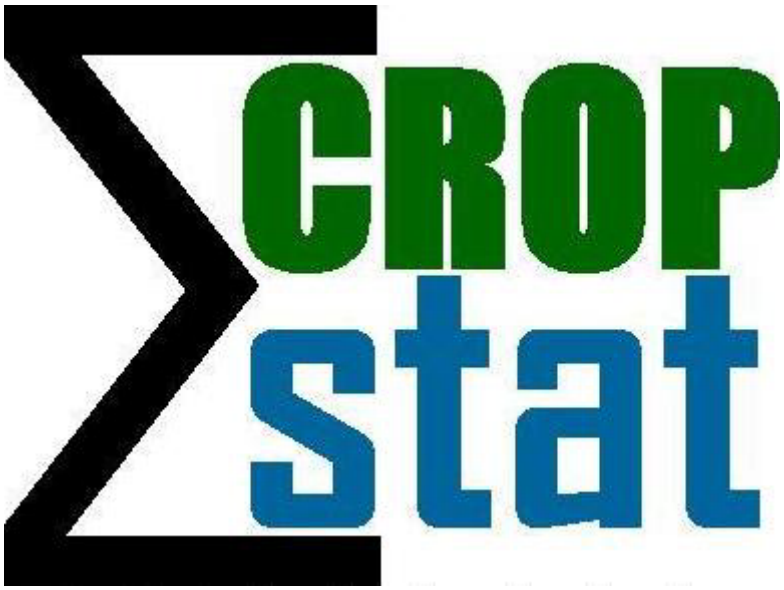


V E R S I O N 7 . 2

Tutorial Manual Part 1

Crop Research Informatics Laboratory
INTERNATIONAL RICE RESEARCH INSTITUTE

A NOTE TO THE READER:
An electronic copy of this tutorial manual comes with the CropStat installer. The CropStat Tutorial may be printed/copied and distributed to any number of users. CropStat is a freeware developed for non-profit use. Hence, selling of either the software or the tutorial is prohibited.

ISBN 978-971-22-0218-6
Revised July 2009

Tutorial Part 1: Contents

Tutorial Part 1: Contents

Tutorial Part 2: Contents

Tutorial Part 2: Contents

INTRODUCTION TO CROPSTAT 7.2

CropStat is a computer program for data management and basic statistical analysis of experimental data. It can be run in any 32-bit Windows operating system. To install the program on your computer, run the file *SETUP.EXE* on the installation disk. To start the program, run the file *CROPSTAT.EXE*, or click the icon that will be installed by the SETUP program.

CropStat has been developed primarily for the analysis of data from agricultural field trials, but many of the features can be used for analysis of data from other sources.

The main modules and facilities are

1. Data management with a spreadsheet
2. Text editor
3. Descriptive statistics and Scatterplot Graphics
4. Balanced analysis of variance
5. Unbalanced analysis (generalized linear models)
6. Linear Mixed Models
7. Combined analysis of variance
8. Analysis of repeated measures
9. Regression and correlation analysis
10. Single-site analysis for variety trials
11. Spatial Analysis
12. Genotype × environment interaction analysis
13. Pattern Analysis
14. Quantitative trait loci analysis
15. Graphics
16. Utilities for randomization and layout, and orthogonal polynomial
17. Analysis of Categorical Data

The CropStat for Windows v6 Tutorial Manual is divided into fifteen (15) main modules. Each module provides the user step-by-step instructions on how to perform certain tasks of interest to CropStat users. Screen images have been included as deemed helpful. Since CropStat provides extensive online documentation, several features of this package are not covered in this tutorial.

Menu items, names of dialogs and form controls are in **bold** letters. Filenames, variables names and directories are *italicized*. Locations of sample datasets that come with the

package or suggested working directories are given in each module. The tutorial assumes, though not required, that CropStat is installed in *C:\PROGRAM FILES*. Hence, the working directory used in this tutorial is *C:\PROGRAM FILES\CROPSTAT7.2*. Datasets used in the discussion and in the exercises for each module are found in *C:\PROGRAM FILES\CROPSTAT7.2\TUTORIAL\TUTORIAL DATASETS* folder.

As a convention, the command file names used are the same as the data file name. For example to perform an analysis on the CropStat data file *SAMPLE.SYS,* a command file will be created named *SAMPLE.GFC*. Discussion on file types and naming conventions are discussed in detail in later modules.

CROPSTAT ENVIRONMENT

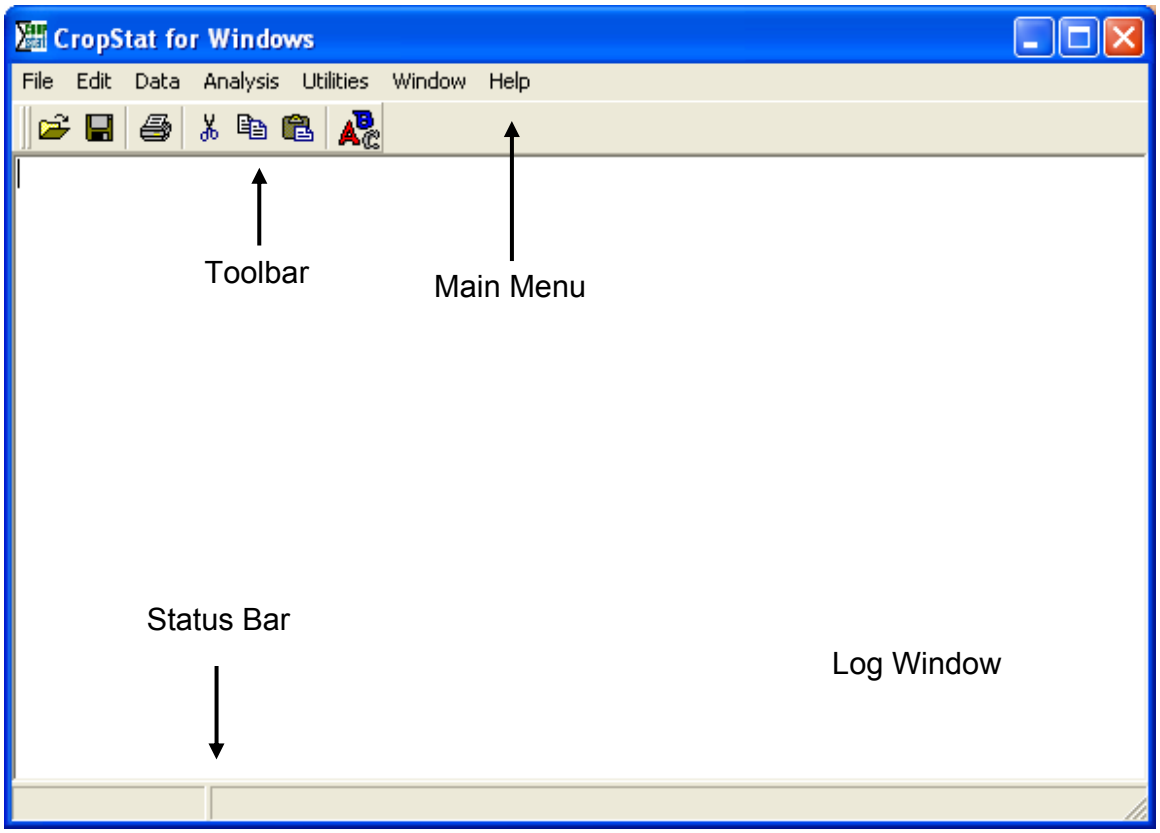
At the end of this section, the user should be familiar with the different features of CropStat such as Windows, Menus, Job Specification Components, and Filename Convention.

I. Windows

There are four types of windows in CropStat.

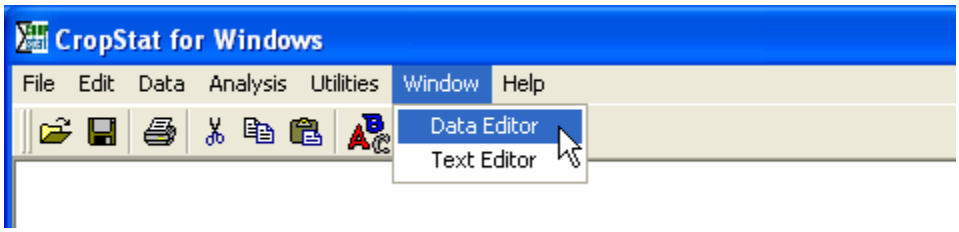
A. Main Window

The Main Window is CropStat's main and log window. It provides the main menu and toolbars.

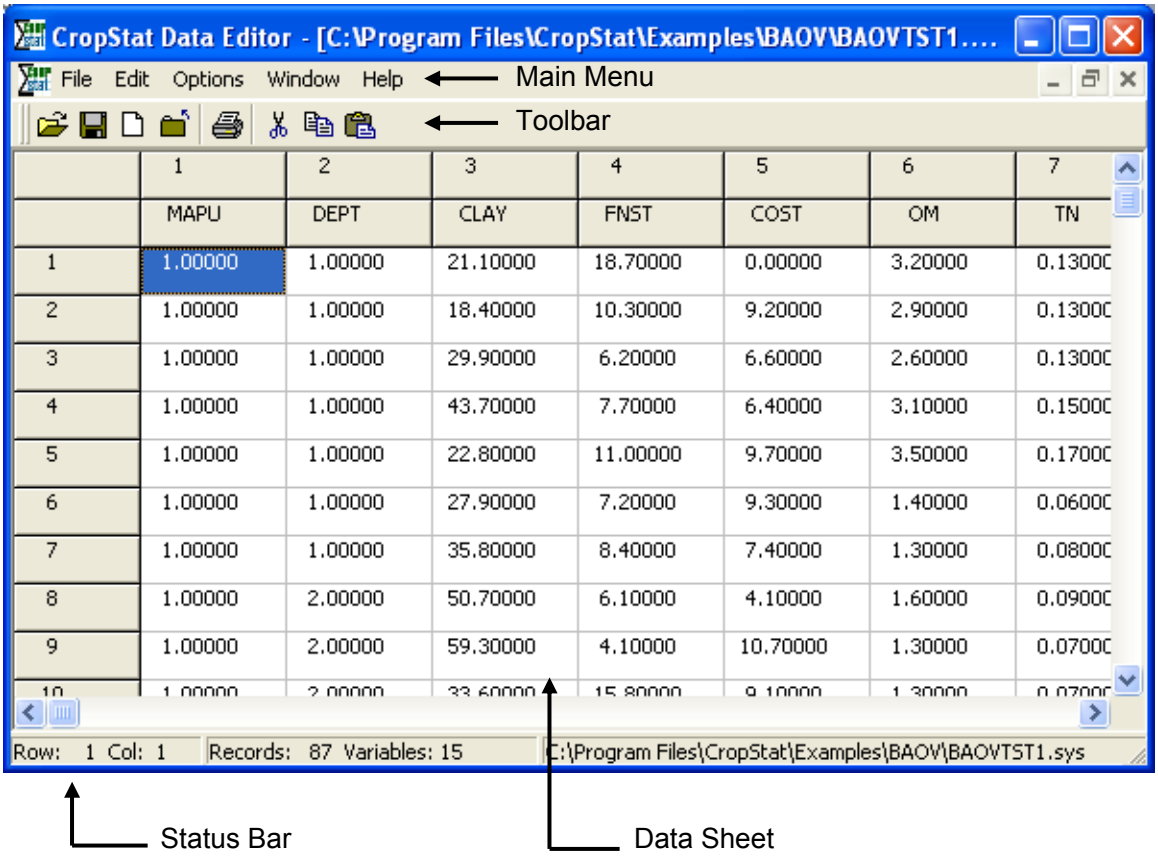


B. Data Editor

The Data Editor displays data in a row-by-column format. Each row is a case, and each column is a variable. You can type new data into an empty data sheet, or you can import data from text file, excel file, and dbase file.



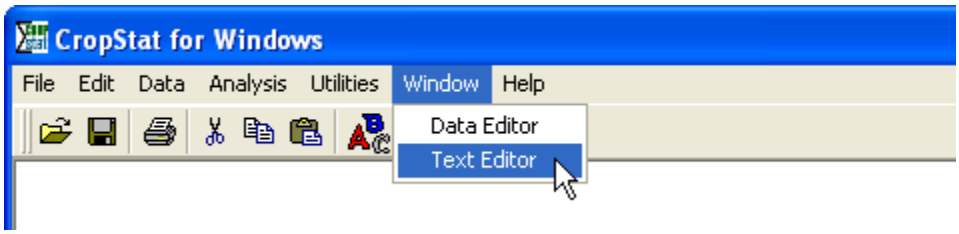
You can open simultaneously two or more data files. The number of data file you can open depends on the memory capacity of your computer.

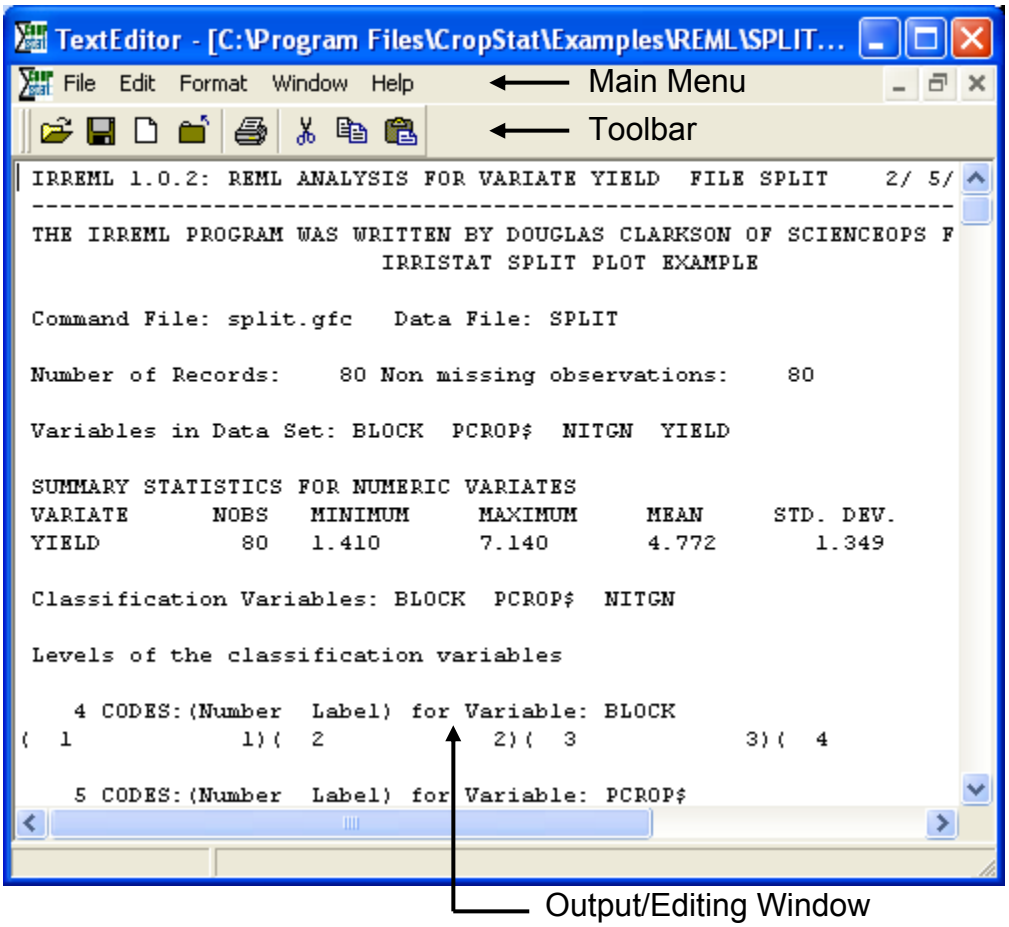


	1	2	3	4	5	6	7
	MAPU	DEPT	CLAY	FNST	COST	OM	TN
1	1.00000	1.00000	21.10000	18.70000	0.00000	3.20000	0.13000
2	1.00000	1.00000	18.40000	10.30000	9.20000	2.90000	0.13000
3	1.00000	1.00000	29.90000	6.20000	6.60000	2.60000	0.13000
4	1.00000	1.00000	43.70000	7.70000	6.40000	3.10000	0.15000
5	1.00000	1.00000	22.80000	11.00000	9.70000	3.50000	0.17000
6	1.00000	1.00000	27.90000	7.20000	9.30000	1.40000	0.06000
7	1.00000	1.00000	35.80000	8.40000	7.40000	1.30000	0.08000
8	1.00000	2.00000	50.70000	6.10000	4.10000	1.60000	0.09000
9	1.00000	2.00000	59.30000	4.10000	10.70000	1.30000	0.07000
10	1.00000	2.00000	33.60000	15.80000	9.10000	1.30000	0.07000

C. Text Editor

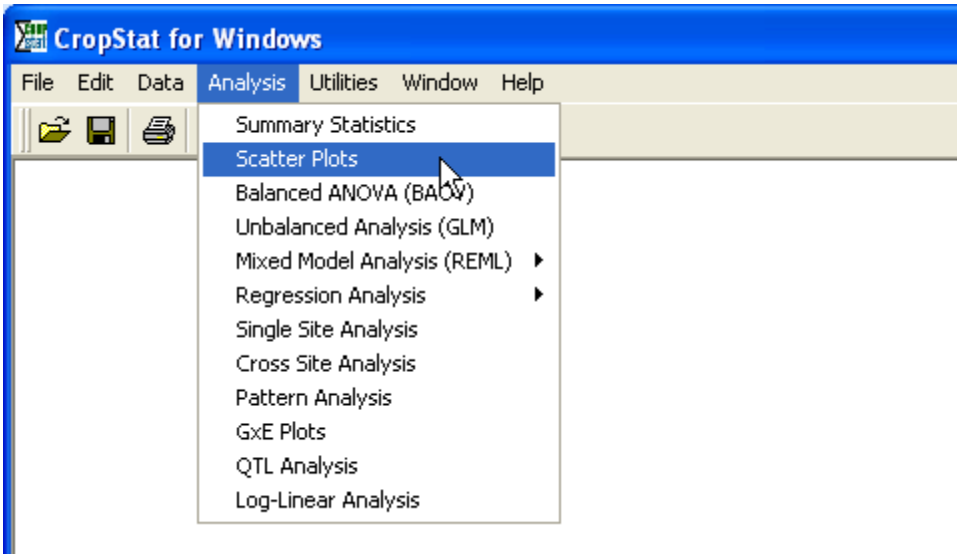
ASCII files can be created and edited in the Text Editor. Command files and statistical results can also be viewed in the editor.

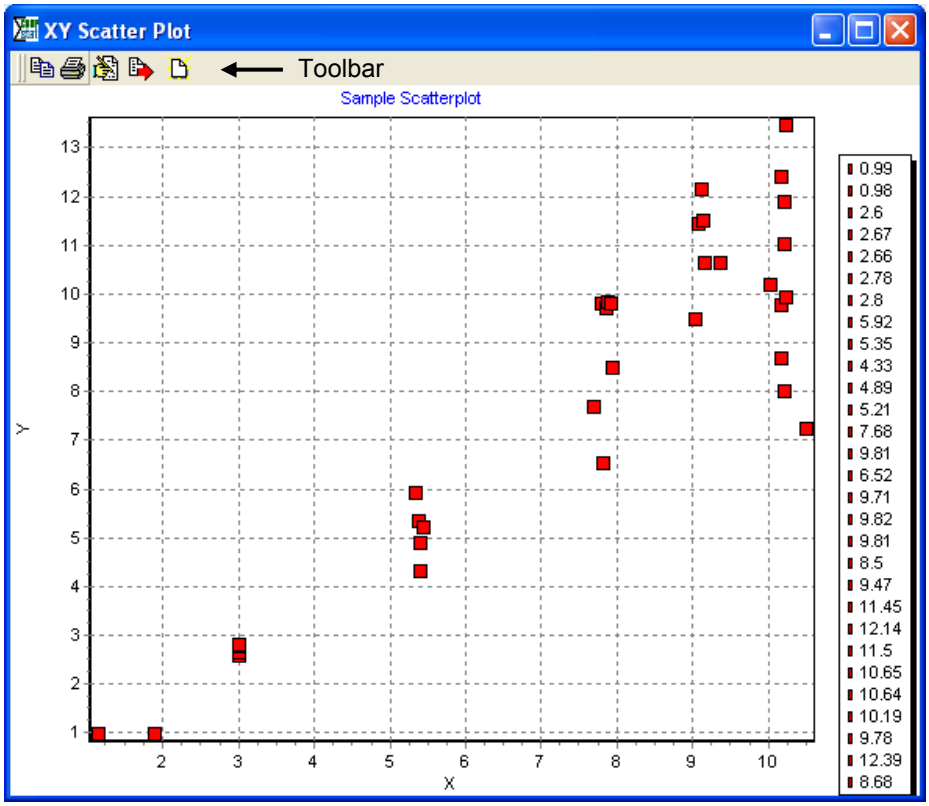




D. Graph Window

The output generated by the Graphics Module appears in the Graph window.

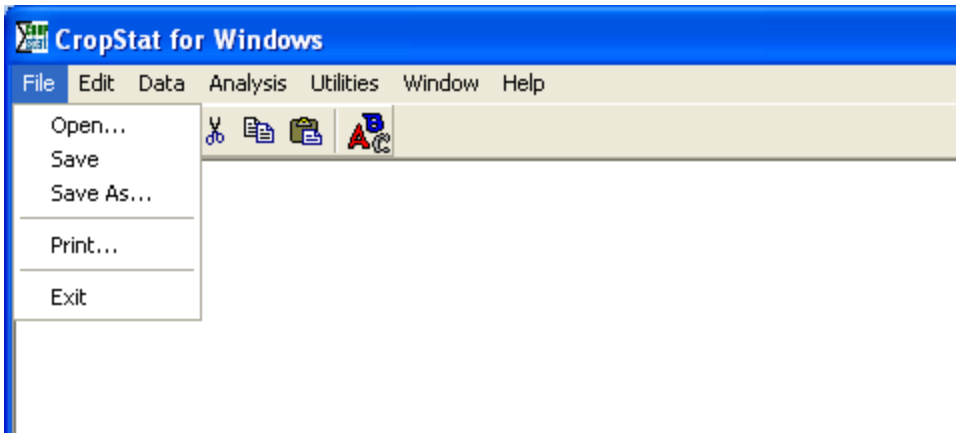




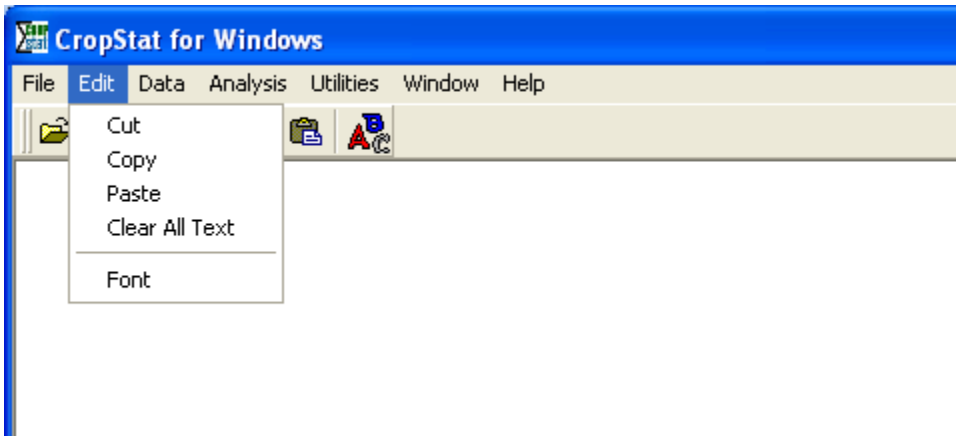
II. Menus

Each CropStat window has its own menu bar that contains menus and selections appropriate to that window.

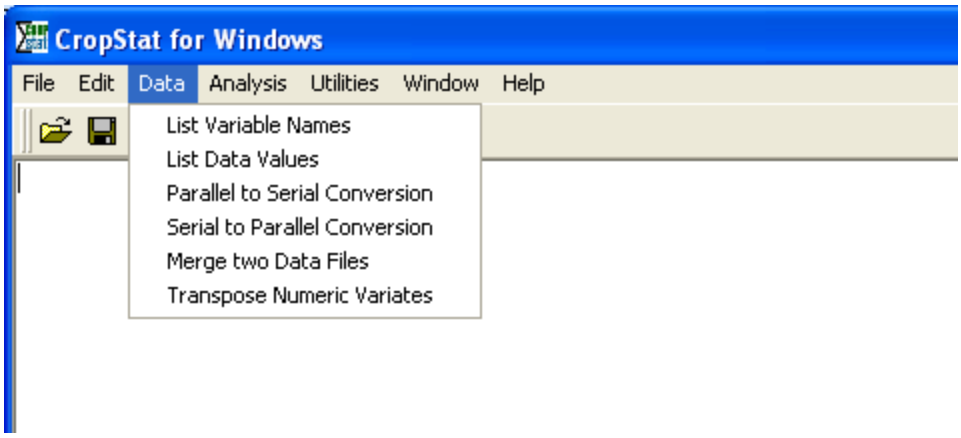
A. Main Window Menus



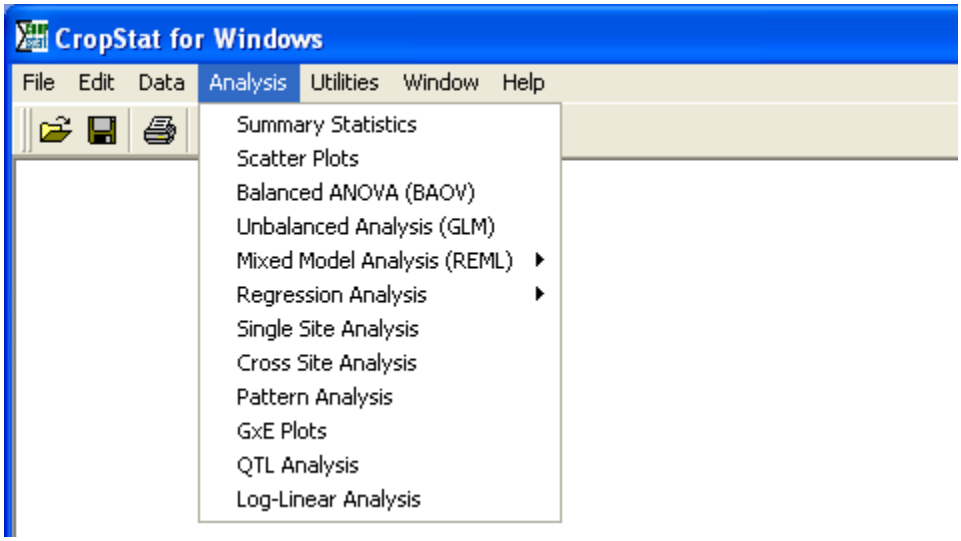
Use the **File** menu to open a file using the Data Editor or the Text Editor (depends on the type of file), save the log file or print the log file.



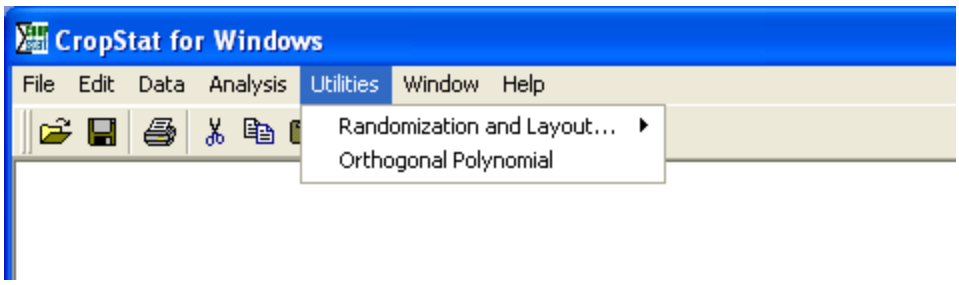
Use the **Edit** menu to cut, copy, paste, or clear the Log window.



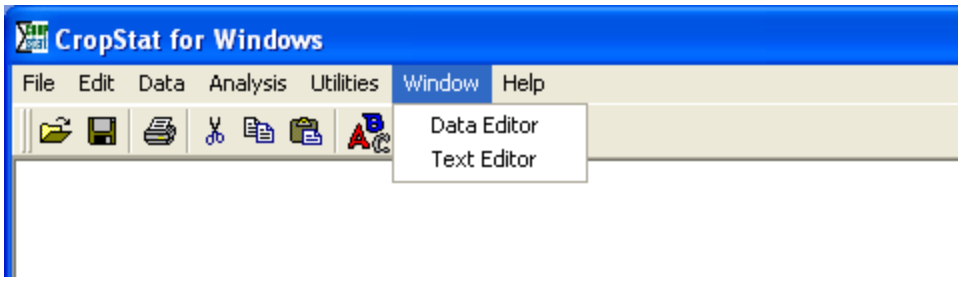
Use the **Data** menu to list variables of a particular SYS file.



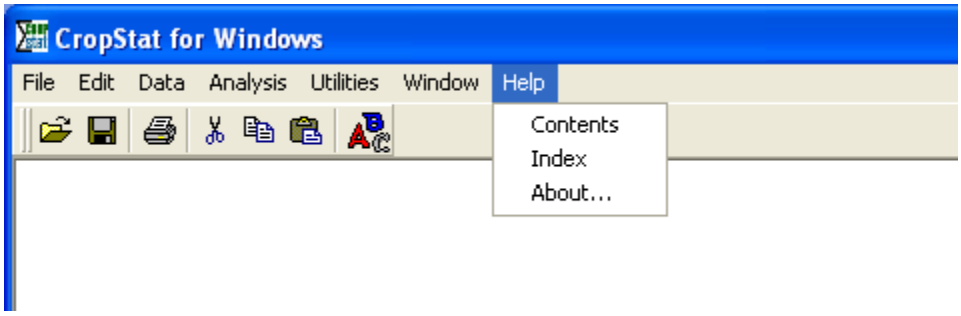
Use the **Analysis** menu to run statistical procedures including analysis of variance, regression, G×E analysis, quantitative trait loci analysis, single-site analysis, pattern analysis and produce summary statistics and scatter plots.



Use the **Utilities** menu to produce experimental design layout and to generate general factorial expected mean squares and coefficients for orthogonal polynomial contrasts.

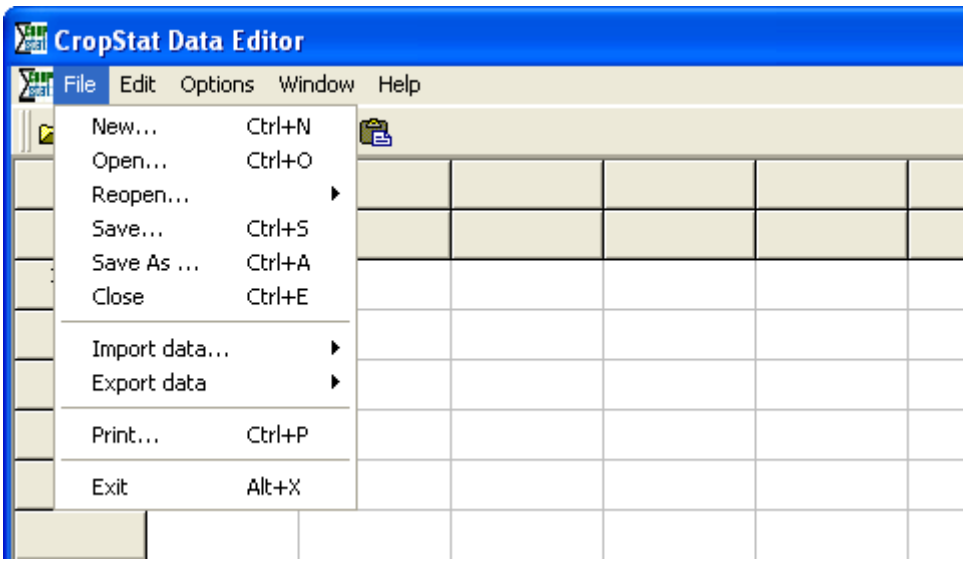


Use the **Window** menu to switch between the different windows in CropStat or rearrange the display of windows.

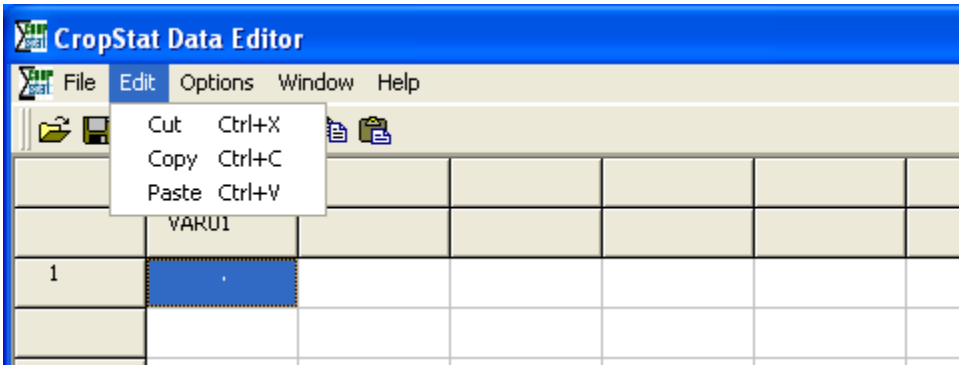


Use the **Help** Menu to access CropStat's online Help system.

