**CACIE Tool #NN** – ***HSSM Builder (build\_hssm.py)***

**Version** **1.1**

**QA**: **QA**

1. **Description and Purpose**

The main purpose of the HSSM Builder tool is to build a Hydrocarbon Spill Screening Model (HSSM) package for MT3D modeling purposes. It processes a single mass flux time series data file or multiple mass flux time series data files by first combining them into a single package. The HSSM Builder tool then shifts mass out of cells that no longer have a saturated layer in the MT3D model to adjacent cells that have a saturated layer and reduces the number of time steps associated with each MT3D model cell. The shifting of mass is consistent with the aqueous flow direction as defined by the MT3D model.

1. **Functional Requirements**

The following are the functional requirements of the HSSM Builder tool.

FR-1: read and process the following MT3D files and save processed files as \*.pkl files for use during subsequent processing of additional data files:

* MT3D Head file (\*.hds),
* Model top definition (\*.ref),
* Layer bottom definitions (\*.ref)
* Layer initial condition files (\*.ref).

FR-2: Identify the top layer where the percent saturation is greater than or equal to user-identified saturation threshold for each time period from MT3D HDS and layer files.

FR-3: Ability to reuse saturated data processed as part of FR-2 in future runs of the application without having to reprocess the MT3D data to rebuild it.

FR-4: User-definable start and end years for the output data (note: HSSM Builder tool converts years to days (365.25 days per year) starting from 0 for generated output files)

FR-5: User-definable flux floor (noise) threshold value

FR-6: User-definable maximum acceptable total mass error

FR-7: Read in and process a single file or multiple data files containing mass flux time series data (multiple files must be the same COPC but can represent different STOMP model domains). Note that the time steps must be the same in each file or it can throw off mass shift and cell summing calculations.

FR-8: Assign a value of “0” (zero) to mass flux values that are less than the user-specified flux floor (noise) threshold (FR-5) for each cell.

FR-9: When processing multiple mass flux time series data files, sum the mass flux of overlapping cells together to form a single time series for each cell. The overlapping cells must have the same time steps in order to be summed together.

FR-10: If a cell is dry (i.e. all cell layers drop below specified saturation) or inactive (i.e. cell water level being <= 0 or layer elevation <= 0), move remaining time series out of cell following the modeled aqueous flow from the cell to adjacent cell(s).

FR-11: Split a cell’s time series between cell layers depending on at which timesteps a layer meets or exceeds the minimum percent saturation threshold, using the highest layer that meets or exceeds the saturation threshold.

FR-12: Reduce the number of timesteps associated with each cell timeseries to within a target range of greater than 50 and less than 200, using the Ramer-Douglas-Peucker algorithm. It is noted that the processing logic prioritizes meeting the acceptable error tolerance between the reduced and original datasets. As a result, the number of datapoints can be less than or greater than the target range of reduced timesteps.

FR-13: Build HSSM package input for MT3D

FR-14: If cell is dry or inactive in day 0, then check if there is at least 10% saturation in any layer. If there is 10% saturation or more, then use it as the saturated layer for day 0 and mass shift starting in day 1. If saturation for all layers is below 10%, then mass shift starting in day 0.

FR-15: Remove those cells that have no mass associated with them in original dataset (i.e. cells with mass = 0 for every timestep in modeling time series in original dataset)

FR-16: optional Function to create stepwise output. This would insert a duplicate day between each step that has the same rate as the previous year. For example:

Day Rate

0 5

365.25 5

365.25 12

730.5 12

73.5 10

Etc.

1. **Software Requirements Specifications**

Programming Language and required modules/libraries:

* Python 3.6
* Numpy 1.16
* Pandas 0.24.2 (incompatible with other versions)
* Scipy 1.4.1
* Flopy 3.3
* Matplotlib 3.1.2

1. **Software Design Description**

Required inputs/generated outputs:

* Inputs:
  + MT3D files:
    - Data file (\*.hds) containing water levels for each cell
    - Data file (top1.ref) containing elevation of the top of the model for each cell
    - Data files (bot\*.ref) containing elevation of the bottom of each layer for each cell
    - Data files (sh\*.ref) containing the initial conditions for the model layers per cell.
  + Pickle files:
    - Data files (\*.pkl) containing preprocessed MT3D data (generated using data from MT3D files listed above).
      * Saturation levels at the final timestep for each cell ([MT3D\_model\_version]\_sat.pkl)
      * Directional mass flow between cells ([MT3D\_model\_version]\_flow.pkl)
      * Layer saturation levels by year for each cell over entire time period ([MT3D\_model\_version]\_yearly\_sat.pkl)

Once generated, these files can be reused instead of having to reprocess the MT3D files between different HSSM builder executions that use the same MT3D model.

