



# Telecoupling Toolbox v1.1b

## User Guide



Author: Francesco Tonini

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

Last Revised: August 15, 2017

# CONTENTS

<b>1</b>	<b>INTRODUCTION.....</b>	<b>4</b>
1.1	Who Should Use the Telecoupling Toolbox?.....	5
1.2	Major Releases .....	7
1.2.1	<i>Version 1.1 (beta).....</i>	7
1.2.2	<i>Version 1.0 (beta).....</i>	7
1.2.3	<i>Version 1.2 (alpha).....</i>	7
1.2.4	<i>Version 1.1 (alpha).....</i>	8
1.2.5	<i>Version 1.0 (alpha).....</i>	8
1.3	Minor Releases.....	8
1.3.1	<i>Version 1.2.1 (alpha).....</i>	8
1.3.2	<i>Version 1.0.1 (alpha).....</i>	8
1.3.3	<i>Version 1.1.2 (alpha).....</i>	8
1.3.4	<i>Version 1.0.1 (alpha).....</i>	8
1.4	The User Guide .....	9
1.5	Available Tools and Telecoupling Typologies.....	9
<b>2</b>	<b>GETTING STARTED .....</b>	<b>11</b>
2.1	Prerequisites .....	11
2.2	Download and unzip the Telecoupling Toolbox repository .....	11
2.3	Install Python libraries for 3rd party external software .....	12
2.4	Install the R-ArcGIS Bridge .....	13
2.5	Add the Telecoupling Toolbox to ArcGIS .....	16
<b>3</b>	<b>TOOLBOX STRUCTURE.....</b>	<b>18</b>
3.1	Overview .....	18
3.2	Systems Toolset.....	18
3.2.1	<i>Add Systems Interactively .....</i>	18
3.2.2	<i>Draw Systems from Table .....</i>	19
3.3	Agents Toolset.....	20
3.3.1	<i>Add Agents Interactively .....</i>	20
3.3.2	<i>Draw Agents from Table .....</i>	21
3.4	Flows Toolset.....	22
3.4.1	<i>Draw Radial Flows.....</i>	22
3.4.2	<i>Add Media Information Flows .....</i>	26
3.5	Causes Toolset.....	28
3.5.1	<i>Add Causes Interactively .....</i>	28
3.5.2	<i>Model Selection (OLS).....</i>	29
3.5.3	<i>Factor Analysis for Mixed Data .....</i>	33
3.6	Environmental Analysis Toolset .....	35
3.6.1	<i>CO<sub>2</sub> Emissions (Wildlife Transfer) .....</i>	35
3.6.2	<i>Forest Carbon Edge Effect (InVEST 3.3.3) .....</i>	37
3.6.3	<i>Crop Production (InVEST 3.3.3) .....</i>	38
3.6.4	<i>Habitat Quality (InVEST 3.3.3) .....</i>	40
3.6.5	<i>Habitat Risk Assessment (InVEST 3.3.3) .....</i>	42

3.6.6	<i>Habitat Risk Assessment Preprocessor (InVEST 3.3.3)</i> .....	49
3.7	Socioeconomic Analysis Toolset .....	51
3.7.1	<i>Cost-Benefit Analysis (Wildlife Transfer)</i> .....	51
3.7.2	<i>Visitation: Recreation and Tourism Storage and Sequestration (InVEST 3.3.3)</i> .....	52
3.8	Future Development.....	53
3.9	References .....	53
3.10	Acknowledgements .....	54
3.11	License & Copyright .....	54
<b>4</b>	<b>TUTORIALS .....</b>	<b>57</b>
4.1	SAMPLE DATASET .....	57
4.2	SYSTEMS TOOLSET .....	59
4.2.1	<i>Add Systems Interactively</i> .....	59
4.2.2	<i>Draw Systems from Table</i> .....	62
4.3	AGENTS TOOLSET .....	64
4.3.1	<i>Add Agents Interactively</i> .....	64
4.3.2	<i>Draw Agents from Table</i> .....	67
4.4	FLOWS TOOLSET.....	69
4.4.1	<i>Draw Radial Flows</i> .....	69
4.4.2	<i>Add Media Information Flows</i> .....	75
4.5	CAUSES TOOLSET .....	83
4.5.1	<i>Add Causes Interactively</i> .....	83
4.5.2	<i>Model Selection (OLS)</i> .....	86
4.5.3	<i>Factor Analysis for Mixed Data</i> .....	89
4.6	ENVIRONMENTAL ANALYSIS TOOLSET .....	96
4.6.1	<i>CO<sub>2</sub> Emissions (Wildlife Transfer)</i> .....	96
4.6.2	<i>Forest Carbon Edge Effect (InVEST 3.3.3)</i> .....	98
4.6.3	<i>Crop Production (InVEST 3.3.3)</i> .....	102
4.6.4	<i>Habitat Quality (InVEST 3.3.3)</i> .....	105
4.6.5	<i>Habitat Risk Assessment Preprocessor (InVEST 3.3.3)</i> .....	108
4.6.6	<i>Habitat Risk Assessment (InVEST 3.3.3)</i> .....	111
4.7	SOCIOECONOMIC ANALYSIS TOOLSET .....	115
4.7.1	<i>Cost-Benefit Analysis (Wildlife Transfer)</i> .....	115
4.7.2	<i>Visitation: Recreation and Tourism Storage and Sequestration (InVEST 3.3.3)</i> .....	119

## 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 socioeconomic 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<sup>1</sup>) 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?

---

<sup>1</sup> <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?

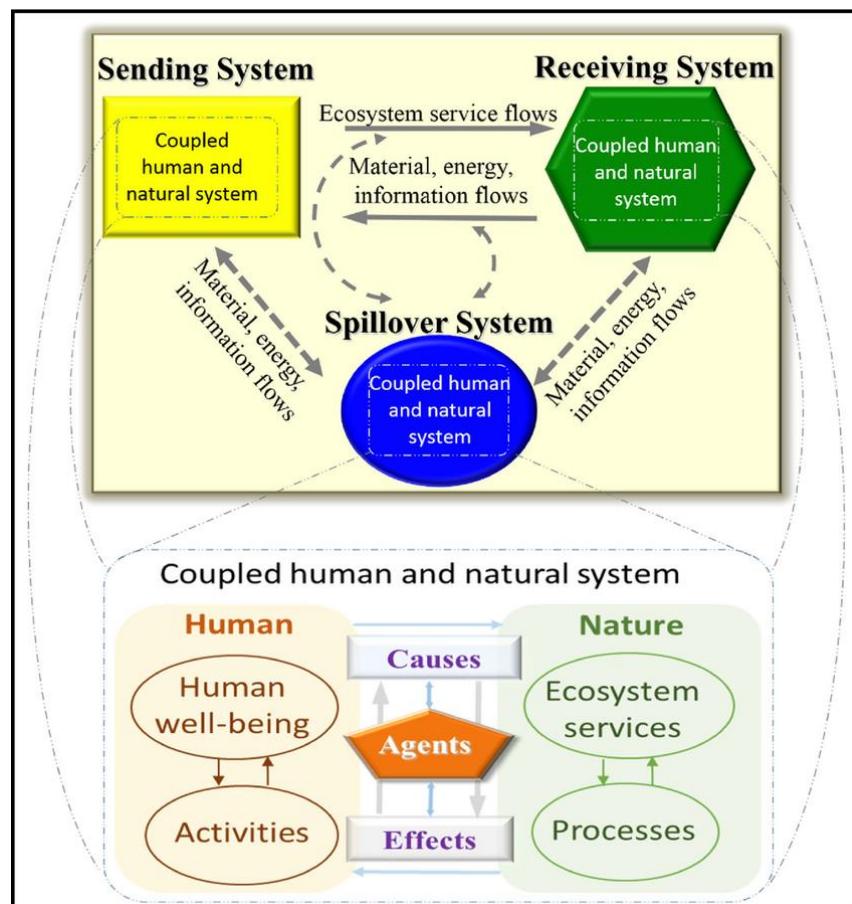


Figure 1. Framing ES in a telecoupling context (adapted from Liu et al. 2013). A “sending” system provides ES to a “receiving” system, and may influence another system (the “spillover” system) in the process. Feedbacks occur among different systems as a result of material energy, and information flows. 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 1.1 (beta)

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

### 1.2.2 Version 1.0 (beta)

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

### 1.2.3 Version 1.2 (alpha)

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

#### 1.2.4 **Version 1.1** (alpha)

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

#### 1.2.5 **Version 1.0** (alpha)

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

### 1.3 Minor Releases

#### 1.3.1 **Version 1.2.1** (alpha)

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

#### 1.3.2 **Version 1.0.1** (alpha)

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

#### 1.3.3 **Version 1.1.2** (alpha)

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

#### 1.3.4 **Version 1.0.1** (alpha)

Fixed bug in the script tools linked to InVEST (3.3.1) that caused any output vector and raster files to incorrectly align with the input layers. The issue seems to be caused by the way ArcGIS interprets the spatial reference (projection string) of output files

saved by the InVEST software. A workaround has been implemented, automatically re-defining the projection of the output layers before showing in the table of content in ArcGIS.

## 1.4 The User Guide

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

## 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 Radial Flows	✓	✓	✓	✓	✓	✓
Add Media Information Flows	✗	✗	✗	✗	✗	✓
Add Causes Interactively	✓	✓	✓	✓	✓	✓
Model Selection (OLS)	✓	✓	✓	✓	✓	✗
Factor Analysis for Mixed Data	⚠	✓	✓	✓	⚠	✗
Visitation Rate (InVEST 3.3.3)	✗	⚠	✗	✗	✗	✗

Habitat Quality (InVEST 3.3.3)	✓	✓	✓	!	✓	!
Crop Production (InVEST 3.3.3)	!	!	✓	!	✓	!
Cost-Benefit Analysis (Wildlife Transfer)	✓	✗	✗	✗	✗	✗
CO2 Emissions (Wildlife Transfer)	✓	✗	✗	✗	✗	✗
Forest Carbon Edge Effect (InVEST 3.3.3)	!	!	✓	✓	✓	!
Habitat Risk Assessment (InVEST 3.3.3)	✓	✓	✓	!	✓	!
Habitat Risk Assessment Preprocessor (InVEST 3.3.3)	Create rating files for input of Habitat Risk Assessment tool					

Figure 2. Script tools available for the Telecoupling Toolbox v1.1b (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 (**ArcGIS comes with Python, so no need to install a standalone version of Python!**)

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

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

### 2.2 Download and unzip the Telecoupling Toolbox repository

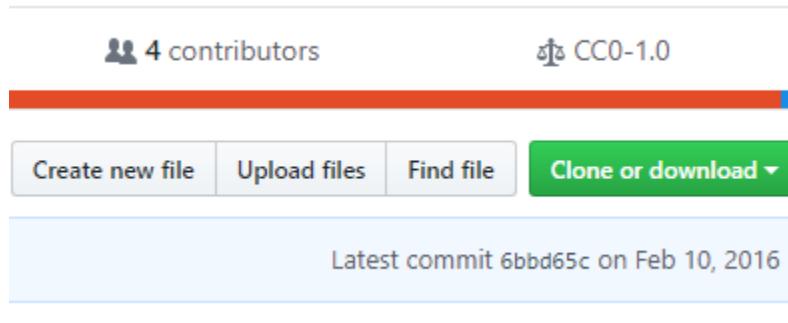
#### Github

Follow these guidelines to download the full repository of the Telecoupling Toolbox<sup>2</sup> publicly available on Github:

1. Find the Clone or Download button on the top right of the main repository page (see below).

---

<sup>2</sup> <https://github.com/f-tonini/telecoupling-toolbox>



2. Click on the button 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.

## 2.3 Install Python libraries for 3rd party external software

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

1. Although your computer may already have a C++ compiler installed, follow this step and **Download** and **install** the Microsoft Visual C++ Compiler for Python 2.7<sup>3</sup>
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](#)<sup>4</sup>
    - Open the CMD prompt on Windows and type: `C:\Python27\ArcGIS10.3\python.exe` followed by the full path to `get-pip.py` (*for example if you downloaded and saved the file on your D:\ drive, the full path would be D:\get-pip.py*)
      - Hit *Enter* to run the command above
      - Open the folder `PyLibs` found inside the (unzipped) telecoupling project folder
      - Double-click on the `ArcGIS103_Py32_libs.bat` file
  - **ArcGIS 10.4.x (standard 32-bit version):**

<sup>3</sup> <https://www.microsoft.com/en-us/download/details.aspx?id=44266>

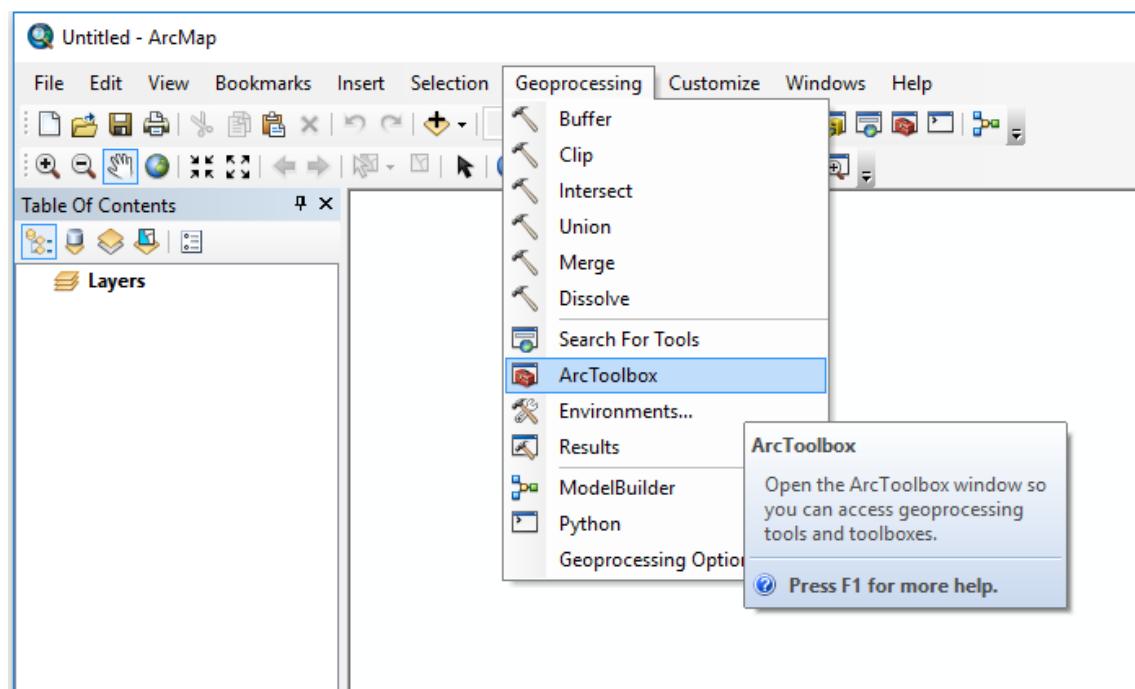
<sup>4</sup> <https://bootstrap.pypa.io/get-pip.py>

- Open the folder **PyLibs** found inside the (unzipped) telecoupling project folder
- Double-click on the **ArcGIS104\_Py32\_libs.bat** file
- **ArcGIS 10.5.x (standard 32-bit version):**
  - Open the folder **PyLibs** found inside the (unzipped) telecoupling project folder
  - Double-click on the **ArcGIS105\_Py32\_libs.bat** file

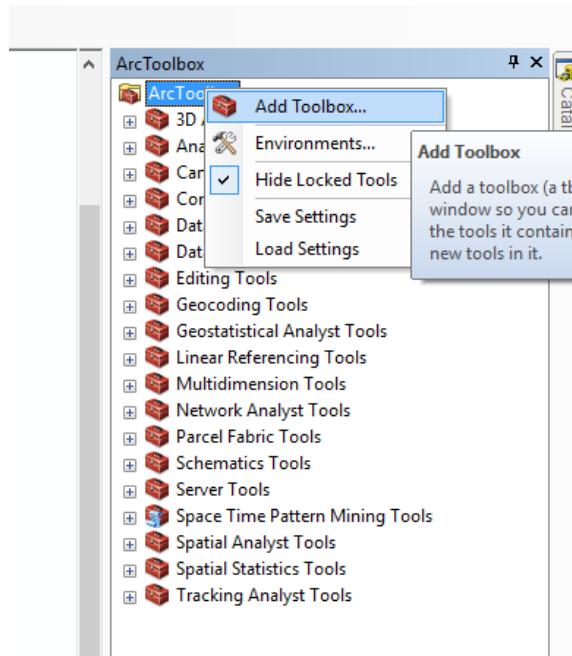
## 2.4 Install the R-ArcGIS Bridge

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

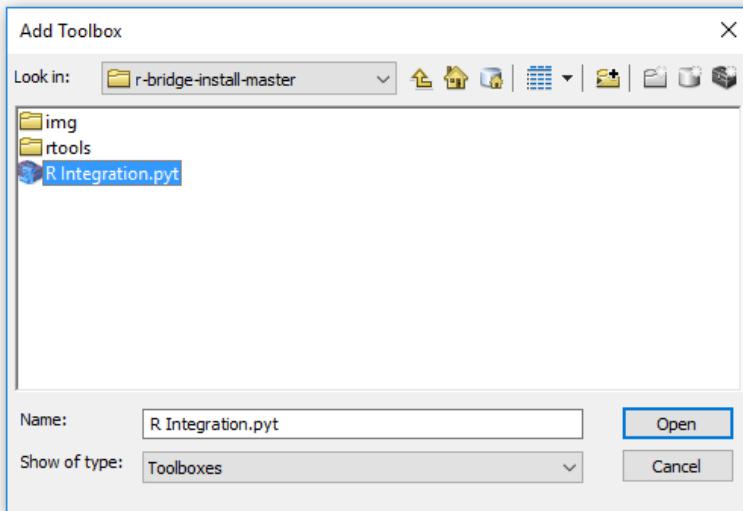
1. Open ArcMap (**NOTE: make sure you have admin rights on your computer or the next steps will not work!**)
2. Find and open the Geoprocessing ArcToolbox window (Menu > Geoprocessing > ArcToolbox)



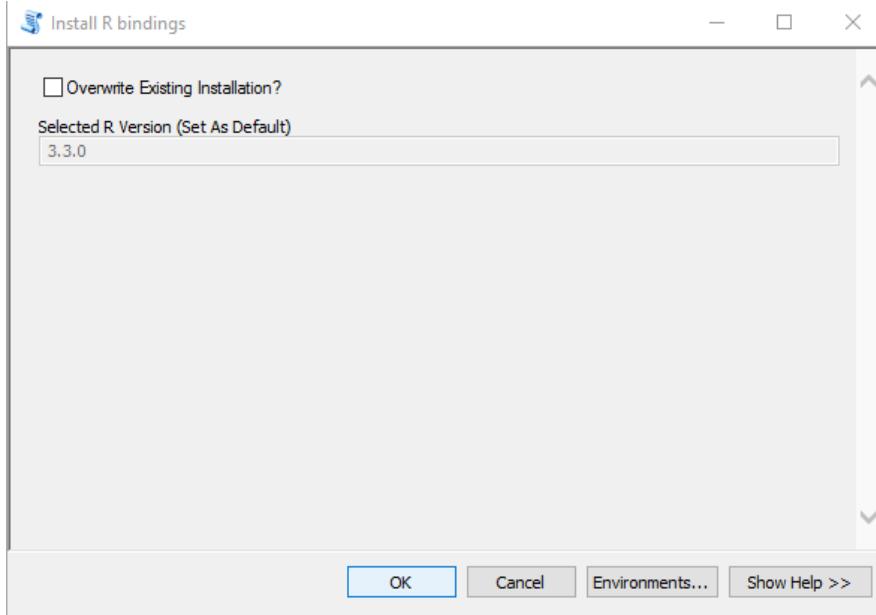
3. Right-click the ArcToolbox folder and select "Add Toolbox..."



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



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



If you need more details and information, ESRI has developed a nice Github webpage<sup>5</sup> with lots of useful documentation on how to install a set of libraries to make sure R and ArcGIS can talk to each other.

## 2.5 Add the Telecoupling Toolbox to ArcGIS

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

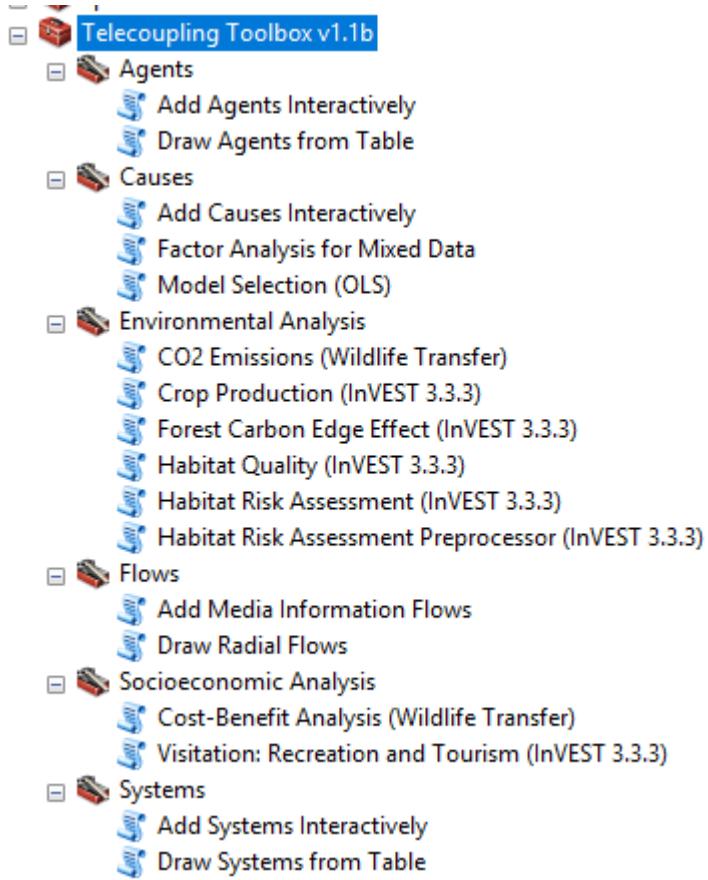
Follow these steps to add the Toolbox to your ArcMap document:

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

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

---

<sup>5</sup> <https://github.com/R-ArcGIS/r-bridge-install>



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

That's it! The Telecoupling Toolbox is now added to the ArcToolbox list and you can start using it with the set of sample data (`SampleData.zip`) that can be downloaded at:

<https://s3.amazonaws.com/telecoupling-toolbox-sample-data/SampleData.zip>

After unzipping the sample data folder, you will see a mix of GIS (vector, raster) data and tables (.csv) needed as input parameters by the script tools.

## 3 TOOLBOX STRUCTURE

### 3.1 Overview

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

### 3.2 Systems Toolset

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

#### 3.2.1 Add Systems Interactively

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

Parameter	Explanation	Data Type
Input_Features	<p>Click on the map to add the spatial location of the system components.</p> <p><i>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</i></p>	Feature Set

	<i>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.</i>	
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.  <b><i>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.</i></b>	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

	<p><b><i>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.</i></b></p> <p><b><i>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.</i></b></p>	
Y_Field	A numerical field in the input table containing the y coordinates (or latitudes) of each system component.  <p><b><i>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.</i></b></p> <p><b><i>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.</i></b></p>	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.  <p><b><i>NOTE: the mouse cursor and its pre-assigned symbology should appear as soon as you move it over the</i></b></p>	Feature Set

	<i>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.</i>	
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.  <b>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.</b>	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.  <i><b>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.</b></i>	Field
Y_Field	A numerical field in the input table containing the y coordinates (or latitudes) of each agent.  <i><b>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.</b></i>	Field

## 3.4 Flows Toolset

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

### 3.4.1 Draw Radial Flows

Draw radial flows between two or more telecoupling system components (i.e. "sending", "receiving", "spillover") on the map. Radial flows have a spoke-like pattern where features and places are mapped in nodal form with one place being a common origin or destination. Flows represent the movement of material, energy, or information. Arrows show the direction of the movement. The tool creates a new feature class containing geodetic line features constructed based on the values in a start x-coordinate field, start y-coordinate field, end x-coordinate field, and end y-coordinate field of a table. If desired, the tool will also compute a point vector layer representing nodes at each destination of the radial flows, using the same end x-coordinate field and end y-coordinate field of the attribute table used to produce flow lines. Points (nodes) at flow

destinations can be symbolized and rendered as desired by each user to show quantities (attributes) of interest. Use this tool if you already have a pre-existing table on file with (at minimum) spatial coordinates of all the origin and destination points for each flow. If you have the amount of material, energy, or information in the table, use it to show differences in magnitude (e.g. line thickness) between flows.

Parameter	Explanation	Data Type
Input_Table	<p>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</p> <p><b><i>NOTE: table attributes should at minimum include XY coordinates of both the origin and destination nodes of the flows.</i></b></p>	Table View
Start_X	<p>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.</p> <p><b><i>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.</i></b></p>	Field
Start_Y	<p>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.</p> <p><b><i>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</i></b></p>	Field

	<i>reproject features on-the-fly, the same will not apply when using the tool as a geoprocessing service consumed within a web application.</i>	
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.  <i>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.</i>	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.  <i>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.</i>	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	String

	<p>lengthy and complex and therefore omitted here. This line type is the default.</p> <p><b>GREAT_CIRCLE</b>—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.</p> <p><b>RHUMB_LINE</b>—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.</p> <p><b>NORMAL_SECTION</b>—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.</p>	
Spatial_Reference (Optional)	The spatial reference of the coordinates found in the input table. On the Spatial Reference Properties dialog box you can Select, Import, or Create a New coordinate system. The default is Web Mercator or the input coordinate system if it is not Unknown.	Spatial Reference
Join_Fields (Optional)	The fields from the join table to be included in the join.	Multiple Value
Create_Destination_Nodes	If checked, a point layer (nodes) is created at each flow destination in addition to a line layer showing flows between locations. These points can be rendered and symbolized according to a specified attribute of interest (e.g. quantities traded).	Boolean

	If <b>unchecked (default)</b> , only a line layer showing flows between locations will be created.	
Output_Flows_Layer_Name	Name of the output flows layer	String
Output_Nodes_Layer_Name (Optional)	Name of the output nodes point layer	String

### 3.4.2 Add Media Information Flows

News media and publication of books and articles heavily contribute to disseminate information on certain topics across the globe. Several online portals let you search through large databases for specific terms or academic publications on a subject of interest. LexisNexis® Academic<sup>6</sup> 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 correspond to the term used to search through the LexisNexis online database.	Feature Set

<sup>6</sup> <http://www.lexisnexis.com/en-us/products/lexisnexis-academic.page>

	<b><i>NOTE: Make sure to select ONLY a SINGLE source location on the map.</i></b>	
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)	<p>The type of geodetic line to construct.</p> <ul style="list-style-type: none"> <li>• 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.</li> <li>• 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.</li> <li>• 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.</li> <li>• 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</li> </ul>	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.