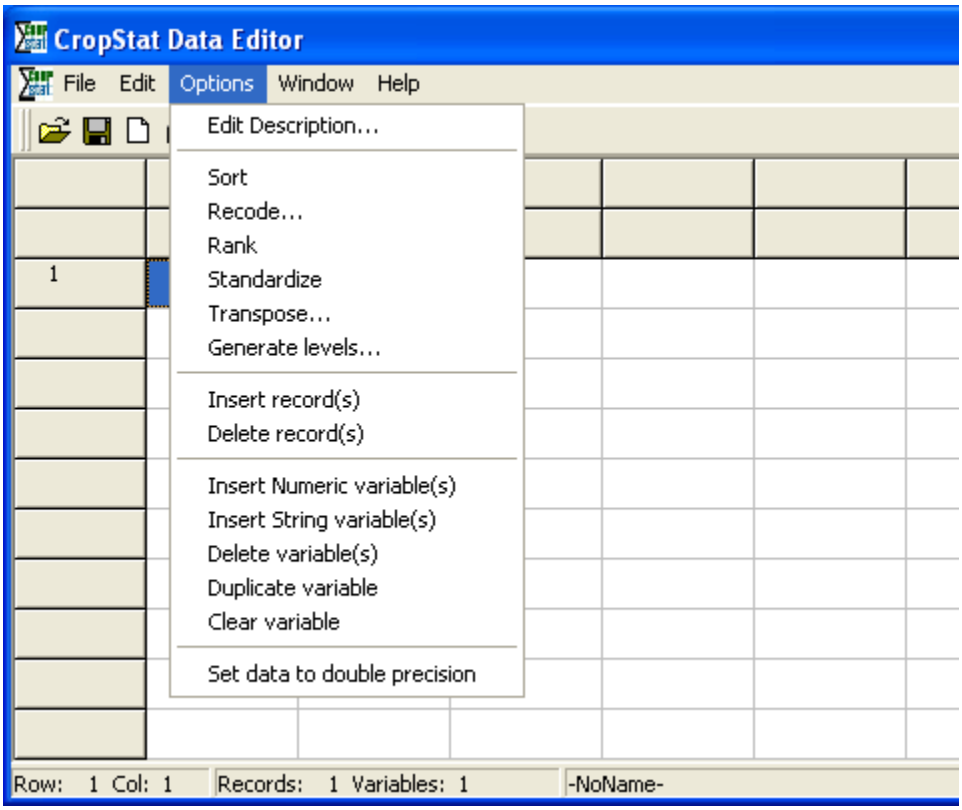
B. Data Editor Menus



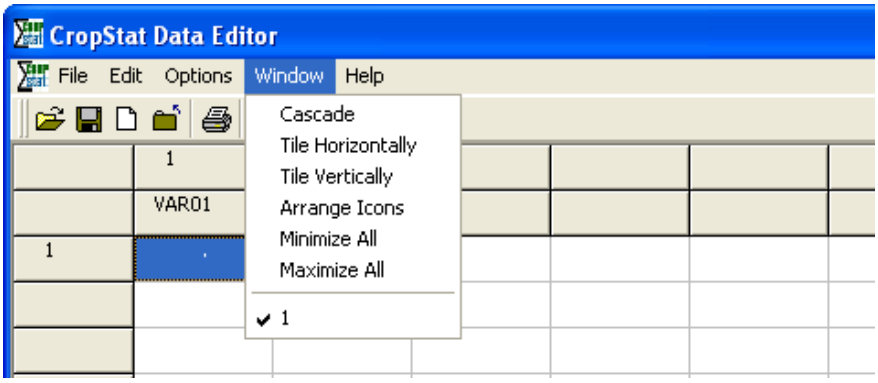
Use the **File** menu to create, save, open, reopen and print data files; to import data from Excel Workbook, Text file, and Dbase file; and to export a SYS file to excel or text file.



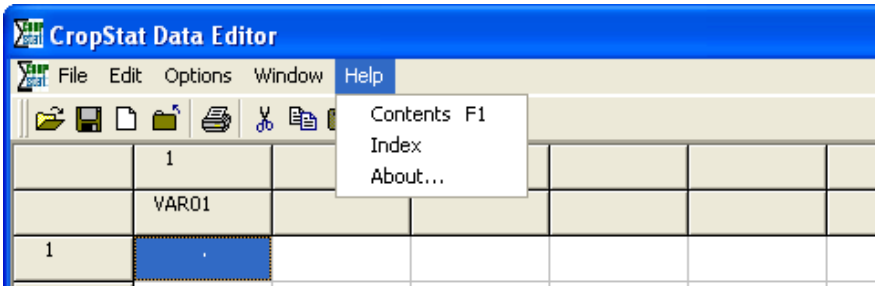
Use the **Edit** menu to cut, copy and paste data in the datasheet.



Use the **Options** menu to edit headings; transform, rank and transpose data; generate levels; insert and delete records; insert, delete, duplicate and clear variables; and set data to single precision.

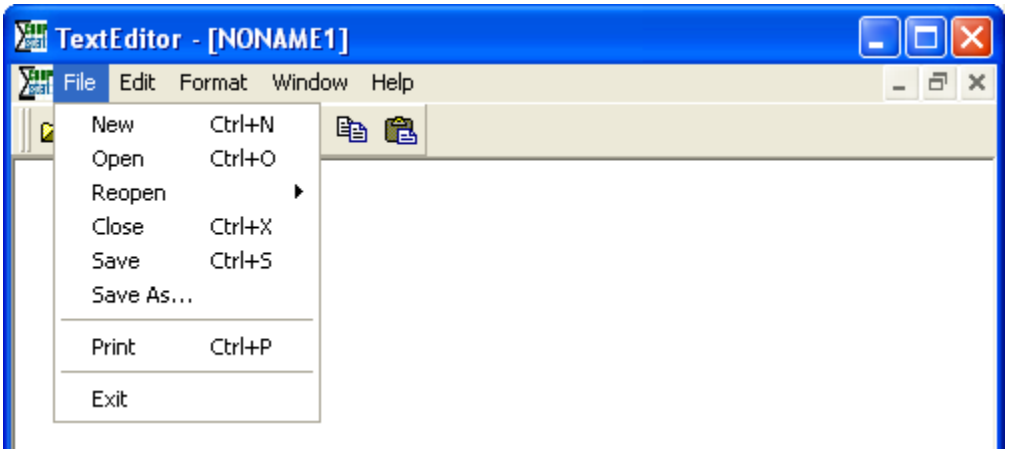


Use the **Window** menu to rearrange one or more windows.

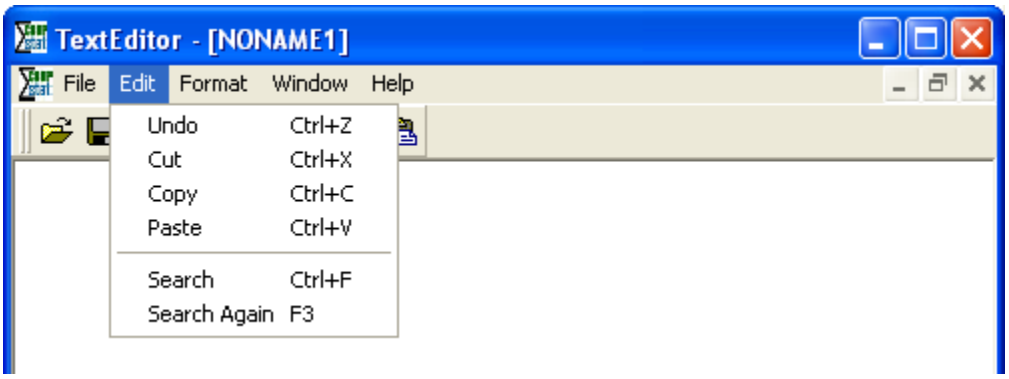


Use the **Help** menu to access CropStat's online Help System.

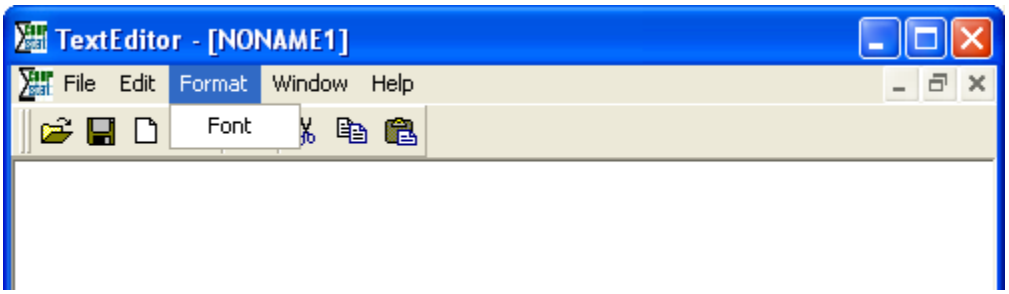
C. Text Editor Menus



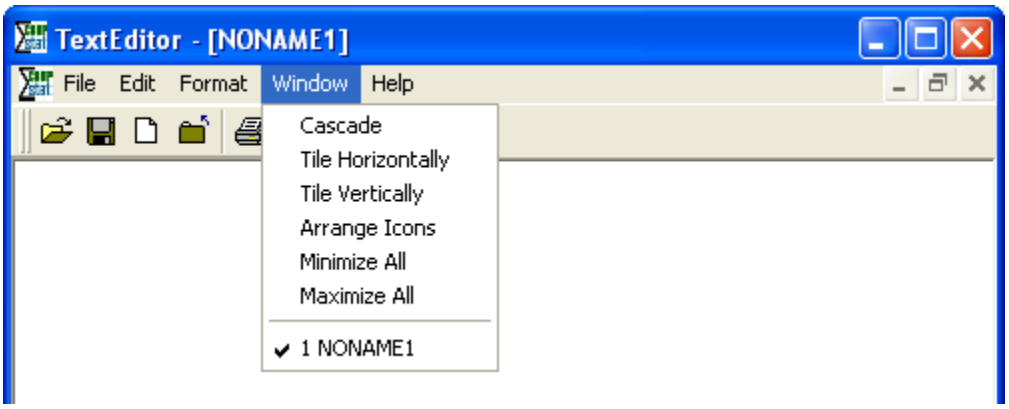
Use the **File** menu to create, open, reopen, print and save text files and to view and print command files and statistical results.



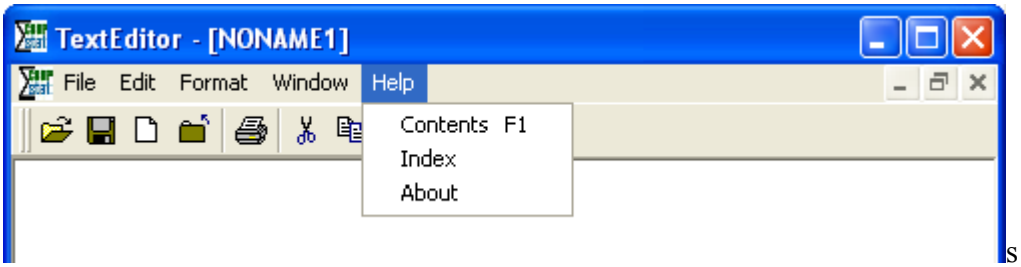
Use the **Edit** menu to cut, copy and paste texts in the text editor. It is also useful for searching a specific text in a textfile.



Use the **Format** menu to change the font characteristics of text files.



Use the **Window** menu to rearrange the display of windows.

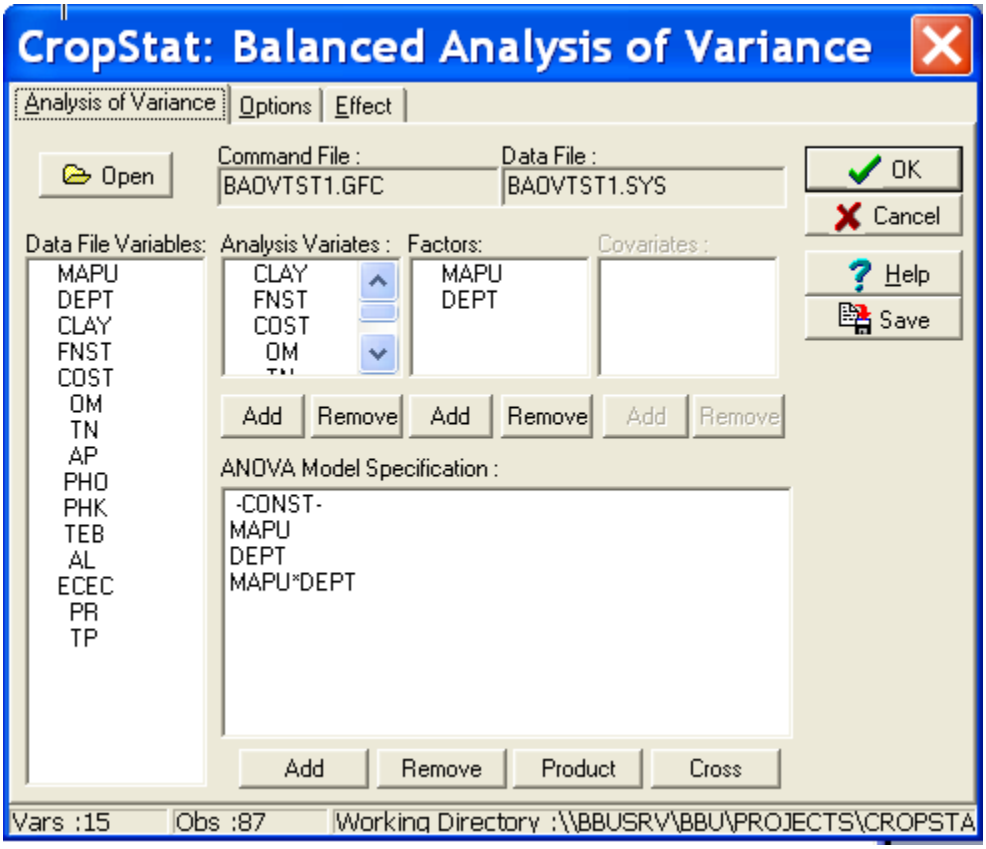


s

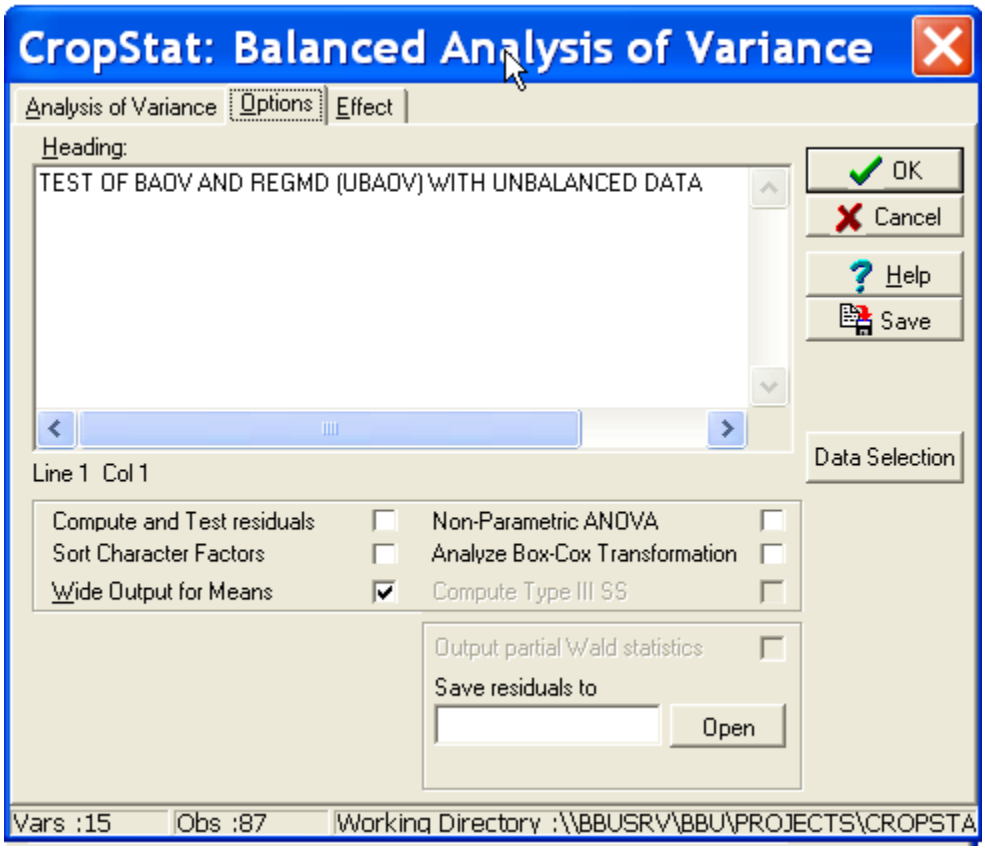
Use the **Help** menu to access CropStat's online Help system.

III. Job Specification Components

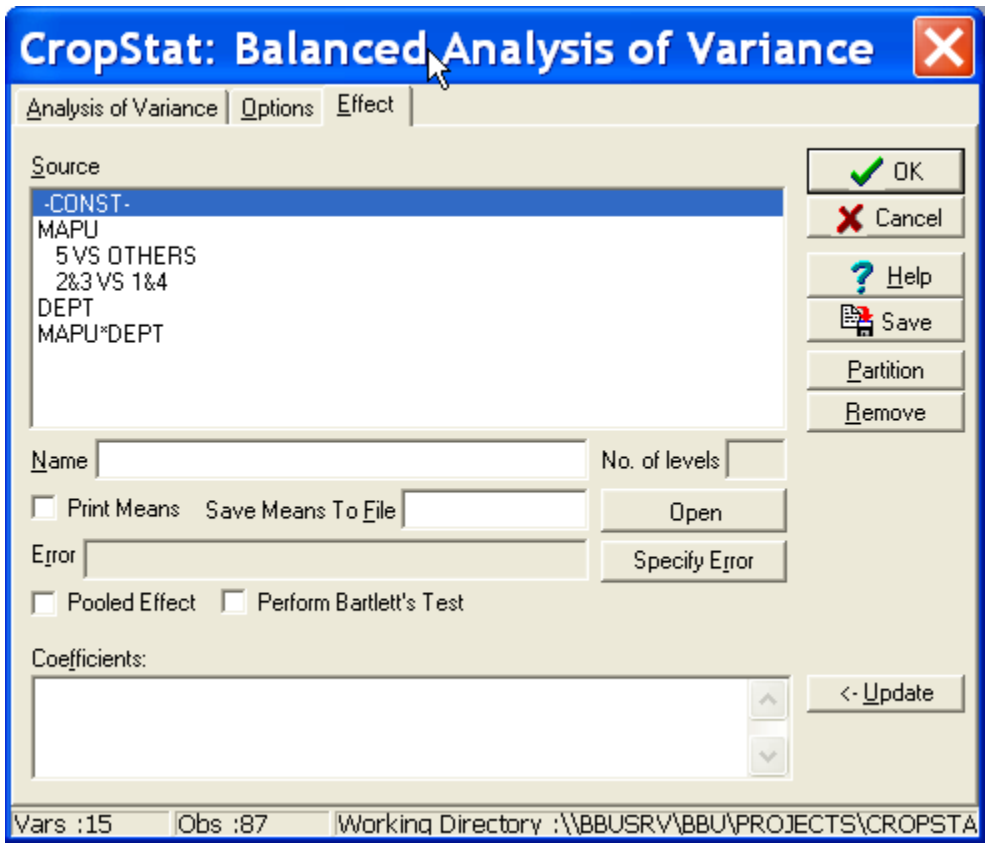
Job specification for statistical commands and graphs involves the completion of one or more pages of a dialog form like the one in the figure below, the Balanced Analysis of Variance dialog.



The Balanced Analysis of Variance dialog is composed of three pages – Analysis of Variance, Options and Effect. The Analysis of Variance Page is used to specify analysis variates, factors and model. The options page is used to invoke test of residuals, non-parametric tests and box-cox transformations. The effect page is useful for partitioning of sum of squares, error specification and testing for homogeneity of variances.



Options Page



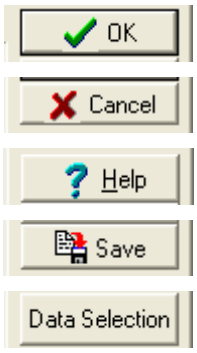
Effects Page

Each page has several components. Some components of job specification pages are only valid for particular settings of other components. Components that are not currently available are dimmed on the screen.

Examples of the different components used in CropStat follow.

1. Command Buttons

   Command buttons are graphical controls that initiate actions. Users can choose buttons by clicking the mouse while the pointer is over the button.



Click this button to perform an action.

Click this button to cancel the requested action.

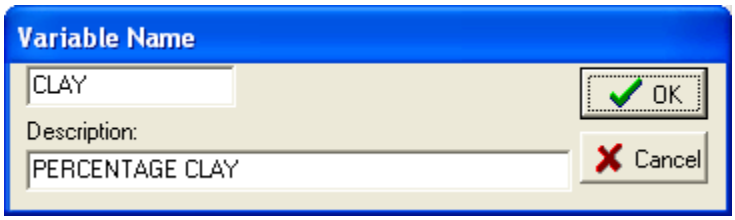
Click this button to get online help.

Click this button to save the file.

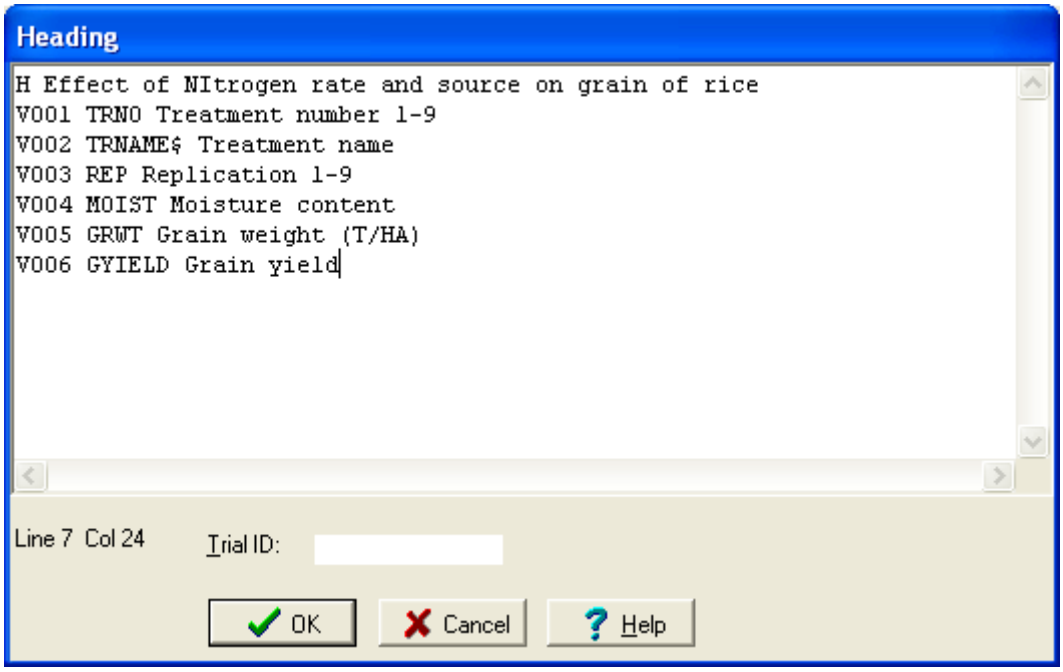
Click this button to select data subsets.

2. Text (Edit) Boxes

   Text boxes are edit controls into which the user types information.



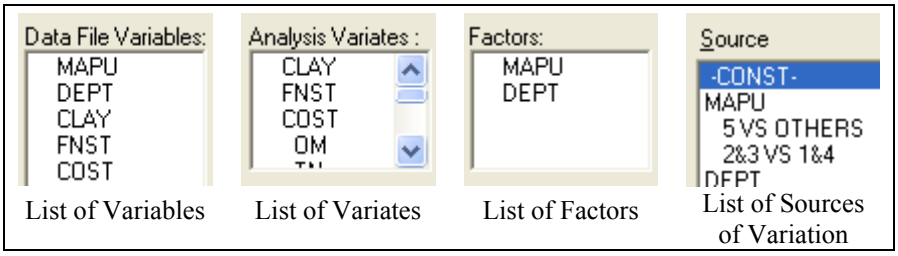
Single line textboxes



Multiple lines textbox

3. List Boxes

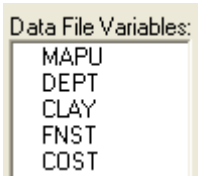
List boxes are used to display choices for the user.



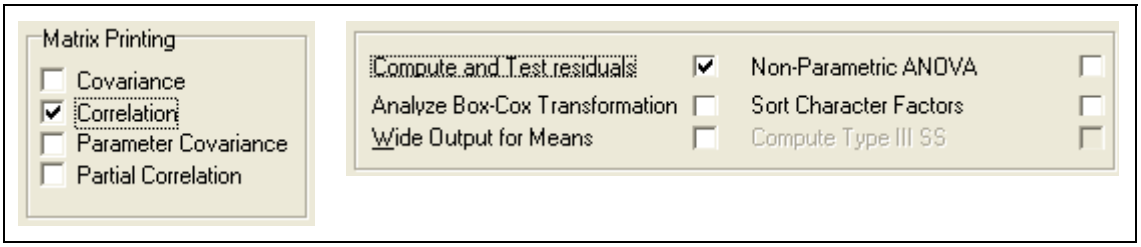
List of Variables | List of Variates | List of Factors | List of Sources of Variation

Examples of List Boxes found in the ANOVA Dialog box. For illustrative purposes the list boxes were reduced in size.

In CropStat, there are two types of list – source list and target list. The target list contains items to be included in the analysis. Input to the target list comes from the source list.

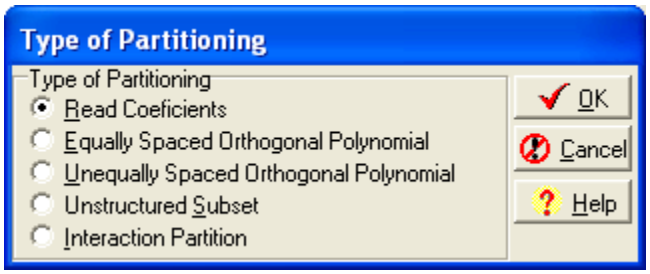
Selecting variables from source variable list to target variable list		
1.Select the variable from the data file variable list by clicking on the variable 	**2.** Click Add button under the target list (analysis variate list) to include the selected variate 	3.To remove a variable from the target list (in this case, factor list), click Remove button 

4. Check Boxes
   Check boxes control individual choices that are either turned on or off. When the choice is turned on, the check box shows a checkmark in it. When the choice is turned off, the check box is blank. The user can toggle the state of a check box by clicking on the box or the label with the mouse or by pressing the Select key (SPACEBAR) when the check box has the focus.



5. Option Buttons

An option button represents a single choice in a limited set of mutually exclusive choices. When an option button choice is selected, the circle is filled; when the choice is not selected, the circle is empty.

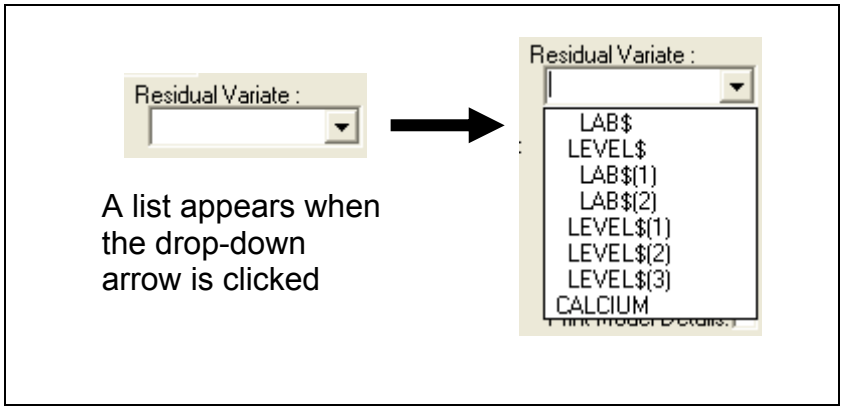


6. Spin Boxes

Spin boxes accept only a limited set of discrete, ordered input values. The user can type a new value into the text box, click the UP ARROW key to increase the value, or click the DOWN ARROW key to decrease the value.

7. Drop-down Lists

A drop-down list is a standard list that can be toggled between the closed and open state. One of the options in the list can be highlighted and moved to the text line by double clicking with the mouse.



IV. Filename Conventions

CropStat uses several files during the processing of a job. Typically these may be a binary data file or a text file of input data, one or more command files specifying particular jobs, and one or more output files.

CropStat keeps track of these different files by using fixed extensions for each file type. These fixed extensions are as follows:

1. .SYS — binary data files entered with CropStat data editor or output by ANOVA, GXE, Single-site or pattern analysis module.
2. .ASC, .TXT, .PRN — ASCII files which can be entered or modified with a text editor.
3. .DGN – Command file to specify request for randomization and layout for plant breeding trials.
4. .RND – Command file to specify request for randomization and layout for factorial experiments.
5. .DES – Describe command files to specify request for summary statistics.
6. .GFC – General command files to specify data selection and analysis requirements for balanced/unbalanced analysis of variance and mixed model analysis.
7. .GLM – Generalized Linear Model command files to specify data selection and analysis requirements for logistic regression.
8. .LLN – Log-linear command files to specify data selection and analysis requirements for log-linear analysis.
9. .REG — Regression command files to specify data selection and analysis requirements for regression.
10. .SSA — Command files to specify the design parameters and data and analysis requirements for single-site analysis.
11. .GXE — Command files for cross site, AMMI, and stability analysis.
12. .PTN — Command files for pattern analysis.
13. .QTL — Command files for quantitative trait loci analysis.
12. .OUT — Output files from a particular analysis.

Output files (.OUT) will take the filename of your Command File. If the user wishes to change the default output filename, it should be saved with your specified name from the text editor.

Command files can be reloaded and modified or reprocessed by opening them with the appropriate job specification. This allows iterative refinement of an analysis or a sequence of analysis to be performed without re-specifying the whole job.

The basic or root filename consists of up to eight alphanumeric characters and may be freely chosen by the user. However, a systematic choice of names will help keep track of jobs. For example, filenames could indicate the crop, program, and season in the first six characters leaving two for site specification. For example, WPVT04 could indicate wheat preliminary variety trials for 2004/05.

RANDOMIZATION AND LAYOUT

At the end of the tutorial, the user should be able to generate randomized field plans for single- and multi-factor designs.

I. Factorial Treatments in Randomized Complete Blocks

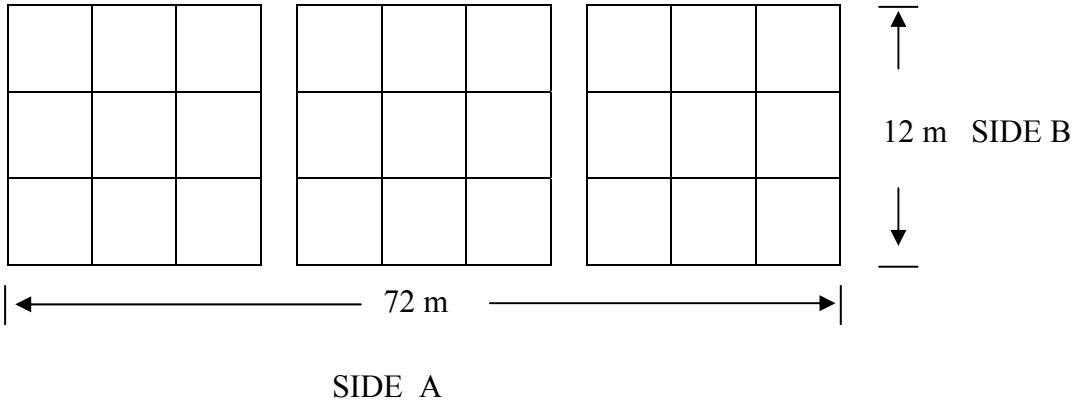
Sample Problem

Generate an experimental layout for the following trial:

Title: Effect of Nitrogen Rate and Source on Grain Yield of Rice

The experiment has 3 nitrogen rates of 3 nitrogen sources to give 9 fertilizer treatments with 3 replications in a randomized complete block design.

Division of experimental plots is as follows:



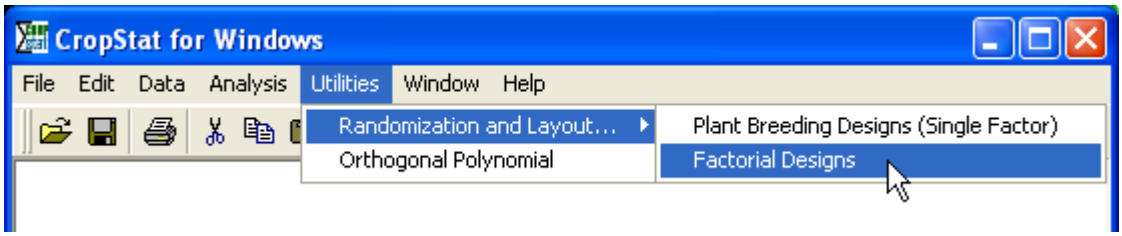
Plot size: 8 m × 4 m

Steps

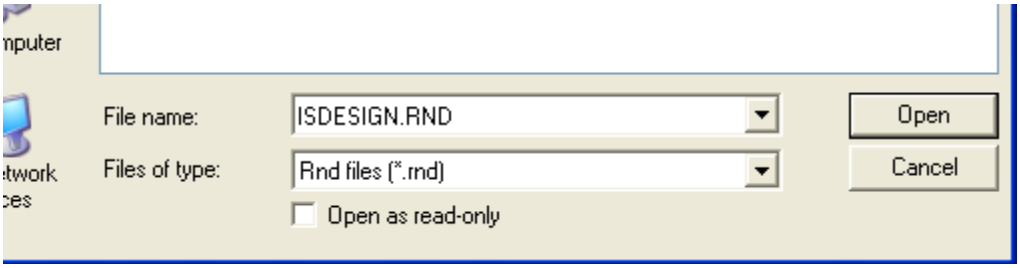
- Create your own working directory C:\MY CROPSTAT\ that will be used for the entire tutorial in saving the data file, command file, and output file.

- Inside your own working directory C:\MY CROPSTAT, create a subfolder RANDMOZATION AND LAYOUT.