* + Mass Flux Time series data
    - Typically STOMP data files (\*.csv) for an individual COPC by model domain
  + Config file
    - JSON-formatted file containing the following user-defined input parameters:
      * isPickled (Boolean): true: \*.pkl files exist; false: \*.pkl files need to be processed.
      * pickleDir (string): directory location of the \*.pkl files
      * satFile (string): name of the \*.pkl file containing the saturation levels at the final timestep for each cell
      * flowFile (string): name of \*.pkl file containing the directional mass flow between cells
      * yearlySatFile (string): name of \*.pkl file containing the layer saturation levels by year for each cell
      * hds (string): path/name of MT3D hds file
      * top\_ref (string): path/name of the file containing elevation of the top of the model for each cell
      * bot\_ref (string): array of file path/names containing elevation of the bottom of each layer of the model for each cell
      * hds\_init\_conditions (string): array of file path/names containing MT3D initial conditions
      * input (string): directory location for COPC mass time series files (all files need to be in the same time steps or any overlapping cells will not sum properly, and mass shifting out of cells will not merge with the new cell location properly)
      * output (string): directory to build HSSM package and verification files in; defaults to “/output” if not provided.
      * find\_sat\_layer\_by\_year (Boolean): true: split cells between layers as saturated layer changes; false: use the layer that is saturated for the entire model.
      * sat\_lvl (float): the minimum percent saturation value (percentage in decimal format, i.e., enter 0.75 to represent 75% saturation) for a layer to be considered saturated (i.e. not dry); defaults to 0.75 if not provided
      * max\_i (integer): maximum rows in the MT3D model grid
      * max\_j (integer): maximum columns in MT3D model grid
      * max\_k (integer): maximum layers in MT3D model grid
      * start\_year (integer): first year to start pulling data from inputs
      * end\_year (integer): last year to pull data from inputs
      * tolerance (float): maximum acceptable relative total mass error (percentage, i.e., enter 0.1 to represent 0.1% total mass error); defaults to 1E-02 if not provided
      * mass\_shift (Boolean): true: shift mass out of cells that go dry (no layers >= sat\_lvl) into other cells; false: leave mass in dry cells; defaults to false if not provided
      * data\_reduction (Boolean): true: reduce number of time steps in data; false: retain all times steps in data; defaults to true if not provided
      * flux\_floor (float): threshold value below which flux or mass values are considered to be equal to zero (i.e., below noise threshold); defaults to 1E-15 if not provided. Anything below this may be dropped out of model
      * max\_tm\_error (float): maximum acceptable total mass error after data reduction; defaults to 2.25E+15
      * HSSpath (string): extended path for file names in hss config file (mt3d.hss); defaults to “/hss” if not provided,
      * copc: name of COPC being processed
      * graph\_name (string): prefix for file names of generated graphs; e.g. “tc-99\_graphs” will equate to “tc-99\_graphs\_x-y(kz).png where x, y and z resolve to the i, j and cell layer address in the MT3D grid.
      * units (string): units applicable to data being processed; choice of pCi (radionuclides) or µg (chemicals); default is pCi if not provided
      * stepwise (bool): true: create stepwise version of the package as well as the normal package. Defaults to false: only make the normal non-stepwise package.
* Outputs:
  + Log files (\*.log) documenting different stages of building the HSSM package
  + \*.dat files for each cell, and for some cases, multiple layers of a cell (i{xx}j{yy}k{z}.dat)
  + Mt3d.hss file, config file for mt3d data.
  + PNG files with plots of the original and reduced datasets (flux and cumulative mass) versus time
  + CSV files, used for verification and recording data processing steps:
    - 01\_all\_cells\_after\_cell\_merge.csv
      * Snapshot of data after all inputs have been loaded and overlapping cells have been merged (i.e., mass flux summed)
    - 02\_all\_cell\_by\_day.csv
      * Snapshot of data after the mass flux has been converted from yearly to daily rates
    - 03\_all\_cell\_by\_day\_dry\_cell\_shifted.csv
      * Snapshot of data after mass has been shifted out of all dry cells
    - cell\_error.csv
      * Total error by cell after reduction
    - cell\_error\_by\_layer.csv
      * total error by cell layer after reduction
    - cell\_map.csv
      * what input files contributed to each cell
    - cumulative\_data\_set\_by\_year.csv
      * cumulative data for debugging
    - dry\_cells.csv
      * list of the cells that went dry and the cells that received the shifted mass from the dry cells
    - dry\_cell\_flux\_shift\_itteration\_{step number}.csv
      * data after it has been shifted to a new cell but before it has been merged with the existing data (there may be multiple iterations for the same grid address). There will be one file for each iteration, as it will iterate through until all mass has been shifted out of dry cells.
    - flux\_mass\_shift\_mapping.csv
      * shows how much mass moved from one cell to another
    - full\_data\_set.csv
      * all \*.dat files combined into a single csv
    - interpolated\_results\_cell.csv
      * each cell interpolated back into full data for comparison against the original full data set
    - interpolated\_results\_cell\_by\_layer.csv
      * each cell layer interpolated back into full data for comparison against the original full data set
    - rate\_error\_check.csv
      * shows where the absolute relative percent error of the reduced flux versus the original flux is greater than 0.1% for same time step
    - saturation.csv
      * Final saturated layer for each active cell
    - yearly\_saturation.csv
      * Saturated layers of each active cell through time

Tool Runner:

The following is the shell script configuration that will be passed as an argument to the Tool Runner for qualified runs:

pylib/hssmbuilder/build\_hssm.py {config.json}

Code Review:

Code walkthrough was performed by Sara Lindberg on 05/26/2020 and 06/23/2020. The code relies on standard python libraries as noted in Section 3 and additional code modules that are specific to the HSSM Builder tool only. No impacts to other repository tools or shared library dependencies were identified for the HSSM Builder tool.

