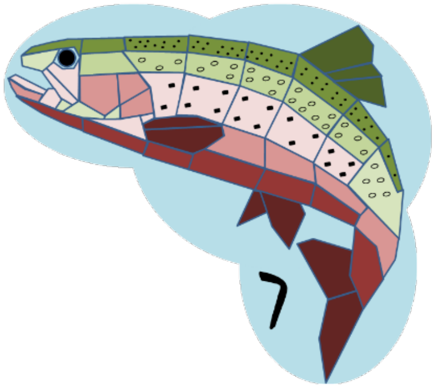
InSTREAM 7.x Cell Merger Tool:
Description and Application Guide



Prepared by:

Steven F. Railsback
Lang Railsback & Associates
Arcata, CA

Last updated: April 24, 2026

Revision Log

Date	**Section**	**Revision**	**Author**
April 24, 2026	All	Complete first version	SFR

Contents

1 Purpose of the Cell Merger Tool

The InSTREAM/InSALMO cell merger tool is intended to help users reduce the number of habitat cells and, therefore, model execution times, by combining cells that are sufficiently similar. Section 22.1 of the InSTREAM User Manual[1] discusses methods for delineating habitat cells. Several of those methods can result in larger numbers of smaller cells than necessary, and the time it takes for InSTREAM or InSALMO to execute increases sharply with the number of cells. Especially, using the mesh of a fine-resolution hydraulic model as InSTREAM cells, and the hexagonal cell generation approach of Dudley (2018; described in the User Manual), are likely to produce many small cells.

This tool is free and open-source, and available on the InSTREAM 7 web site: www.humboldt.edu/ecological-modeling/instream-and-insalmo/instream-7-and-insalmo-7. It consists of a NetLogo file named something like `CellMerger.nlogox` that, like InSTREAM versions 7.4 and higher, works in NetLogo version 7.0.3 and higher, and a separate parameter file named something like `parameters-CellMerging.nls`. The tool also requires the GIS shapefile and depth and velocity input files for the InSTREAM or InSALMO application it is used for.

2 How the Tool Works

The cell merger tool's general approach is to identify adjacent cells that are sufficiently similar to each other in habitat characteristics, and then to merge those cells into one. Three steps are used; the first two are conducted by the tool's NetLogo program when its `merge-cells` procedure is executed by clicking on the Interface button so labeled. The following subsections 2.1 to 2.3 describe these three steps in detail.

First, all model cells, in randomized order, compare their habitat variable values to those of all their adjacent cells (except for adjacent cells that have already been tagged for merger with another cell). The cell making this comparison is referred to as the *root cell*, and a cell being evaluated for merging is referred to as an *evaluation cell*. All the original model cells are, one at a time, treated as the root cell that evaluates whether it can merge with any of its adjacent cells; however, that evaluation is skipped if the root cell has already been tagged to merge with another cell.

Each time an evaluation cell meets all the criteria for merging with root cell, then:

1. The evaluation cell is tagged to be merged with the root cell by setting its cell identification code (cell variable *cell-ID*) to that of the root cell, and
2. The evaluation cell then recursively examines all *its* adjacent cells to see if they are sufficiently similar to the root cell to also be merged. This recursion continues until there are no more adjacent cells that can be merged.

---

[1] Railsback, S. F., B. C. Harvey, and D. Ayllón. 2023. InSTREAM 7 user manual: model description, software guide, and application guide. PSW-GTR-276, USDA Forest Service, Pacific Southwest Research Station, Albany, California. https://doi.org/10.2737/PSW-GTR-276. The most recent version is available at: https://www.humboldt.edu/ecological-modeling/instream-and-insalmo/instream-7-and-insalmo-7.

Because the order in which cells are treated as root cells is randomized, the cell merger tool can produce different results each time it is used. However, the user can optionally set a random number generator seed that causes cells to be treated as root cells in the same order each time.

Second, the NetLogo program creates a new GIS shapefile that still contains the polygons of the original unmerged shapefile, but with new habitat variable values assigned to them. All polygons to be merged together are assigned the same values of all the habitat variables that are stored in the shapefile. This shapefile is written in the same file directory as the cell merger NetLogo file and consists of three files named `MergedCells.dbf`, `MergedCells.shp`, and `MergedCells.shx`. The NetLogo merger code also produces new hydraulic input files that contain depth and velocity lookup tables for the new merged cells.

