

Bent-Hysplit Workflow Tool User's Guide

Version 1.0

University at Buffalo, SUNY

December 14, 2018

Contents

1	Overview	3
2	Workflow Control	4
2.1	Load Configuration File	4
2.2	Save Configuration File	4
2.3	Enable Pretest	4
2.4	Start	4
2.5	Abort	4
2.6	Volcano Name	4
2.7	Current Conditions	5
2.8	Temperature	5
2.9	Eruption Type	6
2.10	Eruption Mass	6
2.11	Eruption Duration	6
2.12	Water Fraction	6
2.13	Particle Density	6
2.14	Run Hours	6
2.15	Save Hours	6
2.16	Number of Particles	6
2.17	Grid Latitude Range	6
2.18	Grid Longitude Range	6
2.19	Grid Height	6
2.20	Grid Resolution	6
2.21	Vent Radius Range	7
2.22	Vent Velocity Range	7

2.23	Grain Mean Range	7
2.24	Grain Standard Deviation Range	7
2.25	Radiosonde Files	7
2.26	Meteo Data Files	7
2.27	Num d_run Sample Directories	8
2.28	Maxwalltime	8
3	Workflow Execution	8
4	Workflow Results	10
5	Troubleshooting	11

1 Overview

The Bent-Hysplit Workflow Tool (benthysplitwf) extends capabilities of the Vhub bent (volcanic plume model) and puffin (bent coupled with the ash cloud model, puff) tools (<https://vhub.org/resources/bent> and [/puffin](https://vhub.org/resources/puffin)), and produces NetCDF formatted files containing information on the probability of ash at specific heights at specific times and locations, following a volcanic eruption. This tool incorporates the abilities of the Air Resources Laboratory HYSPLIT dispersion model, <https://www.ready.noaa.gov/HYSPLIT.php>, for the time evolution of a volcanic ash cloud advected in the atmosphere.

A key feature of the Bent-Hysplit Workflow Tool is the ability to handle uncertainty in Bent and HYSPLIT input parameters, requiring an ensemble of task executions which are encapsulated and executed as a workflow. Please see the **Workflow Execution** section for more information.

A Graphical User Interface (GUI) (Figure 1) is presented when the user clicks on the tool's Launch Tool button. The GUI provides the user interface for controlling and running a Bent-Hysplit workflow.