1. **Requirements Traceability Matrix**

The requirements traceability matrix for the HSSM Builder tool is presented in Table 1.

| **Table 1. HSSM Builder Tool Requirements Traceability Matrix** | | |
| --- | --- | --- |
| **Functional Requirement ID** | **Acceptance Test ID** | **Test Case** |
| QA Level | CACIE- build\_hssm.py-AT-1 | Installation Test |
| FR-1 | CACIE- build\_hssm.py-AT-1 | Verify the MT3D head file, the model top definition file, the layer bottom definitions files, and the layer initial conditions files are read and processed and pkl files generated |
| FR-2 | CACIE- build\_hssm.py-AT-1 | Generate saturation and yearly saturation pkl files |
| FR-3 | CACIE- build\_hssm.py-AT-2 | Use existing saturation and yearly pkl files (already generated as part of AT-1) |
| FR-4 | CACIE- build\_hssm.py-AT-1 | Verify that the initial and final timesteps for the output data files correspond to the user-defined start year and end year in the JSON-formatted config file |
| FR-5 | CACIE- build\_hssm.py-AT-2 | Verify the flux floor threshold value is equal to the user-defined flux floor threshold in the JSON-formatted config file |
| FR-6 | CACIE- build\_hssm.py-AT-1 CACIE- build\_hssm.py-AT-2 | Verify that the maximum acceptable relative and total mass error of the reduced dataset is less than or equal to the maximum acceptable mass error value (FR-12) in the JSON-formatted config file |
| FR-7 | CACIE- build\_hssm.py-AT-1 | Read in and process multiple mass flux time series data files |
| FR-7 | CACIE-build\_hssm.py-AT-2 | Read in and process a single mass flux time series data file |
| FR-8 | CACIE-build\_hssm.py- AT-2 | Verify that mass flux is set to “0” (zero) when the original mass flux values is less than the flux floor threshold (FR-5) |
| FR-9 | CACIE-build\_hssm.py- AT-1 | Verify that the mass flux is summed for each cell that is included in more than one input file when processing multiple mass flux time series data input files |
| FR-10 | CACIE-build\_hssm.py-AT-1 | Verify that when a cell is dry, the mass associated with the cell is transferred to adjacent cell(s) |
| FR-11 | CACIE-build\_hssm.py-AT-1 | Split the time series for a cell between layers according to the timesteps when a layer meets or exceed the saturation threshold |
| FR-12 | CACIE-build\_hssm.py-AT-1  CACIE-build\_hssm.py-AT-2 | Verify the number of timesteps in the reduced time series datasets are within the target range of 50 to 200. If the reduced dataset is not within the target range, verify that the difference between the reduced and original datasets is within the acceptable error tolerance. |
| FR-13 | CACIE-build\_hssm.py-AT-1 | Verify that the HSSM package input file is generated |
| FR-14 | CACIE-build\_hssm.py-AT-3 | Verify cells that go inactive have the mass shifted to other cells. |
| FR-15 | CACIE-build\_hssm.py-AT-3 | Verify cells that have no mass associated with them in original dataset are removed from loaded data |
| FR-16 | CACIE-build\_hssm.py-AT-5 | Creates a stepwise package |

1. **Test Plan and Cases**

The installation test plan for the HSSM Builder tool is presented in Table 2. The acceptance test cases are presented in Table 3, Table 4, and Table 5.

| Table 2  **HSSM Builder Installation Test Plan** | | | |
| --- | --- | --- | --- |
| **HSSM Builder Installation Testing**  **CACIE-hssm\_build.py– IT-1** | | **Date:** | |
| **Tool Runner Log File Location for this test:**  **[PUT LINK TO THE DIRECTORY HERE]** | | **Test Performed By: [FIRST & LAST NAME]** | |
| **Testing Directory: [PROVIDE LINK TO TESTING DIRECTORY]** | | | |
| **Test Step** | **Test Instruction** | **Expected Result** | **Test Result  (Pass/Fail)** |
| Tools Code Repository Directory: | | | |
| Navigate to the testing directory | | | |
| 1 | Invoke Tool runner and test installation of the tool by entering the following command as follows:  *./runner\_run\_IT-1\_hssm\_builder.sh* | | |
| 2 | Verify Tool Runner is invoked and executed. | Tool runner log generated |  |
| 3 | Verify HSSM Builder tool is invoked | Verify the following is output in command window:  flopy is installed in <python environment>  Invalid inputs: File not found, exiting script. |  |

| Table 3  **HSSM Builder Acceptance Test Plan Case 1** | | | |
| --- | --- | --- | --- |
| **HSSM Builder Acceptance Testing**  **CACIE-build\_hssm.py – AT-1** | | **Date:** | |
| **Tool Runner Log File Location for this test:**  **OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-1** | | **Test Performed By:** | |
| **Testing Directory: OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-1** | | | |
| **Test Step** | **Test Instruction** | **Expected Result** | **Test Result  (Pass/Fail)** |
| Navigate to the Testing Directory | | | |
| 1 | Start application *runner\_run\_AT-1.sh* | | |
| 2 | Wait for the tool to finish running. Note: depending on system, network lag, etc this can take anywhere from a couple of hours to a full day to run. it takes a significant amount of time to process all the cells in the P2R Grid through time. | | |
| 3 | Verify the MT3D head file, the model top definition file, the layer bottom definitions files, and the layer initial conditions files are read and processed and saved as \*.pkl files per FR-1 | The following files were generated:  P2Rv8.3\_flow.pkl P2Rv8.3\_sat.pkl |  |
| 4 | Verify that the uppermost layer where the percent saturation is greater than or equal to user-identified saturation threshold for each time period from MT3D HDS and layer files per FR-2 | The following file was generated:  P2Rv8.3\_yearly\_sat.pkl |  |
| 5 | Verify multiple mass flux timeseries data files are read and processed per FR-7 | Check log/build\_mass\_data\_log\_YYYYMMDD.log and verify both input files were processed, and output files correspond to the cells in the input files |  |
| 6 | Verify flux columns for cell 66-101 from both data\_set1\_1 and data\_set1\_2 were summed together in the generated file “misc/01\_all\_cells\_after\_cell\_merge.csv” per FR-9 | Column 66-101 values (01\_all\_cells\_after\_cell\_merge.csv) = Column 66-101 values (data/data\_set1\_1.csv) + Column 66-101 values (data/data\_set1\_2.csv) |  |
| 7 | Verify mass shifted from column 39-95 per FR-10 | Should have been spread between multiple columns. Roughly half going into 36-98 until year 2043 then split between 35-98 and 36-99., roughly 1 quarter goes into 40-94, and 1 quarter into 41-95. (use output/misc/flux\_mass\_shift\_mapping.csv and output/misc/03\_all\_cell\_by\_day\_dry\_cell\_shifted.csv to verify) |  |
| 8 | Verify 66-101 was split between layers 4, and 5 per FR-11 | Check below files exist:  Output/i66j101k4\_hss.dat  Output/i66j101k5\_hss.dat  verification of data:  Output/log/AT-1\_66-101(k4).png  Output/log/AT-1\_66-101(k5).png  Output/log/AT-1\_66-101\_all\_layers.png |  |
| 9 | Verify the initial and final timesteps of the reduced dataset in generated output files correspond to the start and end year in the JSON-formatted config file per FR-4 | Use output/misc/ 02\_all\_cell\_by\_day.csv. This file contains the data after all files have been loaded into a dataframe. Only the pertinent data to be used will be in it (IE data from start date to end date)  AT-1/config\_AT-1\_olive.json:  "start\_year": 2018,  "end\_year": 12020, |  |
| 10 | Verify reduced data plot corresponds to the original dataset plot (Figures (\*.png files) are located in AT-1/log/);  Reduced number of points within target range of 50 to 200 and/or difference is within acceptable error tolerance (error and reduced dataset set summary is located in misc/cell\_error.csv) per FR-6 (error) and FR-12 | Figures (\*.png files) are located in AT-1/output/log/  abs(Total Mass Error) < Total Mass Error Threshold  AT-1/config\_AT-1\_olive.json:  "max\_tm\_error": 2.25e7, (pCi)  abs(% error) < maximum acceptable tolerance  AT-1/config\_AT-1\_olive.json: "tolerance": 0.1, (IE .1%) |  |
| 11 | Verify the Verify that the HSSM package input file is generated per FR-13 | Check to see that the following file was generated:  Output/mt3d.hss |  |

