MICHIGAN STATE
UNIVERSITY

Telecoupling Toolbox v1.0a
User Guide

Author: Francesco Tonini

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

Last Revised: October 18, 2016

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?

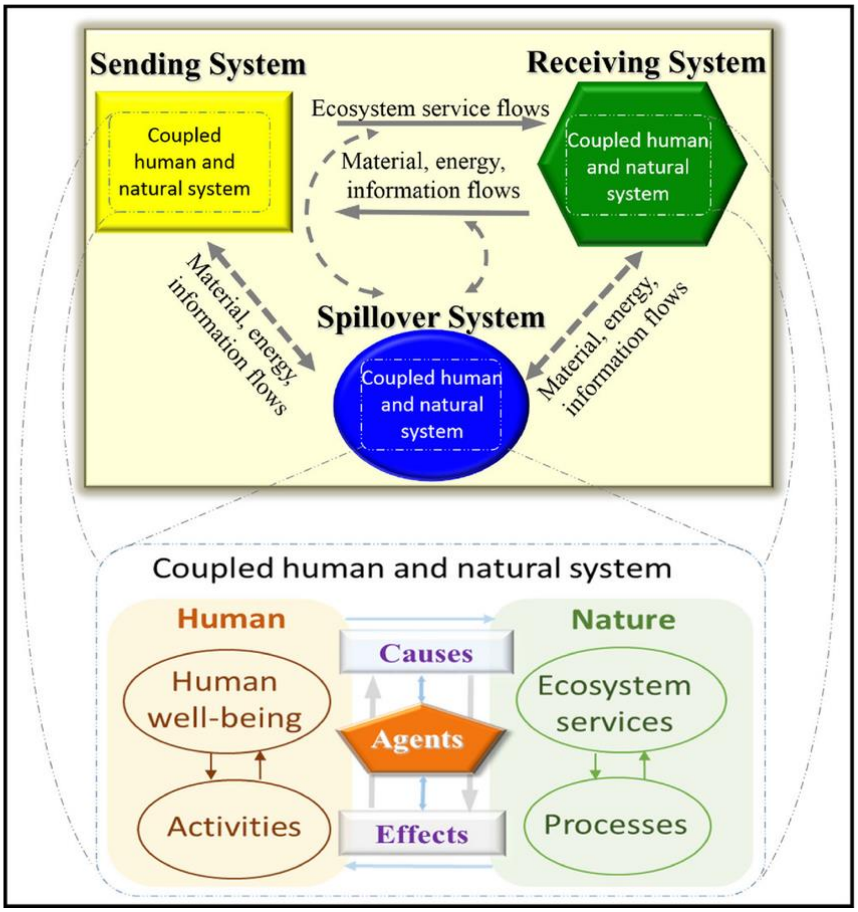


Figure 1. Framing ES in a telecoupling context (adapted from Liu et al. 2013). A “sending” system provides ES to a “receiving” system, and may influence another system (the “spillover” system) in the process. Feedbacks occur among different systems as a result of material energy, and information flows. 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 1.0 (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

No minor releases at this time.

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.

1.5 Available Tools and Telecoupling Typologies

The current version of the toolbox includes script tools that can be applied to several different case studies of telecoupling (Fig. 2), others that only apply to a specific case study, and some that can be applied to multiple telecouplings but require caution depending on data availability or the specific case.

	Wildlife Transfer	**Tourism**	**Agricultural Trade**	**Industrial Trade**	**Conservation Subsidies**	**Information Dissemination**
Add Systems Interactively	✓	✓	✓	✓	✓	✓
Draw Systems from Table	✓	✓	✓	✓	✓	✓
Add Agents Interactively	✓	✓	✓	✓	✓	✓
Draw Agents from Table	✓	✓	✓	✓	✓	✓
Draw Distributive Flows	✗	✗	✗	✗	✗	✗
Draw Radial Flows	✓	✓	✓	✓	✓	✓
Draw Radial Flows & Nodes	✓	✓	✓	✓	✓	✓
Add Media Information Flows	✗	✗	✗	✗	✗	✓
Add Causes Interactively	✓	✓	✓	✓	✓	✓
Model Selection (OLS)	✓	✓	✓	✓	✓	✗
Factor Analysis for Mixed Data	⚠	✓	✓	✓	⚠	✗
Visitation Rate (InVEST 3.3.1)	✗	⚠	✗	✗	✗	✗
Habitat Quality (InVEST 3.3.1)	✓	✓	✓	⚠	✓	⚠
Crop Production (InVEST 3.3.1)	⚠	⚠	✓	⚠	✓	⚠
Cost-Benefit Analysis (Wildlife Transfer)	✓	✗	✗	✗	✗	✗
CO2 Emissions (Wildlife Transfer)	✓	✗	✗	✗	✗	✗
Linear Regression (OLS)	✓	✓	✓	✓	✓	✗
Carbon Storage & Sequestration (InVEST 3.3.1)	⚠	⚠	✓	✓	✓	⚠
Forest Carbon Edge Effect (InVEST 3.3.1)	⚠	⚠	✓	✓	✓	⚠

Figure 2. Script tools available for the Telecoupling Toolbox v1.0a (rows) and examples of telecouplings (columns). Not all script tools can be used in each telecoupling and, in some case, caution is required depending on data availability. The check mark indicates a tool is applicable to the listed telecoupling, the

caution sign suggests careful consideration of tool assumptions and data availability, while the *X* mark indicates a tool is not suitable.

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 (**do NOT install a standalone version of Python! This is installed automatically with ArcGIS**)

ArcGIS Desktop comes with a version of Python and some pre-installed libraries, so you do not need to download and install a separate version of Python, as long as you use the one that comes with ArcGIS.

2.2 Download and unzip the Telecoupling Toolbox repository

Bitbucket
If you have reading access to the Bitbucket repository of the Telecoupling Toolbox[2] follow these guidelines:

1. Find the ***Downloads*** menu on the left of the main overview page.
2. Click on 'Download repository' and save the .zip file on your local computer.
3. Unzip the folder and take a look at the file content and structure. The zipped folder contains a snapshot of ALL current files and documents that are found in this repository.

Google Drive
If you have reading access to the Google Drive repository of the Telecoupling Toolbox[3] follow these guidelines:

[2] https://bitbucket.org/f-tonini/telecoupling-geoapp
[3] https://drive.google.com/drive/u/0/folders/0B0volNNURG7aVnhrWnJ1UFdrc0U

1. Download the entire telecoupling-geoapp folder from Google Drive.
2. Unzip the downloaded folder on your local disk and take a look at the file content and structure. The zipped folder contains a snapshot of ALL current files and documents that are found in this repository.

2.3 Install Python libraries for 3rd party external software

The Telecoupling Toolbox (v1.0a) 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. Download and install the Microsoft Visual C++ Compiler for Python 2.7[4]
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**[5]
     - Open the CMD prompt on Windows and type: `C:\Python27\ArcGIS10.3\python.exe` followed by the full path to `get-pip.py` downloaded file
     - Hit *Enter* to run the command above
     - Open the folder `PyLibs\ArcGIS10.3\Py32bit` found inside the (unzipped) telecoupling toolbox folder
     - Double-click on the `install32.bat` file
   - **ArcGIS 10.3.1 (with 64-bit Background Geoprocessing)**:
     - Download **get-pip.py**[6]
     - Open the CMD prompt on Windows and type: `C:\Python27\ArcGISx6410.3\python.exe` followed by the full path to `get-pip.py` downloaded file
     - Hit *Enter* to run the command above
     - Open the folder `PyLibs\ArcGIS10.3\Py64bit` found inside the (unzipped) telecoupling toolbox folder
     - Double-click on the `install64.bat` file
   - **ArcGIS 10.4.x (standard 32-bit version)**:

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

    - Open the folder `PyLibs\ArcGIS10.4\Py32bit`
    - Double-click on the `install32.bat` file
- **ArcGIS 10.4.x (with 64-bit Background Geoprocessing)**:
    - Open the folder `PyLibs\ArcGIS10.4\Py64bit`
    - Double-click on the `install64.bat` file

2.4 Install the R-ArcGIS Tools

ArcGIS has developed a nice Github webpage[7] with lots of useful documentation on how to install a set of libraries to make sure R and ArcGIS can talk to each other.

Inside the .zip folder of the toolbox you will find another .zip folder called *r-bridge-install-master*. As the instructions in the webpage above point out, unzip the folder and open ArcMap. Go to the toolboxes menu, right-click and select to "add toolbox". Then, find the **R Integration.pyt** toolbox inside the unzipped folder and you should see a new toolbox appearing in the list of existing ones. At this point, follow the instructions from the github webpage above to make sure you have the correct R software version installed and install the appropriate libraries that ArcGIS and R need to talk to each other.

2.5 Add the Telecoupling Toolbox to ArcGIS

You are almost done!

Inside the downloaded (unzipped) Telecoupling Toolbox folder, you will notice an ArcMap file (.mxd) called **TelecouplingApplication.mxd**. If you double-click on it, your ArcGIS will open and show a set of GIS layers used by the application as a basemap. Moreover, if you open the ArcToolbox tab, you should see the current version of the Telecoupling Toolbox already added to the list. Alternatively, you can open a brand new ArcMap document and follow the same procedure shown in the 'Install the R-ArcGIS tools' section to add the Telecoupling Toolbox.

Inside the Telecoupling Toolbox you should see 5 toolsets (***agents***, ***causes***, ***effects***, ***flows***, ***systems***) and a number of python tool scripts inside each one of them. To learn more about what each tool script does and what parameters it takes, please refer to the help documentation that comes with it. To do so, you can either use ArcCatalog, clicking on the 'description' tab for a tool or, alternatively, 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).

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

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

3.3 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***	Feature Set

	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.	
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. 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. ***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_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.	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_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	String

	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.	
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

3.4.2 Draw Radial Flows and Nodes

Draw radial flows between two or more telecoupling system components (i.e. "sending", "receiving", "spillover") and add nodes at locations of flow destinations. 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 two new feature classes: a first one 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; a second one containing point features constructed based on the values in an end x-coordinate field, and end y-coordinate field of a table. 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	Table View

	NOTE: table attributes should at minimum include XY coordinates of both the origin and destination nodes of the flows.	
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. ***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_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	String

	(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.	
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

3.4.3 Draw Distributive Flows

The original Distributive Flow Lines python toolbox (.pyt) was developed by ESRI and updated last in 2014. The following description is taken "as is" from the original tool documentation page:
The tool creates flow lines depicting quantities flowing from a source to destination. The output is a line feature class with a numeric attribute that can be symbolized using graduated symbols. The line weights will increase as they get closer to the source depicting the sum of quantities "flowing" into a line from the "tributaries". If a polygon feature class is used for the source or destinations, centroids will be created and used as starting/ending points. Source, destination, and (optional) impassable and impedance feature classes should all be projected for best results.(Not Geographic Lat Lon Degree units) If the processing extent covers the entire width of the globe, flow lines may "wrap around" to the other side of the map resulting in a confusing representation. To prevent this, decrease both the Left and Right extent processing values by at least one "Cell Size"

The Impassable layer represents features that Flow Lines cannot intersect. These features can be point, line or polygon feature classes. The features are automatically buffered by a width equal to three times the Cell Size parameter and in this buffered area the cost is set very high so that flow lines do not flow into the impassable features. The features are also buffered by a smaller amount and the resulting buffer polygons are used to set the Raster processing Mask environment variable and also burned into a "background" processing layer as NODATA.

The Foreground features all receive a very high cost but in extreme cases flow lines may cross these features. Foreground features must be polygons and represent features flow lines will avoid if possible.

For Example: Country polygons might be used as a Foreground feature class so that when Flow Lines are generated they stay in the ocean area as much as possible and only cross the country polygons to reach otherwise unreachable destination points.

The default cell size parameter is automatically calculated based on the processing extent and units so that the raster processing grid size is 1500 rows or columns in the larger dimension of the processing Extent window. This value can be overridden by the user. Increasing the cell size will decrease processing time and detail in the output Flow lines. Decreasing cell size will increase detail and processing time. Large cell sizes are good for quick preview processing while small cell sizes are best for final results.

Note:

- Source and Destination features should be in a projected coordinate system like UTM for best results.
- If the processing extent covers the entire width of the globe, flow lines may "wrap around" to the other side of the map resulting in a confusing representation. To prevent this, decrease both the Left and Right extent processing values by at least one cell size width.
- The cell size selected in the tool will have some effect on the shape/direction of the flow lines because it is linked to total cost of the path. In general the paths will look fairly similar for a set of inputs if the cell size is not changed by orders of magnitude.

For more information on the tools, its author(s) and its functionality, please refer to https://blogs.esri.com/esri/apl/2013/08/26/flow-map-version-2/.

Parameter	Explanation	Data Type
Source_Feature_Class	Choose a point or polygon feature class or shapefile containing one feature. This will be the source for the flow lines.	Feature Layer
Destinations_Feature_Class	Choose a point or polygon feature class or shapefile containing multiple features. These will be the destinations for flow lines.	Feature Layer
Distributed_Quantity_Field	Select a positive integer field in your destination feature class that represents the quantity to be mapped. If, for example, you were mapping migration, this would be the number of people immigrating/emigrating. If the field is non-integer, a temporary integer field will need to be created by multiplying the numeric field by a multiple of 10.	Field

	After the output line feature class is created divide the output field by the same multiple of 10. Example: If the field to be mapped is Tons and the values range from .01 - 100.74 • Create a temporary field of type LONG called Tons100, for instance • Calculate its values to be Tons * 100 • Run the tool • The output feature class will have an attribute called "grid_code" • Create a new field of type Double , called Tons, in this new dataset. Calculate its values to be grid_code/100	
Impassable_Feature_Class (Optional)	A point, line or polygon feature class or shapefile that flow lines will not cross. These features can be features from an existing feature class or a new feature class, created by the user, and act as barriers to prevent flow lines from going through undesirable locations. These features will be buffered by 3 times the cell size and the buffered area will have a very high cost so the flow lines try to avoid this buffered area. In addition to the buffered area that will be avoided, impassable features will also be buffered by 1.4 times the cell size and the area inside this buffer will be NoData. Flow lines will not cross this inner buffered area. Paths can be dramatically affected by these features.	Feature Layer
Impedance_Feature_Class (Optional)	Flow lines are determined by a combination of four parameters: distance from source, distance from destinations, impedance cost and location of impassable features. Select how much flow lines should try to avoid Impedance features. Higher	Feature Layer

	numbers will cause flow lines to go around these features or cross over these features in narrow locations or the shortest distance to destination features located inside them. This weight is affected by cell size. Varying the cell size by orders of magnitude can cause the path to change radically if Impedance features are used in the model.	
Split_paths_close_to_Source_or_Destinations	Flow lines are determined by a combination of the: distance from source, distance from destinations, impedance cost and location of impassable features. By moving the slider to the right, closer to 10 the paths will split closer to destinations. Moving the slider left, closer to 1, will cause paths to split closer to the Source features. Paths that split closer to the Source feature will look more like a starburst pattern. Paths that split closer to the destination features will look more like a stream drainage pattern. Setting the slider in the middel does not necessarily mean the paths will split halfway between source and destination features.	Long
Impedance_Weight_squared	Flow lines are determined by a combination of four parameters: distance from source, distance from destinations, impedance weight and location of impassable features. Select how much flow lines should try to avoid Impedance features. Higher numbers will cause flow lines to go around these features or cross over these features in narrow locations or the shortest distance to destination features located inside them. This weight is affected by cell size. Varying the cell size by orders of magnitude can cause the path to change radically if impedance features are used in the model.	Long
Processing_Extent	Select the extent flow will be modeled over. At the very least, the processing extent must contain the Source feature point or centroid of the source polygon and at least one destination feature. If the processing extent covers the entire width of the globe, flow lines may "wrap around" to the other side of the map resulting in a confusing representation. To prevent this, decrease both the Left and Right extent processing values by at least one cell size.	String

Cell_Size	The tool uses raster analysis; choose the cell size for modeling flow. The default cell size is set so the number of cells in the largest dimension of the processing extent is 1500. This is a good starting point. Larger cell size will process much faster. Smaller cell sizes will take longer to process. Very large variances in cell size can have dramatic and surprising effects on flow lines. The effectiveness of Source, Destination, and Impedance weights are closely linked to cell size.	Long

3.4.4 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[8] 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
Source_Location	Select on the map the geographic location from which information has emanated out. This location should	Feature Set

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

	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.***	
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. • NORMAL_SECTION—A type of geodetic line which represents a path between any two points on	String

	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 Describe Causes table for each cause class you added. There must be a one-to-one correspondence between the***	Feature Set

	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	Double

	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.	
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 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	Double

	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.	Boolean

	When unchecked, missing values are replaced by their column average.	
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 • "dist N": select the N elements that have the highest distance to the center of gravity	String
Categories_Selection (Optional)	A selection of the categories that are drawn:	String

	• "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	
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 Effects 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.

3.6.1 Carbon Storage and Sequestration (InVEST 3.3.1)

The InVEST Carbon Storage and Sequestration model estimates the current amount of carbon stored in a landscape and values the amount of sequestered carbon over time. First it aggregates the biophysical amount of carbon stored in four carbon pools (aboveground living biomass, belowground living biomass, soil, and dead organic matter) based on land use/land cover (LULC) maps provided by users. A fifth optional pool represents carbon stored in harvested wood products, such as firewood, charcoal, or long-lived timber products. If the user provides a future LULC map, the carbon sequestration component of the model estimates expected change in carbon stocks over time. This portion of the model values the amount of carbon sequestered as an environmental service using additional data on the market value or social cost of carbon, its annual rate of change, and a discount rate. With optional inputs on probability distributions of carbon amount in different pools, the model can perform uncertainty analysis providing standard deviations for carbon estimates and a map showing where sequestration or emissions will occur with confidence. ***NOTE: to have a more detailed explanation on this model and its parameters, please read the official online InVEST documentation***.

Parameter	Explanation	Data Type
Run_Biophysical_Model	The biophysical model and the valuation model may be run independently or together. In order to run the valuation model, either the biophysical model must be run 'Calculate Sequestration' checked, or inputs must be provided under the 'Sequestration Data' section below. When checked (this is the default), the biophysical model component is run. When unchecked, the user must provide parameters found inside the **'Sequestration Data'**section below.	Boolean
Calculate_Sequestration (Optional)	Check to enable sequestration analysis. This requires inputs of Land Cover/Land Use maps for both current and future scenarios.	Boolean
Current_Land_Use_Land_Cover_Raster (Optional)	A GDAL-supported raster representing the land cover of the current scenario.	Raster Layer
Current_Year_of_Land_Cover (Optional)	The calendar year of the current scenario.	Long
Future_Land_Use_Land_Cover_Baseline_Raster (Optional)	A GDAL-supported raster representing the land cover of the future scenario.	Raster Layer
Future_Year_of_Land_Cover (Optional)	The calendar year of the future scenario.	Long
Carbon_Pools (Optional)	A table that maps the land-cover IDs to carbon pools. The table must contain columns of 'LULC', 'C_above', 'C_Below', 'C_Soil', 'C_Dead' as described in the User's Guide. The values in LULC must at least include the LULC IDs in the land cover maps.	File

Current_Harvest_Rate_Map (Optional)	An OGR-supported shapefile containing information about harvested wood products for the current scenario. The field is optional. If supplied, the shapefile must have the fields 'Cut_cur', 'Start_date', 'Freq_cur', 'Decay_cur', 'C_den_cur', 'BCEF_cur'.	Shapefile
Future_Harvest_Rate_Map (Optional)	An OGR-supported shapefile containing information about harvested wood products for the future scenario. The field is optional. If supplied the shapefile must have the fields 'Cut_cur', 'Start_date', 'Freq_cur', 'Decay_cur', 'C_den_cur', 'BCEF_cur'.	Shapefile
Data_for_Uncertainty_Analysis_enable_to_trigger_uncertainty_analysis (Optional)	Check to trigger uncertainty analysis. When unchecked (default), no uncertainty analysis is triggered.	Boolean
Carbon_Pools_Uncertainty (Optional)	A table that maps the land cover ids to probability distributions for carbon pools. Each probability distribution must be specified by a mean and a standard deviation. The table must contain columns of 'LULC', 'C_above_mean', 'C_above_sd', 'C_below_mean', 'C_below_sd', 'C_soil_mean', 'C_soil_sd', 'C_dead_mean', and 'C_dead_sd' as described in the user's guide. The values in LULC must at least include the LULC IDs in the land cover maps.	File
Confidence_Threshold (Optional)	The percent confidence that should be used as a minimum threshold for highlighting areas of confidence in the output data.	Double
Sequestration_Raster (Optional)	The per-pixel sequestered carbon either from a run of the InVEST Carbon Biophysical model, or from a user defined source.	Raster Layer
Start_Year_of_Sequestration_Measurement (Optional)	The calendar year of the current scenario.	Long
Final_Year_of_Sequestration_Measurement (Optional)	The calendar year of the future scenario.	Long

Run_Valuation_Model	In order to run valuation, either the biophysical model must be run with '**Calculate Sequestration**' checked, or inputs must be provided under the '**Sequestration Data**' section above. When unchecked (default), no valuation model is run.	Boolean
Price_in_Terms_of_Metric_Tons_of (Optional)	The units for the price per unit of carbon type.	String
Value_of_Carbon_Price_Metric_Ton (Optional)	The price per unit ton of carbon or C02 as defined in the carbon price units.	Double
Market_Discount_in_Price_of_Carbon (Optional)	The discount rate as an integer percent.	Long
Annual_Rate_of_Change_in_Price_of_Carbon (Optional)	The integer percent increase of the price of carbon per year.	Double

3.6.2 CO_2 Emissions (Wildlife Transfer)

Calculate the amount of CO_2 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 adds a new field (attribute) with the amount of CO_2 (kilograms) given the amount of CO_2 per unit of length (e.g. meters) specified by the user. 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
Units_of_CO2_Emitted_kg_unit	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

3.6.3 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

3.6.4 Forest Carbon Edge Effect (InVEST 3.3.1)

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

3.6.5 Crop Production (InVEST 3.3.1)

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_Folder	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_Folder (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_Information (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

3.6.6 Habitat Quality (InVEST 3.3.1)

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_Contai ning_Threat_R asters	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).	File

	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. 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	File

	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.	
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

3.6.7 Linear Regression (OLS)

Performs global Ordinary Least Squares (OLS) linear regression to generate predictions or to model a dependent variable in terms of its relationships to a set of explanatory variables.
You can access the results of this tool (including the optional report file) from the Results window. If you disable background processing, results will also be written to the Progress dialog box. The present tool has been slightly modified from the original ***Ordinary Least Squares*** 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.

Parameter	**Explanation**	**Data Type**
Input_Feature_Class	The feature class containing the dependent and independent variables for analysis.	Feature Layer
Unique_ID_Field	An integer field containing a different value for every feature in the Input Feature Class.	Field
Dependent_Variable	The numeric field containing values for what you are trying to model.	Field
Explanatory_Variables	A list of fields representing explanatory variables in your regression model.	Multiple Value

Generate_Report_File (Optional)	If checked, the tool generates an optional report PDF file. This report file includes model diagnostics, graphs, and notes to help you interpret the OLS results. If unchecked (default), the tool does not generate any report PDF file.	Boolean

3.6.8 Visitation: Recreation and Tourism Storage and Sequestration (InVEST 3.3.1)

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',	File

	and 'type'. The file paths can be absolute, or relative to the table.	
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 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 Bitbucket[9], a common web-based hosting service for projects that use revision control systems (e.g. Git). Although the present toolbox was developed to work within ESRI's ArcGIS software environment, thus limited to the Microsoft Windows platform, we are currently planning a concurrent transition to a web-based computer application. The major advantage of this transition will be 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.

3.8 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

[9] https://bitbucket.org/f-tonini/telecoupling-geoapp

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

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.

3.9 Acknowledgements

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

3.10 License & Copyright

The telecoupling toolbox is free of cost and licensed under the **ADD HERE**. Some of script tools (see Section 2 for more details) have been modified from or are based on external software tools, e.g. InVEST. In all these cases, we report the type of software license that is publicly disclaimed by the copyright holders and organizations that developed the tools. At this stage.

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.zip*). You can either use the basemap and operational layers provided with the *TelecouplingApplication.mxd* file or start a brand new map and use one of the basemaps freely provided by ESRI (File > Add Data > Add Basemap…) or a global administrative layer that fits your purpose. 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.

<u>NOTE</u>: 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.

<u>NOTE</u>: 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. Each telecoupling has three .csv tables corresponding to systems, agents, and flows ("*_Systems*", "*_Agents*", "*_Flows*") that can be used for the script tools that draw features on the map using a table on disk (see following sections for examples). Inside the *SampleData* folder, you will also find several other subfolders divided by topic which will be used in some of the following tutorials (Fig. 3).

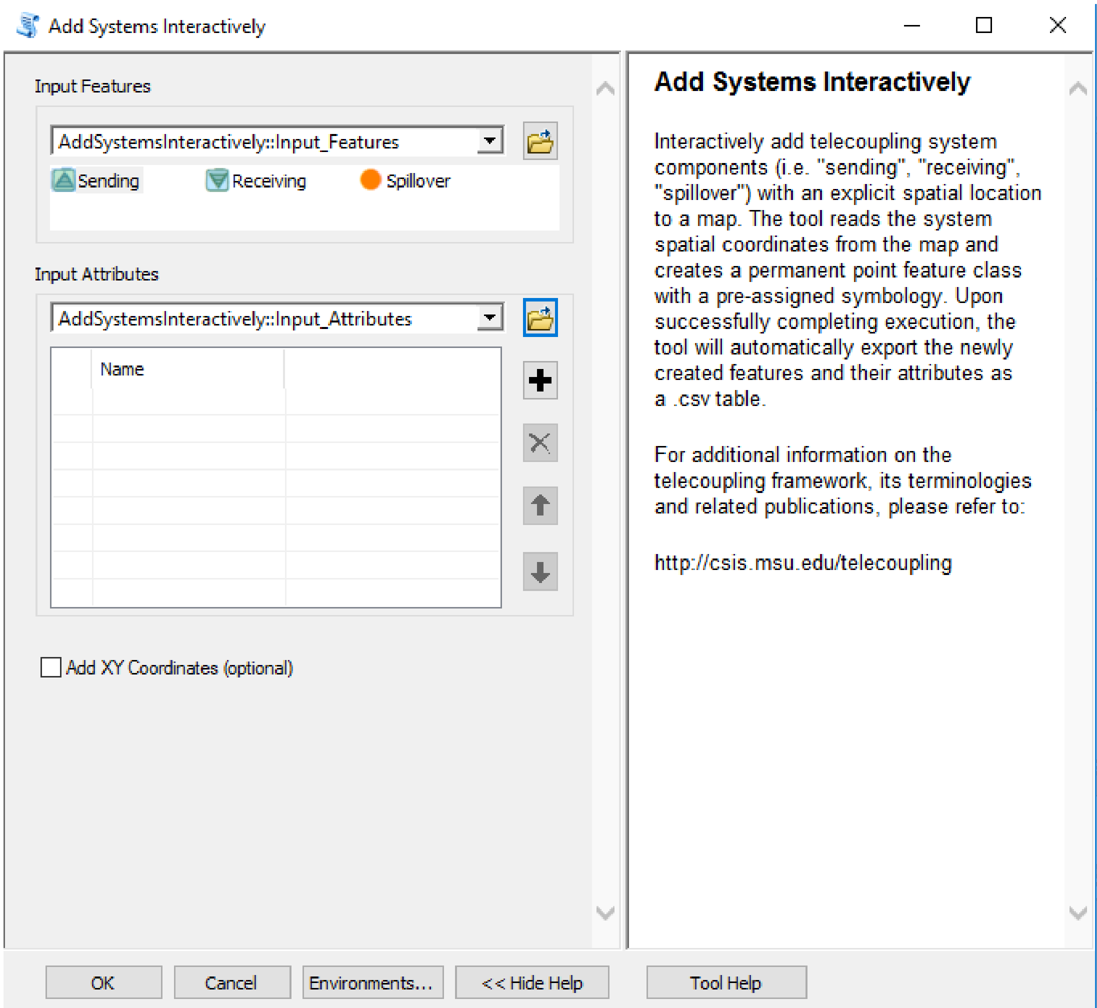
Name	Type
British_Coal_Exports.gdb	File Geodatabase
admin	Folder
carbon	Folder
crop	Folder
gtgp	Folder
habitat_quality	Folder
information	Folder
LULC	Folder
tourism	Folder
trade	Folder
wildlife	Folder
agTrade_Agents.csv	Text File
agTrade_Flows.csv	Text File
agTrade_Systems.csv	Text File
conservation_Agents.csv	Text File
conservation_Flows.csv	Text File
conservation_Systems.csv	Text File
indTrade_Agents.csv	Text File
indTrade_Flows.csv	Text File
indTrade_Systems.csv	Text File
info_Agents.csv	Text File
info_Flows.csv	Text File
info_Systems.csv	Text File
tourism_Agents.csv	Text File
tourism_Flows.csv	Text File
tourism_Systems.csv	Text File
wildlife_Agents.csv	Text File
wildlife_Flows.csv	Text File
wildlife_Systems.csv	Text File

Figure 3. File structure of the SampleData folder provided with the telecoupling toolbox.