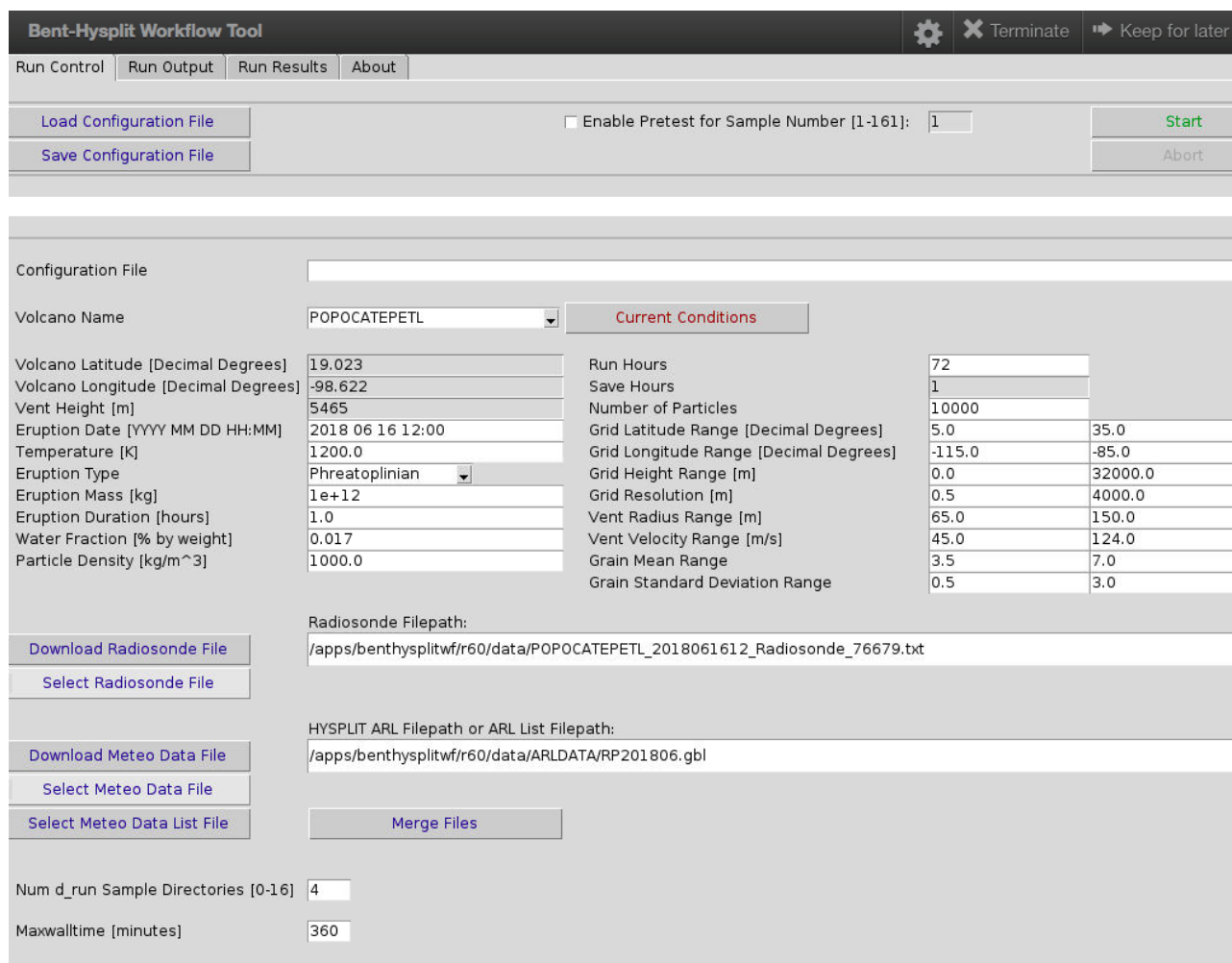


Figure 1: Initial screen of the benthysplitwf tool GUI in vhub.

To run a workflow, go to the **Run Control** tab, update or load configuration parameters for the workflow, and click the

Start button to start the execution of the workflow. When the workflow's execution is complete, output from the workflow is stored in a dated run directory in your home folder and basic output from the workflow is available for display on the **Run Results** tab. See the **Workflow Execution** and **Workflow Results** sections for more information.

In the section that follows, descriptions for the buttons, check boxes, selection lists, and configuration parameter fields of the GUI, for controlling and running a Bent-Hysplit workflow, are provided.

2 Workflow Control

The **Run Control** tab on the GUI provides buttons, check boxes, selection lists, and configuration parameter fields for controlling and running a Bent-Hysplit workflow. In this section, descriptions for these are provided.

2.1 Load Configuration File

Button to load a previously saved configuration file. The name of the loaded configuration file will appear on the GUI next to the **Configuration File** label. If an error occurs while loading the configuration file and the error indicates that an insecure string pickle error occurred, verify that the file is in Unix format by entering the command, `dos2unix filename`.

2.2 Save Configuration File

Button to save the current set of configuration parameters to a configuration file. The name of the saved configuration file will appear on the GUI next to the **Configuration File** label. The configuration file is saved in Unix format.

Note: When the Bent-Hysplit Workflow Tool's **Launch Tool** button is clicked, the GUI loads a default set of configuration parameters. Modify the default set of configuration parameters or load a previously saved configuration file to control a workflow. When the **Start** button is clicked, the current set of configuration parameters is automatically saved to a configuration file named `configuration.txt` in a dated run directory. See the **Workflow Execution** section for more information.