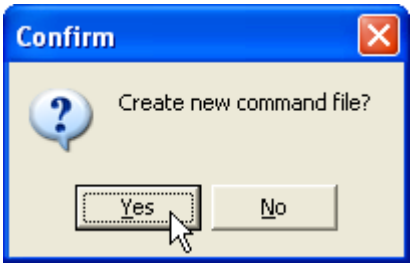
- Click **Utilities ⇒ Randomization and Layout ⇒ Factorial Designs** from the Main Window.



- The **Open a Randomization and Layout File** dialog box will appear. Click the **Look In** box and browse to the C:*MY CROPSTAT\RANDOMIZATION AND LAYOUT*. Inside created folder, type *ISDESIGN* in the **File name** edit box and click **Open**.



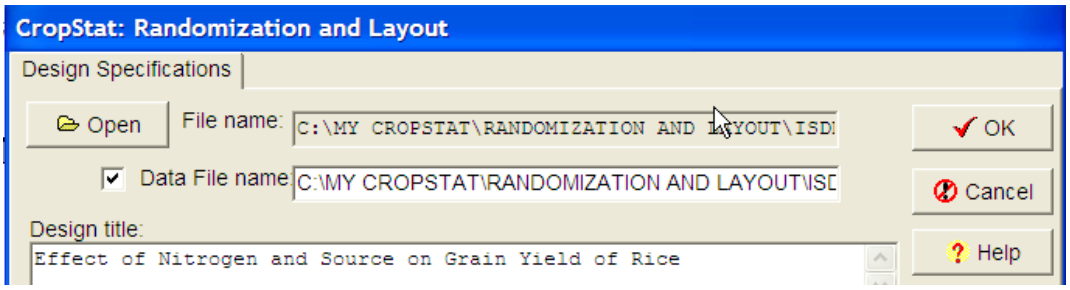
- Click **Yes** in the following message box to confirm creation of a new command file *ISDESIGN.RND*.



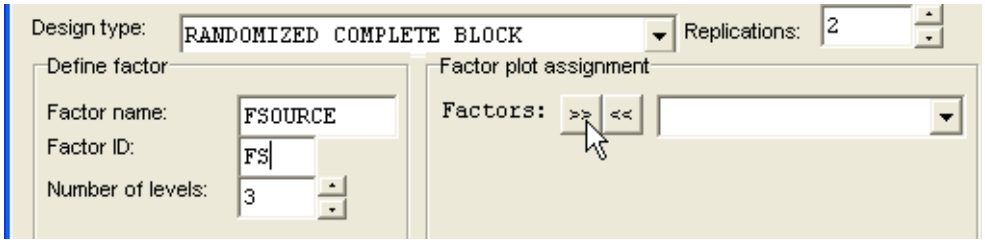
- The **Randomization and Layout** dialog box will appear. In the **Design Title** box, type the following text: *Effect of Nitrogen Rate and Source on Grain Yield of Rice*

- Click the **Design type** drop-down list. Select **Randomized Complete Block**.

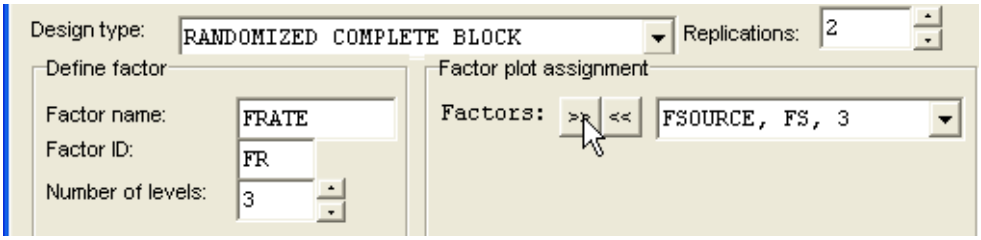
- Click the **Replications** spinbox. Set it to *2*.



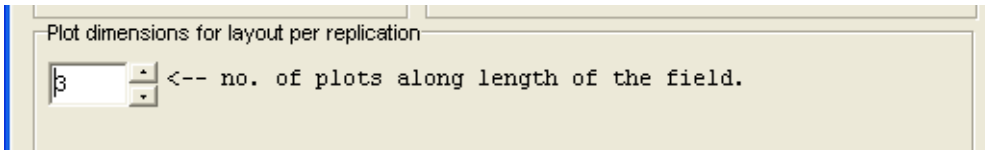
- To specify a treatment factor, click the **Factor name** edit box in the **Define factor** group. Type *FSOURCE*. Factor names can have a maximum of 8 characters. In the **Factor ID** box, type *FS*. Set the **Number of levels** spin box to 3. Click **>>** to add the new factor to the **Factor plot assignment** group.



- Repeat this procedure for *FRATE* with ID *FR* and also having three levels.



- Enter *3* in the **No. of plots along length of the field** spin box in the **Plot dimensions for layout per replication** portion**.** Click **OK** to generate the design.



Sample Output

The following output will appear in the Text Editor. It is also saved as a temporary text file with a default file name RNDLYT.OUT. You may save it to another file to have a permanent copy of the output by selecting **FILE ⇒ SAVE AS** from the Text Editor.

1. Design and factor specifications

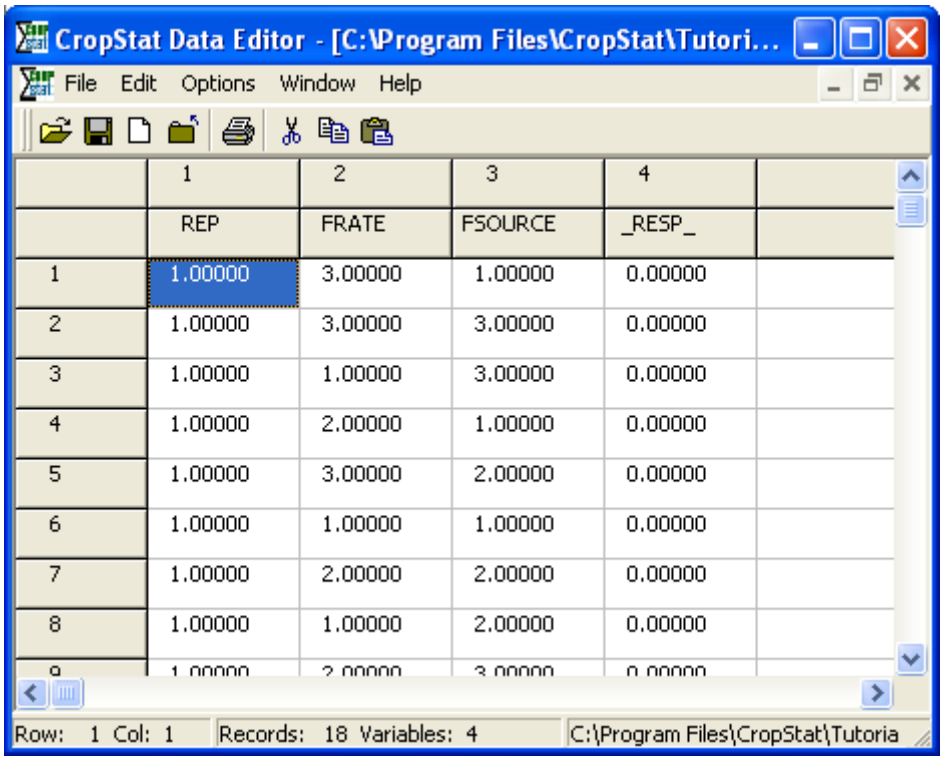
```
RANDOMIZATION AND LAYOUT
========================
FILENAME = "C:\MY CROPSTAT\RANDOMIZATION AND LAYOUT\ISDESIGN.RND"
TITLE = "Effect of Nitrogen Rate and Source on Grain Yield of RIce"
EXPERIMENTAL DESIGN = RANDOMIZED COMPLETE BLOCK
REPLICATIONS = 2
TREATMENTS = 3 x 3
 **** FACTOR(S) ****
  FRATE (FR) = 3 levels
   FRATE (1) = FR1
   FRATE (2) = FR2
   FRATE (3) = FR3
  FSOURCE (FS) = 3 levels
   FSOURCE (1) = FS1
   FSOURCE (2) = FS2
   FSOURCE (3) = FS3
```

2. Experimental layout

```
Experimental layout for file: "C:\MY CROPSTAT\RANDOMIZATION AND LAYOUT\ISDESIGN.RND"
(RANDOMIZED COMPLETE BLOCK)
The following field layout applies to all replications:
(Note: layout is not drawn to scale)
+-----+-----+-----+
|   1 |   2 |   3 |
+-----+-----+-----+
|   4 |   5 |   6 |
+-----+-----+-----+
|   7 |   8 |   9 |
+-----+-----+-----+

REPLICATION NO. 1
-----------------
PLOT NO.  |  TREATMENT ID
       1  |  FR1 FS3
       2  |  FR2 FS3
       3  |  FR1 FS2
       4  |  FR2 FS2
       5  |  FR2 FS1
       6  |  FR3 FS2
       7  |  FR3 FS1
       8  |  FR3 FS3
       9  |  FR1 FS1

REPLICATION NO. 2
-----------------
PLOT NO.  |  TREATMENT ID
       1  |  FR3 FS3
       2  |  FR1 FS1
       3  |  FR3 FS2
       4  |  FR1 FS2
       5  |  FR3 FS1
       6  |  FR1 FS3
       7  |  FR2 FS1
       8  |  FR2 FS3
       9  |  FR2 FS2
```

3. Electronic Field Book

Select **Data Editor** from the **Window** menu. Click **File | Open** and select *ISDESIGN.SYS* in the *C:\MY CROPSTAT\ RANDOMIZATION AND LAYOUT* folder to display the Electronic field book. Data can be entered into this file. Extra response variables can also be added.



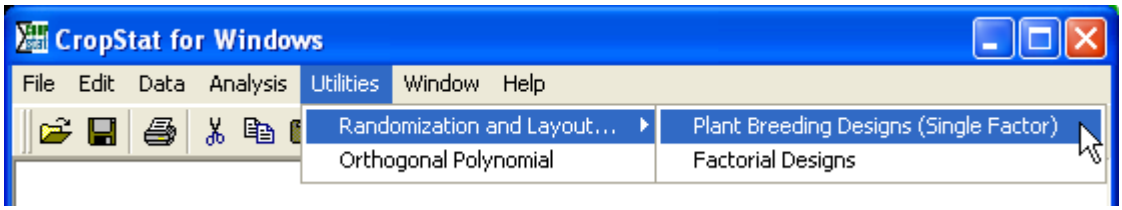
	1	2	3	4	
	REP	FRATE	FSOURCE	_RESP_	
1	1.00000	3.00000	1.00000	0.00000	
2	1.00000	3.00000	3.00000	0.00000	
3	1.00000	1.00000	3.00000	0.00000	
4	1.00000	2.00000	1.00000	0.00000	
5	1.00000	3.00000	2.00000	0.00000	
6	1.00000	1.00000	1.00000	0.00000	
7	1.00000	2.00000	2.00000	0.00000	
8	1.00000	1.00000	2.00000	0.00000	

II. Augmented Latin Square Design

Sample Problem

Generate a randomized layout and field book for an Augmented Latin Square Design with 100 test entries and 5 replicated checks. Field space allows five rows of twenty-five plots each.

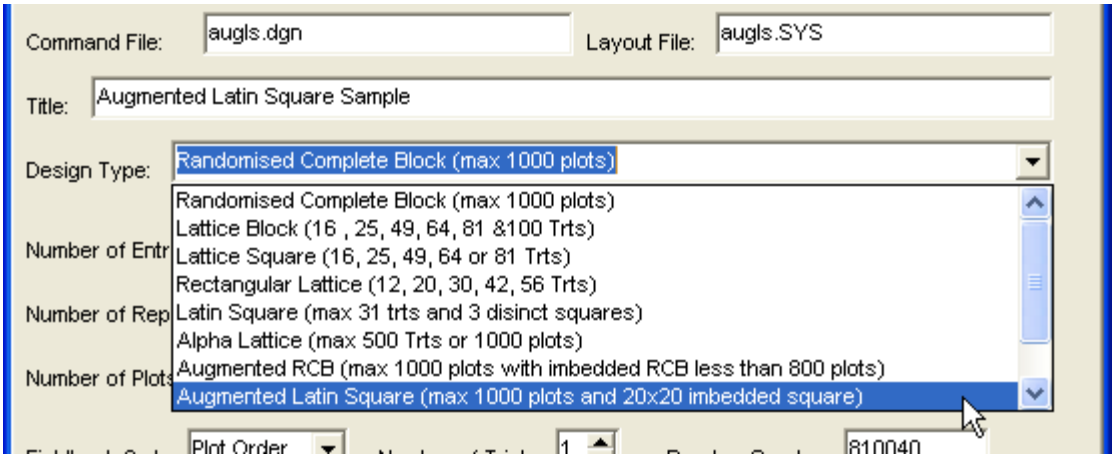
- Click **Utilities ⇒ Randomization and Layout ⇒ Plant Breeding Designs** from the Main Window.



- Open a new PB Design Command File by navigating to the *C:\MY CROPSTAT\ RANDOMIZATION AND LAYOUT* directory and entering a new file name **AUGLS** in the Open Dialogue Box. Click **Open** and then **YES** to create a new command file.

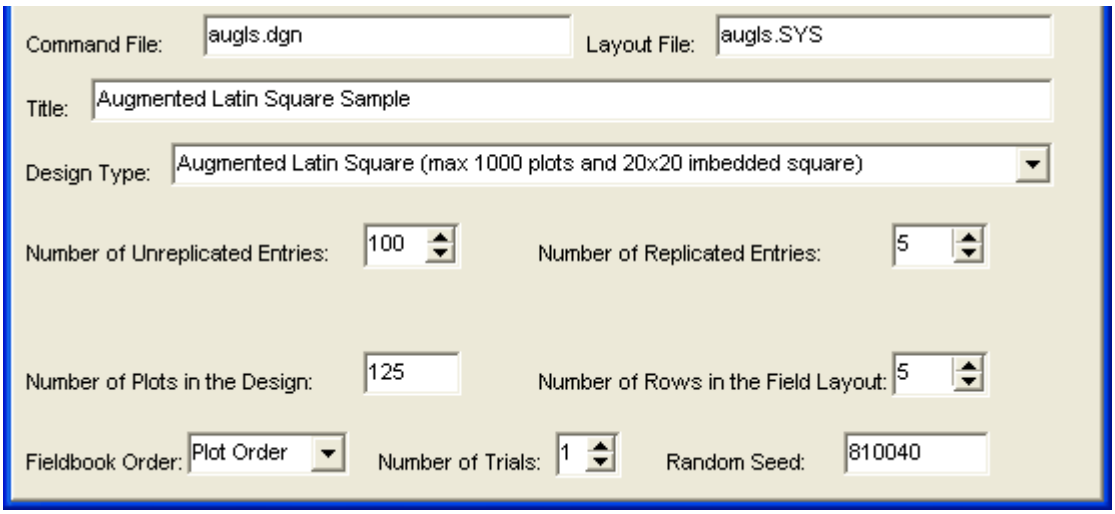
- The command file name is shown and the same name is used as the default for the field book with extension SYS. This could be changed, but for this exercise leave the default. Enter **Augmented Latin Square Sample** in the Title field.

- Select **Augmented Latin Square** from the pull down list of design types.



- Enter **100** for Number of Unreplicated Entries, **5** for Number of Replicated entries and **5** for Number of Rows in the field layout.

- The **Field Book Order** box specifies the order of rows in the field book – plot order or serpentine. Leave the default – plot order. The **Number of Trials** box allows the generation of several trials with the same design but different randomizations. Leave the default 1. The **Random Seed** box allows specification of a previously used random number seed in case it is necessary to re-generate an old randomization. Leave the randomly selected default seed.



- Click **OK** to generate the design and field book.

Sample Output

The text output contains a Treatment Key, indicating the PLOT(s) where each entry (TRT) has been allocated, and a Field plan showing the layout of plot numbers and entry numbers. The replicated entries are the negative entries, -1 to -5.

```
FIELD PLAN FOR AUGMENTED LATIN SQUARE DESIGN 100 UNREPLICATED ENTRIES,   5 REPLICATED
ENTRIES        FILE AUGLS.DGN      2/23/2007  9:56:12 AM
------------------------------------------------------------------------------------
FIELD LAYOUT IS 100 PLOTS
TRIAL DESCRIPTION: Augmented Latin Square Sample

TREATMENTS
TRIAL NUMBER: 1


(Replicated: Negative Entries, Unreplicated: Positive Entries)
PLOT    TRT  AROW  ACOL  LROW  LCOL
   1      1     1     1     1     1
   2     78     1     2     1     1
   3     85     1     3     1     1
   4     -4     1     4     1     1
   5     35     1     5     1     1
   6      7     1     6     1     2
   7     34     1     7     1     2
   8     45     1     8     1     2
   9     37     1     9     1     2
  10     -1     1    10     1     2
  11     40     1    11     1     3
  12     19     1    12     1     3
  13     60     1    13     1     3
  14     53     1    14     1     3
  15     -3     1    15     1     3
  16     -5     1    16     1     4
  17      9     1    17     1     4
  18     42     1    18     1     4
  19     80     1    19     1     4
  20     73     1    20     1     4
   .
   .
   .
 107     -2     5     7     5     2
 108     61     5     8     5     2
 109     69     5     9     5     2
 110     31     5    10     5     2
 111     16     5    11     5     3
 112     -5     5    12     5     3
 113     65     5    13     5     3
 114     12     5    14     5     3
 115     76     5    15     5     3
 116     59     5    16     5     4
 117     99     5    17     5     4
 118     81     5    18     5     4
 119     -4     5    19     5     4
 120     13     5    20     5     4
 121     58     5    21     5     5
 122     -3     5    22     5     5
 123     74     5    23     5     5
 124     43     5    24     5     5
 125     21     5    25     5     5
```

FIELD PLAN FOR AUGMENTED LATIN SQUARE DESIGN 100 UNREPLICATED ENTRIES, 5 REPLICATED ENTRIES FILE AUGLS.DGN 2/23/2007 9:56:12 AM

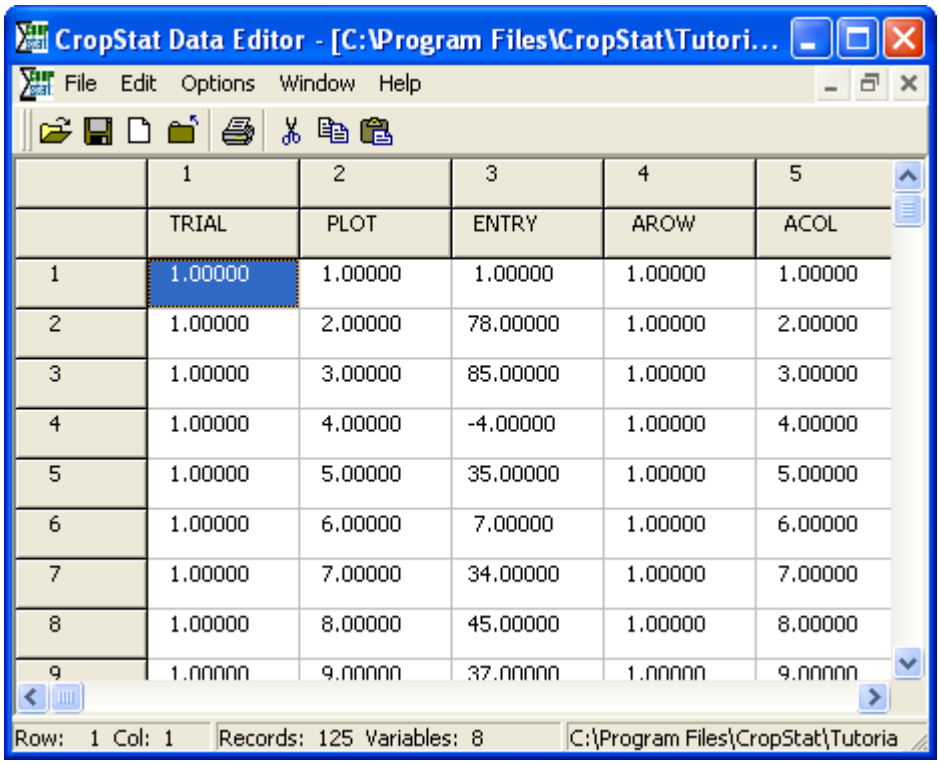
TRIAL DESCRIPTION: Augmented Latin Square Sample RANDOM SEED 810040
TRIAL NUMBER: 1

(CELLS CONTAIN PLOT NUMBERS ON TOP, TREATMENTS/ENTRIES BELOW)
(Replicated: Negative Entries, Unreplicated: Positive Entries)

COL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
ROW																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	1	78	85	-4	35	7	34	45	37	-1	40	19	60	53	-3	-5	9	42	80	73	24	91	93	-2	84
	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
2	26	86	11	48	-2	47	-3	71	82	39	57	72	32	4	-4	63	-1	55	23	56	68	-5	54	100	6
	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
3	2	75	51	-3	52	-5	77	25	14	28	18	29	83	-1	38	98	-2	50	10	67	-4	89	44	5	8
	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
4	70	96	88	-5	33	-4	95	41	17	3	49	15	94	-2	92	22	90	20	66	-3	-1	79	97	64	62
	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125
5	87	46	-1	36	27	30	-2	61	69	31	16	-5	65	12	76	59	99	81	-4	13	58	-3	74	43	21

Sample Field Book

To view the field book, select Window from the main menu, then Data Editor. Select File – Open and select *AUGLS.SYS*. The Field Book file is displayed. TRIAL has a trial number – always 1 since we only asked for one trial. PLOT contains the plot numbers 1-125 in plot order since we did not ask for Serpentine order. NREP indicates replicated checks with value 1 and unreplicated test entries with value 0. ENTRY gives the randomized entry numbers. The last entries, 101 to 105 are the replicated checks. ROW and COL contain the plot coordinates.



	1	2	3	4	5
	TRIAL	PLOT	ENTRY	AROW	ACOL
1	1.00000	1.00000	1.00000	1.00000	1.00000
2	1.00000	2.00000	78.00000	1.00000	2.00000
3	1.00000	3.00000	85.00000	1.00000	3.00000
4	1.00000	4.00000	-4.00000	1.00000	4.00000
5	1.00000	5.00000	35.00000	1.00000	5.00000
6	1.00000	6.00000	7.00000	1.00000	6.00000
7	1.00000	7.00000	34.00000	1.00000	7.00000
8	1.00000	8.00000	45.00000	1.00000	8.00000
9	1.00000	9.00000	37.00000	1.00000	9.00000

III. Simple Balanced Lattice Design

Sample Problem

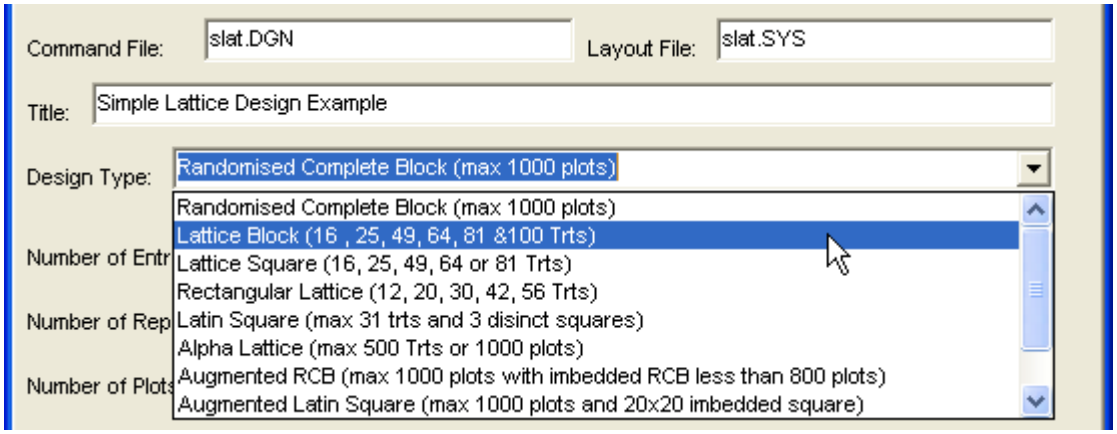
Generate randomized field plans and a field book for three trials with a Simple Lattice Design having 49 entries. (A Simple Lattice is a Lattice Block Design with two replications.)

- Click **Utilities ⇒ Randomization and Layout ⇒ Plant Breeding Designs** from the Main Window.

- Open a new PB Design Command File by navigating to the *C:\MY CROPSTAT\ RANDOMIZATION AND LAYOUT* directory and entering a new file name *SLAT* in the Open Dialogue Box. Click **Open** and then **YES** to create a new command file.

- The command file name is shown and the same name is used as the default for the field book with extension SYS. This could be changed, but for this exercise leave the default. Enter **Simple Lattice Design Example** in the Title field.

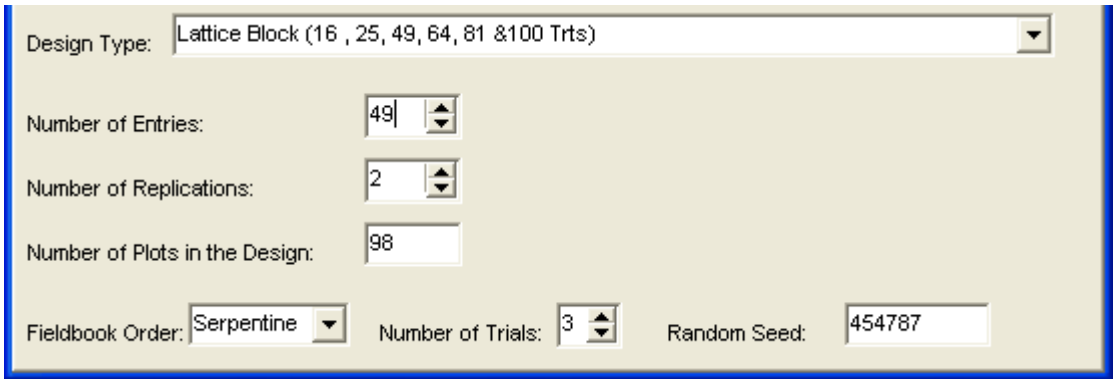
- Select **Lattice Block** from the pull down list of design types.



- Enter **49** for Number of Entries and **2** for Number of Replications.

- Select Serpentine for the Field Book Order. Enter 3 for Number of Trials. The Random Seed box allows specification of a previously used random number seed in case it is necessary to re-generate an old randomization. Leave the randomly selected default seed.

- Click OK to generate the design and field book.



Sample Output

The text output contains Treatment Keys, indicating the PLOT in each REP where each entry (TRT) has been allocated, and Field Plan showing plot numbers and associated entry numbers for each trial. Part of the first trial is shown below.

```
FILED PLAN FOR LATTICE DESIGN, 49 TRTS,  2 REPS  FILE C:\ MY CROPSTAT\ RANDOMIZATION
AND LAYOUT\SLAT.DGN   29/ 9/ 4 15:39
 ------------------------------------------------------------------------- :PAGE   1
 TRIAL DESCRIPTION: SIMPLE LATTICE DESIGN EXAMPLE
 TRIAL NUMBER:  1                    SITE.................        RANDOM SEED  202101


 REP   1 (CELLS CONTAIN PLOT NUMBERS ON TOP, TREATMENTS BELOW)
 PLT   1|  2|  3|  4|  5|  6|  7|
 BLK ---|---|---|---|---|---|---|
     |  1|  2|  3|  4|  5|  6|  7|
   1| 32| 23| 38| 26| 10| 34| 27|
     |---|---|---|---|---|---|---|
     |  8|  9| 10| 11| 12| 13| 14|
   2| 29| 30| 45| 28| 36| 39| 33|
     |---|---|---|---|---|---|---|
     .
     .
     .
     |---|---|---|---|---|---|---|
     | 43| 44| 45| 46| 47| 48| 49|
   7| 21| 13| 40|  4|  5| 25| 41|
     |---|---|---|---|---|---|---|

 REP   2 (CELLS CONTAIN PLOT NUMBERS ON TOP, TREATMENTS BELOW)
 PLT   1|  2|  3|  4|  5|  6|  7|
 BLK ---|---|---|---|---|---|---|
     | 50| 51| 52| 53| 54| 55| 56|
   1| 31|  4| 43| 34| 46| 28| 15|
     |---|---|---|---|---|---|---|
     | 57| 58| 59| 60| 61| 62| 63|
   2| 19| 29| 48| 13| 23| 11| 24|
     |---|---|---|---|---|---|---|
     .
     .
     .
     |---|---|---|---|---|---|---|
     | 92| 93| 94| 95| 96| 97| 98|
   7| 39| 27| 37| 16| 25| 35| 42|
     |---|---|---|---|---|---|---|
```

```
TREATMENT KEY FOR LATTICE DESIGN, 49 TRTS,  2 REPS  FILE C:\MY CROPSTAT\
RANDOMIZATION AND LAYOUT\SLAT.DGN   29/ 9/ 4 15:39
 ------------------------------------------------------------------------- :PAGE   2
 TRIAL DESCRIPTION: SIMPLE LATTICE DESIGN EXAMPLE
 TRIAL NUMBER:  1                    SITE.................        RANDOM SEED  202101


                               REP 1   REP 2
  TRT                          PLOT    PLOT
    1 ......................     39      84
    2 ......................     34      68
    3 ......................     18      69
    .
    .
    .
   48 ......................     33      59
   49 ......................     26      70
```

IV. Alpha Designs

Peterson and Williams devised a new class of incomplete block designs called alpha designs. These designs are in some respect a generalization of Yates' original lattice designs. The main advantage of alpha designs is flexibility; they are available whenever the number of varieties ***v*** is a multiple of block size ***k***, and they can be easily adapted even when it is not.

Alpha designs fall under the category of resolvable designs. An incomplete block design is resolvable when the incomplete blocks can be arranged in complete replication. Non-resolvable designs exist but are less valuable for field trials.

Within each replicate of an alpha design there are ***s*** blocks each of size ***k*** plots. There is thus a two-stage structure for the control of field variation. Firstly, replicates can allow an adjustment for large-scale variation and then within each replicate the blocks provide a second level of adjustment.

Efficiency Factor of an Alpha Design

For a particular set of ***r***, ***s***, and ***k***, many alpha designs are possible, so we must choose the most efficient alpha design. There are many ways of measuring the quality of incomplete block designs but the measure of most relevance where the interest is in comparing all pairs of varieties is the efficiency factor **E**. It is defined as the ratio of $2\sigma^2$ to V, where V is the average pairwise variance or mean variance of all differences between two varieties and σ^2 this error variance in the within-block analysis of the incomplete block design. When choosing an alpha design, the aim is to choose one that has an efficiency factor **E** as large as possible. Designs with maximum **E** among all alpha designs are called alpha optimal.

The number of times a pair of varieties appears together within block of the design is called the concurrence of that pair of varieties. When the concurrence of pairs of varieties is high the pairwise variance is low, but if pairs of varieties do not appear together within blocks at all, the pairwise variance is high. The average pairwise variance will be minimized and **E** maximized if all the concurrence of pairs of varieties are as similar as possible. If the concurrences are either zero or one the design is known as an alpha (0,1)-design.

When the number of plots per block ***k*** is greater than the number of blocks per replicate ***s***, it is not possible to construct alpha (0,1)-designs, the next best thing is to restrict the number of concurrences to two, leading to alpha (0,1,2)-designs.

Availability of Alpha Designs

An advantage of alpha designs is that they are available for a wide range of combinations of parameters. There is no need to restrict resolvable incomplete block designs only to the square and rectangular lattice designs. Efficient designs can be constructed for all combinations of ***r***, ***s***, and ***k*** that would be required in practice.

For some number of varieties there may be a choice of block sizes, for example an alpha design for ***v***=48 could have block sizes of 2, 4, 6, 8, 12, 16 or 24. As a general rule block size should be roughly equal to the square root of the number of varieties. So in the example above for ***v***=48, we would consider designs with ***k***=6 or 8. Often site considerations will dictate the block size; physical constraints on the layout of plots in the field might mean that blocks of six plots are much more natural than blocks of eight plots.

An apparent restriction on the availability of alpha designs is that the number of varieties is the product of the number of blocks per replicate and the block size, i.e. ***v***=***sk***. So if we wanted a resolvable incomplete block design for 17 varieties it would appear difficult. A way around this problem is to allow designs with unequal block sizes. Provided the block sizes differ by no more than one, there is no need to review the model assumption that the pairwise variance between any two points within a block is the same. Alpha designs make it possible to produce efficient designs with block sizes differing by no more than one. This is done by simply deleting variety numbers from an alpha design for a larger number of varieties. So, in the previous example we could delete numbers 18, 19 and 20 to obtain a design for 17 varieties in blocks of size three or four. The largest variety numbers should be deleted before randomization to guarantee that block sizes differ by no more than one.

Sample Problem

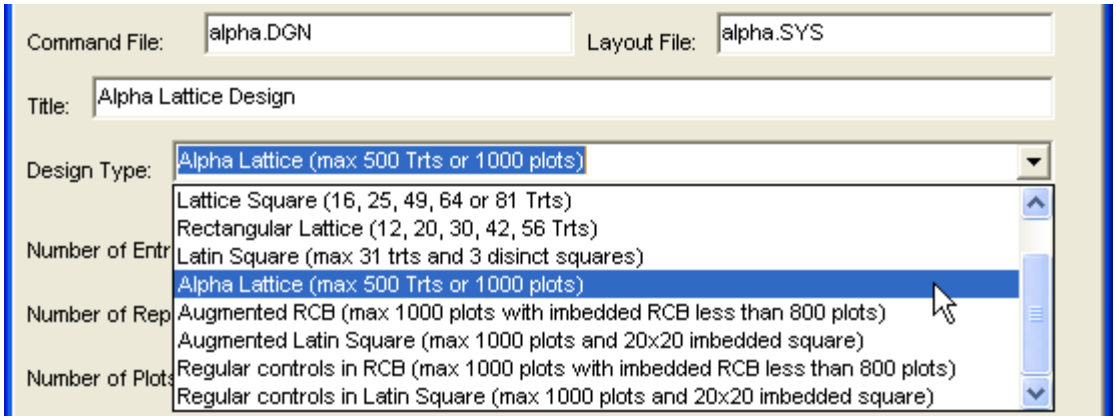
Generate randomized field plans for one trial with an Alpha Lattice Design having seventy-two (72) entries grouped into eight with three replicates.

