

# IRTPRO™

## Contents

1. MCMC graphics .....	1
1.1 Introduction .....	1
1.2 Autocorrelations.....	2
1.3 Trace Plots.....	3
1.4 Running Means .....	3
1.5 Posterior Densities .....	4
1.6 Examples .....	5
1.6.1 3PL model with and without priors .....	5
1.6.2 Two factors CFA fitted to the AACL data.....	19

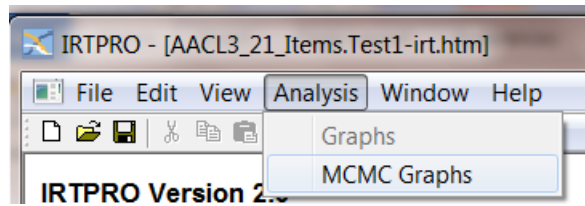
## 1. MCMC graphics

### 1.1 Introduction

The MCMC graphical procedure produces the following four types of plots that can aid further in convergence checks:

- Autocorrelations
- Trace Plots
- Running Means
- Posterior densities

Each of these plots will be briefly discussed in the following four sections and two illustrative examples will be presented in Section 1.6. To display the MCMC graphs, select the **Analysis, MCMC Graphs** option when the output file is displayed (**-irt.htm** file). Note that the estimation method has to be MCMC.



## 1.2 Autocorrelations

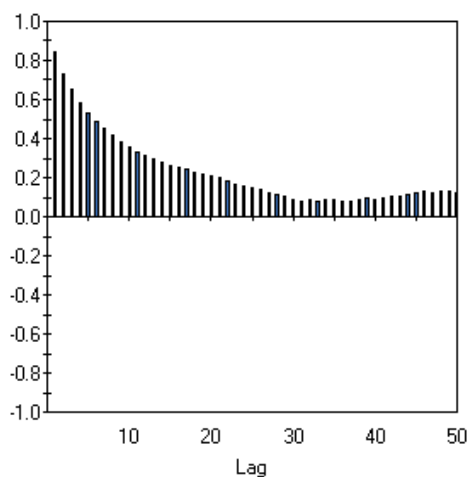
Sample autocorrelation is a statistic that estimates the theoretical autocorrelation. The sample lag- $h$  autocorrelation is given by

$$\hat{\rho}_h = \frac{\sum_{t=h+1}^T (y_t - \bar{y})(y_{t-h} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}.$$

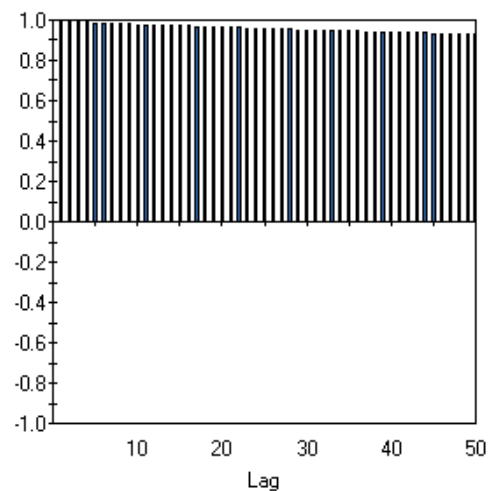
The sample autocorrelation coefficient measures the similarity between MCMC draws as a function of the time separation between them. It should be expected that the  $h$ -th lag autocorrelation is smaller with increase in  $h$  (for example, the 2nd and 30th draws should be less correlated than the 2nd and 4th draws). If the autocorrelation is still relatively high for higher values of  $k$ , it indicates a high degree of correlation between draws and therefore slow mixing.

The display below is an example of an autocorrelation plot that indicates good mixing (left pane) and one that indicates poor mixing (right pane).

Good Mixing



Bad Mixing

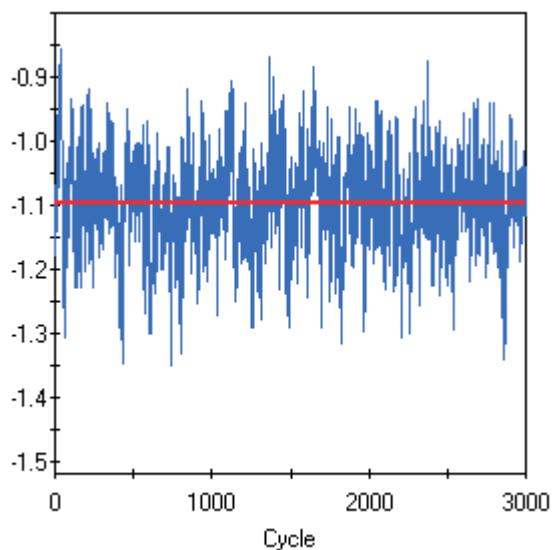


### 1.3 Trace Plots

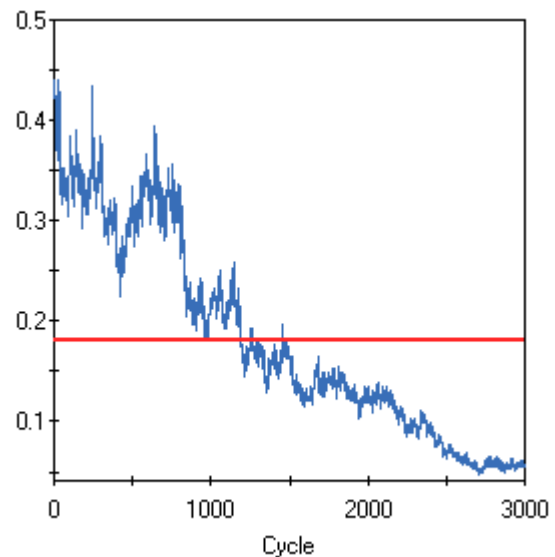
A trace plot shows the values that the relevant parameter took during the runtime of the chain. The mean (parameter estimate) of all the MCMC draws is represented by a horizontal red line.

The display below is an example of a trace plot that indicates good mixing (left pane) and one that indicates bad mixing (right pane).

Good Mixing



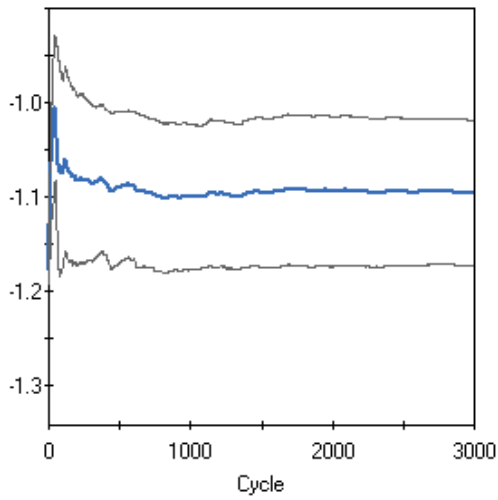
Bad mixing



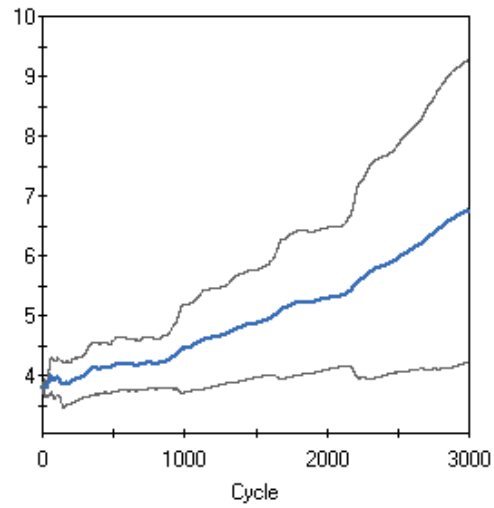
### 1.4 Running Means

Running mean plots are used to check how well the MCMC chains are mixing. The **Running Means** plot updates the means and standard deviations for each new cycle. In other words, once (for example) the mean is known for the first few observations, it is updated using a special algorithm by just adding the value of the next observation, and so on. The plots below show the means (blue line) + or – one standard deviation.

Good Mixing



Poor Mixing

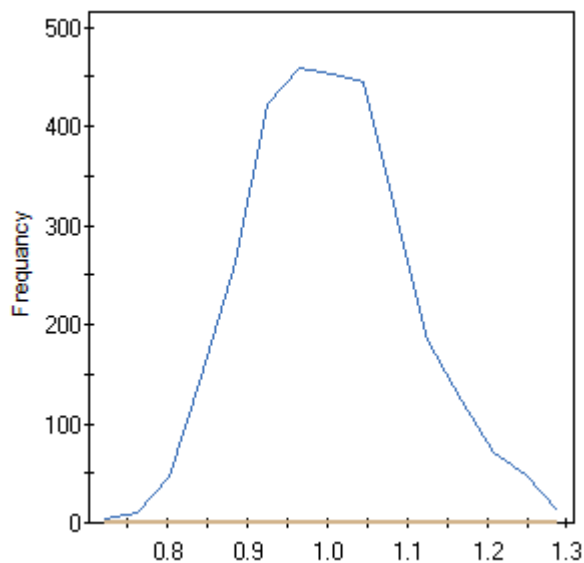


The display above is an example of a running means plot that indicates good mixing (left pane) and one that indicates bad mixing (right pane).