| Table 4  **HSSM Builder Acceptance Test Plan Case 2** | | | |
| --- | --- | --- | --- |
| **HSSM Builder Acceptance Testing**  **CACIE-build\_hssm.py – AT-2** | | **Date:** | |
| **Tool Runner Log File Location for this test:**  **OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-2** | | **Test Performed By:** | |
| **Testing Directory: OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-2** | | | |
| **Test Step** | **Test Instruction** | **Expected Result** | **Test Result  (Pass/Fail)** |
| Navigate to the Testing Directory | | | |
| 1 | Start application *runner\_run\_AT-2.sh* | | |
| 2 | Wait until tool finishes processing the data files.  Note: The P2R model files were processed as part of the AT-1 test case. The AT-2 test case uses the post-processed \*.pkl files already generated in AT-1 which should speed up the process. Processing for this test case should take between 5 – 30 minutes to run. | | |
| 3 | Verify existing pkl files were used per FR-3. | Open: Log/ build\_sat\_data\_log\_20200520.log  Should have entries for reading: -reading saturation pkl file  -reading Flow pkl file  -reading yearly saturation pkl |  |
| 4 | Verify single mass flux timeseries data file is read and processed per FR-7 | Check log/build\_mass\_data\_log\_YYYYMMDD.log and verify the input file was processed  Verify files in output directory correspond to the cells in the input file |  |
| 5 | Verify reduced data plot corresponds to the original dataset plot (Figures (\*.png files) are located in AT-2/log/);  Reduced number of points within target range of 50 to 200 and/or difference is within acceptable error tolerance (error and reduced dataset set summary is located in misc/cell\_error.csv) per FR-6 (error) and FR-12 | Figures (\*.png files) are located in AT-2/output/log/  abs(Total Mass Error) < Total Mass Error Threshold  AT-2/config\_AT-2\_olive.json: "max\_tm\_error": 2.25e7, (pCi)  abs(% error) < maximum acceptable relative tolerance  AT-2/config\_AT-2\_olive.json:  "tolerance": 0.1, (IE .1%) |  |
| 6 | Using misc/cell\_error\_by\_layer.csv verify: | | |
| 7 | Cell/layers that have flux less than the flux floor threshold value of 2.7378507871321015e-06 pCi/day (i.e., 0.001 pCi/year/365.25 days/year) for the entire time series is skipped per FR-5 and FR-8 | These cells have notes in cell\_error\_by\_layer.csv stating: “This cell was skipped as the flux never rises above minimum flux (2.7378507871321015e-06)” |  |
| 8 | Verify percent error is less than 0.1% (or -0.1%) per FR-12 | Use Column “%error”  Notes:  \* only summed\_layers will contain a value as that adds the layers together then compares it to the original data. |  |

| Table 5  **HSSM Builder Acceptance Test Plan Case 3** | | | |
| --- | --- | --- | --- |
| **HSSM Builder Acceptance Testing**  **CACIE-build\_hssm.py – AT-3** | | **Date:** | |
| **Tool Runner Log File Location for this test:**  **OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-3** | | **Test Performed By:** | |
| **Testing Directory: OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-3** | | | |
| **Test Step** | **Test Instruction** | **Expected Result** | **Test Result  (Pass/Fail)** |
| Navigate to the Testing Directory | | | |
| 1 | Start application *runner\_run\_AT-3.sh* | | |
| 2 | Wait until tool finishes processing the data files.  Note: The P2R model files were processed as part of the AT-1 test case. The AT-3 test case uses the post-processed \*.pkl files already generated in AT-1 which should speed up the process. Processing for this test case should take between 5 – 30 minutes to run. | | |
| 3 | Using misc/03\_all\_cell\_by\_day\_dry\_cell\_shifted.csv verify: | | |
| 4 | Verify that the following cells have no mass in them:  49-130,  49-131,  49-132,  50-131,  50-132,  50-133,  51-130,  51-131,  51-132,  51-133,  52-130,  52-132,  53-130,  53-131,  54-131,  54-132 | These cells never reach a saturation greater than 10%. As such no mass will go into them instead it will all be shifted to downgradient cells that do have saturation.  There are more cells than in the original data due to mass shifting into new cells then immediately being shifted out again as the cell is inactive. |  |
| 5 | Verify that the following cells has mass in day 0 and then zero mass for the rest of the time series  50-130,  49-133,  52-131,  53-132 | These cells have less than 75% but greater than 10% saturation before they go inactive at day 0. Per FR-14 mass will remain in this cell for day 0 then be shifted out afterwards. |  |
| 6 | Verify that the follow cells are no longer in the data set:  51-150  91-151 | Cells are removed if they do not have any mass associated with them in the original dataset. As a result, these cells should have been removed prior to 03\_all\_cell\_by\_day\_dry\_cell\_shifted.csv being created per FR-15. |  |

| Table 6  **HSSM Builder Acceptance Test Plan Case 4** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **HSSM Builder Acceptance Testing**  **CACIE-build\_hssm.py – AT-4** | | | **Date:** | | |
| **Tool Runner Log File Location for this test:**  **OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-4** | | | **Test Performed By:** | | |
| **Testing Directory: OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-4** | | | | | |
| **Test Step** | **Test Instruction** | | **Expected Result** | **Test Result  (Pass/Fail)** | |
| Navigate to the Testing Directory | | | | | |
| 1 | Start application *runner\_run\_AT-4.sh* | | | | |
| 2 | Wait until tool finishes processing the data files.  Note: The P2R model files were processed as part of the AT-1 test case. The AT-4 test case uses the post-processed \*.pkl files already generated in AT-1 which should speed up the process. Processing for this test case should take between 5 – 30 minutes to run. | | | | |
| 3 | If HSSM\_builder runs to completion then the error has been fixed. | All pngs are created (if the error is still there then you will be missing some of the AT-4\_{#}-{#}\_all\_layers.png figures) | | |  |