- Click **Utilities ⇒ Randomization and Layout ⇒ Plant Breeding Designs** from the Main Window.

- Open a new PB Design Command File by navigating to the *C:\MY CROPSTAT\ RANDOMIZATION AND LAYOUT* directory and entering a new file name *ALPHA* in the Open Dialogue Box. Click **Open** and then click **Yes** to create a new command file.

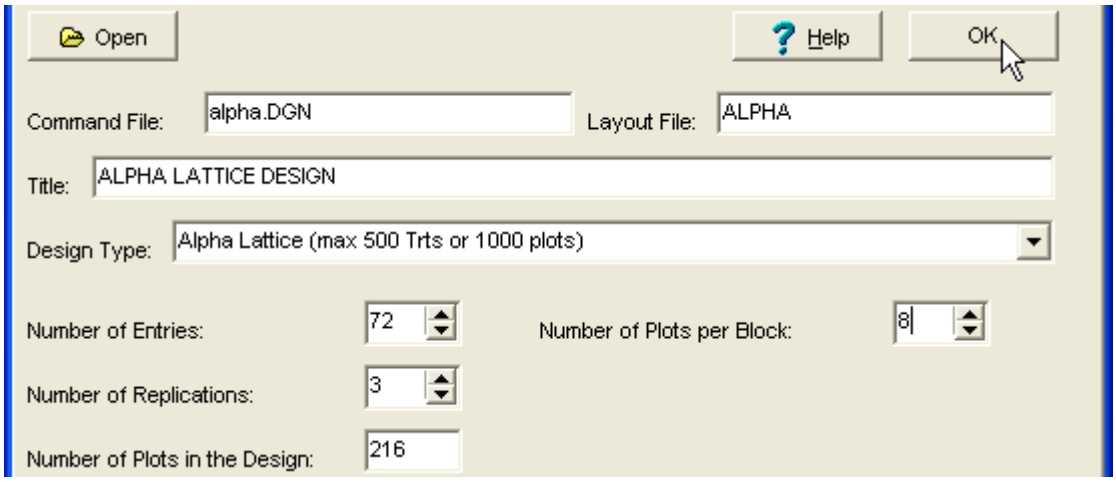
- In the Plant Breeding Dialogue box, enter **Alpha Lattice Design** in the Title field.

From the pull down list of design types select **Alpha Lattice**.



- In the Number of Entries type **72**. Enter **3** for the Number of Replications and **8** for the Number of Plots per Block.

- Click **OK** to generate the design.



Sample Output

```
FIELD PLAN FOR ALPHA LATTICE DESIGN, 72 TRTS,  3 REPS  FILE
S:\Projects\TEST\pbdgn\alpha.DGN    7/ 2/ 7 11:17
 ------------------------------------------------------------------------- :PAGE   1
 TRIAL DESCRIPTION: ALPHA LATTICE DESIGN
 TRIAL NUMBER:  1            SITE.....................        RANDOM SEED  500278

 FIELD PLAN FOR  3 REPS OF  72 TREATMENTS IN  9 BLOCKS/REP OF  8 PLOTS/BLOCK
 EFFICIENCY:  0.8527(100.0%), MAXIMUM CONCURANCE:  1

 REP   1 (CELLS CONTAIN PLOT NUMBERS ON TOP, TREATMENTS BELOW)
 PLT   1|  2|  3|  4|  5|  6|  7|  8|
 BLK ---|---|---|---|---|---|---|---|
    |  1|  2|  3|  4|  5|  6|  7|  8|
   1| 23| 19| 65|  5| 66|  2| 24| 55|
    |---|---|---|---|---|---|---|---|
    |  9| 10| 11| 12| 13| 14| 15| 16|
   2| 44| 68| 71| 18| 26| 72| 37| 21|
    |---|---|---|---|---|---|---|---|
    | 17| 18| 19| 20| 21| 22| 23| 24|
   3| 31| 22| 42| 11| 34|  8| 16|  6|
    |---|---|---|---|---|---|---|---|
    | 25| 26| 27| 28| 29| 30| 31| 32|
   4| 15| 62| 40| 30| 60| 57| 20|  9|
    |---|---|---|---|---|---|---|---|
    | 33| 34| 35| 36| 37| 38| 39| 40|
   5| 53| 50| 64| 28| 67| 46| 69| 52|
    |---|---|---|---|---|---|---|---|
    | 41| 42| 43| 44| 45| 46| 47| 48|
   6| 47| 29| 43| 32| 45| 63| 59| 54|
    |---|---|---|---|---|---|---|---|
    | 49| 50| 51| 52| 53| 54| 55| 56|
   7|  4| 17| 12|  1| 61| 39| 14| 13|
    |---|---|---|---|---|---|---|---|
    | 57| 58| 59| 60| 61| 62| 63| 64|
   8| 58| 48| 38| 41| 25| 56| 70| 33|
    |---|---|---|---|---|---|---|---|
    | 65| 66| 67| 68| 69| 70| 71| 72|
   9| 36| 51| 49| 27| 35|  3|  7| 10|
    |---|---|---|---|---|---|---|---|
```

```
REP   2 (CELLS CONTAIN PLOT NUMBERS ON TOP, TREATMENTS BELOW)
 PLT   1|  2|  3|  4|  5|  6|  7|  8|
 BLK ---|---|---|---|---|---|---|---|
     | 73| 74| 75| 76| 77| 78| 79| 80|
   1| 59| 71| 22| 35|  5| 50|  9| 12|
     |---|---|---|---|---|---|---|---|
     | 81| 82| 83| 84| 85| 86| 87| 88|
   2| 65| 67| 14| 32| 60| 25| 42| 21|
     |---|---|---|---|---|---|---|---|
     | 89| 90| 91| 92| 93| 94| 95| 96|
   3| 58| 47| 28|  1| 26| 55| 16| 49|
     |---|---|---|---|---|---|---|---|
     | 97| 98| 99|100|101|102|103|104|
   4|  8| 66| 41| 57|  3| 39| 64| 29|
     |---|---|---|---|---|---|---|---|
     |105|106|107|108|109|110|111|112|
   5| 10| 46| 30| 31| 19|  4| 38| 72|
     |---|---|---|---|---|---|---|---|
     |113|114|115|116|117|118|119|120|
   6| 45| 40| 11| 69| 27| 56| 17| 18|
     |---|---|---|---|---|---|---|---|
     |121|122|123|124|125|126|127|128|
   7| 37|  7| 20| 70| 23| 54| 13|  6|
     |---|---|---|---|---|---|---|---|
     |129|130|131|132|133|134|135|136|
   8| 62| 68| 48| 52| 34| 51| 24| 63|
     |---|---|---|---|---|---|---|---|
     |137|138|139|140|141|142|143|144|
   9|  2| 61| 33| 53| 36| 43| 44| 15|
     |---|---|---|---|---|---|---|---|

 REP   3 (CELLS CONTAIN PLOT NUMBERS ON TOP, TREATMENTS BELOW)
 PLT   1|  2|  3|  4|  5|  6|  7|  8|
 BLK ---|---|---|---|---|---|---|---|
     |145|146|147|148|149|150|151|152|
   1|  1| 35| 20| 63| 33| 19| 11| 67|
     |---|---|---|---|---|---|---|---|
     |153|154|155|156|157|158|159|160|
   2| 13| 56| 53| 68| 66| 49| 42|  9|
     |---|---|---|---|---|---|---|---|
     |161|162|163|164|165|166|167|168|
   3| 61| 64|  5|  6| 48| 10| 21| 45|
     |---|---|---|---|---|---|---|---|
     |169|170|171|172|173|174|175|176|
   4| 23| 46| 12| 41| 51| 47| 60| 18|
     |---|---|---|---|---|---|---|---|
     |177|178|179|180|181|182|183|184|
   5| 28| 37| 40| 65| 29| 22| 36| 38|
     |---|---|---|---|---|---|---|---|
     |185|186|187|188|189|190|191|192|
   6| 32| 69| 55| 15| 72| 34|  7| 39|
     |---|---|---|---|---|---|---|---|
     |193|194|195|196|197|198|199|200|
   7|  2| 59|  8| 17| 52| 70| 30| 26|
     |---|---|---|---|---|---|---|---|
     |201|202|203|204|205|206|207|208|
   8| 57| 27| 16| 43| 25| 71|  4| 24|
     |---|---|---|---|---|---|---|---|
     |209|210|211|212|213|214|215|216|
   9| 50| 62| 44| 31|  3| 54| 14| 58|
     |---|---|---|---|---|---|---|---|
```

TRT		REP 1 PLOT	REP 2 PLOT	REP 3 PLOT
1	.....	52	92	145
2	.....	6	137	193
3	.....	70	101	213
4	.....	49	110	207
5	.....	4	77	163
6	.....	24	128	164
7	.....	71	122	191
8	.....	22	97	195
9	.....	32	79	160
10	.....	72	105	166
11	.....	20	115	151
12	.....	51	80	171
13	.....	56	127	153
14	.....	55	83	215
15	.....	25	144	188
16	.....	23	95	203
17	.....	50	119	196
18	.....	12	120	176
19	.....	2	109	150
20	.....	31	123	147
21	.....	16	88	167
22	.....	18	75	182
23	.....	1	125	169
24	.....	7	135	208
25	.....	61	86	205
26	.....	13	93	200
27	.....	68	117	202
28	.....	36	91	177
29	.....	42	104	181
30	.....	28	107	199
31	.....	17	108	212
32	.....	44	84	185
33	.....	64	139	149
34	.....	21	133	190
35	.....	69	76	146
36	.....	65	141	183
37	.....	15	121	178
38	.....	59	111	184
39	.....	54	102	192
40	.....	27	114	179
41	.....	60	99	172
42	.....	19	87	159
43	.....	43	142	204
44	.....	9	143	211
45	.....	45	113	168
46	.....	38	106	170
47	.....	41	90	174
48	.....	58	131	165
49	.....	67	96	158
50	.....	34	78	209

```
 51 ......................      66     134     173
 52 ......................      40     132     197
 53 ......................      33     140     155
 54 ......................      48     126     214
 55 ......................       8      94     187
 56 ......................      62     118     154
 57 ......................      30     100     201
 58 ......................      57      89     216
 59 ......................      47      73     194
 60 ......................      29      85     175
 61 ......................      53     138     161
 62 ......................      26     129     210
 63 ......................      46     136     148
 64 ......................      35     103     162
 65 ......................       3      81     180
 66 ......................       5      98     157
 67 ......................      37      82     152
 68 ......................      10     130     156
 69 ......................      39     116     186
 70 ......................      63     124     198
 71 ......................      11      74     206
 72 ......................      14     112     189
```

V. Guidelines

PLANT BREEDING DESIGNS

RCB

- number of treatments x number of replicates ≤ 1000 plots

Lattice Block

Valid No. of Entries	Maximum No. of Replicates	Maximum No. of Plots
16	5	80
25	6	150
49	6	294
64	3	192
81	3	243
100	3	300

Lattice Square

Valid No. of Entries	Maximum No. of Replicates	Maximum No. of Plots
16	6	96
25	6	150
49	6	294
64	4	256
81	3	243

Rectangular Lattice

Valid No. of Entries	Maximum No. of Replicates	Maximum No. of Plots
12	3	36
20	3	60
30	3	90
42	3	126
56	3	168

Latin Square

- maximum of 31 treatments
- maximum of 3 distinct squares
- number of treatments x number of treatments x number of distinct squares ≤ 1000 plots

Augmented Latin Square

- At least 1 unreplicated entry
- Maximum of 20x20 imbedded squares, i.e. maximum of 20 replicated entries
- number of unreplicated entries + (number of replicated entries x number of replicated entries) $\leq$ 1000 plots

Augmented RCB

- number of unreplicated entries + (number of replicated entries x number of blocks) $\leq$ 1000 plots

Alpha Lattice

- minimum of 4 and maximum of 500 entries
- number of replicates should be between 2 and 10
- number of entries x number of replicates $\leq$ 1000 plots

FACTORIAL DESIGNS

- minimum of 2 and maximum of 255 levels per factor
- maximum of 8192 combined levels of factor(s)

Exercise 1

CropStat: Randomization and Layout and Plot Sampling

A. An experiment to determine the response of 4 varieties to 3 different infestation levels will be conducted in transplanted rice. A 3 × 4 factorial in Split Plot Design with 4 replications will be used. Field size is 48 m × 24 m, plot size is 6 m × 4 m. Plant spacing is 20 cm × 20 cm. The three infestation levels are I_1, I_2 and I_3. The four varieties are V_1, V_2, V_3 and V_4. Generate an experimental layout. Save it in file EXER1A.

B. A field with 1000 plots in 20 rows of 50 plots each is available to screen only generation breeding material. The field may have gradients of fertility and water status in two directions. Generate a randomized layout and field book for an Augmented Latin Square with 12 check entries and 856 test entries. Save results in file EXER1B.

C. We want to test 23 advanced breeding lines and 2 standard checks at three locations. Generate randomized layout plans and a field book for three trials of 25 entries in a triple lattice block design (lattice block with 3 reps). Save results in file EXER1C.

DATA AND FILE MANAGEMENT

At the end of the course, the user should be able to

- directly supply the data using the Data Editor through entering the data directly to the new Cropstat tables;
- import the data from text file, print file, excel file, and dbase file;
- create new variable using arithmetic operations, functions, and logical procedures;
- create variable using existing variables with missing values; and
- export a SYS file to excel worksheet or text file.

I. Sample Data Set : Fertilizer Experiment with 9 treatments

A fertilizer experiment was conducted with the following treatments:

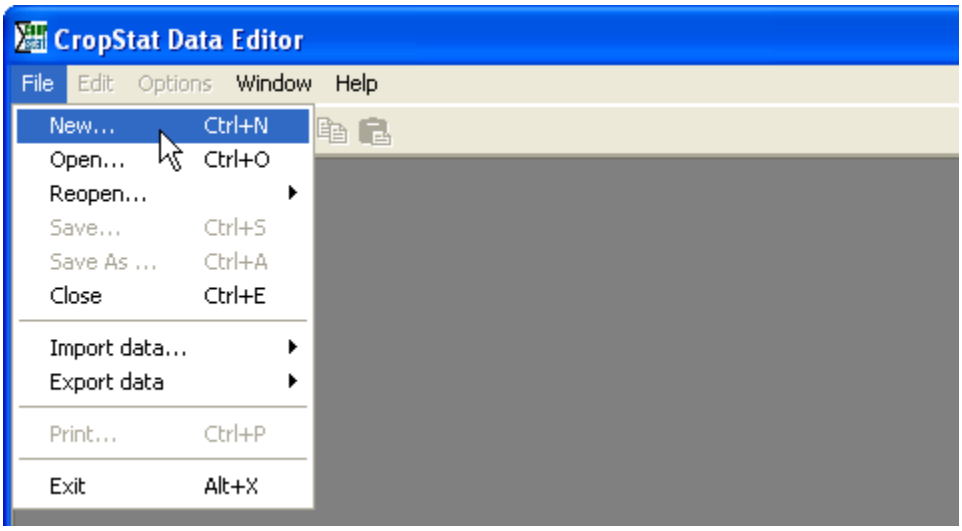
1 = Control
2 = Low N-rate urea
3 = Low N-rate SCU
4 = Low N-rate USG
5 = Low N-rate USG/urea
6 = High N-rate urea
7 = High N-rate SCU
8 = High N-rate USG
9 = High N-rate USG/urea

The data are shown as follows:

TREATMENT	REPLICATION	MOISTURE	GRAIN WEIGHT
Control	1	12.728	2.889
L-N Urea	1	15.911	4.631
L-N SCU	1	17.213	5.283
L-N USG	1	16.379	6.502
L-N USG/Urea	1	16.793	5.426
H-N Urea	1	12.442	5.579
H-N SCU	1	14.942	6.224
H-N USG	1	14.727	6.217
H-N USG/Urea	1	12.643	5.862
Control	2	17.703	5.231
L-N Urea	2	16.223	4.371
L-N SCU	2	15.191	4.421
L-N USG	2	17.827	6.001
L-N USG/Urea	2	13.926	5.649
H-N Urea	2	17.737	6.024
H-N SCU	2	17.607	6.659
H-N USG	2	15.210	6.826
H-N USG/Urea	2	15.387	5.891
Control	3	16.027	3.081
L-N Urea	3	16.215	5.861
L-N SCU	3	16.445	5.576
L-N USG	3	16.001	6.155
L-N USG/Urea	3	14.724	5.361
H-N Urea	3	14.005	5.648
H-N SCU	3	12.941	6.449
H-N USG	3	16.418	7.145
H-N USG/Urea	3	15.257	6.022

II. Data Creation

The first step in data analysis using CropStat is converting the data file into a SYS file through the Data Editor. The data file may come from a text file (ASCII file), excel workbook, dbase file, or directly supplied to the Data Editor. The Data Editor is like a Worksheet wherein the variables are in columns and the observations are in rows.



To open the Data Editor, select **Window ⇒ Data Editor** from the Main Window.

At the start of Data Editor, the menus that are available are File, Window, and Help (in bold letter) while the Edit, Options, and Tools are disabled, but they will be available once you open the file**.**

You can supply data file through the File menu in four different ways:

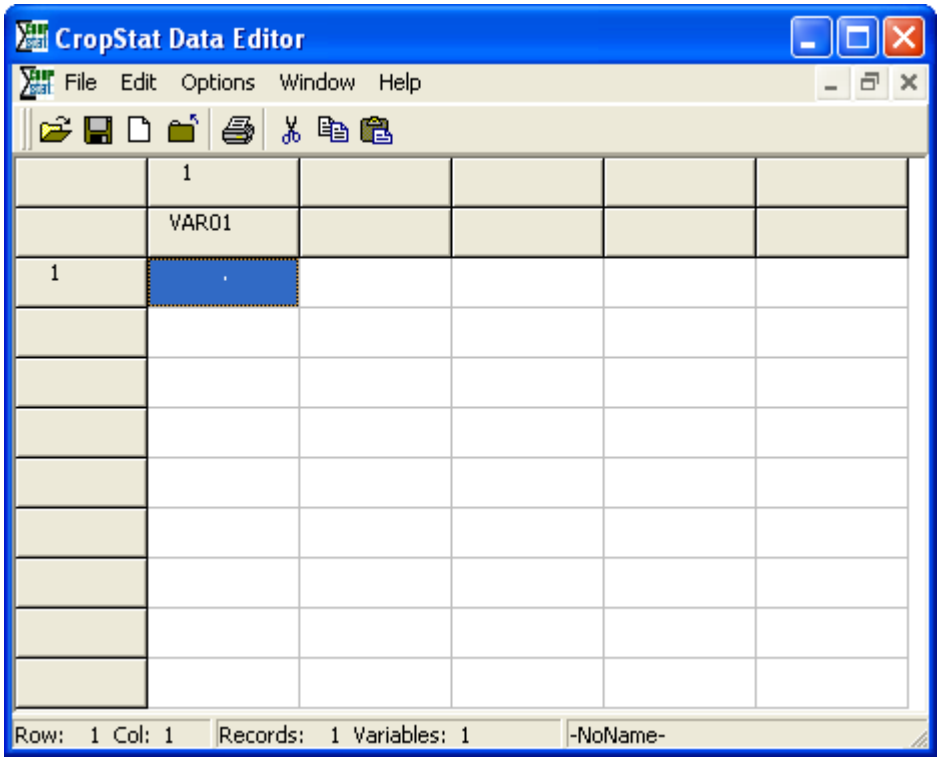
A. **New** to directly supply the data using the Data Editor

B. **Open** to open an existing SYS file

C. **Reopen** to reopen the five most recent opened files

D. **Import** to import from an excel workbook, a text file or a dbase file

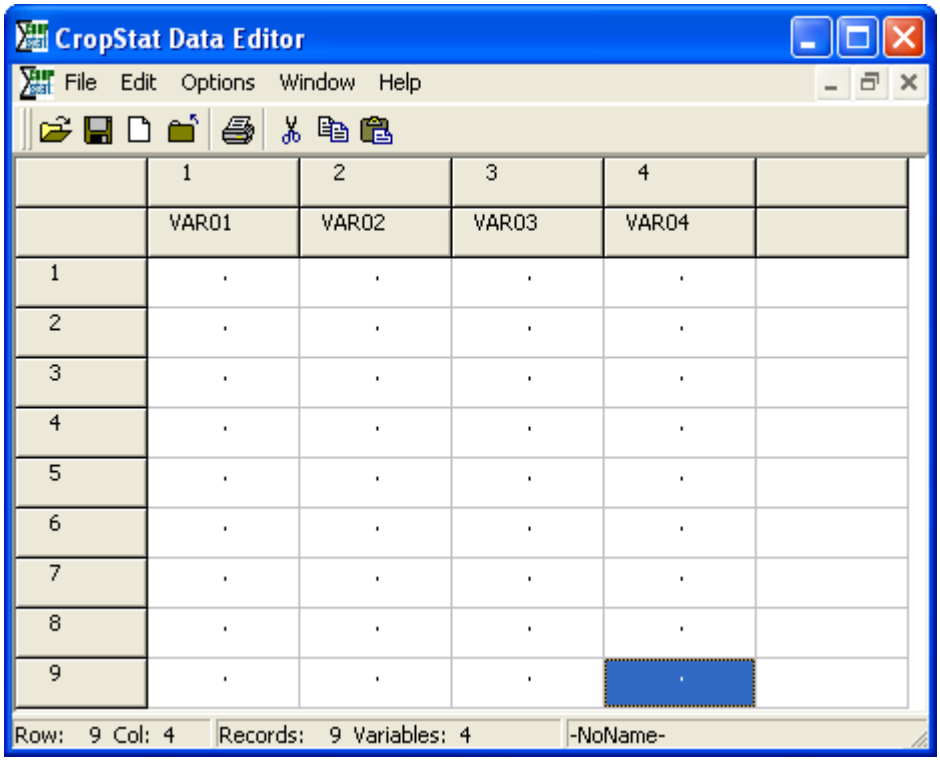
A. New - DATA ENTRY USING THE DATA EDITOR



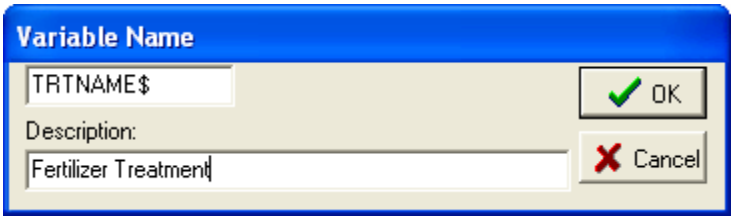
- To create a new CropStat data file, select **File ⇒New** from the Data Editor Window

- By default, CropStat assumes the size of your data set as 1 × 1 (one observation, one variable). Add one more row (record) by pressing the Down Arrow key. Do it repeatedly until you have 27 rows. To add one more column (variable), press the Right Arrow Key. Increase the number of columns to 4. The maximum number of variables and observations the Data Editor can handle will depend on the hardware.

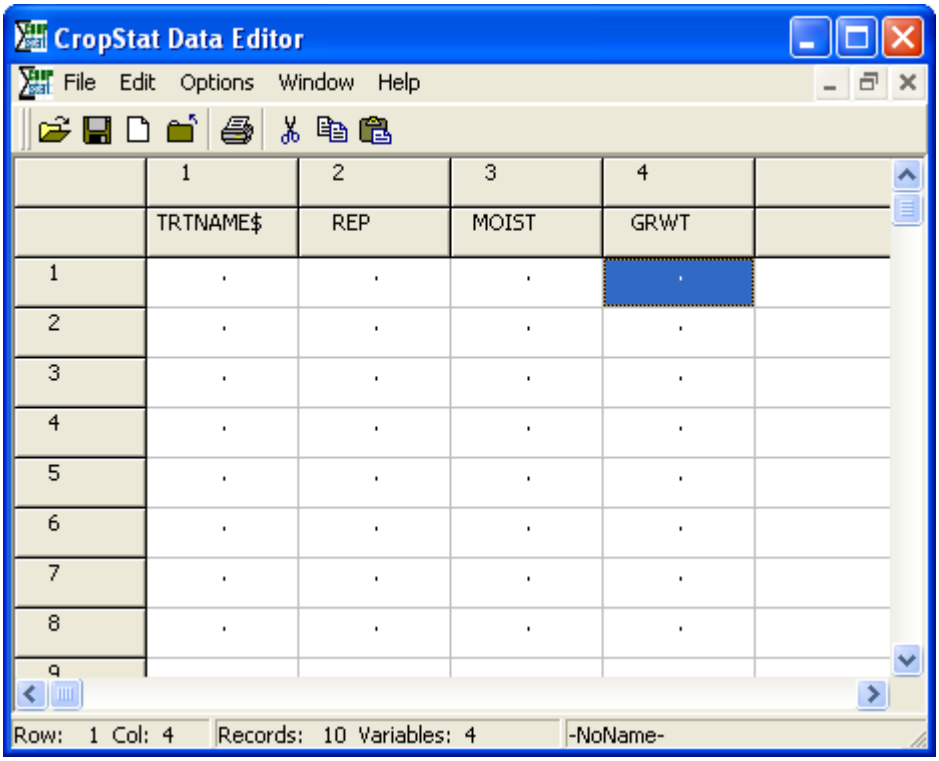
- As new rows and columns are added to the data set, CropStat fills them with missing values which are large negative numbers (-1.E36) but which appear on the screen as dots.



- CropStat provides a default variable name for each column (i.e. VAR01, VAR02, VAR03, etc.). You may change the variable name by pointing your cursor to the variable name you want to change and clicking the left button of the mouse. An input box will prompt you to enter a variable name. It is suggested that you put a variable description for the data documentation.

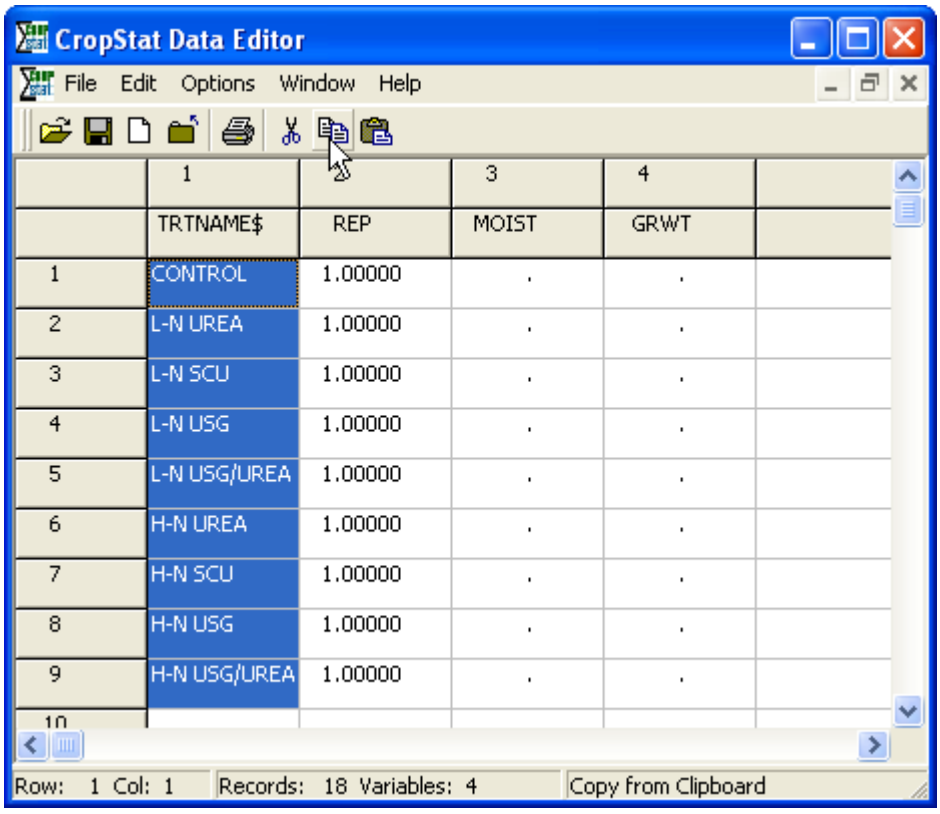


- Assign the following names to each column: *TRTNAME$*, *REP*, *MOIST*, and *GRWT*. Note that a character variable has a $ at the end of its variable name and that all variable names must not be longer than 8 characters (excluding the $).

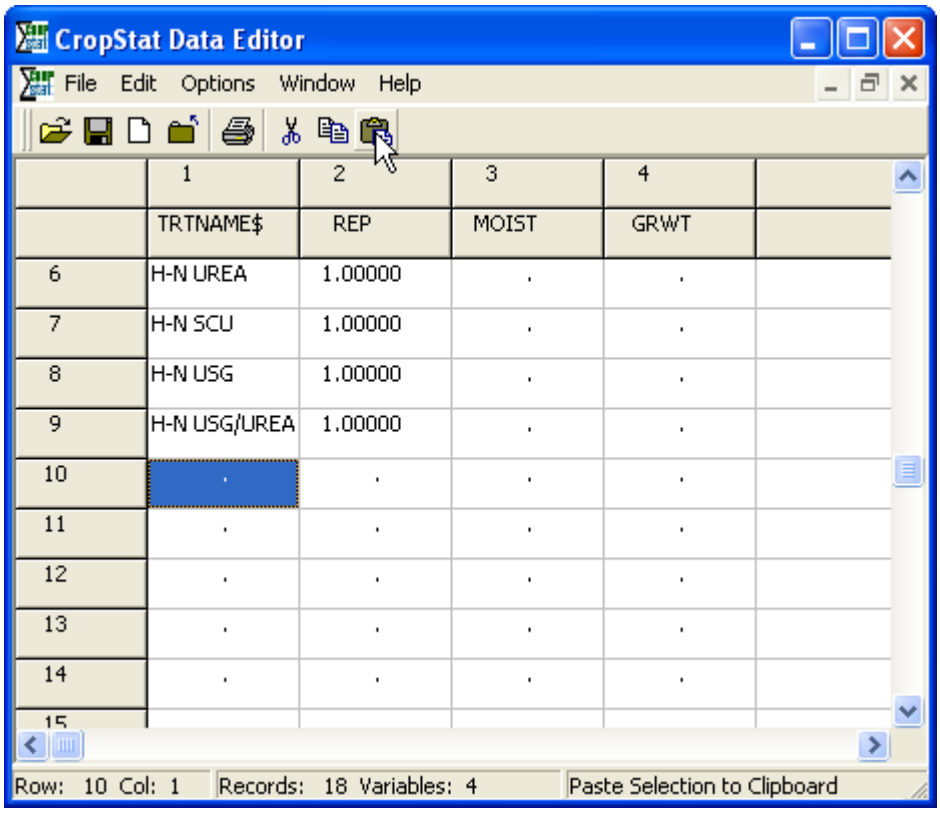


- Type the data values in each cell. To move across columns, press <Tab> after each entry. To move across rows, press the <down arrow key>. Note that values of character variables should not be longer than 12 characters.

- Values of the design variables may be copied and pasted for quick and easy data entry. Highlight the range of values to be copied. Click **Edit ⇒ Copy** in the Data Editor or click the copy icon  .

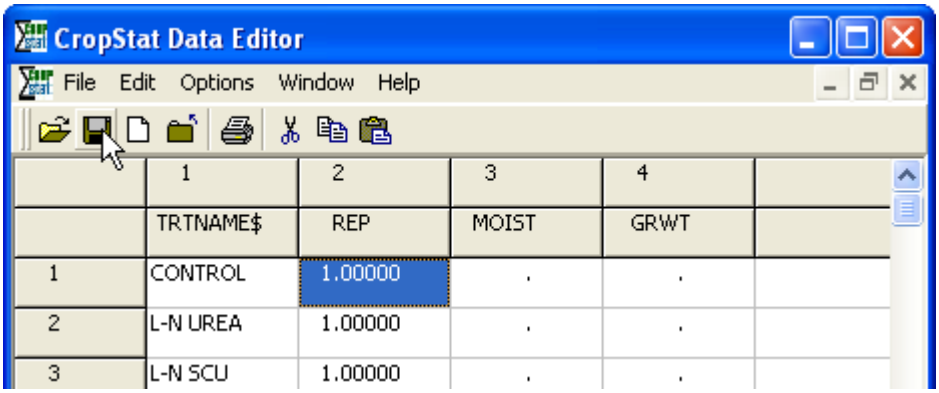


- Click on the cell where the copy will be placed. Select **Edit ⇒ Paste** or click the Paste icon  .