## 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 telecoupling 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	<p>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).</p> <p><b><i>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 spatial location of the cause added and its description attribute.</i></b></p>	Feature Set

Describe_Causes	<p>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.</p> <p><b><i>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.</i></b></p>	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 R<sup>2</sup> 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)	<p>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.</p> <p>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.</p>	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

	<p>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.</p>	
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 continuous 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)	<p>String corresponding to the color which are used.</p> <ul style="list-style-type: none"> <li>• 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.</li> <li>• 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</li> <li>• 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</li> </ul>	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)	<p>A selection of the individuals that are drawn:</p> <ul style="list-style-type: none"> <li>• "coord N": select the N elements that have the highest (squared) coordinates on the 2 first dimensions drawn</li> <li>• "contrib N": select the N elements that have the highest contribution on the 2 first dimensions drawn</li> <li>• "cos2 N": select the N elements that have the highest cos2 on the 2 dimensions drawn</li> <li>• "dist N": select the N elements that have the highest distance to the center of gravity</li> </ul>	String
Categories_Selection (Optional)	A selection of the categories that are drawn:	String

	<ul style="list-style-type: none"> <li>• "coord N": select the N elements that have the highest (squared) coordinates on the 2 first dimensions drawn</li> <li>• "contrib N": select the N elements that have the highest contribution on the 2 first dimensions drawn</li> <li>• "cos2 N": select the N elements that have the highest cos2 on the 2 dimensions drawn</li> <li>• "dist N": select the N elements that have the highest distance to the center of gravity</li> </ul>	
Add_Label_to_Individuals (Optional)	When checked, individuals drawn on the plot are labelled.	Boolean
Output_PDF	Creates a PDF containing graphs generated from the plot function in the FactoMineR package. These graphs show individual and/or variable plots using the first two dimensions extracted from the analysis. For quantitative-only datasets, a graph of the eigenvalues associated with the extracted principal components is shown in the last page. For more information about these graphs, consult the documentation for the FactoMineR package.	File

## 3.6 Environmental Analysis Toolset

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

### 3.6.1 CO<sub>2</sub> Emissions (Wildlife Transfer)

Calculate the amount of CO<sub>2</sub> emissions associated with each flow segment of material transported between two spatial locations. The tool reads the total length of each linear feature created by the Flows tool and multiplies it with the number of trips needed to transfer wildlife units (derived from number of units to transfer divided by a per-trip transportation capacity, e.g. per-flight) and the amount of CO<sub>2</sub> (kilograms) emitted per unit of length (e.g. meters) specified by the user. The tool also allows testing of future CO<sub>2</sub> emission scenarios for wildlife transfer compared to current conditions. Make sure the unit of length corresponds to the coordinate system used (e.g. meters in the case of Web Mercator). Use this tool if you have a general

estimate of how many kilograms of CO<sub>2</sub> per unit of length (e.g. meters) are associated with the type of transportation flow being considered.

Parameter	Explanation	Data Type
Input_Feature	Input Feature representing the linear flows of material between telecoupling systems across the globe.	Feature Layer
Wildlife_Units	The numeric field containing values for the number of wildlife units to be transferred.	Field
Transportation_Capacity	<p>Integer number representing the maximum capacity of the transportation medium (e.g. airplane) used to transfer all wildlife units.</p> <p><b>NOTE: this tool assumes homogeneity of transporation medium (e.g. same type of airplane), hence the estimates do not account for variations in CO<sub>2</sub> emissions caused by multiple carriers or weather conditions (e.g. wind) affecting the transfer.</b></p>	Long
Units_of_CO <sub>2</sub> _Emitted	Specify the approximate amount (in kilograms) of CO <sub>2</sub> emitted per unit of length. Make sure the unit of length corresponds to the coordinate system used (e.g. meters in the case of Web Mercator).	Double
Input_Feature_Scenario_(Optional)	Input Feature representing a future scenario of linear flows of material between telecoupling systems across the globe.	Feature Layer
Wildlife_Units_Scenario_(Optional)	The numeric field containing values for the number of wildlife units to be transferred under a future scenario.	Field
Transportation_Capacity_Scenario_(Optional)	<p>Integer number representing the maximum capacity of the transportation medium (e.g. airplane) used to transfer all wildlife units under a future scenario.</p> <p><b>NOTE: this tool assumes homogeneity of transporation medium (e.g. same type of airplane), hence the estimates do not account for variations in CO<sub>2</sub> emissions caused by multiple carriers or weather conditions (e.g. wind) affecting the transfer.</b></p>	Long
Units_of_CO <sub>2</sub> _Emitted_Scenario_(Optional)	Specify the approximate amount (in kilograms) of CO <sub>2</sub> emitted per unit of length under a future scenario. Make sure	Double

	the unit of length corresponds to the coordinate system used (e.g. meters in the case of Web Mercator).	
--	---	--

### 3.6.2 Forest Carbon Edge Effect (InVEST 3.3.3)

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

Parameter	Explanation	Data Type
Land_Use_Land_Cover_Map	A GDAL-supported raster file, with an integer LULC code for each cell.	Raster Layer
Biophysical_Table	A CSV table containing model information corresponding to each of the land use classes in the LULC raster input. It must contain the fields 'lucode', 'is_tropical_forest', 'c_above'. If the user selects 'all carbon pools' the table must also contain entries for 'c_below', 'c_soil', and 'c_dead'. See the InVEST Forest Carbon User's Guide for more information about these fields.	File
Carbon_Pools_to_Calculate	If 'all carbon pools' is selected then the headers 'c_above', 'c_below', 'c_dead', 'c_soil' are used in the carbon pool calculation. Otherwise only 'c_above' is considered.	String
Compute_Forest_Edge_Effects	If selected, will use the Chaplin-Kramer, et. al method to account for above ground carbon stocks in tropical forest types indicated by a '1' in the 'is_tropical_forest' field in the biophysical table.	Boolean

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

### 3.6.3 Crop Production (InVEST 3.3.3)

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

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

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

Parameter	Explanation	Data Type
Lookup_Table	Filepath to a CSV table used to convert the crop code provided in the Crop Map to the crop name that can be used for searching through inputs and formatting outputs.	File
Crop_Management_Scenario_Map	A GDAL-supported raster representing a crop management scenario.	Raster Layer
Global_Dataset_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
---	---	------

### 3.6.4 Habitat Quality (InVEST 3.3.3)

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

Parameter	Explanation	Data Type
Current_Land_Cover	<p>A GDAL-supported raster file. The current LULC must have its' own threat rasters, where each threat raster file path has a suffix of <b>_c</b>.</p> <p>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.</p>	Raster Layer
Future_Land_Cover (Optional)	<p>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 <b>_f</b>.</p> <p>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. <b>The LULC codes must match the codes in the Sensitivity table.</b></p>	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>_b</b> . If	Raster Layer

	<p>no threat rasters are found, degradation and habitat quality outputs will not be generated for the baseline LULC.</p> <p>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.</p>	
Folder_Containing_Threat_Rasters	The selected folder is used as the location to find all threat rasters for the threats listed in the below table.	Folder
Threats_Data	<p>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).</p> <p><b>THREAT:</b> The name of the threat source and this name must match exactly to the name of the threat raster and to the name of its corresponding column in the sensitivity table. <b>NOTE:</b> 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.</p> <p><b>MAX_DIST:</b> A number in kilometres (km) for the maximum distance a threat has an affect.</p> <p><b>WEIGHT:</b> A floating point value between 0 and 1 for 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.</p> <p><b>DECAY:</b> A string value of either <b>exponential</b> or <b>linear</b> representing the type of decay over space for the threat.</p> <p>See the user's guide for valid values for these columns.</p>	File

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 <b>ACCESS</b> . The <b>ACCESS</b> 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	<p>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: <b>LULC</b>, <b>HABITAT</b>, <b>L_THREAT1</b>, <b>L_THREAT2</b>, ...</p> <p><b>LULC:</b> Integer values that reflect each LULC code found in current, future, and baseline rasters.</p> <p><b>HABITAT:</b> A value of 0 or 1 (presence / absence) or a value between 0 and 1 (continuum) depicting the suitability of habitat.</p> <p><b>L_THREATN:</b> Each L_THREATN should match exactly with the threat names given in the threat CSV file, where the THREATN is the name that matches. This is a floating point value between 0 and 1 that represents the sensitivity of a habitat to a threat.</p> <p>Please see the user guide for more detailed information on proper column values and column names for each threat.</p>	File
Half_Saturation_Constant	A positive floating point value that is defaulted at 0.5. This is the value of the parameter k in equation (4). In general, set k to half of the highest grid cell degradation value on the landscape. To perform this model calibration the model must be run once in order to find the highest degradation value and set k for the provided landscape. Note that the choice of k only determines the spread and central tendency of habitat quality cores and does not affect the rank.	Double

### 3.6.5 Habitat Risk Assessment (InVEST 3.3.3)

The condition of a habitat is a key determinant of the ecosystem services it can provide. For example, multiple stressors including fishing, climate change, pollution and coastal development

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

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

Stressors_Folder	Folder containing stressors to be overlapped with habitats and/or species. This directory should contain ONLY the stressors desired within this model run. All layers must be projected in the same projection.	Folder
Spatially_Explicit_Criteria_Folder (Optional)	<p>If spatially explicit criteria is desired, this input should point to an outer directory for all spatial criteria. A rigid structure <b>MUST</b> be followed in order for the model to run. Within the outer spatial criteria folder, there <b>MUST</b> be the following 3 folders: Sensitivity, Exposure, and Resilience. Vector criteria may then be placed within the desired folder. Each feature in the shapefiles used <b>MUST</b> include a ‘Rating’ attribute which maps to a float or int value desired for use as the rating value of that spatial criteria area.</p> <ul style="list-style-type: none"> <li>• Any criteria placed within the Resilience folder will apply only to a given habitat. They should be named with the form: habitatname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore.</li> <li>• Any criteria placed within the Exposure folder will apply to the overlap between a given habitat and a given stressor. They should be named with the form: habitatname_stressorname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore.</li> </ul> <p>Any criteria placed within the Sensitivity folder will apply to the overlap between a given habitat and a given stressor. They should be named with the form: habitatname_stressorname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore.</p>	Folder
Criteria_Scores_Folder	Folder containing CSV files providing all criteria information for the run of the Habitat Risk Assessment. There are two types of CSVs- habitat overlap CSVs and the stressor buffer CSV. Habitat CSVs will contain not only habitat-specific criteria information, but also all criteria that impact the overlap between that habitat and all applicable stressors. The stressor buffer CSV will be a single file containing the desired buffer for all stressors included in the assessment. Habitat CSVs should be filled out	Folder

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

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

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

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

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

**NOTE:**

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

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

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

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

### 3.6.6 Habitat Risk Assessment Preprocessor (InVEST 3.3.3)

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

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

Use_Spatially_Explicit_Risk_Scores_in_Shapefile	<p>Checked: checking this box indicates that model should use criteria from provided shapefiles</p> <p>Unchecked (default): do NOT use spatial criteria from provided shapefiles</p>	Boolean
Spatially_Explicit_Criteria_Folder (Optional)	<p>Each shapefile in the folder directories will need to contain a 'Rating' attribute to be used for calculations in the HRA model. If spatially explicit criteria is desired, this input should point to an outer directory for all spatial criteria. A rigid structure <b>MUST</b>be followed in order for the model to run. Within the outer spatial criteria folder, there <b>MUST</b>be the following 3 folders: Sensitivity, Exposure, and Resilience. Vector criteria may then be placed within the desired folder. Each feature in the shapefiles used <b>MUST</b>include a 'Rating' attribute which maps to a float or int value desired for use as the rating value of that spatial criteria area.</p> <ul style="list-style-type: none"> <li>• Any criteria placed within the Resilience folder will apply only to a given habitat. They should be named with the form: habitatname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore.</li> <li>• Any criteria placed within the Exposure folder will apply to the overlap between a given habitat and a given stressor. They should be named with the form: habitatname_stressorname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore.</li> </ul> <p>Any criteria placed within the Sensitivity folder will apply to the overlap between a given habitat and a given stressor. They should be named with the form: habitatname_stressorname_criterianame.shp. Criteria names may contain more than one word if separated by an underscore.</p>	Folder

## 3.7 Socioeconomic Analysis Toolset

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

### 3.7.1 Cost-Benefit Analysis (Wildlife Transfer)

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

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

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

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

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

Cell_Size (Optional)	The size of the grid units measured in the projection units of the AOI. For example, UTM projections use meters.	Double
----------------------	--	--------

## 3.8 Future Development

Updated and new versions of the telecoupling toolbox are released periodically where new script tools are added or modifications are made to existing tools to fix errors or improve their functionalities. The source code, sample data, and tool documentation are publicly available on Bitbucket<sup>7</sup>, 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.

## 3.9 References

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

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

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*

<sup>7</sup> <https://github.com/f-tonini/telecoupling-toolbox>

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

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

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

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

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

### 3.10 Acknowledgements

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

### 3.11 License & Copyright

Telecoupling Toolbox (“Software”) is the property of Michigan State University (“MSU”) and is made available solely for educational or non-commercial use.

Copyright © 2017 Board of Trustees of Michigan State University. All rights reserved

Telecoupling Toolbox (“Software”) is the property of Michigan State University (“MSU”) and is made

available solely for educational or non-commercial use. Your receipt and use of the Software is governed by the following terms and conditions.

If you are not willing to accept these terms and conditions, please do not download the Software.

Except for third party materials included therein, MSU retains title to the Software, and you shall not obtain any ownership rights in Software.

Portions of Software include intellectual property of ESRI and its licensor and are used under license. Copyright © 2017 ESRI and its licensors . All rights reserved.

Redistribution and use of Software in source and binary forms, with or without modification, are

permitted provided that redistributions of the Software include these terms and conditions.

YOU UNDERSTAND AND AGREE THAT THE SOFTWARE IS PROVIDED "AS IS", AND MSU DISCLAIMS ALL

WARRANTIES,

EXPRESS OR IMPLIED WITH REGARD TO SOFTWARE, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY,

FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OF THIRD PARTY PROPRIETARY RIGHTS.

MSU ASSUMES NO LIABILITY UNDER THIS AGREEMENT. IN NO EVENT WILL MSU BE LIABLE FOR ANY LOSS OF DATA, LOST PROFITS,

COST OF PROCUREMENT OF SUBSTITUTE TECHNOLOGY OR SERVICES OR FOR ANY SPECIAL, INCIDENTAL, CONSEQUENTIAL OR INDIRECT

DAMAGES ARISING FROM THE USE OF THE SOFTWARE OR OTHERWISE ARISING OUT OF THIS AGREEMENT, HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY,

WHETHER FOR BREACH OF CONTRACT, TORT (INCLUDING NEGLIGENCE) OR OTHERWISE. THIS LIMITATION WILL APPLY EVEN IF MSU HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

THESE LIMITATIONS SHALL APPLY NOTWITHSTANDING THE FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY.

You agree that Software will not be disclosed in violation of any applicable export control regulations.

Please include the following notice of attribution in any publications or presentations reporting on your use of the Software:

Telecoupling Toolbox [version\_number\_here] (<https://bitbucket.org/f-tonini/telecoupling-geoapp>) developed by Dr. Tonini at Michigan State University.

Other than the notice of attribution described above, neither the name of MSU nor the names of Software developers may be used without prior written permission.

If you have any questions about this agreement, please contact:

MSU Technologies

[msut@msu.edu](mailto:msut@msu.edu)

TEC2017-0047

- This toolbox depends on the R Statistical Computing Software:

© 2017 The [R Foundation for Statistical Computing](#). R is free software and comes with ABSOLUTELY NO WARRANTY. See the [COPYRIGHTS](#) file for details.

- This toolbox depends on [ESRI software](#):  
© 2017 ESRI. See the [Software License and Agreement](#) for details.
- This toolbox depends on [InVEST - Natural Capital Project software](#):  
© 2017 NatCap Project. See the [Software License and Agreement](#) for details.

## 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.

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

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

### 4.1 SAMPLE DATASET

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

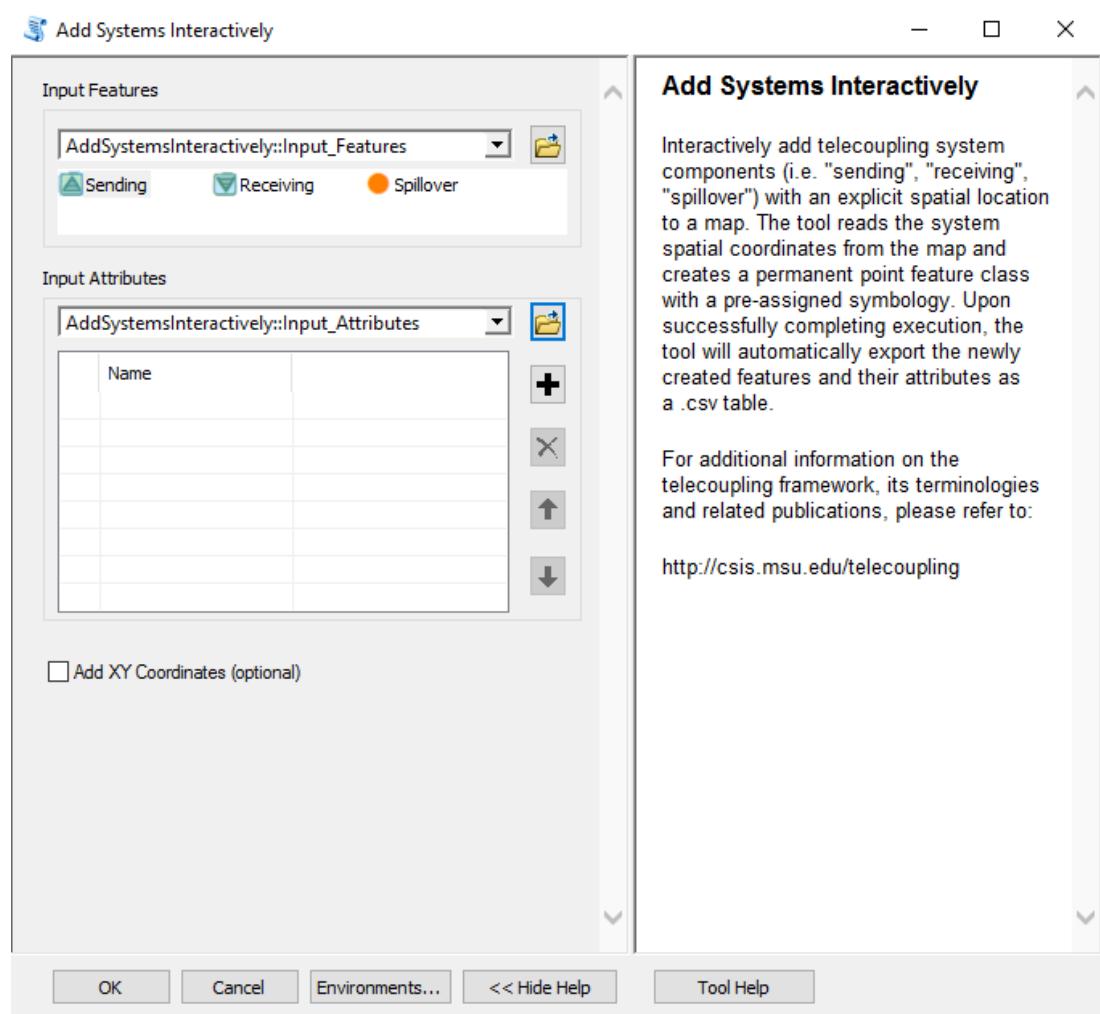
**DISCLOSURE ON DATA USAGE:** the sample datasets were either partially modified, or their format changed in some case from the original datasets to protect privacy and sensitive data when deemed appropriate. Use the sample datasets for learning purposes only and under no circumstances data should be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

 <b>admin</b>	10/21/2016 10:22 ...	File folder
 <b>AGENTS</b>	3/22/2017 4:02 PM	File folder
 <b>carbon</b>	3/22/2017 3:59 PM	File folder
 <b>crop</b>	8/14/2017 11:30 AM	File folder
 <b>FLOWs</b>	7/24/2017 12:36 PM	File folder
 <b>gtgp</b>	7/24/2017 12:19 PM	File folder
 <b>habitat_quality</b>	8/14/2017 2:27 PM	File folder
 <b>habitat_risk_assessment</b>	7/7/2017 12:35 PM	File folder
 <b>information</b>	10/7/2016 1:43 PM	File folder
 <b>landcover</b>	7/25/2017 10:16 AM	File folder
 <b>SYSTEMS</b>	3/22/2017 4:02 PM	File folder
 <b>tourism</b>	10/12/2016 1:24 PM	File folder
 <b>trade</b>	9/21/2016 1:30 PM	File folder
 <b>wildlife</b>	11/9/2016 11:26 AM	File folder

Figure 3. File structure of the SampleData folder provided with the telecoupling toolbox. The *admin* folder contains vector layers of administrative boundaries, such as the Wolong Nature Reserve, China, and other geo-political administrative subdivisions. The *carbon* folder contains text files and images with estimates of carbon (above/below ground) and net primary productivity (NPP) for the Wolong Nature Reserve. The *crop* folder contains files needed to run the crop production tool (Environmental Analysis toolset), such as text files of financial and nutrient data for crops produced in the Wolong Nature Reserve, fertilizer application, irrigation, and global datasets of climate-related data for several crop items. The *gtgp* folder contains vector layers with locations and household survey data recorded in the Wolong Nature Reserve area during 2006 as well as a future “toy” scenario to be used for illustration purposes. The *habitat\_quality* folder contains files needed to run the habitat quality tool (Environmental Analysis toolset), such as text files of sensitivity and threat values for panda species in the Wolong Nature Reserve, zoning districts, and raster data of each human and natural threat to the habitat of the species. The *information* folder contains text files related to the spread of information on the Wolong Nature Reserve, in the form of citations extracted from the Academic LexisNexis portal. The *landcover* folder contains raster data of the forest cover and land-use/land-cover classes for the Wolong Nature Reserve in different years. The *tourism* folder contains text files from a survey conducted at the Panda breeding center in Wolong to identify socioeconomic characteristics of each tourist. Moreover, the folder contains a *visitation\_rate* subfolder to be used for demonstration purposes only with the Visitation: Recreation and Tourism (Socioeconomic Analysis toolset) tool. The *trade* folder contains text files on the household survey within the Wolong Nature Reserve, to assess some trade-related information to the production of local crops. The *wildlife* folder contains text files related to financial information (costs and revenues) from the transfer (loan) of panda species between countries worldwide, as well as a “toy” vector layer with locations and information of each zoo involved with the exchange of pandas from the Wolong Nature Reserve. The folders ‘AGENTS’, ‘FLOWs’, and ‘SYSTEMS’, contain csv table with sample data to be used in the respective tools.

## 4.2 SYSTEMS TOOLSET

### 4.2.1 Add Systems Interactively

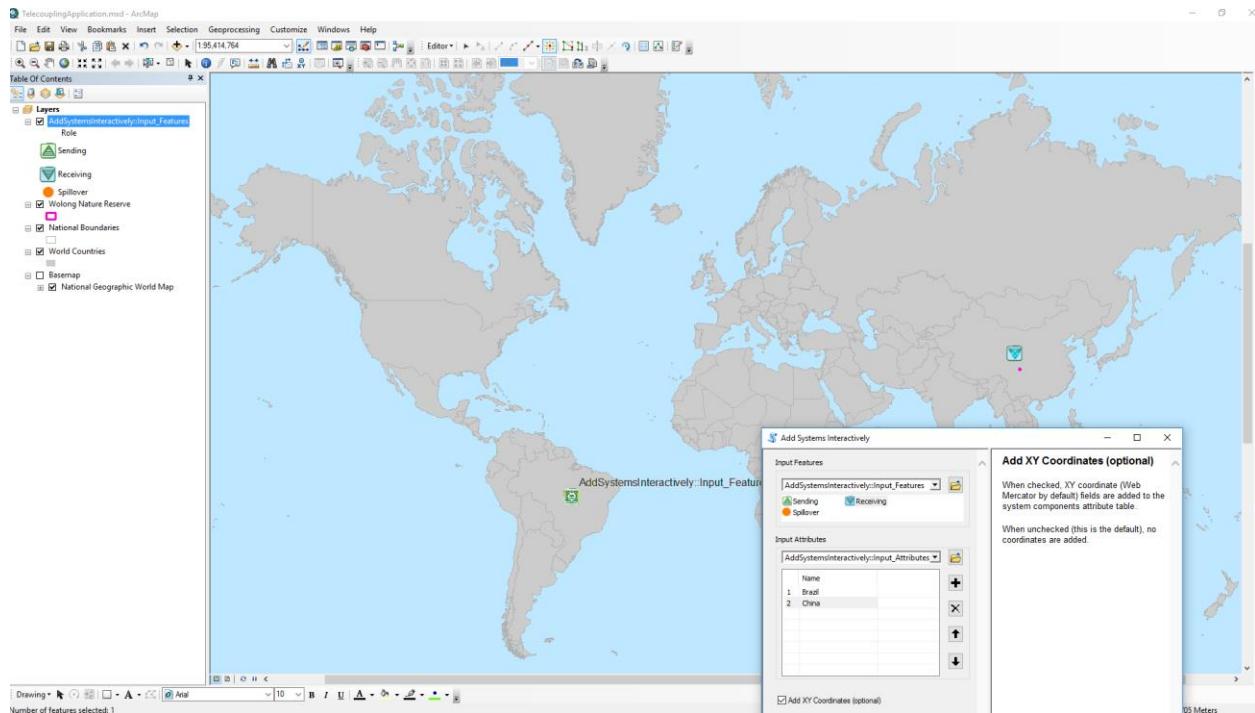


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

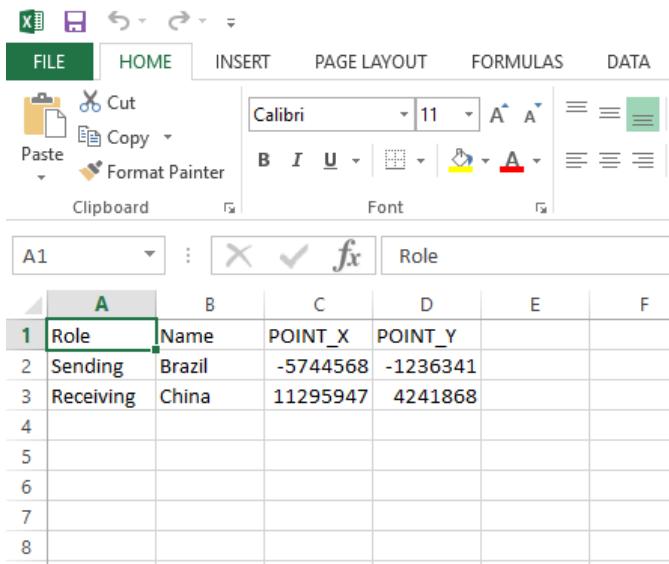
The first parameter asks you to select a category for the system you would like to add to the map (sending, receiving, spillover) with its own symbology. Once you select the desired category, move the mouse cursor over the map and you should see the chosen symbol ready to be placed.

