



**3022 Sterling Circle – Suite 200, Boulder, CO 80301**

**Phone (303) 449-1105 Fax (303) 449-0132**

[www.specinc.com](http://www.specinc.com)

# **HVPS Post Processing Using HVPSview Software**

**User Manual**

**September 2010**


# Using HVPSview Software

The HVPS data processing software is constantly being updated. We urge users to consult with SPEC on the current status of the software before beginning processing and periodically.

1.0 HVPS Pre-Processing	2
1.1 Masked diode filling	5
1.2 Dead diode replacement	5
2.0 HVPSview	7
2.1 HVPSview installation	7
2.2 More pre-processing	8
2.3 Viewing Images	11
2.4 Artifact Removal	11
2.5 Time Series and Particle Size Distribution (PSD) Plots	12
2.5.1 Zooming into a time interval	12
2.5.2 Comparing time series	14
2.5.3 Comparing size distributions	15
2.5.4 Viewing HVPS images from a selected time interval	16
2.5.5 Boolean Filtering	17
3 HVPS data Batch processing (Archiving)	19
3.1 Liquid mode Batch processing	19
3.1.1 Archiving ASCII Files	19
3.1.2 Archiving Images	21
3.1.3 Combining the archived liquid ASCII files	22
3.1.4 Additional Notes	22
3.2 Ice mode Batch processing	23
3.2.1 Archiving Ice ASCII Files	23
3.2.2 Archiving Images	25
3.2.3 Combining the archived Ice ASCII files	26
3.2.4 Additional Notes	26
Appendix A: Derived Parameters, Methods, Equations in HVPSview.....	28
Appendix B: Artifact removal algorithms.....	37

## 1.0 HVPS Pre-Processing

Processing HVPS data first requires the Spec2D Executable, written in C++ and Visual Basic and referred to as Playback. After some initial processing with Playback, the main processing and display of HVPS data is accomplished via HVPSview, an IDL-based program.

After a data file has been recorded, the first step to analyzing the data is to open the Executable program, which is the same program used to record the data file. Make sure that HVPS Mode is enabled, and that 150 microns per slice is set in the Display window. Click on the folder icon . It will prompt you to open a 'baseYYMMDDHHMMSS.HVPS', which is a raw data file. It is binary and the images are compressed. Once a file is chosen, a time period selector box will be displayed.

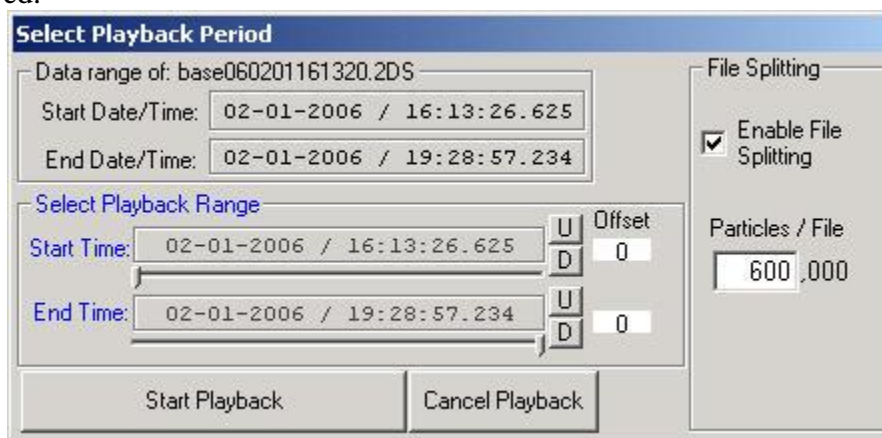


Figure 1:

To select a time period, move the two sliders on the Start Time and End Time. For an exact time period use the U and D buttons to increase and decrease time by one-second intervals. To the right is the File Splitting option. HVPSview attempts to keep information on each particle in memory. When the memory is full, the program slows due to the use of virtual memory. This can result in unacceptably slow processing. This can be remedied by enabling file splitting, which cuts the original large data file into processed data files of manageable sizes, according to a preset number of particles. A typical setting is 300,000 particles per file, but the optimum size will depend on the attributes of the computer used.

Once the time period and file splitting options are chosen, click the 'Start Playback' button. A window will ask whether to create HVPSview pre-processed files, which are called sorted files. These files are also binary with compressed images, but various simplifications to the file structure have been made. For example, the data

is sorted into time sequence. A sorted data file is necessary to run HVPSview. Playback also produces HK\_BINYYMMDDHHMMSS.HVPS, HK\_TXTYYMMDDHHMMSS.HVPS, and PAIRSYYMMDDHHMMSS.HVPS files. The first two contain probe housekeeping data in binary and ASCII forms respectively. Thus, these files allow the user to obtain the housekeeping data independently of HVPSview. The Pair file is now obsolete. Once you click 'yes' you will return to the Main Display window.

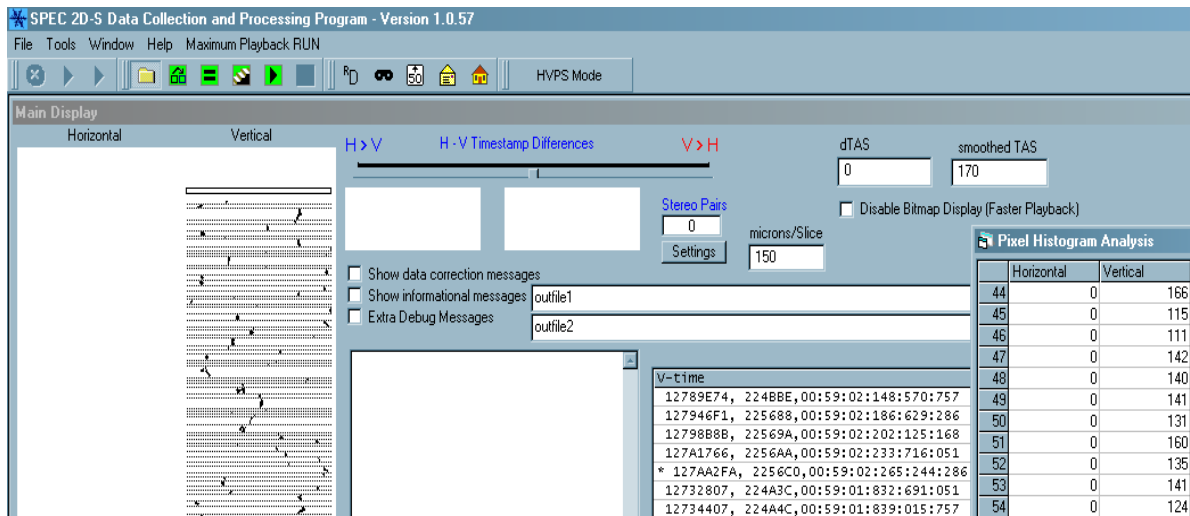


Figure 2


Playback allows the user to look at images while pre-processing the data. By selecting the Maximum Playback RUN and choosing one of the two options, Playback will create the HVPSview files as quickly as possible. This option is to the right of the Help tab. The option to not display timestamps is fastest.

The other option is quite slow. To the right of the Folder icon there are four green buttons. The farthest right is the Play button. This button will play the file back at a speed set by clicking the icon to the right of the folder icon. You can use the Pause button to stop playback. You can step through the data, particle by particle, using the Step button.

On the Main Display, there is a Stereo Pair counter. This area will display horizontal and vertical particles that arrived at nearly the same time. Clicking the Settings button allows the user to set the detection settings:

### 1.1 Masked Diode Filling:

During data collection, a diode may become noisy. If the condition is serious enough, the real time software will detect this and ignore (mask) the diode. This results in a white line running the length of any particle image that includes the mask diode. Examples are shown in Fig. 1 and Fig. 2. Playback has an option to fill these lines within particle images. Click 'tool' then 'mask bit replacement', a window entitled 'Mask Bit Recovery settings' appears. Enable mask filling here and set the maximum number of consecutive masked diodes that can be filled. Filling these artificial white lines is important for the measurement of area, which also affects length when sizing corrections are used. It is even more important to filled masked diodes when they are the edge diode, since that diode is used to determine if an image is complete or not.

-----Image with white stripe

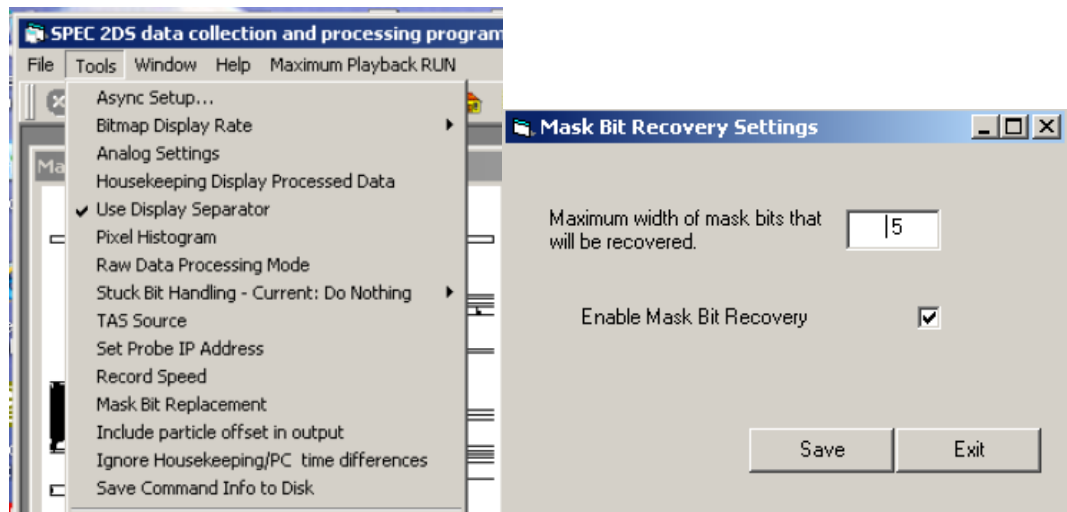


Figure 1: Example of the display used to enable masked diode filling.

## 1.2 Dead diode replacement

Similar to a noisy diode, a dead diode leaves a white line running through particle images. In this case however the real time software cannot detect and mask the diode. Therefore playback has an option to add a diode to the mask during pre-processing. The dead diode must be identified first. This is accomplished via the 'Pixel histogram analysis' window (Fig. 3). The number of counts each diode has had is displayed. Clicking in a numbered box turns it yellow and indicates that that diode will be considered masked for the next time playback is run. That is, adding diodes to the mask is a two step process. First, for the period being processed, you must run playback, monitor the histogram of counts display and click on any diodes that you wanted added to the mask. Next, exit and restart playback, the clicked boxes remain yellow and this time they will be added to the mask for processing.

If a noisy diode is not detected and masked during data collection, it may be detected during these same steps and also added to the mask. This can shorten the time needed for HVPSview to process as it must also remove noisy diode effects.

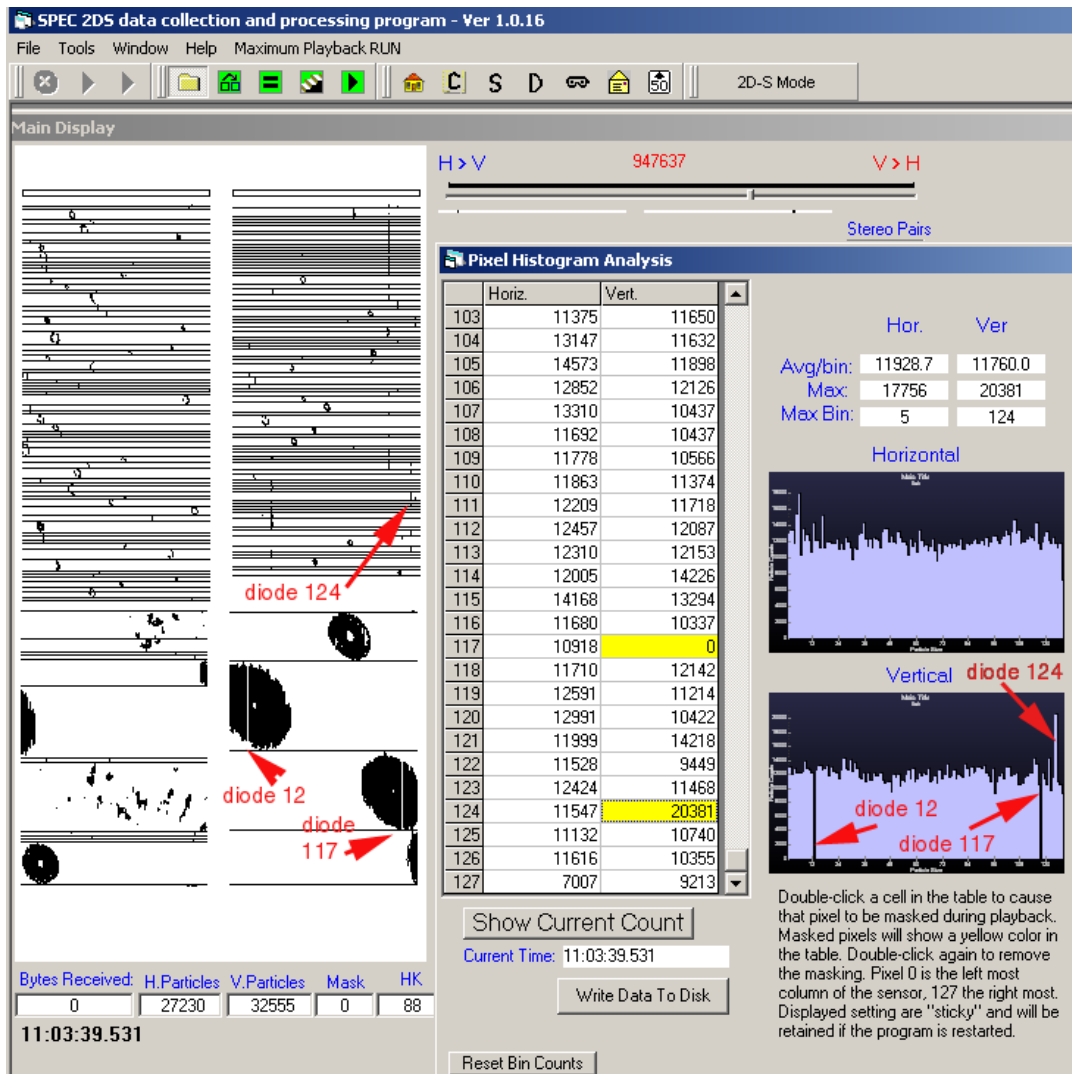


Figure 3: Example of the 'Pixel histogram analysis' window showing a couple dead diodes (V-12 and V-117) and a detectable noisy diode (V-124).