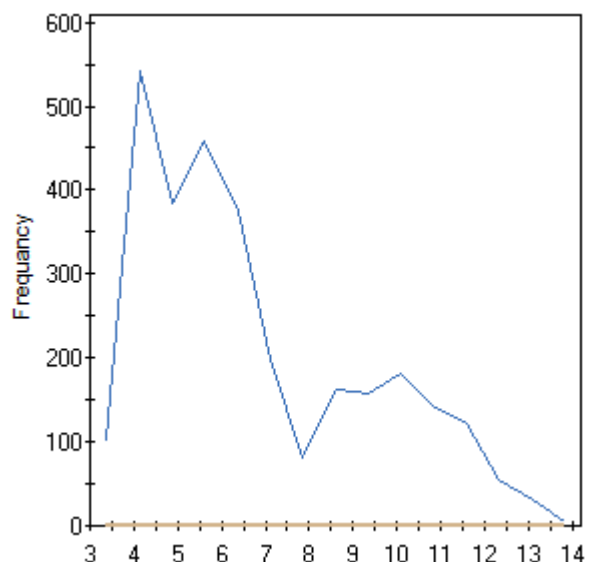
## 1.5 Posterior Densities

A posterior density plot is the histogram of the values in the trace-plot, *i.e.* the distribution of the values of the relevant parameter in the chain. These plots are usually called marginal density plots. The display below is an example of a trace plot that indicates good mixing (left pane) and one that indicates bad mixing (right pane).

Good Mixing



Bad Mixing

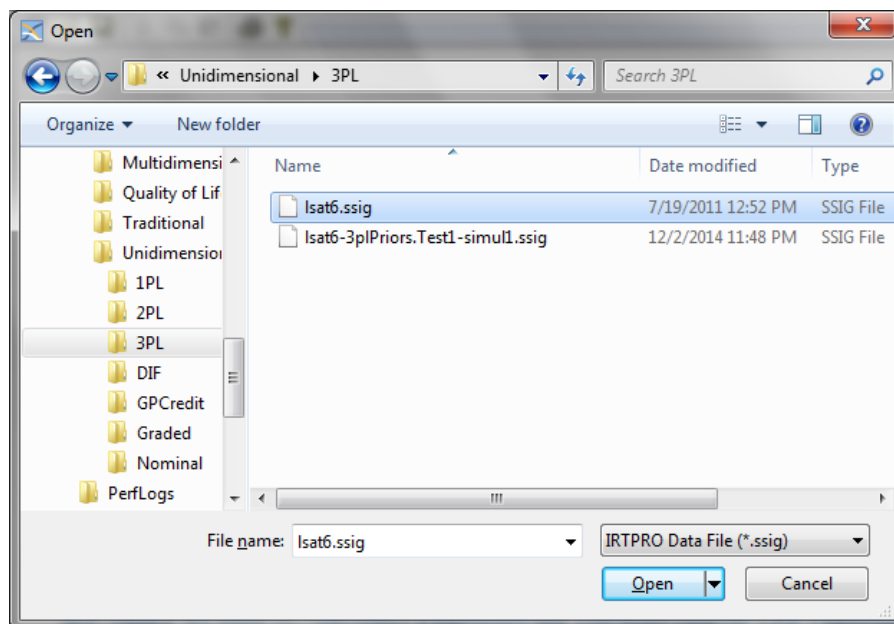


## 1.6 Examples

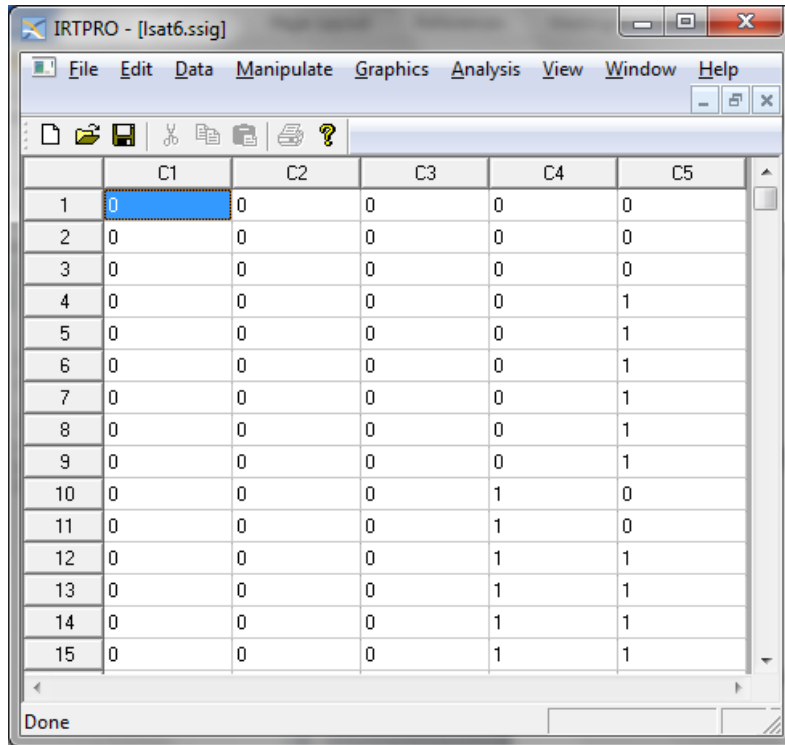
### 1.6.1 3PL model with and without priors

The dataset (**LSAT6.ssig**) considered in this section is described in Thissen (1982) and contains five dichotomously scored items obtained from the Law School Admissions Test (LSAT), Section 6.

To start the analysis, use the **File, Open** file dialog and navigate to the **C:\IRTPRO Examples\Unidimensional\3PL** folder. Change the **Files of type:** selection from its default **IRTPRO Command File (\*.irtpro)** to **IRTPRO Data File (\*.ssig)**. In the **Open** file dialog, select **LSAT6.ssig** and click the **Open** button. See the **Open** files dialog below.

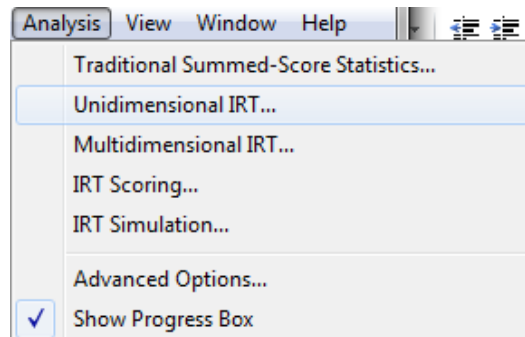


The first 15 cases of this data are shown below.



	C1	C2	C3	C4	C5
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	1
5	0	0	0	0	1
6	0	0	0	0	1
7	0	0	0	0	1
8	0	0	0	0	1
9	0	0	0	0	1
10	0	0	0	1	0
11	0	0	0	1	0
12	0	0	0	1	1
13	0	0	0	1	1
14	0	0	0	1	1
15	0	0	0	1	1

To set up the analyses, select the **Unidimensional IRT...** option from the **Analysis** menu to invoke the **Unidimensional Analysis** widow.



This window has six tabs called **Description**, **Group**, **Items**, **Models**, **Scoring** and **Simulation**. Start with the default tab **Description** and provide a title and comments in the appropriate text boxes as shown below. Note that the default name for the current analysis is Test1. Right-click on the Test1 tab and rename it to 3PL.

The screenshot shows the 'Unidimensional Analysis' dialog box. At the top, the 'Data File' is set to 'C:\IRTPRO Examples\Unidimensional\3PL\lsat6.ssig' with a 'Read file' button. Below this is a '3PL' tab. Inside the tab, there are sub-tabs: 'Description', 'Group', 'Items', 'Models', 'Scoring', and 'Simulation'. The 'Description' sub-tab is active, showing a 'Title' field with the text 'Five items from the LSAT dataset' and a 'Comments' field with the text '3PL models fitted to each item. Estimation method is MCMC', 'Lognormal(0,1) priors on slopes', and 'Beta(4,16) priors on guessing parameters.'. At the bottom of the dialog are buttons for 'Options...', 'OK', 'Cancel', and 'Run'.

Unidimensional Analysis

Data File: C:\IRTPRO Examples\Unidimensional\3PL\lsat6.ssig ... Read file

3PL

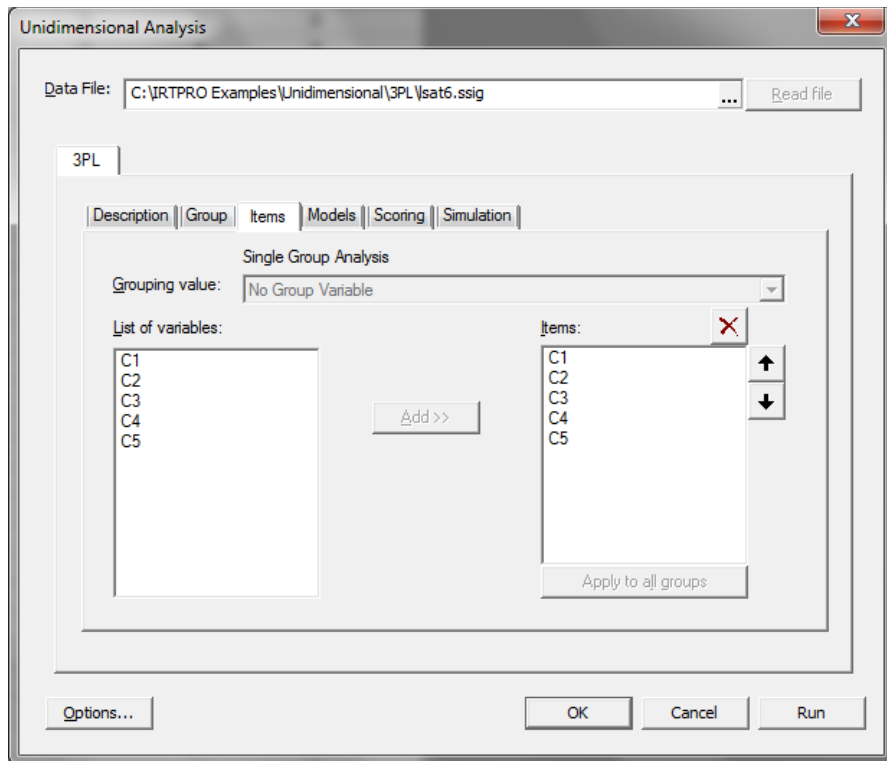
Description | Group | Items | Models | Scoring | Simulation

Title:  
Five items from the LSAT dataset

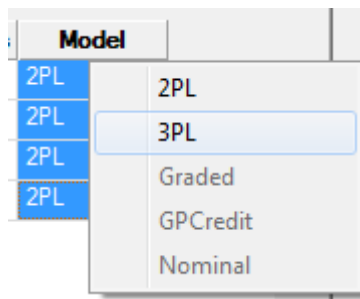
Comments:  
3PL models fitted to each item. Estimation method is MCMC  
Lognormal(0,1) priors on slopes  
Beta(4,16) priors on guessing parameters.

Options... OK Cancel Run

Since this data contains no grouping variable, the **Group** tab is skipped and we proceed to the **Items** tab, where all five items from the **List of variables** are selected. Then use the **Add** button to list these items under **Items**.

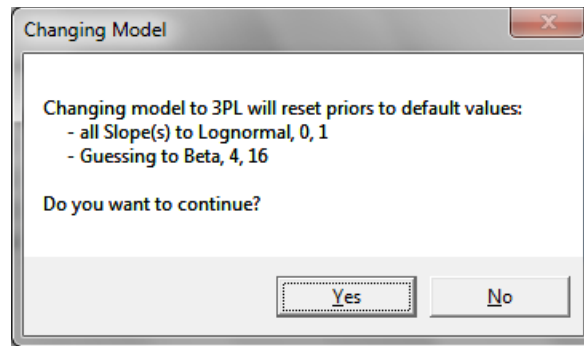


Proceed to the **Models** tab, select all the **2PL** cells, right-click and change 2PL (the default for dichotomous items) to 3PL.

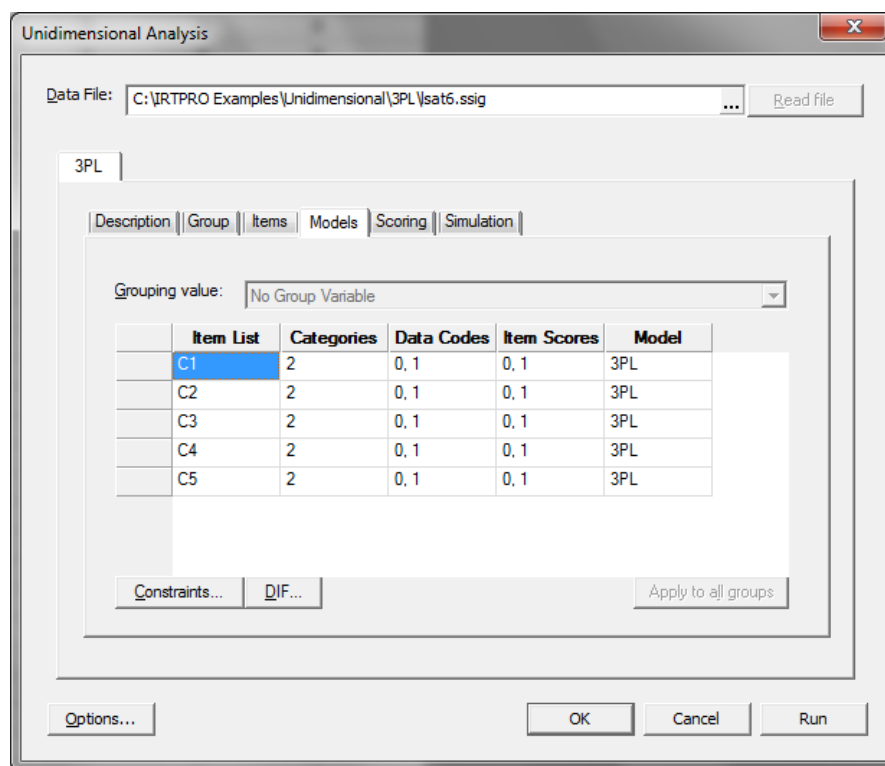


By selecting the **3PL** model, the following message is displayed:

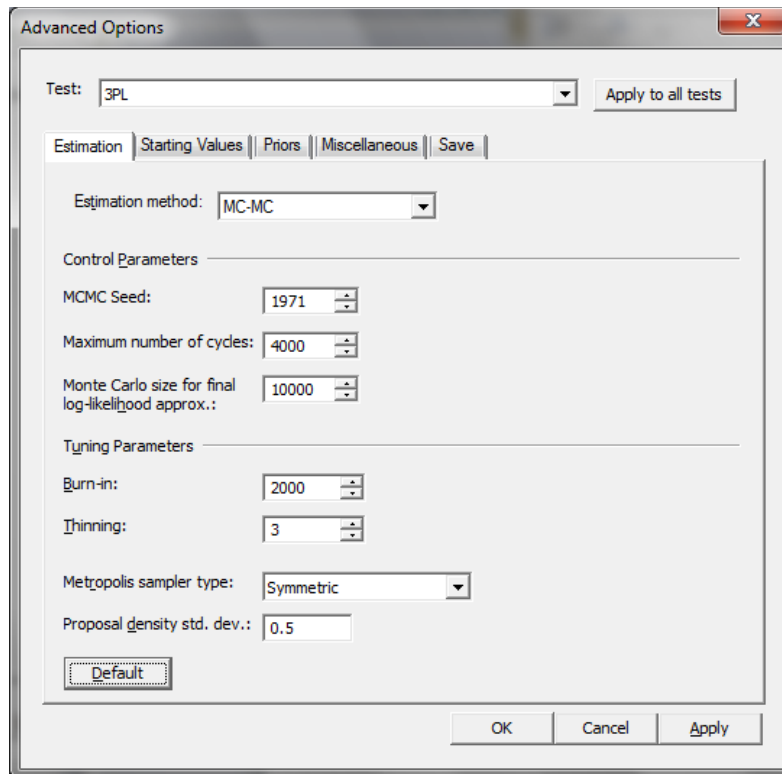




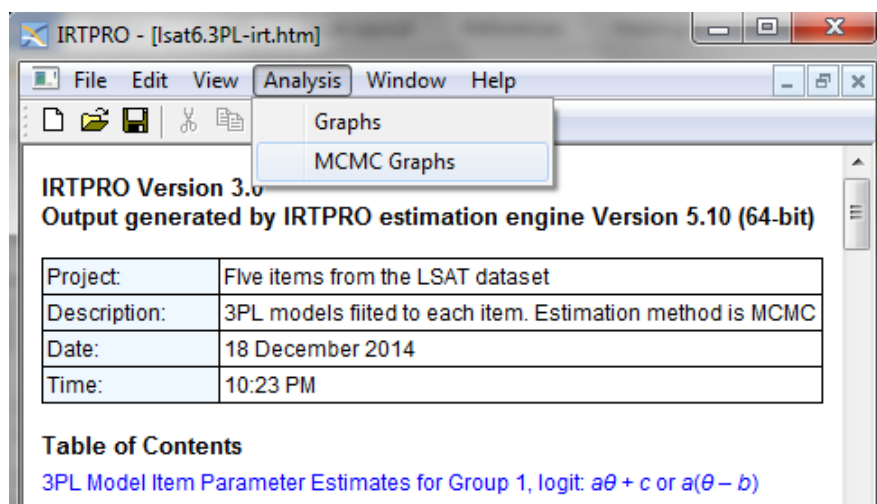
Click the **Yes** button to display the resultant **Models** dialog.