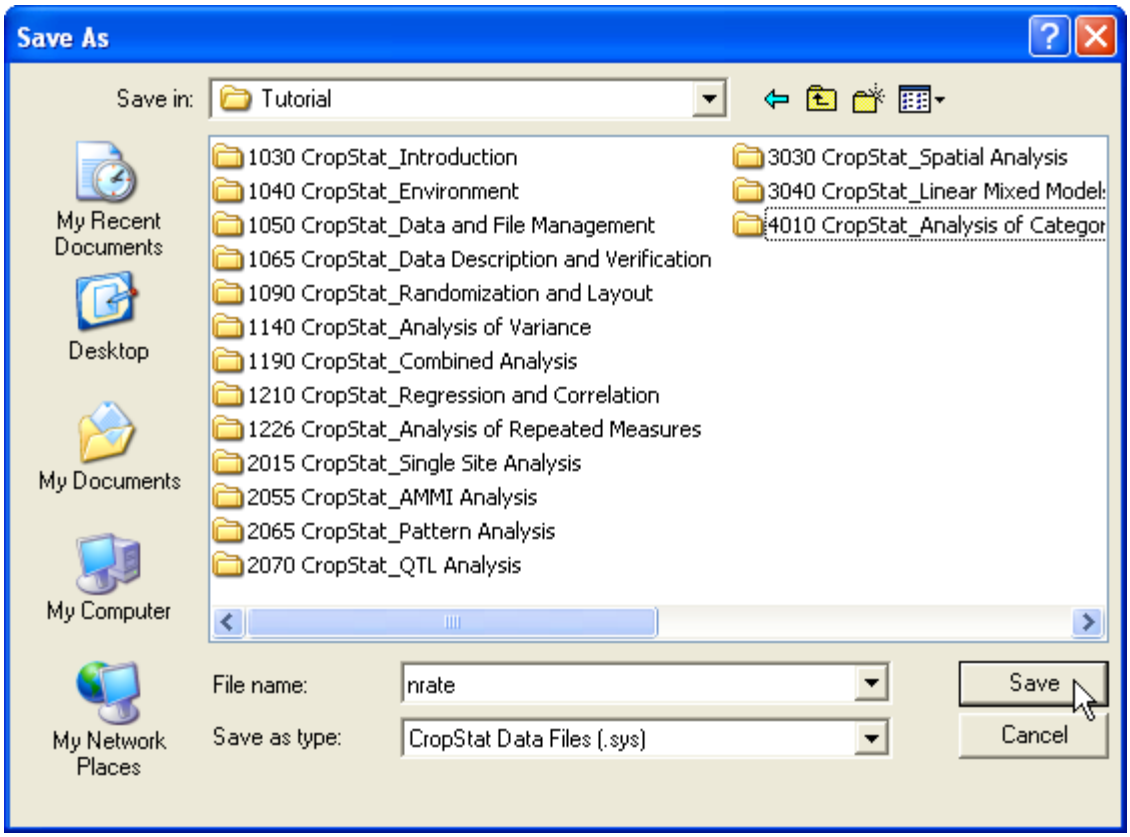


- When data entry is finished, select **Options ⇒ Edit Description** to add information about project details. Type the following text to the **Heading** edit box: *Effect of Nitrogen Rate and Source on Grain Yield of Rice* (***Note***: the variable names and variable description are automatically stored). Click **OK** to close the **Heading** edit box.

- Select the **FILE⇒SAVE** menu option or the save icon  .



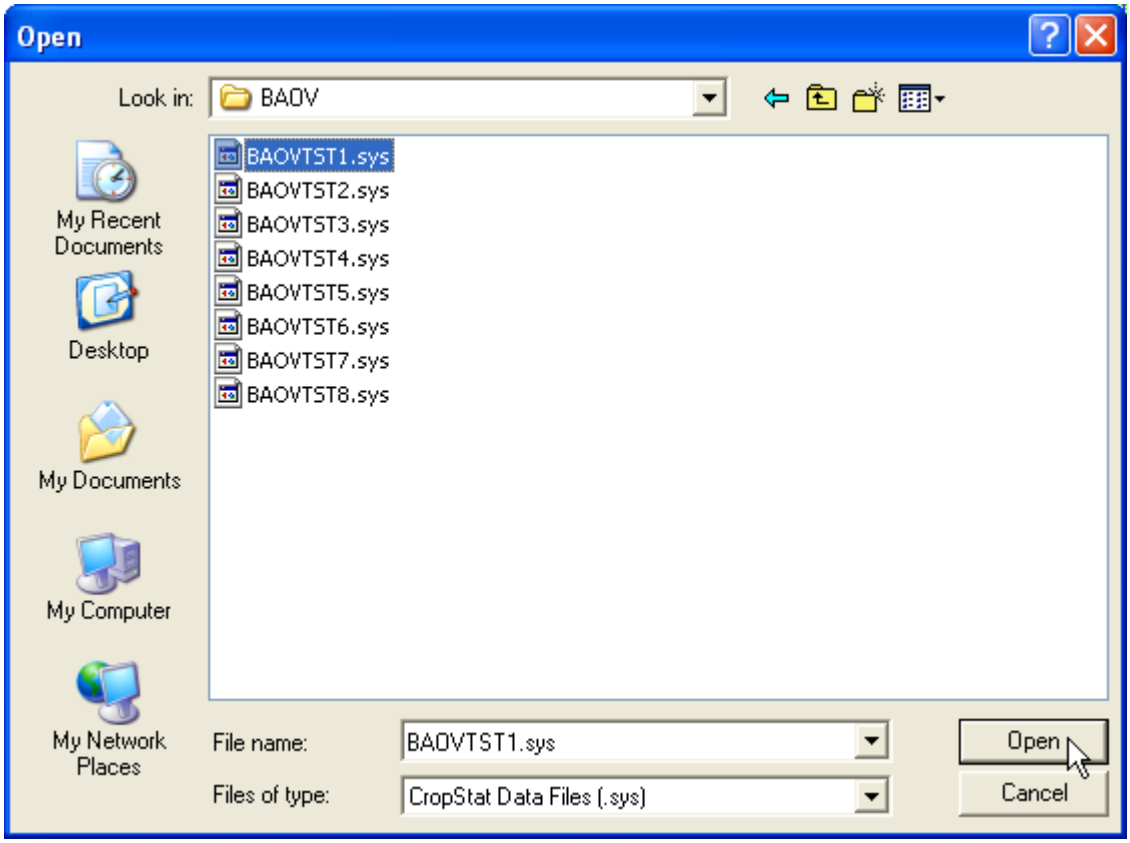
- The **Save As** dialog box will prompt you to enter the name of your **SYS** file. Click the **Save In** box and select data disk drive C. Then double-click (or click Open) **TUTORIAL** in the folder list. Type **NRATE** (the extension SYS will be automatically appended) in the File name window and click **Save.**



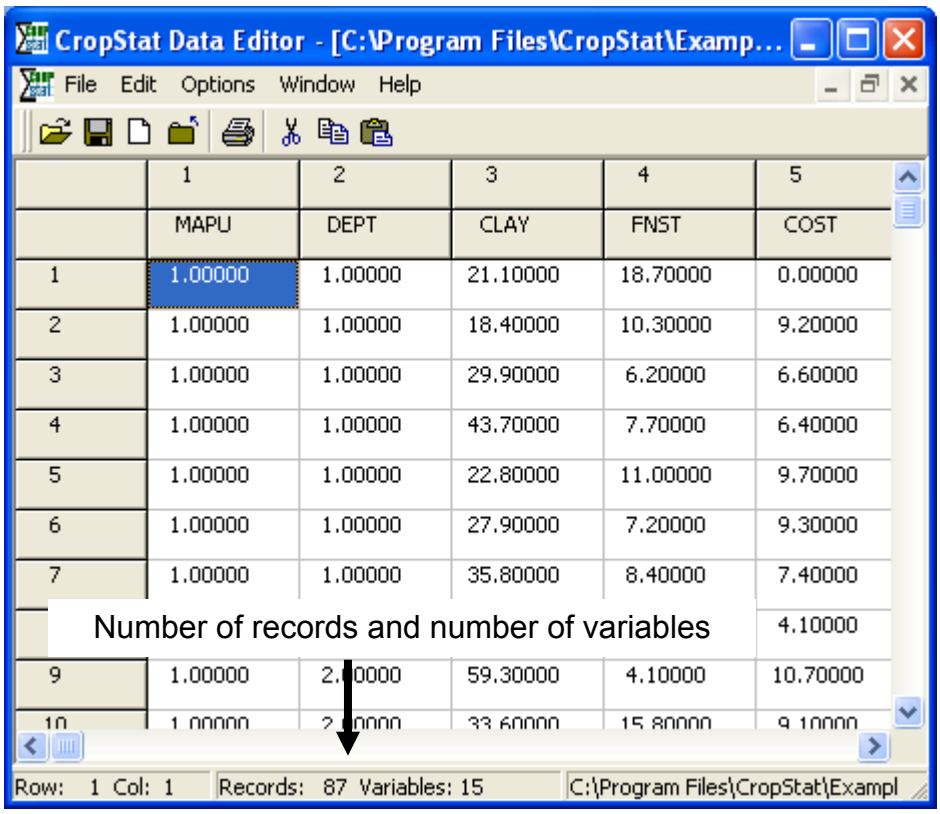
- Once data have been saved, the location of the file and the name of the SYS file are indicated at the lower right corner of the Data Editor.

B. Open – DATA ENTRY BY OPENING AN EXISTING SYS FILE

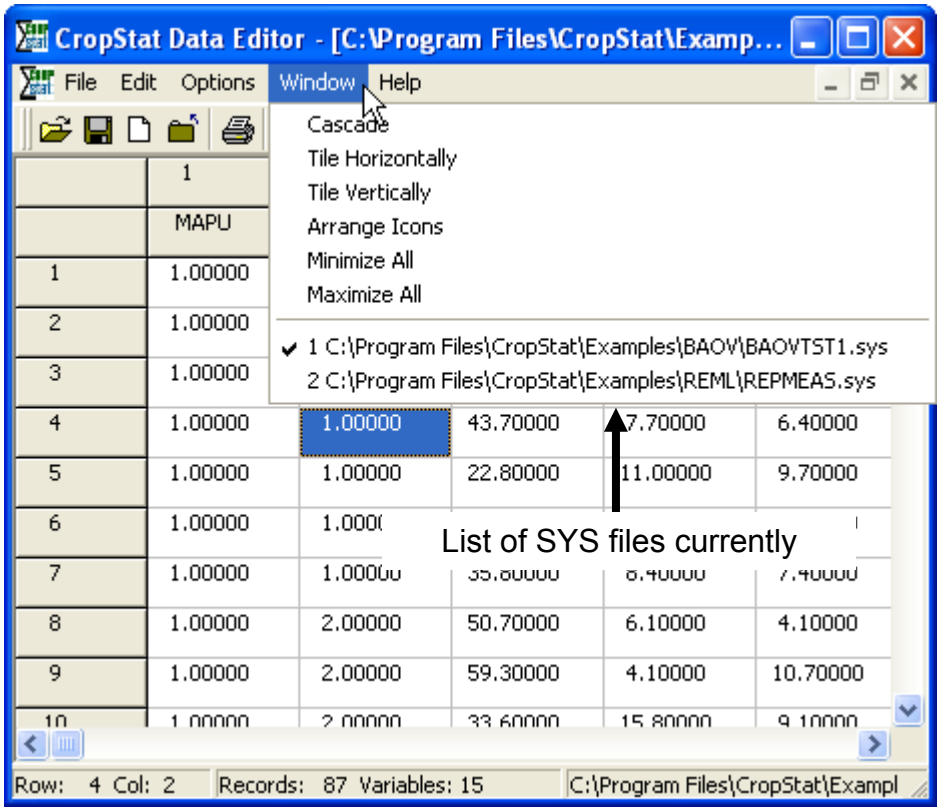
- To browse, edit, add variables and records, delete variables and records, and do some calculation to an existing SYS file, select **File ⇒Open**.
- The Open dialog box will prompt you to select the SYS file you want to open by navigating *C:\PROGRAM FILES\CROPSTAT7.2\EXAMPLES\BAOV*. Select the file and click **Open.**

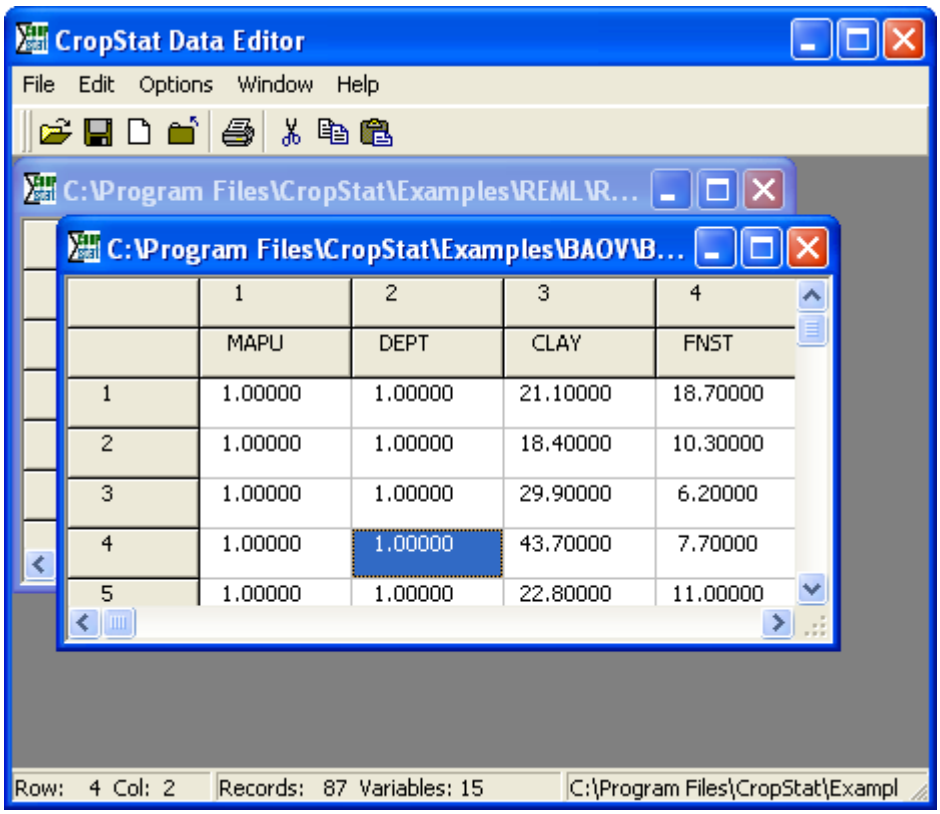


- The Data Editor will show you the number of records, the number of variables, and the location of the file you opened.

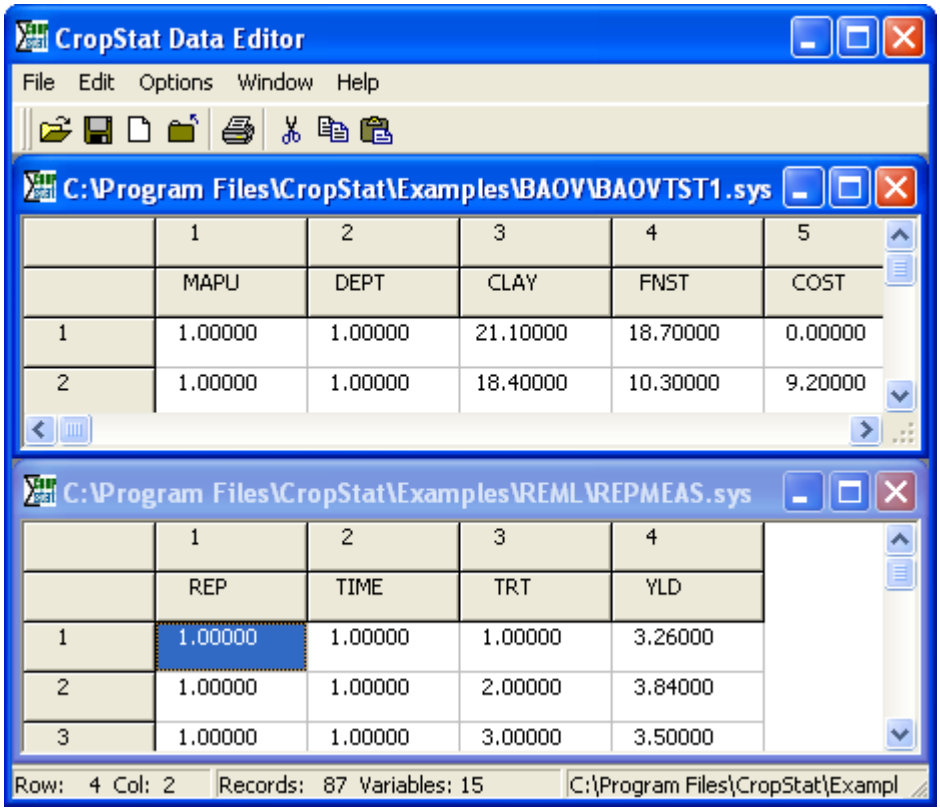


- To open another file (e.g. RepMeas.sys), click on the **File ⇒Open** and select the SYS file. You can open a SYS file stored in a different folder. To view the 2 SYS files, click on Window ⇒ choose from the different view format such as Cascade, Tile Horizontally, and Tile Vertically.

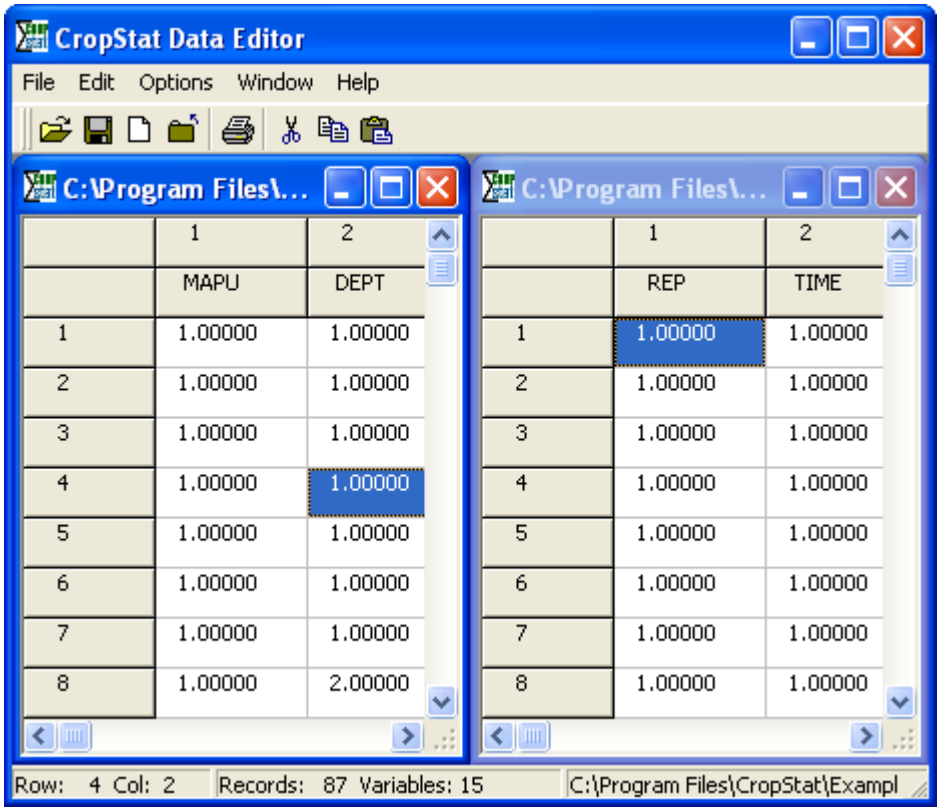




Cascade



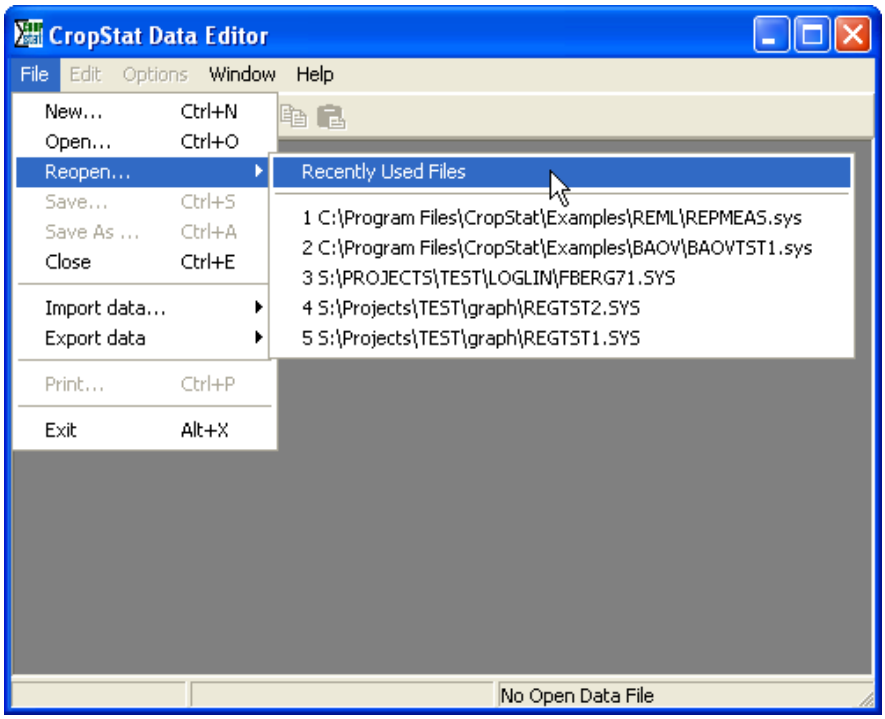
Tile horizontally



Tile vertically

C. Re-open – DATA ENTRY BY REOPENING THE FIVE MOST RECENTLY OPENED FILES

- To reopen a SYS file, select **File ⇒ Reopen.** You can only reopen the five most recently opened files.

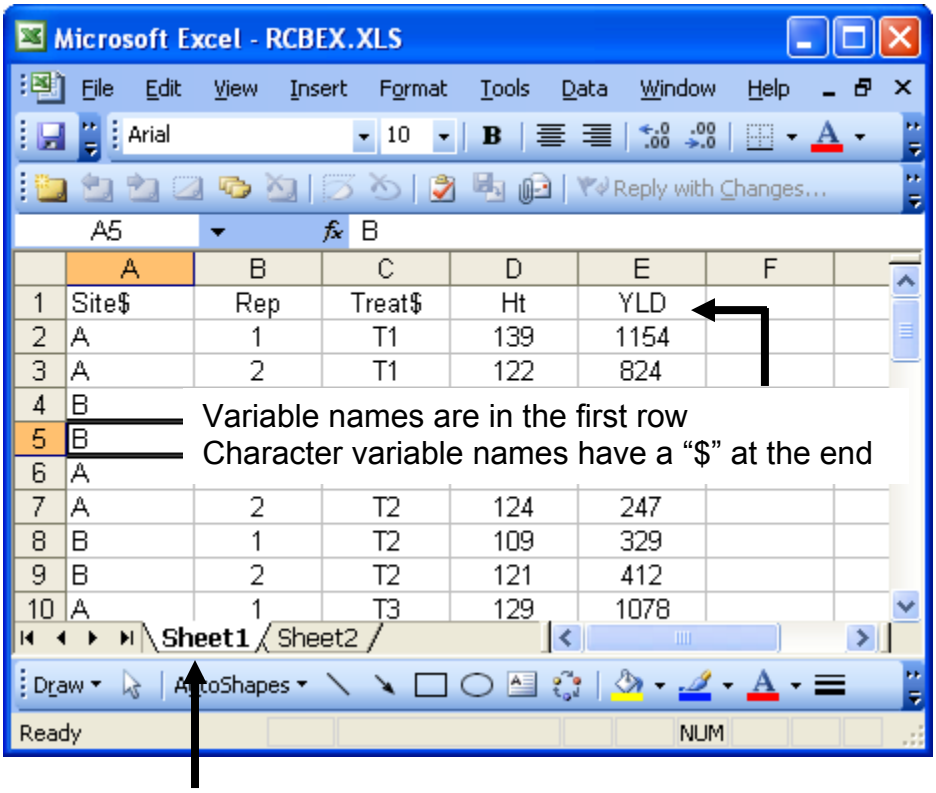


D. Import

- **FROM AN EXCEL WORKBOOK**

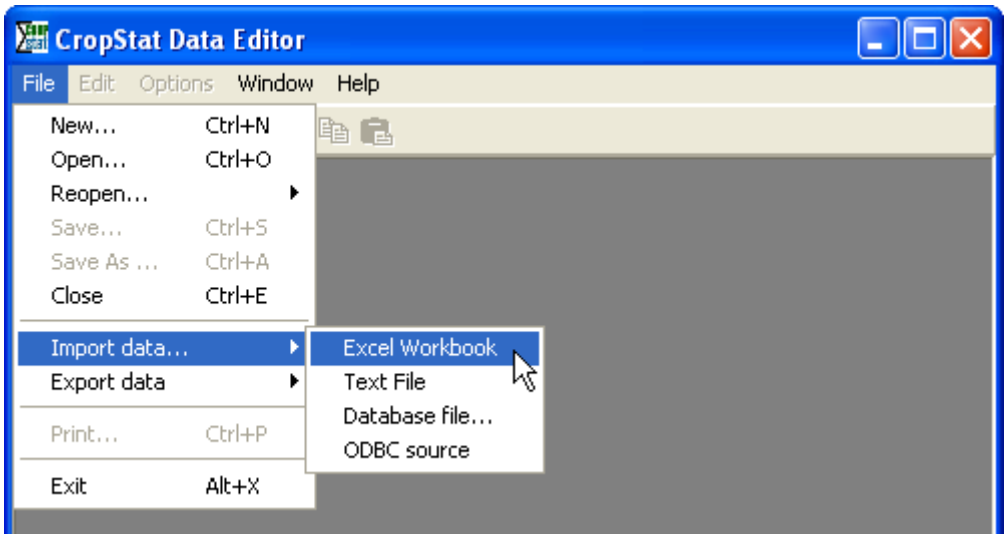
  - **Data Preparation**

From excel prior importing to Cropstat Data editor, the variable names should be in the first row, 8 characters (excluding $) and character names should have a $ at the end.

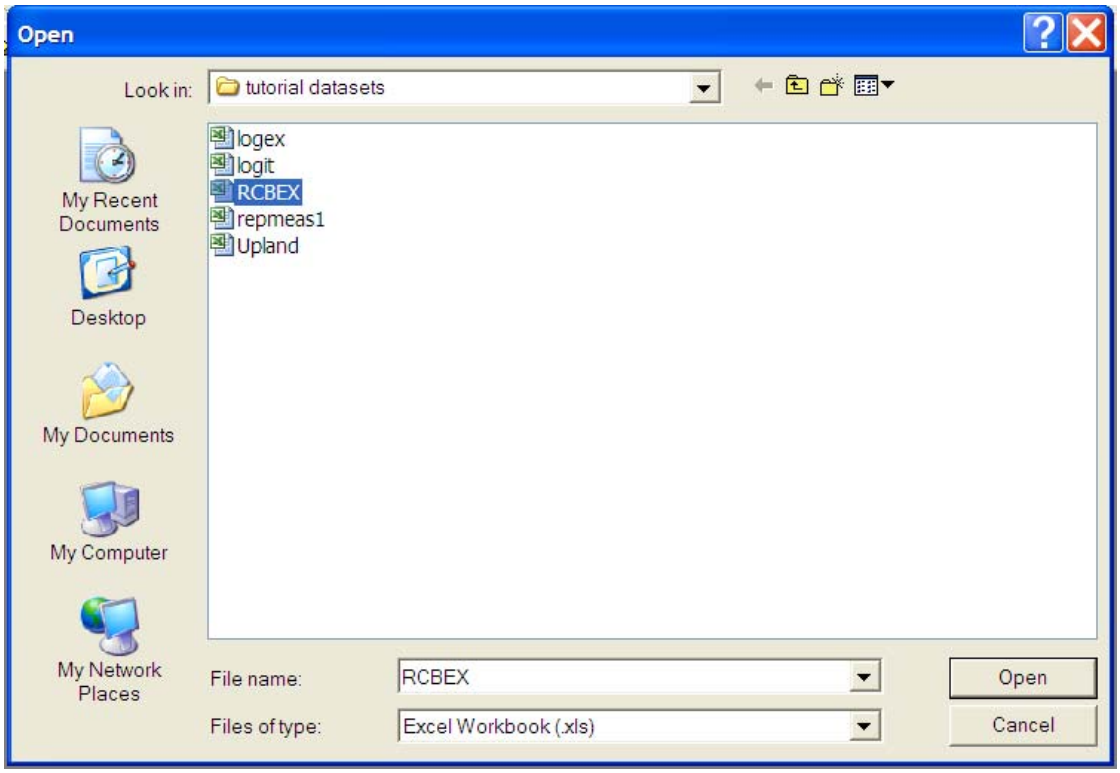


From this example the data is stored in the first worksheet.

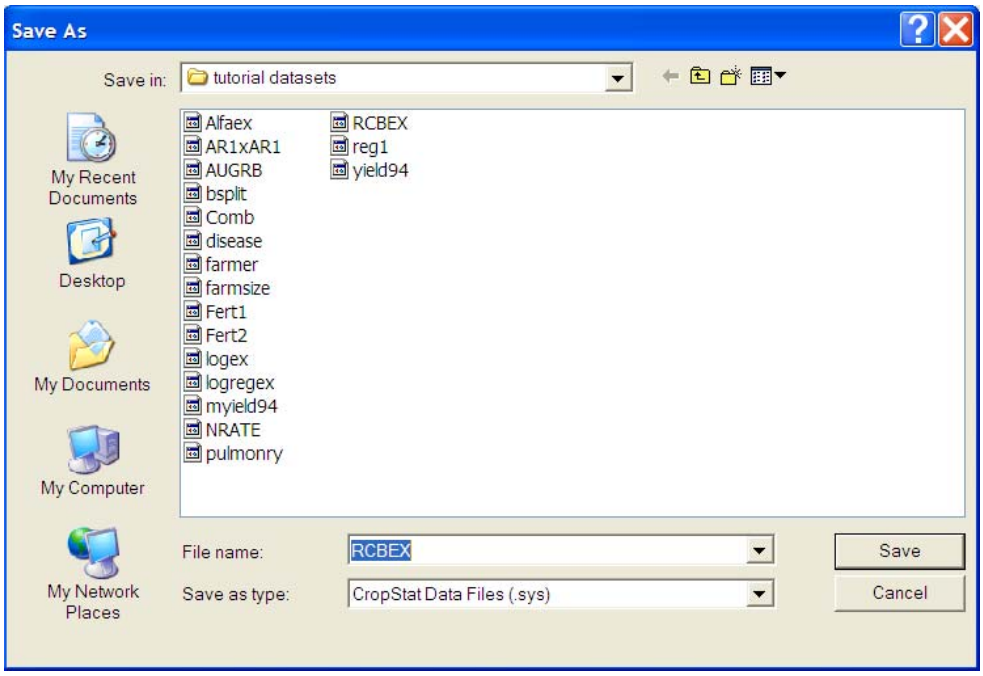
- To import from an excel workbook, select **File ⇒Import data⇒Excel Workbook**.



- The Open dialog box will prompt you to enter the excel file you want to open. Double click on the selected XLS file.



- The **Save As** dialog box will prompt you to supply the name of the SYS file. The default is the filename of the XLS file. Click **Save** to proceed to importation of the file.



- The **Select Worksheet to Import** dialog box will prompt you to select the name of worksheet from where your excel file was saved. CropStat can read all the sheets one at a time of the workbook you opened. Different sheets will be saved in different SYS file.

- Click **Ok** to import the data.

- **FROM A TEXT FILE (ASC, TXT, PRN)**
  - **ASC text file**
    - **Data Preparation**

      The basic raw data and project details can be supplied to CropStat in ASC data files, which are ASCII files with records shorter than 256 characters. They can be created and modified with any text editor. There are two sections to these ASC files — the first contains project and variate descriptions and the second contains the numeric design and response data.

      The first section of the ASC data file for a particular project should contain all the details necessary to analyze the results. The following must be included:

      1. Project title and season;
      2. Identification of researchers;
      3. Description of basic design(s) used;
      4. Relevant site details, name, map reference, planting date, rainfall, soil description, harvest date, relevant extracts from trial diaries, etc;
      5. List of treatments detailing levels and application methods; and
      6. List of missing values with suggested reasons and proposed actions.

      Each factor and data variate in the file should also be described in detail with variate description records, which assign numbers, names, and descriptions to each variate.

```
H 2OTH INTERNATIONAL IRRIGATED RICE YIELD NURSERY - EARLY 1993)
H (21ST IIRYN-E 1993) LOCATION DATA
V001 TRIAL$ TRIAL CODE (REGION/COUNTRY/LOCATION/NO)
V002 REGION$ REGION
V003 COUNTRY$ COUNTRY
V004 SEEDL DATE OF SEEDING (MM/DD/YY)
V006 SOILPH SOIL PH
V007 N(1) NITROGEN APPLICATION ON FIRST SPLIT
V008 N(2)
V009 N(3)
V010 NTOT TOTAL NITROGEN APPLICATION
V011 PTOT TOTAL P APPLICATION
V012 K20 K20 APPLICATION
V013 RAIN(1) RAINFALL DURING FIRST MONTH
V014 RAIN(2)
V015 RAIN(3)
V016 RAIN(4)
V017 RAIN(5)
V018 TRAIN TOTAL RAINFALL
…
```

Example of headings and variate descriptions.

    - **Data types and format of the input data**

Two types of data are supported by CropStat: numeric data and character data. All values in a particular variable must be of one type. Names of character variates must end with a $ in order for CropStat to distinguish between the two data types. Numeric variables are always stored in SYS files as floating point numbers even if they have integer values. Character values can have up to 12 characters and may be used to define subsets of the data or as factor levels to describe the design.

Input data files with extension .ASC are sequential ASCII files with records of up to 400 characters in length. The first part is a heading and variate description section. The heading information is simply entered on records that must start with 'H' for HEADING in the first position and should not extend beyond position 72.

Variate description records have the format:

VNNN MNEMONIC Text …… TO POSITION 72
Text continued …….....TO POSITION 72

They have a 'V' in the first position, a three digit decimal number, NNN in positions 2 to 4, and position 5 is blank. A name or mnemonic for the variable follows as the first word starting in position 6. This name may consist of up to eight uppercase, alphanumeric characters starting with a letter and it must terminate with a $ if the variable is a character variable. The remainder of the line up to position 72 contains a description which may be continued on subsequent records provided columns 1 to 5, or more, are left blank. Variate description and heading records may be interspersed.

The heading and variate description section must end with a record starting with '////' in columns 1 to 4. This record may also have 'TRIAL' in columns 5 to 9, column 10 is blank and a project description code in columns 11 to 20.