2.3 Enable Pretest

Checkbox to enable a pretest of configuration parameters. In particular, **Pretest** can be used to check whether input source parameters, such as eruption plume height, are appropriate. See the **Workflow Execution** section for more information. If the **Pretest** checkbox is unchecked, a full probabilistic set of runs is executed.

2.4 Start

Button to start the execution of a workflow. See the **Workflow Execution** section for more information.

2.5 Abort

Button to abort a previously started workflow. This button allows for regaining control of the GUI. Normally this button will not need to get clicked. See the **Workflow Execution** section for more information.

2.6 Volcano Name

Volcano name selection list. Contains coordinates of vent position as well as volcano name.

2.7 Current Conditions

Button to open the NOAA/CIMSS Volcanic Cloud Monitoring web page, <https://volcano.ssec.wisc.edu>, for the selected volcano. This web page will display on top of the GUI. An example of this web page for the Popocatepetl volcano is shown in Figure 2. To close this web page, enter Ctrl-W or click the x next to the Volcanic Cloud Monitoring tab name in the upper left hand corner. When this web page is closed, the GUI will reappear.

The Volcanic Cloud Monitoring web page features near real-time processing of many geostationary and low-earth orbit satellites covering much of the globe. The satellite imagery and derived product loops contain data for approximately the last 28 days. For more information, visit <https://volcano.ssec.wisc.edu>.

