Last Update on User Manual:

February 29, 1988

```
*********************************************************
*                                                       *
*         FITSIM                      Version 1.63      *
*                                                       *
*                                                       *
*         Chris T. Zimmerle                             *
*                                                       *
*********************************************************
```

The following is a manual for FITSIM and FITDATA and should be placed according to page number as an addition to the KINSIM manual.

Chapter 5 FITSIM: Kinetic Parameter Optimization using KINSIM

Chapter 6 User-Interface Programs for FITSIM

Chapter 7 Implementation of FITSIM

5.1 General Theory & Methods.

Probably the best method to learn the operation of FITSIM is to sit down at a terminal with this manual in hand! Although it may be tempting to use FITSIM immediately and skip through the explanation of KINSIM, it is strongly suggested the user become familiar with the operation of KINSIM first (i.e. see KINSIM documentation). Nearly all operations of KINSIM are required to both use and understand FITSIM. FITSIM is not a replacement or further version of KINSIM, but rather a powerful addition to it, freeing the user from tedious manual simulation.

The program KINSIM was originally coded by Dr. Bruce Barshop as to be a subprogram to allow multiple simulations. This allowed KINSIM to be called by another program, presumably some form of regression fitting routine. This feature would allow progress curves of a reaction to be fit by nonlinear regression to simulations calculated by numerical differentiation of rate equations. This potential has now been realized by the program FITSIM. FITSIM uses KINSIM as a subprocess to allow multiple simulations under the control of a program which optimizes the fit between simulated and real data. FITSIM is actually a control program for information linkage between one of the non-linear regression subroutines contained within FITSIM and the kinetic simulator (KINSIM) subroutine. It serves to hold all intermediate values for the kinetic simulator, to direct the operation of the regression routine, and to have error-handling capabilities.

As most regression procedures do, FITSIM uses the summation of the squared differences between the calculated and experimental data points (Sum of Squares or SSQ) to determine the quality of fit between simulated and experimental progress curves. To adequately determine the kinetic parameters, progress curves under different experimental conditions must usually be done, and the total SSQ determined. The SSQ value is determined by FITSIM using eq. (1) :

$$SSQ = \sum_{s=1}^{n_s} w_2 \left(\sum_{t=1}^{t=m} w_t \left(exp_t - cal_t \right)^2 \right)_s \qquad (1)$$

(1) where n_s is the number of data sets, or reaction time courses, m is the maximum simulation time, while exp_t and cal_t are the experimental and calculated points at the reaction time, t, for a given simulation parameter set, s. Unequal variance in portions of the experimental data can be a factor compensated for by the weighting factor w_t. The term w_s is routinely used to normalize each full time course to a common value, compensating for differences in the number of points or overall scale in each experimental time course. Without this term one reaction time course will often times be weighted over another, resulting in fits which not only are poor for the less weighted data set, but may inadvertently weight the determination of one kinetic parameter over another.

A nonlinear least squares regression equation can be expressed as:

$$Y = F(\beta_0, \beta_1 \ldots \beta_K;\ X_1, X_2, \ldots X_n) + \varepsilon \qquad (2)$$

where ε is the residual error, β's represent all the unkown nonlinear parameters, and X's represent the various independent variables of a function, F .It is clear from this equation that an inadequate number of β parameters will result in a larger ε, while if a model perfectly fits the observed data ε is equal to zero. From the initial estimates of the β parameters, corrections can be applied which result in a better fit of the observed to the simulated data. A better fit is determined as a reduction in the value of ε. The newly calculated β parameters can be refined by applying this procedure repeatedly, a process known as iteration. When these corrections become negligible and no further reduction in the SSQ is possible, convergence has occurred.

While all techniques of nonlinear regression proceed essentially by such a process, the search technique used for parameter optimization, or for the lowest value of ε, can vastly differ. This search process corresponds mathematically to an exploration of a surface relating the error to the parameter values, while parameter optimization consist of finding the minimum in the error surface. When several parameters are involved, information relating the interdependence of the parameters on the error is required to properly choose the next parameter estimate. Such information can be obtained by partial derivatives, which allows the determination of the error gradient vector.

There are several review articles on regression analysis pertinent to properly using FITSIM and a few have been listed in Appendix I. The best of these may be a review article by Mannervick. Generally, a good mechanism will result in convergence in the regression analysis, meaningful parameter values with low standard deviation (less than 50%), and give residuals which show random distribution about the zero level. In most cases, if FITSIM refuses to find a better fit and/or gives large errors in the obtained rate constants, the mechanism and data files should be examined carefully before assuming regression is failing.

5.2 General Operation.

FITSIM is designed such that it can be run either as a subprocess or as batch file (batch running is controlled by the program FITSIMB) since relatively long periods of time may be required for fitting. It is STRONGLY recommended however the user supervises the initial fitting attempts by FITSIM to a new mechanism. A multitude of problems and/or errors and warnings can be output by FITSIM during the initial fitting attempts. While some of these errors are fatal causing execution to stop, some errors can be detected that may be a result of wrong files entered, bad initial rate selection, or others. These are called to the users attention, but execution is not stopped. Such errors could result in a large amount of CPU time being wasted. If FITSIM is ran as a batch process, perhaps a few iterations should be tried at the interactive level. Generally these warnings will occur with the very first iteration tried.

FITSIM can also be linked with the other programs used in plotting and simulating data (i.e. KINSIM, KINCOMP, FITDATA, and SFDISP) by using the SIMUL command file (see Chapter 2). This type of operation is possible only when running the programs under VAX/VMS. Extensive on-line help is available to the user when running FITSIM under the SIMUL command file, and

if your installation runs VAX/VMS, this is probably the best method of using FITSIM.

As described below, there are four essential steps required to use FITSIM. Each step is further explained in Chapter 6 of this manual.

1. Mechanism Entry and Compilation. The user can type a representation of a chemical scheme using conventional chemical formulation. Also encoded are output equation calculations which can allow for complex applications. This feature also allows for fitting of more than one output parameter for a given reaction time course, although a new data set and save set would have to be entered for this type of fitting. Mechanism entry and compilation is identical to that described previously in more detail for KINSIM (Barshop et. al., 1983; Chapters 3-4, this manual). A given mechanism has to be entered and compiled only once, thus this step can be skipped for subsequent runs of FITSIM.

2. Entry of Data Parameters. For each data set, or experimentally observed reaction time course, the kinetic parameters and reactant concentrations pertinent to that data set must be encoded into the form of a KINSIM save set (.SAV). The save sets will include all reactant concentrations, output scales, and the initial kinetic constants. Only the kinetic rate constants of the LAST save set is actually entered into FITSIM. Therefore if different initial rate constants are desired, only the LAST save set needs to be changed. Changing of output factors or integration times, however, requires changing these values in ALL of the save sets.

3. Entry of Regression Options and Parameters. An interactive program, FITDATA, prompts for the necessary input and gives the various regression options. Alternatively these could be coded for manually in a data file by the user. Information required includes the upper and lower limits for the varied kinetic constants, reaction profiles to be fitted and their corresponding data parameters, data weighting method, regression methods and parameters, and convergence terminators. Specific rate constants in a mechanism can also be linked together. This allows the best fit to be obtained while maintaining user-entered ratios between specific rate constants. A more complete description can be found in the FITDATA program section.

4. Fitting. FITSIM initializes itself and then loads in the mechanism, save sets, and data files. To aid in the detection of possible errors in the loading of files, the user-supervised regression can be selected for (see FITDATA section). It is suggested to the user that the OBSERVE qualifier, or possibly regression option 9, be not overlooked as a powerful tool in determining where errors are occurring. This user-input regression routine allows one to see the plots and datasets as FITSIM internally sees them. Often obvious errors in the DATa sets or SAVe sets can be detected in this manner. It is critically important that the proper time scales and output factors are set in the Save sets, otherwise errors can develop. Regression option [8] should also not be overlooked when beginning to fit data to a new mechanism and using linked rate constants. This option allows the user to vary a linkage group to examine whether the data is indeed sensitive to variation of the kinetic parameters in the Linkage group. If the kinetic constants being varied are insensitive to the data being fit, FITSIM operation will be inefficient, and multiple

warning messages will usually occur.

The chosen regression algorithm calculates new values for the kinetic constants, the resultant SSQ calculated, and iteration proceeds until either the convergence parameters are met, or the net reduction in the SSQ reaches some minimum value. At this time iteration stops and the new kinetic parameters for each reaction time course is saved. All information about the regression as well as values of all kinetic constants are printed to a LOG file created with each running of FITSIM.

5.3 Execution of FITSIM.

Required Files. Only one additional file is needed over that of using KINSIM manually to fit experimental data. Data files, Save files, and Mechanism files are in the corresponding formats for these files used in KINSIM. The formats and use of these files are explained more completely in Chapters 3 and 4. The additional file required is referred to as an AUTOSIM file (for AUTOmated SIMulation data file) or *.AUT file which can either be encoded by the user or set-up by the program FITDATA. The format of the AUTOSIM file and the options available in FITSIM are explained in detail in Section 6. This file can be conviently called any name.

Data File Requirements. Any data file in the SF structure read by KINSIM can be used. The structure and requirements for this file can be found elsewhere in this manual. In FITSIM, as with KINSIM, the baseline contained within the SF data file is first subtracted from the data at time t. Data files with negative values in some instances can lead to problems, and the user should select for the correct treatment of negative data values (see Chapter 6). As of the present version of FITSIM, simple interpolation is done between data points. Therefore for good fits to be achieved data files should not be particularly noisy or have large spikes present. Nevertheless, fitting of the data under these circumstances usually still occurs, although errors in the determined kinetic constants may still be large. If such problems do exist, the user is encouraged to edit such spikes out or use an appropriate algorithm to smooth the data.

All values for data files and simulations are in double precision to prevent as much as possible the rounding of subtraction differences between files leading to small differences in the final kinetic constants obtained. Nonetheless, just changing the order of the files in the AUTOSIM.dat file can make small differences in the final values to be obtained. Differences in the final constants obtained and the final SSQ, however, should be quite small and is usually of no concern. For the best possible regression to be achieved, the data files should be entered in order of their initial fits. Those data files with the worst initial fits should be entered first and the data file with the best initial fit entered last. This procedure should help to reduce the amount of rounding error involved in the subtraction of the real from the simulated data files. It should be noted that the first experimental point in data files loaded by FITSIM is assumed to correspond to the first point obtained in a reaction time course (Delta time).

Save File Requirements. The save file (those files usually ending in the extension *.SAV) have the same structure as those used in KINSIM. The

amount of reaction time for each progress curve is determined by the the value found in the save set and NOT the data file. If the data set contains a substantial amount of data near where the reaction is completed the total simulated time should be shortened to the amount of time required for completion of the reaction. Failure to do so may result in over-weighting of the end points, causing an overall poorer fit of the data to the kinetic mechanism being tested.

The time period between points, or the delta time, is also determined by the value given in the SAVE set. This value will also determine the total number of data points used. The number of data points used in a particular progress curve is not equal to the number of points in the data file, but rather is equal to the total simulation time divided by the delta time plus 1. Setting the delta time too small will force excessive computation without any improvement in the accuracy of the kinetic simulation. Therefore it is recommended that the delta time be set generally so that no more than approximately 200 points in a given data set are used. Excessively large delta times, resulting in fewer than 25 points being calculated, should also be avoided.

Start-up. There are three main ways of executing FITSIM; by submission as a batch process, (i.e. through running the command file FITSIMB), interactively by running the EXEcutable file directly, or through the command program SIMUL. The advantage of running the program through SIMUL is that on-line help is available as well as other SIMUL programs such as SFDISP and KINSIM (see chapter 2). The advantage of running the program in batch mode is possible cheaper computing cost and over-night execution of the program. When execting FITSIM directly, or through SIMUL the following screen appears :

```
-------------------------------------
| FITSIM                Version 1.63 |
| Regression FITting by SIMulation |
-------------------------------------
```

Qualifier :

When execting FITSIM directly, a qualifier of AUTOSIM file can be immediately placed after the symbol name and the prompt for the qualifier eliminated. See the description below on input parameters for more details.

