SEDANAL

USERS' MANUAL

Walter F. Stafford
and
Peter J. Sherwood

12 Francis Avenue
Cambridge, MA 02138

stafford@sedanal.org
sherwood@computer.org

28 June 2022
SEDANAL v7.76

[A4]

This will print as-is on A4 paper.
Reduce to 94% for 8.5 x11" paper size

.

TABLE of CONTENTS

1 Introduction and Setup

SEDANAL is a computer program for analysis sedimentation velocity and sedimentation equilibrium data that are produced by the BeckmanCoulter Optima Series XLA/I ProteomeLab analytical ultracentrifuge system or the Beckman Optima AUC, or the Open-Multiwavelength machines available in Germany.

1.1 Background

Starting from the raw data files produced by intensity, absorbance, interference, fluorescence, or multi-wavelength intensity, absorbance and fluorescence experiments, the SEDANAL sedimentation velocity module (**SedVel)** uses numerical solutions to the Lamm equation along with chemical kinetics and equilibrium models to fit parameters to the experimental data. SEDANAL can also produce simulated data in the same format as the raw data files produced by the XLA/I optical systems. SEDANAL fits to time difference curves to completely remove time invariant systematic errors from the data. It also allows preprocessing of the data to essentially eliminate time dependent, radially independent systematic instrumental errors characteristic of the interference optics as well as integral fringe jumps that can occur at the meniscus.

The SEDANAL sedimentation equilibrium module (**SedEq)** fits sums of exponentials to sedimentation equilibrium data to obtain estimates of molecular masses and equilibrium constants for interacting and relative amounts for non-interacting systems.

The output of the SEDANAL **SedVel** curve fitter is a report file containing the root mean squared deviation (rmsd) and best-fit values for the parameters, as well as the experimental and calculated difference curves and the residuals of the fit for each of the difference curves. The user may specify which parameters are to be fitted and which to be held constant on the Control Screen

The output of the SEDANAL **SedEq** curve fitter is a file containing the root mean squared deviation (rmsd) and best-fit values for the parameters, as well as the experimental and calculated equilibrium concentration curves and the residuals of the fit for each of the curves. The user may specify which parameters are to be fitted and which to be held constant.

SEDANAL also writes a detailed Report File containing all the input data, model fitted and parameters returned.

1.2 The SEDANAL software package

The software package is CPU-intensive: it runs best on the fastest possible multi-core CPU with lots of memory and at least 100 Mb of the hard drive space for installation. SEDANAL runs under Windows 7 64 bit, Windows 10 64 bit, and Windows 11. This version does many of the computations in parallel. To take full advantage of parallel computing, the machine should have as many cores as datasets being fit at any one time.

1.3 INSTALLATION: Program and Data Files – paths

1.3.1 The default installation:

SEDANAL expects to find a specific path structure for files, (This can be changed to some extent in the Preferences menu). Most screens now have a "Browse" function , as well.

The top level directory (also called a "folder") may be named anything you wish; we'll call it "SEDANAL". In the "SEDANAL" directory, create three subdirectories called "**ModelEditor**", "**Program**" and "**User_data**".

The file "**ModelInfo.txt**", if it already exists, must be placed in the "**ModelEditor**" directory. If it doesn't exist, for example in a new installation, SEDANAL will create an empty ModelInfo.txt file for you in the "**ModelEditor**" directory. The application "**SEDANAL.exe**" must be placed in the "**Program**" directory along with "**Preferences.txt**" If preferences.txt is omitted, SEDANAL will create a default preferences file. There is also a **DLL** named **"libiomp5md.dll"** (for all versions after 6.01) that must be placed along with SEDANAL.exe in the Program folder. A help file for the on-line help also goes in the Program folder. It is named **"SEDANAL.chm"**. If you downloaded the SEDANAL package from the internet, the SEDANAL.chm file may be blocked. It must be un-blocked by changing its Properties.

The "**User_data**" directory will contain sub-directories in **3 levels**. At the highest level within "**User_data**", any number of experiment directories can be created. Each experiment directory, in turn, **must** contain the data directories in the standard BeckmanCoulter XLA/I date/time format. See Figure 1-1 for an example. Usually one will store control files (which have extension ***.abc**) in an experiment folder. In this case (see example below) it would be the "experiment 1" folder. The run files (with extension ***.abr**) are automatically stored at the level of the XLA/I data files.

1.3.2 Alternate installation methods

The location of both the User_data directory and the ModelEditor directory can be specified in the Preferences:

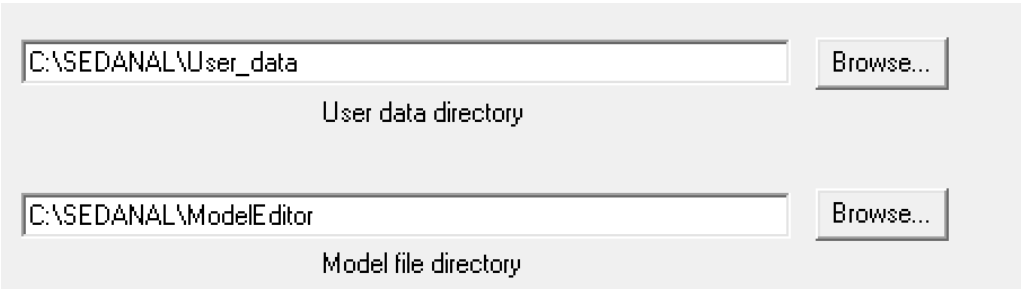


Figure 1-1 Setting up paths to the data files and to the model file.

The "User_data" folder can have any name and be located where-ever you want--like a network server over a VPN, for example. The "ModelEditor" folder also can have any name and be stored anywhere. You might have

multiple “User_data” folders, each with multiple experiment folders and multiple “ModelEditor” folders depending on the analysis du jour. The SEDANAL executable must be located on the local machine, however.

1.3.3 Default path structure for SEDANAL

1.3.3.1 In summary, for getting started, the default directory structure should look like:

C:\SEDANAL\ModelEditor\ModelInfo.txt

C:\ SEDANAL\Program\ SEDANAL.exe
C:\ SEDANAL\Program\Preferences.txt
C:\SEDANAL\Program\

C:\ SEDANAL\User_data\my_experiment#1\mmddyy\hhmmss*.IPn
C:\ SEDANAL\User_data\my_experiment#2\mmddyy\hhmmss*.IPn
C: \SEDANAL\User_data\my_experiment#3\mmddyy\hhmmss*.RAn
C: \SEDANAL\User_data\my_experiment#4\mmddyy\hhmmss*.FLn

1.3.4 A graphical representation of this path structure:

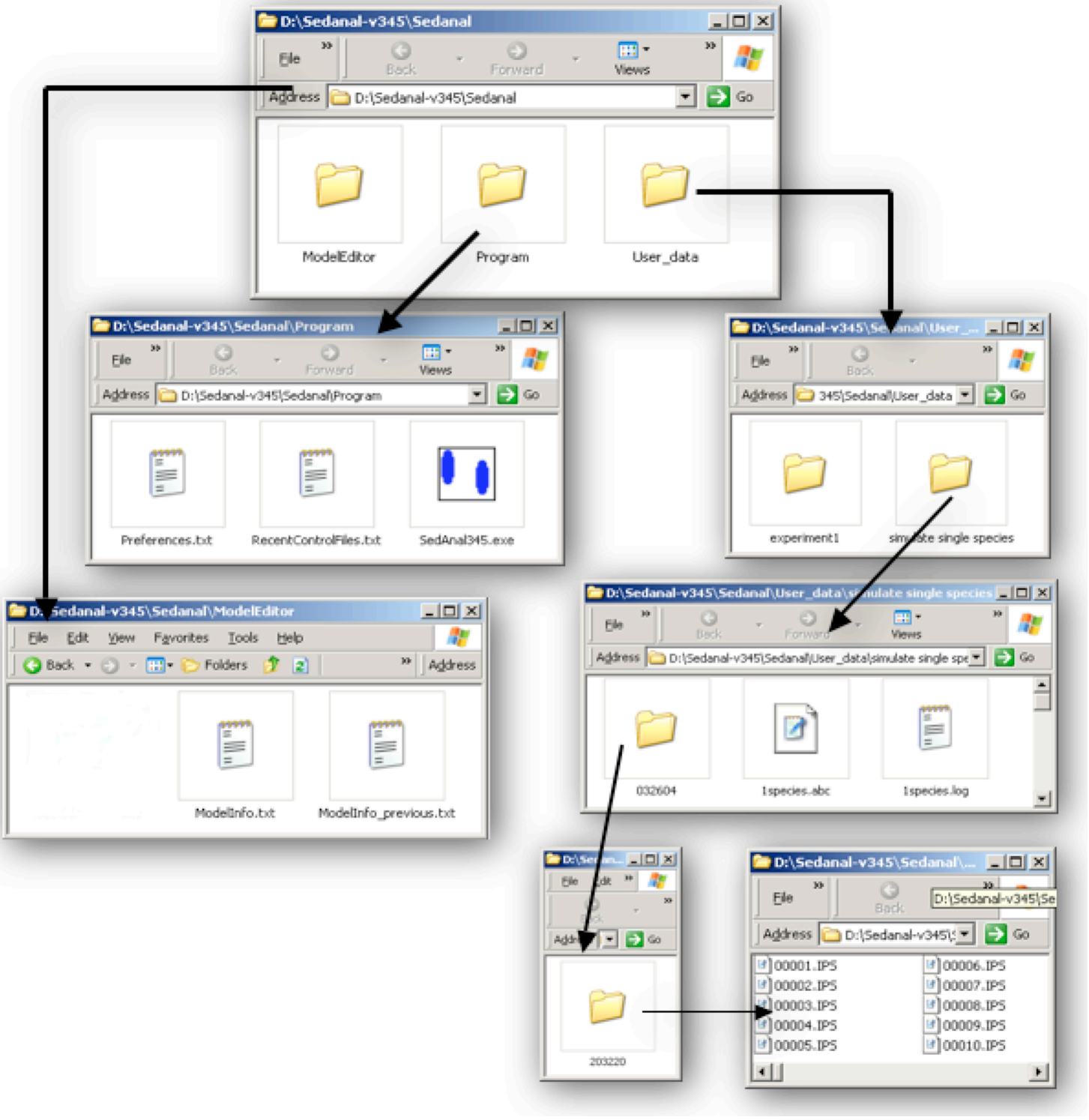


Figure 1-2 Default path structure for SEDANAL initial default installation

2 MAIN MENU

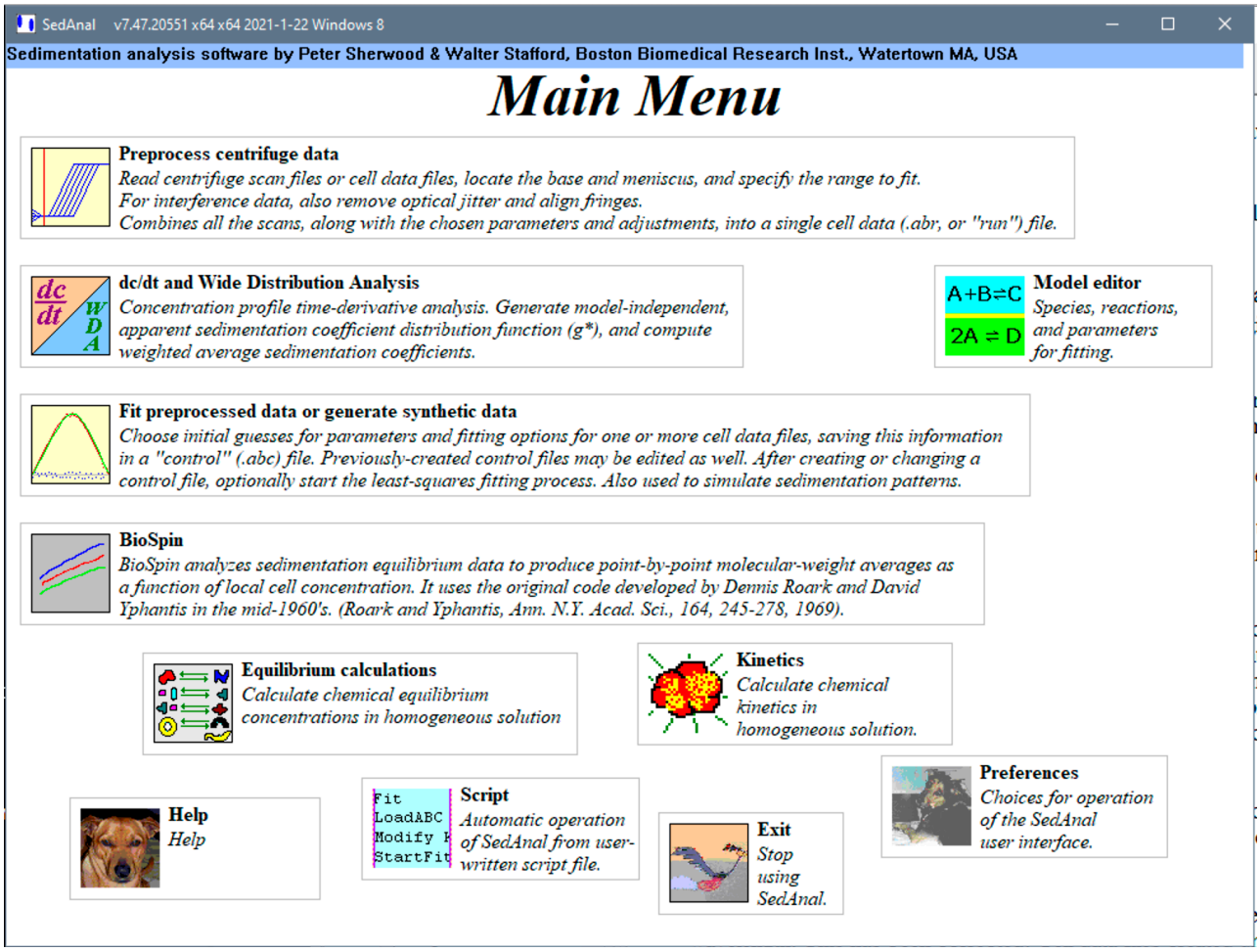


Figure 2-1 MAIN MENU

2.1.1 Preprocessor Screen

- Read in a dataset of scan files
- Designate whether it is a **SedVel** or a **SedEq** run.
- Eliminate systematic error
- Choose meniscus and base
- Choose radial fitting range
- Interpolated optical blank subtraction can be carried out on equilibrium data
- Save preprocessed data in a “cell data file” (*.abr) for fitting, or DCDT and wide distribution analysis (WDA).
- Approach to EQ

2.1.2 ModelEditor

- Enumerate the species in the model
- Create reaction schemes
- Choose parameters to be fit
- Set default limits on parameter values

2.1.3 DCDT and WDA Screens

- Standard time derivative g(s*) analysis (DCDT)
- Multi-speed and single-speed, wide distribution analysis (WDA)
- Both DC/DT and WDA have multi-wavelength capability

2.1.4 Fitter: Control Panel Screen and control file creation

- Read either an existing Control file or Select "New"
- Select either "Analyze data" or "Simulate data"
- Chose a model, enter the parameter values
- Choose datasets to be fit and scans to be included in the fits.
- Enter fitting/simulating parameters (# of points, time increment,...)
- Enter initial guesses
- Fit the model to the data
- The Fitter can be used to fit multiwavelength data, given the extinction spectra of the components.

2.1.5 BIOSPIN

- Computes number, weight, and z-average molar mass as a function of local cell concentration and position in a sedimentation equilibrium run. Also computes M_{Y1} and M_{Y2}, molar mass moments that are independent of the second virial coefficient. Based on BIOSPIN by Roark and Yphantis (1968)

2.1.6 General Purpose Kinetics Simulator

- A general purpose kinetics simulator can be accessed from the Main Menu and can be used to simulate any reaction scheme that can be represented in the ModelEditor.

2.1.7 NOTE: The Main Menu Window can placed in the screen with the following keyboard short-cuts:

The commands are

- **`Ctrl-C`**: Center the main menu both vertically and horizontally in the current monitor.
- **`Ctrl-H`**: Center the main menu horizontally in the current monitor, leaving the vertical position unchanged.
- **`Ctrl-V`**: Center the main menu vertically in the current monitor, leaving the horizontal position unchanged.
- **`Ctrl-R`**: Restore the default position (upper left of the primary monitor) and size.

3 Modules of SEDANAL

3.1 The *ModelEditor*

The Model Editor allows one to create and edit various models and reaction schemes. It is also where the default fitting parameters and their default limits are chosen. (The defaults can be adjusted later if necessary on the Control Screen.)

Models may consist of any combination of reaction schemes ranging from a single species to combinations of non-interacting species, and hetero- and self- associations. It will allow up to 28 thermodynamic components and up to 28 species related by up-to 27 chemical reactions.

When the ModelEditor is first invoked, it starts up with an empty box into which the model name or its aliases may be entered as a text string. The drop-down window will show the other models that have been entered previously. The new model will be added to this list when the "STORE" button is clicked or the "EXIT" button is selected. The user will select one of these models from the Control Screen when fitting data.

3.1.1 THE NAMING TAB

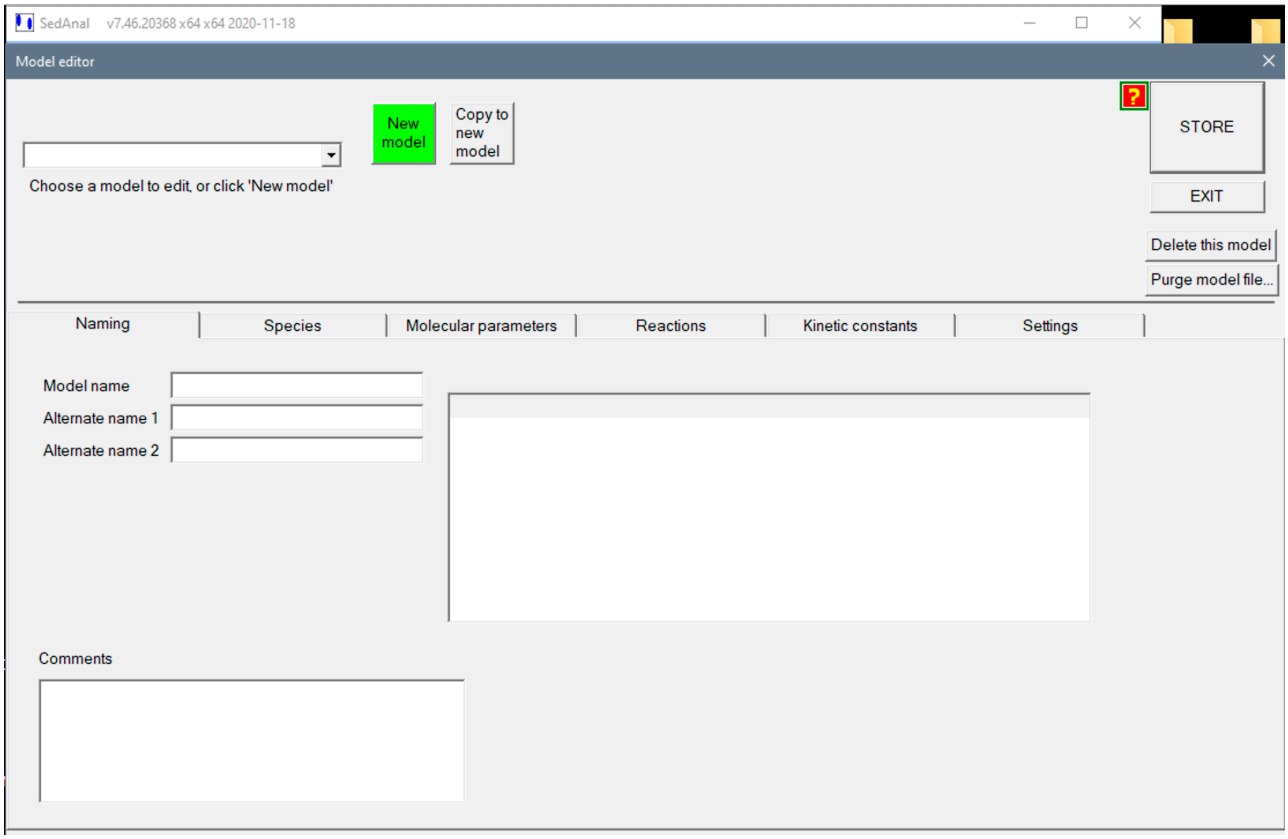


Figure 3-1 The Model Naming tab.

3.1.2 THE SPECIES TAB

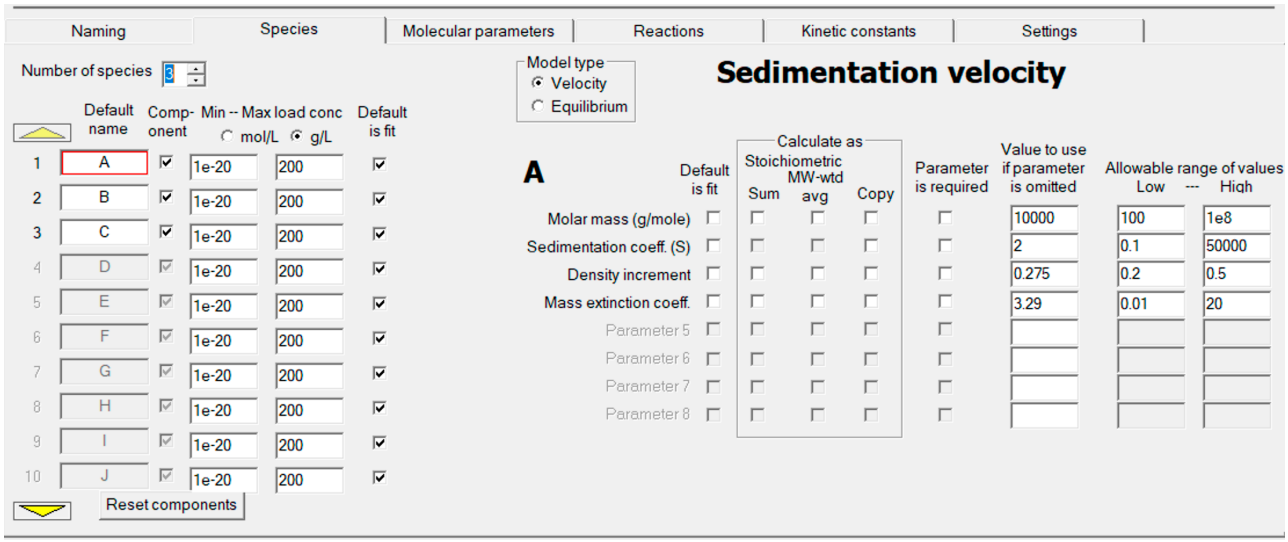


Figure 3-2 The species tab

In this example, we have chosen 3 species, A, B, and C. And since we have not indicated any reactions between them, the Model Editor has designated them as components by putting a check on the boxes next to their names. Examples will be give below for several types of interacting and non-interacting systems.

3.1.3 THE MOLECULAR PARAMETERS TAB

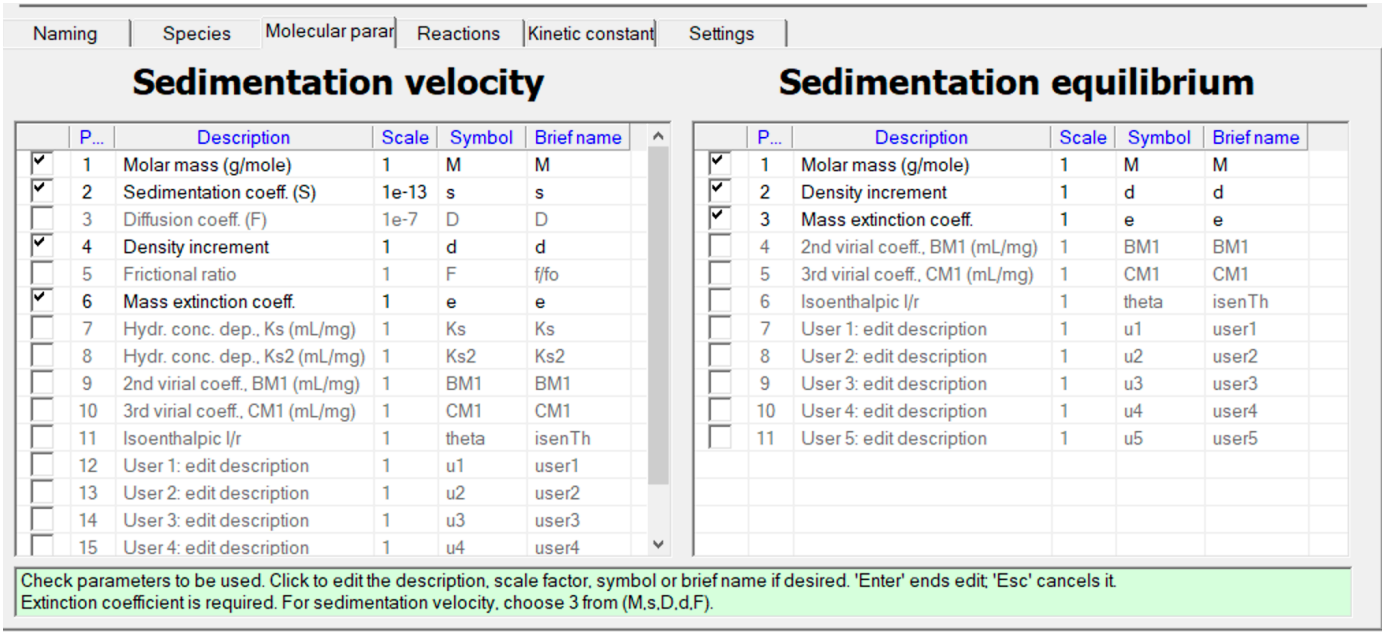


Figure 3-3 Molecular Parameters

The **parameters** include:

- molar mass
- sedimentation coefficient
- density increment, $(d\rho/dc)_T\mu_3$ which for most practical purposes is just $(1-v\rho)$,

- or partial specific volume (actually the product of $v\rho$ as in $(1-v\rho)$, since the density is set to 1.000 internally)
- mass extinction coefficient, either A.U.-$(mg/mL)^{-1}$ or fringes-$(mg/mL)^{-1}$ for a the optical path length of the centerpiece used.
- concentration dependence of the frictional coefficient.,
- k_s, $[s(c)=s_o/(1 + k_s*c)]$ or $[D(c)=D_o/(1 + k_s*c)]$; where k_s, in this case, applies to both s and D, and second order coefficients
- thermodynamic non-ideality expressed through the second virial coefficient, BM_1, as in the expression $(1 + 2BM_1*c)$
- thermodynamic non-ideality expressed through the third virial coefficient, CM_1, as in the expression $(1 + 2BM_1*c + 3M_1*c^2)$
- User defined parameters

The simplest model is a single ideal species. In this case the unknown parameters of interest usually are the sedimentation coefficient and molar mass (or diffusion coefficient). The Model Editor species tab for a single species is shown below. One species has been selected in the species box.

In the list of molecular parameters, the desired boxes are checked, their scale factors indicated and their default limits set.

- In this example, the first parameter is the molar mass and its scale factor has been set to unity. Therefore, future references to this parameter on the control screen must be in units of daltons or grams/mole. If one had entered 1000 in the scale factor field, future references to the molar mass would be in kilodaltons or kg/mole.
- The second parameter is the sedimentation coefficient and has been scaled to $1x10^{-13}$ so that future references to it must be in units of **svedbergs**. At the same time it would also be a good idea to rename the parameter to reflect the new units; for example, you might change it to "Sedimentation coefficient, (S)" as shown below.
- The third parameter is either the density increment or partial specific volume times the density, and the fourth, is the extinction coefficient on the mass concentration scale, (g/L). For absorbance optics this is the specific extinction coefficient multiplied by either 1.2 for the 12 mm path length of the centrifuge cell (**i.e. one has to multiply the usual value (normally corresponding to a 1 cm optical path) by 1.2 for a 12 mm centerpiece before entering it into the box or by 0.3 for a 3 mm centerpiece, for example, etc ...**). For interference optics, it is the number of fringes produced by a 1 mg/mL solution and is approximately 3.29 in a 12 mm centerpiece for typical proteins, but varies depending on whether the macromolecule is a protein, carbohydrate or nucleic acid. Each species may have a different extinction coefficient. SEDANAL will allow data from different optical systems and wavelengths and different centerpieces to be combined for global fitting with a different extinction coefficient to be entered for each species in each dataset (or cell) obtained with a different optical system or wavelength or with a different optical path length.

3.1.4 REACTIONS TAB

The reactions tab (see below) allows the user to specify the reaction scheme relating the several species present in the model.

3.1.5 MODELS

3.1.5.1 One Component - Single Ideal Species Model

The simplest model is a single species with no non-ideality. The necessary parameters are entered into the ModelEditor control screen as shown in Figure 3-4.

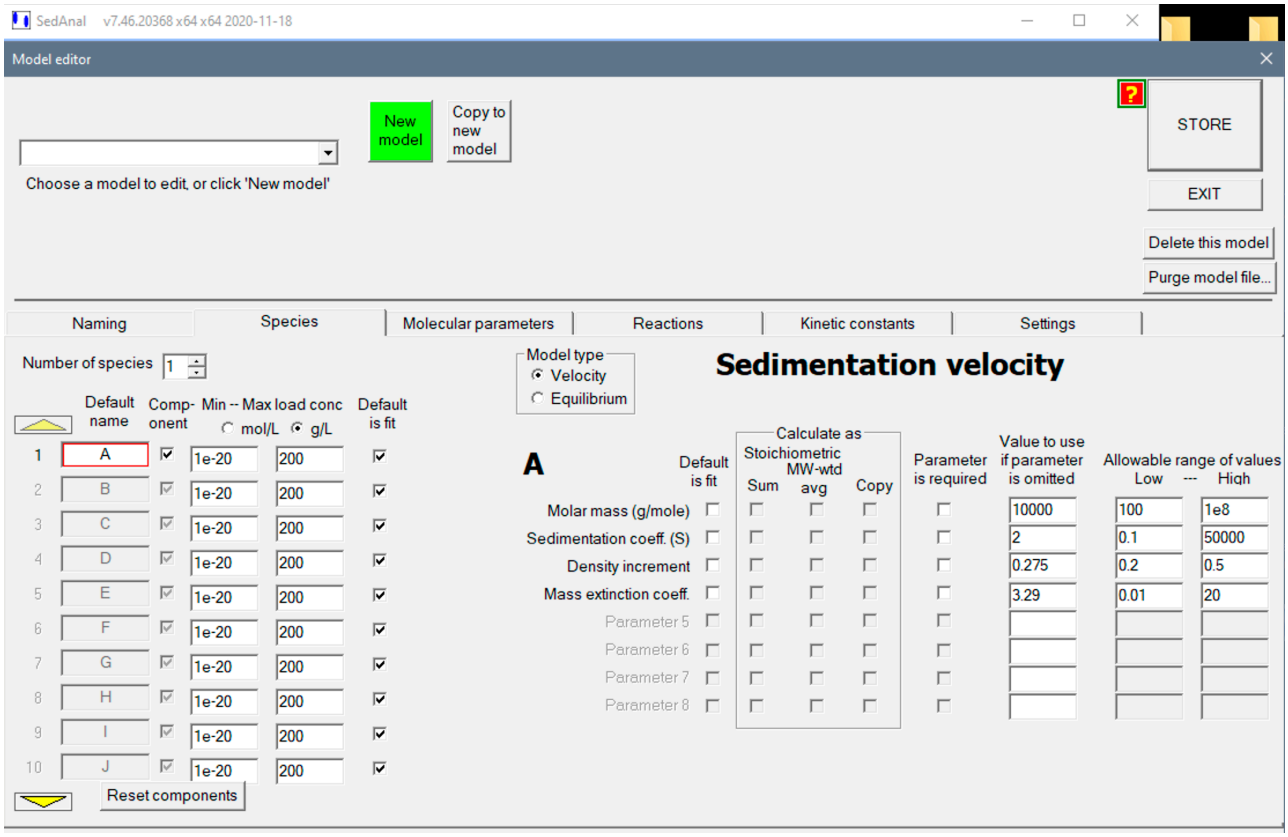


Figure 3-4. **Screen showing default fitting parameters for a One Component - Single Ideal Species Model.**

In this example, (Figure 3-4), the model is named "1 comp 1 species". The number of species is "1", the number of reactions is "0", and the number of parameters is "4". One might want to fit for the molar mass and sedimentation coefficient by default. So, these buttons would be checked. They will appear with a light gray background on the control screen. If they aren't checked here, they will have a blue background on the control screen indicating that the default is to hold them constant. However, their status can be changed later from "hold" to "fit" by right-clicking on the corresponding box on the control screen.

The species having a checked box next to their names are the **components** as determined by the reaction scheme.

One may also select the allowable range for each parameter to be checked during entry of values on the control screen. Allowed ranges to be used during fitting are set by right clicking on the parameter value on the control screen.

NOTE: Pay special attention to these limiting values. Many systems may require that these be changed, especially those that involve organic polymers in non-aqueous solvents and those that involve inorganic compounds or nearly neutrally buoyant or negatively buoyant compounds. They will have to be changed for compounds present in your sample that do not contribute to the optical signal—like background proteins in a fluorescent tracer experiment. Although some of these can be changed on the Control Screen, it is better to change them in the model.

3.1.5.2 One Component - Two Species Ideal Model (e.g. Monomer-dimer, rapidly reversible)

Stoichiometric relationships are established on the Reactions Tab, after the species are indicated on the Species Tab

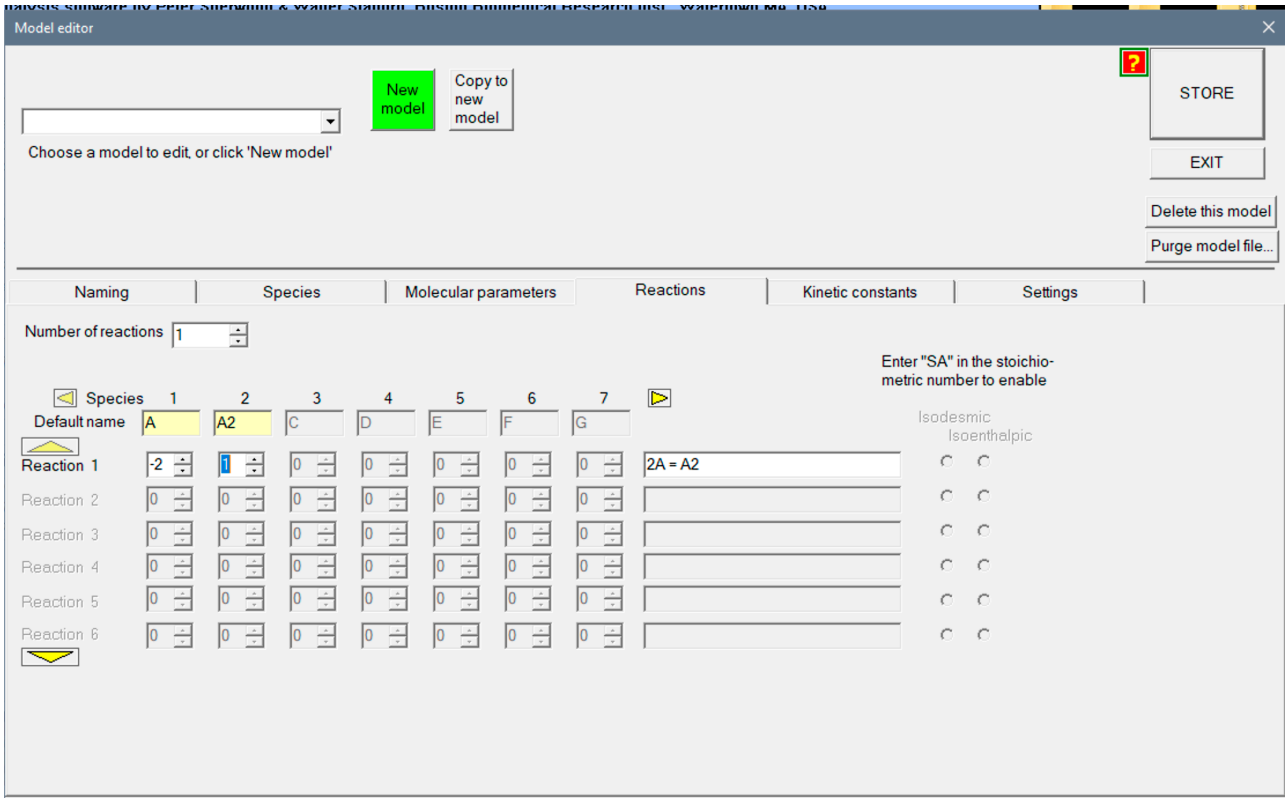


Figure 3-5 One component, two species model.

3.1.5.3 Two Component - Two Species Ideal Model

The next example, "2 comp 2 species", comprises two independent ideal species (i.e. two components) with no interaction between them. Necessary parameters in the ModelEditor control box are shown in **Figure 3-6**.

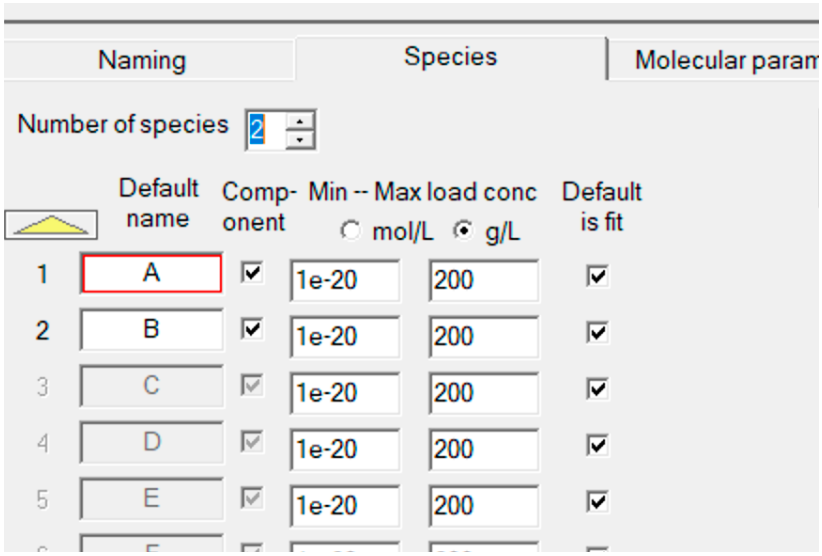


Figure 3-6 Screen showing parameters for the Two Independent Ideal Species Model.

Since we are fitting for A and B as independent species in this example, the model tab for species 2 would be the same as for species 1. If a stoichiometric or other relation between the two independent species is to be specified, it can be done on the control screen using the Equation Editor (The EQN button)

3.1.5.4 Three Component - Three Species Ideal Model

Similarly, 3 component, non-interacting 3 species system, let's call it "3 comp, 3 species", would be designated by choosing 3 species and no reactions. The Model Editor screen would look like **Figure 3-3-7** with each of species tabs the same having molecular mass and sedimentation coefficient checked. As above, if a stoichiometric or other relation between the two independent species is to be specified, it can be done on the control screen using the Equation Editor (The EQN button)

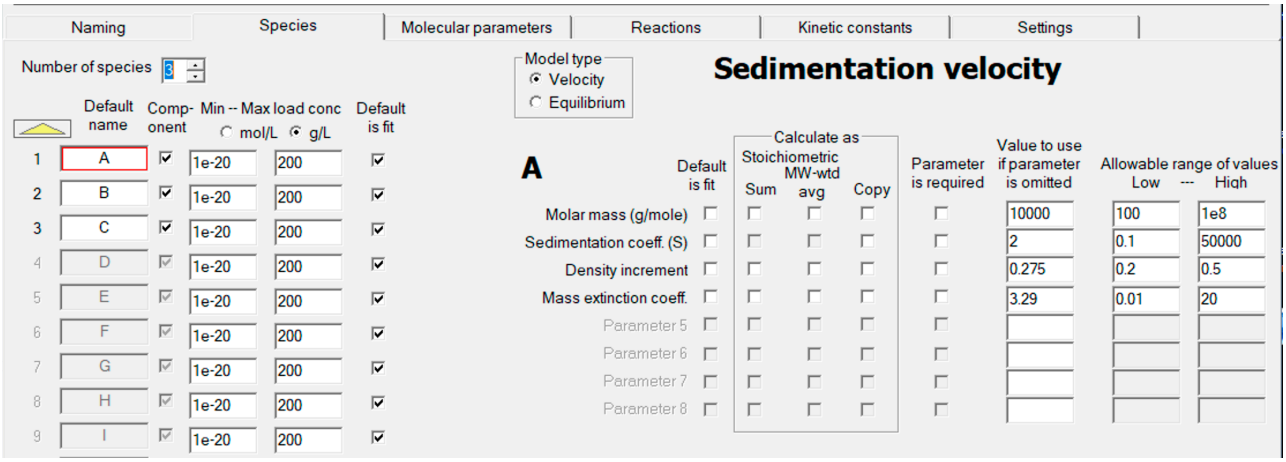


Figure 3-3-7. Model Editor Species tab showing parameters for the Three Independent Ideal Species Model.

3.1.5.5 Two Component - Three Species Interacting Model

In this case, there is simple bimolecular complex formation between species A and species B:

$$A + B = C \qquad K_{eq} = k_f/k_r$$

Where K_{eq} is the association constant, k_f is the forward rate constant and k_r is the reverse rate constant.

In this case, we would have studied A and B separately, and therefore, would not have to fit for their properties. So none of the “Default is fit” boxes are checked for either A or B (**Figure 3-3-8**). In general, in this case, we would be fitting for the sedimentation coefficient of C as well as the association equilibrium constant for complex formation. Note that “Species 3” is not highlighted because it is not an independent component; the two components are A and B on the "species" tabs. (Note: the two components could have been chosen as A and C but that is not convenient in this case.) The species tab for A and B would be filled in the same way (**Figure 3-3-8**).

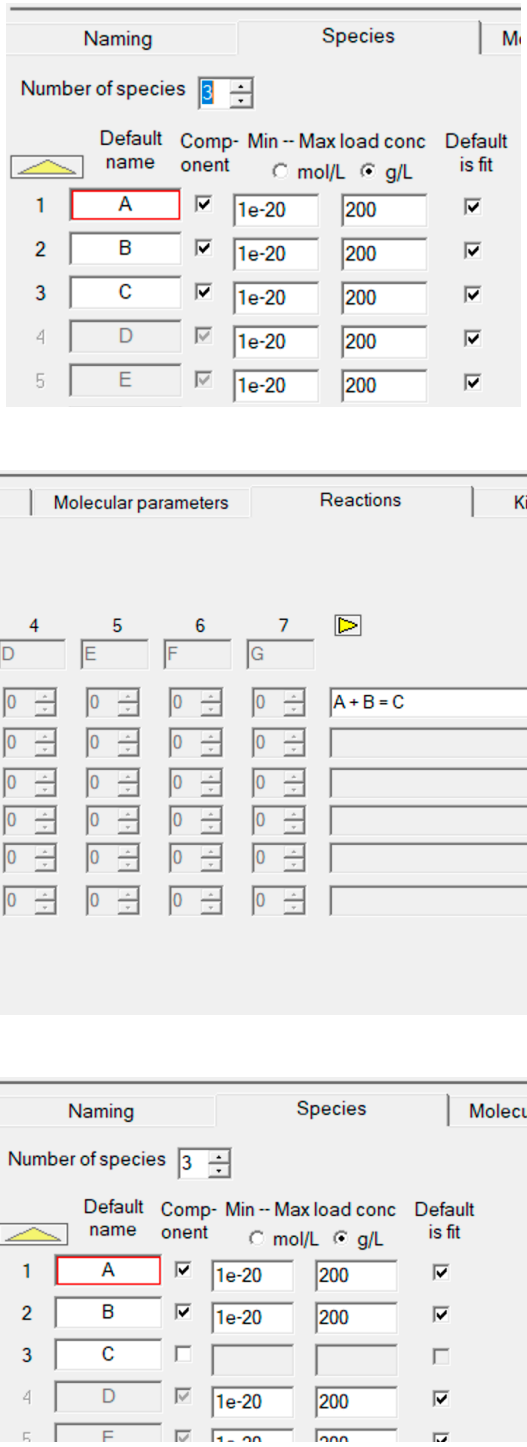


Figure 3-3-8. Screen showing parameters for the Interacting Species Model

For **Species C**, we have (by clicking in the "Species 3" window) the parameters as shown in **Figure 3-3-9**.

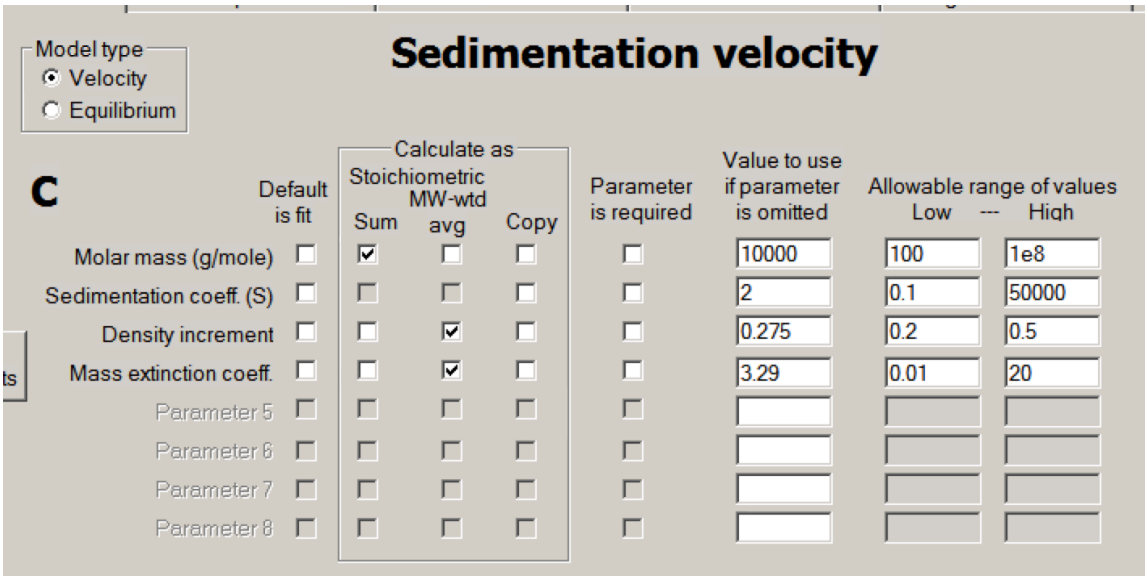


Figure 3-3-9. Species tab showing parameters for Species 3 ("C").

We have checked the "MW-wtg avg" boxes for Density Increment and Mass extinction coefficient to indicate that these parameters are calculated as weight averages of those for species A and B.

In this example, on the row labeled "Reaction 1", the "fit" box for Keq is checked to indicate that we are going to fit for the equilibrium constant by default. For this particular model, SEDANAL uses an analytical solution to the mass action equations. SEDANAL uses analytical solutions for monomer-dimer and A+B=C.

3.1.5.6 Two Component - Four Species Interacting Model

A more complicated example is a 2 component system, comprising of 4 species, and 2 reactions. This might represent an antigen-antibody reaction.

$A + B = C \qquad K_1 = k_{1f}/k_{1r}$

$C + D = D \qquad K_2 = k_{2f}/k_{2r}$

Selection of parameters is shown in **Figure 3-3-10**.

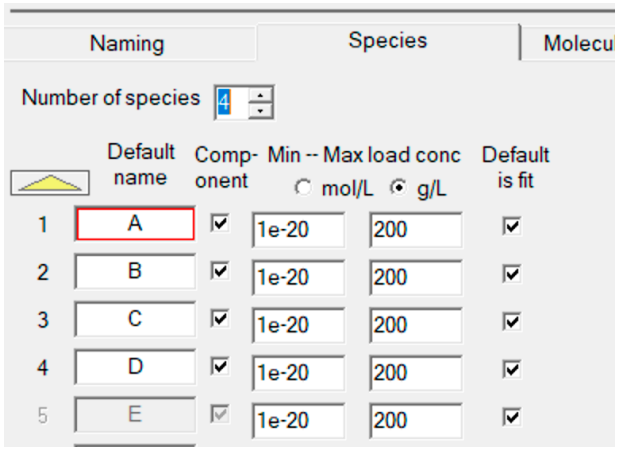


Figure 3-3-10 Selection of species

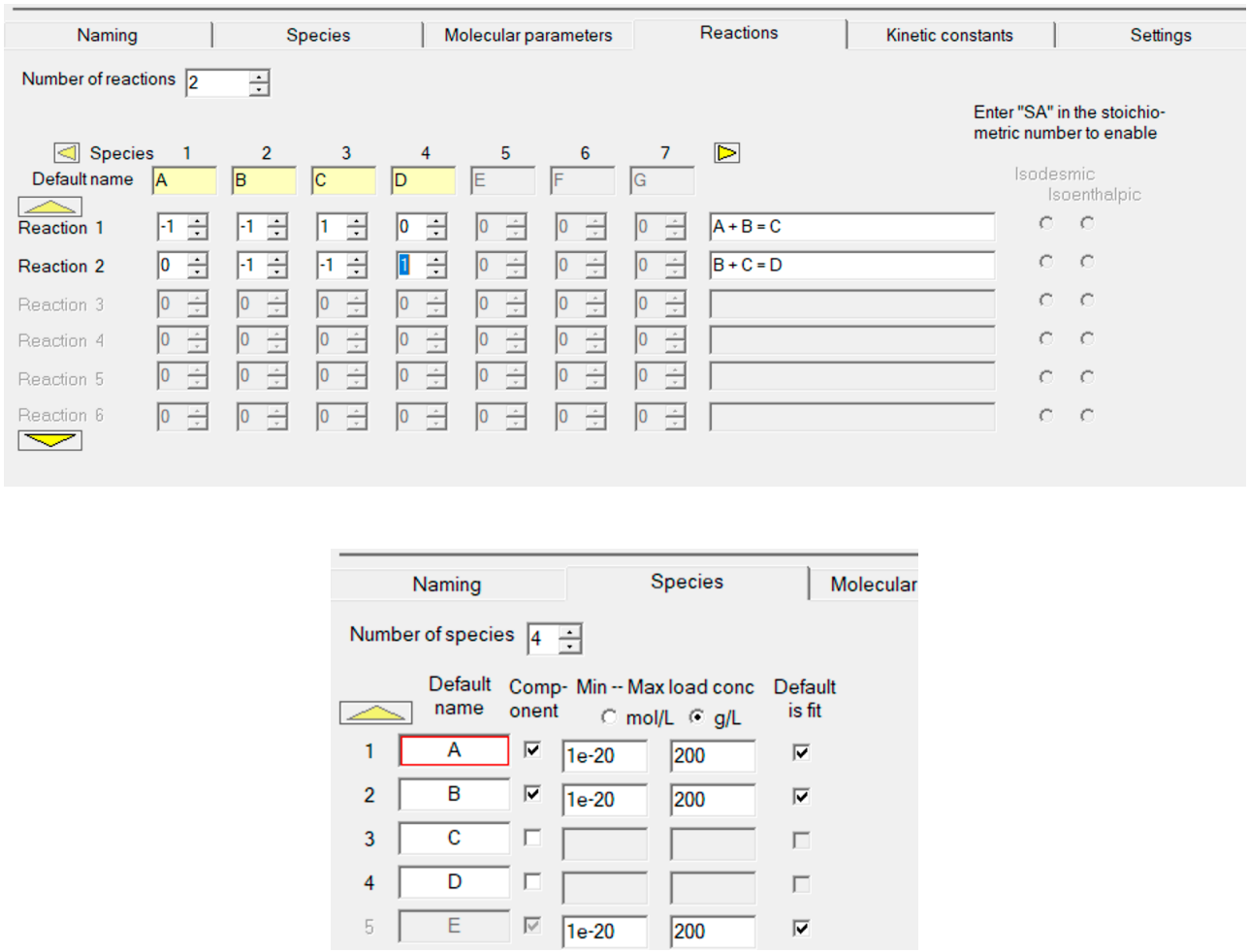


Figure 3-3-11. Parameters for 2 component, 4 species interacting model: A + B = AB, AB + B = AB_2.

When the 4 species are selected on the species tab (**Figure 3-3-10**), all the boxes are automatically checked until the reaction is specified on the reactions tab. After selecting the reaction (**Figure 3-3-11**, **top**), the Model Editor decides which species represent independent components and which are related by a chemical reaction (**Figure 3-3-11**, **bottom).** The number of checked boxed is the number of independent components. SEDANAL fits for the loading concentration of each component.

> **NOTE:** Setting allowable ranges for the parameters shown in the Model Editor affects only the range allowed be inputted on the Control Screen. One may also select the allowable range for each parameter to be used during the fitting procedure by left-shift clicking in the parameter's box on the control screen. If one of the limits is hit, the function evaluator will return a large value for the r.m.s. deviation for that iteration.

3.1.5.7 Three Component - Four Species Interacting Model

An example of a 3 component, 4 species system would be a system such as A+B=C with an impurity or aggregate, D. A and B are the first 2 components and D will be the third. The ModelEditor will correctly understand that D is not participating in a reaction and is, therefore, an additional component.

When this model is chosen (see below), the Control Screen, will show the three components with their appropriate labels. (The species' names can be edited to reflect the three components. For example species C could be renamed, "AB". So the model editor screen would show the third species as AB, as in the figure below.) The figure also shows the tab for species 4 now labeled as D. The Model Editor can determine which species represent independent components and which are not independent.

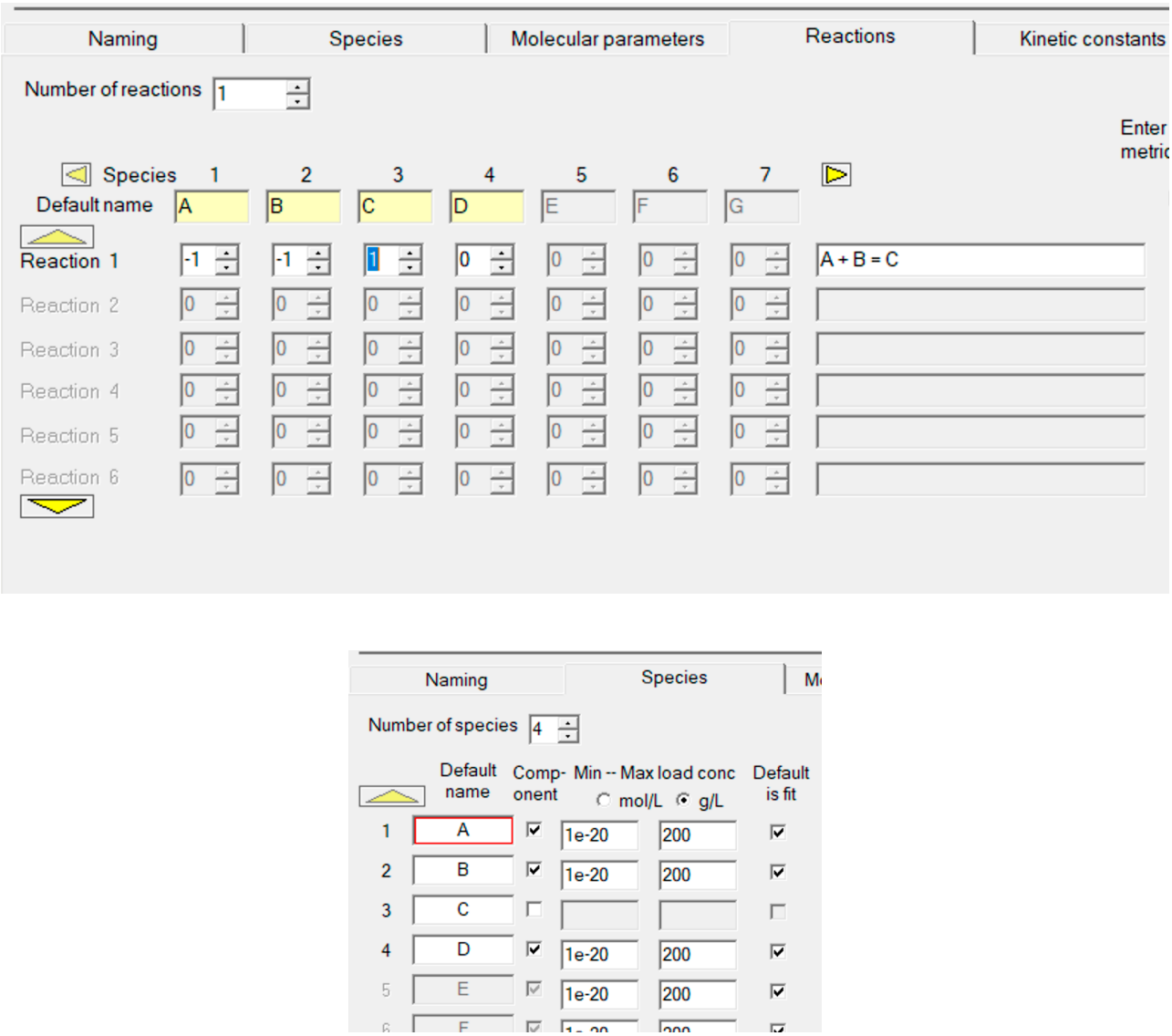


Figure 3-12. Selection of parameters in the control box for three component system

Selection of parameters in the control box for three component system - two interacting species (A+B=C) plus one non-interacting species (D) model (**Figure 3-12**).

3.1.6 MODELS INCLUDING NON-IDEALITY

3.1.6.1 Hydrodynamic Concentration Dependence and Thermodynamic Non-ideality

Non-ideality is treated to first order in concentration through two parameters, k_s, the concentration dependence of the frictional coefficient, and BM_1, the second virial coefficient term. These nonideality parameters

are included in the modeling in the following way for sedimentation and diffusion coefficients: see below: **(Cross Term Non-ideality**:)

$$s(c) = \frac{s^o}{(1 + k_s c)},$$

$$D(c) = \frac{D^o(1 + 2BM_1 c)}{(1 + k_s c)}.$$

The factor $(1+k_s c)$ represents the concentration dependence of the frictional coefficient:

$$f = f_o(1 + k_s c)$$

The factor $(1+2BM_1 c)$ represents the contribution from thermodynamic non-ideality through the second virial coefficient.

It is possible, to add a second order non-ideality terms to allow fitting to data at very high concentrations.

$$s(c) = \frac{s^o}{(1 + k_s c + k'_s c^2)}$$

$$D(c) = \frac{D^o\left(1 + 2BM_1 c + 3CM_1 c^2\right)}{(1 + k_s c + k'_s c^2)}$$

Figure 3-13 Higher order non-ideality terms

3.1.6.2 Single species with non-ideality

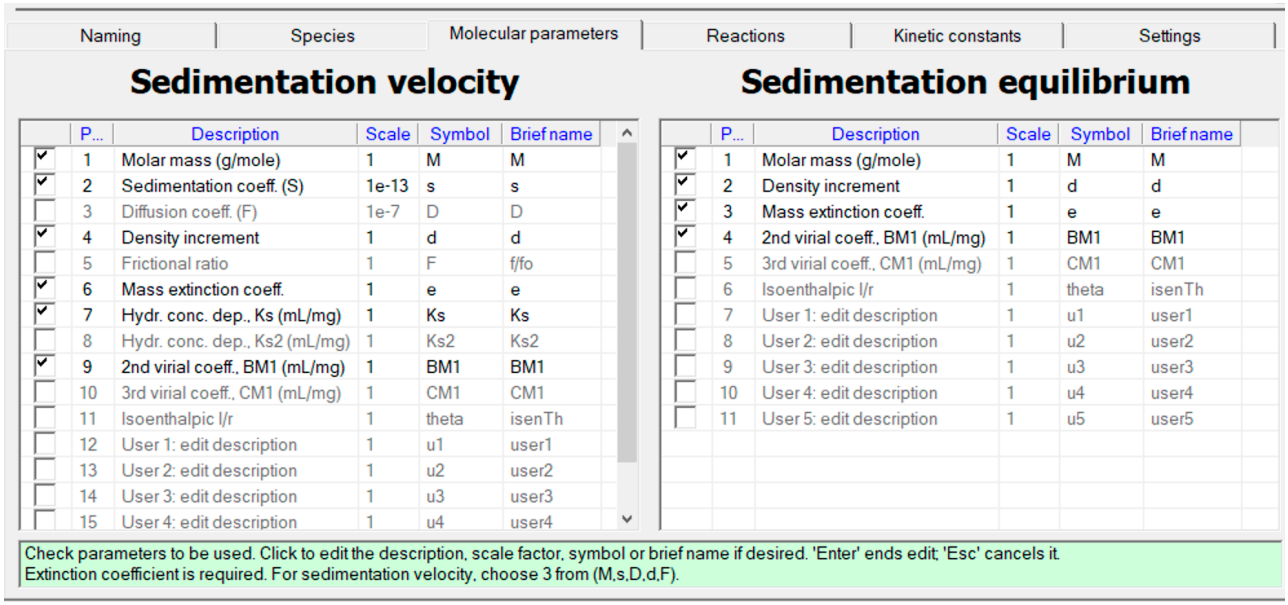
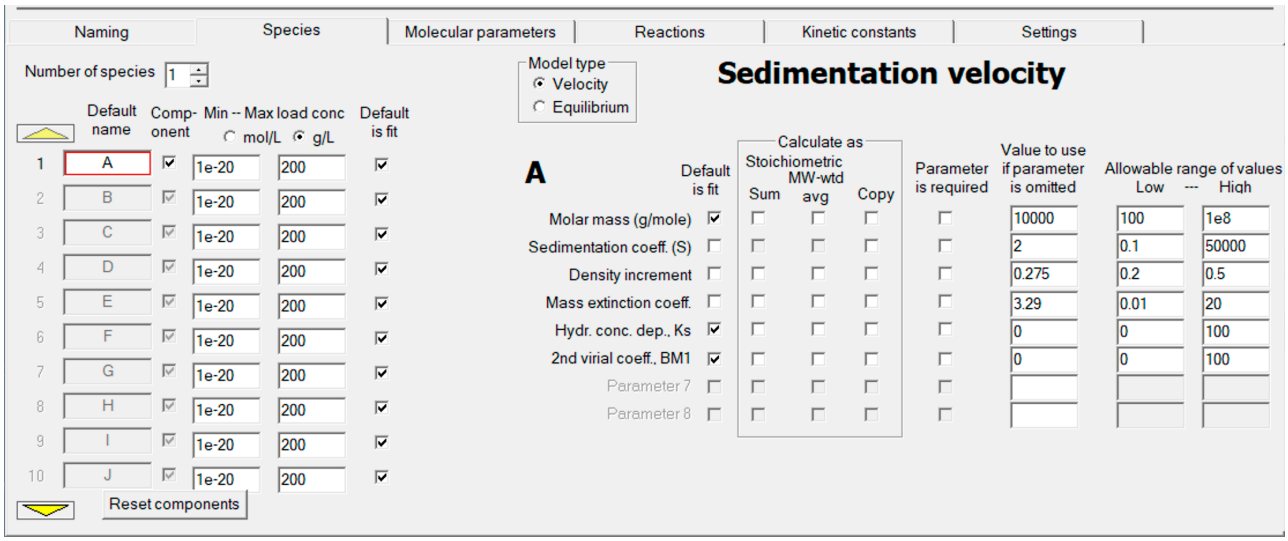


Figure 3-14 Selection of parameters in the control box for single species with non-ideality.

The "Species 1" tab has both Ks and BM1 checked.



3.1.6.3 Multiple Species with Non-ideality

Species "A" and Species "B" would both look the same with the ks and BM1 both checked.

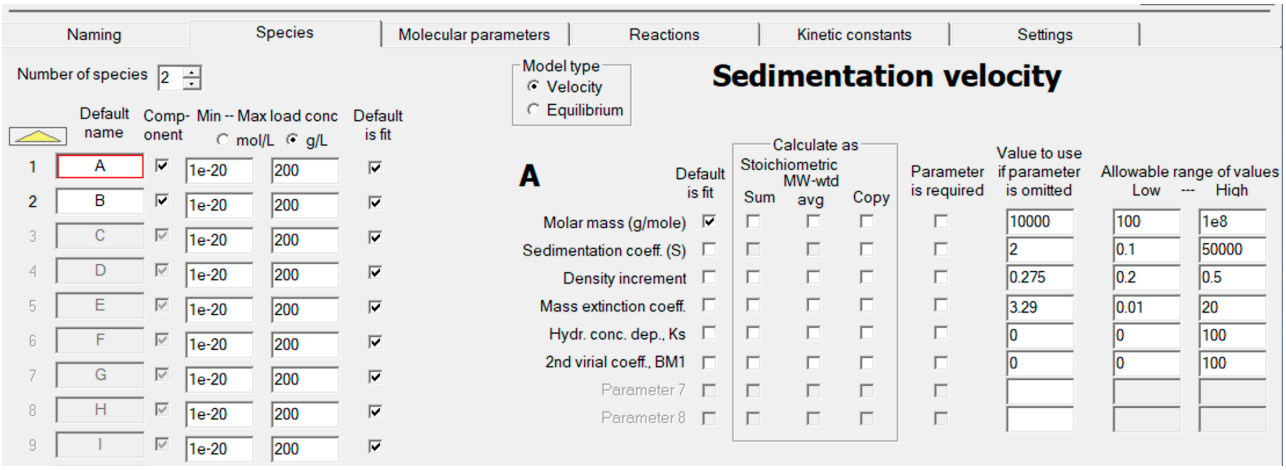


Figure 3-15. Selection of parameters for two species with non-ideality.

If you suspected that the non-ideality was the same for both species A and B, you would check the "copy" box for "Species B" and uncheck the "Default is fit" boxes. This would allow for fitting for a single value of Ks and BM1 for both species Figure 3-16. This might be appropriate on a mass concentration scale.

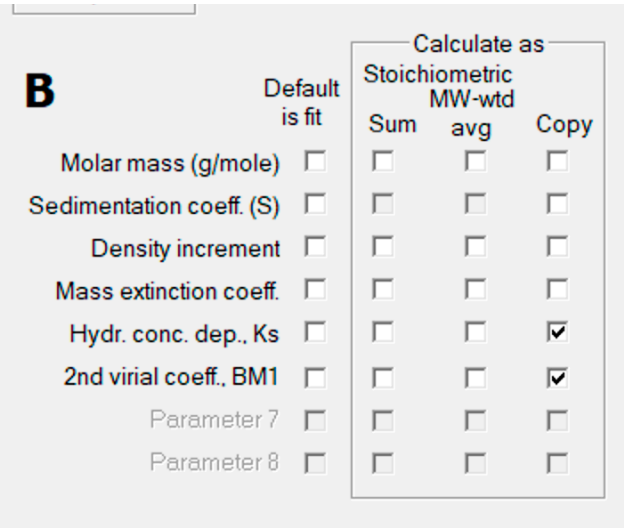
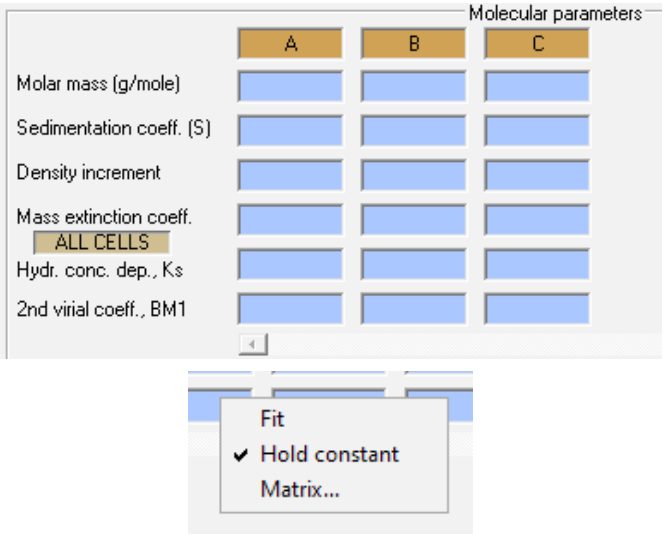


Figure 3-16 **"Copying" values for "Species 2" from "Species 1"**

If one uses only the BM1 and Ks entry boxes on the control screen for each species,

SEDANAL Assigns the non-ideality as follows:



Use either "Fit" or "Hold constant" if fitting for non-ideality with cross term non-ideality represented as follows:

$$B_{11}M_1 = B_{21}M_1 = B_{31}M_1$$
$$B_{12}M_2 = B_{22}M_2 = B_{32}M_2$$
$$B_{13}M_3 = B_{23}M_3 = B_{33}M_3$$

This amounts to requiring that the elements of each **column** of the matrix be equal. This is a reasonable approximation for initial fits to get Ks and BM1 values that are close to the correct ones. This is equivalent to saying that each species has the same effect on all other species. For example it says that the backflow due to species 1 creates a backflow that slows down all other species in proportion to its effective stokes radius and local concentration. The same argument applies to the excluded volume and charge effects (Donnan effects) that species 1 has on the other species in its vicinity in proportion to its local concentration.

NOTE: In most cases, BM1 and Ks are of roughly the same magnitude (in units of cc/g) and both should usually be included in a fit or a simulation. See section below on setting limits to D for the sake of numerical stability. This limit is usually not necessary and generally should be left blank as long as both Ks and BM1 are included in the fit or simulation and are of similar order of magnitude.

3.1.6.4 Cross Term Non-ideality:

(In versions 6.79 and later.)

In a multispecies system, each species can affect each of the others through cross non-ideality coefficients requiring both **Ks and BM1 matrices** to accommodate the cross interactions.

$$s_i = s_i^o \left[\frac{1}{1 + \sum_{j=1}^{N} k_{ij} c_j} \right]$$

$$D_i = D_i^o \left[\frac{1 + \sum_{j=1}^{N} B'_{ij} c_j}{1 + \sum_{j=1}^{N} k_{ij} c_j} \right]$$

The factor in the denominator for both s and D is attributed to the back flow and is expressed as the concentration dependence of the frictional coefficient.

$$f_i = f_i^0 \left(1 + \sum_{j=1}^{N} k_{ij} c_j \right)$$

and the factor in the numerator represents the contribution of the thermodynamic non-ideality expressed through the second virial coefficients. The thermodynamic factor is

$$\left(1 + \sum_{j=1}^{j=N} B'_{ij} c_j \right).$$

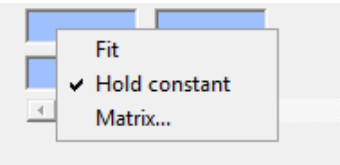
These equations imply the following matrices that SEDANAL will use to express the non-ideality and whose elements are entered on the Control screen:

So for a three species system we would have the following matrices (the order of subscripts is "row, column"):

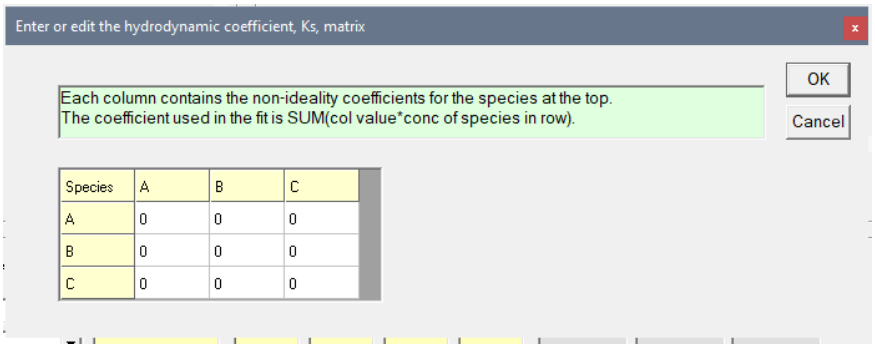
$$\mathbf{k_s} = \begin{vmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{vmatrix}$$

$$\mathbf{B'} = \begin{vmatrix} B'_{11} & B'_{12} & B'_{13} \\ B'_{21} & B'_{22} & B'_{23} \\ B'_{31} & B'_{32} & B'_{33} \end{vmatrix}$$

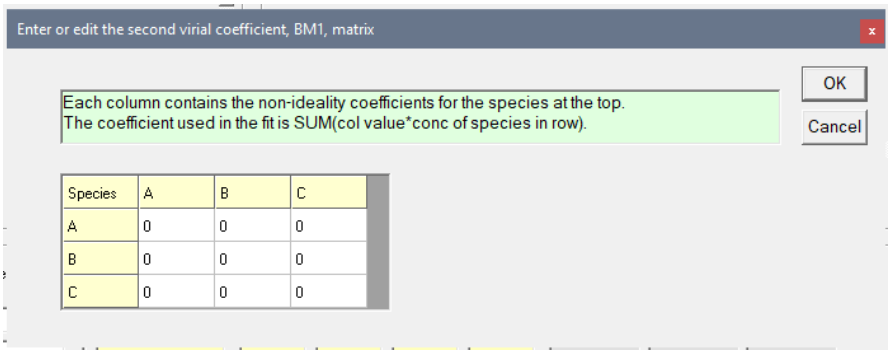
The matrix elements can be entered on the control screen for a non-ideal model by right-clicking on either the Ks or BM1 window:



Click on "Matrix..." to see the matrix entry form. For example, for a three species system, the following matrix window opens to allow entry of the self (diagonal) elements and the cross (off-diagonal) elements of the Ks matrix.



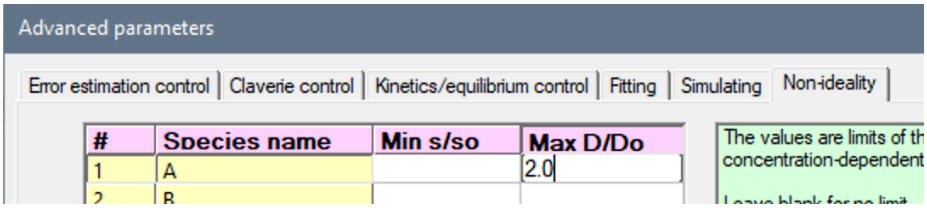
or entry of the BM1 matrix. (where $B'_{ij}=BM1_{ij}$ above)



NOTE: These parameters are read-only and must be determined from a series of separate, pairwise experiments.

3.1.6.4.1 Setting limits to avoid numerical instability in the finite element solutions for non-ideal systems.

Not all systems will require setting of limits, but the more highly non-ideal systems might.



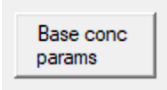
On the Fitting Control Screen, choose Advanced > Non-ideality > and enter 2.0 for D/Do (you may have to adjust this if you get "CHECK GRID" messages).

Leave (s/s_0) blank.