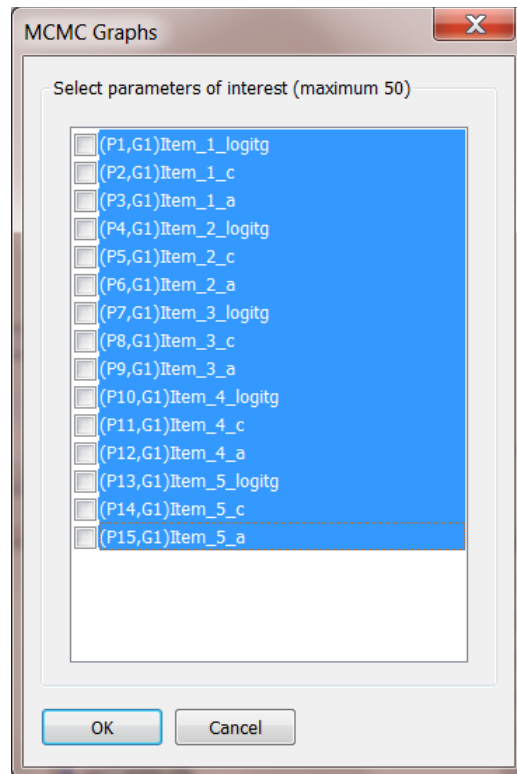
The next step is to change the estimation method from Bock-Aitkin EM to MCMC. This is accomplished by clicking the **Options** button (bottom left in the display above) in order to select MCMC from the **Estimation** tab.



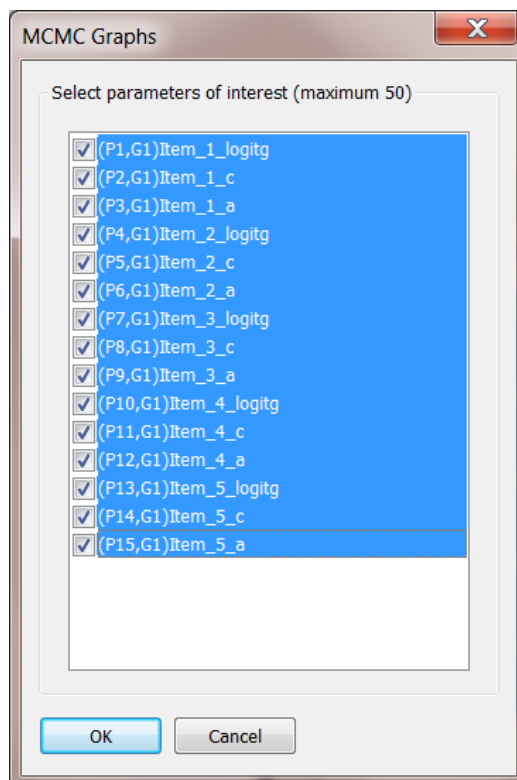
The entry of information for the analysis is now complete, and clicking the **Run** button in the lower right of the **Unidimensional Analysis** dialog will start the MCMC procedure. Once output is displayed (**Isat6.3PL.htm**), select the **Analysis, MCMC Graphs** option.



To select all the items from the resultant menu below, hold the shift key down and then click left on the first and last item label.



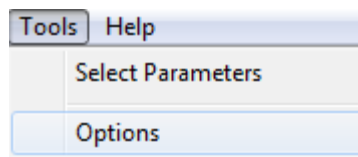
Click on any of the "boxes" on the left of the dialog to obtain the screen shown below. Alternatively, all the items marked blue will also be selected if the space bar is pressed.



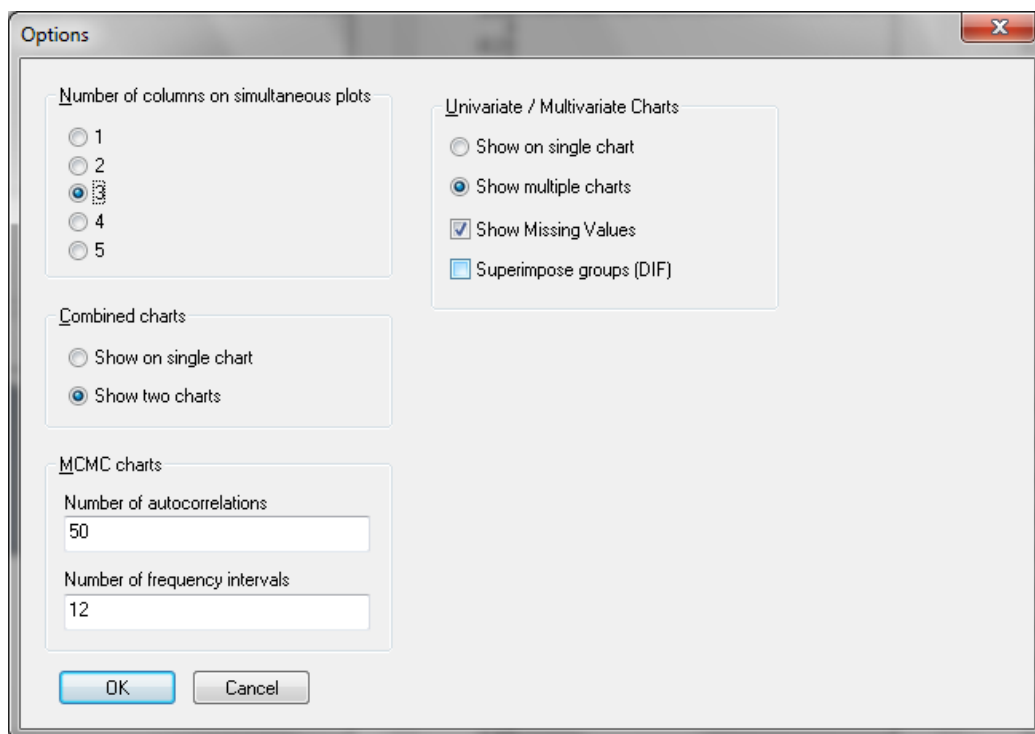
Click **OK** to display the autocorrelations. This is the default type for the four diagnostic plots, these being:

- Autocorrelations
- Trace Plots
- Running Means
- Posterior densities

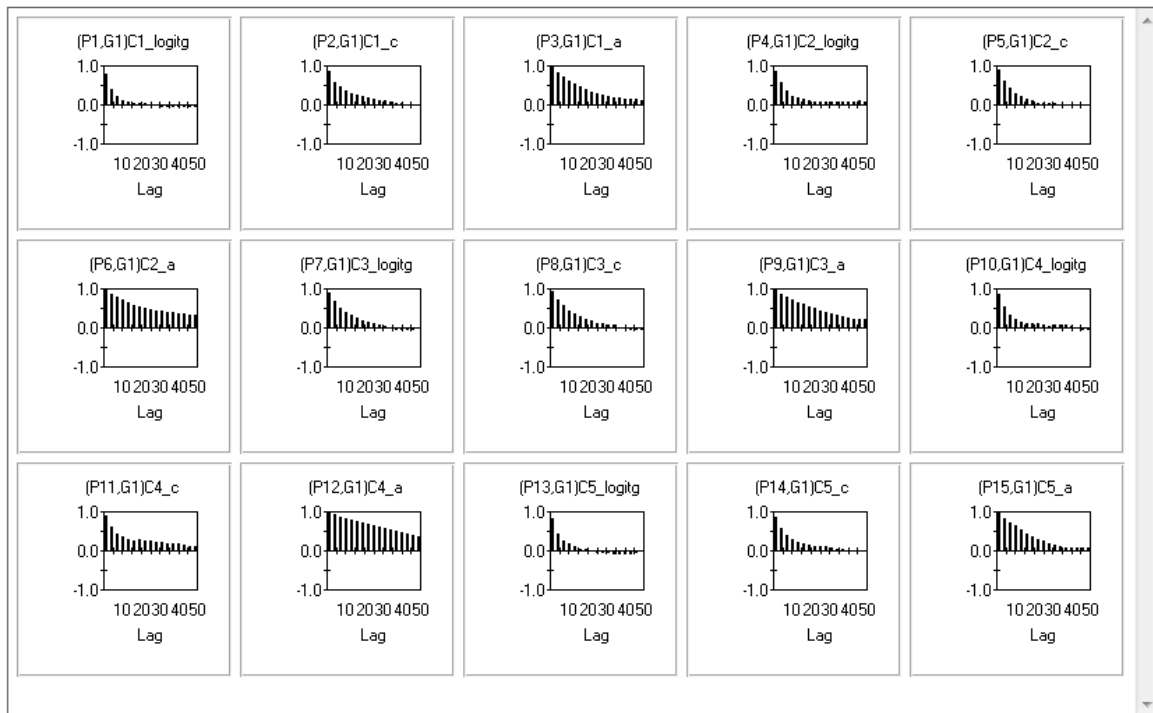
The number of autocorrelation lags, number of frequency intervals and number of columns of simultaneous plots can be controlled by selecting **Tools** from the main menu bar. This menu item is available whenever graphics are displayed.



The **Options** window displayed below contains the default values for the MCMC charts.



The autocorrelations for each of the 15 parameters up to lag 50 is shown below. Generally speaking, it appears as if good mixing has occurred.

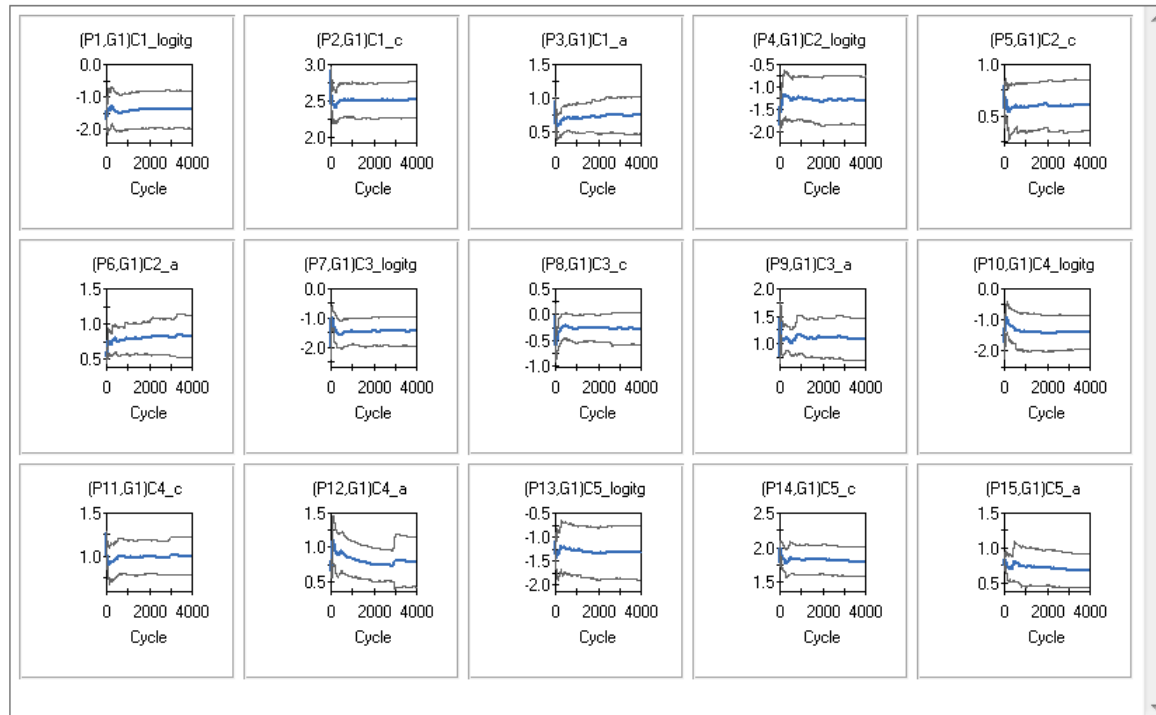


One way to see if the MCMC chain has converged is to see how well the chain is mixing, or moving around the parameter space. If the chain is taking a long time to move around the parameter space, it will take longer to converge. It can be seen how well the chain is mixing through visual inspection of every parameter.

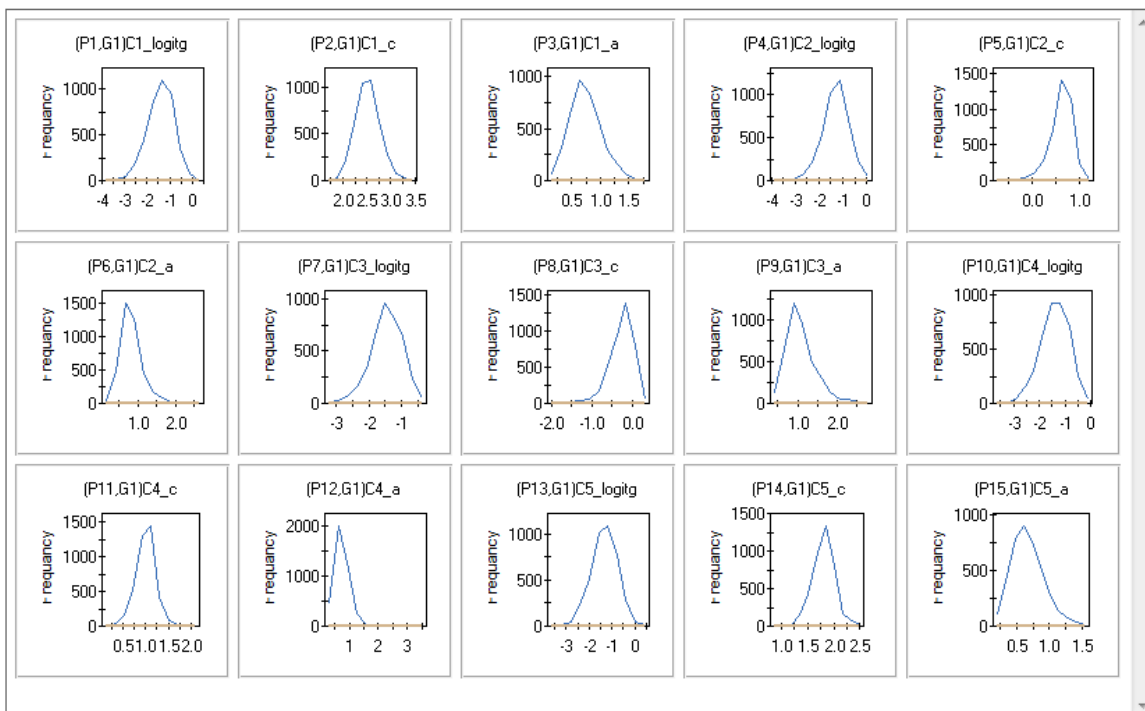
The trace plots presented below indicate that the Markov chain has stabilized and appears constant over the graph. In addition, the chain has good mixing and is "dense" in the sense that it quickly traverses the support of the distribution and is able to explore both tails and the mode areas efficiently.



Each **Running Means** plot (shown below) updates the means and standard deviations for every new cycle. In other words, once (for example) the mean is known for the first 50 observations, it is updated using a special algorithm by just adding the value of the next observation, and so on. The running means are shown in blue while the means  $\pm$  one standard deviation are shown in black.

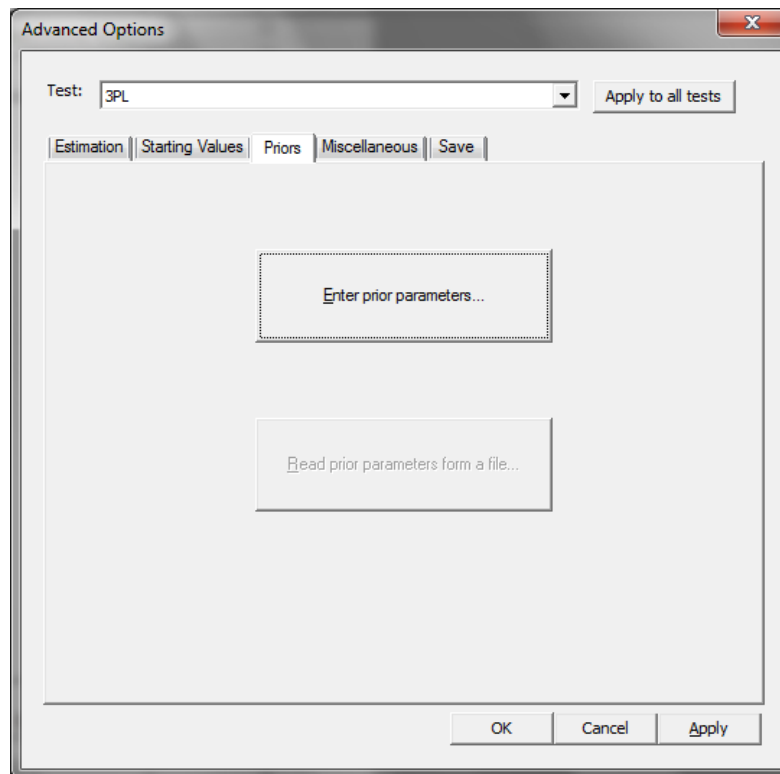


The plots displayed below are usually called marginal density plots. It is the histogram of the values in the trace-plot, *i.e.* the distribution of the values of the relevant parameter in the chain.



### 1.6.1.1 Effect of fitting the 3PL model without priors

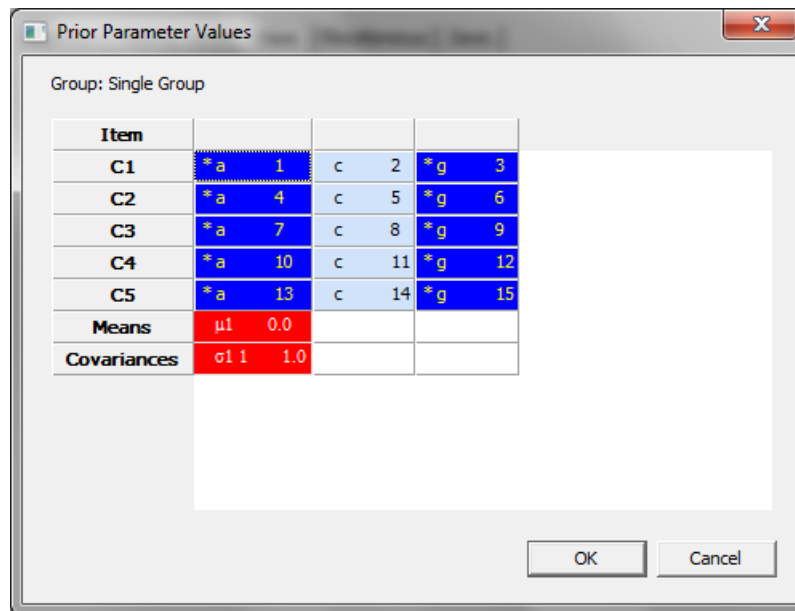
To illustrate, the priors imposed on the slope and guessing parameters were removed and a new MCMC analysis was done. This was accomplished by clicking the **Options** button to display the **Advanced Options** dialog followed by the selection of the **Priors** tab.