5.4 FITSIM options.

Qualifiers. Under VAX/VMS there can be three qualifiers that can be used when invoking the FITSIM program. These qualifiers are placed directly after FITSIM. Any other name placed after FITSIM is assumed to be the name of the regression input file (AUTOSIM or *.AUT format file). If no file name is listed, the user is prompted for a Qualifier. Entering nothing sets the default *.AUT file name which is AUTOSIM.DAT. This feature allows the creation of any number of regression files in the *.AUT format. Names of the regression input file CANNOT, however, correspond to any of the qualifiers listed below.

The OBSERVE qualifier invokes an option similar to specifying the user-regression iteration routine (option [9]). The major differences between the two is the "OBSERVE" qualifier does not create a log file or allow the entering of kinetic constants by the user. This qualifier allows the user to easily check the entering of all files and to visually see the initial plots without re-editing the appropriate *.AUT file. To implement this option simply type 'FITSIM OBSERVE'.

When the OBSERVE qualifier is used, the user is first asked to enter the name of the *.AUT (AUTOSIM) file before execution continues. After this name is entered, each data file entered in the AUTOSIM file is plotted. The user at this time should check carefully the plots presented to determine if FITSIM is accurately reading in the data. For particularly "noisy" data files or data files with spikes, choosing a larger or smaller time increment may result in a smoother curve. At the top of the screen for each plot the total time and the increment time is printed along with the unweighted SSQ between the simulated and real data files. Check these values to insure they are what is expected. It is all to easy to enter the wrong information in the SAVE files.

After all plots are presented, the kinetic constants read in by the LAST save set is printed. These kinetic constants represent the initial and fixed values used by FITSIM. Entering <LF> at this point will present a prompt which ask whether a printout of the entered points is requested. Entering [Y] will present a screen of 6 columns of numbers. These columns correspond to the Dataset number to which the data point has been assigned, the calculated data value at increment T, the simulated data value at increment T, the assigned weight at increment T, the time value corresponding to these data values, and the assigned data point number. These represent all the internal values read in by FITSIM and is useful in checking the data matrices in the program.

After the user indicates whether to see the data the various nonlinear regression parameters selected for are printed to the screen. If there is an error in the limits set for the varied kinetic rate constants this is indicated. Again, the user should check the various nonlinear regression parameters selected for.

The PLOT qualifier is similar to the OBSERVE qualifier except VERSAPLOT files are created to allow hard copies of all fits to be obtained. This qualifier is particularly useful when final fits of the simulations to the data files are needed in hard copy.

The FIT qualifier is also similar to the OBSERVE qualifier except the *.SAV files created by FITSIM, FIT_*.SAV, are used. This option allows the user to conviently view the final fits of the experimental data obtained by FITSIM.

<u>Regression & Output Options</u>. There are numerous options for controlling the nonlinear regression optimization of KINSIM-generated simulations to observed data. The various options are discussed in detail in Chapter 6. One regression option (Variable Check - [8]) puts FITSIM in a mode which is different than the rest, and thus will be explained in more detail here. This mode allows the user to manually enter ratios to change the rate constants, and calculates the difference in the Sum of Squares this change makes. This helps the user in determining whether the rate constant (or constants if linkages are involved) is sensitive to the observed data. A sample screen when this option is invoked for the BIBIORDER test mechanism is shown below :

```
     Kinetic Constant Variation
-------------------------------------------------------------
     The Sum of Squares for all files = 0.68950


     Which variable do you wish to alter < Return to Quit > [1] : 1
     For Variable   1 the FOLD Variation is [ 100.0000 ] :


     For a   100.000 change in Variable   1 :
           The Change in SSQ is 1.2018
           Fold change due to this variation is 1.7431


     Which variable do you wish to alter < Return to Quit > [2] :
```

The total Sum of Squares (SSQ) is added for all the data files entered and the user is requested to pick the variable in which he/she wishes to vary. After selection of the fold change a new SSQ is calculated and the fold change in this variation is also printed. If the fold change in the variable is relatively large (greater than 100) while the fold change in the SSQ is at or very near 1.0, changes in the variable in question has little or no effect on the fit and the user should consider fixing this variable at some value. The variable number corresponds to the kinetic constant listed in the FITSIM log file and in fact may include several rate constants if linkages are involved.

5.5 FITSIM Limits of Operation.

There is undoubtedly a practical limit to the number of parameters which can be varied before FITSIM cannot effectively solve for kinetic parameters which yield the best fit. The maximum upper limit for the number of varied parameters depends in part on both the number and differences in experimental conditions of the data files to be fit, the quality of the data, and how close the initial starting values for the kinetic parameters are to the final ones. In the implementation used here the maximum number of varied parameters is 20, while the maximum number of data files is 8. These numbers can be increased relatively easily however, although at the cost of more memory and CPU time. Generally, the number of varied parameters should approximately equal that, or be less than the number of experimental time courses to be fit. Furthermore, the varied parameters should be checked to insure simulations are sensitive to changes in the varied parameters. If the data is not sensitive to the kinetic parameters varied in FITSIM, little or no improvement in the SSQ will be obtained and the determined parameters will have large errors associated with them.

The number of varied parameters can be reduced by linking rate constants together. If for instance there is a known ratio between two or more rate constants, this ratio can be encoded for in FITSIM. Only the first of the linked rate constants will be varied, all other constants will be calculated from the ratios set by the user. Upper and lower limits to the varied rate constants can also be entered which binds the varied parameters between specific values. The range between the lower and upper

limits should be as low as possible.

The system described here is flexible enough to allow testing of widely different mechanisms by statistical means, although one must decide first which mechanism to test. Under certain conditions in nearly any relatively complex mechanism, a wide range of values for certain rate constants may be allowed. Thus it is possible that incorrect mechanisms could fit a given set of data, provided an insufficiently wide range of experimental conditions are tested. This not only can usually be avoided by selection of the proper range of experimental conditions, but the determined rate constants will have a large error associated with them.

5.6 FITSIM Error & Warning Messages.

This section explains the various errors that are output by FITSIM. Some errors will have obvious solutions, while solutions to other errors may be difficult to determine. The user should have an understanding of his/her data sets and the kinetic constants which are being tested by the data. Especially in complex mechanisms, or when rate linkages are involved, such knowledge may not be obvious. When multiple warning messages are output by FITSIM, or large errors are associated with the determined rate constants, the user should carefully check whether the rate constants can indeed be determined with the data sets being used. Two regression options available to the user (option [8] and [9]; see Chapter 6) are designed to help the user determine whether the type of data collected are correct for obtaining values for the rate constants being varied.

Error Messages due to FITSIM input errors. The following errors occur when input does not match what FITSIM expects to find. These errors occur before iteration begins. They are almost always due to errors in the AUTOSIM file, or inconsistencies in the SAVE sets. Execution is immediately terminated when these errors are discovered.

Fatal Read Error :

Execution is terminated because FITSIM was unable to read the named file or parameter necessary for proper execution of the program. Check to see if the named file or parameter is correctly spelled out within the AUTOSIM file, exist in the correct format, is missing, or is in the directory being accessed by FITSIM.

Grouped Rate is Zero :

Execution is terminated because a rate constant was grouped or linked with another rate constant having an intimate value of zero. This situation must be corrected in order for FITSIM to run.

No Reduction in SSQ, Data not Sensitive to Varied Kinetic Constants :

No reduction in the sum of squares could be obtained upon initial variation of the rate constants. This could be the result of : 1) the data not being sensitive to the rate constant(s) being varied; 2) the starting values for the rate constants are extremely remote from the actual solution; or 3) the data is being converted to zero by the conversion of negative values (see FITDATA program options). Regression option [8] can be used to confirm this error.

Number of Points over Max allowed :

Execution is terminated because the number of points exceeds the limits of FITSIM being able to handle. In the present implementation the maximum number of points is 3000. To correct this error, simply reduce the delta time, or number of points displayed per delta time in one or more of the SAVE sets.

Sum Of Squares is ZERO :

Error when the calculated sum of squares is zero. This could occur for two reasons. Either the fit to begin with is perfect or there is problem with the calculation of the output resulting in zero values.

Warning Messages due to FITSIM input errors.[24m The following Errors occur when input does not match what FITSIM expects to find. These warning messages occur before iteration begins. They are almost Always due to errors in the AUTOSIM file, or inconsistencies in the SAVE Or data sets. Program execution continues.

Input Values Exceed Limits :

This warning message occurs when the input kinetic constants exceed the limits set by the user. This is probably due to user-error in either entering the rate constants or rate limits. Improper running of FITSIM is usually the result when this warning is displayed.

More than 10% of Data is negative :

A warning printed when negative data points compose more than 10% of the total data entered and the absolute values of these points IS NOT being taken. Check to make sure you really expect either a lot of data values near zero, or that are negative. Often times this detects errors in the dataset due to incorrect background subtraction. This can occur when the baseline value used is near the Y-maximum of the data set. As the general plotting program used at WUMS-BCCF (SFDISP) plots absolute values, the fact the data points are negative may not be obvious. If this is the case, correct the baseline value such that is is at or near the Y-minimum, or change the output factors of the mechanism such that negative values are output.

Simulation time exceed that of the actual data :

This warning is printed when FITSIM detects that an insufficient number of data points exist in the experimental data files. The simulation time desired is beyond the longest time point in the experimental data file. Any time points beyond the maximum experimental time is ignored by the program.

# of Data points in Experimental File smaller than expected :

This warning is printed when the number of data points in the experimental file is smaller than what is expected by FITSIM. This may indicate a problem with the data matrix.

Probable No time zero point :

The usual cause for this error is that a time zero point does not exist in the experimental data file. This error can be corrected internally within FITSIM by using the shift time option [9] for the data point at time zero. Using this option, however, will not prevent this warning from being printed.

Regression Warnings. The following warnings occur when the regression algorithm is having difficulty in finding better values for the varied constants. These warnings MAY indicate the final fits do not represent the true global minimum. These warnings may vary depending upon the nonlinear regression algorithm chosen. Those below are for the Marquardt option. Execution in all cases continues, although the warning message is output to the LOG file.

Avoiding a Constraint :

Warning message displayed when at least one of the parameters varied exceeds the limits placed on it. Be sure that the constraint placed on that variable is realistic before allowing execution to continue.

Possible Dangerous Value of Coefficient (X1, X2) :

Warning message printed when changes in the varied parameters are difficult to calculate due to the error surface of the function being nearly flat. The values X1 and X2 give the variable numbers which are causing the problem. By correlating the variable number to the varied kinetic constants, as given in the user log file, the user may identify the kinetic constants causing the error.

This warning may suggest a global minimum will not be reached. Be careful that the result obtained does appear to fit the data. Possible solution to avoid this warning message is to try different initial kinetic rate constants, or to utilize the masking option. This warning appears usually when the on and off rates of a given reaction are allowed to vary simultaneously.

Linked Group (X) has been set to zero :

This warning occurs when linked groups are being used and can occur regardless of the nonlinear regression option chosen. This error occurs if the varied rate constant in a linkage group is set to zero by the regression routine. When this occurs, all linked rates associated with the varied constant will also be set to zero. Such an event may yield changes, in effect, in the actual mechanism being fitted.

Simulation (KINSIM-generated) Warnings. As of version 1.4 FITSIM was designed to continue even under conditions where KINSIM terminates (see Chapter 4). The reasoning behind this was relatively small errors might otherwise terminate the fitting process prematurely. In most cases the following errors occur when FITSIM encounters a terminal KINSIM error, but execution is allowed to continue. The user should, of course, check the final fits to see if correct convergence has occurred. In particular, cases where these errors occur during the final set of simulations should be thoroughly examined as calculational errors were involved in determining the simulations.