!!DESCRIBE EACH DATA AND FOLDERS HERE!!

4.2 SYSTEMS TOOLSET

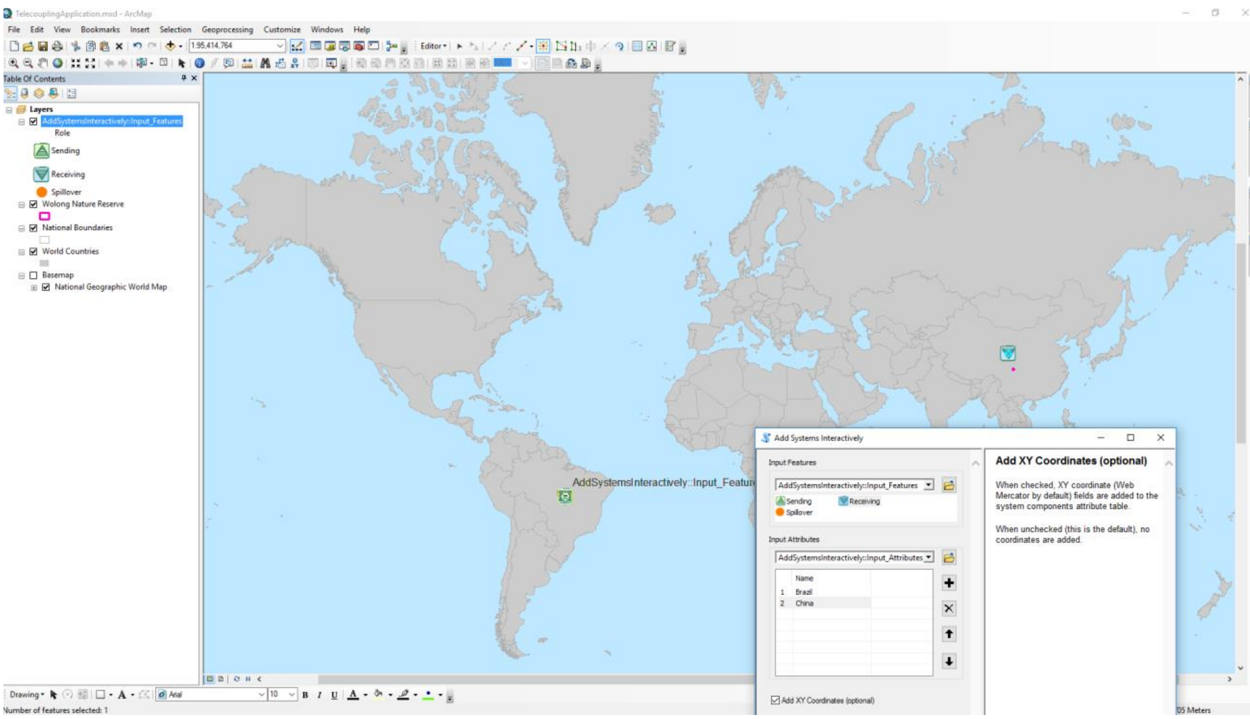
4.2.1 Add Systems Interactively



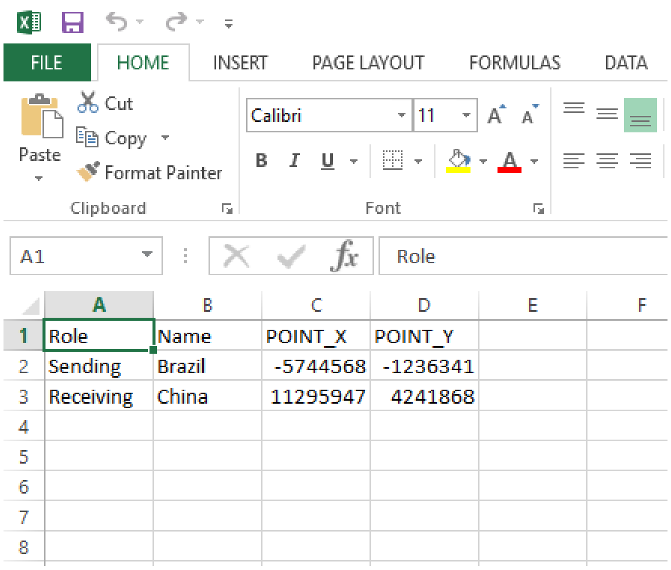
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

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.

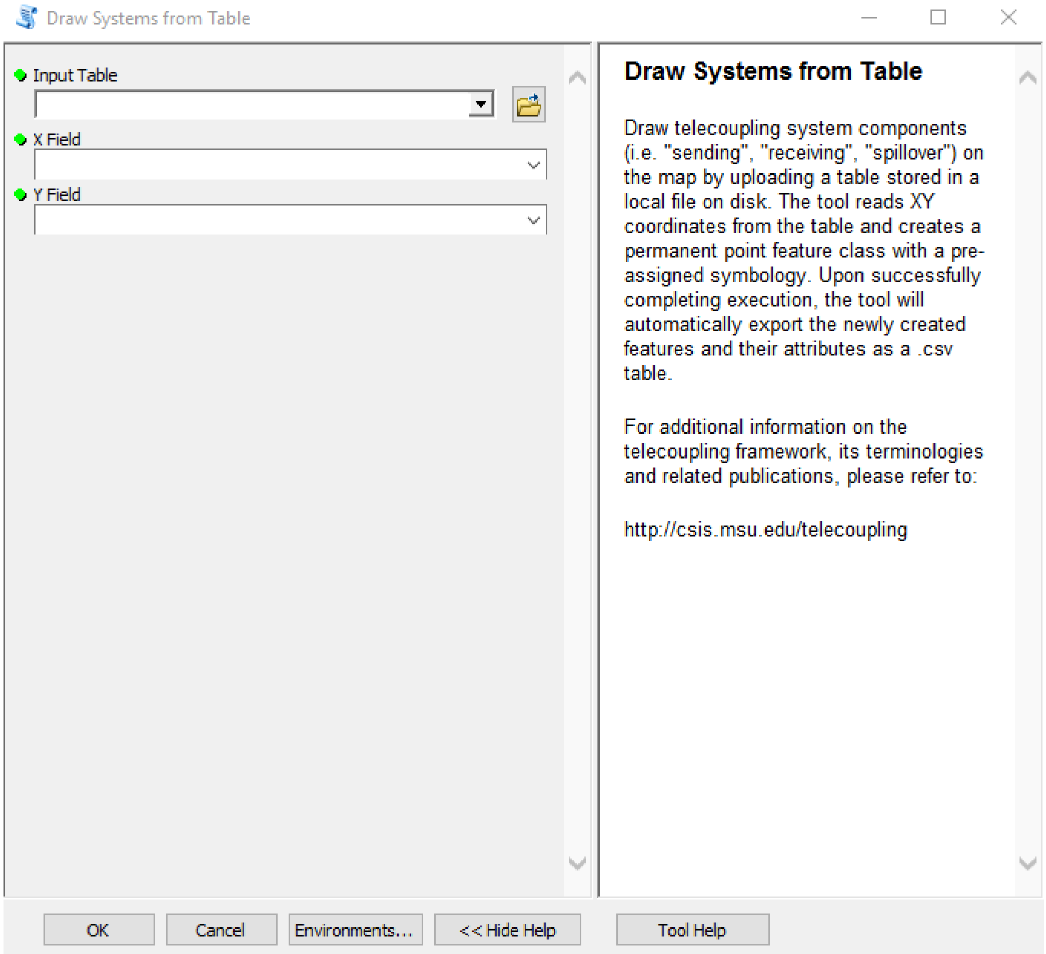


The output .csv table created by the tool in the scratch workspace directory should look something similar to the following figure:



	A	B	C	D	E	F
1	Role	Name	POINT_X	POINT_Y		
2	Sending	Brazil	-5744568	-1236341		
3	Receiving	China	11295947	4241868		
4						
5						
6						
7						
8						

4.2.2 Draw Systems from Table

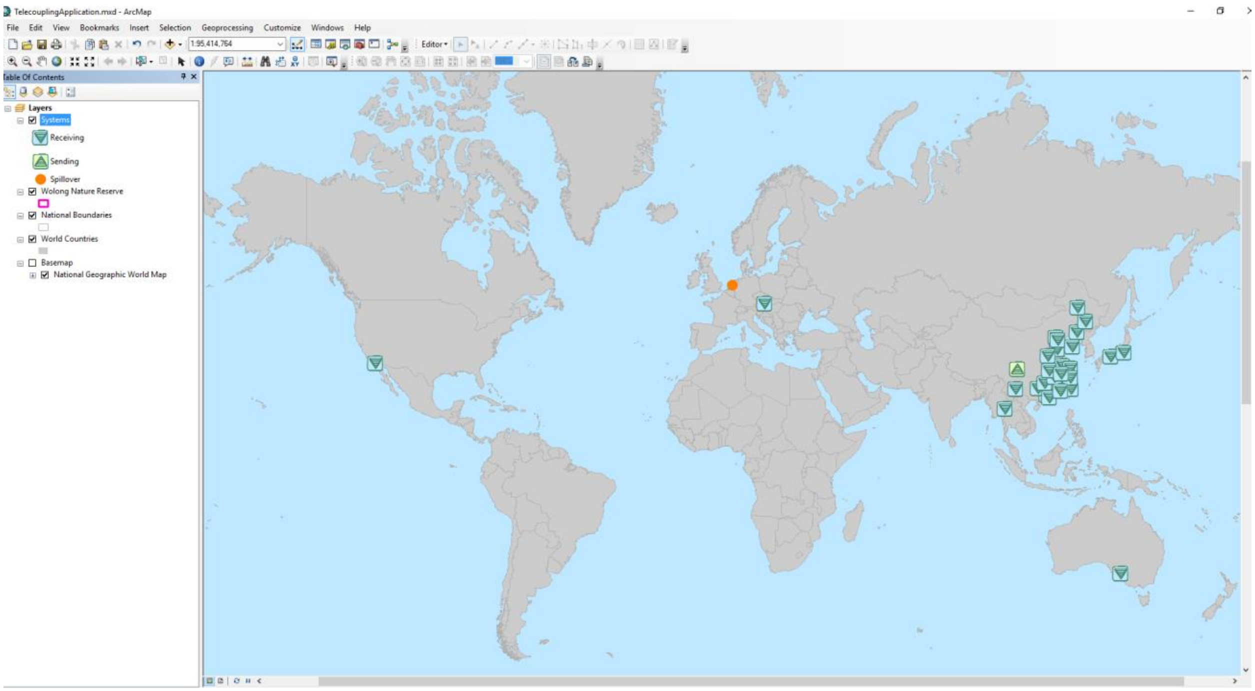


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

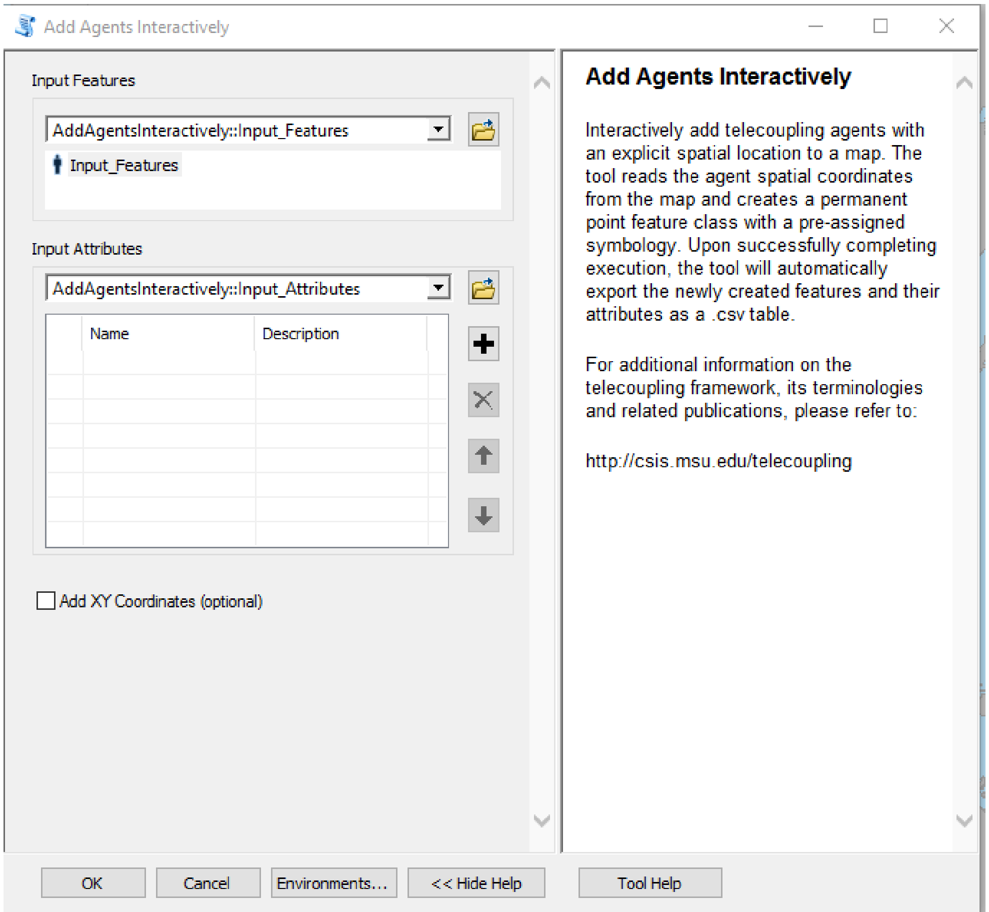
If you are using the sample data provided with the toolbox, you can use any of the .csv tables ending with the word "_Systems" (e.g. */SampleData/wildlife_Systems.csv*).

<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.



4.3 AGENTS TOOLSET

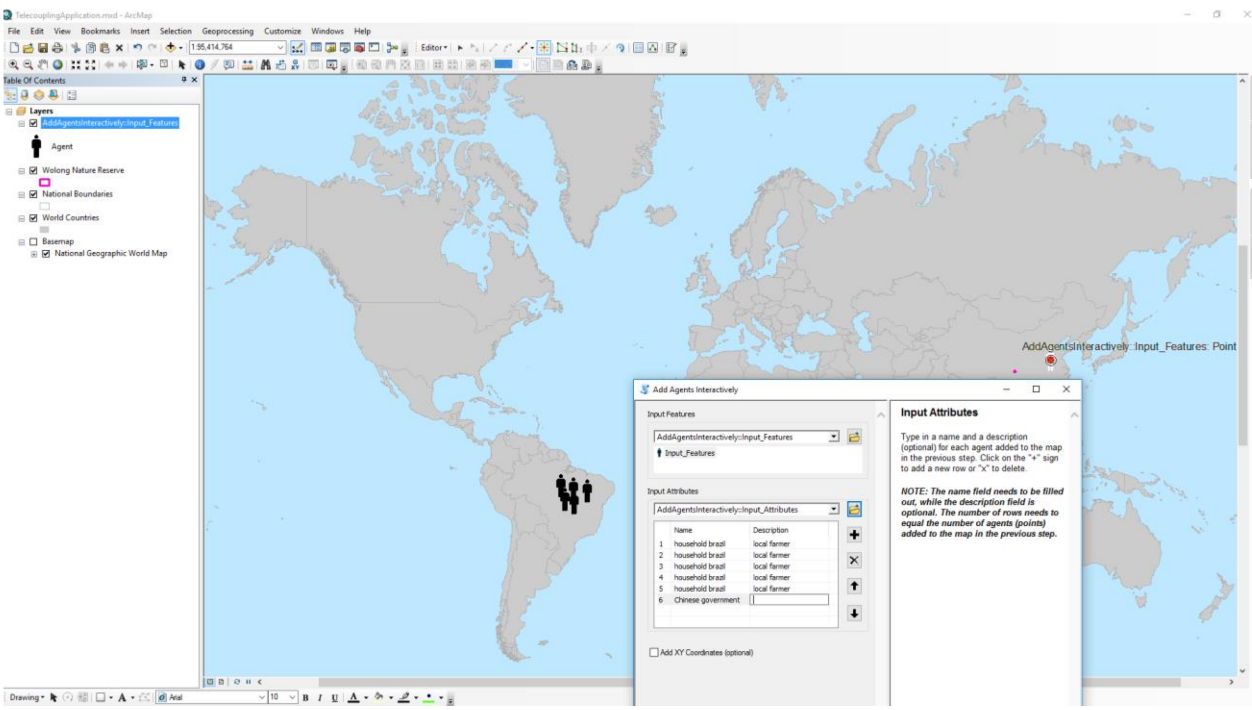
4.3.1 Add Agents Interactively



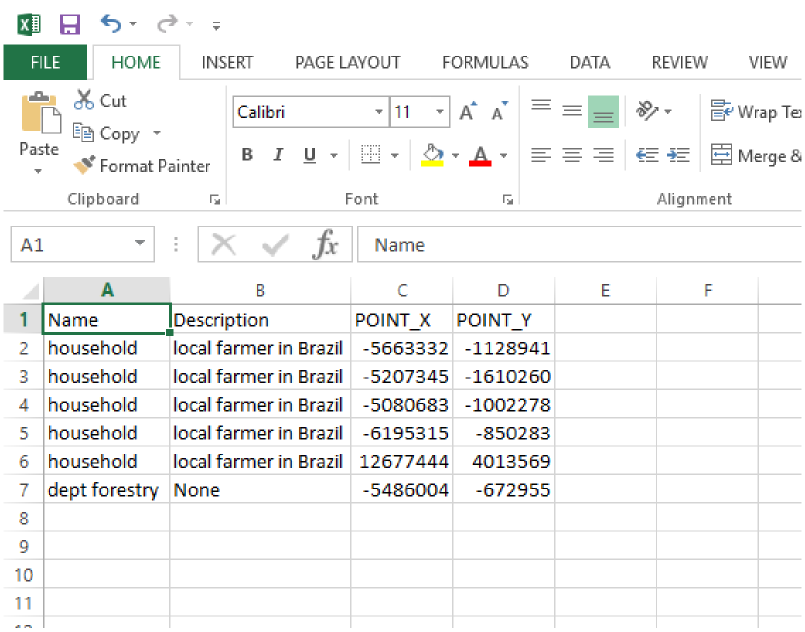
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

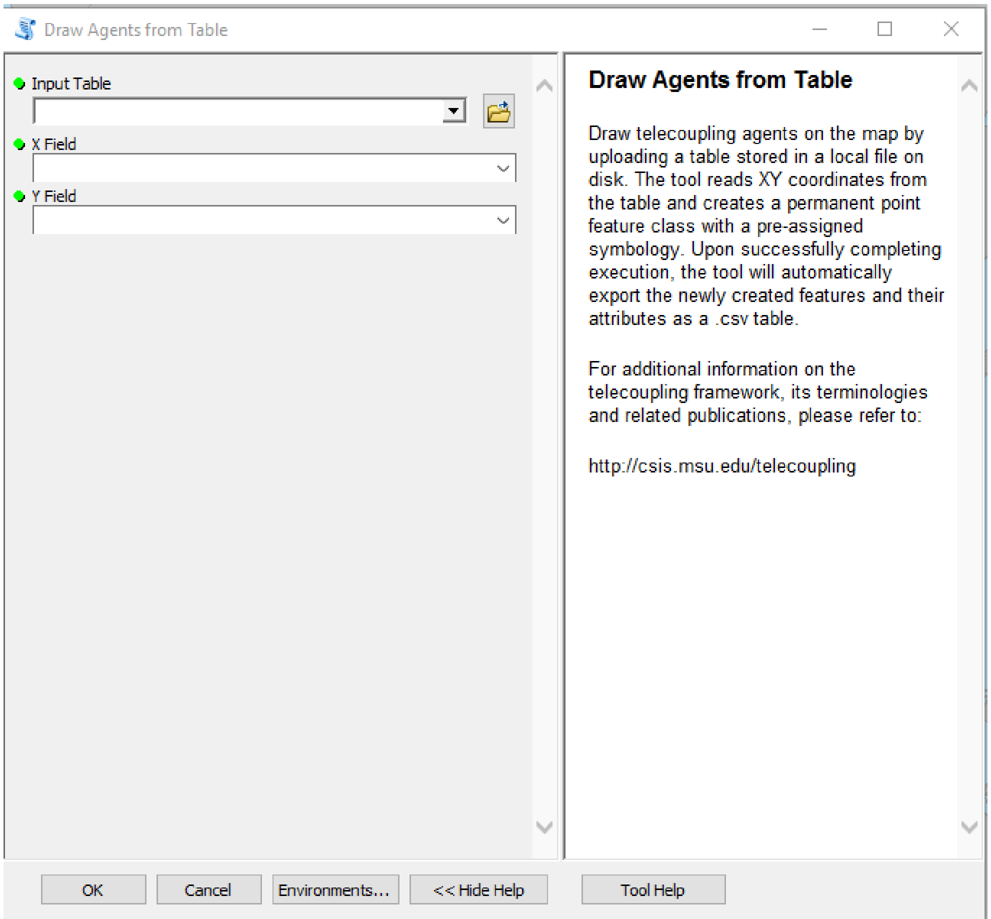


The output .csv table created by the tool in the scratch workspace directory should look something similar to the following:



Name	Description	POINT_X	POINT_Y
household	local farmer in Brazil	-5663332	-1128941
household	local farmer in Brazil	-5207345	-1610260
household	local farmer in Brazil	-5080683	-1002278
household	local farmer in Brazil	-6195315	-850283
household	local farmer in Brazil	12677444	4013569
dept forestry	None	-5486004	-672955

4.3.2 Draw Agents from Table

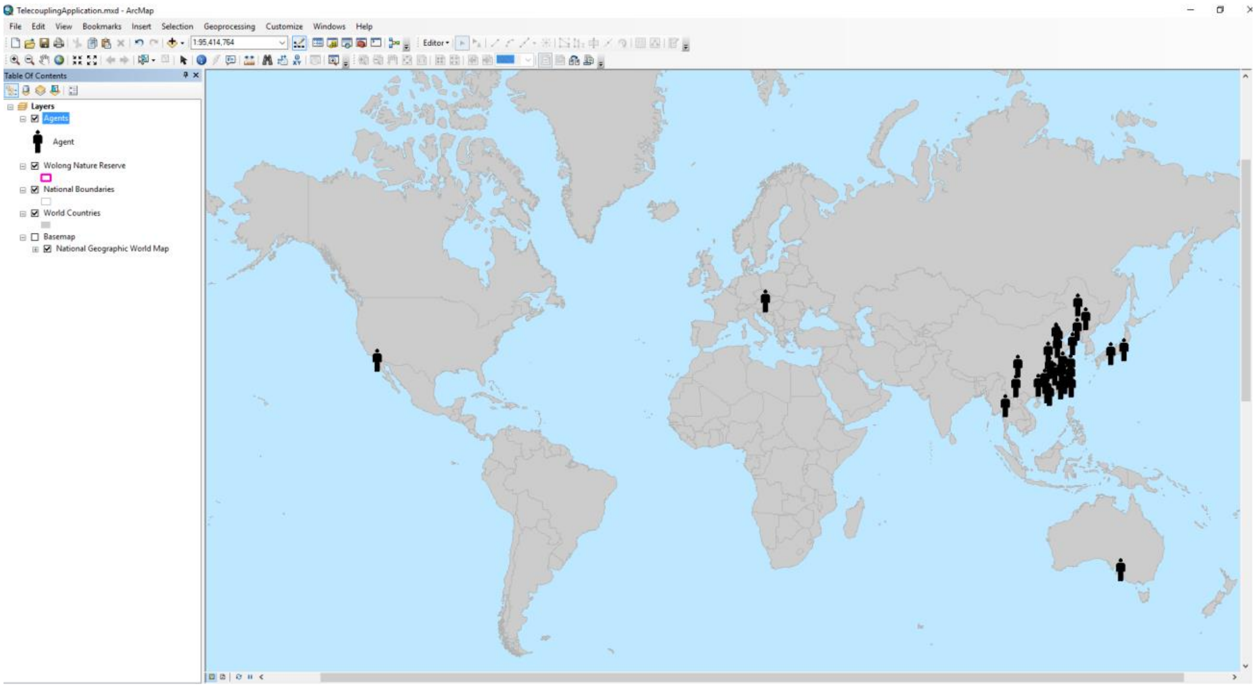


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

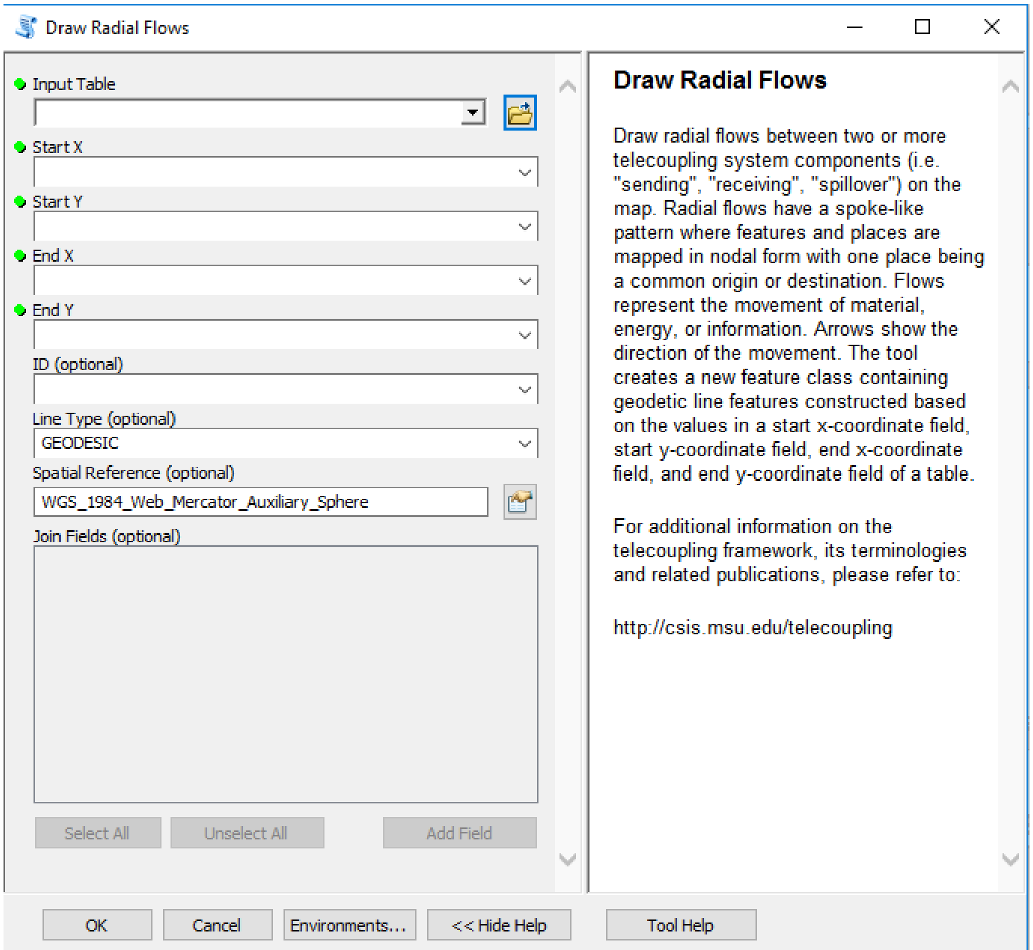
If you are using the sample data provided with the toolbox, you can use any of the .csv tables ending with the word "_Agents" (e.g. */SampleData/wildlife_Agents.csv*).