## 2.0 HVPSview

### 2.1 HVPSview Installation

HVPSview runs on an IDL platform of 6.0 or higher. The projects programs should be stored in a folder called HVPSview-version#. Please note that HVPSview has

been derived from the 2DSview software. The 2D-S has two input channels, Horizontal and Vertical, while the HVPS has only one channel. For various reasons stemming from both the hardware electronic configuration and the software, the single HVPS channel has been called "Vertical" and "V", although the probe usually is mounted horizontally. Please note that some example images below may include a Horizontal channel, as the test data imaged better illustrated a concept, such as masking. HVPS data will only show Vertical data.

To install the HVPSview software, follow these steps:

1. Open IDL
2. Click File→Preferences→Path and set the path to the HVPSview-version# folder first and IDL default second. The window should look similar to the one below. Make sure HVPSview is at the top of the search path.

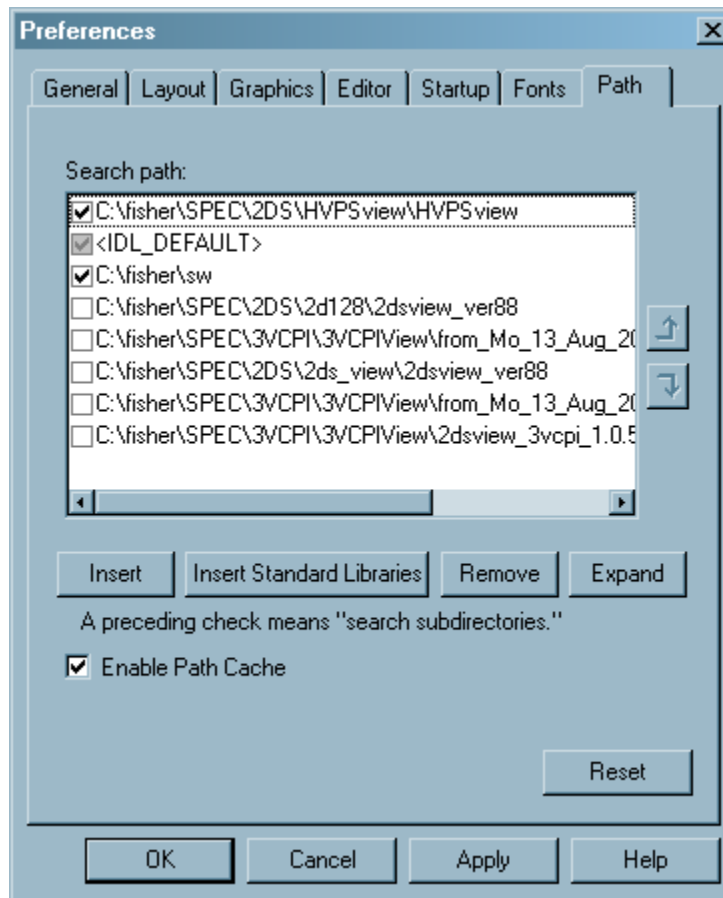
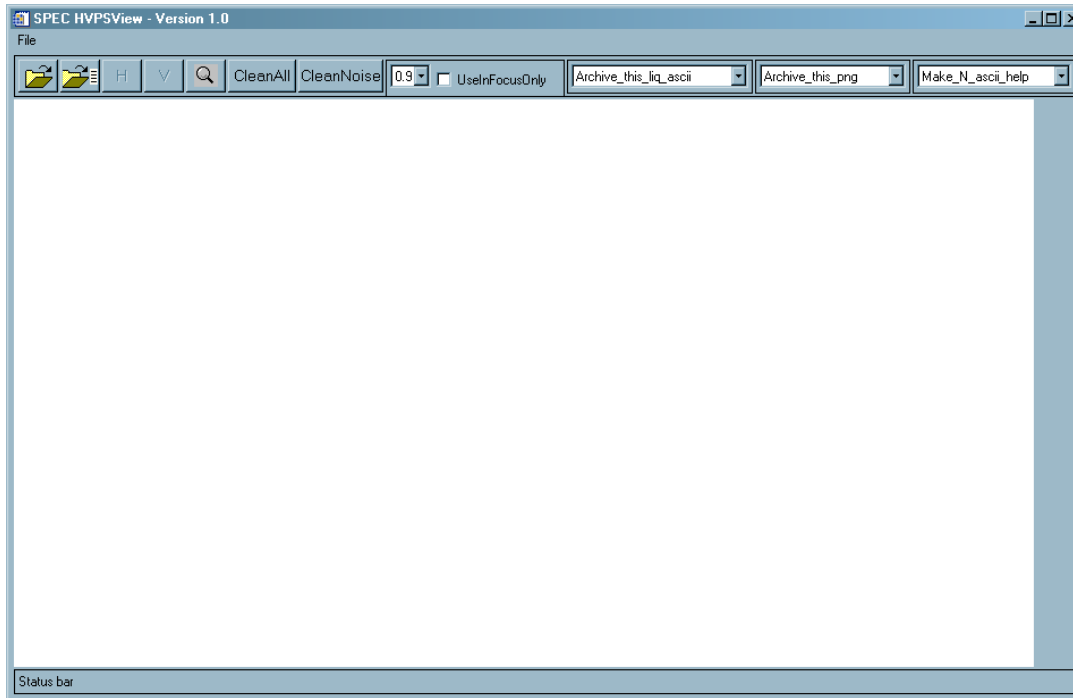


Figure 3

3. Click OK


4. In the command line at the bottom of the window type: @compileHVPS  
→Enter. (NOTE: IDL is not case sensitive, so it is OK to type @compilehvps, and hvpsview.)
5. Type in the command line: HVPSview→ Enter. The following GUI window will open up:



*Figure 4*

## 2.2 More pre-processing

Before performing quantitative analysis some additional pre-processing steps must be done. First .BIN files are needed:

1. Click the left folder icon . This Brings up a list of all the sorted files created during the playback.
2. Select the SORTED file with the desired time interval.
3. Once the file is loaded a window displaying warning messages will pop up. There may be messages regarding time inconsistencies found while processing the sorted file. Usually there are no messages as the case shown in Fig. 6.



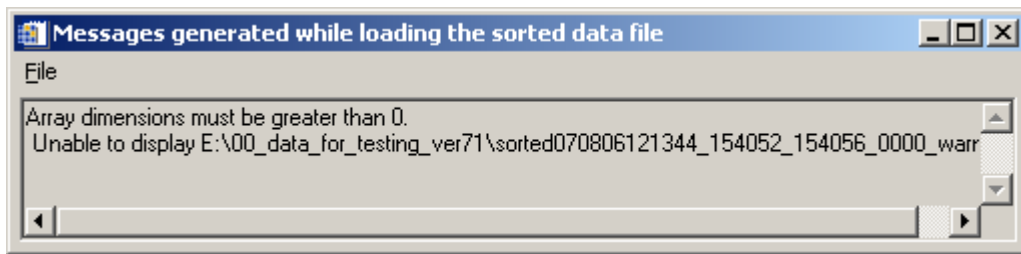



Figure 6: Example of a warning messages window.

Once a file is selected, a status bar in the lower left of the GUI window shows the number of particles counted. The program will automatically generate the files required. All the .BIN files are placed into a subdirectory labeled, "data". This process is only required one time for each sorted file. The program automatically looks in the data subdirectory and uses the appropriate .BIN file if it already exists from a previous running.

\*Note: Due to file size, the time to create .BINS can be quite extensive, but must be done for each SORTED file.

To continue, open a .BIN file:

1. Click the right Folder icon 
2. Open the Data folder
3. Select a \*.BIN for the desired time

Again additional files are created in the data subfolder if they do not already exist. These are 1 Hz data files. There will be several pop-up windows depicting the raw data, and the images are automatically saved in the data folder. These auto-saved files are labeled: horiz\*.png. Figure 7 is an example of the type of window that will appear. The images can be viewed at any point from this point forward. These are created primarily for SPEC programmers to detect and debug idiosyncratic corrupted data problems.

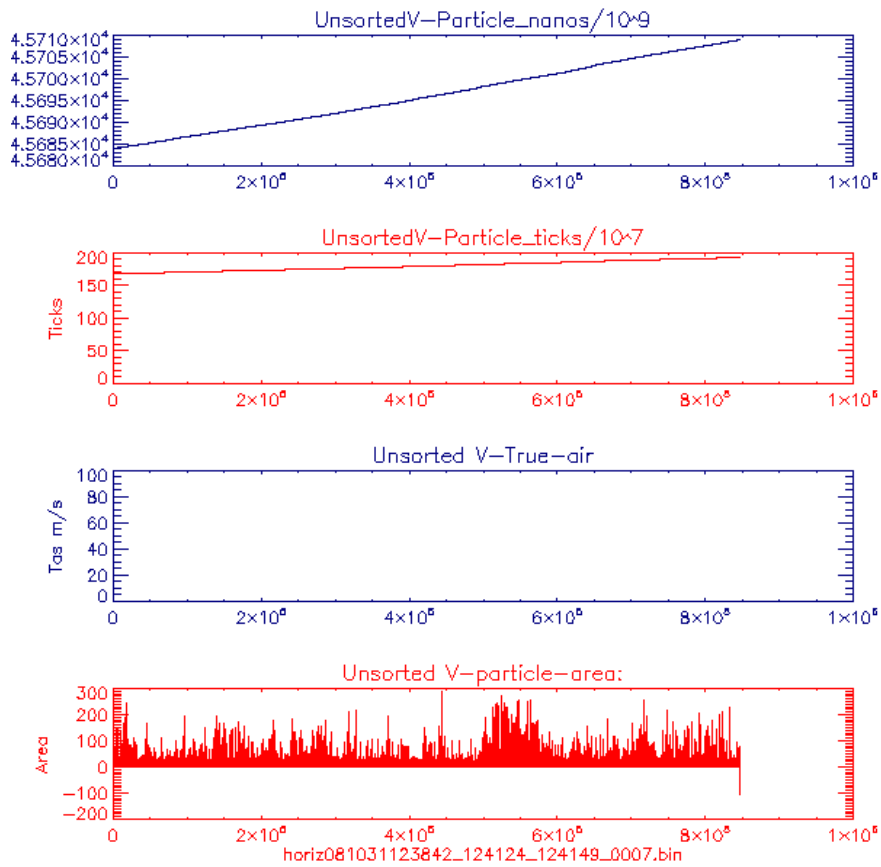


Figure 5

- Again a warning messages window pops up showing warnings of idiosyncratic data events. This window is rarely empty but usually has several to a few dozen events. An example is shown in Fig. 8.

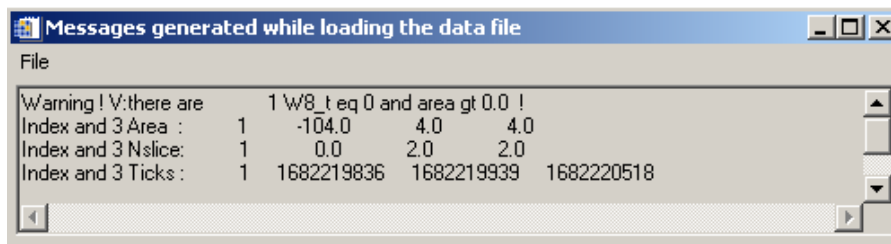
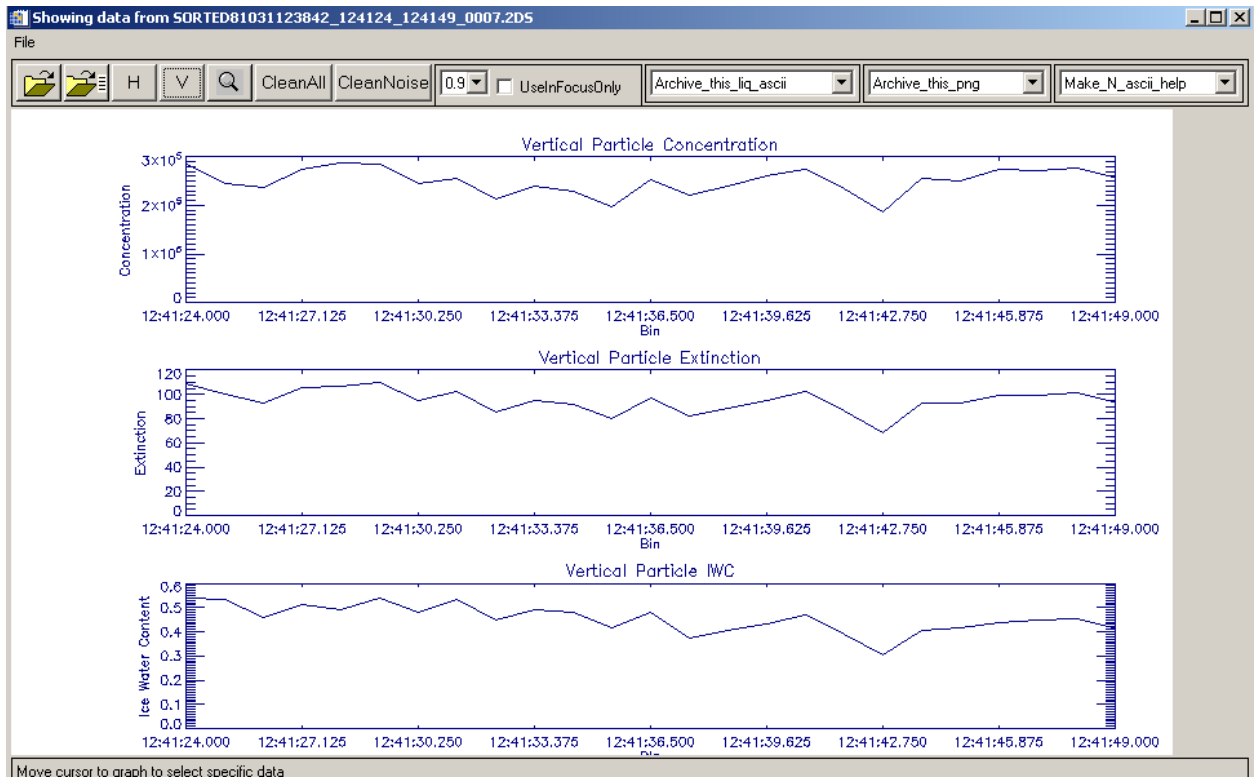


Figure 8: Example of the warning messages window that pops up after loading the bin files.

Once this is completed, three time series are displayed for concentration, extinction, and ice water content. If the file contains more than 300 seconds, the data is coarsely binned into 300 segments for quick look, using the 1 Hz data file that was generated upon opening the .BIN file. If the file is less than 300 seconds then the 1 Hz data is used directly. These 1 Hz data are

for quick look only. The processing uses a crude length parameter ( $L_1$ , see appendix A) and no artifact rejection.