You may also have to set an upper limit on the concentrations near the base:

Advanced > Claverie control > Base conc params

Click on



to get to:

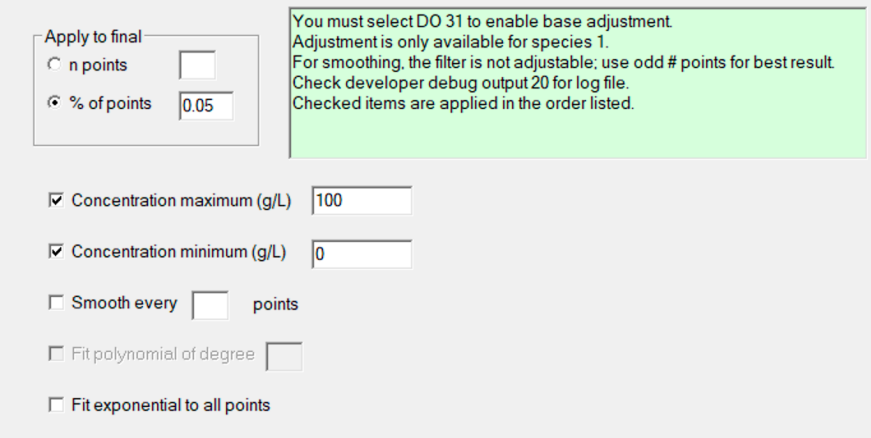


Figure 3-17 Base Concentration Parameters

And be sure to set the “Concentration maximum” to a value several times the loading concentration in g/L. (this limits the concentration and non-ideality contribution at the base of the cell and improves the numerical stability of the finite element calculations. **DO NOT** check the “Fit exponentials to all points” or the Smooth every “n” points boxes. These are experimental and don’t work very well.

3.1.7 INDEFINITE SELF-ASSOCIATION MODELS

The model editor can create isodesmic and isoenthalpic models for use with SEDANAL. Click on the “Number of reactions” to indicate 1 reaction and the type the letters “sa” in the window below the “Default name” for “species 1”, name the reaction either “isodesmic” or "isoenthalpic", click the corresponding button to the right of the row and click the “Store” button.

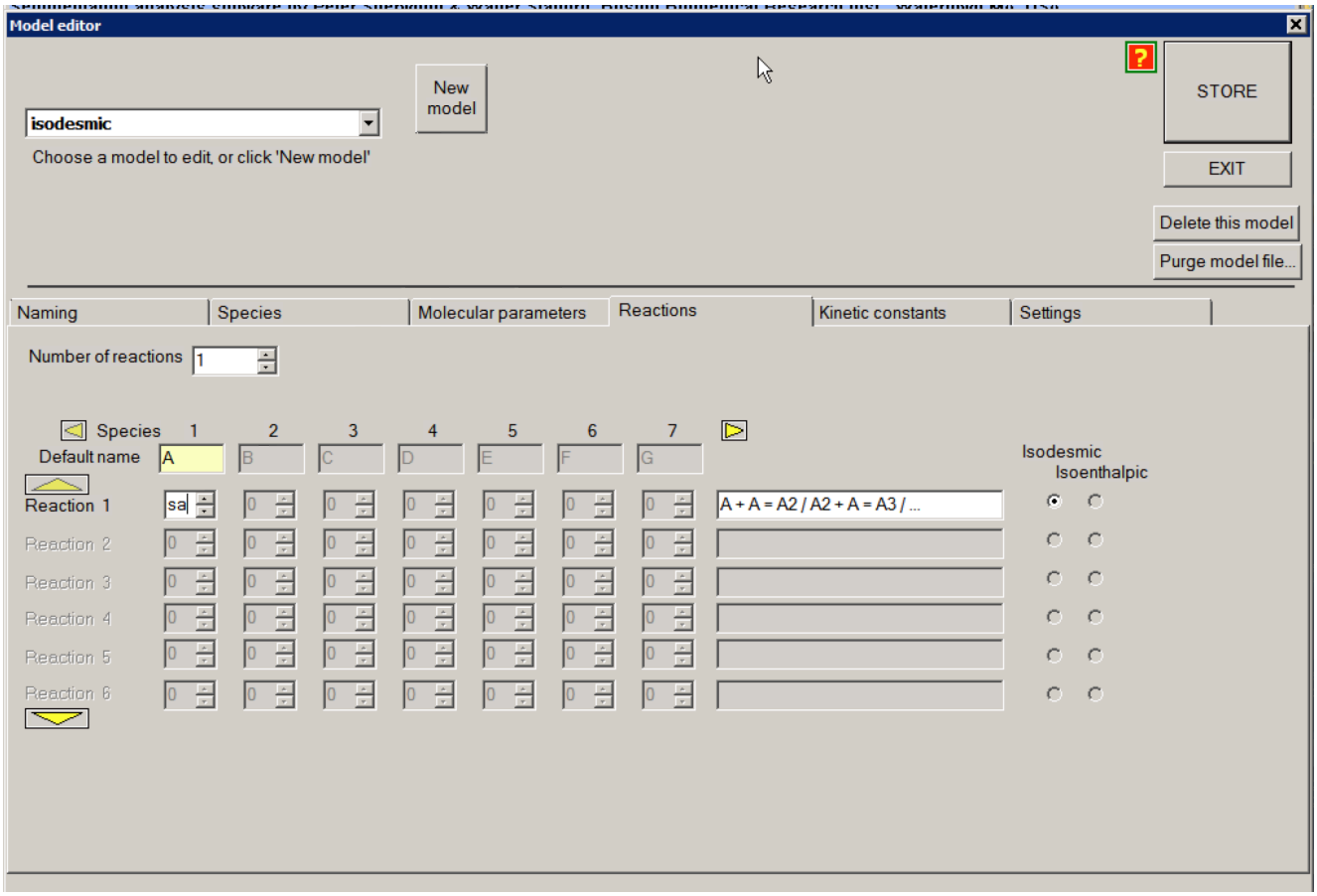


Figure 3-18. Setup for indefinite self-association.

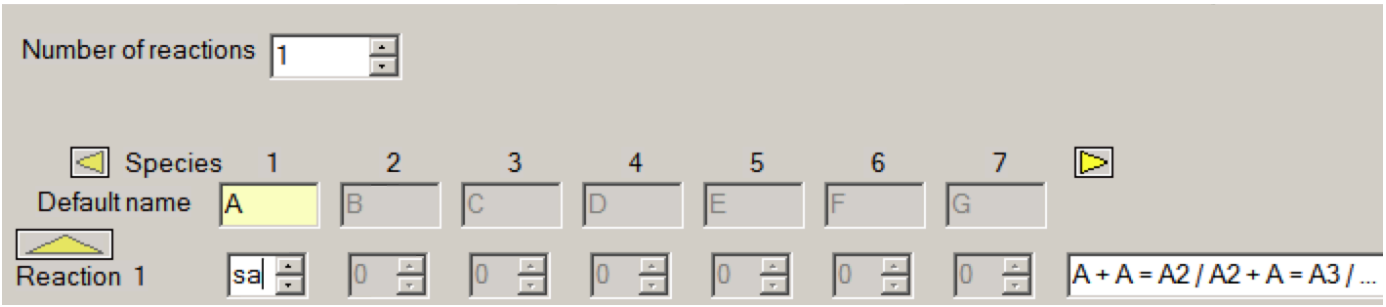


Figure 3-19. Blow-up of upper left of the reactions tab.

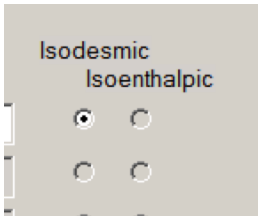


Figure 3-20. Isodesmic case

For the isoenthalpic case (**Figure 3-20**), the ModelEditor window looks the same except that the "Isoenthalpic" button should be clicked to indicate to the fitter that it should use the equations for that case.

3.1.8 Non-ideal Isodesmic systems:

Non-ideality for isodesmic systems is handled that same way as for discreet systems except that SEDANAL use the Fujita-Adams approximation by assuming the non-ideality is a linear function of the total local concentration. Therefore, the Ks and BM1 matrices are not used.

3.1.9 NON-INTERACTING SYSTEMS OF FIXED KNOWN STOICHIOMETRY

Models for non-interacting (i.e. not controlled by mass action) systems composed of components with known molar mass ratios can be treated using the **Equation Editor** on the control screen. Fixed relationships between parameters can be established by typing in FORTRAN-like equations in the **Equation Editor Page** ("**Eqn**" button"). - *More on this later when we get to the control screen*.

3.1.10 EDITING MODELS FOR THE SEDEQ FITTER

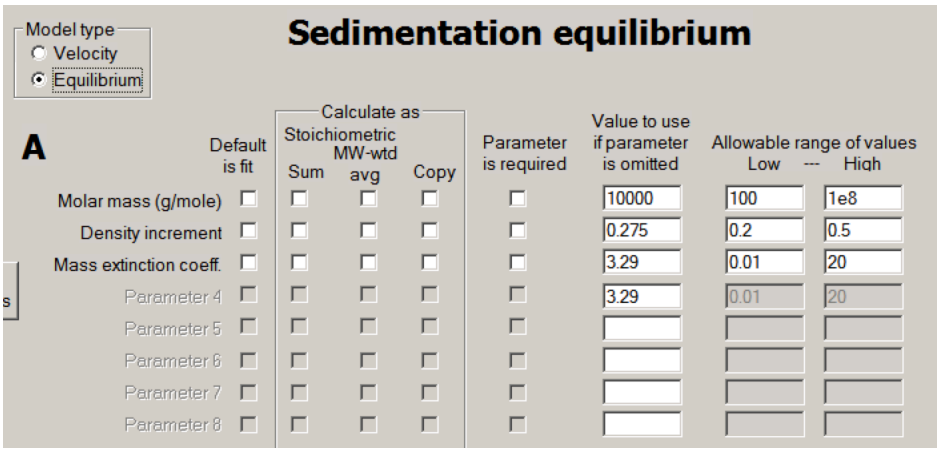


Figure 3-21. SEDANAL will determine whether the data are from SedVel or SedEq

SEDANAL will determine whether the data are from a SedVel or SedEq run from information in the cell data file (*.abr) and display the appropriate boxes on the control screen for a particular model after reading in the cell data file. The type of run is **indicated when preprocessing the data. (This shows the right half of the model editor species tab with the "Equilibrium" button selected.) The model editor has a page for SedVel and a separate page for SedEq in each model selected by clicking the appropriate button above.**

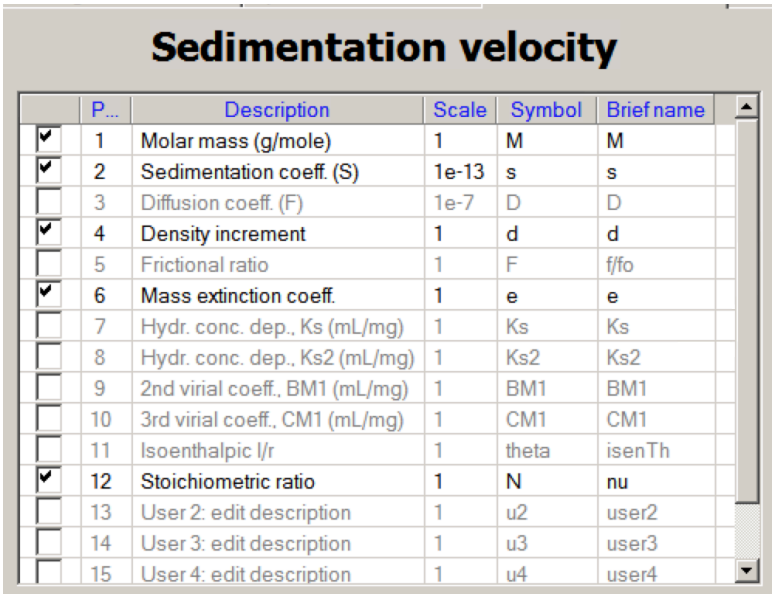
3.1.11 User Auxiliary Parameters for fitting:

The fitting process for sedimentation velocity experiments always uses the molecular parameters molar mass, sedimentation coefficient, density increment (or partial specific volume), and extinction coefficient. For equilibrium experiments, sedimentation coefficients are not used. Other molecular parameters, such as virial coefficients, are optional.

SEDANAL v4.86 introduced a new type of molecular parameter, the ***user auxiliary* parameter**, which is not used directly for fitting. Instead, it is used indirectly via the **Equation Editor**. For example, suppose we want to fit for

the stoichiometric ratios, ν, of species relative to a "monomer" species rather than fitting for their molar masses. For a three species system, the relationships might be $M_2=\nu_1 M_1$ and $M_3=\nu_2 M_1$

First, we create a model in the MODEL EDITOR with a user auxiliary parameter, which we will name N entering the symbol **nu** for them. (Figure 3-22) Now N(*i*) can be used in the Equation Editor.



	P...	Description	Scale	Symbol	Brief name
☑	1	Molar mass (g/mole)	1	M	M
☑	2	Sedimentation coeff. (S)	1e-13	s	s
☐	3	Diffusion coeff. (F)	1e-7	D	D
☑	4	Density increment	1	d	d
☐	5	Frictional ratio	1	F	f/fo
☑	6	Mass extinction coeff.	1	e	e
☐	7	Hydr. conc. dep., Ks (mL/mg)	1	Ks	Ks
☐	8	Hydr. conc. dep., Ks2 (mL/mg)	1	Ks2	Ks2
☐	9	2nd virial coeff., BM1 (mL/mg)	1	BM1	BM1
☐	10	3rd virial coeff., CM1 (mL/mg)	1	CM1	CM1
☐	11	Isoenthalpic l/r	1	theta	isenTh
☑	12	Stoichiometric ratio	1	N	nu
☐	13	User 2: edit description	1	u2	user2
☐	14	User 3: edit description	1	u3	user3
☐	15	User 4: edit description	1	u4	user4

Figure 3-22 Auxiliary parameters—user defined

Next, create a control file, specifying the new model.

The symbol for the auxiliary and other molecular parameters used in the equation editor, N in this example, is given in the model (in the MODEL EDITOR). Symbol names are made up of letters, digits, and underscores ("_"), with the initial character not a digit. **Symbol names are case-sensitive.**

The symbol names must be unique, and cannot conflict with the symbol names used for other fitting parameters; currently these are K, kf, kr, L, r, rw, m, b, y (for equilibrium runs only), p, v and T.

Versions after version 5.78, now include the **Frictional Ratio, F,** as a fitting parameter which allows one to set limits for f/fo requiring it to be in any specified range, like 1.0 to 5.0, and not to allow it to drift below 1.0, the theoretical lower limit of a perfect, unhydrated sphere.

NOTE: When storing a model, the name and aliases (if any) are checked, and storing is disallowed if any other models has the same name or alias. Names are considered the same if they differ only in spacing. For example, `"A+B=C"` and `"A + B = C"` are the same

3.2 MAIN MENU - Other Modules: Access by Main Menu

When SEDANAL is launched, the Main Menu appears with 11 choices, see Figure 3-23.

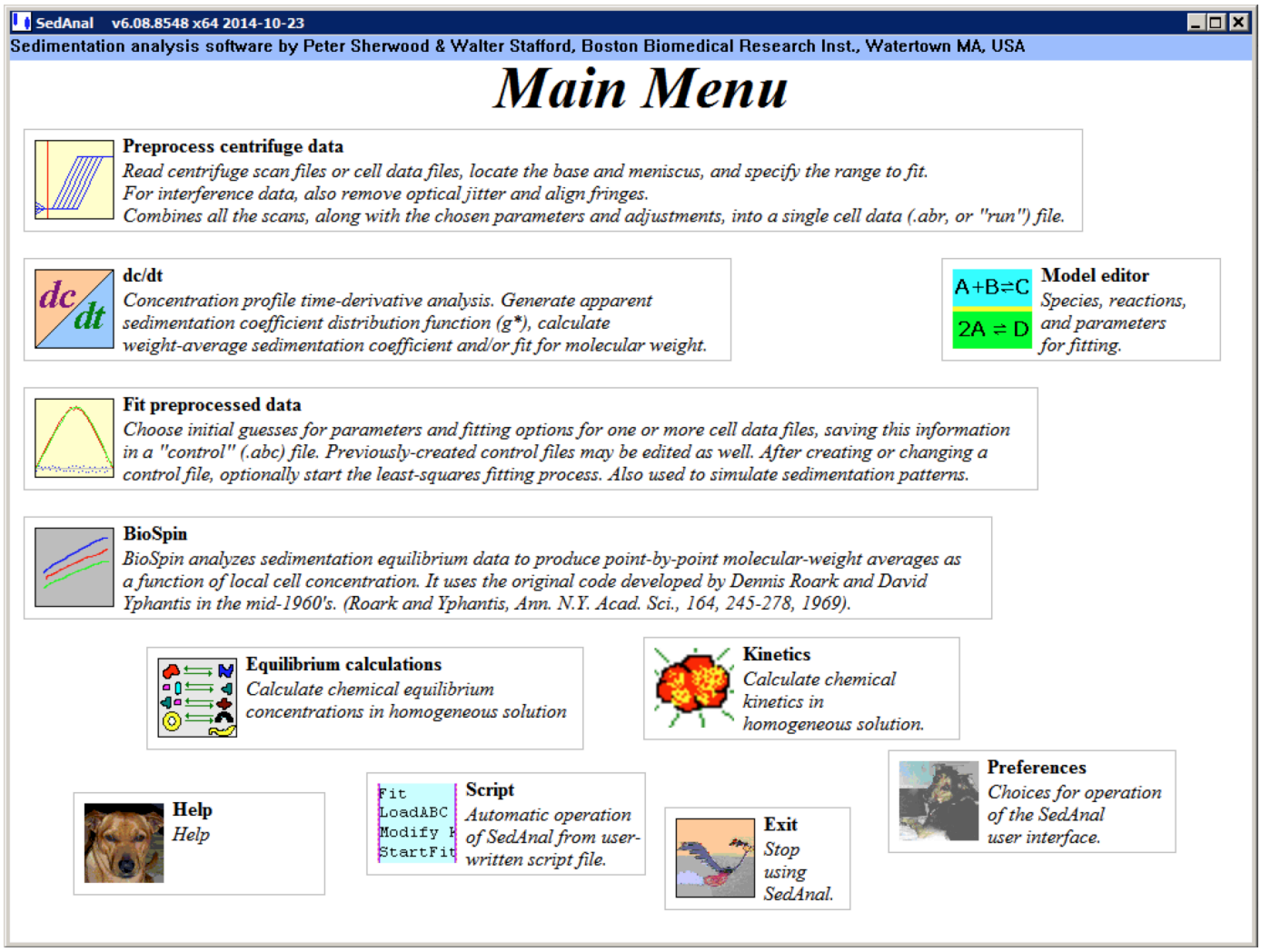


Figure 3-23 **The Main Menu**

3.3 Preprocess Ultracentrifuge Data

The preprocessor screen presents the following window:

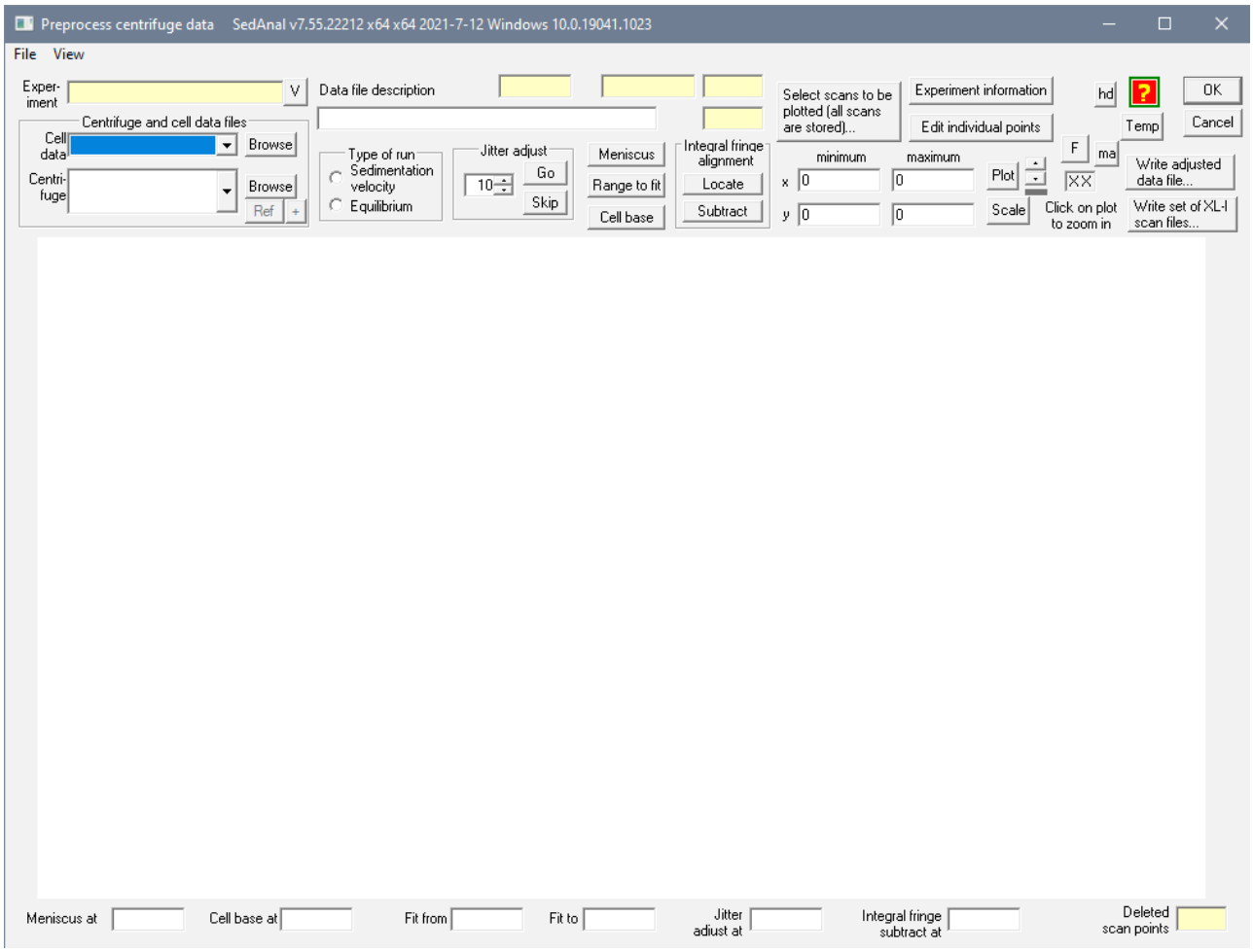


Figure 3-24 PreProcessor Screen

The "View" menu in the upper left corner of the Preprocessor window provides several choices for the display of data: the default values can be set in the Preprocessor Preferences tab.

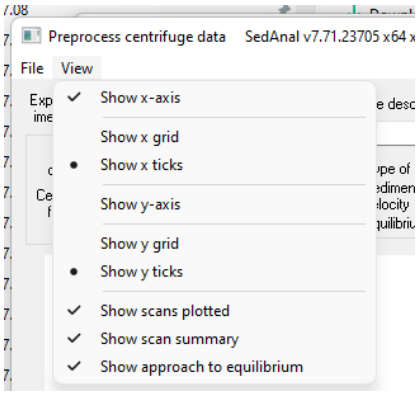


Figure 3-25 View choices

The first step is to select the experiment folder by clicking on the top upper left box to reveal a list of experiments: in the top drop-down window, one must select the experiment folder (Figure 3-26). Then, once the experiment folder has been selected, in the bottom drop-down window, one will see the available datasets and select the dataset to be preocessed (Figure 3-27).

Processing is different for absorance data and interference data. For both absorbance data and interference data, one must select the meniscus, the base and range of data to be fitted. Additionally, for interference data, one must remove optical jitter (time dependent, radially independent noise) by aligning the fringes in the air-air space to the left of (centripetal to) the meniscus and then remove any integral fringe jumps by selecting a spot--either in the plateau, near the meniscus, or near the hinge point depending on the dataset--where the data changes the least from scan to scan.

In addition to absorbance and interference data, the preprocessor recognizes both **intensity** data from the XL-A and the Optima AUC and **fluorescence** data from the Aviv FDS. The preprocessor also recognizes data from the Cölfen type **Multi-wavelength** (Open MWL) absorbance optical systems. Each set of intensity data is presented in the preprocessor as two separate datasets, one from the sample side and one from the reference side. The intensity data are converted to "pseudo-absorbance" data by taking their logarithm before being processed. Pseudo-absorbance data are linearly related to absorbance data but differ from the corresponding absorbance data by an arbitrary additive constant offset.

3.3.1 Loading Data

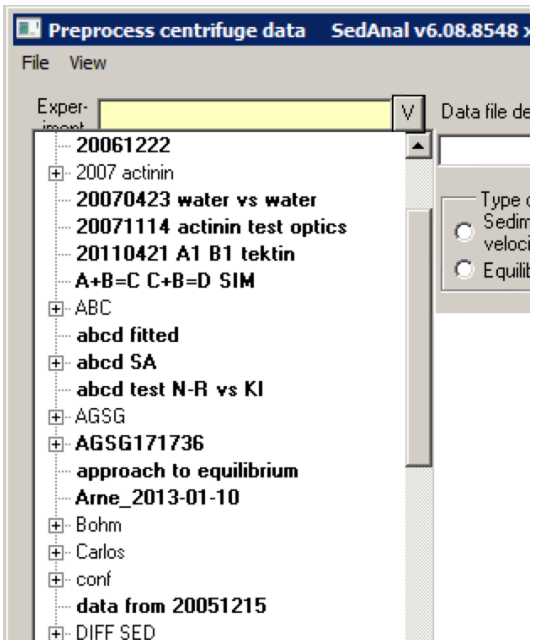


Figure 3-26. The Experiment drop-down window

The Experiment drop-down window displays a list of folders in the User_data directory. Folders with names in bold format contain scan files. Folders in regular font, contain subfolders that contain scan files.

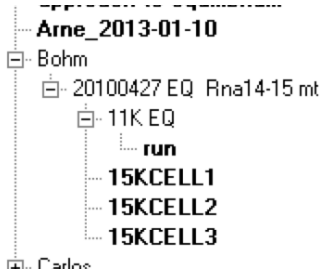


Figure 3-27 List of folders

In this example, we'll select "AGSG171736" which contains XL datasets stored in the standard format. After expanding the folder "AGSG171736" select a set of files such as "15KCELL1", and now click on the "Centrifuge" drop-down window. To reveal the scan files available for fitting.

(N.B: <u>The standard Beckman folder format</u> is

…\User_data\experiment 1\"date"\"time"*.IPn).

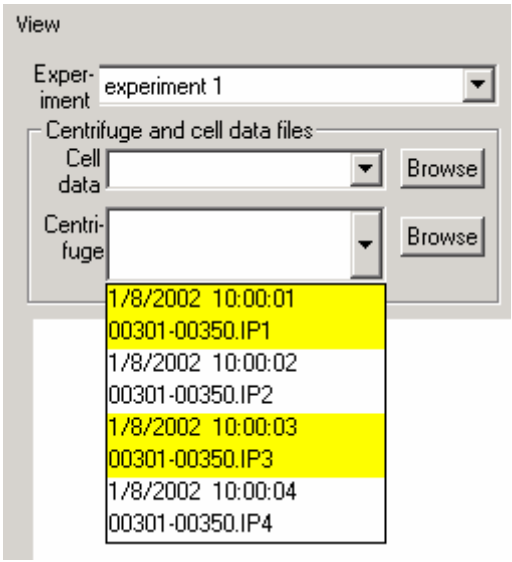


Figure 3-28 **The datasets are shown with alternating yellow and white backgrounds.**

The time and date shown in the "Centrifuge" drop down window are derived from the folder names (in this case from a 4 cell run).

When one of the datasets is selected, the entire dataset will be read in and will be used for fitting. The actual files (i.e. scans) to be used in fitting will be chosen from the control screen by right-clicking on the run file name window (see below)

Note: If a file or files are missing from a dataset, SEDANAL will break the set up into two sets. For example, if the run generated scans 00001.IP1 thru 00099.IP1 but scan 00050.IP1 was missing for some reason, SEDANAL would show two batches of scans from that cell, one from 00001.IP1 thru 00049.IP1 and another from 00051.IP1 thru 00099.IP1.

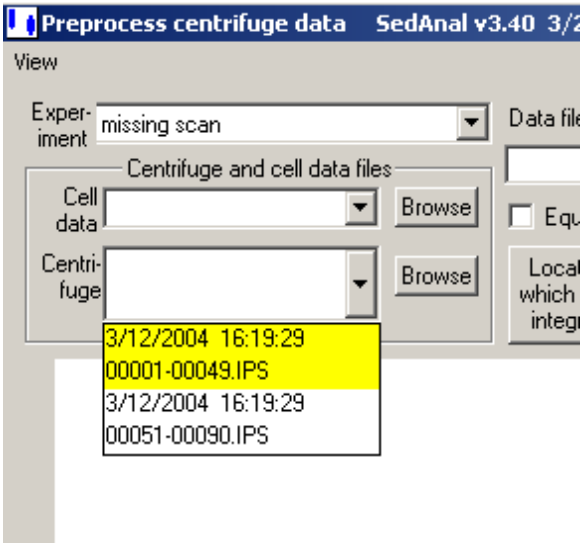


Figure 3-29 missing scan file

In order to rejoin the two parts of the run, a dummy file would be need to be created with the name 00050.IP1. Now the dataset will appear as one contiguous set of files and would be treated as a single dataset. Later, on the control screen one would have to exclude that dummy file from the analysis. The dummy file must at least contain the two header lines at the top. The resulting series of scans (showing only those from 00040.IP1 to 00060.IP1) would look as shown in Figure 3-30 and now the data can be treated as a single set.

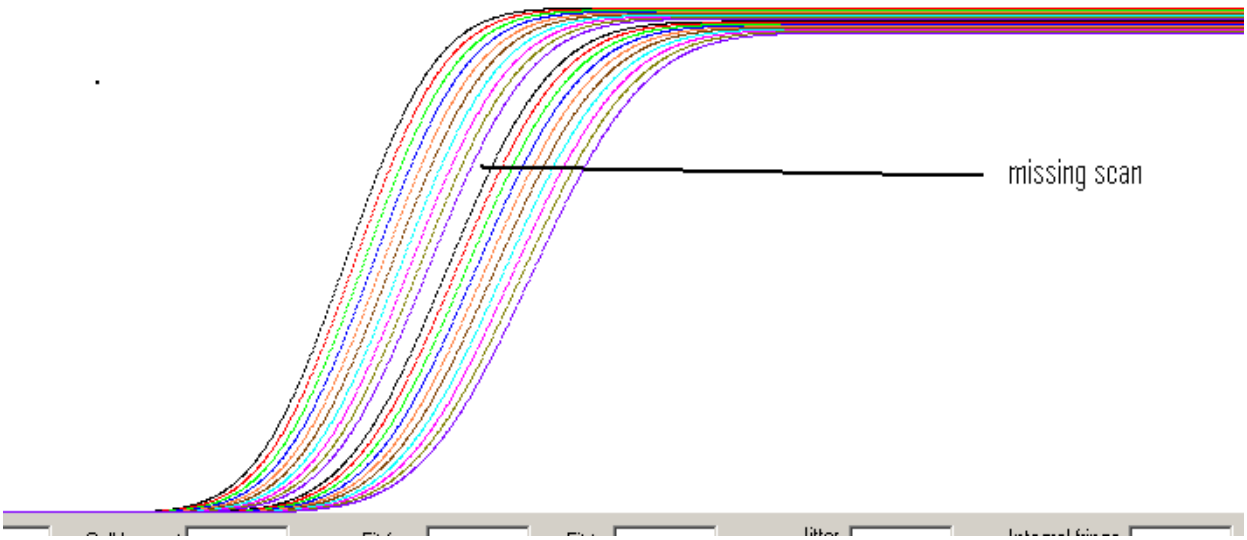


Figure 3-30 Missing scan

Occasionally, the XL-A/I will write an empty file and the data will appear as shown above. The name (i.e. number) of the missing file must be noted and that file must be excluded from the fitting or from the DCDT/WD analysis. Versions 6.01 and later will automatically flag a scan as bad if the file is not complete.

After a dataset has been selected and read in, the screen will look something like this for interference data:

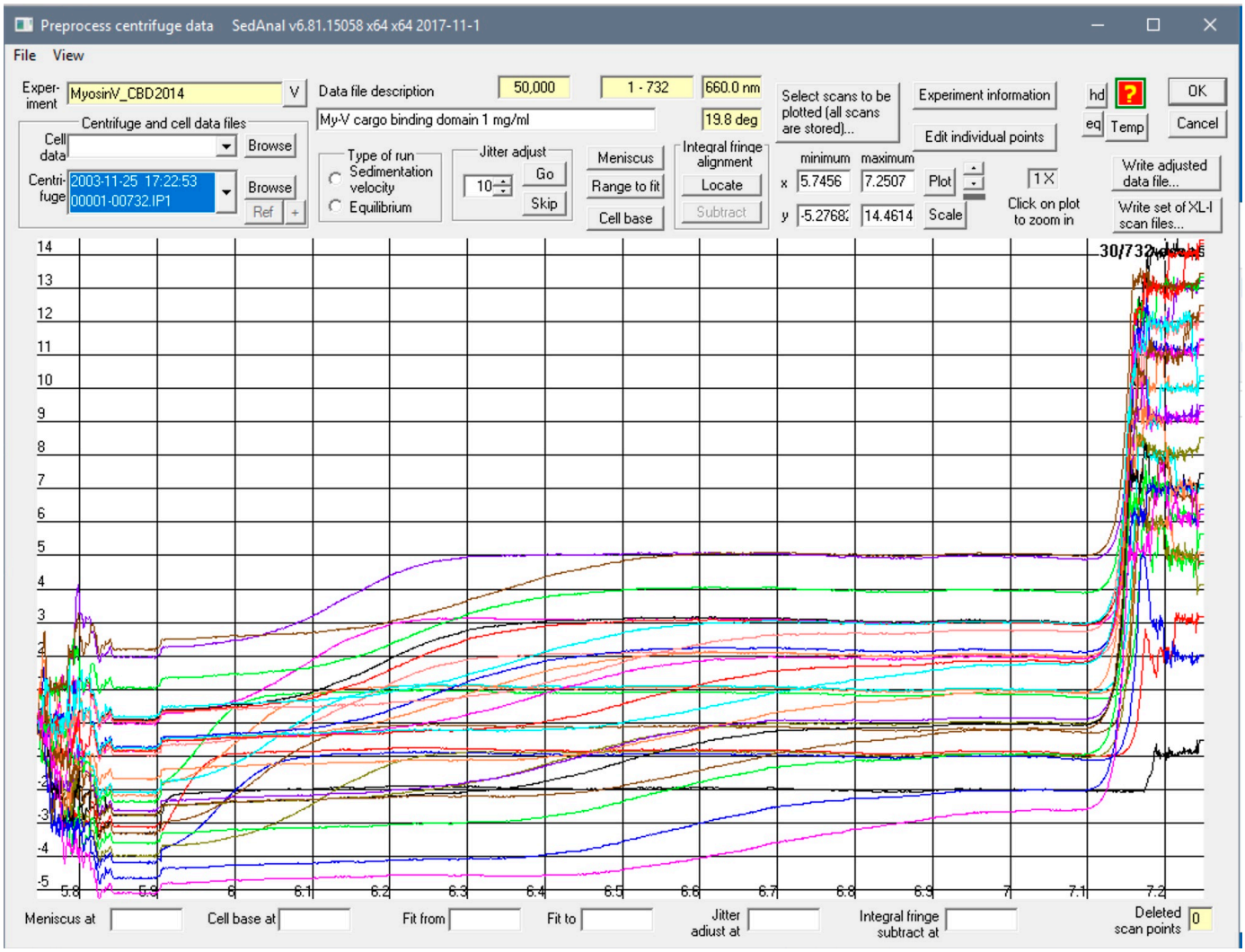


Figure 3-31 Raw Interference data

Initially when the data (in this case, interference data) are read in, they will look like above Figure 3-31 **and will require alignment in the air-air space (jitter adjust) and removal of integer fringe jumps in the solution column that may occur at the meniscus.**

Now indicate what type of run it is:

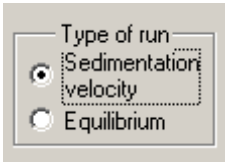


Figure 3-32 **Click the appropriate button for this run type.**

Next, we take care of vertical "jitter" by aligning the fringes in the air-air space. Under the "Jitter adjust" label click on "**Go**" to activate the cursor and click in the air-air space. The number "10" indicates that a range of plus and minus 10 pixels will be used for the alignment. If you need to undo what you have done, click on "reset".

3.3.2 Adjusting Data for off-sets and jitter

Figure 3-33 **The "Go" button: Jitter adjust.**

When the "Go" button is clicked the cursor is activated to indicate the spot at which the jitter adjustment will be carried out. the number "10" in the window next to the "Go" button indicates the number of pixels around the clicked position (i.e. +/- 10 points) that will be used in the jitter correction process. Adjust this number to make the range narrower, as one might for absorbance data.

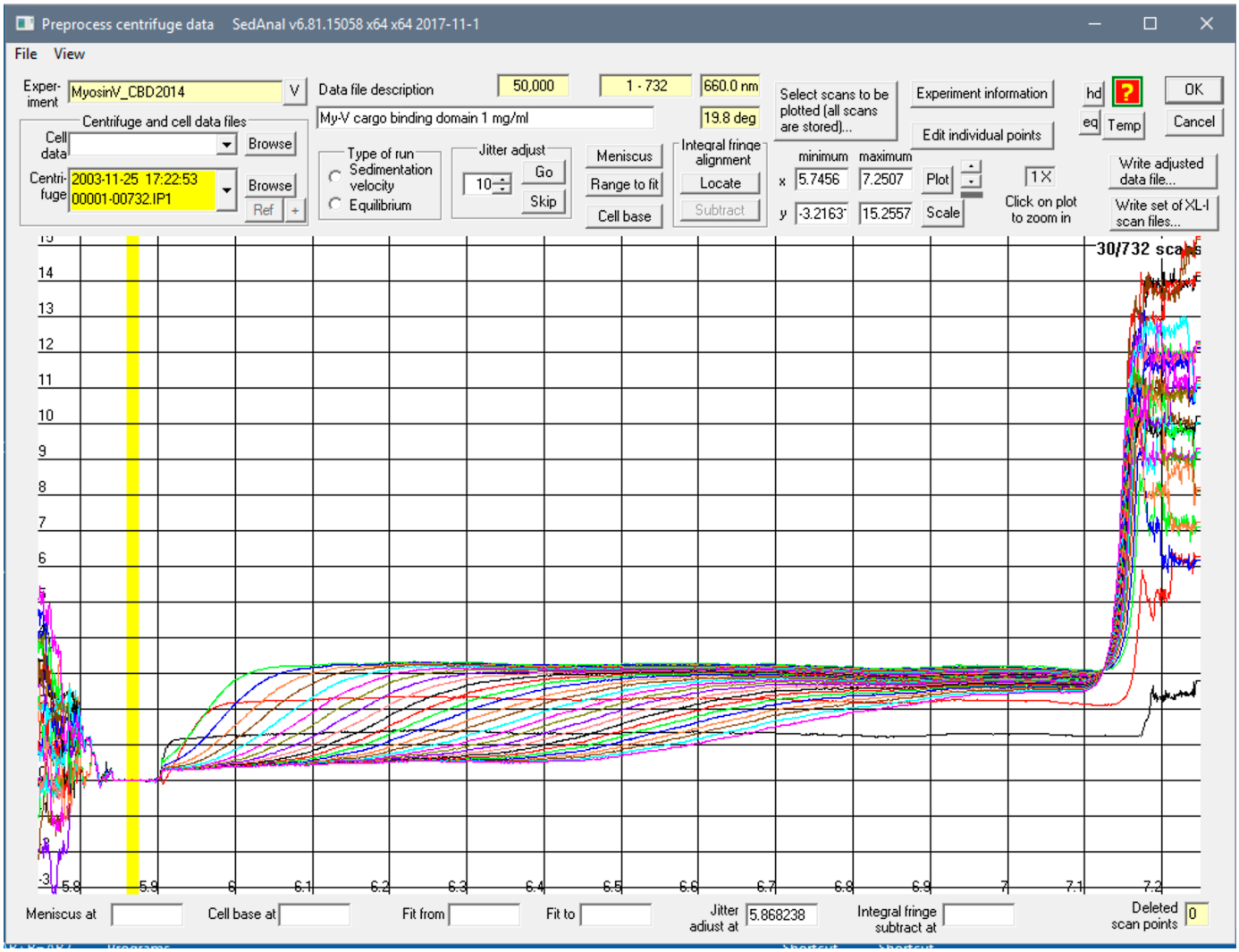


Figure 3-34 **After the jitter adjustments.** The scans are now all aligned in the air-air space.

Once the data have been pre-processed, they will be written into a "run" file (with the extention ".abr") and will be available for re-editing by selecting them in the middle window labeled "Cell Data". All the parameters, meniscus, base etc are stored in the run file and are used by the fitter to perform the fits. Run files can be re-editted at any time, for example, for re-fitting with a different range or better estimate of the base or meniscus positions.

The preprocessing procedure is explained in detail below.

The fringes are aligned first on the air-air space to remove vertical variations due to time dependent instrumental noise. Under "Jitter adjust" click on the “Go” button and then click in the air-air space. The fringe patterns will all be aligned as shown in Figure 3-34.

Now we take care of the integral fringe adjustment:

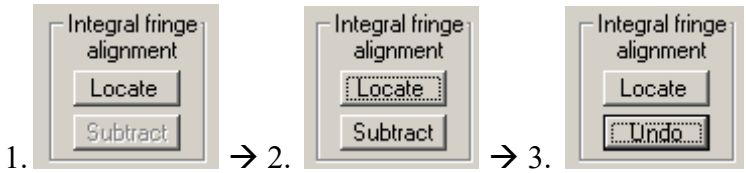


Figure 3-35 Re-adjust the data for integral fringe shifts**.**

The next step is to re-adjust the data for integral fringe shifts introduced by the data acquistion software at the meniscus. Next (1) click on the “Locate” button and select a spot – usually someplace near the bottom in the plateau region – by clicking there. Then (2) click the “Subtract” button. The result will be as shown in Figure 3-36, and the buttons will look like (3) above (Figure 3-35). Clicking the “undo” button will undo the last adjustment.

After manually re-scaling the plot, click on “Meniscus” and ‘click and drag and unclick’ at the meniscus position. Repeat the process to choose the base position by clicking on the “cell base” button and then ‘click and drag and unclick’ at the base position. Note that the base of the normal double sector centerpieces is very close to 7.20 cm and so the actual base postion will probably be at a higher radius than you are at first inclined to choose. (You can verify this by running some simulations: You’ll be surprised.)

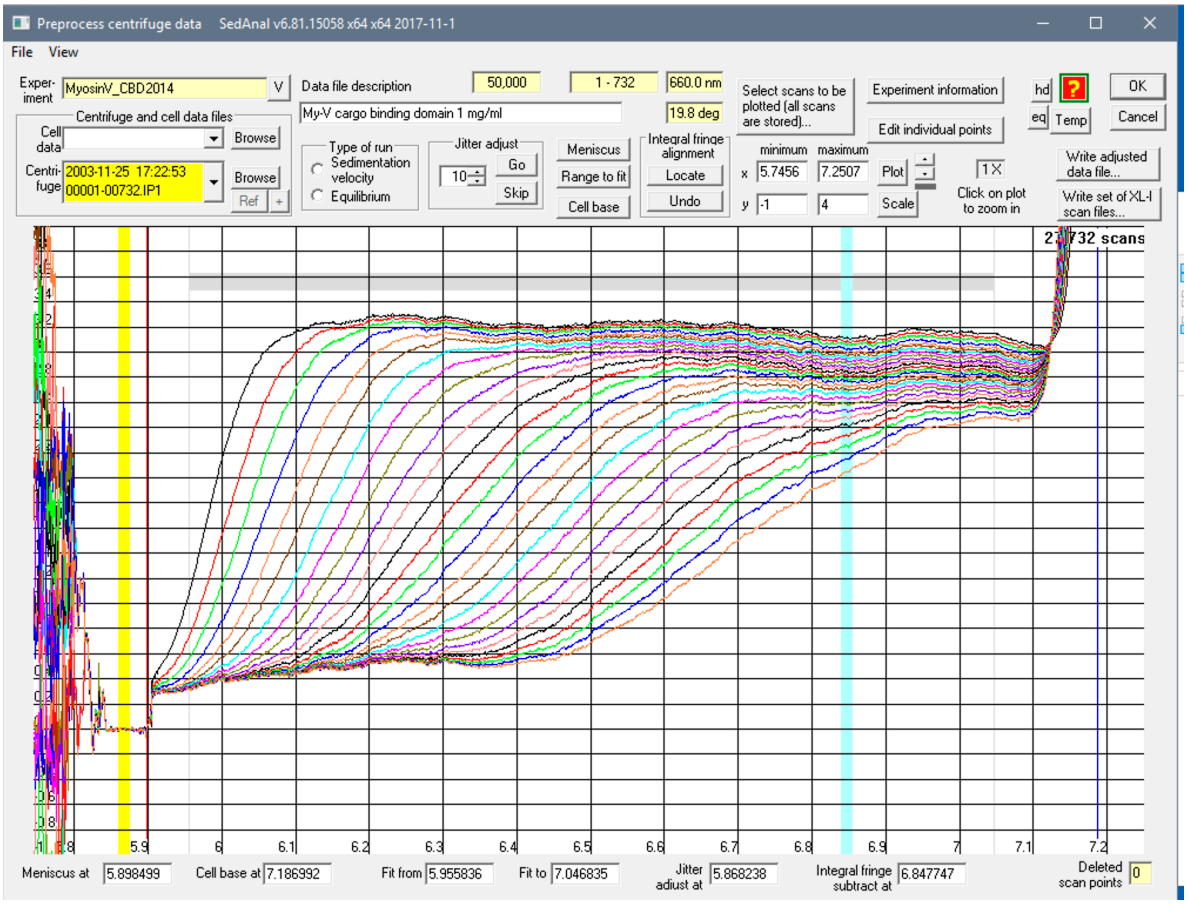


Figure 3-36 **Scans adjusted for integral fringe shifts**

After the fringe adjustments have been completed and after the meniscus, base and range to be fitted have been selected, the screen will look like one presented in with all five windows at the bottom filled in their corresponding values (Figure 3-36).

The plot can be rescaled by manually by selecting the y-axis and x-axis ranges in the four windows at the top of the plot. Click on the "Plot" button to refresh the screen.Figure 3-37

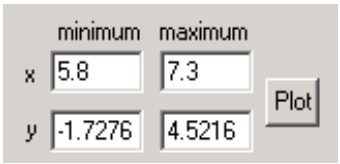


Figure 3-37 Set min and max x and y values manually

The "View" button (Figure 3-38) in the upper left corner of the preprocessor screen allows one to choose various charcteristics of the plot, as well as to show how many of the loaded scans are currently being plotted, which are shown in the upper right corner of the plot.

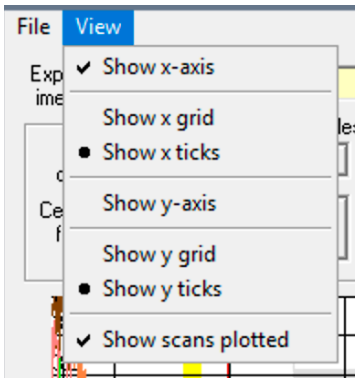


Figure 3-38 Plot format

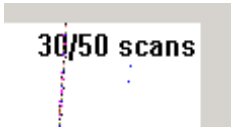


Figure 3-39 Number of scans plotted

The number of scans being plotted is shown in the upper-right corner of the plot.**(**Figure 3-39**)**

3.3.3 Selection of Scans to be displayed.

Click on the "Select scans to be plotted" button to select which scans will be displayed on the screen.

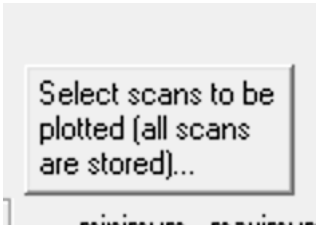


Figure 3-40 Select scans to be plotted

Note that **all** scans will be stored in the "run" file and can be used in the fitting process. Check or uncheck each scan as shown below:Figure 3-42..

A range of scans and the increment can be selected by typing in the small boxes in the middle right of the window (Figure 3-43) as shown in Figure 3-42..below.

3.3.4 Selection of scans to be fitted.

After the meniscus, base, etc. have been selected, it's a good idea to choose an initial range of scans to be fitted at this point, while you have them on the screen. Although these can be selected later in the fitting control window by right clicking on the file name for the cell whose scans are to be chosen (see below), it's easier to do it at this stage

to see what is being chosen. For example, you might decide to fit the following scans after seeing them in the preprocessor (i.e. scans 400 to 499 by 1):

Later, on the fitting screen, this range of scans can be recalled by right-clicking on the data file name and reloading them as they were displayed in the preprocessor. This also gives you opportunity to select the appropriate radial "range-to-fit" for the chosen set. Bad scans can be excluded from fitting and plotting by right clicking on the scan box as descibed in the green box in Figure 3-42. If you want one of the scans to be excluded from analysis, right-click on the corresponding box. A red "X" will indicate which scans have been labeled as "bad" scans that will be excluded from analyses.(Figure 3-41) Right clicking it again will undo the "bad" scan indication.

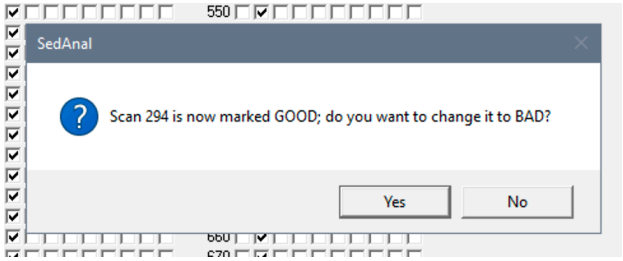


Figure 3-41 Bad scans

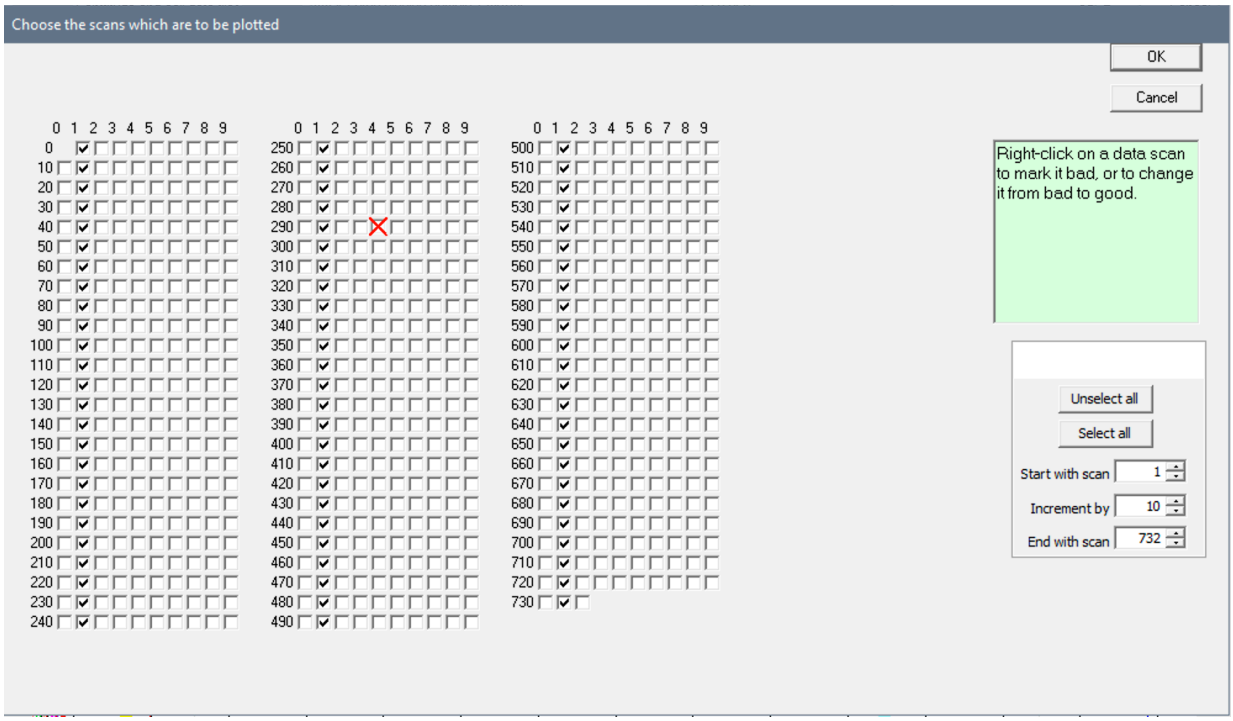


Figure 3-42. Designate scans as "bad scans"

Scan selection for Fitting (Figure 3-43)

Figure 3-43 Select scans for plotting or fitting later

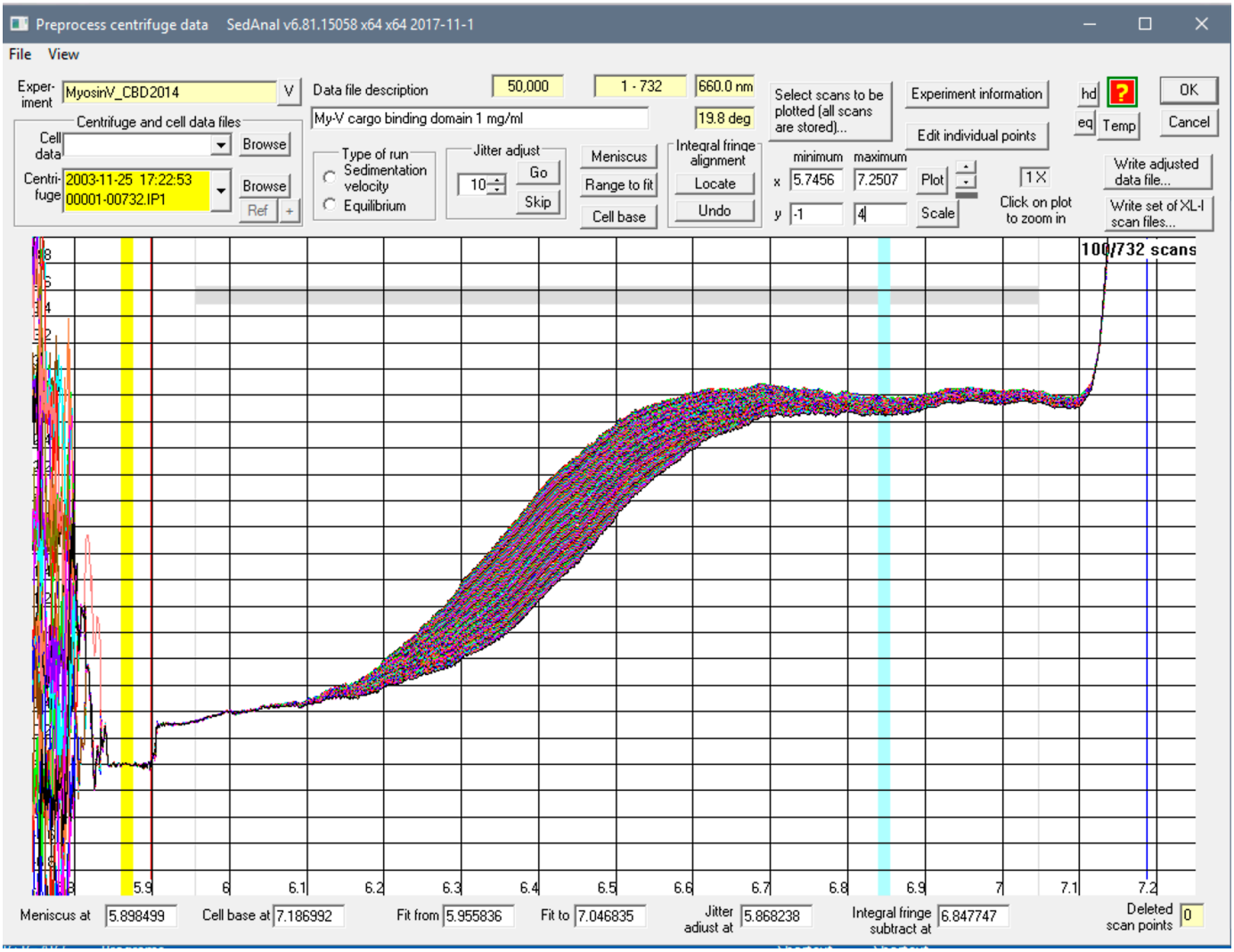
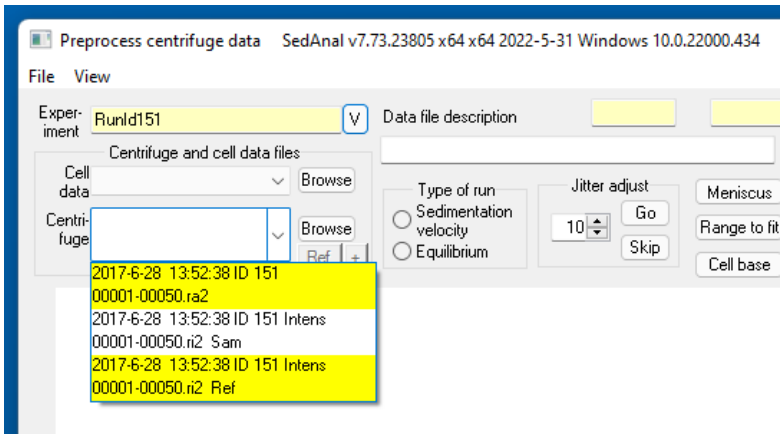


Figure 3-44 Scans plotted are from 400 to 499.

You might want to limit the fitting to just this subset and range because of an aggregate you want to ignore, for example (Figure 3-44).

3.3.5 Processing intensity data:

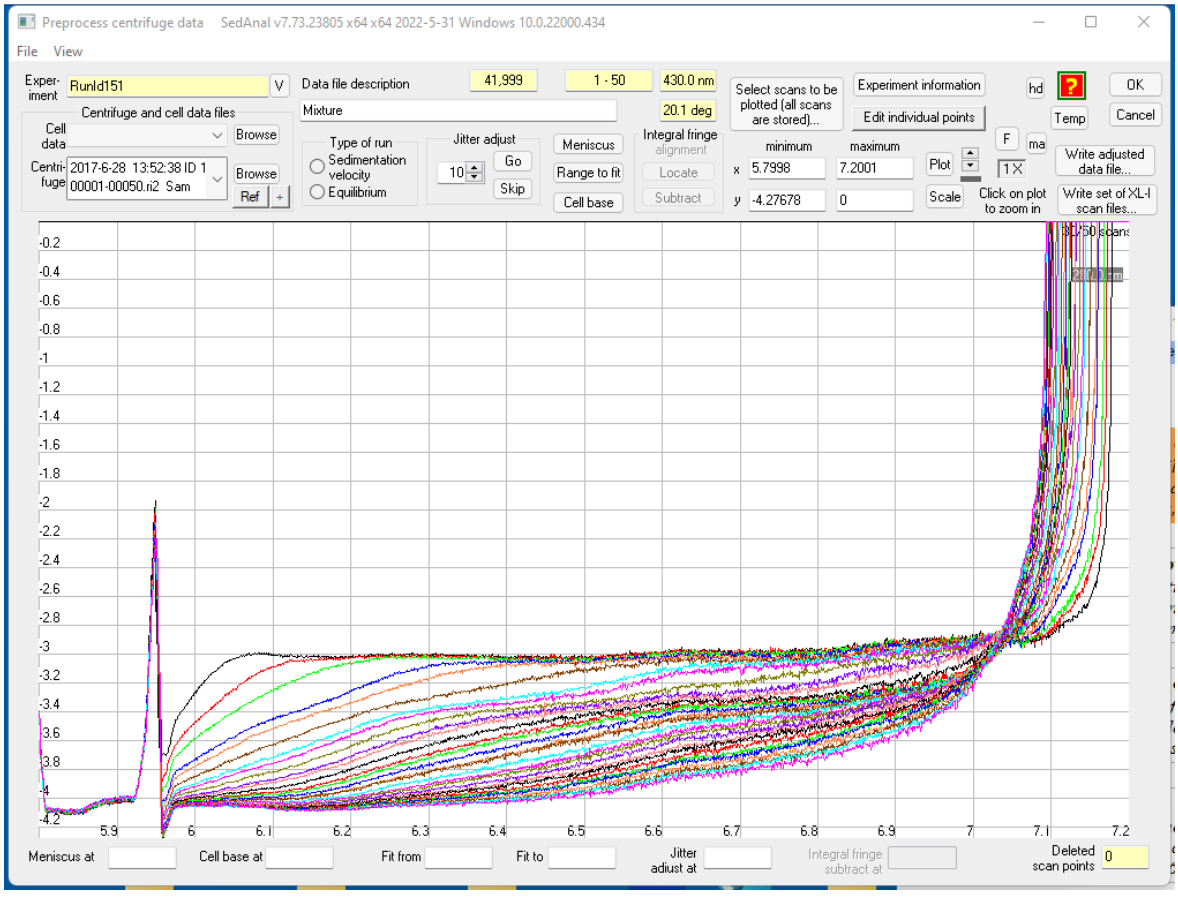
3.3.5.1 Beckman Optima data



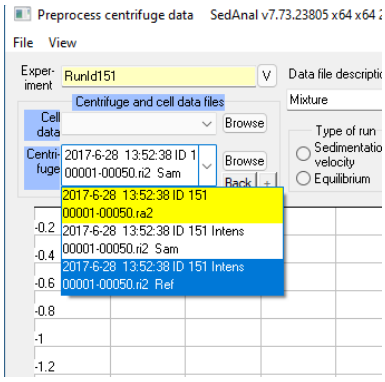
Optima intensity data appears as two sets of data, one from the "sample" sector and one from the "reference" sector. If here is a smple is each of those sectors, each dta set can be treated as a smple (doubling the thoughput). I that case the internsity data is conveted to abrsorbance units by taking the negative logarithm of the scan intenity data. These are then treated as pseudo-absorbance data.

If there is a sample in the "sam" sector and a reference buffer in the "ref" sector, then these intensity scans can be combined and after taking the ratio of the intensities, into absobance scans.

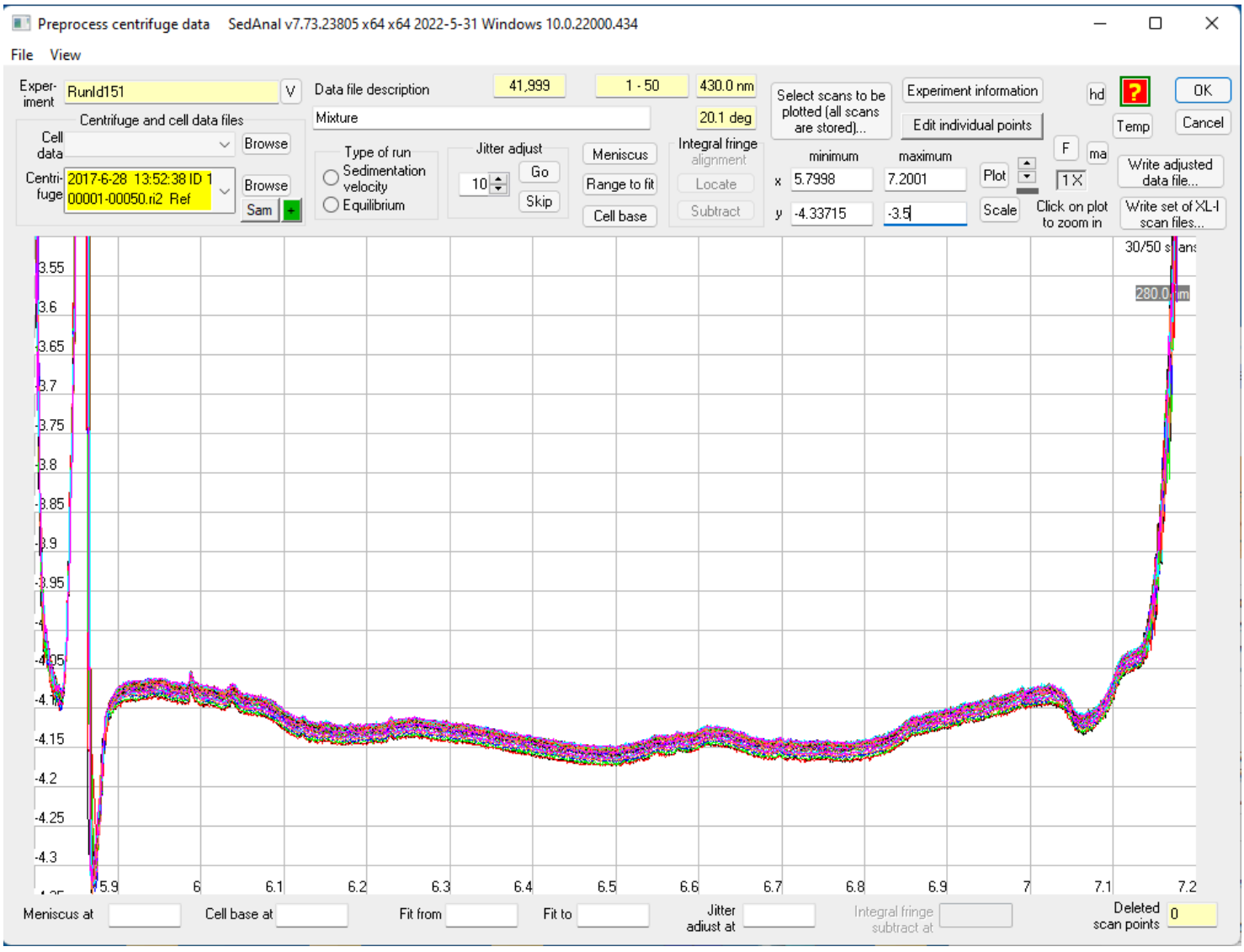
First the "sam" sector is loaded,



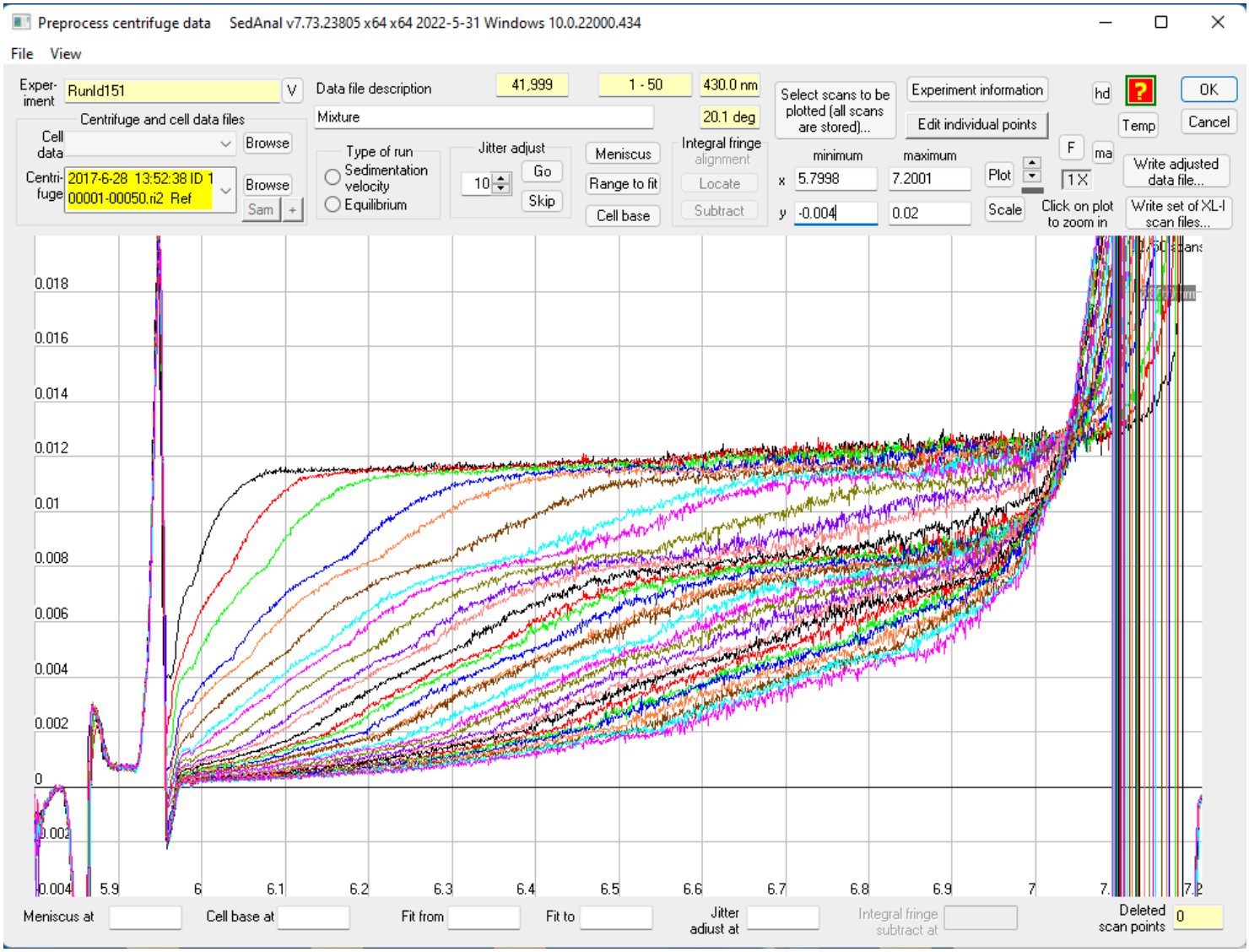
then the "ref" button is clicked to indicatet hat you intend to load reference intensity scans



The you select the "ref" scans dataset and the referene scans are displayed:



After clikcing on the "+" button, the scans are combined and converted to absorbance data.



Optima Multiwavelength intensity data are preocessed the same way.

3.3.6 Summary of the preprocessing procedure:

3.3.6.1 Zoom.

The image on the plot can be expanded by clicking at the point around which you'd like the picture to expand. There are three levels of scale expansion, the fourth click will return the image to its original size.

3.3.6.2 Interference Data.

3.3.6.2.1 Remove optical jitter:

To remove optical jitter, click on the button labeled "Jitter adjust" and then click in a relatively flat region to the left of the meniscus. The fringes will be aligned to each other by least squares fitting the curves to each other over a small region (default is +/- 10 points).

3.3.6.2.2 Remove integral fringe shifts:

Integral fringe shifts can sometime occur at the meniscus. Because there may be a horizontal region of several pixels where the are no fringes, the fringe tracing algorithm can get confused and loose count by one or more integer jumps. These shifts are eliminted by comparing fringe pattterns in a region where there is a shallow gradient like the plateau. To remove integral shifts, click on the "Locate" button under Integral fringe alignment. Then click on a region of the pattern to the right of the mencscus where the pattern changes least between scans; the plateau region is usually a good spot. A green band will appear. Then click on "Subtract" to eliminate the shifts. This is done simply by subtracting or adding integers to the data as required.

3.3.6.3 Absorbance and Interference Data:

To set the meniscus position, first click on the mensicus region to expand the image. Then click on the button labeled "Meniscus" and then click and while holding the mouse down move the black vertical line until it coincides with the meniscus, and release the mouse button. Now a red vertical line will appear at the meniscus position and the radius of the meniscus will appear in a box at the bottom of the screen. The value in the box at the bottom of the screen will track the mouse while the button is held down. You can also type a vlue for th emeniscus postion into this box.

To set the base position, repeat the same process for the location of the base of the cell by clicking on "Cell base" first.

To select the range to be fittted: The range to be fitted involves a "click-and-drag-from-left-to-right-and-release" operation. Click on the "Range to fit" button, then click on the centripetal point and drag the pointer to the centrifugal point. This selects the range of the data to which the model will be fitted.

3.3.6.4 Fluorescence and Intensity (pseudo-absorbance) data.

Since neither of these types of data have a reference solution, they are processed a little differently.

Fluorescence data are processed in the same way as absorbance data; however no background is removed from the scans at this point. The delta-c procedure in DCDT/WDA and the Fitter will remove the time independent background signal when these data are rad in for these types of analysis.

Intensity data are converted to pseudo-absorbance (PsAbs) data by taking the negative logarithm of the intensity signal. Since the PsAbs data have no reference background signal they will contain all the time independent background systematic noise from the optical system. Similarly to the fluorescence data, the PsAba data will have the time independent background signal removed by the delta-c procedure in DCDT/WDA and the Fitter.

Save Preprocessed data.

Once the data have been preprocessed, the data are ready to be fitted. The preprocessed data are stored in a "run" file with the extension ".abr".

Figure 3-45 Choose cell file name.

As shown above, the file name will be of the form “yyyymmdd_<user_suffix>.abr”, where <user_suffix> is supplied by the user when either the “OK” or the “Write adjusted data file” button is clicked. See Figure 3-45, Top: before the user adds the <user_suffix>. Bottom: After the user has added a <user_suffix>. the suffix will be instered between the inintial name of the abr file and the ".abr" extension.

3.3.6.5 Wavelength Data from XL-A.

The advantages of using absorbance optics on the ultracentrifuge is the high sensitivity and specificity; by choosing an appropriate wavelength, you can often observe one species independently of others. This is not the case with interference optics, where all species contribute to the signal.

However, multiwavelength analysis is not recommended on the XL-A: it's slow to scan multiple wavelengths and inaccurate because the monochromator does not always return to the nominal wavelengths chosen. And even when it does, it doesn't always return to the exact true wavelength because of mechanical jitter in the wavelength selection mechanism.

Nevertheless, it can be useful if the wavelengths chosen for eth analysis correspond to either a maximum or minimum in the spectrum of each component. In that case the jitter will have the smallest effect on the inaccuracy of the results.

To differentiate among species, it is convenient to use the multi-wavelength feature of the XL-A or XL-I. This allows you to have every *n*th scan be done at a different wavelength (*n* can be up to 3). This is accomplished by moving the monochromator grating between scans. The nominal wavelength is recorded in the file header. Beckman’s control software allows you to choose among a single wavelength for all scans, or alternate scans at two

or three different wavelengths. On the XL-I, you can also do an interference scan, as well as absorbance at up to 3 different wavelengths, in the same run.

When doing scans at multiple wavelengths, the monochromator reproducibility is about 2-4 nm; the reported wavelength accuracy is unknown. Therefore, it is desirable to avoid wavelengths where the extinction coefficient is changing rapidly with wavelength, although some compromise may be necessary to take advantage of the peak intensities of the Xenon flash lamp.

When preprocessing the scan files produced by the centrifuge, SEDANAL will allow you to separate the scans for each wavelength. This is necessary because the extinction coefficients vary. Normally, the choice of wavelengths is designed to maximize the difference between extinction coefficients at each wavelength for the species of interest.

There are two ways to do the separation using SEDANAL. Remember that the first step in analyzing scans is to store the data in a run file (.abr). As an example, suppose you have collected 100 scans of absorbance data at 280 and 350 nm in cell 2. The XL-A will use 280 nm for the odd-numbered scans (00001.RA2, 00003.RA2, ...), and 350 nm for even-numbered (00002.RA2, 00004.RA2, ...). (Yikes! What in heaven's name were they thinking when they thought up this scheme??)

(1) Create a single run file containing all the scans for the cell. During preprocessing you choose a single meniscus, cell base, and range to fit. On the fitting control screen, you will choose the run file *n* times, selecting sets of scans at each wavelength. For the example, we would analyze two instances of the run file; for the first, we select "scans to be fitted" of 1, 3, ..., 99, and for the second, 2, 4, ..., 100. On the control screen, the two cells will be *linked* (see below) indicating they are physically identical. For the extinction coefficients of the various species, you will put in different values for the two cells, corresponding to the different wavelengths (see below).