Terminal Error DGEAR :

The Tolerance was adjusted up to a value of 0.1 before execution was terminated because the solution still could not be found. Reduce time constants and/or check the mechanism for faults.

Tolerance Increased By 10 :

Execution continues with a tolerance increased by 10. With KINSIM,

this would result in a DGEAR error (see above). This warning is only displayed when tolerance values become greater than 0.001 . It is generally considered that tolerance vales greater than this may yield errors in the kinetic simulations, so be CAREFUL! Generally if this warning is displayed only very occasionally, no problems will occur in the regression procedure. If the tolerance becomes greater than 0.1, execution will terminate with a message of a "Terminal Error DGEAR" (see above).

Simulation (KINSIM-generated) Errors. For a complete explanation of these errors see Chapter 4. When one of these errors occur execution is terminated, however, the present values for the kinetic constants are output to SAVe sets. Usually execution is terminated due to overflow errors which can be corrected simply by changing the time factors in the kinetic simulation which failed.

Other Errors :

FITSIM is still in development in the Frieden lab, and is expected to be in this state for some time. At this time most major bugs appear to be exterminated, and for relatively complex problems, FITSIM appears to work quite well. Any correspondence to us should include the version number of FITSIM which you are presently using (Who knows, we may have found the problem and corrected it before you did !), an EXACT description of the problem, and FITSIM-generated outputs demonstrating the problem you are having. Many times such outputs may be useful to you in finding that incorrectly entered data is the problem, and not the program (see the sections on the various outputs which are available to the user).

Future development and enhancements will be centered around the obtaining of better regression algorithm routines, the ability to fit output factors, and user options to change the regression output to his or her needs. Furthermore, it would be appreciated that we be notified of any useful modifications performed by another user such that we may incorporate these either into future versions of the program, or make them available to other users.

5.7 FITSIM Output.

The output of FITSIM is to a sequential or log file named according to the user-entered name listed in the AUTOSIM file. This file can be subsequently read or printed. If no name is listed the default name for this output file is AUTOSIM.LOG. This file contains the initial values of the varied kinetic parameters, as well as the values at the end of each iteration cycle. Also output in the LOG file for archival purposes are the names of the *.DAT (data) and *.SAV (save sets), a printout of the mechanism fitted, and how the simulation output was calculated. In addition all the regression control options are printed to this file. A sample output is given in Section 5.8 for the BI-BI ordered test files described in section 7.3 .

Convergence is achieved when either changes in the SSQ reaches the set minimum value (Changes in SSQ has reached the MIN. set), or changes in all

other varied parameters reaches the set minimum value (Convergence by Small step size). The convergence criterion met is entered in the LOG file. Generally convergence due to the later reason is a better indication that true convergence has been achieved.

When convergence has been achieved, both the final values and the standard error for each varied parameter are output along with the remaining weighted SSQ for all data files combined, as well as for each individual data file. The SSQ listed for the individual data files are not calculated using weights and is divided by the number of simulation data points. Thus this value represents the average difference between the simulated and experimental data files. Fits are saved into files called with the same name as the entering *.SAV files but have the added prefix FIT_. These files are conveniently displayed after convergence when the output option is greater than 10. This output mode cannot be used when using FITSIM in batch mode.

At the end of iteration an Analysis of Variance (ANOVA) is performed. The value for the maximum deviation of the data from the fit should routinely be checked and should, in most cases, be a small value. If this value is unexpectedly large, check the data sets for a point or points which are outliers. Such points should be discarded as they could alter the final kinetic parameters found by FITSIM. The R-squared value is the ratio of the regression SSQ to the total SSQ. A value of 1 indicates a perfect fit while a value of 0 indicates "perfect" randomness. This value is a method of "standardizing" how good a fit is since the residual SSQ can only be used as a relative value.

Two criteria should be used to decide the "best-fitting model". They are 1) qualitatively, the simplest model or the one with the fewest variables, and 2) quantitatively, the lowest Mean Square Error (MSE). The MSE is equal to the SSQ divided by the degrees of Freedom (DF). MSE is a good indicator for comparison of models with various numbers of variables. As the number of variables in a fitting function or model increases, a better fit of the experimental data by this expanded model, as defined by the SSQ is expected. However when the SSQ is divided by the degrees of freedom, a relative fit is defined (MSE) that allows for the direct comparison of models of various complexity.

Unfortunately neither of these more standard ways of calculating a "goodness of fit" is entirely appropriate for fitting of full time courses by kinetic simulation. Therefore we have introduced a factor Q which may better describe the "goodness of fit". This empirical equation was formulated to numerically grade the quality of the simulated fits to the real data and is described by:

$$Feqk = 2(0.5 - SD*F)$$

where Q is the measure of the fit quality, SD is the standard deviation of the fit and F is the numerical output factor required to standardize the time courses such that the maximum output is equal to the value 1. Therefore Q can range from a value of -1 to 1 where the value of zero would describe data which is perfectly random. Negative values would describe an inverse relationship between the simulated fit and the data. A value of 1 would describe a perfect fit of the kinetic simulations to the data.

5.8 Description of the Output Log File.

Listed below is a sample output LOG file which is made after every run Of FITSIM. The output below is that found for the BIBIORDER mechanism included in the sample files (Section 7). The output which is described below is for an output option of [1]. Key letters for the descriptive comments are enclosed in brackets to the right of the page.

```
=========================================================================
 FITSIM   Simulation Regression Log File                    Version  1.60   (A)
      Log Time : 26-JAN-1988 12:50:47.97                                   (B)
      Log Comment :  FITSIM test using a bi-bi ordered mech
=========================================================================

 Filename for Regression Info : AUTOSIM.DAT
        Name of this LOG file :  bibiorder_test.log
               Mechanism Name :  bibiorder.SIM


 Mechanism Description :


                          Ordered Bi-Bi mechanism TEST


               K+1             K+2           K+3             K+4
          E + A == EA ;  EA + B == EAB ;  EAB == EQ + P ;  EQ == E + Q
               K-1             K-2           K-3             K-4


                              OUTPUT 1 = F1*Q/POT


 There are 8 rate constants in this kinetic mechanism


-------------------------------------------------------------------------

 Files used in this Regression Analysis :                                  (C)

 Data set # 1 :
    Data :  bibiord1.dat                 Parameter :  bibiord1.sav
    Simulation Time :    1200.           Delta Time :    12.00
    Baseline Value :   0.0000E+00        Output # :    1.000
    Integral Tol :   1.0000E-03          Flux Tol :    0.1000
    Max Y-value :    1.000               Number of Pts :  101
    Data Comment : SIMUL OUTPUT FILE


 Data set # 2 :
    Data :  bibiord2.dat                 Parameter :  bibiord2.sav
    Simulation Time :    50.00           Delta Time :    1.000
    Baseline Value :   0.0000E+00        Output # :    1.000
    Integral Tol :   1.0000E-03          Flux Tol :    0.1000
    Max Y-value :    1.000               Number of Pts :   51
    Data Comment : SIMUL OUTPUT FILE
```

```
Data set # 3 :
   Data :  bibiord3.dat                  Parameter :  bibiord3.sav
   Simulation Time :    1000.            Delta Time :    5.000
   Baseline Value :   0.0000E+00         Output # :    1.000
   Integral Tol :   1.0000E-03           Flux Tol :    0.1000
   Max Y-value :    1.000                Number of Pts :  201
   Data Comment : SIMUL OUTPUT FILE


Data set # 4 :
   Data :  bibiord4.dat                  Parameter :  bibiord4.sav
   Simulation Time :    1000.            Delta Time :    10.00
   Baseline Value :   0.0000E+00         Output # :    1.000
   Integral Tol :   1.0000E-03           Flux Tol :    0.1000
   Max Y-value :    1.000                Number of Pts :  101
   Data Comment : SIMUL OUTPUT FILE


-------------------------------------------------------------------------


Report of rate constant Regression Variations :


Fixed constants :                                                        (D)
              #     Index     Label          Value
              1.       1        K+1           10.
              2.       3        K+2           10.
              3.       8        K-4           10.


Free constants :                                                         (E)
   # Index Label    Upper       Lower         Delmn       Value
   1.   2   K-1     1.00E+03    0.10          0.10       1.00E+02
   2.   4   K-2     1.00E+04     1.0          0.10         50.
   3.   5   K+3      1.0        1.00E-04      0.10        0.35
   4.   6   K-3      10.        1.00E-04      0.10       1.00E-02
   5.   7   K+4     1.00E+05     10.           1.0       1.00E+03

There are  0 Variation groups.                                           (F)

=========================================================================
       Nonlinear Regression Parameters Used for this Analysis            (G)
=========================================================================


        Number of Variables to fit :  5
Total Number of Data points Fitted :  454
                       Output Type :  1
       Nonlinear Regression Method :  Marquardt
     Derivative Calculation Choice :  1
     Maximum Number of simulations : 100
             Zero Time Calculation :  Not Used
Minimum SSQ change for Convergence :   0.0000
 Min Parameter change for Converge :   0.0010
            Type of data Weighting :  Identical
Value assumed for Background Noise :  0.00000E+00
 Treatment of Negative data values :  Made zero
                           Masking :  0
```

```
            INITIAL SUM OF SQUARES :     21.559631                        (H)

 ---------------------------------------------------------------------

 CORRECTION = -16.5    -0.751      -6.631E-02  -1.660E-03    158.        (I)

          X =  83.5      49.2       0.284       8.340E-03   1.158E+03     (J)

 END of Iteration    1 Summary :

      SSQ Reduced from   21.559631     to   12.997775
      % Reduction in SSQ this Iteration :   39.71
      # of KINSIM Simulations :      7
      Marq Lambda Value :    1.000
      FOLD Reduction in SSQ :      1.7

--------------------------------------------------------------------

             Iterations 2 -> 12 printed here

--------------------------------------------------------------------

 CORRECTION = 1.523E-03   7.054E-03   6.702E-07   3.108E-04    159.

          X =  10.0        500.       0.250       9.987E-03   4.993E+03

 END of Iteration   13 Summary :

      SSQ Reduced from  6.69092760E-07 to  3.33291549E-10
      % Reduction in SSQ this Iteration :   99.95
      # of KINSIM Simulations :     82
      Marq Lambda Value :   1.0000E-07
      FOLD Reduction in SSQ :     6.47E+10

====================================================================
                       ANALYSIS OF VARIANCE
====================================================================

 STANDARD DEVIATION of Data from Fit =  4.92471E-07
 Weighted STANDARD DEVIATION of Data from Fit =  8.61569E-07

 MAXIMUM DEVIATION of Data from fit = 2.38419E-06                        (K)
 MAXIMUM DEVIATION of Weighted Data from fit =  5.36442E-06

 R-Squared Value =1.0000           # of KINSIM calls = 82
 F-Ratio =  6.34E+13
```

SOURCE	DF	SUM of SQUARES	MEAN SQUARE	(L)
MODEL	5	235.	47.0	
ERROR	449	3.333E-10	7.423E-13	
TOTAL	454	235.		

Approximate STANDARD ERRORS of Parameters :

VAR	INDEX	Label	Value	Standard ERROR
1	2	K-1	9.9999	6.45236E-05
2	4	K-2	500.00	2.36334E-04
3	5	K+3	0.25000	3.98511E-08
4	6	K-3	9.98652E-03	7.53232E-07
5	7	K+4	4993.2	0.38795

SSQ / (# of points) for each fitted file : (M)

1 : 5.77E-09
2 : 1.35E-08
3 : 6.05E-10
4 : 1.07E-09

END LOG OUTPUT ----------

Key to Descriptive Comments :

{A} Version of FITSIM you are using.
{B} Gives the time that program execution began.
{C} Data & Save file descriptions. The Max Y value is for the included data file while the output factor is the numerical factor used in the simulations. Generally the value of these two constants should be similar. The number of points given is for the number generated by the kinetic simulator and is NOT necessarily related to the number of points in the data file.
{D} Gives the rate constants whose values are not varied. The index number corresponds to the variable number assigned to this rate constant by the program. This number is assigned by making the first rate constant in the mechanism equal to 1, the second rate constant equal to 2, and so forth.
Forward rate constants are assigned first.
{E} Gives the rate constants whose values are allowed to vary and the limitations assigned to this variation.
{F} Gives the rate constants whose values are allowed to vary, but are linked to other rate constants. See section 6 for more detail.
{G} Regression options used. See section 6 for more detail on the meaning of these values.