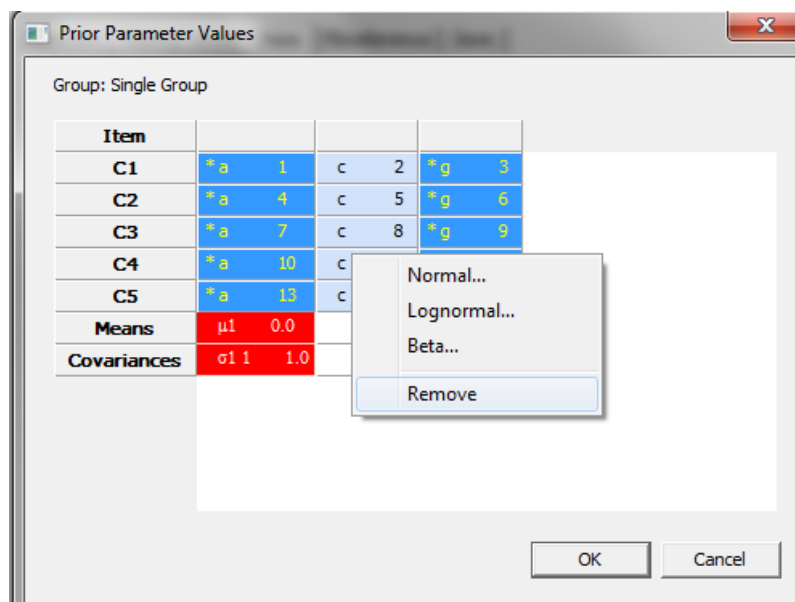
Click the **Enter prior parameters** button to open up the **Prior Parameter Values** dialog.

Cells shown in dark blue indicate parameters that have priors imposed. On the other hand, cells shown in light blue indicate parameters that have no priors imposed on them.





The dialog below shows the effect of selecting all the dark blue cells. Right-click while these cells are selected and choose **Remove** from the pop-up menu.



Because of this action, all cells (except the **Means** and **Variance** which are fixed to 0.0 and 1.0 respectively) indicate parameters with no priors imposed.

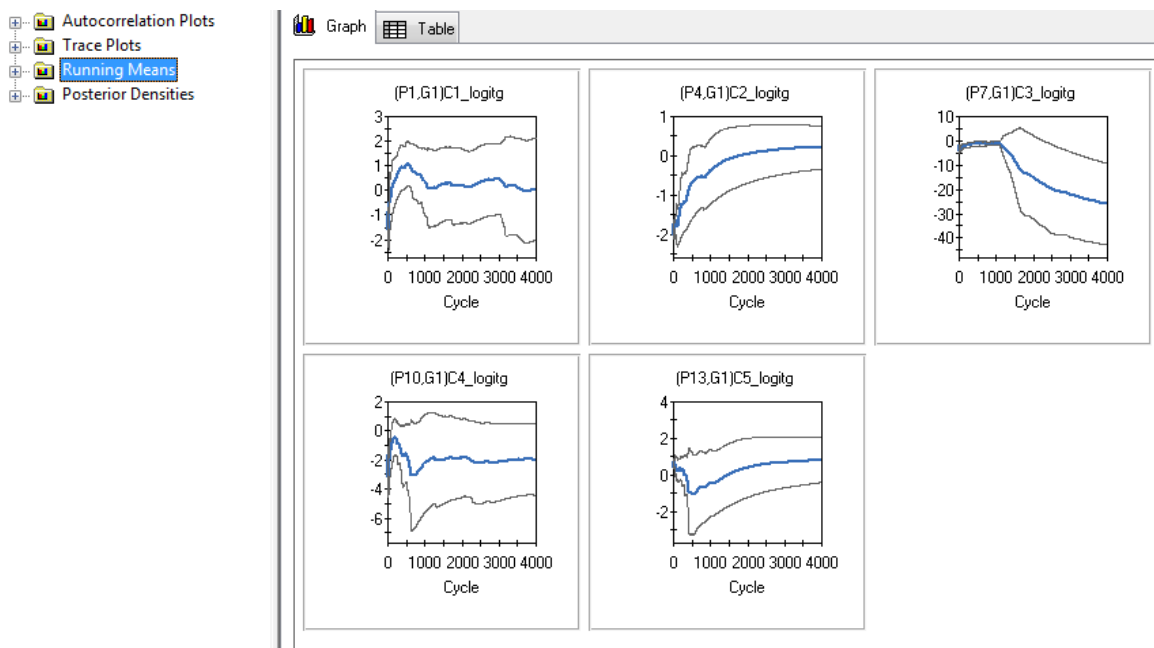
Prior Parameter Values

Group: Single Group

Item				
C1	a	1	c	2
C2	a	4	c	5
C3	a	7	c	8
C4	a	10	c	11
C5	a	13	c	14
Means	$\mu_1$	0.0		
Covariances	$\sigma_1$	1.0		

OK Cancel

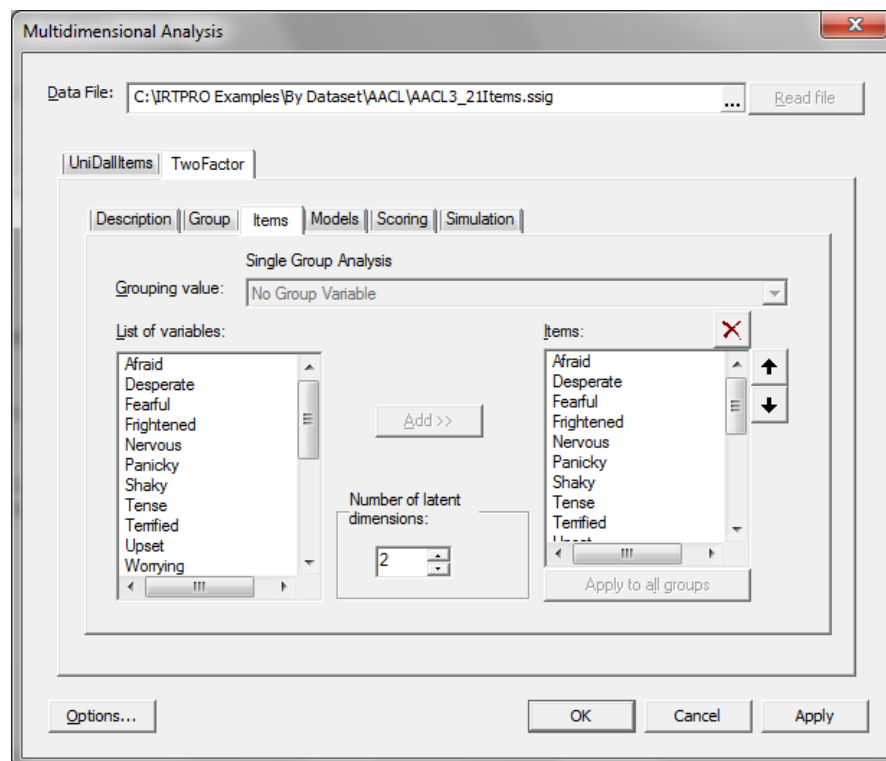
The resultant **Running Means** and **Posterior density** plots for the guessing parameters are shown next and clearly shows problems with attaining convergence, especially so in the case of item number three (C3).



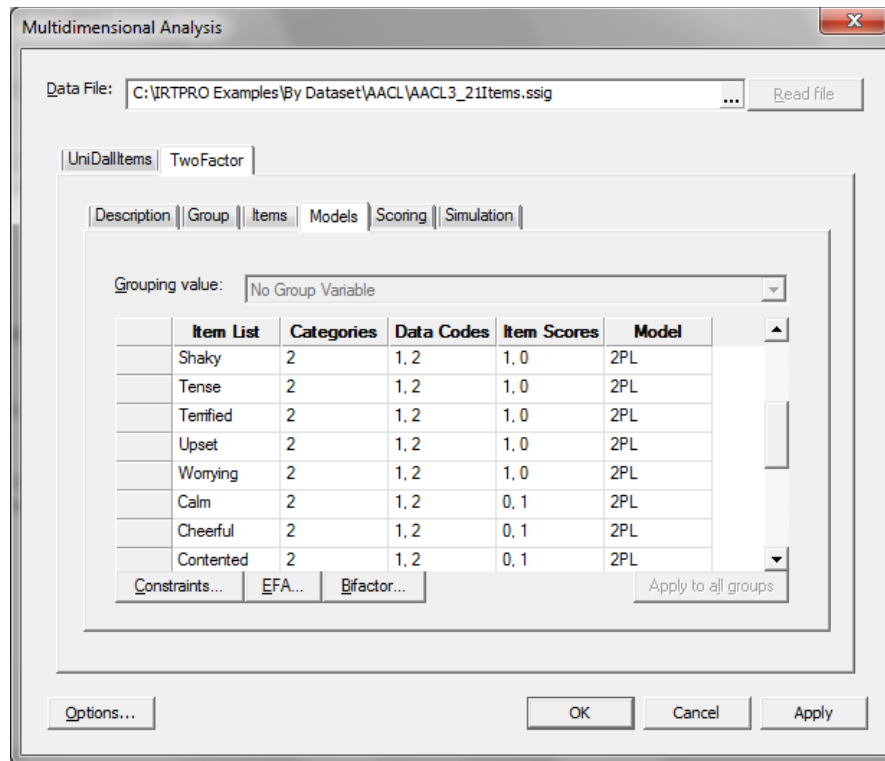
### 1.6.2 Two factors CFA fitted to the AACL data

To obtain a better fit than was obtained with a unidimensional, a two-dimensional model that fits one factor (latent variable) for the "anxiety-plus" items and a second (correlated) factor for the "anxiety-minus" items was considered. Open the file **AACL\_21Items.irtpro** stored in the folder **C:\IRTPRO Examples\by DataSet\AACL**, and then select **Multidimensional IRT ...** under the **Analysis** menu.

