RS Tools User Guide

Table of Contents

Table of Contents

RS Tools toolbox
- Change Log V3.0.3
  - Raster Statistics
  - PixelExtractByPoint
- Change Log V3.0.2
  - MaskNoData (errors only in ArcGIS Pro version)
  - GenerateAnomaly
- Change Log v3.0.1
- Ported ArcGIS 10 version to ArcGIS Pro version
- Change Log v2.3.4
  - Raster Statistics
  - Pixel Extract By Point
- Change Log v2.3.2
  - Raster Statistics
- Change Log v2.3.1
  - Swap Color Map
- Change Log v2.3
  - Pixel Extract by point, Pixel Extract by polygon, & Raster Statistics
  - Create time series composites
- Change Log v2.2
  - General
  - Pixel Extract by point
  - Pixel Extract by polygon
  - Create time series composites
- Change Log v2.1
  - General
  - Pixel extract by point
  - Time Series composite
- Change Log v2.0
  - General
  - Create time series composite
  - Pixel Extract by points
  - Raster statistics by polygon
- Requirements

- Installation
- Connecting to Toolbox
- Input file naming conventions
  - NOAA daily product filenames (L3)
  - NOAA composited filenames (L4)
  - NASA CyAN daily product filenames
- Tools
  - Tool Help
  - Create time series composites
    - Description
    - Usage
  - Pixel extract by points
    - Description
    - Usage
    - Sample output file
  - Pixel extract by polygons
    - Description
    - Usage
    - Sample output file
  - Raster statistics by polygon
    - Description
    - Usage
    - Sample output file
  - Mask no data values
    - Description
    - Usage
  - Swap Raster Colormap
    - Description
    - Usage

- Importing X/Y point data into a feature class
  - Sample X/Y csv file
  - Date/Time attributes

- NASA / NOAA file differences

- Caveats

- Troubleshooting
  - Slow processing
  - Out of memory error

Delete working/intermediate files

RS Tools toolbox

RS_Tools is an ArcGIS Pro Python toolbox. It contains a collection of tools designed specifically for working with satellite data raster files distributed by NOAA/NCCOS and the CyAN project.

Change Log V3.0.4

Raster Statistics

- If a raster has no valid data pixels within zones, a row is added with results set to 0. Previously, only a warning message was displayed and no row added.

Change Log V3.0.3

Raster Statistics

- Fixed bug introduced in v2.3.4 that dropped the zone field column in output CSV

PixelExtractByPoint

- Added additional pixel window extract sizes ('7x7', '9x9', '11x11', '13x13')

Change Log V3.0.2

MaskNoData (errors only in ArcGIS Pro version)

- fixed bug in writing to wrong output location
- fixed bug with scratchWorkspace
- OSError with arcpy.Describe('%s\\Band_1' % raster_fn)

GenerateAnomaly

- creates simple difference

Change Log v3.0.1

Ported ArcGIS 10 version to ArcGIS Pro version

Change Log v2.3.4

Raster Statistics

- Zone Data feature class to layer to fix SelectLayerByAttribute_management (if running from script)
- tweaks to improve processing efficiency

Pixel Extract By Point

- option to reformat output from multi-band raster so each value from each band is added to column "bnum_#" instead of the default which creates a new row for each band (requires pandas therefore ArcGIS v10.4+)

- quote string fields to avoid potential problems with commas

Change Log v2.3.2

Raster Statistics

- Added options to filter by matchup date in polygon shapefile
    - caveats
        - only supporting same day matchups
        - zone field should return a unique polygon
        - does not honor the selection of features in "zone data" feature class

Change Log v2.3.1

Swap Color Map

- New tool to replace internal colormap in raster

Change Log v2.3

Pixel Extract by point, Pixel Extract by polygon, & Raster Statistics

- support for NOAA version2 CI products
- ‘unscaled’ column excluded if not supported

Create time series composites

- Toolbox command added to *lineage* metadata in *.tif.xml output files

Change Log v2.2

General

- Bug fixes

Pixel Extract by point

- options to summarize pixels in 3x3 or 5x5 window
- select multiple attributes for output from feature class
- if matchup option, added time delta field in CSV
- 3x3 or 5x5 window include pixels as columns instead of rows
- band number filtering for multiband geotiffs

Pixel Extract by polygon

- select multiple attributes for output from feature class

Create time series composites

- performance improvements when input rasters located on network

Change Log v2.1

General

- Changed input file selection method. Added option to define method for input file selection:
    - from workspace (current method)
    - from a list of filenames in a text file
- Support for unicode characters in feature classes
- Support for NASA CyAN daily and composite filenames
- Various bug fixes

Pixel extract by point

- Added ability to extract pixels within XxX window centered on point features

Time Series composite

- Added option to combined files across years

Change Log v2.0

General

- various bug fixes

Create time series composite

- added *Do not span months* option input parameter
- added *Copy landmask* option input parameter
- modified output filename format
    - <prefix>_yyyy.mmdd_mmdd.L4.<area>.<prod>.<statistic>[_<suffix>].tif
- if the input files were previously created by *Time Series Composite* (i.e.L4 type files), the selection of files for a composite period is now based on the mid point of start/end date of each L4 file (v1.0 used the start date)

Pixel Extract by points

- added *unscaled_val* column to output CSV file
    - if the product type is identified as a CI, Cicyano or CInoncyano product, the unscaled CI value will be added otherwise the field will be blank
- added *band_num* column to output CSV file
    - for single band input files this field will contain a 1
    - for multi-band input files this field will identify the band for the extract data
- resolved memory leak issue that had potential to produce invalid results
- spatial reference of the X,Y values written to the CSV file set to lon/lat (WGS84)
- removed *Interpolate values* option

Raster statistics by polygon

- added *Minimum valid data value* and *Maximum valid data value* input parameters

- eliminates the need to create a set of input rasters with no data values defined prior to running the tool.

Requirements

ArcGIS Pro 10.2.5 or higher with Spatial Analyst. It may work with earlier versions but has not been tested. This version will **not** work with ArcGIS 10.

Installation

Copy the RS_toolboxPro_v#.#.#.zip to your local disk and unzip all the files within the archive into their own subdirectory. “v#.#.#” refers to the current version number of the toolbox. A suggested location is C:\Users\<username>\Documents\ArcGIS\RS_ToolboxPro_v#.#.#. Assuming this is version 3.0.1 of the toolbox, the subdirectory should contain the following files:

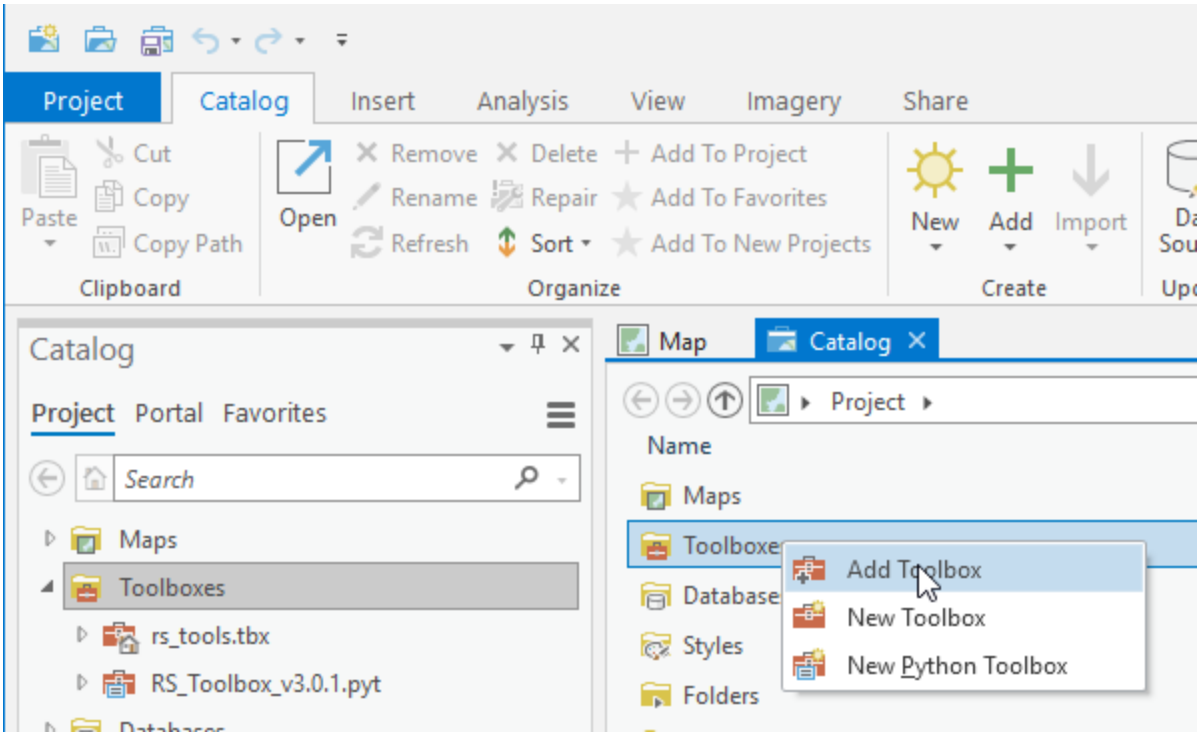
- RS_ToolboxPro_v3.0.1.CreateTimeSeries.pyt.xml
- RS_ToolboxPro_v3.0.1.MaskNoData.pyt.xml
- RS_ToolboxPro_v3.0.1.PixelExtractByPoint.pyt.xml
- RS_ToolboxPro_v3.0.1.PixelExtractByPolygon.pyt.xml
- RS_ToolboxPro_v3.0.1.pyt
- RS_ToolboxPro_v3.0.1.pyt.xml
- RS_ToolboxPro_v3.0.1.RasterStatsByPolygon.pyt.xml
- RS_Tools_Pro_User_Guide.pdf

Connecting to Toolbox

There are a number of ways to connect to the toolbox. First, start ArcGIS Pro and open an existing project or create a new Map project.

The simplest method is to simply navigate to the folder where you unzipped RS_ToolboxPro within the Catalog Pane. If your project is not already connected to the folder, you will first need to connect it.

Alternatively, you can connect the Toolbox directly to a project by adding the RS_Tools under Toolboxes within the Catalog View.



Detail instructions for both methods may be found within the ArcGIS online help.

Input file naming conventions

The tools extract information about the input files from the filenames and therefore the filenames must conform to specific naming conventions as documented below. RS_Tools currently supports files from two sources, NOAA and NASA. Each source has its own unique filenaming convention. Items in bold are used specifically by the tools to control processing and/or output so the actual values are important. The items not in bold may generally contain any string but should not include additional "." or "_".

NOAA daily product filenames (L3)

- <satellite>.**<yyyy>**<jjj>.**<mmdd.HHMM>**[C|S].L3.**<area>**.<ver>[.**<product>**].**tif**
  where

  <satellite> - name of satellite from which the data product was created
  <yyyy> - 4-digit year of data collection
  <jjj> - 3-digit number representing the day of the year of data collection
  <mmdd.HHMM> - month, day, hour and minute of data collection; HHMM may repeat (with "_" separator) if daily file created from multiple same day granules or passes.
  <area> - area name extracted from the first input file in the composite group
  <ver> - software version used to generated original L2 file
  - L3 file version

- product file version
<product> - product name extracted

NOAA composited filenames (L4)

- <prefix>_**<yyyy>**.**<start_mmdd>_<end_mmdd>**.L4.**<area>**.**<product>**..*
  where
  - <prefix> - user specified string
  - <yyyy> - 4-digit year of data collection
    - if *Combine Years* option <start_yyyy>-<end_yyyy>
  - <start_mmdd> - starting month and day of the composite period
  - <end_mmdd> - last month and day of the composite period
  - <area> - area name extracted from the first input file in the composite group
  - <product> - product name extracted from the first input file in the composite group
  - * - any string of characters (may include "." and/or "_")

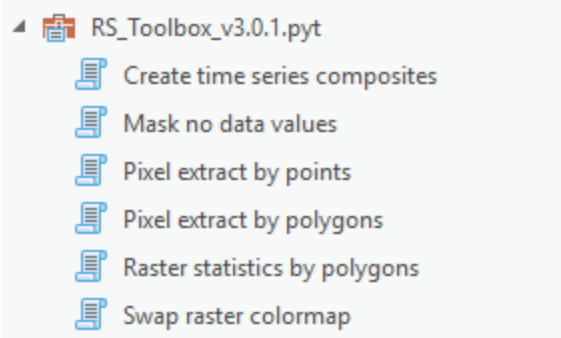
NASA CyAN daily product filenames

- <sensor>**<yyyy><ddd>**.L3m_DAY_CYANF_**<product>**_CYAN_CONUS_300m_**<region>**.tif
  where
  - <sensor> - "M" = MERIS or "L" = OLCI
  - <yyyy> - 4-digit year of data collection
  - <ddd> - 3-digit number representing the day of the year of data collection
  - <product> - product name
  - <region> - geographic region

NASA CyAN composited filenames

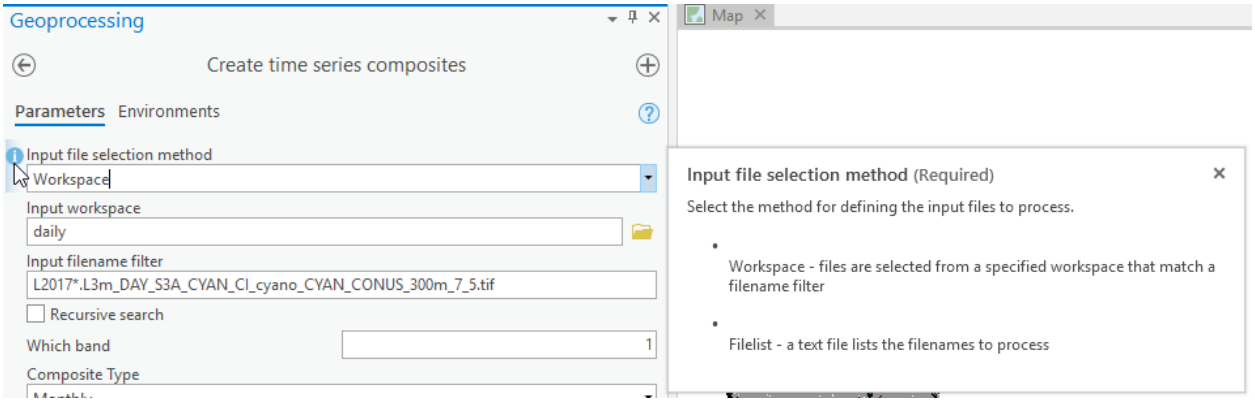
- <sensor>**<start_yyyy><start_ddd><end_yyyy><end_ddd>**.L3m_<period>_CYAN_**<product>**_CYAN_CONUS_300m_**<region>**.tif
  where
  - <sensor> - "M" = MERIS or "L" = OLCI
  - <start_yyyy> - starting 4-digit year of composite period
  - <start_ddd> - starting 3-digit day of year of composite period
  - <end_yyyy> - ending 4-digit year of composite period
  - <end_ddd> - ending 3-digit day of year of composite period
  - <product> - product name
  - <region> - geographic region

Tools

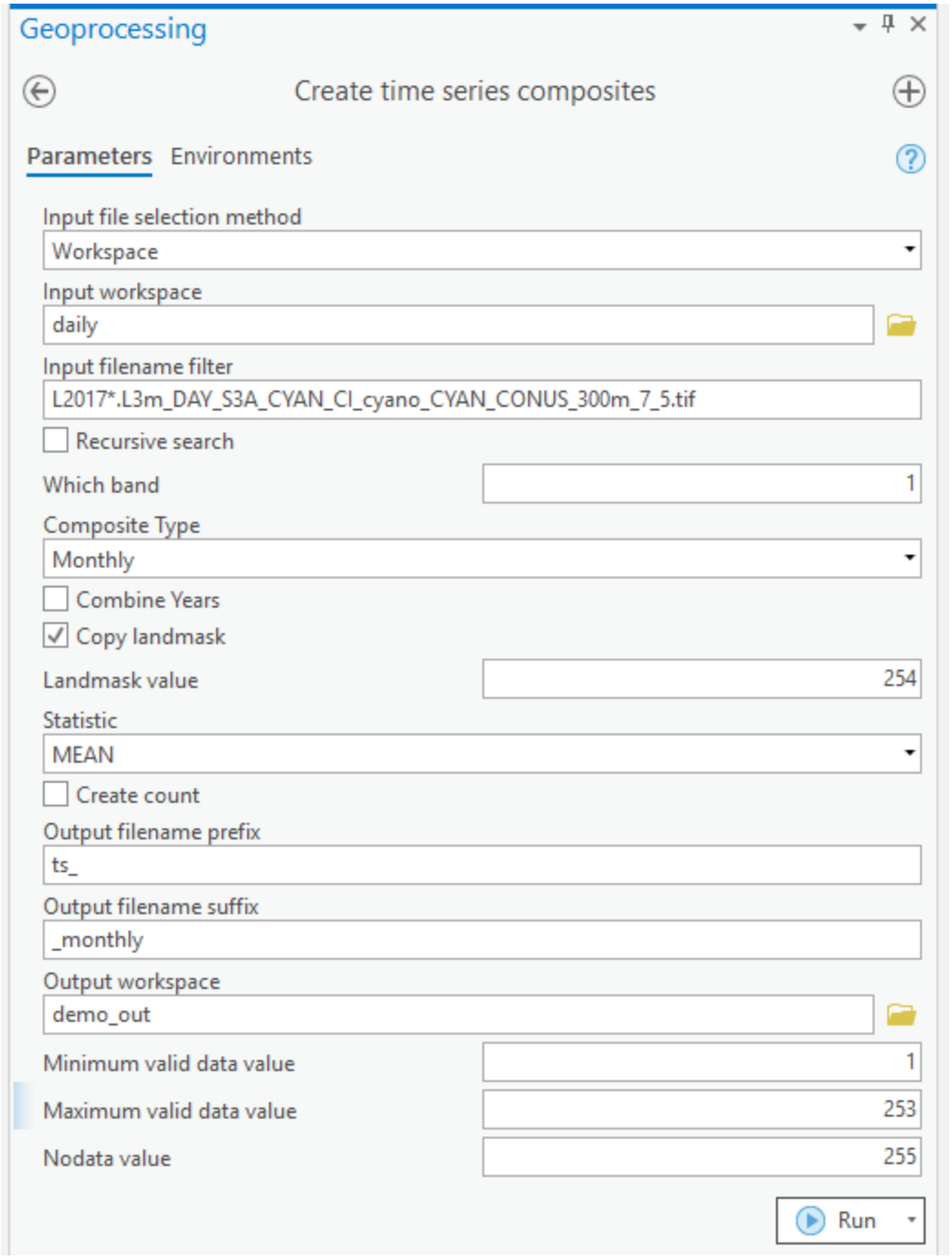


Tool Help

Internal help is available for each tool. To access a description of the tool, hover the cursor over the  icon. Each input parameter has context specific help that may be viewed by hovering over the  icon displayed to the left of each parameter.



Create time series composites



Description

This tool composites a user defined list of geographically coincident raster images based on the specified composite type. The types of compositing available include:

Composite Type	Description

Monthly	Images for the same year and month are composited based on the specified statistic. If *Combine Years* is checked, all images for a month are combined.
Yearly	Images for the same year are composited based on the specified statistic. If *Combine Years* is checked, all images for a month are combined.
Seasonal	Images for consecutive 3-month/year periods are composited based on the specified statistic. The start date defines the beginning of the first period. If *Combine Years* is checked, all images for a 3-month are combined.
Time Period	Images for a user defined time period are composited based on the specified statistic. The start date defines the beginning of the first period. Image composites can only be processed for a single year at a time. If the calculated last time period for the year is less than 80% of the time period, the time window of the previous period will be extended to include the partial period.
All	All images in the list are composited into a single raster based on the specified statistic. Use this option with caution as attempting to composite a large number of files could result in memory issues.

Usage

- All input rasters must have the same raster properties, cover the exact same geographic extent, and have the identical spatial reference system.
- Only a single band from the raster files may be composited. The NOAA/NCCOS supplied raster images are single-band GeoTiffs, so the *which band?* parameter should always be set to “1”.
- See section on Input file naming conventions
- The output composited files are named:
    - <prefix>_<yyyy>.<start_mmdd>_<end_mmdd>.L4.<area>.<product>.<statistic>[_<suffix>]

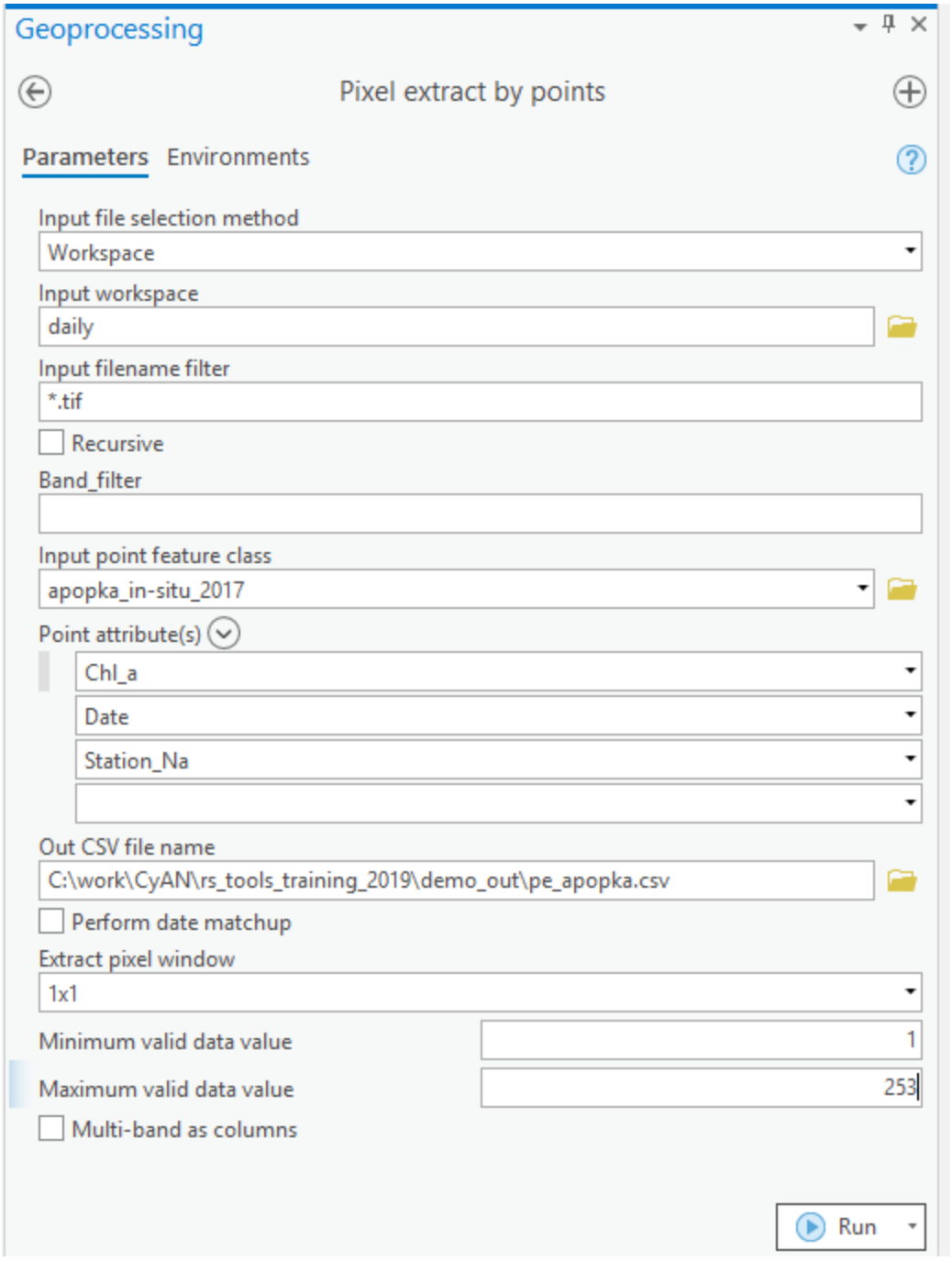
where

<prefix> - user defined output filename prefix
<yyyy> - 4-digit year
<start_mmdd> - starting month and day of the composite period
<end_mmdd> - ending month and day of the composite period
<area> - area name extracted from the first input file in the composite group
<product> - product name extracted from the first input file in the composite group
<statistic> - name of statistic
<suffix> - optional user defined output filename suffix

note: the <yyyy> portion of filenames generated using the “combine years” option will be replaced by <start_yyyy>-<end_year> to represent the range of years contained within the composite

- If the output workspace is a folder, the output format will be GeoTiff and a ".tif" file extension will be appended to the filename.
- If the output workspace is a geodatabase, the "." separator in the filenames will be changed to "__" and no file extension will be appended to the filename.

Pixel extract by points



Description

This tool extracts pixels from all rasters that intersect points within a feature class. By default a row for each point intersecting each raster is written to a CSV file. If the matchup option is checked, rows will be output only if the raster date and the point date fall within the defined time window.

Each row in the CSV file will contain the raster date, point x location, point y location, pixel value, and the raster filename. Optionally selected attribute(s) in the point feature class may be appended to each row.

Usage

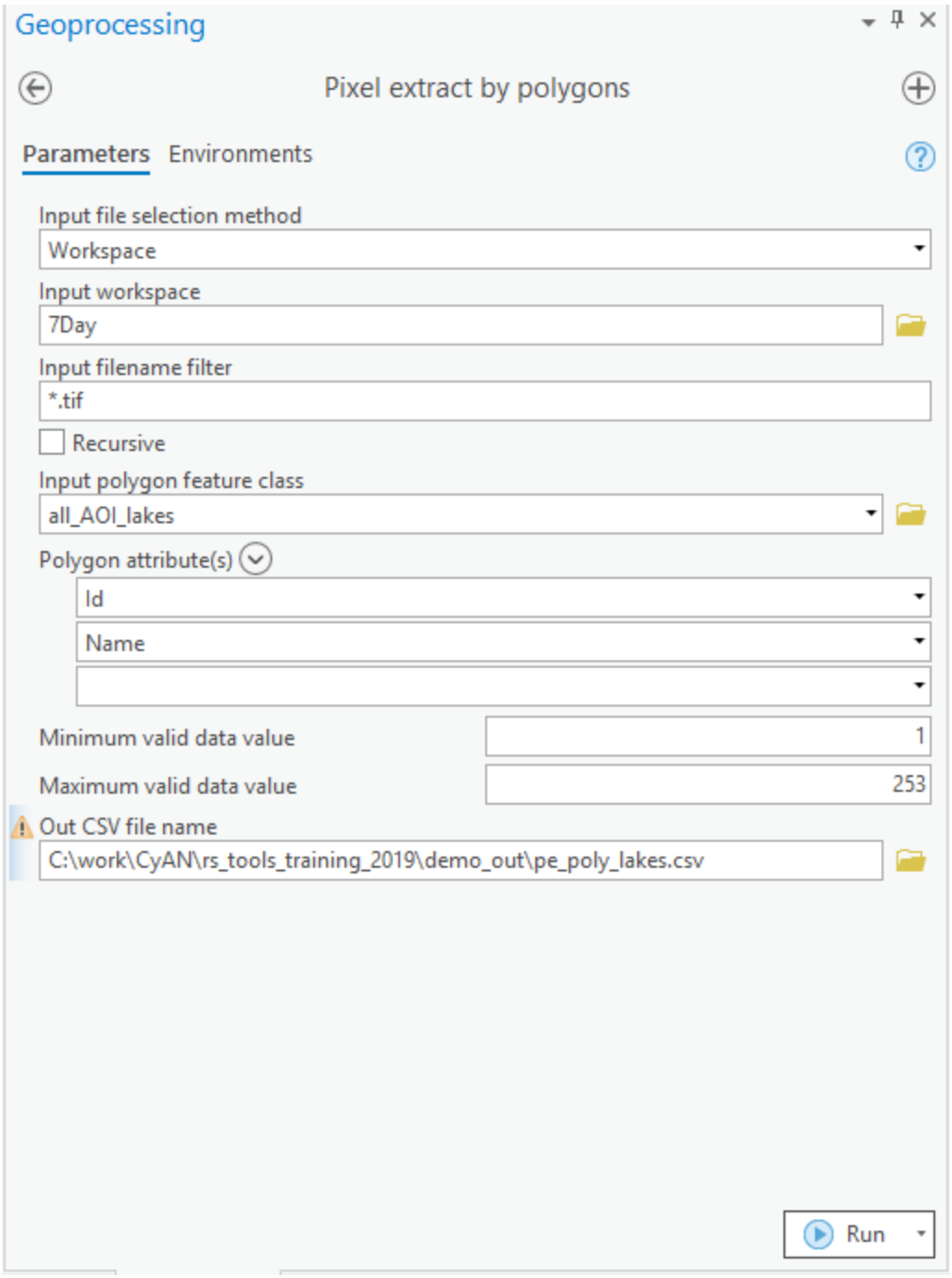
- See section on Input file naming conventions
- If point features within the feature class are selected the extraction will be limited to only those points.
- If the matchup option is checked the feature class must contain an attribute with data type of “date”.
- Note that if the feature class is a coverage or shapefile, ArcGIS truncates the time portion from the attributes defined with a data type of “date”. For example, the datetime 2002-08-20 12:00:00 PM is stored in a coverage or shapefile as 2002-08-20.
- Matchup only uses the date portion from the filename and sampling date if the Time unit is “day(s)”. If the Time unit is “hour(s)”, both the date and time are used. If no time portion has been defined or available, the time will be 00:00:00.
- If the input rasters contain values outside the range of valid data values, set the minimum and/or maximum valid data values to force masking values outside the range. Masked values and nodata values will contain a value of blank within the output file.
- The extracted value is based on the value at the center of the cell if *Extract pixel window* is 1x1. If a larger window is selected all pixels in the window center around the 1x1 cell will be extracted and added as columns unless the “XxX <statistic>” option is selected .
- If *Extract pixel window* option “XxX <statistic>” is selected, the extracted value will be based on the summary statistic within the XxX window. For example, “3x3 max” will return the maximum valid value within the 3x3 window.
- The spatial reference of the X,Y values written to the CSV file is lon/lat (WGS84).
- If the input rasters are a recognized product (i.e. CI, CIcyano, or CInoncyano), an additional column will be added to the CSV file that contains the unscaled product value. The product is identified from the input filename (see section input filenaming conventions).
- The *band_num* column identifies the band corresponding to the extracted data. For single band input rasters it will be “1”.
- If the input rasters are multi-band files, pixel values will be extracted from all bands unless a valid *band_filter* is specified.

Sample output file

Columns A-G are standard in each output file. Column H was included because Perform matchup was checked and the attribute name “result_ug_” was selected in the *Point attribute(s)* field list. Column I was included because the name “sampling_l” was also selected from the *Point attribute(s)* field list.

	A	B	C	D	E	F	G	H	I
1	date	x	y	val	unscaled_	band_nun	file	result_ug_	sampling_l
17	6/21/2010	-122.866	39.06457	62	0.000417	1	C:/work/M	0	CL-01
18	6/21/2010	-122.682	38.96463	133	0.002138	1	C:/work/M	27	CL-03
19	6/21/2010	-122.702	39.01238	136	0.002291	1	C:/work/M	20	CL-04
20	6/23/2010	-122.866	39.06457			1	C:/work/M	0	CL-01
21	6/23/2010	-122.682	38.96463	171	0.005129	1	C:/work/M	27	CL-03
22	6/23/2010	-122.702	39.01238	108	0.001202	1	C:/work/M	20	CL-04

Pixel extract by polygons



Description

This tool extracts pixels from all rasters that intersect polygons within a feature class. By default a row for each raster cell intersecting all polygons from all rasters is written to a CSV file.

Each row will contain the raster date, point x location, point y location, pixel value, and the raster filename. Optionally user selected attributes in the polygon feature class may be appended to each row.

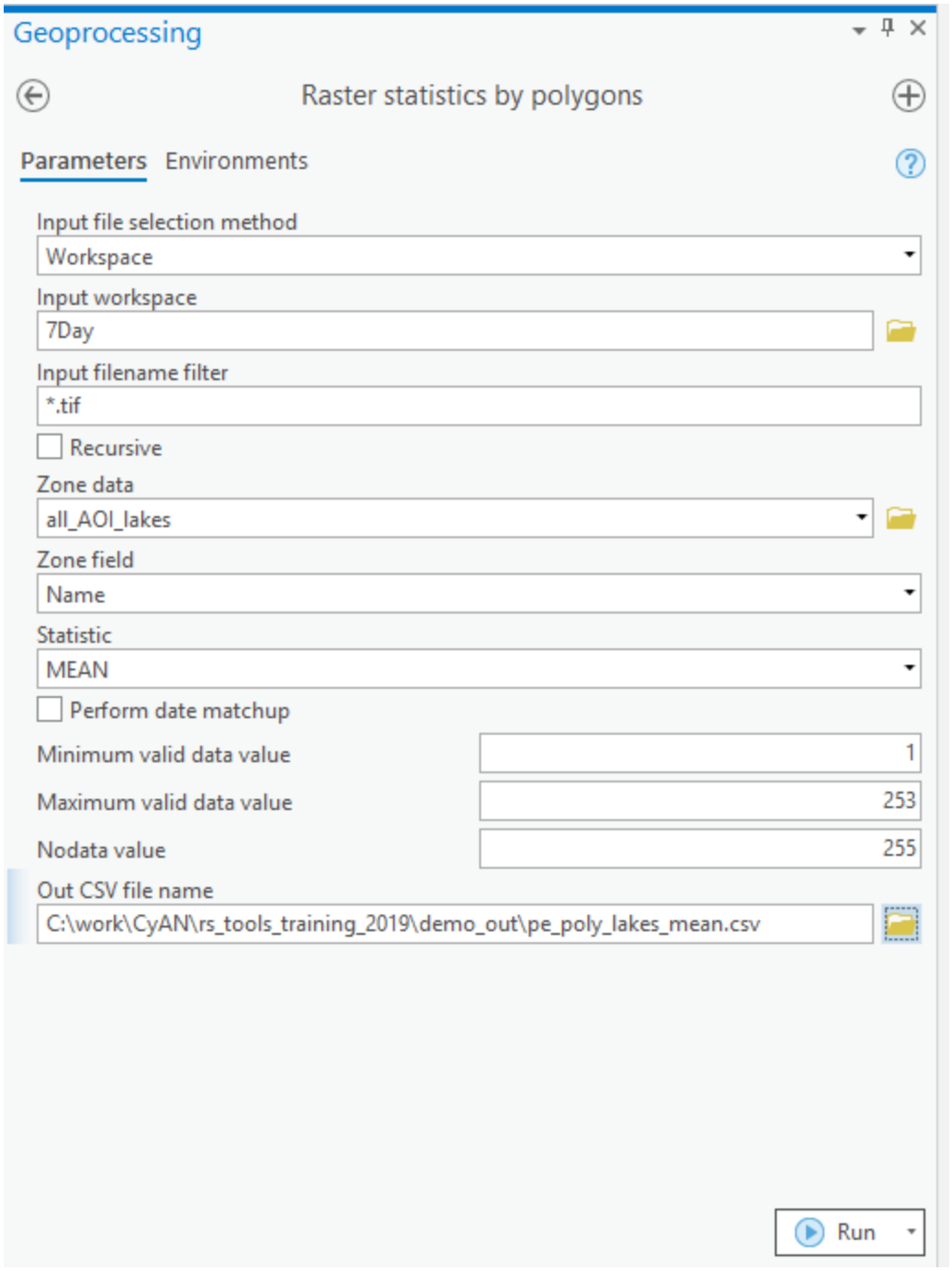
Usage

- See section on Input file naming conventions
- If polygons are selected within the feature class the extraction will be limited to only those polygons.
- If the input rasters contain values outside the range of valid data values, set the minimum and/or maximum valid data values to force masking values outside the range. Masked values will contain a value of *blank* within the output file.
- Pixels defined as *nodata* in the input rasters are not extracted.
- By default the coordinates of the X,Y values written to the CSV file will match the coordinate system of the input raster files. You can override the default behavior by modifying the environment settings.
    - Click the *Environments…* button to open the Environment Settings window and define the output coordinate system under the *Output Coordinates* section.
- If the input rasters are a recognized product (i.e. CI, CIcyano, or CInoncyano), an additional column will be added to the CSV file that contains the unscaled product value. The product is identified from the input filename (see section input file naming conventions).
- Currently does not support multiband rasters

Sample output file

	A	B	C	D	E	F	G
1	date	x	y	val	unscaled_	file	name
14	1/1/2009	-247115	126101.4	1	0.000102	C:/work/N	Clear Lake
15	1/1/2009	-246815	126101.4	83	0.000676	C:/work/N	Clear Lake
16	1/1/2009	-246515	126101.4	1	0.000102	C:/work/N	Clear Lake
17	1/1/2009	-246215	126101.4	4	0.00011	C:/work/N	Clear Lake
18	1/1/2009	-245915	126101.4	16	0.000145	C:/work/N	Clear Lake
19	1/1/2009	-245615	126101.4	90	0.000794	C:/work/N	Clear Lake
20	1/1/2009	-245315	126101.4			C:/work/N	Clear Lake

Raster statistics by polygon



Description

Extract statistics based on zones within a feature class to a CSV file from one or more rasters.

The available statistics (from ArcGIS documentation):

Statistic	Description
Mean	Calculates the average of all cells in the value raster that belong to the same zone as the output cell.
Majority	Determines the value that occurs most often of all cells in the value raster that belong to the same zone as the output cell.
Maximum	Determines the largest value of all cells in the value raster that belong to the same zone as the output cell.
Median	Determines the median value of all cells in the value raster that belong to the same zone as the output cell.
Minimum	Determines the smallest value of all cells in the value raster that belong to the same zone as the output cell.
Minority	Determines the value that occurs least often of all cells in the value raster that belong to the same zone as the output cell
Range	Calculates the difference between the largest and smallest value of all cells in the value raster that belong to the same zone as the output cell.
Std	Calculates the standard deviation of all cells in the value raster that belong to the same zone as the output cell.
Sum	Calculates the total value of all cells in the value raster that belong to the same zone as the output cell.
Variety	Calculates the number of unique values for all cells in the value raster that belong to the same zone as the output cell.

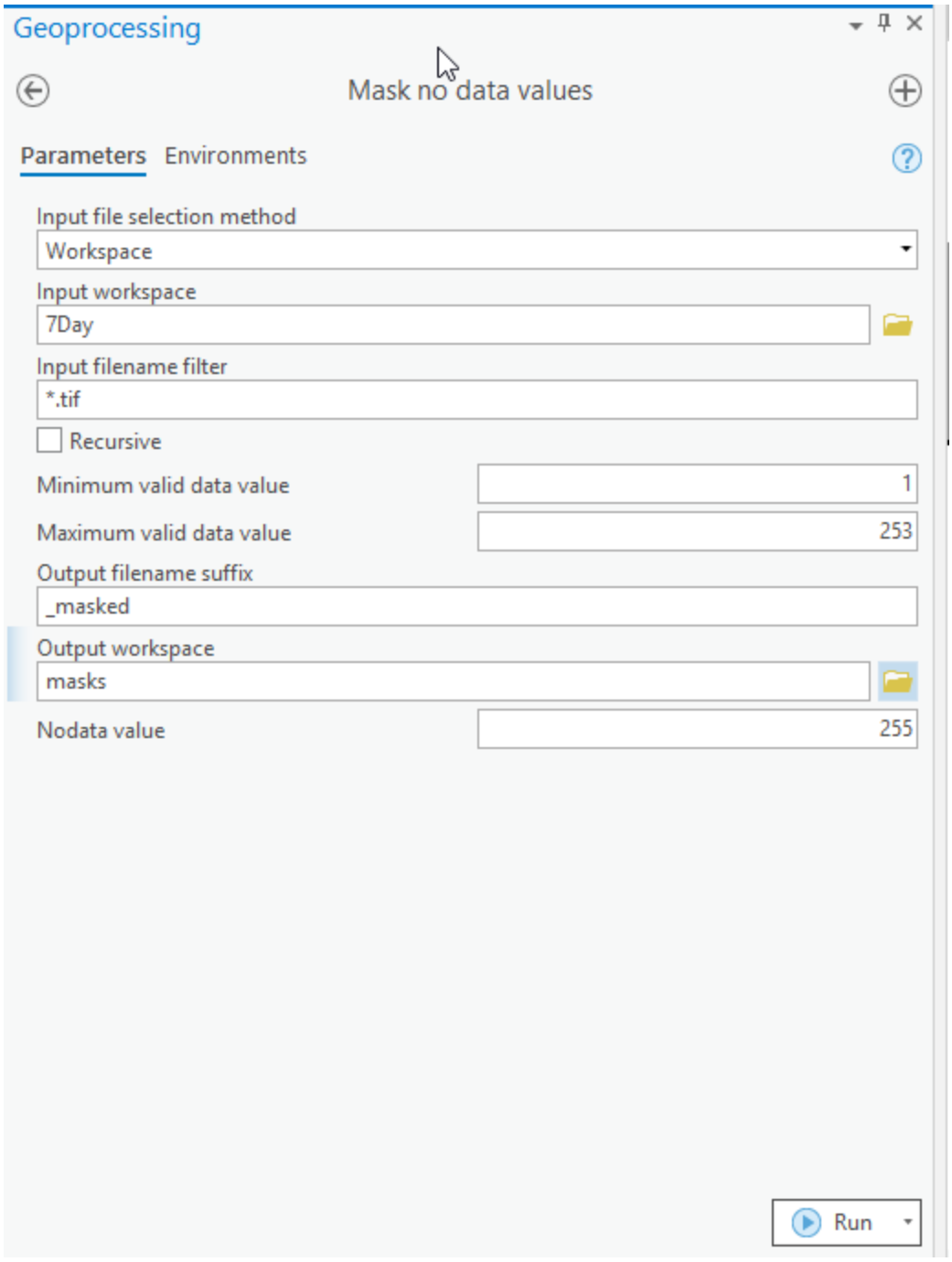
Usage

- See section on Input file naming conventions
- A zone is defined as all areas in the Zone data feature class that have the same value for the selected zone field attribute.
- If features are selected within the Zone data feature class the analysis will be limited to only those features.
- If rasters contain values outside a *valid* data range, set the *Minimum valid data value* and *Maximum valid data value* to force exclusion of values in calculation of statistics.
- An error will be returned for any raster that does not overlap the (selected) zone data.
- If the input rasters are a recognized product (i.e. CI, CIcyano, or CInoncyano), additional column(s) will be added to the CSV file that contains the unscaled product value. The product is identified from the input filename (see the section on Input file naming conventions).

Sample output file

	A	B	C	D	E	F	G	H
1	date	OBJECTID	name	ZONE_CODE	COUNT	AREA	MEAN	file
20	1/31/2010	4	Middle Alkali Lake	4	769	69210000	3.878414	C:\Users\A
21	1/31/2010	5	Honey Lk Wldfwl Mgmt Pnds	5	3027	272430000	0.786587	C:\Users\A
22	2/10/2010	1		1	11201	1008090000	14.41934	C:\Users\A
23	2/10/2010	2	Goose Lake	2	4806	432540000	100.0299	C:\Users\A
24	2/10/2010	3	Upper Lake	3	1564	140760000	88.2759	C:\Users\A
25	2/10/2010	4	Middle Alkali Lake	4	769	69210000	73.45319	C:\Users\A
26	2/10/2010	5	Honey Lk Wldfwl Mgmt Pnds	5	3027	272430000	106.3442	C:\Users\A

Mask no data values



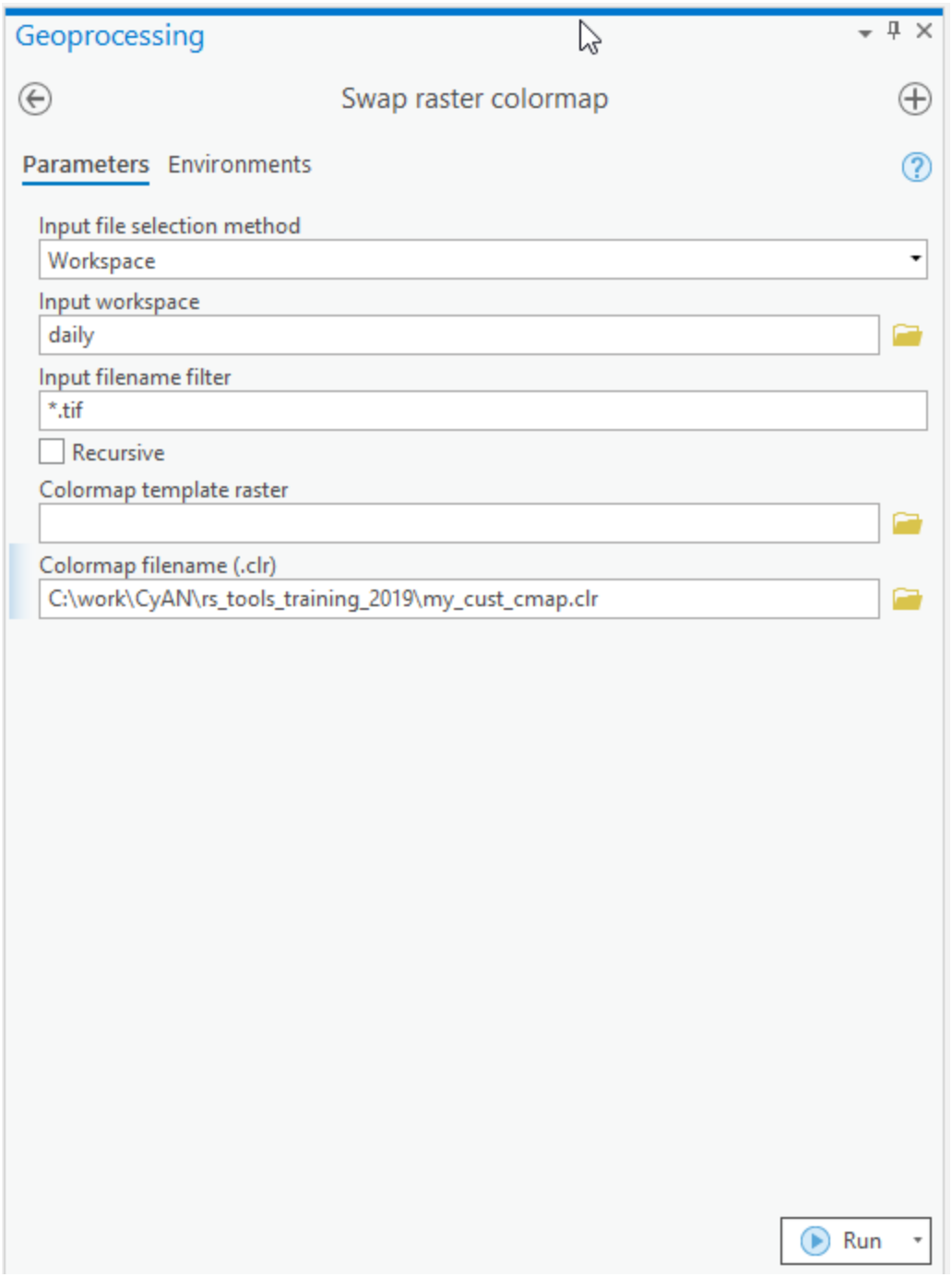
Description

This tool processes the list of rasters and sets pixel values outside the minimum and /or maximum valid data range to the specified *Nodata* value.

Usage

- See section on Input file naming conventions
- Output filenames will match the input filenames but will optionally include the user defined *Output filename suffix*.
- To avoid overwriting the input rasters, either specify an Output filename suffix or if no Output filename suffix is specified be sure the Output workspace is different from the Input workspace.

Swap Raster Colormap



Description

This tool processes the list of rasters and replaces the internal colormap with either the colormap within a template raster or a colormap defined in a colormap (.clr) file.

The default CI product colormap files "*CI_v2_noaa.clr*" and "*CI_v2_nasa.clr*" are included in the *colormaps* subfolder of the RS_Tools distribution.

A colormap file is a space-delimited text file that defines the pixel value and the red, green, and blue color values. The red, green, and blue values can range from 0 to 255. Missing values are assigned a color 0, 0, 0. The first value on each line is the pixel value and remaining values define the RGB code. The following is an example snippet:
0 0 0 0
1 72 25 107
2 72 26 108
...

Usage

- See section on Input file naming conventions
- The input rasters are updated in-place.
- The input rasters must all have a pixelType of unsigned integer.
- A value in the *Colormap template raster* field will override any value in the *Colormap filename (.clr)* field.

Importing X/Y point data into a feature class

A number of tools operate on feature (points or polygons) class data which can be in a standalone shapefile or feature class within an ESRI geodatabase. X/Y point data contained in various text or Excel file formats may be imported into a feature class. Although X/Y text files may be imported directly into ArcGIS, converting the X/Y text file to an Excel file and defining the column formats in Excel prior to importing provides greater control on how ArcGIS interprets each column. Detailed instructions on importing X/Y point data into ArcGIS checkout the following link.

Sample X/Y csv file

date,sampling location,Longitude,Latitude,depth meters,result ug/L
3/16/2010,CL-01,-122.86616,39.06457,0.5,2
3/16/2010,CL-03,-122.68166,38.96463,0.5,2.7
3/16/2010,CL-04,-122.70184,39.01238,0.5,7.1
5/18/2010,CL-01,-122.86616,39.06457,0.5,2

Date/Time attributes

Performing date match-ups with the *Pixel Extract by Point* tool requires the feature class contains an attribute of type *date*. If importing X/Y from an Excel file be sure the column is formatted as a date. If importing X/Y directly from a text file ArcGIS will automatically convert columns to a date if formatted appropriately. Some examples of acceptable date formats are listed below:

8/9/2017
08/09/2017
8/9/2017 23:55:00
8-9-2017
8-9-2017 11:55:00
2017/08/09
2017-08-09

Note that shapefiles do not support the time portion in *date* type attributes. The time component of a *date* type attribute will always be '000000'. To maintain both date and time save the data to a feature class within an ESRI geodatabase instead of a shapefile.

NASA / NOAA file differences

The RS_Tools defaults are defined based on NOAA GeoTiff product files. If using NASA CyAN GeoTiff files the defaults may not always be the best choice. The following table presents the primary differences.

	NOAA	NASA
Valid data range	1-250	1-253
Landmask value	252	254

Caveats

The tools were developed specifically to work with NOAA/NCCOS and NASA CyAN single band integer GeoTiff files. They can potentially be used with other types of inputs although other possible data sources have not been tested. Alternate input rasters would also need to meet the file naming requirements stated earlier in the guide.

Troubleshooting

Slow processing

Some processing just takes time depending on the number of factors including the size of the raster being processed, the number of files being processed, the operation being performed, and the specifications of the computer running the tools. The following suggestions have been documented to improve performance:

1. Write output to local drives rather than network drives
2. Store input data on local drives rather than network drives
3. Be sure your Environment settings for *Current* and *Scratch Workspace* point to a local drive

Out of memory error

Memory leaks have been documented in ArcGIS. These can lead to tools failing with an “out of memory” error. This problem can often be corrected by restarting ArcMap and re-running the tool.

Delete working/intermediate files

Delete files in ArcGIS scratch.gdb and/or scratch folder

Floating point raster input

- **Time series composites** - if “Copy landmask” checked output raster will be converted to 8-bit integer raster.
- **Raster statistics by polygon** - valid selections for Statistic are ALL | MEAN | MAXIMUM | MINIMUM | RANGE | STD | SUM | MIN_MAX | MEAN_STD | MIN_MAX_MEAN.Fails if any other Statistic selected