*Figure 6*

By default, the horizontal array time series are displayed. The arrays can be switched by clicking on the  (vertical) or  (horizontal) icons located on the upper toolbar.

### 2.3 Viewing Images

To view the particle images before artifact rejection has been applied:

1. Choose the desired array either vertical or horizontal
2. Move the cursor over any one of the three plots near the time you want to view
3. Single click, a window similar to Figure 10 will open.

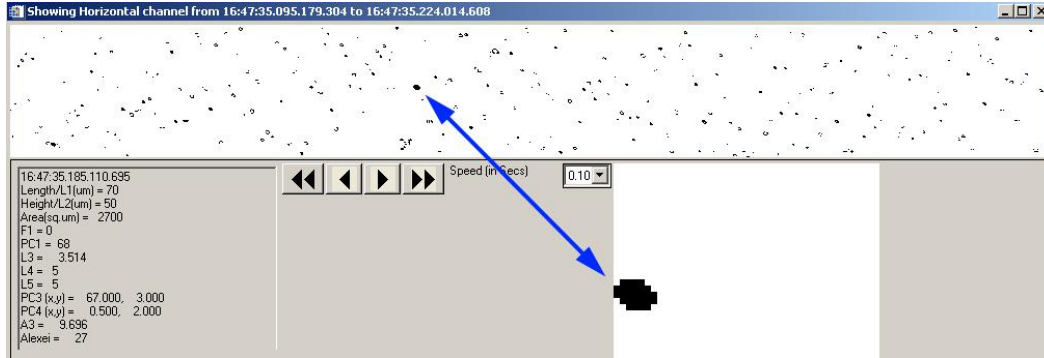




Figure 7

The arrows are used for changing the time interval. The single arrow icons  allow for moving the view one strip at a time. The icons with two arrows  fast forward and rewind. The drop down list to the left of the arrows controls the length of time each strip is in the window while fast forwarding and rewinding. Placing the cursor over single particles displays the size and other parameters in the lower left hand corner. An enlarged image of the particle is displayed, as shown in Figure 10.

## 2.4 Artifact Removal

The .BIN files contain all the data captured by the HVPS Probe. This includes: noise, splashed, and shattered events, as well as valid images. The program has the ability to remove unwanted data. Using the cleaning functions creates concentrated data for the time series and particle size distribution plots. Before quantitative analysis, there are two options for data filtering:

- Clean all: Removes noise, shattering, and splashing.
- Clean noise: Removes noise only.

To use the cleaning functions:

1. Specify the type of cleaning by clicking the corresponding icon on the toolbar.
2. A text box will ask for an aspect ratio, as shown in figure11. If the cloud is known to be liquid enter in 2.0. If the cloud is known to be ice enter a larger value.

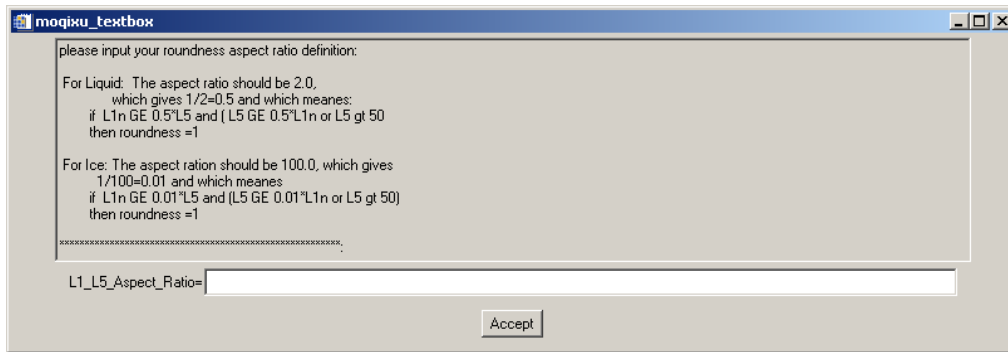


Figure 8

If a cleaning method is not selected, all the images will be used. The resultant data is automatically saved in the directory. If a cleaning function is selected, the program searches for previous cleanings made for each .BIN file. If one is found, you will be prompted with the window shown in Figure 12. If the aspect ratio is the same, select yes. In doing so, the computer will not have to re-clean the data.

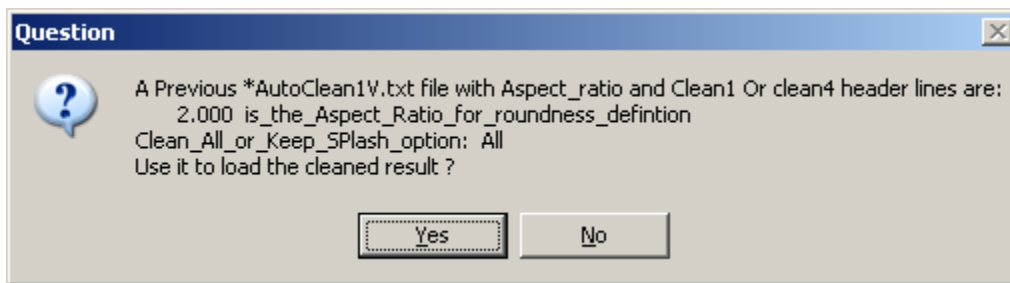


Figure 9

\*Note: The program processes each particle individually based on statistics of many surrounding particles; processing large files can take days.


The In Focus Only feature allows for further data filtering. This removes the particles with holes in them. If the area of the hole is greater than the specified size, the particle is eliminated from the data. For example, if 0.9 is the selected criteria, then only those images with holes having less than 10% of the total area will be used in the process.

To use the In Focus Only Feature:

- 1.) Check the box to the left of "Use In Focus Only" on the toolbar.
- 2.) Select a value from the drop down list (0.9, 0.8, 0.70).

## 2.5 Time Series and Particle Size Distribution (PSD) Plots

### 2.5.1 Zooming into a time interval

- 1.) Click on the magnifying glass icon .

- 2.) Click one of the three time series near the desired start time for analysis.
- 3.) Click the desired end time for analysis.

This creates a smaller time interval to be examined. Choosing small time periods significantly reduces subsequent processing time. Similar to the previous window, this new window initially displays concentration, extinction, and ice water content, read from the 1 Hz data file described above, on single time series plots (both H and V are shown simultaneously here). Figure 13 shows the zoomed-in data window.

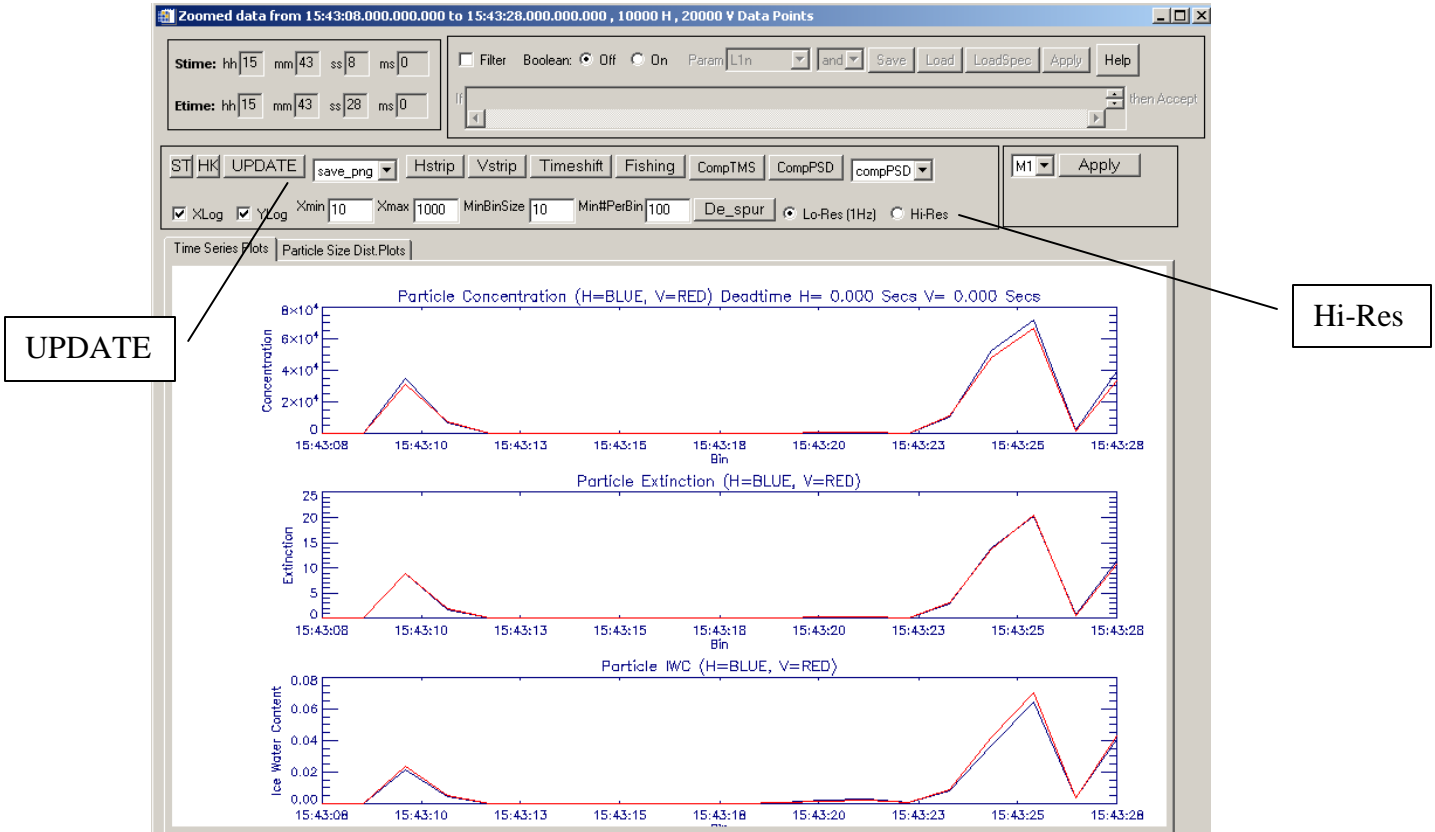


Figure 10

PSDs are plotted as well. To view them, click on the Particle Size Dist. Plots tab. These are also crude quick look products using 1Hz data files. The zoomed-in data window allows for a much more in-depth data analysis.

In order to access these options select Hi-Res, and click UPDATE.

The start and end times can be changed by typing new time values in the Stime and Etime boxes and clicking the UPDATE button.

To save any time series or PSD plot as an image, click the **Save.png** button in the droplist. To save the plot as an ASCII file, click the **Save.Ascii** button. To print a plot, click the **Print** button.

The “ST” button on the left is for displaying stereo images, which is currently not used.

The “HK” is for displaying the Housekeeping data. Figure 14 is an example of the plots obtained by selecting HK. This data is used primarily by SPEC Inc. It can also be used for troubleshooting.



Figure 11

\*Note: Click update after changing any options in the zoomed data window.

## 5.5.2 Comparing time series

Use the **CompTMS** button to compare time series parameters. In this window, different sizing methods, arrays, and parameters can be plotted and compared. Sizing methods are described in Appendix A. ASCII and image files can be saved from this window.

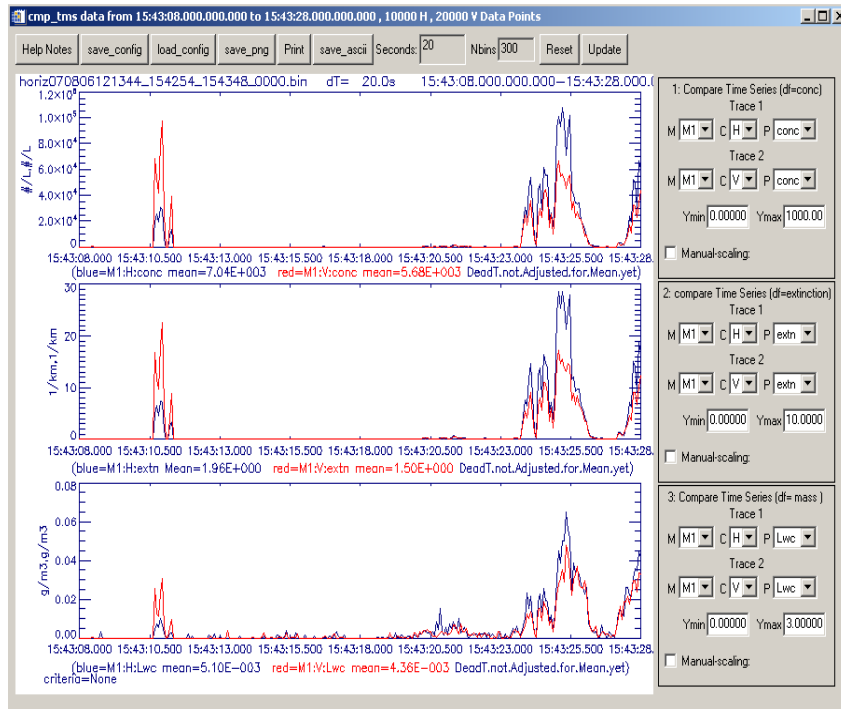


Figure 12

### 2.5.3 Comparing size distributions

Use the **CompPSD** button to compare particle size distributions. In this window, different sizing methods, arrays, and parameters can be chosen, plotted and compare. Sizing methods are described in Appendix A. Furthermore, average values of counts, concentration, extinction and ice water content will be plotted when applicable.



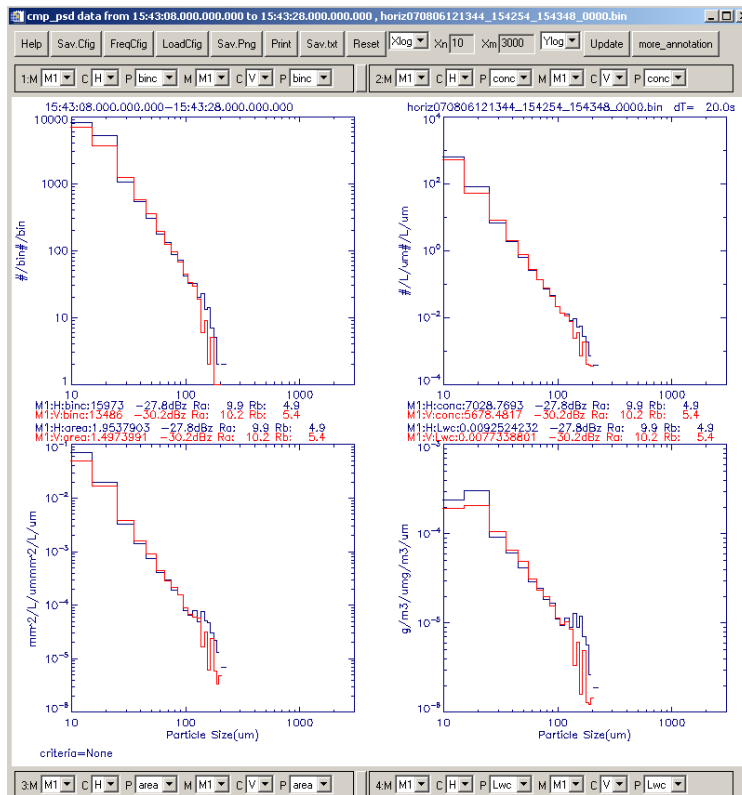



Figure 13

## 2.5.4 Viewing the HVPS images from a selected time interval

- 1.) In the zoomed data window (Figure 13), check the box to the left of 'Filter' if accepted images only are desired (for both display and processing). Unclick the filter box to process and display using all the images accepted and rejected. Rejected images will be highlighted with a yellow background.
- 2.) Click the  button.

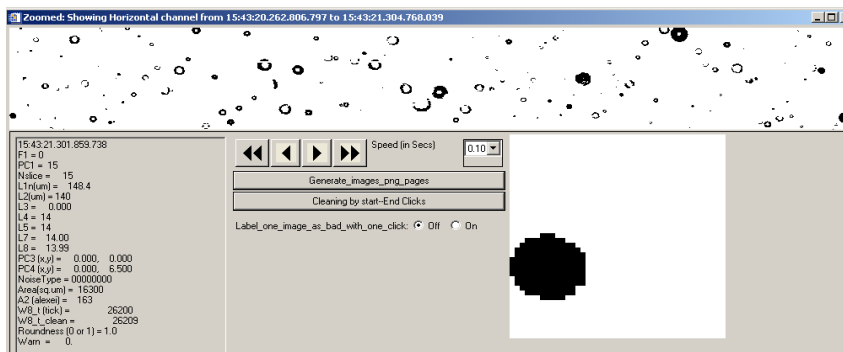


Figure 14

If the filter box is not checked, images will be displayed including: accepted particles, noise, splashes, and shatters. The rejected images are red and are highlighted in yellow as shown in figure 18.

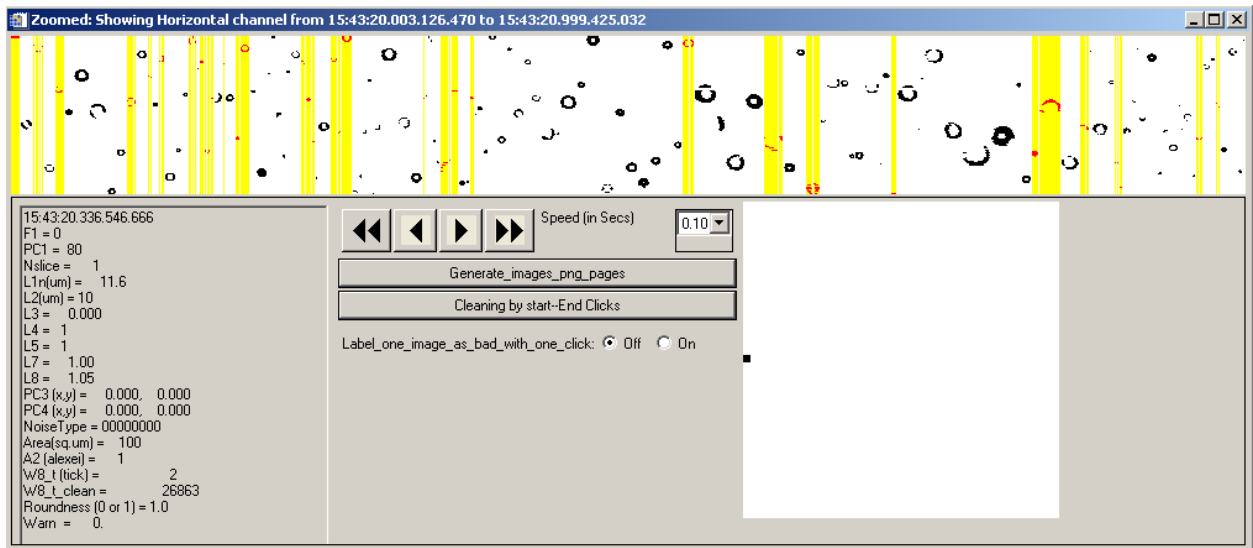
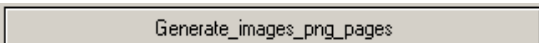


Figure 15

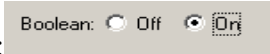
This window has the same functionality as described in section 2.3. Clicking  will cause pages of images to be displayed on the computer screen saved to the data/png directory. Buttons “cleaning by Start-end clicks” and “label\_one\_image\_as\_bad\_with\_one\_click” are for manually removing bad images or noise if the auto-cleaning method doesn’t do a perfect job.

The “Timeshift”, “fishing” and “De\_spur” are special buttons used inside SPEC INC. We use them to look at the laser beam alignment. We also use them to look at the cloud particle special cluster distribution and to study the noise image distribution in the initial stage of the software. They are for development and automatic removal of noise, splashes and shattering. You do not need to use them.

### 2.5.5 Boolean Filtering

The Boolean feature allows for further sub setting of data. It is possible to filter out particles based on size, location, and various other parameters. To use Boolean filtering:

1.) From the zoomed data window select Boolean On as shown:



- 2.) Specify the filtering parameter from the drop down list
- 3.) Use Boolean logic from the drop down list to add filter conditions.
- 4.) If using numerical values, insert a space between the logic and numerical value.
- 5.) Hit UPDATE

The IDL syntax for the Boolean symbols is defined below:

And-And

Or-Or

Not-Not

Eq-Equal to

Ne-Not equal to

Ge-Greater than or equal to

Gt-Greater than

Le-Less than or equal to

Lt-Less than

\*-Multiply

+-Addition

--Subtraction

( )- Used for order of operations

For example, if user types “L1n gt 10” are entered, the filter box is checked and the Update button is clicked, only those accepted and rejected particles with size “L1n greater than 10 pixels” will be selected and displayed in the H-strip and V-strip fields. These will also be used to make the time series and PSD plots. Figure 19 shows an example of when “L1n gt 10” was input for updated processing:

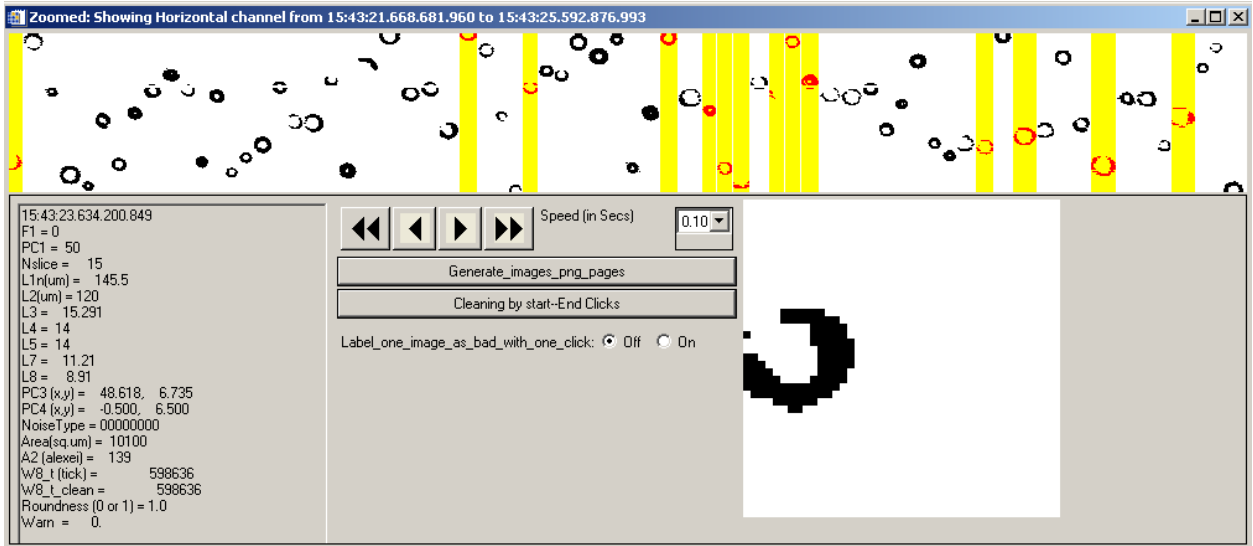


Figure 16

Another way to filter the data in a very specific way is to use what has been named the in-focus only option. To access this, go back to the first main HVPSview window and click the 'use in focus only' check box. You must also select a value from the drop down list. Value options are 0.9, 0.8, and 0.7. This filters the data to use only those images that are at least 90% (80 or 70) shaded. This could be accomplished via the Boolean option as well. The difference is that when using the check box, the sample volume is adjusted using depth of Field values for in-focus particles.

### **3. HVPS data Batch processing (Archiving)**

As described in playback processing section 1.0, after many split sorted files are generated the data can be processed in batch mode, allowing the computer to run over night. When processing batches, all the SORTED files are converted to .BINS and an All Clean filter is applied to them. This allows large quantities of data to be processed and filtered for use at a later date.

The batch processing can be done in Liquid mode or Ice mode. When the user knows if the 2-DS probe is flying through liquid cloud (by looking at the temperature, or by observing the playback image types), the batch processing should be done in Liquid mode. When the probe is flying in the ice cloud, the user should use Ice mode.

### 3.1 Liquid mode Batch processing

#### 3.1.1 Archiving ASCII Files

- 1.) Copy all sorted\*\*\*.HVPS and the associated HK\_bin\*.HVPS, HK\_TXT.HVPS and Log\_TXT\*.HVPS files into the same directory.
- 2.) Get IDL started and follow the steps described in section 1.1.1?
- 3.) The initial GUI screen will appear. Select “archive\_multi\_Liq\_ascii ” from the drop down list. A dialog window will appear for the user to select an input file of type Sorted\*.HVPS, as displayed in Figure 20. Navigate to the directory containing the SORTED\*.HVPS data that requires processing. Click on one of the sorted files, and then click “open” on the lower right corner. That “please select a file” window will disappear.

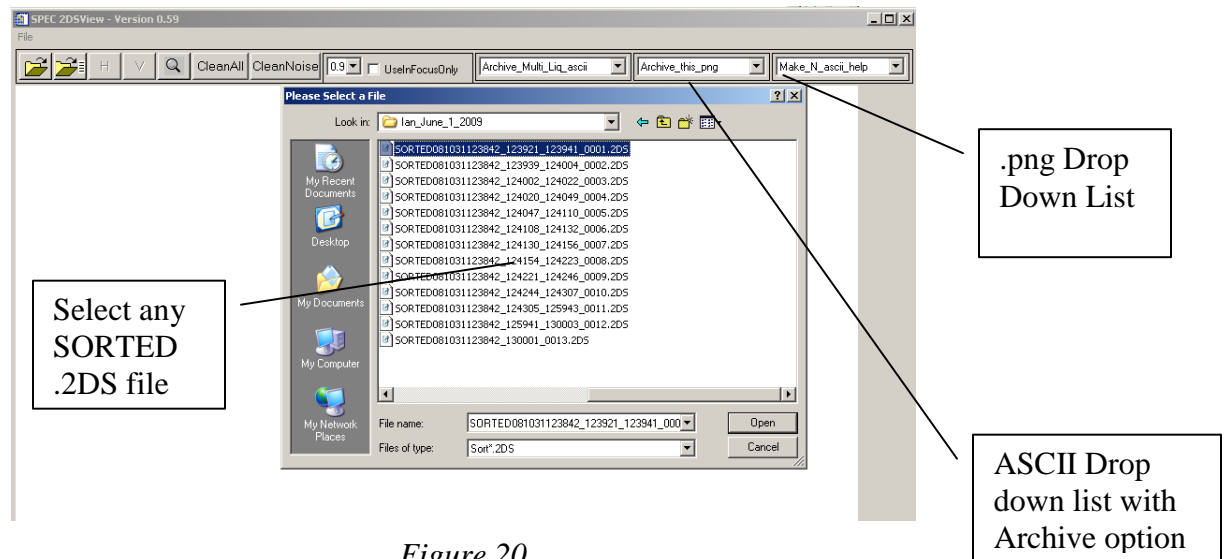


Figure 20

- 4.) Return to the IDL front screen. All of the files for overnight processing will be displayed.
- 5.) If all the files are present, type `y →` Enter in the IDL command line. This will make the computer process all the sorted data files in the selected directory.

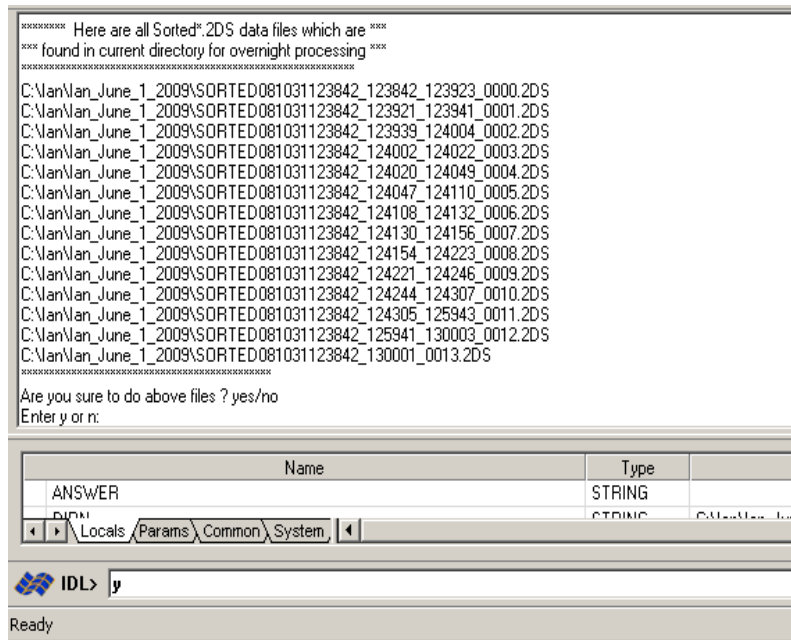


Figure 21

The sorted files will be processed one at a time in order. Some initial plots and warning messages will pop up, be saved, and deleted. The original GUI screen will remain open and the progress is displayed in the status bar. The IDL standard output screen will also display some steps that the archiving process has reached.

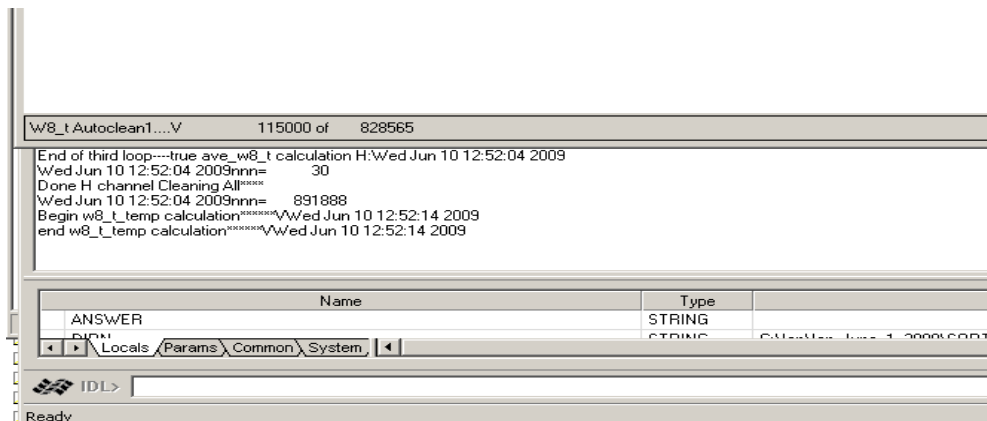


Figure 22

The status bar in Figure 22 displays the following information: 1) The V-channel automatic noise removal is being used. 2) Of the 828565 images, it has examined 115000 particles. 3) The end of the third loop noise removal process for the H channel was done at time: Wed Jun 10, 12:52:04 2009; and etc.

As each SORTED\*\*.HVPS file is processed, many files are generated and saved in sub-directories labeled data1, data2, data3, etc. in the directory where the original SORTED\*\*.HVPS files are stored. These files are initial plots of the file and warning messages encountered when loading the data file, the auto-clean results and the horiz\*\*.bin and vert\*\*.bin files. The most important two files are H\_archive\_HVPS\*\*\*\*\_xxxx\_Ns.txt and V\_archive\_HVPS\*\*\*\*\_xxxx\_Ns.txt. They are located in the subdirectory png under the data1, data2, data3 and etc. These two files contain the 1Hz concentration, extinction, ice water content, time series and particle size distributions. The header lines contain the detail data format of the archive file.

### 3.1.2 Archiving Images

After H\_archive\_HVPS\_\*\*\*Ns.txt and V\_archive\_HVPS\_\*\*\*Ns.txt 1Hz archive files are generated for each sorted\*\*.HVPS file, the HVPSview software can be used on the data. To archive the images:

- 1.) Select “Archive\_Multi\_Liq\_png” from the .png drop down list (see Figure 20).
- 2.) A window will ask how many images per second to archive. This is dependent upon cloud type and computing power. Type the desired number of images per second and press Enter. Only the first images up to that number will be archived each second.

The images are saved in the subdirectory png under the corresponding data1, data2, data3 etc. subdirectory.

### 3.1.3 Combining the archived liquid ascii files

To combine all the smaller archive files into a single data file for the whole flight, a separate IDL program is used. To combine all ascii files follow these steps:

- 1.) Copy all H\_archive\_HVPS\*\*\_Ns.txt into one folder or directory.
- 2.) Start IDL
- 3.) Click File → Open → (Find HVPSview\_ver#) → combine\_Liq\_Ns\_archive\_61bins\_HorV.pro
- 4.) Click Run → Compile combine\_Liq\_Ns\_archive\_61bins\_HorV.pro (Ctrl+F5)
- 5.) Click Run → Run combine\_Liq\_Ns\_archive\_61bins\_HorV.pro (F5)

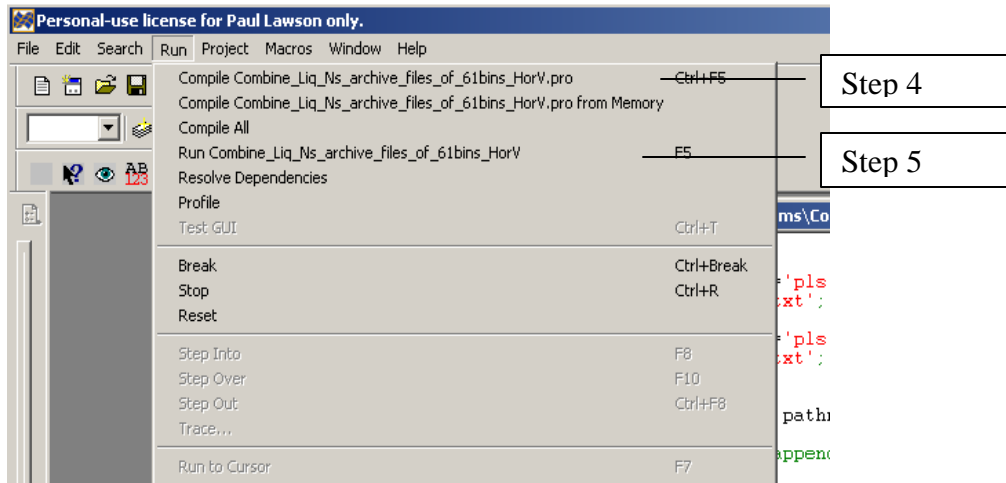


Figure 23

- 6.) You will be prompted to select H or V. This determines whether it is the horizontal or vertical files being combined.
- 7.) Go to the directory containing the files → and select a file.
- 8.) In the IDL command line type y → Enter

From then on, all the H\_archive\_HVPS\_\*\*\*.Ns.txt files will be read and combined in a single archive file, saved in the same directory as the smaller archive files.

### 3.1.4 Additional Notes

To batch archive with the clean option of removing noise only, you should click the “archive\_multi\_Liq\_ascii\_WS” button instead of the “archive\_multi\_Liq\_ascii” button and follow the same instructions above.

To do batch archiving without doing any removal of noise, click the “archive\_multi\_Liq\_ascii\_NC” button instead of the “archive\_multi\_Liq\_ascii” button and follow the same instructions above.

Clicking the “Archive\_this\_liq\_ascii” or “archive\_this\_Liq\_png” in the initial zoomed window will archive the current loaded data file in whatever cleaning state the data is in. If it is cleaned already, it will archive it using only the accepted images. If it is not cleaned, it will use all of the images to do the archive etc.

The “make\_multi\_pbyp\_asciis” is a spare button for future use and it currently has no functionality.

Clicking any of the “Make\_N\_ascii\_Liq”, “Make\_N\_ascii\_ice”, “Make\_N\_ascii\_Liq\_WS”, “Make\_N\_ascii\_ice\_WS”, “Make\_N\_ascii\_NC” will batch-generate many zoomed\*\*\*.txt ascii files that contain information about each



particle in each line for users to do their own data processing and analysis, if they choose to.

## 3.2 Ice mode Batch processing

### 3.2.1 Archiving Ice ASCII Files

1. Copy all sorted\*\*\*.HVPS and the associated HK\_bin\*.HVPS, HK\_TXT.HVPS and Log\_TXT\*.HVPS files into the same directory.
2. Get IDL started and follow the steps described in section 1.1.1
3. The initial GUI screen will appear. Select “archive\_multi\_Ice\_ascii ” from the drop down list. A dialog window will appear for the user to select an input file of type Sorted\*.HVPS as displayed in Figure 20. Navigate to the directory containing the SORTED\*.HVPS requiring processing. Click on one of the sorted files, and then click “open” on the lower right corner. That “please select a file” window will disappear.

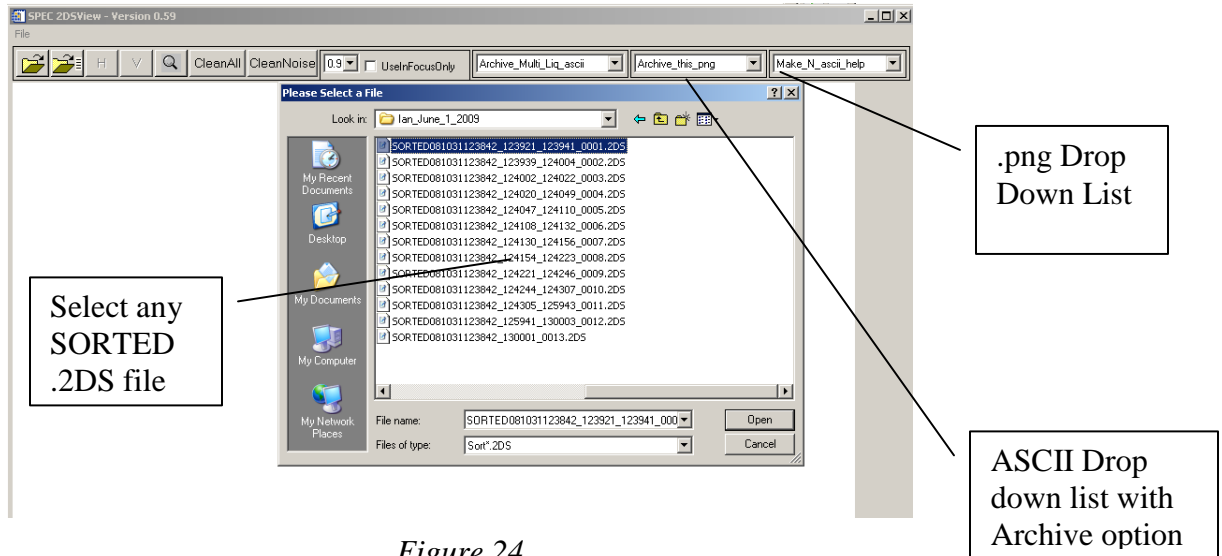


Figure 24

4. A pop up window with the ice archiving options is shown in Figure 25. See Appendix A for a description of the options.

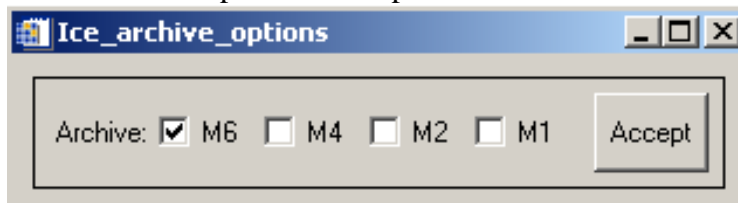


Figure 25

5. Click Accept
6. Return to the IDL front screen. All of the files for overnight processing will be displayed.
7. If all the files are present, type `y` → Enter in the IDL command line. This will make the computer process all the sorted data files in the selected directory.

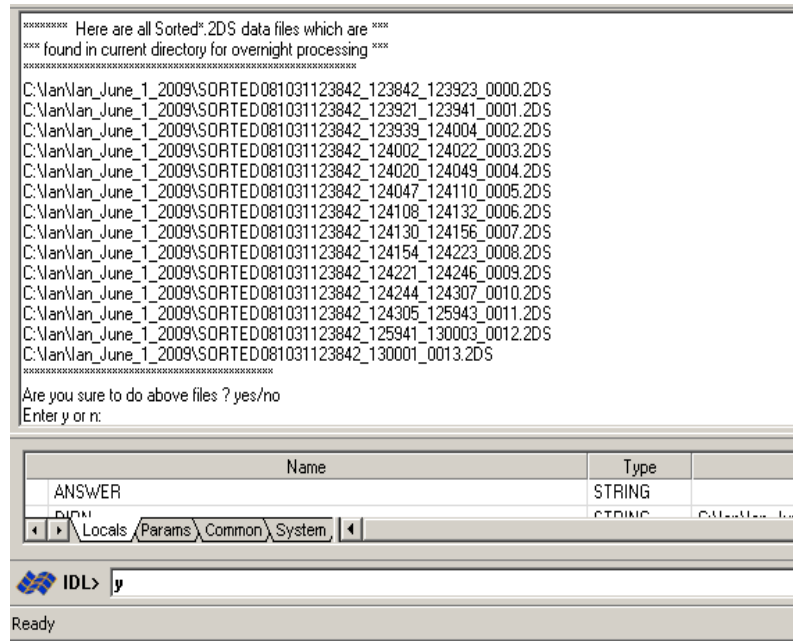


Figure 26

The sorted files will be processed one at a time in order. Some initial plots and warning messages will pop up, be saved, and deleted. The original GUI screen will remain open where the progress is displayed in the status bar. The IDL standard output screen will also print out some of the steps the archiving the process has reached.

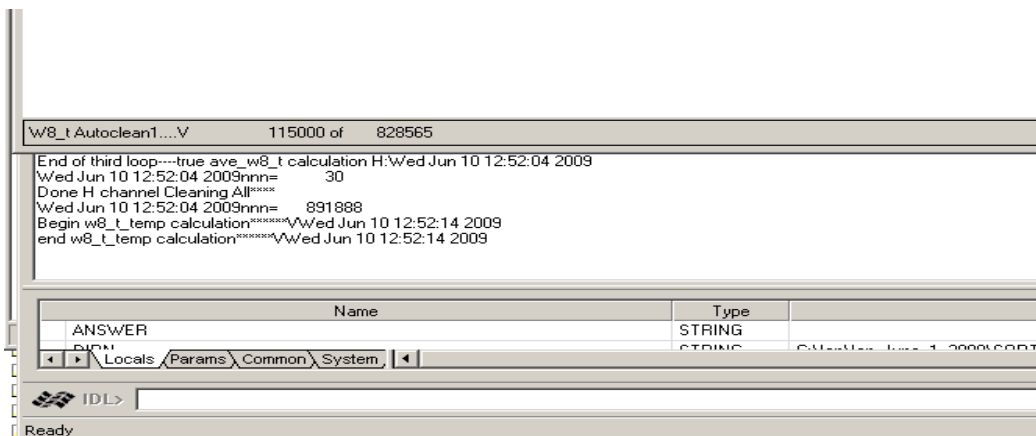


Figure 27

The status bar in figure 27 gives the following information: 1) The V-channel automatic noise removal is being used. 2) Of the 828565 images, it has examined 115000 particles. 3) The end of the third loop noise removal process for the H channel was done at time: Wed Jun 10, 12:52:04 2009.

As each SORTED\*\*.HVPS file is processed, many files are generated and saved in a sub-directory labeled data1, data2, data3, etc. in the directory where the original SORTED\*\*\*.HVPS files are stored. These files are initial plots of the file, warning messages encountered when loading the data file, the auto-clean results and the horiz\*\*\*.bin and vert\*\*.bin files. The most important two files are H\_archive\_HVPS\*\*\*\*\_xxxx\_Ns.txt and V\_archive\_HVPS\*\*\*\*\_xxxx\_Ns.txt. They are located in the subdirectory png under the data1, data2, data3 and etc. These two files contain the 1Hz concentration, extinction, ice water content, times series and particle size distributions. The header lines contain the detail data format of the archived file.

### 3.2.2 Archiving Images

After H\_archive\_HVPS\_\*\*\*Ns.txt and V\_archive\_HVPS\_\*\*\*Ns.txt 1Hz archive files are generated for each sorted\*\*\*.HVPS file, the HVPSview software can be used on the data. To archive the images:

1. Select "Archive\_Multi\_Ice\_png" from the .png drop down list (see Figure 20).
2. A window will ask how many images per second to archive. This is dependent upon cloud type and computing power. Type in the desired number of images per second and press Enter.

The images are saved in the subdirectory .png under the corresponding data1, data2, data3 etc subdirectory.

### 3.2.3 Combining the archived Ice ascii files

To combine all the smaller archive files into a single data file for the whole flight, a separate IDL program is used. To combine all ascii files follow these steps:

1. Copy all H\_archive\_HVPS\*\*\_Ns.txt into one folder or directory.
2. Start IDL
3. Click File → Open → (FindHVPSview\_ver#) → combine\_archive\_TC4\_Steve\_gains\_of\_61bins\_HorV.pro.
4. Click Run → Compile combine\_archive\_TC4\_Steve\_gains\_of\_61bins\_HorV.pro. (Ctrl+F5)
5. Click Run → Run combine\_archive\_TC4\_Steve\_gains\_of\_61bins\_HorV.pro. (F5)

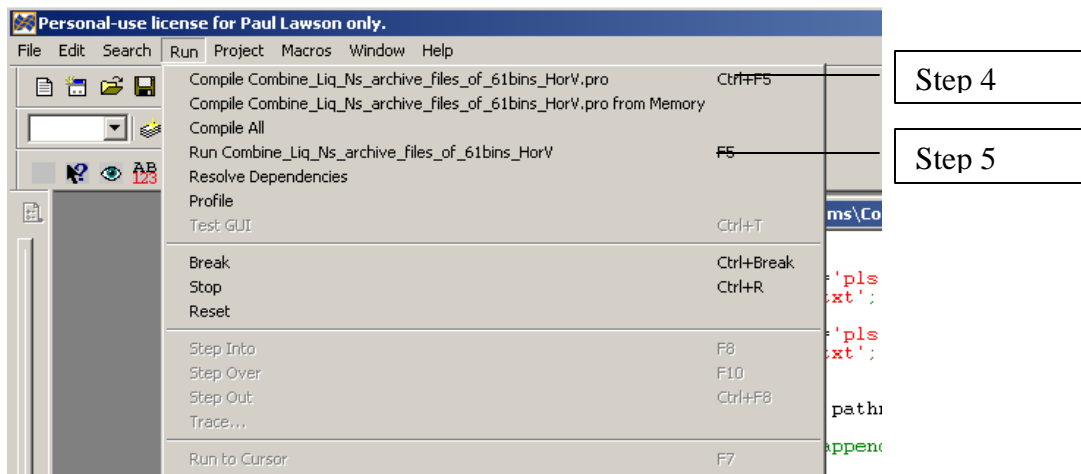


Figure 17

6. A prompt you to select H or V, this determines whether it is the horizontal or vertical files being combined.
7. Go to the directory containing the files → select a file.
8. In the IDL command line type y → Enter

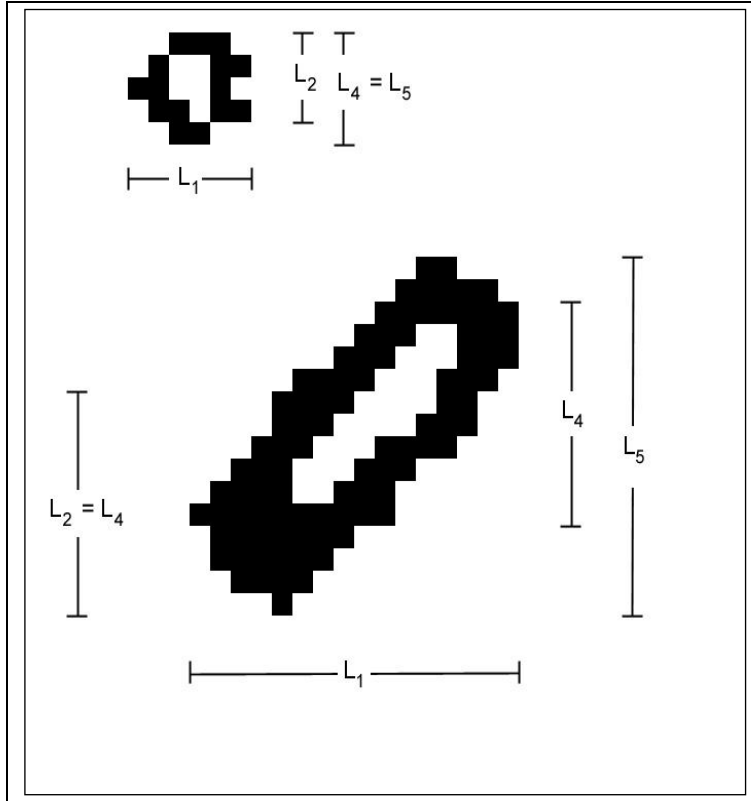
From then on, all the H\_archive\_HVPS\_\*\*\*.Ns.txt files will be read and combined in a single archive file, saved in the same directory as the smaller archive files.

### 3.2.4 Additional Notes

- To batch archiving with the clean option of removing noise only, you should click the “archive\_multi\_Ice\_ascii\_WS” button instead of the “archive\_multi\_Ice\_ascii” button and follow the same instructions as above.

- To do batch archiving without removing noise, you should click the “archive\_multi\_Ice\_ascii\_NC” button instead of the “archive\_multi\_Ice\_ascii” button and follow the same instructions above.
- Clicking the “ Archive\_this\_Ice\_ascii” or “archive\_this\_Ice\_png” in the initial zoomed window will archive the current loaded data file in whatever cleaning state the data is in. If it is cleaned already, it will archive it using only the accepted images. If it is not cleaned, it will use all images to do the archive, etc.
- The “make\_multi\_pbyp\_asciis” is a spare button for future use and it currently has no function.
- Clicking any of the “ Make\_N\_ascii\_Liq”, “Make\_N\_ascii\_ice”, “ Make\_N\_ascii\_Liq\_WS”, “Make\_N\_ascii\_ice\_WS”, “ Make\_N\_ascii\_NC” will batch generate many zoomed\*\*\*.txt ascii files that contain information of each particle, in each line for users to do their own data processing and analysis, if they choose to.

## Appendix A: Derived Parameters, Methods, Equations in HVPSView



**Derived parameters are as follows:** ‘Timestamp’, ‘Index’ of same event’s image in the sorted file, ‘ $A_1$ ’ = number of shaded pixels, ‘ $L_1$ ’ = number of pixels in TAS direction that the event lasted for, ‘ $L_2$ ’ = number of shaded pixels in array direction for the slice during the event for which the number of shaded pixels is maximized, ‘ $L_3$ ’ = the diameter of the circle that just circumscribes the shaded event pixels. ‘ $L_4$ ’ = similar to ‘ $L_2$ ’ except it is the total number of pixels between the shaded end pixels instead of just the shaded pixels. ‘ $L_5$ ’ = is similar to ‘ $L_4$ ’ except it is the distance between the extreme shaded end pixels from all the slices whereas for ‘ $L_4$ ’ it is the end pixels in a given slice. See Figure 1.  $A_3 = (\pi/4)L_3^2$ , ‘ $F_1$ ’ = 0 if neither edge was shaded, 1 if left edge only was shaded, 2 if right edge only was shaded, 3 if both edges were shaded (for any slice during the event), ‘ $Adj_1$ ’ = adjustment factor #1 (defined below), ‘ $Adj_n$ ’ = adjustment factor #n (defined further below), ‘ $PC_1$ ’ = position of particle center on array, determined as the center between the extreme edge points of the slice that determined  $L_2$ . ‘ $PC_3$ ’ = the position of the center of the circumscribing circle. ‘ $PC_4$ ’ = position of particle center on array, determined as the center between the extreme edge points that determined  $L_5$ . For logistic reasons, the derived files will also have TAS included.  $L_6$  = the number of white pixels in the strip(s) that defined  $L_4$ . So  $L_4$  is the strip for which the end

shaded pixels are farthest apart which can be a surrogate for the outer diameter of an out of focus ring.  $L_6$  then is a surrogate for the inner diameter of the ring as it is the number of white pixels for the strip that defined  $L_4$ . If there are two or more strips for which  $L_4$  is maximized then let  $L_6$  be the maximum value for those strips.  $L_7$  is the Alexei scheme corrected size as follows:

First of all use the maximum value of the two inside-perimeter area estimates. Next we simply use the table at the end of Alexei's report. For each image we have a total area ( $A_2 = \max$  of 2 inside perimeter area estimates, from Alexei's Matlab algorithm) and an image area (shaded pixels, also in Alexei's code but we already have this =  $A_1$ ). We take the square root of 1 minus the ratio of the later to the former. We look up the closest value in the appropriate column of the table (that is the Dspot/ Dmax column)(we are using table 40% shadow depth). We find the corresponding value in the 2nd column (that is the Dmax/Do column). We take the inverse of that and multiply by Dmax to obtain Do, which is our corrected size estimate for the bead. If the corrected size (Do) is larger than the original size (Dmax), do not use the correction! use Dmax instead. For round images Dmax might be better estimated from total area but for the bead data we will have to use the along the array size ( $L_4$ ) as the bead's speeds through the probe varied (unless TAS is well adjusted and speeds don't vary too much).

$L_8$  is just as  $L_7$  above except instead of  $D_{\max} = L_4$ , use  $D_{\max} = \sqrt{\frac{4}{\pi} A_2 * \frac{L_4}{L_1}}$ . (in both

$L_7$  and  $L_8$  above we are using  $L_4$  instead of the possibly more logical  $L_5$  because of the tilted arrays giving skewed images...  $L_8$  tries to account for elongated images with the ratio  $L_4: L_1$ )

For both  $L_7$  and  $L_8$  above,  $L_4$  should be used for in-focus images at the 90% level. I. E. after calculating  $L_7$  and  $L_8$ , replace their values with  $L_4$ 's value for the subset of data that is in-focus at the 90% or better level.

**Adj<sub>1</sub>** is calculated as follows:

$$Adj_1 = \frac{SV_{default}}{SV_{1_{actual}}} = \frac{SA_{default}}{SA_{1_{actual}}}$$

$$SV_{default} = TAS \times \Delta t \times SA_{default}$$

$$SV_{1_{actual}} = TAS \times \Delta t \times SA_{1_{actual}}$$

SV stands for sampled volume (Liters), SA for sample area (mm<sup>2</sup>), TAS (m/s) the speed of air through the probe sample area, and Δt (seconds) the time period of the measurement.

$$SA_{default} = 128 \times pxs \times d_{ww}$$

Where ‘pxs’ = 0.01 mm (10 μm) is the optically realized size of the array elements and d<sub>ww</sub> = 63 mm (6.3 cm) is the window to window distance. ‘Strbsz’ is the pixel size in the TAS direction (strobe size) and should also = 10 μm but could differ if the aircraft flies too quickly or an incorrect TAS is sent to the probe during data acquisition.

$$SA_{actual} = (127 + L_1 \times strbsz/pxs) \times pxs \times \min[d_{ww}, d_{of}]$$

where d<sub>of</sub> stands for depth of field and

$$d_{of} (mm) = (F_{dof} \times L_1^2 \times strbsz^2 (\mu m^2)) / 1000.$$

where F<sub>dof</sub> = 5.13 (μm<sup>-1</sup>).

PC<sub>1</sub> can be found as the halfway point between the edge points of the image on the array for the slice that defined L<sub>2</sub>.

PC<sub>3</sub> can be found as the center of the circle that defined L<sub>3</sub>.

### **Methods - Equations:**

**Method 1 (M1):** (TAS direction, all particles): For each time bin (of size Δt),

$$\text{Concentration (\#/L)} = K \times \sum Adj_1 / SV_{default}. \text{ Where K = unit conversion factor.}$$

**Counts time series:** the number of particle events per time bin.

$$\text{Extinction (Km}^{-1}\text{)} = 2 \times K \times \sum (Adj_1 \times Area \times pxs \times strbsz) / SV_{default} \text{ Where K = unit conversion factor.}$$



**IWC: (g/m<sup>3</sup>)** =  $K_1 \times \frac{\sum (Adj_1 \times \min[\alpha \times A^\beta, \rho_i \frac{K_2}{K_1} \frac{\pi}{6} \times (L_1 \times strbsz)^3])}{SV_{default}}$  where A = (Area x pxs x strbsz)/1000000,  $\alpha = 0.115$ ,  $\beta = 1.218$ ,  $\rho_i = 0.917$ , and where  $K_1$  and  $K_2$  are unit conversion factors.

**LWC: (g/m<sup>3</sup>)** =  $K \times \frac{\sum (Adj_1 \times \frac{\pi}{6} \times (L_1 \times strbsz)^3)}{SV_{default}}$ . Where  $K_2$  is unit conversion factor.

**Counts PSD:** For each size bin, the formula for bin height is simply the number of particles in that size bin. Use size bins with width equal to strbsz for now. Bin ( $L_1 \times strbsz$ ) into the size bins. The first size bin has events for which  $L_1 = 1$  and the bin edges are  $0.5 \times strbsz$  and  $1.5 \times strbsz$ , the  $n^{th}$  size bin has events for which  $L_1 = n$  and  $(n - 0.5) \times strbsz$  and  $(n + 0.5) \times strbsz$  as bin edges....

**Concentration PSD (#/L/ $\mu$ m):** For each size bin, the formula for bin height is  $K \times \frac{\sum (Adj_1)}{(bin\_width(\mu m) \times SV_{default})}$  where the sum is over all events in that size bin and Where K = unit conversion factor. Bin as above for counts PSD.

**Area PSD (mm<sup>2</sup>/L/ $\mu$ m):** For each size bin, the formula for bin height is  $K \times \frac{\sum (Adj_1) \times A}{(bin\_width(\mu m) \times SV_{default})}$  where the sum is over all events in that size bin and  $A = (\text{Area} \times \text{pxs} \times \text{strbsz}) / 1000000$  and Where K = unit conversion factor so the units come out. Bin as above for counts PSD.

**Ice Mass PSD (g/m<sup>3</sup>/ $\mu$ m):**

$K_1 \times \frac{\sum (Adj_1 \times \min[\alpha \times A^\beta, \rho_i \frac{K_2}{K_1} \frac{\pi}{6} \times (L_1 \times strbsz)^3])}{(bin\_width(\mu m) \times SV_{default})}$  where  $A = (\text{Area} \times \text{pxs} \times \text{strbsz}) / 1000000$ ,  $\alpha = 0.115$ ,  $\beta = 1.218$ ,  $\rho_i = 0.917$ , and where  $K_1$  and  $K_2$  are unit conversion factors.

**Liquid Mass PSD (g/m<sup>3</sup>/ $\mu$ m):**

$K \times \frac{\sum (Adj_1 \times \frac{\pi}{6} \times (L_1 \times strbsz)^3)}{(bin\_width(\mu m) \times SV_{default})}$  Where K = unit conversion factor.

**Method 2 (M2):** (All in, along array): For each time bin (of size  $\Delta t$ ),

**Concentration (#/L)** =  $K \times \sum Adj_2 / SV_{default}$  where

$$Adj_2 = \frac{SV_{default}}{SV_{2_{actual}}} = \frac{SA_{default}}{SA_{2_{actual}}} \quad \text{where } SA_{2_{actual}} = (127 - L_4) \times pxs \times \min[d_{ww}, d_{of\_2}]$$

and  $Adj_2 = 0$  if either edge pixel is shaded ( $F_1 \neq 0$ ). Where

$$d_{of\_2} (mm) = (5.13(\mu m^{-1}) \times L_4^2 \times pxs^2 (\mu m^2)) / 1000. \quad \text{and where } K = \text{unit conversion factor.}$$

**Counts time series:** the number of particle events per time bin.

**Extinction ( $Km^{-1}$ )** =  $2 \times K \times \sum (Adj_2 \times Area \times pxs \times strbsz) / SV_{default}$  where  $K$  = unit

conversion factor.

**IWC: ( $g/m^3$ )** =  $K_1 \times \frac{\sum (Adj_2 \times \min[\alpha \times A^\beta, \rho_i \frac{K_2}{K_1} \frac{\pi}{6} \times (L_4 \times pxs)^3])}{SV_{default}}$  where

$A = (Area \times pxs \times strbsz) / 10000000$ ,  $\alpha = 0.115$ ,  $\beta = 1.218$ ,  $\rho_i = 0.917$ , and  $K_2$  and  $K_1$  are unit conversion factors.

**LWC: ( $g/m^3$ )** =  $K \times \sum (Adj_2 \times \frac{\pi}{6} \times (L_4 \times pxs)^3) / SV_{default}$ . Where  $K$  = whatever you have

to make it so the units come out.

**Counts PSD:** For each size bin, the formula for bin height is simply the number of particles in that size bin. Use size bins with width equal to  $pxs$  for now. Bin ( $L_2 \times pxs$ ) into the size bins. The first size bin has events for which  $L_2 = 1$  and the bin edges are  $0.5 * pxs$  and  $1.5 * pxs$ , the  $n^{th}$  size bin has events for which  $L_2 = n$  and  $(n - 0.5) * pxs$  and  $(n + 0.5) * pxs$  as bin edges....

**Concentration PSD (#/L/ $\mu m$ ):** For each size bin, the formula for bin height is

$K \times \sum (Adj_2) / (bin\_width(\mu m) \times SV_{default})$  where the sum is over all events in that size bin

and where  $K$  = unit conversion factor.

**Area PSD (mm<sup>2</sup>/L/μm):** For each size bin, the formula for bin height is

$$K \times \sum (Adj_2) \times A / (bin\_width(\mu m) \times SV_{default})$$

where the sum is over all events in that size bin and A=(Area x pxs x strbsz)/1000000 and where K = unit conversion factor.

Bin as above for counts PSD.

**Ice Mass PSD (g/m<sup>3</sup>/μm):**

$$K_1 \times \sum (Adj_2 \times \min[\alpha \times A^\beta, \rho_i \frac{K_2}{K_1} \frac{\pi}{6} \times (L_4 \times pxs)^3]) / (bin\_width(\mu m) \times SV_{default})$$

where A=(Area x pxs x strbsz)/1000000, α = 0.115, β = 1.218, ρ<sub>i</sub> = 0.917, and K<sub>2</sub> and K<sub>1</sub> are unit conversion factors.

**Liquid Mass PSD (g/m<sup>3</sup>/μm):**

$$K \times \sum (Adj_2 \times \frac{\pi}{6} \times (L_4 \times pxs)^3) / (bin\_width(\mu m) \times SV_{default})$$

where K = unit conversion factor.

**Method 3 (M3):** (All in, simple non-adjusted spheres): For each time bin (of size

$$\Delta t), \quad \text{concentration} = K \times \sum Adj_3 / SV_{default} \quad \text{where}$$

$$Adj_3 = \frac{SV_{default}}{SV_{3_{actual}}} = \frac{SA_{default}}{SA_{3_{actual}}} \quad \text{where}$$

$$SA_{3_{actual}} = (127 - L_3) \times \min[pxs, strbsz] \times \min[d_{ww}, d_{of\_3}]$$

and Adj<sub>3</sub> = 0 if either edge pixel is shaded (F<sub>1</sub> ≠ 0). Where d<sub>of<sub>-3</sub></sub> (mm) = (5.13(μm<sup>-1</sup>) × (L<sub>3</sub>)<sup>2</sup> × min[pxs, strbsz]<sup>2</sup> (μm<sup>2</sup>))/1000. and where K = unit conversion factor. Reject events for which L<sub>4</sub> gt 2.1\*L<sub>1</sub> or L<sub>4</sub> lt 0.49\*L<sub>1</sub> and area gt 2.

**Counts time series:** Simply plot the number of accepted particle events per time bin.

$$\text{Extinction (Km}^{-1}\text{)} = \frac{\pi}{2} \times K \times \sum (Adj_3 \times L_3^2 \times pxs \times strbsz) / SV_{default}$$

where K = unit conversion factor. Rejects events for which L<sub>4</sub> gt 2.1\*L<sub>1</sub> or L<sub>4</sub> lt 0.49\*L<sub>1</sub> and area gt 2.

**IWC: (g/m<sup>3</sup>)** = 
$$0.917 \times K \times \sum (Adj_3 \times \frac{\pi}{6} \times (L_3 \times \sqrt{pxs \times strbsz})^3) / SV_{default}$$
 where K = unit conversion factor. Reject events for which L4 gt 2.1\*L1 or L4 lt 0.49\*L1 and area gt 2.

**LWC: (g/m<sup>3</sup>)** = 
$$K \times \sum (Adj_3 \times \frac{\pi}{6} \times (L_3 \times \sqrt{pxs \times strbsz})^3) / SV_{default}$$
 Where K = unit conversion factor. Reject events for which L4 gt 2.1\*L1 or L4 lt 0.49\*L1 and area gt 2.

**Counts PSD:** For each size bin, the formula for bin height is simply the number of particles in that size bin. Use size bins with width equal to  $\sqrt{pxs \times strbsz}$  for now. Bin ( $L_3 \times \sqrt{pxs \times strbsz}$ ) into the size bins. The first size bin has edges of  $0.5 \times \sqrt{pxs \times strbsz}$  and  $1.5 \times \sqrt{pxs \times strbsz}$ , the n<sup>th</sup> size bin has  $(n - 0.5) \times \sqrt{pxs \times strbsz}$  and  $(n + 0.5) \times \sqrt{pxs \times strbsz}$  as bin edges. Reject events for which L4 gt 2.1\*L1 or L4 lt 0.49\*L1 and area gt 2.

**Concentration PSD (#/L/μm):** For each size bin, the formula for bin height is 
$$K \times \sum (Adj_3) / (bin\_width(\mu m) \times SV_{default})$$
 where the sum is over all events in that size bin and where K = unit conversion factor. Bin as above for counts PSD. Reject events for which L4 gt 2.1\*L1 or L4 lt 0.49\*L1 and area gt 2.

**Area PSD (mm<sup>2</sup>/L/μm):** For each size bin, the formula for bin height is 
$$\frac{\pi}{2} \times K \times \sum (Adj_3 \times L_3^2 \times pxs \times strbsz) / (bin\_width(\mu m) \times SV_{default})$$
 where the sum is over all events in that size bin and where K = unit conversion factor. Bin as above for counts PSD. Reject events for which L4 gt 2.1\*L1 or L4 lt 0.49\*L1 and area gt 2.

**Ice Mass PSD (g/m<sup>3</sup>/μm):** For each size bin, the formula for bin height is 
$$0.917 \times K \times \sum (Adj_3 \times \frac{\pi}{6} \times (L_3 \times \sqrt{pxs \times strbsz})^3) / (bin\_width(\mu m) \times SV_{default})$$
 where K = unit conversion factor. Reject events for which L4 gt 2.1\*L1 or L4 lt 0.49\*L1 and area gt 2.

**Liquid Mass PSD (g/m<sup>3</sup>/μm):** 
$$K \times \sum (Adj_3 \times \frac{\pi}{6} \times (L_3 \times \sqrt{pxs \times strbsz})^3) / (bin\_width(\mu m) \times SV_{default})$$
 where K = unit

conversion factor. Reject events for which  $L_4$  gt  $2.1 \cdot L_1$  or  $L_4$  lt  $0.49 \cdot L_1$  and area gt 2.

**Method 4 (M4):** (All In, Ring-Spot Adjusted Spheres A): For each time bin (of size  $\Delta t$ ),

**Concentration** =  $K \times \sum Adj_4 / SV_{default}$  where  $Adj_4 = SV_{default} / SV_{4_{actual}} = SA_{default} / SA_{4_{actual}}$  where

$$SA_{4_{actual}} = (127 - L_7) \times \min[pxs, strbsz] \times \min[d_{ww}, d_{of\_4}]$$

and  $Adj_4 = 0$  if either edge pixel is shaded ( $F_1 \neq 0$ ). Where

$$d_{of\_4} (mm) = (5.13(\mu m^{-1}) \times (L_7)^2 \times \min[pxs, strbsz]^2 (\mu m^2)) / 1000.$$

and where  $K$  = unit conversion factor. Also reject events for which  $L_5$  gt  $2.1 \cdot L_1$  or  $L_5$  lt  $0.49 \cdot L_1$  and area gt 2.

**Counts time series:** Simply plot the number of accepted particle events per time bin.

**Extinction ( $Km^{-1}$ )** =  $\frac{\pi}{2} \times K \times \sum (Adj_4 \times L_7^2 \times pxs \times strbsz) / SV_{default}$  where  $K$  = unit

conversion factor. Reject events for which  $L_4$  gt  $2.1 \cdot L_1$  or  $L_4$  lt  $0.49 \cdot L_1$  and area gt 2.

**IWC: ( $g/m^3$ )** =  $0.917 \times K \times \sum (Adj_4 \times \frac{\pi}{6} \times (L_7 \times \sqrt{pxs \times strbsz})^3) / SV_{default}$  where  $K$  = unit

conversion factor. Reject events for which  $L_4$  gt  $2.1 \cdot L_1$  or  $L_4$  lt  $0.49 \cdot L_1$  and area gt 2.

**LWC: ( $g/m^3$ )** =  $K \times \sum (Adj_4 \times \frac{\pi}{6} \times (L_7 \times \sqrt{pxs \times strbsz})^3) / SV_{default}$  where  $K$  = unit

conversion factor. Reject events for which  $L_4$  gt  $2.1 \cdot L_1$  or  $L_4$  lt  $0.49 \cdot L_1$  and area gt 2.

**Counts PSD:** For each size bin, the formula for bin height is simply the number of particles in that size bin. Use size bins with width equal to  $\sqrt{pxs \times strbsz}$  for now.

Bin ( $L_7 \times \sqrt{pxs \times strbsz}$ ) into the size bins. The first size bin has edges of  $0.5 \cdot$

$\sqrt{pxs \times strbsz}$  and  $1.5 \cdot \sqrt{pxs \times strbsz}$ , the  $n^{th}$  size bin has  $(n - 0.5) \cdot \sqrt{pxs \times strbsz}$  and  $(n$

$+ 0.5) \cdot \sqrt{pxs \times strbsz}$  as bin edges.... Reject events for which  $L_4$  gt  $2.1 \cdot L_1$  or  $L_4$  lt

$0.49 \cdot L_1$  and area gt 2. Do not adjust bin heights by multiplying by the ratio of total events to accepted events.

**Concentration PSD (#/L/ $\mu\text{m}$ ):** For each size bin, the formula for bin height is  $K \times \sum (Adj_4) / (bin\_width(\mu\text{m}) \times SV_{default})$  where the sum is over all events in that size bin and where K = unit conversion factor. Bin as above for counts PSD. Reject events for which  $L_4 \geq 2.1 * L_1$  or  $L_4 \leq 0.49 * L_1$  and area  $\geq 2$ .

**Area PSD ( $\text{mm}^2/\text{L}/\mu\text{m}$ ):** For each size bin, the formula for bin height is  $\frac{\pi}{2} \times K \times \sum (Adj_4 \times L_7^2 \times p_{xs} \times strbsz) / (bin\_width(\mu\text{m}) \times SV_{default})$  where the sum is over all events in that size bin and where K = unit conversion factor. Bin as above for counts PSD. Reject events for which  $L_4 \geq 2.1 * L_1$  or  $L_4 \leq 0.49 * L_1$  and area  $\geq 2$ .

**Ice Mass PSD ( $\text{g}/\text{m}^3/\mu\text{m}$ ):** For each size bin, the formula for bin height is  $0.917 \times K \times \sum (Adj_4 \times \frac{\pi}{6} \times (L_7 \times \sqrt{p_{xs} \times strbsz})^3) / (bin\_width(\mu\text{m}) \times SV_{default})$  where K = unit conversion factor. Reject events for which  $L_4 \geq 2.1 * L_1$  or  $L_4 \leq 0.49 * L_1$  and area  $\geq 2$ .

**Liquid Mass PSD ( $\text{g}/\text{m}^3/\mu\text{m}$ ):**  $K \times \sum (Adj_4 \times \frac{\pi}{6} \times (L_7 \times \sqrt{p_{xs} \times strbsz})^3) / (bin\_width(\mu\text{m}) \times SV_{default})$  where K = unit conversion factor. Reject events for which  $L_4 \geq 2.1 * L_1$  or  $L_4 \leq 0.49 * L_1$  and area  $\geq 2$ .

**Method 5 (M5):** (All In, Ring-Spot Adjusted Spheres B): The same as Method 4 above except replace all  $L_7$  with  $L_8$ .

**Method 6 (M6):** (in focus only, combined): This method is for when there are sufficient particles to use in focus only and when there is need to extend to larger sizes than can be adequately sampled by 'all in' techniques. It uses M2 in-focus-only after CleanAll for sizes smaller than 265. It uses M1 in-focus-only without cleaning for sizes greater than 325. For sizes between 265 and 325, it uses the average of the preceding two methods. A future improvement should be to replace the M1 part with a reconstruction technique.