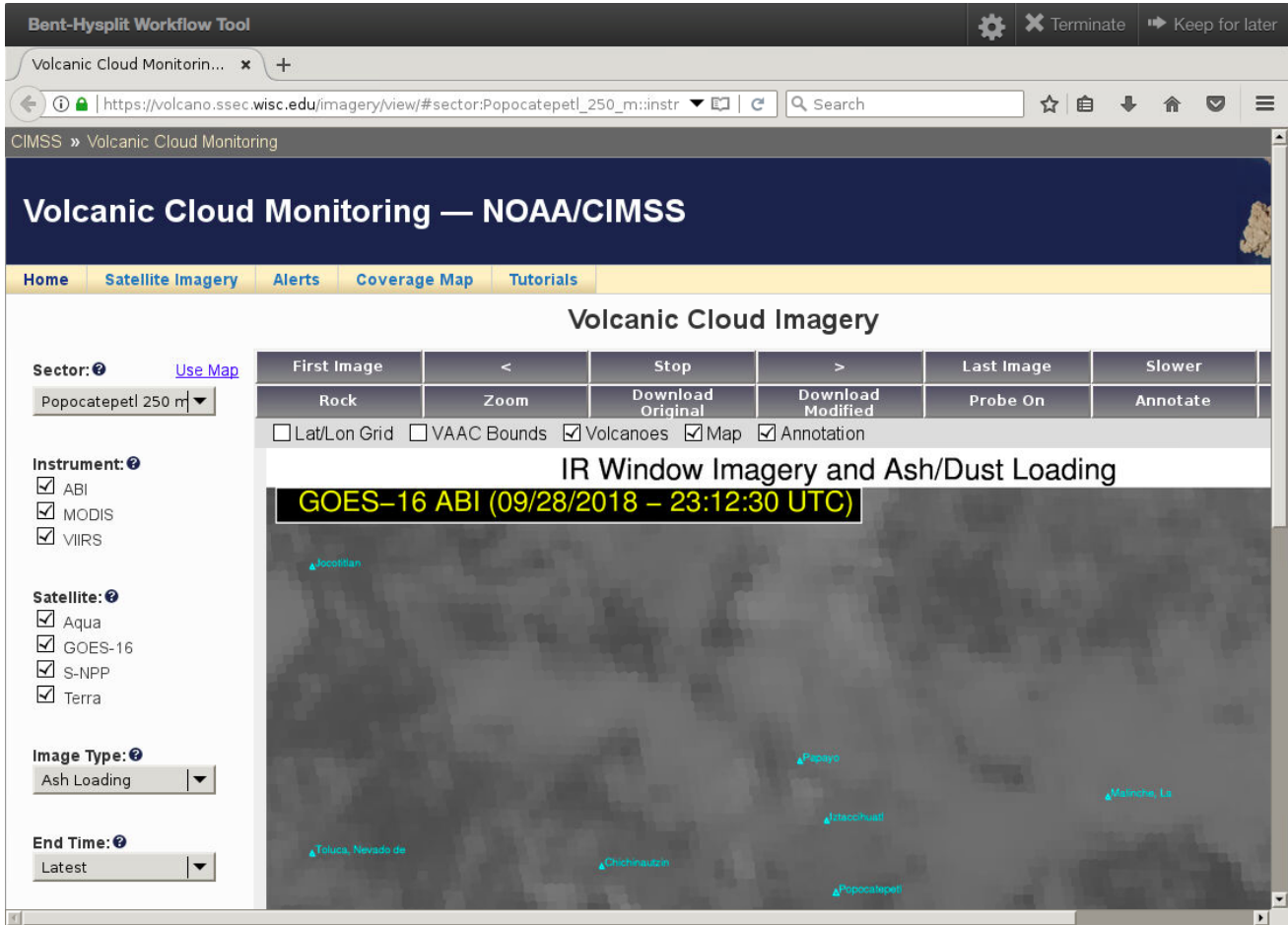


Figure 2

2.8 Temperature

Temperature of the erupting material in the volcanic plume, in Kelvin.

2.9 Eruption Type

Eruption type selection list. Determines total grain size distribution using the data in Bursik and Woods (1991). Currently this is overridden by the uncertain grain mean and grain standard deviation parameters, but can give an idea of the typical particle size range for different eruption types, if the eruption type is known.

2.10 Eruption Mass

Total eruption mass in kilograms. This is a notional mass that is distributed evenly among the HYSPLIT ash cloud particles or parcels. The true total mass erupted is given by $M_{erupt} = \pi\beta r_0^2 u_0 \times t_{erupt} = MER \times t_{erupt}$, where β is density of the erupting mixture, r_0 is vent radius, u_0 is centerline eruption speed (vent velocity), t_{erupt} is eruption duration and MER is mass eruption rate.

2.11 Eruption Duration

Eruption duration in hours.

2.12 Water Fraction

Weight fraction of water in erupting mixture.

2.13 Particle Density

Density of most pyroclasts in kilograms per meters cubed.

2.14 Run Hours

Number of real-time simulation hours, in other words, how long the plume is tracked.

2.15 Save Hours

HYSPLIT write time interval in hours.

2.16 Number of Particles

Number of ash cloud particles or parcels that will be simulated in HYSPLIT.

2.17 Grid Latitude Range

Latitude range over which particles will be simulated in decimal degrees.

2.18 Grid Longitude Range

Longitude range over which particles will be simulated in decimal degrees.

2.19 Grid Height

Height range over which particles will be simulated in meters.

2.20 Grid Resolution

Horizontal and vertical grid spacing in decimal degrees and meters, respectively.

2.21 Vent Radius Range

Range for the uncertain vent radius parameter in meters.

2.22 Vent Velocity Range

Range for the uncertain speed at which the erupting mixture exits the vent in meters per second.

2.23 Grain Mean Range

Range for the uncertain mean grain size parameter in phi units ($\varphi = -\log_2 mm$).

2.24 Grain Standard Deviation Range

Range for the uncertain grain standard deviation parameter in phi units.

2.25 Radiosonde Files

Radiosonde files are column formatted text files containing atmospheric parameters required by Bent for the currently selected volcano and eruption date and time. These parameters are: pressure [hPa], height [m], temperature [C], dew point temperature [C], wind speed [m/s] and wind direction [deg]. A filepath to the currently selected radiosonde file is shown in the **Radiosonde Filepath** field.

This tool can process the weather balloon radiosonde files obtained in HTML format from the University of Wyoming <http://weather.uwyo.edu> web site. The software parses the radiosonde file and creates an input file named weather.txt, extracting the columns required by Bent, creating a density column from the temperature and pressure columns, and, converting the wind speed column from knots to meters per second.

Click the **Download Radiosonde File** button to download a weather balloon radiosonde file from the Wyoming web site for the selected volcano and eruption date and time. Software searches a master location identifier database to determine WMO stations closest to the volcano, and then an attempt is made to retrieve a radiosonde file from the closest WMO station for the selected eruption date and time. When a radiosonde file is successfully retrieved from the Wyoming web site, a prompt will display to select a path for saving the file and the selected path will display in the **Radiosonde Filepath** field.

Click the **Select Radiosonde File** button to select a path to a radiosonde file previously uploaded to Vhub. The selected path will display in the **Radiosonde Filepath** field. The tool's file transfer popup box can be used to upload files to Vhub.

2.26 Meteo Data Files

NOAA Air Resources Laboratory (ARL) files are meteorological data files formatted for use by HYSPLIT.

Click the **Download Meteo Data File** button to download an ARL formatted file from the <ftp://arlftp.arlhq.noaa.gov/pub/archives/reanalysis> web site for the selected eruption date. When an ARL file is successfully retrieved from the web site, a prompt will display to select a path for saving the file and the selected path will display in the **HYSPLIT ARL Filepath or ARL List Filepath** field.

Note: ARL files on the <ftp://arlftp.arlhq.noaa.gov/pub/archives/reanalysis> web site for a particular month and year are uploaded at the beginning of the following month. If an ARL file is not yet available for the selected eruption date, a file not found error will display on the **Run Output** tab.

Click the **Select Meteo Date File** button to select a path to a meteo data file previously uploaded to Vhub. Select an ARL formatted file or a WRF-ARW NetCDF formatted file. If a WRF-ARW NetCDF formatted file is selected, this file will be automatically converted to ARL format and a prompt will display to select a path for saving the converted file. The path to the selected or converted ARL file will display in the **HYSPLIT ARL Filepath or ARL List Filepath** field. The tool's file transfer popup box can be used to upload files to Vhub.

Note: Per ARL request, use of forecast meteorology datasets is not allowed; attempted use of forecast meteorology datasets will result in an error.

HYSPLIT has the ability to work with lists of ARL files. Click the **Select Meteo Data List File** button to select a text file (ending in .txt) containing a list of ARL formatted files or a list of WRF-ARW NetCDF formatted files. If a list of WRF-ARW NetCDF formatted files is selected, the files in the list will be automatically converted to ARL format and prompts will display to select paths for saving the converted files and file list. The name of the selected or converted file list will display in the **HYSPLIT ARL Filepath or ARL List Filepath** field.

Note: Lists are implemented to allow the user to use multiple input NWP files following standard HYSPLIT rules (see the HYSPLIT User Guide).

Click the **Merge Files** button to concatenate the ARL files specified in the selected or created ARL list file into one ARL file. A prompt will display to select a path for saving the created ARL file. The **HYSPLIT ARL Filepath or ARL List Filepath** field will be updated to display the name of the created ARL file.

2.27 Num d_run Sample Directories

Number of d_run sample directories to download from the remote host. See the **Workflow Results** section for more information.

2.28 Maxwalltime

The maximum wall time for running each of the Pegasus tasks. The default value for **Maxwalltime** should provide ample time for running the workflow tasks.

In the section that follows, more details for running a workflow are provided.

3 Workflow Execution

Click the **Start** button on the **Run Control** tab to start the execution of a workflow.

When the **Start** button is clicked, a dated run directory is created in your home folder on Vhub. The name of the dated run directory is set to `benthysplitwf.<current year>-<current month>-<current day>.<current hour>-<current minute>-<current second>`. For example, `/home/vhub/renettej/benthysplitwf.2018-09-09_11-58-22`. A snapshot of the current set of configuration parameters is saved to a configuration file named `configuration.txt` in the dated run directory; these configuration parameters can later be loaded via the **Run Control** tab's **Load Configuration File** button.

The Bent-Hysplit Workflow Tool handles uncertainty in column height or eruption plume rise height through the **Vent Radius** and **Vent Velocity** parameters, which together control mass eruption rate (MER), and the grain size and fine-grain fraction through the **Grain Mean** and **Grain Standard Deviation** configuration parameters. Polynomial Chaos Quadrature (PCQ) sample points, generated using the Conjugate Unscented Transform method, are used in conjunction with ranges for these parameters specified on the **Run Control** tab, to create sets of Bent input values for the parameters. Running Bent with a particular set of input values for these parameters is considered a sample run of Bent. Bent output from each member of the sample run is used in conjunction with other configuration parameters specified on the **Run Control** tab to

create CONTROL and SETUP.CFG files for HYSPLIT; running HYSPLIT with these CONTROL and SETUP.CFG files is considered a sample run of HYSPLIT. HYSPLIT output from all HYSPLIT sample run members is used as input to the tool's uncertainty quantification analysis software. The output of the tool's uncertainty quantification analysis software is a NetCDF formatted file, probmap.nc, containing probabilities of ash concentrations greater than specific levels at specific times and locations. In all, 161 sample runs of Bent and HYSPLIT are performed requiring an ensemble of Bent, HYSPLIT and uncertainty quantification analysis tasks, which are encapsulated and executed as a workflow.

Note that the Bent model allows for eruption column collapse. This means that in a broad vent radius–vent velocity parameter space, numerous runs may result in collapse, causing formation of a low eruption fountain (Fig. 3). If this behavior is not desirable, restrict the range of vent radius and velocity, and run with higher values of vent velocity, as this situation causes more rapid entrainment of air (Fig. 3), quicker generation of buoyancy in the erupting mixture, and more likely results in convective rise to great height.