The dialog below represents the **Description** tab for the test named **TwoFactor**. Note that the number of latent dimensions is set equal to 2.



To view the constraints imposed upon the slope parameters, proceed to the **Models** tab shown below and click the **Constraints** button.



The constraints dialog show that all the  $a_2$ -slopes corresponding to the items Afraid to Worrying are fixed to zero, while the  $a_1$ -slopes corresponding to the items Calm to Thoughtful are set to zero.

The elements of the mean vector and covariance matrix of the latent variables are also model parameters; they are shown at the bottom of the **Item Parameter Constraints** window. In this example, the means and variances are fixed (at 0.0 and 1.0, respectively) to standardize the two latent variables. The covariance between those two standardized variables ( $\sigma_{21}$ ) is estimated – that is the correlation between the two latent variables.

Item Parameter Constraints

Group: Single Group

Item					
Nervous	a1	9	a2	0.0	c
Panicky	a1	11	a2	0.0	c
Shaky	a1	13	a2	0.0	c
Tense	a1	15	a2	0.0	c
Terrified	a1	17	a2	0.0	c
Upset	a1	19	a2	0.0	c
Worrying	a1	21	a2	0.0	c
Calm	a1	0.0	a2	23	c
Cheerful	a1	0.0	a2	25	c
Contented	a1	0.0	a2	27	c
Happy	a1	0.0	a2	29	c
Joyful	a1	0.0	a2	31	c
Loving	a1	0.0	a2	33	c
Pleasant	a1	0.0	a2	35	c
Secure	a1	0.0	a2	37	c
Steady	a1	0.0	a2	39	c
Thoughtful	a1	0.0	a2	41	c
Means	$\mu 1$	0.0	$\mu 2$	0.0	
Covariances	$\sigma 1$	1			
	$\sigma 2$	1	43	$\sigma 2$	2
				1.0	

☐ Set parameters equal across groups

OK Cancel

After inspection of the item parameter constraints, click **OK** to return to the **Models** tab and then click the **Options** button (bottom left of the **Multidimensional Analysis** window). Select the **Estimation** tab and change the estimation method to **MCMC**.

Advanced Options

Test: TwoFactor Apply to all tests

Estimation | Starting Values | Priors | Miscellaneous | Save

Estimation method: MC-MC

Control Parameters

MCMC Seed: 1971

Maximum number of cycles: 4000

Monte Carlo size for final log-likelihood approx.: 10000

Tuning Parameters

Burn-in: 2000

Thinning: 3

Metropolis sampler type: Symmetric

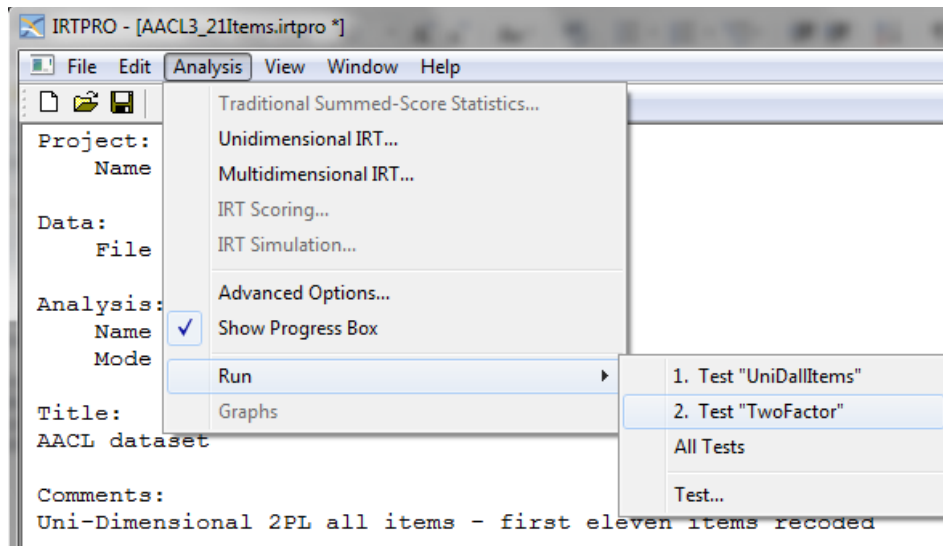
Proposal density std. dev.: 0.5

Default

OK Cancel Apply

A somewhat crude, yet reasonably effective, method dealing with autocorrelation is to only keep every  $k$  draws from the posterior and discard the rest; this is known as thinning the chain. The disadvantage is that information is being discarded; thinning can never be as efficient as using all the iterations. As shown above the **Thinning parameter** is set equal to 3. In IRTPRO, this specification implies that a random number of draws in the interval  $(0;3)$  are discarded.

Click **OK** to return to the **Multilevel Analysis** window and then click **OK** again to return to the syntax window. Select **Analysis, Run, 2.Test "TwoFactor"**.




One way to assess convergence is to assess the autocorrelations between the draws of the Markov chain. We would expect the  $k$ -th lag autocorrelation to be smaller as  $k$  increases (our 2nd and 50th draws should be less correlated than our 2nd and 4th draws). If autocorrelation is still relatively high for higher values of  $k$ , this indicates a high degree of correlation between our draws and slow mixing.

The parameter estimates are shown below and it is noted that the slope and intercept parameter estimates for the item Frightened are relatively large.

IRTPRO - [AACL3\_21\_Items.Test1-irt.htm]