(2) Create multiple run files containing only the scans at a particular wavelength. Each wavelength is treated independently. During preprocessing you choose a meniscus, cell base, and range to fit for each wavelength. When fitting the data, you use the *n* run files, and no scan selection is needed, although you may want to do so to speed processing or restrict the part of the run being analyzed. For the example, we have SEDANAL create two separate run files; the first containing the 280 nm data, the second, 350 nm data. These cells will also be *linked* (see below), because they have identical loading concentrations. For the extinction coefficients of the various species, you will put in different values for each of the two cells (see below), corresponding to the different wavelengths.

The difference between these choices is only how many run files are created from the original scans, whether the scans need to be separated by wavelength, and whether different menisci, etc, can be used for different wavelengths.

The XL-A monochromator does not reposition itself to the exact specified wavelength for successive scans. In addition, the measurement of the actual wavelength, stored in the header of the scan file, is subject to error. SEDANAL has a tolerance for wavelength matching: two scans are considered to be at the same wavelength if they are within this tolerance. The value of this tolerance is set in Preferences->Preprocessor, by checking the box "Preprocessor should offer the option to separate scans..." and entering a tolerance below. The default tolerance is 4.0 nm, with the option enabled.

When the preprocessor is reading in a set of scans, it looks at the wavelengths, and if it detects more than one, it will offer the choice to create separate run files, assuming this option has been enabled in Preferences, as described above.

If the wavelength tolerance is too small, there may be more different wavelengths than intended. In this case, SEDANAL will list the first six wavelengths seen, and suggest increasing the wavelength tolerance.

The wavelength tolerance can be set in the Preferences under Preprocessor:

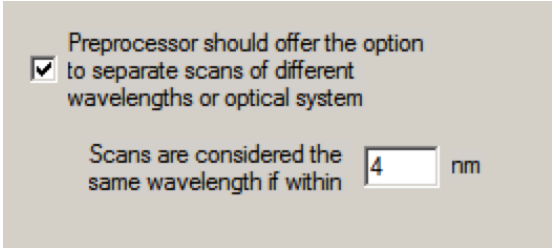


Figure 3-46. **Main menu > Preferences > Preprocessor**

3.3.6.6 Very large multi-wavelength (MWL) datasets containing several hundred wavelengths -

When loading multi-wavelength data from either Helmut Cölfen's 2D detector, or from Kristian Schilling's 2D detector, a new window will open allowing the user to select which of the many possible wavelengths (up to 2048) to process in the Preprocessor. (J. Walter et al. Anal. Chem. 87(6):3396-403, 2015).

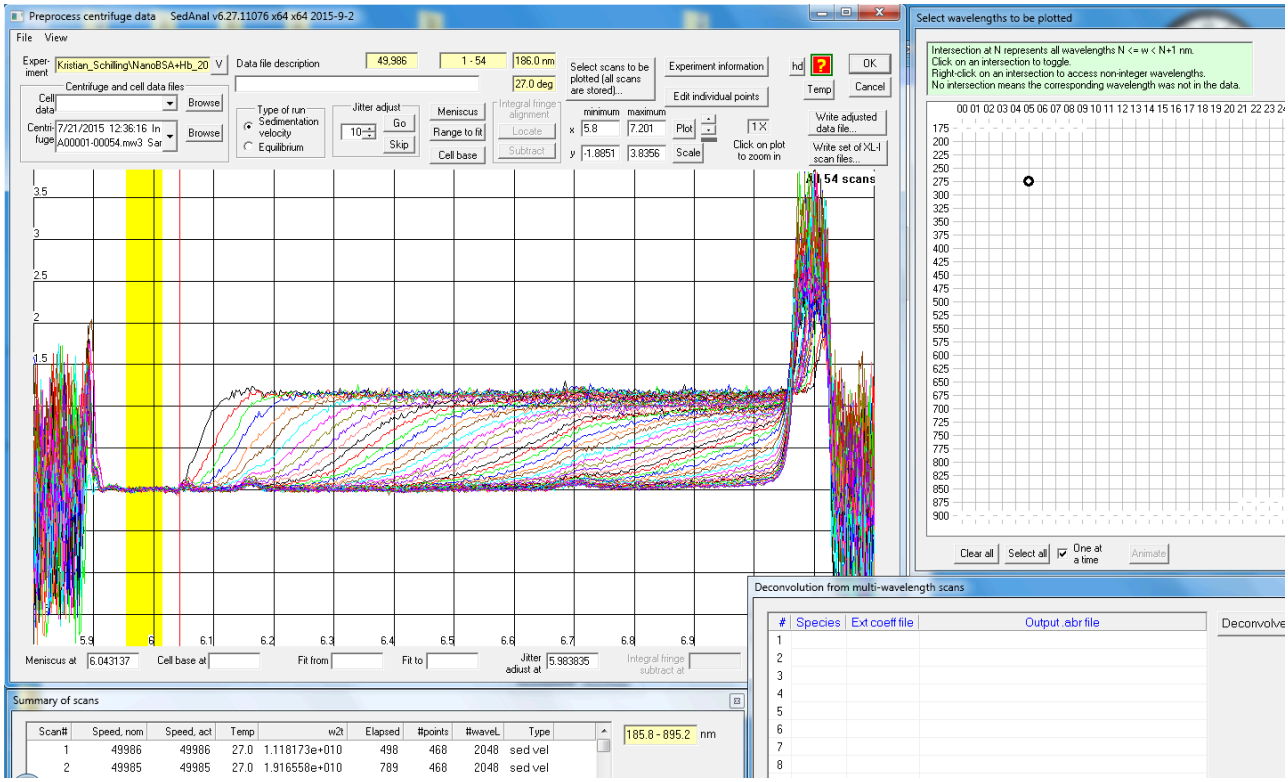


Figure 3-47 Scans from several wavelengths

Scans from multiple wavelengths can by plotted here but usually only one is necessary to choose the meniscus, base and range-to-fit, and take care of vertical jitter. After adjusting the jitter, only the meniscus position is necessary for either DCDT or WD analysis. But base radius and range-to-fit are required for the Fitter

3.3.6.7 Deconvolution of Concentration Profiles from Multi-Wavelength Data

If the user has the extinction spectra of the components that compose a mixture, the Preprocessor can read those data in and then can deconvolute the MWL data into component concentration profiles that are stored as separate abr files for further processing by either WDA or the Fitter.

3.3.6.8 Preprocessing multi-speed datasets for Wide Distribution Analysis (WDA):

When loading a multi-speed dataset, the preprocessor will stop after loading the scans from each speed and allow you to choose a radius at which to align the fringes in the air-air space and to choose a meniscus position at that speed. The integral fringe subtraction will be done after the scans from the last speed have been loaded.

After the scans from the first speed have been loaded the boxes at the top of the screen will look something like Figure 3-48.

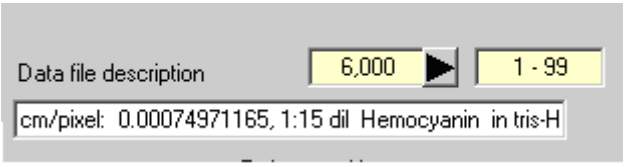


Figure 3-48 load speed set

One clicks on the little black arrow to advance the preprocessor to the next speed's dataset where the next meniscus and vertical alignment in the air-air space are carried out. Neither the ""Range to fit" nor "Cell Base" needs to be set for DCDT or WDA data.

Figure 3-49 Load more speed sets

After all the speeds have been processed, carry out the integral fringe alignment and click "OK" to save the abr file. Next proceed to the DCDT/WDA button on the Main menu.

If it has not been set in the preferences as the default (which is recommended), click the "Wide distribution" button (left side above the dcdt window) to enter the WDA mode. Then load the abr file. Enter 1.0 for the smoothing window. The data from each speed will appear in a different color (Figure 3-50). See below for more on WDA.

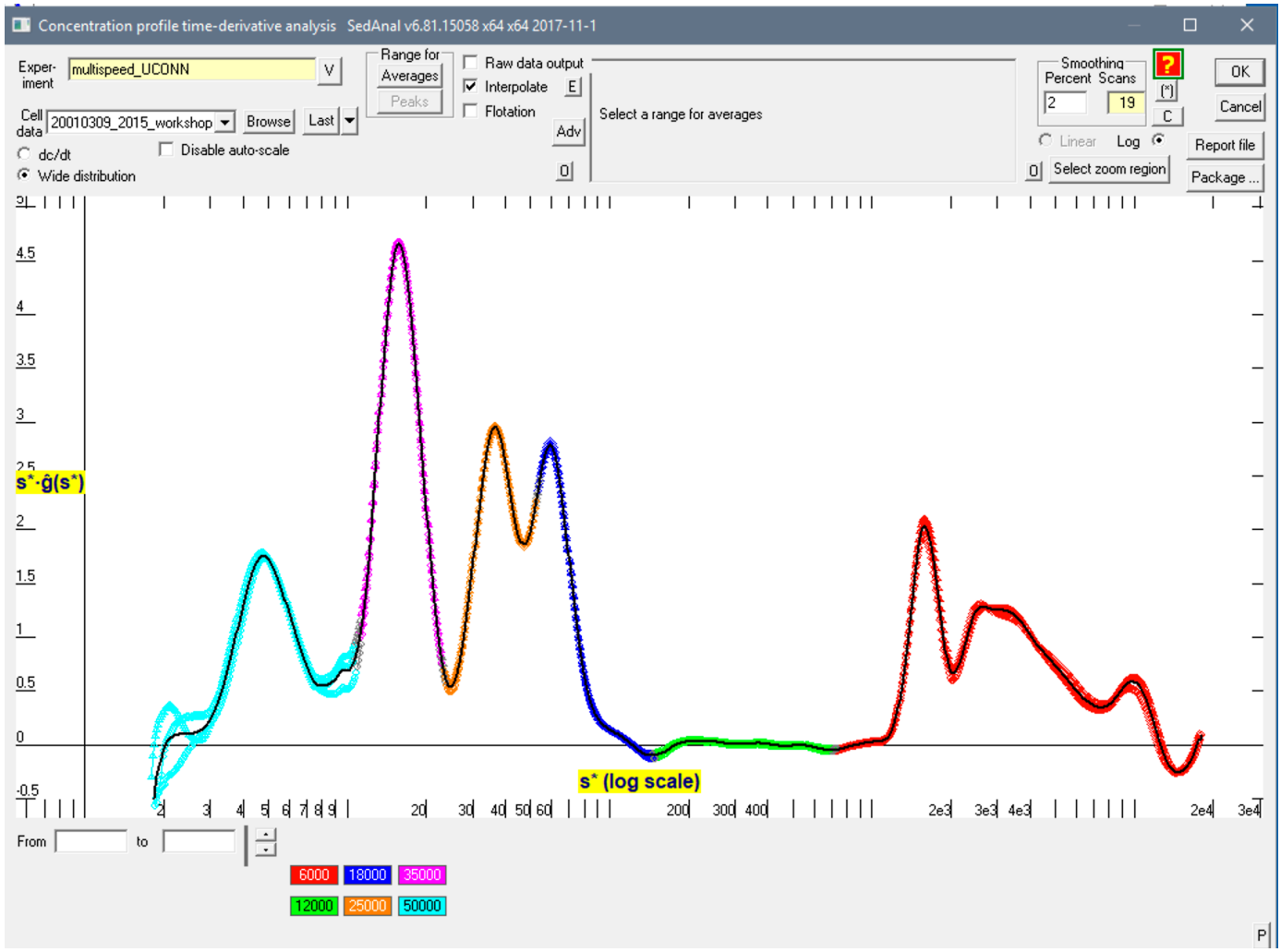


Figure 3-50 Multi-speed Wide Distribution Analysis (WDA).

3.3.6.9 Preprocessing Multiwavelngth data from the Optima

Preprocessing multiwavelength data from the Optima is essentially the same as for the OpenMWL machines but usually with many fewer wavelengths. Components can be deconvoluted in the same way when a spectrum for each component has been measured with by WDA or in a spectrophotometer.

3.3.6.10 Preprocessing Sedimentation Equilibrium data

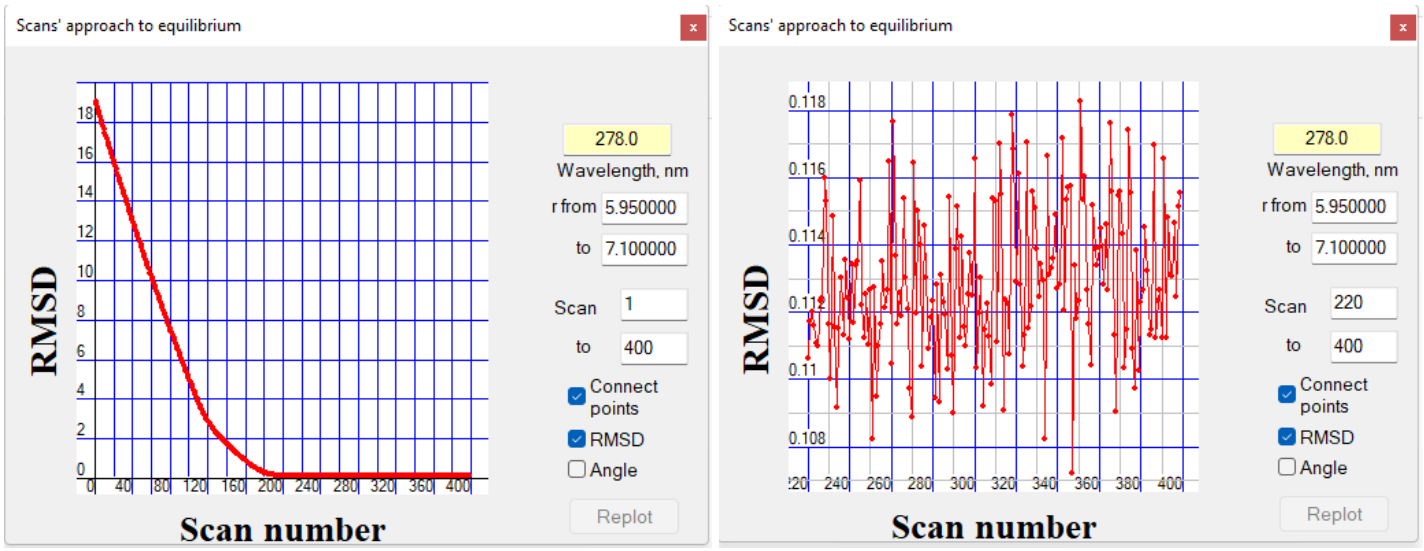
The scan or scans to be precessed for SedEq analysis should be put in a folder on the same “date/time” path as for sedimentation velocity. Only the scan(s) to be precessed should be made visible in that folder. SEDANAL will allow the user to load an equilbrium scan and an optical baseline scan (a background scan taken usually with water vs water to be used to correct for systematic errors that arise from irregularities in the optical system). SEDANAL will subtract the optical background scan from the equilibiurm data scan. This is mandatory for interfernce data and often helpful for absorbance data if there is a scratch on a window or dirt on the optics. SEDANAL will perform an interpolated baseline subtraction using the optical basline scan to correct the run scan at corresponding radial points.

3.3.6.10.1 Testing for Equilibrium

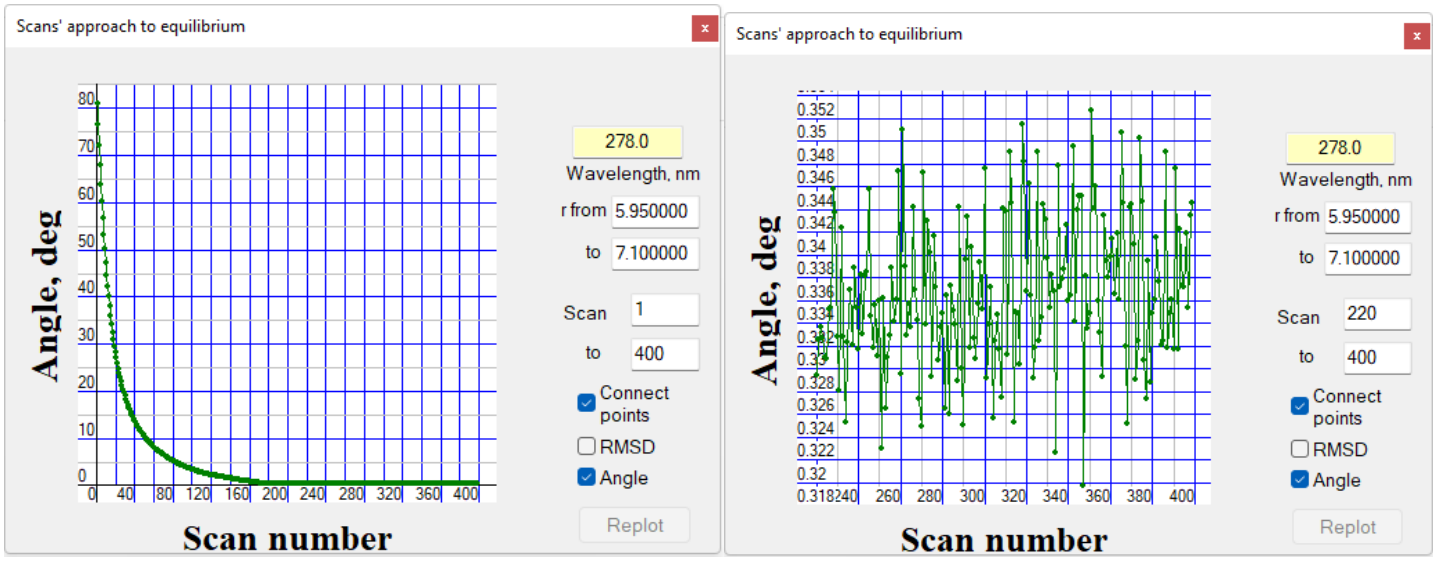
Displaying approach to equilibrium

The Preprocessor displays the approach to equilibrium when either a set of centrifuge files, or a cell data (.abr) file is loaded. It does this by comparing each scan to a reference scan (normally the last one). The comparison is displayed numerically in the scan summary window, and graphically in a new Scans' approach to equilibrium window.

Two plots can be shown: either (1) the root mean squared deviation (RMSD) between the reference scan and each earlier scan is computed as the average of the squared deviations, corrected for baseline shifts (i.e. vertical jitter).



or (2) the user can choose to plot the "angle" between the last scan and the others by taking the **dot product** between each earlier scan and the last scan.(Ninety degrees mean orthogonal and zero degrees means they are the same.)



The left plot shows the entire span of the run, while the right window show, in this case, scans 200 to 400. It is clear that equilibrium has been reached effectively by about scan 220.

Figure 3-51 Displaying approach to equilibrium

The "r from" and "to" radii are chosen automatically, in this priority:

1. If the user has entered radii, those are used.
2. Else if the meniscus and/or cell base has been chosen, it will be used after adjustment . The adjustment is to add 0.02 cm to the meniscus, and subtract 0.002 cm from the cell base.
3. Else if the range to fit has been chosen, it will be used
4. Otherwise, the center 96% of the range of radii for all scans will be used as an estimate.

The reference scan and starting scan for comparison are taken to be the initial and final scan, respectively. The initial parameters and appearance of the graph are controlled by eight preference variables, but at present (v7.55), these are not editable by users.

The Scans' approach to equilibrium window can be resized by dragging on its edges.

For **multi-speed data**, the approach to equilibrium is calculated **separately for each speed**:

This feature is useful. for example, when using a multispeed equilibrium method to verify whether or not equilibrium had been achieved before each of the speed changes.

3.3.6.10.2 Equilbrium run in standard double sector centerpiece.

Here is an example from a run performed in a standard double sector cell with a loading volume of 0.25 mL (Figure 3-52).

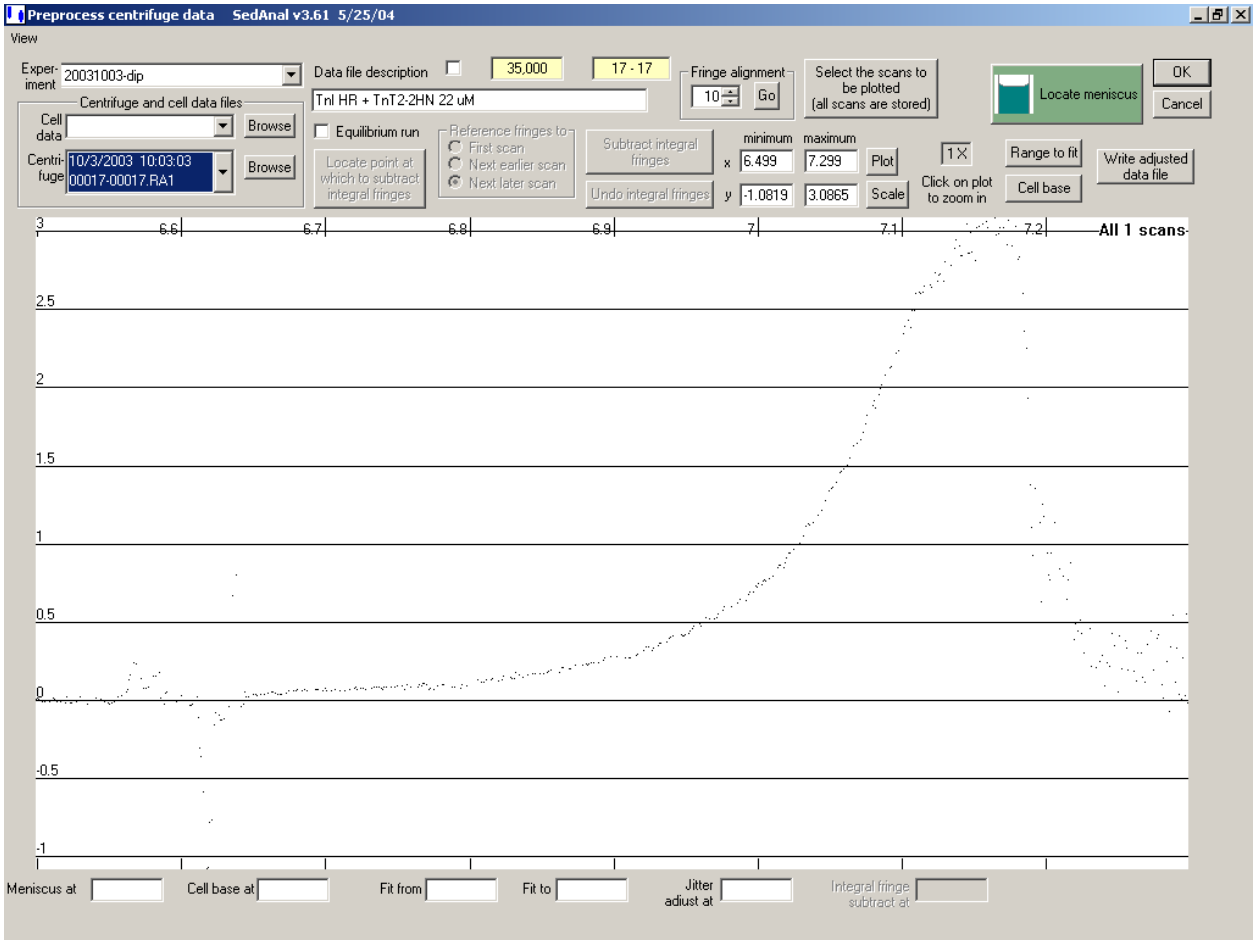


Figure 3-52 Equilibrium scan

After selecting the meniscus, base and range-to-be-fitted, the screen will look like in Figure 3-53, (**Do not forget to click the "Equilibrium" check box**).

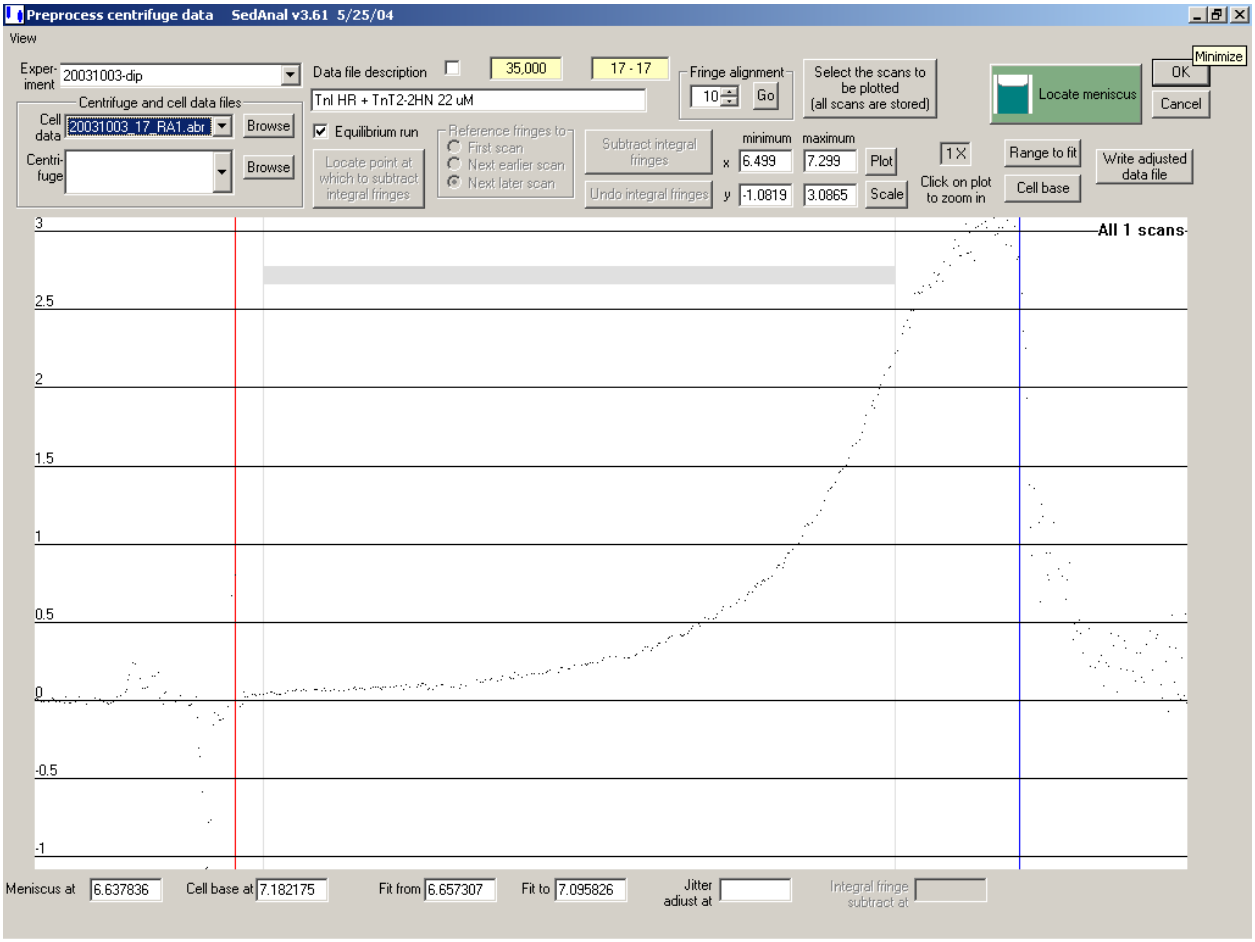


Figure 3-53 Scan with meniscus, base, and range to fit selected

If an optical blank is to be subtracted from the data run, select it now in the Centrifuge data file drop-down window. Three curves will apear, the equilibrium data, the blank, and the corrected data.

3.3.6.10.3 Multi-channel data with 6 channel centerpieces

After loading a dataset from a 6 channel centerpiece, we select the meniscus, base and range to fit for each channel and repeat until three cell data files are produced one for each channel pair.

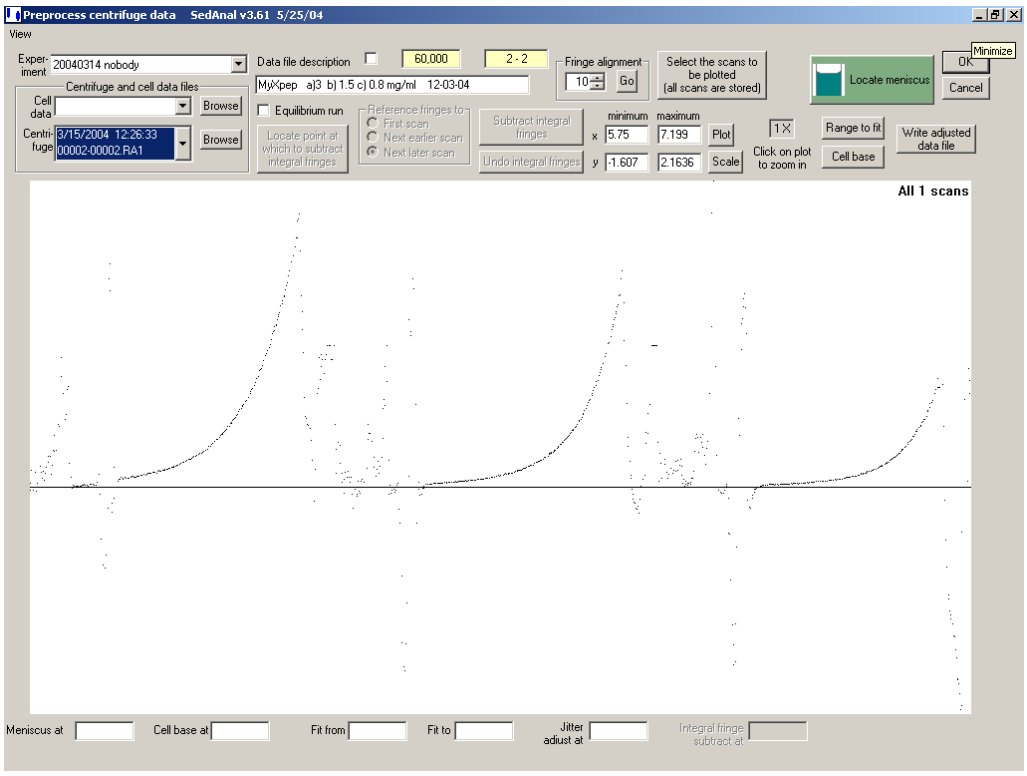


Figure 3-54 **Data from 6 channel Yphantis type equilibrium centerpeice.**

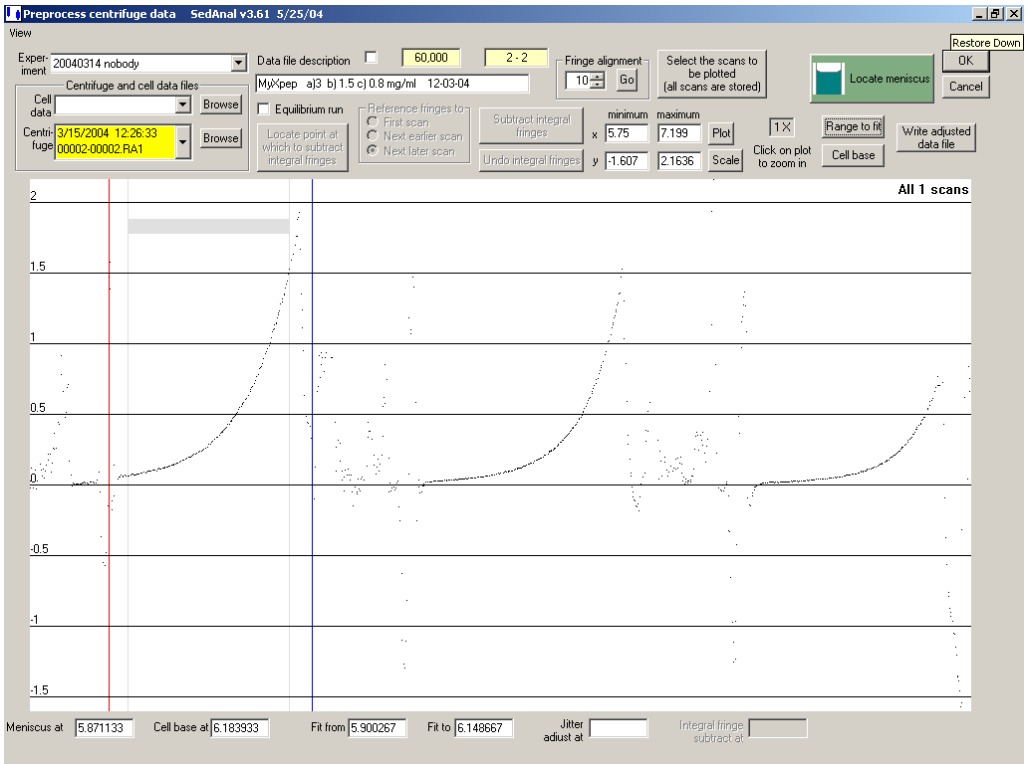
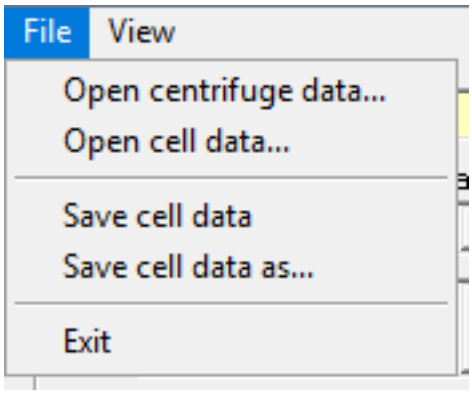


Figure 3-55 Selecting channel A. Repeat for remaining channels

After the meniscus, base, and range-to-fit have been chosen for the first channel, the processed data can be saved with a name like 20040314_60K_CHA, for example. And then these data must be reloaded and the process is

repeated for the other channels and each is saved as a separate cell data (*.abr) file using similar names (eg. 20040314_60K_CHB, and 20040314_60K_CHC, etc). Rather than reloading the file three times, use the **"Save cell data as..."** feature of the **File** menu in the upper left corner of the menu bar:



3.4 Fit Preprocessed Data

3.4.1 The Control Screen

When "Fit Preprocessed Data" is selected from the Main Menu, the default control screen will appear.

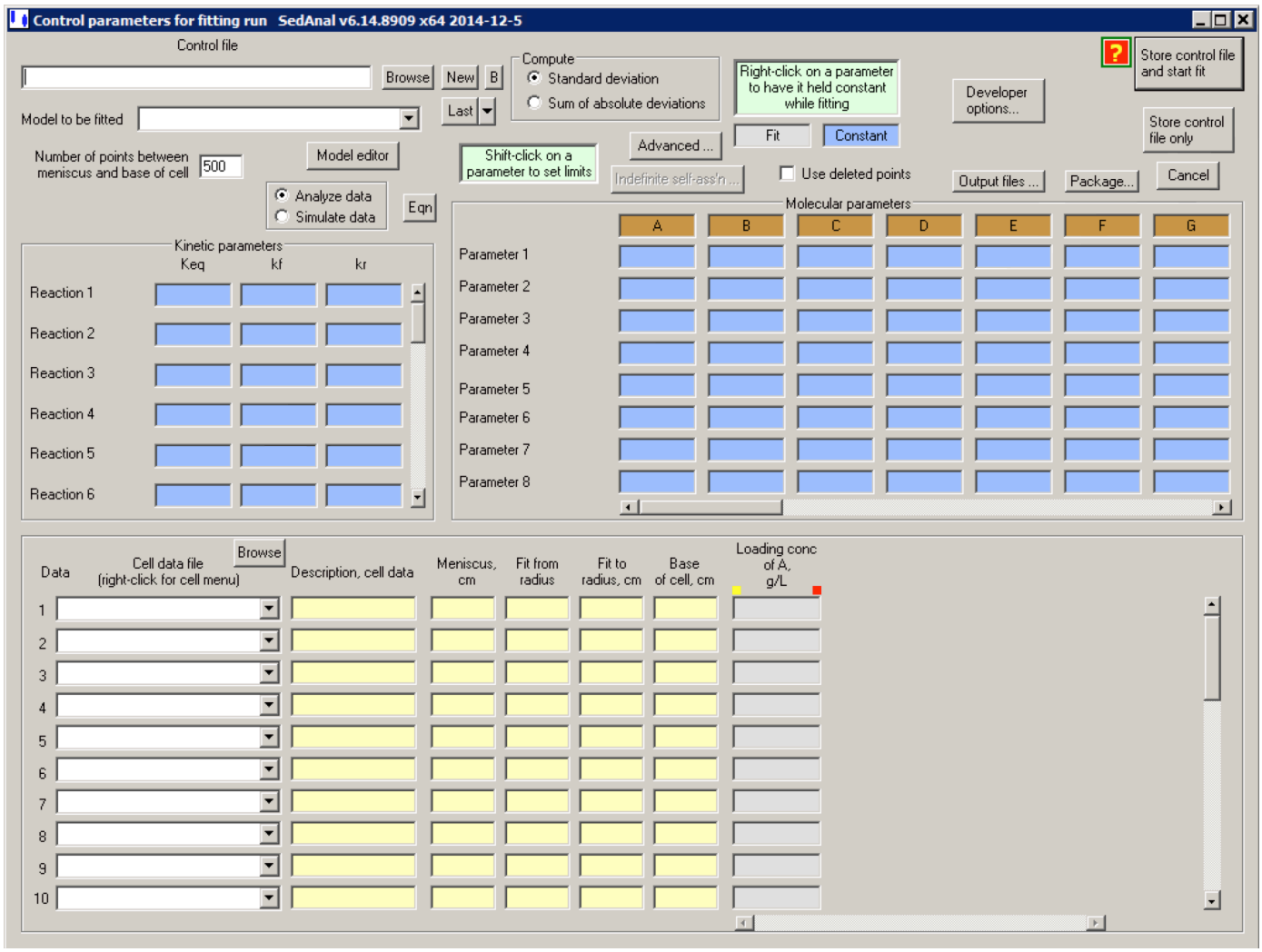


Figure 3-56 The main control screen for the fitting preprocessed data

3.4.1.1 Create a New Control File

To start a new fit, click on “New” and a dialog box asking you to specify an “experiment” directory will appear. The directory you specify next will be the location in which the control file will be stored with extension “.abc” (Figure 3-57).

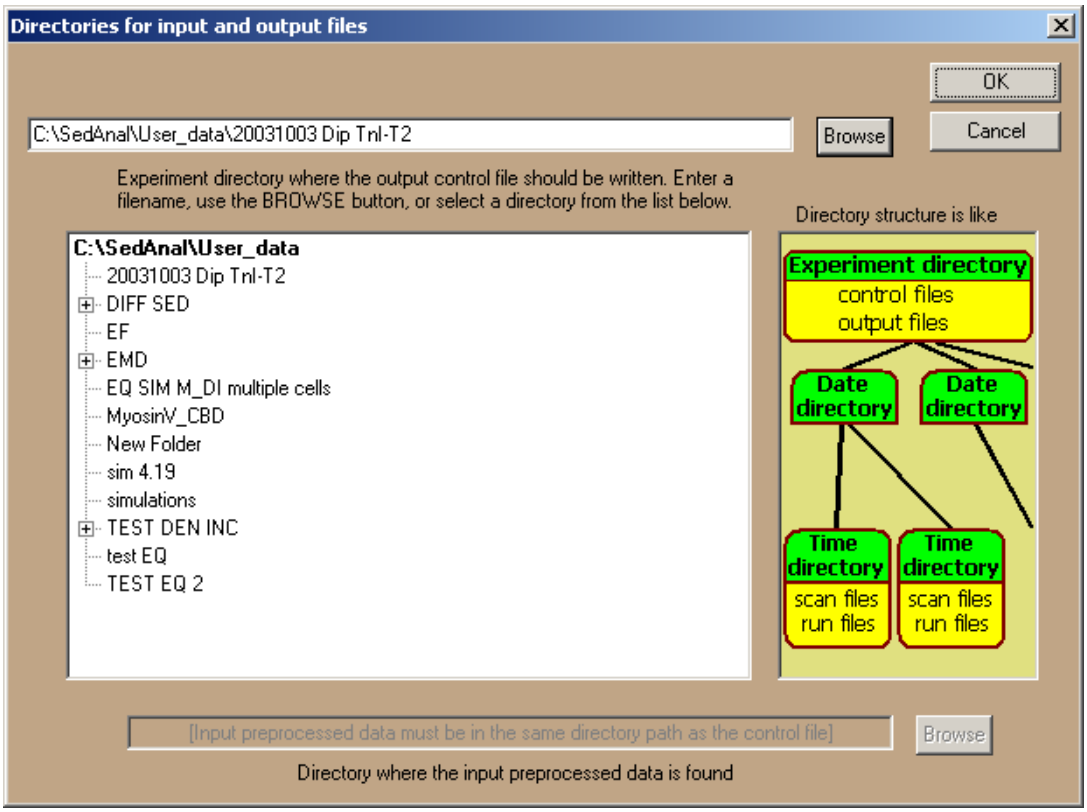


Figure 3-57 Dialog box for output directory

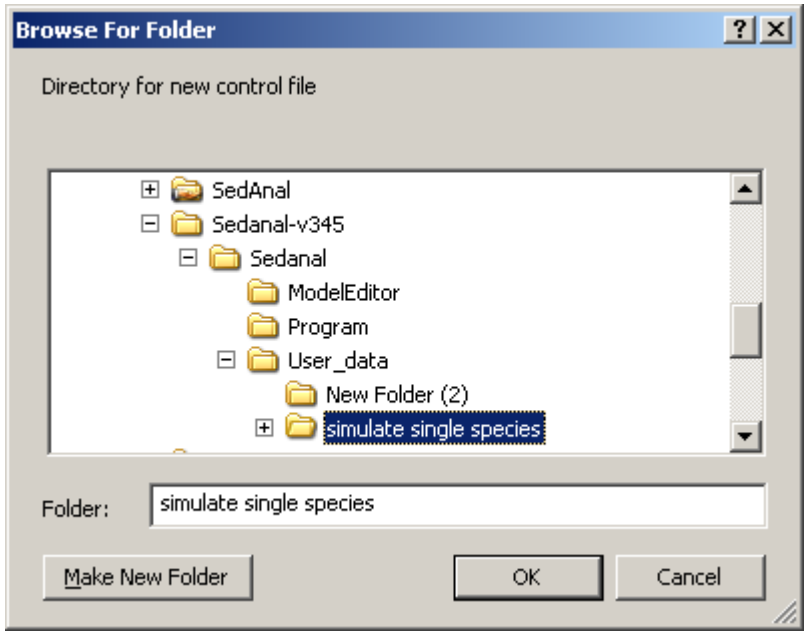


Figure 3-58 Select experiment folder

Click on “Browse” and select the experiment directory in which the control file will be stored, here: “simulate single species’

The control file name will appear in the Control file box. It will be named “New_Control_Filename” by default and should be renamed immediately at this point before storing it.

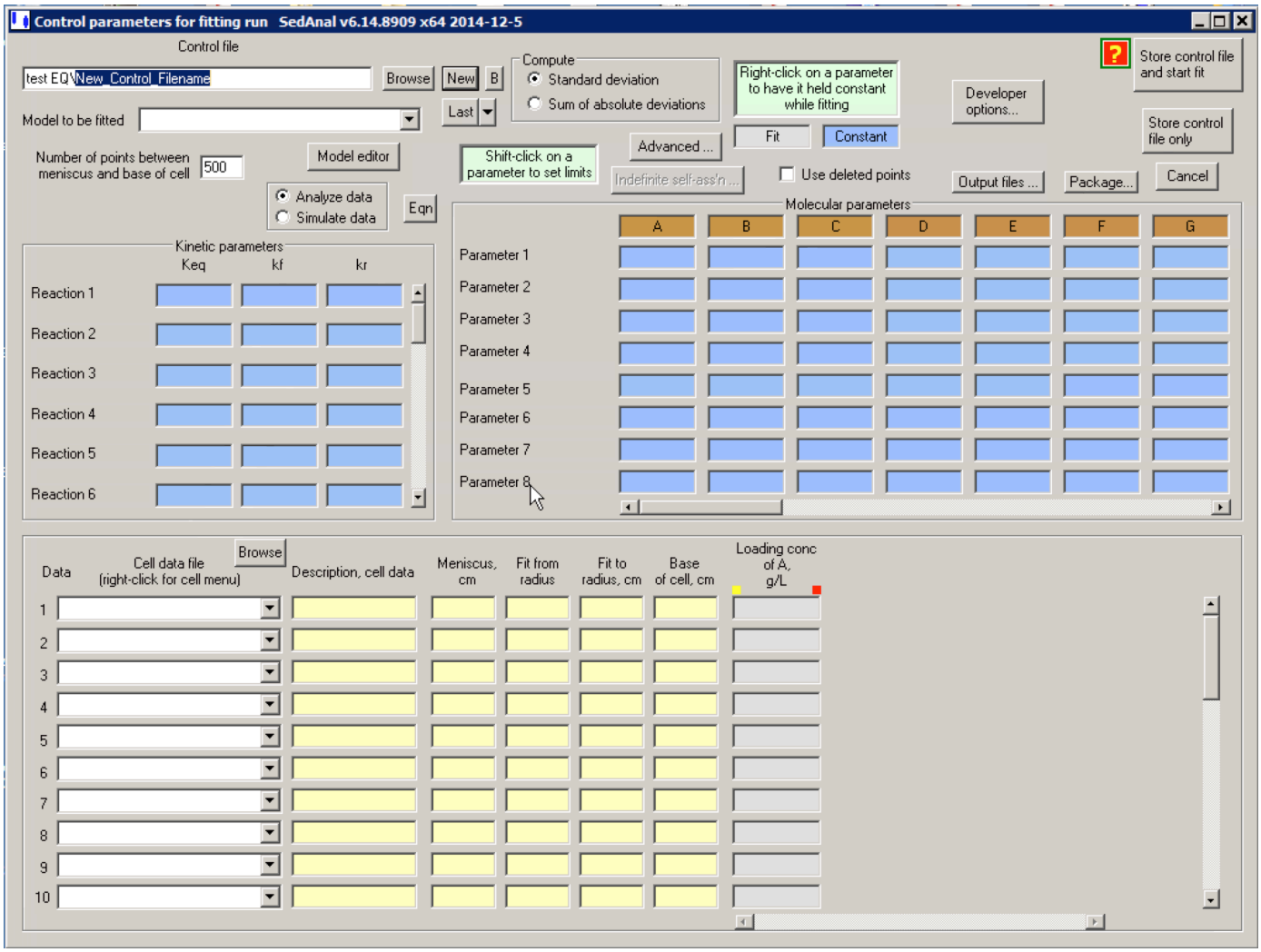


Figure 3-59 The main control screen for the fitting preprocessed data

3.4.1.2 Choose between Analyze Data and Simulate Data

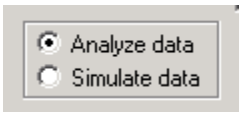


Figure 3-60. Now indicate whether you are going to "Analyze data" (the default for fitting) or "Simulate data" by clicking the appropriate button, and select the model from the "Model to be fitted" dropdown window.

Figure 3-60 Analyze vs simulate

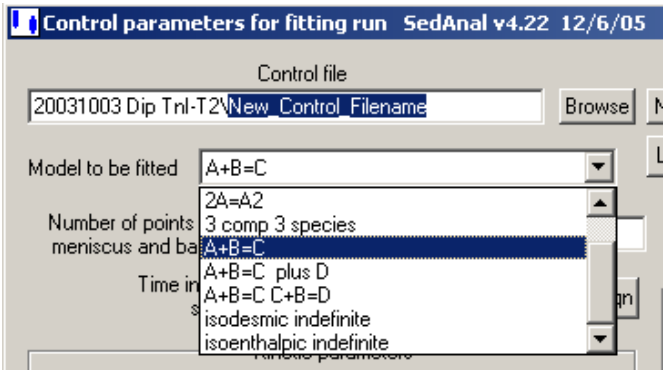


Figure 3-61 Select model

Now the Control Screen will change to correspond to the model chosen Figure 3-60.

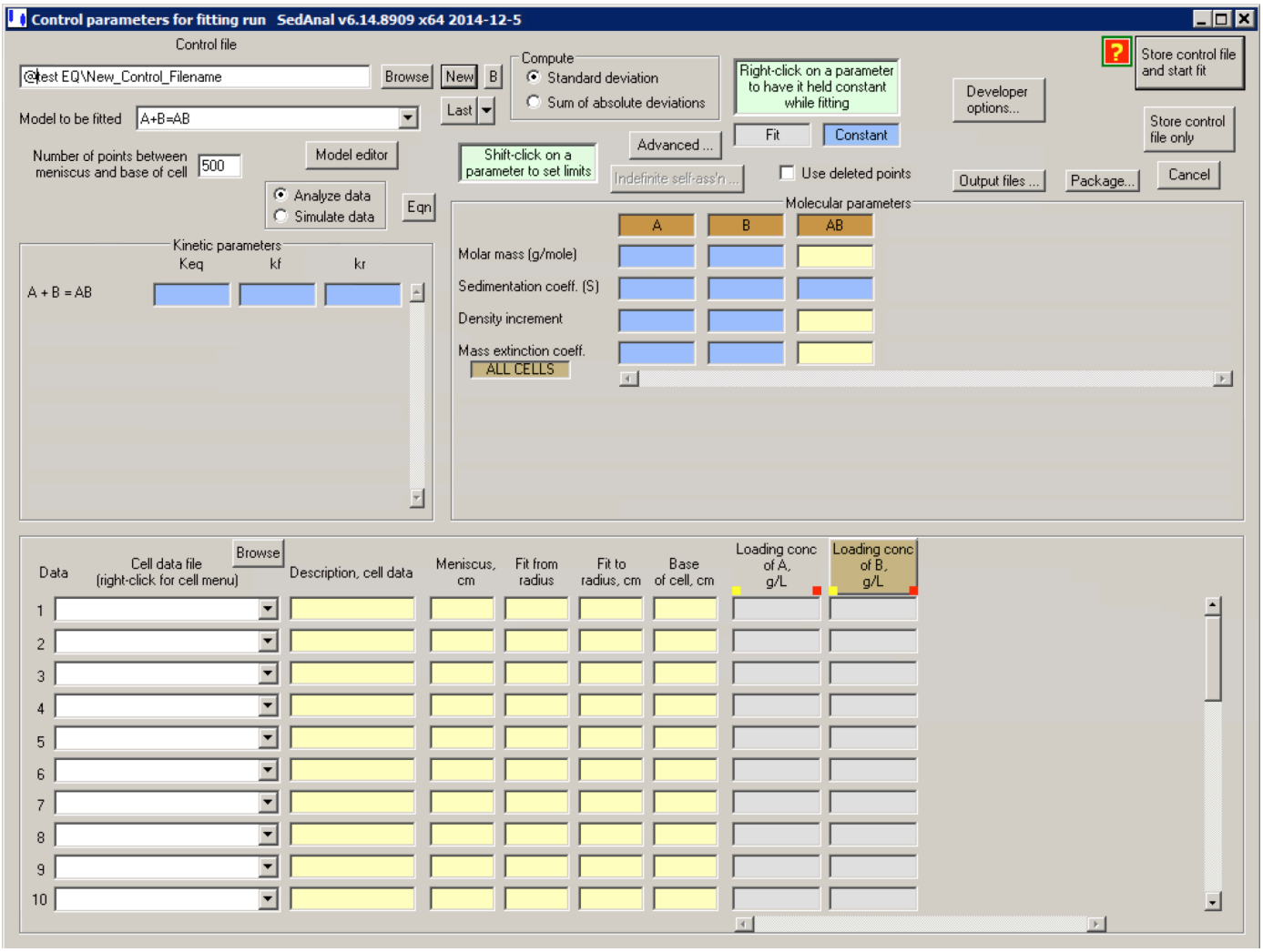
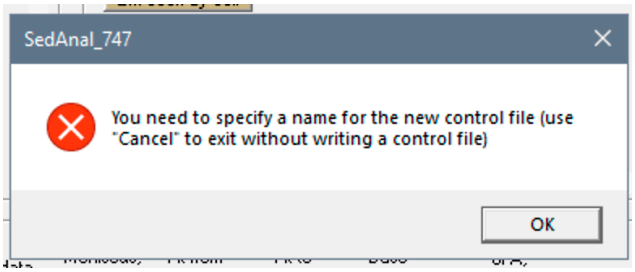


Figure 3-62. The main control screen for the fitting preprocessed data. As the model is selected the appropriate boxes will appear on the top portion of the control screen.

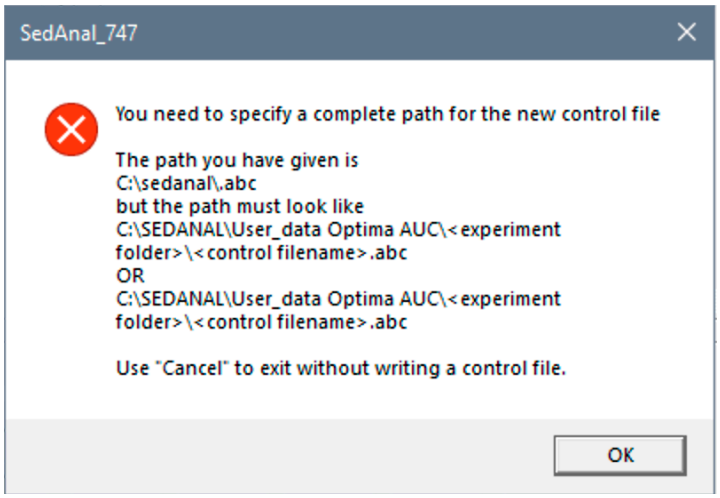
3.4.1.3 How to Save the Control file if you haven't started with the "New" button.

If you have filled in all the parameters but had not pressed "New" before you started, you will see this message:



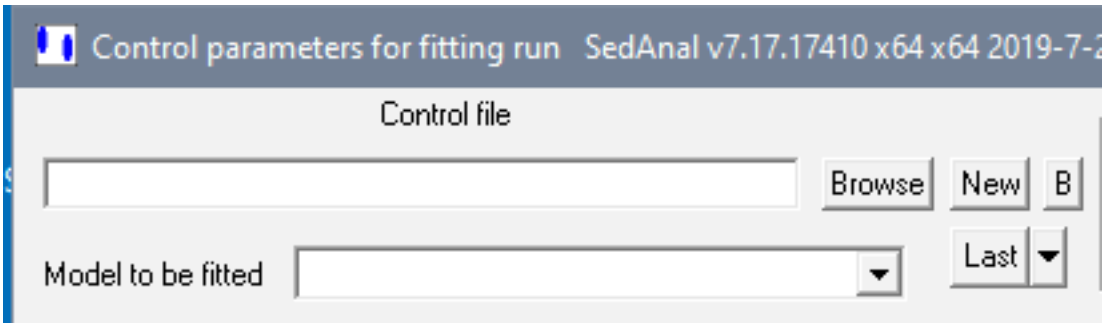
You can still save the new control file by typing the full path to the folder, into which you want to store the new control file, into the path window just to the left of the "Browse" button.

At that point, if you make a mistake in the path, you will see this message:



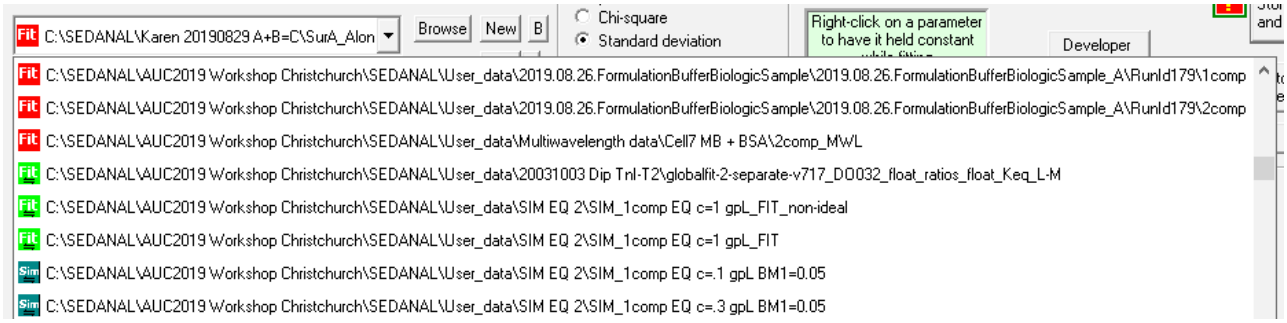
After you have entered the correct path, SEDANAL will save the new control file.

3.4.1.4 Reload a previous control file



There are two ways to load an existing control file: (1) To reload the last control file for the fit you just finished; click on the "Last" button. The previous abc file will be loaded onto the screen and to update the best fit parameters, click on the "B" button and the best fit values of the parameters that were allowed to float in the last fit will be loaded in RED font. (2) To load a previous abc file from a fit performed at some other time, click on the button with the down arrow to see a list of abc files that were processed previously (The number of files in this list can be set in the Preferences.)

When you click on the down arrow you will see a list like this:



The icon at the start of each line identifies the type of fit: SV, SE, or SIMulation

3.4.1.5 Loading cell data ("abr") files:

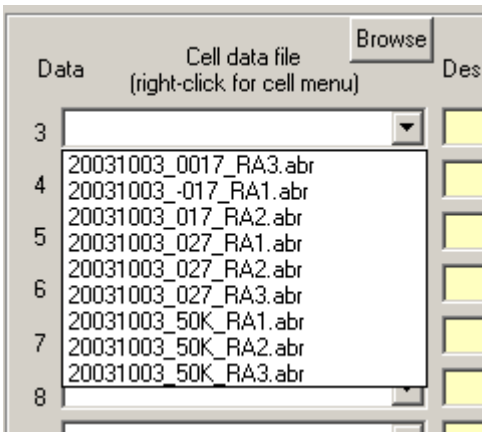


Figure 3-63 Select cell data file

Figure 3-63. **Select the cell data (*.abr) file from the "Cell data" drop-down window.** When the Cell data file is selected, the type of run (velocity or equilibrium) will be recognized and the parameter boxes on the control screen updated accordingly. The default screen assumes velocity run until the abr file is loaded.

If the cell data files are NOT stored on the default path, you will have to "Browse" for the files. This could happen, for example, if the runs were done at different times and/or stored in different folders.

3.4.1.6 The control screen is divided into several regions.

3.4.1.6.1 TOP LEFT:

In the top left region of the control screen, one supplies a name for the **Control file.** Below that window is a drop-down window from which a **model** may be chosen. Below that is a box for entering the **number of points** to be

used between the meniscus and the base of the cell by the Lamm equation solver:. (Other parameters used by the Lamm equation solver can be set under the **Advanced > Fitting** tab (or from the Preferences > Control extended > Advanced > Fitting); these include the **time interval** to be used by the ODE (Ordinary Differential Equation) solver, and the maximum **number of iterations** to be performed by the non-linear least squares function minimizer ... more on this later.) The Equation Editor can be accessed by clicking on the "**Eqn**" button. And the Model Editor can be accessed by clicking on the "Model Editor" button.

3.4.1.6.2 MIDDLE-LEFT:

In the mid-left region, in the box labeled "**Kinetic Parameters**", kinetic rate constants (k_f, k_r) or equilibrium constant (Keq) may be entered by the user. If an equilibrium constant is chosen, the internal routine uses 0.01 sec^{-1} for the reverse rate constant and fits for the forward rate constant only. The default value for the reverse rate constant may be changed on the **"Advanced > Kinetics/equilibrium control"** tab. If nothing is entered in the box labeled Keq, one must supply an initial guess for both the forward and reverse rate constants. Other parameters used by the Kinetics ODE solver also may be adjusted in the **"Advanced > Kinetics/equilibrium control"** tab.

3.4.1.6.3 MIDDLE-RIGHT:

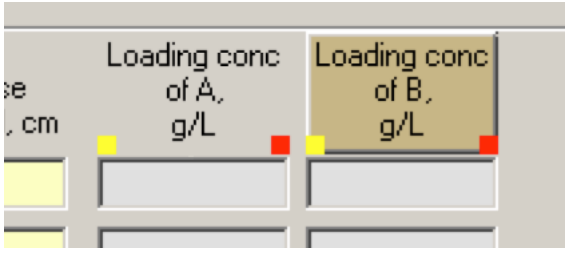
In the middle right region, in the box labeled **"Molecular parameters"**, one enters the molecular parameters for each of the species participating in the reaction scheme, as well as those which might be extra components. These parameters include molar mass or diffusion coefficient or frictional coefficient, sedimentation coefficient, density increment, and mass scale extinction coefficient for the optical path length corresponding to the centerpiece in use.

3.4.1.6.4 LOWER:

The Lower region allows one to select the datasets that will be fitted either singly (if only one is chosen) or globally, if several (up to thirty-two datasets in v6.14) are chosen. The datasets are chosen from the drop-down windows on the left and will be found in the **Experiment** folder in which one had preprocessed the datasets under consideration. The dataset (*abr) files live at the “time” directory level i.e. the lowest level along with the scan files.

The best guesses for the loading concentrations in molar or mass units are entered in the boxes at the right end of the dataset line of boxes. These values should usually be allowed to float (you may think you know what they are but you really don't). One may either float or hold a common value for the ratio of $[B]_o/[A]_o$ or $c_{B,o}/c_{A,o}$. In most cases in which a dilution series is being analyzed, this ratio should be the same for all cells. To constrain the ratio to be the same value for all cells, left click on the text above the ratio boxes.

To switch between mass and molar concentrations click on the little red/blue squares to the bottom right of the column headers. To switch between molar or mass concentrations and molar or mass ratios, click on the yellow square on the bottom left of the column header. (Note: The yellow square at the bottom left of the first column header does nothing.)



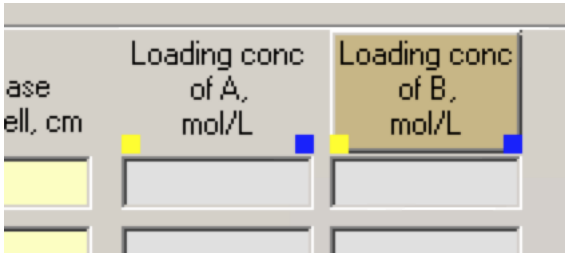


Figure 3-64. Switching between molar and mass concentrations

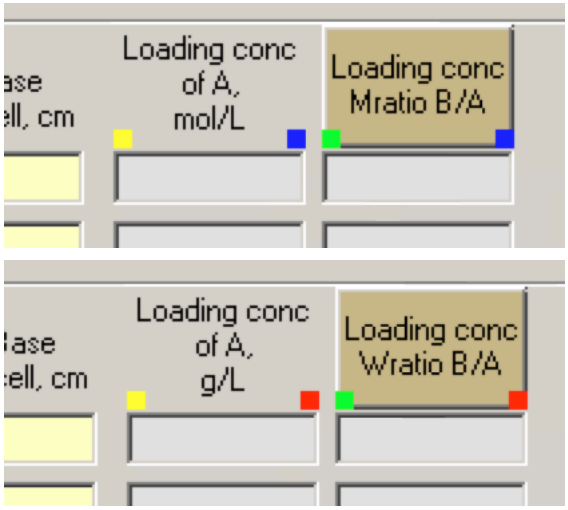
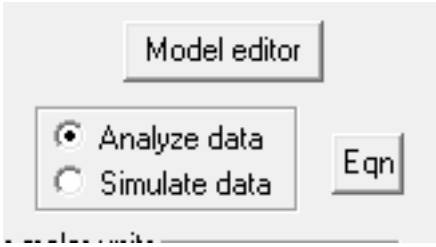


Figure 3-65. Switching between mole ratios and weight (i.e. mass) ratios.

3.4.1.7 The Equation Editor

Other relationships between parameters not established by the Model Editor can be established with the Equation Editor.

The equation editor window can be opened by clicking on the button labeled **"Eqn"**.



Relationships established by the **ModelEditor** will appear in green font. Relationships established by the **Control screen** will appear in blue font. Custom relationships that are entered by the **user** will appear in grey font and labeled as "Custom". (see below).

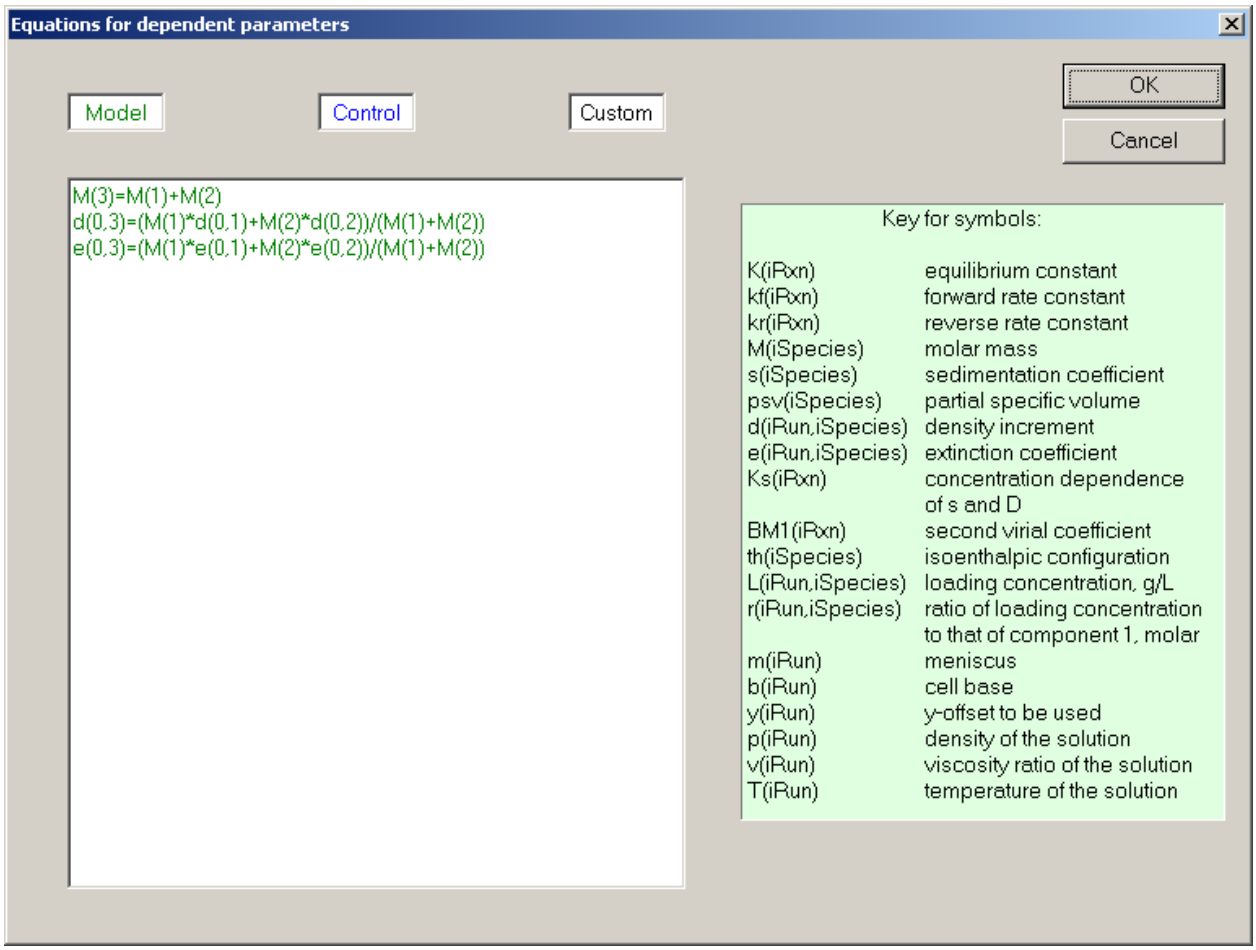


Figure 3-66. Equation Editor

In this example, for the model A+B=C, the equations are generated by the ModelEditor and show the relationships established by clicking on the boxed on the tab for species 3, "C". That is, the molar mass is the sum of those of species A and B. The density increment and the extinction coefficient of "C" are both expressed as mass weighted averages of those of species A and B.

Other relationships between the parameters can be entered manually. For example, consider the case of a 2 component, 2 species system comprising a non-interacting monomer and dimer. One would choose the model "2 comp 2 species" and then enter the relationship relating the molar mass of the two species.

Relationships in **Green** are established by the ModelEditor; those in **Blue** by the control screen and those in **Grey** by you. So, for the system A+B=AB we might have a case in which we had performed an experiment in which we had done a dilution series (4 cells) at a constant ratio of $[B]_o/[A]_o$ loading concentrations and combined those data with another experiment in which we had varied those ratios (2 more cells). Moreover, we've used both absorbance and interference optics for 2 of the cells. That gives us a total of 8 run files to combine in a global fit. Assuming we know the extinction coefficients accurately (both in units of A.U.-mL-mg and of fringes-mL-mg for a 12 mm optical path length), we can require that the loading concentrations in the cells in which both optics were used be the same by linking the cells on the control screen. Or we can use the equation editor to establish those constraints if the 2 cells in question were not adjacent to each other in the list of run files on the screen. If we wanted to link dataset 1 (absorbance data) to dataset 4 (interference data), because they are the same cell and therefore must have the same loading concentrations, we would enter L(4,1) = L(1,1) in the Equation editor.

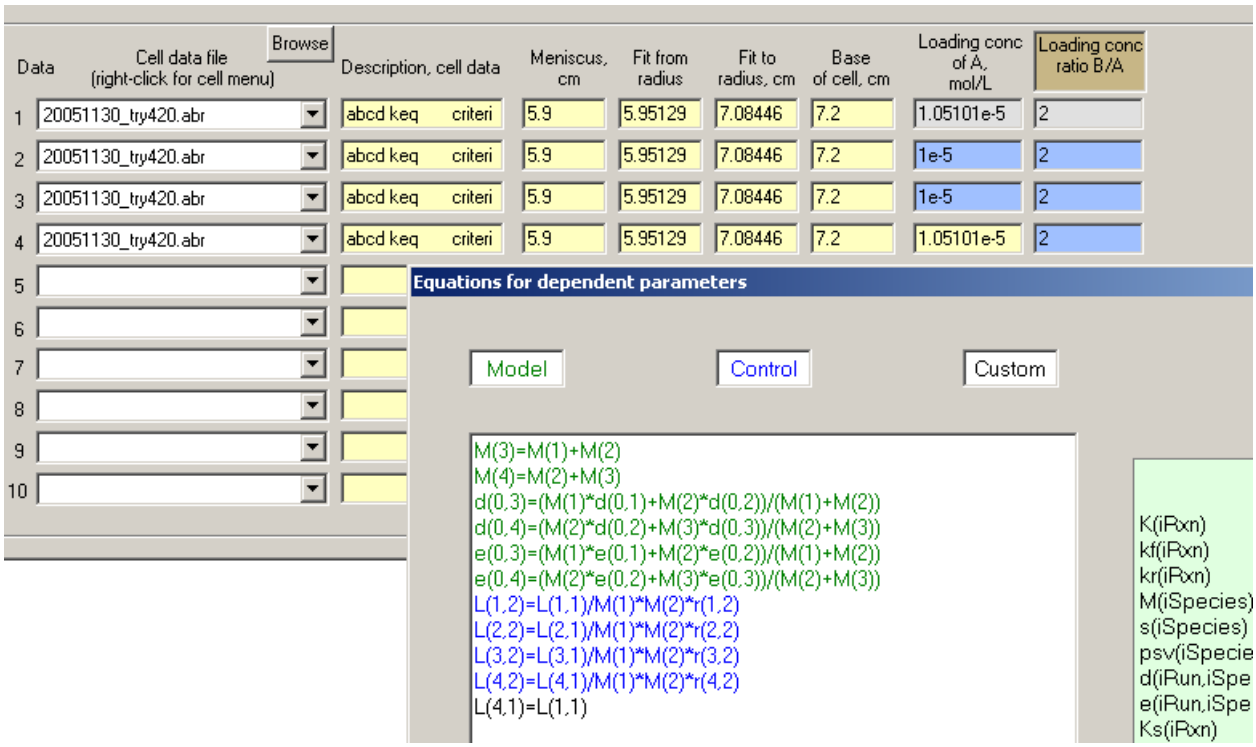


Figure 3-67. **Using the equation editor to link parameters in different cell datasets.**

3.4.1.8 Start the fitting:

To start the fitting after all the initial guesses and other constraints have been entered, the fitting is started by clicking the "**Store control file and start fit**" button.

3.4.1.9 Advanced Parameters Button

3.4.1.9.1 Error Estimation Control

The first set of advanced parameters is under the "Error estimates" tab and allows choice of either a Bootstrap with replacement, a Monte Carlo type of error analysis, or search for confidence limits by computing F-statistics. (F-statistics is selected in Figure 3-68.)

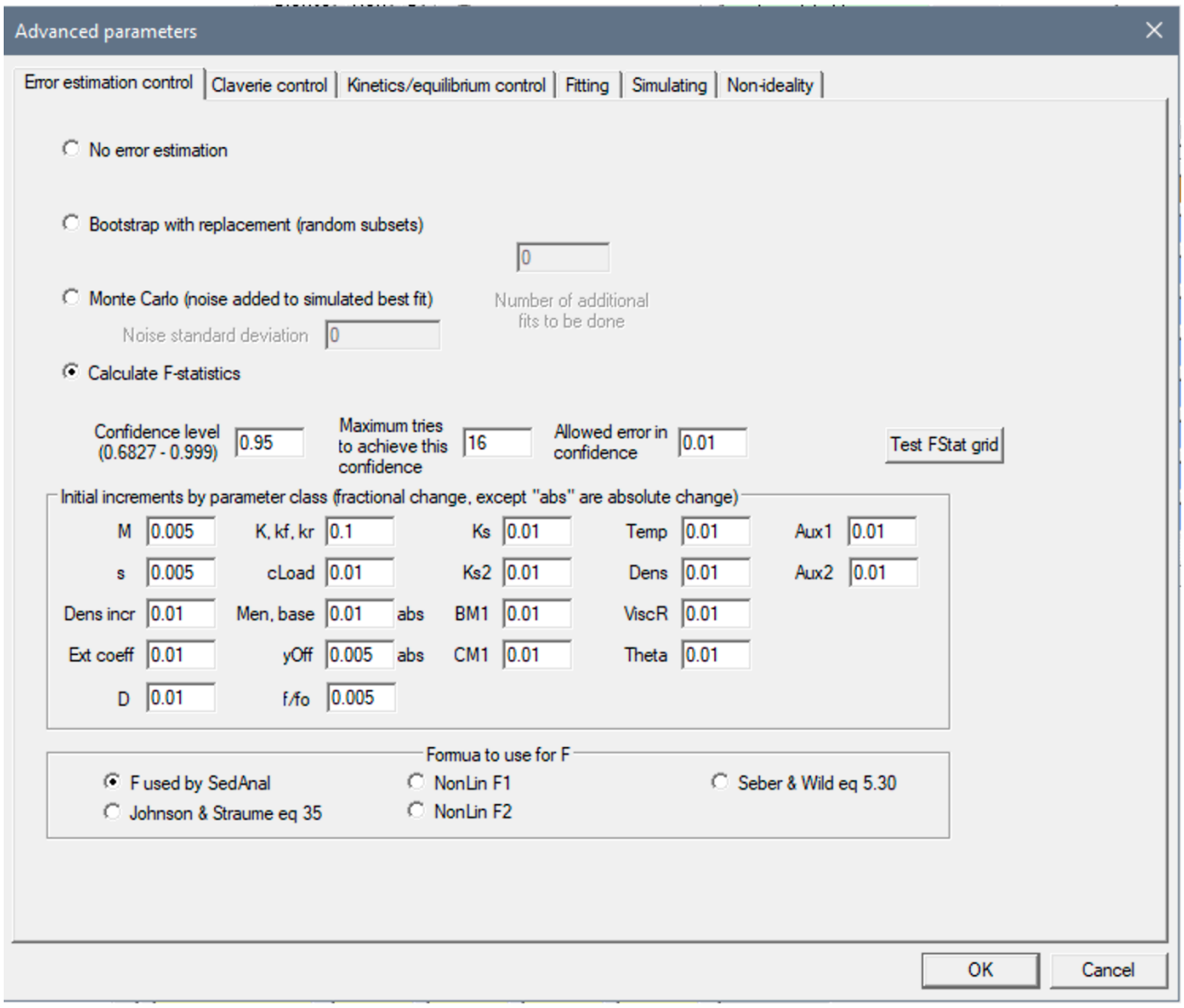


Figure 3-68**. Advanced window.**

3.4.1.9.2 Lamm Equation Solutions (Claverie control):

The second set of advanced parameters under the "Claverie control" tab allow adjustment of the performance of the finite element procedure for solving the Lamm equation and concerns the adjustable time increment used for each step and the distribution of grid points along the radial axis. The example below says essentially, if the

fractional change in concentration does not exceed 0.02 at any point in the cell over that last time interval, then increase the time interval by a factor of 1.02. A smaller rate of change can be set for the first few steps in the box labeled "Sedimentation time step initial fractional change".