Third, the user creates a new shapefile with merged cells, in GIS software. This step is done manually in software such as ArcGIS and QGIS, by applying the "dissolve" function to the `MergedCells` shapefile. This third shapefile and the new hydraulic input files are then used in subsequent model runs.

2.1 Merging criteria

The following criteria are used to determine whether a cell (the "evaluation cell") should be merged with a root cell. All but three of the criteria are controlled by parameters that allow the user to adjust how similar cells must be to merge. The criteria consider all the cell habitat variables that are set by model input.

Several of the criteria are based on the "fractional difference in value" between a root cell and an evaluation cell. This fractional difference is equivalent to percent difference except that: (a) it is not multiplied by 100, and (b) the absolute value is used so it is always positive. For a variable X with values of X_r and X_e for the root and evaluation cells, the fractional difference in value between cells is $abs\{(X_r - X_e)/X_r\}$.

Other criteria are based on the "difference in value", which is the absolute value of the root cell's value minus the evaluation cell's value. The "difference in value" is: $abs\{X_r - X_e\}$.

2.1.1 Not already merged

An adjacent cell is not considered for merging with a root cell if it is already tagged for merging with a different root cell. Likewise, as the code considers each cell as a potential root cell, it skips any cells that have already been tagged for merging with another root cell.

2.1.2 Same reach

Cells cannot be merged if they do not belong to the same stream reach.

2.1.3 Maximum cell size

Evaluation cells cannot be merged with a root cell if the resulting new cell (including the evaluation cells already merged with the root cell) would have an area exceeding the parameter ***max-merged-cell-area*** (m^2). This criterion allows the user to limit how big merged cells are.

The value of this parameter normally should be higher for larger streams. However, its value can have little effect when the other merger criteria limit merged cell size. Users can initially use a high value and reduce it if needed to avoid excessively large cells.

2.1.4 Centroid inside cell

Evaluation cells cannot be merged if the resulting new cell (including evaluation cells already merged) would have its centroid outside the boundaries of the new cell. This criterion is intended to avoid oddly shaped cells that would be poorly represented by InSTREAM's use of a cell's centroid as its location when calculating distances between cells.

2.1.5 Distance to escape cover

Two criteria are used to determine whether cells have sufficiently similar values for distance to escape cover. Two separate criteria are needed to avoid numerical issues (e.g., division by zero) when distance to escape cover is near zero and because predation risk in InSTREAM typically is relatively sensitive to low values of this distance but less sensitive to high values.

The first criterion is used when distance to escape cover has low values. If either the root cell or the evaluation cell has a value ≤ 50 cm, then the cells can be merged only if their difference in values is less than the parameter ***dist-escape-tol-mag*** (cm). Because of InSTREAM's sensitivity to small values of distance to escape cover, useful values of this parameter should be small, such as 50 cm or less.

Otherwise (both cells have values > 50 cm), the cells can be merged only if the fractional difference between values is less than the parameter ***dist-escape-tol-frac***. Because InSTREAM is less sensitive to higher values of distance to escape cover, this parameter could have values up to ~0.5 (50% difference between root and evaluation cell).

2.1.6 Velocity shelter availability

Two cells can be merged only if their difference in values of fraction of area providing velocity shelter is less than the parameter ***frac-vel-shelter-tol***. InSTREAM's results can be quite sensitive to velocity shelter availability, so this parameter should have a relatively low value, e.g., ≤ 0.3.

2.1.7 Spawning gravel availability

Two cells can be merged only if their difference in values of spawning gravel availability (as fraction of cell providing gravel) is less than the parameter ***frac-spawn-tol***. Model results are typically not especially sensitive to spawning gravel availability, so values of ***frac-spawn-tol*** somewhat greater than 0.3 could be acceptable.