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.*



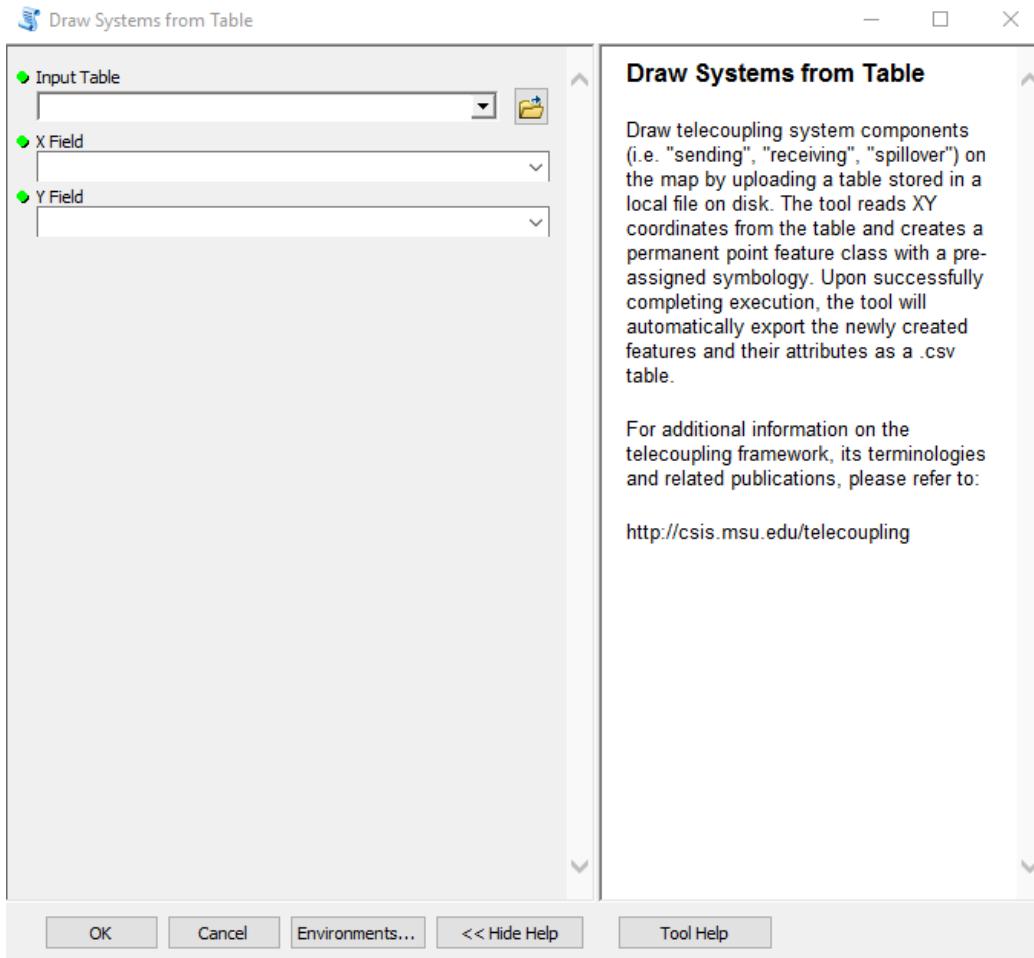
The output .csv table created by the tool in the scratch workspace directory should look something similar to the following figure:



A screenshot of a Microsoft Excel spreadsheet titled "Role". The table has columns labeled A through F. Column A contains row numbers 1, 2, and 3. Column B contains "Role" and "Name". Column C contains "POINT\_X" and "POINT\_Y". The data is as follows:

	Role	Name	POINT_X	POINT_Y	
1	Sending	Brazil	-5744568	-1236341	
2	Receiving	China	11295947	4241868	
3					
4					
5					
6					
7					
8					

#### 4.2.2 Draw Systems from Table



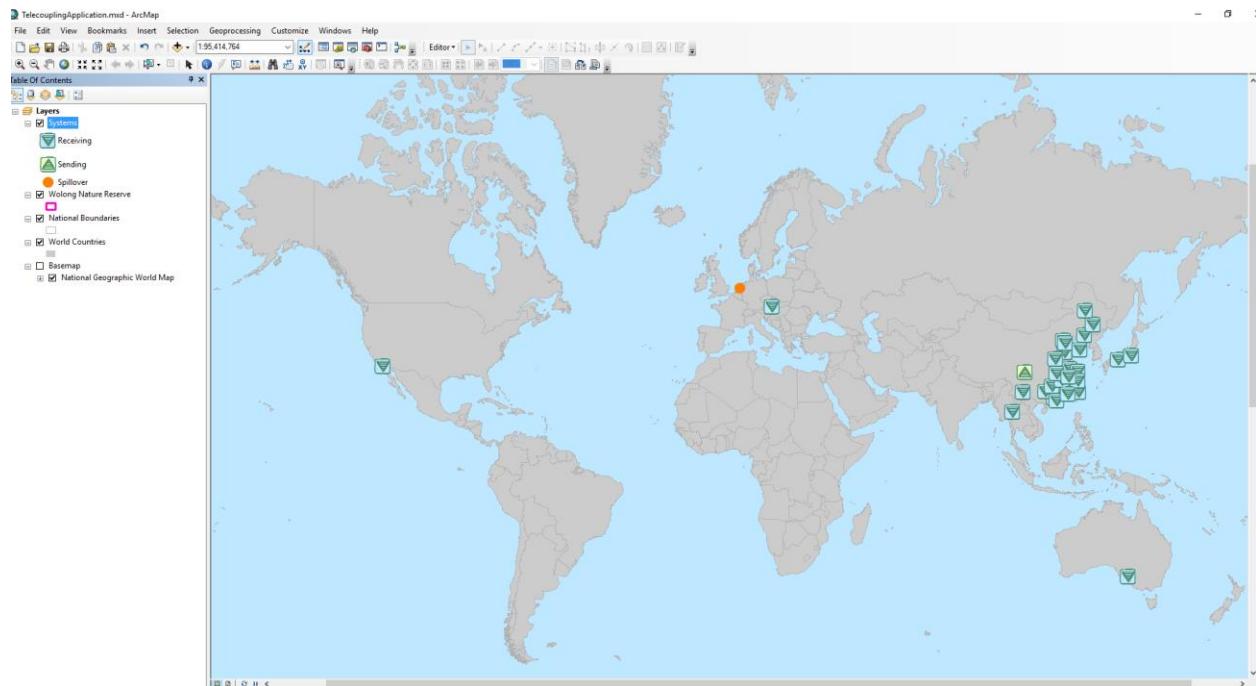
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

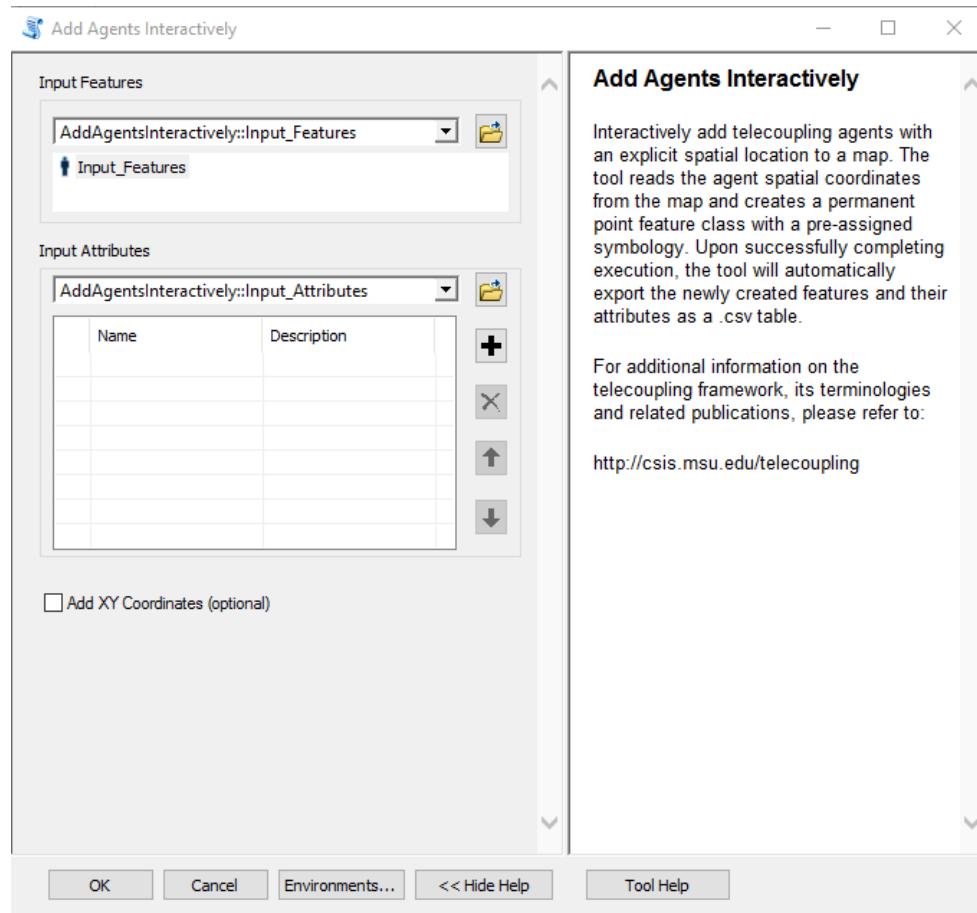
If you are using the sample data provided with the toolbox, you can use any of the .csv tables ending with the word “\_Systems” (e.g. `/SampleData/SYSTEMS/wildlife_Systems.csv`).

***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.



## 4.3 AGENTS TOOLSET

### 4.3.1 Add Agents Interactively

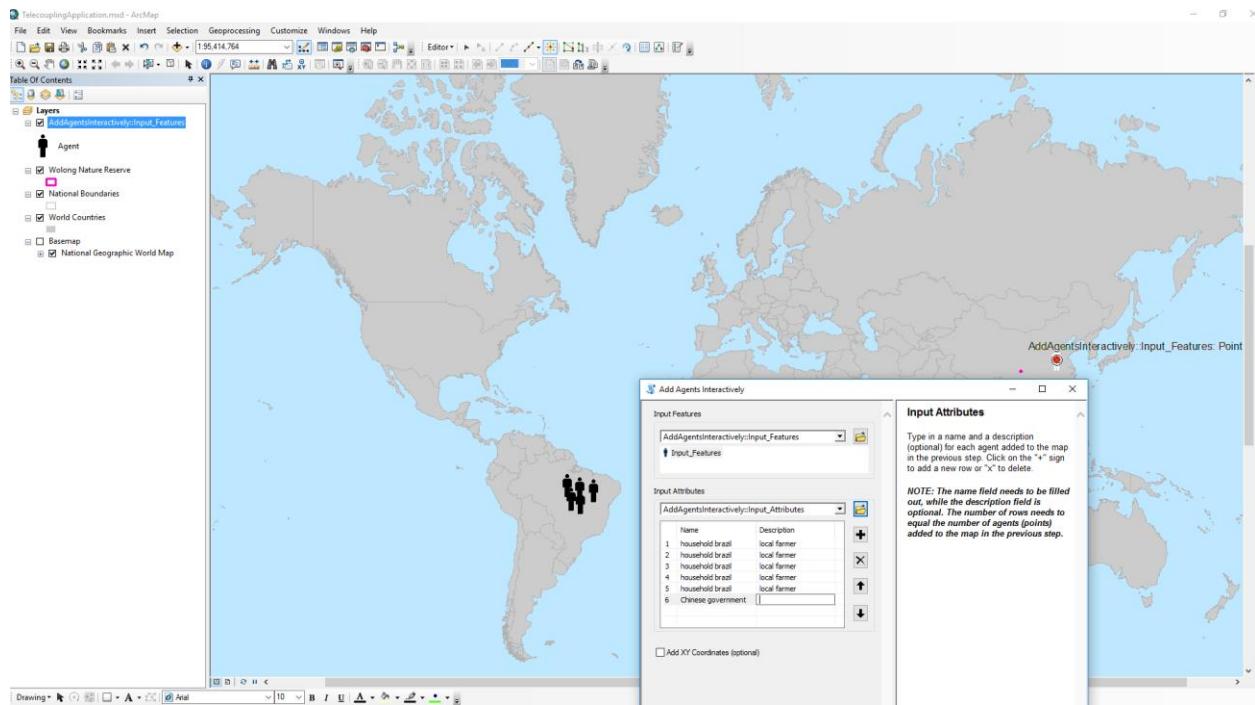


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

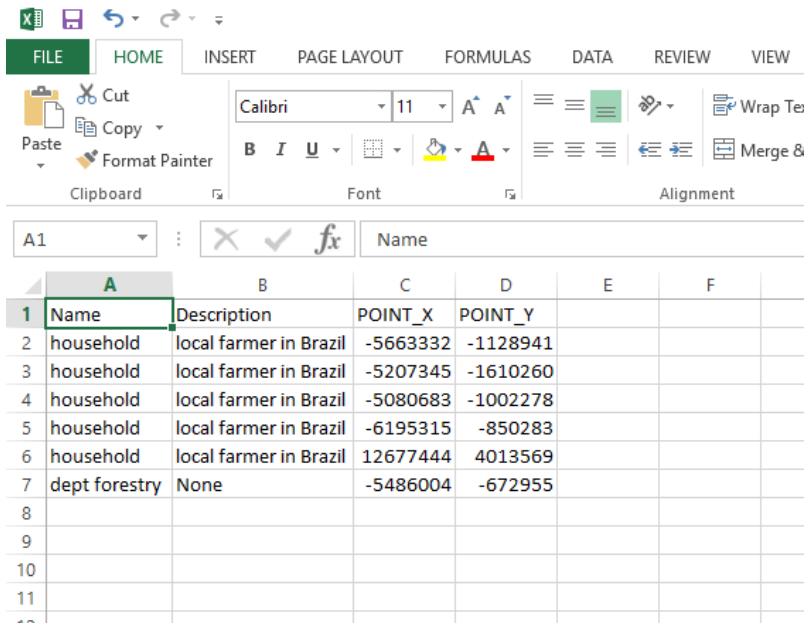
The first parameter asks you to select agents (e.g. household, organization, governmental departments) you would like to add to the map. Move the mouse cursor over the map and you should see an agent symbol ready to be placed. Make sure to zoom in/out to the extent that best captures your agent size 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.

***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.



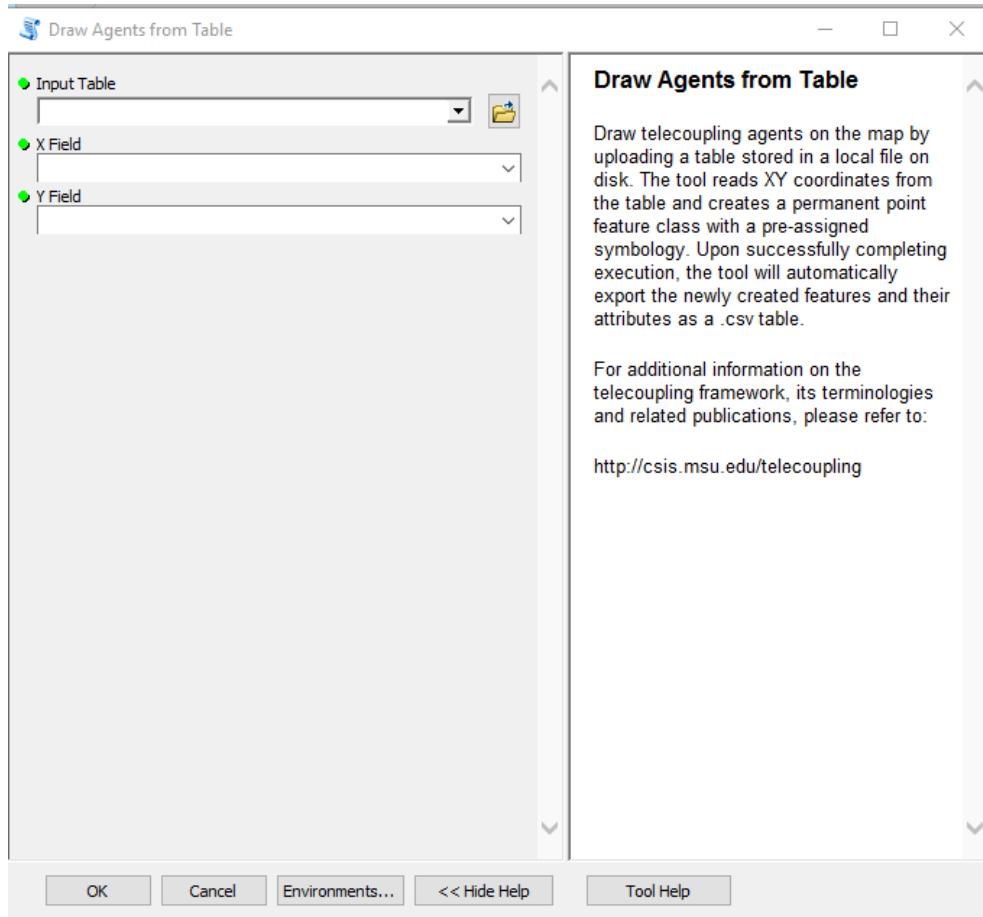
The output .csv table created by the tool in the scratch workspace directory should look something similar to the following figure:



The screenshot shows a Microsoft Excel spreadsheet titled 'Name'. The table has columns labeled A, B, C, D, E, and F. Column A contains row numbers from 1 to 7. Column B contains names like 'household' and 'dept forestry'. Column C contains descriptions. Columns D and E contain numerical coordinates. Column F is empty.

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

#### 4.3.2 Draw Agents from Table

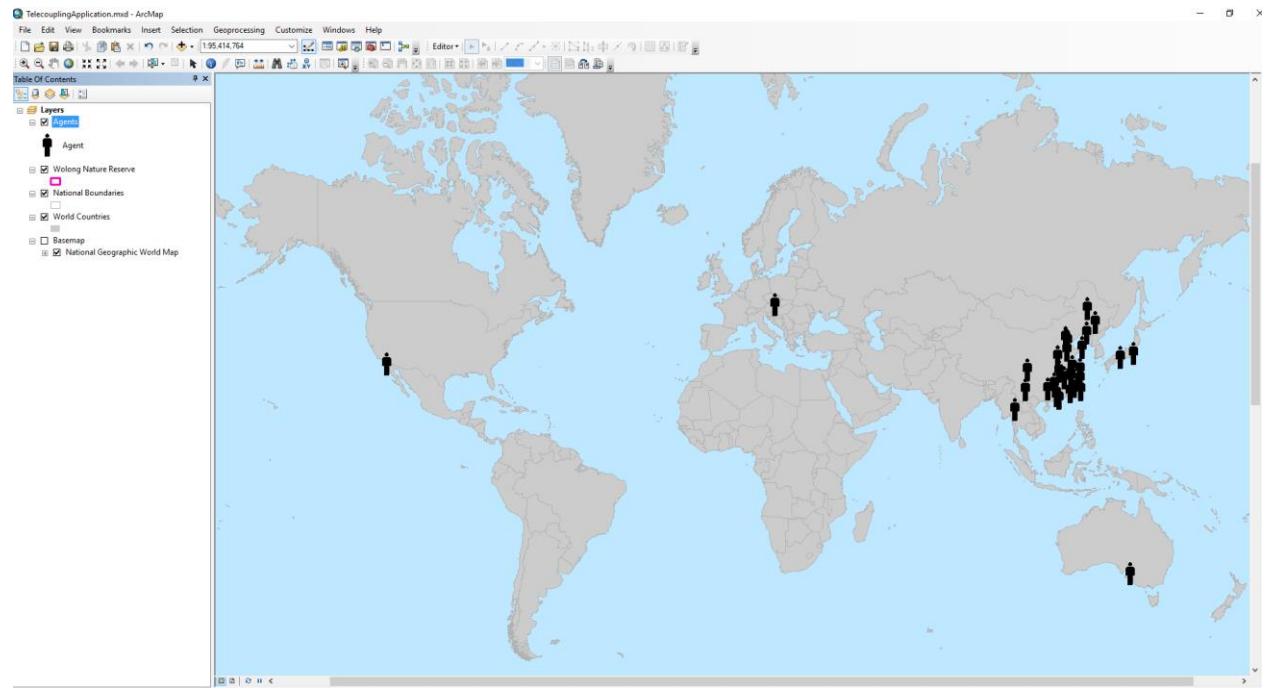


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

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

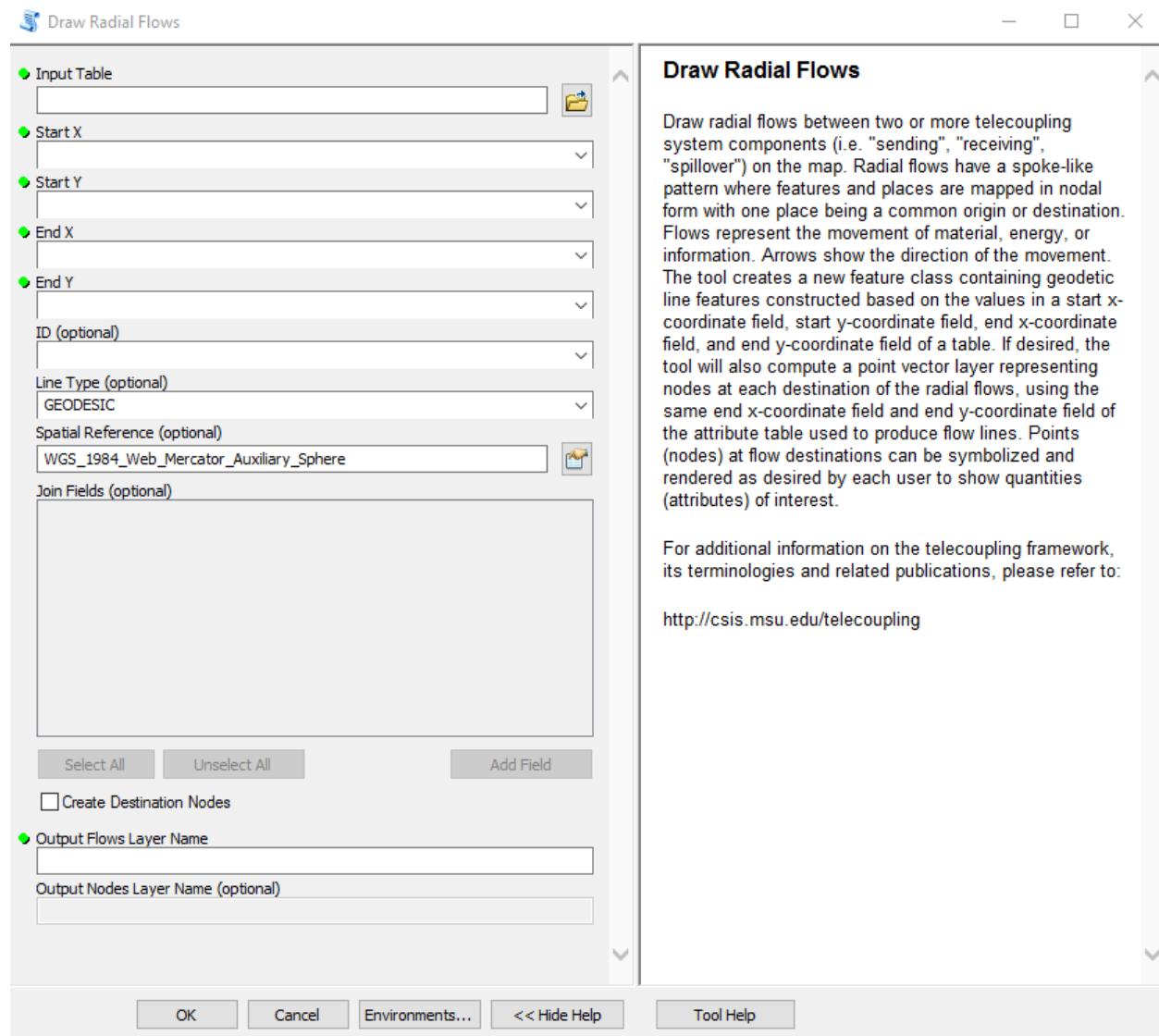
If you are using the sample data provided with the toolbox, you can use any of the .csv tables ending with the word “\_Agents” (e.g. */SampleData/AGENTS/wildlife\_Agents.csv*).

**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.



## 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.

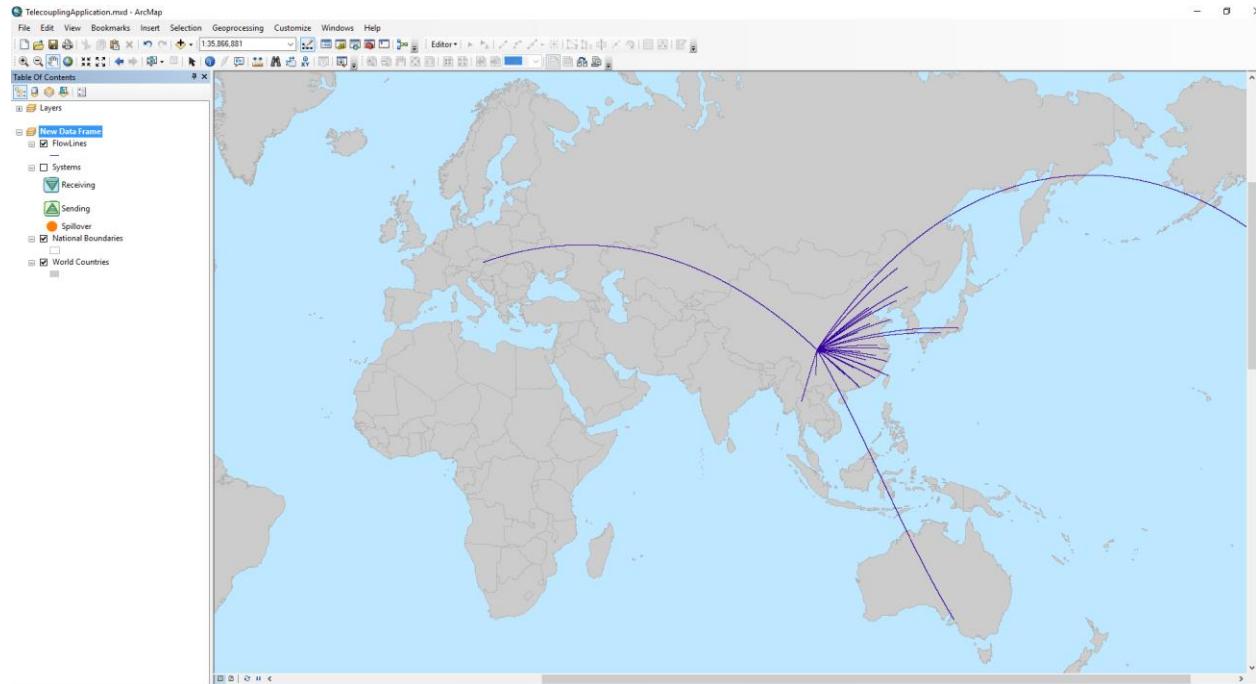
The first parameter asks you to select a table on disk containing records on start and end points of the flow. The next four parameters need the field name (column) of the input table containing

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 joined to the output flow attribute table.

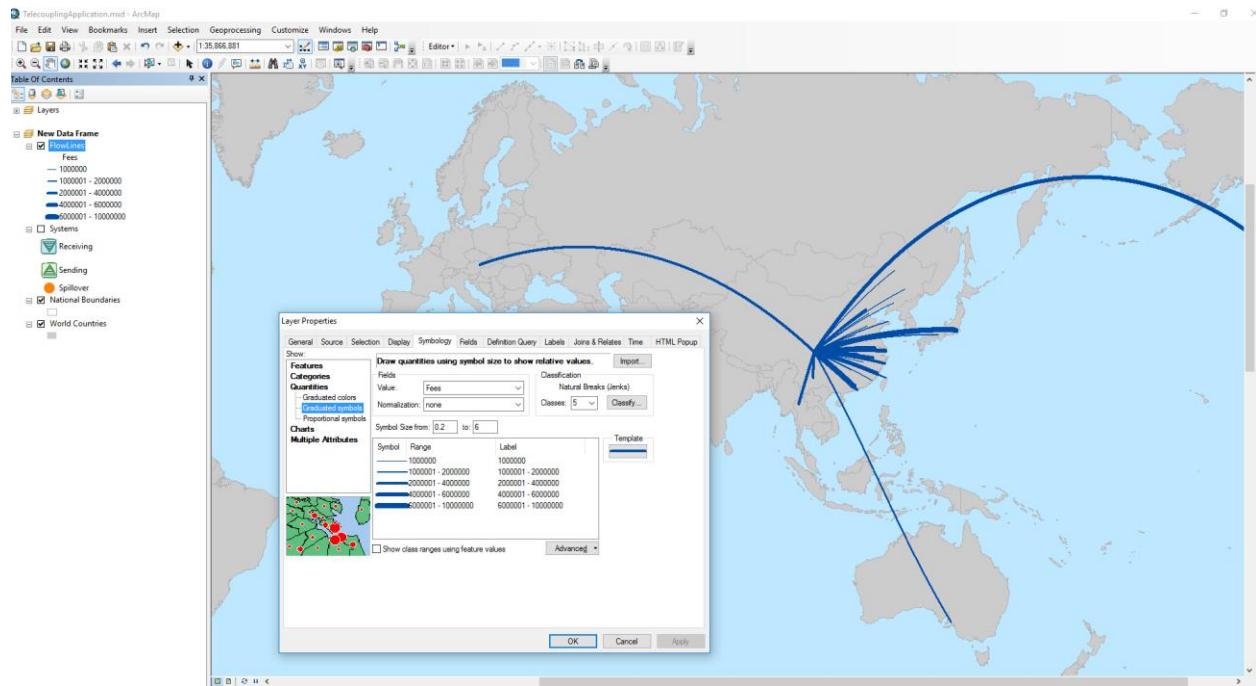
Once you click OK to run the tool, ArcGIS will create a permanent output line feature class and add it to the map.

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/FLOWS/wildlife\_Flows.csv). Use the column called “FID” as ID parameter, select WGS 1984 as a spatial reference for the spatial coordinates found in the sample table, and choose one or more additional columns that show the quantity of the flow.

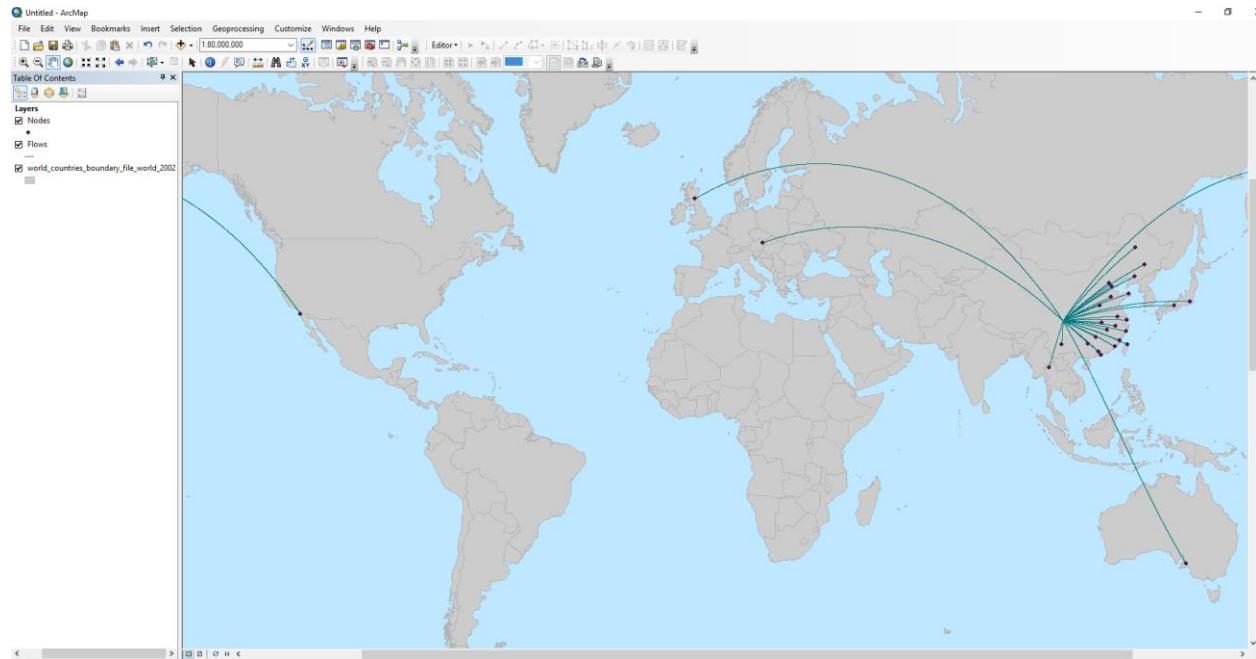
***NOTE: In order to change the default workspace location where your output files and layers are created, click on the “Environments...” button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.***



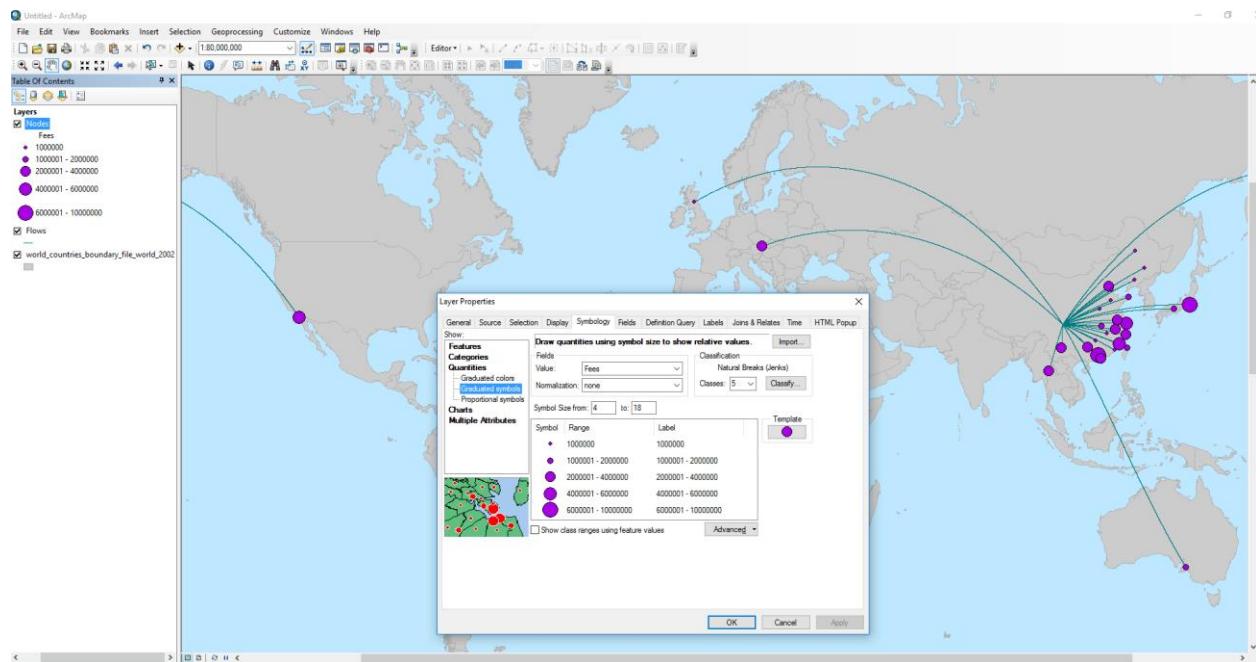
Once the output flow feature class has been created, the lines will all have the same color and thickness. If you have selected additional variables to represent the flow amount, you can use one of them to improve the visualization of the lines. To do so, right-click on the output flow layer found in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



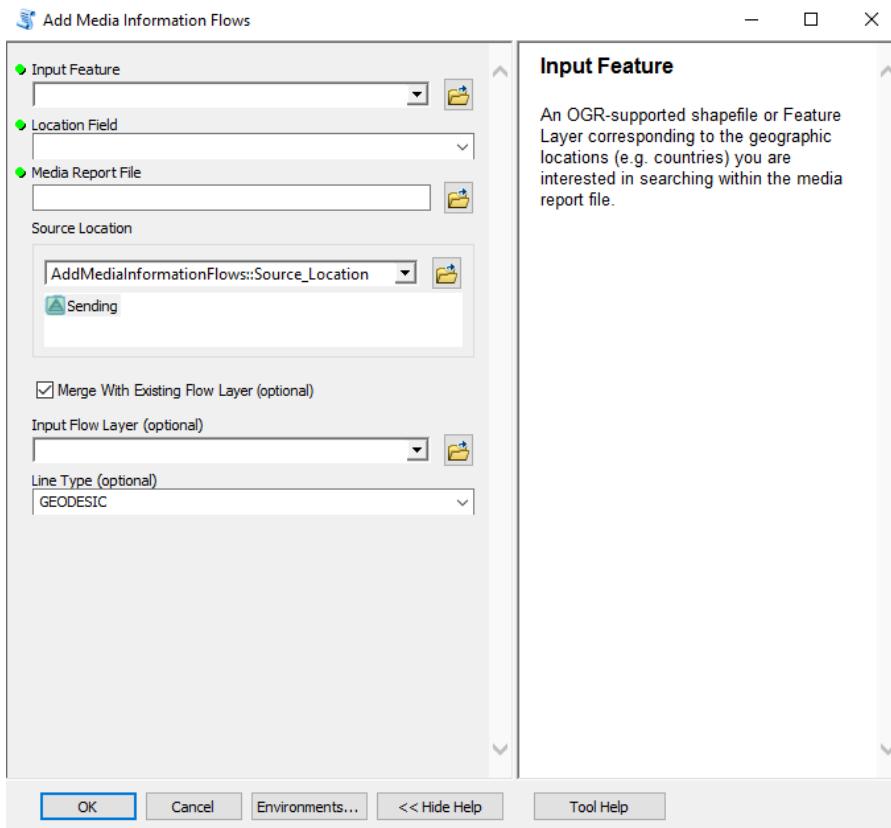
If in addition you desire to create a point layer with nodes at flow destinations, repeat all the steps above, making sure you check the box next to “Create Destination Nodes”. Give a name to the layers and click OK. ArcGIS will create a permanent output point feature class representing destination points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



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



#### 4.4.2 Add Media Information Flows



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

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

Go to the LexisNexis website<sup>8</sup>, and you should see a window similar to the next figure:

The screenshot shows the LexisNexis Academic search interface. At the top, there is a browser header with icons for back, forward, refresh, and home, followed by the URL 'www.lexisnexis.com/hottopics/lnacademic/'. Below the URL is a toolbar with links to 'Apps', 'Bookmarks', 'Blogs', 'Data Science', 'GeoDev', 'GIS', 'Jobs', 'Programming', 'Telecoupling\_MSU', 'Spotify Web Player', and a 'Source Directory' section with 'Find or Browse' and 'Create Permanent' options. A note below the toolbar states 'Use of this service is subject to Terms and Conditions'. On the left, a vertical sidebar has links for 'What's New', 'Video Tutorials', 'Research Guides', 'Download Content List', and 'Academic Knowledge Center'. A blue 'Tools' button is highlighted. The main search area has a red header bar with the search term 'wolong nature reserve' and a 'Search' button. Below the search bar is an 'Advanced Options' dropdown. A promotional box for the 'Social Sciences Student Survey!' is present. The 'Hot Topics Links' section lists current news items: ISIS, Zika Virus, 2016 Presidential Election; Brexit, New York and New Jersey Bombings, National Football League; Columbia and FARC, Supreme Court, Nobel Prizes. At the bottom, three buttons provide links to 'Search the News' (U.S. and World News), 'Look up a Legal Case' (Federal and state cases), and 'Get Company Info' (Over 80 million companies).

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:

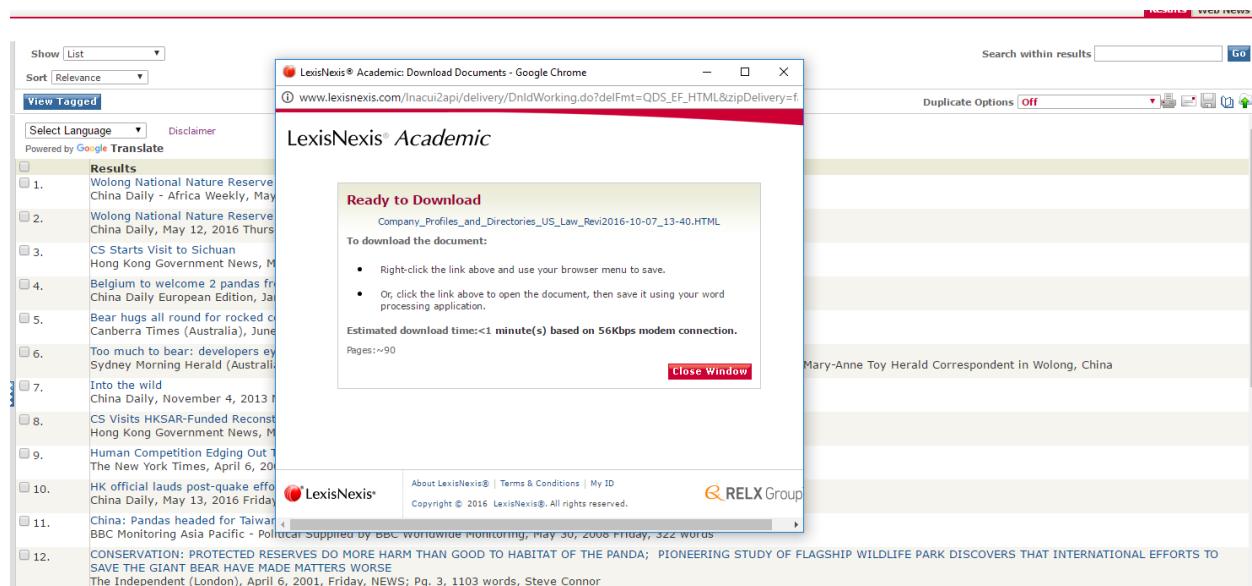
<sup>8</sup> <http://www.lexisnexis.com/hottopics/lnacademic/>

The screenshot shows a search results page from the Telecoupling Toolbox. The results are filtered by relevance and show 1-25 of 906 records. The results list includes various news articles from sources like China Daily, Hong Kong Government News, and The New York Times, all related to the Wolong National Nature Reserve.

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

The screenshot shows a "Download Documents" dialog box from LexisNexis Academic. It allows users to download selected documents in various formats (HTML, Word, PDF, etc.). The "Format" dropdown is set to "HTML". The "Document Range" section has "Select Items" selected and contains the range "1-50". The "Page Options" section includes "Cover Page" and "Add a Brief Note (appears on cover page)". The "Font Options" section includes "Times New Roman" and "Search Terms in Bold Type". The "Download" button is highlighted.

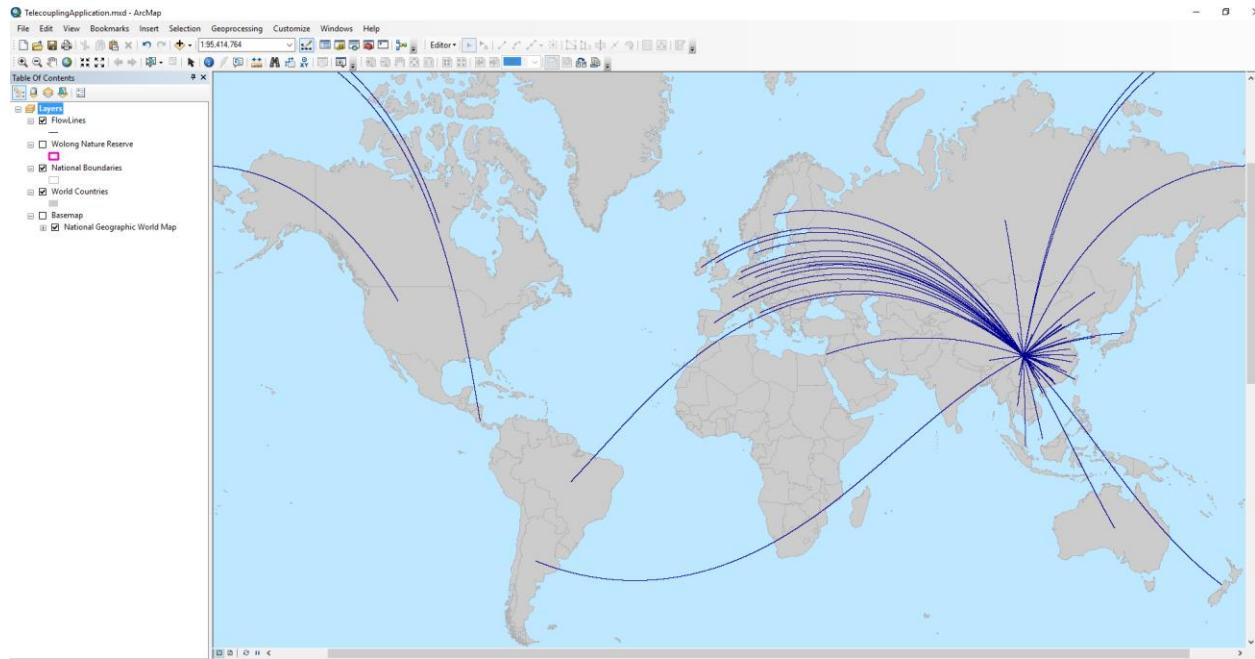
In the “Format” drop-down menu, select HTML. The document range tells you how many records have been found and whether you want to save all of them (default) to the HTML file. If you have too many records, an error message should come up asking to reduce the number of selected documents. In this example,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

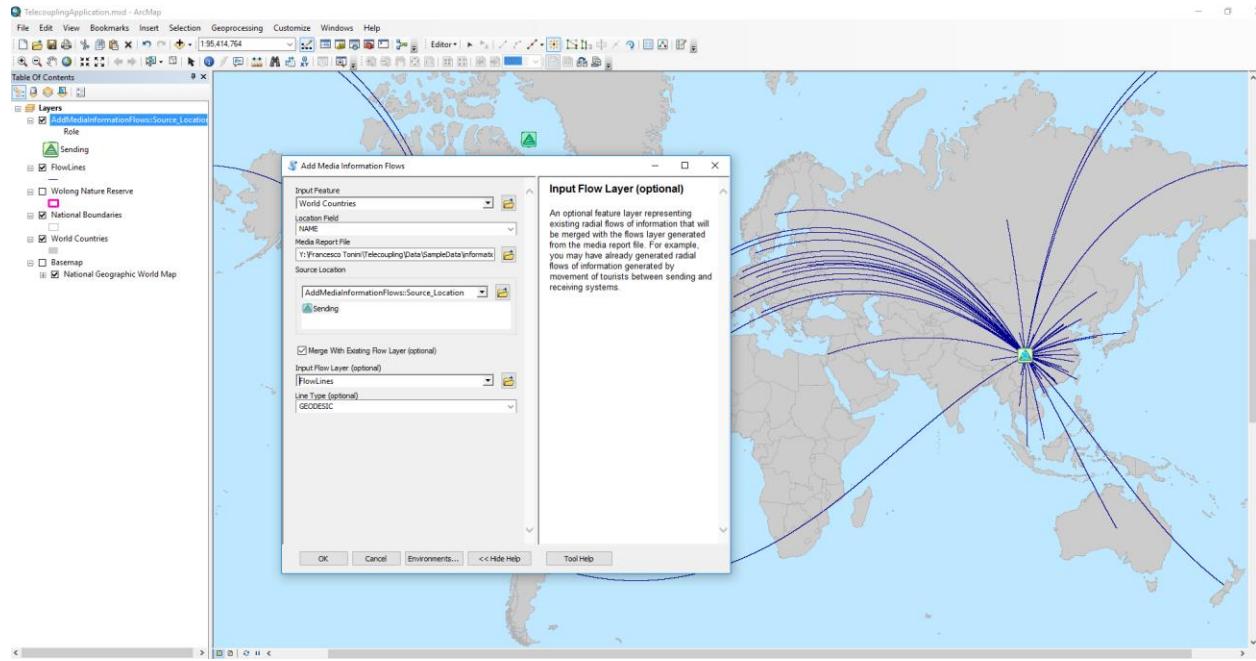
If you decided to skip the LexisNexis online word search, continue from this point on. For this tutorial, we are going to merge the media flows obtained from the HTML report file with an existing flow layer. To do this, use the “[Draw Radial Flows](#)” tool, and use the “`./SampleData/FLOWS/info_Flows.csv`”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

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

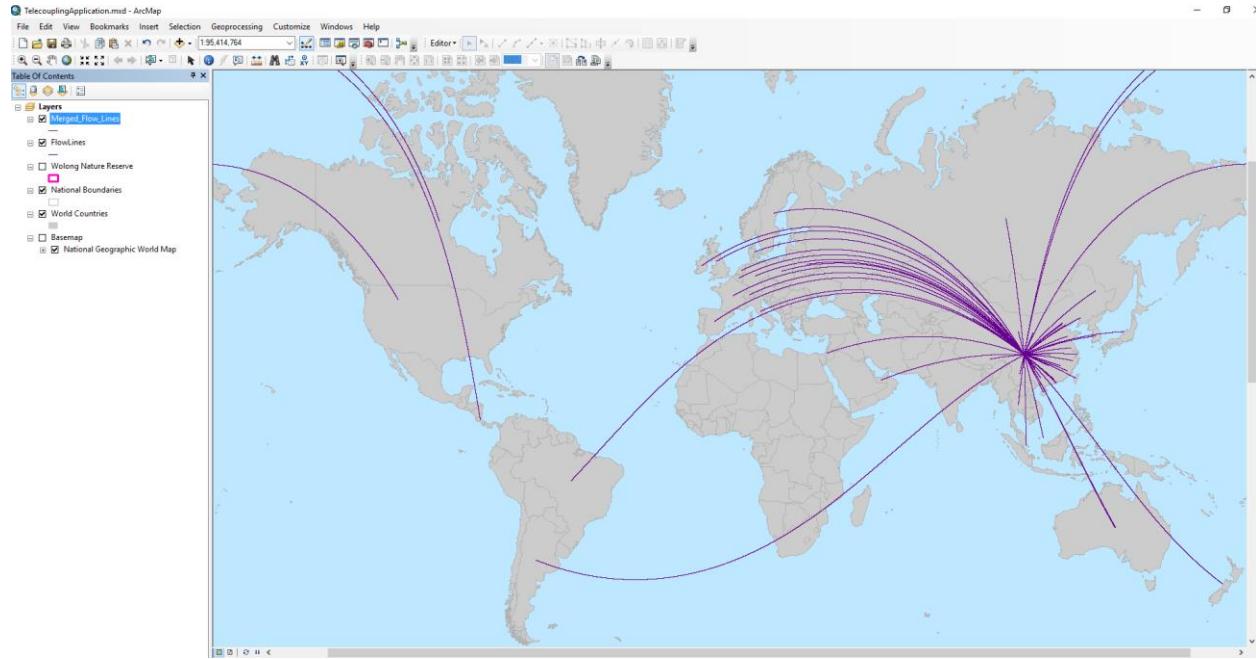


Now we are ready to use the “Add Media Information” tool. If you are using the *TelecouplingApplication.mxd* file, select “World Countries” as the input feature layer for the first parameter of the tool. If you are using a brand new map, then select the “*./SampleData/admin/World\_countries\_2002.shp*” 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.

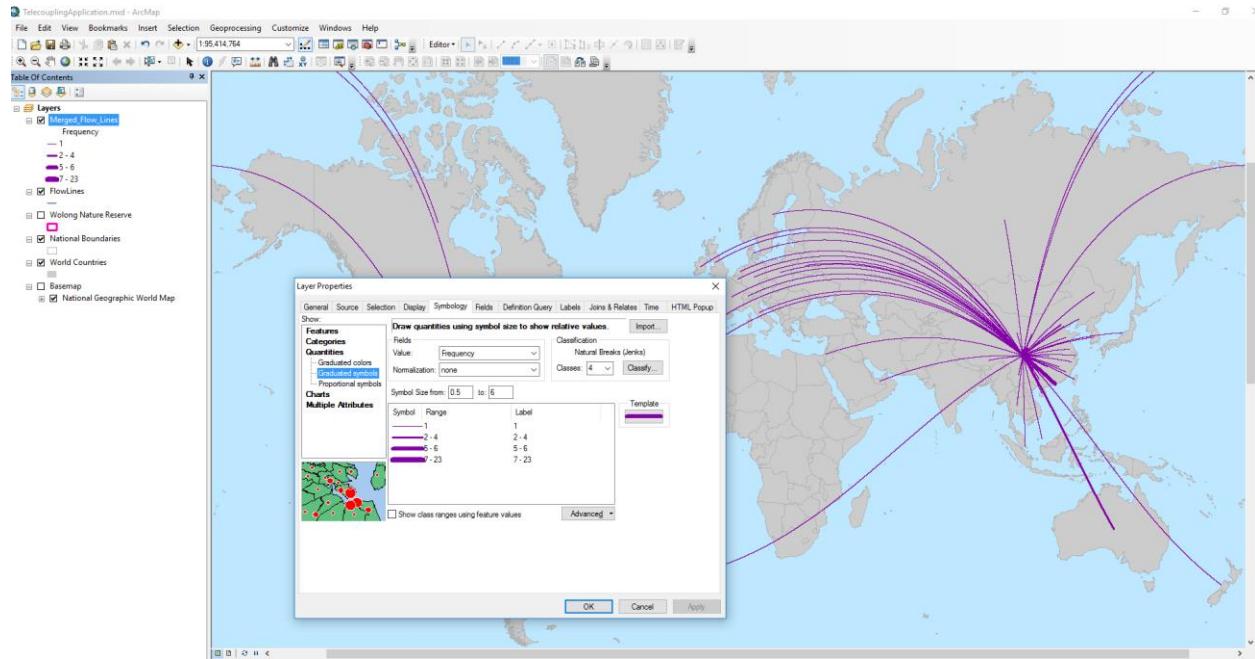
**NOTE:** In order to change the default workspace location where your output files and layers are created, click on the “Environments...” button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.



Once 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.

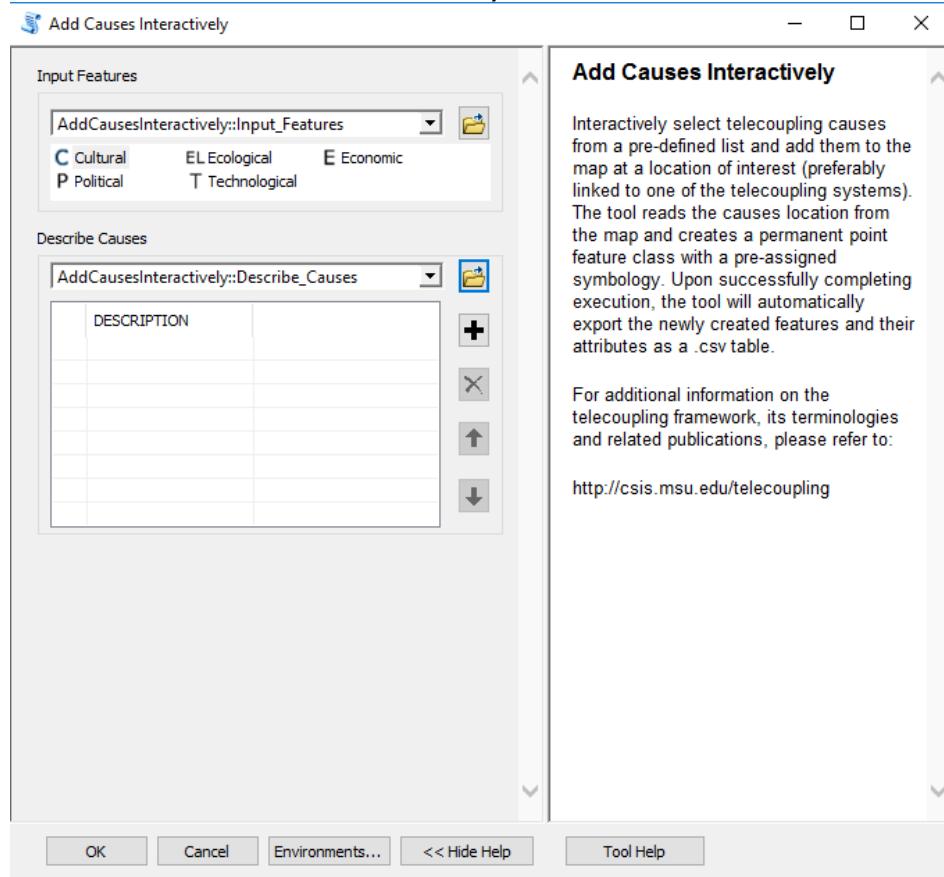


Once the output flow feature class has been created, the lines will all have the same color and thickness. If you have selected additional variables to represent the flow amount, you can use one of them to improve the visualization of the lines. To do so, right-click on the output flow layer found in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



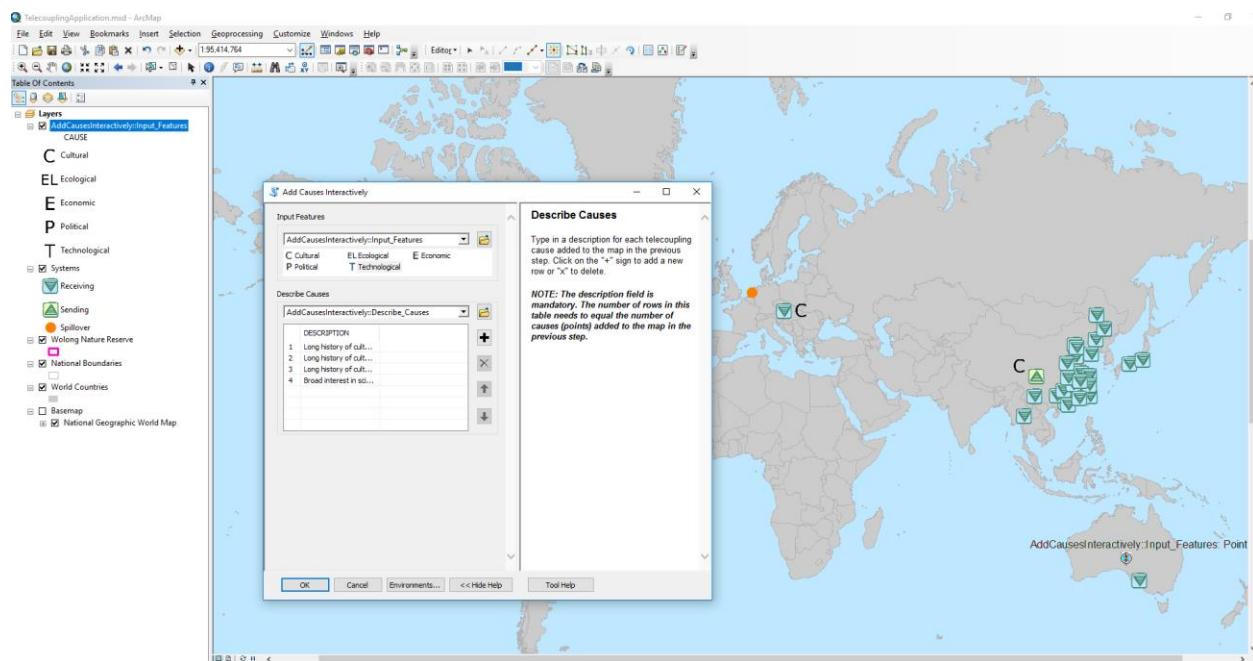
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

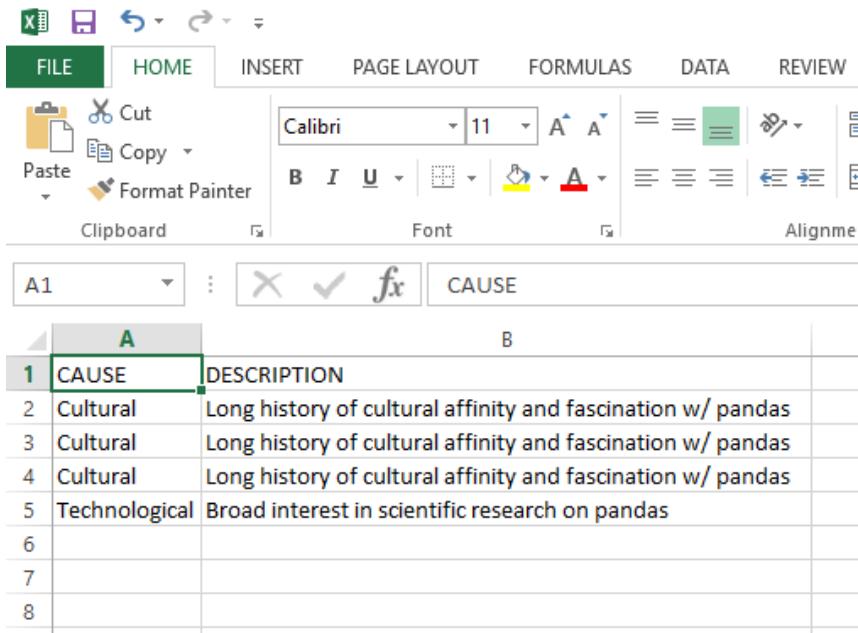
Alternatively, the factor instigating the flow may be solely economic given the anticipated return

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 causal categories and their description you added to the map.

***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.*



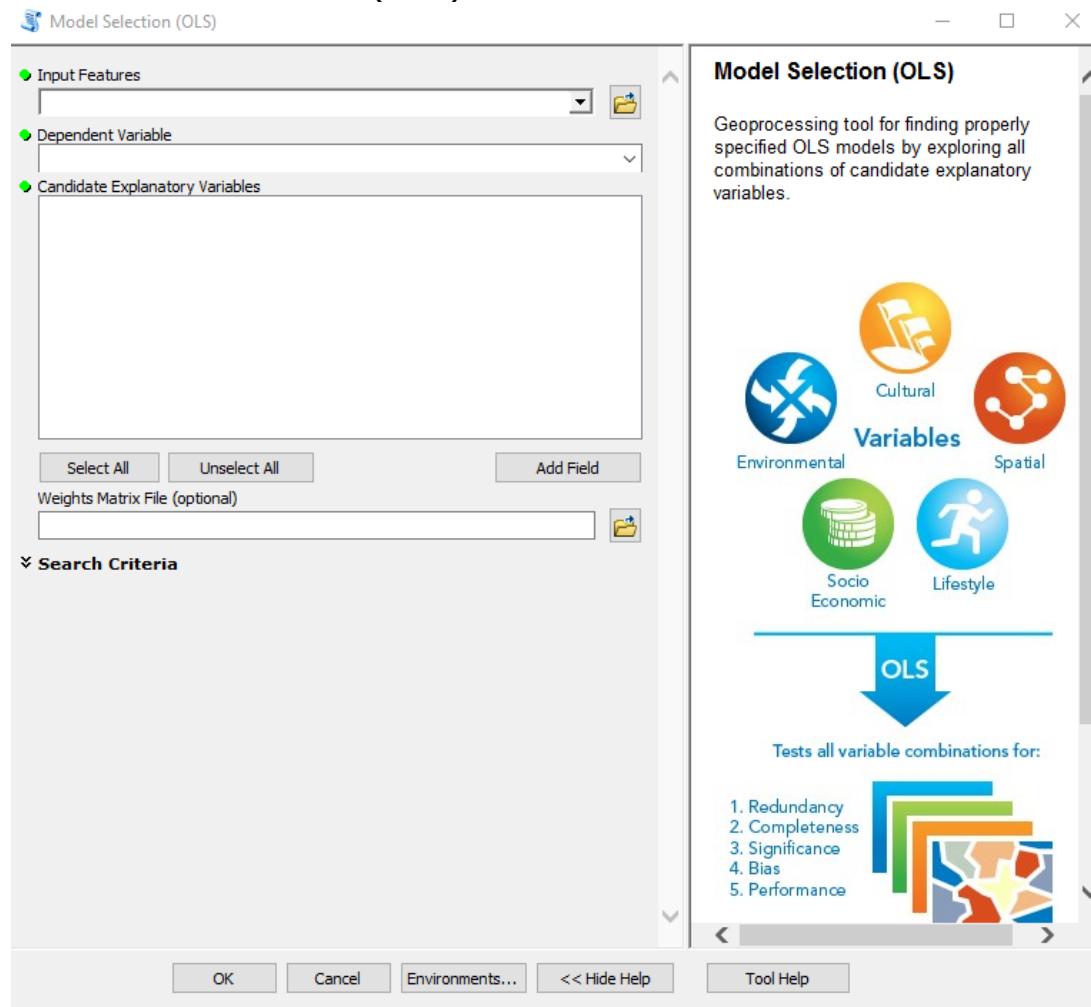
The output .csv table created by the tool in the scratch workspace directory should look something similar to the following figure:



The screenshot shows a Microsoft Excel spreadsheet titled "CAUSE". The table has two columns: "A" (Cause) and "B" (Description). The first row is a header with "1" and "CAUSE". Rows 2 through 5 list causes under "A" and their corresponding descriptions under "B". Row 6 is blank, and rows 7 and 8 are also blank.

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	

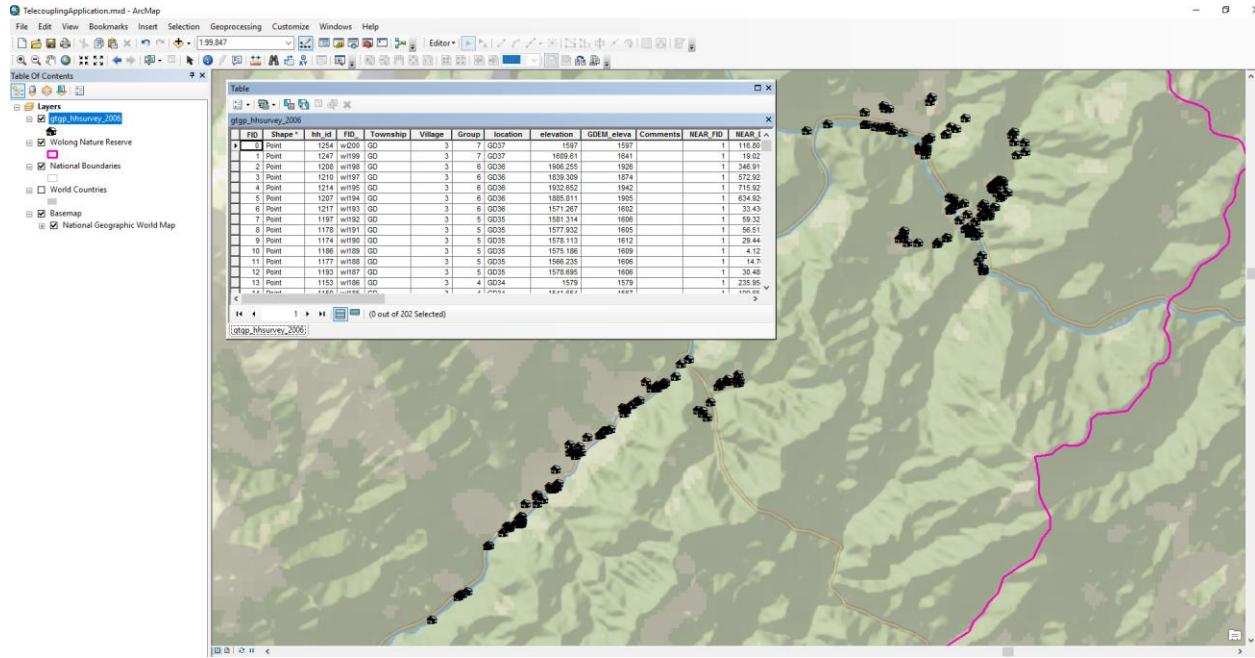
#### 4.5.2 Model Selection (OLS)



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
-----

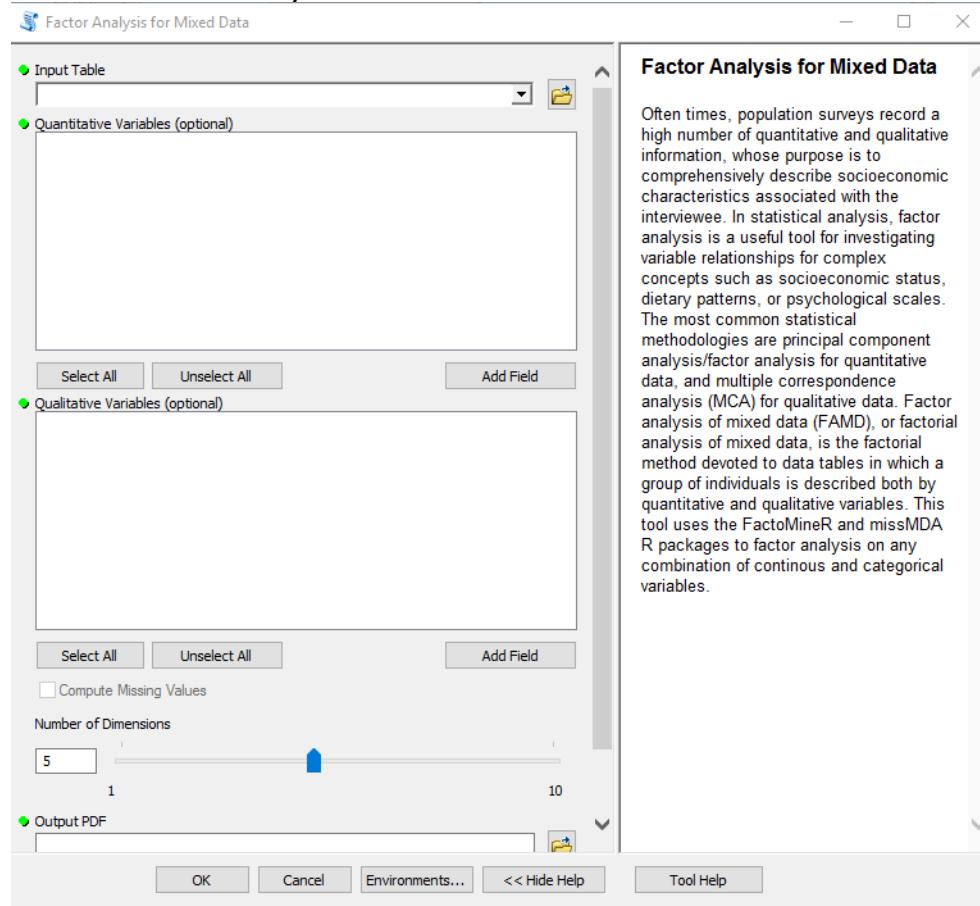
Summary of Multicollinearity
Variable VIF Violations Covariates
GTGP 1.16 0 -----
GTBP 1.31 0 -----
TOURISM_PA 1.09 0 -----
NUMBER_OF 1.02 0 -----
LOG_DISTA 1.12 0 -----

```

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.

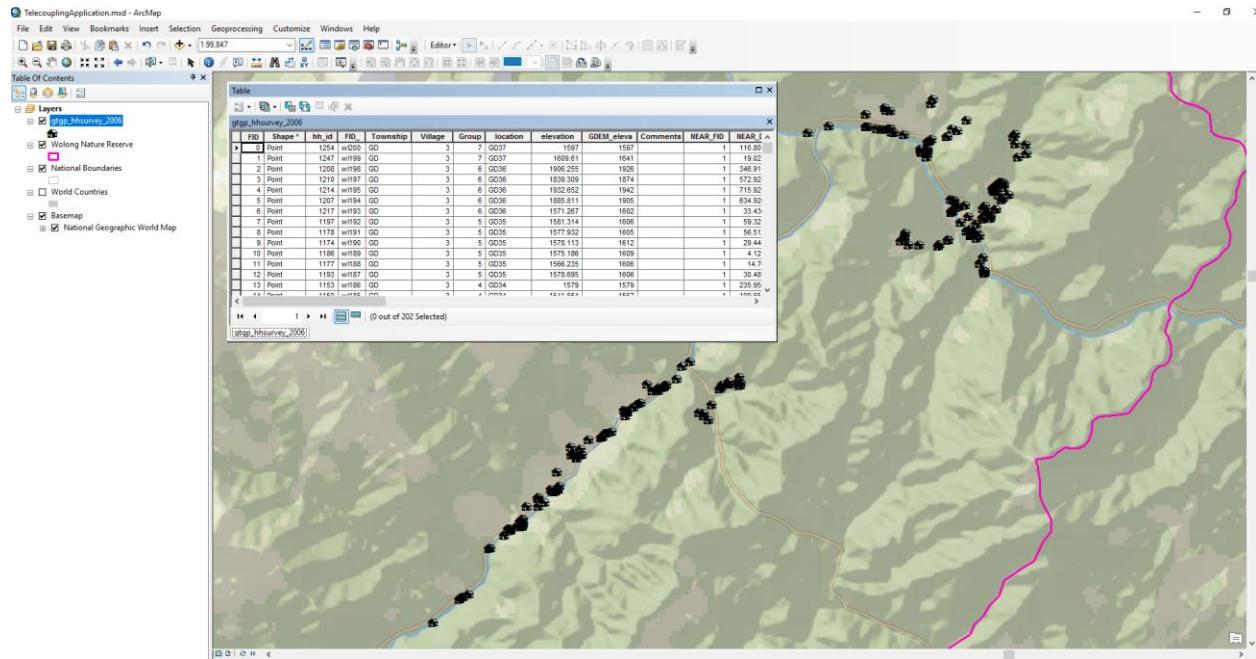
NOTE: In order to change the default workspace location where your output files and layers are created, click on the “Environments...” button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

### 4.5.3 Factor Analysis for Mixed Data

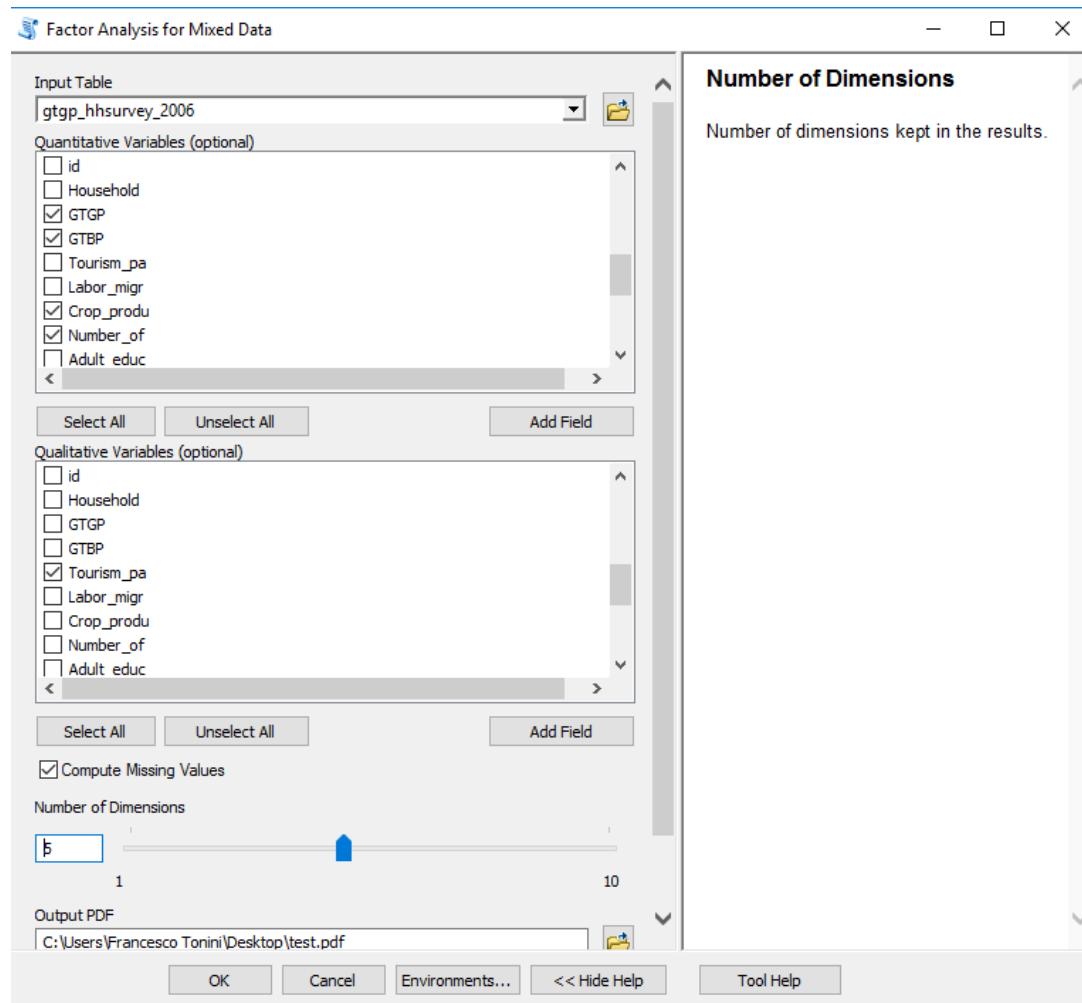


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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



**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.

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:

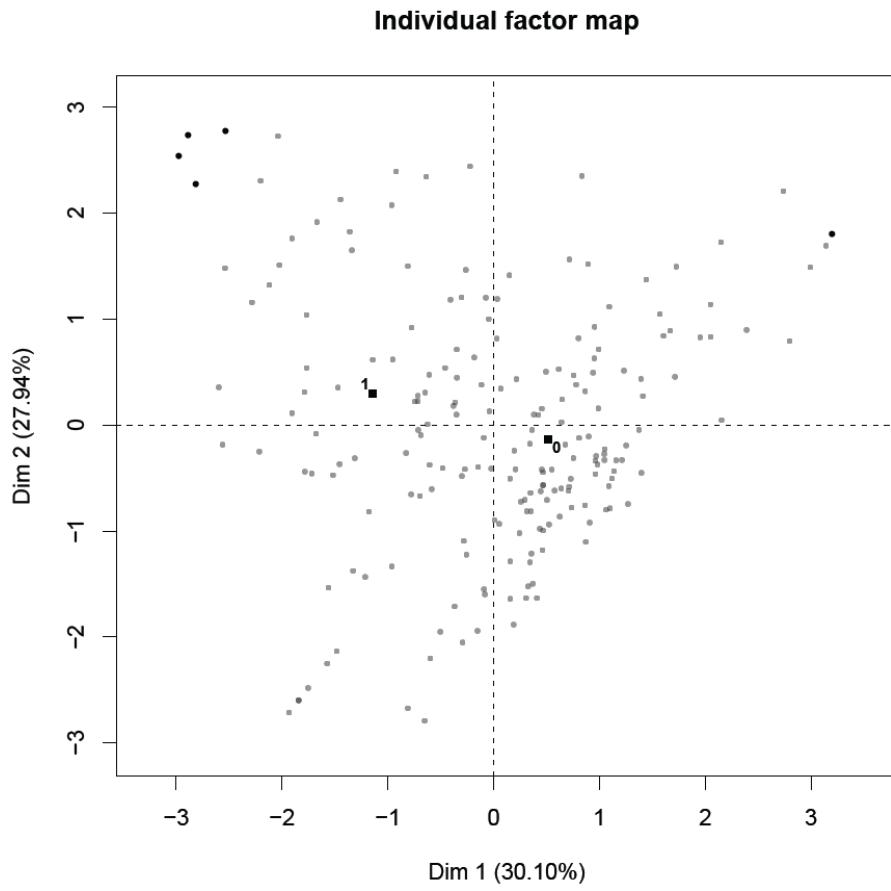
## Messages

```

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

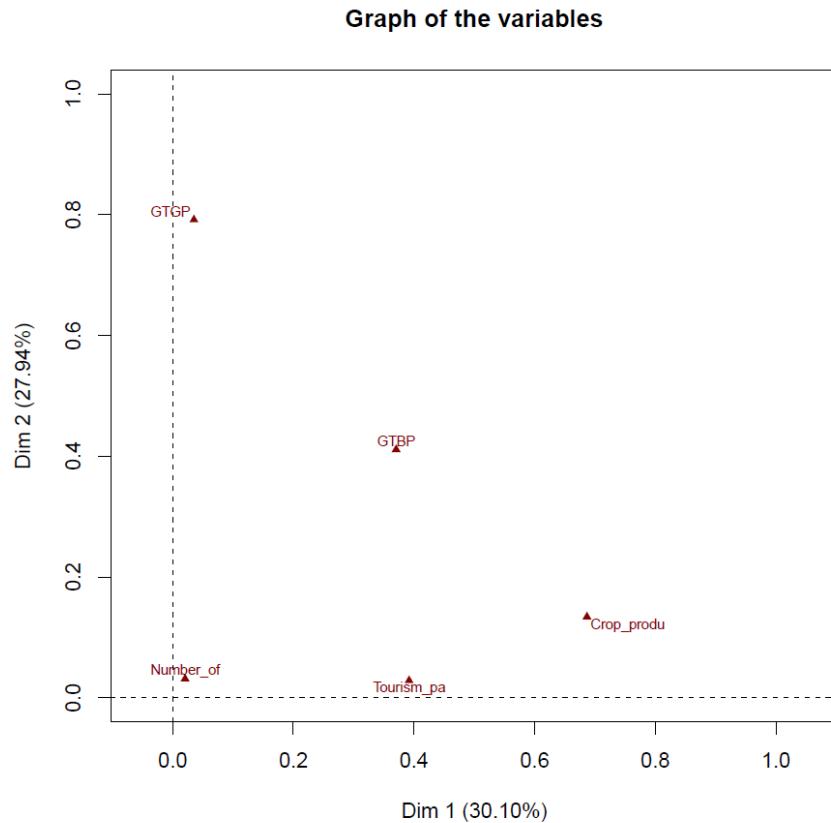
```

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.



---

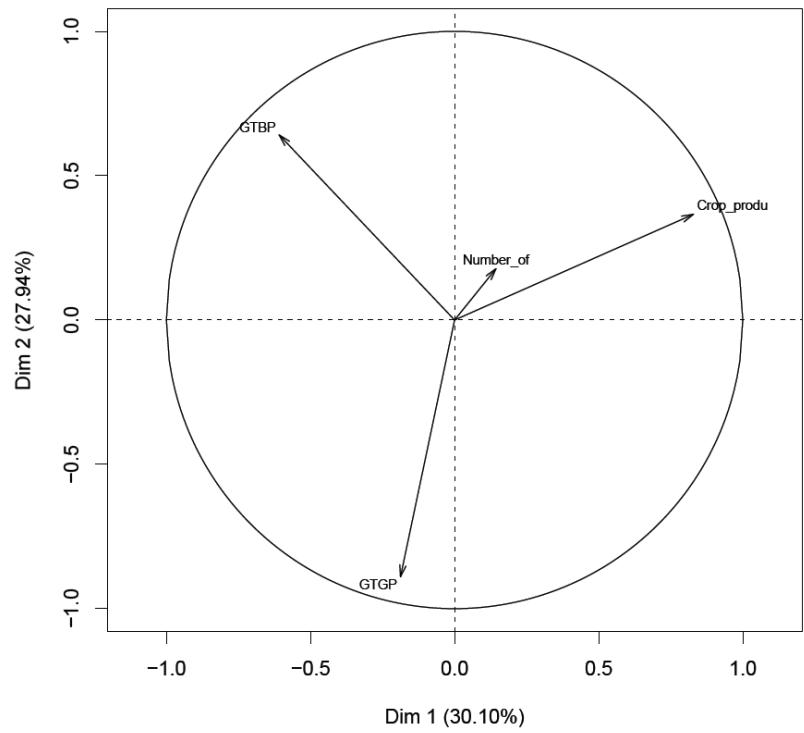
The graph of the variables shows that “GTGP” is strongly associated with the second dimension, while “crop\_produ” and “tourism\_pa” have a medium-to-high association with the first dimension.



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

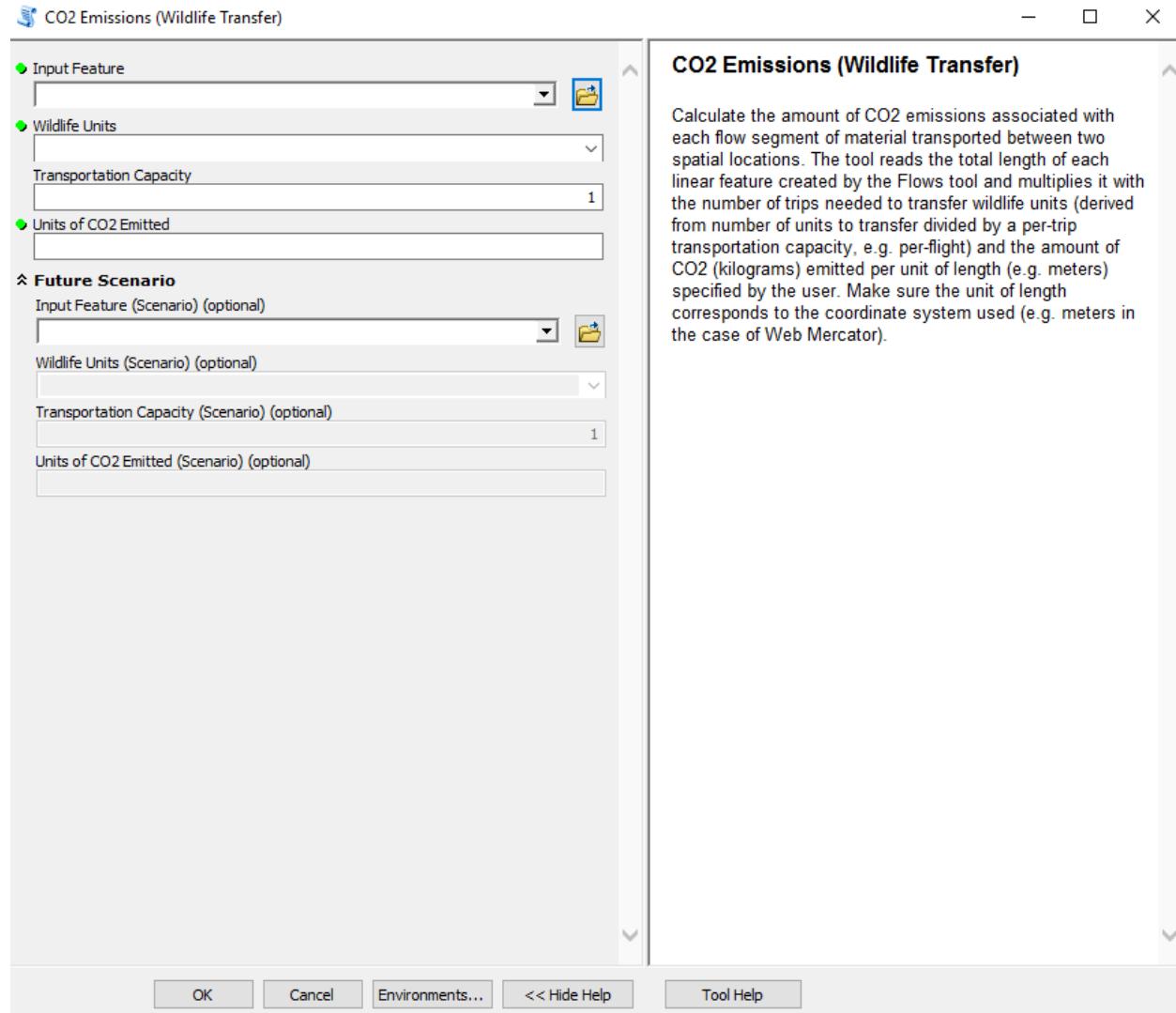
likely that they will have less land dedicated to agriculture and more willing to enroll it into the GTGP.

**Graph of the quantitative variables**



## 4.6 ENVIRONMENTAL ANALYSIS TOOLSET

### 4.6.1 CO<sub>2</sub> Emissions (Wildlife Transfer)



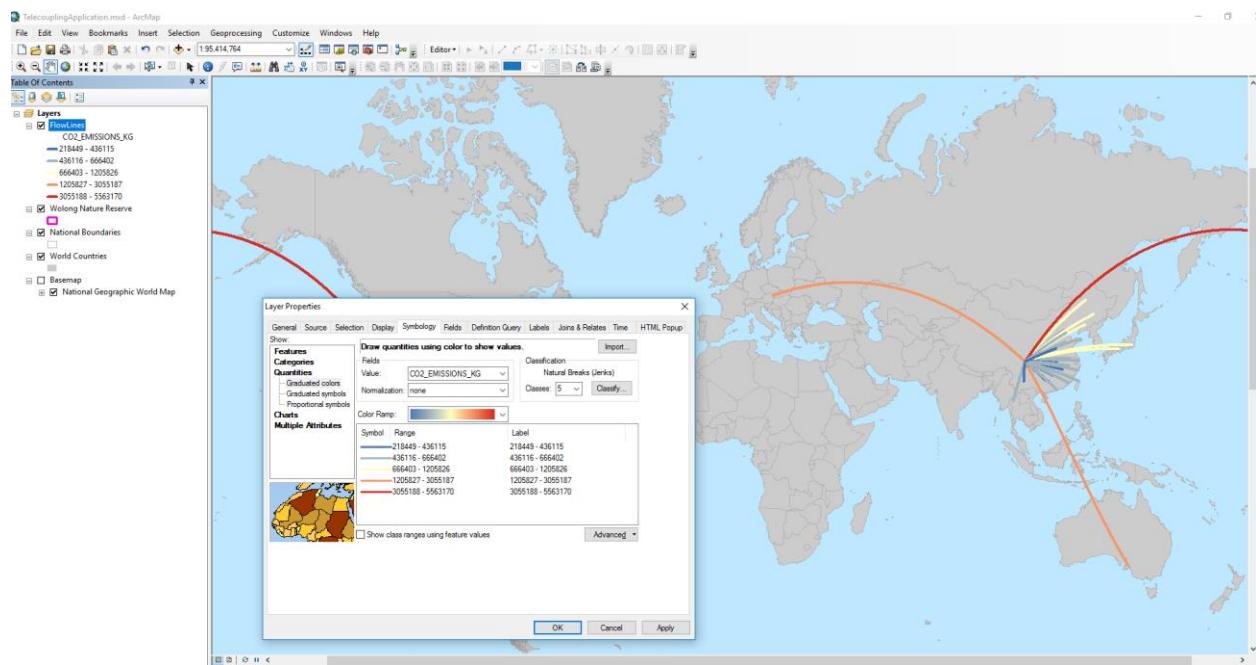
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

This tool uses the flow layer computed with the [Draw Radial Flows](#) tool and calculates the amount of CO<sub>2</sub> emissions based on the length of the flow lines, number of units to transfer,

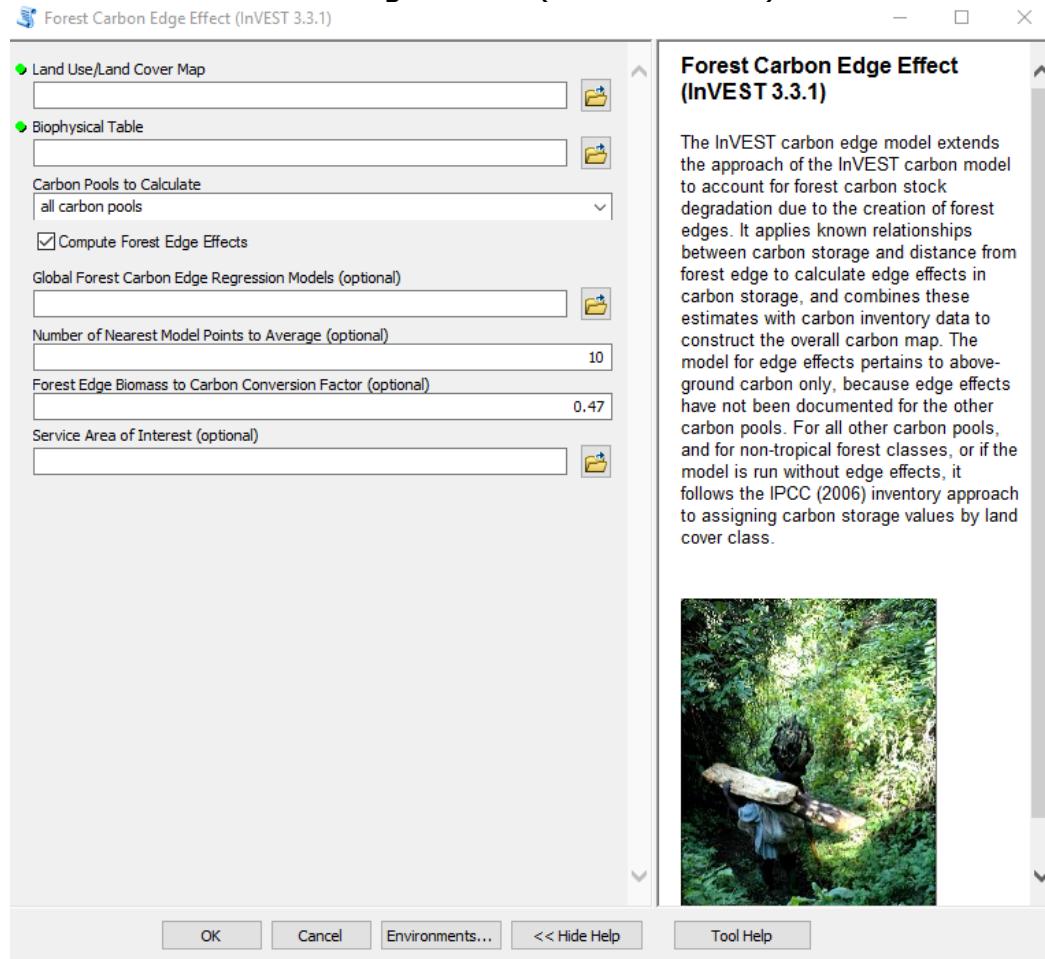
maximum transportation capacity of the medium used, and an estimated amount of CO<sub>2</sub> (kilograms) emitted per unit-length (e.g. meters). For this example, make sure you first computed the radial flows and select that layer as input for the first parameter here. Then, choose a field from the input layer that represents the number of wildlife units that are to be transferred. In the capacity parameter (default = 1) specify the maximum capacity of the transportation medium in terms of units of wildlife. Finally, type in an estimated amount of CO<sub>2</sub> 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<sub>2</sub> emissions.

**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 can now symbolize the flows by the amount of CO<sub>2</sub> emitted in the atmosphere, as shown in the next figure:



## 4.6.2 Forest Carbon Edge Effect (InVEST 3.3.3)

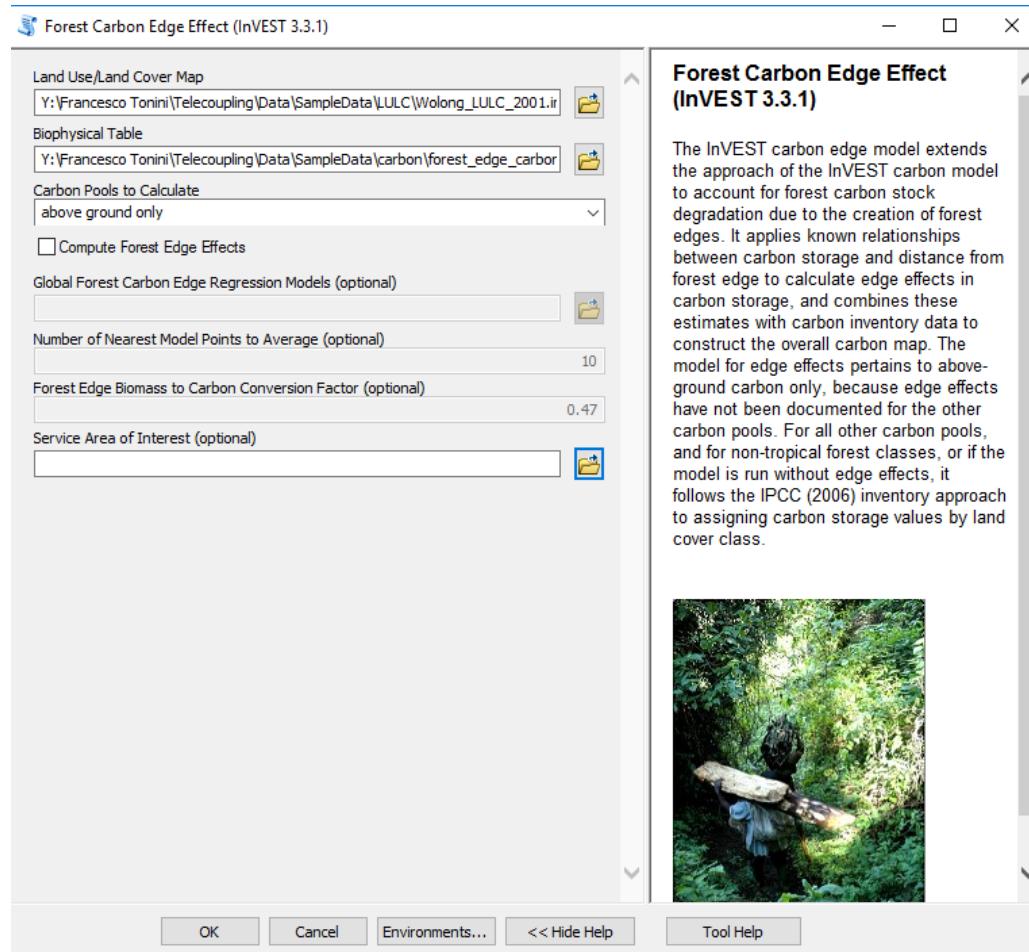


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

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

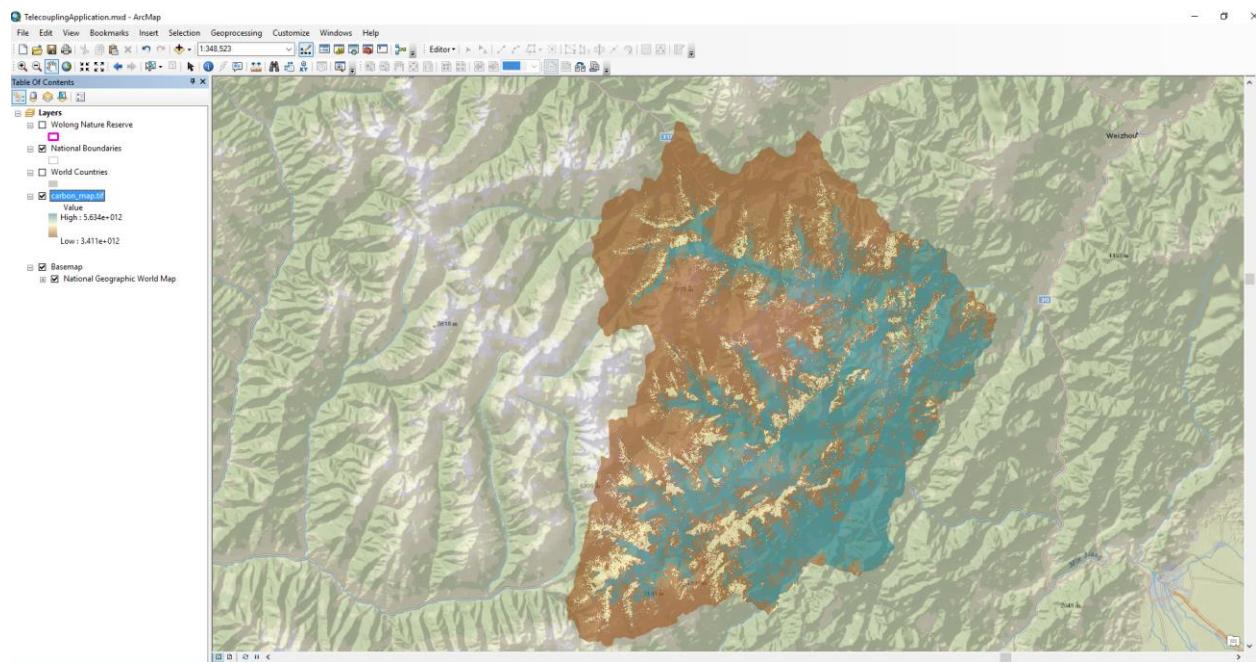
This tools is linked to the InVEST 3.3.3 forest carbon edge effect model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this example, let's use the land use/land-cover raster files from 2001 for the Wolong Nature Reserve area (“`./SampleData/landcover/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 is used by the biophysical model to determine 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.

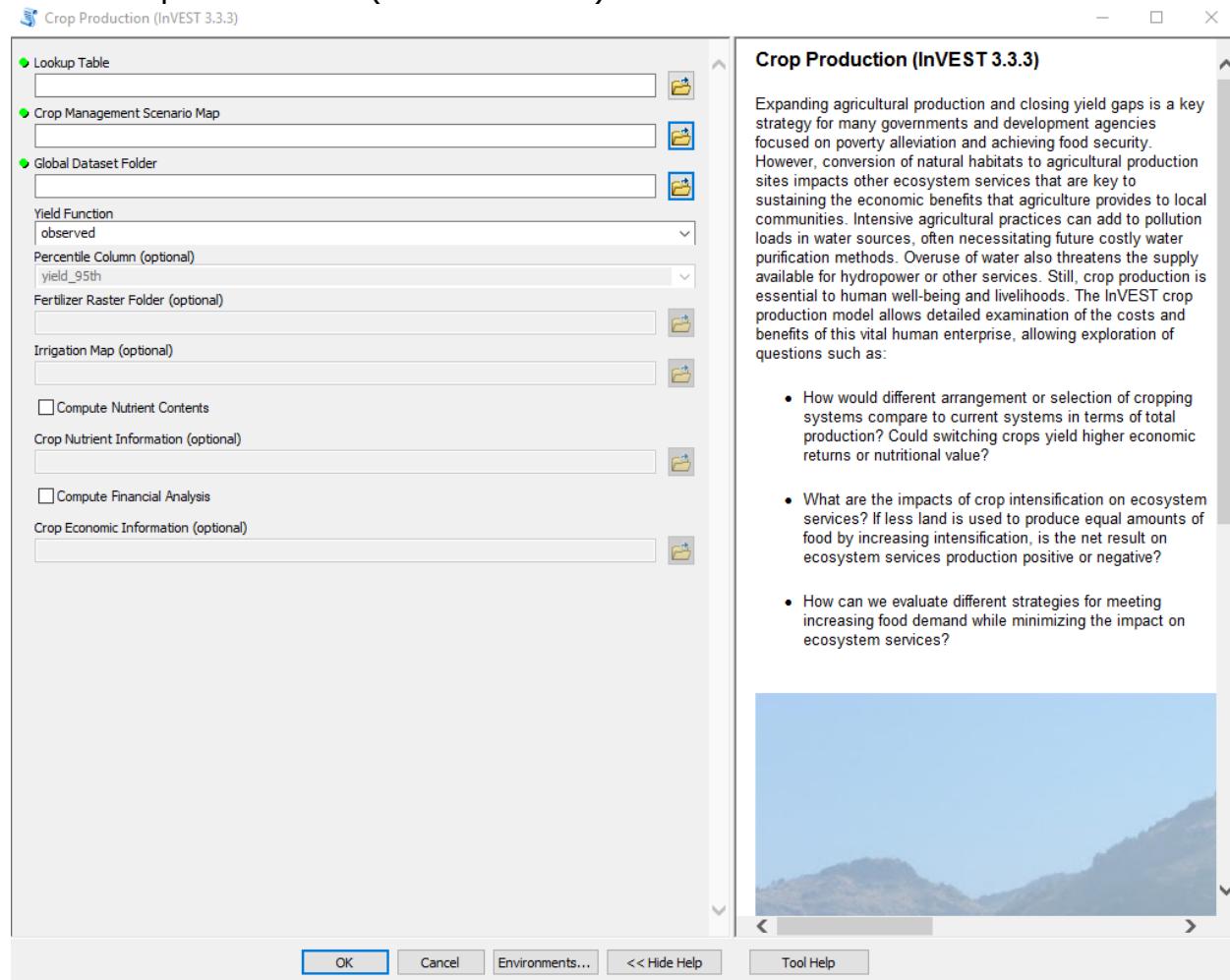


**NOTE:** In order to change the default workspace location where your output files and layers are created, click on the “Environments...” button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.

Click OK to run the tool. Once completed, you should have an output raster with estimated 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:



### 4.6.3 Crop Production (InVEST 3.3.3)



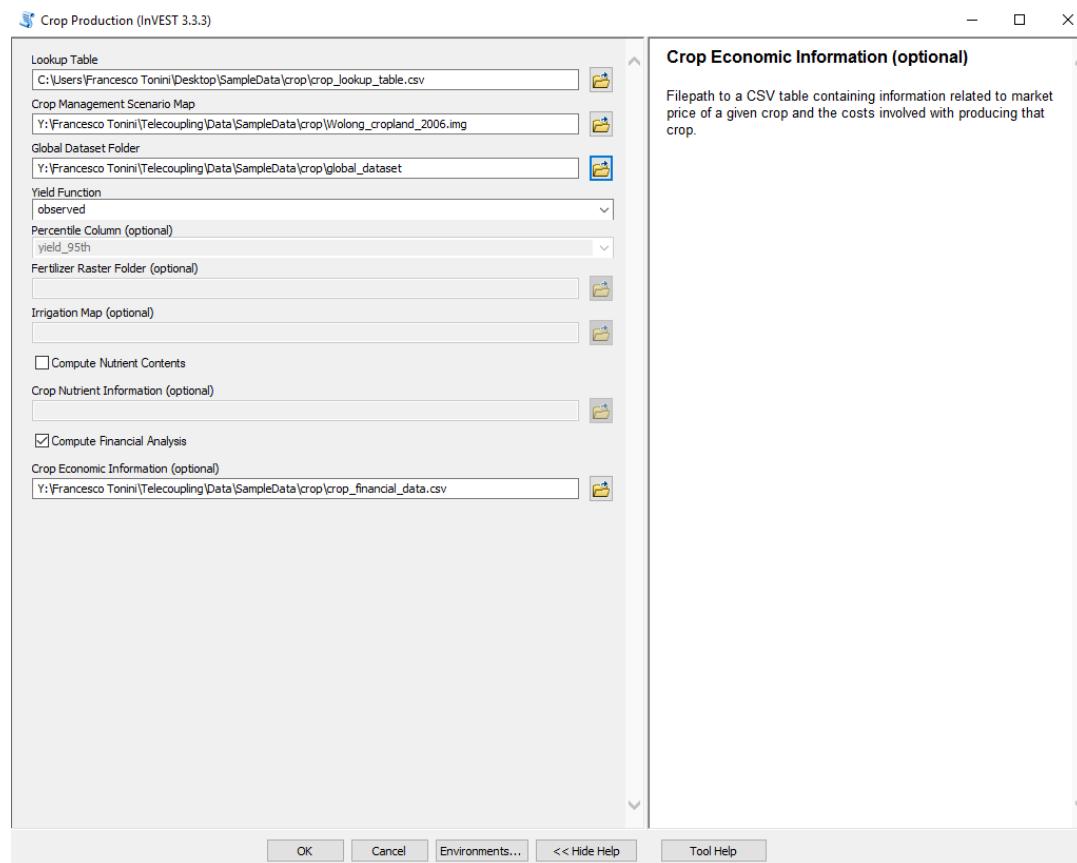
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

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

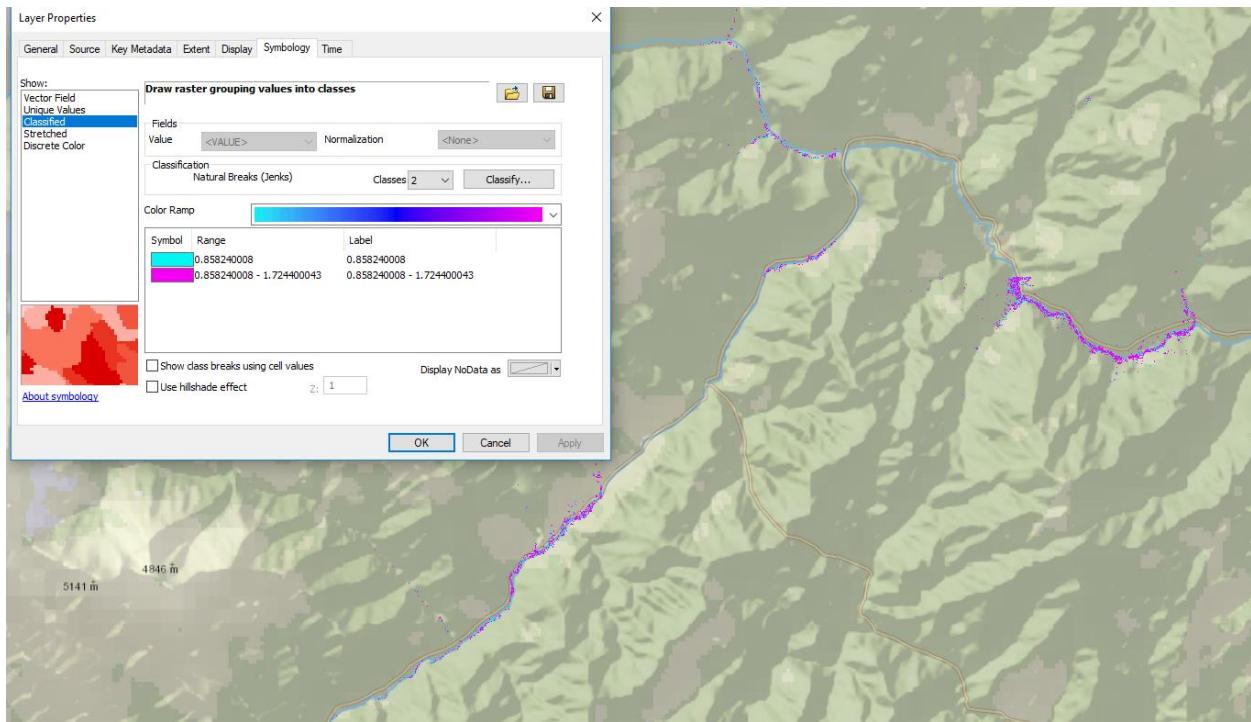
This tool is linked to the InVEST 3.3.1 crop production model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this example, let's use the crop

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).

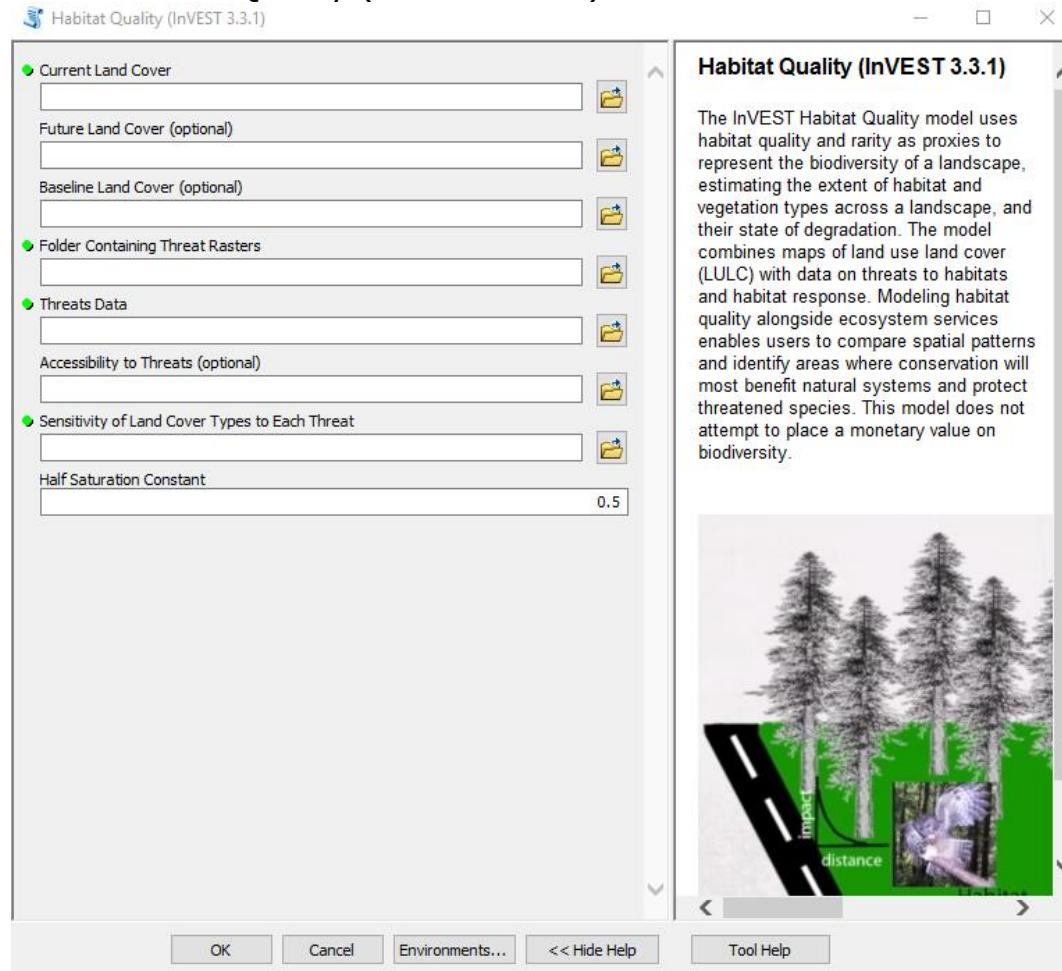
**NOTE:** *In order to change the default workspace location where your output files and layers are created, click on the “Environments...” button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.*



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. Click on OK and you should see something like the following image:



#### 4.6.4 Habitat Quality (InVEST 3.3.3)

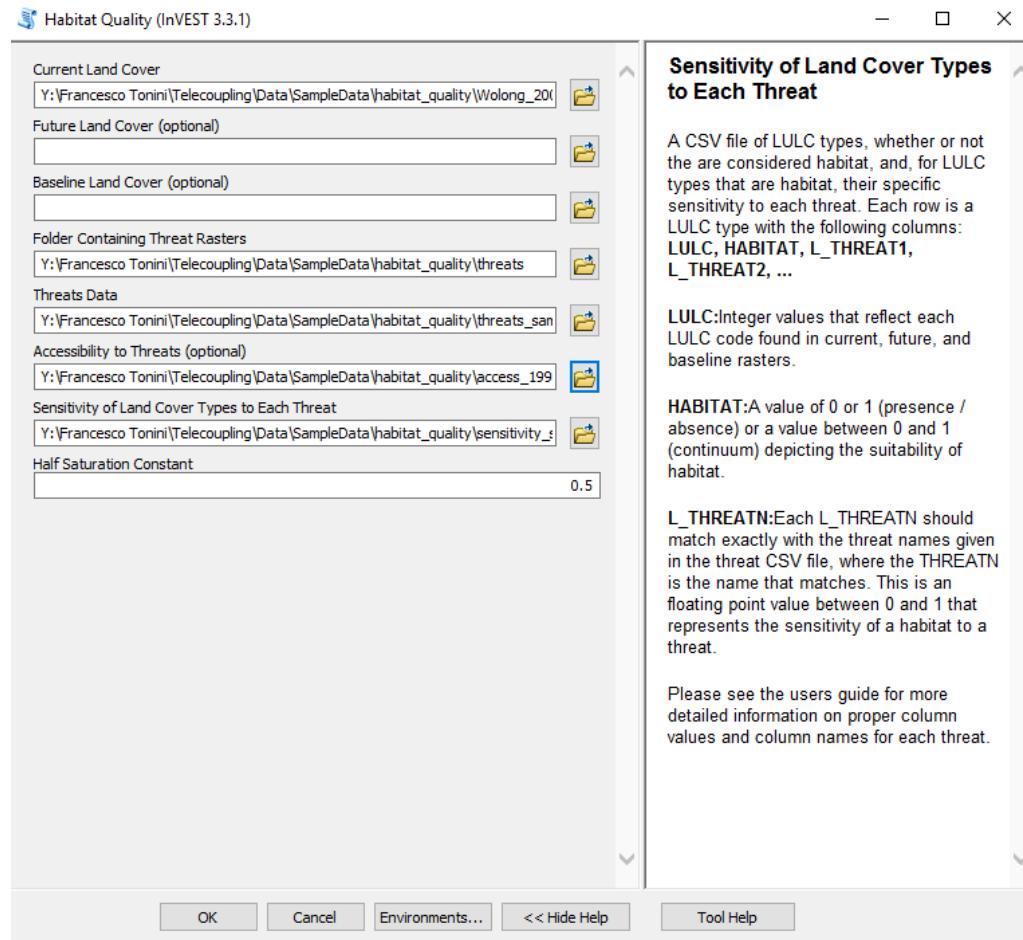


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some cases,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

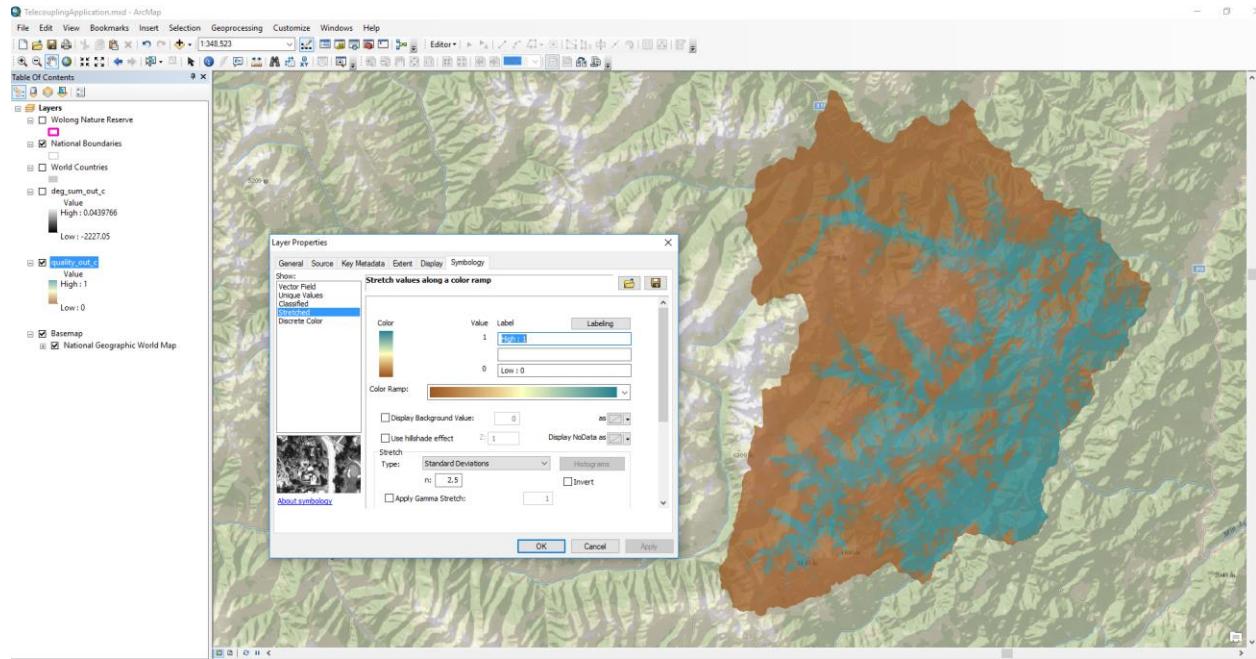
This tool is linked to the InVEST 3.3.3 habitat quality model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this tutorial, we will use the Wolong Nature reserve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

“*./SampleData/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.

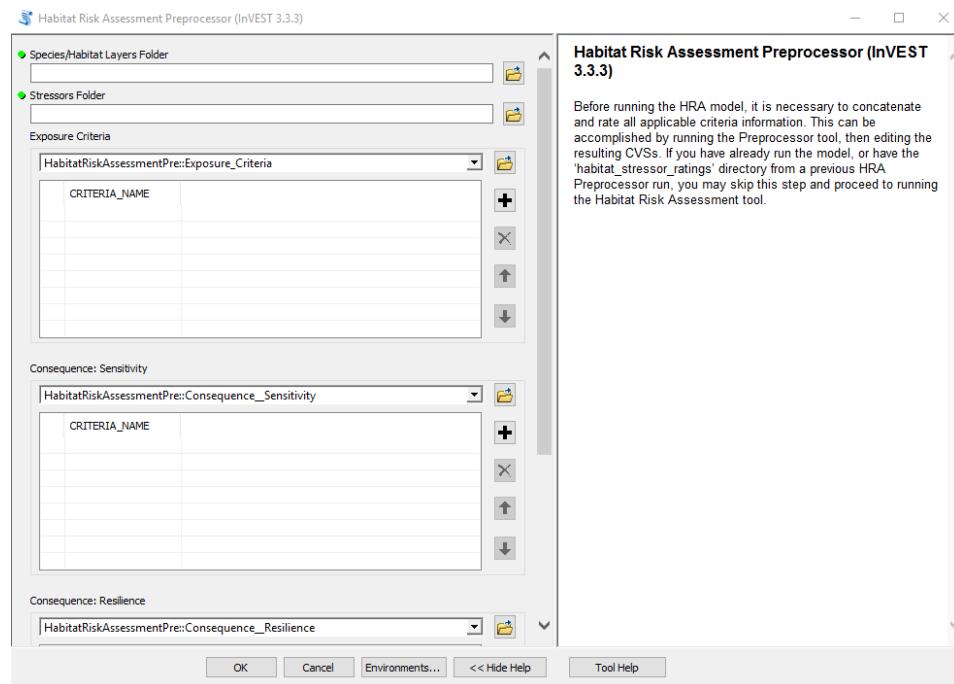
**NOTE:** *In order to change the default workspace location where your output files and layers are created, click on the “Environments...” button at the bottom of the tool interface, then expand the workspace option and set a different path to both the current and scratch folder parameters.*



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):

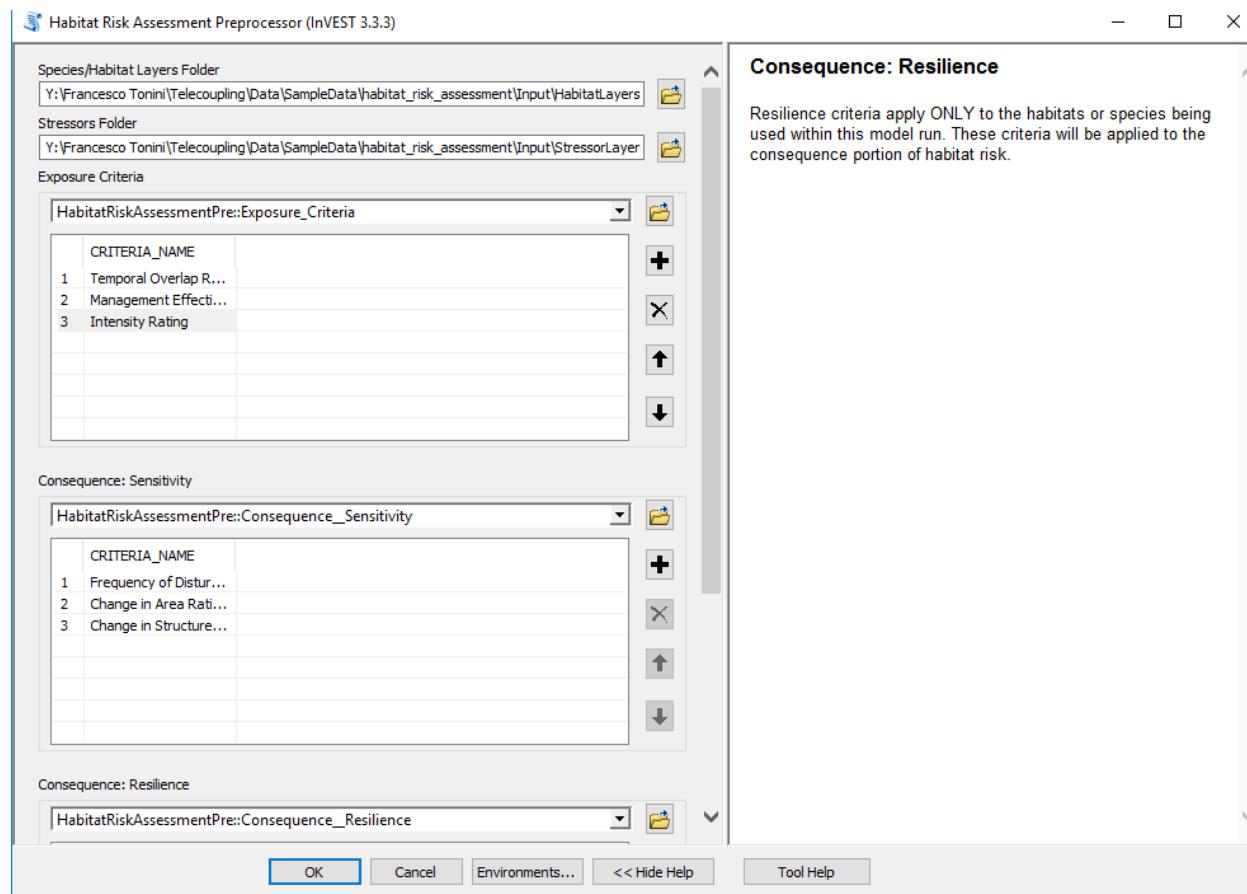


#### 4.6.5 Habitat Risk Assessment Preprocessor (InVEST 3.3.3)



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

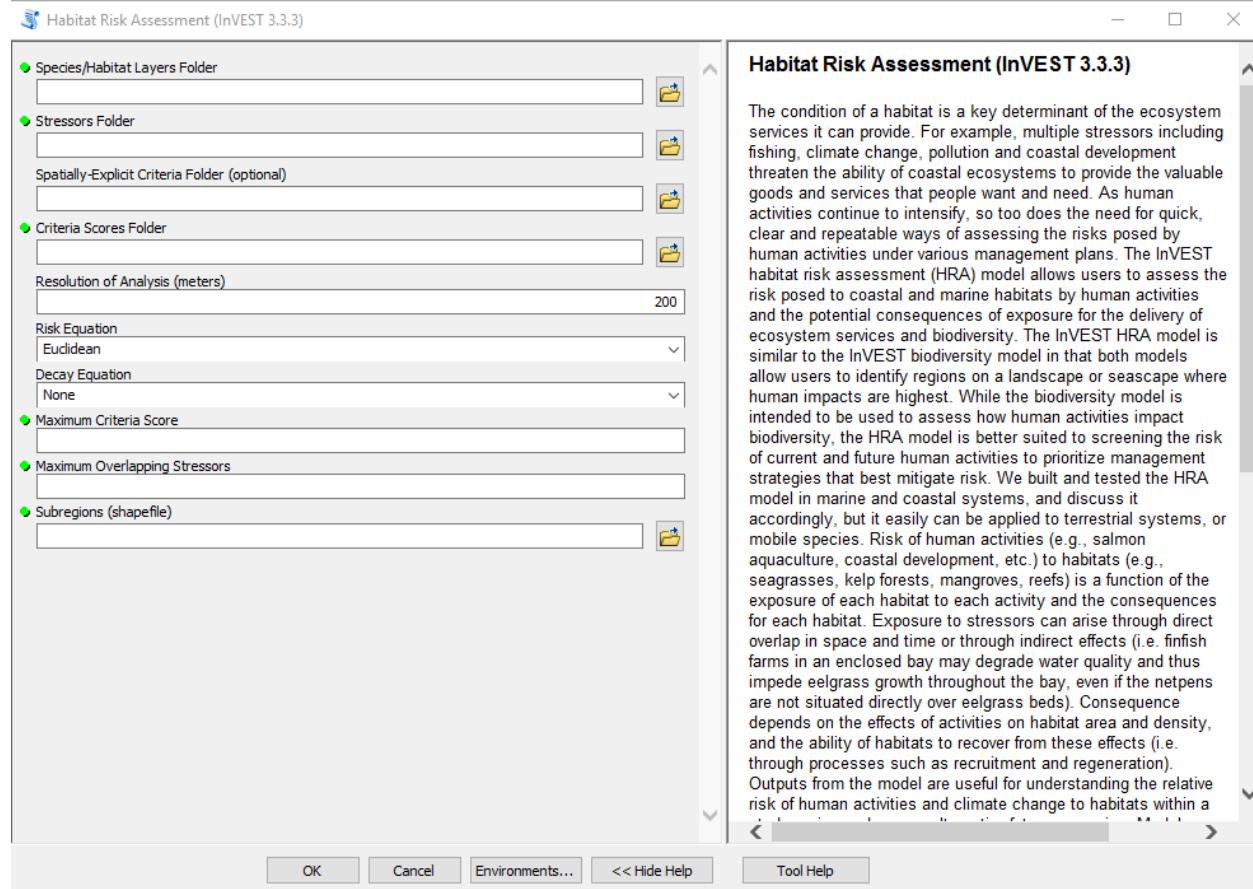
This tools is linked to the InVEST 3.3.3 habitat risk assessment preprocessor model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this tutorial, we will use the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



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

Name	Date modified	Type
dir_names.txt	8/14/2017 3:45 PM	Text Document
eelgrass_ratings.csv	8/14/2017 3:45 PM	Microsoft Excel C...
hardbottom_ratings.csv	8/14/2017 3:45 PM	Microsoft Excel C...
kelp_ratings.csv	8/14/2017 3:45 PM	Microsoft Excel C...
softbottom_ratings.csv	8/14/2017 3:45 PM	Microsoft Excel C...
stressor_buffers.csv	8/14/2017 3:45 PM	Microsoft Excel C...

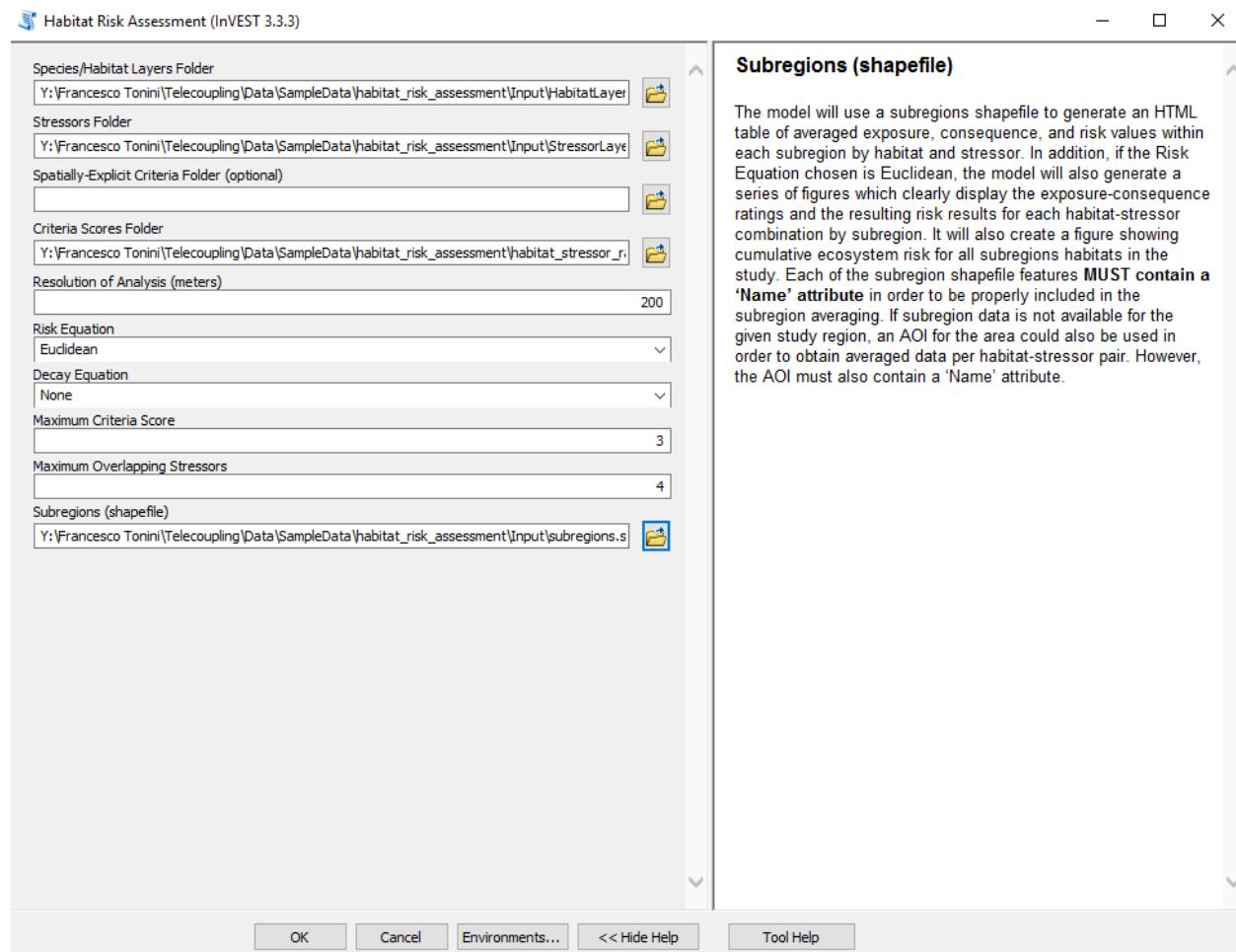
#### 4.6.6 Habitat Risk Assessment (InVEST 3.3.3)



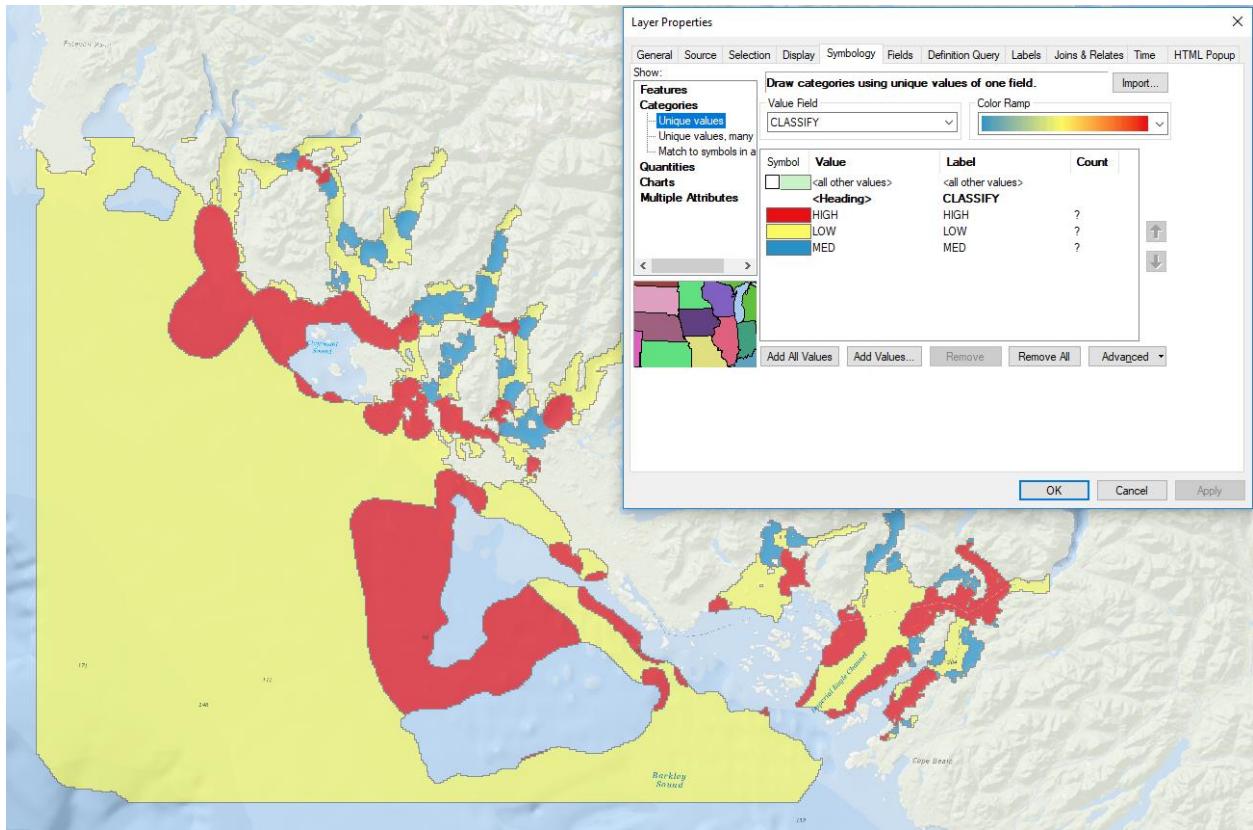
Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

This tools is linked to the InVEST 3.3.3 habitat risk assessment model. All input parameters that are specified in the ArcGIS tool interface, will be sent to the appropriate InVEST model which then return the results produced back to the main application. In this tutorial, we will use the sample data provided by the InVEST development team. This tools/model will generate a series of risk shapefiles for each habitat/species provided as input, as well as a zipfile with a number of HTML plots and tables. This tutorial assumes you either ran the preprocessor tool to produce scoring criteria for each habitat-stressor pair or that you took the sample data provided and customized the criteria and scores on your own to be used in this tool. The species/habitat layer folder parameter needs you to specify a folder on disk containing a shapefile for each habitat/species of interest. Select the folder inside the sample data folder

(*./habitat\_risk\_assessmen/Input/HabitatLayers*) and do the same for the stressors folder parameter, which contains a shapefile for each human stressor affecting one or more of the habitat/species layers specified above (*./habitat\_risk\_assessmen/Input/StressorLayers*). As a criteria scores folder parameter, select the folder that was generated as output in the preprocessor tool. If you have skipped the preprocessor tutorial, then select the folder provided inside the sample data (*./habitat\_risk\_assessment/habitat\_stressor\_ratings\_sample*). Choose the resolution of the analysis that will be used internally by the model to generate raster files for habitat-stressor risks. Pick a resolution that makes sense for your analysis. In this case, we will choose 200 meters. Leave default values for both risk and decay equation parameters (see official InVEST model guide for more details on these parameters and how they affect the final results, [http://data.naturalcapitalproject.org/nightly-build/invest-users-guide/html/habitat\\_risk\\_assessment.html](http://data.naturalcapitalproject.org/nightly-build/invest-users-guide/html/habitat_risk_assessment.html)). For maximum criteria score type the number 3, and for maximum overlapping stressor type the number 4. The maximum criteria score MUST reflect the scoring system used in your criteria scoring input files. Similarly, the maximum overlapping stressor number MUST equal the number of stressor shapefiles chosen for the analysis. Finally, select a shapefile that represent one or more management areas (subregions) for the case study (*./habitat\_risk\_assessment/Input/subregions.shp*).

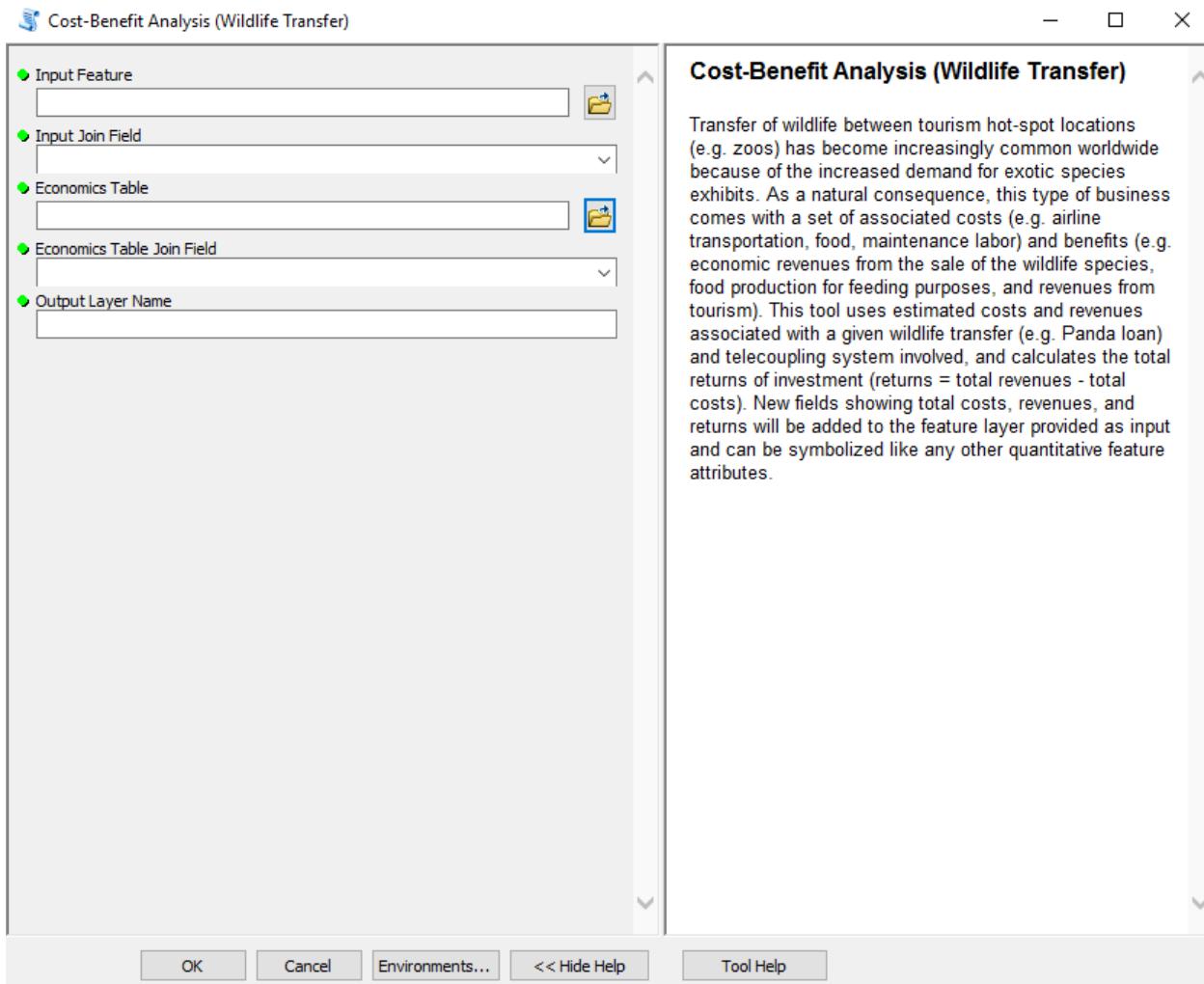


Click OK to run the tool. Once completed, you should have a list of shapefiles with estimated risk for each habitat/species layer provided as input, plus a zipped file with HTML plots and tables. The risk in each shapefile has been classified into LOW, MEDIUM, HIGH internally by the model. The following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



## 4.7 SOCIOECONOMIC ANALYSIS TOOLSET

### 4.7.1 Cost-Benefit Analysis (Wildlife Transfer)



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

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

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

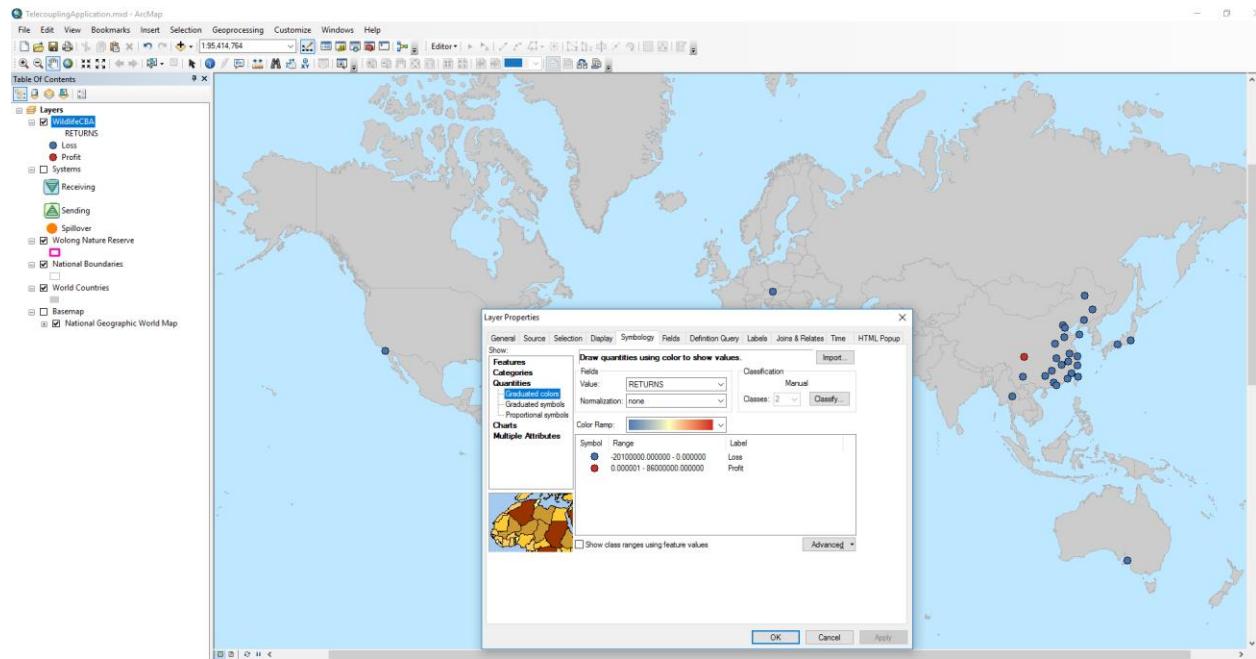
**different number of records between the two tables, the join operation should still complete successfully with a warning message.** Specify the economics table provided with the sample dataset above and use NAME as a join field. You do NOT need to have the same column labels (NAME) like in this case,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

**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.

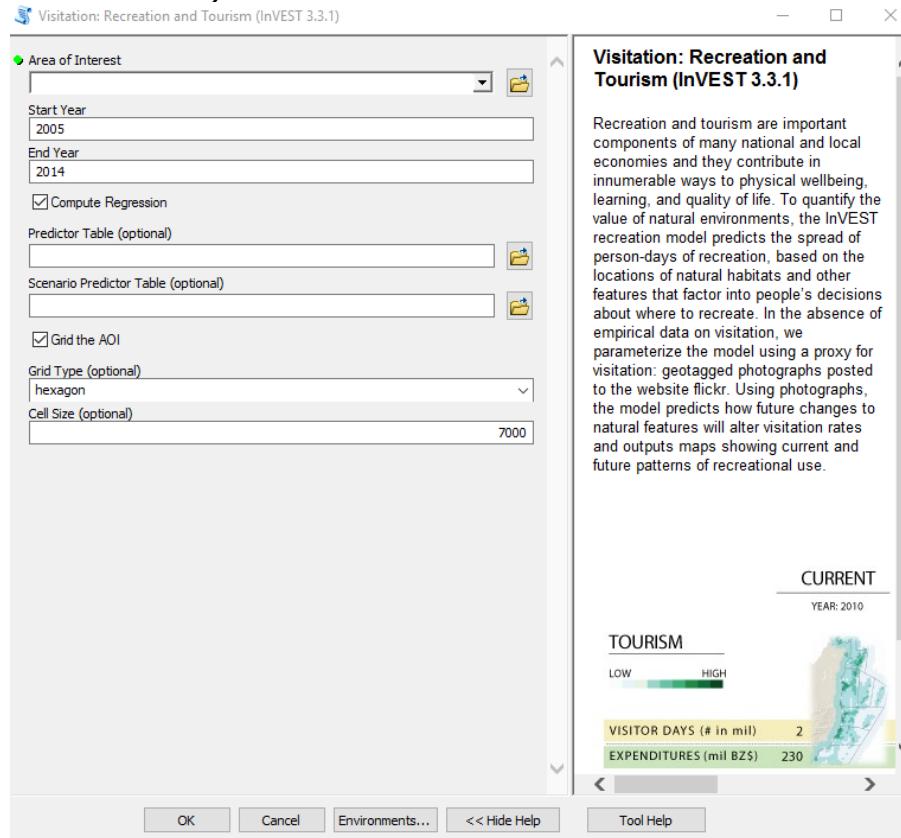
Click OK to run the tool and ArcGIS will return a new feature class in output with three new attributes specifying COSTS, REVENUES, and RETURNS.

	Shape *	NAME	Role	LON	LAT	POINT_X	POINT_Y	COSTS	REVENUES	RETURNS
▶	Point	Wolong	Sending	103.058077	31.001193	11472372.6538	3632904.0781	0	86000000	86000000
▶	Point	Adelaide	Receiving	138.601	-34.929	15428992.7434	-4154236.7088	1210000	0	-12100000
▶	Point	Beijing	Receiving	116.39	39.91	12956475.5334	4852872.331	1410000	0	-14100000
▶	Point	Benxi	Receiving	123.766667	41.3	13777642.3476	5056692.8088	1110000	0	-11100000
▶	Point	Chiang Mai	Receiving	98.9986	18.795	11020473.7412	2130815.2905	1310000	0	-13100000
▶	Point	Fuzhou	Receiving	119.31	26.08	13281528.4465	3008992.6645	1510000	0	-15100000
▶	Point	Guangzhou	Receiving	113.26	23.13	12608045.5272	2647747.5275	2010000	0	-20100000
▶	Point	Guilin	Receiving	110.29	25.27	12277426.6396	2908944.6543	1310000	0	-13100000
▶	Point	Hengyang	Receiving	112.6	26.9	12534574.6633	3110983.6162	1110000	0	-11100000
▶	Point	Hongkong	Receiving	114.159	22.278	12708121.7495	2544935.3432	1410000	0	-14100000
▶	Point	Jilin	Receiving	126.57	43.87	14089707.9497	5445346.3744	1110000	0	-11100000
▶	Point	Jinan	Receiving	116.98	36.67	13022154.033	4393208.3209	1110000	0	-11100000
▶	Point	Kobe	Receiving	135.2	34.68	15050395.1553	4120479.0079	1210000	0	-12100000
▶	Point	Kunming	Receiving	102.705	25.04	11433068.3019	2880658.5234	1310000	0	-13100000
▶	Point	Nanchang	Receiving	115.00	29.69	12000700.5024	2234070.0006	1110000	0	-11100000

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



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

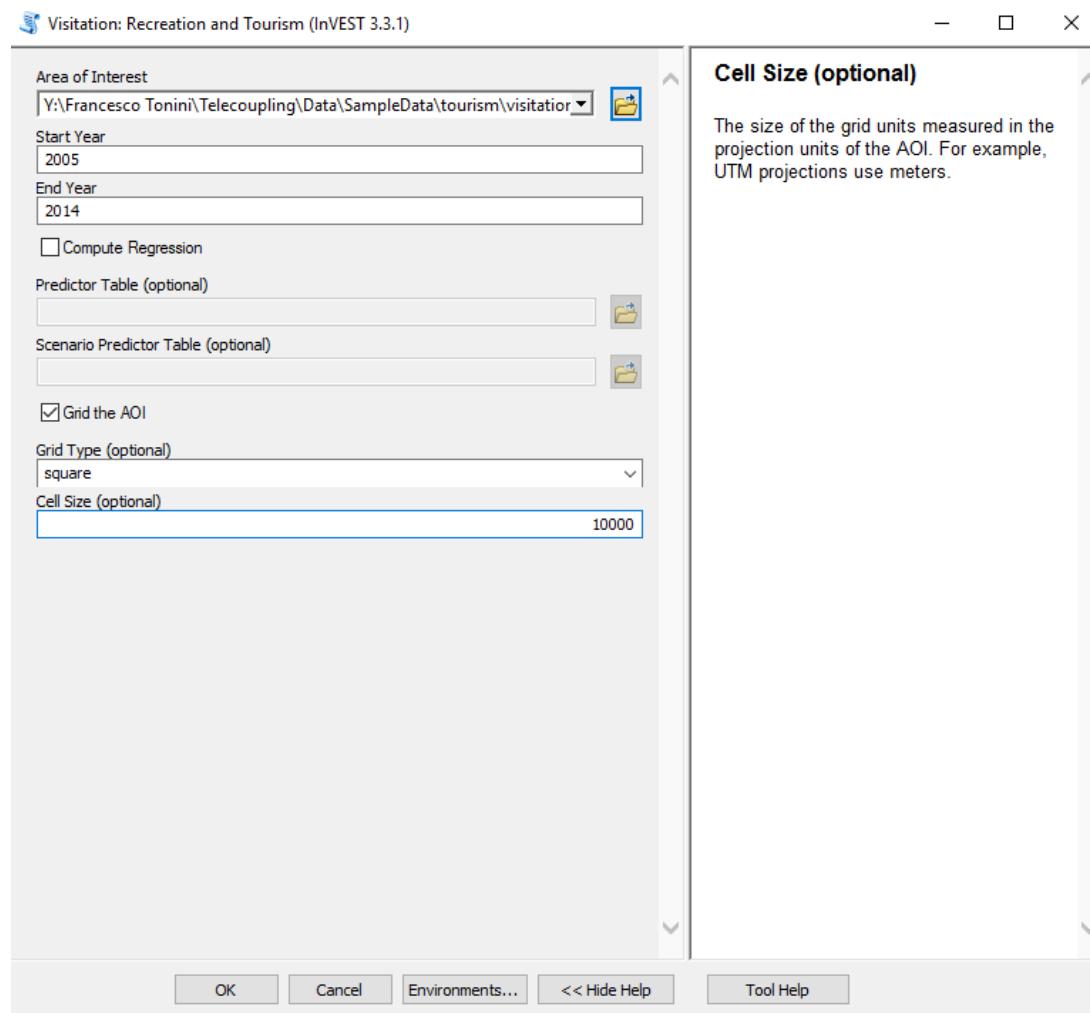


Once you open the tool (by double-clicking on it) make sure the help window stays open (bottom of the tool interface to hide/open it). The help is there to provide general information about the tool as a whole, as well as describing in detail each single parameter after clicking on it. In some case,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

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

As area of interest select the shapefile of the Wolong Nature reserve boundaries (“./SampleData/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:

