# The Wibs AnalysiS Program
# User Manual

N.H. Robinson

September 6, 2012

niall.robinson@manchester.ac.uk

# Contents

# Introduction

The WIBS AnalysiS Program (WASP) is a complete software suite for off-line analysis of data from the Waveband Integrated Bioaerosol Sensor (Wibs). It load and applies corrections to the logged data to provide quantitative products such as concentration time series and size distributions.

# 1 A step-by-step guide to analysis

## 1.1 Loading data

The *Load* tab allows the raw data recorded by the Wibs to be loaded and averaged into time series which can the be analysed in different ways in the subsequent tabs.

### 1.1.1 Load settings

For datasets with dual gain mode data, the *Gain Mode* box contains radio buttons to specify which gain mode data to load. Only data from this mode will be used in all subsequent analysis.

The *Load Parameters* box contains the settings that affect loading of the Wibs data.

***Time Res (min)* value** sets averaging time for Wasp time series

***Flow (l/min)* value** sets what the instrument sample flow rate was

***Min size (um)* value** sets a threshold minimum size, below which measurements are rejected as noise

***Custom Bins* button** allows the user to edit the size, AF, and fluorescence bin values

***Fit to time grid* check box** forces the time series to start at an integer time point

***Recalc sizes* check box** recalculates the particle sizes from scattering measurements. Edit the `calcSizeFromScatter()` function to change the size calculation. This function can be found in the Igor Procedure window (Ctrl+M).

***Remove saturated* check box** discards any measurement which contains saturation in any fluorescence channel

***File ending* text box** defines the Wibs data file ending ("txt" or "csv")

The *Load Dist.* box defines which measured variables are loaded into distributions. Note that size is permanently selected, meaning that all measurements are resolved as size resolved data. The user can also choose to load FL1, FL2, FL3, or AF resolved data.

The *interp* and *avg* radio buttons in the *BL* box set weather the fluorescence baseline and threshold are interpolated between forced trigger measurements, of just the average of all the measurements (see next section). This choice can be made by inspection of the

amount of structure in the baseline time series displayed after loading the forced trigger data. If a manual baseline has been set then these radio buttons have no effect.

### 1.1.2 Fluorescence baseline

A fluorescence baseline (BL) must be set before any data is loaded into Wasp. This baseline allows Wasp to account for any fluorescence exhibited by the interior of Wibs detection region, rather than the particles themselves. It is determined by measuring fluorescence when no particles are present, so called *forced trigger* measurements. These measurements are also used to set fluorescence level above which a particle is considered to be fluorescent.[1] The baseline and threshold values can be set in two ways:

***Load BL* button** lets the user specify a data folder containing Wibs data files which contain forced trigger measurements (files may also contain normal measurement data). After the forced trigger data is loaded, a plot will appear showing a time series of baseline and threshold values.

***Manual BL* button** lets the user manually enter baseline and fluorescence threshold values.

### 1.1.3 Loading data

Once the baseline is set and the load parameters have been entered, the data may be loaded using the *Load data* button. Particles are classified into different types as they are loaded. The default particle types are:

**All** which includes all particles (that have satisfied the Load Parameters)

**Fl** which includes all particles which have fluorescence above the fluorescence threshold in at least one channel

**NonFl** which includes all particles which do not have fluorescence above the fluorescence threshold in any channel

**FL1** which includes all particles which have fluorescence above the fluorescence threshold in FL1

**FL2** which includes all particles which have fluorescence above the fluorescence threshold in FL2

**FL3** which includes all particles which have fluorescence above the fluorescence threshold in FL3

Once the blue progress bar indicates that the loading has completed, the concentration time series may be inspected using the *View series* button. The user may select which type of time series to view from the list on the left, and may shift+select to display multiple series.

[1] The fluorescence baseline is defined as the mean forced trigger fluorescence plus three sigma.

## 1.2 To Do lists

### 1.2.1 Creating new to do lists

Different periods of interest may be defined by setting *To Do lists* using the right click menu on some data plots, and the controls in the *To Do* box on the right of the panel. A period of interest may be defined by left clicking+dragging to draw a marquee on any data plot with a time x-axis. The marquee details can then be input to the *To Do* box controls by right clicking on the marquee and selecting *Set ToDo list*. The *Start pt* and *End pt* boxes update to display the start and end point of the to do list. A suitable name can be entered in the *Name* box, and a new to do list created by clicking the *New* button, which is then added to the *To Do* drop down list. Any subsequent data analysis or product is only performed for data included in the to do list selected in the *To Do* drop down list. To do lists can be deleted by selecting them from the drop down list and clicking *Del*.

To do lists may also be defined according to particle size. A to do list may be created by drawing a marquee on any plot with a size x-axis. The size range is then highlighted in the size bin list in the *To Do* box. If a to do list is set using a plot with a time x-axis and a size y-axis, then the to do wave will be defined by both time and size. To do wave details may also be entered manually in to the *To Do* box controls.

Different to do lists may be logically combined two at a time using the *Name* field. Entering `<to do A> AND <to do B>` and clicking the *New* button creates a to do list of measurements in both to do list A and to do list B (with the name entered in a pop-up box). Using the logic word `OR` creates a to do list of measurements in either to do list A or to do list B, and `NOT` creates a to do list of measurements in to do list A that are not also in to do list B.