{H} The initial WEIGHTED sum of squares before attempts to reduce this value commences.

{I} The correction to the varied rate constant. The freely varied rate constants are printed in numerical order followed by those rate constants which are linked.

{J} The value of the varied rate constant at the iteration number stated. If a terminal error occurs after this point, the values for the kinetic rate constants correspond to the lowest SSQ thus far achieved.

{K} Maximum deviation of the data from the fit. This is useful in determining if an outlier may exist in a data set.

{L} A summary of the errors and fitting parameters. The Mean squared value for the error is useful in comparing fits.

{M} The values printed represent an "average" error between a time point in the experimental data set and the simulated data set. A smaller value generally corresponds to a better overall fit.

Chapter 6 User Interface Programs for FITSIM

6.1 General Description & Use.

There are two auxiliary programs for FITSIM. The first program, FITDATA, is a user friendly interface to FITSIM which creates an AUTOSIM file. The AUTOSIM file is necessary for passing various regression parameters, file names, and save sets to FITSIM. Alternatively the file can be manually encoded by the user in the format shown in Table 6-1 and the need for FITDATA eliminated. At the present time, this is the only alternative for users not operating FITSIM under VAX/VMS. No transportable version as of yet exist for the program FITDATA. FITSIM, however, should detect most errors that occur due to the encoding of wrong input information.

The second program, FITSIMB, is a short VAX/VMS command program which allows the running of FITSIM as a batch file. This program should only be used after the user becomes familiar with the operation of FITSIM.

6.2 FITDATA File Parameter Inputs.

In this section the prompt given by FITDATA will be underlined followed by an explanation of the option. Anything in brackets listed below represents the default setting. Pressing <ENTER> at this point will load the default setting. First, executing the program FITDATA will clear the screen and the following printed :

```
 -------------------------------------
| FITDATA                Ver. 1.32    |
| FITting by SIMulation DATA entry    |
 -------------------------------------
```

1. Enter a NEW mechanism (Y/N) .

Default is [Y]. If a new mechanism is entered [Y] should be selected. If [N] is selected the user is prompted for the name of an AUTOSIM (*.AUT) format file to be loaded into memory [AUTOSIM.DAT] an a name for the new file. If an "old" mechanism is loaded the default settings become the retrieved values for each of the regression options. When loading an "old" autosim.dat file all regression options can be changed and a new AUTOSIM.dat file created. These options are indicated by (*) below. Loading a previous autosim-type file also allows review of the parameters coded for in the AUTOSIM file in a more readable manner.

2. Available FITSIM Regression choices [1] : (*)

[0] - Gauss-Newton Algorithm. In the present implementation, this uses the same routine as [1], but sets the Marquardt lambda parameter to zero.

[1] - Marquardt Algorithm. Under most conditions this should be the best algorithm to use. However, if this is particularly slow, one of the other option methods might be tried.

[2] - Stept. This is regression by "brute force" and is usually quite slow. Each variable is changed by specific user-entered increments until a smaller sum of squares is achieved. This regression method is not recommended and mainly serves as a stub for a future regression option.

[8] - Variable Check. This allows the user to quickly alter the kinetic constants that are varied in FITSIM by a user-entered factor. Although this option is similar to [9], this option was designed to allow the user to examine whether the data is sensitive to the varied kinetic constants entered though the AUTOSIM data file. Although this is relatively easy to determine with simple mechanisms, when complex mechanisms with linked rate constants are used in FITSIM, such determination can be difficult. A more complete description of this option along with a sample output can be found in Chapter 5.6 .

[9] - User-Supervised. This is "regression" by user-entered changes in the various rate constants. All files entered are plotted along with the respective simulation calculated by the corresponding SAVe set. At the end of each set the user is presented with the table of kinetic constants. At this time the user may enter different values for the constants. This mode of FITSIM can be used as a check routine to make sure the files and save sets are being entered correctly by FITSIM. This regression option can also be used to quickly examine sets of fits after regression has been terminated. In this case the SAVe sets should be replaced by FIT_*.SAV sets.

The default regression routine is [1], and the value for this option is listed in the first line of the AUTOSIM file.

3. Mechanism Entry> Enter file name :

The mechanism to which the data will be fitted to. This mechanism is edited and compiled exactly as described for KINSIM. Once this mechanism is entered, it can be conviently displayed by the keystroke <Control-C>. There is no default mechanism name. The mechanism name is listed in Line 2 of the AUTOSIM file. When the mechanism has been entered into memory correctly a prompt of <Mechanism Loaded> appears.

4. Rate/Equilibrium Constant Variation.

k_{+1} = 0 k_1 = 0
k_{+2} = 0 k_2 = 0

ENTER : Varied = 0, Fixed = -1, Linked = 1,2,...n (by group)

All rate constants in the mechanism are listed and are initially assumed to be varied (value = 0). To fix a given constant enter the value [-1] when the cursor is on the line next to the rate constant in question. Relationships, or linkages, between rate constants can be coded for also. If for example a particular rate constant is always 2-fold over another, this linkage between these rate constants can be entered in FITDATA. To link rate constants enter a positive integer starting with [1] for all rate constants in a given linkage group. One group of linked rate constants can be identified by entering the number [1], while another group of linked rate constants can be identified by entering the number [2]. Every group of linked rate constants should have a unique integer number. Linked rate constants allow specific ratios between k_+ and k_- to be maintained. FITSIM will only vary the first rate constant in a given linkage group and calculate the remaining rate constants on the basis of the ratios entered (see Linkage). To allow a parameter to freely vary leave the value at the default, [0]. It should be noted that unless the data is sensitive to the rate constant being varied, FITSIM will not work. If, in complex mechanisms, you are unsure whether the data is sensitive to the rate constant, use regression option [8] to check.

4a. Rate Constant Linkages.

Enter the PROPORTIONAL relationship for the Linked rates in group 1
k+1 = 1.00000E+00
k-1 = 1.00000E+00
NOTE: Constant with Lowest Proportional Value is varied.

If rate constant variation options greater than zero are entered, all rate constants with the identical integer value are assumed linked, and proportions for these linkages must be entered. The value entered is the value which when multiplied by the first rate constant in the group gives the initial value for the linked rate constant in question. Thus if the two rate constants k_{+1} and k_{-1} were linked and the initial values were 1 and 100 respectively, the linkage for k_{-1} would be 100 (i.e. k_{-1} is 100 fold greater than k_{+1}). Furthermore, in the Linkage group menu any ratio can be used (i.e., instead of 100, k_{-1} could be equal to 1, while k_{+1} would be set to 0.01). As FITDATA automatically scales the linkages, with the lowest being 1, when manually making an AUTOSIM file the first constant in the linked group must have a ratio of 1 (see Table 6-2).

5. Rate / Equilibrium Constant Limits : (*)

k_{+1} Upper = 1.00E+10 Lower = 0
ke_{-1} Upper = 1.00E+10 Lower = 0

All rate constants which are allowed to vary (as determined in the previous option) are listed. Only the first rate constant in a given

linkage group is listed, since this is the only rate constant actually allowed to vary. Thus with linkage groups be careful that the limits chosen are good for all of the rate constants linked with the one shown. The upper default is assumed to be 1*10k10e while the lower default is assumed to be zero. It is generally a good idea to enter limits to a given varied parameter when some idea of the parameter value is known. This helps in eliminating the possibility of convergence at a local, rather than global, minimum. The lower limit cannot be lower than 0 (the default setting) since a negative rate constant has no physical meaning. Initial settings are 10k+1e0 for the upper limit and zero for the lower limit.

6. Error Parameters.

This menu option appears only for regression using STEPT, otherwise this input menu is bypassed and defaults used. DELMN refers to the minimum variation in the varied parameter Xeik which is required for convergence in that parameter to be achieved. This is still used in MARQ regression, but this parameter is changed with every iteration, and is based on the percentage change in Xeik with each iteration. DELTX refers to the initial change in the varied parameters Xeik and is only used in the STEPT regression. DELTX sets the value for the initial change in the varied rate constants. The default option for DELTX for the Marquardt algorithm is 10%. It is possible variation of this number may allow better fitting of the data.

7. Number of Data Sets to be fitted. (*)

The number of datasets to be fit is entered. Each data file is assumed to be one progress curve. In the present implementation up to 8 data files can be entered.

8. Entry of Data File Names Save sets, and output factors. (*)

The user is prompted for the name of the save set (file must end in *.SAV) and the corresponding data file name. It is **IMPORTANT** to note that the time scale, number of data points, output factors, and Y-scale is determined by the Save set and **NOT** by the Data set. This feature allows parts of a data file to be dropped, and also allows the output scale to differ from data file to data file.

As only the kinetic constants of the last save set is actually entered into FITSIM, if different initial rate constants are desired only the LAST save set needs to be changed. Changing of output factors or integration times requires changing these values in ALL of the save sets. The amount of reaction time for each progress curve is determined by the time listed in the save set and **NOT** the data file. If the data file contains data near where the reaction is completed the total simulated time should be shortened to the amount of time required for completion of the reaction. Failure to do so may result in over-weighting of the end point, resulting in an overall poorer fit of the data to the kinetic mechanism being tested. If the total simulation time exceeds the reaction time of the experimental data file, the weight of the time points exceeding the total reaction time is given a weight of zero. Thus simulation times past the total reaction

time given by the data is ignored.

It is generally a good idea to check the entering save sets by displaying the initial simulated and real data fits by using the user-supervised [9] regression option. The most common user error to date has been the improper saving of SAVE set parameters resulting in various FITSIM errors.

As of version 1.4 all internal FITSIM calculations are in double precision. This prevents, for the most part, changes in the final fits being brought about by changing the order of how the data files are read in. Very small differences in the final fits may still be detected if the order of how the data files are read in is changed. The most accurate final fit will be achieved if the user enters those data files with the best initial fit before those with the worst initial fit. Under nearly all circumstances, however, the order of data file entry is not significant.

After both the data file and corresponding save set have been entered the user is asked for which simulation output factor is to be used. The user should select the output number of the mechanism to which the data files should be compared. Depending upon the mechanism, upto eight different outputs can be selected for in the present implementation.

9. Default Options chosen.

Default is [N] for No. If [Y] is chosen the defaults for options 10-20 listed below are chosen. The defaults have been selected on the basis of the most COMMON regression applications, but by no means are appropriate for all types of progress curve data. NOTE: This option is NOT written to the AUTOSIM file.

10. Maximum # of Iterations. (*)

Default is [100]. Maximum number of calls to KINSIM to calculate a simulated fit. The number of calls per iteration is generally equal to the number of varied parameters..

11. Output Choices. (*)

[-1] - Error output only.
[_] - Initial parameter values and SUMSQ and final parameter values and SUMSQ.
[1] - Initial and final values plus intermediate values after each iteration are output, along with an Analysis of Variance.
[2] - All output associated with [1] plus output to the AUTOSIM.LOG file of the final values for the experimental, fitted, weights, and residuals of the final fits. Output of the lower triangle of the correlating matrix is also printed.
[3] - Prints all intermediate values and calculated derivative.
[11] - Same as [1] but with final plots to terminal.
[12] - Same as [2] but with final plots to terminal.

The default value is an output choice of [11]. The addition of 10 to any of the above choices allows output of the final fits to the video terminal. Therefore output choice [1] along with output of the final fits to the

terminal is coded for by entering the value [11]. This allows the user to conviently "eyeball" the fits. NO regression analysis should ever be trusted unless one either looks at the fits or the residuals. Ideally the residuals should be scattered around 0 if a good fit was achieved.