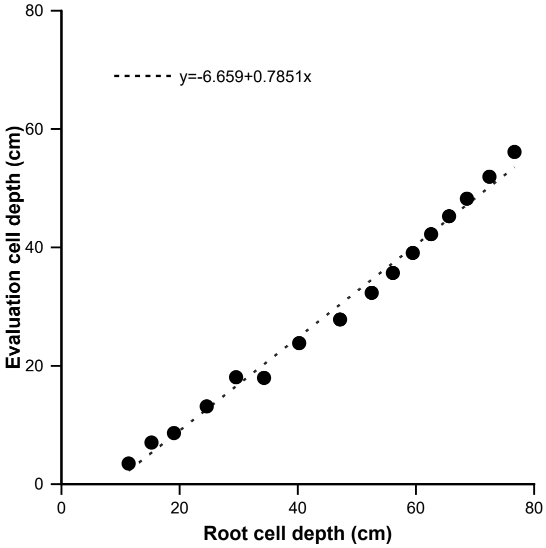
2.1.8 Hiding place availability

This criterion compares cell values of the variable for (integer) number of hiding places. To accommodate cells of different sizes, the comparison uses density of hiding places, evaluated as number of hiding places divided by cell area. A cell can be merged with a root cell only if the difference in values between cells in hiding place density is less than the parameter ***hiding-place-density-tol*** (which has units of places per m^2).

To set a value of this parameter, it is useful to think about how many hiding places it represents in an area such as 10 m^2. A value of 0.1 translates to ±1 hiding places per 10 m^2.

2.1.9 Depth—flow relationship

Cells can be merged only if they have similar depths and velocities across all simulated flows. Three criteria are used to evaluate depth similarity. These criteria are based on linear regression of the model's hydraulic lookup table input. That input specifies cell depth and (separately, considered below) velocity at each of a number of flows. Because the same flows are used for all cells, we can regress the depths of two cells against each other. If two cells have similar depths at all flows, the regression relation between their depth input will have a slope near 1.0, a Y intercept near 0.0 cm, and a squared regression coefficient R^2 near 1.0. The following figure illustrates this regression: the points are the depths (cm) at (X) the root cell and (Y) the evaluation cell, at each of the flows in the depth input file, which in this case ranged from 8.0 to 70 m^3/s. (InSTREAM uses depth and velocity input in units of m and m/s, but converts the input to cm and cm/s. Hence, all values considered here are in cm.)



The similarity between two cells in depth—flow relationship is evaluated using these steps:

1. If all depth values, over all flows, are zero for both cells, then depth does not affect whether the cells can be merged; therefore, depths for the two cells are assumed perfectly correlated by setting slope to 1.0, intercept to 0.0, and R^2 to 1.0.
2. If instead only one of the cells has all depths equal to zero, then depths of the two cells are assumed completely uncorrelated, with slope, intercept, and R^2 all set to 0.0.
3. If both cells have depth equal to zero at one or more of the lowest flows, those zeros are removed from the regression input.
4. If the resulting regression input has only two flows with non-zero depths, then the slope and intercept are calculated directly from the line defined by those two flows. The slope and intercept are calculated using the root cell for the X axis and the evaluation cell as the Y axis. However, if either cell has the same value for both flows, then the cell depths are assumed completely uncorrelated.
5. Otherwise (three or more flows with non-zero depths for at least one of the two cells), least-squares regression is used to calculate the slope, intercept (cm), and R^2.
6. The cells can merge only if:

a. the absolute value of (1.0 – slope) is less than the parameter ***depth-slope-tol-frac***, and
b. the absolute value of the intercept is less than the parameter ***depth-int-tol-mag*** (cm), and
c. the value of R^2 exceeds the parameter ***depth-min-R2***.

The value of ***depth-slope-tol-frac*** can be considered as the average fractional difference in depth between cells across all flows; a value of 0.2 means that cells are similar enough if depths are within 20% of each other. The value of ***depth-int-tol-mag*** can be considered the difference between cells in depth (cm) at near-zero flows; having a low value may be more important for sites where flows can be near zero. The value of ***depth-min-R2*** should be between ~0.7 and ~0.9; however, if the other two criteria are met for adjacent cells, their depths are likely to be closely correlated so this parameter may have little effect.

2.1.10 Velocity—flow relationship

The velocity—flow relationship is compared between cells exactly as the depth—flow relationship is, except that the three criteria parameters are: ***vel-slope-tol-frac***, ***vel-int-tol-mag*** (cm/s), and ***vel-min-R2***.