The boxes at the bottom of this tab allow termination of the finite element algorithm in the event of numerical instability :viz:

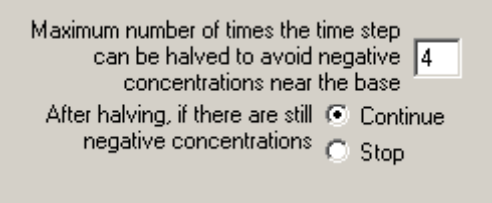


Figure 3-69 criterion for stopping F.E.M.

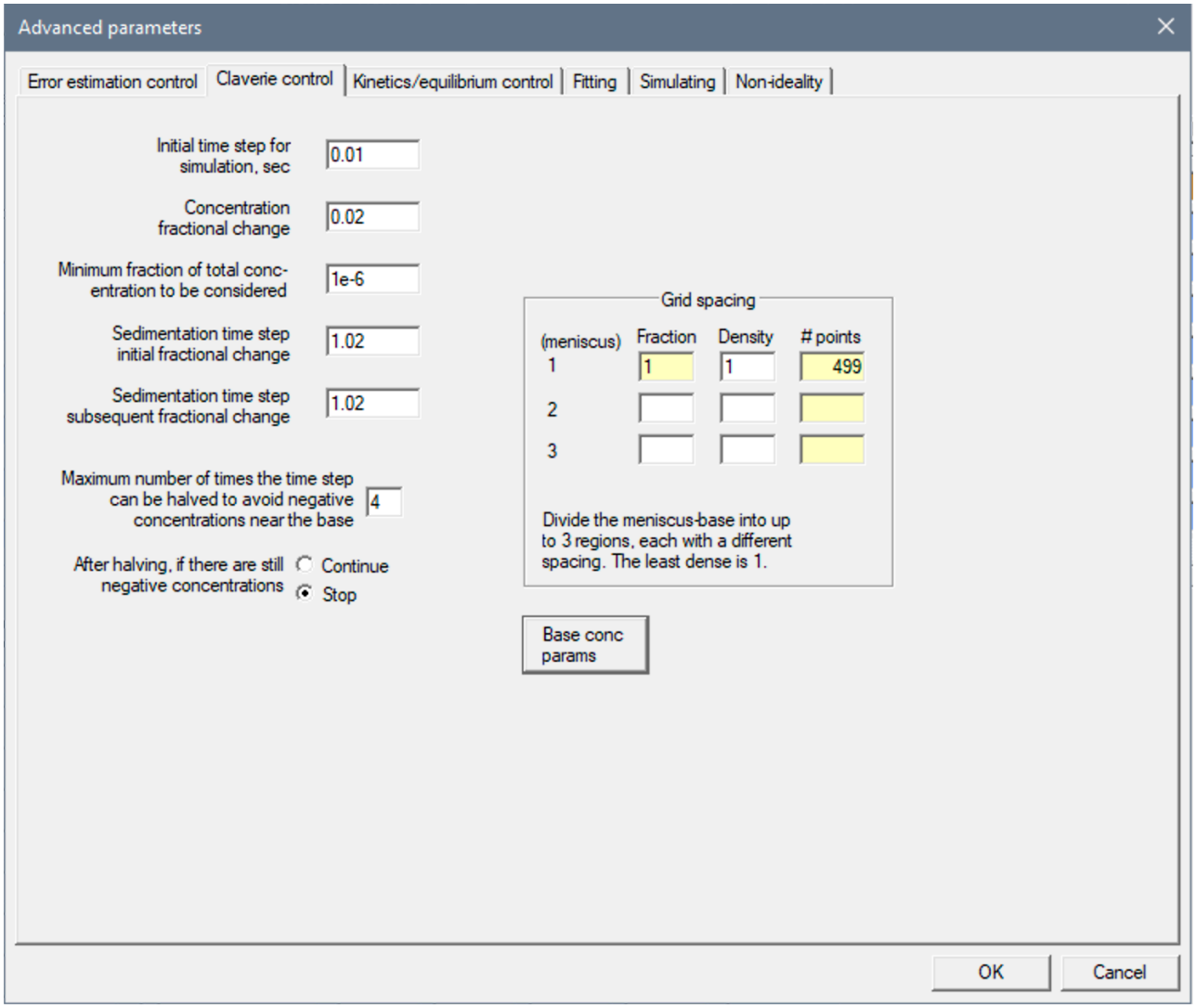


Figure 3-70**. Control of the finite element calculations.**

The grid spacing can be controlled as in the following example which shows the cell divided up into three regions spanning 0.1, 0.7 and 0.2 of the distance from the meniscus to the base. The point density and actual number of points used are shown in the designated boxes.

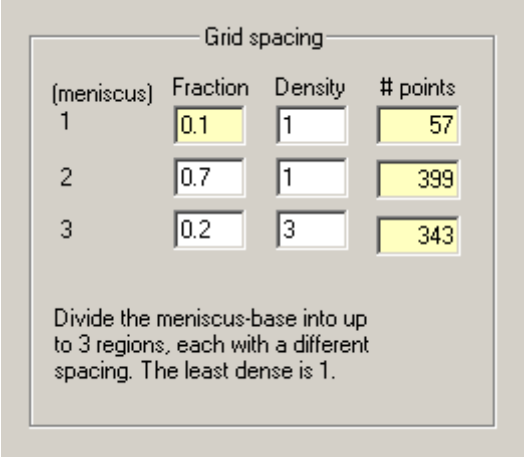
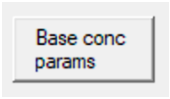


Figure 3-71 Specify grid spacing: three zones.

Limit concentrations near the base



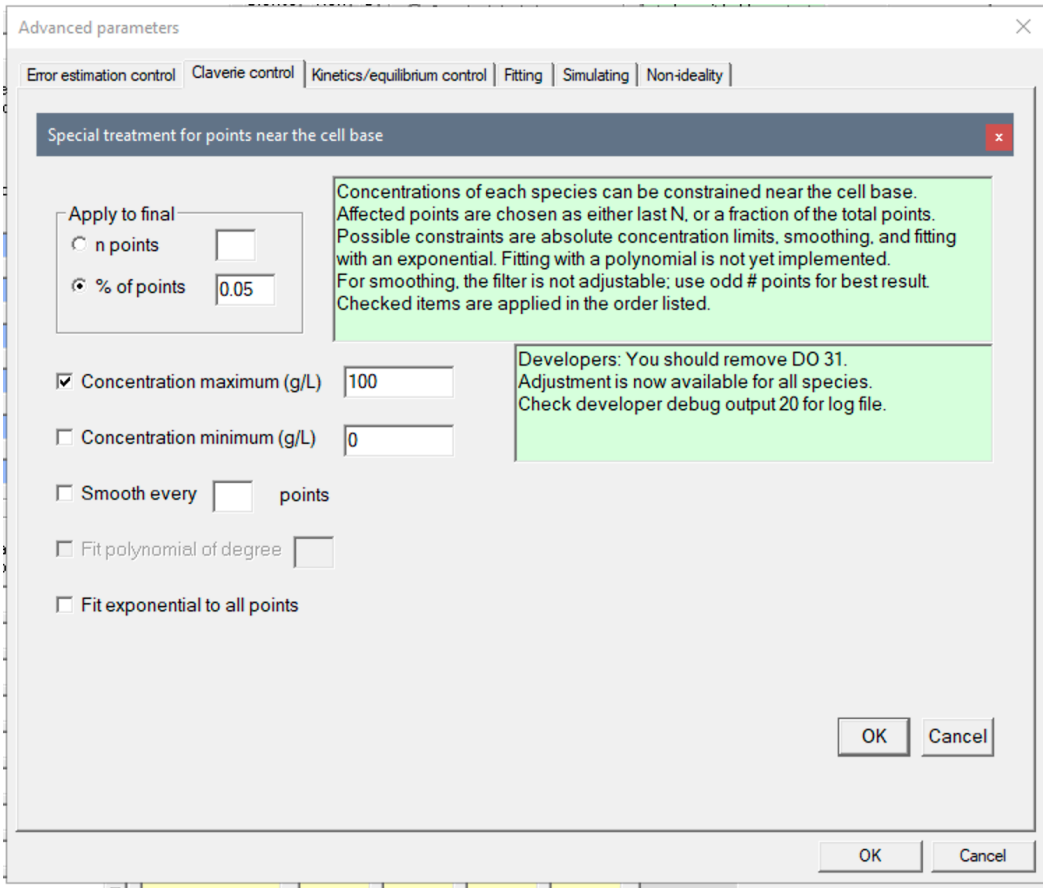


Figure 3-72 Limit concentrations near base

3.4.1.9.3 Kinetic integrator control

The third set of advanced parameters under the "Kinetic integrator control" tab allows one to adjust the performance of the ODE (Ordinary Differential Equation) solver. The current version of SEDANAL uses either the Bulirsch-Stoer algorithm (with Richardson rational polynomial extrapolation) or the Euler extrapolation method to solve the differential equations describing the kinetics. For those cases in which instantaneous equilibrium can be assumed, the equilibrium equations may be solved directly using the Newton-Raphson

successive approximation iterative method. One method, kinetic or equilibrium solution, could be faster than the other depending on the model.

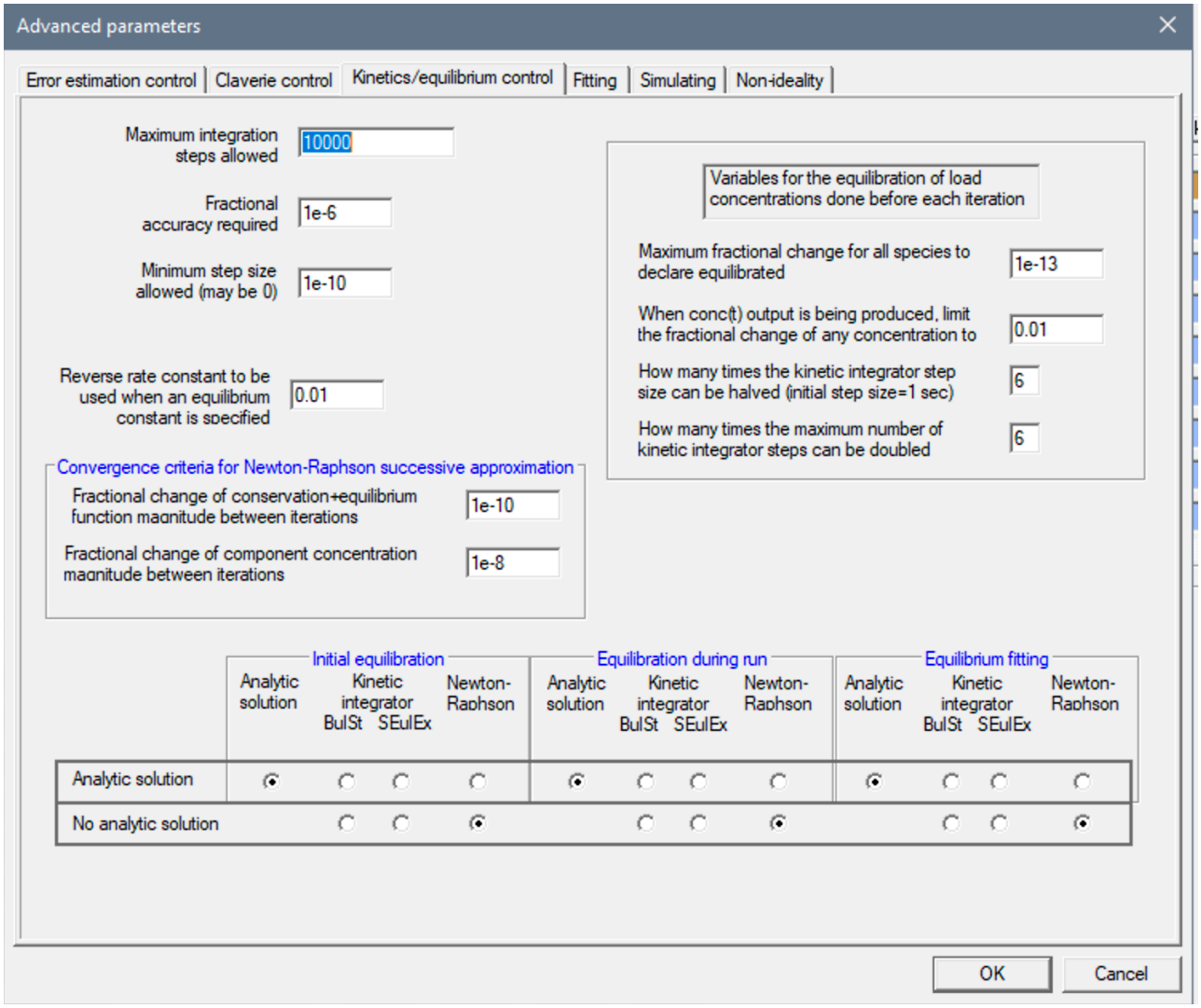


Figure 3-73**. Control of the kinetic integrator.**

After each time step of the Lamm equation solution, the updated concentrations of all the species are recomputed according to the kinetic rate constants. If the kinetic methods are too slow and you are sure that you r system is not kinetically limited, use the Newton-Raphson method; it may run faster. You will have to determine this on a case-by-case basis. For fitting equilibrium runs, use only Newton-Raphson.

3.4.1.9.4 Fitting Tab

First choose the fitting method, either simplex or Levenberg-Marquardt minimization. For simplex, the mode for setting up the initial simplex is controlled by selecting the forth tab “Fitting”. In the upper right-hand area, The top button is to vary only the diagonal elements of the simplex array in the positive direction. The middle selection is experimental – don’t use it. The last one (bottom button) varies all the elements by small positive and negative amounts. It seems to work best and is recommended. The convergence criteria are entered in the bottom set of

windows. Whether the fit is stopped, or perturbed and restarted is indicated in the "Perturb fit and re-do" section of this screen. Other settings here affect the display.

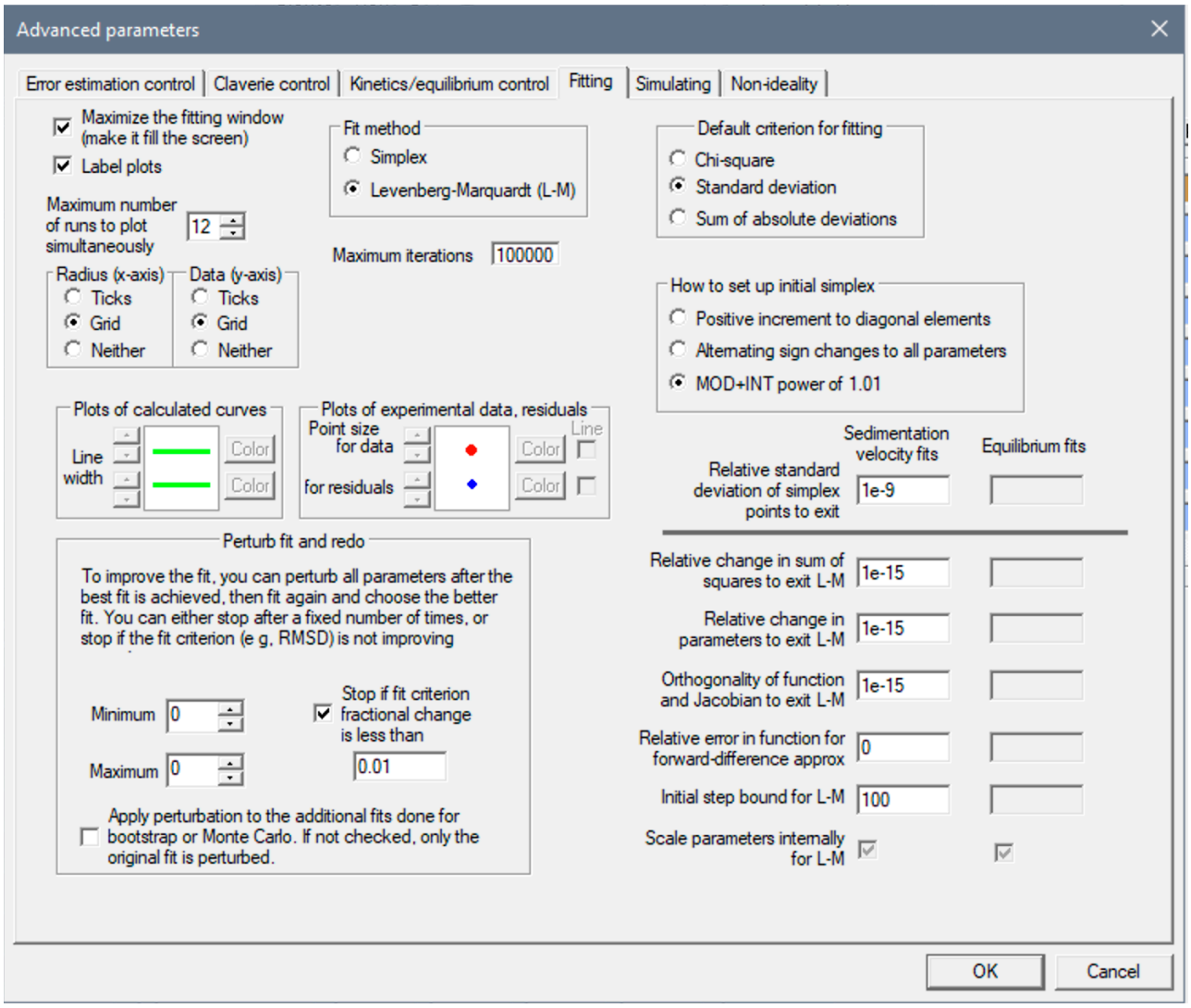
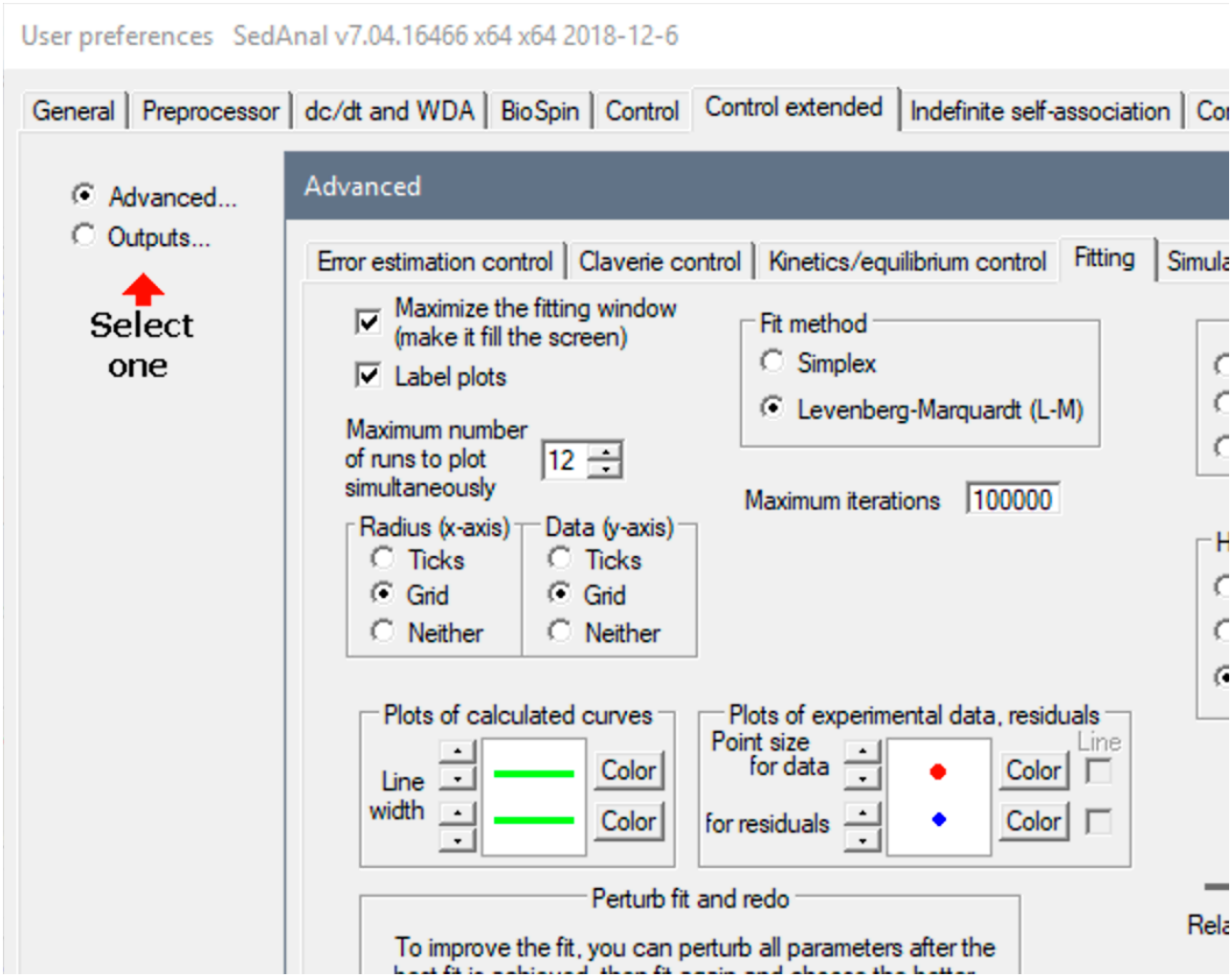


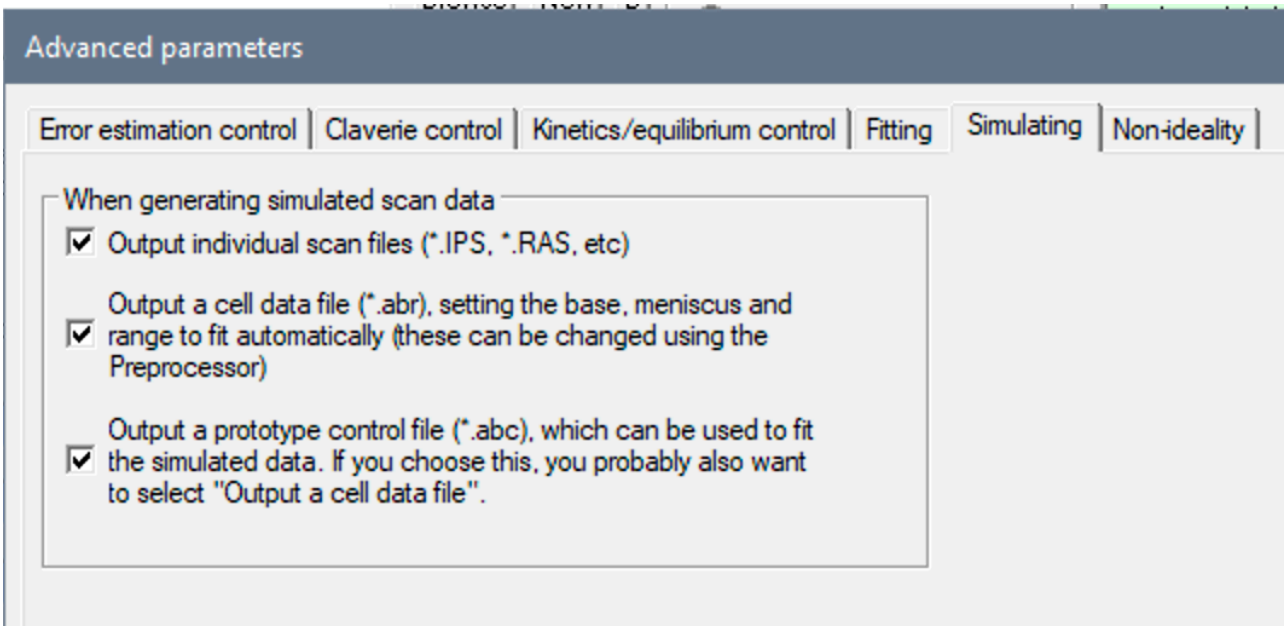
Figure 3-74. **Choosing fitting parameters.**

3.4.1.9.5 Plot color and symbol size

Plot color and symbol size are set in the Main Menu > Preferences under the control Extended tab:



3.4.1.9.6 Select output for Simulations



3.4.1.9.7 Set limits on s and D under non-ideal conditions.

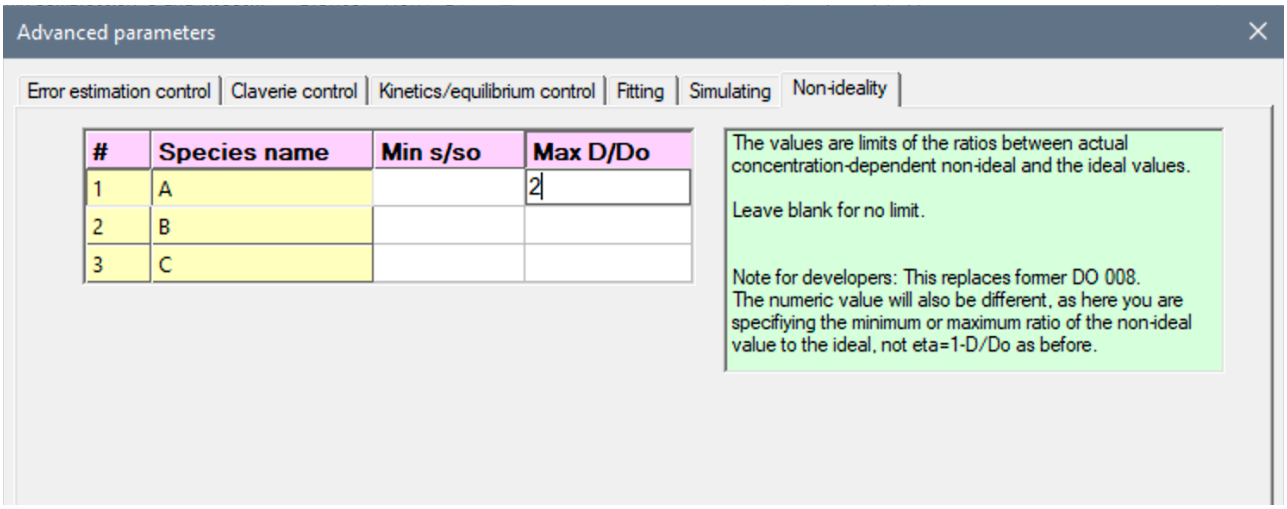


Figure 3-75. Set limits on D and s for non-ideal systems

Leave s/so blank and set D/Do to 2.0. If needed, s/so should be set to a fraction like 0.5.

3.4.1.10 Concentration dependence of s and D without cross terms.

Hydrodynamic concentration dependence of both s and D through the frictional coefficient is expressed through the coefficient k_s defined as follows: $f=f_{(c=0)}(1+k_s c)$.

Thermodynamic concentration dependence of D is expressed through the second viral coefficient, BM_1, as follows. $D=D_{(c=0)}(1+2BM_1\ c)/(1+k_s c)$ One may either fit for these parameters or hold them constant at preset, "known" values.

Therefore, the hydrodynamic concentration dependence for s is given by

$$s(c) = s_o \left[\frac{1}{1+k_s c}\right]$$

while the total concentration dependence for D is given by

$$D(c) = D_o \left[\frac{1+2BMc}{1+k_s c}\right]$$

This approach has been superseded by Cross term fitting using the Ks and BM1 matrices (cf **Cross Term Non-ideality**:)

3.4.1.11 Pressure dependence of density and viscosity:

3.4.1.11.1 Compressibility

When dealing with compressible solutions, one can enter the isothermal compressibility value in the **Experiment information** window in the **preprocessor**: [the units of compressibility are cm^2/dyne (a.k.a. Ba^{-1} = $bayre^{-1}$).]

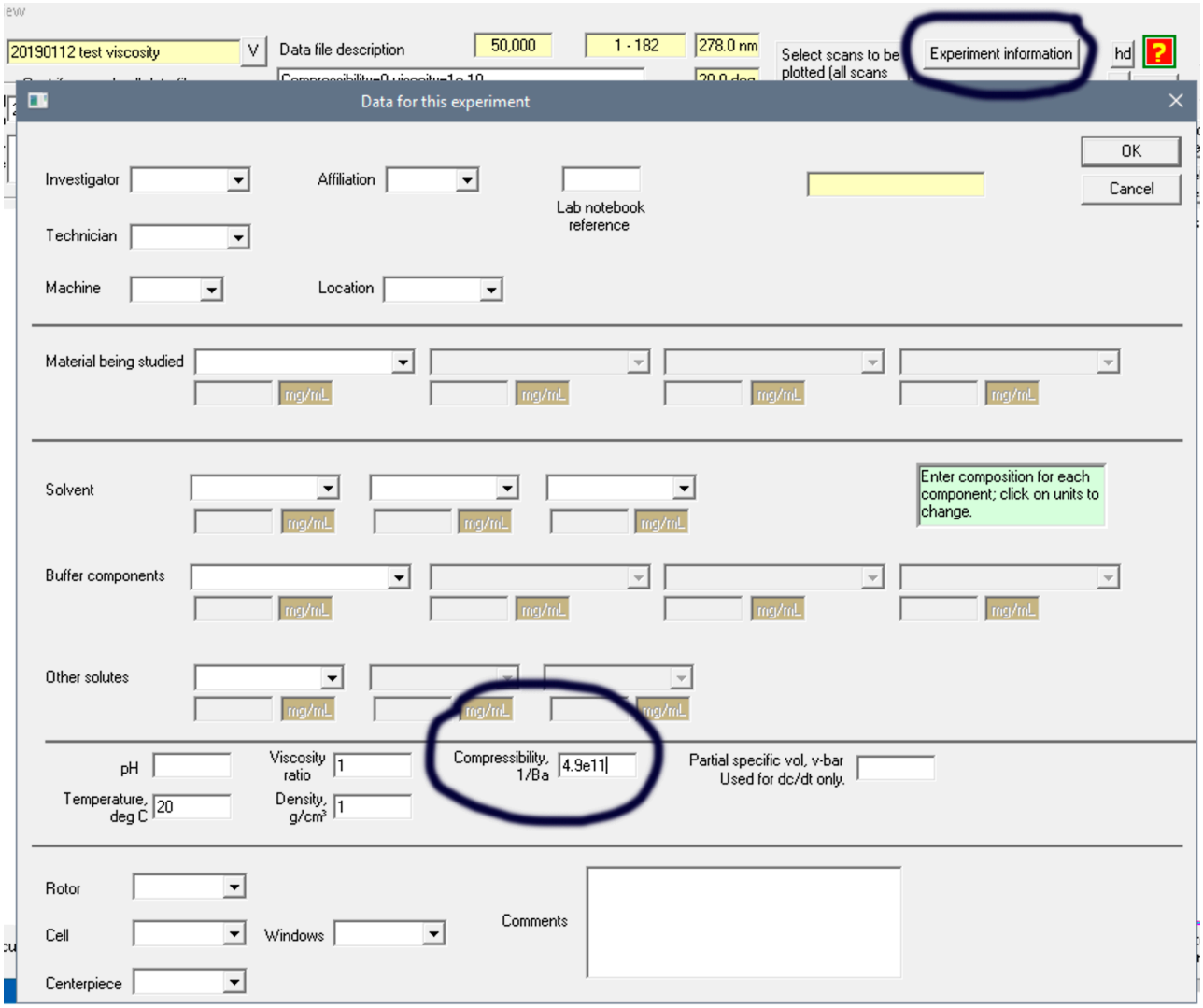


Figure 3-76 Setting Compressibility for each cell.

Load the abr file and click on Experiment information and enter the compressibility as indicated.

We have generally followed the equations described by: Schuck,P., "A model for sedimentation in inhomogeneous media. II. Compressibility of aqueous and organic solvents Bioph. Chem. 108 (2004) 201–214

3.4.1.11.2 Pressure dependence of viscosity

If the pressure dependence of solvent viscosity is to be taken into account. The values for the pressure-viscosity coefficient can be entered as a developer option: select option 024 and enter the viscosity-pressure coefficient under Parameter 1. (Version 7.05: this will be moved to experiment info in later versions)

Figure 3-77 Pressure dependent viscosity

3.4.1.12 Selecting datasets to be fitted:

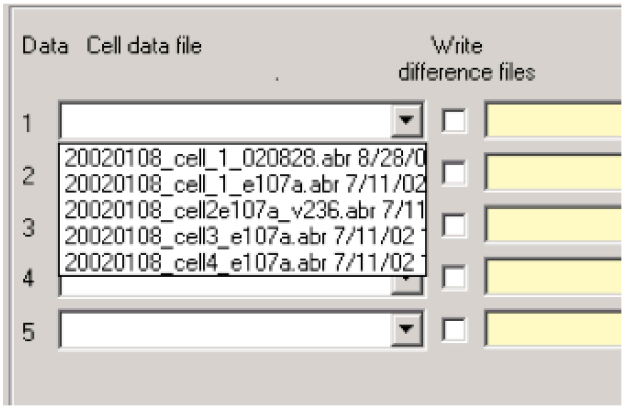


Figure 3-78. Preprocessed datasets.

Once the experiment directory has been chosen, all the available preprocessed run files (*.abr) will be visible in the dropdown windows.

When the cells are chosen, one must enter guesses for the loading concentration of each component.

Selected scans within a cell may be fitted. The default is all scans; to change this, right-click on the run file name (lower left), a small window appears that allows you to either choose the scans to fit or to enter a weighting factor for that dataset.

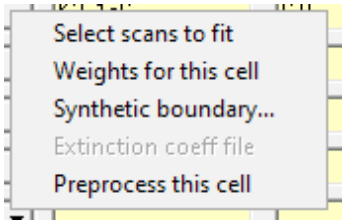


Figure 3-79 Specifying scans to fit

If "Select scans to fit" is chosen, a window similar to the one for selecting scans to be plotted in the preprocessor will appear. You may select the scans to be used for fitting in that window. The result is written into the control file. There is an option in the lower right of the scan-select dialog box to use the same scans for fitting that had been chosen for plotting in the preprocessor (Figure 3-80).

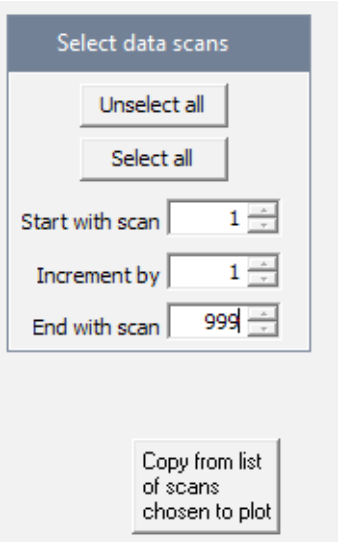


Figure 3-80 select scans to be fitting or copy those used in the preprocessor

3.4.1.12.1 Removing a dataset from the fit.

To remove dataset from the fit, highlight the name in the “Cell data file” window (requires two left clicks) and press the “Delete” (not the “Backspace”) key.

3.4.1.12.2 Weighting Factors

In the Fitter the use of weighting factors and user settable values for weighting factors can be set for each cell data file by right-clicking on the cell data file name on the control screen. To enter weighting factors for a particular dataset, right-click on the cell data file name

(the *.abr file) and select "weights for this cell"

(NOTE: THIS HAS BEEN UPDATED TO USE STANDARD DEVIATIONS, instead of inverse variance, as of version 6.92)

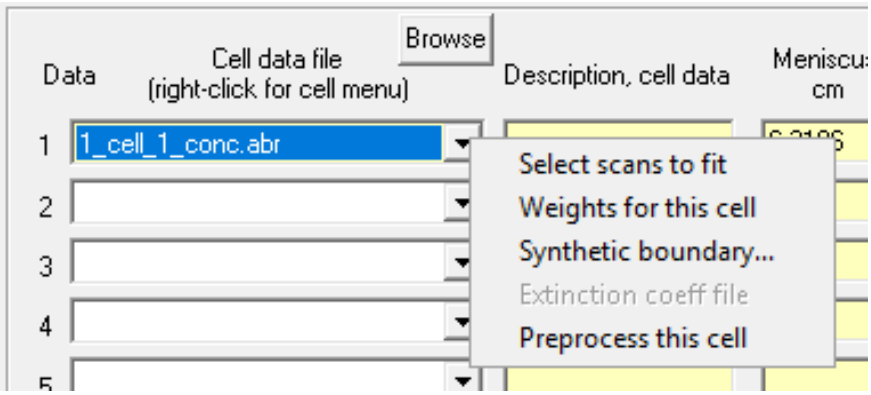


Figure 3-81. Selecting "Weights for this cell” brings up the following window.

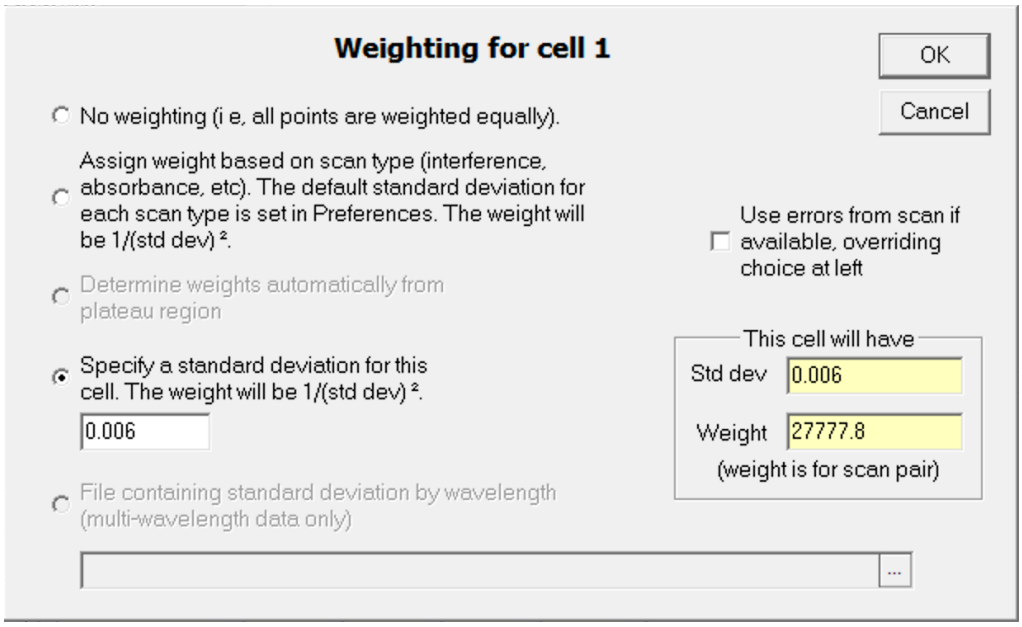
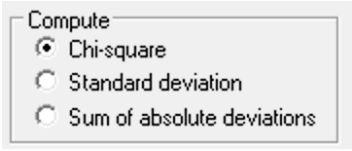


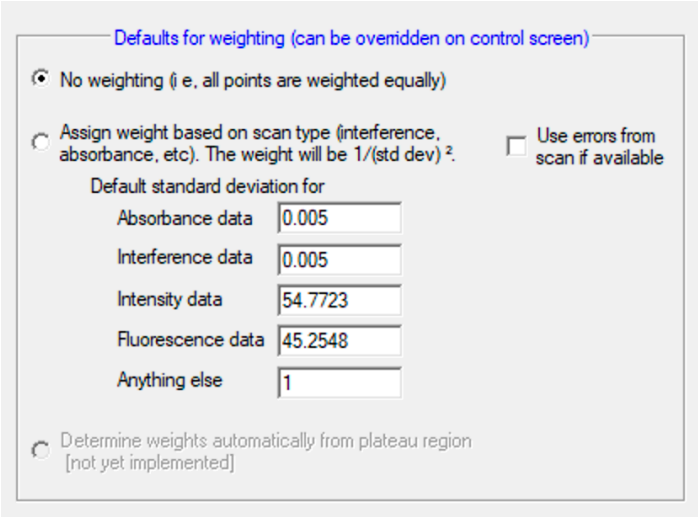
Figure 3-82 Select weighting for each cell

If datasets from different optical systems are to be combined in a global fit, they must be weighted according to the magnitude of the data and the noise on the data. If the **inverse of the variance** ($1/\sigma^2$) of the data is used as the weighting factor for the squared residuals, then data from different datasets can be combined. In the example above, the value of 27777.8 is the square of the inverse of 0.006, the typical order of the standard deviation of either interference (0.004) or absorbance data (0.006) from the XL-A/I. **When weighting factors derived this way are used, the reduced chi-squared values instead of sums of squares of the residuals are minimized. The results are equivalent: in both cases the procedure is called Maximum Likelihood Estimation (MLE).**

NOTE:: **SEDANAL displays the reduced chi-squared values on the fitting screen instead of the rms deviation when inverse variances are used as weighting factors and "Chi-square" is selected at the top of control screen.**



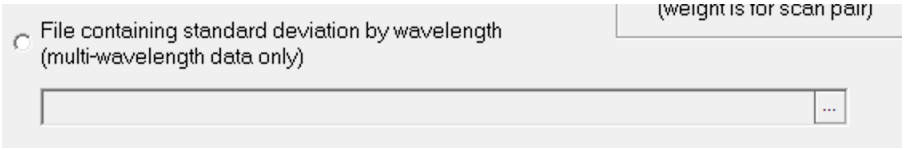
Default values can be set in the Preferences under the Control tab:



These can be changed on the control screen for each cell.

3.4.1.13 Multi-Wavelength Weighting Factors by Wavelength

When a multiwavelength dataset has been loaded, the following weight option becomes available:



allowing the user to load a file containing the standard deviations of the data as a function of wavelength.

3.4.1.13.1 Standard deviation by wavelength file format

The standard deviation files may have any extension (e g, .txt or .csv), but the contents will be interpreted as ASCII text. SEDANAL looks for lines like

wavelength stdDev

The *stdDev* is the standard deviation to be used for all points at that wavelength, in the same units as the scans contained in the corresponding cell data (.abr) file. The standard deviation is used to weight scan points at the given wavelength.

The format is identical to that for extinction coefficient spectrum files.

The *wavelength* may contain a decimal point, so either `253` or `253.0` is valid, and has the same meaning. The unit for wavelength is nanometer (nm), and the values are rounded to the nearest 0.1 nm.

Lines beginning with `//` are comments, and will be ignored.

The wavelengths may appear in any order, and need not be consecutive. Within the range of wavelengths contained in the standard deviation file, the standard deviation will be linearly interpolated to the scan data wavelength. Any scan data with wavelength outside that range will be ignored during fitting (weight = 0). Only the first 2500 wavelengths will be included. Wavelengths should not be < 200 nm.

3.4.1.14 Selecting parameters for the finite element solutions to the Lamm equation.

One must select the grid density, i.e. the number of points between the meniscus and the base, to be used for the fitting. The data are interpolated onto this grid and the finite element solutions to the Lamm equation are generated on this grid. The time interval is the initial time increment used by the finite element procedure. This initial time increment can be adjusted according to the parameters enter under the “Advanced” menu.

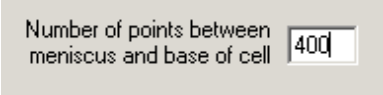


Figure 3-83 Select number of points for F.E.M for fitting

For initial fits to a new system, it is recommended to choose 200 points between the meniscus and the base. This will make the exploratory fits relatively fast. The number of points should be increased to 400 or 800 later to refine the fits for the most accurate estimation of fitted parameters. Higher numbers of points may be necessary if the sample generates steep gradients. The maximum is 10,000 points

3.4.1.15 Extinction coefficients: Global fitting with multiple optical systems -

If several optical systems and/or wavelengths have been used, different mass extinction coefficients can be used for each cell. **(Note: the extinction coefficients to be entered must correspond to a 12 mm path length: i.e. one must multiply the usual (mg/mL for 1 cm) UV extinction coefficients by 1.2). Or if 1 or 3 mm centerpieces were used, the usual 1 cm value must be multiplied by either 0.1 or 0.3, respectively, to correspond to actual path length used.**

If all cells were run with interference optics, then the extinction coefficients would be entered by choosing the ”ALL CELLS” option which is the default:

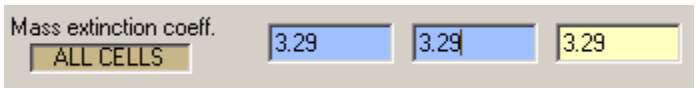


Figure 3-84 ALL CELLS

Left-clicking on the text “ALL CELLS” will allow one to enter different extinction coefficients for species for each cell.

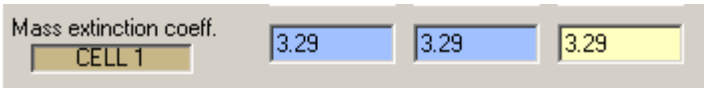


Figure 3-85. Left-clicking on the text “ALL CELLS”

Each click advances to the next cell. If interference optics have been used for cell 1, one would input the number of fringes per 1 mg/mL:

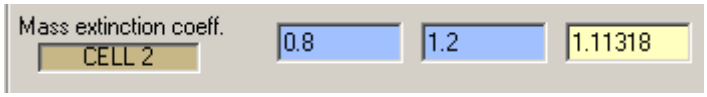


Figure 3-86. Set extinction coefficients for each cell

NOTE: Starting with version 7.40, SEDANAL allows a matrix of extinction coefficients to be filled in for each species in each cell. For example, (cf Figure 3-87) for the system $2A=A_2$; $2A_2=A_4$ and two cells, with, say, a 12mm and a 3 mm centerpiece, respectively, we would have the following matrix that would allow us to either fit or hold the extinction coefficients. (In this case, for example, we have assumed that the extinction coefficients do not change upon association.)

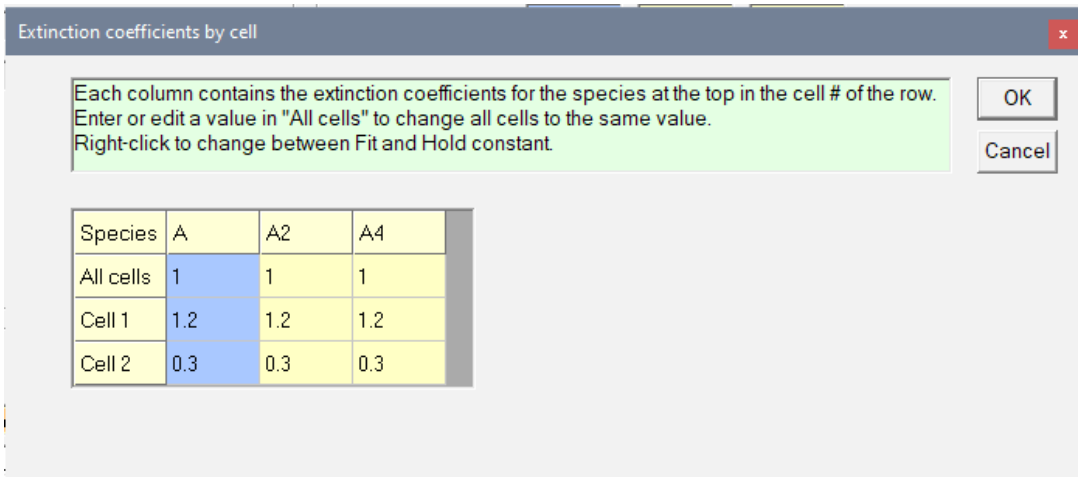


Figure 3-87 Extinction Coefficient Matrix

If absorbance at 280 nm had been used, one would click again on the text and enter the mass extinction coefficients of each species at 280 nm: where the last box is computed as a weight average based, in this case, on the model A+B=C. If the data were taken with different optical systems on the same cell, then the loading concentrations for those data will be the same. SEDANAL allows one to constrain the loading concentration to be identical for those datasets.

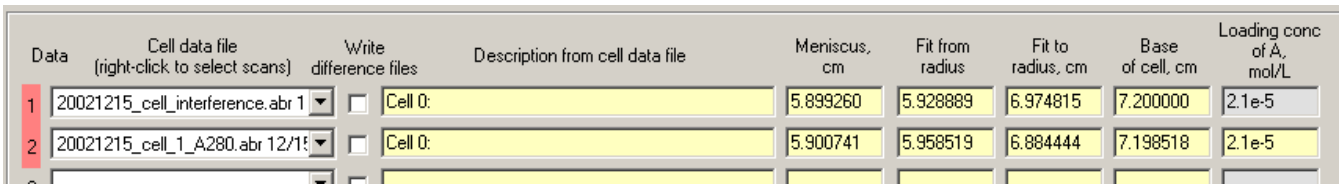


Figure 3-88 data from "cell data files" 1 and 2

For example, if the data from "cell data files" 1 and 2 were taken from the same cell, their loading concentration will be identical. One may click on the Number "2" on the far left of the row to "slave" that dataset to dataset # "1".

If cell "3" had been taken at another wavelength on the same cell, then cell data files "2" and "3" would both be "slaved" to cell data file "1" by clicking on the number "3" at the left (not shown).

NOTE: Many of the relationships discussed in the following section can now be established in the Equation Editor which is especially useful for cells not adjacent in the list shown in Figure 3-89 **and for more complicated relationships between cells.**

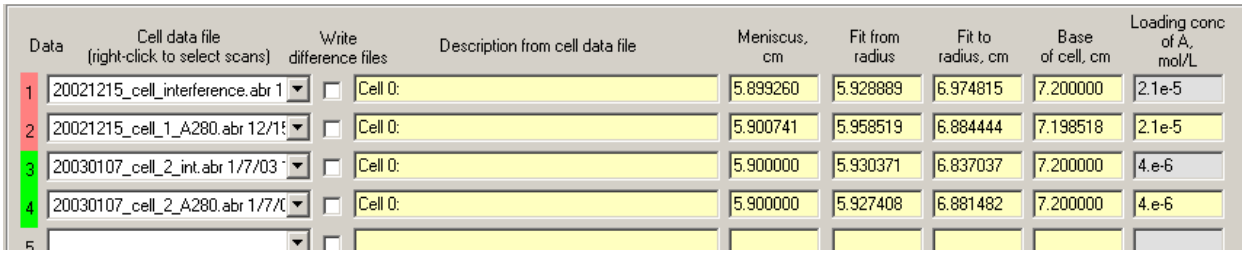


Figure 3-89. Linking cells that have the same sample but different optical system.

For example, if cell data files "1" and "2" are from one cell and cell data files "3" and "4", then click on the row for cell data file "2" to slave it to cell data file "1". Then click on the row for cell data file "4" to slave it to cell data file "3". Cell dataset that are slaved are highlighted in the same color.

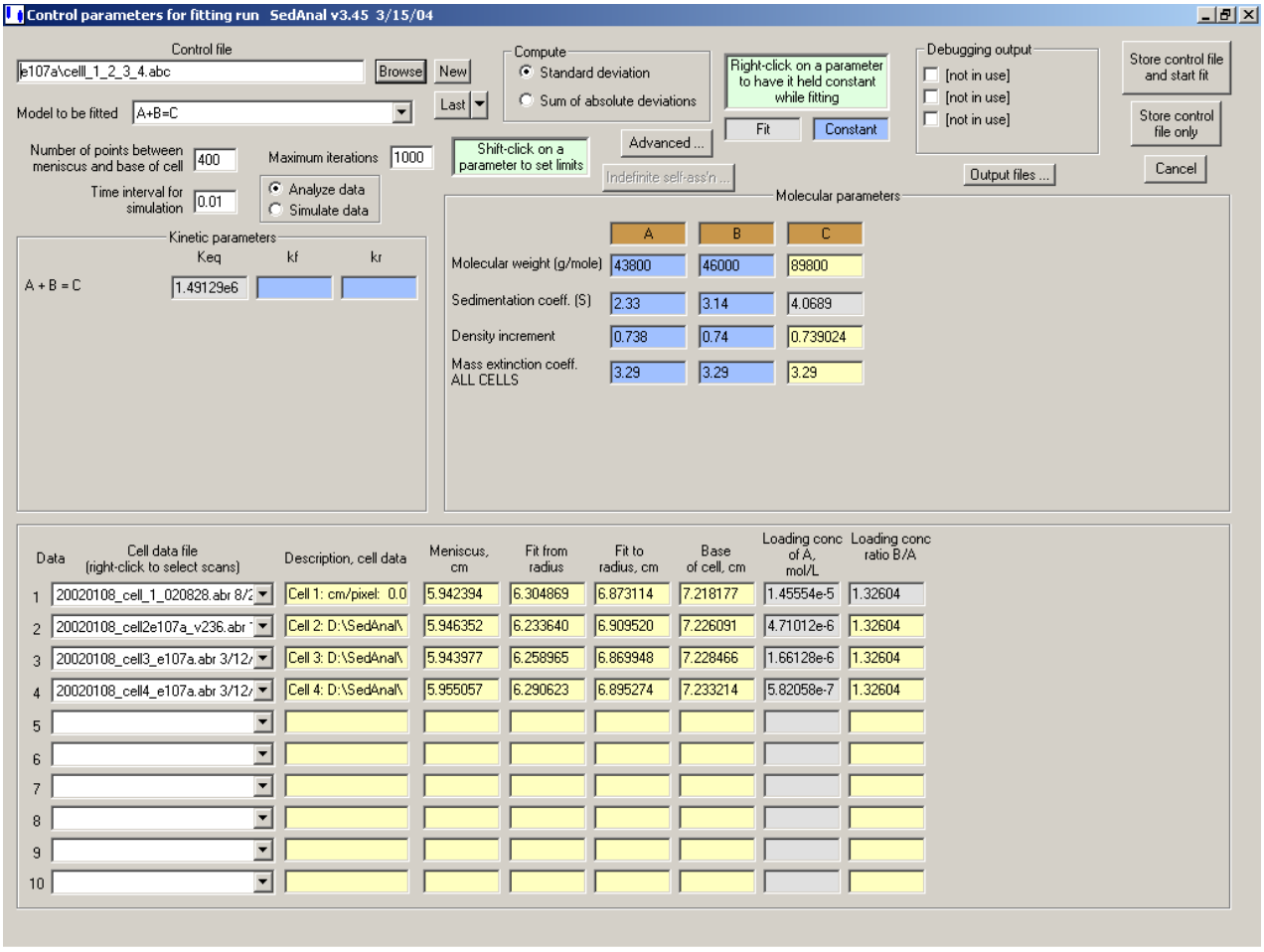


Figure 3-90 The main control screen for the fitting preprocessed data with all parameters filled in.

When all the parameters have been filled in, the screen will look as shown above (Figure 3-90).

NOTE: The loading concentrations can be expressed in moles per liter. The loading concentration of component 2 can be entered in terms of the molar ratio of B to A: $[B]_o/[A]_o$. For the 3 component system, the last column will be the ratio of $[C]_o/[A]_o$. This will allow one to specify that the ratio be fit as a global value for all runs, and is used in cases in which the runs are part of a dilution series for which these ratios might not be known but which are usually expected to be the same for all cells.

3.4.1.16 Local vs. global parameters - changing

To change the ratio from a local to a global parameter left-click on the text “Loading conc ratio B/A” or “Loading conc ratio C/A” above the column of boxes. The first click will fill all the boxes with the value entered into the top box and propagate that value to the other cells and change to background color to yellow to indicate that these are not being fit separately.

All cells will use the same ratio in the fitting process.

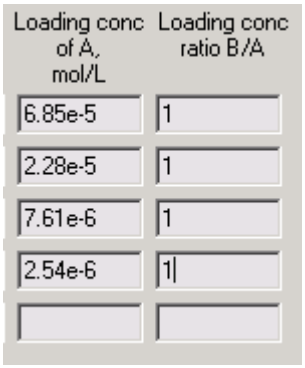


Figure 3-91 The loading concentration ratio is entered as 1.0 in this example.

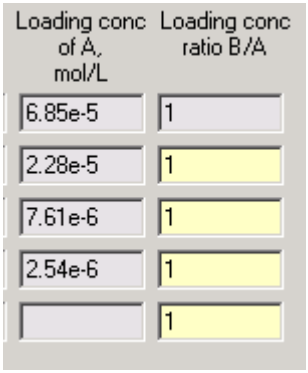


Figure 3-92. A single left-click on the text "Loading conc ratio B/A" propagates the top 1 into the boxes below with a yellow background to indicate that these values are derived from cell #1 and will be kept equal to and vary with the value for cell #1 as a single global parameter.

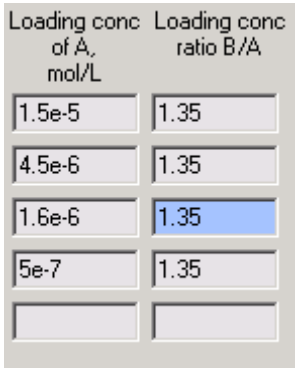


Figure 3-93. A second click on the text "Loading conc ratio B/A" turns the backgrounds back to gray to indicate that these values will be now be allowed to float independently for all cells:

Figure 3-94. Right clicking on one of the gray boxes will turn its background blue and now its value would be held constant during the fit, if that were desired:

3.4.1.17 Indefinite Self-association

3.4.1.17.1 Isodesmic case:

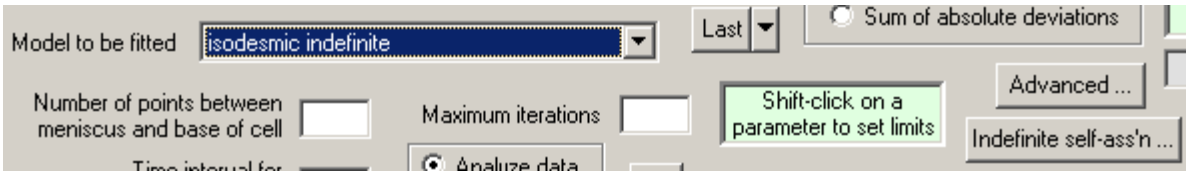
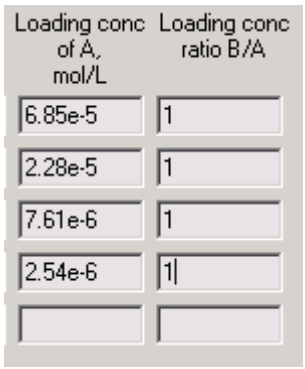


Figure 3-95 Isodesmic model selection.

On the control screen, if an isodesmic model has been chosen (Figure 3-95), a button labeled “Indefinite self-ass’n…” will become active. Click the “Indefinite self-ass’n…” button to enter the coefficients for a polynomial representing the relationship between the sedimentation coefficient of the oligomers and the degree of polymerization (Figure 3-96).



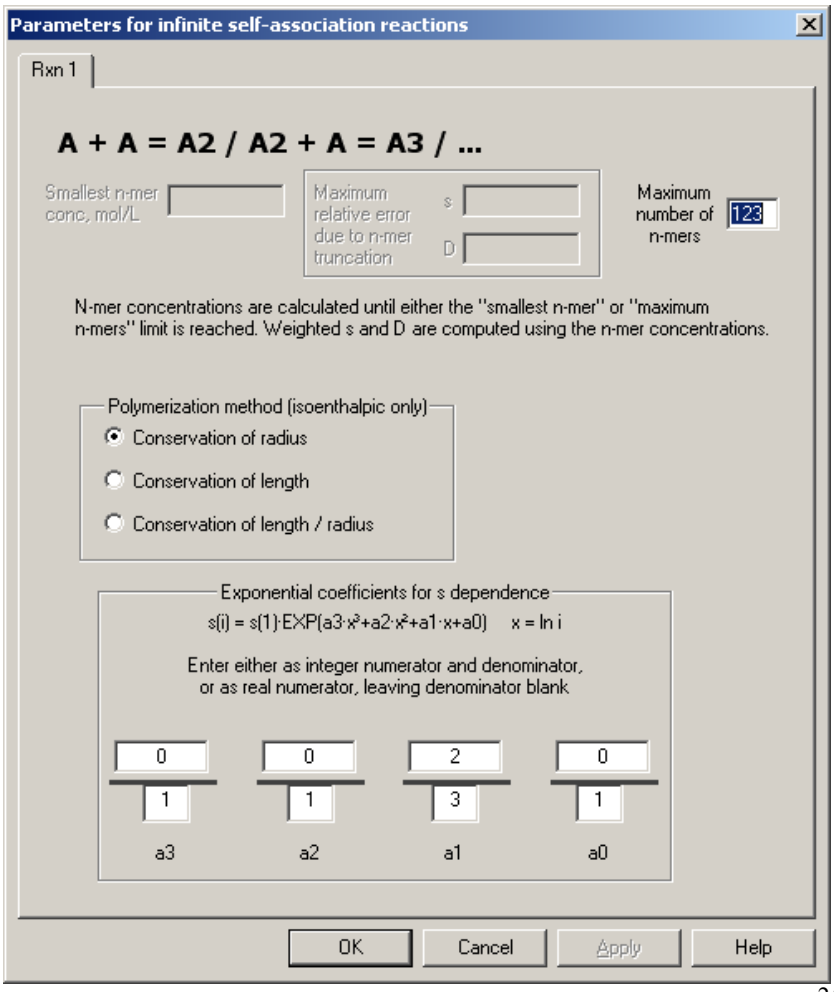


Figure 3-96. This example is using the s(i-mer) = s(monomer)* $(i)^{2/3}$.

Other coefficients, for end-to-end polymerization for example, can be entered from bead modeling or other theoretical considerations. The coefficients can be entered as rationals in the numerator box or as fractions by specifying both numerator and denominator.

The coefficients, a0, a1, etc… represent the coefficients of the following polynomial

ln(s(i-mer)) = a0 + a1*ln(i) + a2*(ln(i))2 + a3*(ln(i))**3.**

The coefficients are obtained by fitting a polynomial to the values of ln(s(i)) as a function of ln(i). Usually **a0** will be set to **ln(s(monomer))**. The values of **s(i-mer)** can be obtained from the analysis of bead modeling, for example.

3.4.1.17.2 Isoenthalpic indefinite self-association:

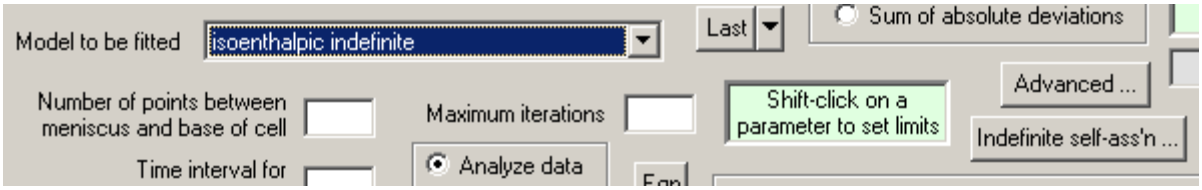


Figure 3-97 Isoenthalpic indefinite self-association

On the control screen (Figure 3-97), if an isoenthalpic model has been chosen, a button labeled “Indefinite self-ass’n…” will become active:

Click the “Indefinite self-ass’n” button to enter the coefficients for a polynomial representing the relationship between the sedimentation coefficient of the oligomers and the degree of polymerization. Then select the polymerization method according to how the entropy terms will be handled. [For details, see Ronald Chatelier, “Indefinite isoenthalpic self-association of solute molecules”, Biophysical Chemistry, 28, 121-128 (1987)]

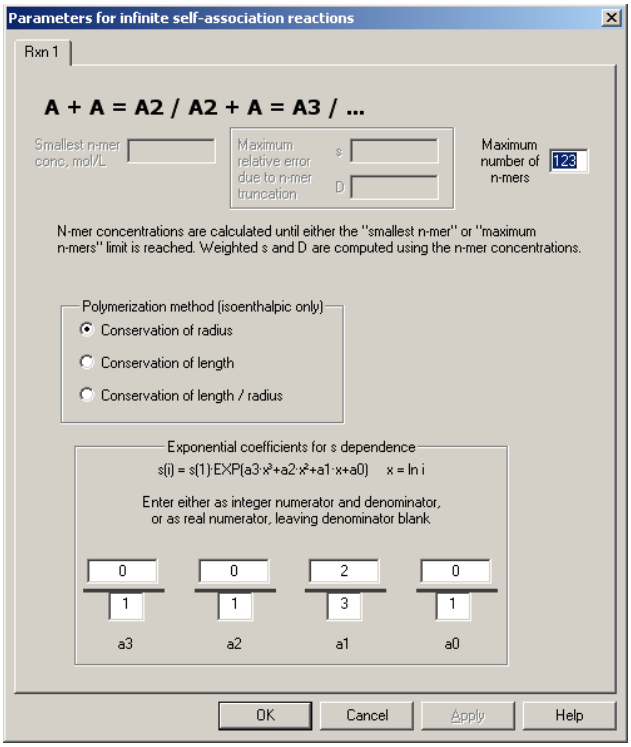


Figure 3-98. **Entering parameters for indefinite self-association**

3.4.1.18 Constraining the range of parameter values during fitting.

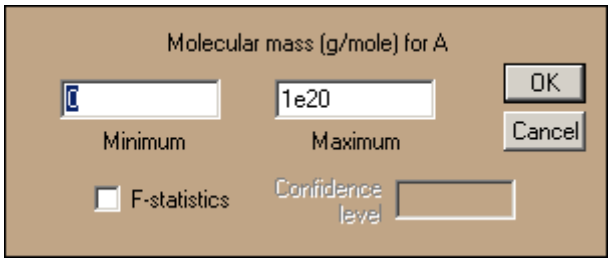


Figure 3-99. The range of allowable values

The range of allowable values for a parameter may be set by **shift-left-clicking** (Figure 3-99) on the value in the control screen (Limits set by the ModelEditor can be changed here.):

3.4.2 Exiting the Control Screen to start fitting

Figure 3-100. Start the fit

After all the appropriate boxes are filled in, the fit can be started by clicking the button labeled "Store control file and start fit" or if you just want to store this control file and maybe open up another one, click the button labeled (you guessed it) "Store control file only".

3.4.2.1 Fitting Screen

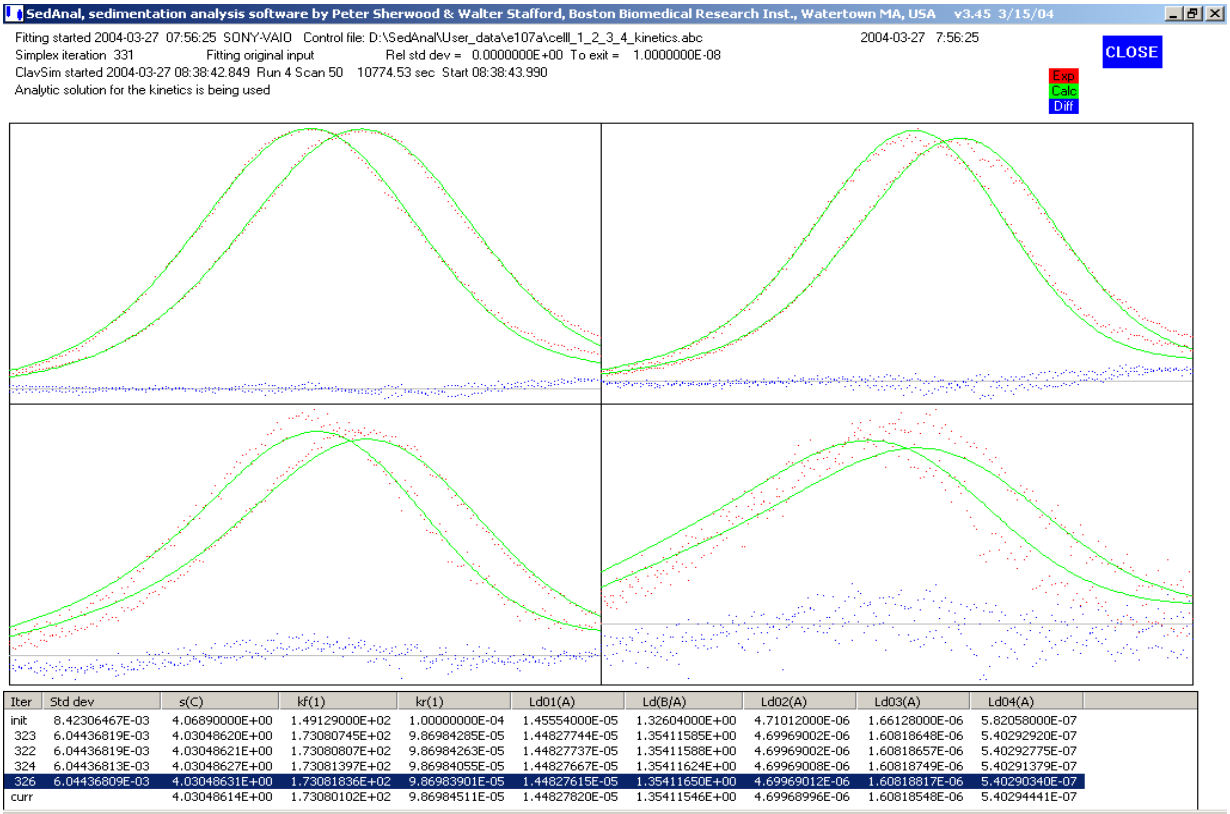


Figure 3-101. Fitting screen display for 4 cell global fit to the system A+B=C

Here (Figure 3-101) is an example of what the fitting screen looks like before convergence, after a few iterations, when performing a 4 cell global fit to the system A+B=C:

During the fitting process, the screen is updated every 10 seconds or each time a new minimum is reached. SEDANAL will print at the bottom of the screen 7 lines of text: (1) column titles, (2) the initial guesses, (3-6) the last 4 best minima, (7) the current guess being processed. The middle region of the screen displays 2 sets of three plots for each cell. The screen above shows 4 cells. In the upper left panel, the first set of 3 plots is from the first difference curve and the second set is from the last difference curve of the dataset. The intermediate difference curves are not displayed; these can be found in the output file (with extension *.min). Within the window for each cell are plotted (1) the difference data being fitted (red dots), (2) the best fit to that difference data (green line) and (3) the residuals for the current best fit (blue dots).

After the fitting process, files are written that contain the data, the fit and the residuals. These files are named "**<control_filename_prefix>__0nMin00m.txt**" where "**n**" is the cell number and "**m**" is the sequence number corresponding to the number of times that control file has been used to start a fit.

The first few lines of a "Min" file look like this:

```
SedAnal v6.14.8909 x64 2014-12-5                              2015-01-06  12:35:54  CPW-7
Control file: C:\SedAnal\User_data\MyosinV_CBD\3comp-6.14-m-di-tri.abc                                    2015-01-06  12:35:54
Run  1         Minimum 3  Iteration 2
Cell data file: C:\SedAnal\User_data\MyosinV_CBD\112503\172253\20031125_-v5.01.abr
*%* Std dev =  5.1568286675930438E-003 Params:  3.9058793100983734E+004  7.7846811277466599E-001  9.3326679883548050E-002
  6.7149280901115313E-003  2.6510978620094980E+000  3.9976541193779744E+000
  5.5219066726371402E+000
Param codes:    4    1    7  101    7  102    7  103   30    1   30    2   30    3
      #     Radius   delta-C(obs)  delta-C(calc)          DEV  Radius   delta-C(obs)  delta-C(calc)          DEV  Radius   delta-C
    300   5.96540    2.16808E-02    1.43009E-02    7.37997E-03   5.96540    1.23464E-02    7.86488E-03    4.48154E-03   5.96540    6.
    301   5.96620    2.09711E-02    1.47110E-02    6.26003E-03   5.96620    8.48934E-03    8.09546E-03    3.93877E-04   5.96620    1.
    302   5.96690    1.85509E-02    1.50797E-02    3.47121E-03   5.96690    6.27854E-03    8.30293E-03   -2.02438E-03   5.96690    3.
    303   5.96760    2.10810E-02    1.54514E-02    5.62961E-03   5.96760    6.90851E-03    8.51220E-03   -1.60369E-03   5.96760    2.
    304   5.96840    2.49410E-02    1.58922E-02    9.04880E-03   5.96840    8.45549E-03    8.76077E-03   -3.05280E-04   5.96840    2.
    305   5.96910    2.22208E-02    1.62785E-02    5.94231E-03   5.96910    1.10585E-02    8.97865E-03    2.07983E-03   5.96910    5.
    306   5.96990    1.98708E-02    1.67400E-02    3.13076E-03   5.96990    8.56546E-03    9.23938E-03   -6.73923E-04   5.96990    9.
```

The "Min" file header contains information about the fit. Most of the lines are self-explanatory.

The global root mean square deviation for the fit is labeled "Std dev =" The fitted parameters corresponding to the last minimum are listed after the word "Params:" These will be labeled more clearly in a future version.

After the first two columns, "#" and "Radius", the min file contains, in groups of three columns, the entire set of difference curves from the experimental data, "delta-C(obs)", as well as all the difference curves generated by the fitting procedure, "delta-C(calc)" and then the deviations between the data and the fit for each difference curve, "DEV".

"#" is the point number

"Radius" is the radius
"delta-C(obs)" is the first difference curve from the experimental data,
"delta-C(calc)" is the first difference curve generated for the same times
"DEV" is the residual at that point for the first difference curve

From this point on, the last three columns are repeated for each difference curve:
delta-C(obs) delta-C(calc) DEV delta-C(obs) delta-C(calc) DEV …etc…

These data can be imported into the plotting program of your choice.

3.4.2.2 Screen Dumps

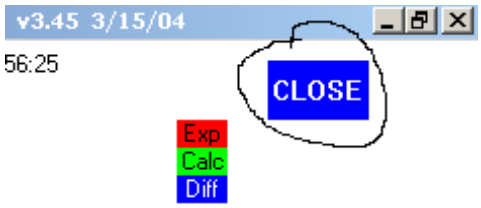


Figure 3-102. Screen Dumps

The screen can either be dumped directly to a printer at any time during the fitting process by right clicking on the "Stop/Close" button (Figure 3-102): Doing so will cause a printer dialog window to open. The screen can be dumped as a bit map to a file by shift-right-clicking on the Stop/Close button. This is done silently and no dialog window will appear. The bit map will be found in a bmp file with same name and sequence number as the report file.