The data records follow after the '////' record. All variable values are given for each plot in turn, although the order of the plots is not important. Data for each plot must start on a new record. If there are too much data for each plot to fit on one 80-character record, it is continued on subsequent lines.

The variate numbers in the variate description records must correspond to the order of the variate values for all plots on the data records. There are two options for data format: a) free format in which every data value must be separated by a comma or at least one space and where decimal points must be entered unless the values are integers; and b) fixed format where a format specification follows the '////' record, before the data, and indicates

the columns for each variate and the position of any omitted decimal points.

- **Free format data records**

  In free format, variate values may be separated by single commas and/or any number of spaces. Values for each unit may occupy any number of lines provided that a line break does not divide a single data value and that each plot starts on a new record.

  Character variables are similarly separated from other variables. If the character value is to contain a space or a comma, it must be contained in double or single quotes as well as be separated from other variables by a comma or by spaces. The quotes are stripped off from the data value when it is read. If the value is to contain one type of quote mark, the whole value must be contained in the other type. Character values shorter than 12 characters are padded with spaces, those longer are truncated, and the extra letters are lost.

```
H SAMPLE1.ASC is an example
H of a free format data file for CropStat
V001 PLOT Plot number 1-6
V002 REP Replication 1-2
V003 TRT Treatment 1-3
         1=CONTROL,  2=NEW TEST A,  3=NEW TEST B
V004 OBSV Observation (g/plot)
V005 TRTNAME$ Treatment name
////TRIAL SAMPLEFREE
1 2 1 123.34  CONTROL
2 1 2 341.12  'NEW TEST A'
3 1 3 172.14  'NEW TEST B'
4 2 1 183.12  CONTROL
5 2 2 721.34  'NEW TEST A'
6 2 3 425.81  'NEW TEST B'
```

Free format data records.

- **Fixed format data records**

  In fixed format data, the format statement follows the rules for FORTRAN format statements. It may occupy up to five lines following the //// record and must be enclosed in parentheses. It may only include real field descriptors, rFx.y, skip fields, rX, and character field descriptors, rAx, separated by commas or slashes.

  The field descriptor rFx.y indicates r consecutive fields of x characters with y digits after a decimal point. The r may be omitted if it is 1, and any explicitly punched decimal point in the data overrides the y value. The field rX indicates that r characters must be skipped on input. It is useful for skipping text data which have been entered into the data file and which

are not wanted in a character variable. In rAx fields, r is a repeat count as for F descriptors and the x is a field width. If x < 12, the characters are truncated. The slash (/) separator indicates the start of a new record.

The following example specifies exactly the same data as the previous example for a free format data file.

```
H SAMPLE2.ASC is an example of
H a fixed format data file for CropStat
V001 PLOT Plot number 1-6
V002 REP Replication 1-2
V003 TRT Treatment 1-3
V004 OBSV Observation (g/plot)
V005 TRTNAME$ Treatment name
////TRIAL SAMPLEFIXD
(F1.0,F2.0,F1.0,F6.2,2X,A10)
1 1112334  CONTROL
2 1234112  NEW TEST A
3 1317214  NEW TEST B
4 2118312  CONTROL
5 2272134  NEW TEST A
6 2342581  NEW TEST B
```

The following example gives another format for the same data where the first text for each plot has been skipped with the 11X descriptor and the second has been read with the A10 descriptor.

```
H SAMPLE3.ASC is an example of
H a fixed format data file for CropStat
V001 PLOT number 1-6
V002 REP Replication 1-2
V003 TRT Treatment 1-3
V004 OBSV Observation (g/plot)
V005 TRTNAME$ Treatment name
////TRIAL SAMPLEFIXD
(F1.0,1X,F1.0,11X,F1.0/2X,F6.0,A10)
1 1 VALENCIA  1
  123.34CONTROL
2 1 CROSS 1x3 2
  342.12NEW TEST A
3 1 VARIETY A 3
  172.14NEW TEST B
4 2 VALENCIA  1
  183.12CONTROL
5 2 CROSS 1x3 2
  721.34NEW TEST A
6 2 VARIETY A 3
  425.81NEW TEST B
```

- **Missing values.**
  Missing values are flagged in the SYS file by large negative numbers (-1.E36). In free format input, they can be indicated in the ASC file by putting an asterisk '*' in the field corresponding to the missing value. In fixed format, a value such as –1.E36 must be placed in the appropriate field.

```
////                and          ////
1 1 2 * 100                     (F1.0,2F2.0,2F6.0)
                                1 1 2-1.E36 100
```

The example would both flag the fourth variate value in the first plot as missing. Note, however, that format F5.x reads -1E36 as -1.0E(36-x). To indicate a missing value in F5.2 format, you must enter -1E38 or use a wider field and enter the decimal point explicitly.

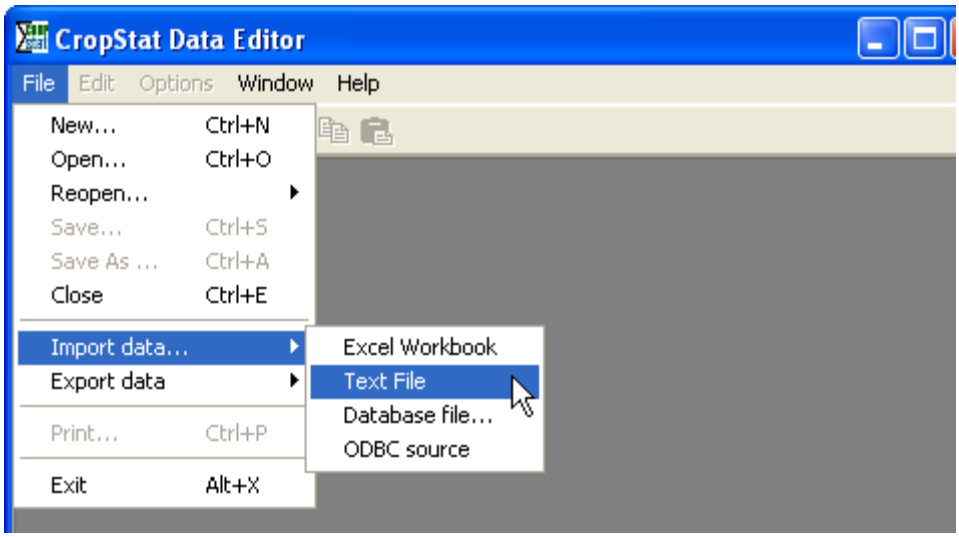
- **Importing ASC files into CropStat**

The sample data set (NRATE) was formatted into an ASC, as shown below. This file is stored in *C:\PROGRAM FILES\CROPSTAT7.2\ TUTORIAL\TUTORIAL DATASETS* folder with the filename NRATE.ASC.

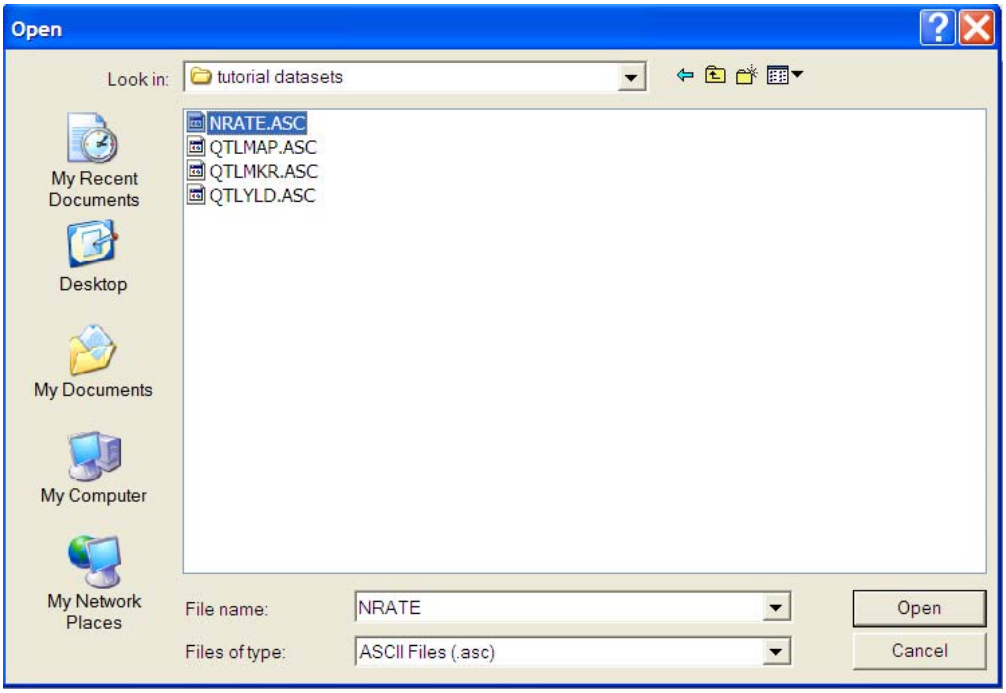
```
H FERTILIZER EXPERIMENT
H SAMPLE DATA SET
V001 TRTNO          TREATMENT NO. (1-9)
V002 TRTNAME$       TREATMENT NAME
V003 REP            REPLICATES (1-3)
V004 MOIST          MOISTURE CONTENT (14%)
V005 GRWT           GRAIN WEIGHT
////
1 CONTROL                1    12.728   2.8890
2 'L-N UREA'             1    15.911   4.6310
3 'L-N SCU'              1    17.213   5.2830
4 'L-N USG'              1    16.379   6.5020
5 'L-N USG/UREA'         1    16.793   5.4260
6 'H-N UREA'             1    12.442   5.5790
7 'H-N SCU'              1    14.942   6.2240
8 'H-N USG'              1    14.727   6.2170
9 'H-N USG/UREA'         1    12.643   5.8620
1 CONTROL                2    17.703   5.2310
2 'L-N UREA'             2    16.223   4.3710
.
.
.
3 'L-N SCU'              3    16.445   5.5760
4 'L-N USG'              3    16.001   6.1550
5 'L-N USG/UREA'         3    14.724   5.3610
6 'H-N UREA'             3    14.005   5.6480
7 'H-N SCU'              3    12.941   6.4490
8 'H-N USG'              3    16.418   7.1450
9 'H-N USG/UREA'         3    15.257   6.0220
```

To import this data file, follow these steps:

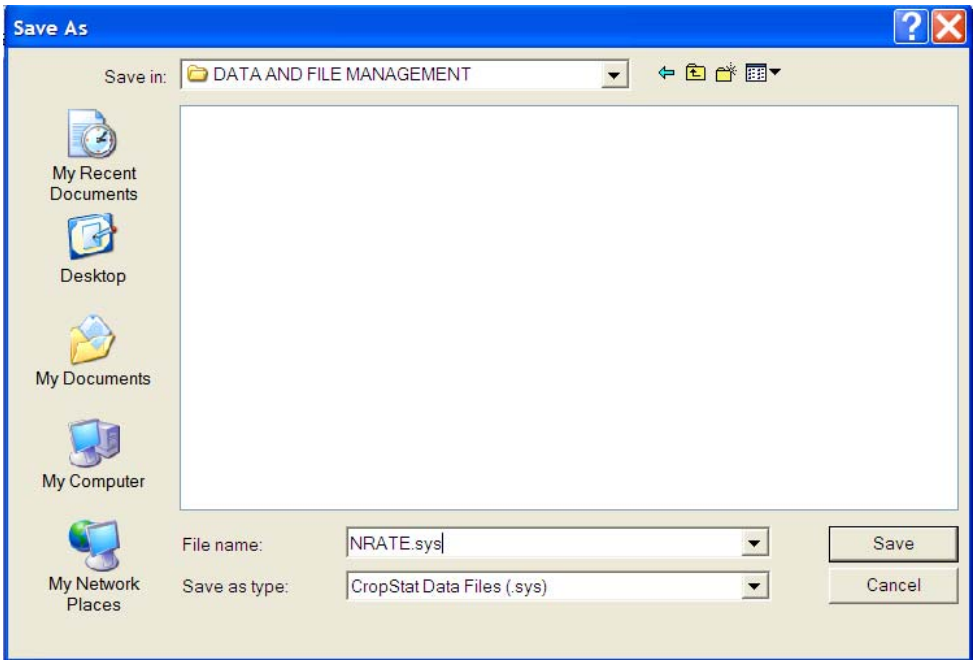
- Select **File⇒Import data⇒Text file** from the Data Editor Window.


CropStat Data Editor
File
Edit
Options
Window
Help
New... Ctrl+N
Open... Ctrl+O
Reopen...
Save... Ctrl+S
Save As ... Ctrl+A
Close Ctrl+E
Import data...
Export data
Print... Ctrl+P
Exit Alt+X
Excel Workbook
Text File
Database file...
ODBC source

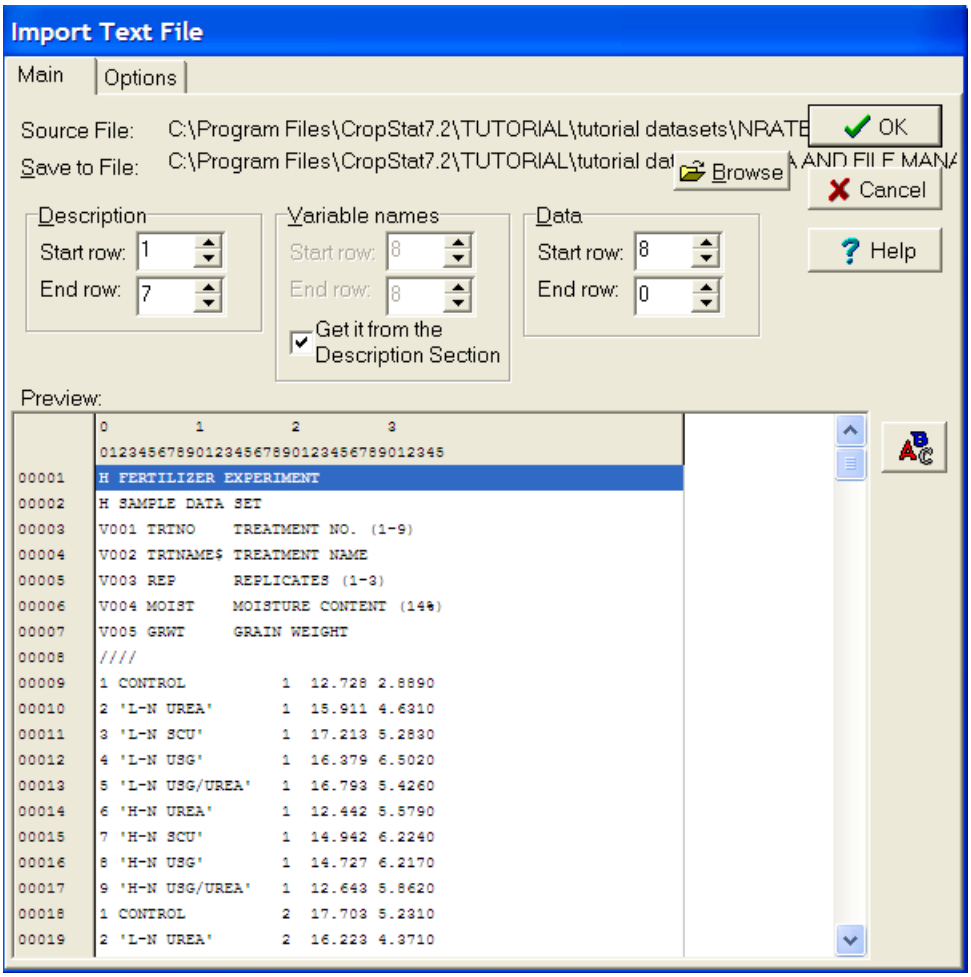
- The **Open** dialog box will prompt you to select the filename you want to open. To select the ASC file, click the down arrow under the **Files of type** and select **ASCII Files (asc)**. Click on the *Nrate.asc* and click **Open.**



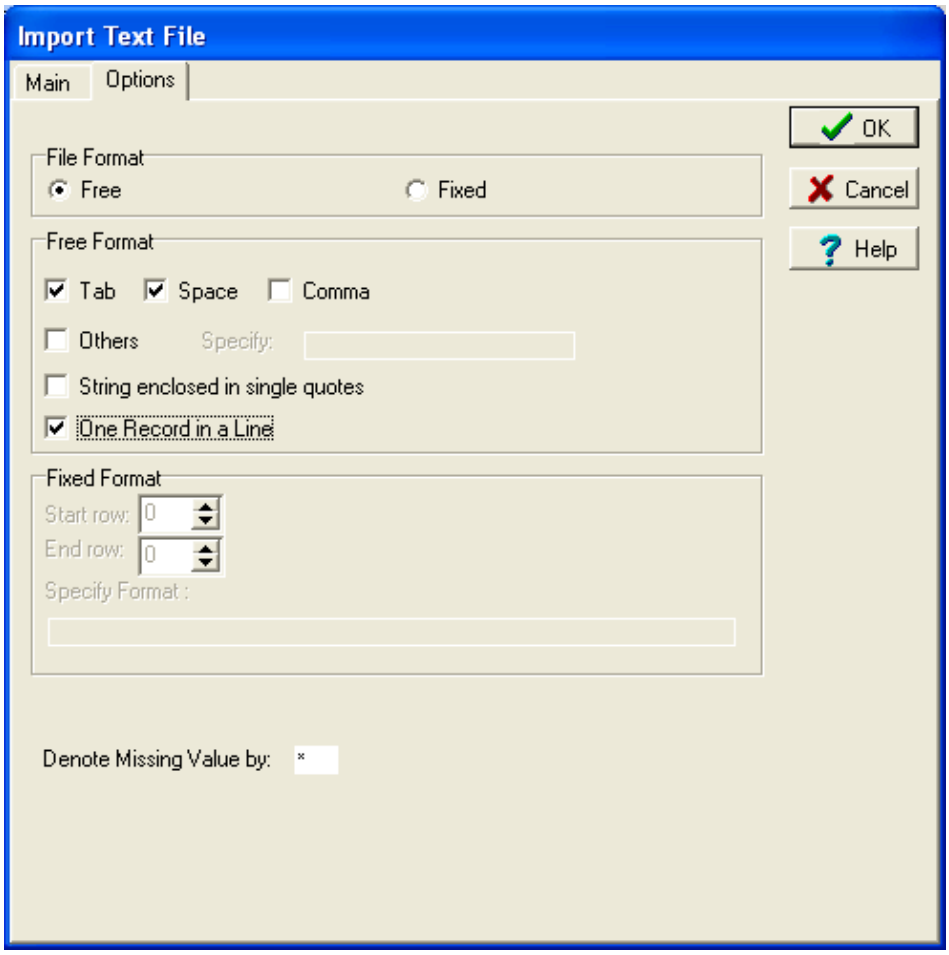
- The **Save As** dialog box will prompt you to supply the name of the SYS file. The default is the filename of the ASC file. Browse in the directory C:\MY CROPSTAT and create a new folder DATA AND FILE MANAGEMENT. Click **Save** to import file inside the created folder.



- The **Import Text File** dialog will appear. The default of the Description Rows is from the first line up to ////. Data row will start after //// up to the end (End row = 0 means last row with data). The ASC file is formatted in such a way that the rows for the descriptions and variable names and data are automatically supplied. You don't need to change any row number (i.e., changing any row number will lead to file reading error).



- You may edit the delimiter of the Free Format, the symbol used for missing values, or you may want to read a Fixed Format ASC file by clicking the **Options** menu from the Import Text File Window.

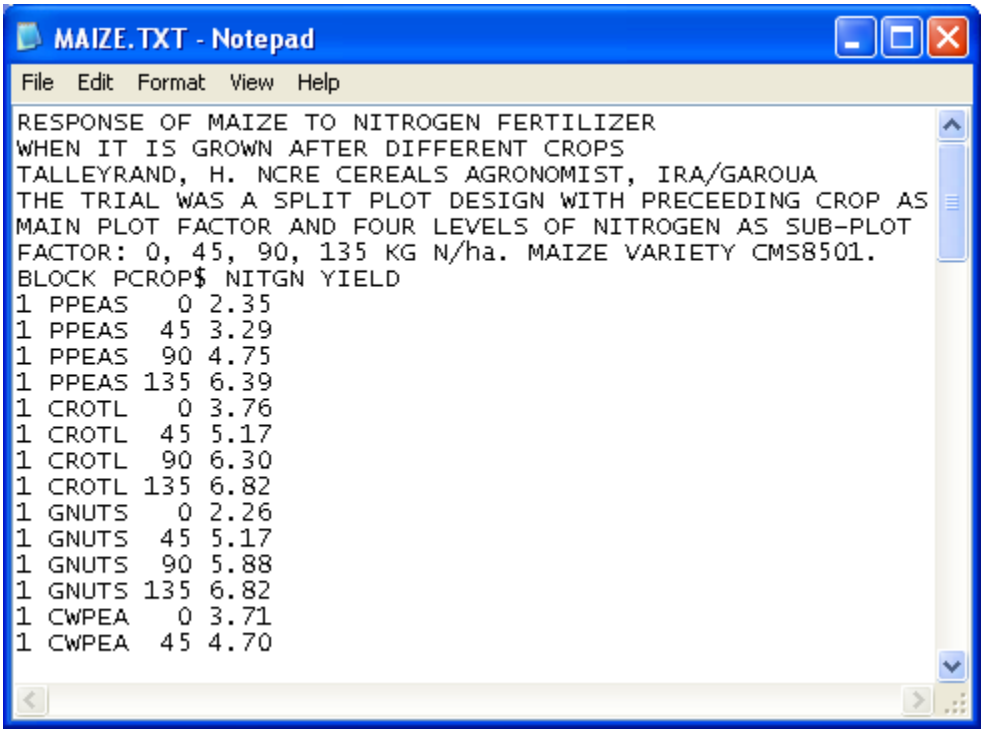


- Click **OK** to import the data.

- *Free format text file*

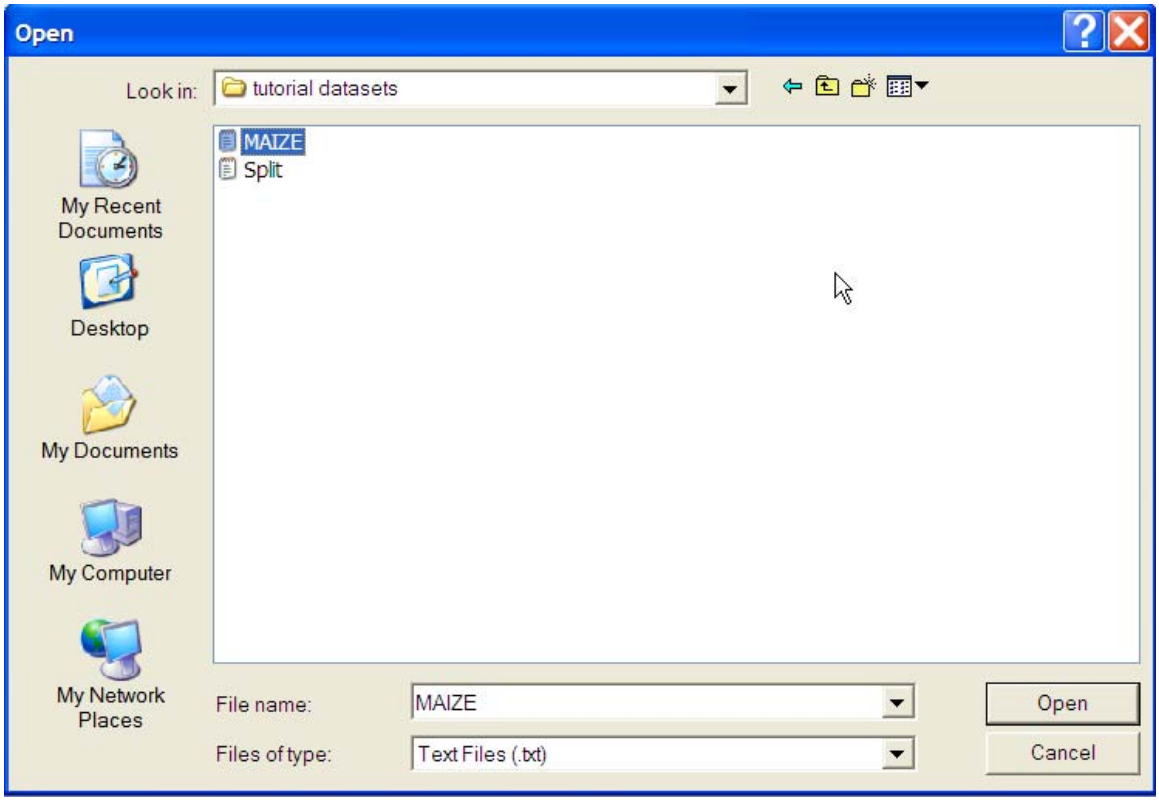
  - **Data Preparation**

    It is a good practice to put headings, description of the experiment, design of the experiment and other pertinent information about the data on the text file. Headings can be placed in one or more rows either at top or below the data, as long as the rows are adjacent. Variables can be placed in one or more adjacent rows, but should have a separate row from the heading and data. Data can be placed after the heading and variable names. Shown below is the text format of the Maize.txt. The Headings are at rows 1-8, variable names at row 10 and data at rows 11-90.

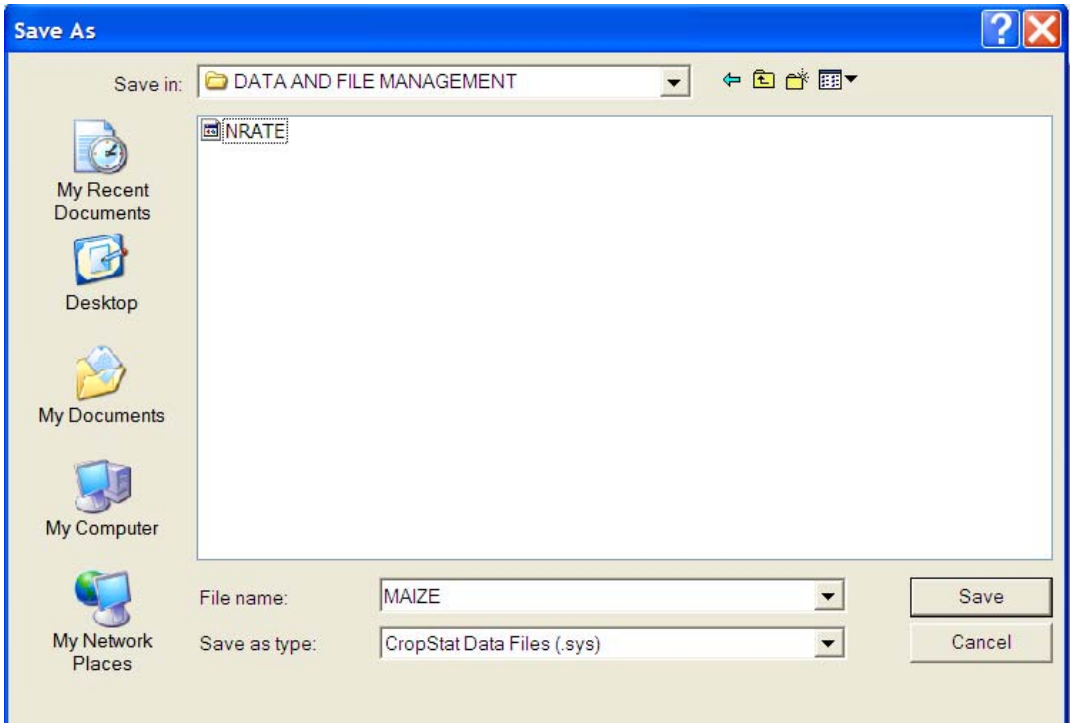


  - The file *MAIZE.TXT* in the *C:\PROGRAM FILES\CROPSTAT7.2\ TUTORIAL\ TUTORIAL DATASETS* folder will be used in this module. To import this data file, use the following steps:

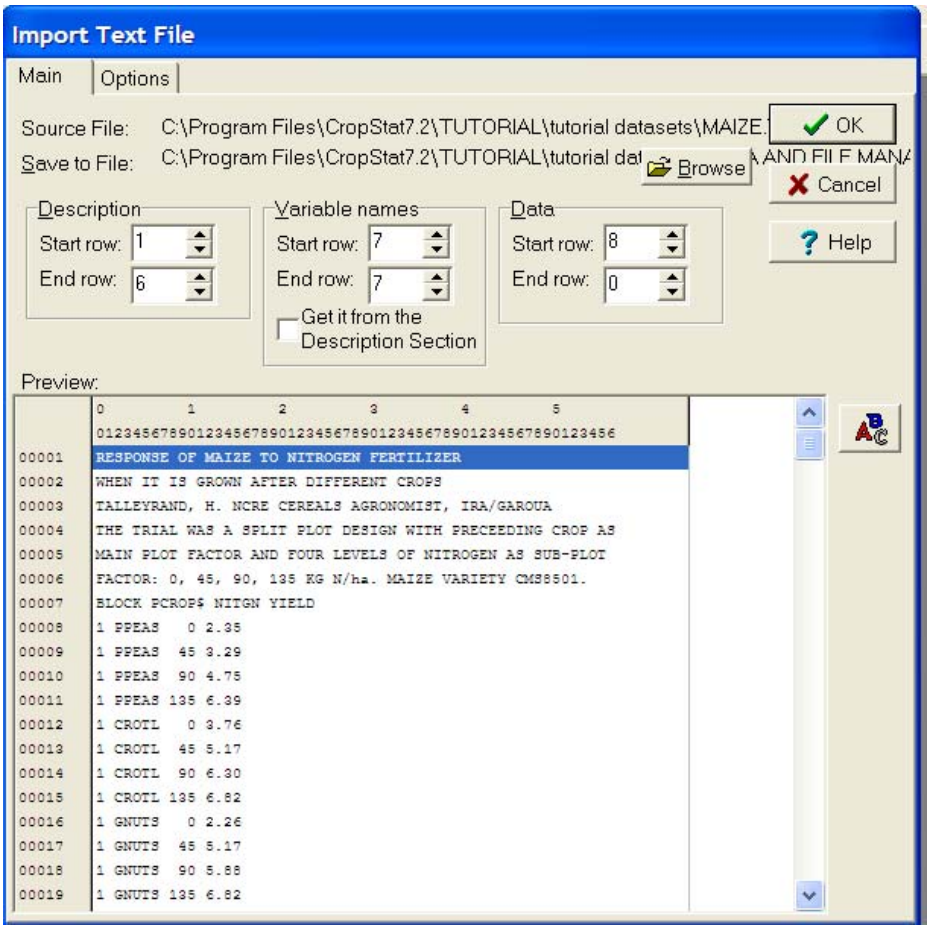
- Select **File⇒Import data⇒Text file** from the Data Editor Window.
- The **Open** dialog box will prompt you to select the filename you want to open. Click on the *Maize.txt* and click **Open.**



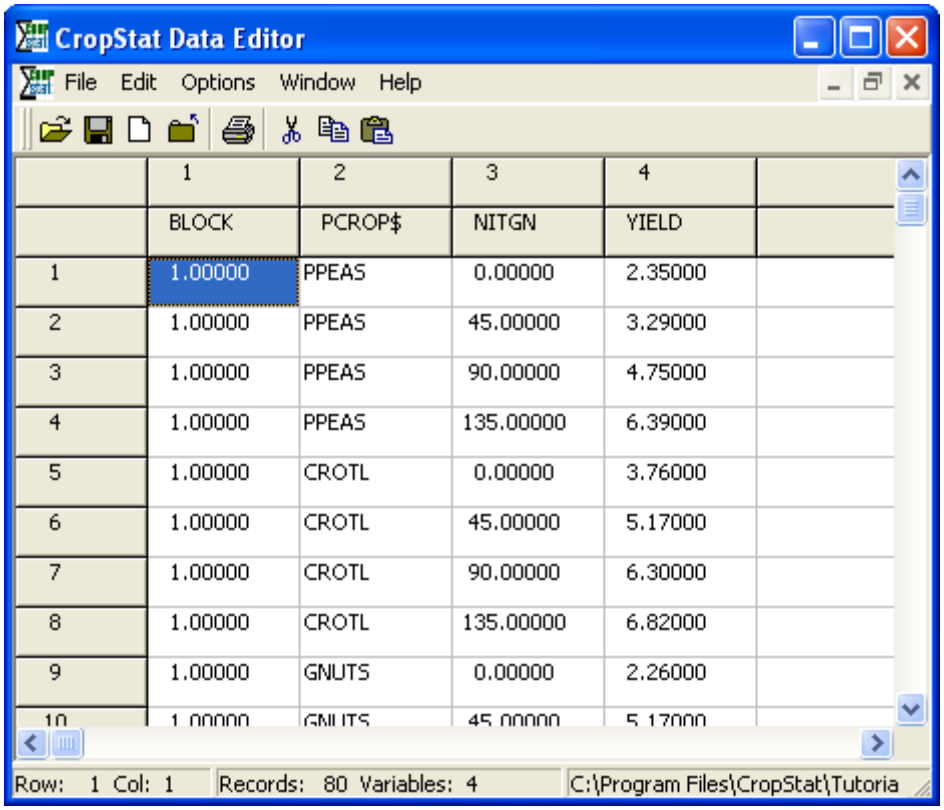
- The **Save As** dialog box will prompt you to supply the name of the SYS file. The default is the filename of the TXT file. Click **Save** to import the file. The MAIZE.SYS must be saved inside the created folder DATA AND FILE MANAGEMENT.



- The **Import Text File** dialog will appear. Unlike in the ASC file, all rows are set to zero. You have to set the Description, Variable names, and Data rows. The row number is located at left side of the Preview. You can see that Description starts at row 1 and ends at row 6, Variable names start at row 7 and end at row 7710 and Data start at row 8. A zero at the end row of the Data means that the program will read all the data.



- Click **OK** to import the data.