<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



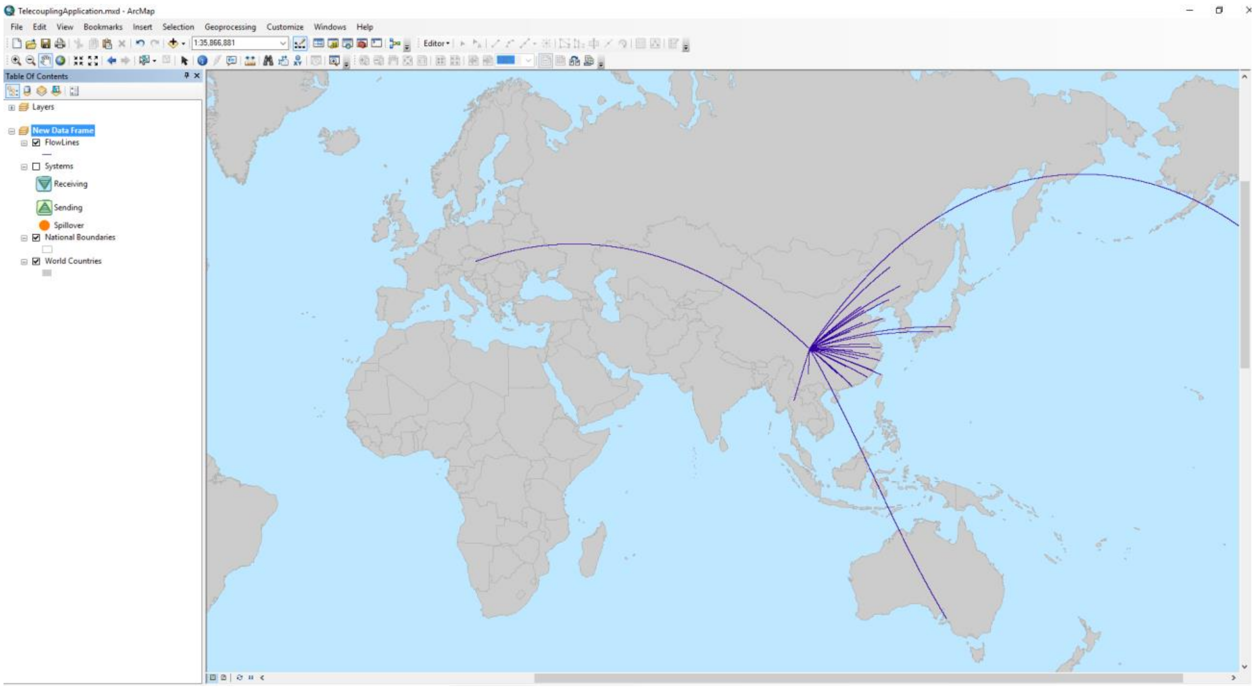
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

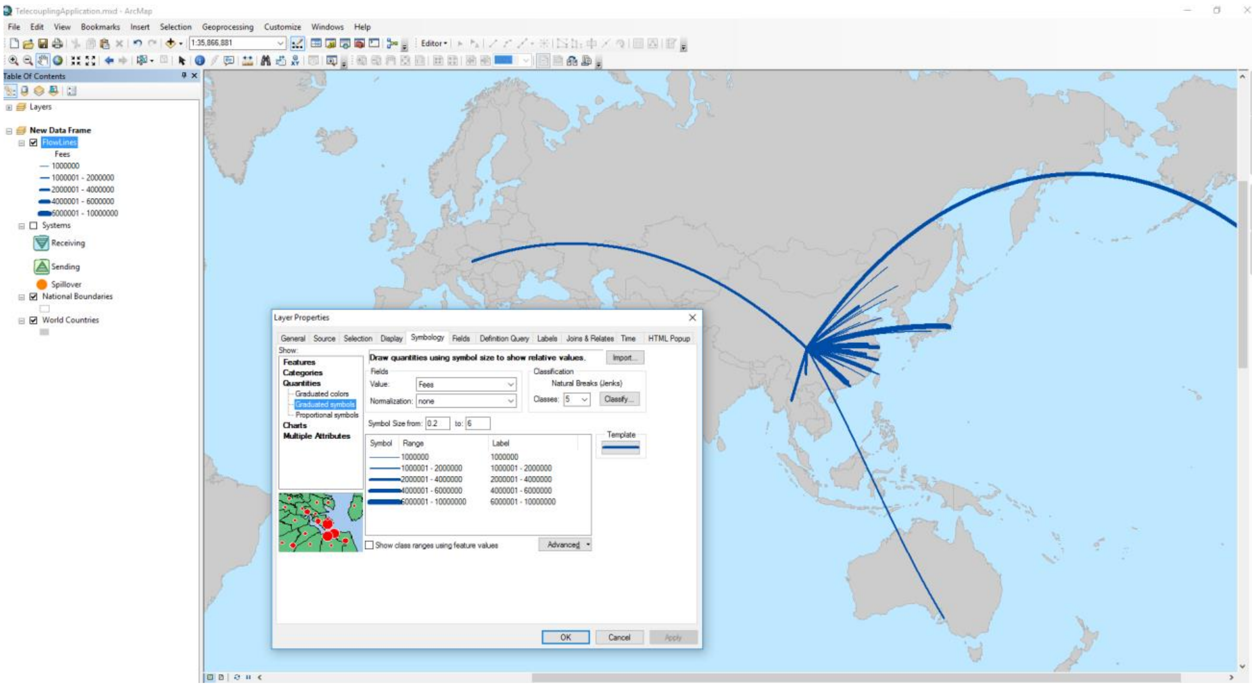
Once you click OK to run the tool, ArcGIS will create a permanent output line feature class and add it to the map.

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/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.

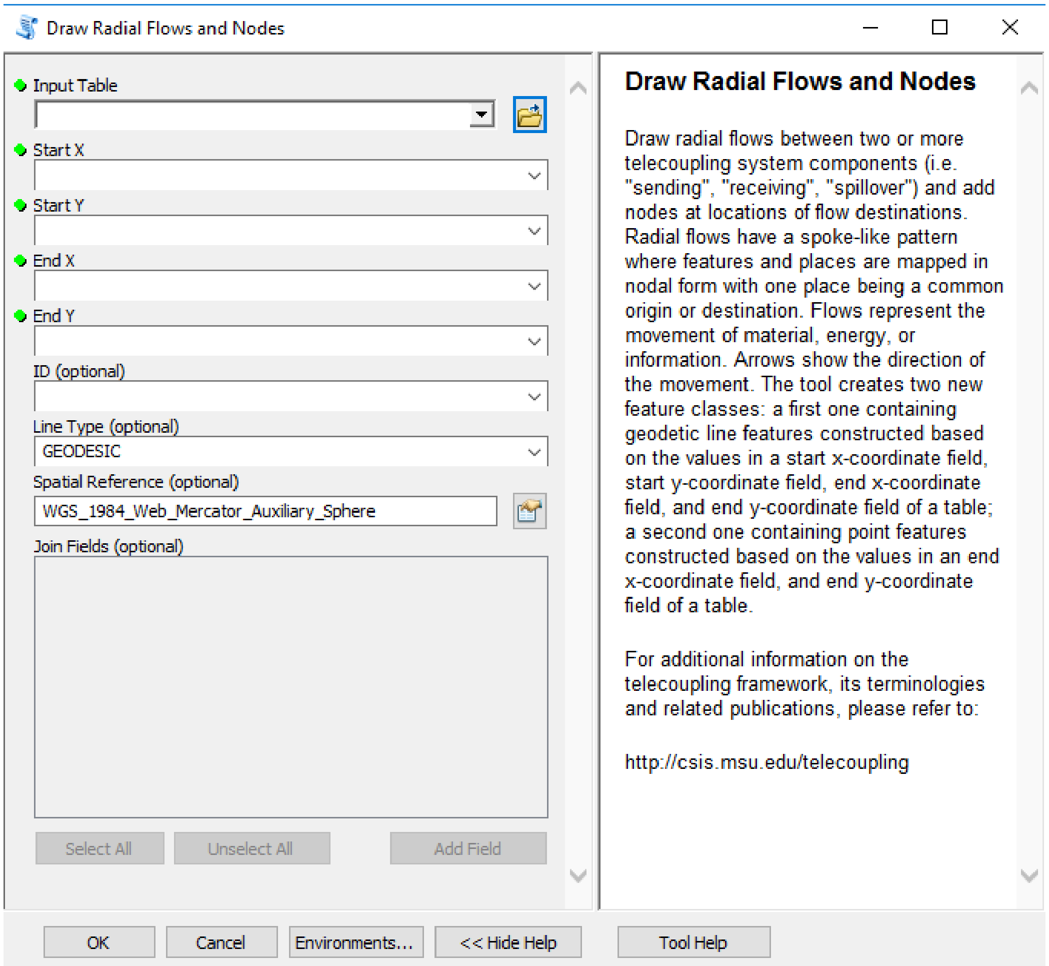
NOTE: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



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.


Layer Properties
General
Source
Selection
Display
Symbology
Fields
Definition Query
Labels
Joins & Relates
Time
HTML Popup
Draw quantities using symbol size to show relative values.
Value:
Fees
Normalization:
none
Natural Breaks (Jenks)
Classes:
5
Symbol Size from:
0.2
to:
6
1000000
1000001 - 2000000
2000001 - 4000000
4000001 - 6000000
6000001 - 10000000
OK
Cancel
Apply

4.4.2 Draw Radial Flows and Nodes



This tools is identical to the previous one, with the difference that nodes (point locations) are created at each destination point of the flows between systems and can be used to symbolize the amount of material (e.g. traded goods, transaction fees) at the receiving system.

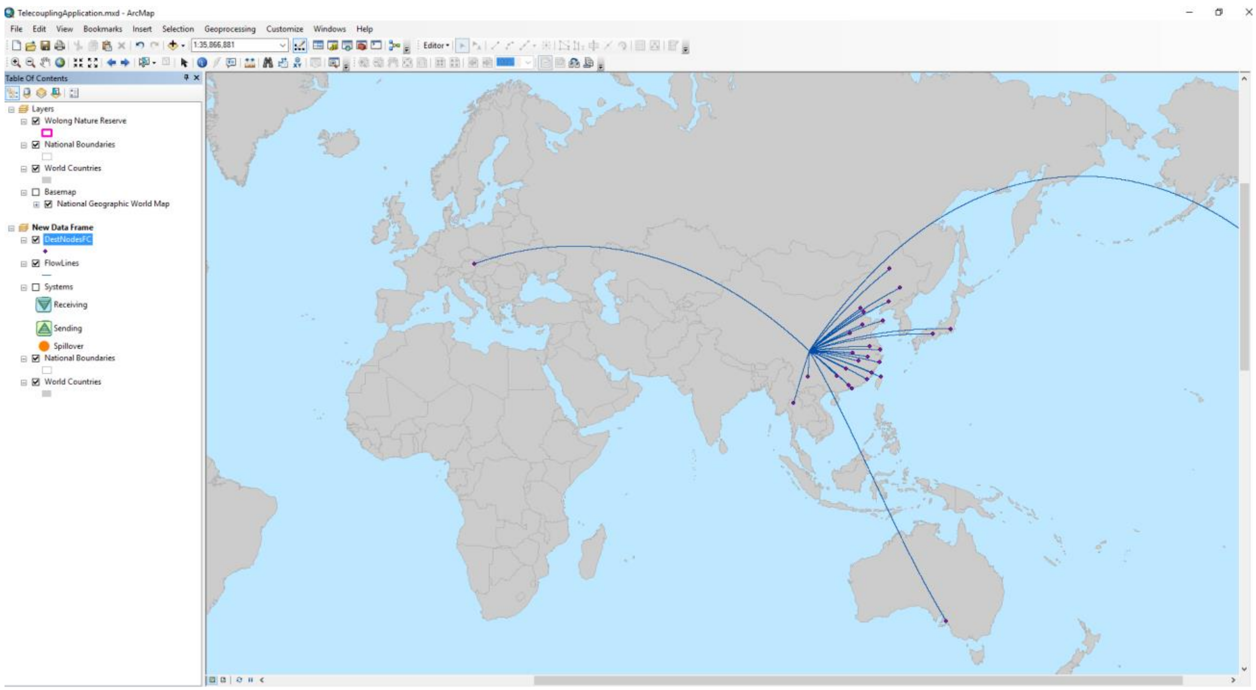
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

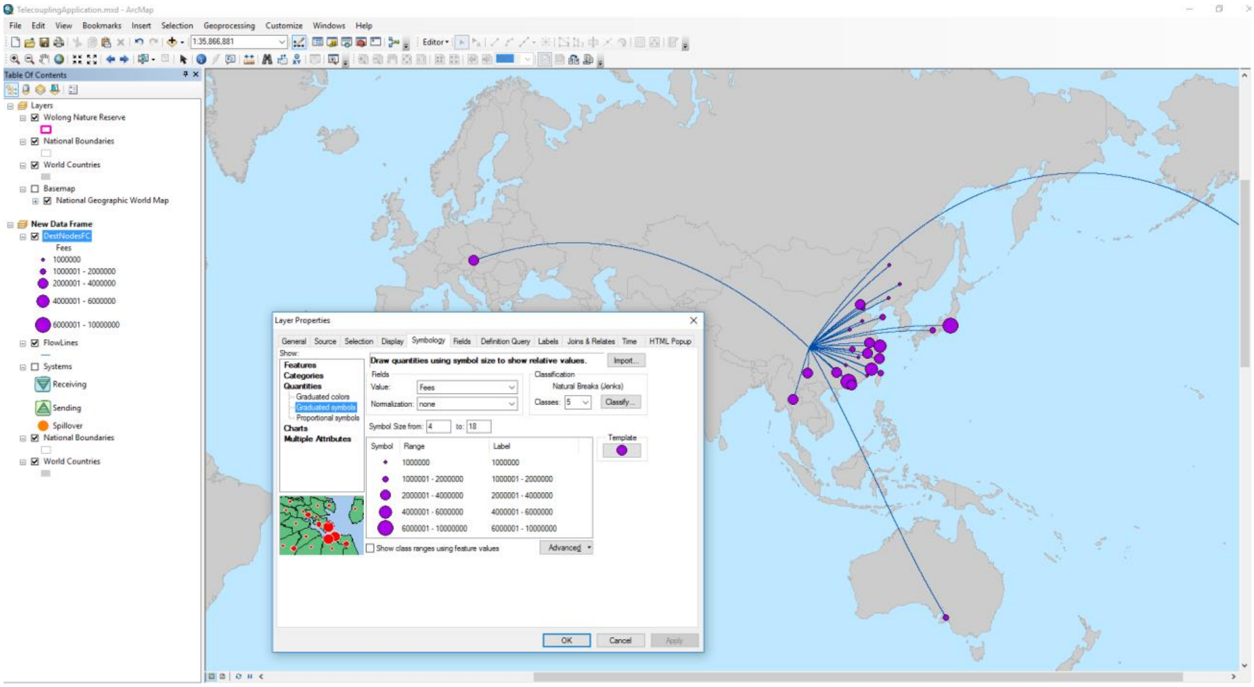
Once you click OK to run the tool, ArcGIS will create a permanent output line feature class representing flows between systems, a point feature class representing destination points of the flows between systems, and add both to the map.

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/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.

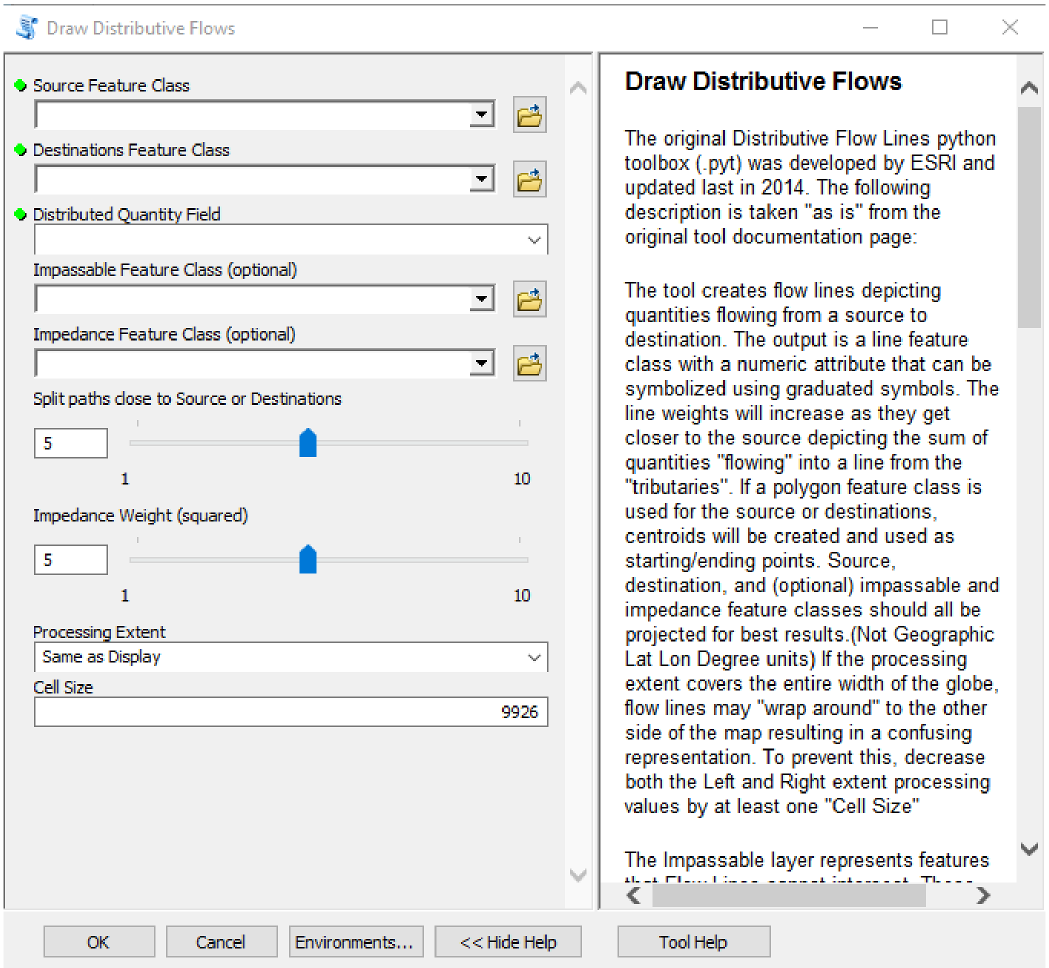
NOTE: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



Once the output flow and nodes feature classes have been created, the lines and points will all have a uniform color and thickness/size. If you have selected additional variables to represent the flow amount, you can use one of them to improve the visualization of the lines or nodes. For example, 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 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 nodes have been symbolized proportionally to the amount of fees paid for the wildlife transfer transaction.



4.4.3 Draw Distributive Flows



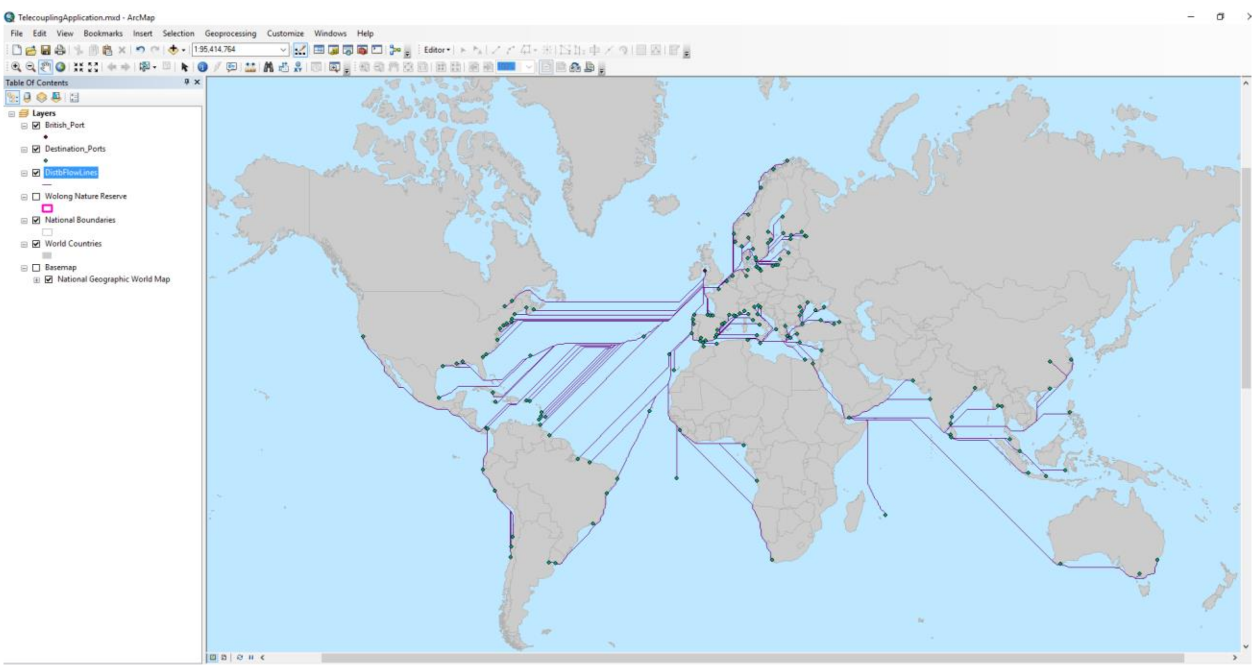
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

This tool is meant to compute and map the distribution of commodities or some other flow that diffuses from one or only a few origins to multiple destinations (e.g. maritime routes for traded goods). Therefore, the distributive flow is not appropriate for flows of material transported via airplane routes or flows of information where the spread is better represented via networks. To show how this tool works, we will use a dataset on British coal exports provided by ESRI in the form of a file geodatabase (*./SampleData/British_Coal_Exports.gdb*).

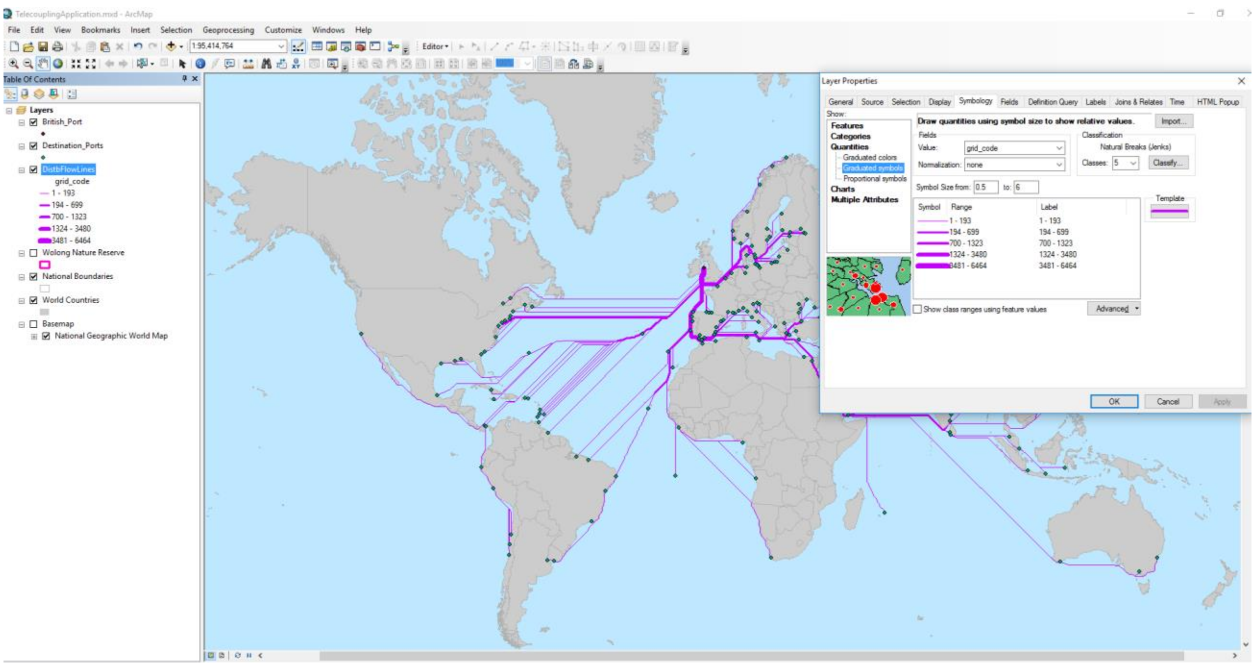
The first and second parameters need points or polygon layers representing the source and destination of the distributive flow. In this example, select the “British_port” and the “Destination_ports” feature classes within the British Coal geodatabase, respectively. Use CoalTonnage as quantitative attribute from the input features table to distribute its flow over the map. The impassable and impedance parameters are a fairly similar concept: the first type of features cannot be crossed by any flow lines and represent a hard barrier, while the latter should be avoided if possible but flows can cross them if necessary to reach a destination location. For this example, you can choose the “Countries” layer as impassable or impedance feature class and play around with the split paths and impedance weight parameter values to test out differences. Choose a cell size that is appropriate for your case study depending how fine or course the output distributive flow layer should be. The smaller the cell size (higher resolution), the longer the tool will take to run. For this example, you can also accept the default value provided.

<u>NOTE 1</u>: Make sure that this geoprocessing tool is running in foreground (right-click script tool and select the option to “always run in foreground”) or you will receive an execution error. This issue will be addressed in forthcoming versions of the toolbox.

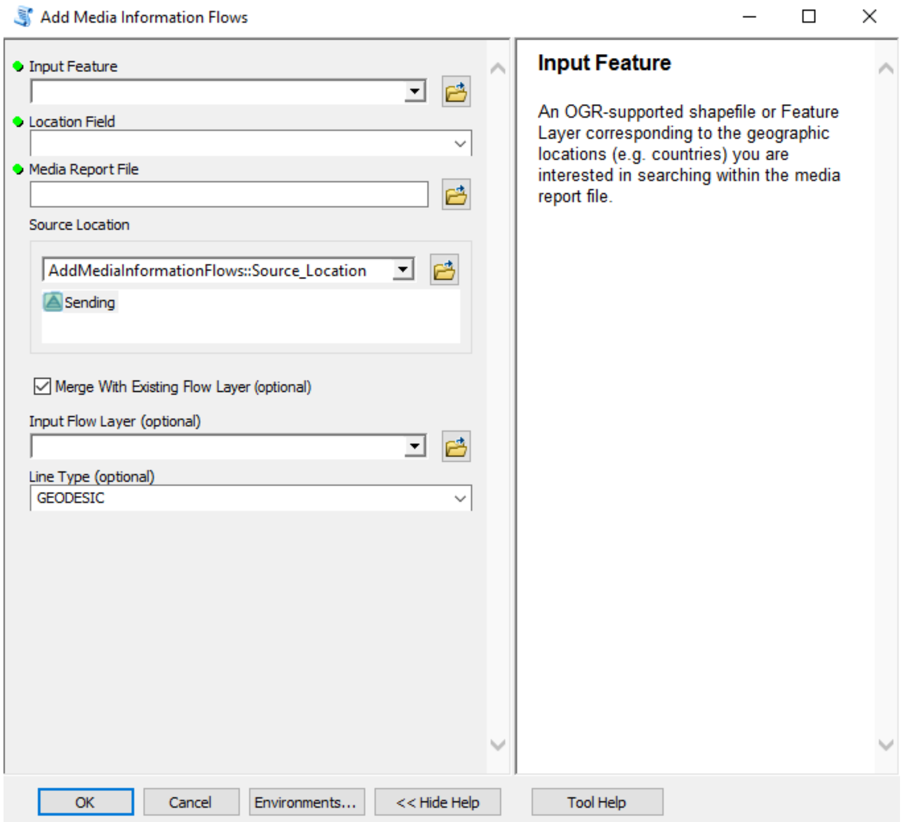
<u>NOTE 2</u>: In order to change the default workspace location where your output files and layers are created, click on the “Environments…” button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



Once the output distributive flow layer has been created, similarly to what described in the radial flows tool you can symbolize the lines based on values of a quantitative or qualitative attribute from the input feature layer. For example, you may want to proportionally increase the line thickness of the flows based on the coal tonnage per grid cell. 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. "grid_code"). Choose how many bins you want to group your values and a desired color. The next figure shows you an example using 4 value classes of the grid_code variable.