3.4.2.2.1 Toggling Plots During Fitting

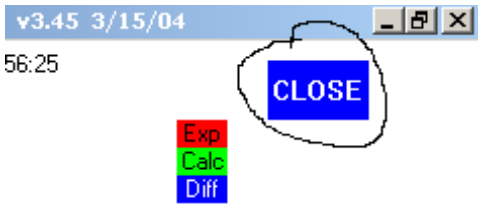


Figure 3-103 Toggling plots during fitting

As long we have this little picture (Figure 3-103) handy … One can turn on or off any of the plots on the fitting screen by clicking the little red, green or blue boxes sporting the text "Exp", "Calc" or "Diff" corresponding to the plot of the experimental data, the best fit, or the residuals, respectively.

3.4.2.2.2 Displaying the Residuals Bitmap

Residuals can be displayed in either grey scale or on a color scale by clicking on the "R" button on the upper right hand corner of the fitting window (Figure 3-104):

Figure 3-104 Displaying residuals

The residuals window shows the difference between observed and calculated difference curves (the residuals). The colors can be scaled either relative to the minimum and maximum values or to absolute minimum and maximum user specified values by clicking on the "abs" button.

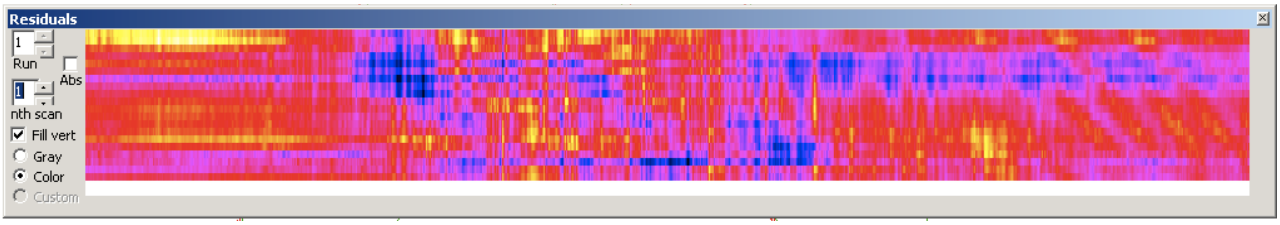


Figure 3-105 Residual plots

The plot in the residuals window has x = the radial point, y = the scan number, and the color = the residual value. Only one cell is shown at a time. The range of x is the user-selected range to fit. The color mapping is determined in one of three different ways, and scaled to the range of data, z, in the run:

Gray scale: *intensity* = $255\ ((z - z_{min}) / (z_{max}–z_{min}))$, so z_{min} gives black and z_{max} gives white.

Color: A 16-bit look-up table is used (65,535 colors) to convert $p = 2^{16}\ ((z - z_{min}) / (z_{max}–z_{min}))$ to a color. The table is shown below. Again, z_{min} gives black and z_{max} gives white.

Custom: Same as color, but the look-up table is user-specified, read in from the file `ResidualColors.txt`. The format of `ResidualColors.txt` is

```
// Optional comments
0  0  0  0
1  0  1  0
2  1  0  0
...
65534
65535
```

Each line is *p Red Green Blue*, where the color values are an 8-bit intensity 0–255.

Horizontal and/or vertical scroll bars appear if the plot window is not large enough to accommodate the data (current size is 930 radial points and 132 scan pairs). Clicking the scroll bar arrows moves one pixel (i.e. one radial point or one scan pair), while clicking the bar outside the thumb moves half the size of the plot window.

The run to be shown is selected with a thumbwheel. Only a single run's residuals are visible.

Normally, every scan is displayed (one scan per row of pixels). This can be changed with the nth scan thumbwheel to show only every n^{th} scan in adjacent rows.

3.4.2.3 Fitting to Sedimentation Equilibrium Data:

3.4.2.3.1 Loading the dataset - load equilibrium datasets before choosing the model

Open the control screen and click "New' and set the path and create a new control file. Select one of the cell data files. (Selecting a cell data file before a model is selected is not mandatory but is less confusing than the other way around). Selecting the cell data file tells the Fitter whether this is SedEq or SedVel data and arranges the Molecular Parameters boxes appropriately when the model is chosen (**Figure 3-106**).

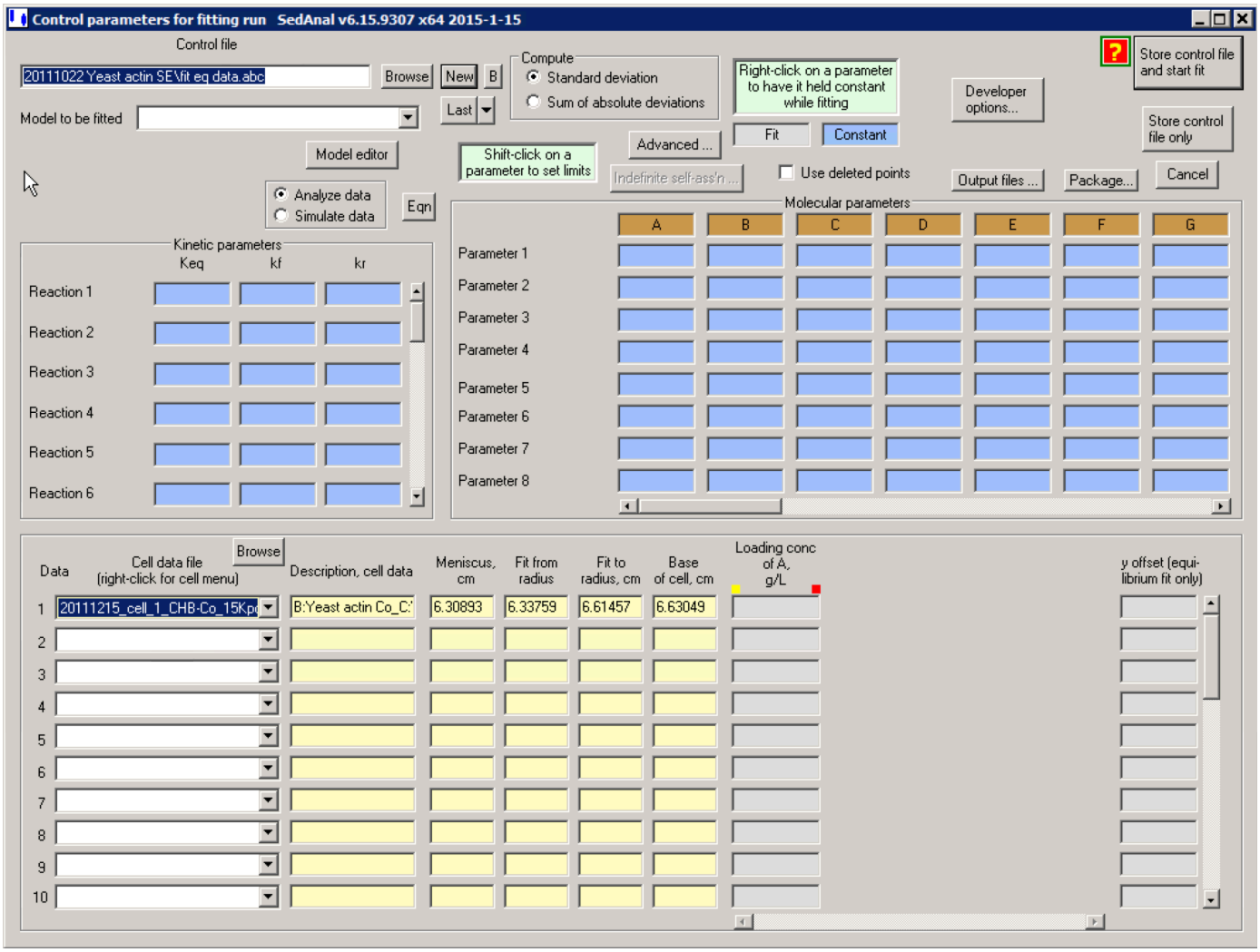


Figure 3-106 The main control screen for the fitting preprocessed data

3.4.2.3.2 Choose a model

Choose a cell data file first (**Figure 3-107**) and then choose a model (in this case for a single species) and now the control screen changes to the appropriate sedimentation equilibrium model. (i.e. no window for **s** is shown under Molecular parameters). The fitting screen reads the abr file to see whether this is a sedimentation velocity or an equilibrium run before loading the model.

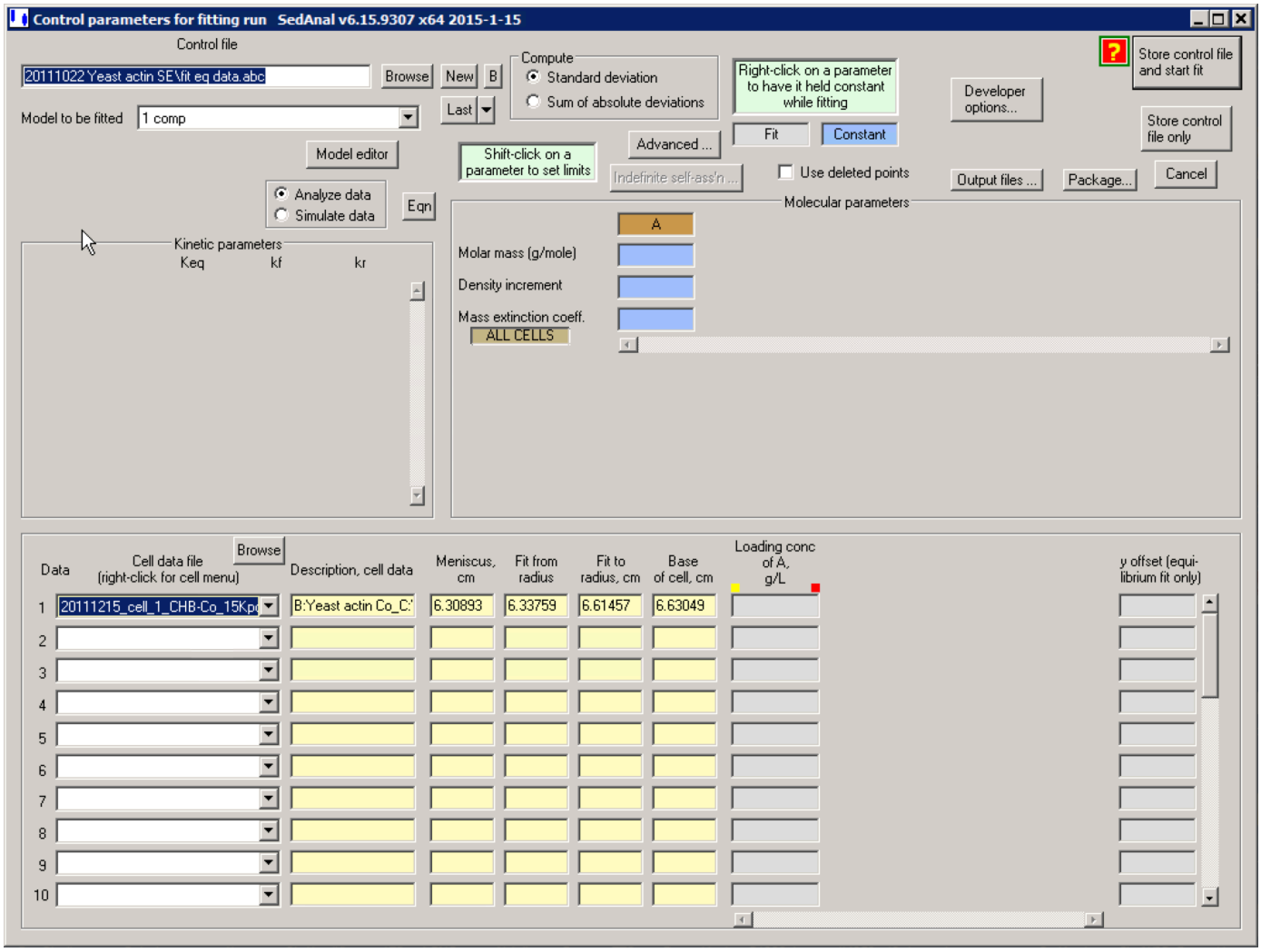


Figure 3-107. The main control screen for the fitting of preprocessed data for a sedimentation equilibrium run.

For an A+B=C model the screen would change to (Figure 3-108). In addition to the molar mass, density increments, extinction coefficients, an additional "local" cell parameter, y-offset, has been added to account for vertical offsets in the data. It is absolutely necessary to include this parameter in fit of interference data since the fringe

displacements are known only to within an arbitrary additive constant. Often it is necessary to include a y-offset to fit absorbance data if there are buffer mismatches or optical system calibration issues.

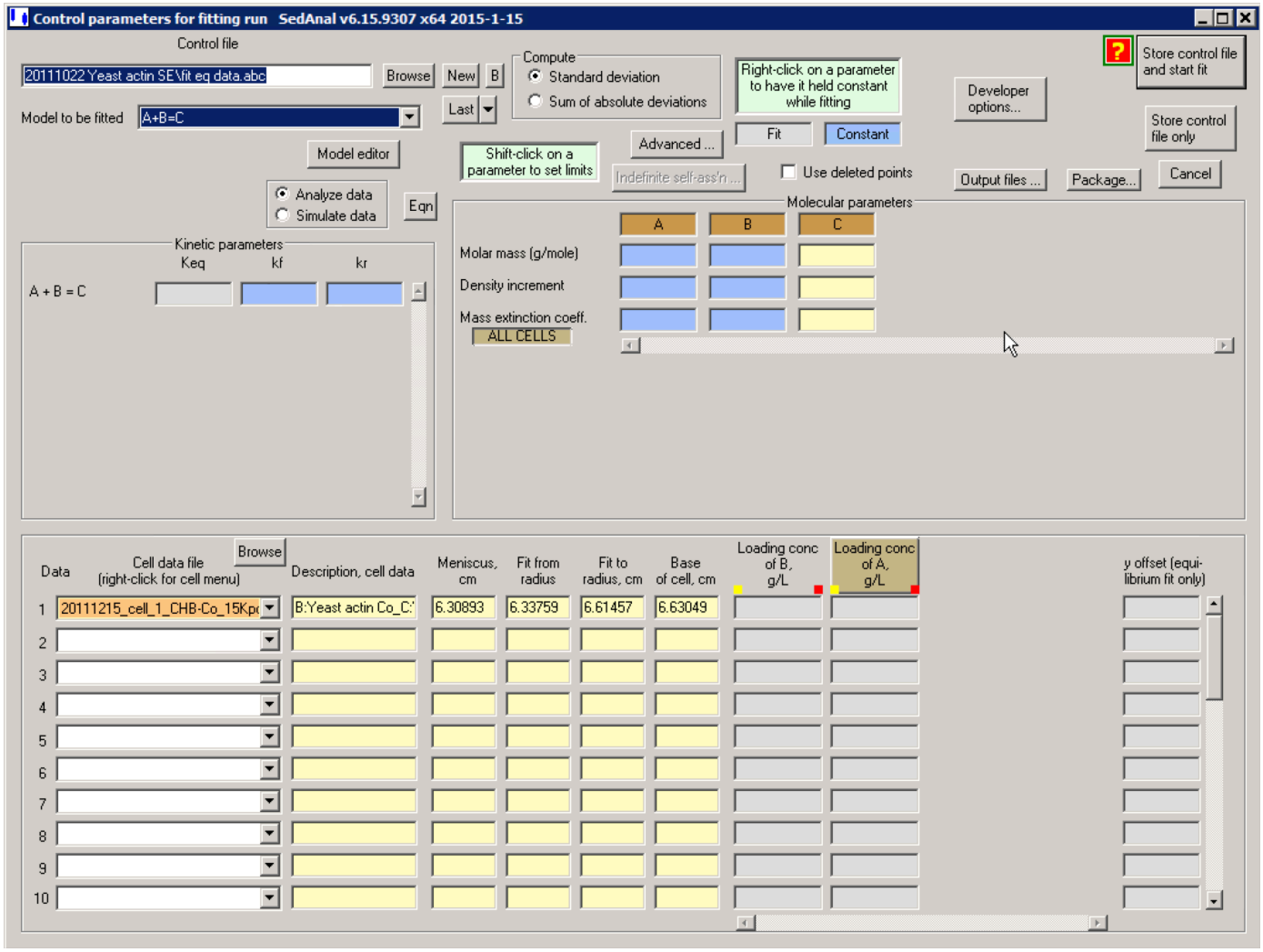


Figure 3-108. **The control screen for the fitting an equilibrium run for the model A + B =C.**

3.4.2.3.3 Enter initial guesses and other parameters

After filling in guesses for the parameters for a single species fit, the screen looks like the one below (**Figure 3-109**).

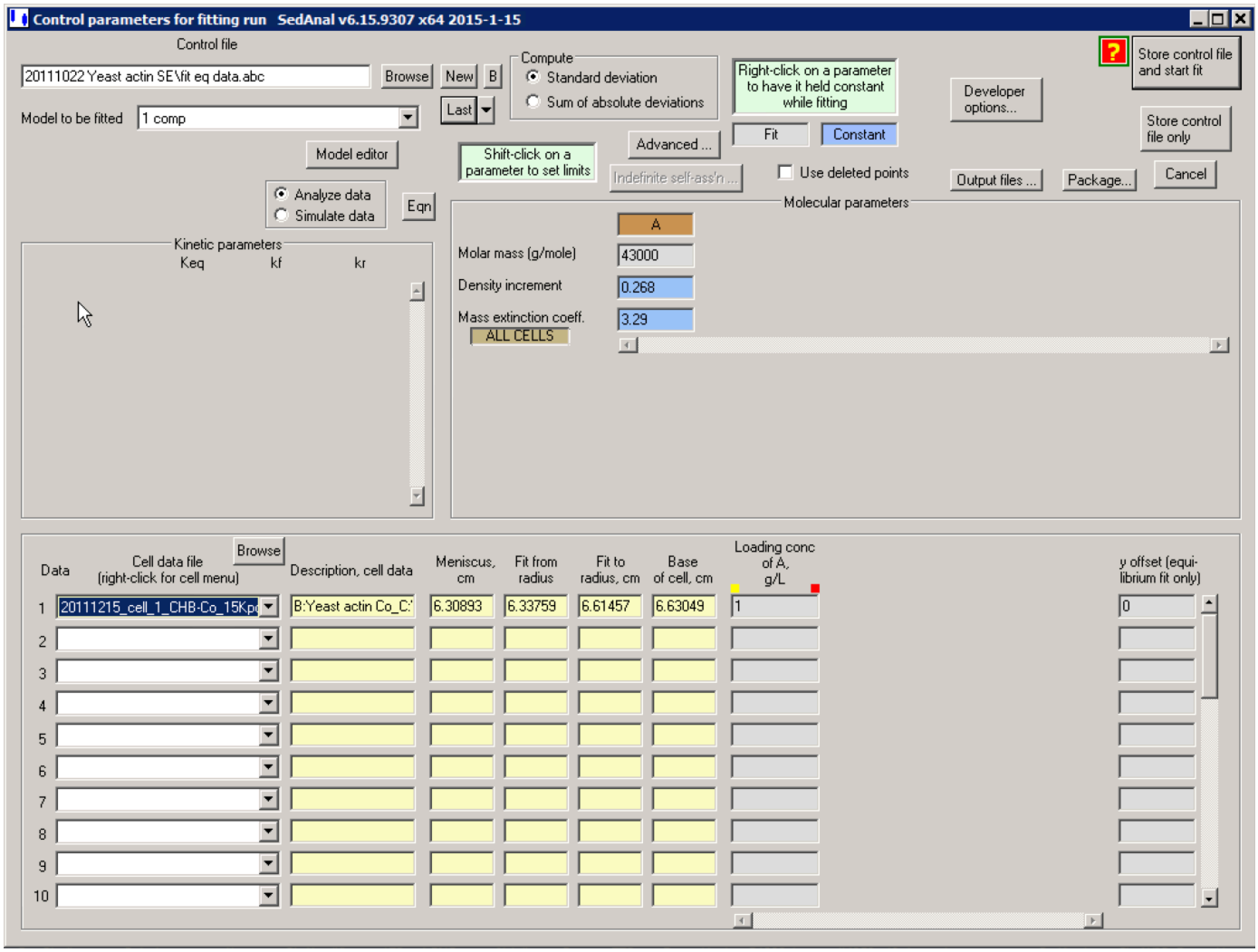


Figure 3-109. The main control screen for the fitting of preprocessed data for a sedimentation equilibrium run. Here we will float the molar mass, the loading concentration and the y-offset.

3.4.2.3.4 Start the fit

Now, after pressing "Store … and start fit", we get to the following fitting screen at convergence (**Figure 3-110**).

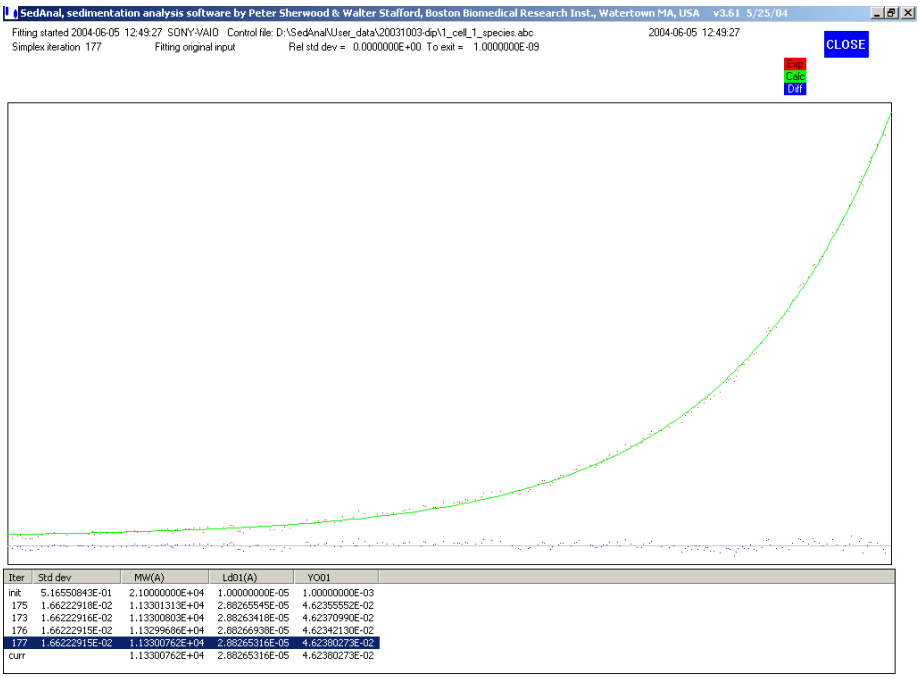


Figure 3-110. Final fit to equilibrium data after convergence.

For a multi-cell fit of a 3 component model to 3 datasets spanning 3 loading concentrations at one speed we might see (Figure 3-111).

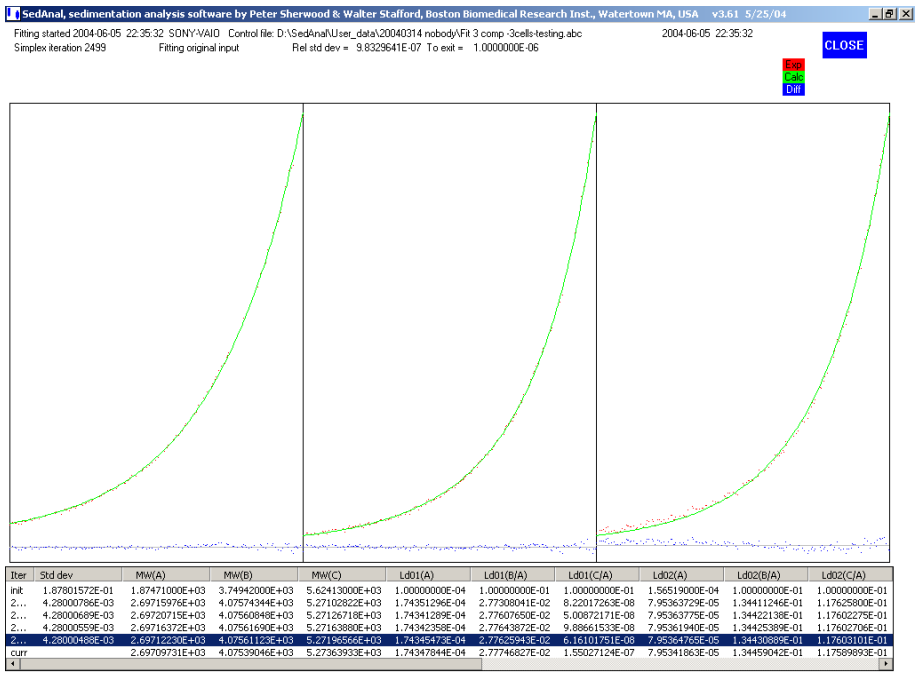


Figure 3-111**. Three-cell global fit to sedimentation equilibrium data**

3.4.2.3.5 Global fit to a dilution series with linked cells

Global fit to 9 datasets from a run at 3 speeds and 3 loading concentrations to the Model A+B=C:

The control screen was set up as follows with linked cells (note colored bars on the left side) and, since this is a dilution series, with fixed molar ratios of $[B]_o/[A]_o$

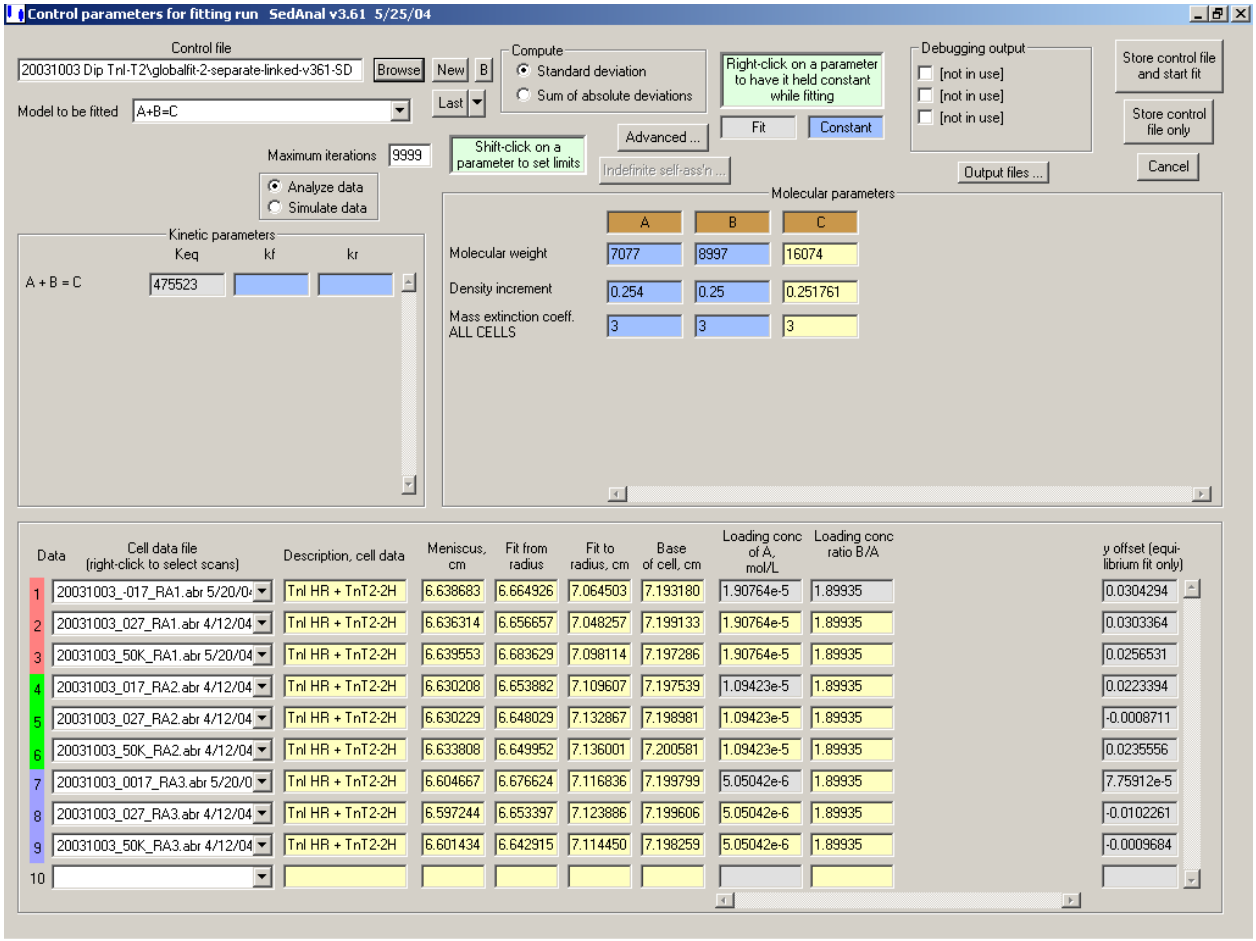
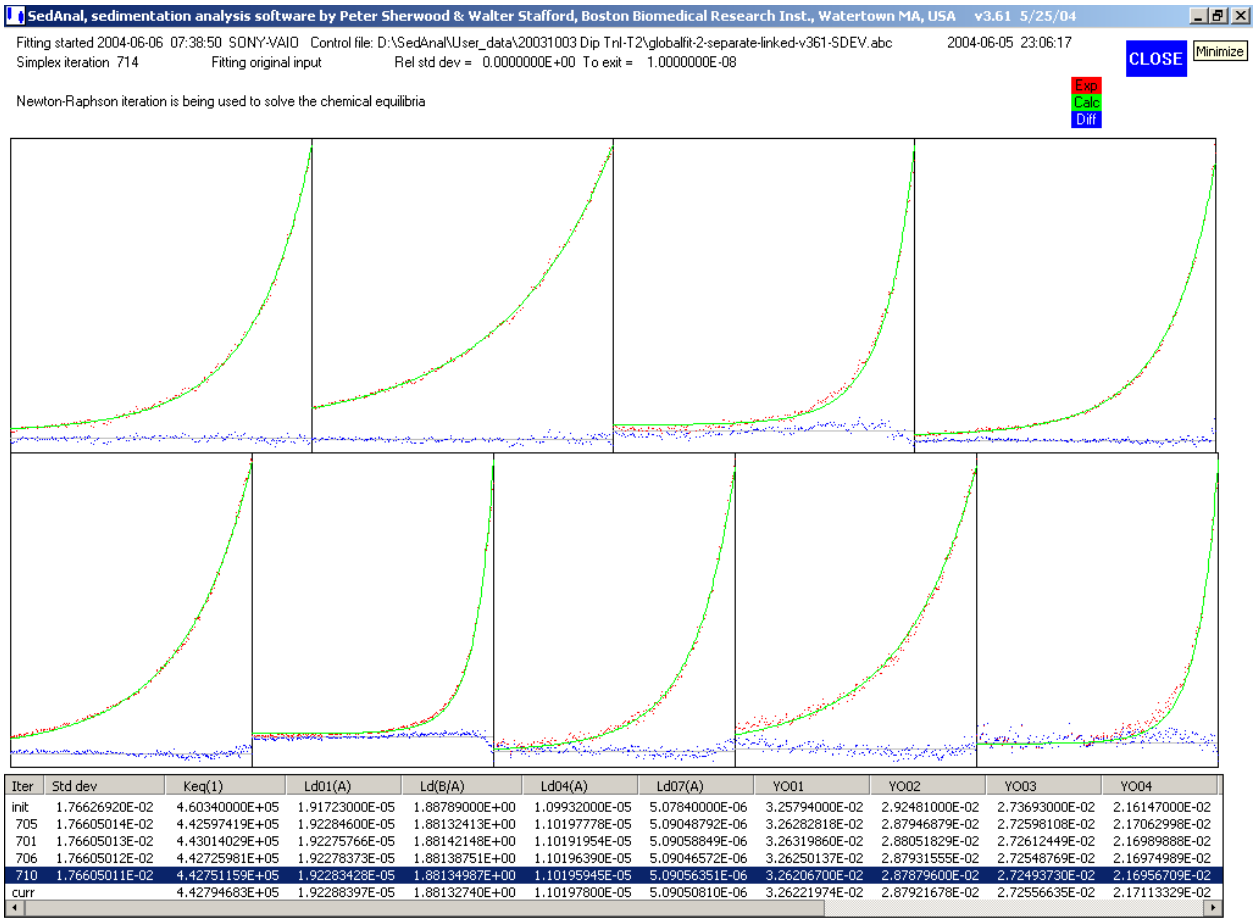


Figure 3-112**. The main control screen for the fitting preprocessed data for sedimentation equilibrium run.**

This constrained fit (i.e. requiring the cell loading concentration to be the same for the same cell at different speeds) resulted in a fairly good fit but having some systematic errors in a few of the datasets.



Iter	Std dev	Keq(1)	Ld01(A)	Ld(B/A)	Ld04(A)	Ld07(A)	YO01	YO02	YO03	YO04
init	1.76626920E-02	4.60340000E+05	1.91723000E-05	1.88789000E+00	1.09932000E-05	5.07840000E-06	3.25794000E-02	2.92481000E-02	2.73693000E-02	2.16147000E-02
705	1.76605014E-02	4.42597419E+05	1.92284600E-05	1.88132413E+00	1.10197778E-05	5.09048792E-06	3.26282818E-02	2.87946879E-02	2.72598108E-02	2.17062998E-02
701	1.76605013E-02	4.43014029E+05	1.92275766E-05	1.88142148E+00	1.10191954E-05	5.09058849E-06	3.26319860E-02	2.88051829E-02	2.72612449E-02	2.16989888E-02
706	1.76605012E-02	4.42725981E+05	1.92278373E-05	1.88138751E+00	1.10196390E-05	5.09046572E-06	3.26250137E-02	2.87931555E-02	2.72548769E-02	2.16974989E-02
710	1.76605011E-02	4.42751159E+05	1.92283428E-05	1.88134987E+00	1.10195945E-05	5.09056351E-06	3.26206700E-02	2.87879600E-02	2.72493730E-02	2.16956709E-02
curr		4.42794683E+05	1.92288397E-05	1.88132740E+00	1.10197800E-05	5.09050810E-06	3.26221974E-02	2.87921678E-02	2.72556635E-02	2.17113329E-02

Figure 3-113 Fitting with linked cells

3.4.2.4 Since the molar masses and density increments are known for the two components, only the loading concentrations, y–offsets and global equilibrium constant were fit.

3.4.2.4.1 Global fit with <u>un-linked cells</u>

Now, not linking the cells (i.e. not linking the cell loading concentrations) gives a better fit. This because the cell loading concentration depends quite strongly on having estimated the correct position of the base of the cell. In

this fit we are still holding the molar ratio of the loading concentrations of B/A to be the same for all cells with the overall value allowed to float

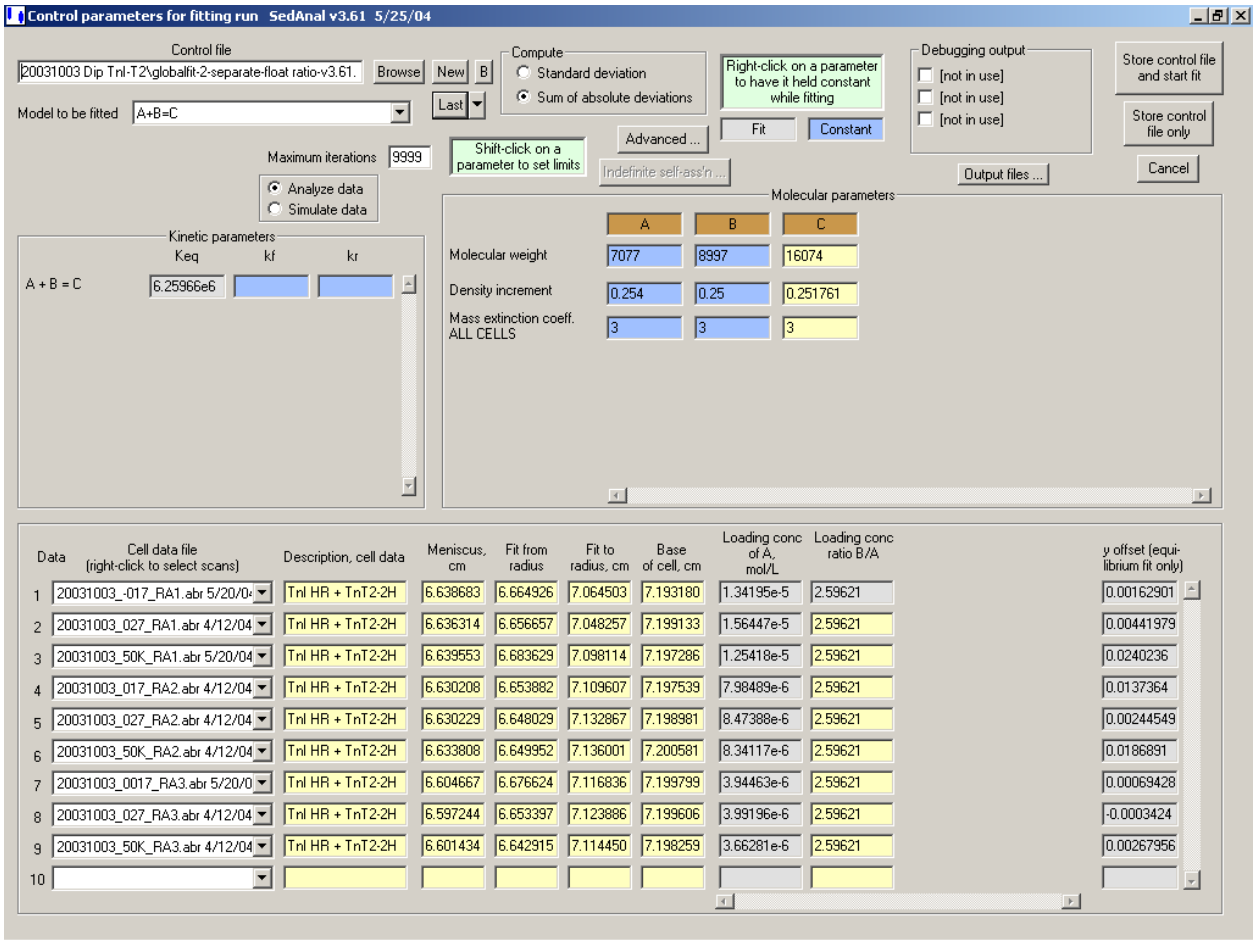


Figure 3-114**.** The main control screen for the fitting preprocessed data for sedimentation equilibrium run.

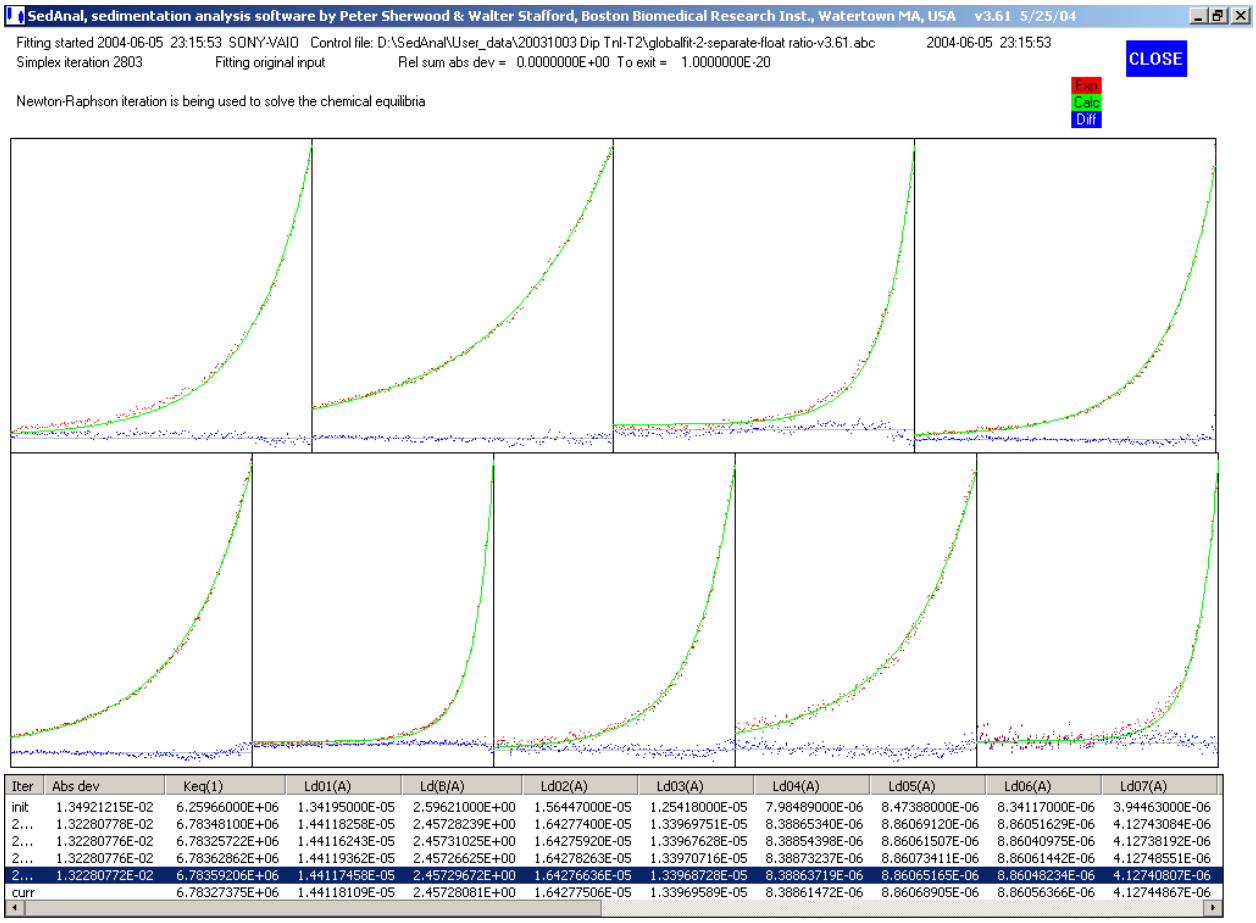


Figure 3-115. Fit with unlinked cells.

Returning to the case of linked cells but allowing the radius of the base of each cell to be a fitting parameter: we get a better linked-cell fit than the first one but still not as good letting all the loading concentrations float.

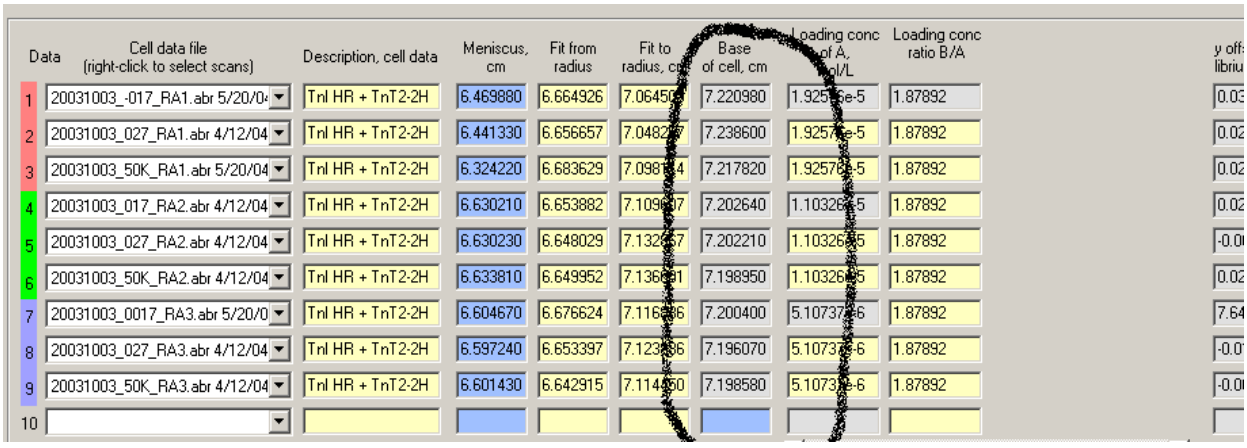
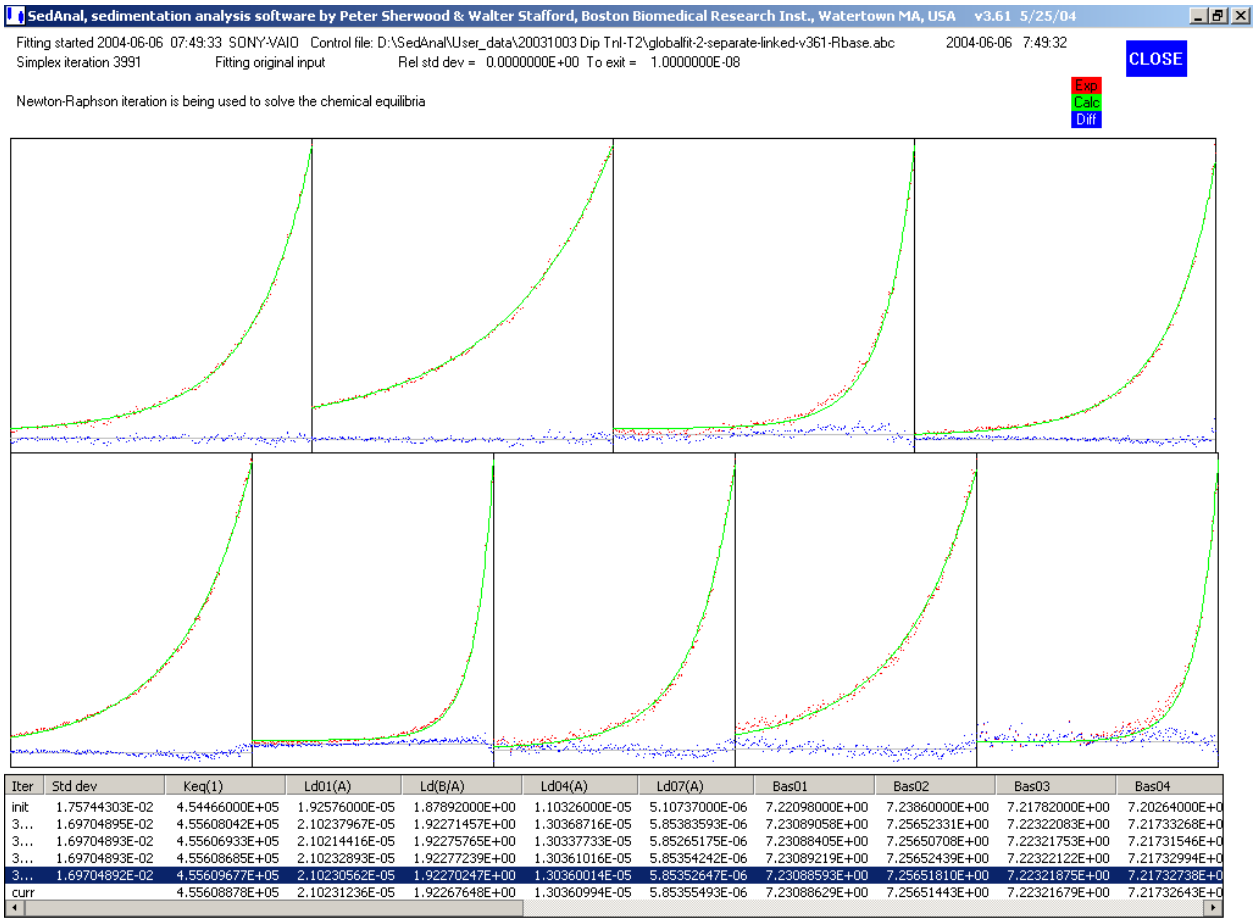


Figure 3-116 Fitting for the base radius



Iter	Std dev	Keq(1)	Ld01(A)	Ld(B/A)	Ld04(A)	Ld07(A)	Bas01	Bas02	Bas03	Bas04
init	1.75744303E-02	4.54466000E+05	1.92576000E-05	1.87892000E+00	1.10326000E-05	5.10737000E-06	7.22098000E+00	7.23860000E+00	7.21782000E+00	7.20264000E+0
3...	1.69704895E-02	4.55608042E+05	2.10237967E-05	1.92271457E+00	1.30368716E-05	5.85383593E-06	7.23089058E+00	7.25652331E+00	7.22322083E+00	7.21733268E+0
3...	1.69704893E-02	4.55606933E+05	2.10214416E-05	1.92275765E+00	1.30337733E-05	5.85265175E-06	7.23088405E+00	7.25650708E+00	7.22321753E+00	7.21731546E+0
3...	1.69704893E-02	4.55608685E+05	2.10232893E-05	1.92277239E+00	1.30361016E-05	5.85354242E-06	7.23089219E+00	7.25652439E+00	7.22322122E+00	7.21732994E+0
3...	1.69704892E-02	4.55609677E+05	2.10230562E-05	1.92270247E+00	1.30360014E-05	5.85352647E-06	7.23088593E+00	7.25651810E+00	7.22321875E+00	7.21732738E+0
curr		4.55608878E+05	2.10231236E-05	1.92267648E+00	1.30360994E-05	5.85355493E-06	7.23088629E+00	7.25651443E+00	7.22321679E+00	7.21732643E+0

Figure 3-117 Fitting for R_b. with linked cells

Since a better fit was obtained without linking the cells, it may be that some material had pelleted or gelled at the bottom at the higher speeds.

Allowing separate ratios of B/A for each cell gives an even better fit in the unlinked case.

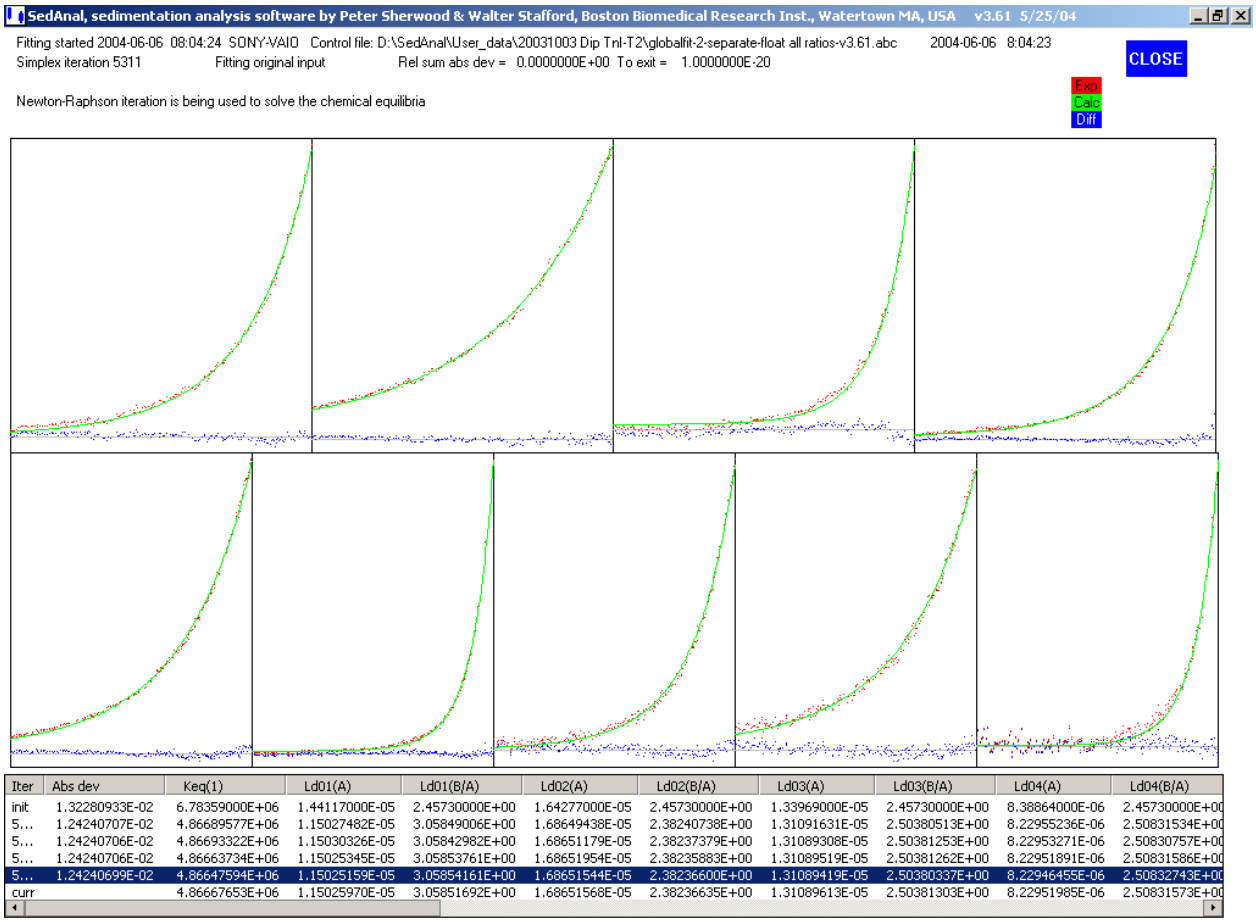
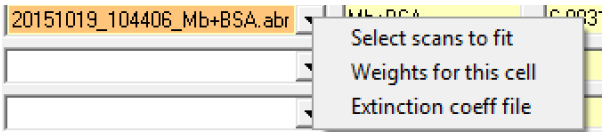


Figure 3-118. Allowing the ratio B/A to vary

Note that although this fit is better, the equilibrium constant found in this case is about 1 order of magnitude higher than that value found with the more constrained fits.

3.4.2.5 Global Fitting of Multi-wavelength Data

The large, 4-dimensional, MWL datasets can be fit directly if you supply the extinction spectra for each of the components in the mixture. After loading an abr file that was written by the PreProcessor, a model can be selected, and parameter guesses entered as usual. The extinction spectra files, and wavelengths to be used for fitting, must be specified by right-clicking on the abr filename window and selecting “Extinction Coeff file” from the drop-down list:



The following window appears that allows you to browse to choose the path to the extinction coefficient files to be used.

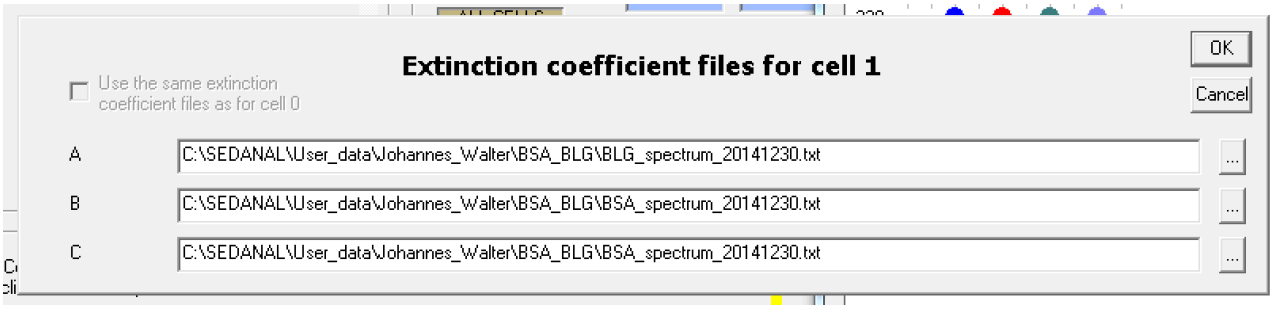


Figure 3-119 Select paths to the extinction coefficient spectra

Wavelengths to be used for fitting are selected by clicking on the intersections in the wavelength selection window:

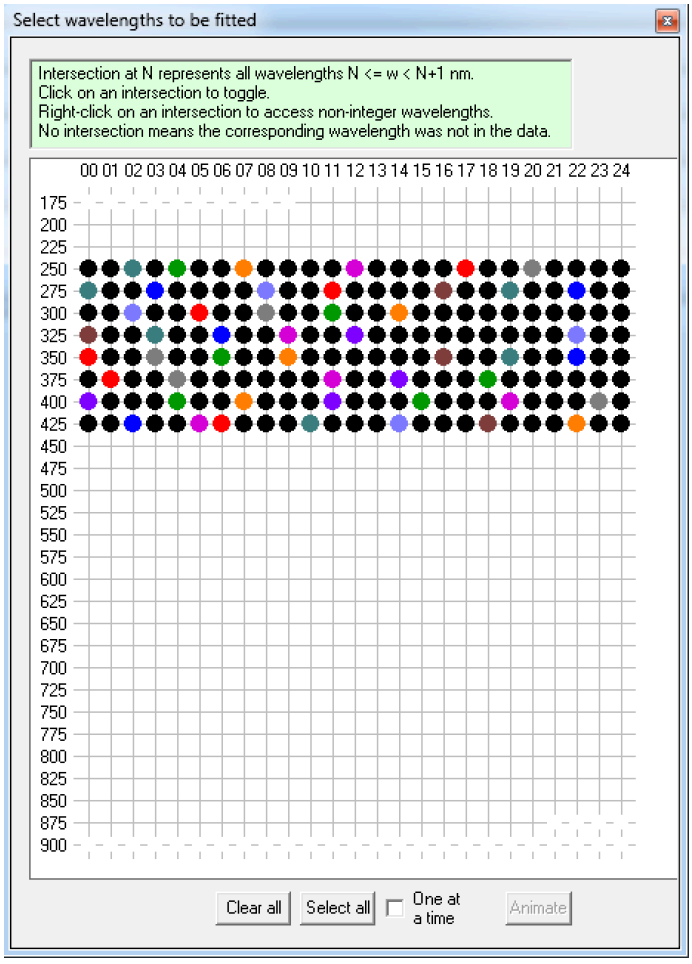
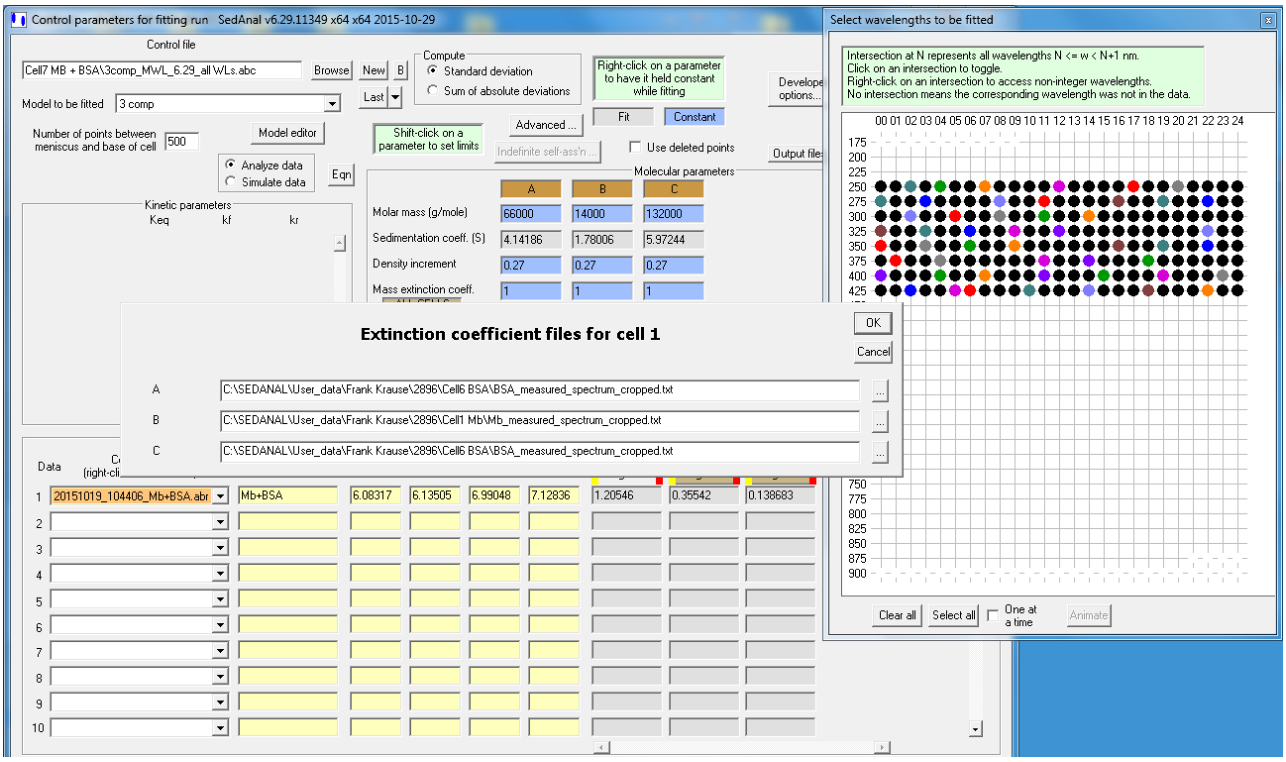
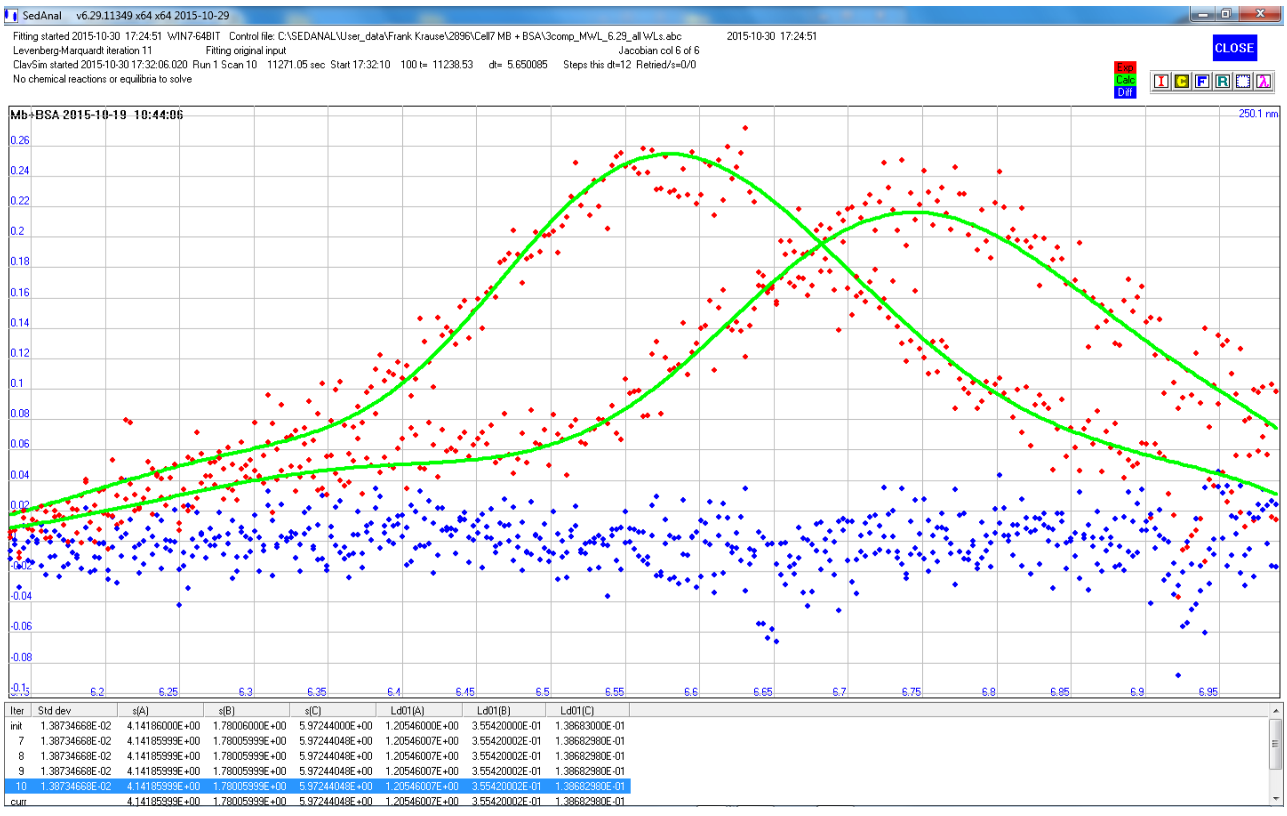


Figure 3-120. Select wavelengths to be used in the fit



When we start the fit, we see a curve corresponding to the lowest wavelength (in this case 250.1 nm)



For monitoring the fitting process, any particular wavelength can be displayed by clicking on the wavelength, "lambda", button in the upper right corner of the screen and clicking on the wavelength matrix at the desired intersection. All the wavelengths selected on the control screen are used in the global fitting procedure.

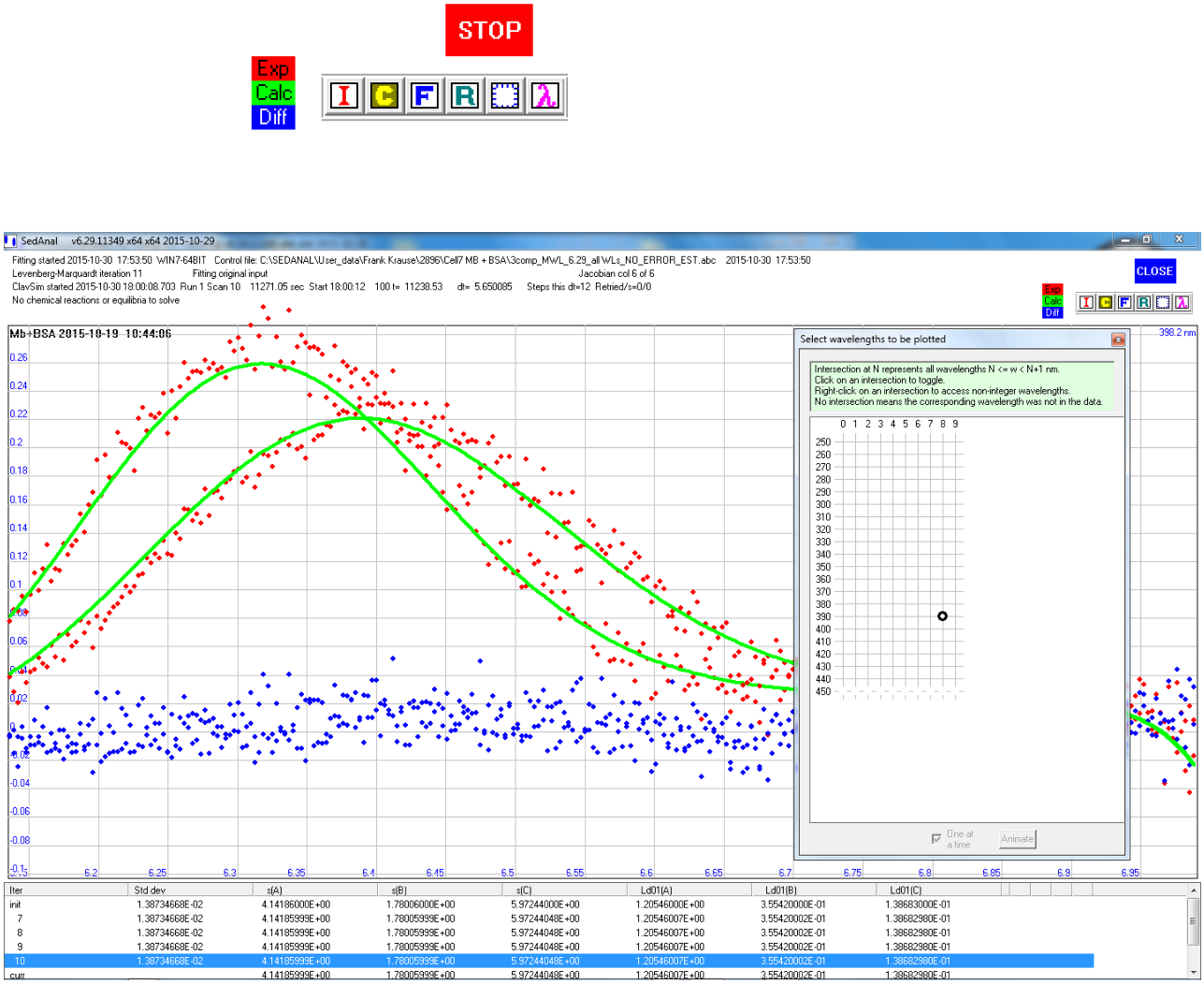


Figure 3-121 Select wavelength whose residuals are to be displayed

3.4.2.6 Fitting Multiwavelength Data at a Single Wavelength

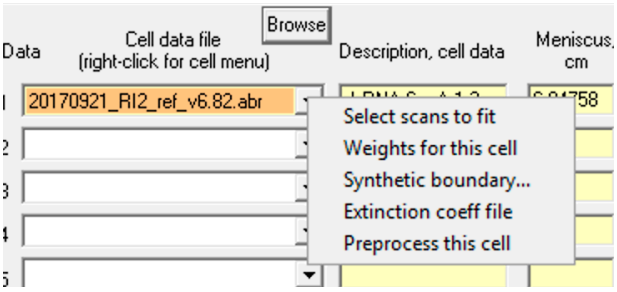
Regarding the Extinction Coefficient file:

Weight extinction coefficients are adjusted to the 1.2 cm path--or whatever path, like 3mm, that you used.

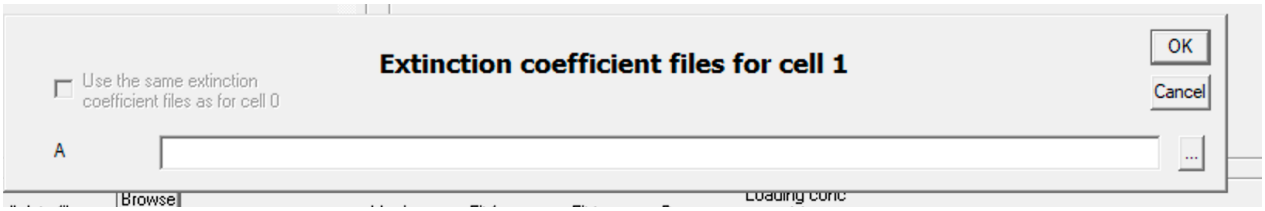
The file can be named anything you want with ".txt" extension

The spectrum files can be stored anywhere you want. It is convenient to store them in a folder one level up from the data folders in case I need them for other experiments on the same material.

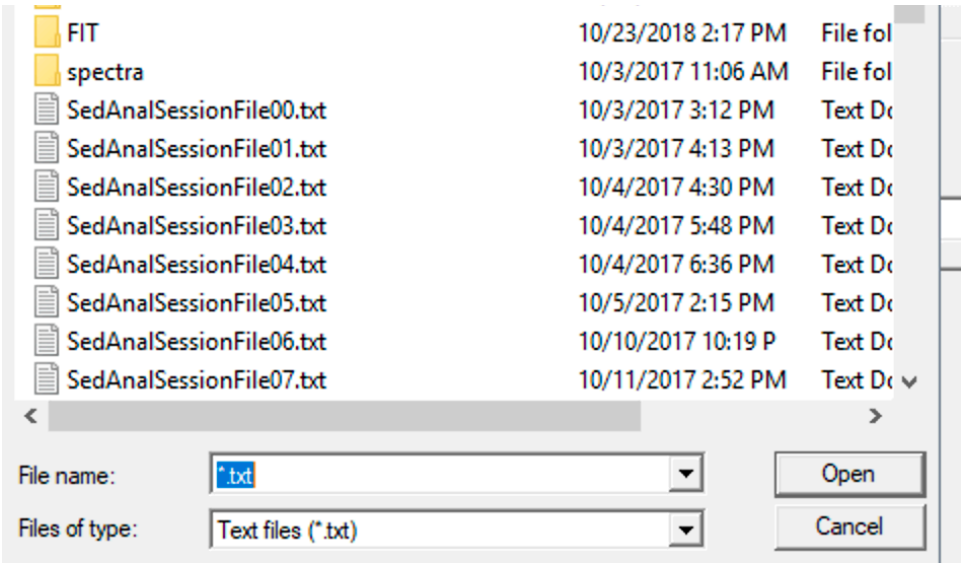
To load the spectrum file, right-click on the cell data file window in the Fitter to reveal the dropdown menu:

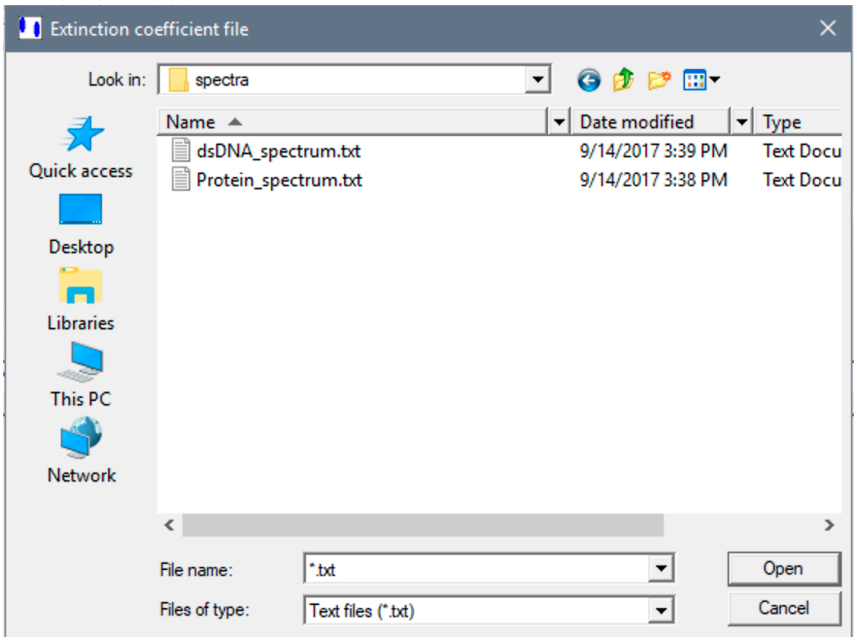


In the dropdown window, click on "Extinction coeff file" and then click on the "..." button to browse for the file.



My extinction coefficient files are in the “spectra” folder in this example: like so





Select the file you want and click "open":

A spectrum file can look like this:

```
Spectra output from dc/dt   File version 1
Peak 1 (0.4798-0.7797)
Input file: C:\SEDANAL\User_data Optima AUC\20170905.abr
WL,nm    Area
220.0 0.0593433
260.0 0.0326965
280.0 0.0632744
```

The important part is the **two columns of numbers, wavelength and extinction coefficient**; all other text is ignored (This particular file is spectral output directly from DCDT/WDA i.e. it hasn't been converted to actual extinction coefficients yet; it needs to normalized at some wavelength at which the extinction coefficient is known, and multiplied by 1.2 cm or other pathlength like 0.3 cm).

For single wavelength, it would look like

```
WL,nm    abs
280.0    1.200
```

I.e. only one wavelength need be entered into the file.