	1	2	3	4	
	BLOCK	PCROP$	NITGN	YIELD	
1	1.00000	PPEAS	0.00000	2.35000	
2	1.00000	PPEAS	45.00000	3.29000	
3	1.00000	PPEAS	90.00000	4.75000	
4	1.00000	PPEAS	135.00000	6.39000	
5	1.00000	CROTL	0.00000	3.76000	
6	1.00000	CROTL	45.00000	5.17000	
7	1.00000	CROTL	90.00000	6.30000	
8	1.00000	CROTL	135.00000	6.82000	
9	1.00000	GNUTS	0.00000	2.26000	
10	1.00000	GNUTS	45.00000	5.17000	

- **A DBase file**
  - To import a Dbase file (dbf), follow the steps in importing an excel file.

III. Data Transformation

A. New variables can be created from existing variables using the Data Editor. These new variables can be created using CropStat expressions. The CropStat data editor supports numeric, string, or logical (Boolean) expressions for transforming or creating data variables.

1. Numeric expressions

A Numeric Expression can incorporate the standard math operations (+, -, *, /, %) as well as the POWER operator, parenthesis, and unary minus.

The primitive elements of a numeric expression are numeric constants, numeric variables, and functions that return a numeric value. A numeric constant is a decimal number, unless a $ character is used as its prefix.

The numeric operators, in decreasing evaluation order, are:

- Parenthetical operation [that is, expression inside ( )]
- Unary minus (-) and arithmetic not
  - Unary minus (e.g. -VAR1)
  - Arithmetic not (e.g. NOT 0)
- POWER (e.g. VAR1 power 2)
- Multiplication operators
  - Multiply (*)
  - Divide (/)
  - Modulus (%) [Example: 10 % 3 = 1]
- Additive operators
  - Add (+)
  - Subtract (-)

Operators are left associative.

2. String expressions

String Constants are delimited either by single or double quotes; the matching quote is determined by the first quote. This makes it easy to create strings that include the "other" quote character. Strings can also include the quote character by doubling it.

e.g. - "This string has a single quote right here: ' "

e.g. - 'And this one has a double quote here: " '

3. Logical (Boolean) expressions

Logical expressions return a value - TRUE or FALSE.

The supported logical operators are AND, OR, XOR, NOT, and parenthesis.

The operators in decreasing evaluation order are:

- Parenthesis
- NOT (~)
- AND
- XOR
- OR

4. Libraries

- **Math Library**

  The Math Library includes the following constants, functions, and procedures:

  Constants

Pi	- 3.14....
MISSING	- data numeric missing value = -1e36

  Functions:

  a) Exp(*argument*)

  The EXP function raises the constant *e*, approximately given by 2.71828, to the power supplied by the numeric *argument*.

  e.g. exp(002) will return a value of 7.38905

  b) Sin(*argument*)

  The SIN function returns the sine of the numeric *argument*.

  e.g. sin(0.5) will return a value of 0.47943.

  c) Random(*seed*)

  The RANDOM function returns a uniform random number within the range 0 to the *seed* value.

  e.g. random(100) will return a value between 0 and 100.

  d) Cos(*argument*)

  The COS function returns the cosine of the numeric *argument*.

  e.g. cos(0.5) will return a value of 0.87758.

e) Ln(*argument*)

The LN function returns the natural logarithm of the numeric *argument*.

e.g. ln(1) will return a value of 0.

f) Log10(*argument*)

The LOG function returns the common logarithm of the numeric *argument*.

e.g. log(52) will return a value of 1.716003.

g) Abs(*argument*)

The ABS function returns a nonnegative number equal in magnitude to that of the numeric *argument*.

e.g. abs(-48) will return a value of 48.

h) Arctan(*argument*)

The ARCTAN function returns the arctangent (inverse tangent) of the numeric *argument*. The value returned is in radians.

e.g. arctan(0.5) will return a value of 0.46365.

i) Sqr(*argument*)

The SQR function returns the square of the numeric *argument*.

e.g. sqr(4) will return a value of 16.

j) Sqrt(*argument*)

The SQRT function returns the square root of the *argument*.

e.g. sqrt(4) will return a value of 2.

k) Tan(*argument*)

The TAN function returns the tangent of the numeric *argument*.

e.g. tan(0.5) will return a value of 0.54630.

l) Cotan(*argument*)

The COTAN function returns the cotangent of the numeric *argument*.

e.g. cotan(0.5) will return a value of 1.83049.

m) Arcsin(*argument*)

The ARCSIN function returns the arcsine of the numeric *argument*.

e.g. arcsin(0.5) will return a value of 0.52360.

n) Arccos(*argument*)
The ARCCOS function returns the arccosine of the numeric *argument*.
e.g. arccos(0.5) will return a value of 1.04720.

o) Trunc(*argument*)
The TRUNC function returns the integer part of the numeric *argument*.
e.g. trunc(5.6789) will return a value of 5.

p) Round(*argument*)
The ROUND function returns a value rounded to the nearest integer of the numeric *argument*.
e.g. round(5.6789) will return a value of 6.

- **Character and String Library**

The Standard String Library supports the following procedures and functions:

Constants:

MISSING$ - data string missing value = ""

5. Functions*:*

a) Chr(*argument*)
The CHR function returns the character equivalent of the numeric *argument*.
e.g. chr(65) will return the value of 'A'.

b) Ord(*argument*)
The ORD function returns the ordinal number of the ASCII character *argument*.
e.g. ord('A') will return the value of 65.

c) Copy(*argument, position, n*)
The COPY function returns a portion of an expression you specify in *argument*. The portion begins with the character specified by *position* and is the number of characters specified by *n*.
e.g. copy ('hello', 2, 3) will return the value 'ell'.

d) Length(*argument*)
The LENGTH function returns the length of the *argument*.
e.g. length('ABCDEFG') will return the value of 7.

e) Insert(*argument, string, position*)

The INSERT function inserts an expression you specify in *string* into the *argument* starting at the specified *position*.

e.g. insert('ho', 'ell', 2) will return the value 'hello'.

f) Delete(*argument, length, position*)

The DELETE function deletes a portion of an expression you specify in *argument*. The portion begins with the character specified by *position* and is the number of characters specified by *n*.

e.g. delete('hello', 3, 2) will return the value 'ho'.

g) Pos(*argument, string*)

The POS function returns the position of the *argument* from the specified string. If *argument* is not found, a value of 0 is returned.

e.g. pos('ell', 'hello') will return the value 2.

h) Trim(*string*)

The TRIM function deletes all leading and trailing spaces from the *string*.

e.g. trim(‘ hello ‘) will return a value ‘hello’.

i) TrimRight(*string*)

The TRIMRIGHT function deletes all trailing spaces from the *string*.

e.g. trim(‘ hello ‘) will return a value ‘ hello’.

j) TrimLeft(*string*)

The TRIMLEFT function deletes all leading spaces from the *string*.

e.g. trim(‘ hello ‘) will return a value ‘hello ’.

6. Assignment statement

The ASSIGNMENT statement assigns a value to a variable.

Syntax

<Variable> := expression

Examples

a) VAR1 : = 2*4
b) VAR2$: = _MISSING$_

7. If-Then-Else statement

The IF statement is used to perform an operation according to a set of rules that is correct (evaluated to be TRUE) when the IF statement is executed.

Syntax

IF (*Conditional expression*) [THEN]
    ... commands to do if conditional expression is evaluated to be TRUE
[ELSE
    ... commands to do if conditional expression is evaluated to be FALSE]
ENDIF

Where Conditional expression is a logical expression that can be evaluated to a Boolean value.

Every IF statement must end with the ENDIF keyword. An optional ELSE keyword defines the end of the statement block that should be evaluated when the conditional expression is evaluated to be TRUE, and the start of the statement block that should be executed if the conditional expression is evaluated to be FALSE.

Examples

a)
```
if (VAR1 =1) then
    VAR2 := sin(VAR3)
endif
```

b)
```
if (VAR1 =1) then
    VAR2 := sin(VAR3);
    VAR4 := VAR2*VAR5
endif
```

c)
```
if (VAR1 =1) then
    VAR2 := sin(VAR3)
else
     if (VAR1 =2) then
         VAR2 := 0
     endif
endif
```

d)
```
if ((VAR1 = _MISSING_) or (VAR2 = _MISSING_)) then
     VAR3 := _MISSING_
else
     VAR3 := VAR1/VAR2
endif
```

e)
```
 if (DOSE$ ='High       ') then
```

```
            DAM1  := DAM -10
        endif
f)      if (Trim(DOSE$)='High') then
                DAM1 := DAM-10
        endif
```

Note

i. Token inside [] is optional (not required for syntax).

ii. The conditional expression should be inside the parenthesis ( ).

iii. In conditional expressions, the variable name and the conditional symbols (<,>,<>,>=,=<) should be separated by at least one space. E.g. if (VAR1 =1)

iv. You can use only one ELSE in an IF-THEN statement. But IF-THEN can be nested within ELSE. If there are many conditions, then the user should consider using the SWITCH statement which is described.

v. Commands are not case-sensitive but variable names should be capitalized.

vi. Examples given in (e) and in (f) performs the same operation. It should be noted that in example (e) you need to enter char values as a 12 byte right padded constant while in (f) the TRIM function was used.

8. Switch .. EndSwitch statement

The SWITCH Statement is used to choose one option from a list of options, based on an expression that is evaluated first.

Syntax

```
SWITCH expression [OF]
      CASE expression1 : ....
      ENDCASE
      [CASE expression2 : ....
       ENDCASE ...]
      [ELSE ....
       ENDCASE]
ENDSWITCH
```

Where expression is the expression that expression1, expression2 … will be tested against. When a match between expressionN to expression is found, the statements specified until the ENDCASE keyword, are executed, and execution continues after the ENDSWITCH keyword.

If no expression from expression1 ... expressionN was matched to expression, and an ELSE option is specified, the statements between the ELSE and the ENDCASE keywords will be executed.

Example

```
(equivalent to example c)
switch VAR1 of
    case 1 : VAR2 := sin(VAR3)
    endcase
    case 2 : VAR2 := 0
    endcase
endswitch
```

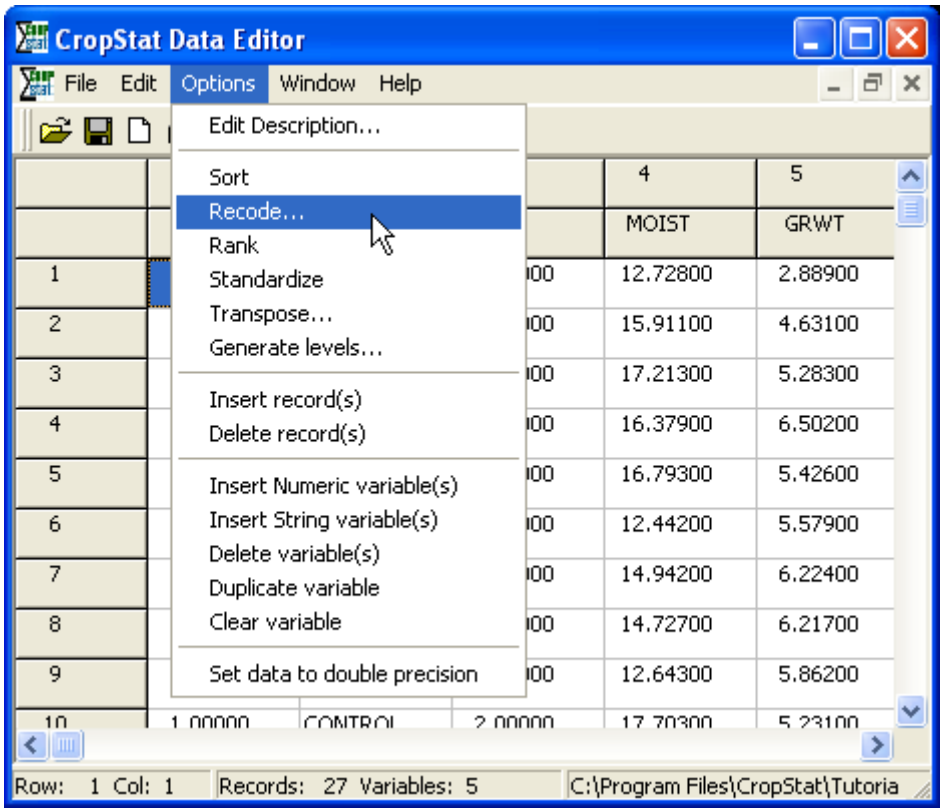
Note: Switch statements cannot be nested.

B. Creating new variables using existing variables

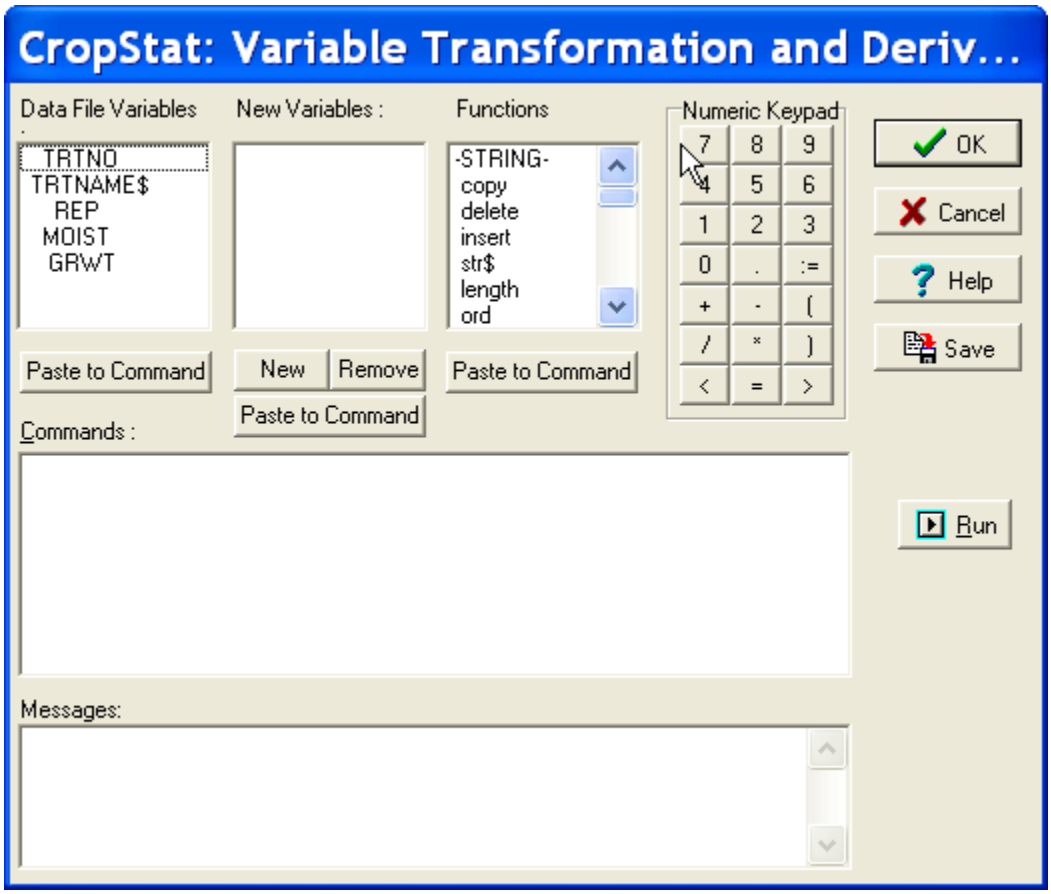
Using the *NRATE.SYS*, we would like to create a new variate called *GYIELD*, where

GYIELD = GRWT*(100 - MOIST)/86

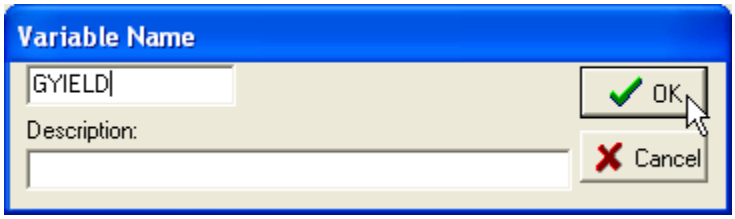
- Select **File ⇒ Open** from the Data Editor.
- Open or Reopen the NRATE.SYS from the *CROPSTAT TUTORIAL* folder inside the subfolder TUTORIAL DATASETS.
- Click **Options ⇒ Recode**.



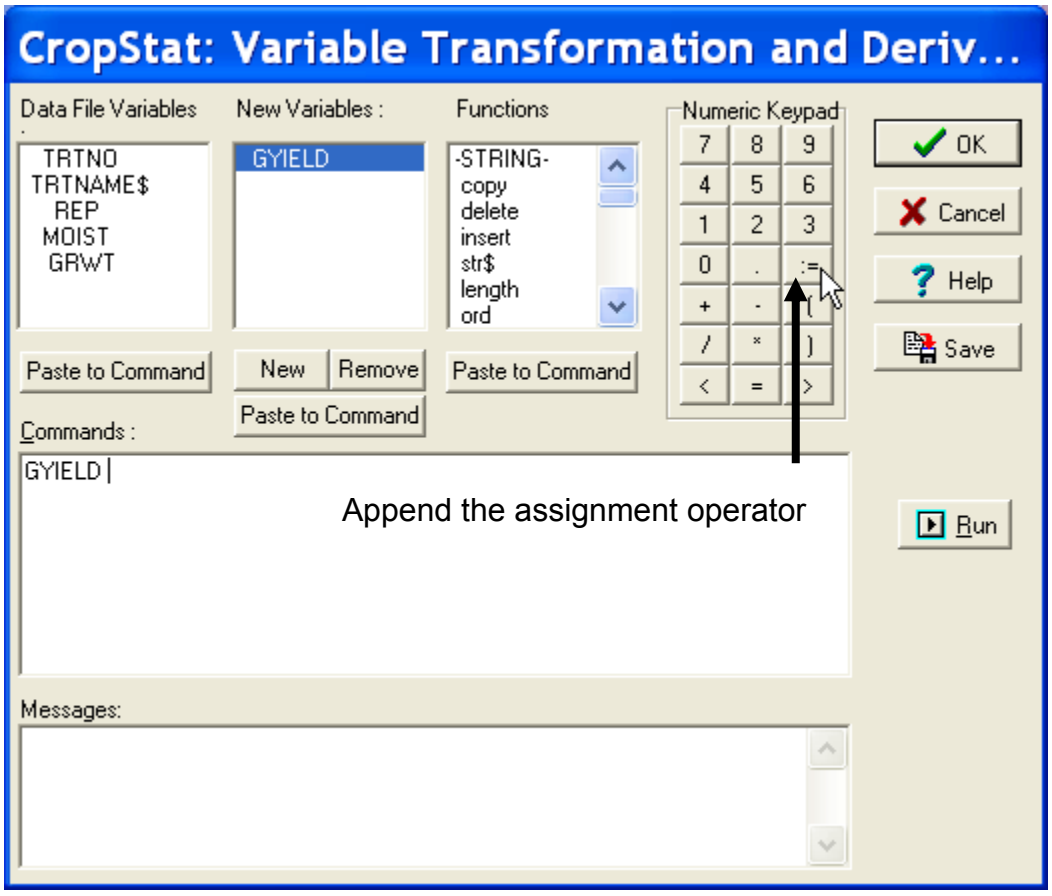
- The **Recode** dialog box will appear.



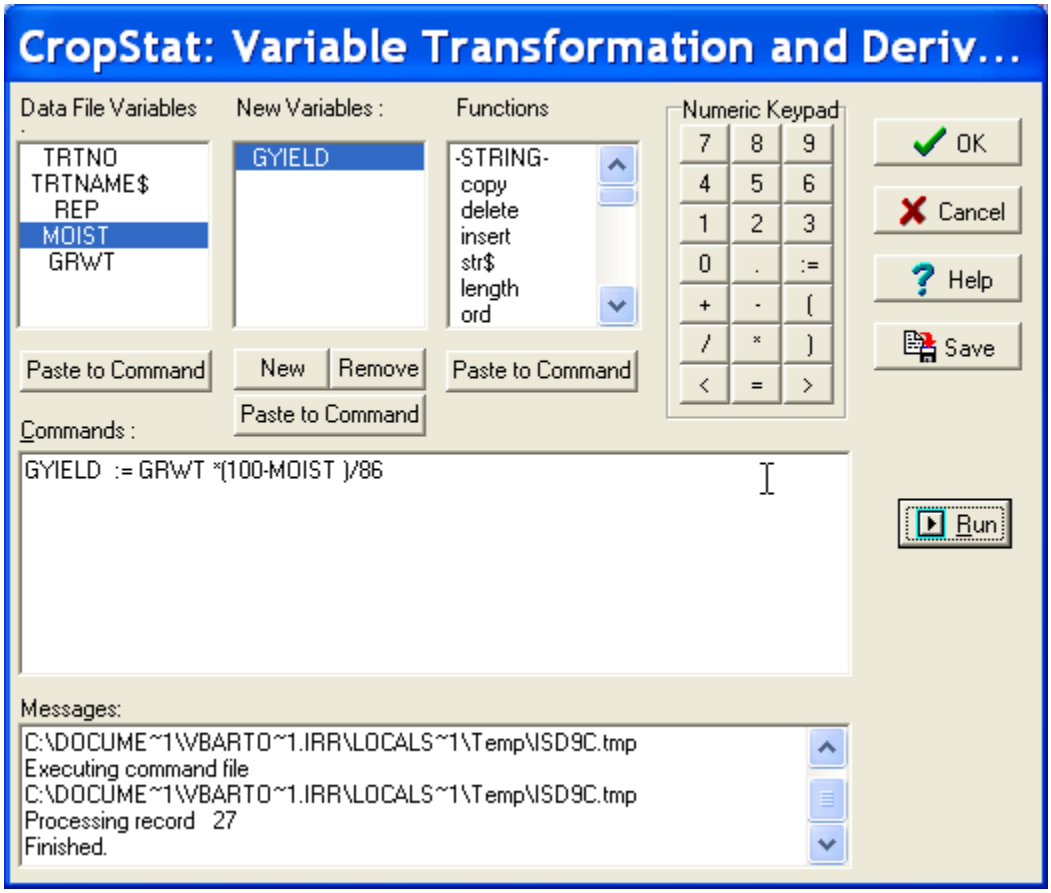
- Select **New** button under **New Variables** list. Type *GYIELD* as variable name. Click **OK**.



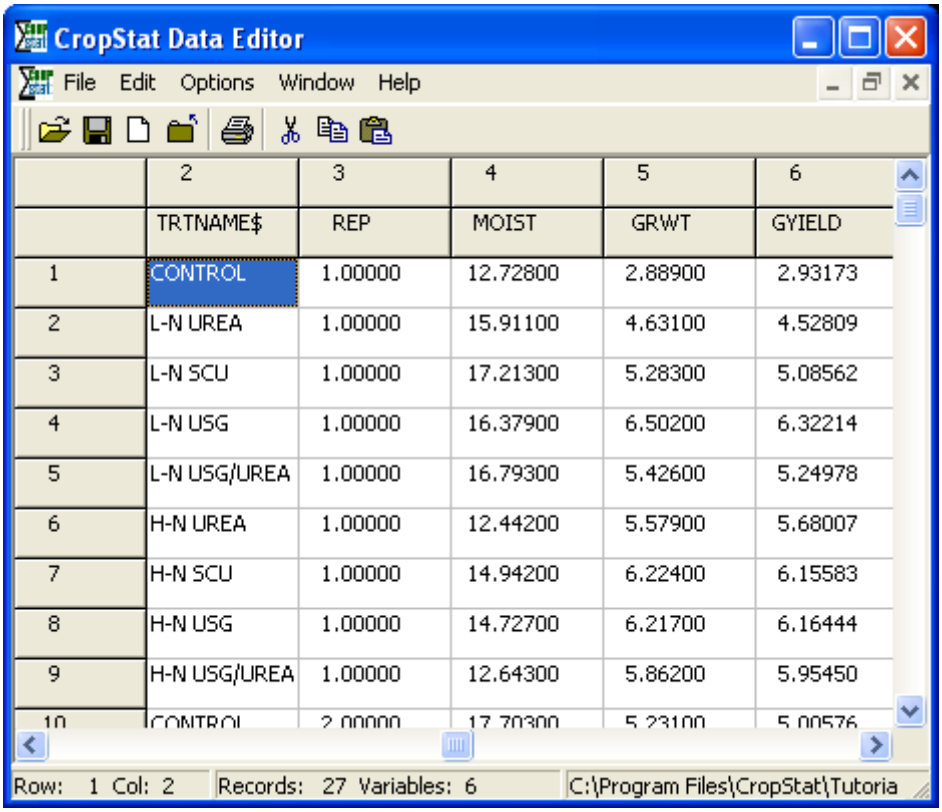
- Click *GYIELD* from **New Variables** list.
- Append **:=** to the list using the numeric keypad or by typing it on the keyboard.



- Select *GRWT* from **Data File Variables** list.
- Append ***(100-*** using the numeric keypad or by typing it on the keyboard.
- Select *MOIST* from **Data File Variables** list.
- Append ***)/86*** using the numeric keypad or by typing it on the keyboard.
- Click the **Run** button**.** If you see the message **Finished** at the lower box, it means that the recoding is successful. However, if you see an Error message, it means that you have to check the Commands. Click Run again.



- Click **OK** to close the dialog box.
- The result will be variable *GYIELD* with values equal to GRWT*(100-MOIST)/86.
- Select **File ⇒ Save As.** Click the **Save in** box and go inside the *C:\MY CROPSTAT\DATA FILE AND MANAGEMENT* and save *NRATE.SYS* to effect the changes.

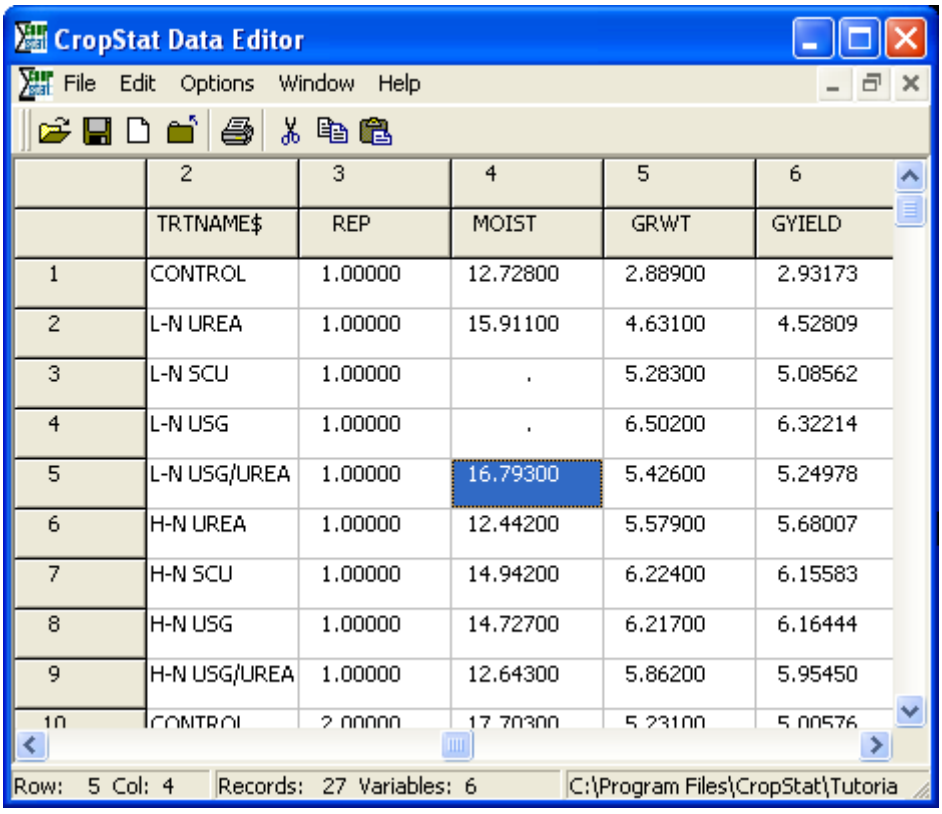

CropStat Data Editor
File Edit Options Window Help
2 3 4 5 6
TRTNAME$ REP MOIST GRWT GYIELD
1 CONTROL 1.00000 12.72800 2.88900 2.93173
2 L-N UREA 1.00000 15.91100 4.63100 4.52809
3 L-N SCU 1.00000 17.21300 5.28300 5.08562
4 L-N USG 1.00000 16.37900 6.50200 6.32214
5 L-N USG/UREA 1.00000 16.79300 5.42600 5.24978
6 H-N UREA 1.00000 12.44200 5.57900 5.68007
7 H-N SCU 1.00000 14.94200 6.22400 6.15583
8 H-N USG 1.00000 14.72700 6.21700 6.16444
9 H-N USG/UREA 1.00000 12.64300 5.86200 5.95450
10 CONTROL 2.00000 17.70300 5.23100 5.00576
Row: 1 Col: 2
Records: 27 Variables: 6
C:\Program Files\CropStat\Tutoria

C. Creating variables using existing variables with missing values

- To illustrate data transformation with missing values, reopen the NRATE.SYS and set to missing, by replacing the value with a space (point on the specific cell and press the spacebar, the space will be replaced by a dot (.) once you click on the other cell), the following observations:

  1. Variable *MOIST* of L-N SCU Rep 1

  2. Variables *MOIST* of L-N USG Rep 1



	2	3	4	5	6
	TRTNAME$	REP	MOIST	GRWT	GYIELD
1	CONTROL	1.00000	12.72800	2.88900	2.93173
2	L-N UREA	1.00000	15.91100	4.63100	4.52809
3	L-N SCU	1.00000	.	5.28300	5.08562
4	L-N USG	1.00000	.	6.50200	6.32214
5	L-N USG/UREA	1.00000	16.79300	5.42600	5.24978
6	H-N UREA	1.00000	12.44200	5.57900	5.68007
7	H-N SCU	1.00000	14.94200	6.22400	6.15583
8	H-N USG	1.00000	14.72700	6.21700	6.16444
9	H-N USG/UREA	1.00000	12.64300	5.86200	5.95450
10	CONTROL	2.00000	17.70300	5.23100	5.00576

- Data Editor will not automatically recompute the *GYIELD* though you change the value of *MOIST* or *GRWT*.

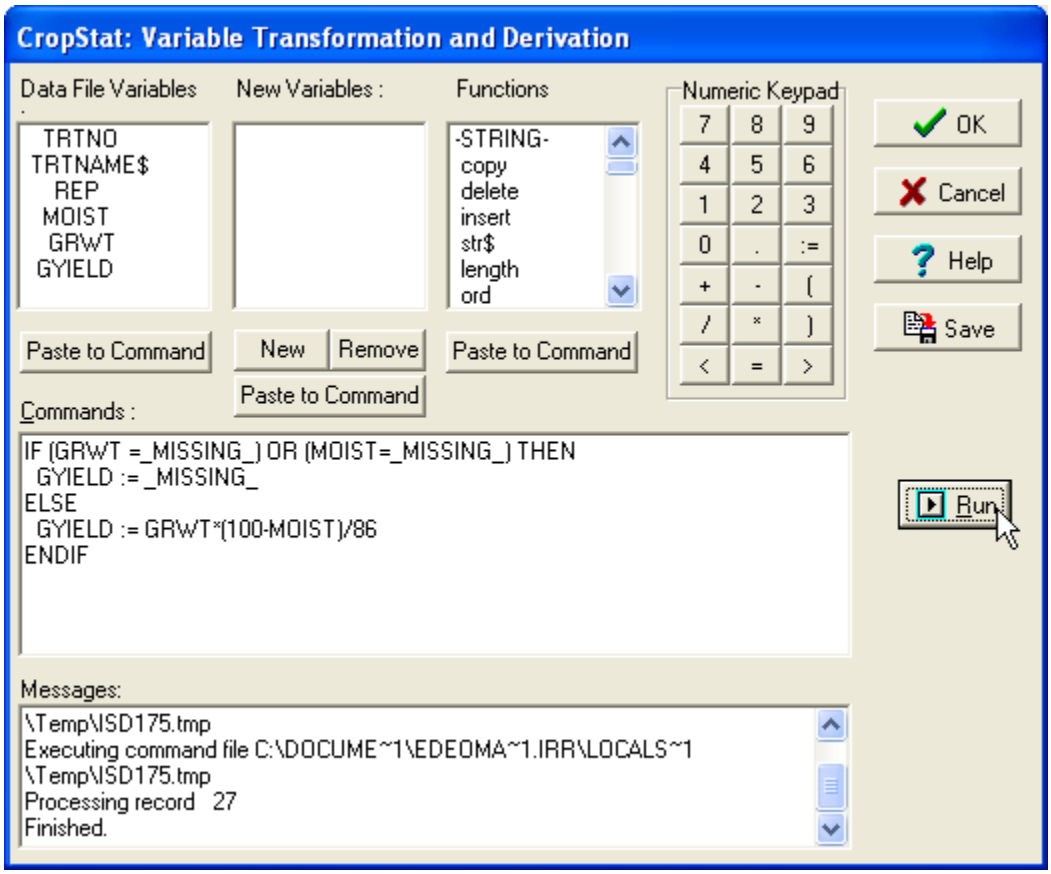
- To recompute, click the **Options** and **Recode**.