### 1.2.2 Blacklisting

Any periods where data is considered to be spurious may be removed from all subsequent analysis by drawing a marquee and selecting *Blacklist*, similarly to the to do wave creation described above. This can be performed based on time period or size range.

## 1.3 Single particle data

In addition to the time/size/fluorescence/AF binned data loaded using the *Load* tab, single particle data (SPD) may also be loaded using the *Scatter SPD* tab. Note that only data in the selected to do list are loaded. The controls along the top of this tab can be used to load and display this single particle data:

***% SPD to load* value** sets the proportion of particles to load as single particle data. Note that loading too many single particle data may cause the program to run out of memory.

***raw* check box** load data exactly as it is recorded in the data file, without applying corrections such as the minimum size threshold and subtracting the fluorescence baseline.

***Load* button** Loads the single particle data

Once the data are loaded, the variable names are displayed in the two list boxes in the middle of the tab

***Delete* button** deletes the variables selected in the left hand box

***Plot selected* button** scatter plots the variables selected in the left hand box, against the variable selected in the right hand box. Note that time is displayed as a number[2]. Multiple variables can be selected in the left hand box by shift+selecting, but only one variable can be selected in the right hand box.

Additional series (for instance relative humidity) can be added to the list of available variables using the controls at the bottom of the tab:

***Select a data wave* button** allows the user to select the data wave

***Select a time wave* button** allows the user to select the corresponding time wave

***Import series* button** then adds a time series of these waves to the variable list

## 1.4 Time series

Different time series can be calculated and displayed using controls in the *Series* tab. The particle type to display is selected from the list on the left of the tab, with shit+selecting allowing the selection of multiple particle types which lead to the simultaneous generation of multiple time series. The type of time series (total number/surface area/volume) may be chosen using the radio buttons. The buttons at the bottom of the panel generate and display the according time series:

***Time series* button** displays simple time series

***Diurnal time series* button** displays diurnal time series box and whisker plots

***Image plot* button** displays size resolved (3D) time series

***Diurnal image plot* button** displays size resolved (3D) diurnal average time series

## 1.5 Distributions

Average distributions may be generated using the controls in the *Distris* tab. The particle types to display may be selected from the list on the left, with shift+clicking allowing the selection of multiple types. The radio buttons can then be used to define the type of distribution, including: number, size, and volume vs size distributions; number vs fluorescence channel distributions; and number vs AF distributions. Note that in order to calculate fluorescence and AF distributions, the data must be selected to be loaded in the *Load* tab settings. Average distributions (for all data in the selected to do list)

[2]Igor Pro time i.e. seconds since Jan 1st 1904

may then be generated using the *View Distribution* button, with plots showing mean (solid line), median (dashed line), 25th/75th percentile and 10th/90th percentile (shaded regions). Size/time resolved data may be displayed using the *View image plot button* (similarly to the button in the *Series* tab).

## 1.6 Cluster Analysis

Cluster analysis can be performed using the controls under the *CA* tab. This analysis is performed on the single particle data loaded using the *Scatter SPD* tab. The variable to analyse (including externally loaded variables) may be selected by shift+clicking in the list box on the left. The analysis can then be performed using the other tab controls:

***View input stats* button** displays "quick-look" data to allow the user to assess the potential for a significant clustering solution using the given data. A matrix of plots detailing the relationship between the selected variables is displayed. This consists of scatter plots of each variable with each other variable, and variable histograms. If the data can clearly be seen to cluster here, then a very robust cluster analysis solution will be reached, however, a lack of apparent particle clusters does not preclude a significant solution.

**distance metric radio buttons** select which clustering distance metric to use

***Norm to scat* check box** normalises the fluorescence measurements to the scattering signal

***Do cluster analysis* button** performs the cluster analysis. The time taken for this step goes as the square of the number of particles, and can take a very long time.

***Calc stats* button** calculates statistical scores for each cluster analysis solution (i.e. number of clusters), up to the number of solutions defined in the *up to sol* box

***View solution* button** calculates the cluster centroids for the solution (i.e. number of clusters) set in the box. If the *Plots* check box is not selected then the centroids are calculated using the data means. If it is checked, then a Gaussian fit is made to data variable histograms. These fits are sequentially checked by the user to see if they consider them to be reflective of the data, and, if passed, the mode centre is used. Otherwise the mean is used. When this process is complete, as table is displayed showing the cluster centroids for each variable, and the number of constituent members. The relative standard deviations can be viewed by pressing the arrow in the top right corner of the table (to change matrix layers).

***Assign remaining* button** reloads the data, assigning each particle to the most similar cluster. Only clusters specified in the semi-colon separated list are used. Once this step is complete, the cluster time series are added to the particle type lists, allowing regular analysis to be performed.

# 2 Program functions

## 2.1 waspPanel.ipf

This file contains all the routines that set up, initiate, and operate the Wasp panel.

## 2.2 waspLoad.ipf

This file contains all the routines that load Wibs data.

## 2.3 waspStats.ipf

This file contains all the routines that deal with principle component analysis and cluster analysis.

## 2.4 waspMechanics.ipf

This file contains tool functions that are used in the other three files.