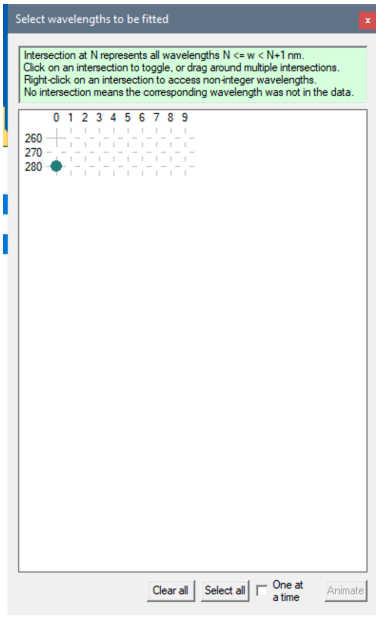
If you have done the experiment at, say, 280 nm but have only extinction coefficient values at, say, 279 nm and 282nm, SEDANAL will interpolate to get the value you need at 280 nm. Therefore, the entries in the file do not have to correspond exactly to the wavelengths being fitted.

SEDANAL will obtain the value at 280 nm by interpolating the value from these data:
(or you can do it yourself)

```
WL,nm     abs
279.0     1.210
281.0     1.230
```

However, with only one or a few wavelengths, exact values are relatively easy to generate.

You must select the wavelengths at which you want to fit:



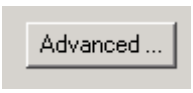
In this particular example, that sample had only two wavelengths, 260 nm and 280 nm, and was to be used to analyze a mixture of protein and RNA at 260 and 280 nm. To fit the pure protein in this cell, only 280 nm is selected.

3.4.2.7 Error Analysis: Boot-strap with Replacement, Monte Carlo and F-statistics

3.4.2.7.1 Error Estimation: Parameter standard deviations.

3.4.2.7.1.1 Bootstrap with replacement

SEDANAL will carry out a specified number of so called "**boot-strap with replacement**" or fitting operations to estimate the standard deviations of the estimates of the fitted parameter values. Essentially the way it works is that the original data, consisting of N data points, are randomly sampled, with replacement, N times to produce a new dataset that is then fit using the best fit parameters values as the starting guesses. Each fit is carried out until convergence is reached and the new fitted parameters are written into a table. After a specified number of boot-strap fits, the standard deviation of each parameter is computed. In the case of SEDANAL, for which the fitting times can be very long, it is practical to carry out only a limited number of boot-strap fits. Ideally one would prefer to repeat the boot-strap operation 500 to 1000 times and generate a distribution function for each parameter from which true confidence limits can be computed. However, it is feasible to carry out only a limited number of boot-strap fits and so an abbreviated boot-strap procedure has been implemented to compute only the standard deviation of each parameter. The ability to compute a standard deviation from a smaller number of samples of a potentially non-normal distribution relies on the Central Limit Theorem which basically states that the distribution of repeated estimates of the parameters will be normally distributed no matter what the shape of the actual distribution of the parameters. This will be an approximation that will give one a reasonably good idea of the degree to which any given parameter is determined by the data being fitted.



To specify the number of boot-strap operations to be performed, one enters the number in the boot-strap tab under the **Advanced** button of the Control Screen (Figure 3-122).

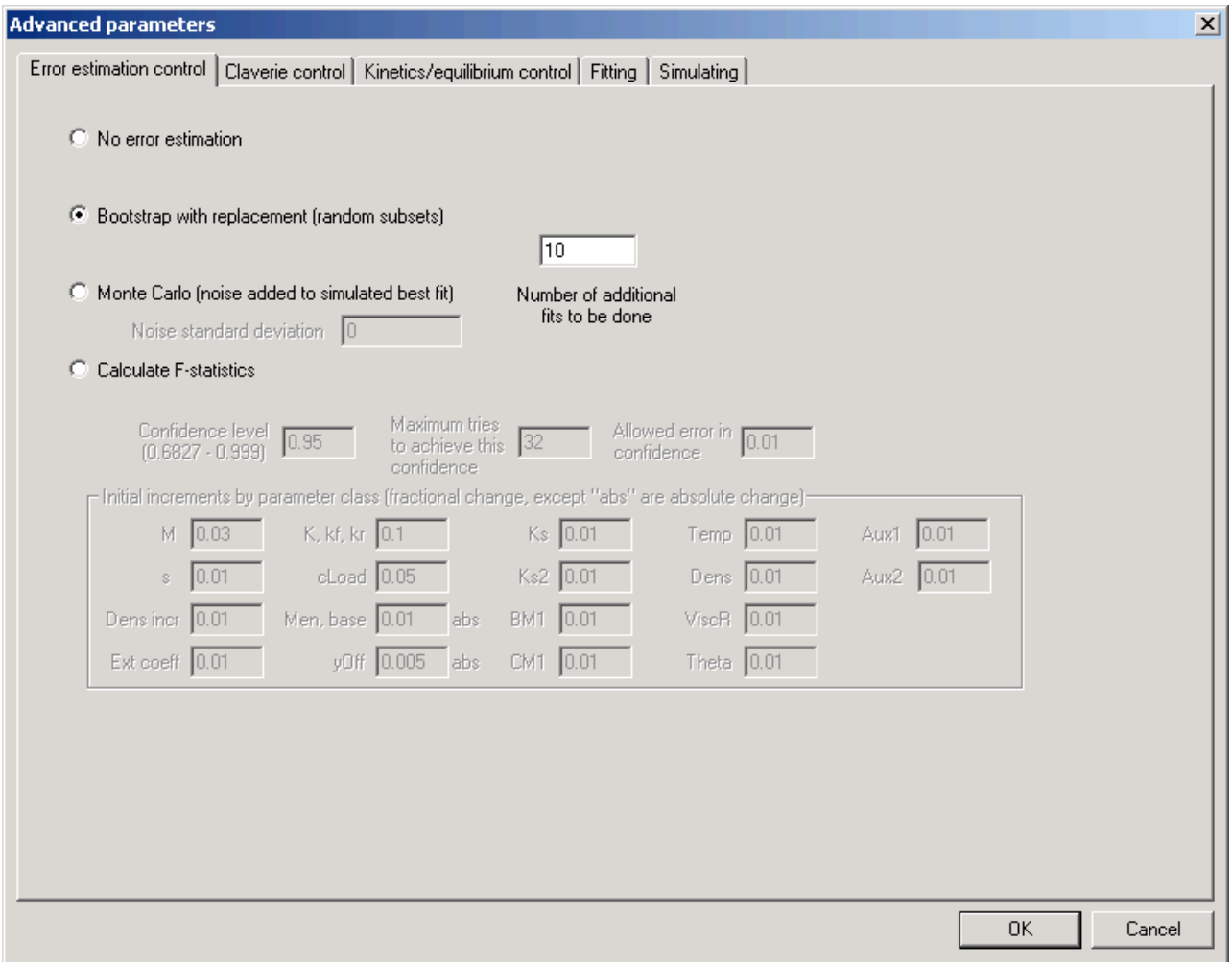


Figure 3-122**. Error estimation control**

3.4.2.7.1.2 Monte Carlo Analysis

The **Monte Carlo analysis** works similarly except that the best fit model is simulated first and then multiple different noise sets are added and the simulated data with noise are refit each time.

For a Monte Carlo analysis, the Monte Carlo button is clicked the number of simulations is entered and the standard deviation of the normally distributed random noise that will be added to the data is entered:

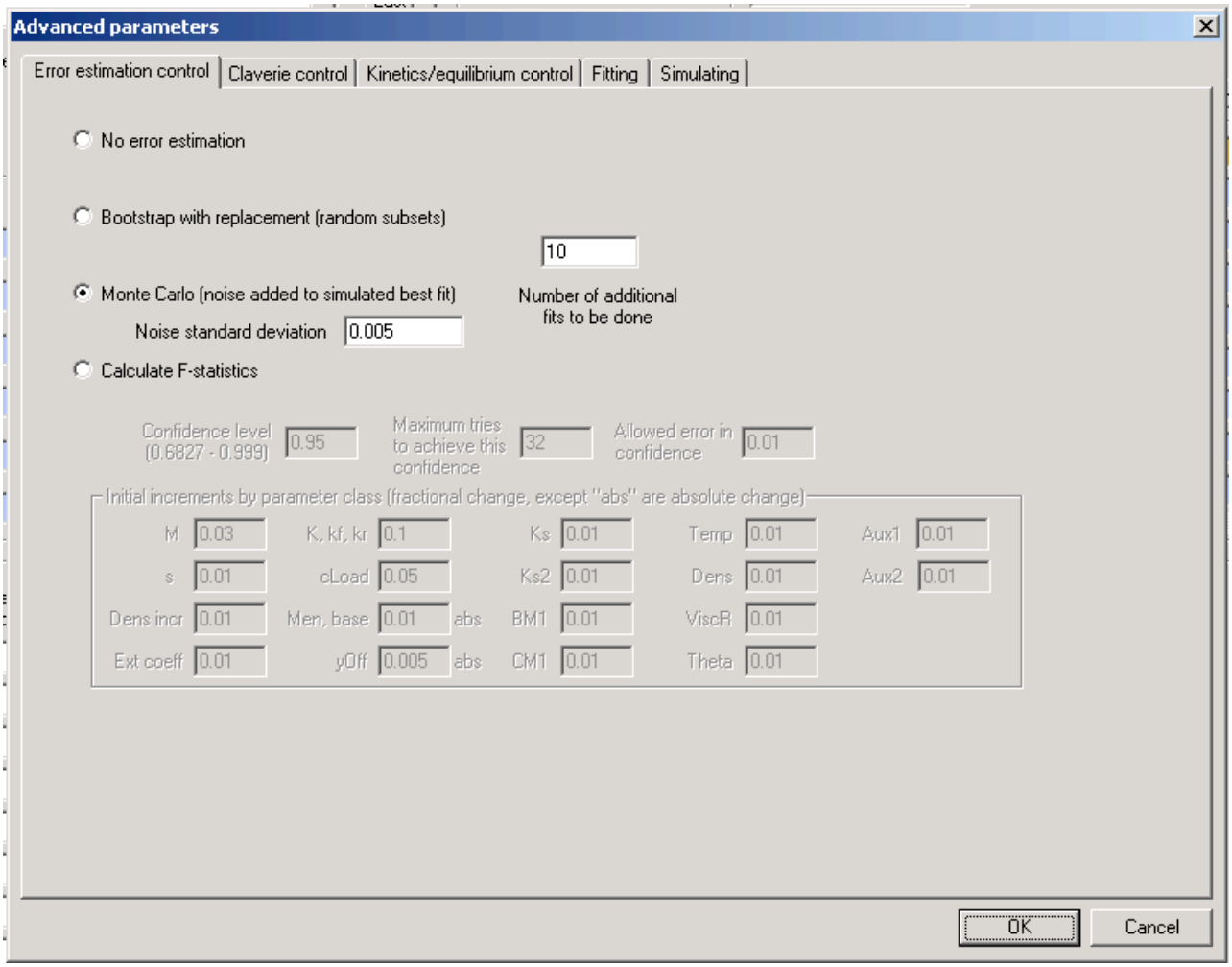


Figure 3-123**. Error estimation - Monte Carlo**

N.B.: Monte Carlo analysis is also useful when designing experiments. One can simulate particular cases using the values of the expected parameters to see the effect of noise on the ability to obtain reliable values of the fitted parameters.

The convergence level of each fit, including each boot-strap fit, can be set under the “Fitting” tab under the Advanced button.

3.4.2.7.2 Error Estimation: Parameter Confidence Limits

3.4.2.7.2.1 F-Statistics.

3.4.2.7.2.1.1 General F-stat preferences

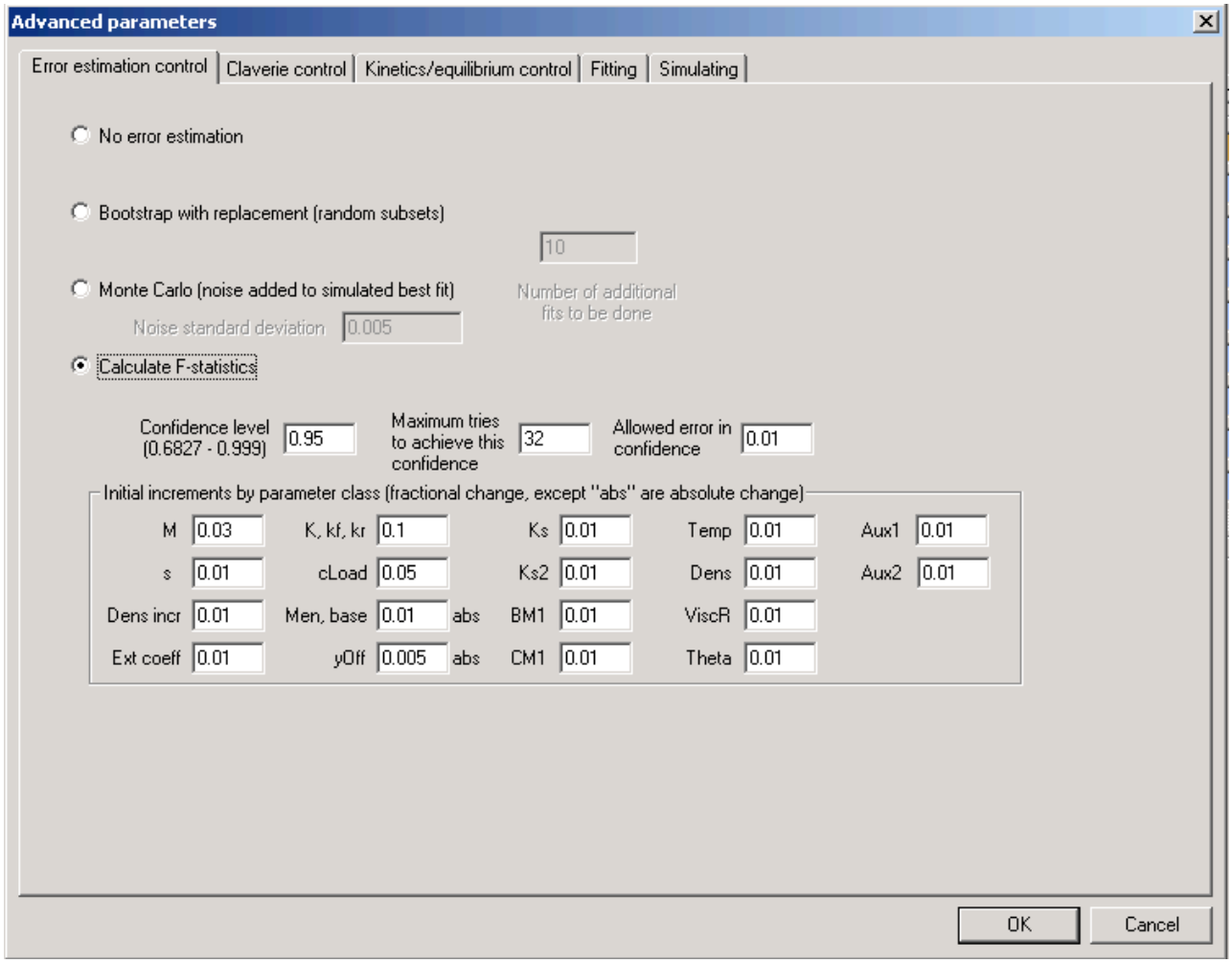


Figure 3-124**. Error estimation: F-statistics.**

To turn on the ability to compute F-Statistics for **any** parameter, under the advanced menu, error estimation tab, click the "Calculate F-statistics" button. Indicate the desired default confidence level, the maximum number of steps to take in each direction and the tolerance, i.e. allowed error in the confidence level, to stop the search of parameter space. In the boxes along the bottom of this window, you can tailor the size of the search steps to be taken for each type of variable.

3.4.2.7.2.1.2 Turning on F-stats for a particular parameter

To indicate that F-statistics are to be calculated for a **particular** parameter, left-shift-click on the parameter window on the control screen. A window will appear; check the F-statistics box and enter the desired confidence level for that parameter.

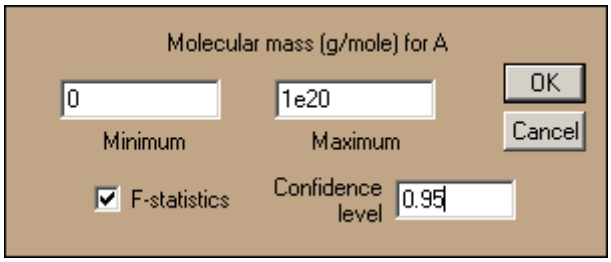


Figure 3-125 Turning on F-stats and choosing the CL for an individual parameter

Confidence limits can be determined by computing F-statistics in the following way (cf. M.L. Johnston and M. Straume, Meth. Enz. 1994). After the main fitting has finished, confidence intervals of any desired magnitude can be computed for any of the parameters by shift-left-clicking on the parameter box on the control screen, checking the "F-statistics" box and inputting the confidence level (CL), e.g. 0.95. F-Statistics are computed by stepping each parameter, holding it constant and refitting the dataset. The new rmsd is used to compute F = [rmsd(new fit)/rmsd(minimum)]^2. When F increases to the value corresponding to the CL and number of degrees of freedom, that value of the parameter is the 95% confidence level for the parameter. The parameter is varied in both directions to yield the (usually) asymmetric confidence intervals. When a fit has finished, SEDANAL will compute the confidence intervals at the level indicated (95% in this case) for those variables for which F-Stats have been enabled. F-Stats must also be enabled in the Error Estimation control tab under the "Advanced ..." button on the control screen or under the "Control Extended tab" on the preferences screen that can be accessed from the Main Menu.

The results of the F-Stat can be written to a text file for further analysis. Under the "Outputs" button on the control screen, select the "Log files" tab and click the button for the F-Stat log (Figure 3-126):

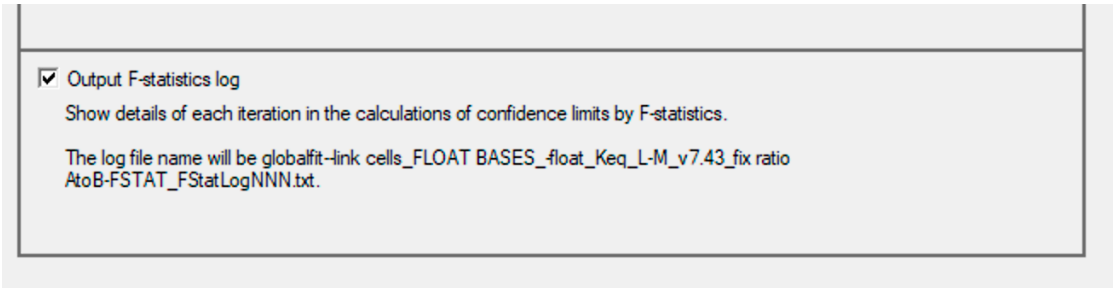


Figure 3-126 Output the F-Stat log file

3.4.2.7.2.1.3 Displaying F-stat progress

To display the F-stat progress, click on the **"F"** button on the upper right corner of the fitting window:

The following floating window (**Figure 3-127**) will open allowing you to monitor the F-Stat progress.

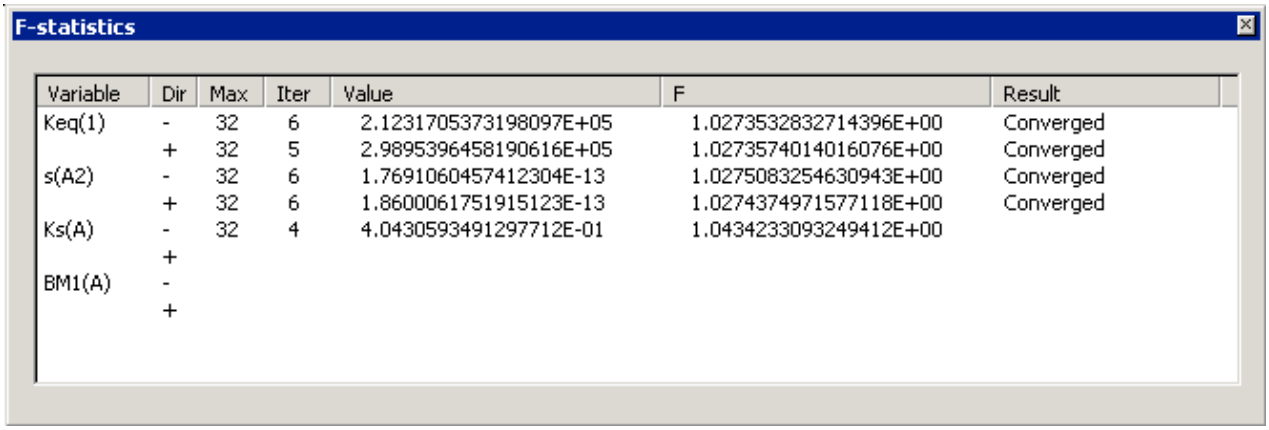


Figure 3-127. F-statistics monitoring window.

3.4.2.8 Fitting for Extinction Coefficients

By combining data from interference optics and absorbance optics, we can determine extinction coefficients of pure components. Two datasets can be combined from a single run such that the interference optics provides an estimate of the concentration, while the concentrations of the two cells are linked (i.e. requiring them to be the same) while the extinction coefficient is allowed to float as a fitting parameter. The most important requirement is that the "extinction coefficient" (i.e. numbers of fringes per g/L) for the interference signal is known or calculated from its composition, taking into account the amino acid composition, glycosylation, and nucleotide content, and any other absorbing moieties or adducts.

3.4.2.9 Fitting flotation data

Flotation data can be fit if one enters a negative value for the sedimentation coefficient of a floating species **and also** sets its density increment to a negative value. Models can also be constructed that allow for different species to sediment and float in the same sample. While DCDT and WDA can handle flotation data, they cannot be used to analyze sample which contain both sedimenting and floating species.

3.5 DCDT and WDA

Time derivative analysis, g(s*), and multi-speed wide distribution analysis (WDA) can be carried out by selecting the "dc/dt" button on the main menu (Figure 3-23). The "Concentration profile time-derivative analysis" window will then appear.

3.5.1 DCDT

A run file is loaded by clicking in the "Experiment" window to specify which experiment is to be analyzed followed by clicking in the "Cell data" window to select a particular cell data (.abr) file (**Figure 3-128**). When the cell data file is selected, all the data in the "abr" file will be presented. Usually the run will span a large time interval and the data plotted may look rather strange. The scans to be analyzed are selected in the same way they are in the Fitting control screen, by right clicking on the Cell data file window. The numbers of scans to be used should be selected consistent with the molar mass of the species present. If too many scans are included, the g(s*) peaks will be artificially broadened. After the desired scans have been selected, the plots should look similar to **Figure 3-129**.

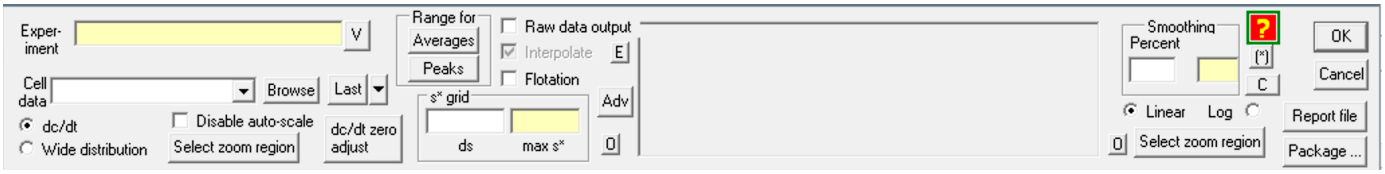


Figure 3-128 DCDT header from v7.43

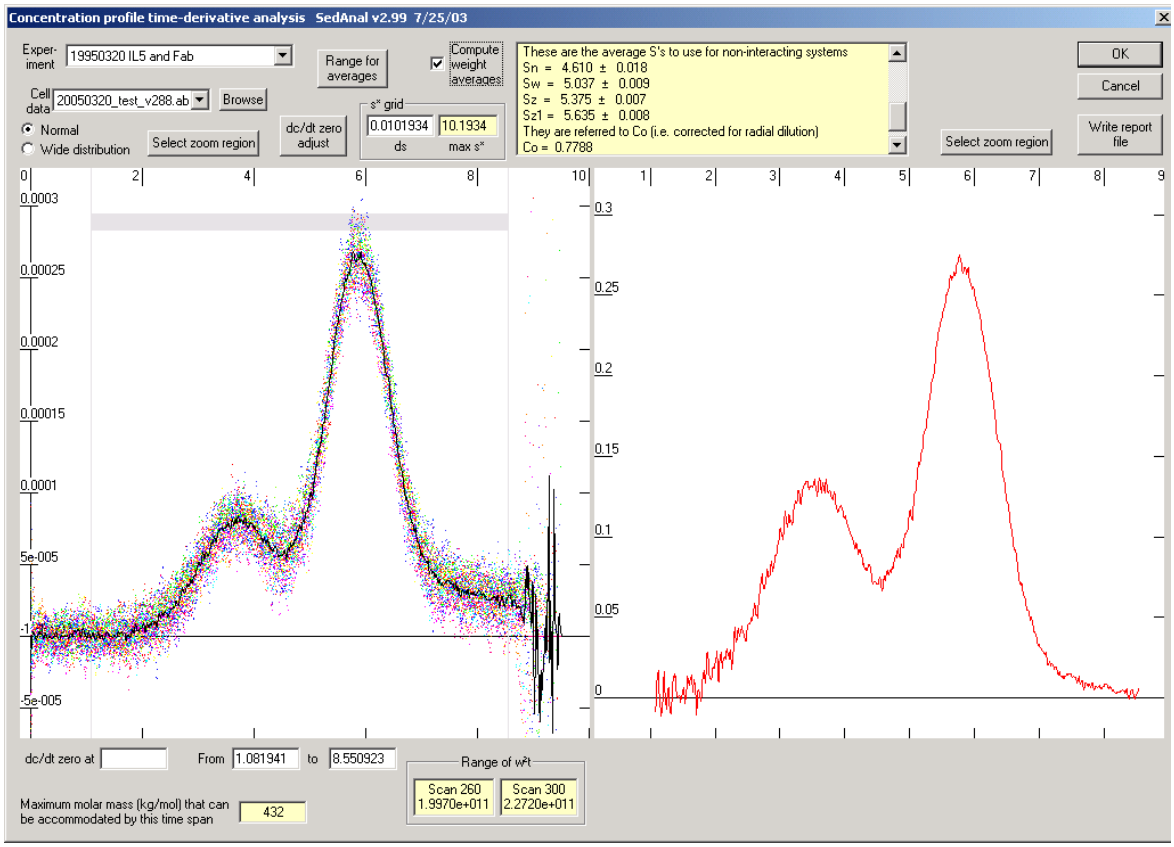


Figure 3-129 (left) dc/dt. (right) g(s*)

The left window contains a plot of "dc/dt vs s*" for every scan pair used in the analysis. A range of data can be selected in the left window to remove undesirable stray data points that might interfere with the analysis. The range chosen at this point will be used for the final "g(s*) vs s*" plot displayed in the right window, as well as in the computation of the various averages of the sedimentation coefficient, which will be displayed in the top window after the "Compute weight averages" box is checked.

The s*-axis scaling is selected in the s*-grid window. There are a thousand s* points plotted across the window. For example, if a value of 0.01 is selected for the s*-grid spacing, the value of s* will span from 0-10S. After the value of s*-grid has been edited, you must right click the Cell data window and then click on the OK button to refresh the screen. Left clicking on the Cell data window will reload the entire dataset.

The plots may be re-scaled by clicking on the "**Select zoom region**" button and using the mouse to select a rectangular region to be re-plotted. The previous zoom levels are stored in a stack. A previous zoom level can be regained by typing CTRL-Z and a later zoom level, by typing CTRL-Y .

Figure 3-130. The weight average sedimentation coefficient

The weight average sedimentation coefficient may be computed by selecting a range over which to do the average by first clicking on the button labeled “Range for averages”, selecting a range on the dc/dt plot by clicking and dragging and un-clicking.

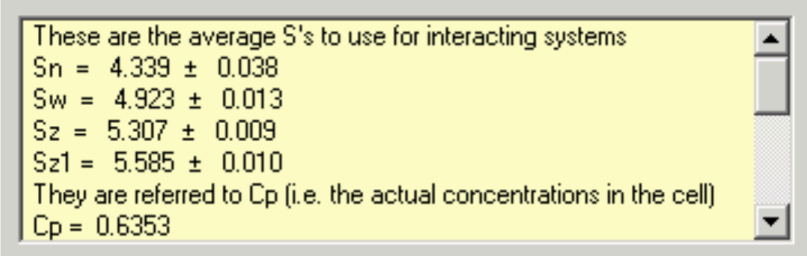
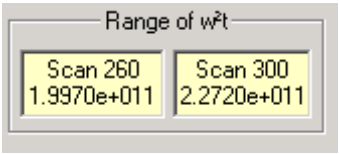


Figure 3-131 The corresponding averages appear highlighted in yellow in the adjacent box.

Note: When averages are computed, the report will also contain the normalized values of g(s*) and s*g(s*).



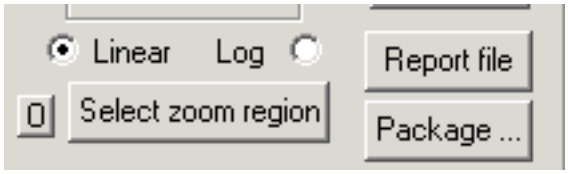
The temporal range of the data is displayed in a window at the bottom of the screen.

Figure 3-132 Maximum molar mass for DCDT

Also displayed is the maximum molar mass of a macromolecule whose diffusion coefficient could be reliably determined from the dataset (i.e. the set of scans) chosen by fitting the g(s*) vs. s*pattern to a gaussian.

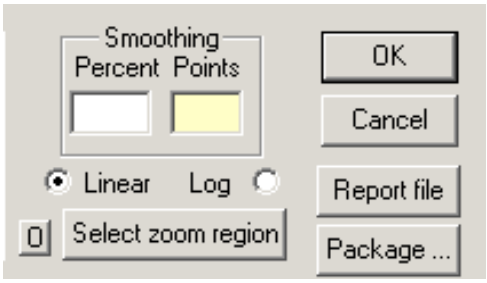
3.5.1.1 Switching between g(s*) vs s* and s*•g(s*) vs ln(s*)

It is possible to change the y and x-axis scales. On the s*g(s*) vs ln(s*) scale, the vertical off-set is a constant and independent of the value of ln(s*).

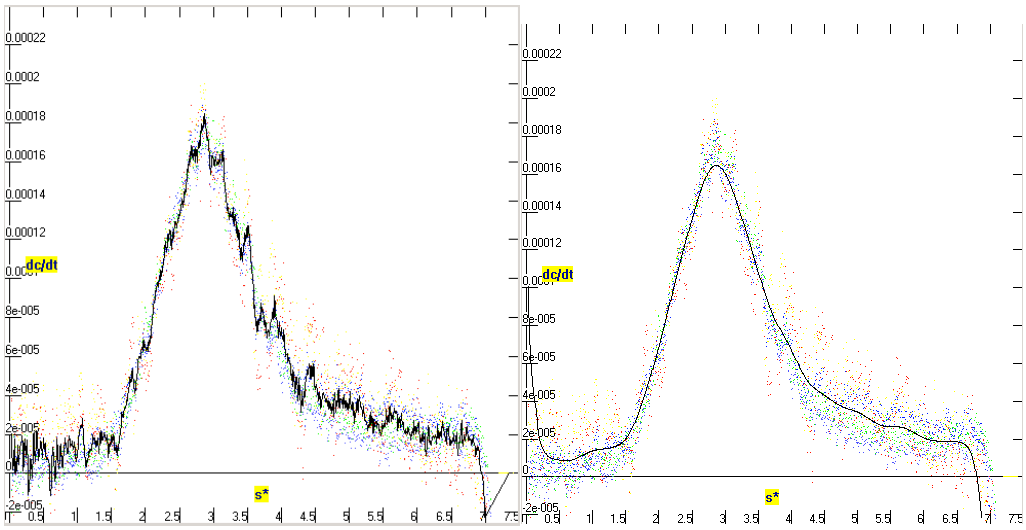


3.5.1.2 Smoothing of g(s*) vs s* for presentation purposes

Some smoothing can be applied to g(s*) curves.

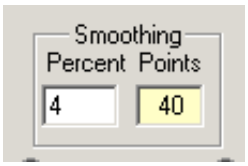


With no smoothing a plot might look noisy; like:



No smoothing and with 4% smoothing

Figure 3-133 Smoothing DCDT analysis

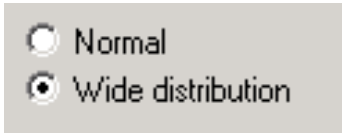


Adding 4% (i.e. using 4% of the total data span in the sliding fit) smoothing gives the plot on the right. The unsmoothed individual dc/dt plots are shown as dots in the background. This degree of smoothing distorts the plot very little while removing most of the noise.

3.5.2 Wide Distribution Analysis (WDA)

Data from a either a single speed run or a multi-speed experiment is processed as described above. Before loading the run file (.abr file), click on the “Wide distribution” button

Dialog box



The following two windows (Figure 3-134) will appear.

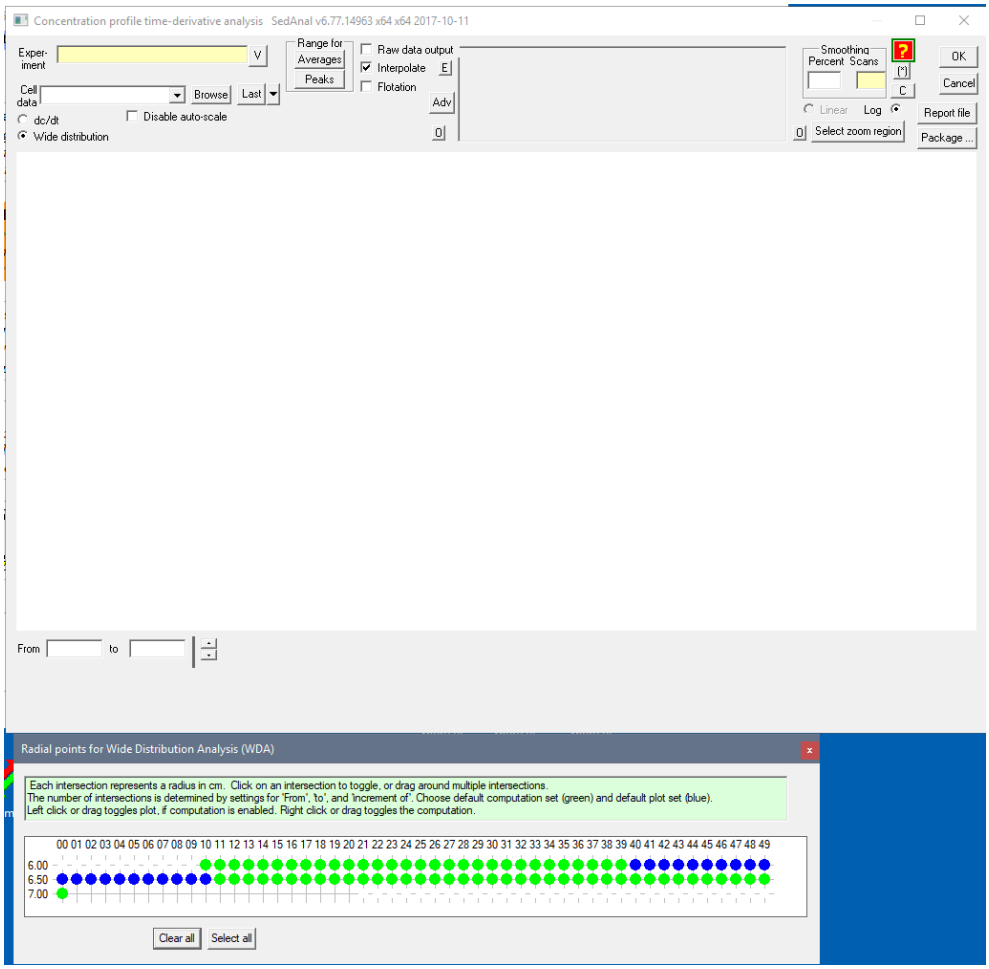


Figure 3-134 The radius control box for performing wide-distribution analysis.

The bottom window shows which radii will be used for WD analysis and allows the user to change the selection of radii. This box may be just below the bottom of the WD Window and not visible on some monitors if the resolution is set too low. The minimum recommended resolution is 800 vertical, 100%, and at this setting the tool bar can be only one layer. At this resolution the title bar will be just visible at the bottom of the screen.

3.5.2.1 Choice of radii for WD Analysis

The choice of which radii to use for WDA depends on whether or not the run was allowed to clear the slowest sedimenting material is allowed to clear the point of observation.

First, we'll explore how WDA is performed. Given a set of scans, first we choose a radius at which to "sit" and observe the passage of sedimenting material. The figure below shows the passage of the boundary at each radius.

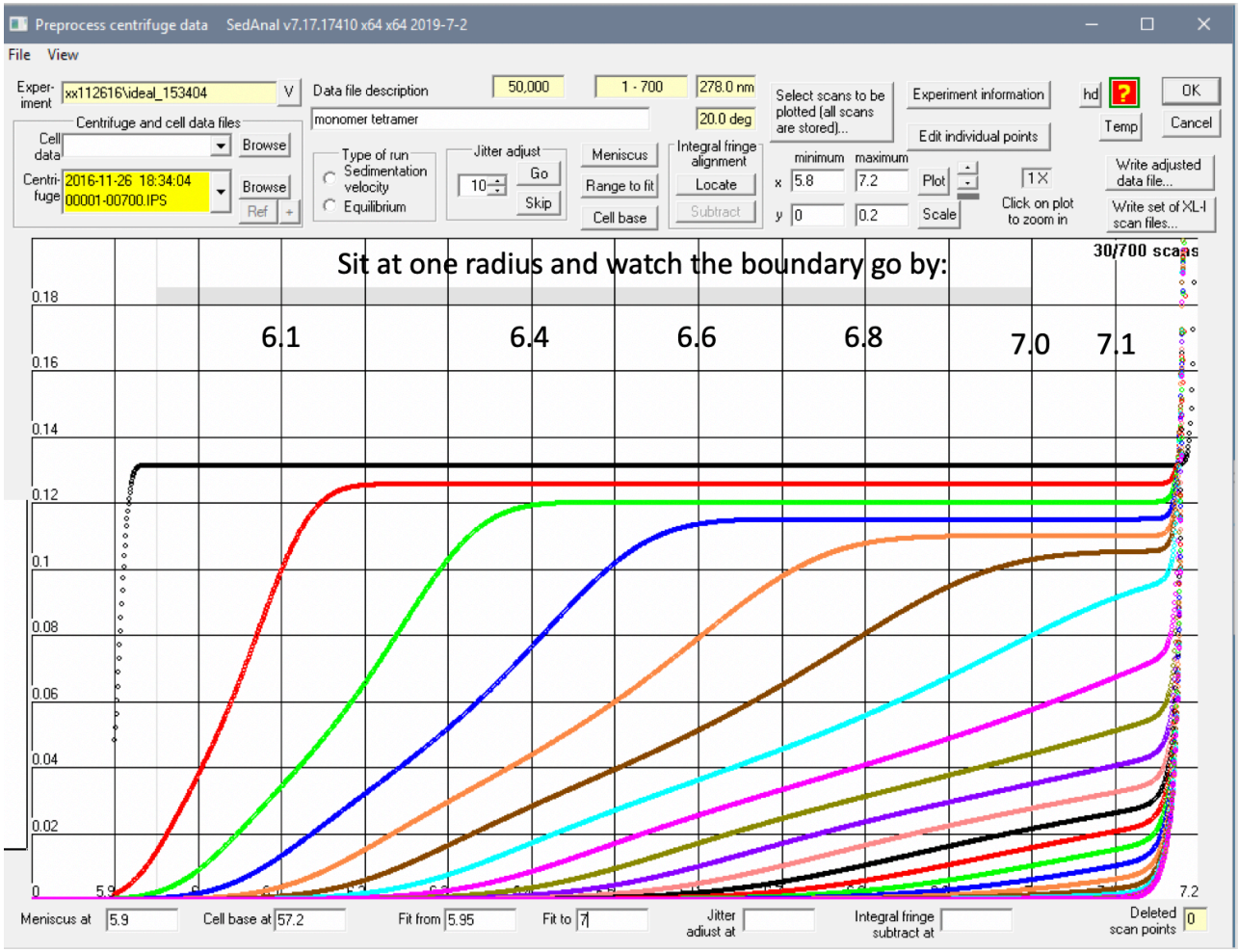


Figure 3-135 A set of scans from a sedimentation velocity experiment.

Separation is roughly proportional to the first power of the time (distance traveled) while the spreading due to diffusion is roughly proportional to the square root of the time (distance traveled). Therefore, one can expect that the overall resolution to increases as the square root of the distance traveled. See Figure 3-136 below showing the increase in resolution as the radius at which the observations are made is increased from 6.1 cm to 7.1 cm. If one were to run until the slowest species is just at the bottom, a radius in the middle of the cell--say, around 6.5 cm--would allow essentially all the material to pass that point by the time the run is stopped. It would be best for most types of analysis to run nearly twice as long as that to assure that the slowest material has cleared the cell completely. In this case, a radius closer to bottom--say, 6.9 -7.1 cm-- would be appropriate allowing high resolution.

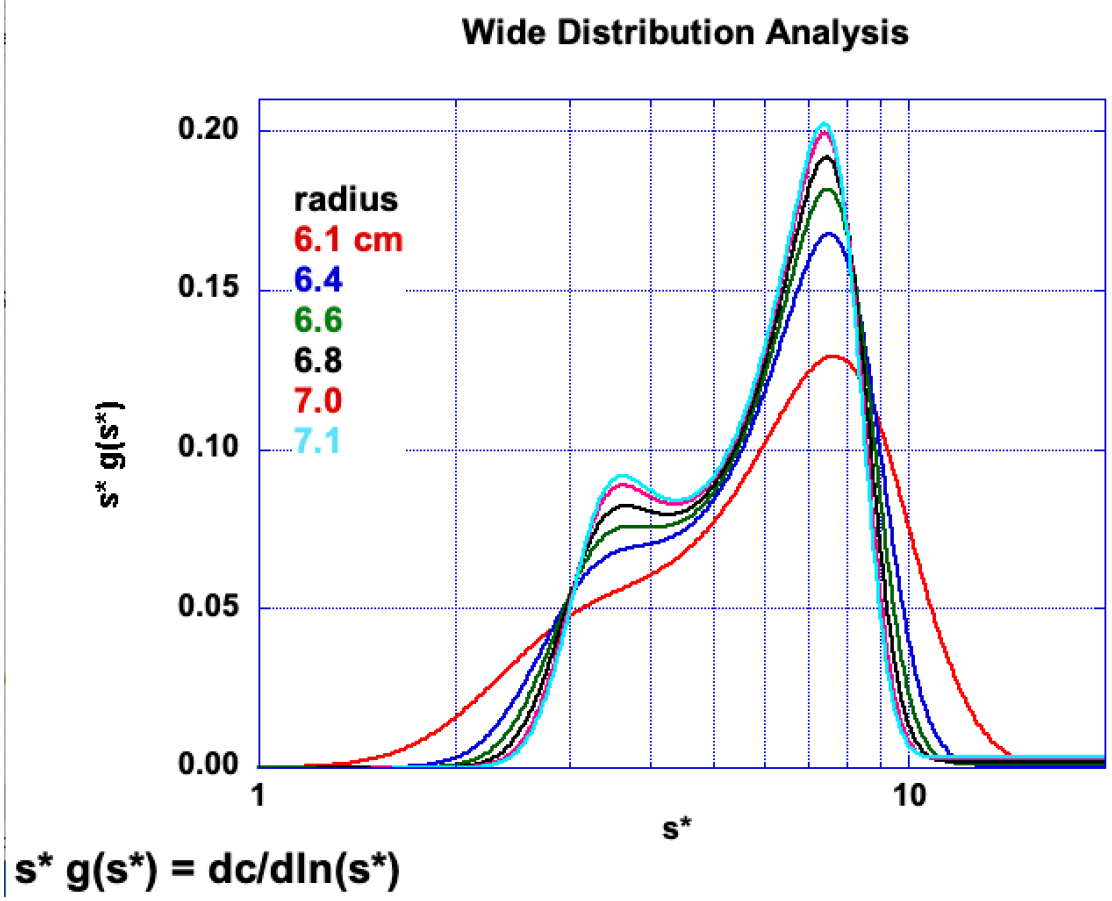


Figure 3-136 WDA resolution as a function of radius.

After choosing a run file, and selecting a smoothing range (see below), a plot will appear from data taken at a set of radii selected for plotting in the preferences: in this case a single radius of 6.6 cm was selected (**Figure 3-137**).

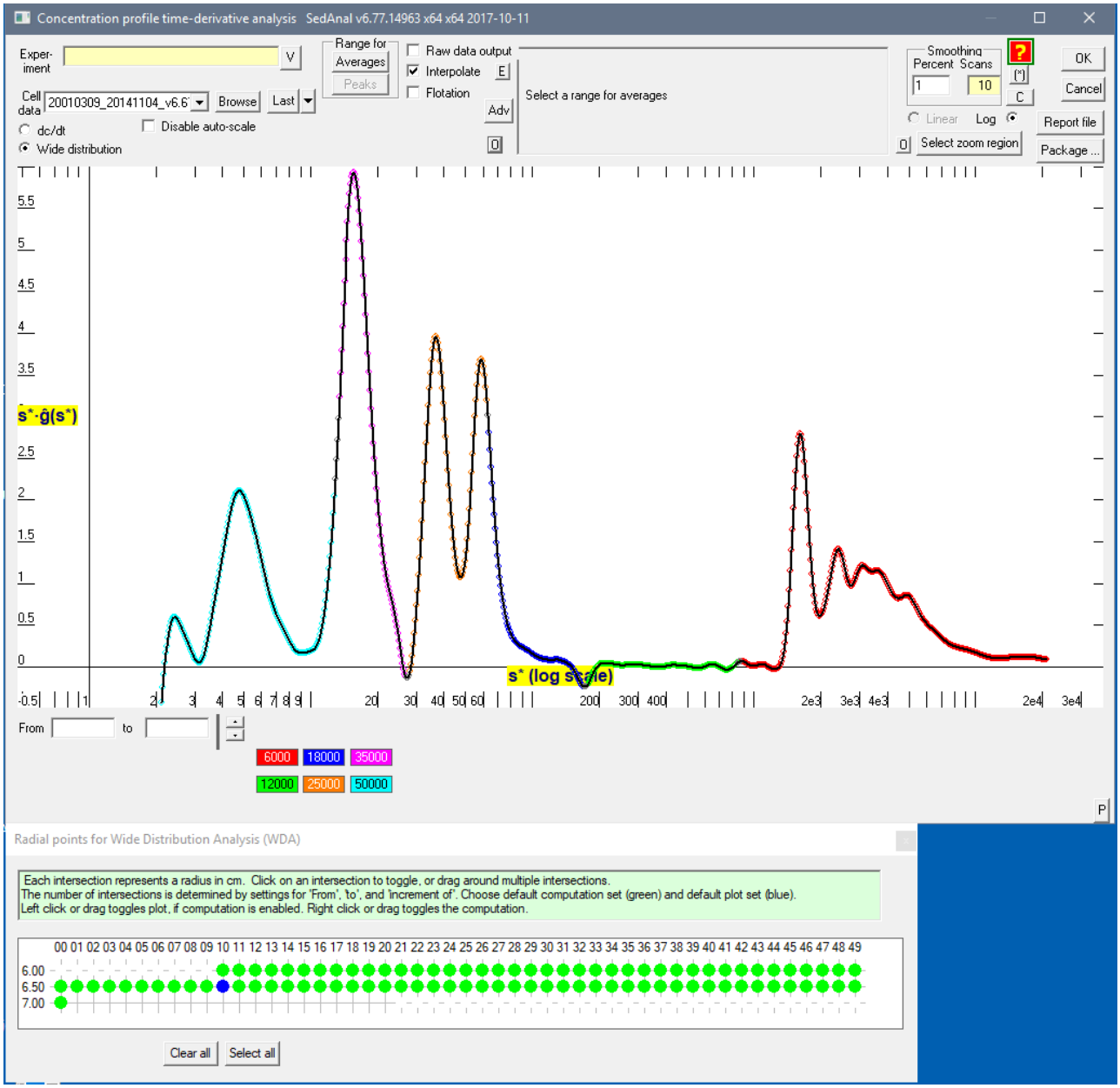


Figure 3-137. Results of WD analysis.

This was a multispeed run at the speeds shown above: 6000, 12000, 18000, 25000, 35000, and 50000 RPM indicated by the different colors on the plot. For this run a range in s of 2S to 20,000S is displayed in the plot. (This was a mixture of hemacyanin and polystyrene beads.)

If more than one radius is selected, WDA curves from all radii are plotted and the average of the plots from all radii is shown as a solid black line. By clicking on the little “O”, one can toggle between, (1) just the individual plots, (2) just the average, or (3) both sets of plots. The s*g(s*) vs ln(s*) are interpolated onto an equally spaced grid of ln(s*) before being smoothed or averaged. Clicking on the "E" will add error bars to the averaged curve (solid black line). **NOTE:** The error bars reflect the standard error of estimate of the average from each radius at each value of s* and are not influenced by the smoothing.

If one unchecks the Interpolate button, only the individual, WDA plots can be displayed by clicking on the little "O".

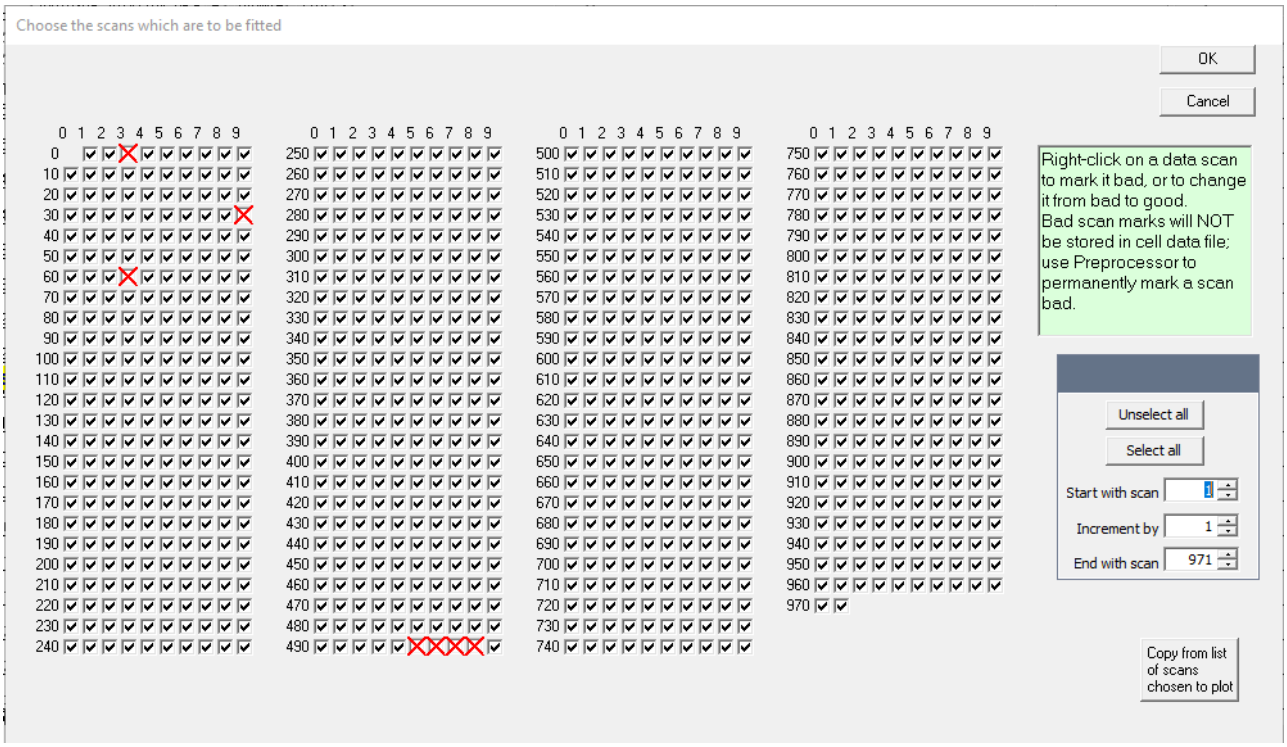


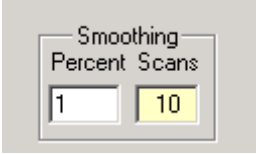
Figure 3-138. Control box for picking bad scans in WDA.

3.5.2.2 Bad Scans

Right-click on the numbers of any bad scans (**Figure 3-138**) (There are 4 in this example: #'s 3, 39, 63, & 495-8) to eliminate them from the analysis. Bad scans can be identified in the preprocessor.

3.5.2.3 Smoothing after numerical differentiation

Some smoothing must be performed on the data since the derivative is computed by taking simple central differences of the c(r,t) vs s*(r,t) data. Smoothing is performed with three passes of a moving box-car filter using a window (i.e. number of points) expressed as a percentage of the total span of the data. This algorithm was chosen because it has very small leakage of high frequency components that are common to many other smoothing algorithms, and it's frequency response can be easily controlled by varying the size of the window (Stafford, 1994).



Control dialog box for choosing the size of the smoothing window is expressed as a percentage of the entire span of the data. In this case, 1% of 1000 s values on the x-axis, uses 10 points on the x-axis.

Note: after entering the percent (i.e. "1" in this case), press enter to have it take effect.

The example given above (Figure 3-133 and **Figure 3-137**) is for a mixture of hemacyanin and polystyrene beads. The data are plotted as a differential distribution function, dc/dln(s*) vs ln(s*), on a logarithmic scale of s*. (N.B. dc/dln(s*) = s*g(s*))

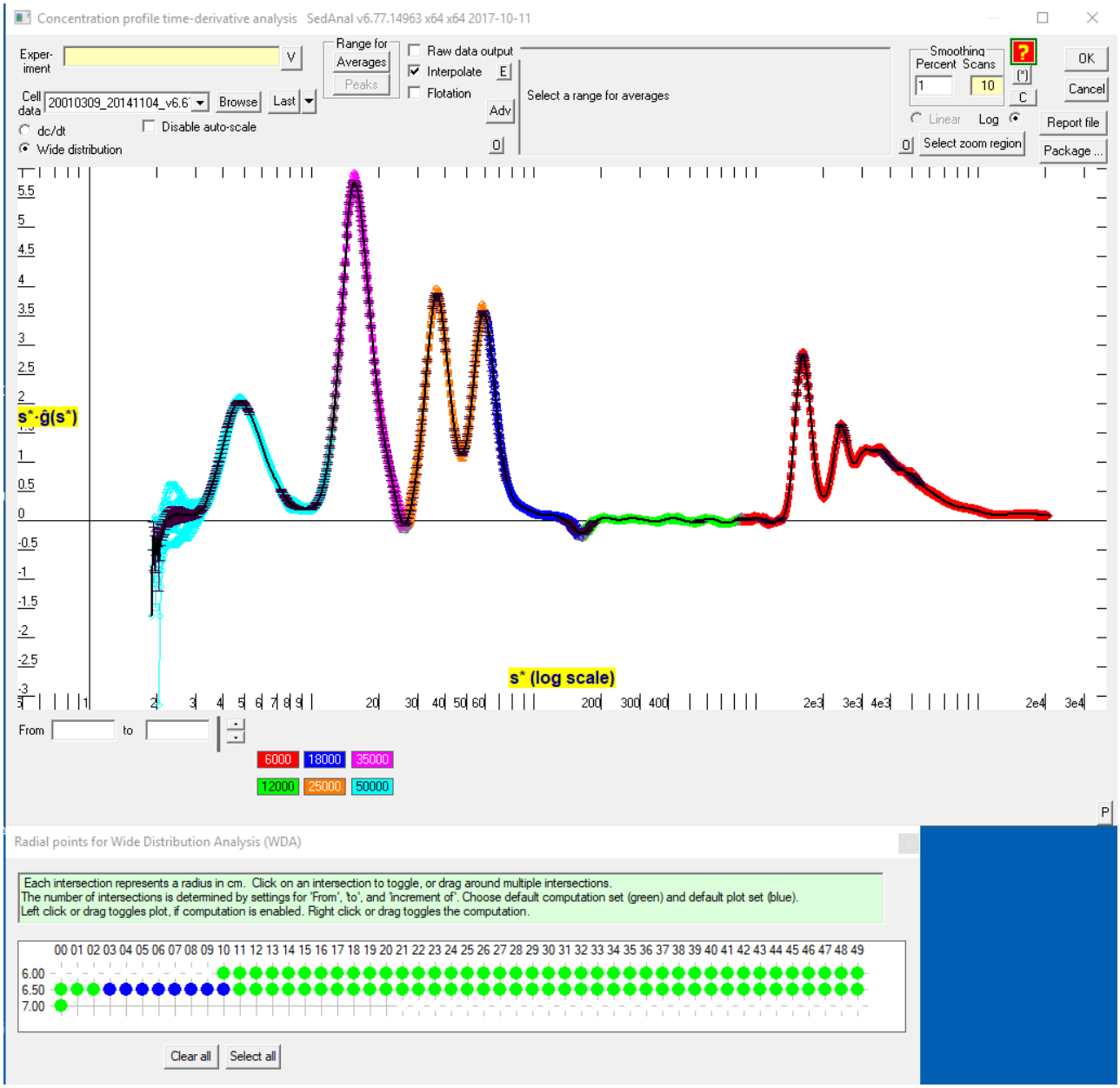


Figure 98. WD analysis for the mixture of hemacyanin with polystyrene beads (Stafford and Braswell, 2004).

Here, plots from radii 6.53, 6.54, 6.55, 6.56, 6.57, 6.58, 6.59, and 6.60 cm are superimposed. The black line is the average of the curves from those individual radii. By clicking on the "E" button, the error bars representing the standard error of the mean are plotted at each averaged point.

Plots from higher radii will have higher resolution than those from a lower radius.

The range of these data is from 2.7 S to over 20,000 S.

The color of the points in the plots corresponds to the speed at which the data were acquired. The color code is given below the plot. and the list of radii are displayed in the radial points window.

3.5.2.4 Adding additional radii for WD Analysis:

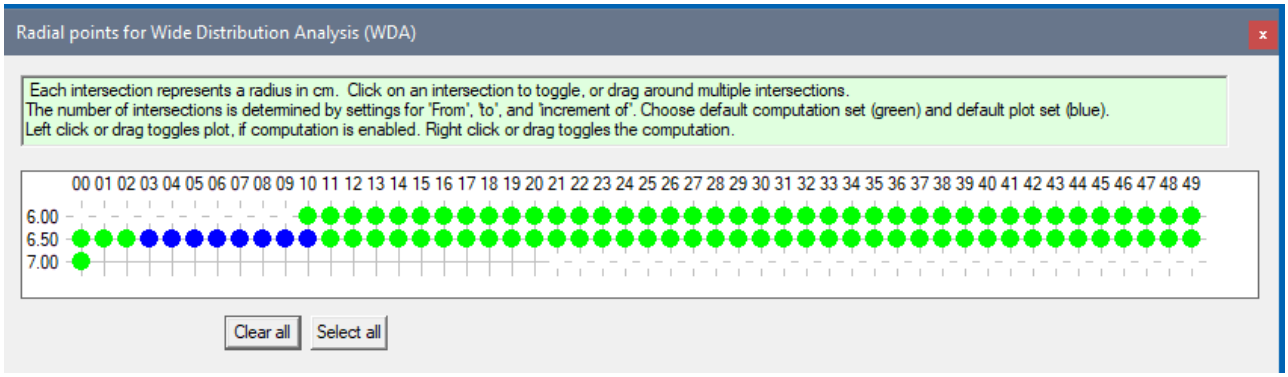


Figure 3-139 Radii available for WD analysis

These values can be changed here and/or in the Preferences -> dcdt/wda**.**

3.5.2.5 Averaging overlayed WDA curves.

The particular WDA curves that are displayed will be averaged over the radii represented by the blue dots and are written to a separate *_WDA-ave.txt output file when the "Write report file" button is clicked. The values of radius corresponding to the green dots will be included in the normal *_WDA.txt output file. The output file names will have the form "Name_of_abr_file_WDA.txt" and "Name_of_abr_file_WDA-ave.txt", respectively.

3.5.3 Flotation

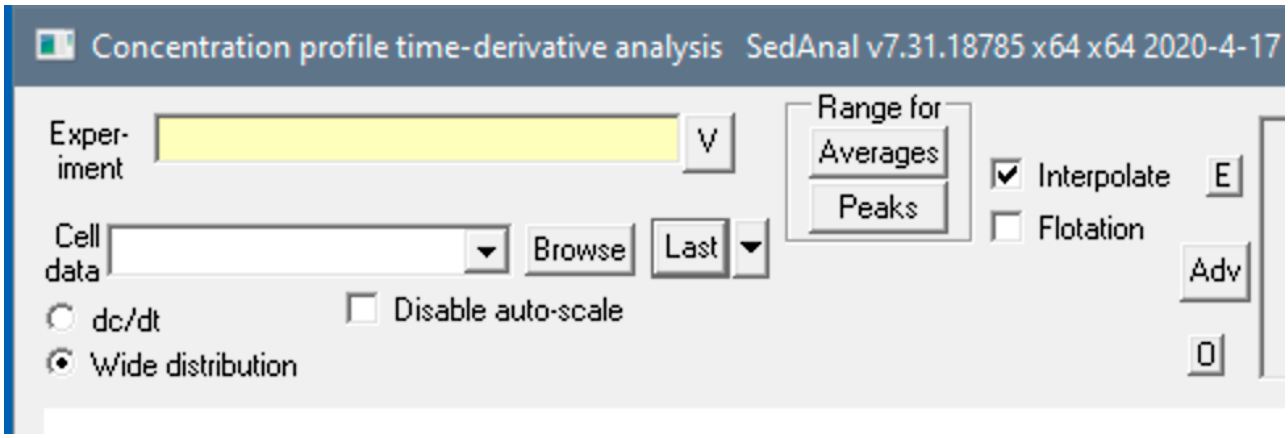


Figure 3-140. Flotation button in DCDT/WD

3.5.4 The Adv button:

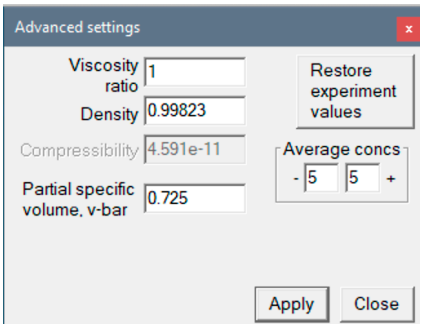


Figure 3-141 The Adv Button

Under the Adv button, we can add correction factors for the sedimentation coefficient axis of the WD plot to convert the values to $s_{(20,w)}$. Also, we can specify the range of radial values at each chosen radial point to average. In the example, shown in Figure 3-141, at each radius selected, WD will select 5 points in either direction and average them with the central point and use that averaged value for the signal at the central point. For example, if the user entered 0 and 0, only the central point would be used. The number of points to se to be used depends on the point density of the scans data, and the amount of noise that is acceptable. The values of 5/5 are suitable for the interreference optical system.

3.5.5 Time Derivative Analysis of Multi-wavelength Data: DCDT and WD.

After a MWL (multi-wavelength) run has been processed in the Preprocessor and stored in a run file (*.abr file), it can be analyzed by either DCDT or WDA. There are several ways in which these large datasets consisting of several hundred scans from several hundred wavelengths can be treated: (1) Scans from individual wavelengths may be analyzed separately by either DCDT or WDA. (2) Spectra of unknown components corresponding to well separated peaks in the DCDT or WDA patterns can be extracted from each peak, (3) Given the extinction spectra for each of the components in a mixture, the data can be deconvoluted in the Preprocessor into a set (abr files) of concentration profiles for each constituent component (Walters, J. et al., 2015 Anal. Chem., 87, 3396-3403).

3.5.5.1 WDA: Extracting Spectra.

After preprocessing and saving the abr file, load it into WDA. At first, select only a few wavelengths to speed the process up. Next select each peak whose spectrum you want to extract by clicking on the "Peaks" button as shown in **Figure 3-142**. This a click-and-drag procedure (**Figure 3-143**) much like selecting the range-to-fit in the preprocessor. Repeat this process for each peak whose spectrum you want to extract. Then click on "Select all" in the Wavelength window (**Figure 3-144**). Now all the absorbance vs wavelength values for each peak will be listed in the "Integration" window. The spectra can be written to spectrum files as shown in Figure 3-147**.**

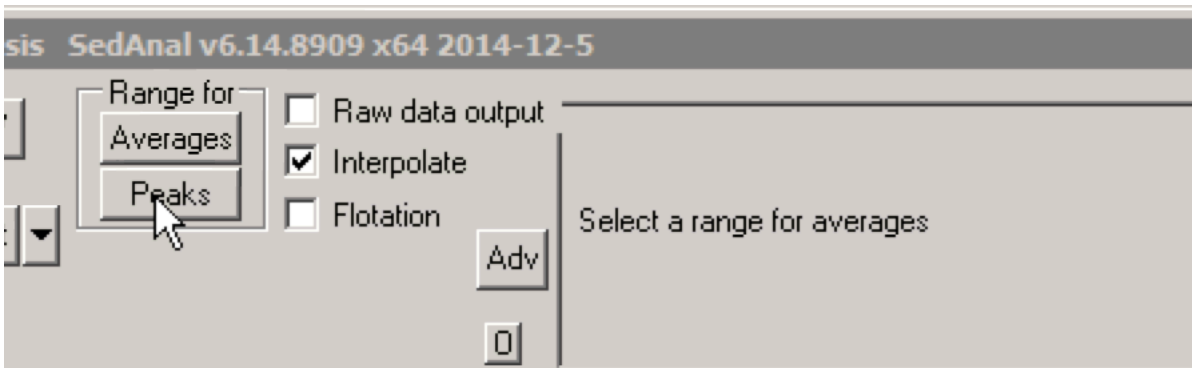


Figure 3-142. Selecting peaks for spectrum extraction:

Click on the "Peaks" button, and click-and drag to select the peak.

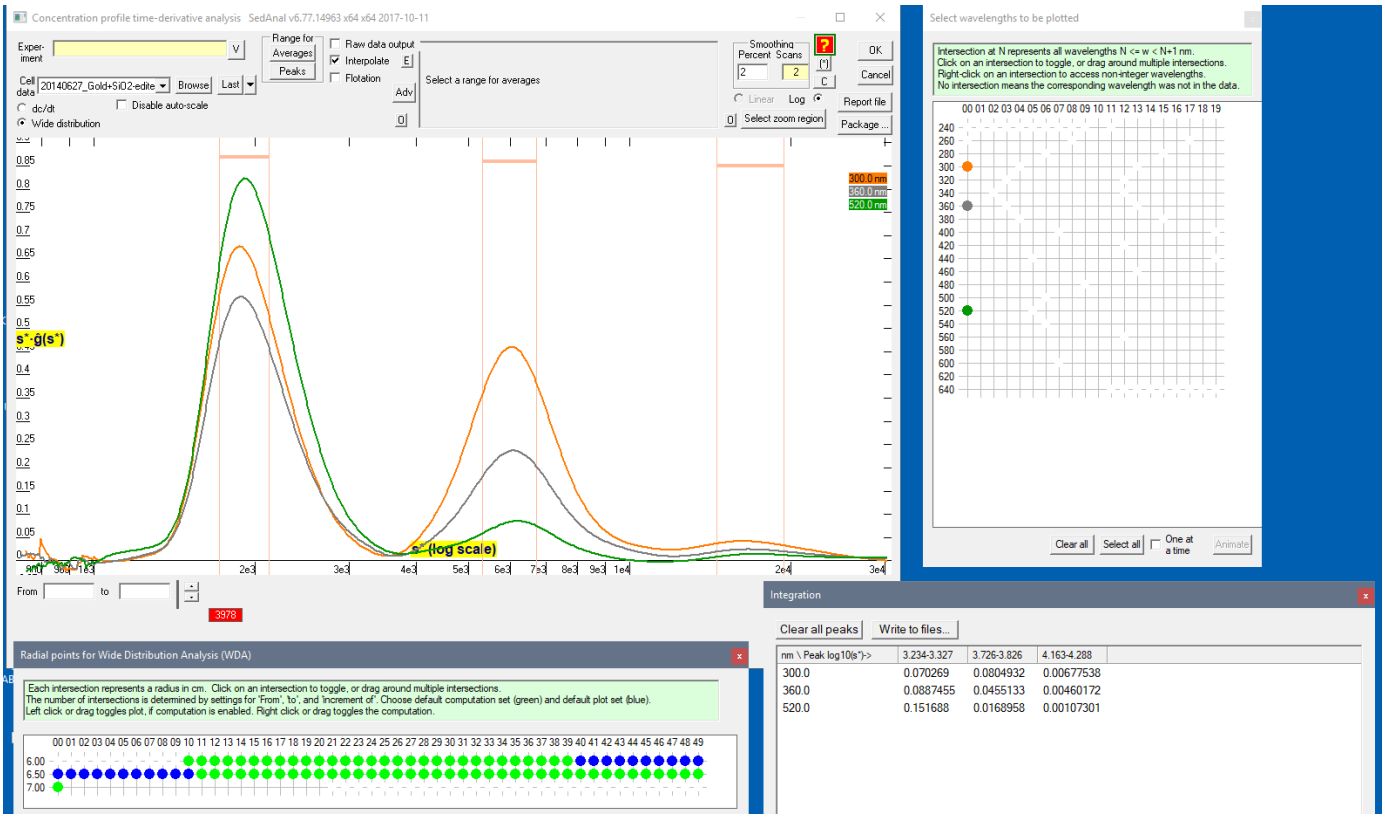


Figure 3-143 Three peaks selected for only three wavelengths.

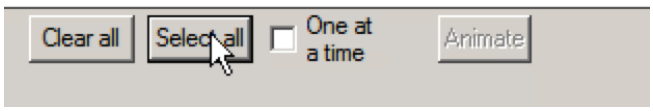


Figure 3-144 Click the "Select all" button to plot all the WD curves for all wavelengths.

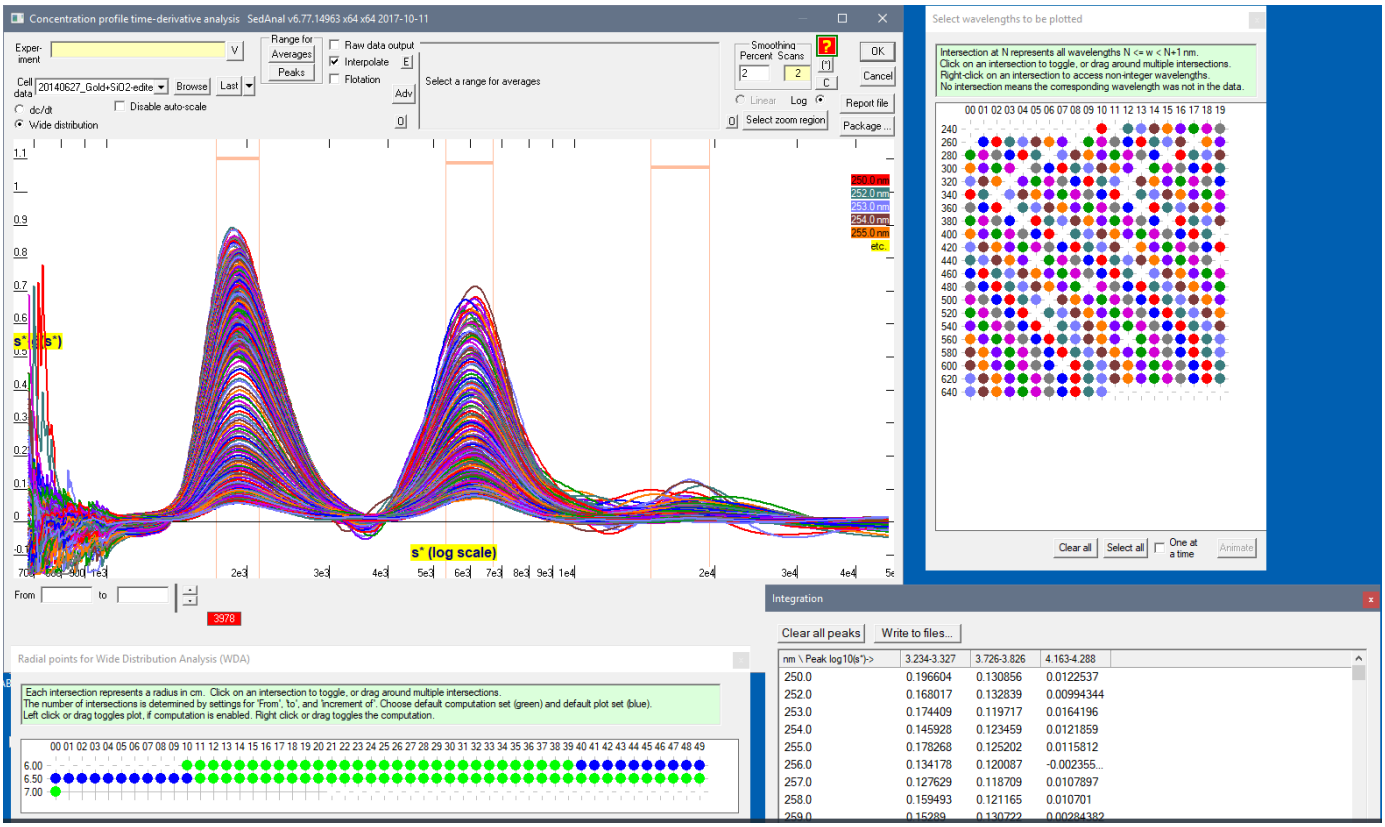


Figure 3-145. WD curves for all wavelengths.