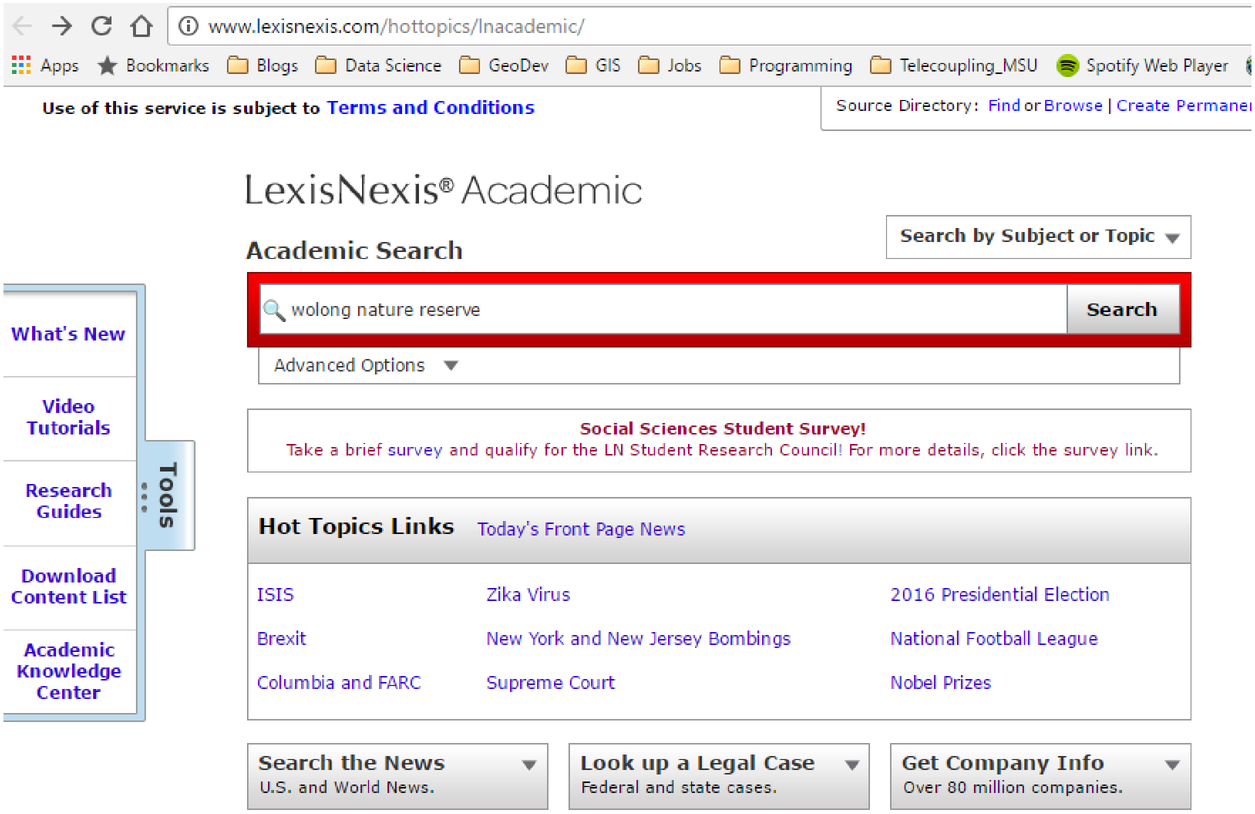
4.4.4 Add Media Information Flows



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/information/LexisNexis.html*" file provided with the sample dataset.

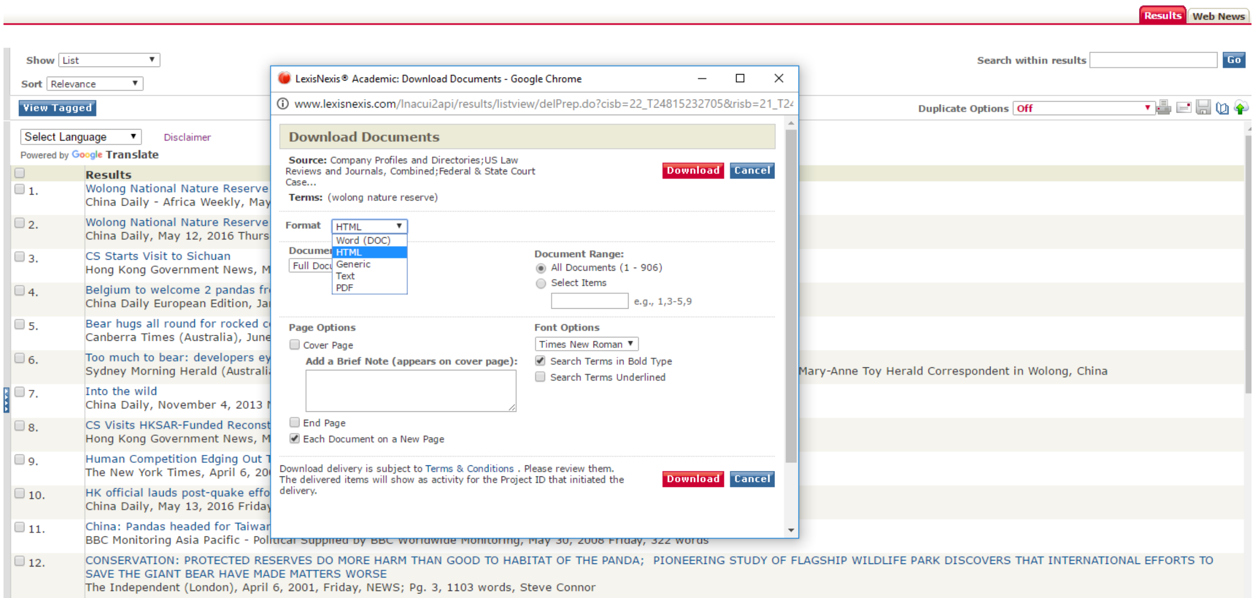
Go to the LexisNexis website[10], and you should see a window similar to the next figure:



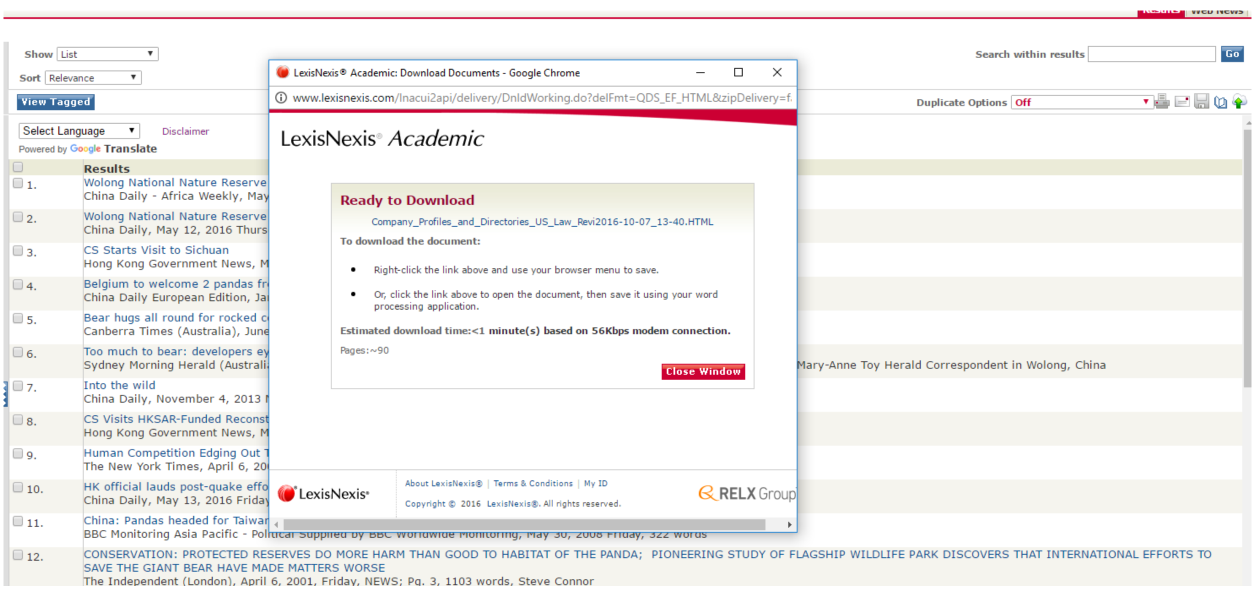
Inside the search box, type something like “wolong nature reserve”, and click the search button. The result page should list a number of records that were found to contain the searched term:

[10] 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:



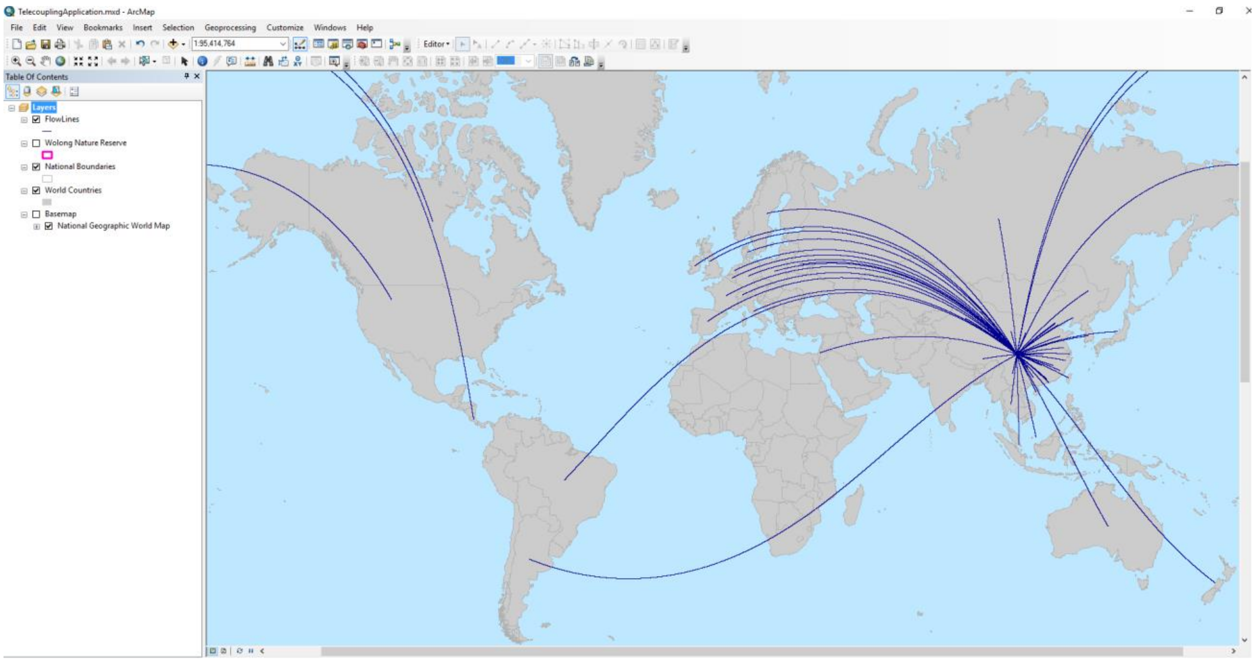
In the “Format” drop-down menu, select HTML. The document range tells you how many records have been found and whether you want to save all of them (default) to the HTML file. If you have too many records, an error message should come up asking to reduce the number of selected documents. In this example, 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.

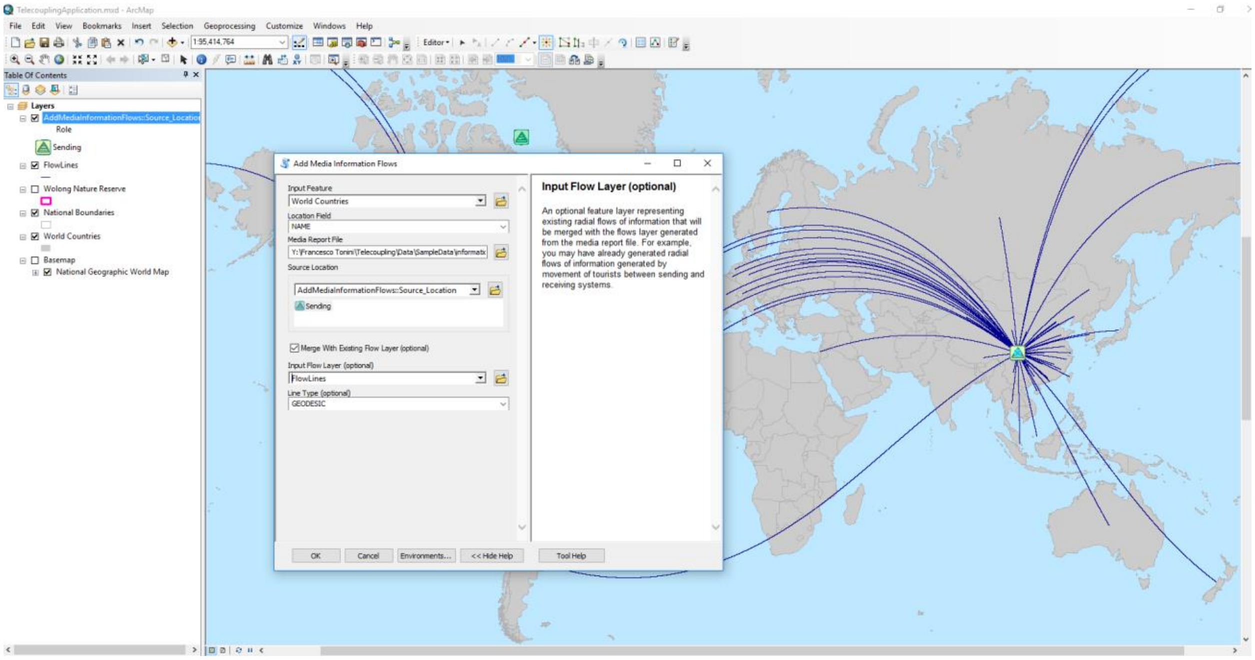
If you decided to skip the LexisNexis online word search, continue from this point on. For this tutorial, we are going to merge the media flows obtained from the HTML report file with an existing flow layer. To do this, use the "[Draw Radial Flows]" tool, and use the "*./SampleData/info_Flows.csv*" table 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.**

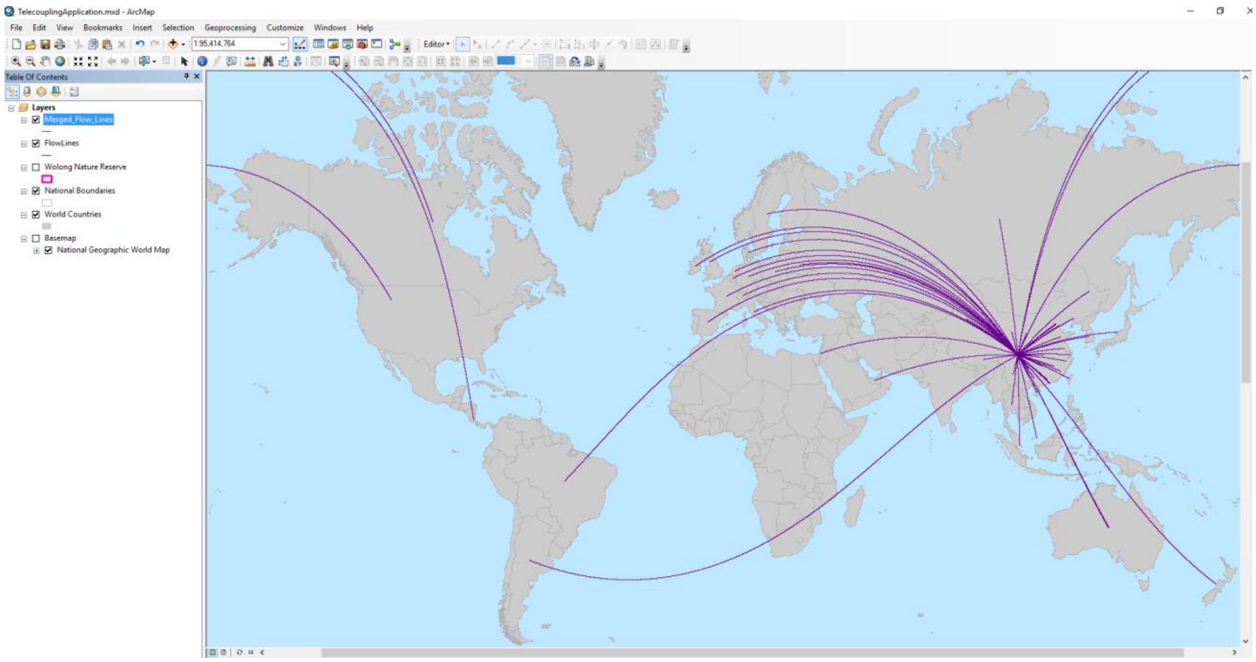


Now we are ready to use the “Add Media Information” tool. If you are using the *TelecouplingApplication.mxd* file, select “World Countries” as the input feature layer for the first parameter of the tool. If you are using a brand new map, then select the “*./SampleData/admin/World_countries_2002.shp*” 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/information/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.

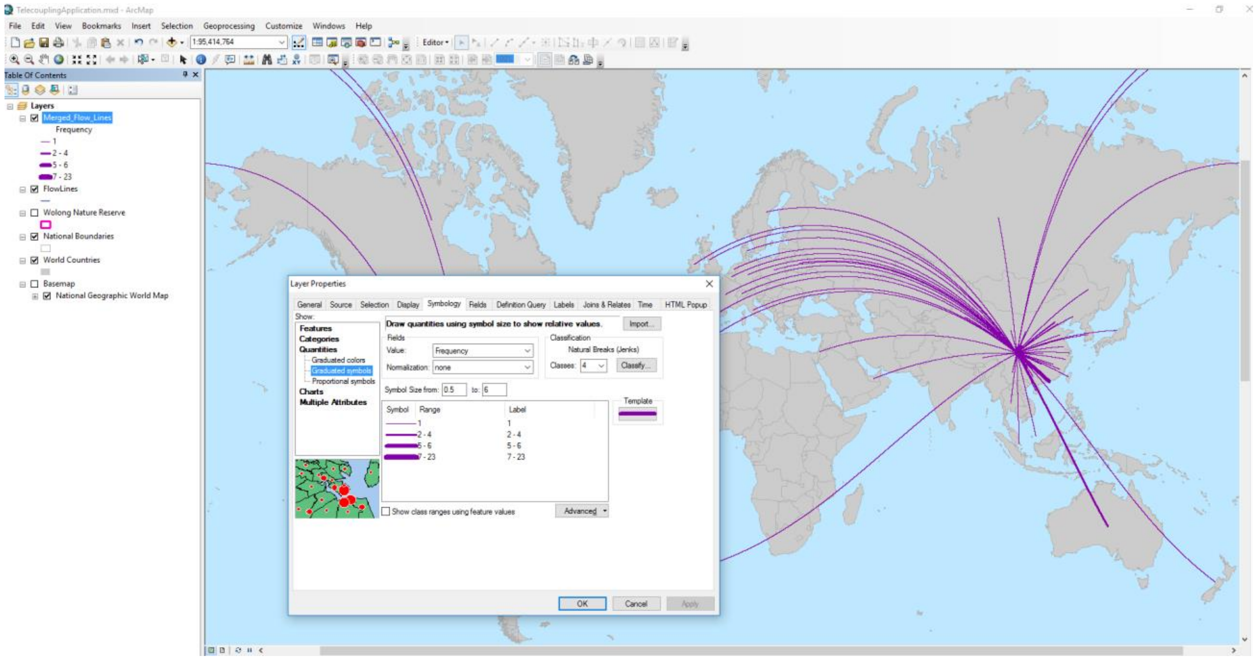
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



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.

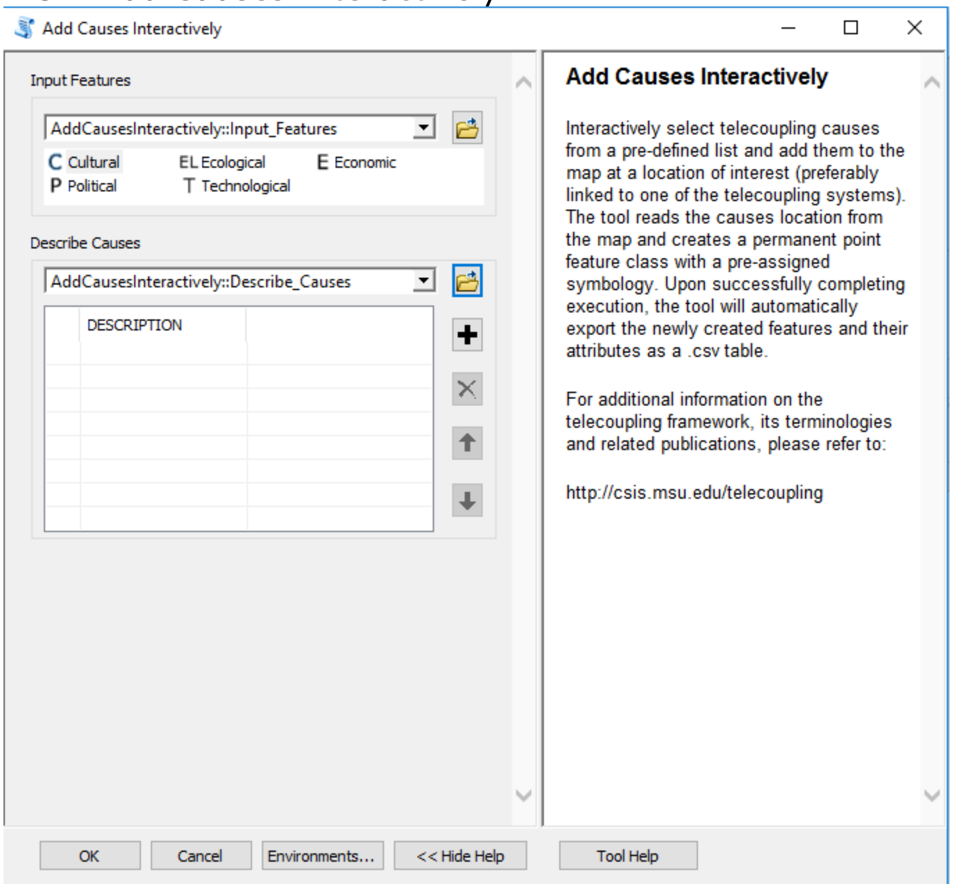


Once the output flow feature class has been created, the lines will all have the same color and thickness. If you have selected additional variables to represent the flow amount, you can use one of them to improve the visualization of the lines. To do so, right-click on the output flow layer found in 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

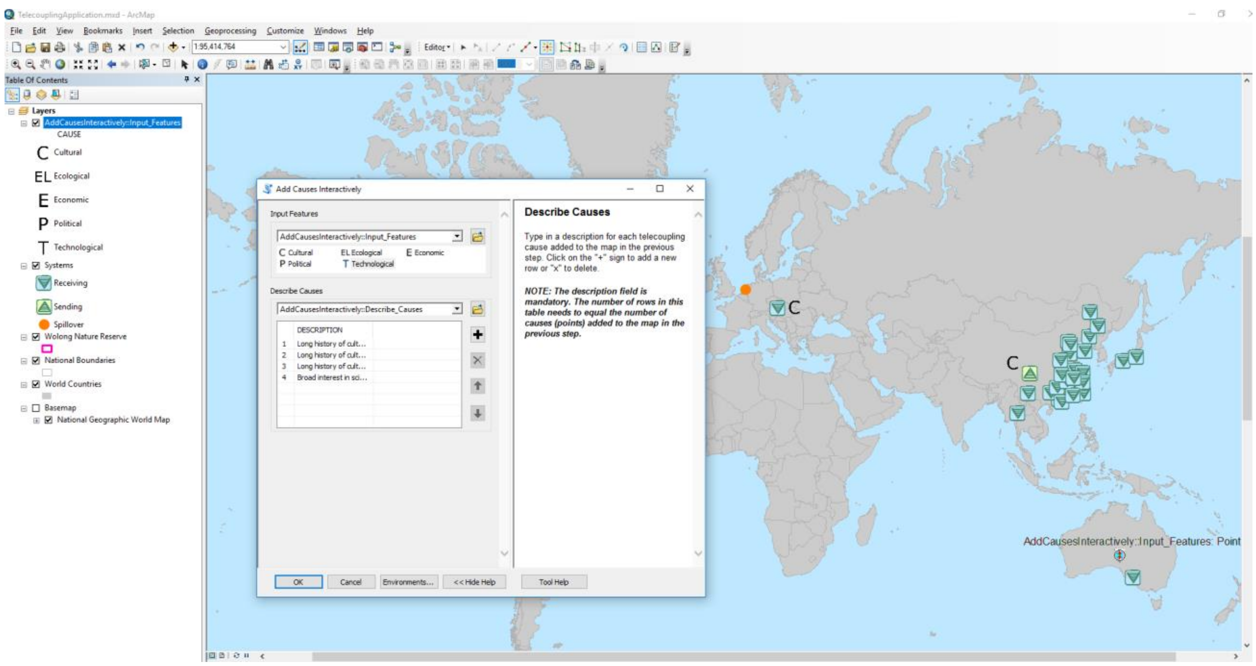
4.5.1 Add Causes Interactively



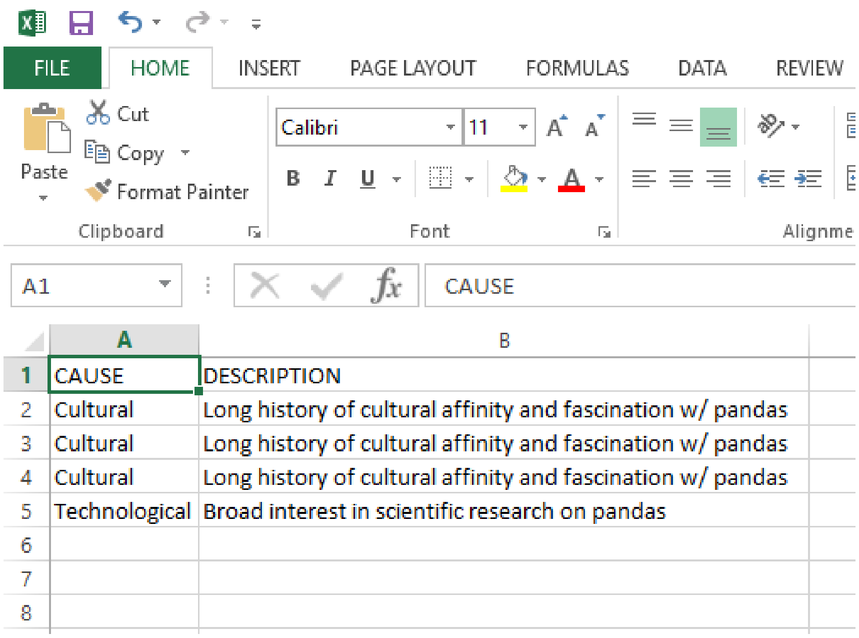
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

