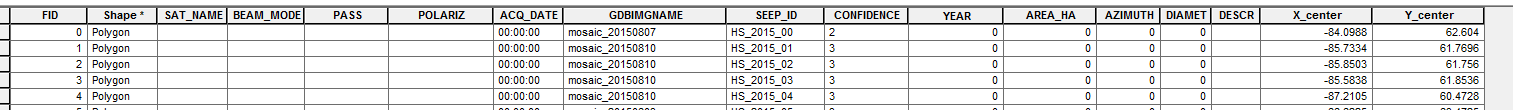
User’s Manual – GEM1 to GEM2 Oil Seep Detection and Analysis Integration

This document explains the use of the “GEM1 to GEM2 Toolbox” python toolbox in the context of spatial-temporal analysis of dark targets. These dark targets are identified as possible natural oil seeps by RADARSAT-2 imagery. An explanation on how to install and use the GEM2 Toolbar developed for the visualisation process stage.

### Workflow Requirements

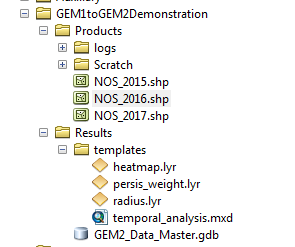
1. In order to use this toolbox and analysis process, Step 1 through Step 5 of the GEM1 dark features extraction process must be complete with one master shapefile for the year of analysis (e.g. NOS2017.shp). The GEM1 toolbox can be found in the GEM\_z820 drive at: [\\w-bsc-a0235555\Code\GEM1-toolbox](file:///\\w-bsc-a0235555\Code\GEM1-toolbox). The NOS shapefile must have the following structure in order to properly integrate with GEM2:



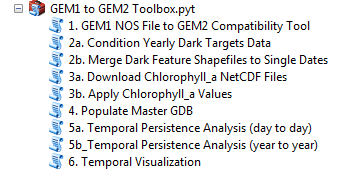
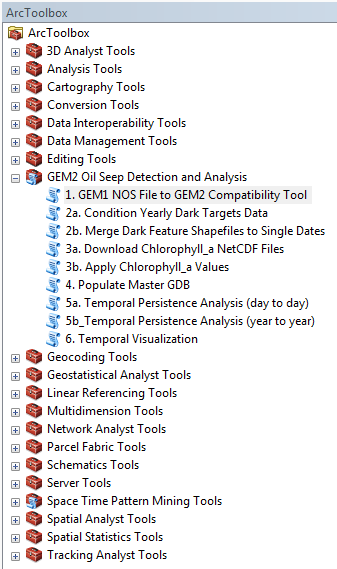
*GDBIMGNAME*: the field must be populated with the following structure: ‘mosaic\_YYYYMMDD’ with YYYYMMDD signifying the day of acquisition of the image being interpreted.

*SEEP\_ID*: the field must have the following structure: ‘XX\_YYYY\_##’. XX can be any two characters chosen for the region, YYYY is the year of analysis, and ## is a unique identifier number ,starting at 00, for each particular dark feature.

1. The “**Results**” folder containing the map document, layer and master geodatabase templates, and the “joinAttrFromCSV” python script, must be located in the same directory as the “**Products**” folder. As of November 2017, the **“Results\_template”** folder is found in the GEM\_z820 drive at: [\\w-bsc-a0235555\Code\GEM2-Oil Seep Detection and Analysis\Results\_Template](file:///\\w-bsc-a0235555\Code\GEM2-Oil%20Seep%20Detection%20and%20Analysis\Results_Template).



1. The “GEM1 to GEM2 Toolbox” python toolbox and its various tools can be launched either from the Catalog window in ArcMap, or can be added as a toolbox in the ArcToolbox window as seen in the following figures.

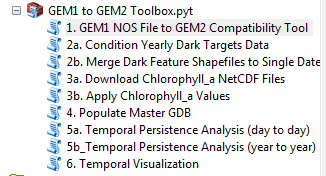
1. The various tools often overwrite files and layers during geoprocessing. **Overwrite** must be enabled in the Geoprocessing options in ArcMap in order to prevent overwrite issues.
2. Some geoprocesses in this toolbox require the **Spatial Analyst extension** to be enabled in order to execute.

## GEM1 to GEM2 Toolbox.pyt

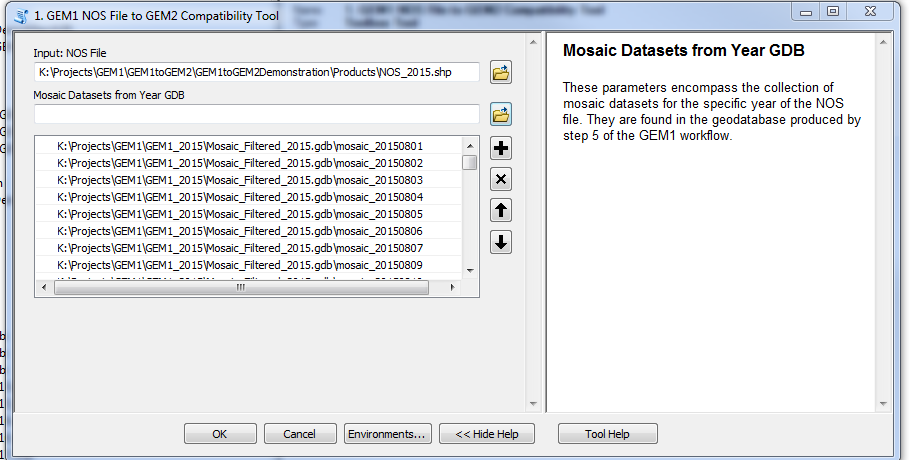
### Part 1: Modify and condition GEM 1 visualisation to GEM2 standards

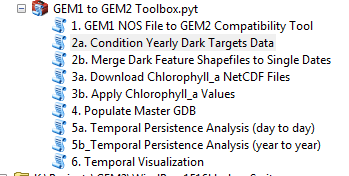
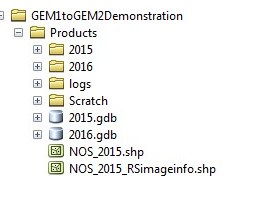
The first step of the analysis is to condition the dark targets data in order to clean up the input shapefiles for the spatial-temporal analysis and convert them to feature classes organized as single day acquisitions in separate yearly dark targets geodatabases.

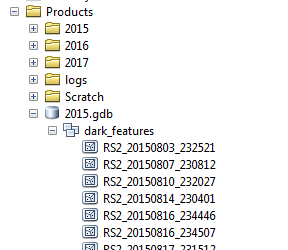
1. Launch ‘1. GEM1 NOS File to GEM2 Compatibility Tool’ tool script from the GEM1 to GEM2 Toolbox.



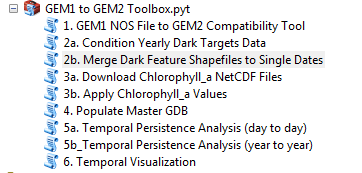
1. Add the GEM1 visual interpretation results shapefile (e.g. NOS\_YYYY) to the Input: NOS File parameter (see following figure). Also, add all mosaic datasets from the GEM1 Step 5 output geodatabase.



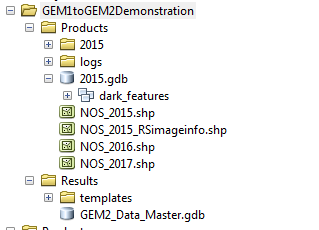
1. Launch “2a. Condition Yearly Dark Targets Data” tool script.
2. Enter the NOS file that’s been modified by ‘1. GEM1 NOS File to GEM2 Compatibility Tool’ (e.g. NOS\_2015RSimageinfo.shp’ in the following figure).
3. Click “OK” to run the tool. Once complete, the shapefiles in the year folder (e.g. “2015”) will be populated to GEM2 annotation and conditioned and organized in a new geodatabase of the same name (e.g. “2015.gdb”), placed in the same folder. These features are not conditioned by date, but rather by image.



1. Launch ‘2b. Merge Dark Feature Shapefiles to Single Dates’ tool script.



1. Select the ‘dark\_features’ dataset within the appropriate year of analysis (e.g. 2015/dark\_features’. This will merge the dark feature by swath to dark features by day.

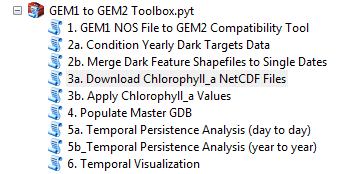


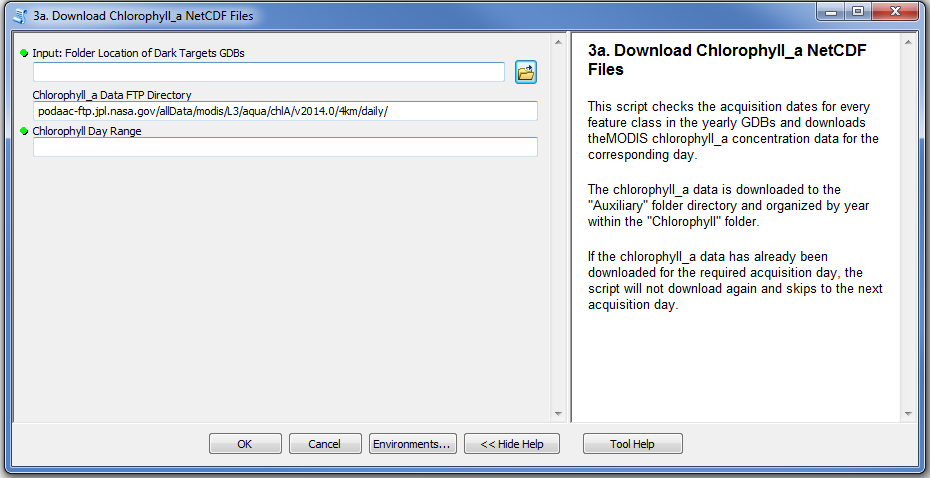
1. Run the tool again (steps 1 to 7) for every year (NOS file) to be analysed in the “Products” folder. The previous steps were designed to prepare the visual interpretation results to be interpreted for chlorophyll and persistence with the same methods as the GEM2 automated outputs. The data has now been conditioned and is ready for subsequent analysis.

### Part 2: Download Chlorophyll\_a NetCDF Files and Apply Chlorophyll\_a Values

The next sub steps in data conditioning are to acquire and apply attribute values pertaining to the concentration of chlorophyll\_a in which the dark targets are located. This is available as part of the daily acquisitions of the MODIS Aqua sensor platform. The concentration values are measured in mg/m-3. The chlorophyll data is smoothed with a 5x5 pixel moving window, unless otherwise specified. Additionally, the chlorophyll data can be analysis for a number of days before and after the dark feature is present on an image if chlorophyll is not located on the day of acquisition.

1. Launch “3a. Download Chlorophyll\_a NetCDF Files” tool script.

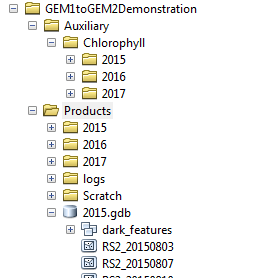




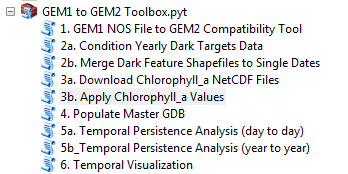
1. Enter the location of the “Products” folder in the first parameter.

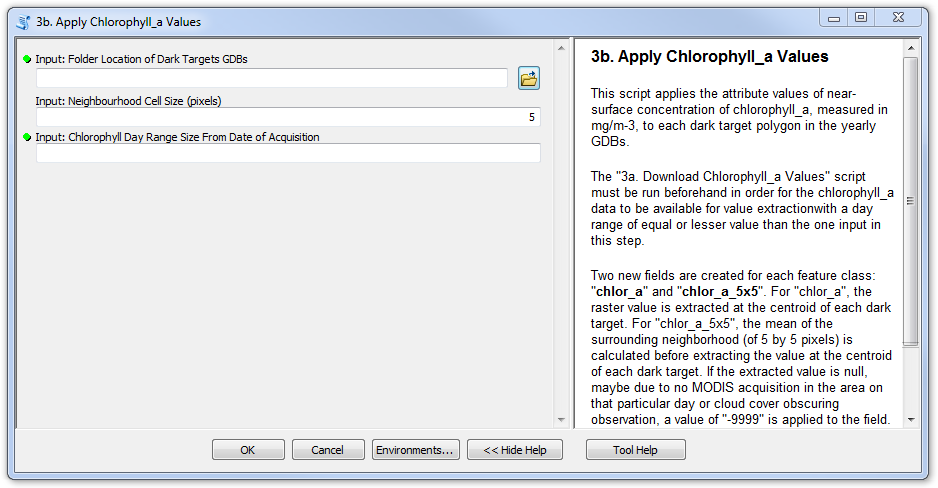


1. The second parameter contains the ftp site address where the chlorophyll\_a data is located and available for download. The default value will suffice until such time that the ftp site is updated or removed.
2. The third parameter is the range of days before/after the day acquisition that will be analysed for an intersection between chlorophyll and the dark features.
3. Click “OK” to run the tool. The script will iterate through all the geodatabases located in the “Products” folder and download the same day +/- the day range of MODIS acquisition corresponding to the acquisition days of the dark targets feature classes. An “Auxiliary” folder is created in the same directory as the “Products” folder, where a “Chlorophyll” folder is created containing the chlorophyll\_a rasters organized by year.



1. Launch “3b. Apply Chlorophyll\_a Values”





1. Enter the location of the “Products” folder in the first parameter.



1. The second parameter holds the value of the neighbourhood size that will be used with which to determine the mean pixel value for a focal statistics calculation. In most cases, the default value of 5 (for a 5x5 pixel neighbourhood window) will suffice.
2. The third parameter can be equal or lesser than the day range selected in the previous tool script.
3. Click “OK” to run the tool. The script will extract the neightbourhood mean statistic over value over each dark targets centroid. This creates two new attributes for each feature class, “chloro\_dayRange” which is the amount of days +/- the day of acquisition where there is a chlorophyll reading, and “chlor\_a\_5x5” for the mean value of the pixel neighbourhood at the specified size. When “chloro\_dayRange is negative, it means that the date of chlorophyll is before the day of analysis. If the “chloro\_dayRange” is positive it is after the day of acquisition.



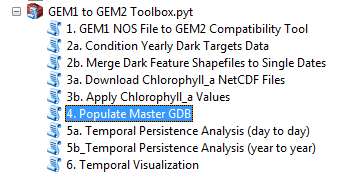
These tools (3a and 3b) can be run again in order to download and update any new feature classes that have been created by the “2a. Condition Yearly Dark Targets Data” tool. The tools will iterate over the entirety of the “Products” folder every time they are executed, however if a raster file is detected inside the “Auxiliary/Chlorophyll” folder, it will not be downloaded again. As well, if the “chlor\_a” field is detected as already created inside a feature class, the tool will assume that the value extractions have already been performed and will skip to the next feature class. This allows for data updates with new acquisition feature classes without needed to recalculate the entirety of the data collection.

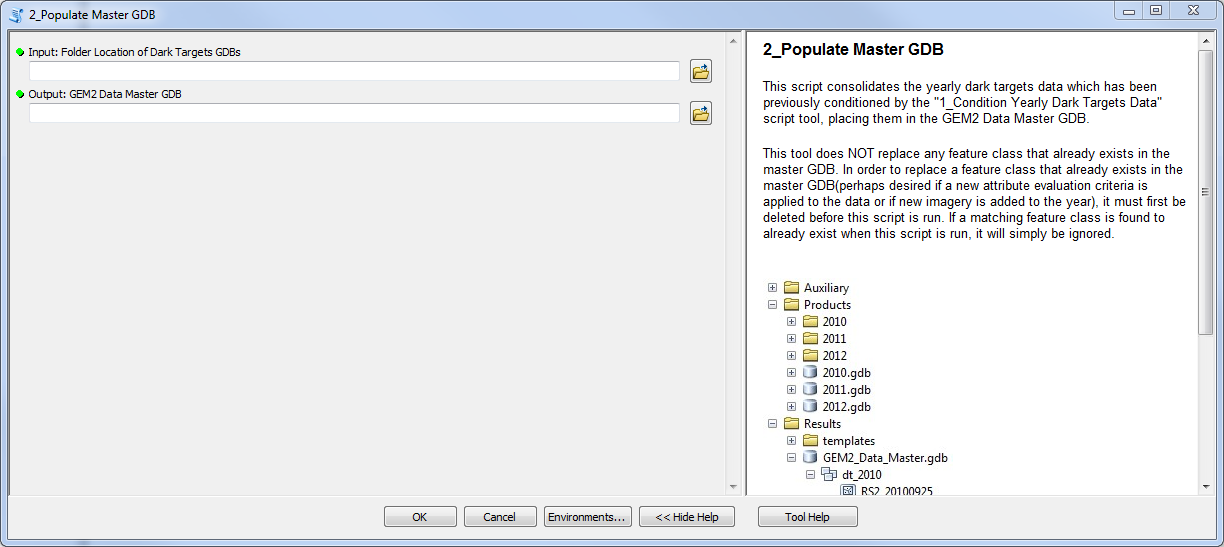
### Step 3: Populate Master GDB

*Note: The script tools for the following steps were slightly modified from the original GEM2 workflow written by David Hennessy; however, the changes were minor and do not affect the output or structure of GEM2 temporal analysis.*

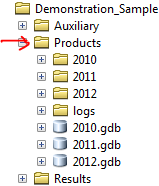
The next step consists of consolidating the conditioned data, which is currently organized in separate yearly geodatabases. This tool will organize the entire collection of data into a single master geodatabase in order to perform the subsequent spatial-temporal analysis.

1. Launch “4\_Populate Master GDB” tool script.

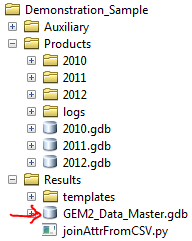




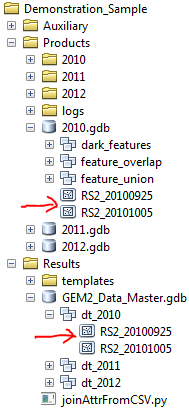
1. Enter the location of the folder in which the geodatabases are located in the first parameter. In most cases this will be the “Products” folder, however may be another folder in case a data update of future acquisitions was conditioned separately.



1. Enter the location of the “GEM 2 Data Master GDB”. In all cases, this geodatabase will be located in the “Results” folder, however the location must still be specified in case the first parameter (“Products” folder) happens to be located in a separate folder hierarchy altogether.



1. Click “OK” to run the tool. The script will iterate through all the geodatabases in the “Products” folder and copy every feature class from its yearly geodatabase to a corresponding feature dataset in the master geodatabase.

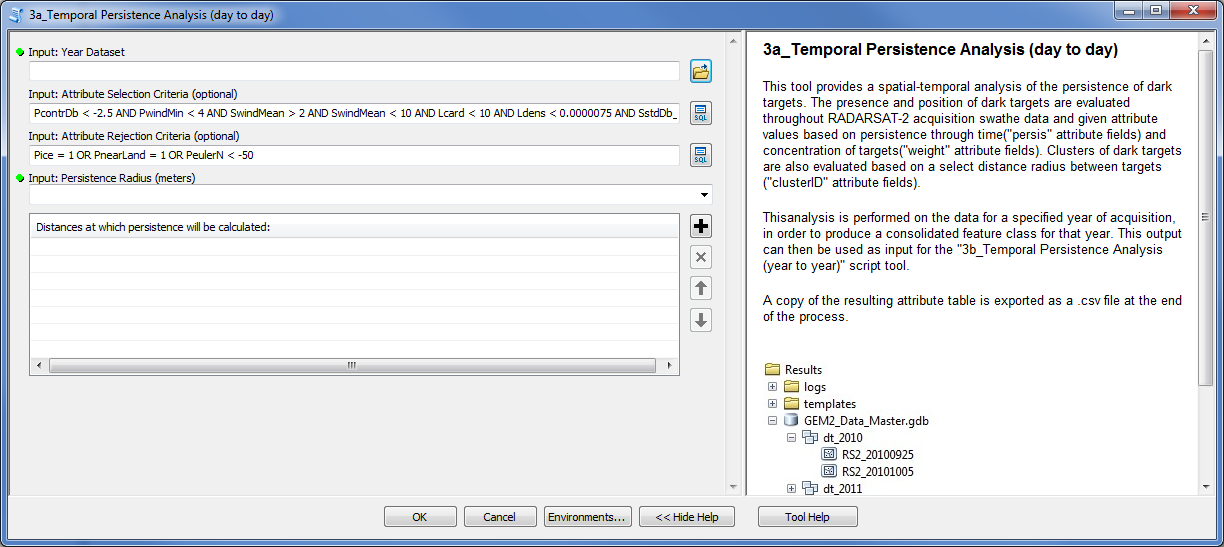


This tool can be executed every time new data becomes available and needs to be incorporated into the master geodatabase. The script will verify if a feature class already exists in the master geodatabase before copying it. If the feature class already exists, the script will not copy and skip to the next one. This preserves the integrity of the master geodatabase, and any feature class that needs to be overwritten will need to be purposefully deleted before running this tool to update it.

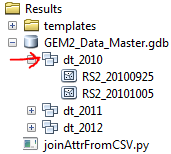
### Step 4: Temporal Persistence Analysis (day to day)

With the data properly conditioned and organized by acquisition day in the master geodatabase, the first part of the spatial-temporal analysis can be performed. This tool takes every feature class (acquisition day) and performs a temporal analysis for that year. This produces a consolidated feature class with persistence values for the specified year, which will then be subsequently used as an input feature class for the subsequent year to year analysis.

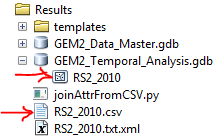
1. Launch “5a.\_Temporal Persistence Analysis (day to day)



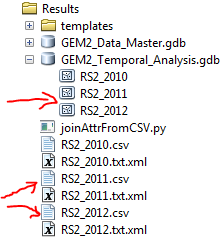
1. Enter the desired Year Dataset from the master geodatabase as the first parameter.



1. The second parameter is optional and contains the SQL expression that is used as the filter criteria with which to narrow down the number of dark targets to consider in the spatial-temporal analysis. This is based on selecting the attribute values which can help to indicate a higher likelihood that a dark target is a natural oil seep.
2. The third parameter contains the SQL expression that is used to remove any undesired dark targets that may be selected by the second parameter. This parameter is also optional.
3. Enter the desired distances, in metres, with which the spatial portion of the persistence analysis will be performed in the fourth parameter. Any number of distances can be entered in this parameter, for which the persistence values will be calculated. If a distance of ‘0’ is entered, the analysis will be a direct intersection of polygons rather than buffered with a distance.
4. Click “OK” to run the tool. Every acquisition day feature class is filtered to select only those dark targets that meet the selection criteria and analyzed together to calculate three persistence values (persistence, weight and cluster). The resulting analysis is output in a single consolidated feature class for the selected year, containing all dark targets. This includes those dark targets that were originally filtered out, however those targets will have ‘null’ persistence values. This result is placed in a newly created (if not already done) “GEM2\_Temporal\_Analysis” geodatabase, which contains the analysis results and provides a workspace in which the geoprocessing occurs. In addition, once the output feature class has been created, its attribute table is exported as a CSV file, which can then be accessed and manipulate with other software for further analysis of the data.



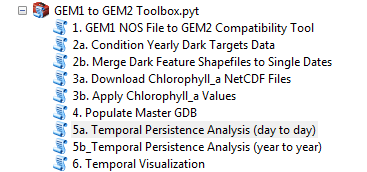
1. Run the tool again (steps 1 to 6) for every Year Dataset in the master geodatabase in order to create the necessary inputs for the next step in analysis.

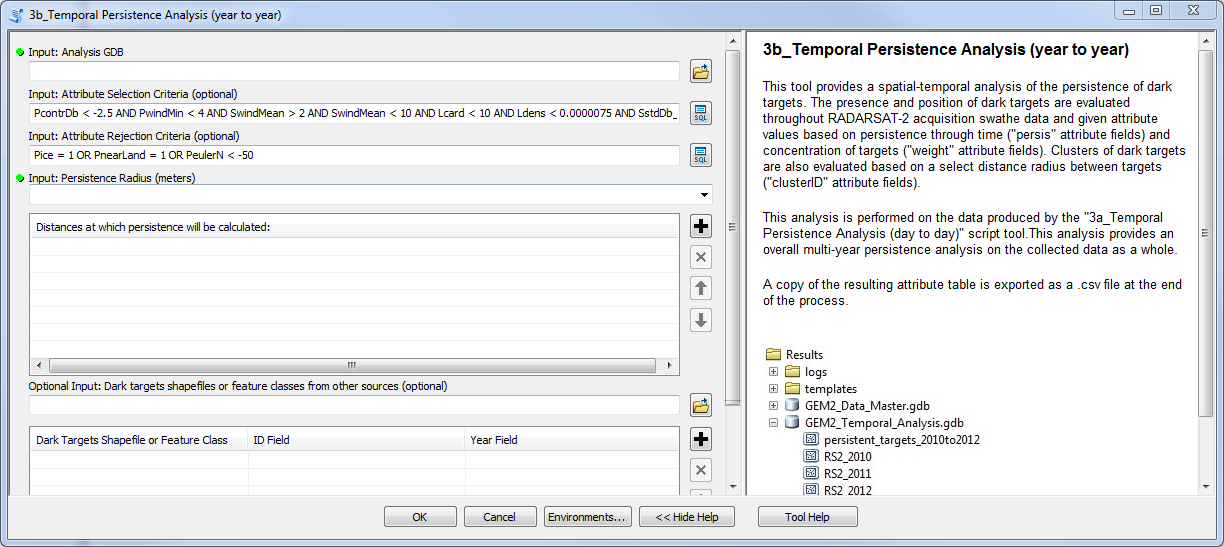


### Step 5: Temporal Persistence Analysis (year to year)

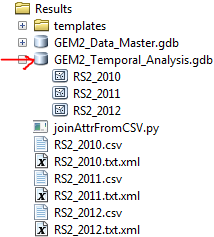
With the spatial-temporal analysis performed individually for every year, the data is now ready for the overall year-to-year persistence analysis.

1. Launch “3b\_Temporal\_Persistence\_Analysis (year to year)

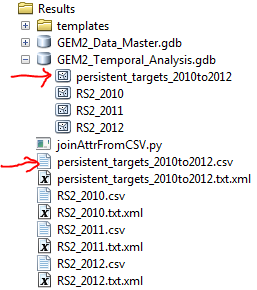




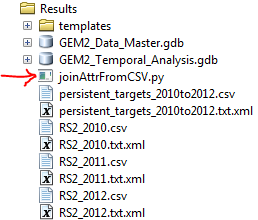
1. Enter the “GEM2\_Temporal\_Analysis” geodatabase as the input for the first parameter. The script will look inside the geodatabase and use all feature classes whose name starts with “RS2\_” to be used as input for the spatial-temporal analysis.



1. The selection and rejection criteria, as well as the persistence radius, parameters are the same as for tool “5a\_Temporal\_Persistence\_Analysis (day to day). Please refer to steps 3 to 5 in the previous section or the tool help for instructions on these parameters.
2. The last parameter is an optional input which can ingest external shapefiles or feature classes to also consider in the spatial-temporal analysis. A shapefile or feature class of dark targets can be selected with this parameter as long as a field identifying the year that the dark target was acquired is included in the attribute table. The shapefile or feature class is selected, with the field containing the unique identified, as well as the year field, subsequently specified.
3. Click “OK” to run the tool. The same spatial-temporal analysis is performed, however this time looking at calculating persistence value over the years rather than over the days within the year. A single “persistent\_targets” feature class is output, with the time span of the analysis appended at the end. A corresponding CSV file is also exported for further analysis if required.



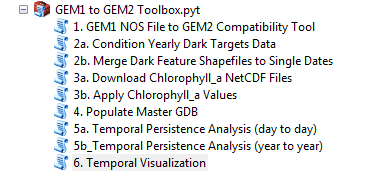
At this point, further analysis performed with data in the CSV file may lead to additional columns of data to be produced and to be reintroduced into the output analysis feature classes for further criteria filtering and analysis. This utilizes the “joinAttrFromCSV.py” script located in the “Results” folder. The steps on how to perform this additional analysis are described in a separate guide.

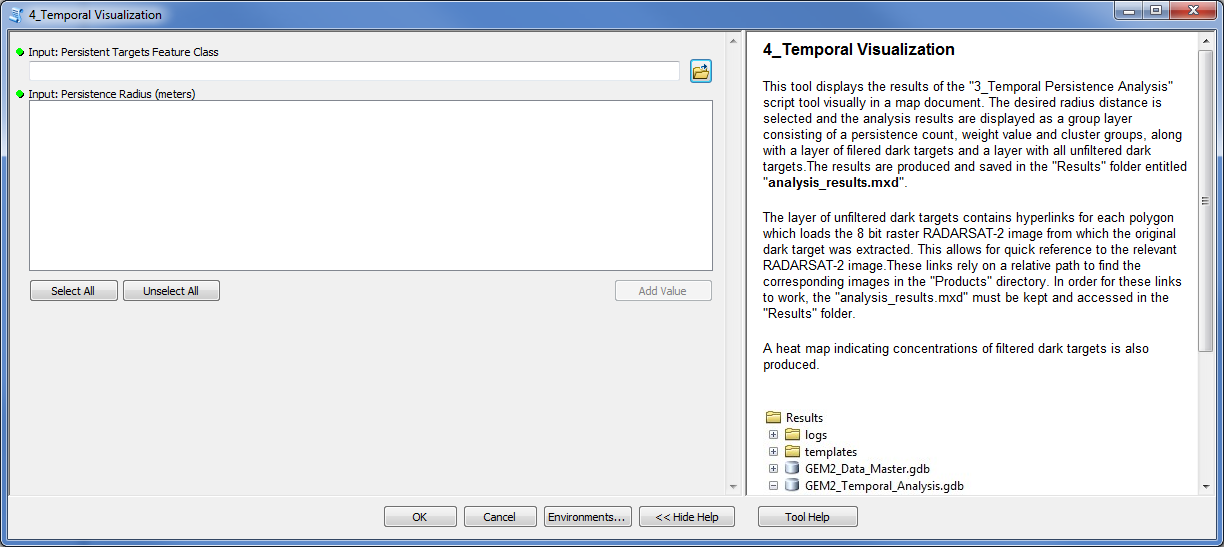


**Temporal Visualization**

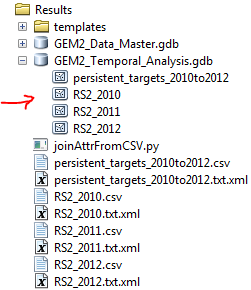
The final tool in the toolbox is used to produce a visual depiction of the persistence values that were calculated with analysis tools 3a and 3b. The script applies the output feature class, and its persistence values, to the template layers and map document located in the “templates” folder under “Results”. The map document template is copied and modified to reflect the attribute values of the desired output feature class.

1. Launch “6. Temporal Visualization” tool script.

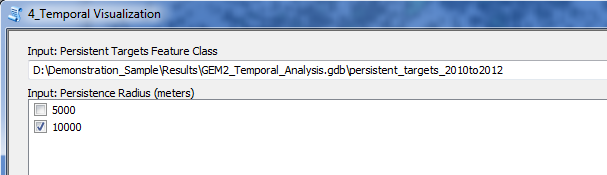




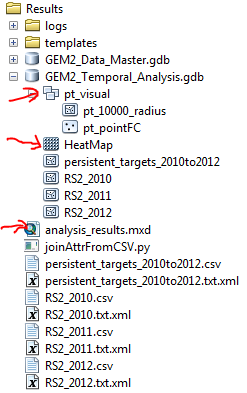
1. Enter the output analysis feature class for which the persistence values will be depicted as the first parameter. This will be one of the feature classes located in the “GEM2\_Temporal\_Analysis” geodatabase, be it the result from the day to day analysis (3a) or the year to year analysis (3b).



1. Upon selecting a feature class in the first parameter, a list of available persistence radii will appear as a selection in the second parameter. This reflects the buffer distances at which the analysis was performed and for which persistence values were calculated. Any number of persistence radii can be checked to include in the visual product (included as a group layer).



1. Click “OK” to run the tool. The input feature class is applied to the templates in the “templates” folder in order to depict the persistence values with a consistent format and symbology. A new dataset is created in the “GEM2\_Temporal\_Analysis” geodatabase, containing the required data used to build the visual product, as well as a heat map. The visual product is built and can be accessed through the newly copied “analysis\_results.mxd” map document in the “Results” folder.



Once created, the analysis results can be opened as a map document for further interpretation. The interpretation guide for the visual product and the persistence values is available as a separate document.