At this point, you will want to save these spectra to a file for each peak:

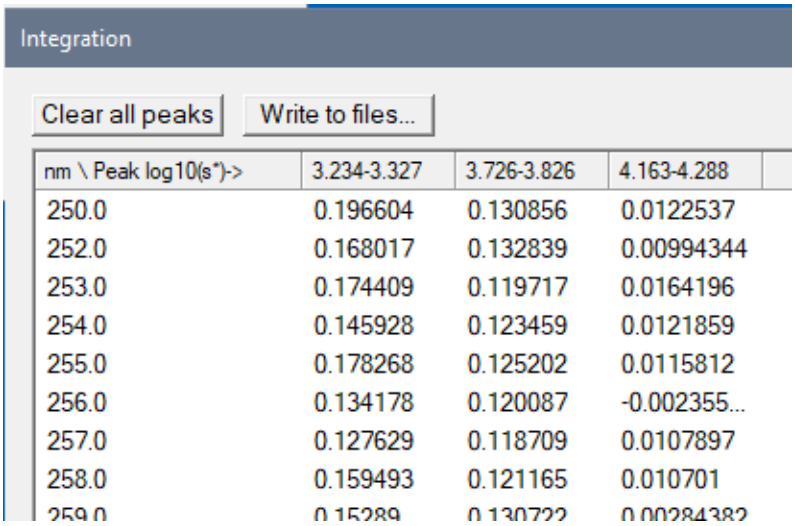


Figure 3-146 Spectra gathered from the WD plot

Click on "Write to files ..."; the following window will appear

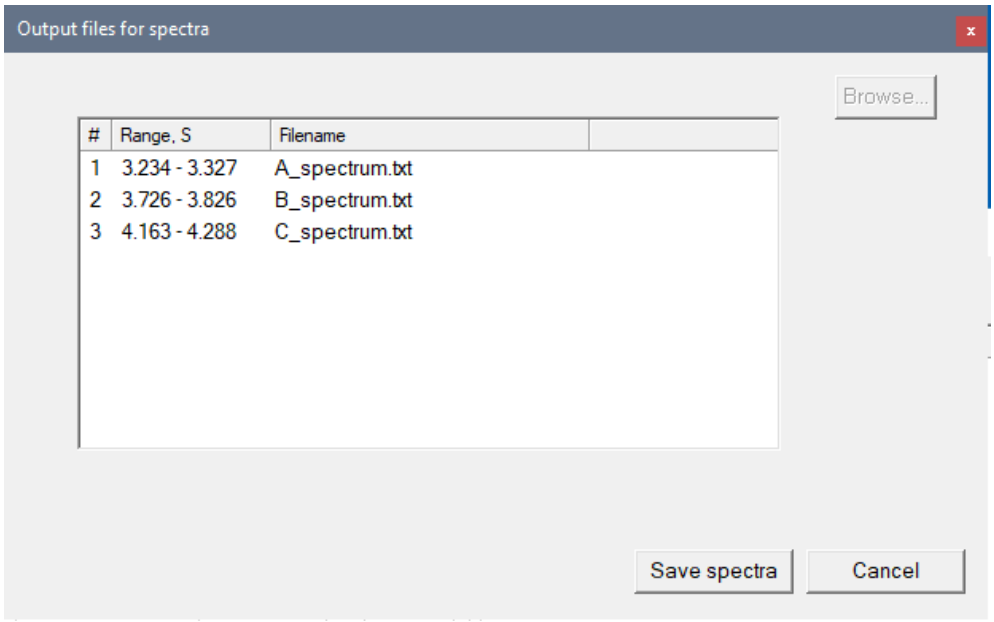


Figure 3-147 Spectrum output dialog box

The default file names can be renamed:(Figure 3-148)

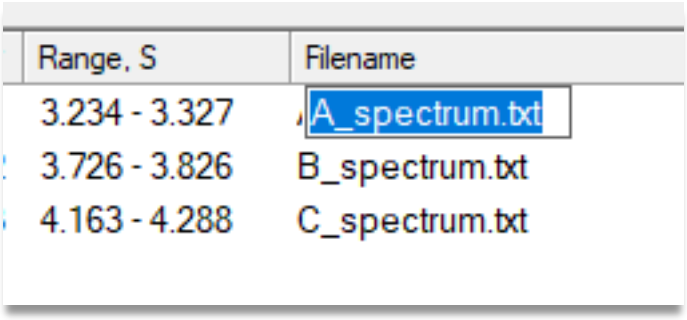


Figure 3-148 Specifying a filename for spectrum files

After renaming the files, click on "Save spectra" to write out the spectra to individual files

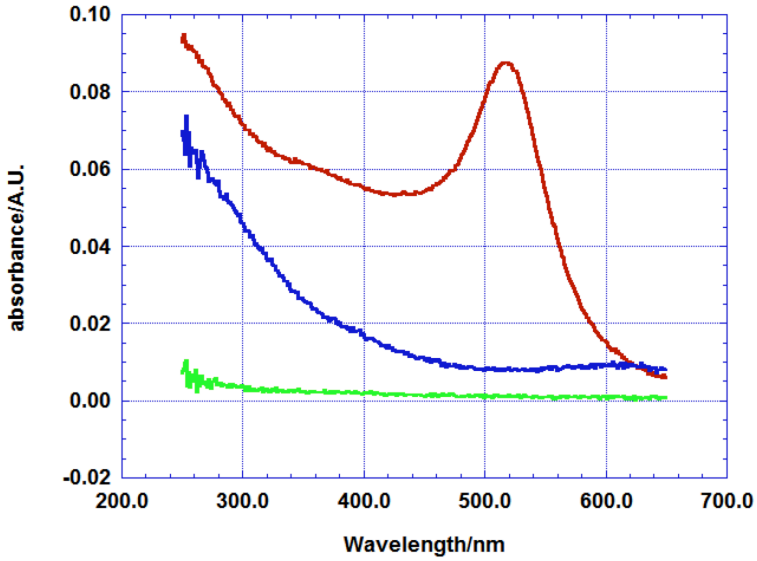


Figure 3-149. Spectra of the three peaks selected from WD analysis. Plotted in a separate plotting program.

3.5.5.2 Least Squares deconvolution of component (constituent) concentration profiles.

In the Preprocessor, after loading a multi-wavelength dataset file, the individual constituent concentration profiles can be deconvoluted into a separate run file for each constituent component.

After loading the data and selecting the meniscus, base, and range-to-fit, etc. We can enter paths to the extinction spectra and click on "Deconvolve" button and see the green or blue progress bar indicating that the deconvolution is proceeding (Figure 3-150):

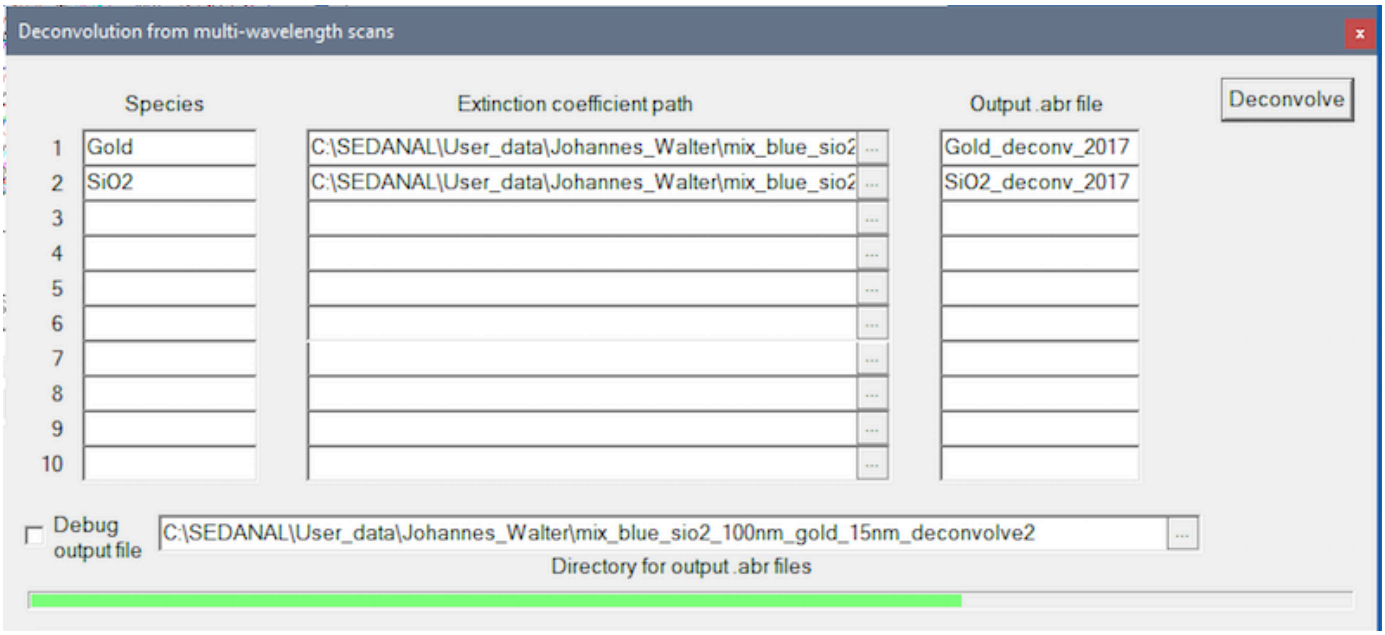


Figure 3-150 Deconvolution of constituent concentrations.

3.5.5.3 Multi-wavelength Fluoresence Intensity Data

When Multi-wavelength intensity data are loaded into the preprocessor, the scans will appear as inverted pseudo-absorbance. The sign of the data can be changed by clicking on the "F" button that is located in the upper righthand region of the Preprocessor window - as shown:

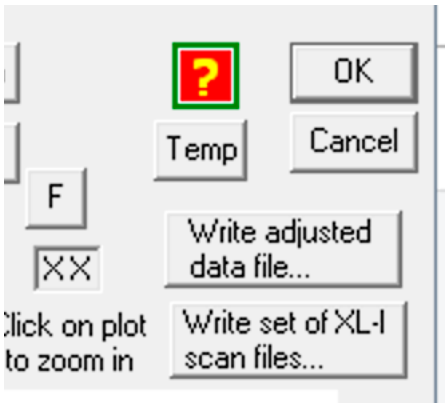


Figure 3-151 Location of the "F" button

From this point on, the scans are treated as pseudo-absorbance scans and saved in the abr file as positive scans.

3.5.5.4 Deconvoluting Multi-wavelength Fluoresence Intensity Data

After the data have been "flipped", the deconvolution process is the same as for ordinary absorbance data. See above in paragraph 3.5.5.2

3.6 Synthetic Boundary and Band Sedimentation

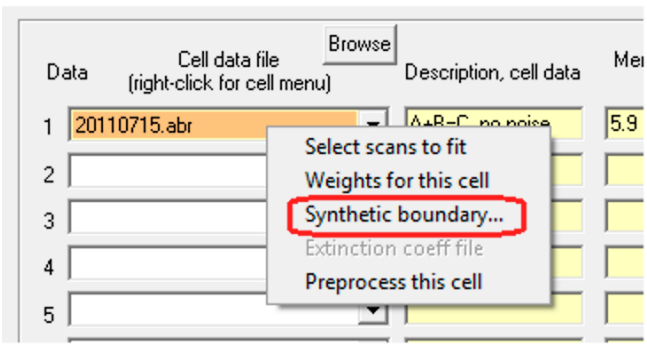
3.6.1 Fitting or simulating synthetic boundary experiments

3.6.1.1 Fitting

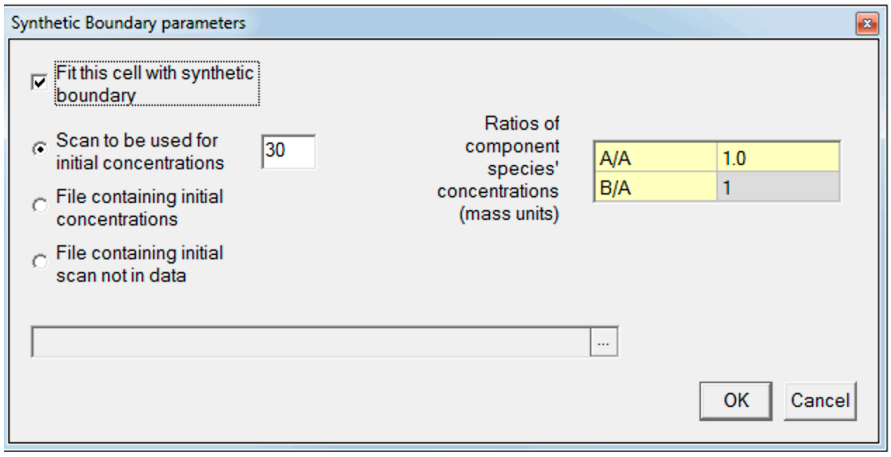
Each run may be treated as a synthetic boundary experiment or a normal sedimentation velocity fit.

From your set of scans, choose the one that is to be used for the initial concentrations (the *I-scan*). Typically, it will be the first scan, but it may be later. This must be done separately for each run.

Several species can be accommodated, since the concentration is calculated from the signal, the extinction coefficient and the signal weighted ratios of the species.



To identify a cell as a synthetic boundary or band sedimentation run, on the control screen, right click on the cell data file to get the context menu. Click "Synthetic boundary"... to bring up the Synthetic boundary parameters window.



Check the box, and select one of the two locations from which the initial concentration profile will be obtained, either from a scan from the cell data, or from a file containing a concentration profile.

For fitting cell data (normal sedimentation velocity fit), you can choose the number of the I-scan. For simulating, this option is not available. For an I-scan that is not in the set being analyzed, use the third option, and browse to the I-scan file.

The coloration of the B/A field shows that the ratio is to be fitted, and 1.0 is just the initial guess.

The format of the file containing a concentration profile is the same for single-species and multi-component SB simulations or fits, and is given below.

Remove the I-scan and any prior scans from scans to fit. This must be done separately for each run.

Each time the grid or parameters change, the I-scan is interpolated onto the simulation grid, and the concentrations for species 1 calculated as the interpolated signal / extinction coefficient for species 1 for the run. These calculated concentrations are used as the initial concentration profile for species 1 for the Claverie FEM simulation of scans.

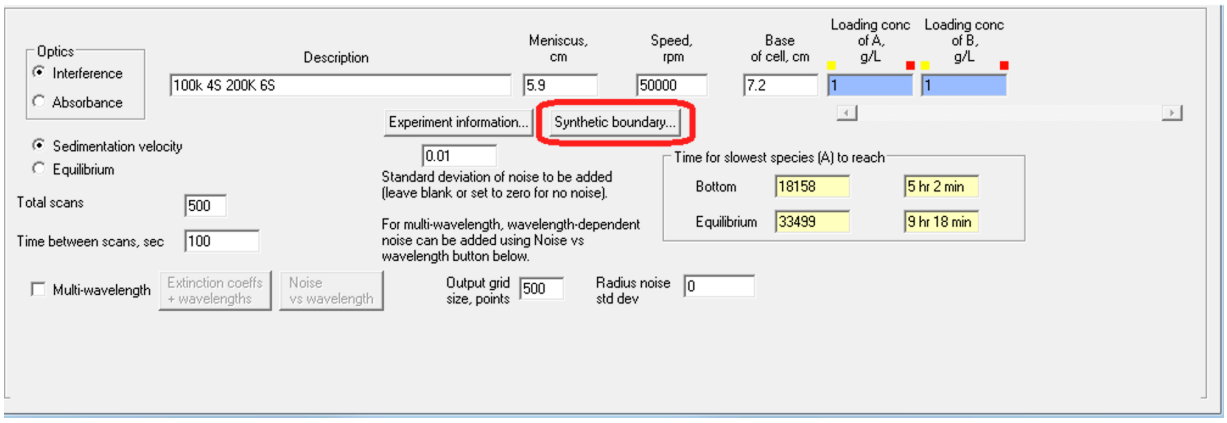
The "simulation clock" is started at the time of the I-scan, rather than at 0, to make the simulated scans appear at the same time relative to the I-scan. For example, if the I-scan was taken at 40 sec, and the next scan was taken at 60 sec, the Claverie simulation will run for 60−40 = 20 sec to produce the simulated scan to be compared with the second experimental scan.

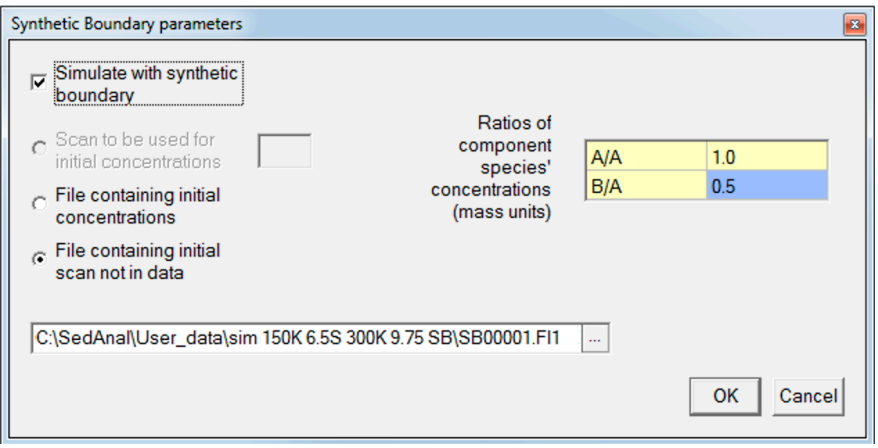
Because the loading concentration does not affect the fit directly, you should not attempt to fit for it.

The report will show [SB] after the load concentration in the **Parameters** section for a synthetic boundary cell.

3.6.1.2 Simulating

To identify a simulation as a synthetic boundary run, on the control screen click Synthetic boundary...





Check the box, and identify a file containing a either an I-scan, or a concentration profile (the Scan to be used choice will be grayed, as it is not available for simulations).

The coloration of the B/A ratio shows that the ratio is fixed at 0.5.

You can either type or paste the path into the box, or browse using the ... button.

3.6.1.2.1 Concentration file format

A concentration file contains the mass concentrations of all *nc* component species at each of *nr* radii. The number of radii and their values need not correspond to either the experimental data or the simulation grid; values will be linearly interpolated as needed. This means, for example, that a step function at *r*=6 cm could be described by points at meniscus, 5.999, 6.001, and base.

```
SEDANAL concentrations v1   // [optional comment]
radius1   c1   c2   ...   cnc
radius2   c1   c2   ...   cnc
.
.
radiusnr   c1   c2   ...   cnc
```

The initial line in the file must be exactly as shown, except that the "//" is only needed if there is a comment following. Any number of spaces may separate values.

Concentrations of the species which are not component species (if any) will be computed from the mass-action equilibria (reactions). Note that the component concentrations are *not* the "total component", but the equilibrium concentrations of the species which have been chosen as components.

Here is an example of a concentration file which has a linear gradient of component 1 from 0 g/L at meniscus=5.9 cm to 5 g/L at the base=7.2 cm, and a 1.5 g/L pulse of component 2 between 6.0 and 6.1 cm.

```
SEDANAL concentrations v1   // illustration of gradient + pulse
5.9     0.0     0.0
5.9999  0.3846  0.0
6.0001  0.3846  1.5
6.0999  0.7692  1.5
6.1001  0.7692  0
7.2     5.0     0.0
```

3.7 PREFERENCES

The Preferences screen allows one to choose the names of the directories used by SEDANAL, the default number of scans to show in the initial plot in the Preprocessor, and the maximum number of control files (*.abc) that are shown in the "Last" drop-down list on the Control Screen (Figure 3-152, and Figure 3-153). Also, various default and other parameters can be set for other SEDANAL functions like DCDT, BIOSPIN, the control screen, etc. …

3.7.1 General Preferences

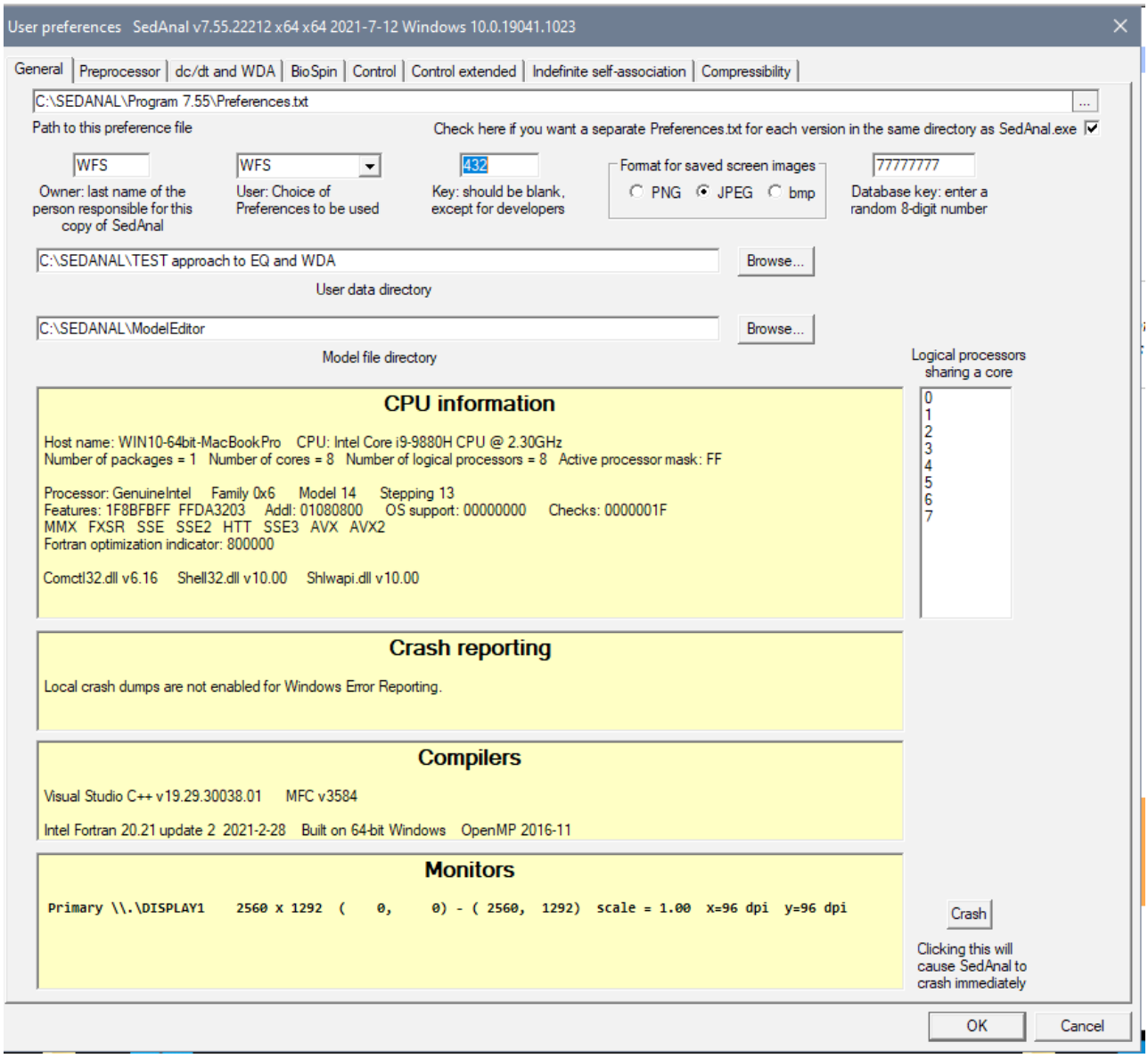


Figure 3-152 Setting User_data and ModelEditor files paths.

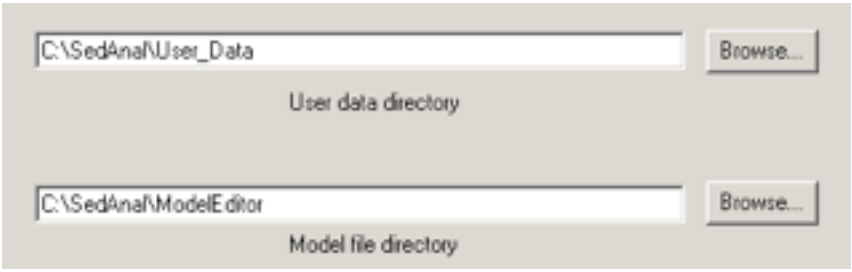
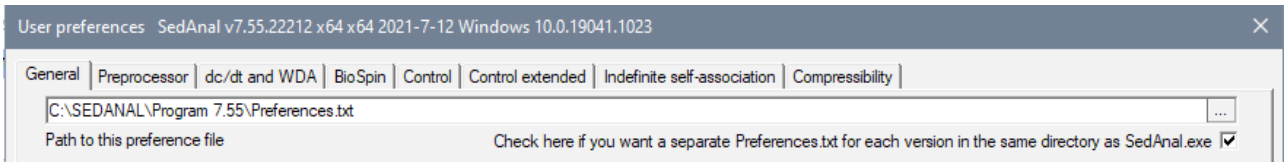


Figure 3-153. Browse to set new paths.

3.7.1.1 Location of Preferences.txt files

When SEDANAL is started, the preferences file is read, and unrecognized lines pop up a message ("The preference file line *nn* begins with the unrecognized symbol..."). This message has been expanded to show the complete path to the file being read. Unrecognized lines usually come from reading a file using an earlier version of SEDANAL than the version that created it.

Users now have more flexibility in setting the location of their preference files by choosing the path to the preferences file to be used:

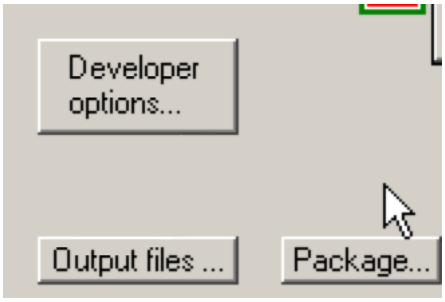


3.7.1.2 Multiple users and multiple Preference files:

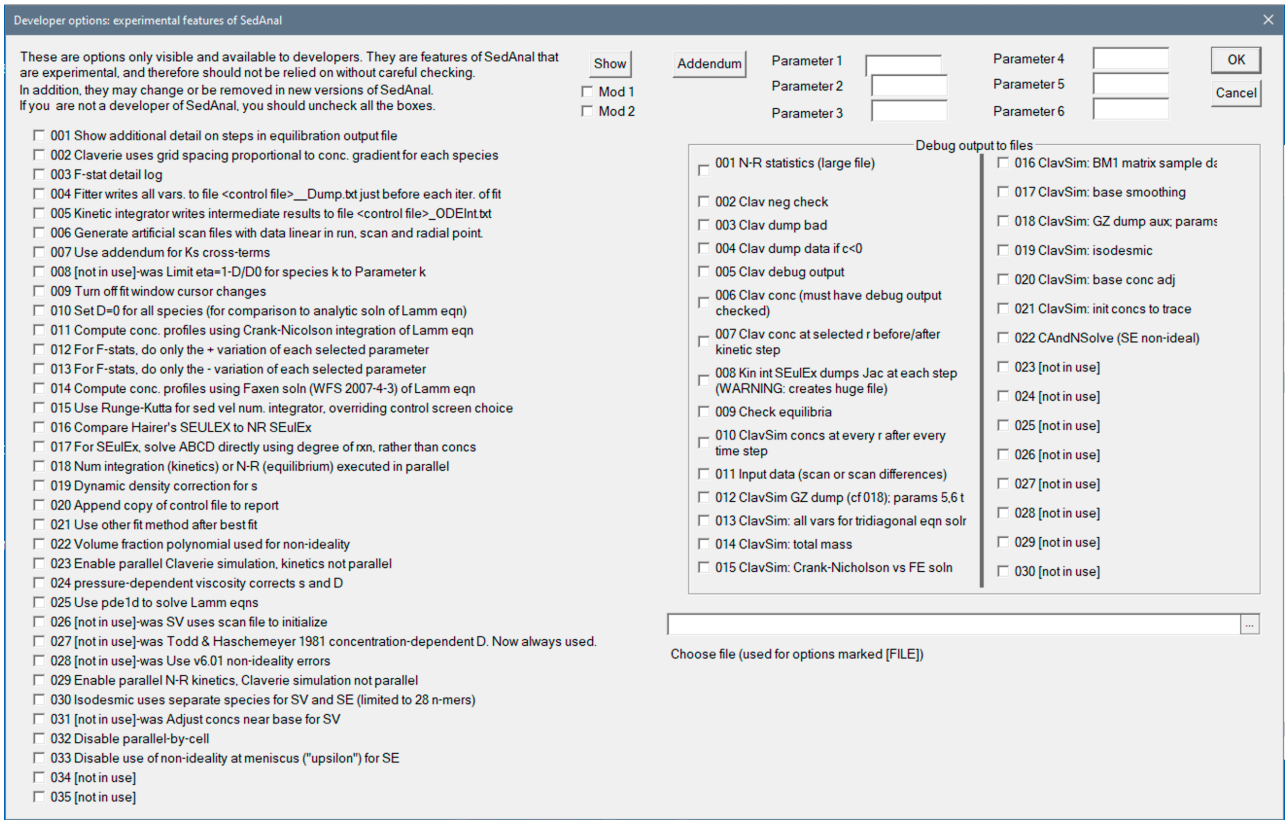
3.7.1.3 Developer Options

Developer Options can be activated by entering “432” in the “Key: should be blank except for developers” window.

On the Fitting screen, the following button will appear if Developer Options are activated:



When the Developer Options button is pressed, the following rather daunting window will appear. You are now looking at many of the secret, deep inner workings of SEDANAL. Many of the items on this list are experimental, some are defunct but several are useful and after a little more testing will be available on the other screens.



For example, items 012 and 013 are useful for doing the positive and negative sides of an F-stat search separately. Also item 019 allows SEDANAL to handle dynamic density distributions, ρ(r,t), calculated by adding up the density contributions from each of the species present in a model.

$$\rho(r,t) = \sum_{i=1}^{i=N} \left(\frac{\partial \rho_i}{\partial c_i} \right)_{\mu_{j\neq i},T,P} c_i(r,t)$$

cf. Schuck P (2004) A model for sedimentation in inhomogeneous media. I. Dynamic density gradients from sedimenting co-solutes. Biophys Chem 108(1–3):187–200.

3.7.2 Preprocessor preferences

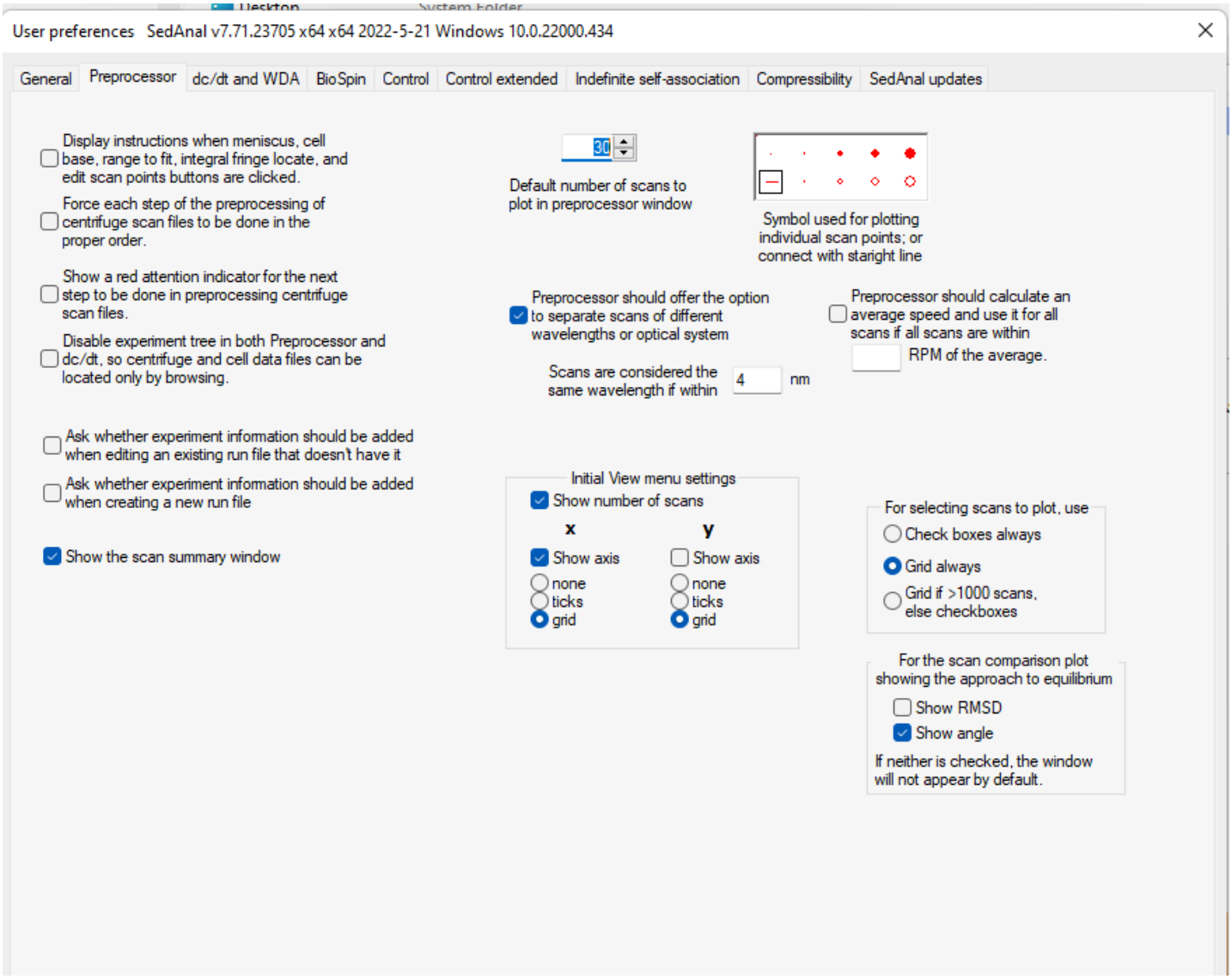
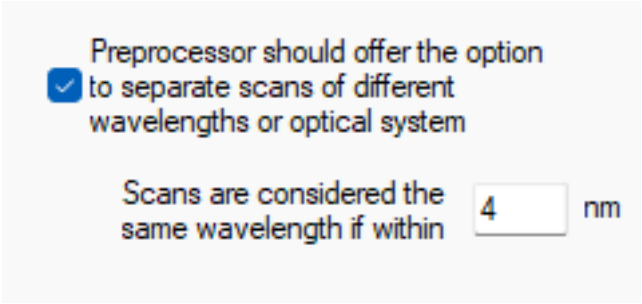


Figure 3-154. After you have used SEDANAL a few times, the top three items on the left hand side should be unchecked for genreal use. They can be checked for beginners to force the work flow in the PP.

The middle section:



allows the user to separate scans based on wavelength. The XL-A has some variation in the wavelength when a single wavelength has been chosen. In the example above, if the variation is less than or equal to 4 nm, SEDANAL will treat the scans as if they were from a single wavelength. And if the difference is greater than 4 nm, the scans are treated as coming from different wavelengths as a separate set of scans.

3.7.3 DCDT and WDA preferences

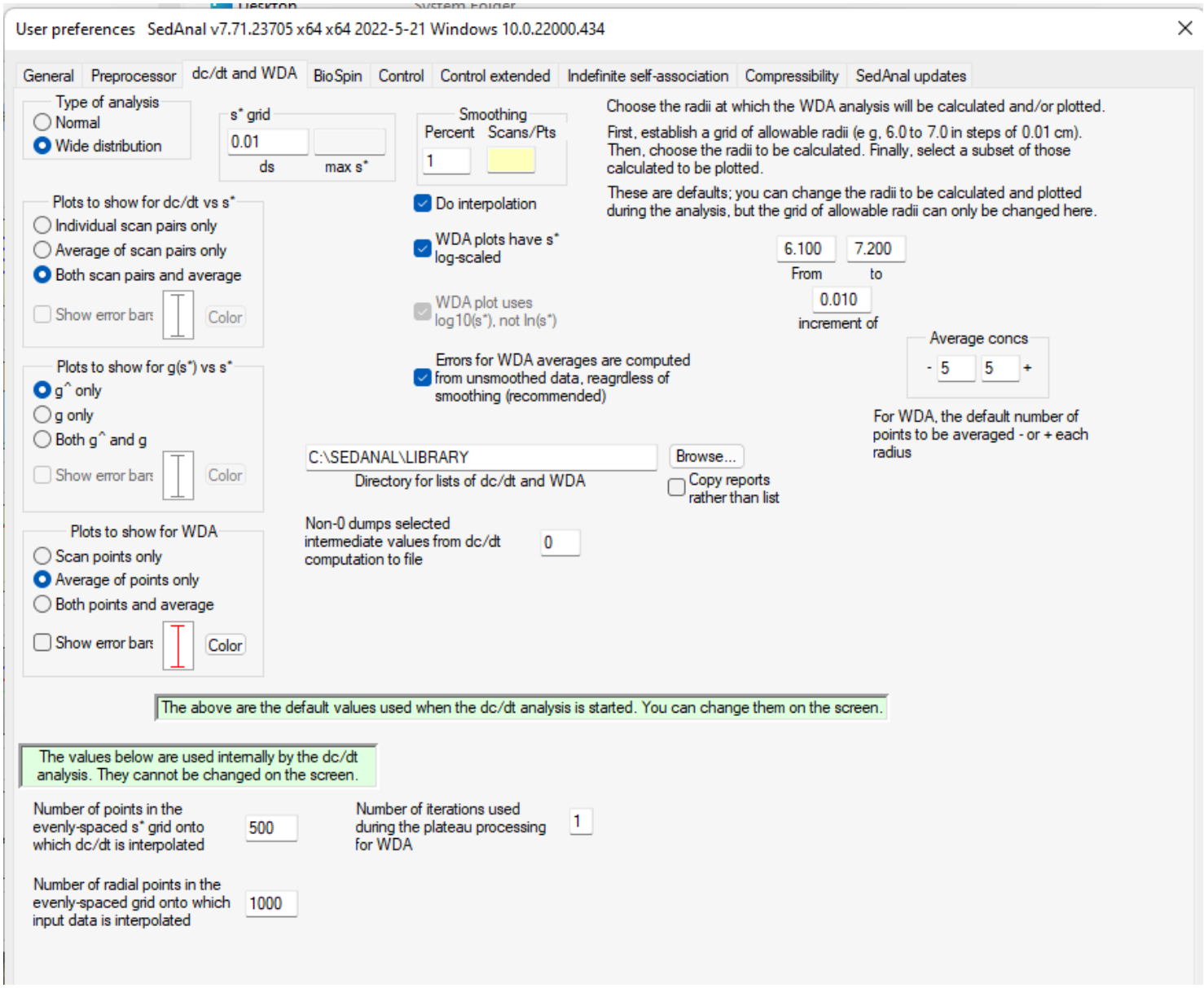
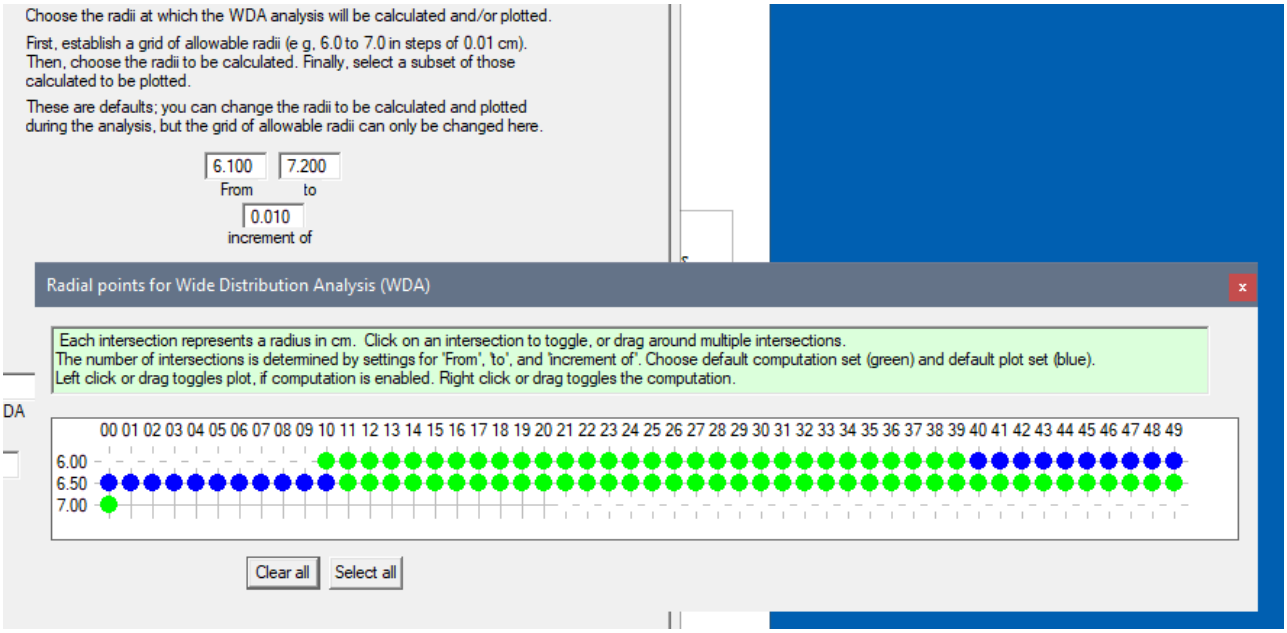


Figure 3-155. Setting parameters for dc/dt and WDA.

Setting parameters for dc/dt and WDA. It is recommended that a default value of smoothing percent = 2 be used for the initial look at WDA data. Smoothing of WDA is required because the derivative is computed using simple centered divided first differences. For DCDT, set it to zero on the DCDT screen for the initial look. Smoothing can obscure outliers. You should be aware that a single outlier (i.e. a delta function), with enough smoothing, will be transformed into a gaussian shape (review your course on convolution.) You don't want that to happen.

The bottom window with green and blues dots is for selecting the radii to be used in the WD analysis. The three little windows in the upper right corner are used to establish the radial grid that will be used by WDA. Those values determine which of the radii in the lower window are represented by crosses and are those radii that can be used for WD analysis. The green dots, which are selected by right-click-n-dragging over the intersection crosses, represent those radii that are used in the computation. Those results are written to the report file ending in "_WDA.txt"; similarly, the blue dots represent those radii that are used for the average WDA curve and are displayed in the screen. The blue dot results are written a report file ending in "_WDA-avg.txt"



BIOSPIN preferences

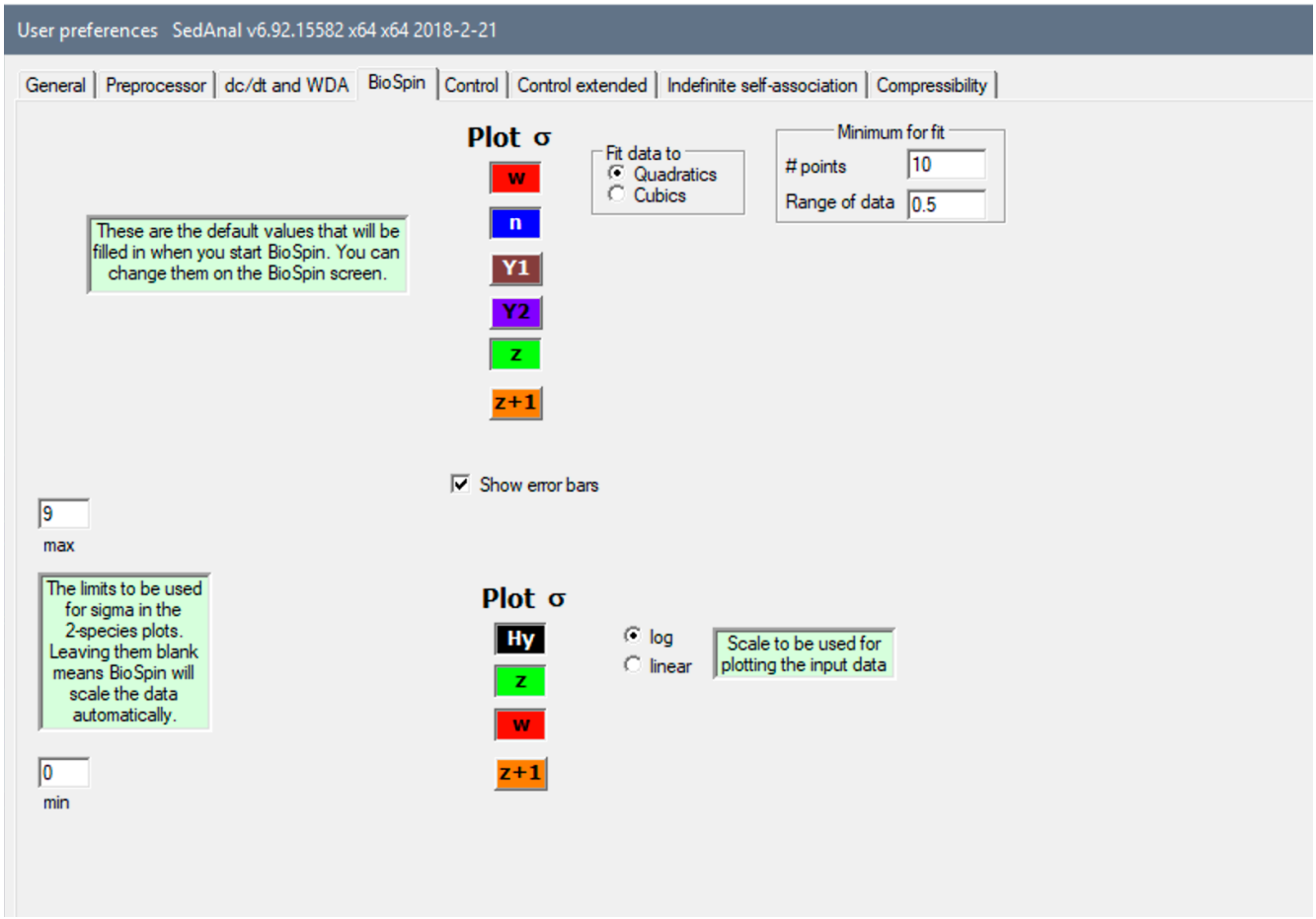


Figure 3-156. Preferred settings for BIOSPIN

3.7.4 Weighting Factor preferences - "Control" Tab

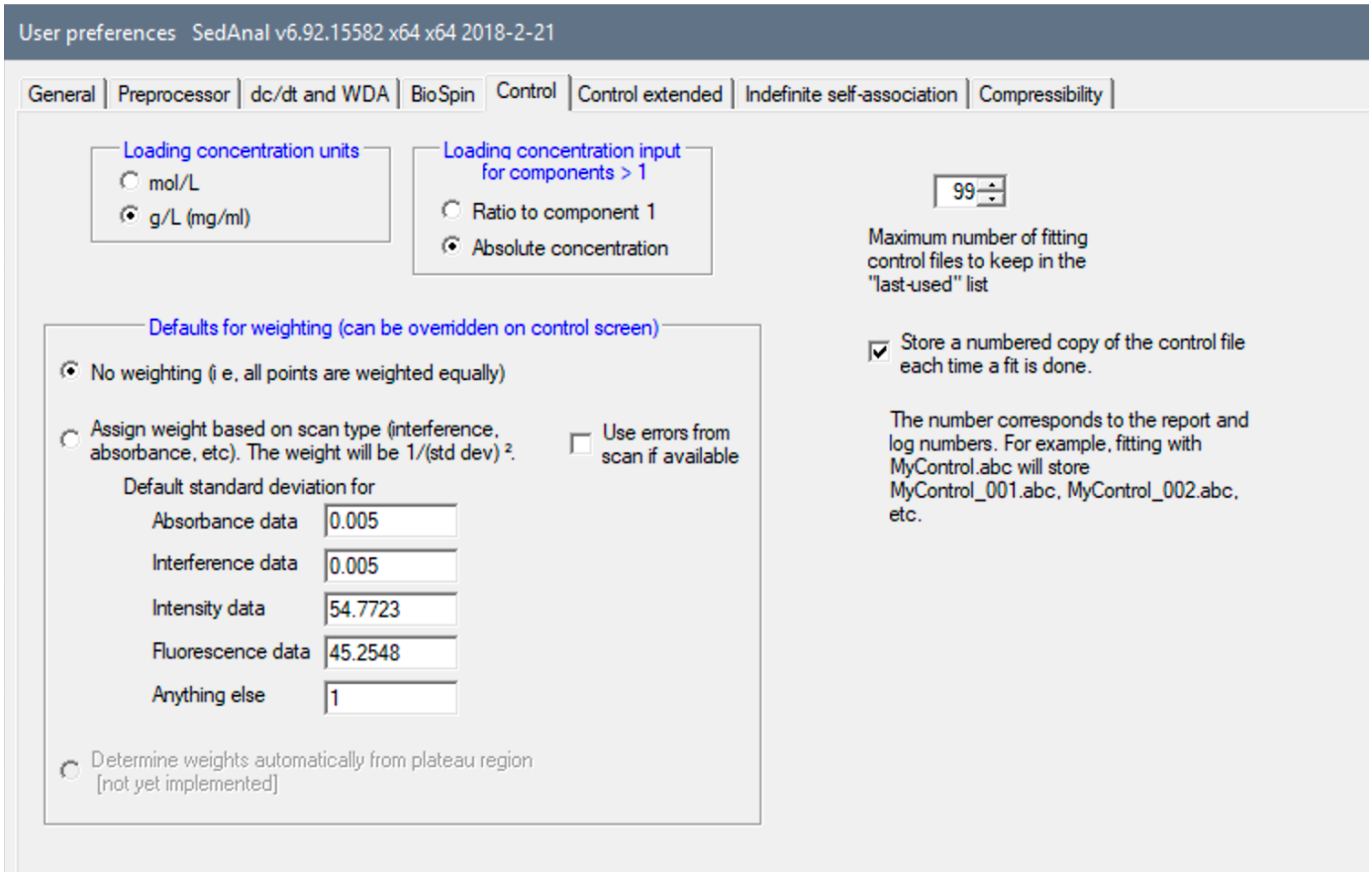


Figure 3-157. Parameters for the control screen: setting defaults for use of weighting factors.

Parameters for the control screen (**Figure 3-157**): setting defaults for use of weighting factors. In the preference be sure to **un-click the box labeled "Use errors from scan if available" as a default**. This can be selected later after you have loaded a run file in the Fitter and you want to use the errors in the third column of the absorbance scan file. The values shown above correspond to values of the inverse variance ($1/\sigma^2$) typically observed for the various types of data: you should change these values according to your data type and quality. (Do not use values from the "plateau region" that was an idea that was never implemented.)

In the case shown here, using fluorescence data from the FDS machine, the noise (i.e. standard deviation of the noise) on the data was +\- 0.5397 fluorescence units. The inverse square of that is 3.433 and so that value is used as a weighting factor for that dataset. If these data had been combined with absorbance data whose noise is about +/- 0.006, a value of 27778 (=1.0/(0.006*0.006) would be entered for that dataset.

NOTE:

But since version 6.92, the user can enter the standard deviation and SEDANAL will compute the inverse variance for you.

3.7.5 Confidence Limits -- Error estimation control: Advanced, Control Extended tab

NOTE: The Advanced tab can be accessed also from the Fitting Control screen. Choices made there will apply only to the control file under consideration. Choices made here in Preferences will become defaults but may be changed from the control screen for a particular run.

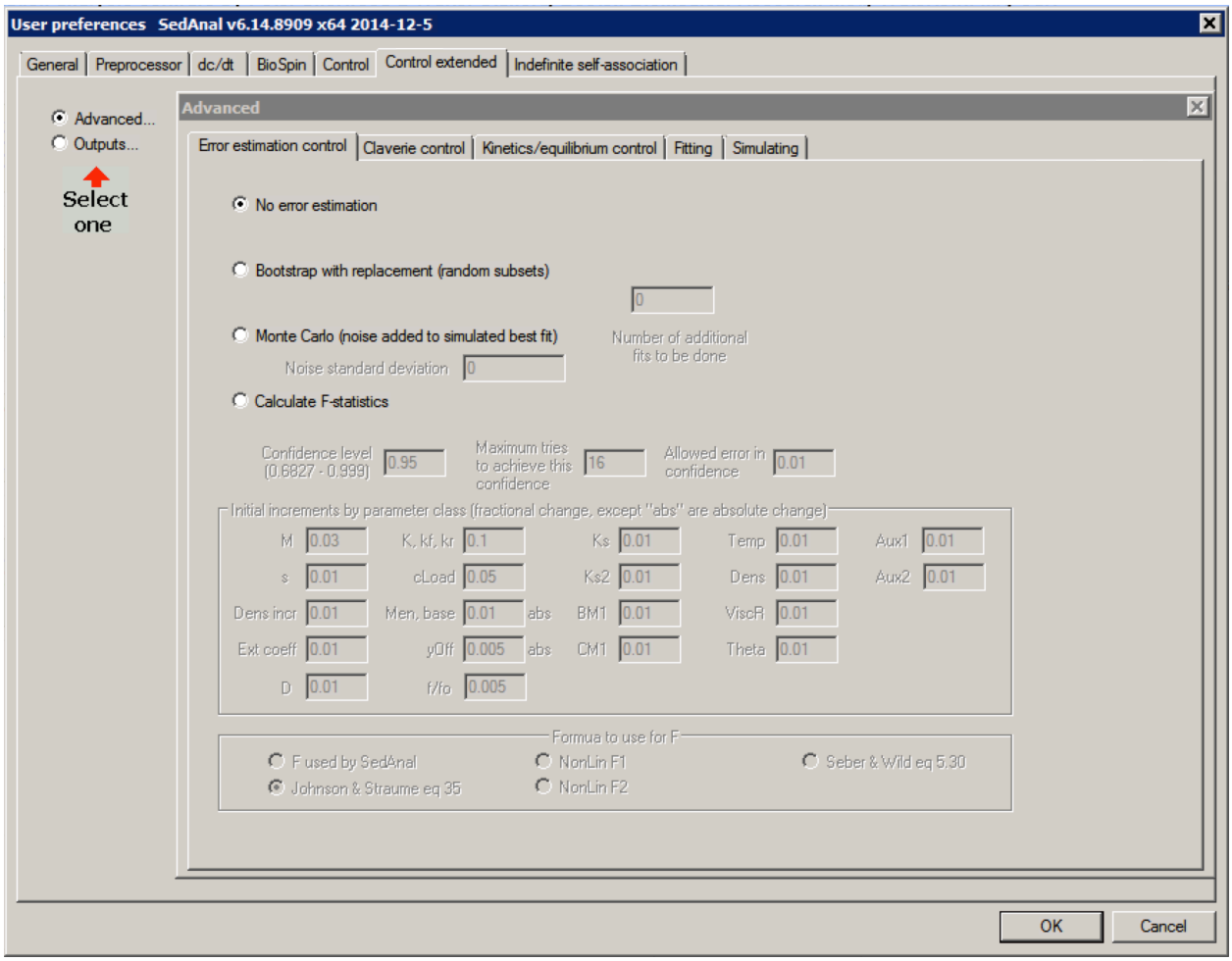


Figure 3-158. "Control extended" parameters allow setting of the default parameters found under the "Advanced…" button on the control screen.

3.7.6 "Claverie Control".

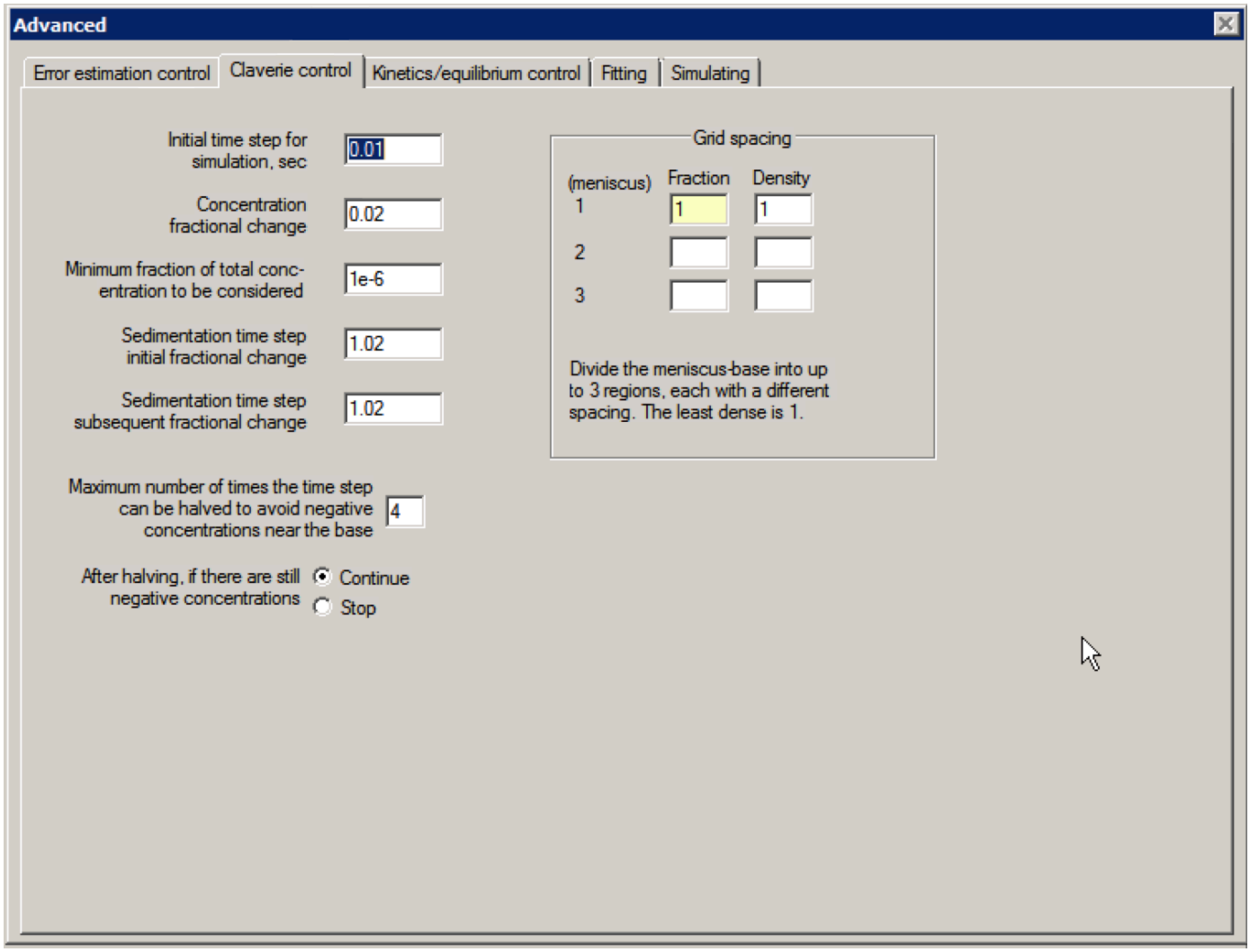


Figure 3-159 Setting Finite Element controls.

3.7.7 Kinetics/equilibrium Controls.

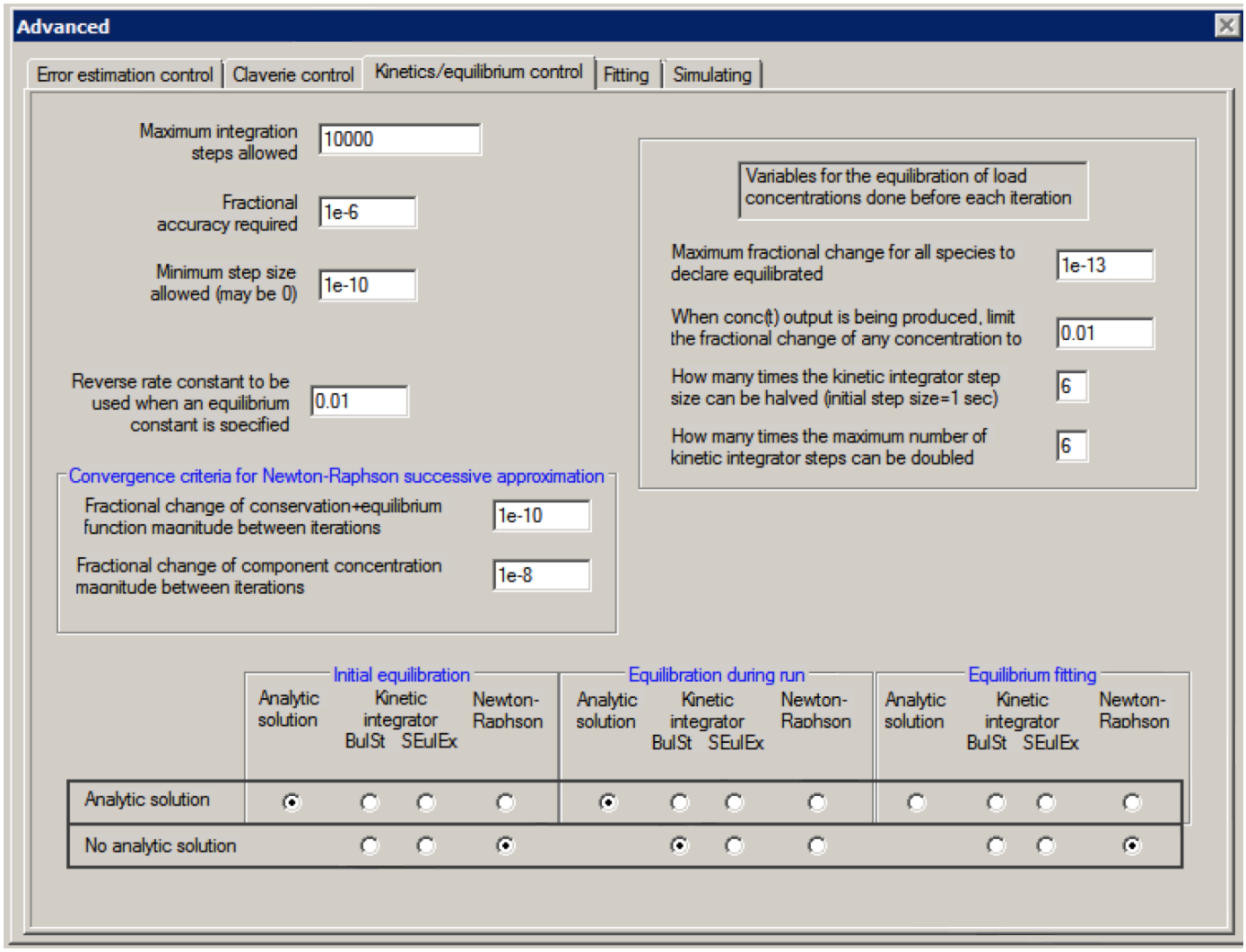


Figure 3-160. Setting for kinetics and slowly equilibrating systems.

Settings for a system whose re-equilibration is kinetically limited, choose either Kinetic integrator: BulSt or SEulEx, may be used. For any particular problem, one will be faster than the other: try both.

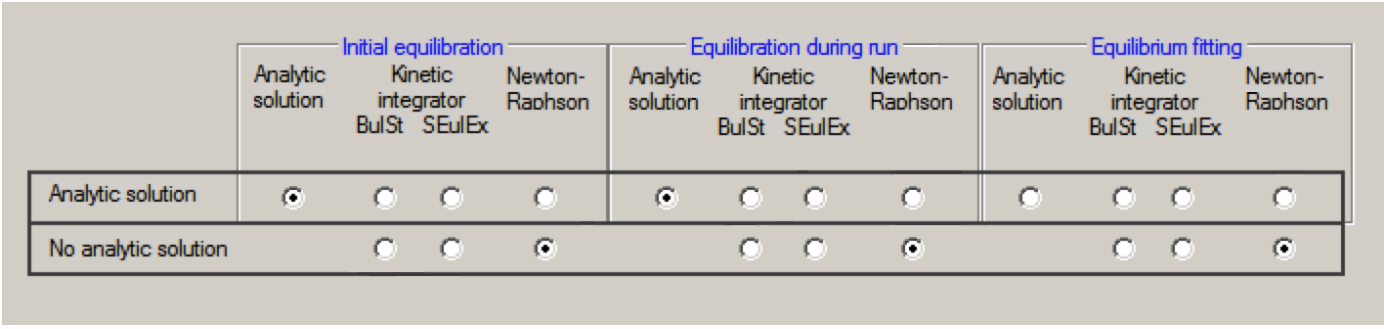


Figure 3-161 Settings for a system that is in instantaneous equilibrium.

3.7.8 The Fitting Control tab

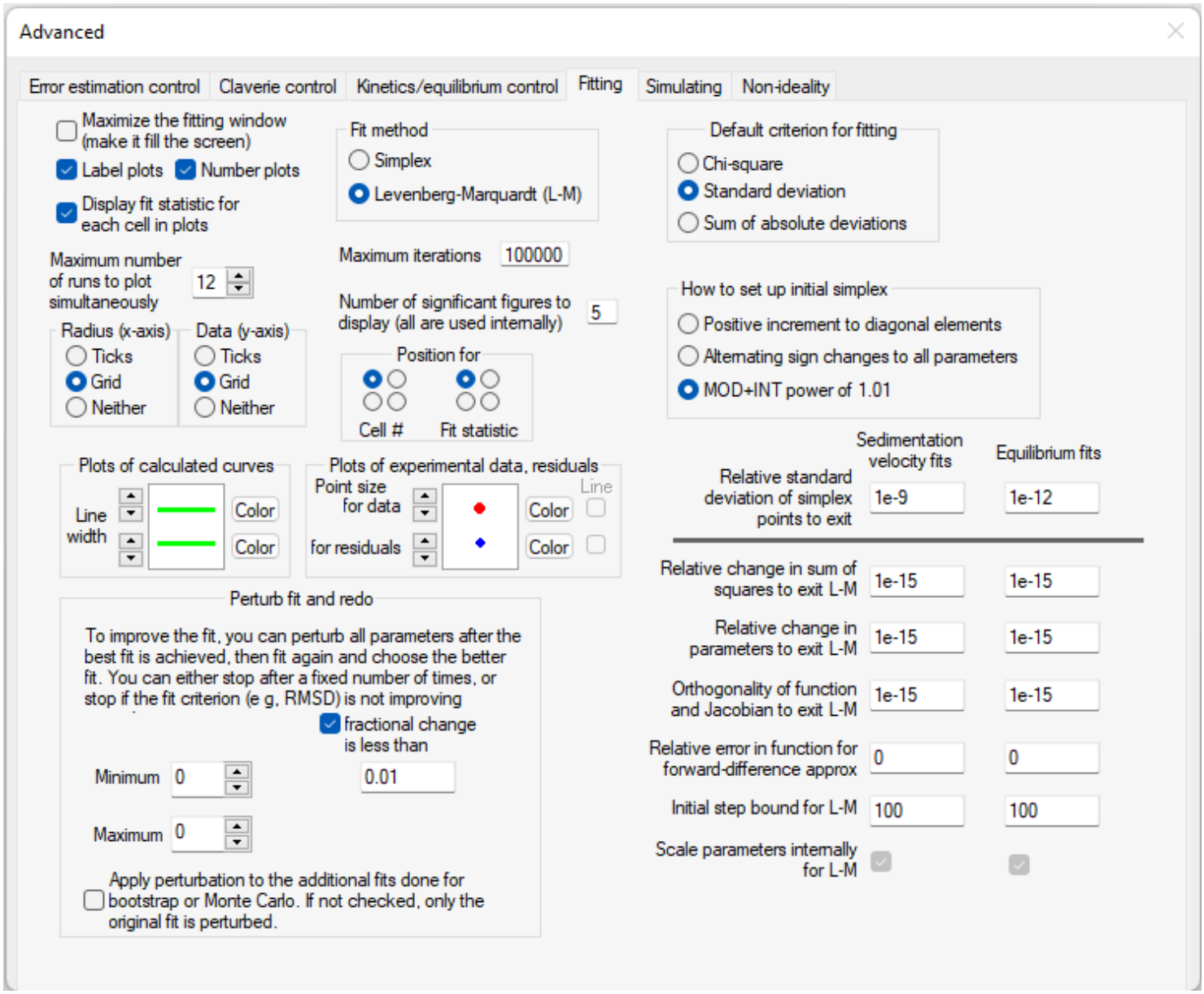


Figure 3-162. Choice if default fitting parameters.

The choice between the default fitting method: Simplex vs Levenberg-Marquardt is made here.

NOTE: The Advanced tab can be accessed also from the Fitting Control screen. Choices made there will apply only to the control file under consideration. Choices made here in Preferences will become defaults but may be changed from the control screen for a particular run.

3.7.9 Simulating Output Choices

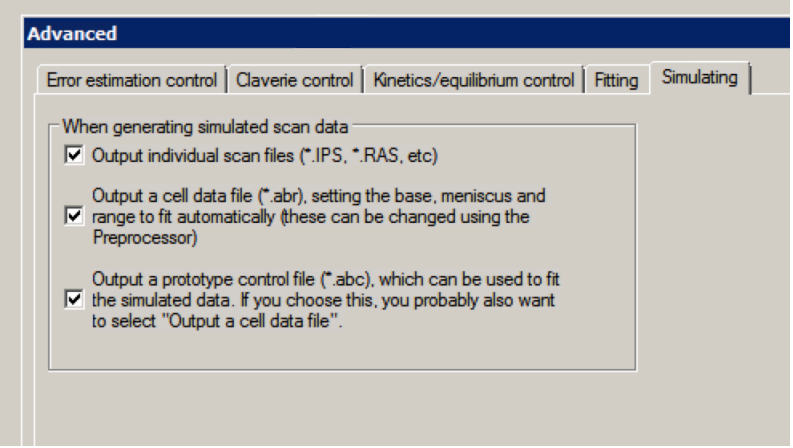


Figure 3-163 Choice of outputs when simulating (choose them all)

3.7.10 Reports: Output, Control Extended tab

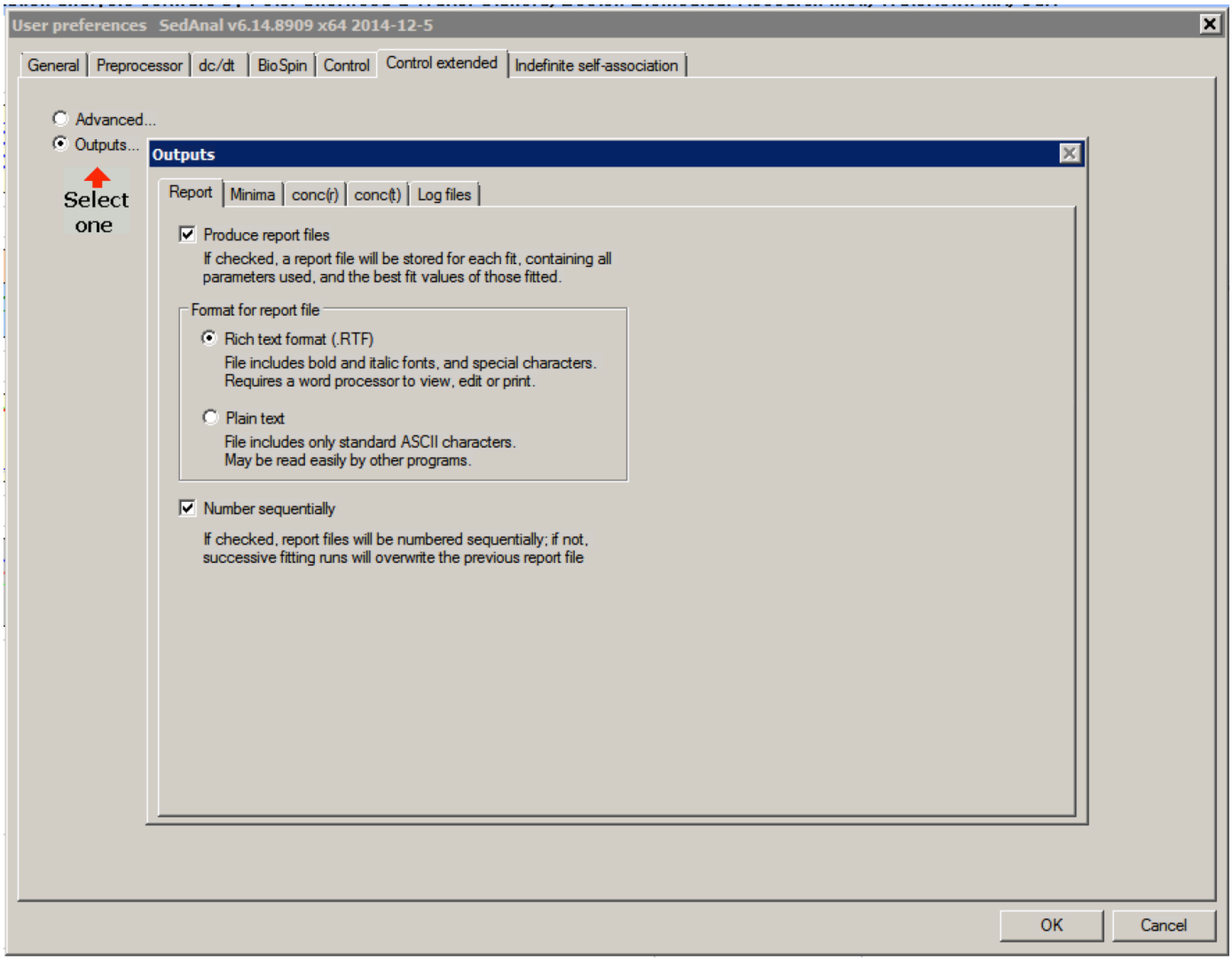


Figure 3-164. Control extended parameters also allow setting of the default parameters found under the "Outputs…" button on the control screen.

3.7.11 The other “Output …” tabs

Control Extended-> Outputs->Minima:

Minimum files: "Minima" tab

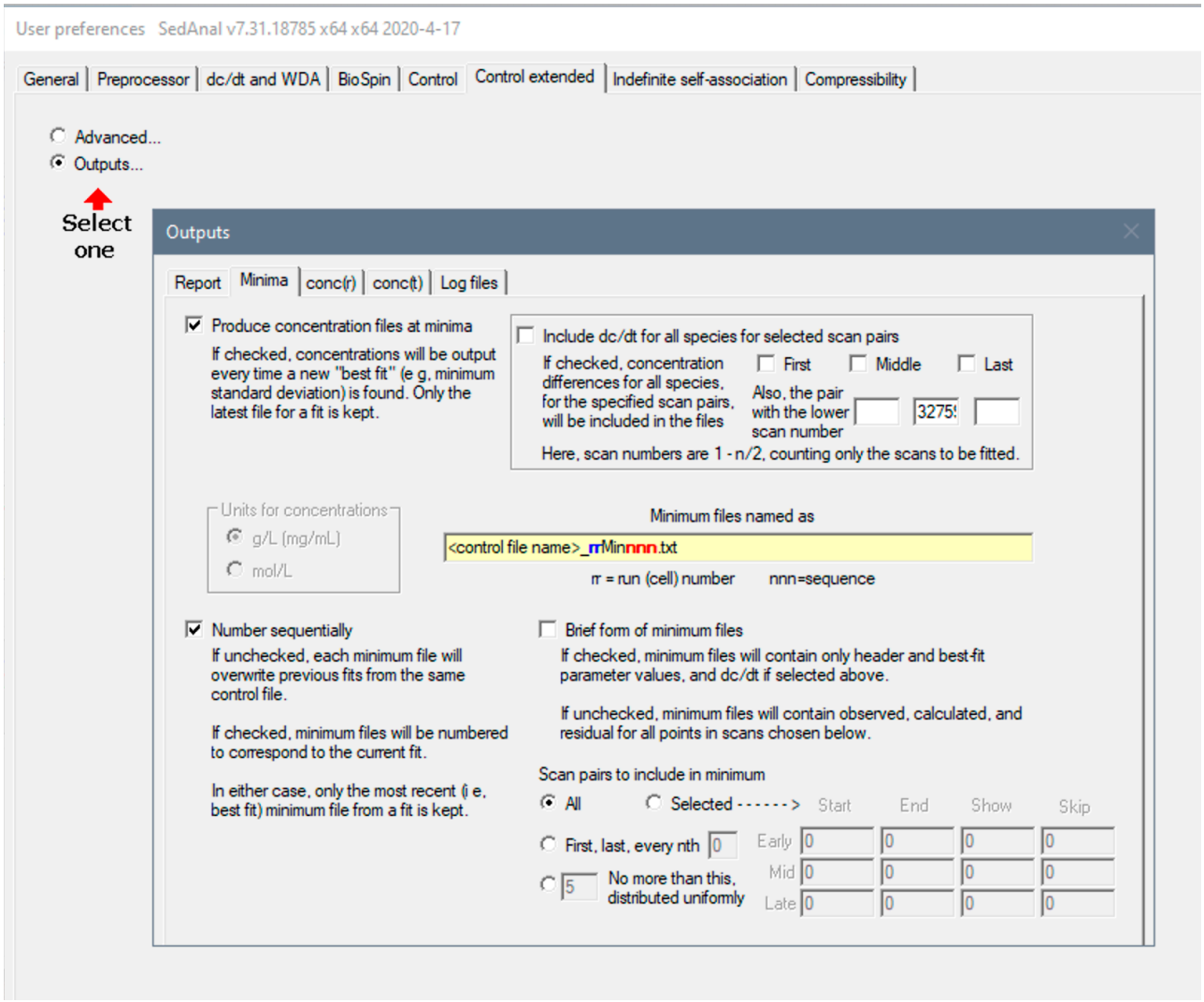


Figure 3-165. Outputs: choose the “min” file formats

At each new minimum during fitting, SEDANAL will write a "minimum" file that contains the last best fit, and the data with the residuals etc... for plotting the results. The number of scans to be included in the “min” files can be indicated in several ways by selecting the output format in the lower right section of the **Figure 3-165.** This format can be chosen to accommodate your plotting program. Data from all delta-C scan pairs or Selected scan pairs, or first and last and evenly spaced, or a specified number evenly spaced.