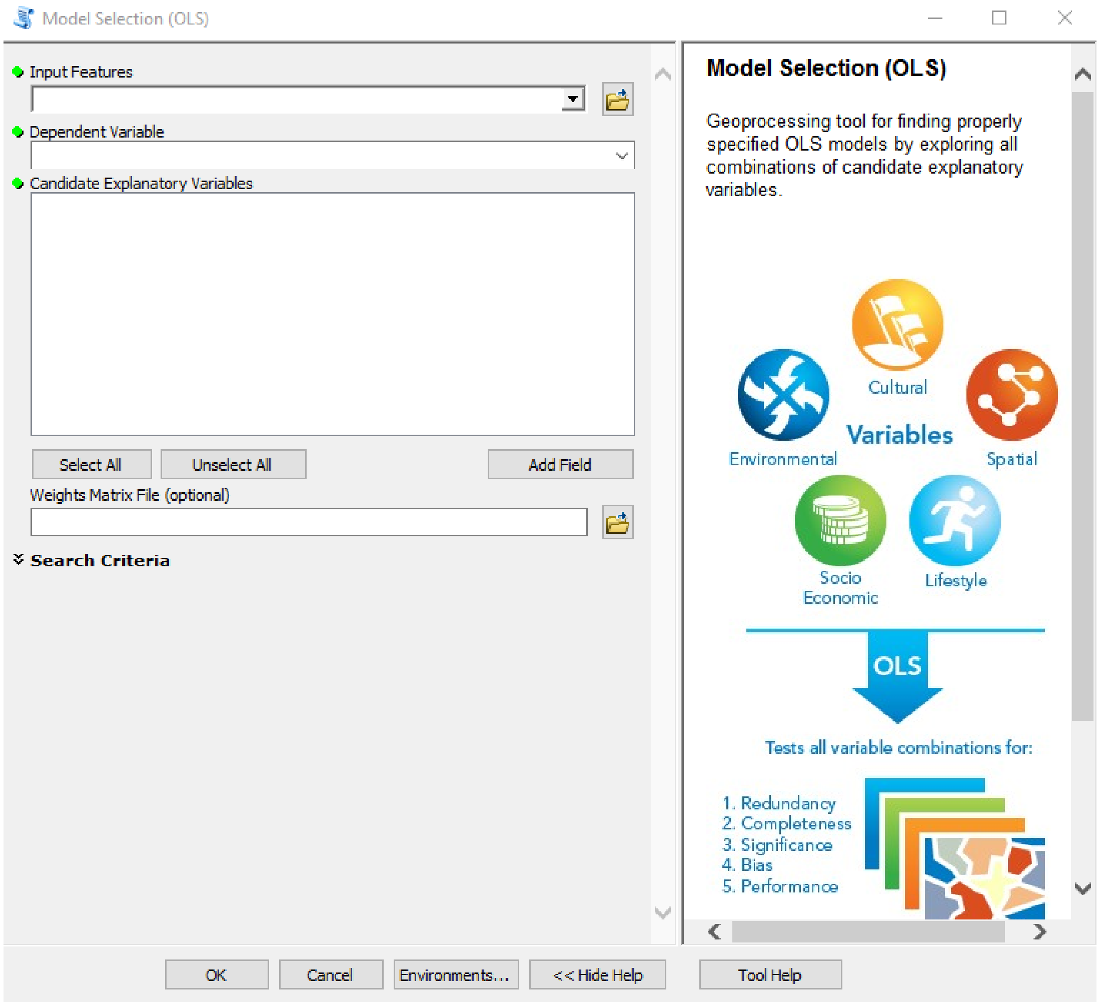


The output .csv table created by the tool in the scratch workspace directory should look something similar to the following figure:



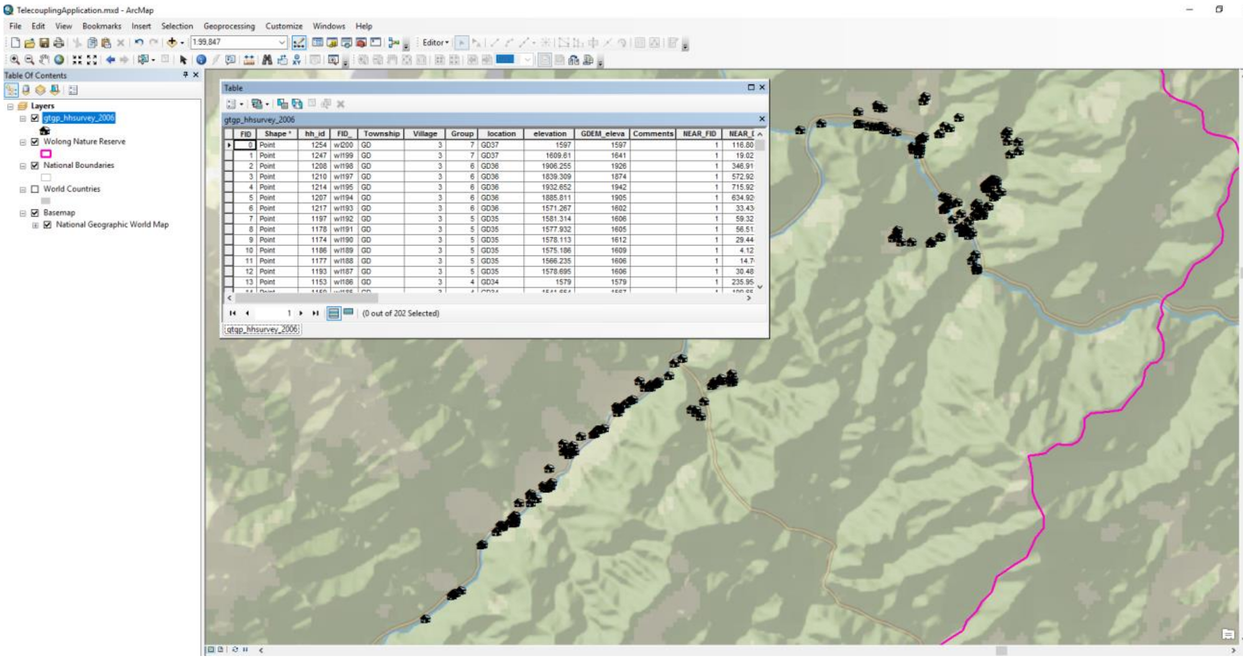
	A	B
1	CAUSE	DESCRIPTION
2	Cultural	Long history of cultural affinity and fascination w/ pandas
3	Cultural	Long history of cultural affinity and fascination w/ pandas
4	Cultural	Long history of cultural affinity and fascination w/ pandas
5	Technological	Broad interest in scientific research on pandas
6		
7		
8		

4.5.2 Model Selection (OLS)

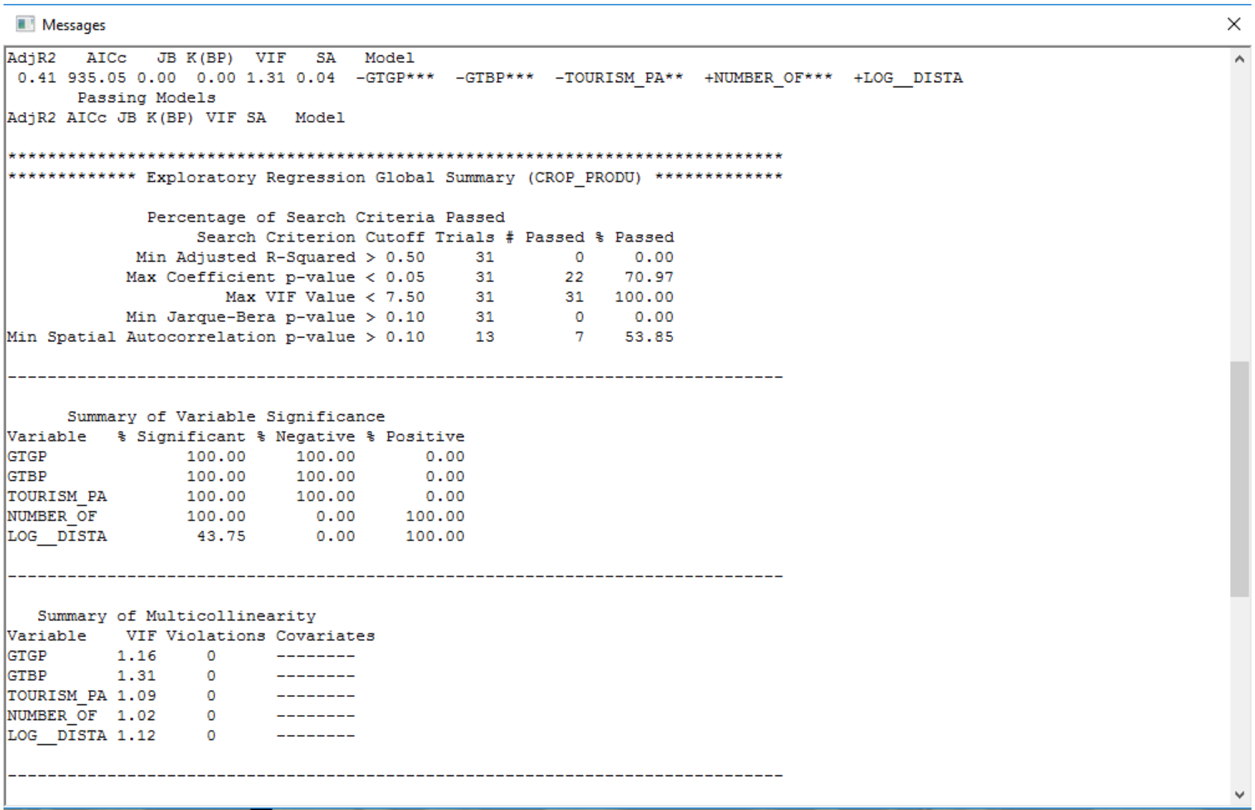


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

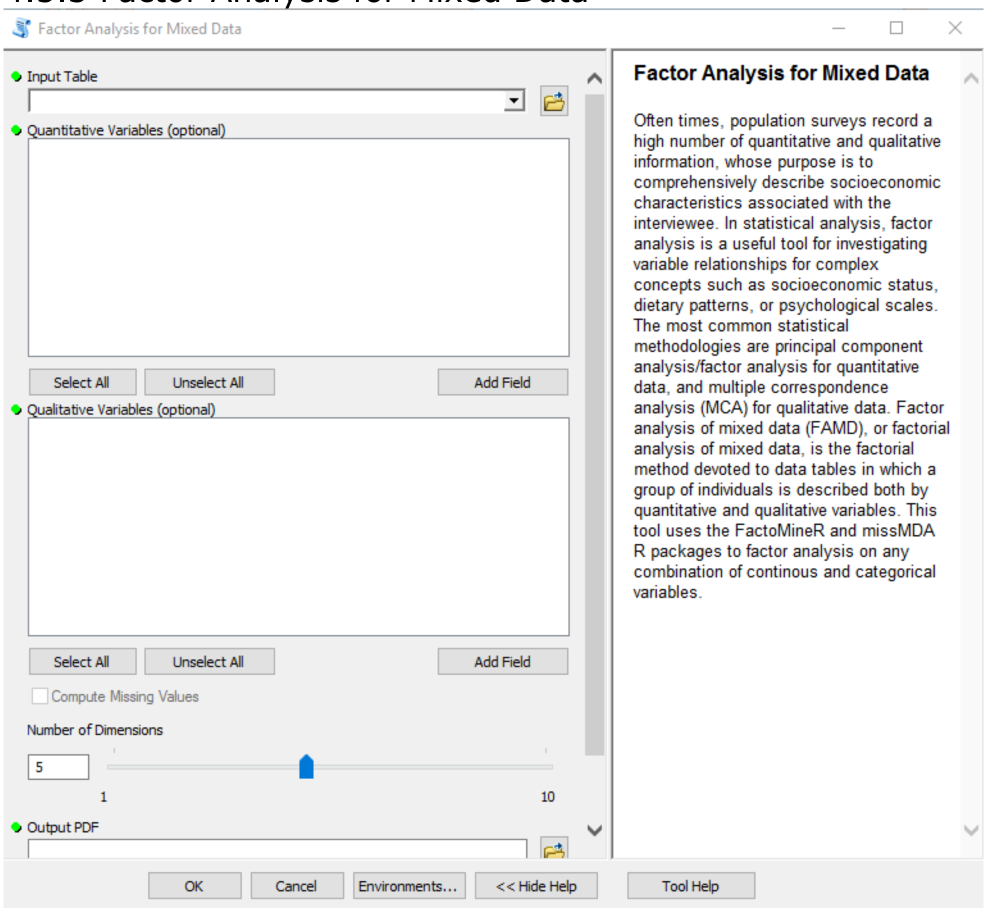


```
Messages
AdjR2   AICc   JB K(BP)  VIF   SA   Model
 0.41 935.05 0.00  0.00 1.31 0.04  -GTGP***  -GTBP***  -TOURISM_PA**  +NUMBER_OF***  +LOG__DISTA
      Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model

******************************************************************************
************* 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.

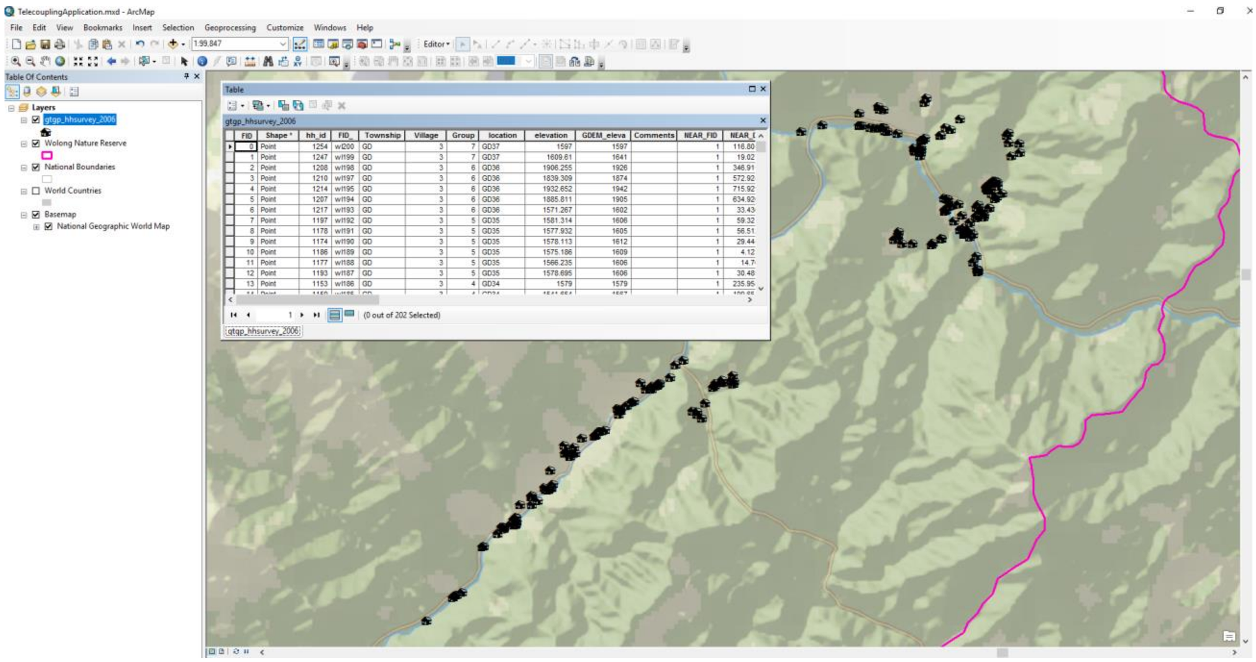
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

4.5.3 Factor Analysis for Mixed Data

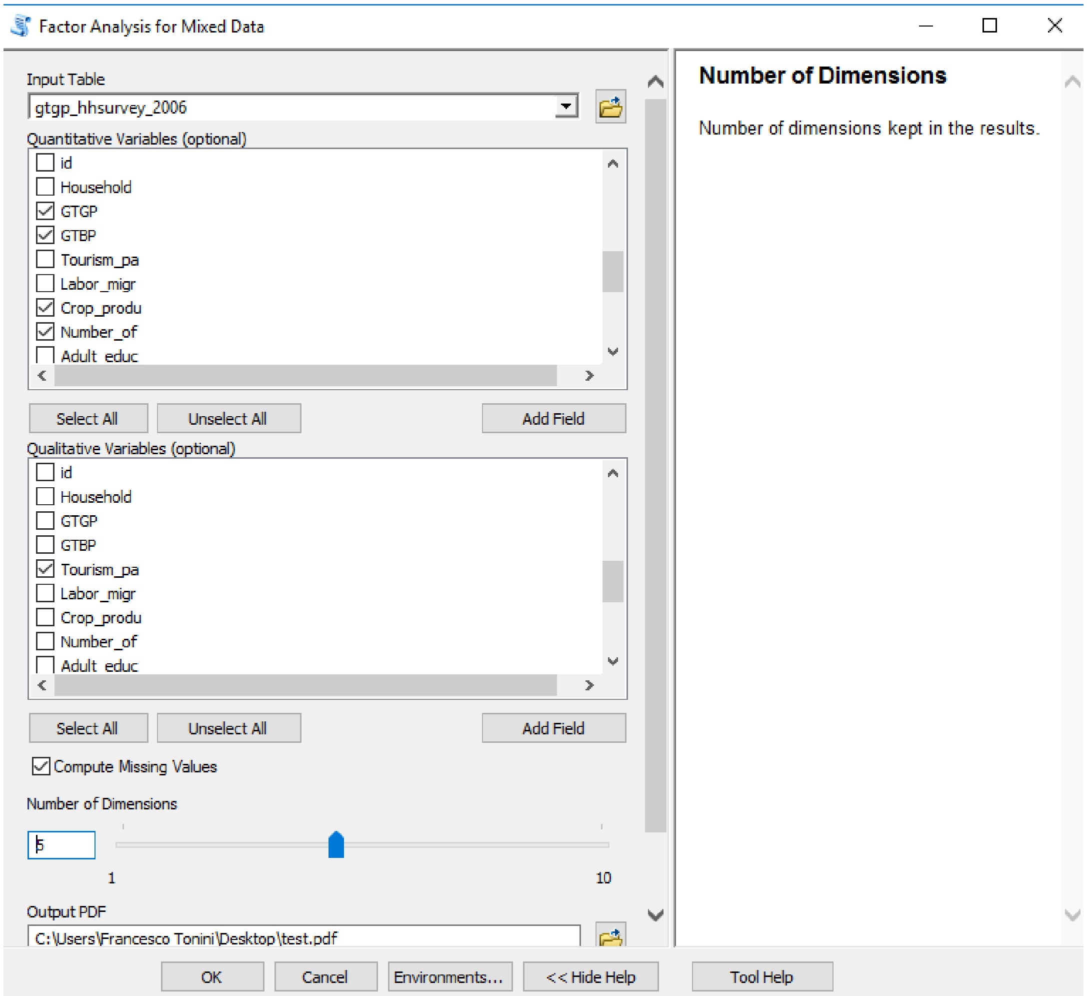


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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



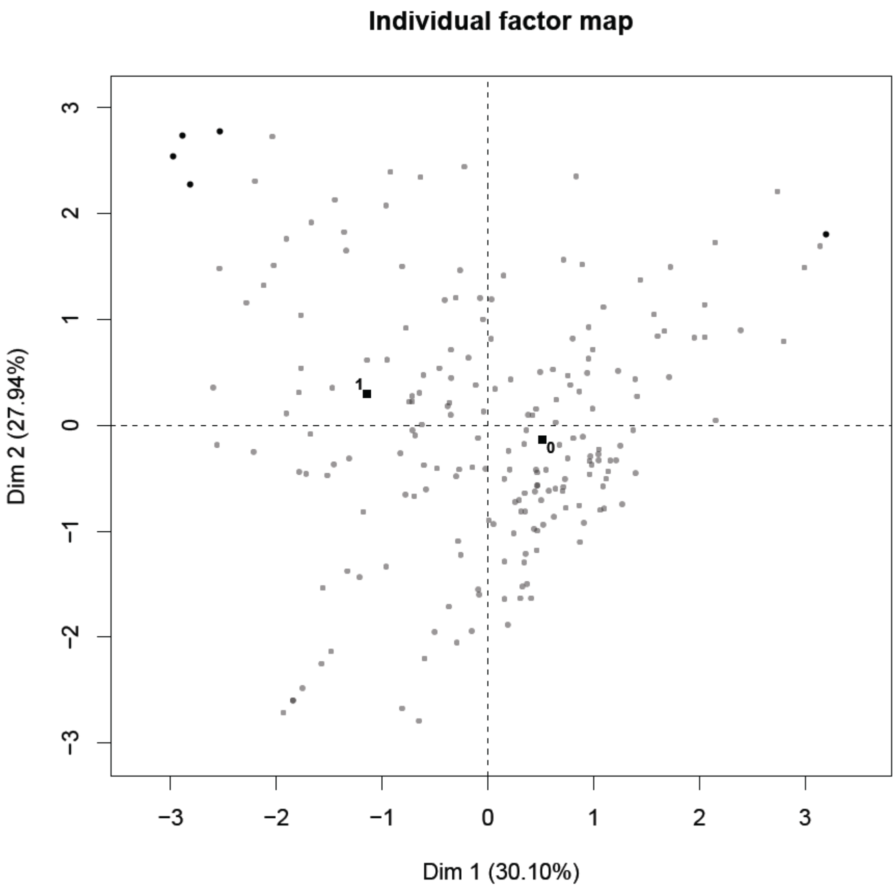
<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

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:

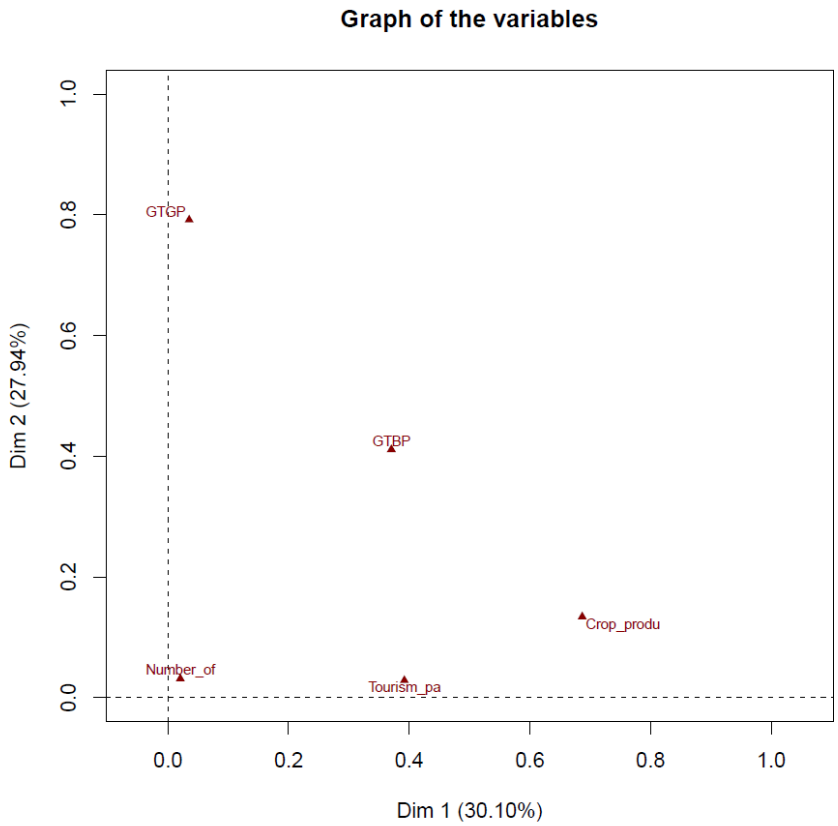
Messages

```
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
```

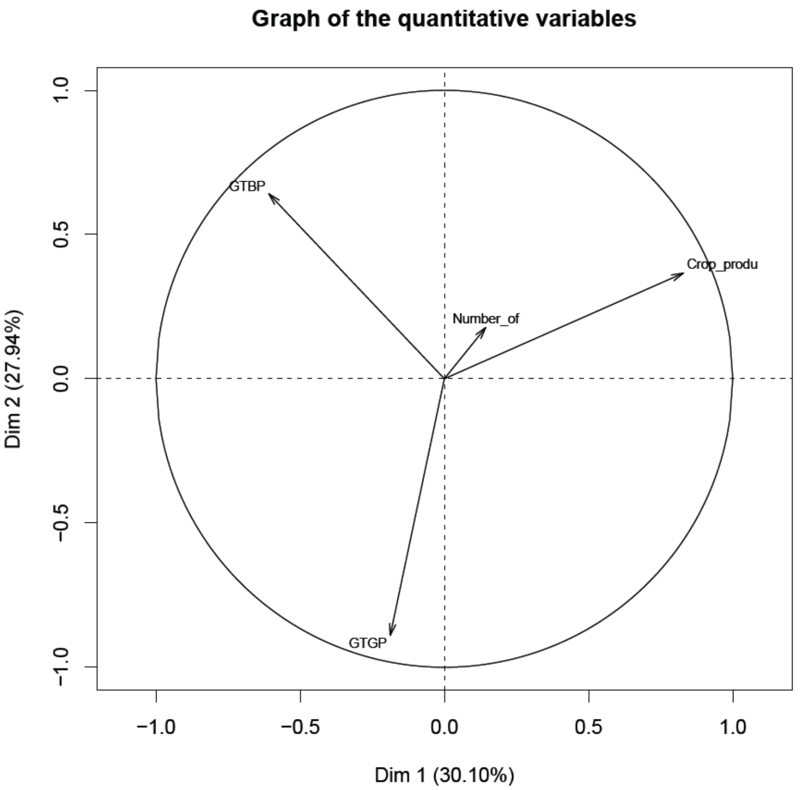
Within each extracted dimension, higher values indicate a higher contribution of an individual or variable to the composition of that dimension. If distinct groups of individual units (e.g. households) or variables exist, higher scores will show for different units/variables in separate dimensions. In 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.

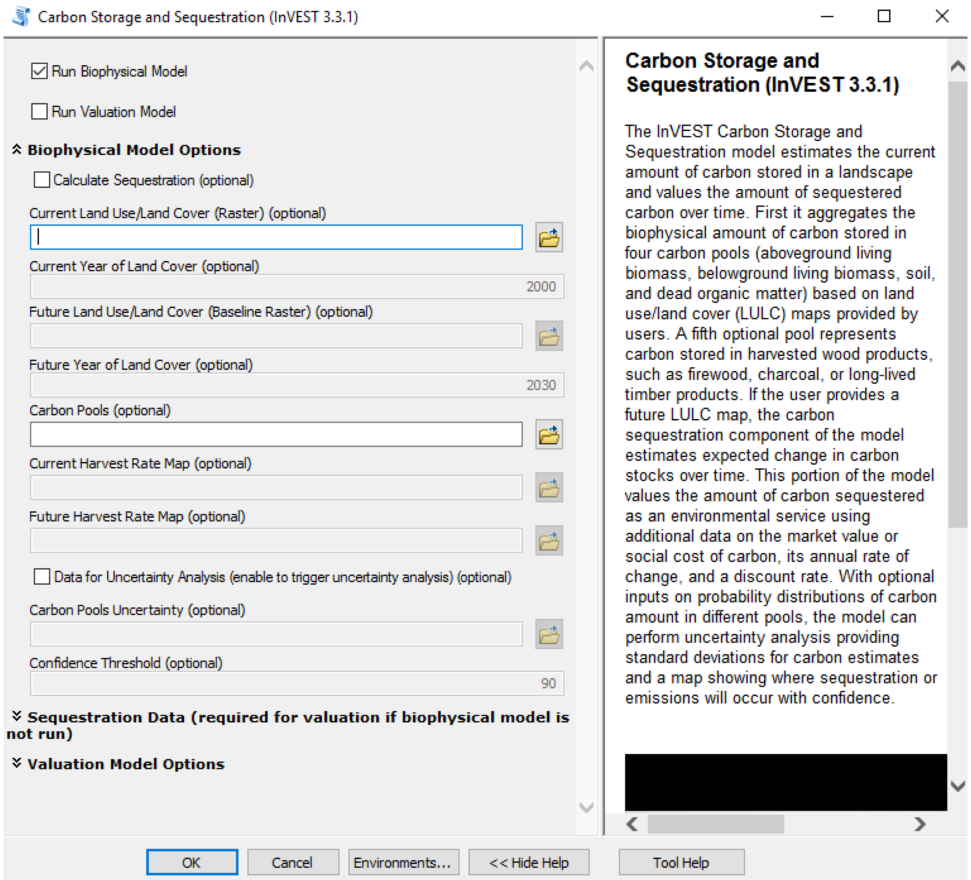


The graph of the quantitative variables on the unit circle, tells which quantitative variables are mostly correlated with each other as well as with the first two dimensions. The "GTBP" has a negative correlation with the first dimension and a positive one with the second one, while the angle almost 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 EFFECTS TOOLSET

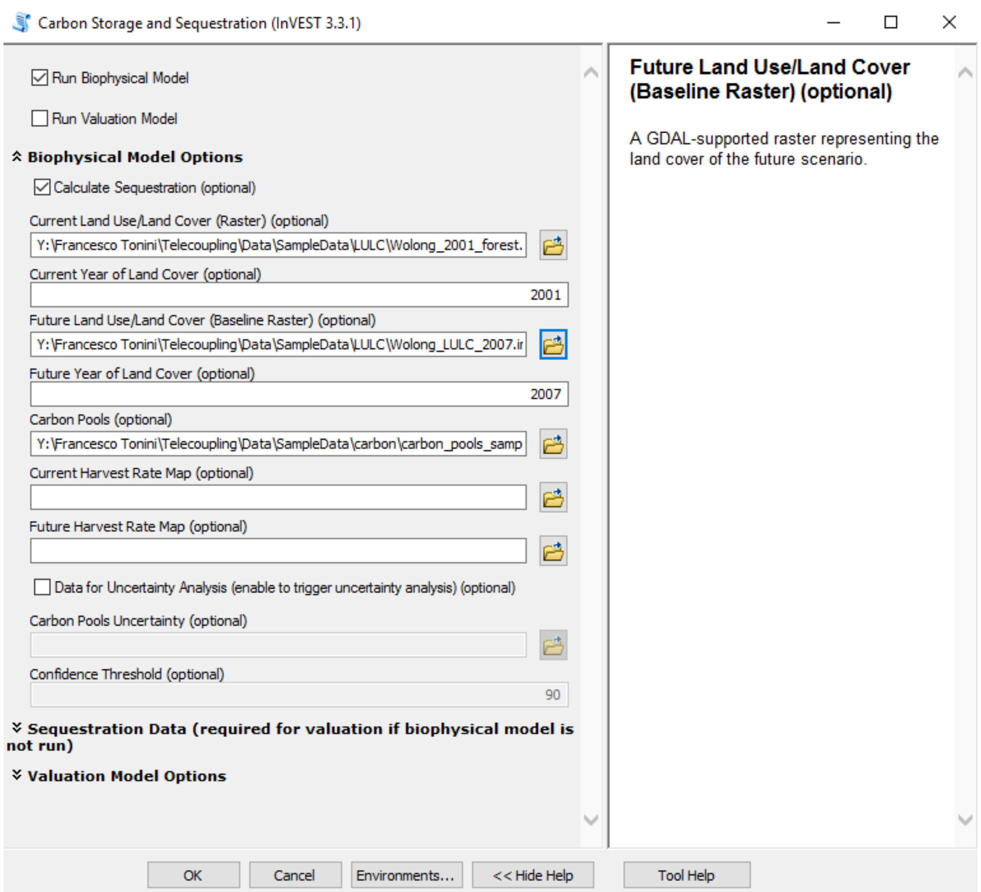
4.6.1 Carbon Storage and Sequestration (InVEST 3.3.1)



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.

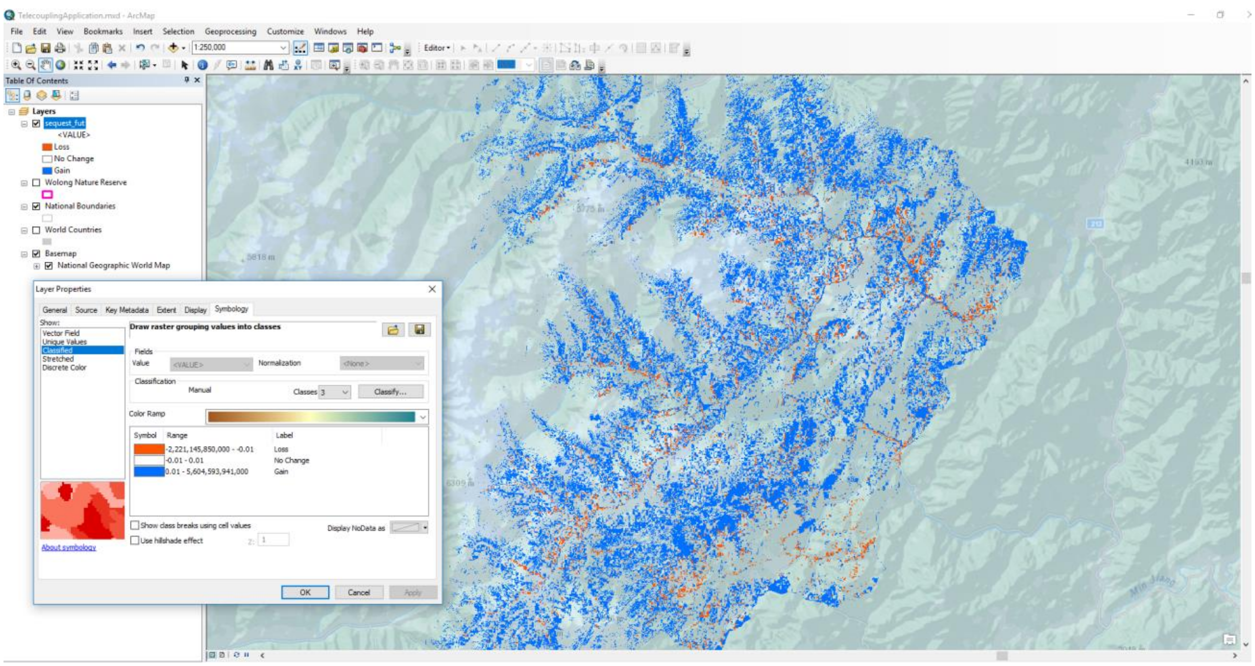
This tools is linked to the InVEST 3.3.1 carbon sequestration model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this example, let's use the land use/land-cover raster files from 2001 and 2007 for the Wolong Nature Reserve area ("*./SampleData/LULC/Wolong_LULC_2001.img*", "*.../Wolong_LULC_2007.img*") to calculate carbon sequestration between two years (2001-2007). The tool also needs a carbon pool table input ("*./SampleData/carbon/carbon_pools_samp_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.

NOTE: although the tool shows "Carbon Pools" parameter as (optional), this is a bug in ArcGIS which currently does not allow to turn off/on the parameter type (required vs optional) when selecting other options in the tool.

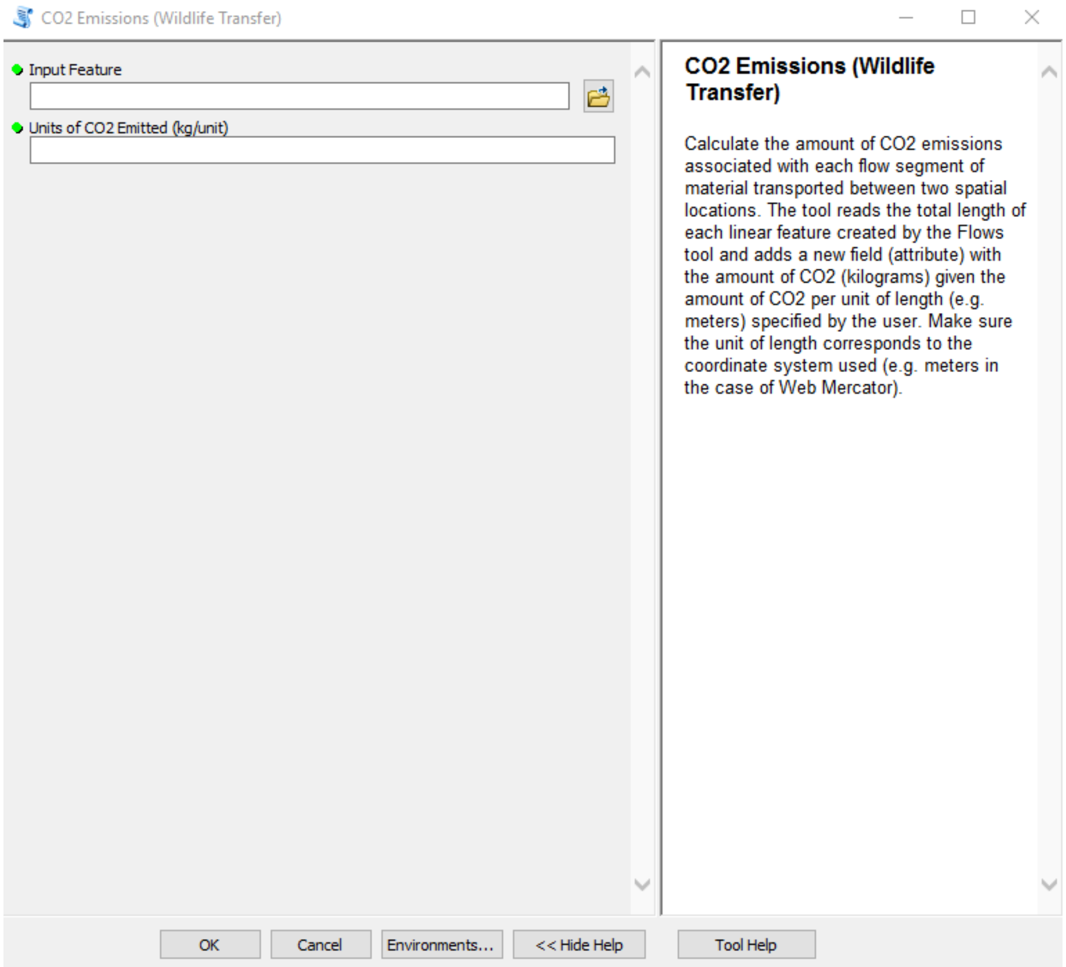


<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

Leave all other options as default and click OK to run the tool. Once completed, you should have three different output rasters. Current carbon values (2001) and future (2007) ones, as well as carbon sequestered between the two years provided. Negative differences correspond to losses in carbon while positive values corresponds to gains. To better visualize gain and losses we recommend to use the symbology tab in the raster layer properties (right-click on the layer). You can use "Classified" as a symbology and choose 3 classes, then click on "Classify" to manually define your intervals. In this case, let's set up the first interval to "-0.01" and the second to "0.01". This way, any very small negative or positive value that may and artifact of the biophysical model computation can be grouped into a separate class. Then, double-click on the color assigned to this middle class and select "no color" (or something very neutral). You can also use the Label column to change the text that will show in your map table of content. To make it more useful, let's change the label of the negative class to "Loss" and the positive values class to "Gain". Then, change the neutral one to "No Change". Click on OK and you should see something like the following image:



4.6.2 CO_2 Emissions (Wildlife Transfer)

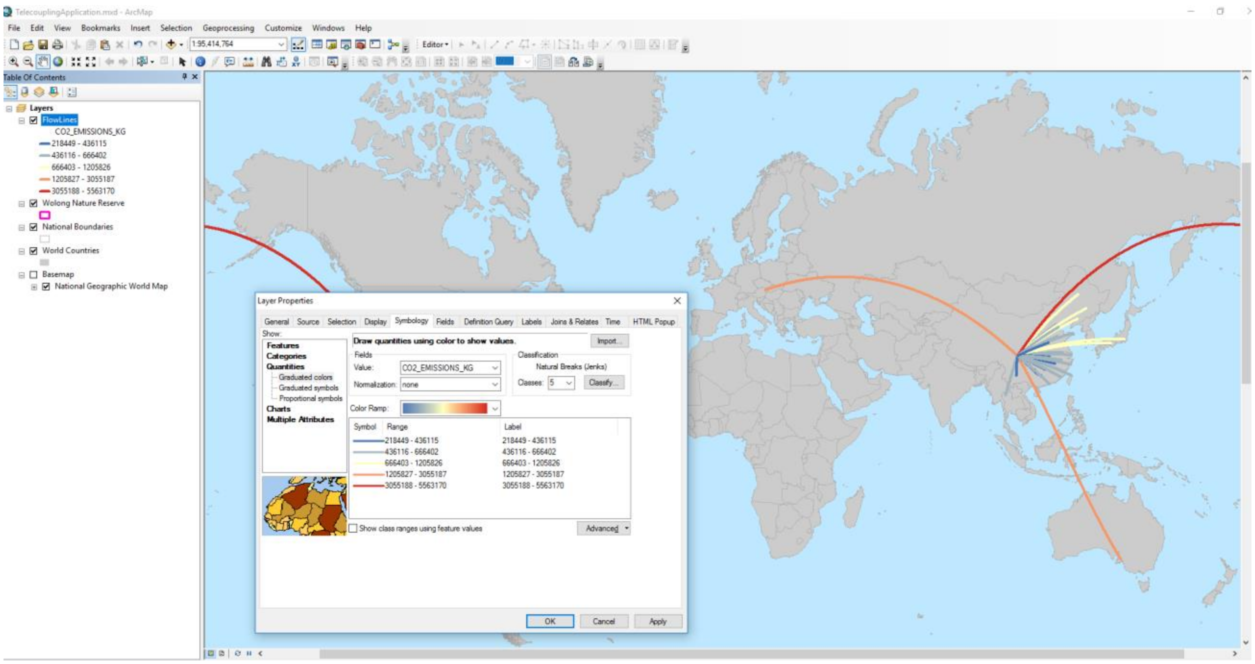


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

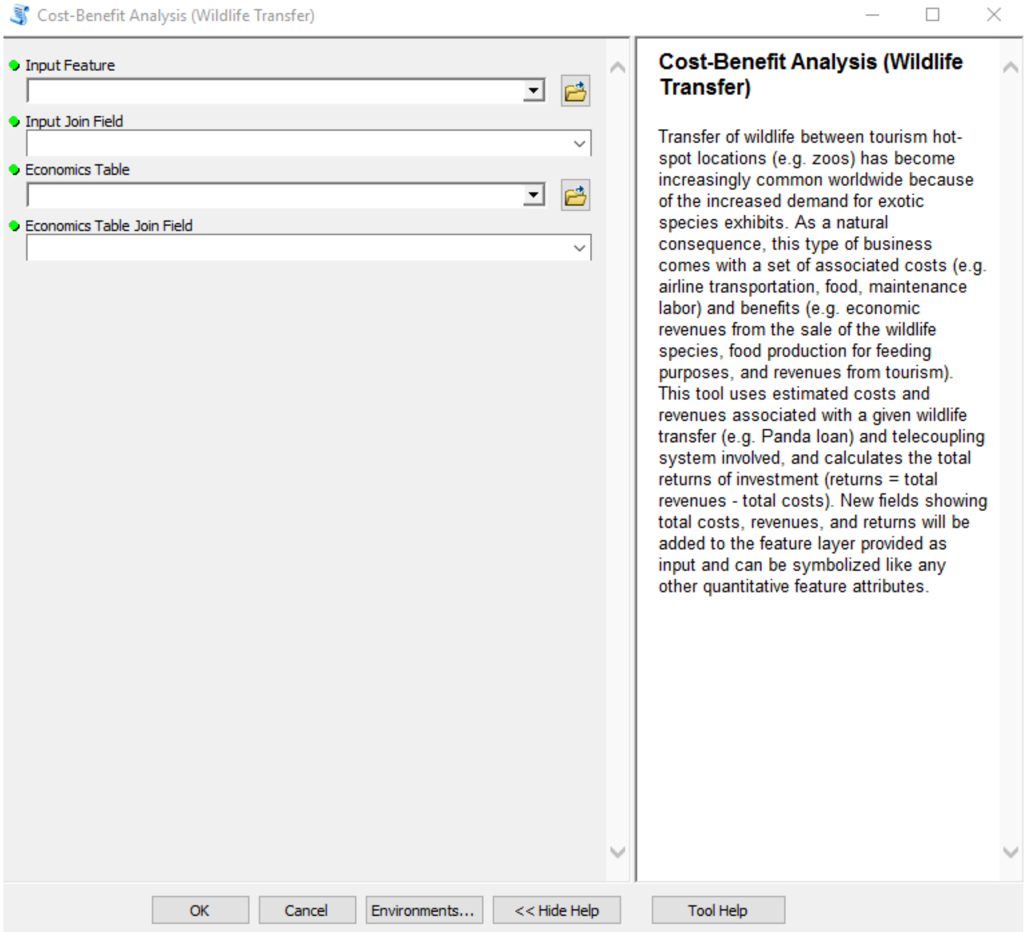
This tool uses the flow layer computed with the Draw Radial Flows tool and calculated the amount of CO_2 emissions based on the length of the flow lines. For this example, make sure you first computed the radial flows and select that layer as input for the first parameter here. Then, type in an estimated 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:

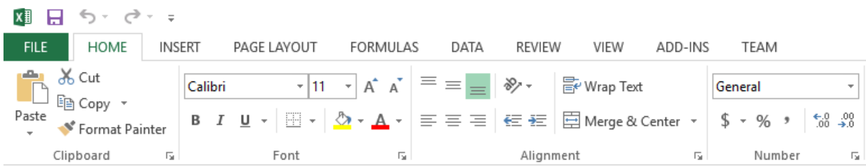


4.6.3 Cost-Benefit Analysis (Wildlife Transfer)



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 case, 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/wildlife/wildlife_financial_data.csv*":



	A	B	C	D	E	F	G	H
1	NAME	cost_per_ai	cost_feedin	cost_transp	cost_maint	revenue_fr	revenue_fr	revenue_from_food_prod
2	Wolong	0	n/a	0	n/a	86000000	n/a	n/a
3	Adelaide	2000000	100000	0	10000000	n/a	n/a	n/a
4	Beijing	4000000	100000	0	10000000	n/a	n/a	n/a
5	Benxi	1000000	100000	0	10000000	n/a	n/a	n/a
6	Chiang Mai	3000000	100000	0	10000000	n/a	n/a	n/a
7	Fuzhou	5000000	100000	0	10000000	n/a	n/a	n/a
8	Guangzhou	10000000	100000	0	10000000	n/a	n/a	n/a
9	Guilin	3000000	100000	0	10000000	n/a	n/a	n/a
10	Hengyang	1000000	100000	0	10000000	n/a	n/a	n/a
11	Hongkong	4000000	100000	0	10000000	n/a	n/a	n/a
12	Jilin	1000000	100000	0	10000000	n/a	n/a	n/a
13	Jinan	1000000	100000	0	10000000	n/a	n/a	n/a
14	Kobe	2000000	100000	0	10000000	n/a	n/a	n/a
15	Kunming	3000000	100000	0	10000000	n/a	n/a	n/a
16	Nanchang	1000000	100000	0	10000000	n/a	n/a	n/a
17	Nanjing	3000000	100000	0	10000000	n/a	n/a	n/a
18	Qiqihar	1000000	100000	0	10000000	n/a	n/a	n/a
19	San Diego	5000000	100000	0	10000000	n/a	n/a	n/a
20	Shanghai	6000000	100000	0	10000000	n/a	n/a	n/a
21	Taiwan	2000000	100000	0	10000000	n/a	n/a	n/a
22	Tianjin	1000000	100000	0	10000000	n/a	n/a	n/a
23	Uneco Park	10000000	100000	0	10000000	n/a	n/a	n/a
24	Vienna	3000000	100000	0	10000000	n/a	n/a	n/a
25	Washingtor	2000000	100000	0	10000000	n/a	n/a	n/a
26	Wenling	3000000	100000	0	10000000	n/a	n/a	n/a
27	Wuhan	2000000	100000	0	10000000	n/a	n/a	n/a
28	Xiamen	1000000	100000	0	10000000	n/a	n/a	n/a
29	Xiuning	3000000	100000	0	10000000	n/a	n/a	n/a
30	Xixiakou	2000000	100000	0	10000000	n/a	n/a	n/a
31	Zhengzhou	1000000	100000	0	10000000	n/a	n/a	n/a
32	Holland	0	0	0	0	0	0	100000
33								

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.

Table

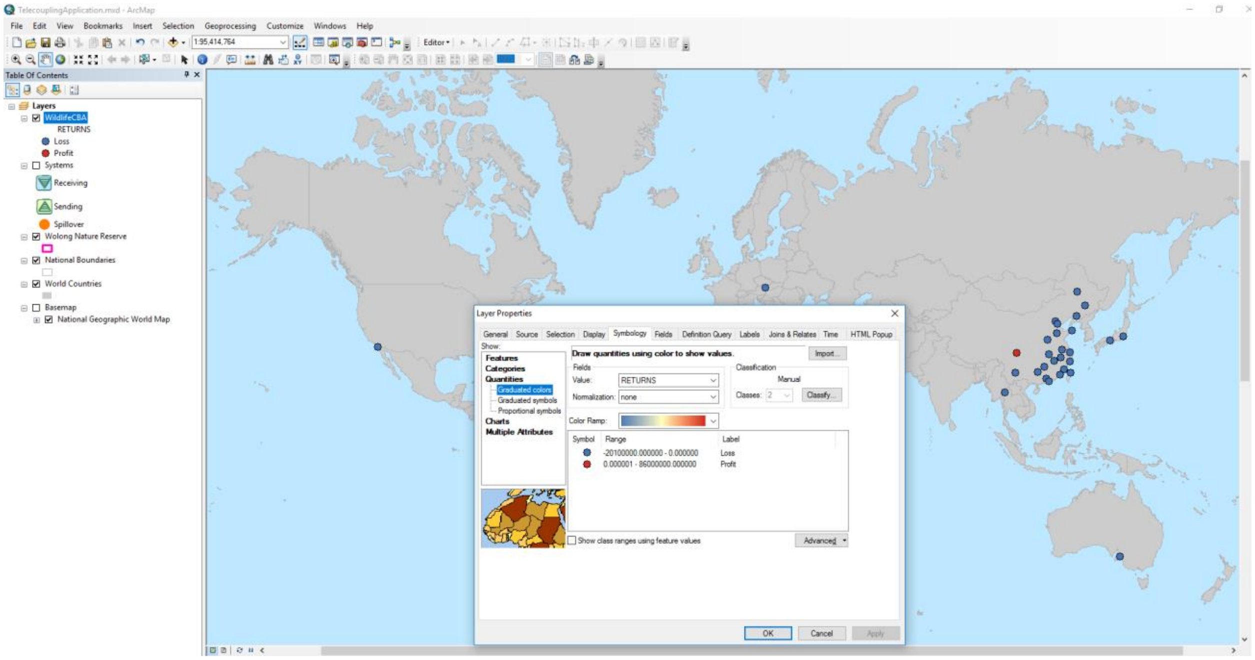
WildlifeCBA

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

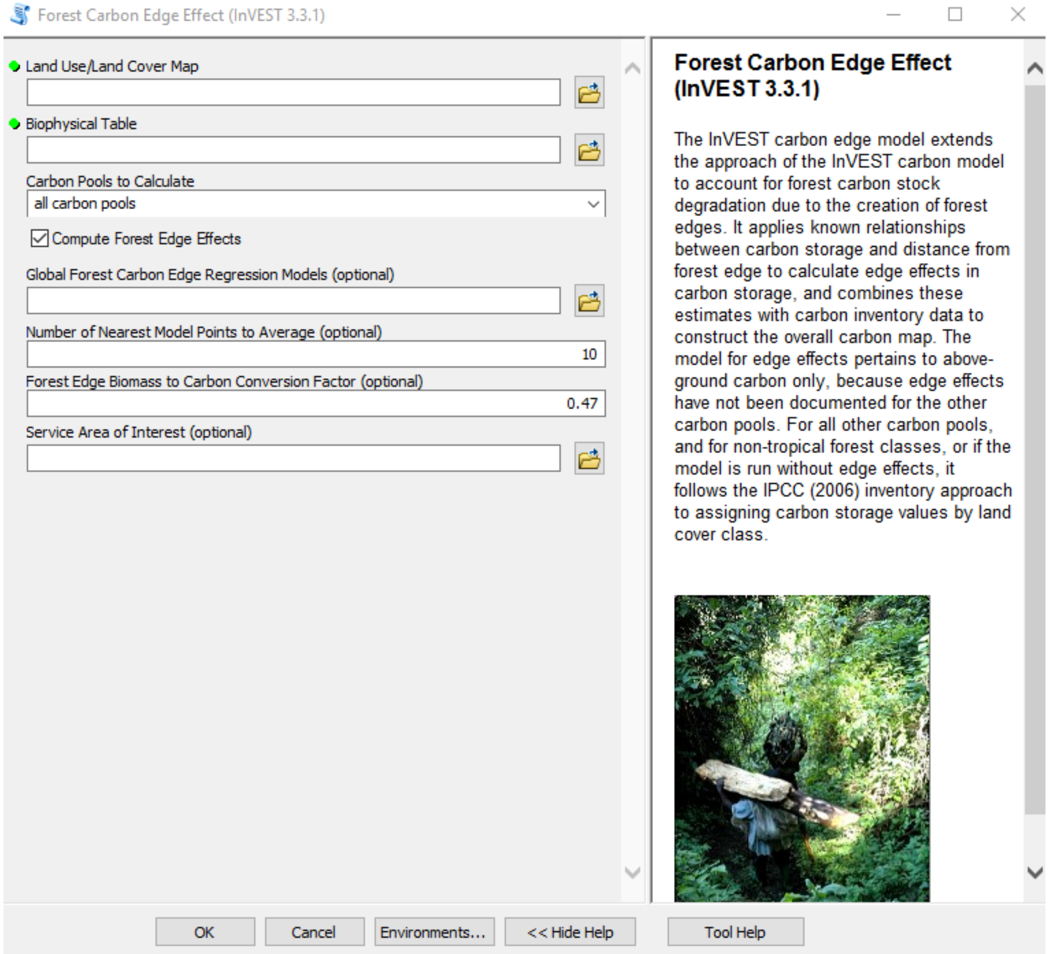
1 (0 out of 30 Selected)

WildlifeCBA

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:


TelecouplingApplication.mxd - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
Table Of Contents
Layers
WildlifeCBA
RETURNS
Loss
Profit
Systems
Receiving
Sending
Spillover
Wolong Nature Reserve
National Boundaries
World Countries
Basemap
National Geographic World Map
Layer Properties
General Source Selection Display Symbology Fields Definition Query Labels Joins & Relates Time HTML Popup
Show:
Features
Categories
Quantities
Graduated colors
Graduated symbols
Proportional symbols
Charts
Multiple Attributes
Draw quantities using color to show values.
Import...
Fields
Value: RETURNS
Normalization: none
Classification
Manual
Classes: 2
Classify...
Color Ramp:
Symbol Range Label
-20100000.000000 - 0.000000 Loss
0.000001 - 86000000.000000 Profit
Show class ranges using feature values
Advanced
OK Cancel Apply

4.6.4 Forest Carbon Edge Effect (InVEST 3.3.1)

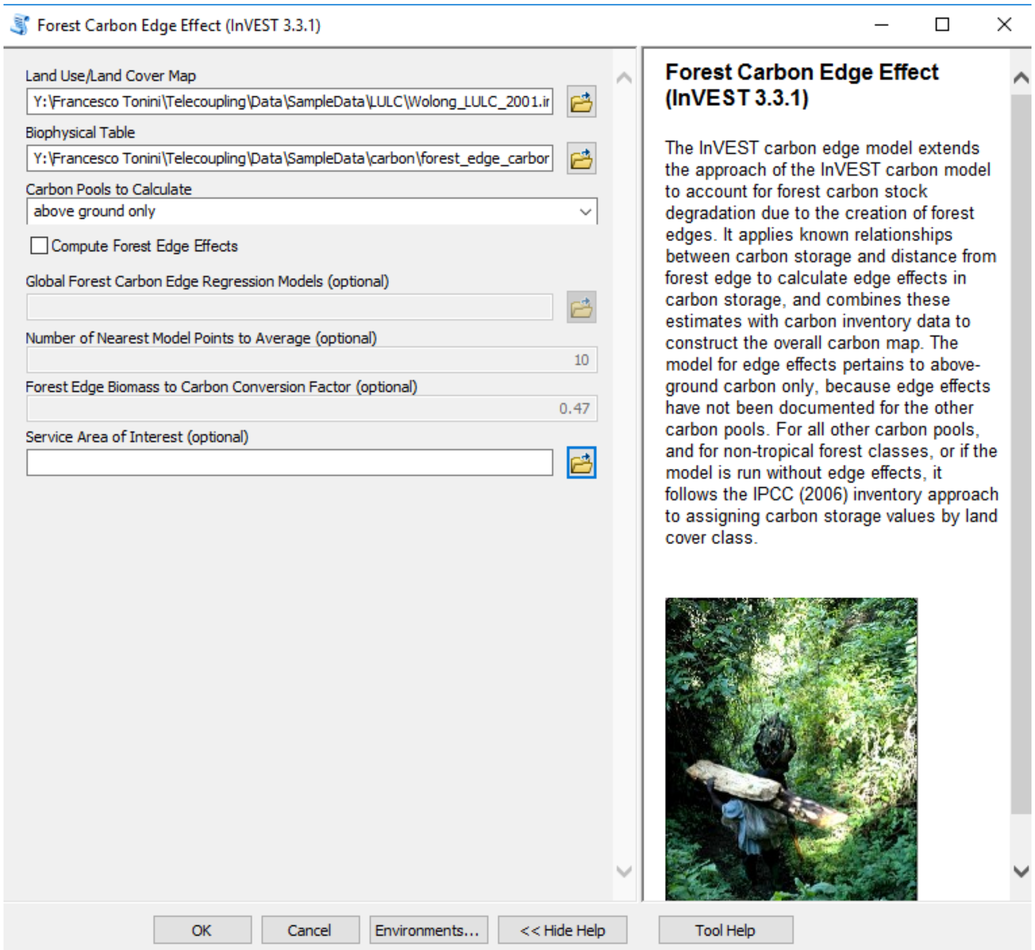


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

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

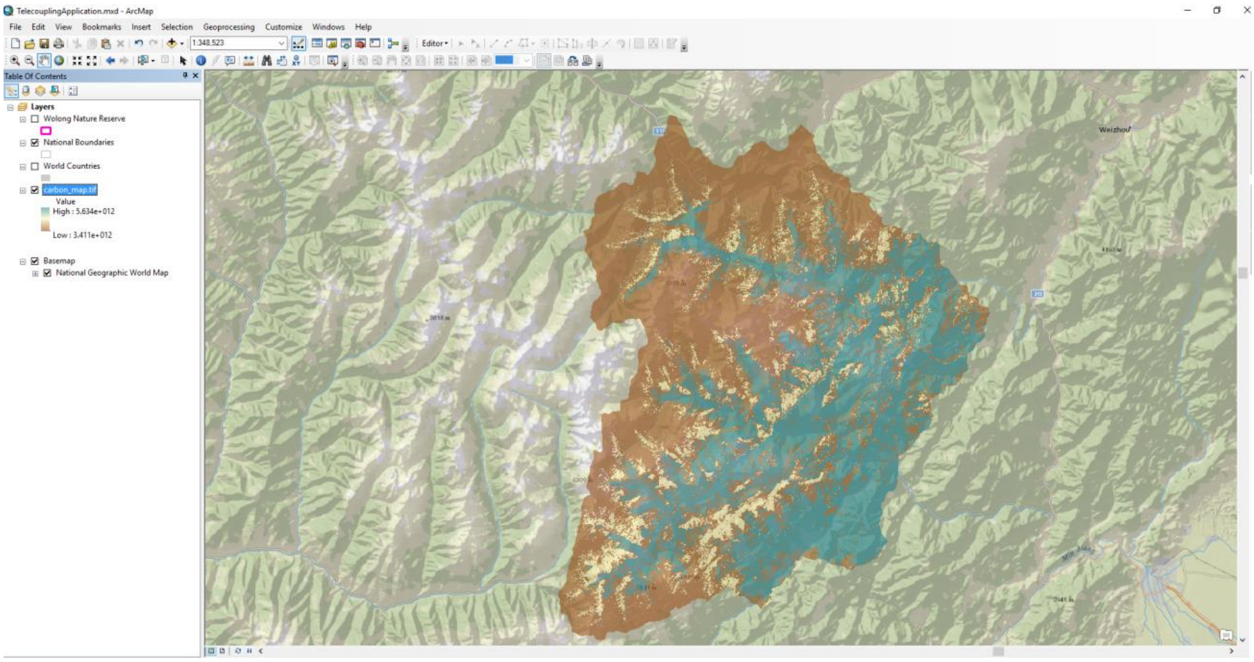
This tools is linked to the InVEST 3.3.1 forest carbon edge effect model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this example, let's use the land use/land-cover raster files from 2001 for the Wolong Nature Reserve area ("*./SampleData/LULC/Wolong_LULC_2001.img*") to calculate above-ground carbon amounts in 2001. The tool also needs a carbon pool table input ("*./SampleData/carbon/*

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.

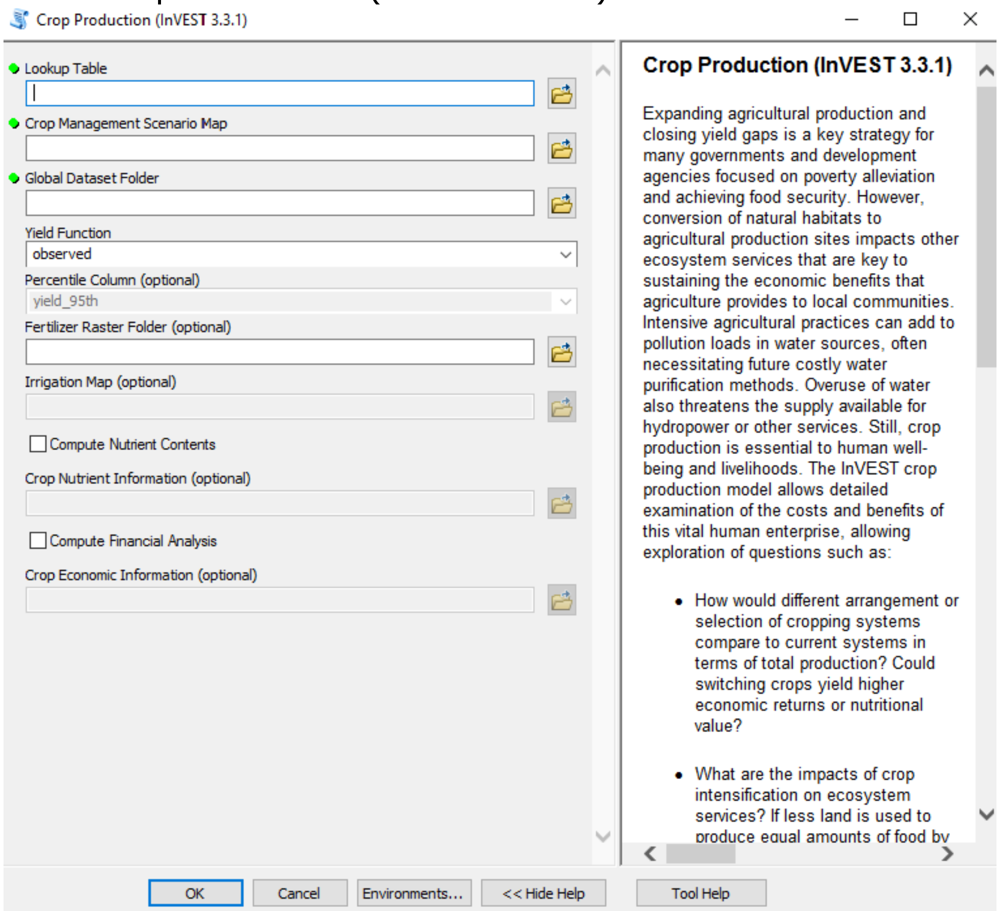


<u>NOTE</u>: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

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

4.6.5 Crop Production (InVEST 3.3.1)

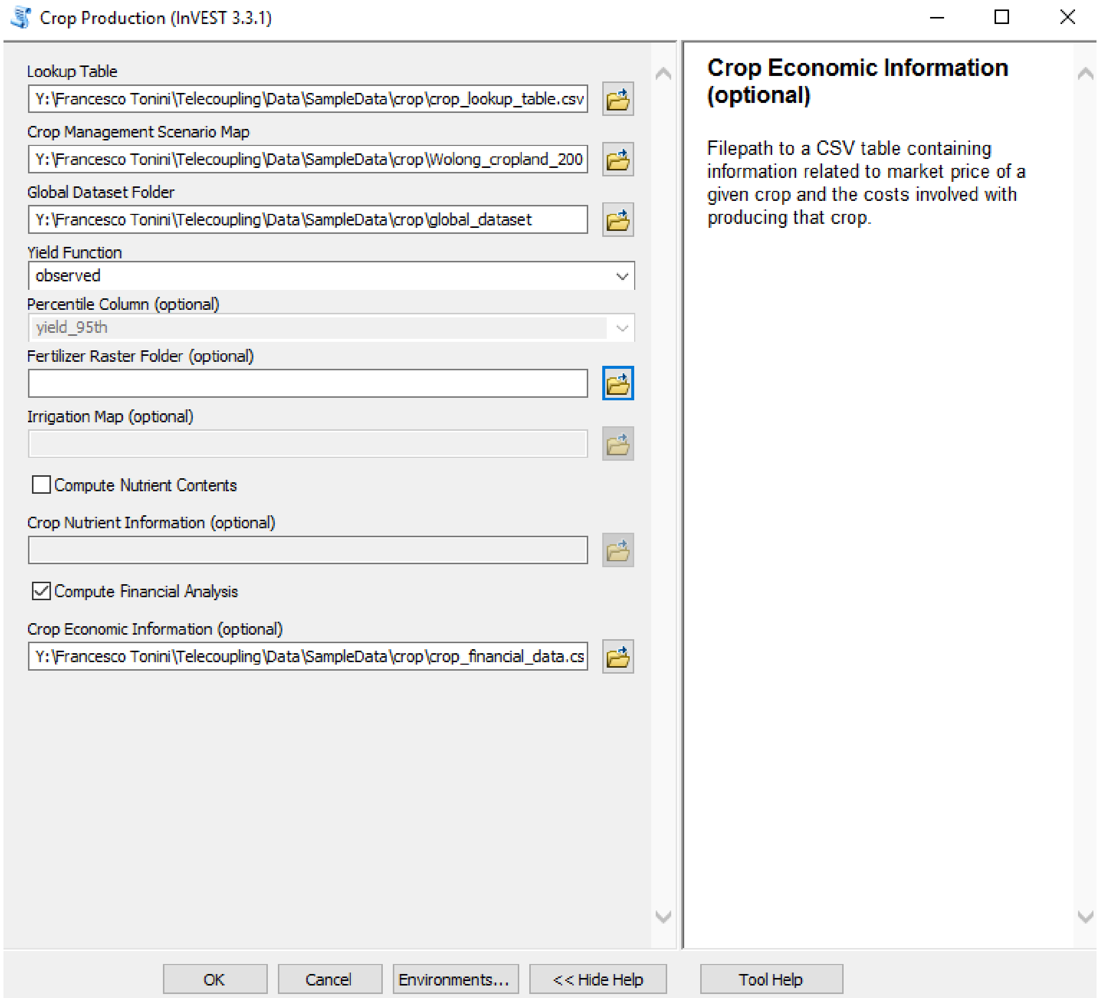


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

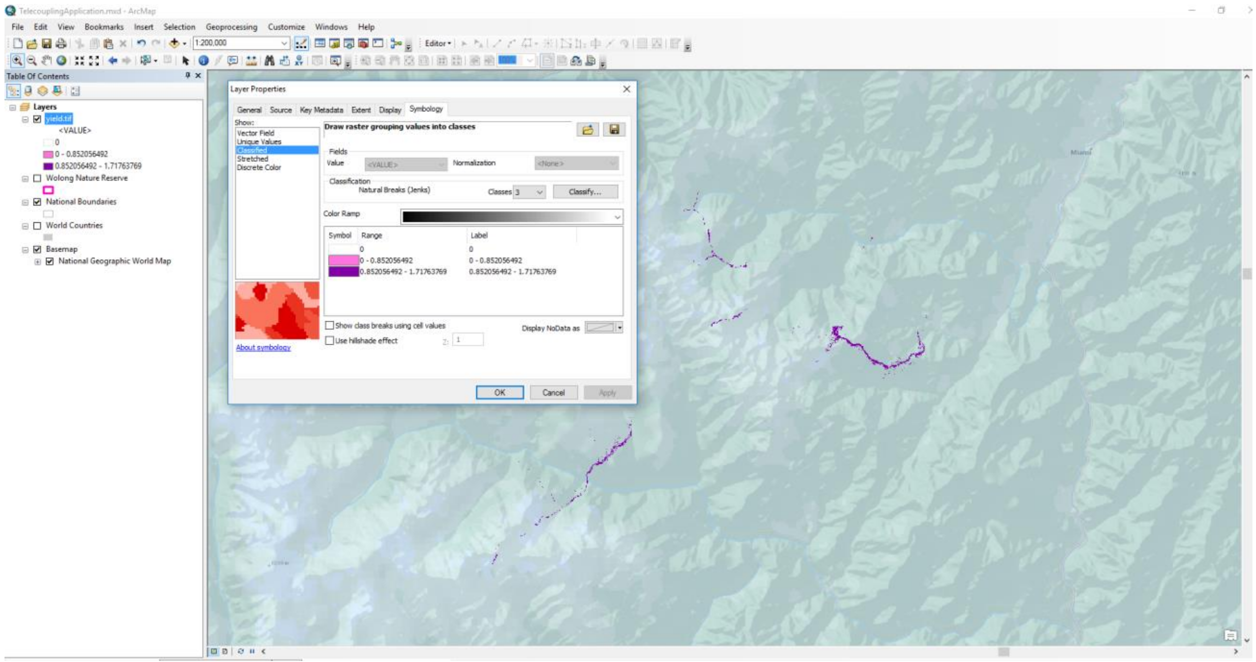