12. Marq Lambda to Initialize. (*)

The starting value of lambda, a regression parameter used for the Marquardt algorithm. Generally speaking, the better the initial fit, the lower the initial value. This chosen value for lambda does not affect the final answer, only how quickly the computer may converges. When lambda = 0 the Gauss-Newton method is selected while when lambda approaches infinity the steepest descent method is selected. The default value is [1.0]

13. Weighting Choices. (*)

[-1] -Special weighting. Two ranges in which the points are heavily weighted. This may be the best approach for curves which have two inflection points. If this option is selected, input is required for the two ranges and the amount of weight given to the points within this range.

[0] - Identical weighting. All points in the curve are given identical weight, the value of one. This option is best used on noisy data sets where the data is not a relatively smooth function.

[1] - Proportional weighting. All points are weighted to the inverse of their value. This particular weighting may be useful if a large range in the data occurs. This has the affect of weighting more heavily the data points which are smaller in absolute value.

[2] - Proportional squared weighting. All points are weighted to the inverse square of their value. This is often assumed to be approximately equal to the variance of the data value. This has the effect of weighting even more heavily the data points which are smaller in absolute value than option [1].

[3] - Weighting by Slope. This may be the preferred method of weighting for data files which appear to be relatively smooth functions and contain no spikes in the data. This option makes the weight approximately equal to the slope of the tangent.

The weights assigned will probably NOT be what the user might expect since weighting is carried out automatically by FITSIM to account for differences in the number of points and Y-ranges between data sets. This has the affect of weighting each curve identically, a necessary condition for proper regression. However the points within a given dataset can be weighted with respect to one another. The default is identical weighting [0].

14. Convergence Percentage. (*)

The largest percentage change in the varied parameters which allow convergence. If all parameters change by a percentage less than or equal to the value after a given iteration, convergence is achieved, and iteration stops. This value in some part determines the accuracy of the

final parameters. The default is [1%]. Values too small can cause unnecessary iteration without improvement in the final fit, while values too large can cause false convergence to be achieved.

15. Absolute Value for Background Noise. (*)

The largest absolute value difference between the baseline and the experimental data which is ignored by FITSIM. It is assumed that values lower than this number do not reflect any real difference between the baseline and the data, and thus no attempt to fit these points is made. This value depends both on the absolute scale of the curves being fit, and the "noise" within the data. This option allows for trivial differences at either the beginning or end of a time course to be ignored in the fitting process. The default is [0].

16. Percentage Change in the SUMSQ to Stop Iteration. (*)

When the percentage change in the sum of squares (SUMSQ) between iterations is smaller than this value, iteration stops and convergence is assumed. When convergence occurs in this manner convergence in reality may not have taken place, and the values found may thus not be the best fit values. The default is [1%]. Any larger values for this option are highly discouraged, and indeed for "best" fits this value may need to be smaller than the default.

17. Treatment of Negative values. (*)

Default value is [1]. This option allows modification of negative data values according to the schemes listed below.

[0] - No change, negative values are allowed and fitted.
[1] - Convert all negative vales to zero before fitting.
[2] - Convert all negative values to their absolute value. This is what the programs KINSIM and SFDISP do to displayed data files. Although screen fitting to the data may actually appear better, the best fits are not necessarily found to the actual data by this option. Therefore be careful when selecting this option and know approximately how many negative data points you have to begin with !

18. Calculation of Derivatives. (*)

This determines how the partial derivatives are calculated. A value of [1] will calculate the partial derivatives by a first-order approximation while a value of [2] will calculate partial derivatives by a second-order approximation. The default value for this option is [1].

19. Data Point at Time Zero. (*)

[0] - The time zero point in the simulations is assumed equal to zero, and is not compared to any equivalent point in the data. The

time zero value is NOT figured into the calculation of the SSQ. This is the best option if one is not sure which option should be chosen. This option should also be selected when data files contain a time zero point, but the user does want this value to included into SSQ calculations.

[1] - The first data point read by FITSIM is assumed equal to the data point at time zero minus the baseline value. The time zero point IS entered into the calculation of the SSQ. This option should be selected for files in which a time zero point is recorded.

[2] - The time zero data point is assumed equal to the value zero. This value IS included in the calculation of the SSQ. This may be the best option for files in which the baseline is recorded BEFORE the reaction is started.

[9] - This option should be selected for when files without a recorded time zero point are analyzed. This is usually indicated by a warning message printed to the LOG file. This option causes a shift of the time data by one delta time period and a time zero point added to the data. This time zero value is, however, subsequently ignored in SSQ calculations by the program.

This determines how the data point at time zero is either calculated or treated by the regression algorithm. This option is necessary due to the fact all datasets may not have a time zero point, a point which is calculated by the KINSIM subroutine. For instance at WUMS-BCCF datasets do not contain a time zero point in the passed array, although a baseline value is recorded. The default value for this option is [0].

20. Output Number Variation. (*)

Entering the value 1 (ON) allows variation of the entered output factors. No variation occurs in the output factor when the value [0] is chosen, which is the default value. If the output is allowed to vary, the user is asked to give the percentage of variation allowed in this parameter. Variation of the output factors is not recommended as progress curves are not, in reality, sensitive to specific values for this constant if other kinetic parameters are varied. This option should be avoided at this time. This option is included simply as a departure point for a future version of FITSIM which will allow such variation. Default value for this option is [0].

21. Masking Option. (*)

This option determines the handling of errors in which the error gradient is so shallow, determination of the directional change in a varied constant is difficult. When this condition occurs, the warning <Possible Dangerous Value of Coefficient XX> is printed. This occurs most often when both the k_+ and k_- constants for a given reaction are varied. When Masking is on (value = 1) the next iteration will mask the k_+ constant and only variation of k_- will occur. When masking is turned on, a message is printed in the log file stating this fact. This masking occurs for only one iteration, after which normal iteration of both constants will occur. This process will often allow a better determination of the k_+ and k_- constant values. When Masking is off (value = 0), iteration proceeds

normally. With most analysis, masking is never needed; thus whether this option is used or not will make no difference.

22. LOG file name. (*)

The name of the LOG file for this particular run. The default name is AUTOSIM.LOG .

23. LOG Comment. (*)

Default is None. A descriptive comment can be entered for this run which will be printed in the log file. Useful for comparing different runs.

6.3 FITDATA Error Messages.

Cannot OPEN file :
File cannot be opened. File either does not exist or is of the wrong format. For data file names the user is asked if he wishes to retain this name regardless. This option is included because it is sometimes easier to change the AUTOSIM file manually rather than rerun FITDATA from the beginning.

Improper Input, Reenter :
Input was not as expected, value needs to be re-entered correctly.

Too many variables, Reduce number :
In the present implementation of FITSIM a maximum of 20 variables is allowed. If attempts to vary more than this number occur, this message is printed and the user is forced to vary fewer rate constants before any further input is allowed.

High number of variables :
In the present implementation attempts to vary more than 10 rate constants at a time is probbaly doomed to failure. Attempts to vary more than 9 rate constants and less than 20 trigger this warning message. The user can continue if he wishes, however.

6.4 Batch Processing Program.

This program submits FITSIM as a batch process when analysis requires long times. This command process is undoubtedly specific for the VAX/VMS system at WU, but can be modified easily for use on other VAX/VMS systems. The user is cautioned, however, that unless correct regression parameters and kinetic constants are selected, computer time can be wasted unnecessarily by this method. Execution of this command process brings the following screen :

```
------------------------------------------------------------
- FITSIM BATCH Submission -                          Ver 1.0
------------------------------------------------------------


**************************************************
*  AUTOSIM File MUST be in the present directory *
*  WARNING : Output # Cannot be greater than 9   *
**************************************************


Name of FITSIM Read file (*.AUT) [ AUTOSIM.AUT] :


Number of FITSIM job submitted [1] :


Present Time :
      29-Feb-1988   15:49:31


Sumit When ? (NOW=0:0, Overnight=<CR>, SPAWN)

Name for Batch Log file in TMP [FITSIM] :
```

The inputs required by the user are straight forward and consist of the *.AUT file to be read, the number of the job submitted (important for submitting multiple jobs at night), the submission time, and the name for the log file created in conjunction with the batch process.

6.5 Description of the AUTOSIM file.

Listed is a detailed description of the file format made using the program FITDATA and choosing the default regression options. The sample listed represents the minimum size file required for proper execution of FITSIM. The example below can use the bibiorder sample mechanism described in the next chapter.

Line #	Data in File:				Description:
{01)	1				**Regression Routine used**
{02}	bibiorder.SIM				**Mechanism name**
{03}	4				**# of Fixed parameters**
(4)	1	3	6	8	**Fixed kinetic constants**
{05}	4				**# of varied paramters**
{06}	2	4	5	7	**varied kinetic constants**
{07}	0				**# of linked grps present**
{08)	1000.0	10000	1.0000	100000	**Upper Limits of Parameters**
{09}	0.1000	l.0000	0.0100	10.000	**Lower Limits of Parameters**
{10)	0.1000	0.1000	0.1000	0.1000	**Minimum Change**
{11}	1.00E-06	1.00E-06	1.00E-06	1.00E-06	**Initial Change**
{12}	4				**# of datasets to fit**
{13}	bibiord1.sav				**Save set for datafile #1**
	bibiord1.dat				**Datafile name #1**
	1				Output # used for datafile #1
	bibiord2.sav				
	bibiord2.dat				
	1				
	bibiord3.sav				
	bibiord3.dat				
	1				
	bibiord4.sav				**Save set which loads constants**
	bibiord4.dat				
	1				
{14}	100				**Max. # of Iterations**
{15}	11				**Output Choices**

Line #	Data in File:	Description:
{16}	1.000000	**Marquardt Lambda to Initialize**
{17}	0	**Weighting Choice**
{18}	1.000000	**Convergence Percentage**
{19}	0.000000	**Value assumed to be BCKGRND noise**
{20}	1.0000000E-04	**SUMSQ change to assume convergence**
{21}	1	**Convert Negative to Positive**
{22}	0	**Calculation of Derivatives**
{23}	1	**Data point at Time Zero**
{24}	0	**Output Number to Fit**
{25}	0	**Masking**
{26}	BIBIORD.LOG	**Name of the LOG file output**
{27}	FITSIM Test of BIBIOrd Mec	**Comment printed to LOG file**

Example with the inclusion of Linked Rate Constants:

{7a}	1		**# of linked grps present**
{7b}	2		**# of rate constants in linked grp**
{7c}	4	1.00	**Rate constant #, linkage value**
{7d}	6	1.00	**Rate constant #, linkage value**

@Kinetic constants are numbered internally in FITSIM simply in the order they are read in. Forward rate constants (k_+) are numbered before back rate constants (k_-).

Chapter 7. Implementation of FITSIM

7.1 Implementation of FITSIM.

FITSIM is NOT a module which is added to KINSIM and then linked. Rather FITSIM, at least in its present implementation (Version 1.63), is a separately compiled program from KINSIM. Several program modules make up the FITSIM program. Compiling and linking these programs is simplified in VAX/VMS by invoking a command procedure named MAKFITSIM. This procedure will prompt the user for any necessary information and then link and/or compile FITSIM.

There are several KINSIM program modules necessary for compilation of FITSIM. They are the modules UTIL, FILEIO, GETVALS, QIOAST, QATRAN, QASCRN, TTOUTBUF, TRMSPC125, TRMSPC240, and TRMSPC640. These modules are explained in detail in section E1. Additional (new) modules necessary to compile FITSIM are listed below.

Substituting the QATRAN module for the QASCRN module is needed to compile a "transportable" version which is a much less system-dependent form of on-line editing of the data. While with KINSIM this changed several aspects of input/output, with FITSIM only the printing of the kinetic constants using the OBSERVE qualifier or regression option [9] are affected by this change.