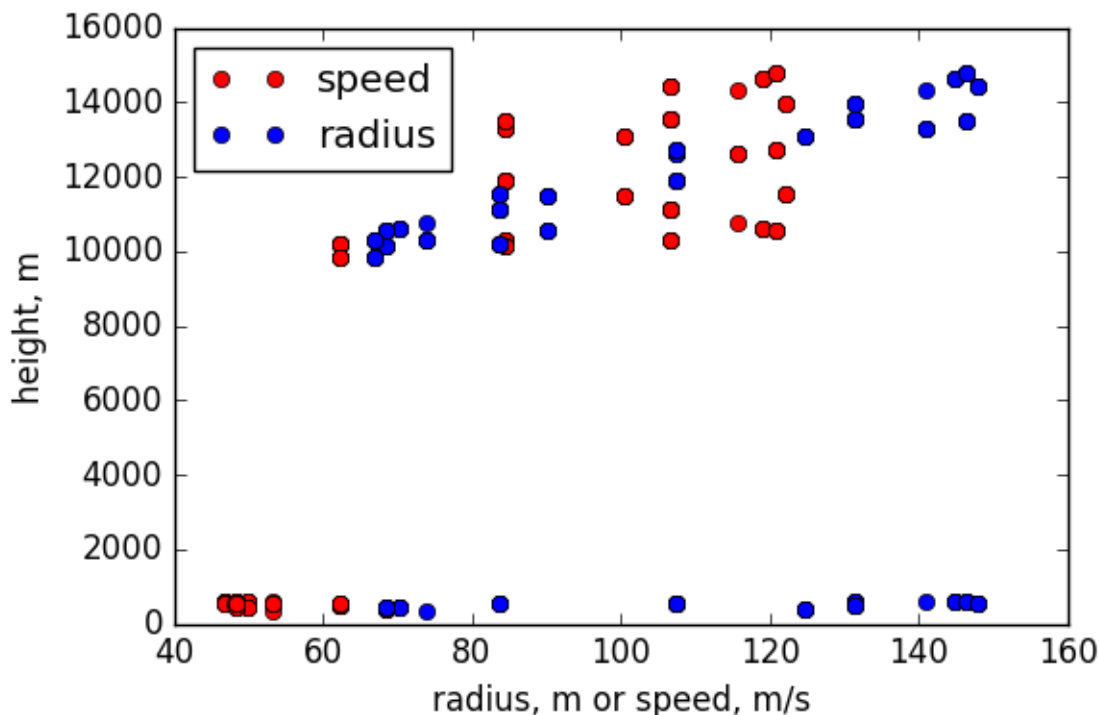


Figure 3

The Pegasus Workflow Management System is used to execute the Bent-Hysplit workflow. Pegasus facilitates the automation and management of the Bent-Hysplit workflow tasks, including, staging jobs for the workflow tasks, submitting the jobs to run on a remote host, handling data flow dependencies, overcoming job failures and downloading workflow output from the remote host to Vhub.

When the **Start** button is clicked and the **Enable Pretest** check box is checked, Bent and HYSPLIT tasks are run locally on Vhub for the selected sample only. **Pretest** enables testing a set of configuration parameters before running a workflow via Pegasus.

Because the most common observation that is used to constrain source characteristics is the plume rise height (eruption column height), the most basic use of **Pretest** is to check a single or a set of sample runs to determine if the eruption column height is within some acceptable range of observations. This check can be made by finding the line in the **Run**

Output tab labeled “ERUPTION PLUME HEIGHT”. If the heights are not satisfactory, then adjustments to the vent radius and vent velocity can be made and tested until a satisfactory set of values is obtained. Future versions of the tool will report this value more immediately.

Eruption plume height and other model outputs, as well as inputs, can be viewed and evaluated by looking in the `d_run_xx` subdirectory, in the dated run directory. In particular, the Bent output file `umbr_xx.dat` contains information on radial umbrella cloud growth rate and upwind stagnation distance, as well as plume height. The output in this file can be used to see if model umbrella cloud growth rate matches observed growth rate, in cases where an umbrella cloud forms and is visible on satellite imagery. The stagnation distance is the distance to which the umbrella cloud grows in the upwind direction, and is used by HYSPLIT as the radius of the initial ash cloud. The `hysplit.nc` file is a CF-compliant netCDF file that shows the propagation of the ash cloud for that specific run. This can be viewed in a McIDAS or Panoply viewer as an additional check on whether the test model ash cloud conforms to available observations.

Pretest can also be used to obtain the values for the neutral buoyancy height, by looking at the line labeled “NEUTRAL BUOYANCY HEIGHT” in the **Run Output** tab or in `umbr_xx.dat`, and the mass loading, by multiplying the value in the line labeled “PARTICLE FLUX AT HB” by the eruption duration.

When the **Start** button is clicked and the **Enable Pretest** check box is *not* checked, a Pegasus-plan with associated DAX file is submitted to Pegasus, via the HUBzero Rapture interface, to start the execution of the workflow with a full complement of bent-HYSPLIT runs that characterize the input parameter space, based on investigation of the pretest runs, and the following adjustments to the input variable values.

Bent input values, derived from the PCQ sample points and the **Run Control** tab, are consolidated into Bent input files that are uploaded by Pegasus to the remote host. The names of the Bent input files are set to `puffin_sample_<sample number>.inp`. For example, `puffin_sample_001.inp`. HYSPLIT input values derived from the **Run Control** tab are consolidated into files that are uploaded by Pegasus to the remote host. These are `.hysplitrc` and `.hysplitargs`.