3.7.11.1 Concentration as a function of radius output.

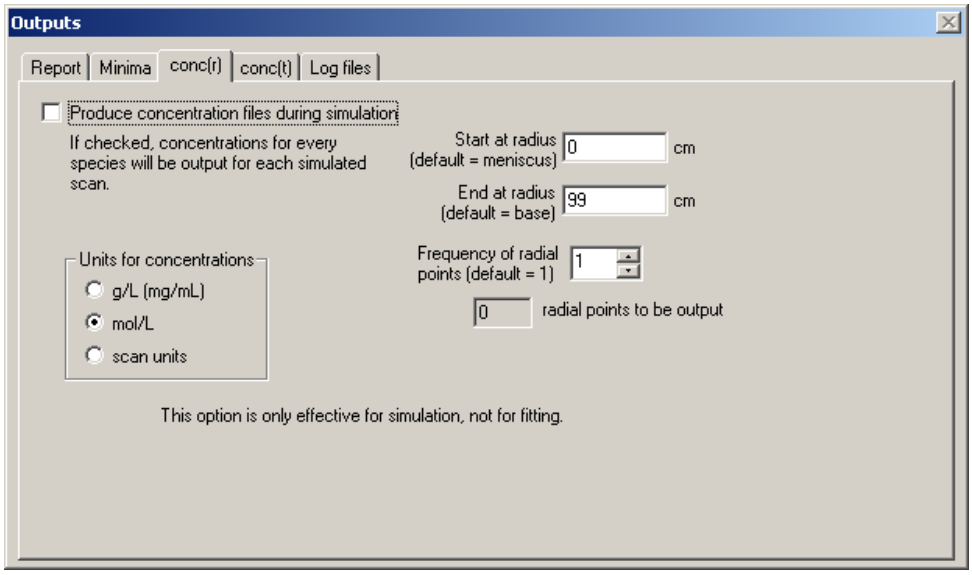


Figure 3-166**. Ancillary output files**

Ancillary output files can be written during simulation that will contain the concentration of each species as a function of radius at each time point.

3.7.11.2 Concentration as a function of time for the initial equilibration step

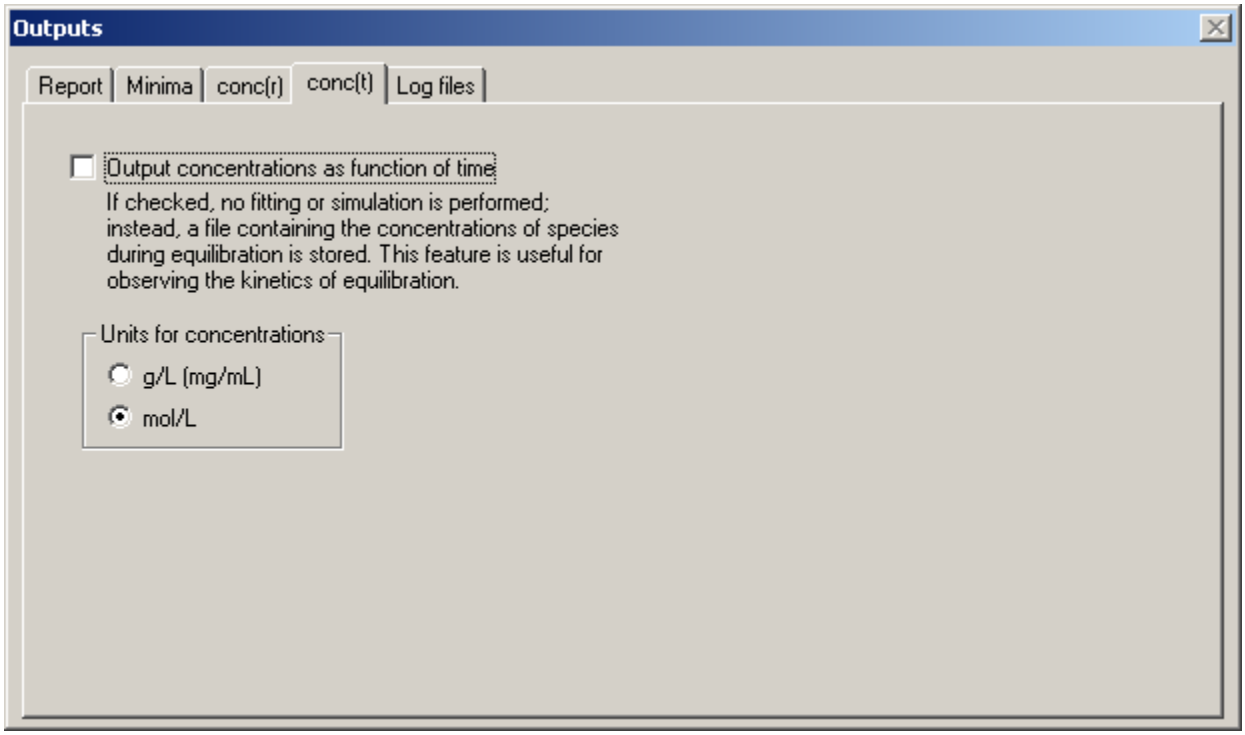


Figure 3-167. Kinetics data for the model under consideration.
This has been superseded by the Kinetics function accessible from the Main Menu.

3.7.11.3 Disposition of various log files

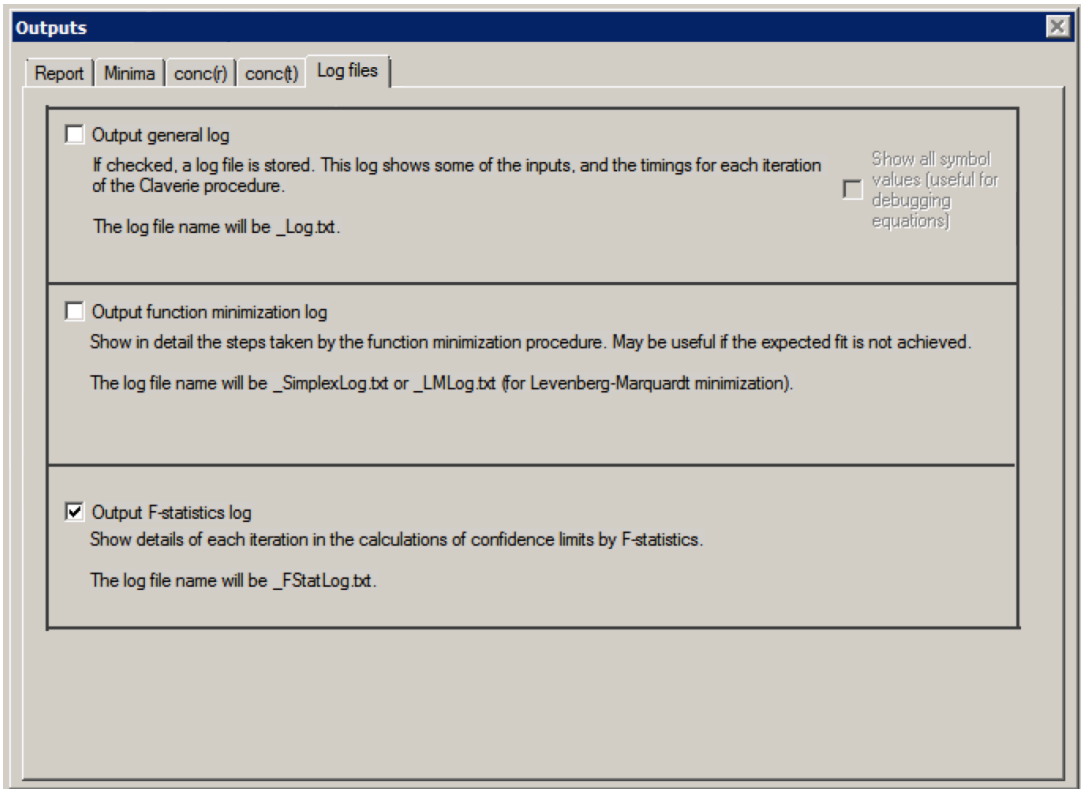


Figure 3-168. Disposition of various log files

In this example, the "Output F-statistics log" is checked. Check this. The other two are for debugging purposes.

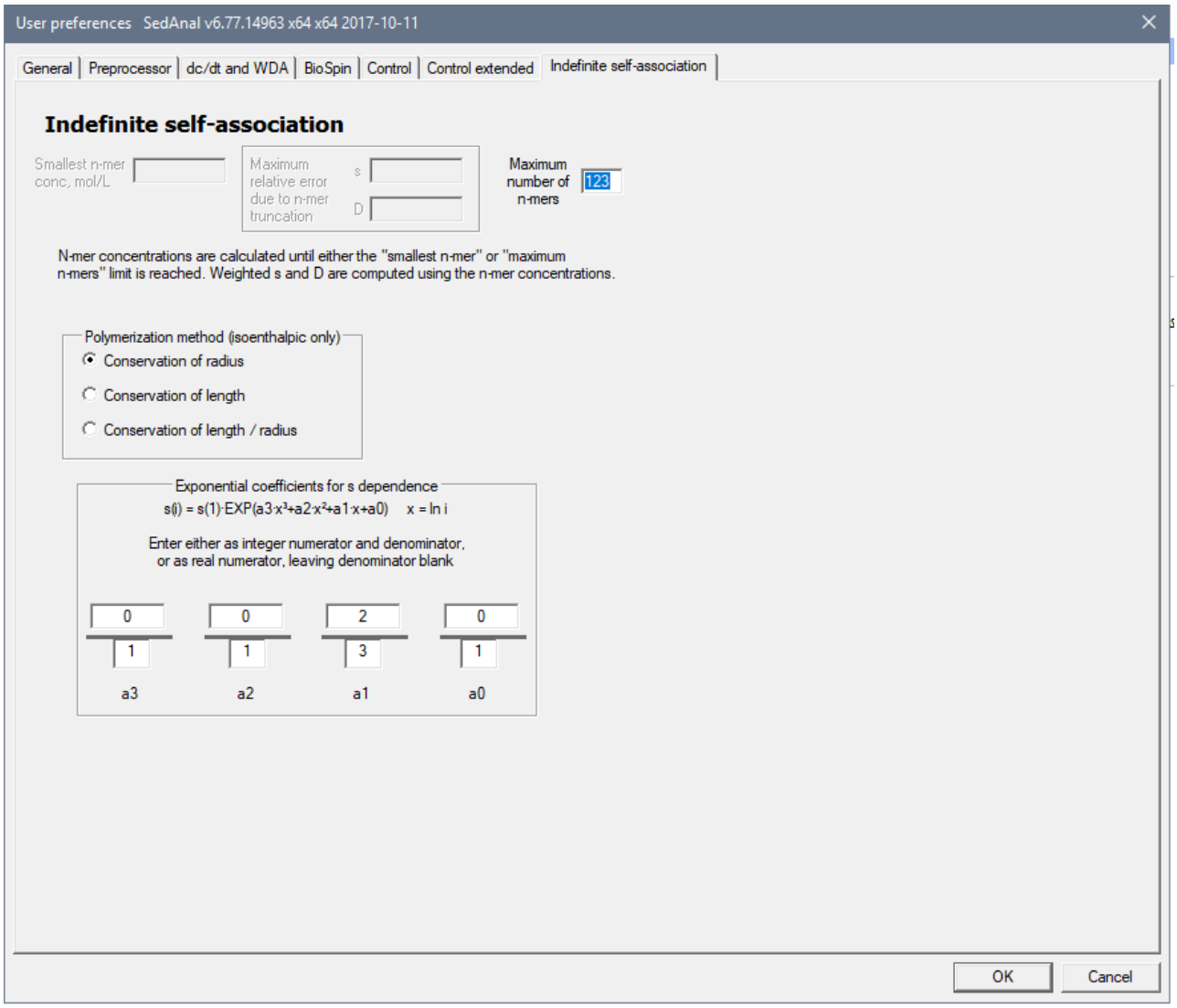


Figure 3-169**. Setting the default parameters for indefinite self-association reactions.**

Setting the default parameters for indefinite self-association reactions. The coefficient for a1 is 2/3 in this example and signifies that the values of the sedimentation coefficient of each of the oligomer are based on the $s_i=M_i^{2/3}$ relationship for oligomers that all have the same frictional ratio: this has been used traditionally. However, the other coefficients allow one to select any relationship between the sedimentation coefficients of the oligomers and can be determined using bead models, for example, or by other appropriate means**.**

3.7.12 Fitting preferences

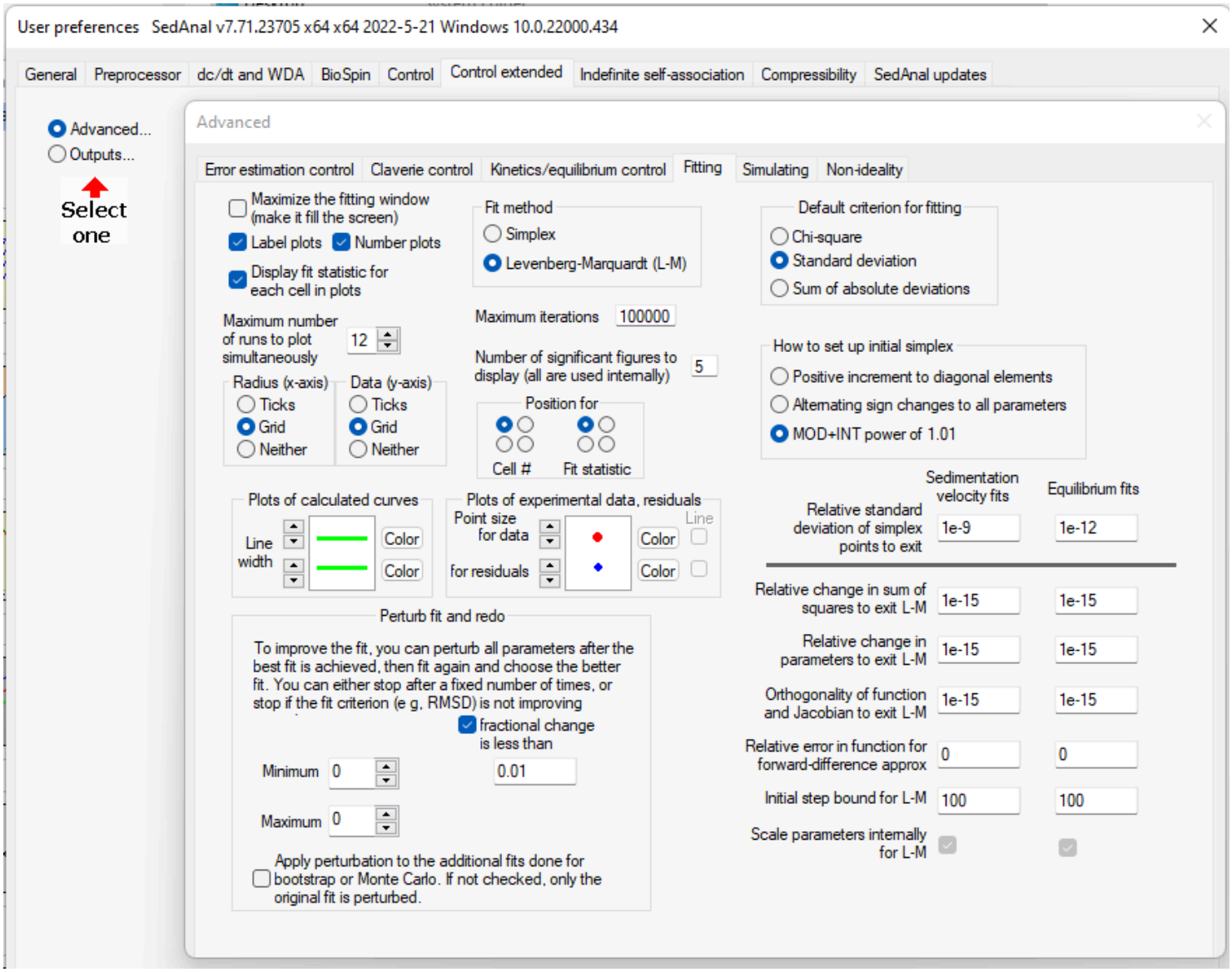


Figure 3-170. **Fitting preferences.**

The "MOD+INT power of 1.01" button causes each element of the initial simplex to be filled with a different number. The "Positive increment …" button alters only the diagonal elements of the initial simplex. Don't use the "Alternating sign …" button; we don't remember what it does.

3.7.13 Simulation Preferences

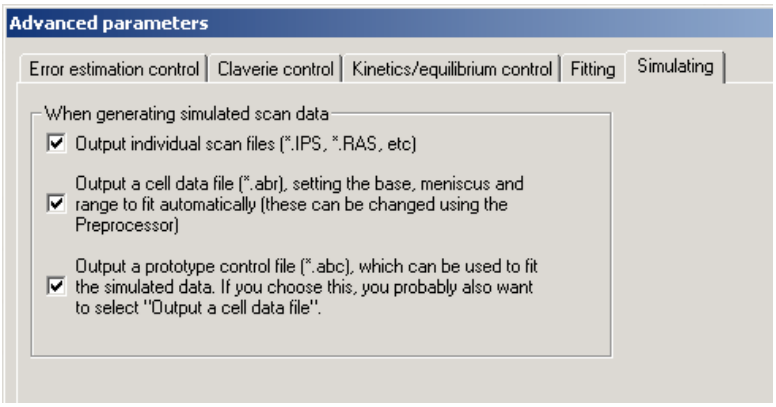


Figure 3-171. **Fitting and Simulating Preferences** - choosing output files: check them all.

3.8 BIOSPIN

BIOSPIN is the program originally developed by Dennis Roark and David Yphantis to compute point-by-point molar mass averages as a function of the local cell concentration from sedimentation equilibrium data (D.E. Roark & D.A. Yphantis, 1969). It has been implemented as a subprogram of SEDANAL. The reader is referred to the original literature for further information. In addition to the point-wise number, weight, Z- and Z+1 molar mass averages, it also computes several of the so-called “charge-independent” molar mass averages that are independent of the second and higher virial coefficients: the Y1, Y2, Y3, etc… averages. (Yphantis & Roark, 1972). The charge independent averages are useful for analyzing systems under non-ideal conditions. The original BIOSPIN user manual can be downloaded from https://sedanal.org/biospin_manual.pdf.

3.9 SIMULATING DATA

To simulate XLA/I or Multi-wavelength data, select "**Fit preprocessed data**" from the Main Menu. When the Control Screen appears, click the "new" button and either create a new "experiment" folder or select an existing “experiment” folder in which the simulated data files will be stored. Then click the button labeled "Simulate data"; the screen will change to the simulator screen. Now select a model from the drop-down "Model to be fitted" list. Enter the number of points to be used for the grid between meniscus and base. Enter the molecular parameters in the "Molecular parameters" boxes.

In the bottom panel, select the optical system; indicate whether it is a velocity or equilibrium experiment: enter a comment; specify the meniscus, speed, base and loading concentration. For a velocity run enter the total number of scans and the time between scans. Now enter the magnitude of normally distributed random noise to add to the simulated data.

The window labeled "Bottom time, sec" will display the total time for the slowest species to reach the bottom of the cell. You can adjust the total number of scans and time between scans as you desire.

When all the boxes have been filled, click "Store control file and start fit" to start the simulation and to save the simulation control file. The simulated data files will have names of the form 000nnn.IPs or 00nnn.RAs and will be in XLA/I standard format. The simulated data, *.IPs or *.RAs files, will be stored in standard time and date

subdirectories of the experiment directory which was selected when you clicked on the "New" button. The time and date used for the path name will be the same as the time and date of the simulation run. The simulated XLA/I data files are recognized and can be processed like any other XLA/I data files by SEDANAL or other programs.

Simulation can generate a cell data (run, or .abr) file, as well as the usual scan files. The run file's name is YYYYMMDD_simHHMMSS.abr, and it has the meniscus and cell base from the simulation, and the range to fit is (meniscus+0.05) - (base-0.1) (i e, 5.95-7.1 for m=5.9 and b=7.2). It can also write a prototype control file that can be used to fit the simulated data. To turn on either of these features go to the "Advanced …" button and click on the "Simulating" tab and select the appropriate buttons. A prototype fitting control file can also be produced by clicking the appropriate button below:

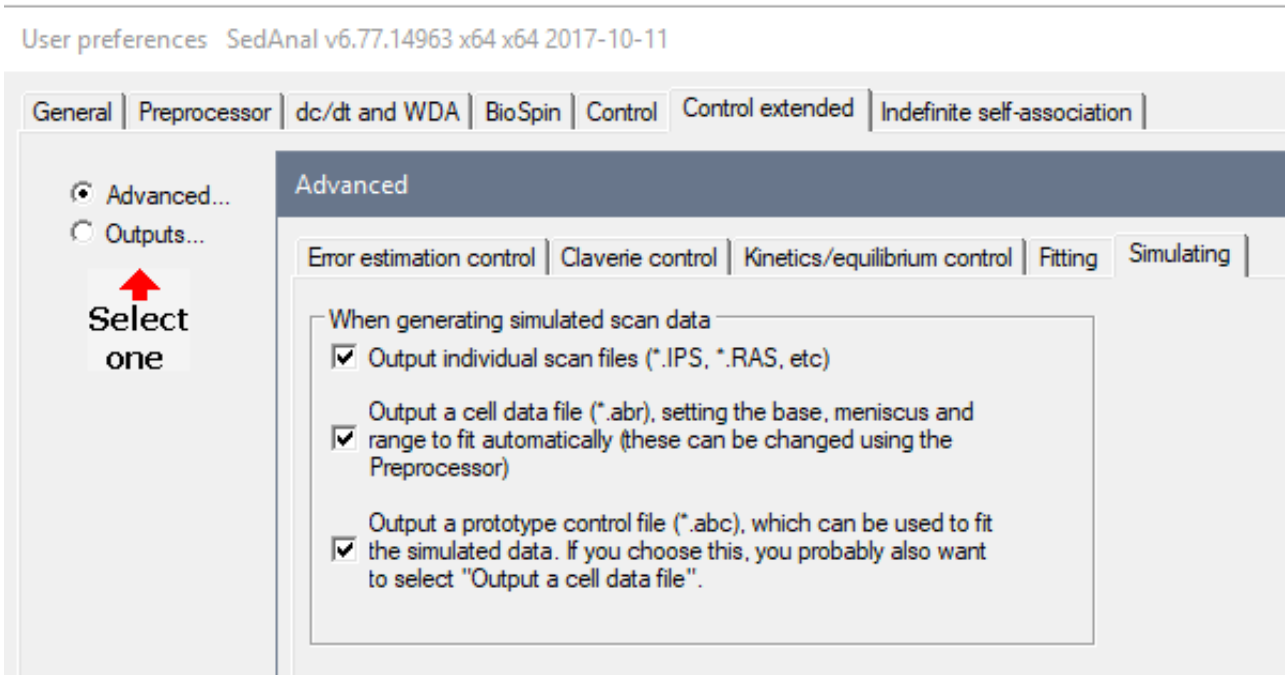


Figure 3-172. Choosing output files: Preferences -> Control extended -> Simulating. Choose them all.

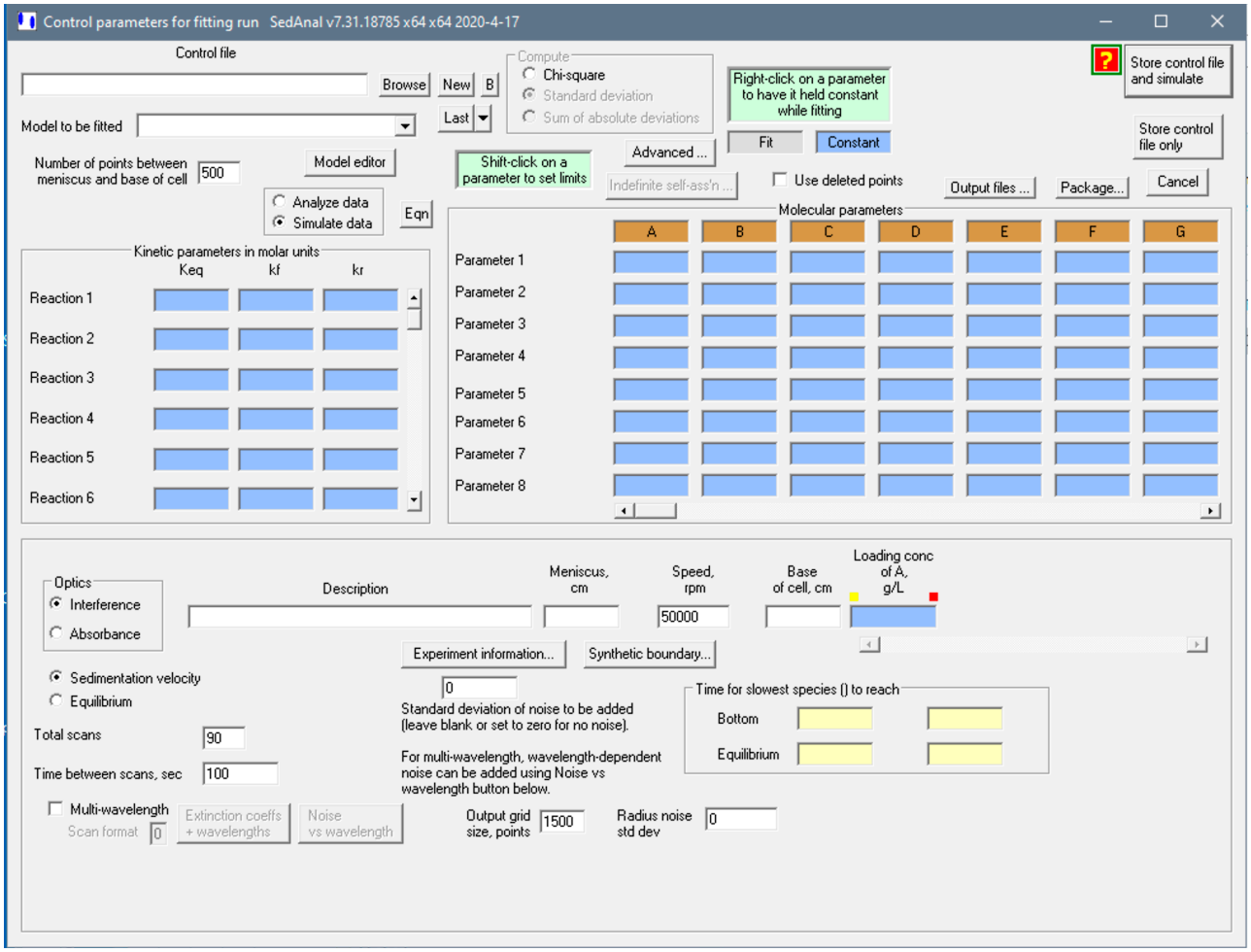


Figure 3-173 Control screen after selecting "simulate"

Other features:

3.9.1 Saving a "package".

A package consists of all the files required to reproduce the current fit. It is useful for debugging purposes if an "anomaly" occurs and SEDANAL is misbehaving. The package is a zip file containing the control file, the run files, the Model file needed to reproduce the fit, and several other relevant files. A package is also a convenient vehicle for sharing a dataset with another researcher. A Package can be generated by clicking on the "Package" button on the Fitting Control Screen:

Figure 3-174**. Generate a Package**

3.9.2 Output of initial reaction time course.

(See paragraph 2.10- "*Kinetics Simulator*" for a more user-friendly and comprehensive way of computing the kinetics.)

Kinetics: To write a file of the time course of the initial equilibration of a particular model, click on the "Output files button and select the "conc(t)" tab and check the "Output concentrations as a function of time" box. Then select the units for the concentrations

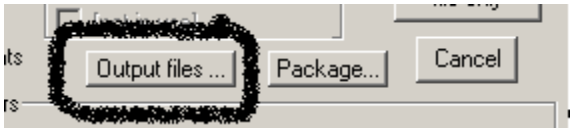


Figure 3-175**. Output files**

3.9.3 Outputting the time course of the initial equilibration step

When the "Store control file and …" button is clicked, only the kinetics of the model are calculated and a file named after the abc and report files is written. No fitting will be done in this case.

If the abc file is named xxxxxx.abc, the report file will be named xxxxxx_Report001.rtf and the kinetic time course will be found in a filed named: xxxxxx _Equilibration.txt. Remember **to unclick the "Output concentrations as a function of time" box when you are done.**

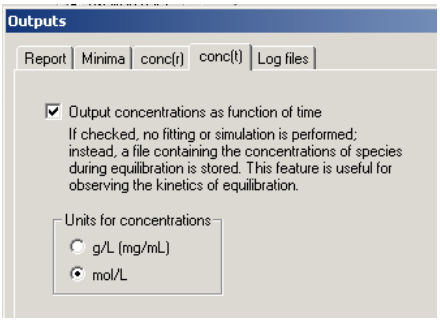


Figure 3-176**. This feature has been replaced by the Kinetics Simulator described in Section 3.10**

3.10 Kinetics Simulator

A general purpose kinetics simulator has been added to SEDANAL and is accessed from the Main Menu (Figure 3-23) by clicking the Kinetics button. The kinetics simulator can handle any model that can be represented in the ModelEditor. When the kinetic simulator screen is first opened, a blank screen showing the possible combinations of 28 species and 27 reactions. One should click on "New" to create a new kinetics control file, or select a previously used control file from either the drop down menu or by browsing for it.

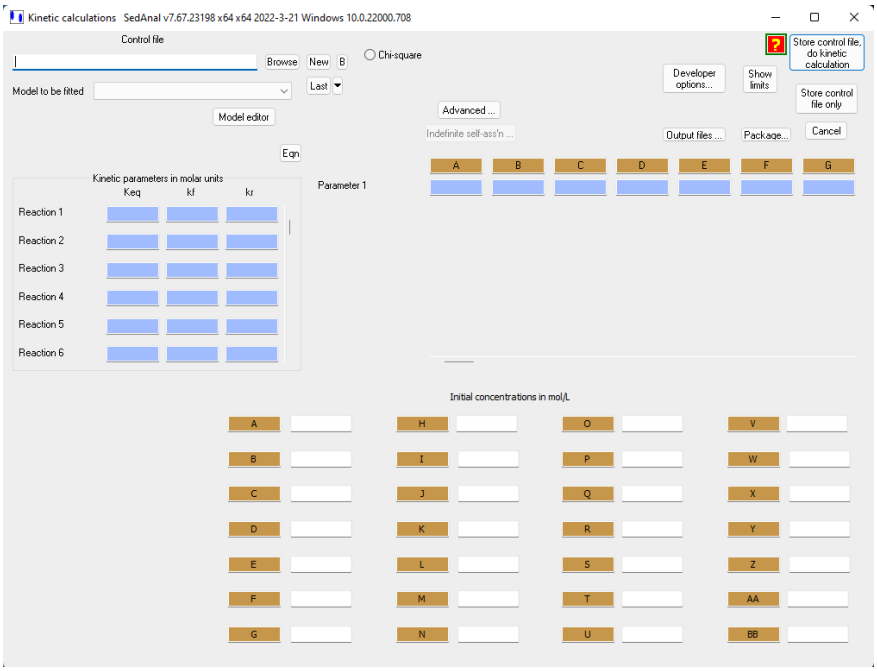


Figure 3-177 Kinetics Control Screen

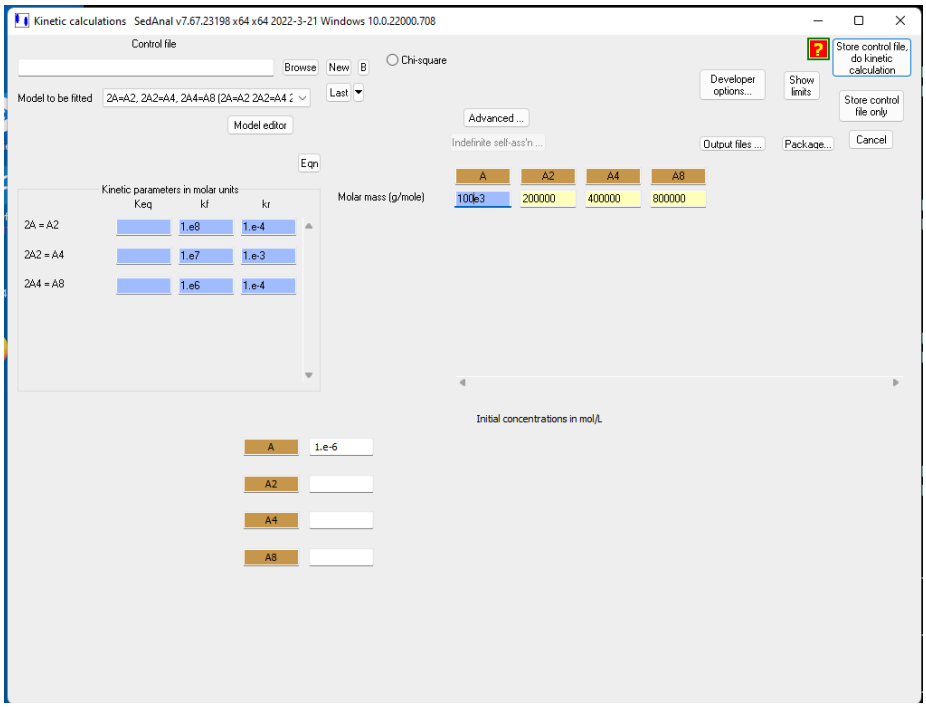


Figure 3-178 Kinetics control screen for the Model 2A=A2; 2A2=A4; 2A4=A8

Molar masses of A and B are 20,000 and 50,000 g/mol, respectively. The forward and reverse rate constants for the first and second reactions are 2 x 10^4 $M^{-1}s^{-1}$, 0.01 s^{-1}, and 0.5 x 10^4 $M^{-1}s^{-1}$, 0.01 s^{-1}, respectively. The initial concentrations of A and B are 1 x 10^{-6} and 2 x 10^{-6}M, respectively**.**

Clicking "Store control file and do kinetic calculation" results in the following display.

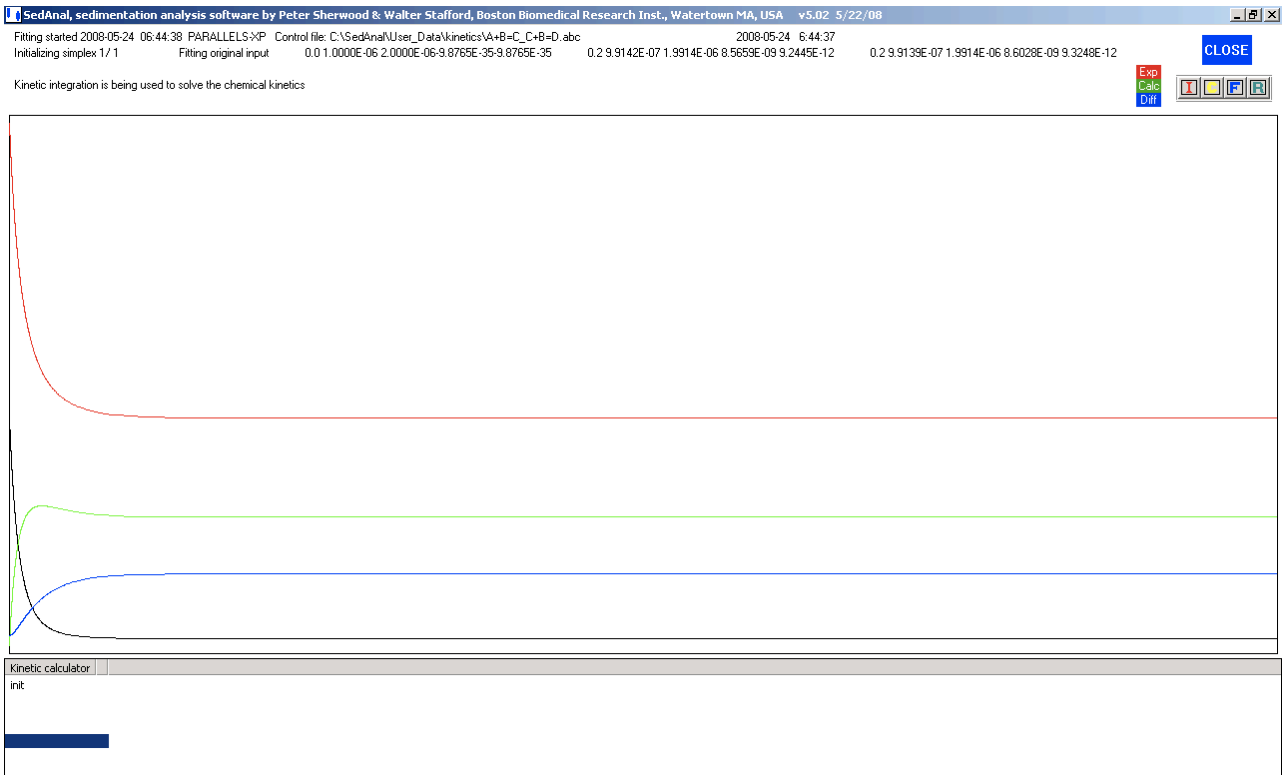


Figure 3-179 Screen dump of kinetic simulation

These data have been written to a file whose file name will have the following form:

New_Control_Filename_Equilibration.txt.

The text "New_Control_Filename" will be replaced by the name you gave the control file after you pressed "New" or changed the control file name after loading an existing abc file with either the "Last" or "Browse" buttons.

These simulated data can be read into your favorite plotting program to produce a plot something like this:

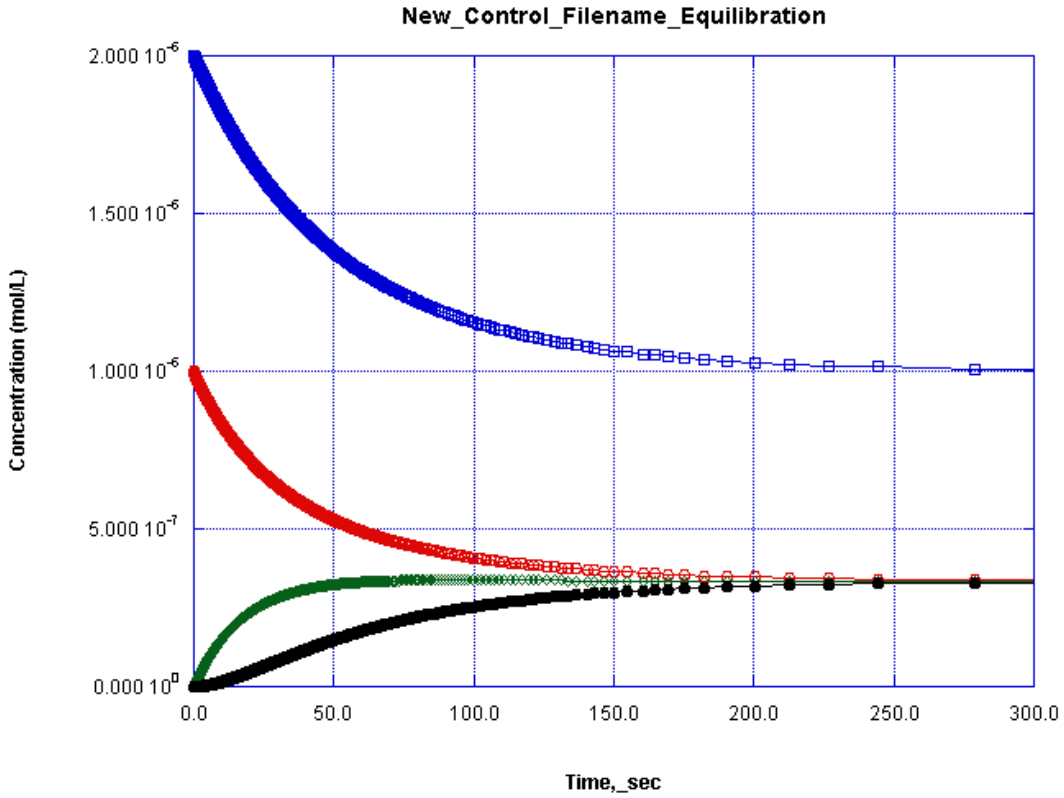


Figure 3-180 Plot of molar concentrations as a function of time.

Plot of molar concentrations as a function of time for the system A+B=C; C+B=D (Figure 3-180). The forward and reverse rate constants for both the first and second reactions are 1 x 10^4 $M^{-1}s^{-1}$, 0.01 s^{-1}, respectively. The initial concentrations of A and B are 1 x 10^{-6} and 2 x 10^{-6}, respectively.

Here is another example with the same equilibrium constants but the second reaction forward and reverse rate constants are one tenth those the previous example.

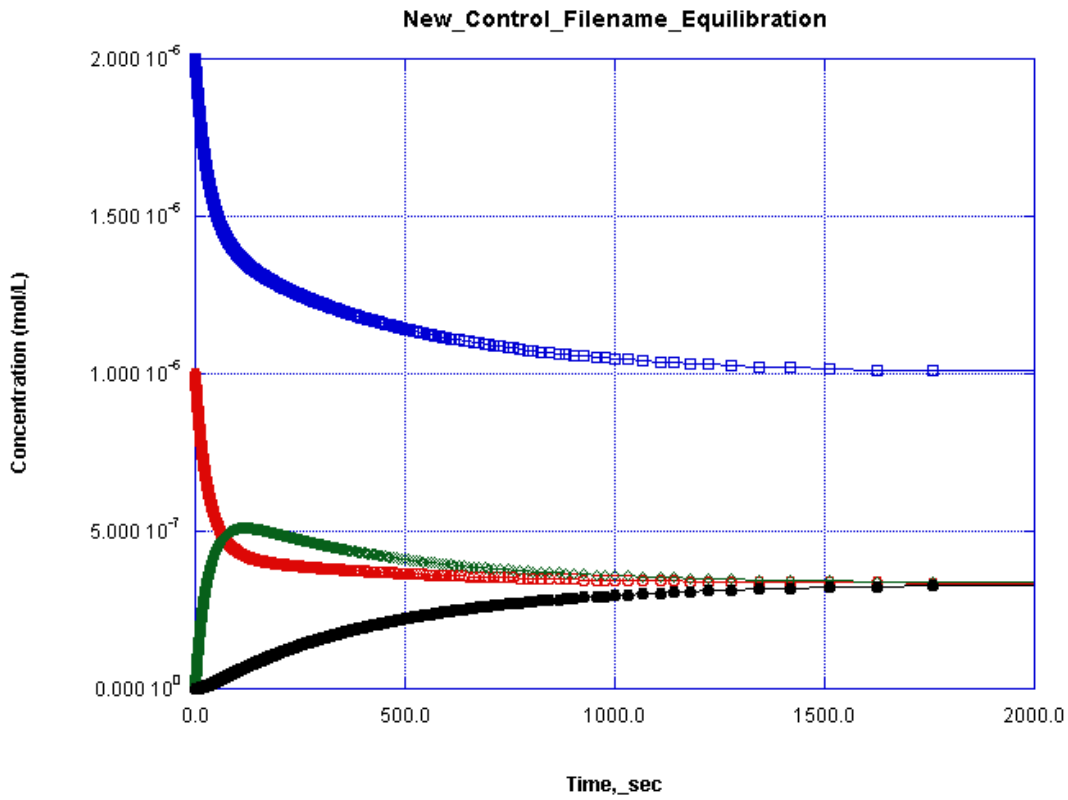


Figure 3-181 Plot with same equilibrium constants; but rate constants for the second step are 1/10 those of the first example above (Figure 3-180).

Another plot (**Figure 3-181**) with same equilibrium constants but forward and reverse rate constants for the second step are 1/10 those of the first example above. The forward and reverse rate constants for the first and second reactions are 1 x 10^4 $M^{-1}s^{-1}$, 0.01 s^{-1}, and 1 x 10^3 $M^{-1}s^{-1}$, 0.001 s^{-1}, respectively. The initial concentrations of A and B are 1 x 10^{-6} and 2 x 10^{-6}, respectively

.

3.11 Scripting sedimentation fitting and simulation

The script functionality is rudimentary at present, to give us a chance to experiment with it and find out what is useful and what isn't. Not all the error-checking or UI features are present.

The script files must be in the User_data directory, and have an extension of .abs. Script files are created and modified with a text editor.

3.11.1 Scripting commands:

The initial line of the script must be exactly

```
SEDANAL script 1
```

The following commands may be included in a script:

Command	Meaning	Example
`MENU FIT`	Equivalent to clicking Fit preprocessed data on the main menu.	`MENU FIT`
`LOAD` *controlFile*	Specify a control file to be loaded; *controlFile* can be either the name of a control file (see below), or `$LAST`.	`LOAD MyExperiment\ABCD_Fit`
`MODIFY` *expression*	Change the value for a kinetic, molecular, or cell parameter from the value in the control file, and as previously modified. The symbols and format of the expression are the same as in the equation editor.	`MODIFY K(1)=K(1)/10`
`FIT`	Equivalent to clicking Store control file and start fit on the control screen. See note below.	`FIT`
blank lines	These lines are ignored.	[this line intentionally left blank]
comment lines	These lines are ignored, except for checking spelling, grammar, and clarity of style.	`// This is a comment`

A typical script file might look like:

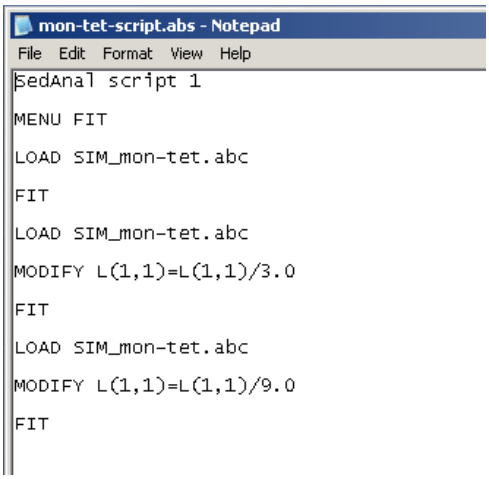


Figure 3-182. typical script

This particular script was used to simulate a dilution series of a monomer-tetramer system creating three abr files for 1 to 3 serial dilutions of the monomer tetramer system that was specified in the "SIM_mon-tet.abc" file. This script will load the original abc file and perform the simulation; then reload it and simulate again with the loading concentration divided by 3; and then repeat with the loading concentrations divided by 9. This is very useful for generating large datasets for exploratory studies of interacting systems to help in designing experiments.

3.11.2 Specifying control files in scripts

The control file is identified by the following rules:

1) If it's just a name, the complete path is completed by trying the directory of the most recent control file loaded, then the most recent control file stored for this instance of SEDANAL, then the most recent control file stored for any instance of SEDANAL. The idea is that SEDANAL tries to guess what you mean if you just say, e g, `ABCD_Fit`.

2) If it is a directory followed by a name (e g, `MyExperiment\ABCD_Fit`), the directory is assumed to be in the user data directory.

3) Anything else is a complete path (e g, `C:\SEDANAL\User_data\MyExperiment\ABCD_Fit`).

4) If the control file name does not end in `.abc`, the extension is added. Thus, `ABCD_Fit` and `ABCD_Fit.abc` mean the same.

3.11.3 Storing scripting control files

To avoid modifying the original control file, when you have `FIT` in a script, the control file is stored with the suffix "`_script001`". For example, if the script in the example column above is run, the name of the control file with K_1 a factor of 10 smaller will be `ABCD_Fit_script001.abc`, and `ABCD_Fit.abc` will not be changed.

When a script file has been loaded, the following screen will appear:

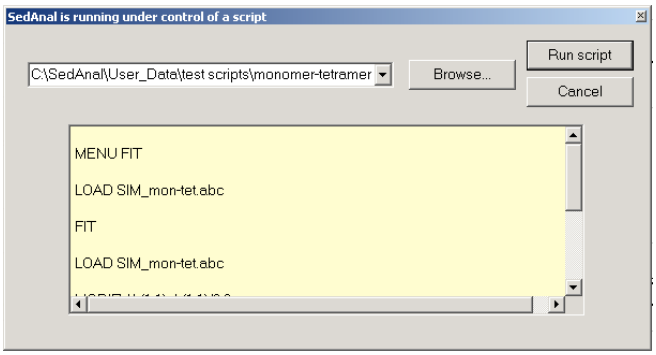


Figure 3-183. Script control

The script is initiated by clicking on "Run script".

3.12 Keyboard Shortcuts

There are several keyboard shortcuts that work from the Main Menu.

Key	Action
p or P	Preprocess centrifuge data
d or D	dc/dt
f or F	Fit preprocessed data
l or L	Reload the previous fit (same as L + LAST)
b or B	BIOSPIN
e or E	Equilibrium calculations
q or Q	Preferences
k or K	Kinetics
s or S	Script

h or H	Help
m or M	Model editor
x or X	Exit

3.13 Help

The help files can be accessed from the main menu. By either selecting "Help" from the main menu or by clicking on the "yellow-question-mark-on-a-red-background" found on most screens.

You will see a window like this:

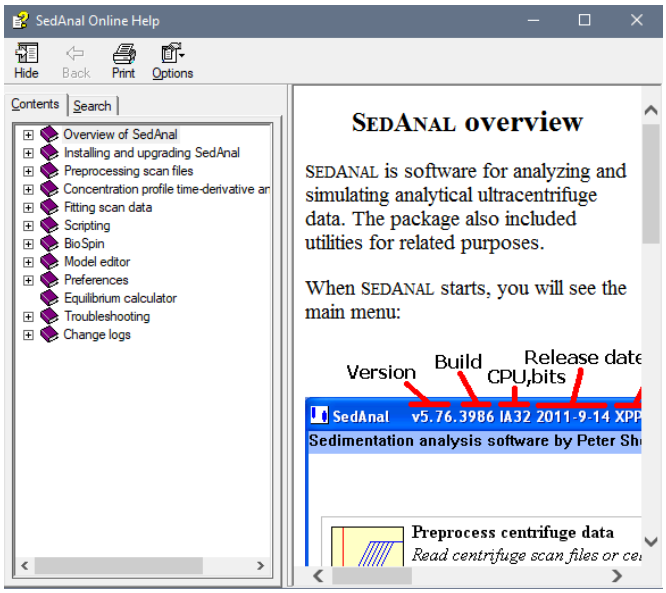


Figure 3-184 SEDANAL On-line Help

Be sure to peruse the Change Logs:

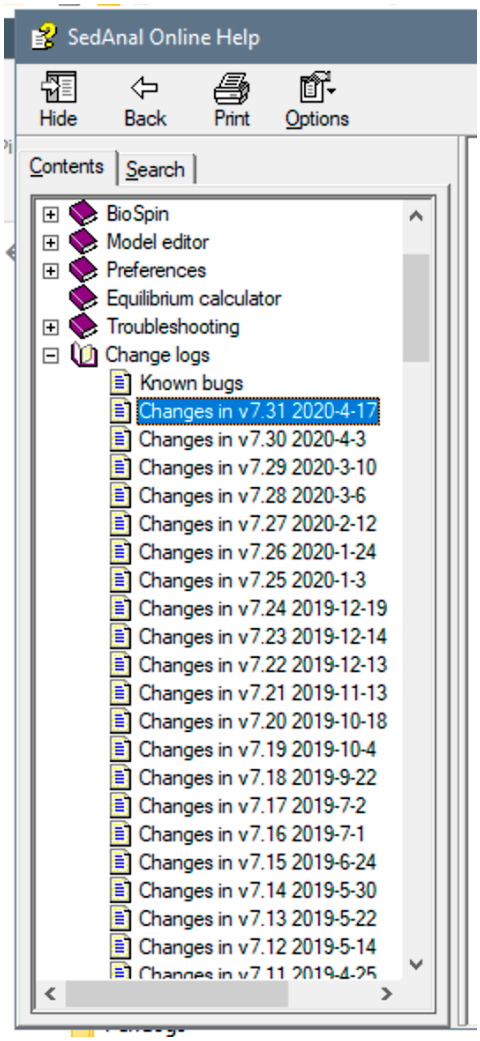


Figure 3-185 Change Log

3.14 Exit

Self-explanatory … bye…

4 Background Theory

The curve fitting portion of SEDANAL is designed to analyze sedimentation velocity and sedimentation equilibrium data from interacting systems composed of multiple macromolecular components.

First, a few words about **components**, **species** and **reactions**:

In the context of this discussion, a component is an electroneutral macromolecule or several molecules related by chemical reaction that can be added or removed, at least conceptually, from a solution independently of other molecules. A component may be composed of several macromolecular species. Species interact through chemical reactions. In general, the number of components is equal to the number of species minus the number of chemical reactions. At constant temperature and pressure (Actually, we are assuming here that the solution is incompressible.), the number of degrees of freedom is equal to the number of components. For any given values of the equilibrium constants, the composition of the solution (i.e. the amount of each species) is determined by the total concentration of each component at each position in the centrifuge cell.

An example of a single component, multi-species system is a monomer-dimer, rapidly reversible self-associating system. The monomer and dimer are in equilibrium with each other and, therefore, cannot be added to or removed from the solution independently. This is why they cannot be separated on a gel filtration column, for example. As soon as some dimer and monomer become separated either the monomer reassociates to form more dimer or the dimer dissociates to form more monomer. At any given temperature and pressure, the amount of monomer and dimer present is determined completely by the total concentration of the component and the equilibrium constant for the dimerization reaction. This is a one component system composed of two species with one chemical reaction between them. If this were a monomer-dimer-tetramer system, it would be a one component, three species system with two chemical reactions between them. It would have one degree of freedom such that the amounts of monomer, dimer and tetramer at each point in the boundary would be uniquely determined by the total macromolecular concentration at that point.

Similarly, consider a system composed of two components, A and B that interact to form a complex, C. Let C interact with another molecule of B to form D. This is a two component system which is composed of four species that are related by 2 chemical reactions. The system has two degrees of freedom, the total concentrations of A and B, respectively.

$$A + B = C \qquad K1=k_{1f}/k_{1r}$$
$$C + B = D \qquad K2=k_{2f}/k_{2r}$$

However, if D were composed of 1 mole of A and 2 moles of B but was not in equilibrium with C and B, it would be a third component. This system would have 3 degrees of freedom, the total concentrations of A, B that participate in the reactions and the total concentration of D. In principal, D could be removed from the system by gel filtration without being reformed from C and B. D might be a covalently cross-linked aggregate, for example. So …

Number-of-components = number-of-species - number-of-reactions

4.1 Sedimentation Velocity Theory:

4.1.1 Concentration time-difference curves.

Consider the signal obtained from the centrifuge. Whether it is fringe displacements, absorbance or fluorescence, it is a function of both time and radius, call it S(r,t). After S(r,t) has been preprocessed to remove optical jitter and integral fringe shifts, it is composed of the contribution from the true concentration distribution, C(r,t), which is also a function of time and radius, as well as a background optical (systematic error) component, B(r), that is time independent and a function of radius only. The signal, S(r,t), also has stochastic noise included with it. We have

$$S(r,t) = \alpha C(r,t) + B(r) + noise$$

where α is the conversion factor (also known as the extinction coefficient) between concentration (g/L) and either fringes, absorbance or fluorescence.

We can remove completely the time independent background component, B(r), of the signal by subtracting any two experimental curves, say at times t_1 and t_2.

$$S(r,t_2) = \alpha C(r,t_2) + B(r) + noise$$
$$S(r,t_1) = \alpha C(r,t_1) + B(r) + noise$$
$$\text{- -}$$
$$\Delta S(r,t_1,t_2) = \Delta\alpha C(r,t_1,t_2) + 0 + \sqrt{2}\, noise$$

The time difference curves, $\Delta S(r,t_1,t_2)$, are proportional to the concentration difference curves, $\Delta C(r,t_1,t_2)$ but have no time independent systematic error. The irregular, time dependent, systematic, error from the optics (often referred to as jitter) will have been removed at the preprocessing stage. The data to be fitted, $\Delta S(r,t_1,t_2)$, have only stochastic errors which makes them suitable for least squares fitting.

SEDANAL fits to the time difference curves $\Delta S(r,t_1,t_2)$ for any set of parameter guesses by generating concentration curves, $C(r,t_1)$ and $C(r,t_2)$, corresponding to times t_1 and t_2, and subtracting them to form $\alpha\Delta C(r,t_1,t_2)$. The root mean square residual is computed as the triple sum over all the points, scans and cells.

$$RMSD = \left[\frac{1}{LMN} \sum_{k=1}^{L} \sum_{J=1}^{M} \sum_{i=1}^{N} [\Delta S_{k,j}(r_i, t_j, t_{j+M}) - \alpha \Delta C_{k,j}(r_i, t_j, t_{j+M})]^2 \right]^{1/2}$$

where k is the cell number, L is the number of cells, j is the difference curve index, M is the number of difference curves and 2M is the total number of scans, j and j+M are the indices of the scans being subtracted, N is the number of radial points in each difference curve, and i is the radial point index.

SEDANAL can also minimize reduced Chi-Square for cases in which datasets are combined from different optical systems that have different relative signal strengths and noise levels. In this case we, minimize the sum of the squares of the weighted residuals, as normalized by the variance of the noise on the data. For example, noise levels on interference data are typically on the order of +/- 0.005 fringes, while the noise on absorbance data is in the range of +/- 0.004-0.006 A.U. and noise on fluorescence data might be +/- 200 arbitrary signal units. So when

data from, say, absorbance and fluorescence optics are combined, we will calculate the normalized residual where σ is the standard deviation of the raw data.

$$\chi^2_{red} = \frac{1}{LMN} \sum_{k=1}^{L} \sum_{J=1}^{M} \sum_{i=1}^{N} \left[\frac{\Delta S_{k,j}(r_i, t_j, t_{j+M}) - \alpha \Delta C_{k,j}(r_i, t_j, t_{j+M})}{\sigma_{i,j,k}} \right]^2$$

Summed over points, scans and runs.

Choose a series of scans to be analyzed (e.g. 14 scans) with the Preprocessor

Subtract scans in pairs to generate Δc curves to eliminate time independent background (e.g. C_8-C_1, C_9-C_2...)

Fit to the Δc curves

(from which time independent systematic errors have been eliminated by the subtraction)

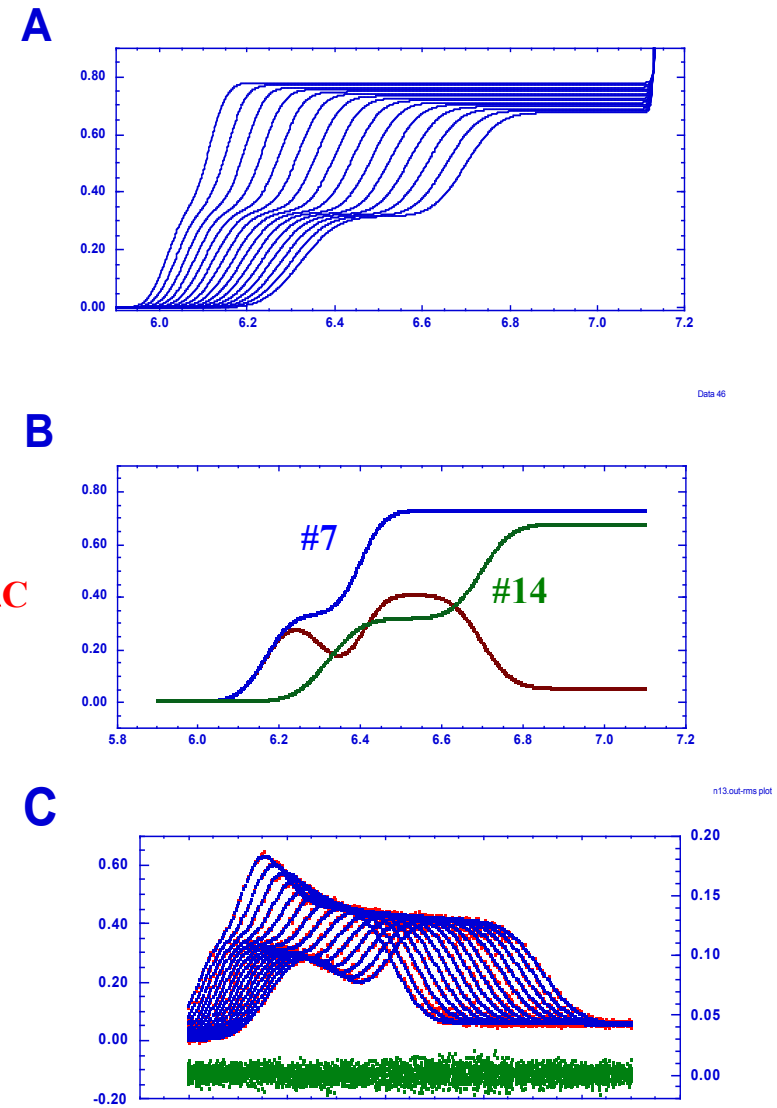


Figure 4-1 Schematic of the procedure used by SEDANAL for fitting time difference curves

4.1.3 Procedure used by SEDANAL for fitting time differnce data.

The figure above (**Figure 4-1**) schematically outlines the basic procedure used by SEDANAL to fit the time difference data, Δc vs. radius.

The rmsd or chi-square is minimized with respect to the model parameters using the simplex directed search method of Nelder and Mead (1965), or the Levenberg-Marquardt (1944, 1963) method.

The numerical solutions to the Lamm equation are generated using the finite element method as described by Claverie, (1975, 1976), with corrections and enhancements introduced by Todd and Haschemeyer (1981, 1983).

The general fitting approach, using numerical solutions to Lamm equation and the Claverie procedure, is based on the method of Todd and Haschemeyer, (1981)

The kinetic differential equations are solved using the Bulirsch-Stoer (BulSt) algorithm with Richardson extrapolation. (Numerical Recipes in FORTRAN), or by the Semi-implicit Euler extrapolation (SEulEx). (Numerical Recipes in FORTRAN)

Bulirsch, R. and Stoer, J. §2.2 in Introduction to Numerical Analysis. New York: Springer-Verlag, 1991.

Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; and Vetterling, W. T. ``Richardson Extrapolation and the Bulirsch-Stoer Method." §16.4 in Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed. Cambridge, England: Cambridge University Press, pp. 718-725, 1992.

Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; and Vetterling, W. T. ``Stiff Sets of Equations, Semi-implicit Euler extrapolation." §16.6 in Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed. Cambridge, England: Cambridge University Press, pp. 718-725, 1992.

Todd, G. P. and Haschemeyer, R. H. ", "General solution to the inverse problem of the differential equation of the ultracentrifuge", Proc. Natl. Acad. Sci. U.S.A., 78(11) 6739--6743, 1981.

Todd, G.P. and Haschemeyer, R. H. "GENERALIZED FINITE ELEMENT SOLUTION TO ONE-DIMENSIONAL FLUX PROBLEMS" Biophysical Chemistry 17 (1983) 321-336

4.2 Sedimentation Equilibrium Theory

4.2.1 Ideal Case:

For a thermodynamically ideal, incompressible system we have for each species that

$$c(r) = c(ref)\exp\left(\sigma\left(\xi - \xi_{ref}\right)\right)$$

where σ is defined as $\sigma = M\left(1-v\rho\right)\omega^2/RT$, $\xi=r^2/2$ and $\xi_{ref}=r^2_{ref}/2$ where r_{ref} is an arbitrary reference radius usually chosen as the first data point. The value of c_{ref} can be related to the loading concentration by invoking conservation of mass and noting that the concentration at the beginning of the run is uniform and equal to the loading concentration, C_o. Invoking conservation of mass, the total amount of macromolecule, T, in the cell is given by multiplying the concentration by the volume:

$$T = h\theta\int_{\xi_m}^{\xi_b} c_o\, d\xi = h\theta c_o\left(\xi_b - \xi_m\right)$$

since this quantity doesn't change with time the following must also be true at equilibrium:

$$T = h\theta\int_{\xi_m}^{\xi_b} c\left(\xi\right) d\xi = h\theta\int_{\xi_m}^{\xi_b} c_{ref}\exp\left(\sigma\left(\xi - \xi_{ref}\right)\right) d\xi$$

$$T = \frac{h\theta c_{ref}}{\sigma}\left[\exp\left(\sigma\left(\xi_b - \xi_{ref}\right)\right) - \exp\left(\sigma\left(\xi_m - \xi_{ref}\right)\right)\right]$$

Now we can equate the two relationships, and after dividing both sides by $h\theta$, write:

$$c_o\left(\xi_b - \xi_m\right) = \frac{c_{ref}}{\sigma}\left[\exp\left(\sigma\left(\xi_b - \xi_{ref}\right)\right) - \exp\left(\sigma\left(\xi_m - \xi_{ref}\right)\right)\right]$$

and

$$c_o = \frac{c_{ref}}{\sigma\left(\xi_b - \xi_m\right)}\left[\exp\left(\sigma\left(\xi_b - \xi_{ref}\right)\right) - \exp\left(\sigma\left(\xi_m - \xi_{ref}\right)\right)\right]$$

and so c_{ref} as a function of the loading concentration c_o is given by

$$c_{ref} = \frac{c_o \sigma\left(\xi_b - \xi_m\right)}{\left[\exp\left(\sigma\left(\xi_b - \xi_{ref}\right)\right) - \exp\left(\sigma\left(\xi_m - \xi_{ref}\right)\right)\right]}$$

C_o can be computed from C_{ref} if the position of the meniscus and the base of the cell can be determined with satisfactory accuracy.

For non-interacting systems each species' loading concentration can be related to its value of c_{ref} in this way, independently of the other species. (In this case, each species is also an independent component.)

For reversibly interacting systems, on the other hand, the law of mass action has to be taken into account. This leads to somewhat more complicated, but still easily solvable relationships between the loading concentrations each component and the values of c_{ref} for each species in equilibrium.

We will consider two reversible systems as examples: a monomer-dimer self-association and a simple one-to-one heterodimer association.

4.2.1.1 Monomer-dimer:

$$2A = A_2 \qquad k_{1,2} = \frac{\left[c_2\right]}{\left[c\right]^2}$$

$$T = h\theta \int_{\xi_m}^{\xi_b} c\left(\xi\right) d\xi = h\theta \int_{\xi_m}^{\xi_b} c_{1,ref} \exp\left(\sigma_1\left(\xi - \xi_{ref}\right)\right) + c_{2,ref} \exp\left(2\sigma_1\left(\xi - \xi_{ref}\right)\right) d\xi$$

$$c_o\left(\xi_b - \xi_m\right) = \frac{c_{1,ref}}{\sigma_1}\left[\exp\left(\sigma\left(\xi_b - \xi_{ref}\right)\right) - \exp\left(\sigma\left(\xi_m - \xi_{ref}\right)\right)\right] + \frac{c_{2,ref}}{2\sigma_1}\left[\exp\left(2\sigma_1\left(\xi_b - \xi_{ref}\right)\right) - \exp\left(2\sigma_1\left(\xi_m - \xi_{ref}\right)\right)\right]$$

Now, invoking the Law of Mass Action, we can substitute for $c_{2,ref}$ and write:

$$c_o\left(\xi_b - \xi_m\right) = \frac{c_{1,ref}}{\sigma_1}\left[\exp\left(\sigma\left(\xi_b - \xi_{ref}\right)\right) - \exp\left(\sigma\left(\xi_m - \xi_{ref}\right)\right)\right] + \frac{k_{1,2} c_{1,ref}^2}{2\sigma_1}\left[\exp\left(2\sigma_1\left(\xi_b - \xi_{ref}\right)\right) - \exp\left(2\sigma_1\left(\xi_m - \xi_{ref}\right)\right)\right]$$

This can be solved easily for $c_{1,ref}$ with the quadratic equation for the monomer-dimer system. For higher oligomers and polydisperse self-association, the equations must be solved by successive approximation.

4.2.1.2 Hetero-Association:

Now for the hetero-association,

$$A + B = AB \qquad K_{AB} = \frac{c_{AB}}{c_A c_B}$$

Conservation of mass for each component, A and B, requires

$$c_{A,o} = \frac{c_{A,ref}}{(\xi_b - \xi_m)\sigma_A}\left[\exp\left(\sigma_A\left(\xi_b - \xi_{ref}\right)\right) - \exp\left(\sigma_A\left(\xi_m - \xi_{ref}\right)\right)\right] \quad +$$

$$\left[\frac{M_A}{M_{AB}}\right]\frac{k_{AB}c_{A,ref}c_{B,ref}}{(\xi_b - \xi_m)(\sigma_A + \sigma_B)}\left[\exp\left((\sigma_A + \sigma_B)\left(\xi_b - \xi_{ref}\right)\right) - \exp\left((\sigma_A + \sigma_B)\left(\xi_m - \xi_{ref}\right)\right)\right]$$

$$c_{B,o} = \frac{c_{B,ref}}{(\xi_b - \xi_m)\sigma_B}\left[\exp\left(\sigma_B\left(\xi_b - \xi_{ref}\right)\right) - \exp\left(\sigma_B\left(\xi_m - \xi_{ref}\right)\right)\right] \quad +$$

$$\left[\frac{M_B}{M_{AB}}\right]\frac{k_{AB}c_{A,ref}c_{B,ref}}{(\xi_b - \xi_m)(\sigma_A + \sigma_B)}\left[\exp\left((\sigma_A + \sigma_B)\left(\xi_b - \xi_{ref}\right)\right) - \exp\left((\sigma_A + \sigma_B)\left(\xi_m - \xi_{ref}\right)\right)\right]$$

This pair of equations can be solved for $c_{A,ref}$ and $c_{B,ref}$, given values of $c_{A,o}$ and $c_{B,o}$, easily by rearrangement into an analytical expression. More complicated stoichiometries can be solved easily by successive approximation (The Newton-Raphson method works nicely in the general case). Success in fitting multi-component systems globally to multiple datasets requires accurate knowledge of the meniscus and base positions (especially the base position) in order to estimate the correct values of c_{ref} from the conservation of mass relationships.

4.2.2 Non-ideal Case:

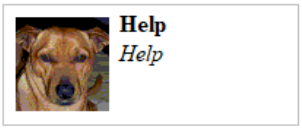
For a thermodynamically, non-ideal, incompressible system we have :

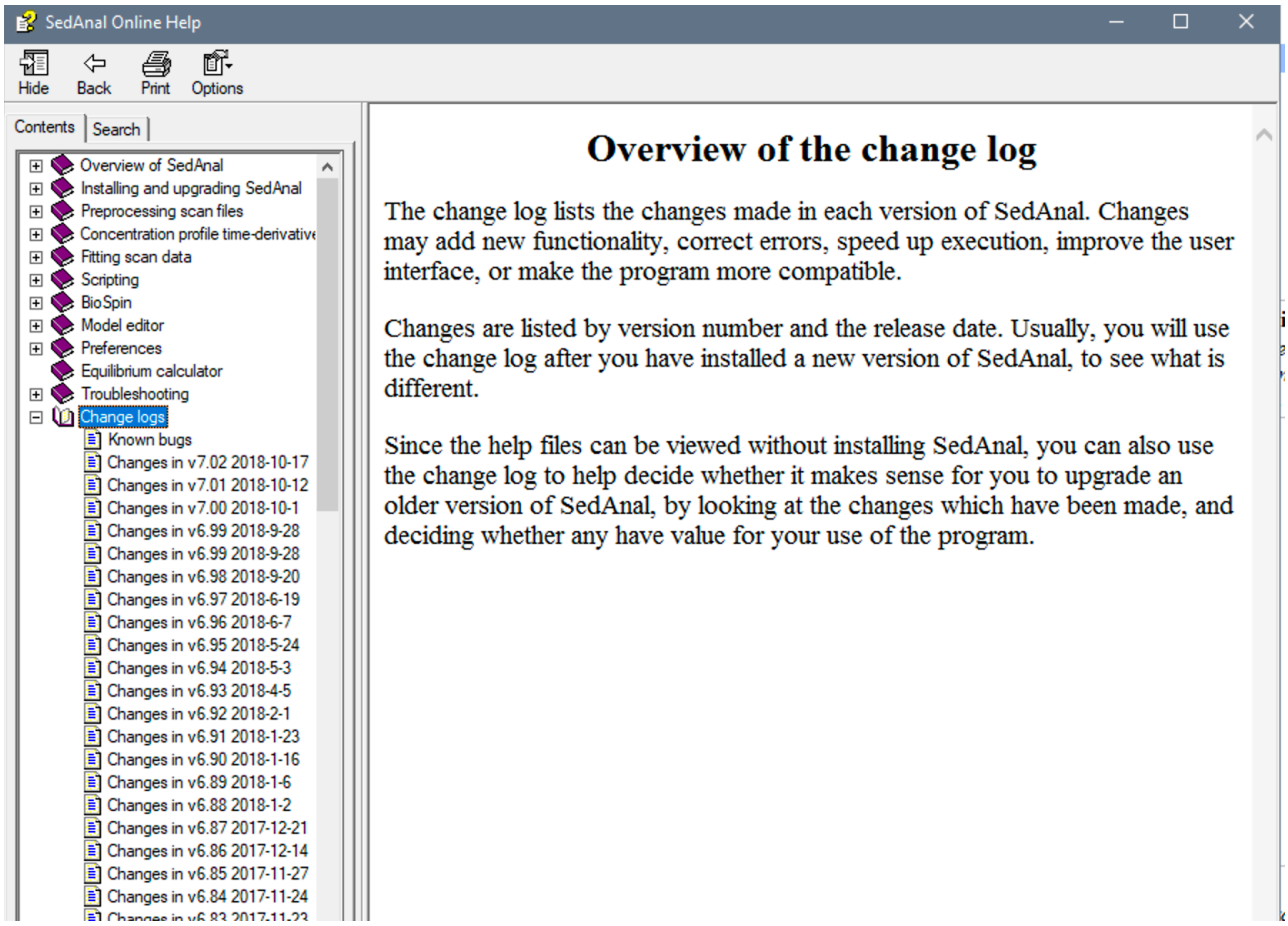
This has been updated- see above.

5 APPENDIX A: HELP FILE and CHANGE LOG.

See the built-in **Help file** that accompanies SEDANAL. It can be accessed either from the **Main menu** or from any screen that has the yellow-question-mark-on-a-red-background symbol:

.





6 INDEX

7 LIST OF FIGURES