********** FORTRAN modules **********

FITSIM : Collection of routines which reads in regression control information, as well as the mechanism, experimental data, and saved parameter files. Weighting of points and weight corrections for each experimental time course is calculated here. Error checking of the data is also performed. On most systems little modification of this module is expected except in the case of opening the LOG file and using two VAX RTL procedures.

These RTL procedures in the FITSIM module can be eliminated if necessary as they are not crucial for program execution. These RTL procedures find the system time and allow the passing of the AUTOSIM format file name. If the later procedure is eliminated only a file name of AUTOSIM.dat will be accepted. These RTL procedures are pointed out in the FORTRAN listing.

FUNC : Collection of routines which packs the varied kinetic parameters into an array as well as performs the calls to KINSIM. Probably quite system-independent.

FUNCTION : Various function routines needed by both KINSIM and FITSIM. String handling and kinetic constant labeling routines. Also contains the data block for the operation of KINSIM.

MARQ1 : Main module holding the Marquardt regression routine. Math operations should be quite system-independent. It is possible output statements may have to be modified.

MARQERR : Routines for initializing and computing the final fits and errors for the marq regression routines. This should be system independent.

FUNCTION : Contains various function routines used in the FITSIM and FITDATA programs.

SIMAUTO : Modified version of SIMUL module found in KINSIM.

SOLVEAUTO : Modified version of the SOLVE module of KINSIM. Modifications deal with allowing simulation calculations to continue under certain cases even if integral tolerance requirements are not met.

STEPIT : Regression routine of Chandler (1965) modified for easier use. This can simply be a stub if the user does not wish to incorporate this routine into FITSIM.

WRITEMES : Collection of routines which outputs various messages to the user.

GRAPHICAUTO: Modified version of GRAPHIC module found in KINSIM. Modification deals with the displaying of data files.

LODMECAUTO : Modified version of LODMEC module in KINSIM.

PROMPTAUTO : Very slightly modified version of PROMPT module in KINSIM. At this time only the routines for the entry of equilibrium and rate constant entry are used. Perhaps at a future date these routines could be used in an enhanced data entry version of the program FITDATA. For now, most of the contained subroutines serve only as lengthly stubs.

STRINGAUTO : Modified version of STRINGS module in KINSIM.

********** PARAMETER INCLUDE FILES **********

ALLRUN.CMN - Storage constants & matrices
FNAME.CMN - Contains the file name to be opened in FITSIM.
KGROUP.CMN - Information about the interdependence between rate constants and holds the information of which rate constants are fixed and which are varied.
MCNSRV.CMN - Information about mass conservation. Change in KINSIM version 3.4
MREGRESS.CMN - Information important for regression by the Marquardt algorithm used in FITSIM.
REGRESS.CMN - Information for the control of the regression routines.
RESIDU.CMN - Sum of square information and parameters.

The module construction of FITSIM allows relatively easy linking with other nonlinear regression routines. Indeed it is expected that future versions of FITSIM may support more powerful regression routines as they become available. These additional regression routines could simply be incorporated as CALL statements in the main FITSIM module. Two common blocks can be included in these routines and simple relationships between the variables in FITSIM and the regression routine in question constructed. The Common block RESIDU contains the arrays for the experimental points, simulated points, and point weights. It also includes the total number of points and the number of data points in each data set. The Common block REGRESS contains the variables holding information for the control of regression routines. Included is the number of varied parameters and their values. Also included is the constraints of the varied parameters and the values below which convergence is allowed. These common blocks also contain comments which give the definition of all variables pertinent to the linkage of regression routines with KINSIM.

Included in this documentation are listings obtained by running the files for the BI-BI ordered simulations included with FITSIM. These should be ran as a check for proper operation of the program.

Other notes of Interest:

The Marquardt nonlinear regression subroutine included here has been the best nonlinear regression routine used to date. However, without extensive modification FITSIM has been designed to allow the inclusion of other non-linear regression routines. If another non-linear regression routine is used it must have the following characteristics :

1) Derivatives must be calculated internally by the program and not supplied by the user. Although this in a strict sense is not true since a user-entered derivative can be accessed (see FORTRAN listing for the module MARQ1), each change in the kinetic mechanism would require a change in the kinetic mechanism.

2) All experimental and fitted data points must be entered into the subroutine at the same time. The regression routine must NOT make calls to the external function (KINSIM) for each data point. It should be noted the RNLIN/DRNLIN regression routines of Version 10 of the IMSL statistics library cannot be incorporated into FITSIM due to this. Extensive rewriting of KINSIM and FITSIM would be required for theses type of regression routines to be incorporated.

3) The regression routine must allow weights and limits to be entered. Weights are critically important for the proper execution of FITSIM and MUST be used. Limits are needed to prevent the occurrence of negative rate constants and other impossible values. Negative rate constants, of course, have no physical meaning.

4) Listed below are some other files of interest when compiling and testing FITSIM.

********** ANCILLARY FILES **********

SIMUL.COM - VAX/VMS command file which links KINSIM, FITSIM, SFDISP, KINCOMP, and FITDATA together. To run this program under VAX/VMS requires the user to define three logical names for directories where the Executable versions of KINSIM, FITSIM, KINCOMP, and FITDATA are contained. See the program listing for details.

SIMUL.HLB - VAX/VMS help file for the SIMUL command procedure. This is a library file which uses the powerful HELP feature under VAX/VMS. The file VAX_HELP.txt is a text version of this file which could be modified by the user.

********** FILES FOR TESTING FITSIM **********

CONFOR* - Files for running the Conformational change mechanism for testing the operation of FITSIM. Included are *.SAV sets describing the Time frame, kinetic constants, and initial concentrations; *.DAT sets which are X vs. Y full time courses; AUTOSIM.DAT file which is read by the FITSIM program

BIBIORD* - Files for running the Bi-Bi ordered test mechanism as described in the documentation. Included files are the same as above.

********** VAX/VMS EXECUTABLE FILES **********

FITSIM640.EXE - The executable version of FITSIM for the VT640.
FITSIM240.EXE - The executable version of FITSIM for the VT240 (terminal must be placed in the VT100 mode for proper operation) and the VT125.
FITDATA125.EXE - The executable version of FITDATA for all DEC terminals (This routine has no graphics output; When running on a VT240 terminal should be placed in the VT100 mode for proper operation.)

7.2 Testing of FITSIM Implementation.

To test the implementation of FITSIM we suggest that simulations generated by KINSIM be used to test it. There are two test files which are included with FITSIM. One set of files is called CONFOR and represents a mechanism as described by Zimmerle, Patane, and Frieden (1987). Another set of files is called BIBIORD and represents kinetic runs of a

hypothetical enzyme following a BI-BI ordered mechanism. The BIBI files are described in more detail below :

MECHANISM Descriptor For BIBIORDER:

```
            K+1             K+2           K+3             K+4
      E + A == EA ;  EA + B == EAB ;  EAB == EQ + P ;  EQ == E + Q
            K-1             K-2           K-3             K-4
```

OUTPUT 1 = F1*Q/POT

RATE CONSTANTS :

K+ 1 =	1.000D+01	K- 1 =	1.000D+01
K+ 2 =	1.000D+01	K- 2 =	5.000D+02
K+ 3 =	2.500D-01	K- 3 =	1.000D-02
K+ 4 =	5.000D+03	K- 4 =	1.000D+01

OUTPUT CONSTANTS :

F1	=	1.000D+00
POT	=	1.000D+02

Initial Guesses for FITSIM for the Datafiles BIBORD1 - BIBIORD2 :

RATE CONSTANTS

K+ 1 =	1.000D+01	K- 1 =	1.000D+02
K+ 2 =	1.000D+01	K- 2 =	5.000D+01
K+ 3 =	3.500D-01	K- 3 =	1.000D-02
K+ 4 =	1.000D+03	K- 4 =	1.000D+01

Experimental Data file BIBIORD1 :

SIMULATION PARAMETERS

DELTA TIME	=	1.000D+00	ITERATIONS/POINT	=	2
RUN TIME	=	1.200D+03	YMAX	=	1.000D+00
FLUX TOLERANCE	=	1.000D-01	INTEGRAL TOLERANCE	=	1.000D-03

CONCENTRATIONS

E	=	1.000D+00
A	=	1.000D+02
EA	=	0.000D+00
B	=	1.000D+02
EAB	=	0.000D+00
EQ	=	0.000D+00
P	=	0.000D+00
Q	=	0.000D+00

Experimental Data File for BIBIORD2 :

SIMULATION PARAMETERS

```
DELTA TIME     =  1.000D+00      ITERATIONS/POINT   =          1
RUN TIME       =  5.000D+01      YMAX               =  1.000D+00
FLUX TOLERANCE =  1.000D-01      INTEGRAL TOLERANCE =  1.000D-03
```

CONCENTRATIONS

```
E          =  1.000D+00
A          =  1.000D+01
EA         =  0.000D+00
B          =  1.000D+03
EAB        =  0.000D+00
EQ         =  0.000D+00
P          =  0.000D+00
Q          =  0.000D+00
```

Experimental Data File BIBIORD3 :

SIMULATION PARAMETERS

```
DELTA TIME     =  5.000D+00      ITERATIONS/POINT   =          1
RUN TIME       =  1.000D+03      YMAX               =  1.000D+00
FLUX TOLERANCE =  1.000D-01      INTEGRAL TOLERANCE =  1.000D-03
```

CONCENTRATIONS

```
E          =  1.000D+00
A          =  1.000D+03
EA         =  0.000D+00
B          =  1.000D+01
EAB        =  0.000D+00
EQ         =  0.000D+00
P          =  0.000D+00
Q          =  0.000D+00
```

Experimental Data File BIBIORD4 :

SIMULATION PARAMETERS

```
DELTA TIME     =  5.000D+00      ITERATIONS/POINT   =          1
RUN TIME       =  1.000D+03      YMAX               =  1.000D+00
FLUX TOLERANCE =  1.000D-01      INTEGRAL TOLERANCE =  1.000D-03
```

CONCENTRATIONS

```
E          =  1.000D+00
A          =  1.000D+03
EA         =  0.000D+00
B          =  1.000D+01
EAB        =  0.000D+00
EQ         =  0.000D+00
P          =  5.000D+04
Q          =  0.000D+00
```

Using the OBSERVE qualifier of FITSIM the user can examine whether the program is calculating initial values correctly. Listed on the next few pages are the values printed when the BIBI test files are loaded into FITSIM. These values should be checked after first compiling the program to check for proper execution.