File Edit View Analysis Window Help



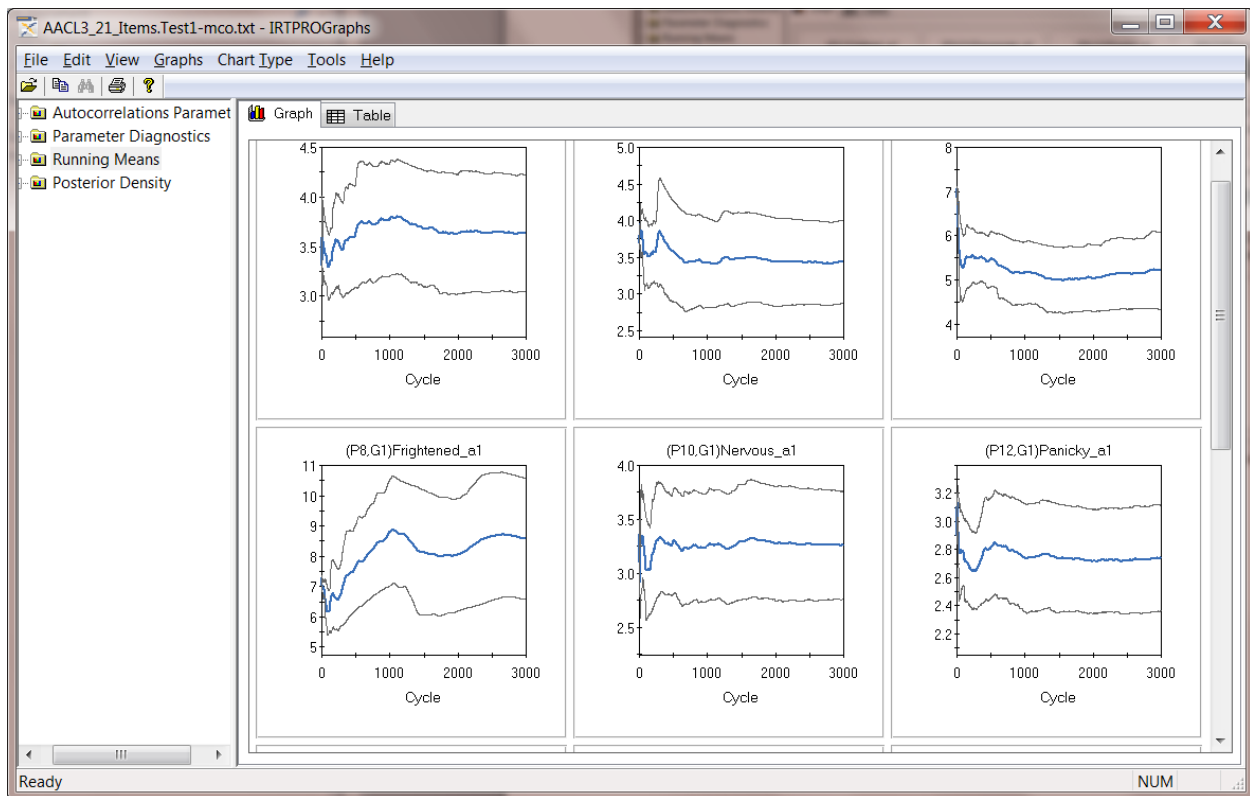
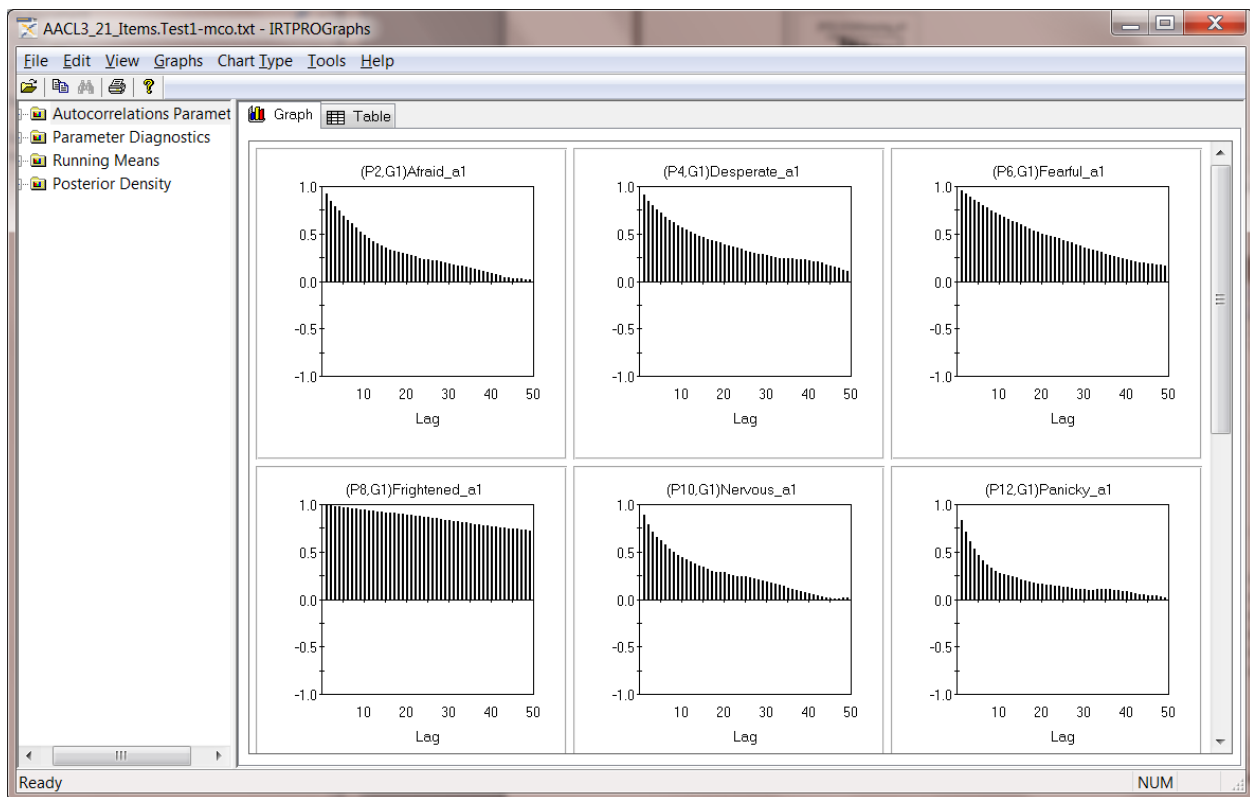
Item	Label		$a_1$	s.e.	$a_2$	s.e.	c	s.e.		
1	Afraid	2	3.63	0.59	0.00	----	1	4.26	0.60	
2	Desperate	4	3.43	0.56	0.00	----	3	3.90	0.58	
3	Fearful	6	5.21	0.86	0.00	----	5	5.55	0.87	
4	Frightened	8	8.58	1.99	0.00	----	7	10.34	2.42	
5	Nervous	10	3.26	0.50	0.00	----	9	1.54	0.30	
6	Panicky	12	2.74	0.38	0.00	----	11	2.64	0.35	
7	Shaky	14	2.35	0.38	0.00	----	13	2.85	0.36	
8	Tense	16	3.02	0.44	0.00	----	15	0.58	0.22	
9	Terrified	18	4.64	1.00	0.00	----	17	8.84	1.67	
10	Upset	20	2.15	0.33	0.00	----	19	2.26	0.30	
11	Worrying	22	4.35	1.01	0.00	----	21	-0.38	0.29	
12	Calm		0.00	----	24	1.61	0.26	23	-1.11	0.19
13	Cheerful		0.00	----	26	2.36	0.37	25	-0.95	0.23
14	Contented		0.00	----	28	3.15	0.47	27	-1.14	0.29
15	Happy		0.00	----	30	3.91	0.70	29	-1.81	0.42
16	Joyful		0.00	----	32	2.95	0.57	31	1.20	0.30
17	Loving		0.00	----	34	1.60	0.26	33	-0.59	0.18
18	Pleasant		0.00	----	36	2.86	0.51	35	-2.92	0.39
19	Secure		0.00	----	38	2.28	0.36	37	-0.92	0.22
20	Steady		0.00	----	40	2.19	0.38	39	-1.58	0.26
21	Thoughtful		0.00	----	42	1.60	0.28	41	-1.87	0.23

Done

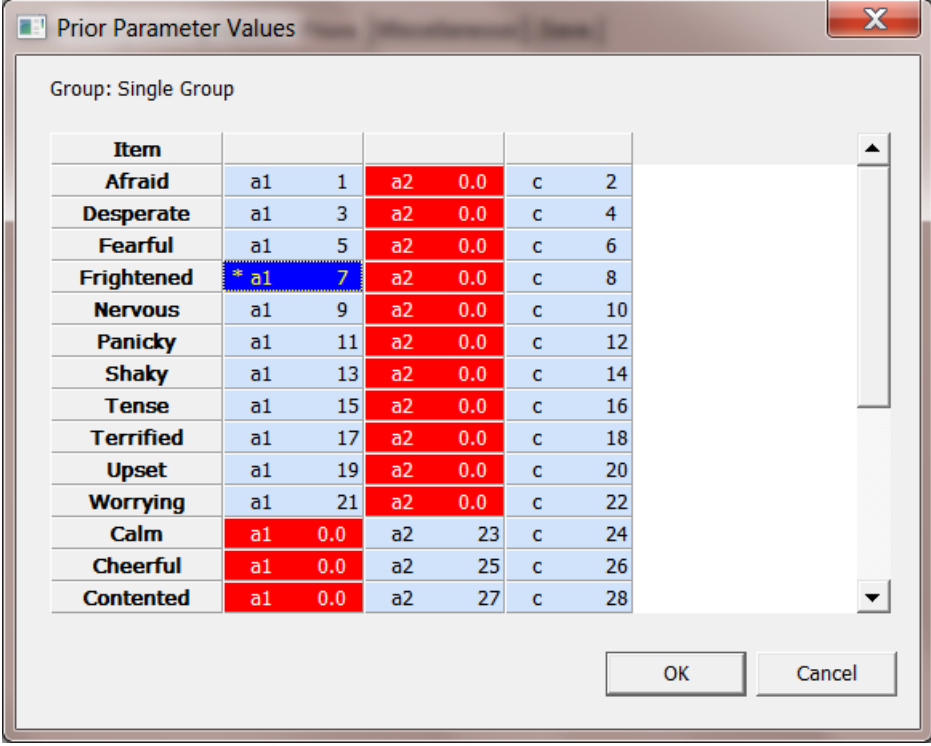
NUM

To investigate further, MCMC diagnostic plots of the slope parameters for the first factor were requested. Both the autocorrelation and running means plots indicate that stable parameter estimates for the item Frightened is difficult to obtain.





One can expect noise at the beginning of the plots; however, if trends are evident, there might be a problem. To address this problem, a normal prior with mean 0 and standard deviation of 3 was imposed on Frightened:



Group: Single Group

Item	a1		a2		c	
Afraid	a1	1	a2	0.0	c	2
Desperate	a1	3	a2	0.0	c	4
Fearful	a1	5	a2	0.0	c	6
Frightened	* a1	7	a2	0.0	c	8
Nervous	a1	9	a2	0.0	c	10
Panicky	a1	11	a2	0.0	c	12
Shaky	a1	13	a2	0.0	c	14
Tense	a1	15	a2	0.0	c	16
Terrified	a1	17	a2	0.0	c	18
Upset	a1	19	a2	0.0	c	20
Worrying	a1	21	a2	0.0	c	22
Calm	a1	0.0	a2	23	c	24
Cheerful	a1	0.0	a2	25	c	26
Contented	a1	0.0	a2	27	c	28

OK Cancel

The imposition of a prior seemed to have solved the problem of poor mixing as shown by the **Autocorrelation**, **Running Means** and **Posterior density** plots given next.

The last plot is a set of pie charts representing the distribution of the first six items. From this display it is noted that, compared to the other items, the percentage values in category 1 is rather low.