2.2 Evaluation of habitat variables for merged cells

Once a root cell has identified the other cells it can merge with, it calculates the values of the habitat variables that the combined cell will have, using the following methods. These variables are all polygon properties of the GIS shapefile that defines the habitat cells.

The combined cell is given the cell identifier (cell variable *cell-ID*) of the root cell.

Cell area is set to the total area of all the merged cells.

The values of distance to escape cover, velocity shelter availability, and spawning gravel availability are all set to the area-weighted mean of all the merged cells.

The number of hiding places in the combined cell is set to the sum of the number of places in all the merged cells.

The relations between flow and depth, and flow and velocity, for the combined cell are averages of those for its constituent cells. The cell merger's NetLogo code writes new depth and velocity input files for use with the new cells. Those files are named `MergedDepthInput.csv` and `MergedVelInput.csv`. They each contain one row for each of the merged cells, identified via the root cell's value of *cell-ID*, plus rows for the cells that were not merged with any others. For each of the flows in the original hydraulic input files, the depth or velocity of the merged cell is calculated as the area-weighted mean of depth/velocity in the merged cells.

2.3 Creation of a new shapefile

The cell merger tool writes a new shapefile, named `MergedCells`, that contains all the original cell polygons, but the polygons to be merged have new values of their properties. All the polygons that are to be merged to create one cell have the same property values, those determined using the methods described in Sect. 2.2. The user must then convert the

`MergedCells` shapefile into a new shapefile in which the polygons of merged cells are actually joined into new, larger polygons.

Users should note that NetLogo's GIS extension, which is used to create the `MergedCells` shapefile, may modify polygon property names by dropping linking characters such as "_" and "-". This behavior occurs in NetLogo versions at least through 7.0.3.

A new shapefile with merged cell polygons can be created via these steps:

1. Open the `MergedCells` shapefile in GIS software.
2. Use the GIS's "dissolve" function or tool to create a new layer that combines polygons having the same value of the cell identification code property. Use options that set the property values of the combined polygons to the values of the first of the merged polygons. This method of setting property values appears to be the default in QGIS. In ArcGIS, this method is selected by setting the statistics field to `FIRST`; note that this changes the property value names.
3. Save the resulting layer as a new shapefile.

2.4 The test output file

The cell merger's NetLogo program has a switch on its interface labeled `merge-test-output?`. If that switch is set to TRUE then each execution of the `merge-cells` procedure writes a file of output that can be used to test and understand the merging process. That test file is always named `MergeTestOut.csv` and it is overwritten each time `merge-cells` is executed with `merge-test-output?` set to TRUE.

The test output file has two sections. The first section contains one row for each time a root cell considers an evaluation cell for merger. It reports the habitat variables of both cells, and whether they can be merged. If they cannot be merged, the row reports why: the first of the criteria in Sect. 2.1 that was not met.

The second section appears after the first; it starts with the first blank row in the file. (In Excel, the start of the second section can be found by selecting cell A1, then typing Control-down arrow.) That section contains a block of rows (separated by blank rows) for each group of cells that can be merged into a new, larger, cell. Each row reports the variables of one of the cells that was merged and the variables of the merged cell.

3 Recommended Application Procedure

The cell merger tool is implemented as a NetLogo file that uses the same habitat and hydraulic input as InSTREAM and InSALMO; this file is normally named something like `CellMerger.nlogox`. The following steps are recommended for using it.

3.1 Set file names and locations

The first step is to set the parameter values that tell the cell merger code where the InSTREAM shapefile and hydraulic input files are, and that provide the shapefile's property names for cell habitat variables. Sect. 14.1.2 of the InSTREAM User Manual explains how to do this, by opening the cell merger’s NetLogo file, clicking on its “Code” tab, and (if necessary) clicking on the “Included files” button to open a tab with the parameter file.

3.2 Set preliminary parameter values

Preliminary values for the parameters that control which cells are merged need to be reviewed and potentially changed. The values of these parameters are set via “sliders” on the Interface tab of the NetLogo code (illustrated below); parameter values can be changed by moving the sliders or by typing values into the slider’s value box.

max-merged-cell-area 100 m2
dist-escape-tol-mag 50 cm
dist-escape-tol-frac 0.20
frac-vel-shelter-tol 0.20
frac-spawn-tol 0.20
hiding-place-density-tol 0.20 /m2
vel-slope-tol-frac 0.20
vel-int-tol-mag 5 cm/s
vel-min-R2 0.80
depth-slope-tol-frac 0.20
depth-int-tol-mag 10 cm
depth-min-R2 0.80

3.3 Conduct a test merge

Next it is useful to execute a test cell merge using these steps:

- Click on the “setup” button on the Interface, and wait until it finishes executing (the button’s color reverts to its original color). After setup, the NetLogo patches (square grid cells) in each habitat cell are set to a random color.
- Click on the “merge-cells” button. Output written to the Command Center says how many cells can be merged. The patches underneath merged cells are given the same color, so you can see the approximate shape and size of the resulting cells.
- Repeating the “setup” and “merge” steps several times to get an idea of how much stochastic variation there is in the process. (Random variation among merge executions

occurs when the Interface input widget labeled `rand-num-seed` has a value of 0; if it instead has a positive value, each merger will process cells in the same order so results will be the same if no parameter values were changed. Changing the value of `rand-num-seed` results in cells being processed in a different order.)

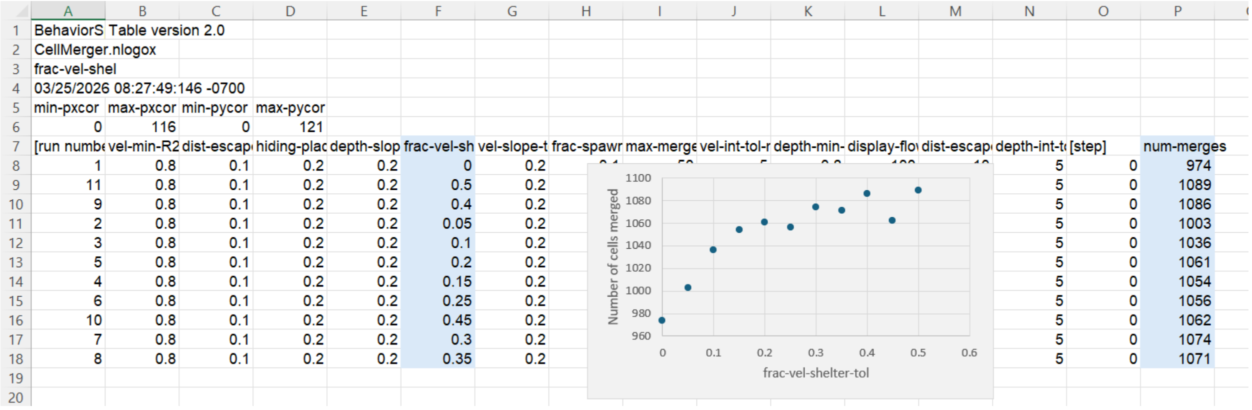
- Optionally: if the switch labeled "merge-test-output?" on the Interface is turned on, the "merge-cells" procedure will write an output file `MergeTestOut.csv` that reports, for each time a root cell considers an evaluation cell, the variable values of both cells and which (if any) criterion prevents merger.

3.4 Analyze sensitivity to merger parameters

Analyzing how sensitive the merge results are to each of the cell merger's parameters is recommended because parameter sensitivity is likely to vary strongly among study sites, and understanding it helps select parameter values that allow more cells to be merged without unnecessarily reducing habitat variability. The merger's NetLogo file includes 12 BehaviorSpace experiments (see Sect. 15 of the InSTREAM User Manual) that each can be used to see how the number of merged cells varies with the value of one parameter, while the other parameters are held at the values on their Interface slider.

When running these sensitivity experiments, it is important to make sure that all the parameter sliders are set to useful values. Normally they should all be set to "standard" values that the user expects to use for the final merger unless the sensitivity experiment indicates that other values might be better; doing so indicates how strongly the parameter being analyzed will affect the final merger. However, it could also be useful to set all the parameters to extremely high values (except the depth and velocity R^2 parameters, which would be set to 0.0); this kind of experiment indicates how strongly the analyzed parameter affects mergers in the absence of other criteria.