## **Appendix B: Removing Spurious HVPS Events**

**HVPS raw data include spurious effects. These are primarily from instrument noise and from splashing of precipitation. Algorithms used to clean the data of the majority of these spurious effects while retaining the majority of the valid images are described here.**

**There are 5 quasi independent steps to the algorithm implemented via two loops through the data:**

### **First Loop:**

- 1) Test for roundness.**
- 2) Test for splashing events based on black and white area considerations.**
- 3) Test for noise via line and dot patterns.**
- 4) Test for noise via statistics of particle center locations.**

### **Second Loop:**

- 5) Test for splashing events based on inter-event-distances if the probe is in precipitation.**

**Each step will be described in detail after defining terms:**

**Four measures of image length are shown in figure A.1.** ' $L_1$ ' is the number of slices (pixels in the direction of travel) for which the event lasted. ' $L_2$ ' is the number of shaded diodes (pixels in the direction along the array) for the slice for which the same quantity is maximized. ' $L_4$ ' is the number of diodes between, and including the shaded end diodes, for the slice that maximizes the same quantity. ' $L_5$ ' is the distance between (and including) the shaded end diodes considering all of the slices together.

Other size parameters are a corrected size for out of focus images based on Korolev's work ( $L_7$ ), the number of shaded pixels for the entire image (summed over all slices) ( $A_s$ ) and an estimate of total number of pixels, shaded or not, for the entire image ( $A_t$ ). The algorithm for  $A_t$  is part of the Korolev size correction method.

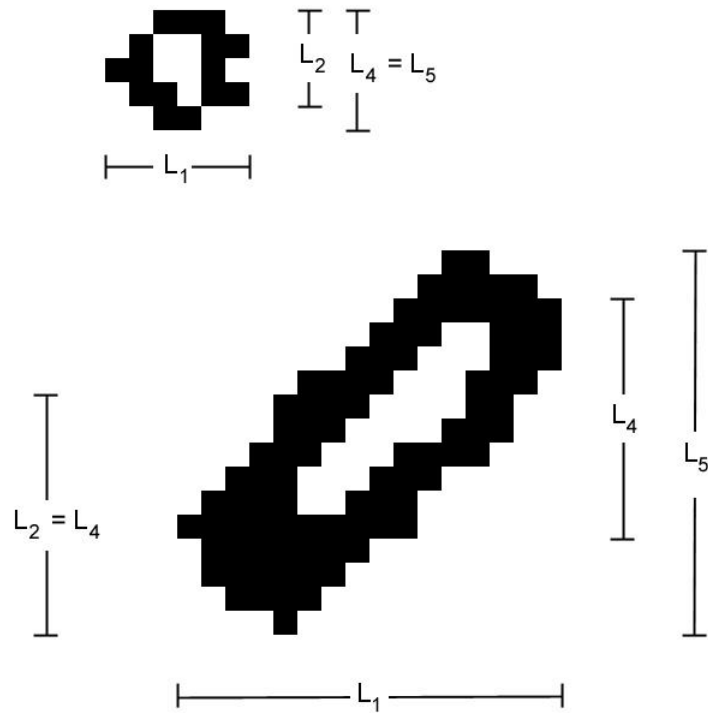


Figure A.1: Two example particle images designed to demonstrate the four measures of image size described in the text.

1. An image is determined to be round if  $L_1 \geq 0.5 \times L_5$  and  $L_5 \geq 0.5 \times L_1$ . If an image is not round, it is rejected. An exception is made for the very large images that do not fit within the array. If  $L_5 > 50$  (500  $\mu\text{m}$ ) then an image is determined to be round if  $L_1 \geq 0.5 \times L_5$ . I. E. an image is determined to be round if  $L_1 \geq 0.5 \times L_5$  and ( $L_5 \geq 0.5 \times L_1$  or  $L_5 > 50$ ).
2. The larger a valid image is the greater the percentage of shaded pixels. Splash effects often create large images that have lower percentages of shaded pixels than valid images. The follow criteria are used to eliminate such spurious images. If any of these five criteria are met the image is rejected.
  - ( $L_5 > 10$  or  $L_1 > 10$ ) and ( $A_t > 3.0 \times A_s$ )
  - ( $L_5 > 15$  or  $L_1 > 15$ ) and ( $A_t > 2.5 \times A_s$ )
  - ( $L_5 > 20$  or  $L_1 > 20$ ) and ( $A_t > 2.0 \times A_s$ )



- $(L5 > 35 \text{ or } L1 > 35) \text{ and } (At > 1.5 \times As)$
- $(L5 > 2.2 \times L7) \text{ and } (L1 > 10.0)$

3. Figure B.2 shows some examples of noise generated images appearing in line plus dot patterns.

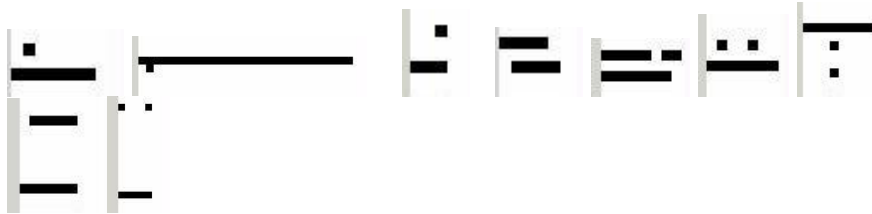


Figure B.2: Examples of line plus dot patterns caused by noisy diodes.

These are eliminated using the following criteria:

- $(L1 = As) \text{ and } (L2 = 1) \text{ and } (L1 > 4)$
- $(As \leq 1.35 \times L1) \text{ and } (L4 = L5) \text{ and } (L1 > 4) \text{ and } (L2 = 2)$
- $(L1 > 10) \text{ and } (L1 > 0.75 \times As) \text{ and } (L1 \leq 1.5 \times As)$
- $(L4 = L5) \text{ and } (At > 0.9 \times L1 \times L5) \text{ and } (L2 = 2) \text{ and } (L2 \neq L4)$
- $(L4 = L5) \text{ and } (At > 3.0 \times As) \text{ and } (L2 = 2)$
- $(L4 = L5) \text{ and } (At > 4.0 \times As)$

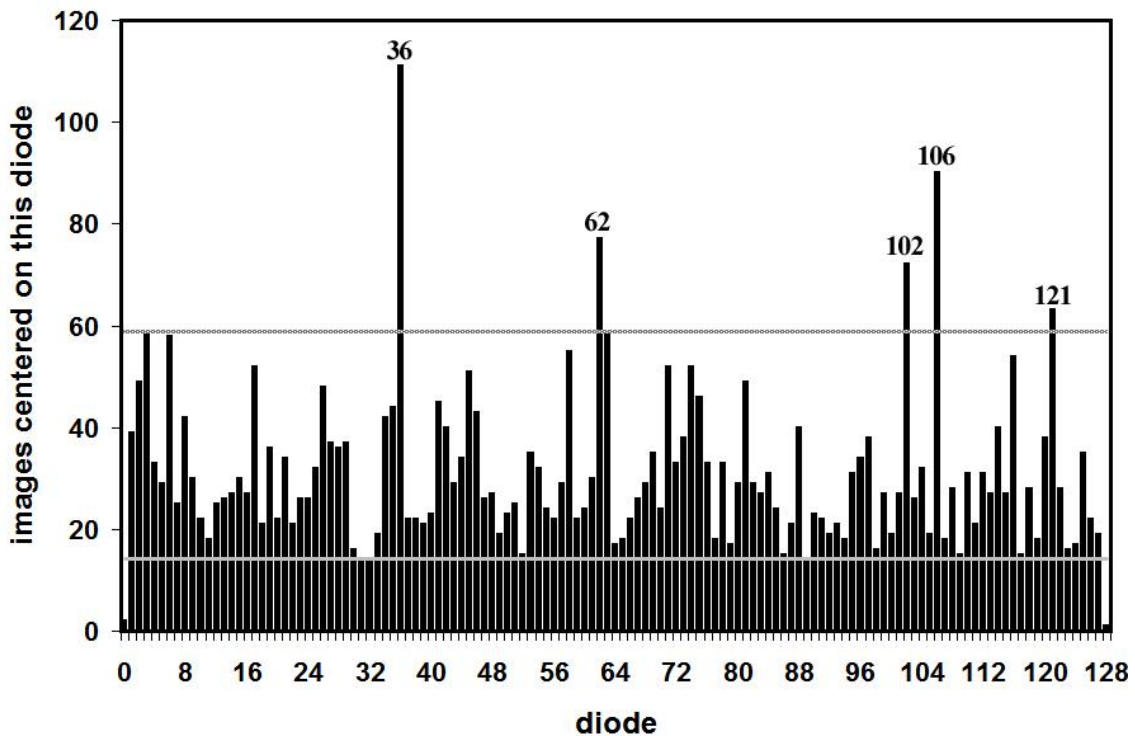
If any of these six criteria are met the image is rejected.

4. The final step, of this first cleaning loop, is another noisy diode removal method, based on the statistics of image center locations calculated over 4000 images approximately centered on the image being evaluated. When a diode is noisy there are more image centers located on that diode than are located on quiet diodes. A diode is labeled bad when it has more particle center locations than the threshold value ( $TH = \max(M + 5 \times \sqrt{M}, 1.5)$ ). Where M is the mean number of image centers per diode across the array, for diodes having more than a minimum number of image center counts.

This minimum number is  $Mt - 3 \times \sqrt{Mt}$ , where  $Mt$  is the mean number of image centers per diode across the array, for all diodes. If less than 33 diodes satisfy the requirement of having more than  $Mt$  counts then  $M$  is calculated from all diodes.  $Mt$  and  $M$  are recalculated after bad diodes are identified, ignoring those diodes, and the process repeated until no diodes exceed the threshold. An image whose center falls on a bad diode is rejected unless it meets one of the following criteria:

- i.  $(L1 \geq 15)$  and  $(L5 \geq 15)$  and  $(As > 0.7 \times At)$
- ii.  $(L1 \geq 4)$  and  $(L5 \geq 4)$  and  $(As \geq 0.5 \times At)$  and  $(L2 > 0.25 \times L5)$  and  $(L1 < 50)$
- iii.  $(L2 = L4)$  and  $(L2 = L5)$  and  $(L2 \geq 2)$  and  $(As \geq 0.5 \times At)$

Figure B.3 shows an example of noisy diode data and the distribution of image centers across the array.



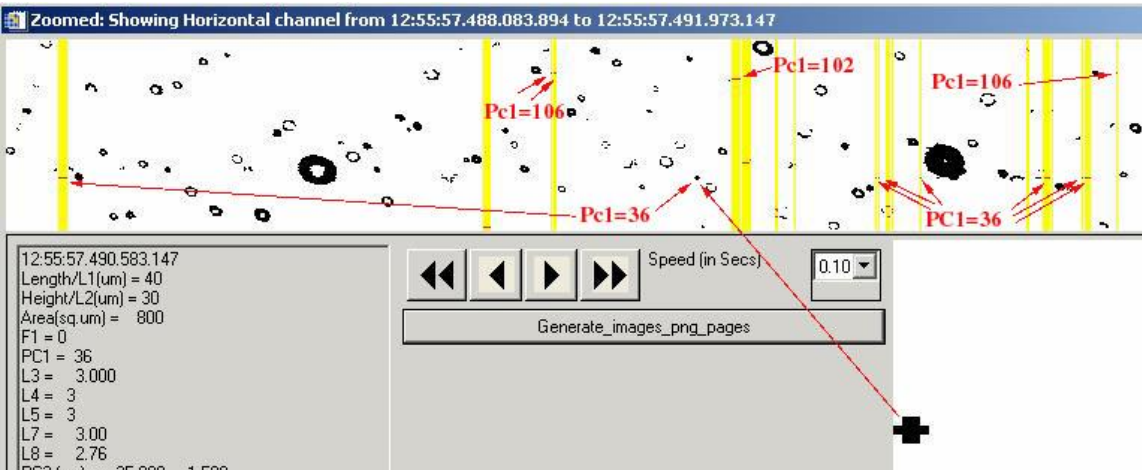


Figure A.3: Shows an example of noisy diode data intermixed with good particle data. The images highlighted in yellow are rejected. Labels in red point to images centered on diodes determined to be bad (too noisy) by the criteria described in the text and exemplified in the particle center location distribution shown above.

1. This final step is another loop through the data but is applied only if the current precipitation status is 'yes'. The current precipitation status is 'yes' if there is one or more particle(s) with L7 gt 100 microns in the 10000 particles, accepted by the previous 4 cleaning steps, centered approximately on the current particle. It is approximately centered because the precipitation status is updated only every 100 particles<sup>2</sup>, instead of for every particle, to improve speed. Another parameter, the current mean inter-particle time (aveW8), is similarly updated every 100 particles<sup>1</sup> and calculated as the average inter-particle time between the 10000 images, accepted by the previous 4 cleaning steps, centered approximately on the current particle. All the particles in the previous second are used instead of 5000 particles if there are less than 5000 particles in the previous second, similarly for the following 5000 particles or second. If  $\text{aveW8} > 40000$  then  $C_1 = 2000$  and  $C_2 = 8000$ . If  $\text{aveW8} < 40000$  then  $C_1 = 0.05 \times \text{aveW8}$  and  $C_2 = 0.2 \times \text{aveW8}$ . A particle is rejected if the precipitation status is 'yes' and either: a) the particle's inter-particle time or the following particle's inter-particle time is less than  $C_1$  or b) both the particle's inter-particle time and the following

<sup>1</sup> The precipitation status and current mean waiting time are updated every particle, instead of every 100, if the valid particle rate is slower than 1000 per second, averaged over the surrounding 0.2 seconds.

particle's inter-particle time are less than  $C_2$ . Finally, a variable,  $adj_0$ , is calculated for each image. It is 0 for rejected images, 1 for accepted images if precipitation status is 'no'. When precipitation status is 'yes', then for accepted images, we calculate a new true average inter-particle time (TaW) = the average inter-particle time of particles accepted by all the steps minus  $C_1$ . Then calculate

$$k = 1 - e^{-2C_1/TaW} + e^{-2C_1/TaW} \left(1 - e^{-C_2/TaW}\right)^2$$

$k$  is limited to a maximum of 0.9 and then  $adj_0 = \frac{1}{1-k}$ .  $adj_0$  is used elsewhere, in the algorithms for calculating concentrations etc., to increase the weight of each accepted particle to account for those good particles rejected by the inter-particle time criteria.

$k$  is the fraction of events eliminated by the criteria given a true waiting time distribution with mean of TaW.  $1 - e^{-C_1/TaW}$  is the probability of, or fraction of events with, inter-arrival times less than  $C_1$ .  $1 - e^{-2C_1/TaW}$  is the fraction of events eliminated due to the first criteria, leaving  $e^{-2C_1/TaW}$  not eliminated by the first criteria. Of those, a fraction  $\left(1 - e^{-C_2/TaW}\right)^2$  are eliminated by the second criteria.