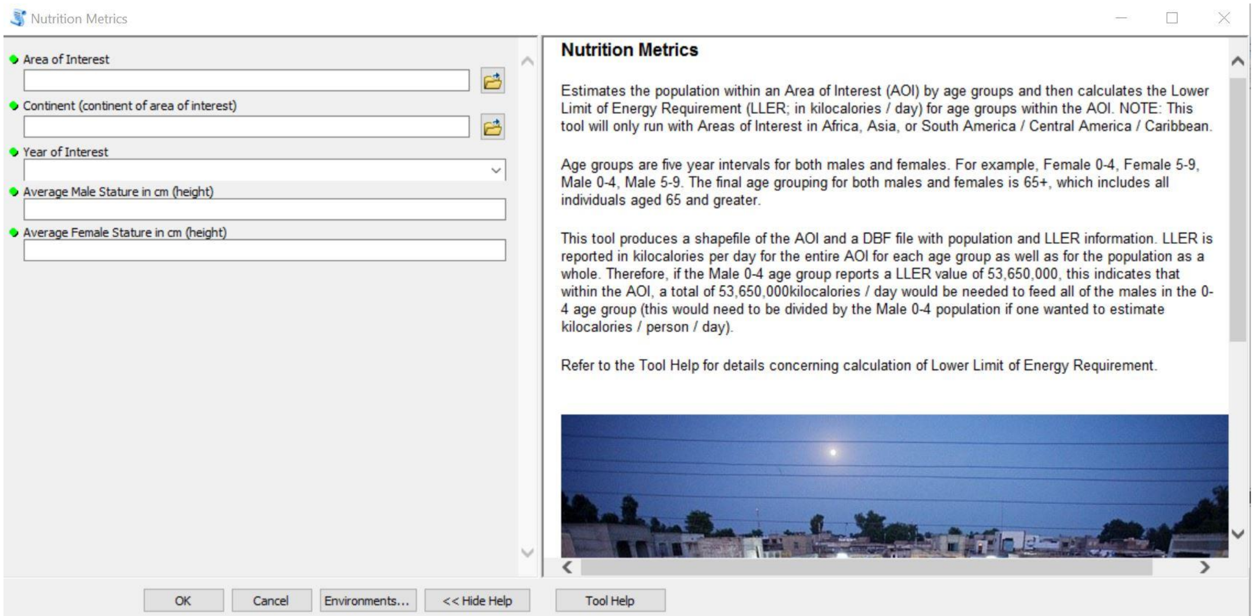
Click OK to run the tool. The tool produces a copy of the shapefile that was used for the Reporting Unit Feature. The attribute table of the returned shapefile contains information on population count, the area of the Reporting Unit ID Field, and the population density within the Reporting Unit ID Field:

Table

[redacted]\FC_layer

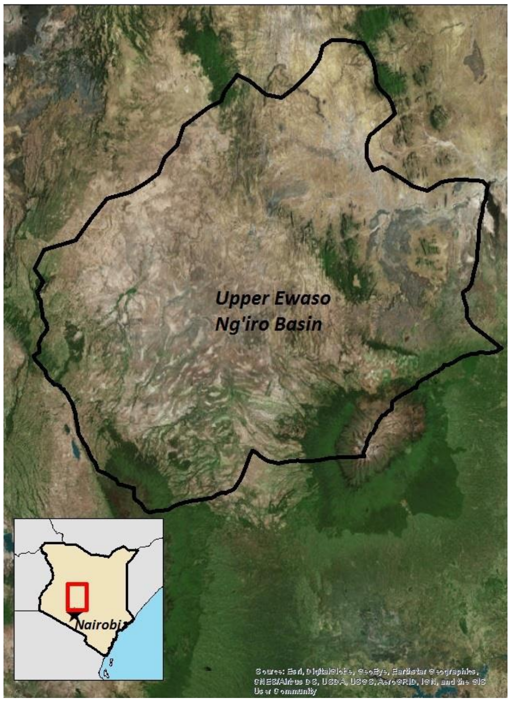
FID	Shape *	popCount	AREAKM2	popDens
0	Polygon	54611.99996	2440.097053	22.381077
1	Polygon	527611.999218	329.703574	1600.261691
2	Polygon	265204.999631	623.822958	425.12863
3	Polygon	273991.000907	1291.558035	212.139907
4	Polygon	206260.99985	352.992788	584.32072
5	Polygon	210312.000679	4711.538592	44.637648
6	Polygon	95419.000661	5599.645984	17.040184
7	Polygon	80559.001801	4735.791807	17.010672
8	Polygon	39015.944583	3110.868462	12.541818
9	Polygon	122897.001305	2378.050352	51.67973
10	Polygon	184135.009453	6648.654823	27.695077
11	Polygon	76687.00017	5567.21893	13.774741
12	Polygon	86163.999965	5194.282735	16.588238
13	Polygon	145319.999604	6373.44436	22.800858
14	Polygon	136640.001051	5575.930947	24.505325
15	Polygon	117202.941278	9435.604329	12.42135
16	Polygon	119229.999148	103.798745	1148.665127
17	Polygon	399802.007991	19585.286982	20.413385
18	Polygon	145502.000957	12724.636772	11.434668
19	Polygon	57106.000183	83.485714	684.021223
20	Polygon	134703.00012	7338.107544	18.35664
21	Polygon	70324.339131	5598.796203	12.560618
22	Polygon	88432.00256	6073.925272	14.559284
23	Polygon	46052.999846	110.407546	417.118225
24	Polygon	67062.999891	4822.282757	13.906899
25	Polygon	75529.000887	141.679958	533.095871
26	Polygon	175124.001697	17095.248393	10.244016
27	Polygon	31798.528848	4619.486725	6.883563
28	Polygon	130638.010453	16080.892721	8.123803
29	Polygon	64065.001615	7264.093032	8.819408
30	Polygon	65392.999849	5347.253603	12.229268
31	Polygon	73183.001484	5578.662877	13.118377