Each experiment, when executed, produces an output file that can be opened in spreadsheet software for analysis. Analysis typically is as simple as inserting an XY plot of how the number of merged cells (*num-merges*) varies with the parameter value. In the example below, the parameter ***frac-vel-shelter-tol*** was varied from 0.0 to 0.5 in steps of 0.05; the plot indicates that the number of merged cells rises sharply as this parameter increases from 0.0 to 0.2, but increases little more at higher values.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	BehaviorS	Table version 2.0														
2	CellMerger.nlogox															
3	frac-vel-shel															
4	03/25/2026 08:27:49:146 -0700															
5	min-pxcor	max-pxcor	min-pycor	max-pycor												
6	0	116	0	121												
7	[run numbe	vel-min-R2	dist-escap	hiding-plac	depth-slop	frac-vel-sh	vel-slope-t	frac-spawr	max-merge	vel-int-tol-r	depth-min-	display-flo	dist-escap	depth-int-t	[step]	num-merges
8	1	0.8	0.1	0.2	0.2	0	0.2							5	0	974
9	11	0.8	0.1	0.2	0.2	0.5	0.2							5	0	1089
10	9	0.8	0.1	0.2	0.2	0.4	0.2							5	0	1086
11	2	0.8	0.1	0.2	0.2	0.05	0.2							5	0	1003
12	3	0.8	0.1	0.2	0.2	0.1	0.2							5	0	1036
13	5	0.8	0.1	0.2	0.2	0.2	0.2							5	0	1061
14	4	0.8	0.1	0.2	0.2	0.15	0.2							5	0	1054
15	6	0.8	0.1	0.2	0.2	0.25	0.2							5	0	1056
16	10	0.8	0.1	0.2	0.2	0.45	0.2							5	0	1062
17	7	0.8	0.1	0.2	0.2	0.3	0.2							5	0	1074
18	8	0.8	0.1	0.2	0.2	0.35	0.2							5	0	1071
19																
20																

When reviewing these sensitivity analysis graphs, pay careful attention to the Y axis scale. If the axis scale is adjusted automatically (as Excel does by default), parameters with strong and weak effects on the number of cells merged can produce similar-looking graphs. In the above example, it looks like low values of ***frac-vel-shelter-tol*** strongly reduce the number of mergers, but the actual effect (e.g., 1060 to 975 mergers as the parameter decreases from 0.2 to 0.0) is quite small.

3.5 Select and use final parameter values

Next, the sensitivity analysis results are used to select final parameter values, and those values are used in a final application of the cell merger. Save the resulting shapefile (`MergedCells`).

However, several test merges (varying the parameters most strongly affecting how many cells are merged) may be needed to find a good compromise between the resulting number of cells and the similarity of merged cells. Simply re-click the "setup" button, adjust parameter values, and re-click the "merge-cells" button. The new shapefile `MergedCells` and the new depth and velocity input files `MergedDepthInput.csv` and `MergedVelInput.csv` (and the merge test output file, if used) will be re-written each time.

3.6 Create a new GIS shapefile

The `MergedCells` shapefile created by the cell merger must be processed in GIS to create a final shapefile with merged cell polygons. Use the methods described at Sect. 2.3.

3.7 Assemble a new InSTREAM/InSALMO project

The final step is to assemble a new InSTREAM or InSALMO project that uses the new shapefile and new hydraulic input files. A "project" is a file directory tree containing the model code, input files, and output files (see Sect. 11 of the InSTREAM User Manual). Modifying an existing project to use the merged cells includes the following steps:

- Replace the shapefile with the new one containing merged cells.
- Rename the new hydraulic input files `MergedDepthInput.csv` and `MergedVelInput.csv` so the names are meaningful for the project, and replace the previous depth and velocity input files with them.
- Edit the InSTREAM parameter file to update the names of the shapefile (parameter ***GIS-file-name***) and its new property names (parameters ***GIS-property-for-cell-ID*** etc.); and the new hydraulic input files (parameters ***depth-file-names*** and ***velocity-file-names***).