| Table 7  **HSSM Builder Acceptance Test Plan Case 5** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **HSSM Builder Acceptance Testing**  **CACIE-build\_hssm.py – AT-5** | | | **Date:** | | |
| **Tool Runner Log File Location for this test:**  **OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-5** | | | **Test Performed By:** | | |
| **Testing Directory: OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-5** | | | | | |
| **Test Step** | **Test Instruction** | | **Expected Result** | **Test Result  (Pass/Fail)** | |
| Navigate to the Testing Directory | | | | | |
| 1 | Start application *runner\_run\_AT-5.sh* | | | | |
| 2 | Wait until tool finishes processing the data files.  Note: The P2R model files were processed as part of the AT-1 test case. The AT-4 test case uses the post-processed \*.pkl files already generated in AT-1 which should speed up the process. Processing for this test case should take between 5 – 30 minutes to run. | | | | |
| 3 | If HSSM\_builder runs to completion | Check that folder output\_h-3/step\_format exists and has files:  i100j51k4\_hss.dat  mt3d.hss | | |  |
| 4 | Check format of step\_format/i100j51k4\_hss.dat | Data should be in a step format ie:  day1 0 rate1  day2 0 rate1  day2 0 rate2  day3 0 rate2  day3 0 rate3  etc | | |  |

1. **Acceptance Test Report**

Results of the acceptance testing are in Appendix A. The two test cases are described as follows:

* Acceptance Test 1 is in Table A-1 of Appendix A. The test executed the HSSM Builder Tool and the outputs were evaluated to validate the functional requirements as specified in Requirements Traceability Matrix in Table 1.
* Acceptance Test 2 is in Table A-2 of Appendix A. The test executed the HSSM Builder Tool and the outputs evaluated to validate the functional requirements as specified in Requirements Traceability Matrix in Table 1.
* Acceptance Test 3 is in Table A-3 of Appendix A. The test executed the HSSM Builder Tool and the outputs evaluated to validate the functional requirements as specified in Requirements Traceability Matrix in Table 1.

Details of these tests, when they were conducted, by whom, and if they Passed or Failed are in each table of Appendix A.

1. **User Guide**

To run this tool, you will need to pass 1 argument into it.

1. Config file: Refer to Section 4 of this software management plan for a full description

It is recommended that a shell script is used to execute the tool. The recommended structure of this shell script is shown below.

TOOL=<path/to/build\_hssm.py>

INPUTFILE=<path/to/config/file>

python $TOOL $ INPUTFILE

1. **Tool Versions**

This section details changes incorporated into each version of the HSSM Builder tool.

* 1.0 – Tool was developed.
* 1.1 – New Functionality: Cells going inactive or starting inactive to be mass shifted instead of skipped similar to dry cells; cells without any mass associated with them in the original dataset are deleted from loaded data.
* 1.2 – Fixed 1 issue and created 1 new function:
  + Fixed when saturation layer changes for a cell in the last year of the time series and the layer it is moves to is dropped due to lack of mass, it creates an issue in the checking files when rebuilding the data to compare to the original data.
  + Created optional function to create a stepwise output.

# 

# Appendix

**Completed Acceptance Test Cases**

**Tool Runner Log**

| Table A-1  **HSSM Builder Acceptance Test Plan Case 1** | | | | |
| --- | --- | --- | --- | --- |
| **HSSM Builder Acceptance Testing**  **CACIE-build\_hssm.py – AT-1** | | | **Date: 06/22/2020** | |
| **Tool Runner Log File Location for this test:**  **OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-1** | | | **Test Performed By: S. Tomusiak** | |
| **Testing Directory: OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-1** | | | | |
| **Test Step** | **Test Instruction** | **Expected Result** | | **Test Result  (Pass/Fail)** |
| Navigate to the Testing Directory | | | | |
| 1 | Start application *runner\_run\_AT-1.sh* | | | |
| 2 | Wait for the tool to finish running. Note: depending on system, network lag, etc this can take anywhere from a couple of hours to a full day to run. it takes a significant amount of time to process all the cells in the P2R Grid through time. | | | |
| 3 | Verify the MT3D head file, the model top definition file, the layer bottom definitions files, and the layer initial conditions files are read and processed and saved as \*.pkl files per FR-1 | The following files were generated:  P2Rv8.3\_flow.pkl P2Rv8.3\_sat.pkl | |  |
| 4 | Verify that the uppermost layer where the percent saturation is greater than or equal to user-identified saturation threshold for each time period from MT3D HDS and layer files per FR-2 | The following file was generated:  P2Rv8.3\_yearly\_sat.pkl | |  |
| 5 | Verify multiple mass flux timeseries data files are read and processed per FR-7 | Check log/build\_mass\_data\_log\_YYMMDD.log and verify both input files were processed, and output files correspond to the cells in the input files | |  |
| 6 | Verify flux columns for cell 66-101 from both data\_set1\_1 and data\_set1\_2 were summed together in the generated file “misc/01\_all\_cells\_after\_cell\_merge.csv” per FR-9 | Column 66-101 values (01\_all\_cells\_after\_cell\_merge.csv) = Column 66-101 values (data/data\_set1\_1.csv) + Column 66-101 values (data/data\_set1\_2.csv) | |  |
| 7 | Verify mass shifted from column 39-95 per FR-10 | Should have been spread between multiple columns. Roughly half going into 36-98 until year 2043 then split between 35-98 and 36-99., roughly 1 quarter goes into 40-94, and 1 quarter into 41-95. (use output/misc/flux\_mass\_shift\_mapping.csv and output/misc/03\_all\_cell\_by\_day\_dry\_cell\_shifted.csv to verify) | |  |
| 8 | Verify 66-101 was split between layers 4, and 5 per FR-11 | Check below files exist:  Output/i66j101k4\_hss.dat  Output/i66j101k5\_hss.dat  verification of data:  Output/log/AT-1\_66-101(k4).png  Output/log/AT-1\_66-101(k5).png  Output/log/AT-1\_66-101\_all\_layers.png | |  |
| 9 | Verify the initial and final timesteps of the reduced dataset in generated output files correspond to the start and end year in the JSON-formatted config file per FR-4 (AT-1/config\_AT-1\_olive.json) | Use output/misc/ 02\_all\_cell\_by\_day.csv. This file contains the data after all files have been loaded into a dataframe. Only the pertinent data to be used will be in it (IE data from start date to end date)  AT-1/config\_AT-1\_olive.json:  "start\_year": 2018,  "end\_year": 12020, | |  |
| 10 | Verify reduced data plot corresponds to the original dataset plot (Figures (\*.png files) are located in AT-1/log/);  Reduced number of points within target range of 50 to 200 and/or difference is within acceptable error tolerance (error and reduced dataset set summary is located in misc/cell\_error.csv) per FR-6 (error) and FR-12 | Figures (\*.png files) are located in AT-1/output/log/  abs(Total Mass Error) < Total Mass Error Threshold  AT-1/config\_AT-1\_olive.json:  "max\_tm\_error": 2.25e7, (pCi)  abs(% error) < maximum acceptable tolerance  AT-1/config\_AT-1\_olive.json: "tolerance": 0.1, (IE .1%) | |  |
| 11 | Verify the Verify that the HSSM package input file is generated per FR-13 | Check to see that the following file was generated:  Output/mt3d.hss | |  |

**Tool Runner Log**

| Table A-2  **HSSM Builder Acceptance Test Plan Case 2** | | | |
| --- | --- | --- | --- |
| **HSSM Builder Acceptance Testing**  **CACIE-build\_hssm.py – AT-2** | | **Date: 06/22/2020** | |
| **Tool Runner Log File Location for this test:**  **OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-2** | | **Test Performed By: S. Tomusiak** | |
| **Testing Directory: OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-2** | | | |
| **Test Step** | **Test Instruction** | **Expected Result** | **Test Result  (Pass/Fail)** |
| Navigate to the Testing Directory | | | |
| 1 | Start application *runner\_run\_AT-2.sh* | | |
| 2 | Wait until tool finishes processing the data files.  Note: The P2R model files were processed as part of the AT-1 test case. The AT-2 test case uses the post-processed \*.pkl files already generated in AT-1 which should speed up the process. Processing for this test case should take between 5 – 30 minutes to run. | | |
| 3 | Verify existing pkl files were used per FR-3. | Open: Log/ build\_sat\_data\_log\_20200520.log  Should have entries for reading: -reading saturation pkl file  -reading Flow pkl file  -reading yearly saturation pkl |  |
| 4 | Verify single mass flux timeseries data file is read and processed per FR-7 | Check log/build\_mass\_data\_log\_YYMMDD.log and verify the input file was processed  Verify files in output directory correspond to the cells in the input file |  |
| 5 | Verify reduced data plot corresponds to the original dataset plot (Figures (\*.png files) are located in AT-2/log/);  Reduced number of points within target range of 50 to 200 and/or difference is within acceptable error tolerance (error and reduced dataset set summary is located in misc/cell\_error.csv) per FR-6 (error) and FR-12 | Figures (\*.png files) are located in AT-2/output/log/  abs(Total Mass Error) < Total Mass Error Threshold  AT-2/config\_AT-2\_olive.json: "max\_tm\_error": 2.25e7, (pCi)  abs(% error) < maximum acceptable relative tolerance  AT-2/config\_AT-2\_olive.json:  "tolerance": 0.1, (IE .1%) |  |
| 6 | Using misc/cell\_error\_by\_layer.csv verify: | | |
| 7 | Cell/layers that have flux less than the flux floor threshold value of 2.7378507871321015e-06 pCi/day (i.e., 0.001 pCi/year/365.25 days/year) for the entire time series is skipped per FR-5 and FR-8 | These cells have notes in cell\_error\_by\_layer.csv stating: “This cell was skipped as the flux never rises above minimum flux (2.7378507871321015e-06)” |  |
| 8 | Verify percent error is less than 0.1% (or -0.1%) per FR-12 | Use Column “%error”  Notes:  \* only summed\_layers will contain a value as that adds the layers together then compares it to the original data. |  |

**Tool Runner Log**

| Table A-3  **HSSM Builder Acceptance Test Plan Case 3** | | | |
| --- | --- | --- | --- |
| **HSSM Builder Acceptance Testing**  **CACIE-build\_hssm.py – AT-3** | | **Date: 06/22/2020** | |
| **Tool Runner Log File Location for this test:**  **OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-3** | | **Test Performed By: S. Tomusiak** | |
| **Testing Directory: OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-3** | | | |
| **Test Step** | **Test Instruction** | **Expected Result** | **Test Result  (Pass/Fail)** |
| Navigate to the Testing Directory | | | |
| 1 | Start application *runner\_run\_AT-3.sh* | | |
| 2 | Wait until tool finishes processing the data files.  Note: The P2R model files were processed as part of the AT-1 test case. The AT-3 test case uses the post-processed \*.pkl files already generated in AT-1 which should speed up the process. Processing for this test case should take between 5 – 30 minutes to run. | | |
| 3 | Using misc/03\_all\_cell\_by\_day\_dry\_cell\_shifted.csv verify: | | |
| 4 | Verify that the following cells has mass in day 0 and then zero mass for the rest of the time series  50-130,  49-133,  52-131,  53-132 | These cells have less than 75% but greater than 10% saturation before they go inactive at day 0. Per FR-14 mass will remain in this cell for day 0 then be shifted out afterwards. |  |
| 5 | Verify that the following cells has mass in day 0 and then zero mass for the rest of the time series  50-130,  49-133,  52-131,  53-132 | These cells have less than 75% but greater than 10% saturation before they go inactive at day 0. Per FR-14 mass will remain in this cell for day 0 then be shifted out afterwards. |  |
| 6 | Verify that the follow cells are no longer in the data set:  51-150  91-151 | Cells are removed if they do not have any mass associated with them in the original dataset. As a result, these cells should have been removed prior to 03\_all\_cell\_by\_day\_dry\_cell\_shifted.csv being created per FR-15. |  |

**Tool Runner Log**

| Table A-4  **HSSM Builder Acceptance Test Plan Case 4** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **HSSM Builder Acceptance Testing**  **CACIE-build\_hssm.py – AT-4** | | | **Date:** | | |
| **Tool Runner Log File Location for this test:**  **OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-4** | | | **Test Performed By:** | | |
| **Testing Directory: OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-4** | | | | | |
| **Test Step** | **Test Instruction** | | **Expected Result** | **Test Result  (Pass/Fail)** | |
| Navigate to the Testing Directory | | | | | |
| 1 | Start application *runner\_run\_AT-4.sh* | | | | |
| 2 | Wait until tool finishes processing the data files.  Note: The P2R model files were processed as part of the AT-1 test case. The AT-4 test case uses the post-processed \*.pkl files already generated in AT-1 which should speed up the process. Processing for this test case should take between 5 – 30 minutes to run. | | | | |
| 3 | If HSSM\_builder runs to completion then the error has been fixed. | All pngs are created (if the error is still there then you will be missing some of the AT-4\_{#}-{#}\_all\_layers.png figures) | | |  |

| Table A-5  **HSSM Builder Acceptance Test Plan Case 5** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **HSSM Builder Acceptance Testing**  **CACIE-build\_hssm.py – AT-5** | | | **Date:** | | |
| **Tool Runner Log File Location for this test:**  **OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-5** | | | **Test Performed By:** | | |
| **Testing Directory: OLIVE//CAVE/CA-CIE-Tools-TestEnv/v4-4\_HSSM\_Package\_builder/AT-5** | | | | | |
| **Test Step** | **Test Instruction** | | **Expected Result** | **Test Result  (Pass/Fail)** | |
| Navigate to the Testing Directory | | | | | |
| 1 | Start application *runner\_run\_AT-5.sh* | | | | |
| 2 | Wait until tool finishes processing the data files.  Note: The P2R model files were processed as part of the AT-1 test case. The AT-4 test case uses the post-processed \*.pkl files already generated in AT-1 which should speed up the process. Processing for this test case should take between 5 – 30 minutes to run. | | | | |
| 3 | If HSSM\_builder runs to completion | Check that folder output\_h-3/step\_format exists and has files:  i100j51k4\_hss.dat  mt3d.hss | | |  |
| 4 | Check format of step\_format/i100j51k4\_hss.dat | Data should be in a step format ie:  day1 0 rate1  day2 0 rate1  day2 0 rate2  day3 0 rate2  day3 0 rate3  etc | | |  |