While a workflow is running, workflow status information is displayed on the **Run Output** tab, including the percentage of the workflow’s execution that is complete. This output is also written to a log file that is copied to the dated run directory when the workflow’s execution is complete. The name of this file is `benthysplitwf_GUI_log.txt`.

4 Workflow Results

The working directory for the Bent-Hysplit Workflow Tool is created in your home folder’s `data/sessions` directory when the Bent-Hysplit Workflow Tool is launched. For example, `/home/vhub/renettej/data/sessions/49256`.

When a workflow’s execution is complete, input files, including files previously uploaded to the remote host by Pegasus, are moved from the tool’s work directory to the dated run directory. See the **Workflow Execution** section for more information on the dated run directories. In addition, workflow output downloaded by Pegasus from the remote host to the tool’s work directory are moved to the dated run directory. The downloaded files include `probmap.nc` and `probmap.txt`, which is a text file containing `probmap.nc` statistics and summary information. For **Pretest**, the `probmap.nc` and `probmap.txt` files are not created.

The contents of `probmap.txt` are available for display on the **Run Results** tab. Also available for display are animated gif files showing ash concentrations for each level.

The `probmap.nc` file is compliant with NetCDF Climate and Forecast (CF) Metadata Conventions (<http://cfconventions.org/cf-conventions/cf-conventions.html>), so it can be read with numerous NetCDF analysis tools, including UW-SSEC McIDAS, UCAR-Unidata IDV and NASA-GISS Panoply.

For each Bent and HYSPLIT sample run, a directory is created on the remote host containing interim Bent and HYSPLIT output files. For example, the Bent `gui.puff` output file and the HYSPLIT `*.srm` output files are contained in these directo-

ries. See the **Workflow Execution** section for more information on Bent and HYSPLIT sample runs. Due to the disk space available on Vhub for most Vhub users, only a select number of randomly selected sample run directories are downloaded from the remote host by Pegasus and moved to the dated run directory. The names of the sample run directories are set to `d_run_sample_<sample number>`. For example, `d_run_sample_001`. The **Num d_run Sample Directories** field on the **Run Control** tab is used to specify the number of `d_run` sample directories to download from the remote host. For **Pretest**, the `d_run` sample directory for the selected sample is created locally and moved to the dated run directory.

In addition to files generated on the remote host by the workflow, Pegasus status and analysis information is returned in a file called `pegasus.analysis` and copied from the tool's work directory to the dated run directory. An example `pegasus.analysis` file is displayed below. The contents of the `pegasus.analysis` file are also available for display on the **Run Results** tab. If a workflow error occurs, see the `pegasus.analysis` file for details. Other Pegasus files are also copied to the dated run directory. These are, `pegasusjobstats.csv`, `pegasussummary-time.csv` and `pegasussummary.csv`.

```

1
2 *****Summary*****
3
4 Submit Directory      : /tmp/1537993556_00024219_01/work
5 Total jobs           :      266 (100.00%)
6 # jobs succeeded     :      266 (100.00%)
7 # jobs failed        :         0 (0.00%)
8 # jobs held          :         0 (0.00%)
9 # jobs unsubmitted   :         0 (0.00%)
10

```

When the Bent-Hysplit Workflow Tool is terminated, the tool's working directory is renamed to an expired sessions directory. Expired sessions directories are prefixed with the name `expired`. For example, `/home/vhub/renettej/data/sessions/expired-49256`. It is recommended that you periodically remove expired session directories and dated run directories that are no longer used to prevent running out of disk space on Vhub.

5 Troubleshooting

If workflow execution is aborted, check the `pegasus.analysis` file for detailed information on potential set-up or execution errors.

Perhaps the most common error is generated when the volcanic eruption column is high, and Bent sends data to HYSPLIT to run with a plume height above the top of the model domain, as set by **Grid Height**. The error message is "Start height above model domain..." Often, increasing **Grid Height** will eliminate the error and allow normal completion of the workflow.