Dataset	Dataset Y	Simulated Y	Weight	Time value	Point #
1	0.000E+00	0.000E+00	0.000E+00	0.000E+00	1
1	1.981E-02	3.968E-02	1.12	12.0	2
1	3.948E-02	7.926E-02	1.12	24.0	3
1	5.901E-02	0.119	1.12	36.0	4
1	7.840E-02	0.158	1.12	48.0	5
1	9.764E-02	0.197	1.12	60.0	6
1	0.117	0.236	1.12	72.0	7
1	0.136	0.275	1.12	84.0	8
1	0.154	0.313	1.12	96.0	9
1	0.173	0.352	1.12	108.	10
1	0.192	0.390	1.12	120.	11
1	0.210	0.427	1.12	132.	12
1	0.228	0.464	1.12	144.	13
1	0.246	0.501	1.12	156.	14
1	0.264	0.537	1.12	168.	15
1	0.281	0.572	1.12	180.	16
1	0.299	0.607	1.12	192.	17
1	0.316	0.640	1.12	204.	18
1	0.333	0.672	1.12	216.	19
1	0.350	0.703	1.12	228.	20
1	0.366	0.732	1.12	240.	21
1	0.383	0.758	1.12	252.	22
1	0.399	0.783	1.12	264.	23
1	0.415	0.804	1.12	276.	24
1	0.431	0.822	1.12	288.	25
1	0.447	0.838	1.12	300.	26
1	0.462	0.850	1.12	312.	27
1	0.477	0.860	1.12	324.	28
1	0.492	0.868	1.12	336.	29
1	0.507	0.873	1.12	348.	30
1	0.521	0.877	1.12	360.	31
1	0.535	0.880	1.12	372.	32
1	0.549	0.882	1.12	384.	33
1	0.563	0.883	1.12	396.	34
1	0.577	0.884	1.12	408.	35
1	0.590	0.885	1.12	420.	36
1	0.603	0.885	1.12	432.	37
1	0.616	0.886	1.12	444.	38
1	0.628	0.886	1.12	456.	39
1	0.640	0.886	1.12	468.	40
1	0.652	0.886	1.12	480.	41
1	0.664	0.886	1.12	492.	42
1	0.675	0.887	1.12	504.	43
1	0.686	0.887	1.12	516.	44
1	0.697	0.887	1.12	528.	45
1	0.708	0.887	1.12	540.	46
1	0.718	0.887	1.12	552.	47
1	0.728	0.887	1.12	564.	48
1	0.738	0.887	1.12	576.	49
1	0.747	0.887	1.12	588.	50
1	0.757	0.887	1.12	600.	51
1	0.766	0.887	1.12	612.	52
1	0.774	0.887	1.12	624.	53
1	0.783	0.887	1.12	636.	54
1	0.791	0.887	1.12	648.	55
1	0.799	0.887	1.12	660.	56

Dataset	Dataset Y	Simulated Y	Weight	Time value	Point #
1	0.806	0.887	1.12	672.	57
1	0.813	0.887	1.12	684.	58
1	0.820	0.887	1.12	696.	59
1	0.827	0.887	1.12	708.	60
1	0.833	0.887	1.12	720.	61
1	0.840	0.887	1.12	732.	62
1	0.846	0.887	1.12	744.	63
1	0.851	0.887	1.12	756.	64
1	0.857	0.887	1.12	768.	65
1	0.862	0.887	1.12	780.	66
1	0.867	0.887	1.12	792.	67
1	0.871	0.887	1.12	804.	68
1	0.876	0.887	1.12	816.	69
1	0.880	0.887	1.12	828.	70
1	0.884	0.887	1.12	840.	71
1	0.888	0.887	1.12	852.	72
1	0.891	0.887	1.12	864.	73
1	0.895	0.887	1.12	876.	74
1	0.898	0.887	1.12	888.	75
1	0.901	0.887	1.12	900.	76
1	0.903	0.887	1.12	912.	77
1	0.906	0.887	1.12	924.	78
1	0.908	0.887	1.12	936.	79
1	0.910	0.887	1.12	948.	80
1	0.912	0.887	1.12	960.	81
1	0.914	0.887	1.12	972.	82
1	0.916	0.887	1.12	984.	83
1	0.918	0.887	1.12	996.	84
1	0.919	0.887	1.12	1.008E+03	85
1	0.921	0.887	1.12	1.020E+03	86
1	0.922	0.887	1.12	1.032E+03	87
1	0.923	0.887	1.12	1.044E+03	88
1	0.924	0.887	1.12	1.056E+03	89
1	0.925	0.887	1.12	1.068E+03	90
1	0.926	0.887	1.12	1.080E+03	91
1	0.927	0.887	1.12	1.092E+03	92
1	0.928	0.887	1.12	1.104E+03	93
1	0.929	0.887	1.12	1.116E+03	94
1	0.929	0.887	1.12	1.128E+03	95
1	0.930	0.887	1.12	1.140E+03	96
1	0.930	0.887	1.12	1.152E+03	97
1	0.931	0.887	1.12	1.164E+03	98
1	0.931	0.887	1.12	1.176E+03	99
1	0.932	0.887	1.12	1.188E+03	100
1	0.932	0.887	1.12	1.200E+03	101
2	0.000E+00	0.000E+00	0.000E+00	0.000E+00	102
2	2.336E-02	3.407E-02	2.25	1.00	103
2	4.697E-02	6.853E-02	2.25	2.00	104
2	7.057E-02	0.103	2.25	3.00	105
2	9.417E-02	0.137	2.25	4.00	106
2	0.118	0.172	2.25	5.00	107
2	0.141	0.206	2.25	6.00	108
2	0.165	0.241	2.25	7.00	109
2	0.188	0.275	2.25	8.00	110
2	0.212	0.309	2.25	9.00	111

Dataset	Dataset Y	Simulated Y	Weight	Time value	Point #
2	0.236	0.343	2.25	10.0	112
2	0.259	0.378	2.25	11.0	113
2	0.283	0.412	2.25	12.0	114
2	0.306	0.446	2.25	13.0	115
2	0.330	0.480	2.25	14.0	116
2	0.353	0.514	2.25	15.0	117
2	0.377	0.548	2.25	16.0	118
2	0.400	0.582	2.25	17.0	119
2	0.424	0.616	2.25	18.0	120
2	0.447	0.649	2.25	19.0	121
2	0.470	0.682	2.25	20.0	122
2	0.494	0.715	2.25	21.0	123
2	0.517	0.748	2.25	22.0	124
2	0.540	0.781	2.25	23.0	125
2	0.564	0.812	2.25	24.0	126
2	0.587	0.843	2.25	25.0	127
2	0.610	0.872	2.25	26.0	128
2	0.633	0.899	2.25	27.0	129
2	0.656	0.922	2.25	28.0	130
2	0.679	0.940	2.25	29.0	131
2	0.702	0.953	2.25	30.0	132
2	0.725	0.963	2.25	31.0	133
2	0.748	0.970	2.25	32.0	134
2	0.770	0.976	2.25	33.0	135
2	0.793	0.979	2.25	34.0	136
2	0.815	0.982	2.25	35.0	137
2	0.836	0.984	2.25	36.0	138
2	0.857	0.985	2.25	37.0	139
2	0.878	0.986	2.25	38.0	140
2	0.897	0.987	2.25	39.0	141
2	0.915	0.987	2.25	40.0	142
2	0.930	0.988	2.25	41.0	143
2	0.943	0.988	2.25	42.0	144
2	0.954	0.988	2.25	43.0	145
2	0.962	0.988	2.25	44.0	146
2	0.969	0.988	2.25	45.0	147
2	0.975	0.988	2.25	46.0	148
2	0.980	0.988	2.25	47.0	149
2	0.983	0.989	2.25	48.0	150
2	0.986	0.989	2.25	49.0	151
2	0.988	0.989	2.25	50.0	152
3	0.000E+00	0.000E+00	0.000E+00	0.000E+00	153
3	2.034E-02	0.110	0.562	5.00	154
3	4.034E-02	0.216	0.562	10.0	155
3	6.000E-02	0.316	0.562	15.0	156
3	7.931E-02	0.409	0.562	20.0	157
3	9.829E-02	0.495	0.562	25.0	158
3	0.117	0.574	0.562	30.0	159
3	0.135	0.646	0.562	35.0	160
3	0.153	0.709	0.562	40.0	161
3	0.171	0.764	0.562	45.0	162
3	0.188	0.811	0.562	50.0	163
3	0.205	0.850	0.562	55.0	164
3	0.222	0.883	0.562	60.0	165
3	0.238	0.909	0.562	65.0	166

Dataset	Dataset Y	Simulated Y	Weight	Time value	Point #
3	0.255	0.930	0.562	70.0	167
3	0.270	0.946	0.562	75.0	168
3	0.286	0.959	0.562	80.0	169
3	0.301	0.969	0.562	85.0	170
3	0.316	0.976	0.562	90.0	171
3	0.331	0.982	0.562	95.0	172
3	0.345	0.986	0.562	100.	173
3	0.359	0.989	0.562	105.	174
3	0.373	0.992	0.562	110.	175
3	0.387	0.993	0.562	115.	176
3	0.400	0.995	0.562	120.	177
3	0.413	0.996	0.562	125.	178
3	0.426	0.997	0.562	130.	179
3	0.438	0.997	0.562	135.	180
3	0.450	0.998	0.562	140.	181
3	0.462	0.998	0.562	145.	182
3	0.474	0.999	0.562	150.	183
3	0.486	0.999	0.562	155.	184
3	0.497	0.999	0.562	160.	185
3	0.508	0.999	0.562	165.	186
3	0.519	0.999	0.562	170.	187
3	0.530	1.00	0.562	175.	188
3	0.540	1.00	0.562	180.	189
3	0.550	1.00	0.562	185.	190
3	0.560	1.00	0.562	190.	191
3	0.570	1.00	0.562	195.	192
3	0.580	1.00	0.562	200.	193
3	0.589	1.00	0.562	205.	194
3	0.599	1.00	0.562	210.	195
3	0.608	1.00	0.562	215.	196
3	0.616	1.00	0.562	220.	197
3	0.625	1.00	0.562	225.	198
3	0.633	1.00	0.562	230.	199
3	0.642	1.00	0.562	235.	200
3	0.650	1.00	0.562	240.	201
3	0.658	1.00	0.562	245.	202
3	0.666	1.00	0.562	250.	203
3	0.673	1.00	0.562	255.	204
3	0.681	1.00	0.562	260.	205
3	0.688	1.00	0.562	265.	206
3	0.695	1.00	0.562	270.	207
3	0.702	1.00	0.562	275.	208
3	0.709	1.00	0.562	280.	209
3	0.715	1.00	0.562	285.	210
3	0.722	1.00	0.562	290.	211
3	0.728	1.00	0.562	295.	212
3	0.735	1.00	0.562	300.	213
3	0.741	1.00	0.562	305.	214
3	0.747	1.00	0.562	310.	215
3	0.753	1.00	0.562	315.	216
3	0.758	1.00	0.562	320.	217
3	0.764	1.00	0.562	325.	218
3	0.769	1.00	0.562	330.	219
3	0.775	1.00	0.562	335.	220
3	0.780	1.00	0.562	340.	221

3	0.785	1.00	0.562	345.	222
3	0.790	1.00	0.562	350.	223
3	0.795	1.00	0.562	355.	224
3	0.800	1.00	0.562	360.	225
3	0.804	1.00	0.562	365.	226
3	0.809	1.00	0.562	370.	227
3	0.813	1.00	0.562	375.	228
3	0.818	1.00	0.562	380.	229
3	0.822	1.00	0.562	385.	230
3	0.826	1.00	0.562	390.	231
3	0.830	1.00	0.562	395.	232
3	0.834	1.00	0.562	400.	233
3	0.838	1.00	0.562	405.	234
3	0.842	1.00	0.562	410.	235
3	0.846	1.00	0.562	415.	236
3	0.849	1.00	0.562	420.	237
3	0.853	1.00	0.562	425.	238
3	0.856	1.00	0.562	430.	239
3	0.860	1.00	0.562	435.	240
3	0.863	1.00	0.562	440.	241
3	0.866	1.00	0.562	445.	242
3	0.869	1.00	0.562	450.	243
3	0.872	1.00	0.562	455.	244
3	0.875	1.00	0.562	460.	245
3	0.878	1.00	0.562	465.	246
3	0.881	1.00	0.562	470.	247
3	0.884	1.00	0.562	475.	248
3	0.887	1.00	0.562	480.	249
3	0.889	1.00	0.562	485.	250
3	0.892	1.00	0.562	490.	251
3	0.895	1.00	0.562	495.	252
3	0.897	1.00	0.562	500.	253
3	0.900	1.00	0.562	505.	254
3	0.902	1.00	0.562	510.	255
3	0.904	1.00	0.562	515.	256
3	0.907	1.00	0.562	520.	257
3	0.909	1.00	0.562	525.	258
3	0.911	1.00	0.562	530.	259
3	0.913	1.00	0.562	535.	260
3	0.915	1.00	0.562	540.	261
3	0.917	1.00	0.562	545.	262
3	0.919	1.00	0.562	550.	263
3	0.921	1.00	0.562	555.	264
3	0.923	1.00	0.562	560.	265
3	0.925	1.00	0.562	565.	266
3	0.927	1.00	0.562	570.	267
3	0.928	1.00	0.562	575.	268
3	0.930	1.00	0.562	580.	269
3	0.932	1.00	0.562	585.	270
3	0.933	1.00	0.562	590.	271
3	0.935	1.00	0.562	595.	272
3	0.936	1.00	0.562	600.	273
3	0.938	1.00	0.562	605.	274
3	0.939	1.00	0.562	610.	275
3	0.941	1.00	0.562	615.	276
3	0.942	1.00	0.562	620.	277
3	0.944	1.00	0.562	625.	278