NOTE: At this very moment, this InVEST model is NOT recommended for real decision-making as it needs more validation. However, feel free to use it for exploratory and testing purposes.

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/crop/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/crop/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/crop/global_dataset*"). Check the box to compute financial analysis and specify the crop economic information table ("*./SampleData/crop/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/crop/crop_nutrient_data.csv*).

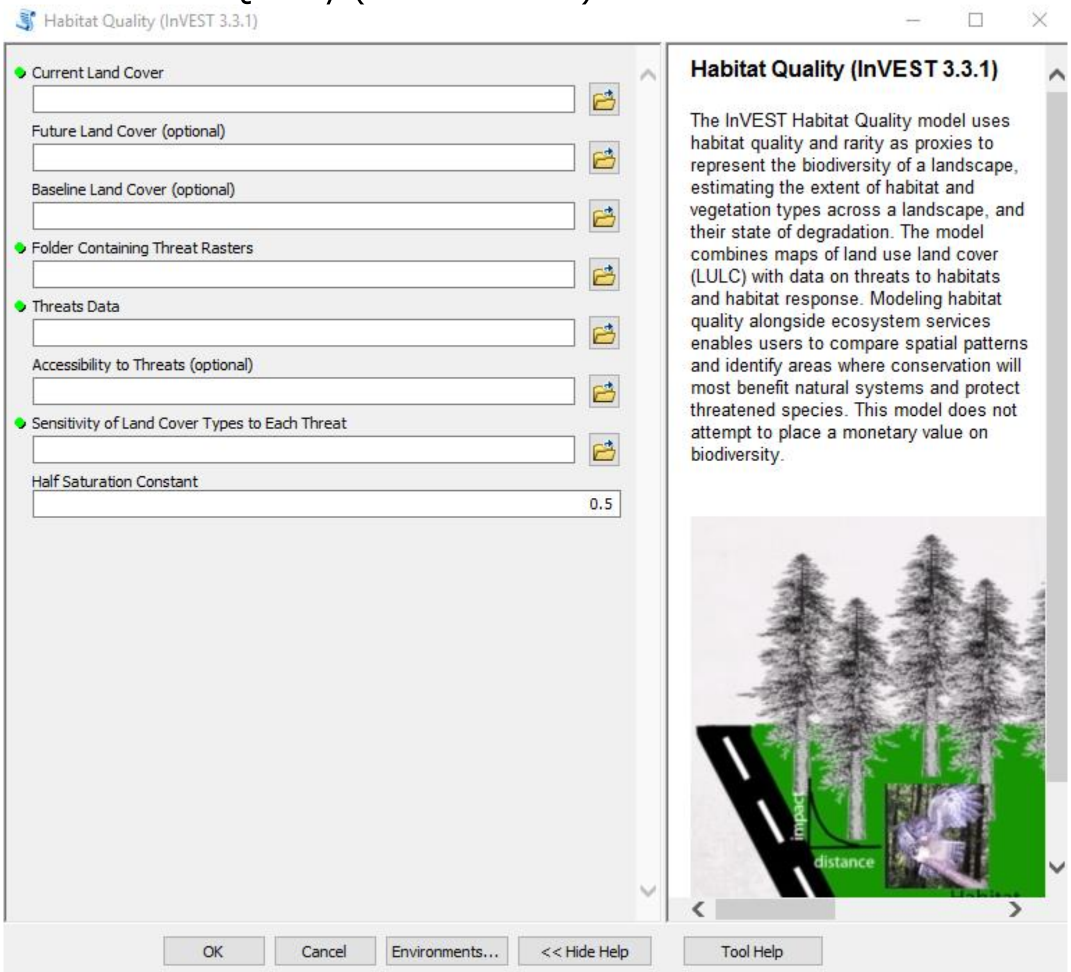
NOTE: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



Click OK 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 and assign “no color” to the class with values equal to 0. 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
1:200,000
Editor
Table Of Contents
Layers
yield.tif
<VALUE>
0
0 - 0.852056492
0.852056492 - 1.71763769
Wolong Nature Reserve
National Boundaries
World Countries
Basemap
National Geographic World Map
Layer Properties
General Source Key Metadata Extent Display Symbology
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 3
Classify...
Color Ramp
Symbol Range Label
0 0
0 - 0.852056492 0 - 0.852056492
0.852056492 - 1.71763769 0.852056492 - 1.71763769
Show class breaks using cell values
Use hillshade effect
Z: 1
Display NoData as
About symbology
OK Cancel Apply

4.6.6 Habitat Quality (InVEST 3.3.1)

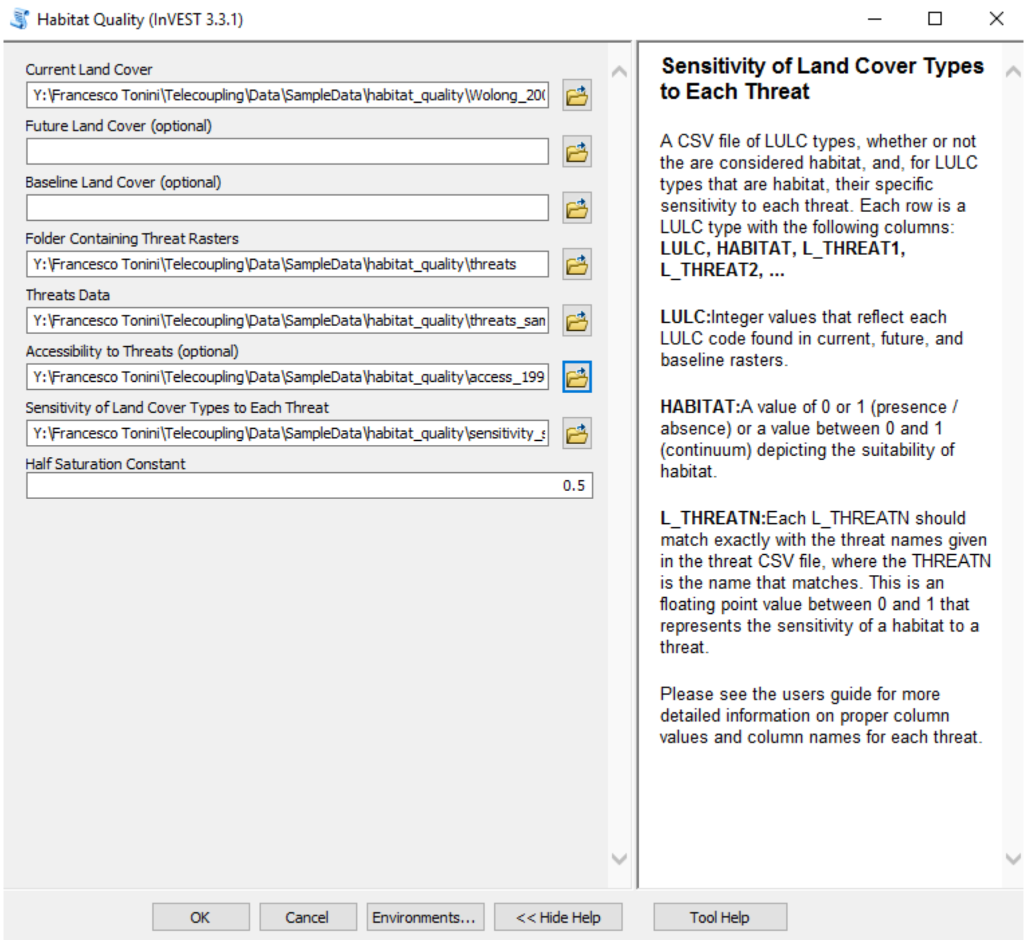


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

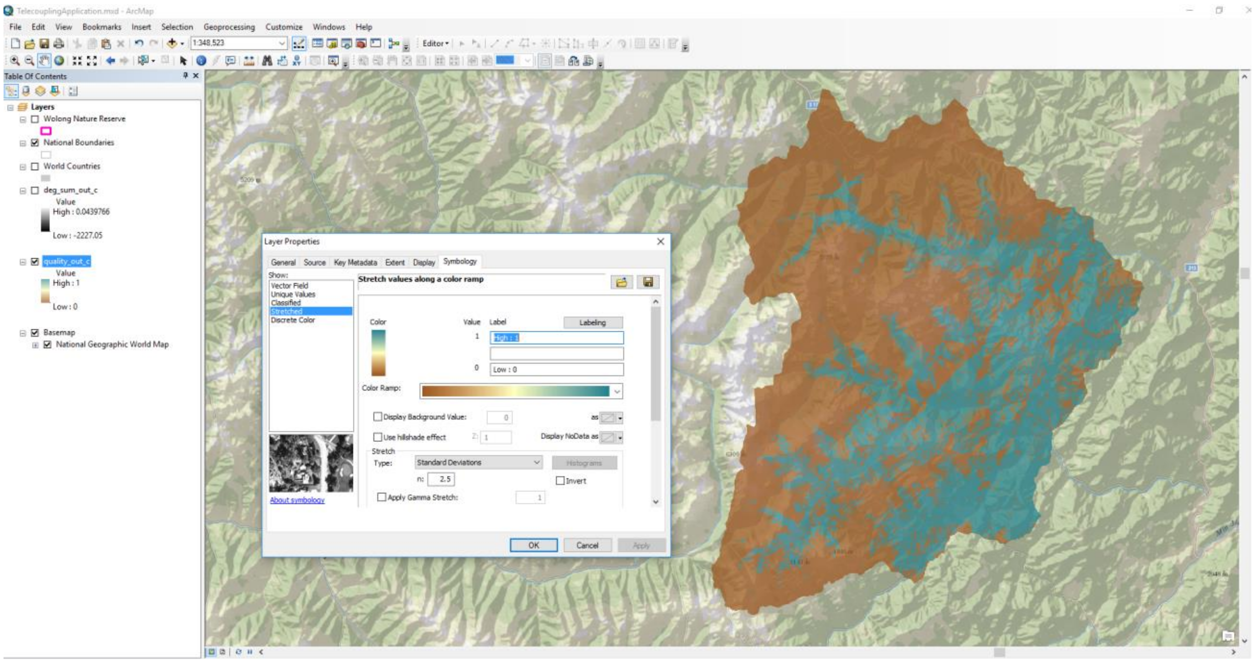
This tools is linked to the InVEST 3.3.1 habitat quality model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this tutorial, we will use the Wolong Nature reserve 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/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/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/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/habitat_quality/access_1998.shp*"). To define the relative sensitivity of each habitat type to each threat, select the table "*./SampleData/habitat_quality/sensitivity_samp.csv*". Leave the half saturation constant with the default value.

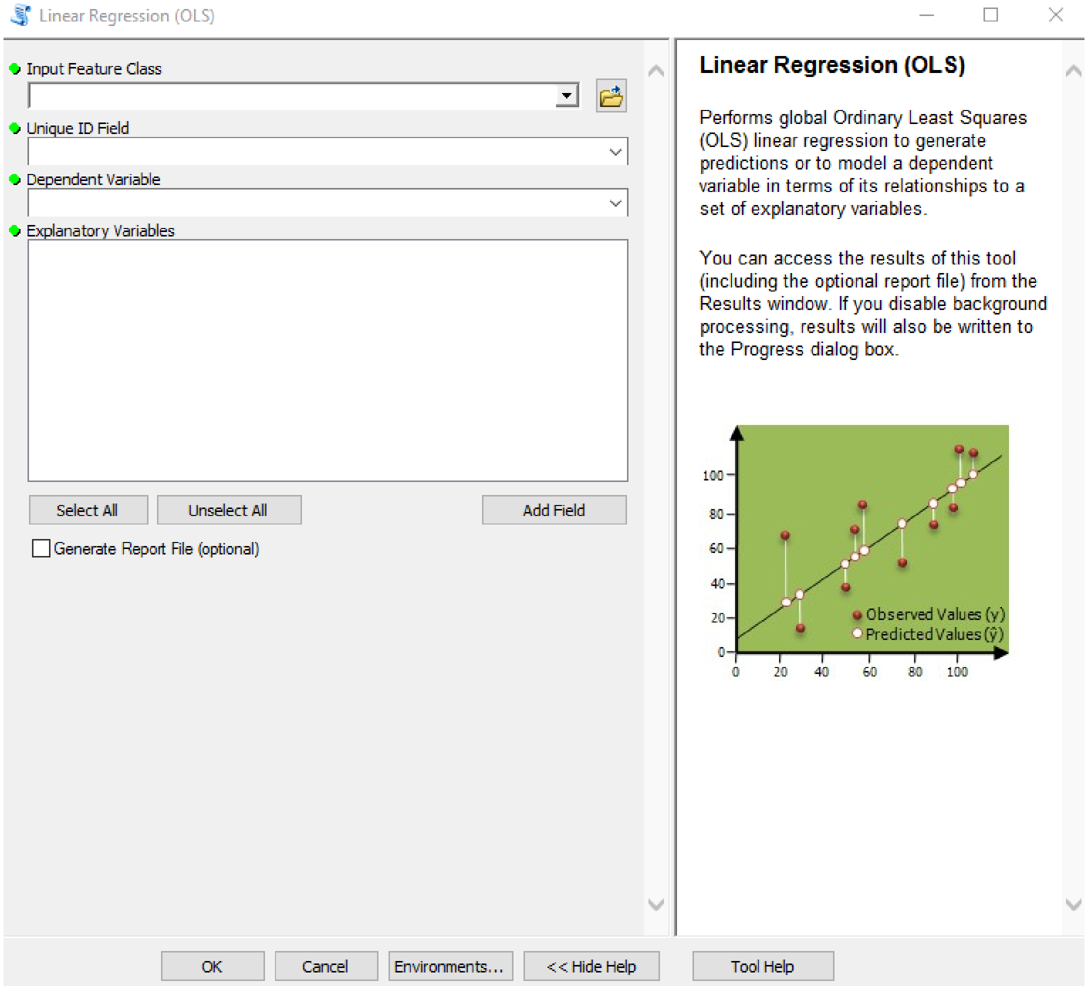
NOTE: In order to change the default workspace location where your output files and layers are created, click on the "Environments…" button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



Click OK 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):


TelecouplingApplication.mxd - ArcMap
Table Of Contents
Layers
Wolong Nature Reserve
National Boundaries
World Countries
deg_sum_out_c
quality_out_c
Basemap
National Geographic World Map
Layer Properties
Stretch values along a color ramp
Color Ramp:
Standard Deviations
About symbology
OK
Cancel
Apply

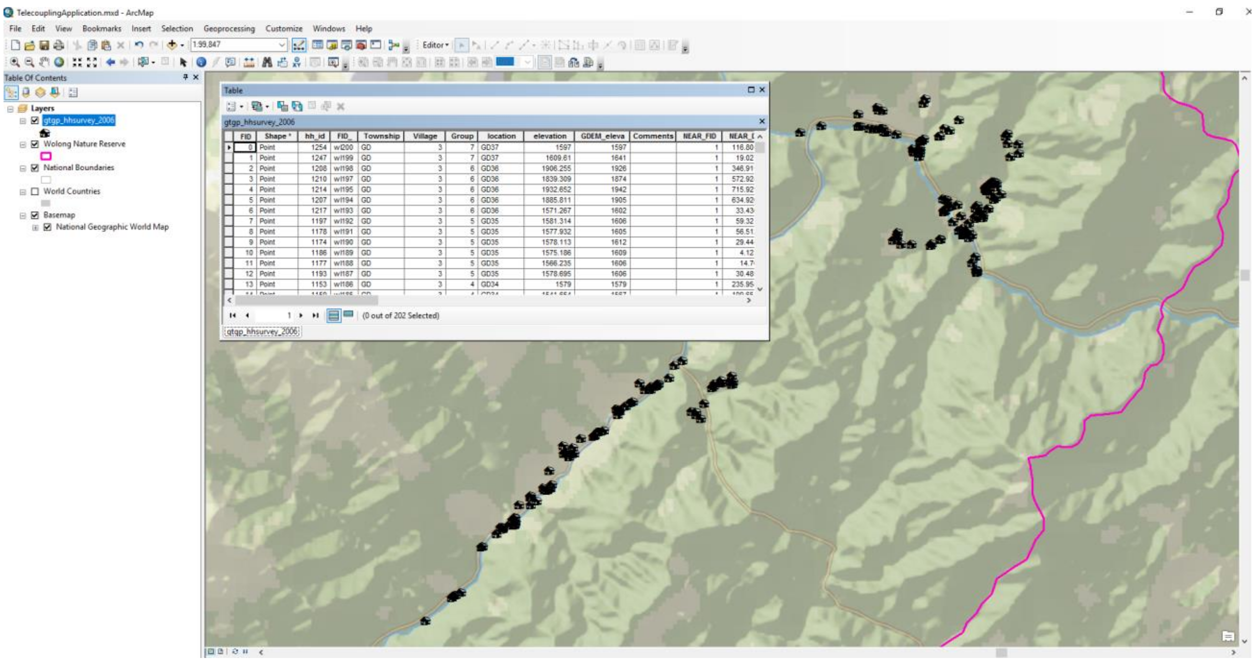
4.6.7 Linear Regression (OLS)



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.

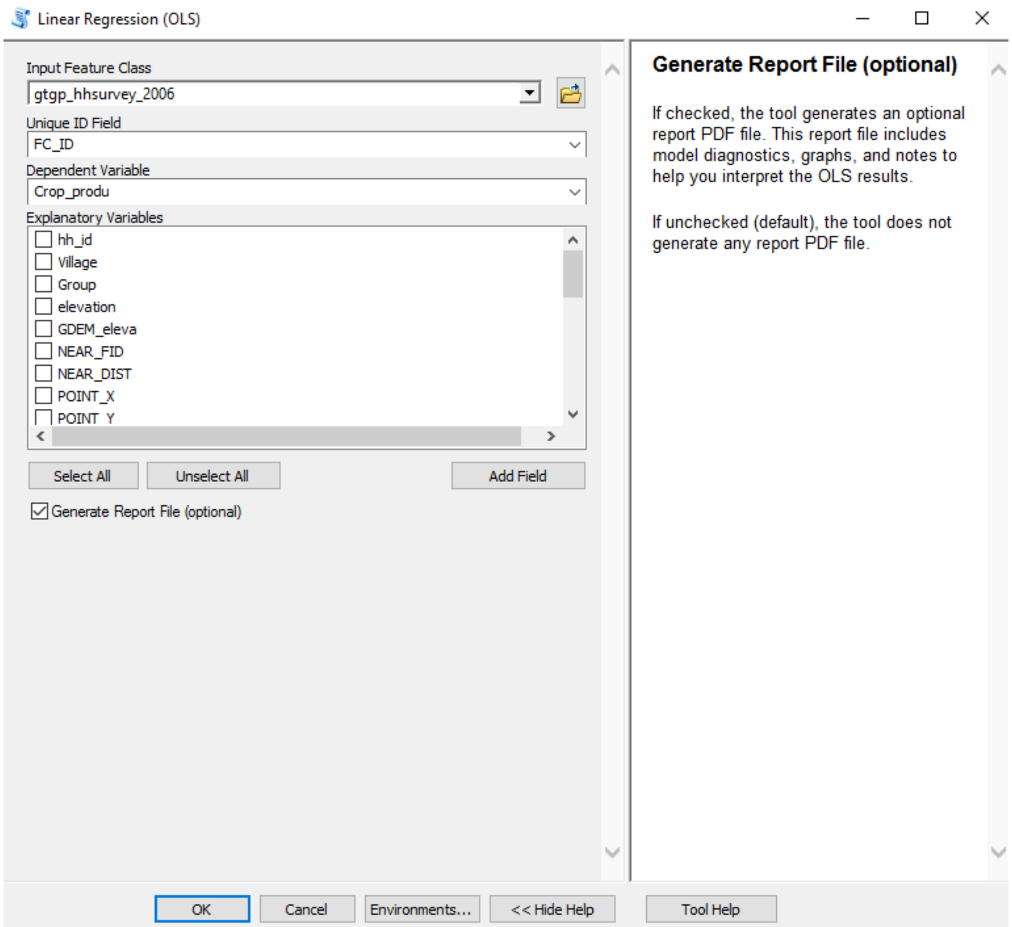
This tools, slightly modified from the original ESRI Ordinary Least Squares script tool, should be used to estimate the contribution of a set of factors (explanatory variables) that are most associated with a quantitative variable of interest, e.g. amount of cropland owned by local households in China. ***NOTE: Ideally, you should run this model after running the Model Selection (OLS) tool under the Causes toolset, because you would have a better idea on what factors (explanatory variables) should be used with the Linear Regression tool to estimate and predict the dependent variable under future scenarios (i.e. using future values for the chosen***

factors that best explain the dependent variable). 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. 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/gtgp/gtgp_hhsurvey_2006.shp*").

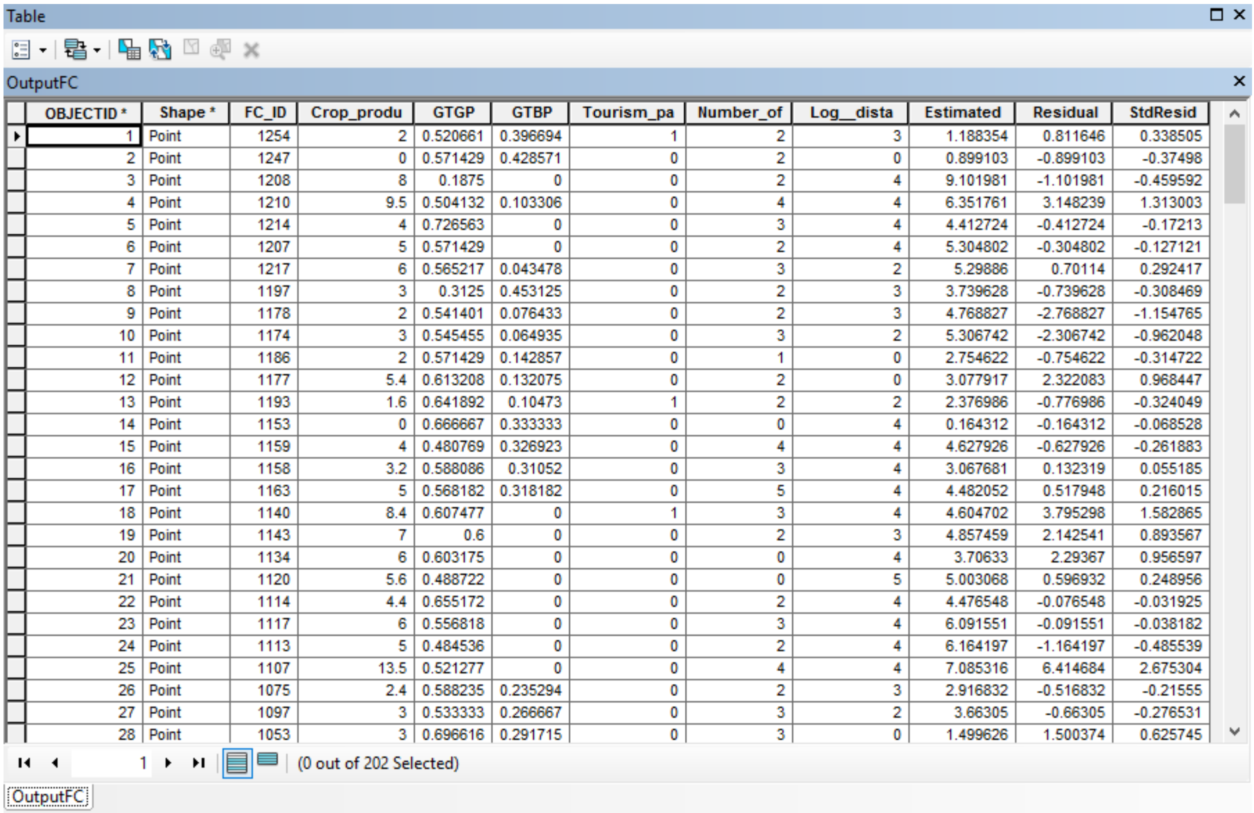


This shapefile describes spatial location of local households in the Wolong Nature Reserve, China, and a number of socio-economic variables recorded in 2006 with a targeted household survey including cropland size, household age, median income, educational level, and others. You may be interested in 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). Check the box (optional) to create a PDF report as well. Click OK and run the tool.

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.



You should receive a warning message about possible spatial autocorrelation between the observations. Linear regression assumes all observations are independent, so you should always check for significant spatial autocorrelation using tools available in ArcGIS (e.g. Global Moran's I). The model still runs and estimates the coefficients of all factors specified as well as the dependent variable of interest. However, keep in mind that model estimates may be inflated by the presence of spatial autocorrelation, in which case this tool may not be the most appropriate. ArcGIS return an output feature class that has the same attributes as the input dataset, with the addition of three columns at the end with estimated values of cropland owned, model residuals (difference between observed and estimated values of the dependent variable), and standardized model residuals.



Table

OutputFC

OBJECTID *	Shape *	FC_ID	Crop_produ	GTGP	GTBP	Tourism_pa	Number_of	Log__dista	Estimated	Residual	StdResid
1	Point	1254	2	0.520661	0.396694	1	2	3	1.188354	0.811646	0.338505
2	Point	1247	0	0.571429	0.428571	0	2	0	0.899103	-0.899103	-0.37498
3	Point	1208	8	0.1875	0	0	2	4	9.101981	-1.101981	-0.459592
4	Point	1210	9.5	0.504132	0.103306	0	4	4	6.351761	3.148239	1.313003
5	Point	1214	4	0.726563	0	0	3	4	4.412724	-0.412724	-0.17213
6	Point	1207	5	0.571429	0	0	2	4	5.304802	-0.304802	-0.127121
7	Point	1217	6	0.565217	0.043478	0	3	2	5.29886	0.70114	0.292417
8	Point	1197	3	0.3125	0.453125	0	2	3	3.739628	-0.739628	-0.308469
9	Point	1178	2	0.541401	0.076433	0	2	3	4.768827	-2.768827	-1.154765
10	Point	1174	3	0.545455	0.064935	0	3	2	5.306742	-2.306742	-0.962048
11	Point	1186	2	0.571429	0.142857	0	1	0	2.754622	-0.754622	-0.314722
12	Point	1177	5.4	0.613208	0.132075	0	2	0	3.077917	2.322083	0.968447
13	Point	1193	1.6	0.641892	0.10473	1	2	2	2.376986	-0.776986	-0.324049
14	Point	1153	0	0.666667	0.333333	0	0	4	0.164312	-0.164312	-0.068528
15	Point	1159	4	0.480769	0.326923	0	4	4	4.627926	-0.627926	-0.261883
16	Point	1158	3.2	0.588086	0.31052	0	3	4	3.067681	0.132319	0.055185
17	Point	1163	5	0.568182	0.318182	0	5	4	4.482052	0.517948	0.216015
18	Point	1140	8.4	0.607477	0	1	3	4	4.604702	3.795298	1.582865
19	Point	1143	7	0.6	0	0	2	3	4.857459	2.142541	0.893567
20	Point	1134	6	0.603175	0	0	0	4	3.70633	2.29367	0.956597
21	Point	1120	5.6	0.488722	0	0	0	5	5.003068	0.596932	0.248956
22	Point	1114	4.4	0.655172	0	0	2	4	4.476548	-0.076548	-0.031925
23	Point	1117	6	0.556818	0	0	3	4	6.091551	-0.091551	-0.038182
24	Point	1113	5	0.484536	0	0	2	4	6.164197	-1.164197	-0.485539
25	Point	1107	13.5	0.521277	0	0	4	4	7.085316	6.414684	2.675304
26	Point	1075	2.4	0.588235	0.235294	0	2	3	2.916832	-0.516832	-0.21555
27	Point	1097	3	0.533333	0.266667	0	3	2	3.66305	-0.66305	-0.276531
28	Point	1053	3	0.696616	0.291715	0	3	0	1.499626	1.500374	0.625745

1 (0 out of 202 Selected)

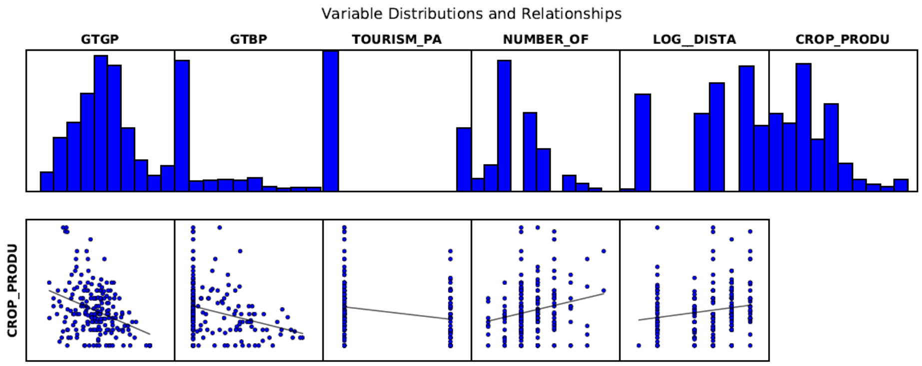
OutputFC

If you open the Results window (Menu > Geoprocessing > Results) and right-click on the message icon, click on View and you can see the estimated coefficients by the regression model:

Messages

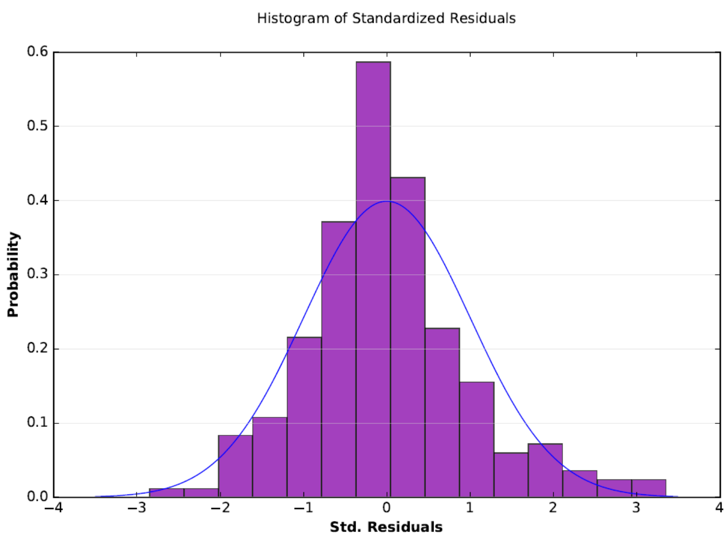
```
Executing: LinearRegressionOLS gtgp_hhsurvey_2006 FC_ID Crop_produ GTGP;GTBP;Tourism_pa;Number_of;Log__dista TRUE
Start Time: Wed Oct 12 10:46:19 2016
Running script LinearRegressionOLS...