4.7.4 Nutrition Metrics

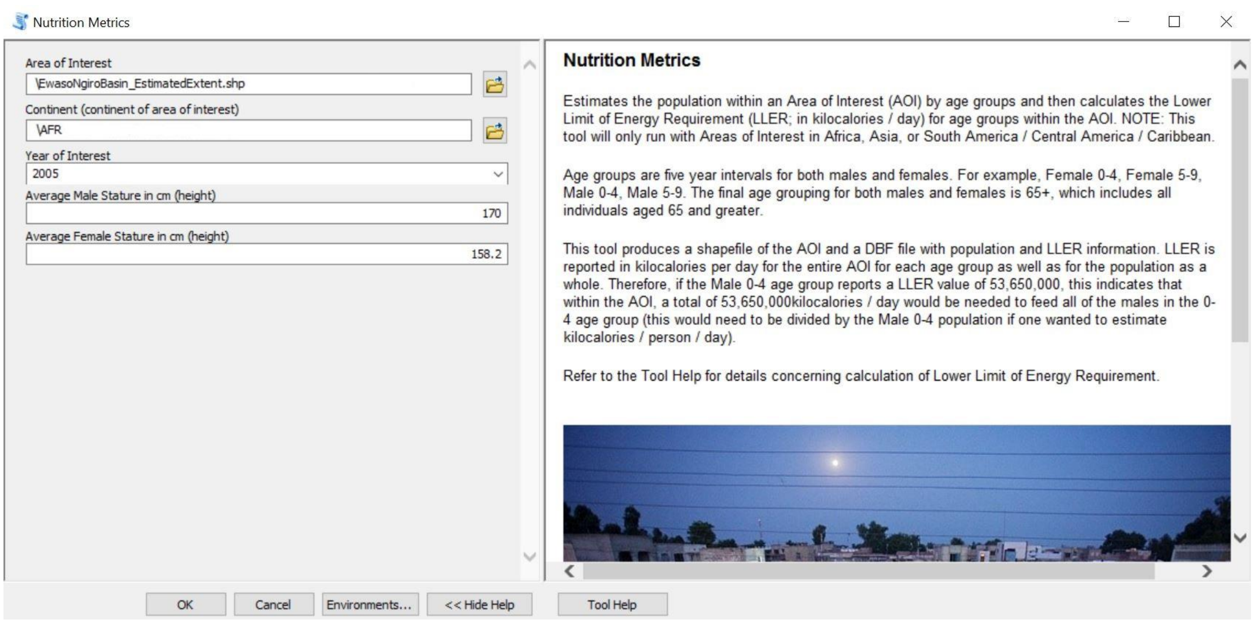


After opening the tool, you will have several parameters to populate before the tool will run. As the area of interest, select the shapefile of the Upper Ewaso Ng'iro basin, a river basin on the northern and northwestern slopes of Mount Kenya ("*./SampleData_ArcGIS_v1.7.3b/Socioeconomic Analysis/Nutrition Metrics/EwasoNgiroBasin_EstimatedExtent.shp*").

The continent dialogue box requires a folder containing rasters representing the population and age groupings necessary to calculate calorie demand for the area of interest. Select the "AFR" folder ("*./SampleData_ArcGIS_v1.7.3b/Socioeconomic Analysis/Nutrition Metrics/AFR*"). *Note: folders containing rasters for Africa, Asia, and Latin America are downloaded with the Telecouling Toolbox sample data*. For the year of interest, the options are limited to 2000, 2005, 2010, 2015, and 2020. For this exercise, choose 2005. The average adult male and female heights within the Upper Ewaso Ng'iro basin are not known, but the average heights for the country of Kenya are available (in centimeters). Enter 170 for average male stature (in cm) and 158.2 for average female stature (in cm).



Your populated dialogue box should look like this:



Click OK and run the tool. The tool will produce a copy of the shapefile that was used as the AOI as well as a DBF file. The attribute table of the returned shapefile and the DBF file contain information on population size and the lower limit of energy requirements (LLER) for each age group. LLER is reported in kilocalories per day for each age group. For example, in the attribute table below, the male 25-29 age group requires a total of 67,039,375 kcal/day. This is approximately 2,350 kcal/person/day in the male 25-29 age group (calculated as follows: LLER of the male 25-29 age group / total population in the male 25-29 age group).

Table

outputTable

OID	Id	ageGroup	Pop	LLER
0	0	f0004	52384.8	48336488.2336
1	0	m0004	52091.1	51593453.7294
2	0	f0509	47460.4	71028382.4453
3	0	m0509	47145	74282002.0225
4	0	f1014	36699.4	61094741.5981
5	0	m1014	36441.6	64188158.1978
6	0	f1519	31223.1	40747761.1555
7	0	m1519	31018	71355930.8918
8	0	f2024	29044.1	38122084.8242
9	0	m2024	28864.9	67775015.9285
10	0	f2529	28758.1	37746739.0623
11	0	m2529	28551.6	67039375.1727
12	0	f3034	24402.6	45595267.648
13	0	m3034	24197.8	56255155.8197
14	0	f3539	18115.6	33848295.9805
15	0	m3539	17951.3	41733236.031
16	0	f4044	14783.7	27622729.6194
17	0	m4044	14651.4	34061616.5127
18	0	f4549	12375	23122170.9967
19	0	m4549	12262.3	28507424.7806
20	0	f5054	9074.62	16955555.9906
21	0	m5054	9009.32	20944914.4654
22	0	f5559	7341	13716369.8866
23	0	m5559	7271.35	16904467.3412
24	0	f6064	5852.8	9626037.46526
25	0	m6064	5800.18	11035863.4883
26	0	f65pl	10703.9	17604617.6552
27	0	m65pl	10597.7	20164020.2762
28	0	total	654072.871094	1111007877.22

1 (0 out of 29 Selected)

outputTable