Dataset	Dataset Y	Simulated Y	Weight	Time value	Point #
3	0.945	1.00	0.562	630.	279
3	0.946	1.00	0.562	635.	280
3	0.948	1.00	0.562	640.	281
3	0.949	1.00	0.562	645.	282
3	0.950	1.00	0.562	650.	283
3	0.951	1.00	0.562	655.	284
3	0.952	1.00	0.562	660.	285
3	0.954	1.00	0.562	665.	286
3	0.955	1.00	0.562	670.	287
3	0.956	1.00	0.562	675.	288
3	0.957	1.00	0.562	680.	289
3	0.958	1.00	0.562	685.	290
3	0.959	1.00	0.562	690.	291
3	0.960	1.00	0.562	695.	292
3	0.961	1.00	0.562	700.	293
3	0.962	1.00	0.562	705.	294
3	0.963	1.00	0.562	710.	295
3	0.964	1.00	0.562	715.	296
3	0.965	1.00	0.562	720.	297
3	0.965	1.00	0.562	725.	298
3	0.966	1.00	0.562	730.	299
3	0.967	1.00	0.562	735.	300
3	0.968	1.00	0.562	740.	301
3	0.969	1.00	0.562	745.	302
3	0.969	1.00	0.562	750.	303
3	0.970	1.00	0.562	755.	304
3	0.971	1.00	0.562	760.	305
3	0.971	1.00	0.562	765.	306
3	0.972	1.00	0.562	770.	307
3	0.973	1.00	0.562	775.	308
3	0.973	1.00	0.562	780.	309
3	0.974	1.00	0.562	785.	310
3	0.975	1.00	0.562	790.	311
3	0.975	1.00	0.562	795.	312
3	0.976	1.00	0.562	800.	313
3	0.977	1.00	0.562	805.	314
3	0.977	1.00	0.562	810.	315
3	0.978	1.00	0.562	815.	316
3	0.978	1.00	0.562	820.	317
3	0.979	1.00	0.562	825.	318
3	0.979	1.00	0.562	830.	319
3	0.980	1.00	0.562	835.	320
3	0.980	1.00	0.562	840.	321
3	0.981	1.00	0.562	845.	322
3	0.981	1.00	0.562	850.	323
3	0.982	1.00	0.562	855.	324
3	0.982	1.00	0.562	860.	325
3	0.982	1.00	0.562	865.	326
3	0.983	1.00	0.562	870.	327
3	0.983	1.00	0.562	875.	328
3	0.984	1.00	0.562	880.	329
3	0.984	1.00	0.562	885.	330
3	0.984	1.00	0.562	890.	331
3	0.985	1.00	0.562	895.	332
3	0.985	1.00	0.562	900.	333

Dataset	Dataset Y	Simulated Y	Weight	Time value	Point #
3	0.986	1.00	0.562	905.	334
3	0.986	1.00	0.562	910.	335
3	0.986	1.00	0.562	915.	336
3	0.987	1.00	0.562	920.	337
3	0.987	1.00	0.562	925.	338
3	0.987	1.00	0.562	930.	339
3	0.988	1.00	0.562	935.	340
3	0.988	1.00	0.562	940.	341
3	0.988	1.00	0.562	945.	342
3	0.988	1.00	0.562	950.	343
3	0.989	1.00	0.562	955.	344
3	0.989	1.00	0.562	960.	345
3	0.989	1.00	0.562	965.	346
3	0.989	1.00	0.562	970.	347
3	0.990	1.00	0.562	975.	348
3	0.990	1.00	0.562	980.	349
3	0.990	1.00	0.562	985.	350
3	0.990	1.00	0.562	990.	351
3	0.991	1.00	0.562	995.	352
3	0.991	1.00	0.562	1.000E+03	353
4	0.000E+00	0.000E+00	0.000E+00	0.000E+00	354
4	3.659E-02	0.138	1.12	10.0	355
4	7.177E-02	0.250	1.12	20.0	356
4	0.106	0.338	1.12	30.0	357
4	0.138	0.405	1.12	40.0	358
4	0.169	0.453	1.12	50.0	359
4	0.199	0.488	1.12	60.0	360
4	0.228	0.511	1.12	70.0	361
4	0.255	0.527	1.12	80.0	362
4	0.282	0.538	1.12	90.0	363
4	0.307	0.545	1.12	100.	364
4	0.331	0.549	1.12	110.	365
4	0.354	0.552	1.12	120.	366
4	0.377	0.553	1.12	130.	367
4	0.398	0.555	1.12	140.	368
4	0.418	0.555	1.12	150.	369
4	0.438	0.556	1.12	160.	370
4	0.456	0.556	1.12	170.	371
4	0.474	0.556	1.12	180.	372
4	0.491	0.556	1.12	190.	373
4	0.507	0.556	1.12	200.	374
4	0.522	0.556	1.12	210.	375
4	0.537	0.556	1.12	220.	376
4	0.551	0.556	1.12	230.	377
4	0.564	0.556	1.12	240.	378
4	0.577	0.556	1.12	250.	379
4	0.589	0.556	1.12	260.	380
4	0.601	0.556	1.12	270.	381
4	0.612	0.556	1.12	280.	382
4	0.623	0.556	1.12	290.	383
4	0.633	0.556	1.12	300.	384
4	0.642	0.556	1.12	310.	385
4	0.652	0.556	1.12	320.	386
4	0.660	0.556	1.12	330.	387
4	0.669	0.556	1.12	340.	388
4	0.677	0.556	1.12	350.	389

Dataset	Dataset Y	Simulated Y	Weight	Time value	Point #
4	0.684	0.556	1.12	360.	390
4	0.691	0.556	1.12	370.	391
4	0.698	0.556	1.12	380.	392
4	0.705	0.556	1.12	390.	393
4	0.711	0.556	1.12	400.	394
4	0.717	0.556	1.12	410.	395
4	0.722	0.556	1.12	420.	396
4	0.728	0.556	1.12	430.	397
4	0.733	0.556	1.12	440.	398
4	0.738	0.556	1.12	450.	399
4	0.742	0.556	1.12	460.	400
4	0.746	0.556	1.12	470.	401
4	0.751	0.556	1.12	480.	402
4	0.755	0.556	1.12	490.	403
4	0.758	0.556	1.12	500.	404
4	0.762	0.556	1.12	510.	405
4	0.765	0.556	1.12	520.	406
4	0.769	0.556	1.12	530.	407
4	0.772	0.556	1.12	540.	408
4	0.775	0.556	1.12	550.	409
4	0.777	0.556	1.12	560.	410
4	0.780	0.556	1.12	570.	411
4	0.782	0.556	1.12	580.	412
4	0.785	0.556	1.12	590.	413
4	0.787	0.556	1.12	600.	414
4	0.789	0.556	1.12	610.	415
4	0.791	0.556	1.12	620.	416
4	0.793	0.556	1.12	630.	417
4	0.795	0.556	1.12	640.	418
4	0.797	0.556	1.12	650.	419
4	0.798	0.556	1.12	660.	420
4	0.800	0.556	1.12	670.	421
4	0.802	0.556	1.12	680.	422
4	0.803	0.556	1.12	690.	423
4	0.804	0.556	1.12	700.	424
4	0.806	0.556	1.12	710.	425
4	0.807	0.556	1.12	720.	426
4	0.808	0.556	1.12	730.	427
4	0.809	0.556	1.12	740.	428
4	0.810	0.556	1.12	750.	429
4	0.811	0.556	1.12	760.	430
4	0.812	0.556	1.12	770.	431
4	0.813	0.556	1.12	780.	432
4	0.814	0.556	1.12	790.	433
4	0.814	0.556	1.12	800.	434
4	0.815	0.556	1.12	810.	435
4	0.816	0.556	1.12	820.	436
4	0.817	0.556	1.12	830.	437
4	0.817	0.556	1.12	840.	438
4	0.818	0.556	1.12	850.	439
4	0.818	0.556	1.12	860.	440
4	0.819	0.556	1.12	870.	441
4	0.819	0.556	1.12	880.	442
4	0.820	0.556	1.12	890.	443

Dataset	Dataset Y	Simulated Y	Weight	Time value	Point #
4	0.820	0.556	1.12	900.	444
4	0.821	0.556	1.12	910.	445
4	0.821	0.556	1.12	920.	446
4	0.822	0.556	1.12	930.	447
4	0.822	0.556	1.12	940.	448
4	0.822	0.556	1.12	950.	449
4	0.823	0.556	1.12	960.	450
4	0.823	0.556	1.12	970.	451
4	0.823	0.556	1.12	980.	452
4	0.824	0.556	1.12	990.	453
4	0.824	0.556	1.12	1.000E+03	454

7.3 Implementation of FITDATA.

Compiling and linking the FITDATA program is simplified in VAX/VMS by using a command procedure named MAKFITDATA. This procedure will prompt the user for any necessary information and then will link and/or compile FITDATA. There is no "transportable" version of this program which is presently available (linking FITDATA with the module QATRAN instead of QASCRN will not work). However, the AUTOSIM.DAT file can be created manually, following the guidelines listed in Table 6-1 and the need for FITDATA bypassed.

7.4 Literature References for FITSIM.

Barshop, B.A., Wrenn, R.F., and Frieden, C. (1983) Analysis of Numerical Methods for Computer Simulation of Kinetic Processes : Development of KINSIM - A Flexible, Portable System. Anal. Biochem. 130, 134-145.

Cleland, W. (1979) Statistical Analysis of Enzyme Kinetic Data. Meth. Enzymol. 63, 103-139.

Frieden, C. & Goddette, D.W. (1983) Polymerization of Actin and Actin-like systems : Evaluation of the Time Course of Polymerization in Relation to the Mechanism. Biochemistry 22, 5836-5843.

Mannervick, B. (1982) Regression Analysis, Experimental Error, and Statistical Criteria in the Design and Analysis of Experiments for Discrimination Between Rival Kinetic Models. Meth. Enzymol. 63, 103-139.

Motulsky, H.J. and Ransnas, L.A. (1987) Fitting curves to data using nonlinear regression : a practical & nonmathematical review. FASEB J. 1, 363-374.

Zimmerle, C.T., Patane, K., and Frieden, C. (1987) Divalent Cation Binding to the High and Low Affinity Sites on G-actin. Biochemistry 26, 6545-6552.

UPDATE NOTICE

KINSIM ERROR in Version 3.3*

A slight error is present in this version that will affect the input of KINSIM simulation outputs into FITSIM version 1.5 and beyond. The user would probably only notice this error if these outputs were used to check the regression routines of FITSIM. If you are using Version 3.4 of KINSIM or later, the program has been corrected to eliminate this error and you may disregard this note.

The outputs generated by KINSIM contain points from time zero to the final time listed minus the delta time. It is NOT from time zero to the final time listed in the "T" option of KINSIM. Thus the number of output points are equal only to the total simulation time divided by the delta time. The correct value for the number of output points should be this value plus 1. This error occurred due to an incorrect data file structure.

This error has been corrected in FITSIM and the simulation outputs do range from time zero to the final time listed. Unfortunately earlier KINSIM output files are not compatible with this format, and this incompatibility will result in fitting errors if these files are used in FITSIM. This incompatibility will result in a warning message in the LOG file. The errors in the final fits will generally be small. The save files output by FITSIM can be loaded by KINSIM with no apparent problem (although the Agreement option will not give the correct value).

The above error can be corrected by shifting thetime frame by one delta T. This can easily be done by using the shift option [9] of FITSIM's handling of the time zero point. See section 6.2 for further details.