                                      Summary of OLS Results
  Variable Coefficient [a] StdError t-Statistic Probability [b] Robust_SE  Robust_t Robust_Pr [b]  VIF [c]
 Intercept        9.012874 0.826588  10.903709       0.000000*  1.019067  8.844243     0.000000* --------
      GTGP       -9.890328 1.071110  -9.233717       0.000000*  1.235279 -8.006553     0.000000* 1.160764
      GTBP       -8.742180 1.279920  -6.830254       0.000000*  1.101758 -7.934757     0.000000* 1.305346
TOURISM_PA       -0.985820 0.380601  -2.590163       0.010307*  0.380605 -2.590137     0.010308* 1.092297
 NUMBER_OF        0.642247 0.131232   4.893996       0.000003*  0.143436  4.477583     0.000015* 1.024602
LOG__DISTA        0.164763 0.106159   1.552042       0.122277   0.115133  1.431065     0.154014  1.117571


                                            OLS Diagnostics
Input Features:              gtgp_hhsurvey_2006   Dependent Variable:                          CROP_PRODU
Number of Observations:                     202   Akaike's Information Criterion (AICc) [d]:   935.046362
Multiple R-Squared [d]:                0.425195   Adjusted R-Squared [d]:                        0.410532
Joint F-Statistic [e]:                28.997043   Prob(>F), (5,196) degrees of freedom:          0.000000*
Joint Wald Statistic [e]:            132.957360   Prob(>chi-squared), (5) degrees of freedom:    0.000000*
Koenker (BP) Statistic [f]:           30.591311   Prob(>chi-squared), (5) degrees of freedom:    0.000011*
Jarque-Bera Statistic [g]:            24.218679   Prob(>chi-squared), (2) degrees of freedom:    0.000006*


Notes on Interpretation
 *  An asterisk next to a number indicates a statistically significant p-value (p < 0.01).
[a] Coefficient: Represents the strength and type of relationship between each explanatory variable and the dependent variable.
[b] Probability and Robust Probability (Robust_Pr): Asterisk (*) indicates a coefficient is statistically significant (p < 0.01); if the Koenker (BP)
Statistic [f] is statistically significant, use the Robust Probability column (Robust_Pr) to determine coefficient significance.
[c] Variance Inflation Factor (VIF): Large Variance Inflation Factor (VIF) values (> 7.5) indicate redundancy among explanatory variables.
[d] R-Squared and Akaike's Information Criterion (AICc): Measures of model fit/performance.
[e] Joint F and Wald Statistics: Asterisk (*) indicates overall model significance (p < 0.01); if the Koenker (BP) Statistic [f] is statistically
significant, use the Wald Statistic to determine overall model significance.
[f] Koenker (BP) Statistic: When this test is statistically significant (p < 0.01), the relationships modeled are not consistent (either due to non-
stationarity or heteroskedasticity).  You should rely on the Robust Probabilities (Robust_Pr) to determine coefficient significance and on the Wald
Statistic to determine overall model significance.
[g] Jarque-Bera Statistic: When this test is statistically significant (p < 0.01) model predictions are biased (the residuals are not normally
distributed).
```

All factors except distance to roads seems to be statistically significant with different magnitude and positive/negative association to the dependent variable (cropland owned). If you selected the option to generate a PDF report file, under the Results window you should be able to find the report. Double-click on it to open it. The report has several pages, some of which contains textual information that is also reported in the message window (see figure above), and other graphical output that helps interpret the results of the linear regression model.

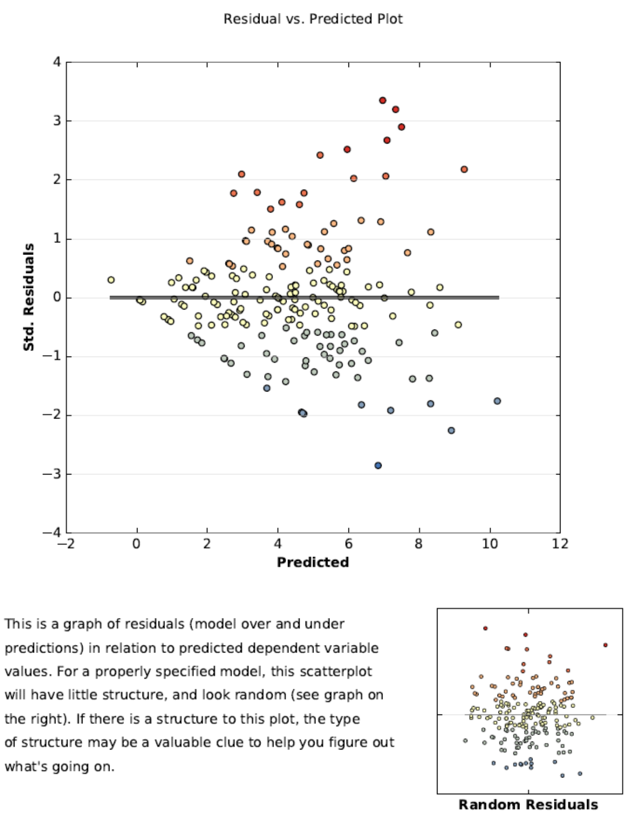


The above graphs are Histograms and Scatterplots for each explanatory variable and the dependent variable. The histograms show how each variable is distributed. OLS does not require variables to be normally distributed. However, if you are having trouble finding a properly-specified model, you can try transforming strongly skewed variables to see if you get a better result.

Each scatterplot depicts the relationship between an explanatory variable and the dependent variable. Strong relationships appear as diagonals and the direction of the slant indicates if the relationship is positive or negative. Try transforming your variables if you detect any non-linear relationships. For more information see the Regression Analysis Basics documentation.

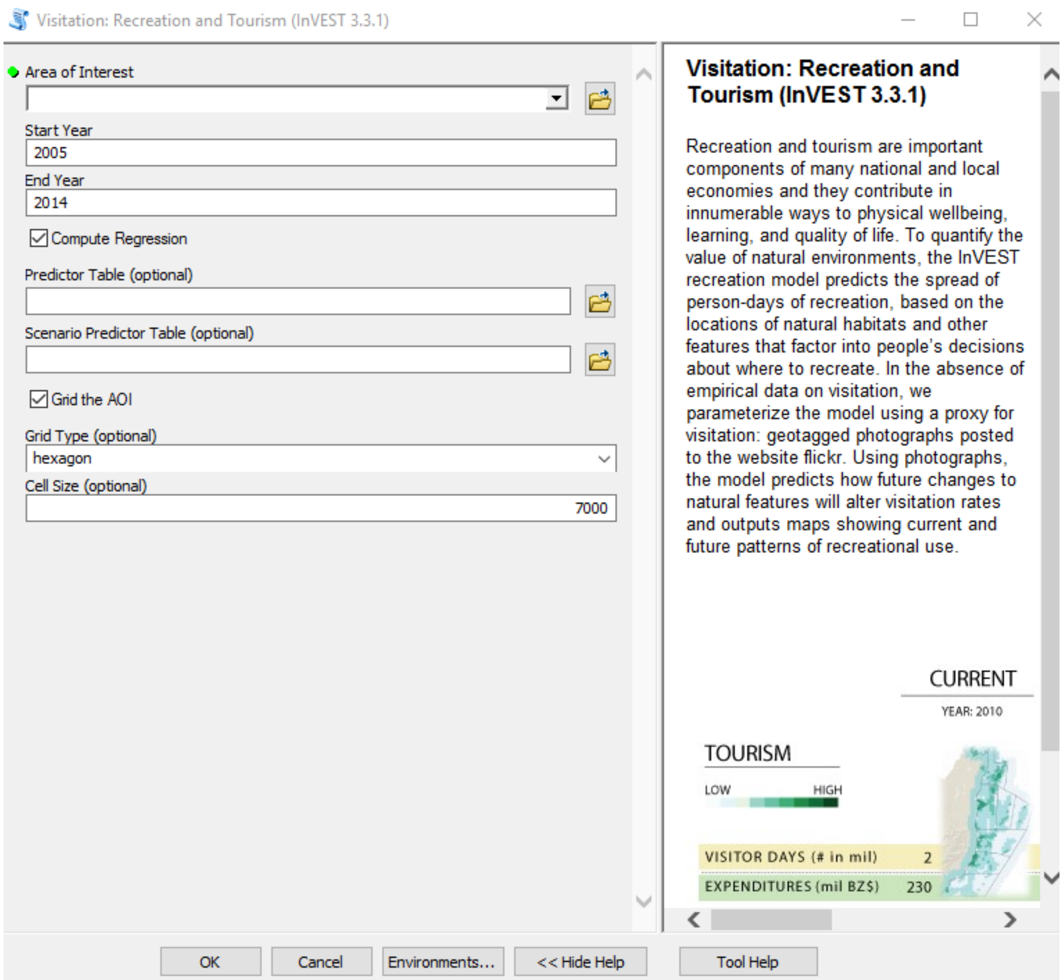


Ideally the histogram of your residuals would match the normal curve, indicated above in blue. If the histogram looks very different from the normal curve, you may have a biased model. If this bias is significant it will also be represented by a statistically significant Jarque-Bera p-value (*).



The first image shows bi-variate plots between cropland owned and each factor separately, as well as the histogram of their observed values. The second figure shows the histograms of standardized residuals with a Gaussian curve overlaid on top of it. We will not give much detail here on how to interpret results of linear regression models, as it is assumed the user has some experience with it. However, the more Gaussian distributed the residuals are the better your data will follow one of the linear regression model assumptions. The third figure shows a plot of predicted values versus standardized residuals. Ideally the values will be scattered around with no distinct pattern.

4.6.8 Visitation: Recreation and Tourism Storage and Sequestration (InVEST 3.3.1)

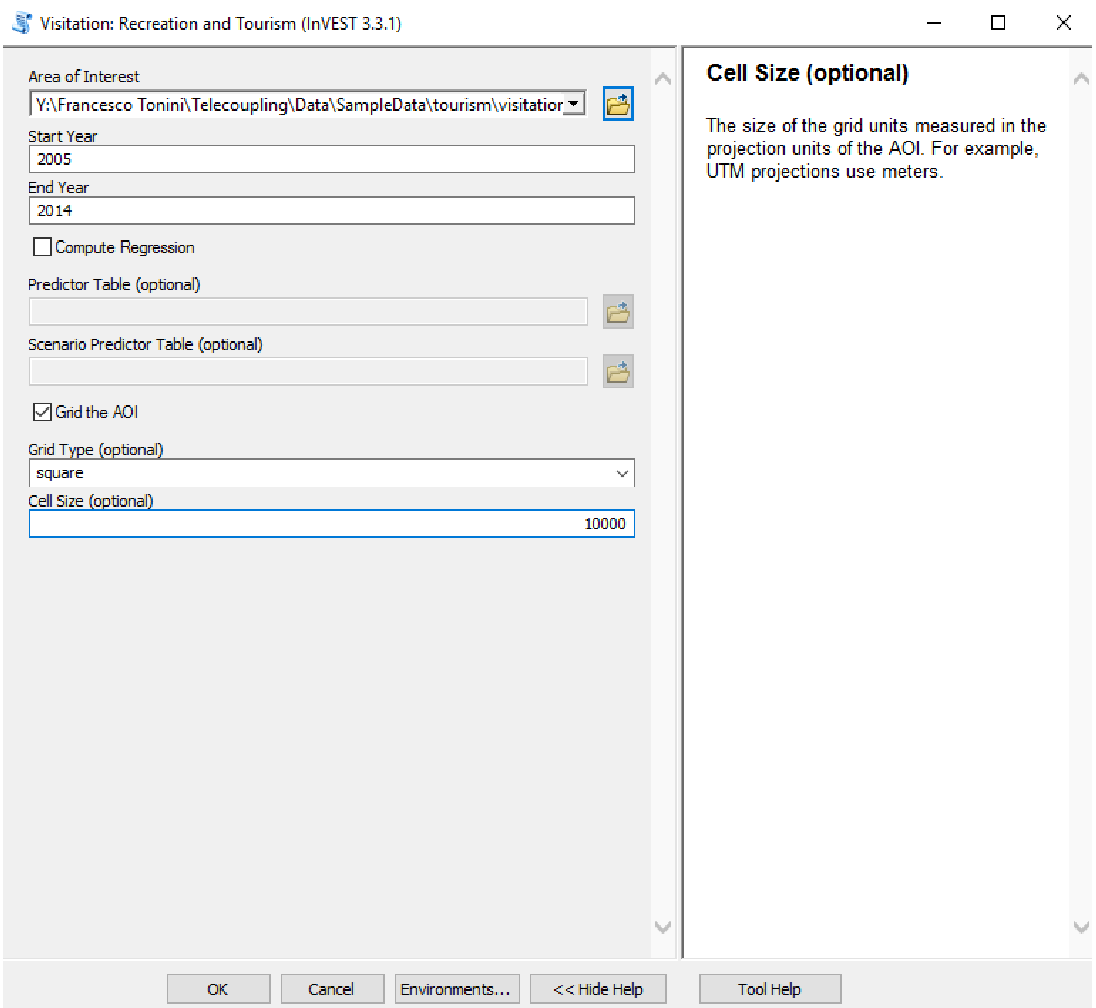


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

This tools is linked to the InVEST 3.3.1 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/tourism/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: