Help for StreamStats Batch Processor

Version 5.02 patch 7

5/22/2020

Changes

- The output GlobalWatershed feature class now only contains fields for the requested basin characteristics instead of all of them
- Name field changed to 100 characters in all cases. This field is populated by the user defined local_id field, which must be 100 characters or less.
- DRNAREA, CONTDA, and NONCONTDA are always computed if 'Compute Basin Chars' or 'Compute Flow Stats' checked
- Basin Characteristics ordered according to calculation needs (drain areas, simple calcs, then derived calcs)

Known Bugs:

- Shapefile output option is not available
- Option to select output projection is not considered
- Need to add PC (Percent Complete) and Disclaimer where appropriate for Flow Statistics

Version 5.0 8/1/2019

Changes (Only back end – bp_job changes)

- Bug fix: under certain conditions,flow stats were not computing if the value of the basin characteristic was zero
- As of this version, a different process will start to be used for calculating basic characteristics. The results are expected to be the same, but rounded to USGS specifications before being stored in the globalwatershed . Those basin characteristics calculated with the original processing will continue to be stored to six decimal places.

Known Bugs:

- Shapefile output option is not available
- Option to select output projection is not considered
- Schema mismatches showing up - looking into generating schema on the fly
- Need to add PC (Percent Complete) and Disclaimer where appropriate
- ORREG2 in OR required for Peak Flows, but not automatically added in web interface. It needs to be added manually

Version 4.8 5/25/2019

Changes (Only back end – bp_job changes)

- Bug fix: study areas with the format XX_XXXX were failing due to programming logic

- Added jobID to bp_queue to aid in debugging
- Fixed bug where user uploads custom projection with slightly different params than data

Known Bugs:
- Shapefile output option is not available
- Option to select output projection is not considered
- Schema mismatches showing up - looking into generating schema on the fly
- Need to add PC (Percent Complete) and Disclaimer where appropriate

Version 4.7 3/12/2019

Changes (Only back end – bp_job changes)
- If the Local ID Field does not exist in the shapefile, one is selected - it no longer fails
    - Search order for selected field: Name, site_no, SITE_NO, STA_ID, FID, OBJECTID, field in shapefile
- Delineation cleanup - when shape field is null, the record is skipped instead of entire job failing
- Latitude, Longitude fields in GlobalWatershedPoint now output correctly
- Regional Study Areas (DRB, CRB, RRB) now compute

Known Bugs:
- Shapefile output option is not available
- Option to select output projection is not considered
- Schema mismatches showing up - looking into generating schema on the fly
- Need to add PC (Percent Complete) and Disclaimer where appropriate

Version 4.6 2/12/2019

Changes (Only calculations on the back end – bp_job – were changed)
- The following files are now included with the output
    - README.txt (Includes all citations cited in Flows)
    - USGS Disclaimer.txt
    - RelateBatchProcessorOutput.pdf
- Rounding for Non AK **Drainage Areas** follows https://water.usgs.gov/admin/memo/SW/sw12.07.html
- Rounding for AK drainage areas follows
  USGS Alaska Science Center drainage area rounding procedures
  Last updated 11/12/15 by Janet Curran

  Rounding example:

		OSW memo 12.07 guidelines			Alaska drainage area precision		
raw drainage area (sq mi)	range (sq mi) (equal to or greater than lower bound, less than upper bound)	rounded value (sq mi)	place rounded to	effective significant figures	rounded value (sq mi)	place rounded to	effective significant figures
0.456789	0-1	0.46	0.01	2	0.5	0.1	1
4.56789	1-10	4.57	0.01	3	4.6	0.1	2
45.6789	10-100	45.7	0.1	3	45.7	0.1	3
456.789	100-1000	457	1	3	457	1	3
4567.89	1000-10000	4568	1	4	4570	10	3
45678.9	10000-100000	45679	1	5	45680	10	4
456789	>100,000	456789	1	6	456800	100	4

- Rounding for flows (same as version 4.5)
    - As per email 2/1/2018
    - Q<10 cfs, round to nearest 0.01 cfs
    - 10<Q<100 cfs, round to nearest 0.1 cfs
    - 100 cfs < Q <1000 cfs, round to nearest 1 cfs
    - 1,000 cfs < Q <10,000 cfs, round to nearest 10 cfs
    - 10,000 cfs < Q <100,000 cfs, round to nearest 100 cfs
    - Or in other words, use 3 significant digits but don't go beyond 0.01 cfs.
    - Also, we should be rounding to an even on the half. For example:
        - 12.35 rounds to 12.4
        - 12.45 rounds to 12.4

- Rounding for other Basin Characteristics use the number of decimals defined in the XML, and python rounding=decimal.ROUND_HALF_EVEN to round to the nearest even when the reminder is exactly .5
- Basin Characteristics going into NSS are rounded using the above rounding rules

Known Bugs:

- Shapefile output option is not available
- Option to select output projection is not considered
- Schema mismatches showing up - looking into generating schema on the fly
- Need to add PC (Percent Complete) and Disclaimer where appropriate
- ss_bp will allow user to select Regional Study Areas. However, computations for CRB and DRB are not performed – code does not go to a particular state to perform the delineation
- Latitude, Longitude GlobalWatershedPoint are not in the correct coordinate system

Version 4.5 12/14/2018

Changes (Only calculations on the back end – bp_job – were changed)

- Rounding consistent with interactive application and WRI pub 01-4044

https://pubs.er.usgs.gov/djvu/WRI/wrir_01_4044.pdf

- Fix Basic Characteristic failures
- SQL connection timeout addressed
- General error reporting improvements

Known Bugs:

- Need to add PC (Percent Complete) and Disclaimer where appropriate
- Need to add full citation
- More rounding refinements needed
- Remove unused fields from output globalwatershed

Version 4.4 11/28/2018

Changes (Only Flow Statistic calculations on the back end – bp_job – were changed)

- Fixed to only output results for selected Flow Types
- Fixed to only consider regression regions for selected Flow Types
- Fixed weighting to return values that add up to 100% for study areas
- Some performance issues fixed
- All fields in CHARACTERISTICS and FLOWSTATS tables changed to TEXT to allow differentiation between NULL and Zero
- CHARACTERISTICS table: fields modified
    - RegionName added
    - AreaPercent added
    - AreaSqMi added
    - Label changed to StatLabel
    - StatName added
    - Error removed
    - Min changed to MinLimit
    - Max changed to MaxLimit
- FLOWSTATS table fields modified
    - RegionName added
    - PercentWeight changed to AreaPercent
    - AreaAverage removed
    - AreaSqMi added
    - Error removed
    - PIl added
    - PIu added
    - SEp added
    - SE added
    - CitationID added

Known Bugs:

- README file for field definitions, citations, USGS disclaimer not completed
- Shapefile output option is not available
- Option to select output projection is not considered

Version 4.3 – 10/29/2018

Changes

- Error reporting should be better
- Returns partial results when some points fail
- Flow Types are now displayed for only those types where there are basin characteristics available within the Study Area. Because NSS is Regression Region based, and the batch processor is Study Area based, there may be points within uploaded file (Study Area based) where a particular selected Flow Type is not available.

Known Bugs:

- ss_bp will allow user to select Regional Study Areas. However, computations for CRB and DRB are not performed – code does not go to a particular state to perform the delineation
- Weighting for flow statistics in Kentucky and possibly other states returns values > than 100%
- It is still difficult to know for sure whether a value is truly zero or the result of a characteristic not being computed.
- Flow Stats performance issues identified but not fixed
- All Flow Types computed regardless of user selection
- Shapefile output option is not available
- Option to select output projection is not considered

Version 4.2 - 10/17/2018

The batch processing tool will accept up to 200 points on a State or Regional Study basis.

To run: https://streamstatsags.cr.usgs.gov/ss_bp/

The input shapefile is a **point** file of pour points. It is best to first download the stream grids at: https://streamstatsags.cr.usgs.gov/StreamGrids/ and be sure that your points fall near (generally no more than a couple grid cells from) the streams in this grid. If a state stream grid is missing, let us know by emailing support@streamstats.freshdesk.com

The “Local ID Field” should be a field in the input shapefile that contains a unique identifier for the site/delineation. The contents of this field are carried over into the name field of the resulting datasets and may be used to relate back to the input shapefile.

Enhancements:

- Will accept Regional Study data as well as State data
- Those Basin Characteristics required for Flow Statistics are automatically selected
- Problems seen in previous versions due to instability have been considered and hopefully resolved.

Known Bugs:

- Error reporting is cryptic
- Weighting for flow statistics in Kentucky and possibly other states returns values > than 100%
- It is still difficult to know for sure whether a value is truly zero or the result of a characteristic not being computed.

Version 4.1

Internal changes only.

Version 4.0

We have been having a lot of problems with the batch processor. When it computes things, it seems to do them right, but for reasons we have yet to figure out, it often drops out some results. Also, sometimes it returns partial results with either null or zero values for some parameters, but valid results for others. If you re-submit sites that did not give complete results, you can generally get complete results, although it often takes several iterations. It also can be difficult to distinguish true zeros from dropped results for parameters that can actually be zero. The records near the end of the result datasets tend to be the ones with partial results.

At this time, our only option with the batch processor is to submit jobs, inspect the results, select the sites that did not get complete results, write those out to a new shapefile, and submit that one. Once you get all the results you need, you can use the ArcToolbox Append tool to append the results to a single output database. We have gotten very accustomed to doing this, and it works OK, but sometimes takes quite a few iterations. This may be pretty tedious, but that's what we have to work with right now. Also, please do not submit multiple jobs at once. Please wait till one job finishes before submitting another. On weekends or overnight, it is OK to submit a couple jobs at one time, though. If you need to do more than this, please email us to discuss options.

For a long time we felt like we couldn't trust the results, but we've tested enough that we are convinced the results are reliable, if you get them. The only thing that still is really difficult is distinguishing true zero values from dropped results. When in doubt, remove the results and run the site again.

The other thing you will need to be aware of is the snapping to stream grid cells. You will need to check carefully to see that the basins delineated for each site are actually the intended basins. If you get a tiny

watershed that does not look right, for example if it is a single grid cell, you may be able to move your point a little to fall on the stream grid.

Finally, although we try very hard to get it right, the software and data do contain errors. It is important to remember that the user bears responsibility for checking the watershed delineations and basin characteristic computations. If you ever see results that appear to be in error or do not make sense, please report it to us so we can check and see if something needs to be fixed.