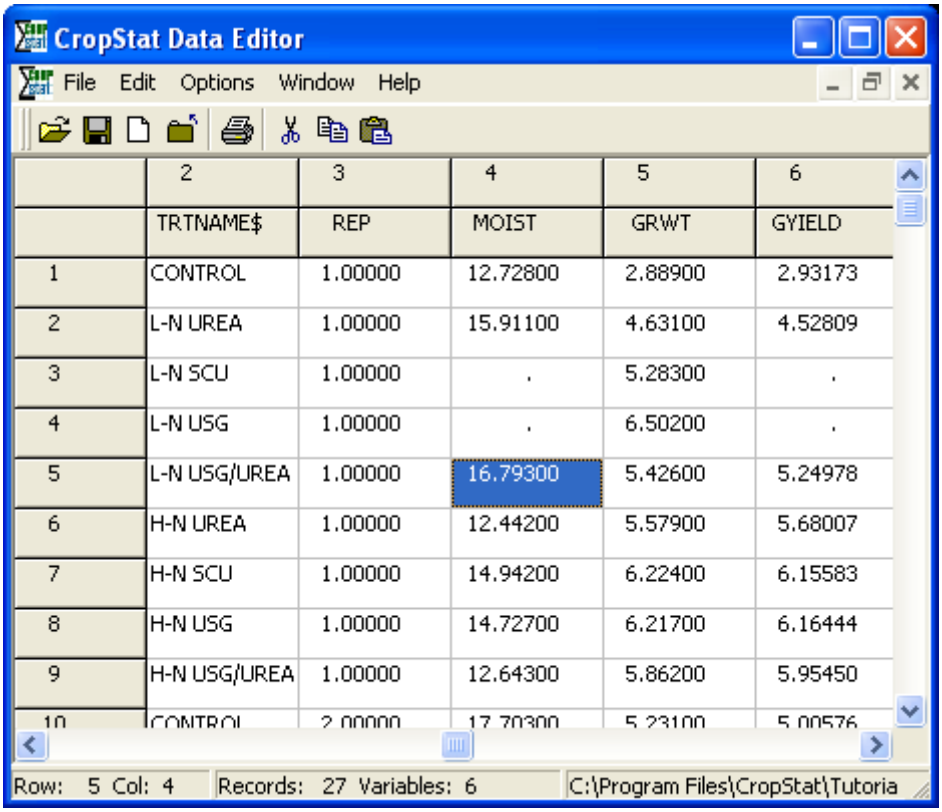
- Since you set two missing values, a statement should be added on the Command box to set the *GYIELD* as missing if *MOIST* or *GRWT* is missing. Failing to do this will give a wrong value of *GYIELD*. If both variables are missing, *GYIELD* will be automatically set to missing.

- In the **Commands** box, type the following:

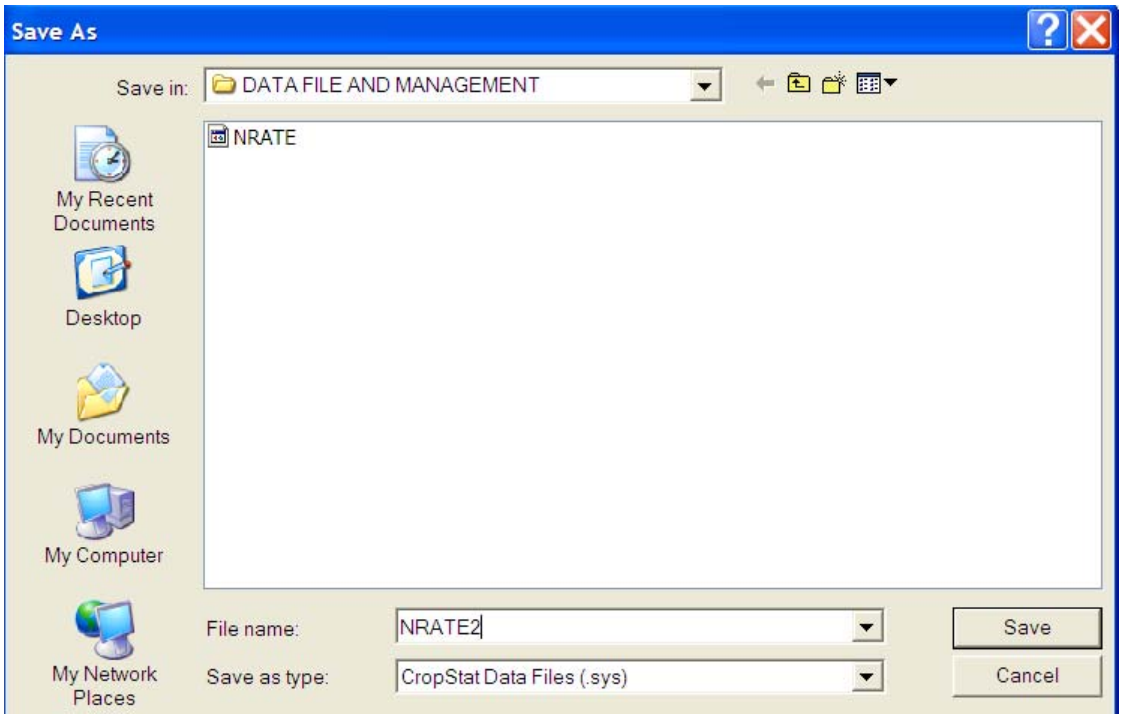
```
IF (GRWT =  _MISSING_) OR (MOIST =  _MISSING_) THEN
        GYIELD := _MISSING_
ELSE
        GYIELD := GRWT*(100-MOIST)/86
ENDIF
```

- Click **Save** to save the commands.
- Click **Run** to recompute and click **OK.**



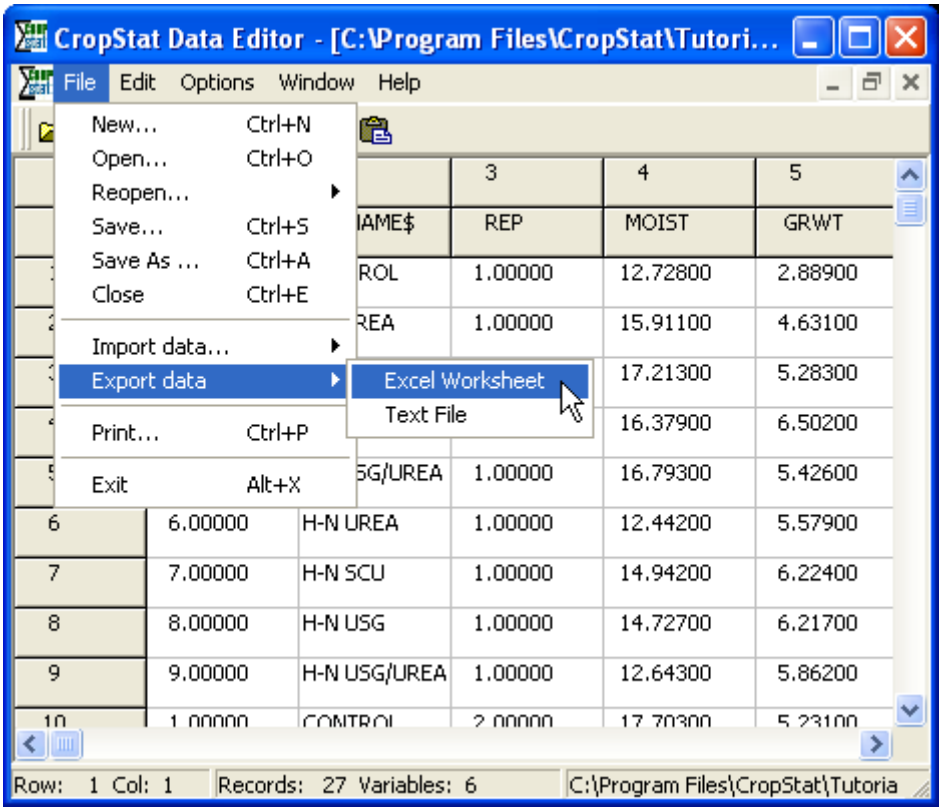


- Select on **File ⇒ Save as** NRATE2.SYS. Click the **Save in** box and go inside your working folder *C:\MY CROPSTAT\DATA FILE AND MANAGEMENT*.

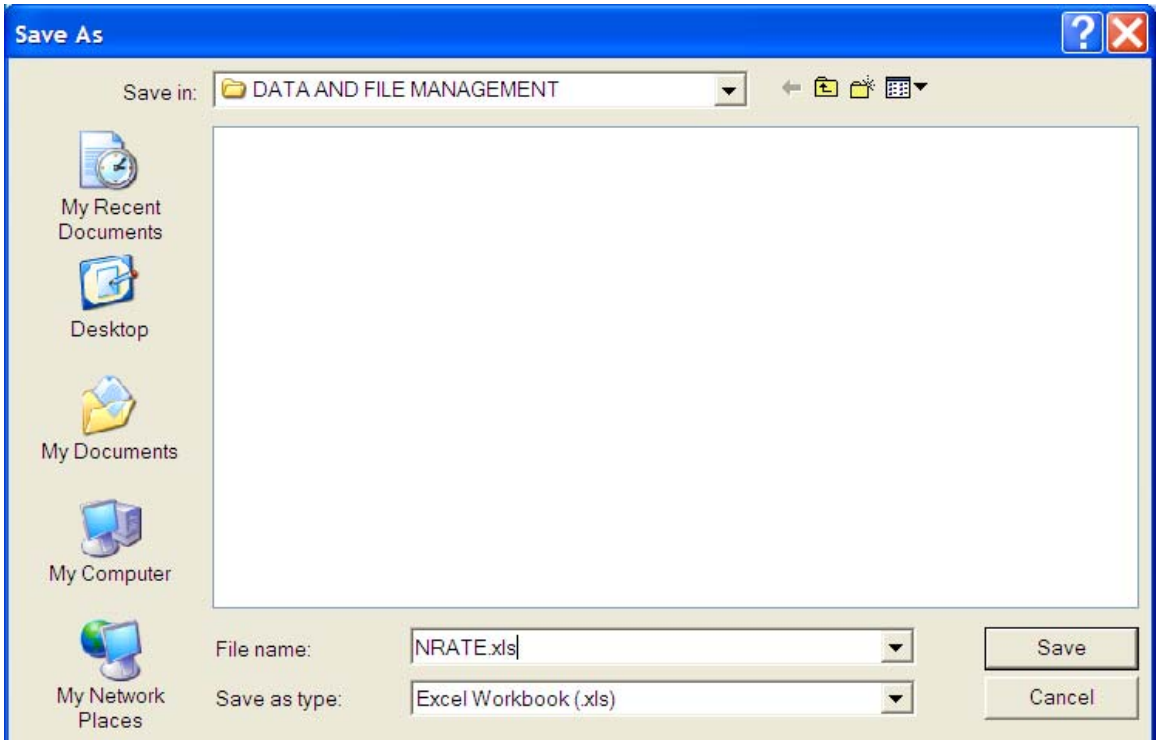


IV. Data Export

- SYS file can be exported (saved) to an Excel Worksheet or a Text File.
- Only an open SYS file can be exported.
- **Open** or **Reopen the** NRATE.SYS inside your working folder *C:\MY CROPSTAT\DATA FILE AND MANAGEMENT*. Click **File ⇒ Export Data⇒Excel Worksheet (or Text File).**



- The **Save As** dialog box will prompt you to supply the filename of an Excel Workbook. Type NRATE and click **Save**.


Save As
Save in:
DATA AND FILE MANAGEMENT
My Recent Documents
Desktop
My Documents
My Computer
My Network Places
File name:
NRATE.xls
Save as type:
Excel Workbook (.xls)
Save
Cancel

Exercise 2

CropStat: Data and File Management

Yield data (kg/ha) from a factorial experiment in split plot design is shown below.

Variety	Replication 1	2	3	4
	Infestation level 1			
1	3209	3007	3197	4280
2	3044	2840	3014	3223
3	3877	4129	3950	4880
4	3688	3873	3229	4047
	Infestation level 2			
1	3733	4113	3654	3765
2	3888	3559	3757	4006
3	3767	4624	4212	4411
4	3334	3215	3122	3688
	Infestation level 3			
1	3238	3293	3880	3868
2	2930	3964	3909	3886
3	2837	3546	3917	3879
4	2452	2495	2174	3693

2.1 Enter the data using the Data Editor. Save it in file EXER2.

2.2 RECODE the variates YIELD into SYIELD = SQRT(YIELD).

2.3 Describe SYIELD using the Summary Statistics under the Analysis Menu.

DATA DESCRIPTION AND VERIFICATION

At the end of the tutorial, the user should be able to generate summary statistics and check for outliers in the dataset.

I. Data Editor

Sample Problem

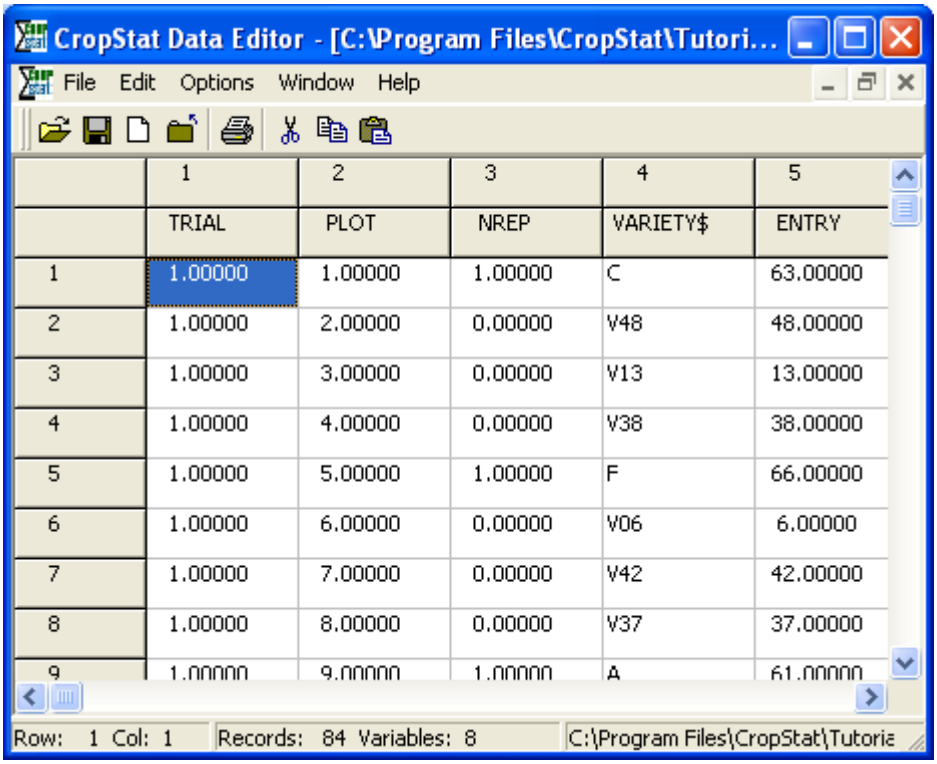
Although there is no substitute for dual data encoding or proof reading to verify data entry, summary statistics and data sorting can be used to check for outlying values. These outliers however, may not be data entry errors. View data in the **Data Editor** and use the sort facility to check for outlying values.

Steps

- Select **Data Editor** from the **Window** menu on the main CropStat Window.

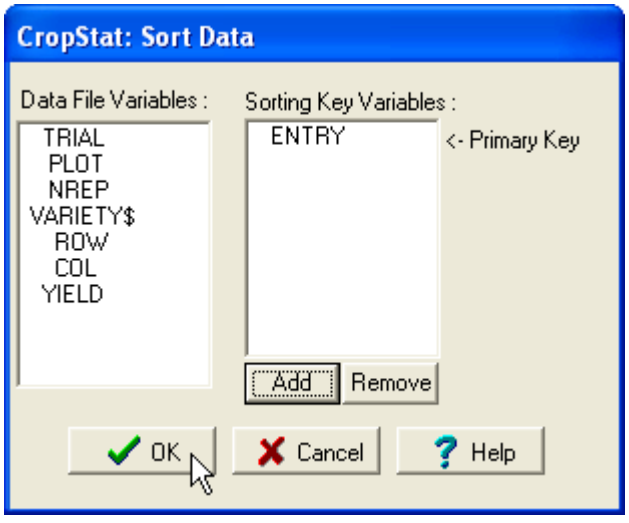
- Select **File ⇒ Open.** Navigate to the *CROPSTAT TUTORIAL* folder and select *AUGRB.SYS* found in the *TUTORIAL DATASETS* folder. Click **Open** button. You should see the following on display.

- Select **File ⇒ Save As.** Click the **Save in** box and go inside the directory *C:\MY CROPSTAT* and create a subfolder DATA DESCRIPTION AND VERIFICATION. Save AUGRB.SYS inside this folder.

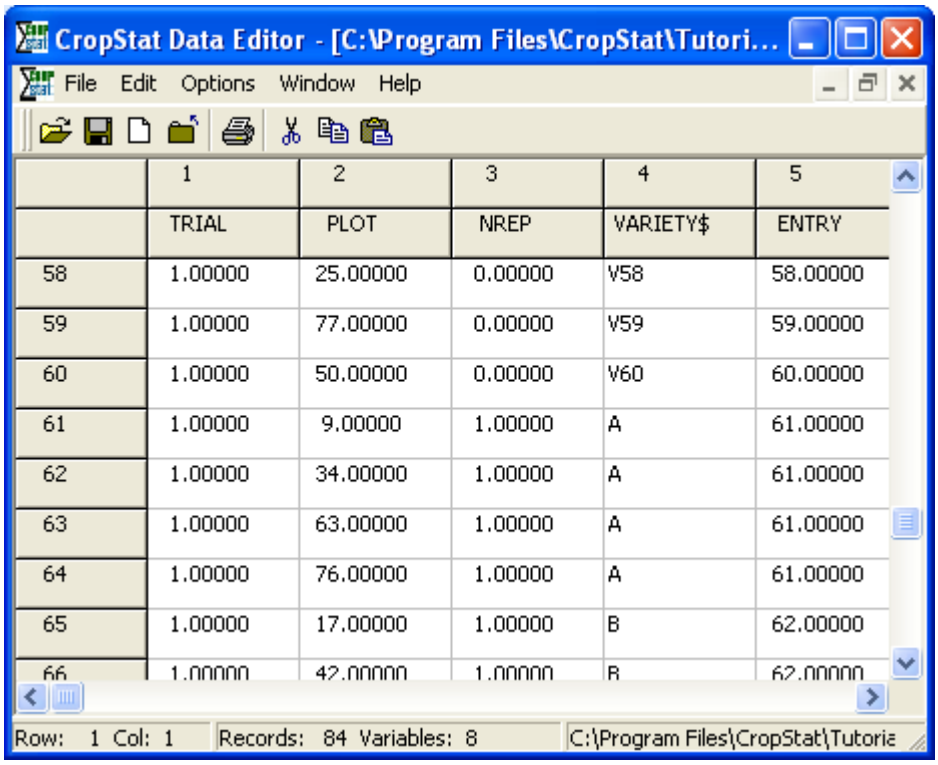


- Notice that the observations are arranged according to plot number. The next task is to sort the dataset in terms of the values of *ENTRY*. To do this, select **Options ⇒ Sort.** The **Sort Data** dialog appears.

- Specify the **Sorting Key Variable** by selecting *ENTRY* in the **Data File Variables** list and clicking the **Add** button. Click **OK**.

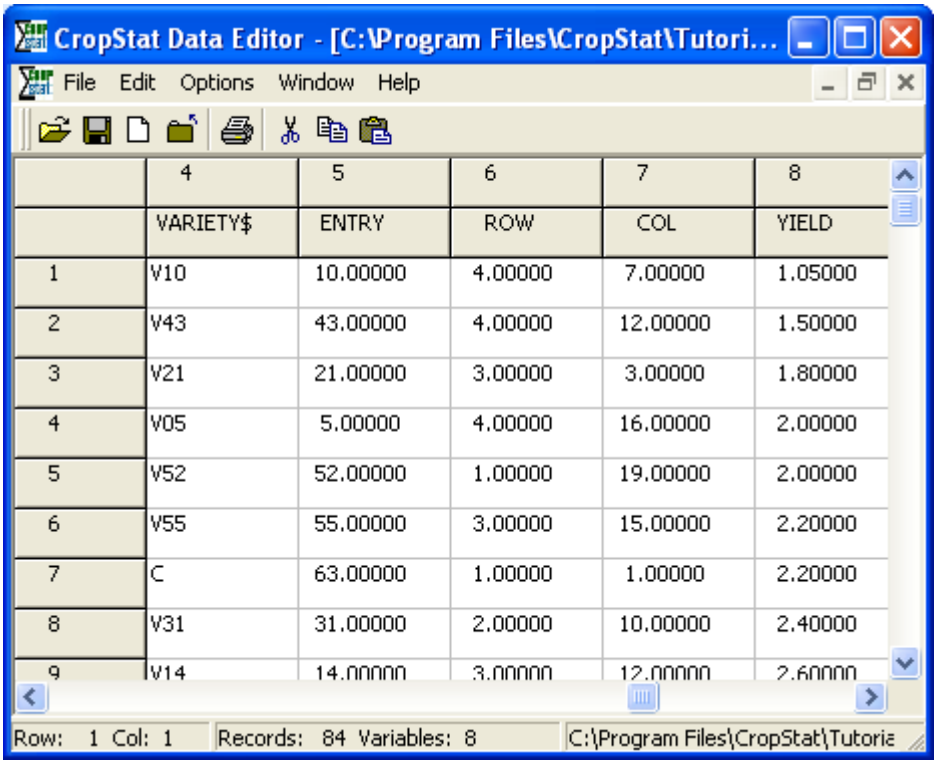


- Check the data again and notice that the observations are now arranged according to entry number. Note that all entries 1 to 60 and A to F are present and that only A to F are replicated.



	1	2	3	4	5
	TRIAL	PLOT	NREP	VARIETY$	ENTRY
58	1.00000	25.00000	0.00000	V58	58.00000
59	1.00000	77.00000	0.00000	V59	59.00000
60	1.00000	50.00000	0.00000	V60	60.00000
61	1.00000	9.00000	1.00000	A	61.00000
62	1.00000	34.00000	1.00000	A	61.00000
63	1.00000	63.00000	1.00000	A	61.00000
64	1.00000	76.00000	1.00000	A	61.00000
65	1.00000	17.00000	1.00000	B	62.00000
66	1.00000	42.00000	1.00000	B	62.00000

- Sort the file on other columns, for example, *YIELD*. View the result and see the range of yield values.



	4	5	6	7	8
	VARIETY$	ENTRY	ROW	COL	YIELD
1	V10	10.00000	4.00000	7.00000	1.05000
2	V43	43.00000	4.00000	12.00000	1.50000
3	V21	21.00000	3.00000	3.00000	1.80000
4	V05	5.00000	4.00000	16.00000	2.00000
5	V52	52.00000	1.00000	19.00000	2.00000
6	V55	55.00000	3.00000	15.00000	2.20000
7	C	63.00000	1.00000	1.00000	2.20000
8	V31	31.00000	2.00000	10.00000	2.40000
9	V14	14.00000	3.00000	12.00000	2.60000

- Close the **Data Editor** without saving the latest sort order.

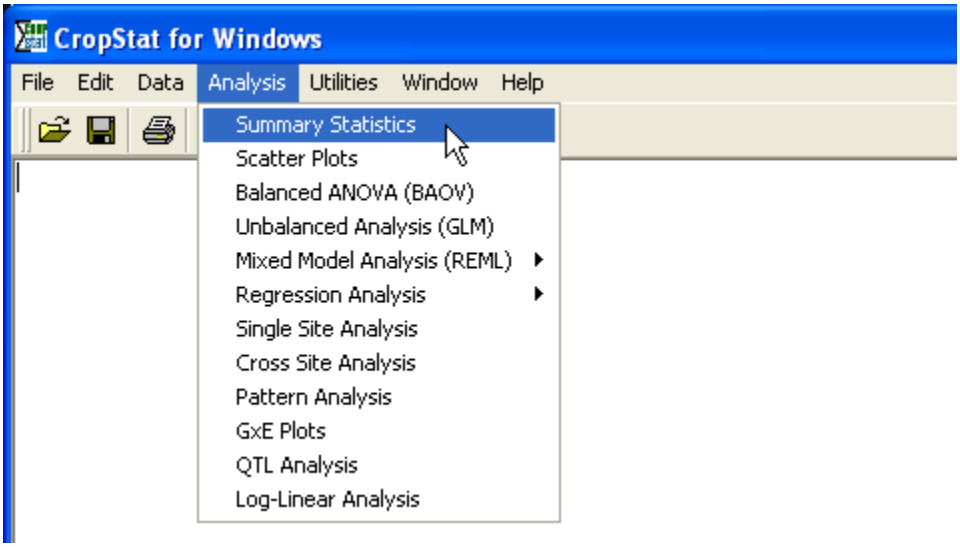
II. Summary Statistics

Sample Problem

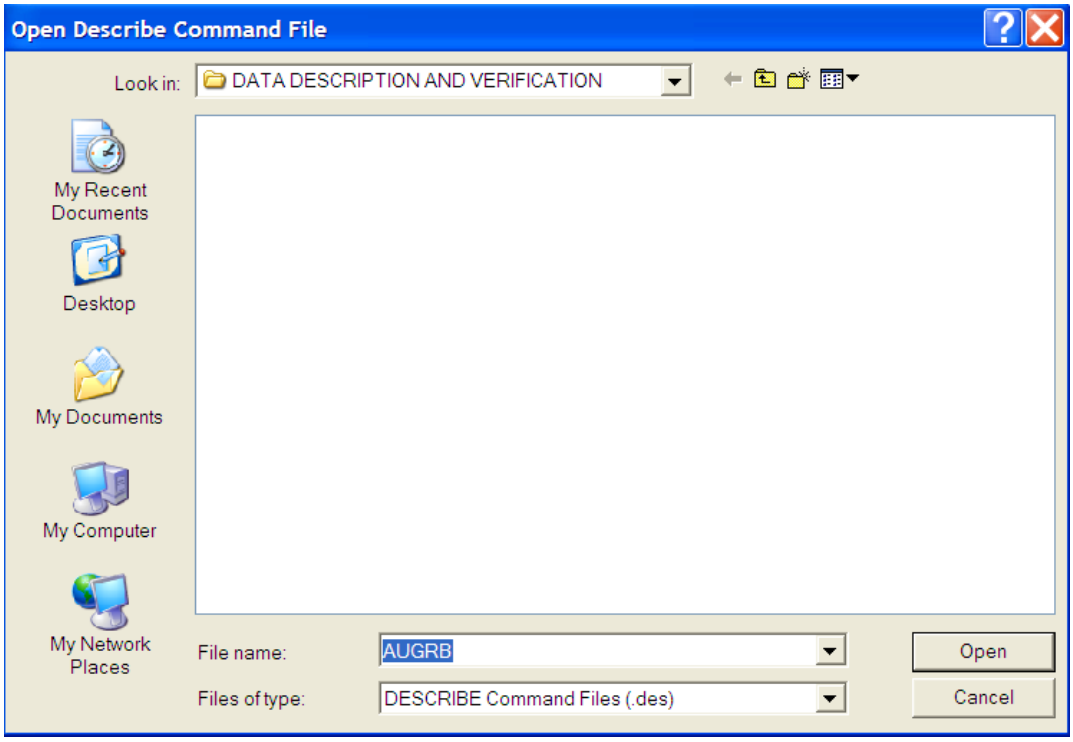
Use summary statistics and tabular counts to check for outlying data values.

Steps

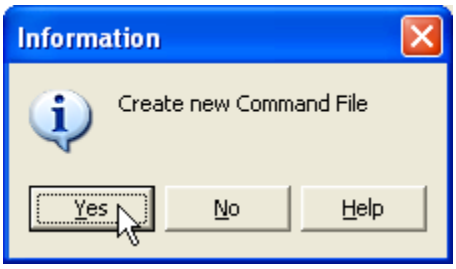
- Select **Summary Statistics** from the **Analysis** menu of the **Main Window**.



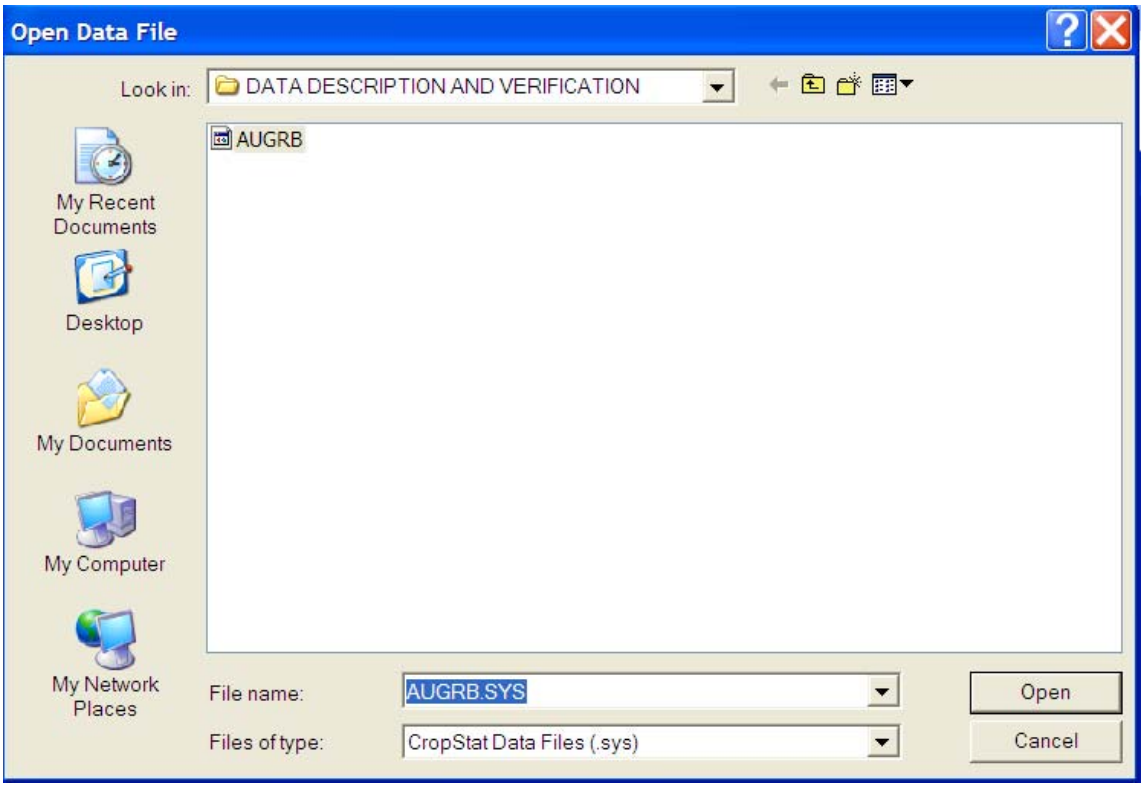
- Click the **Look in** box and go inside your working folder *C:\MY CROPSTAT\ DATA DESCRIPTION AND VERIFICATION* folder. Enter *AUGRB* as the name for a new data description command file. Note that the file extension for a data description command file is .DES.

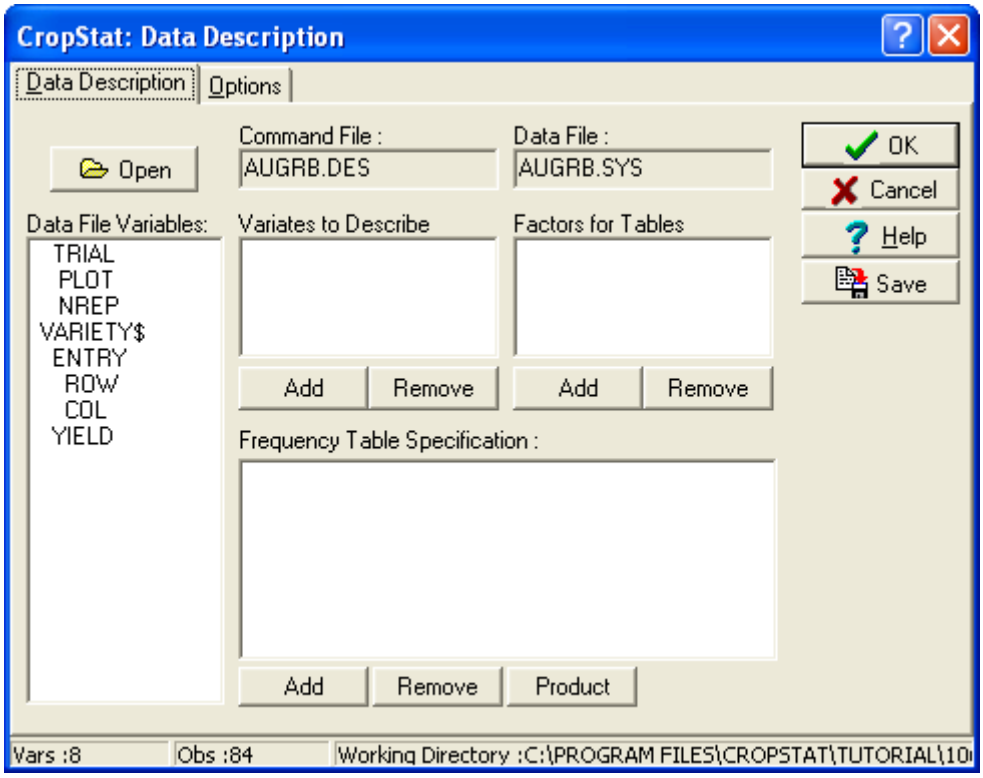


- Click **Open** and **YES** to create a new command file.

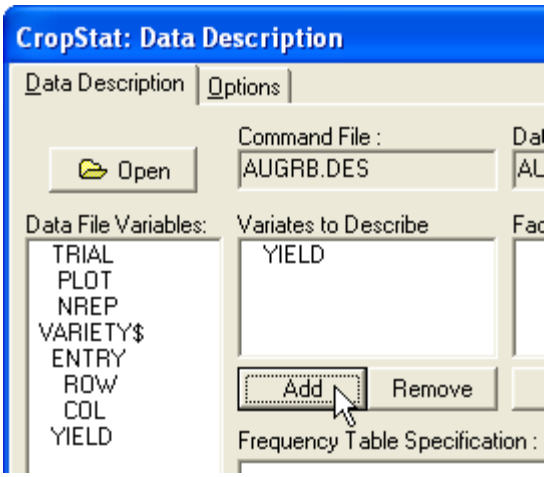


- Select *AUGRB.SYS* as the data file and click **Open** to display the **Data Description** dialog.

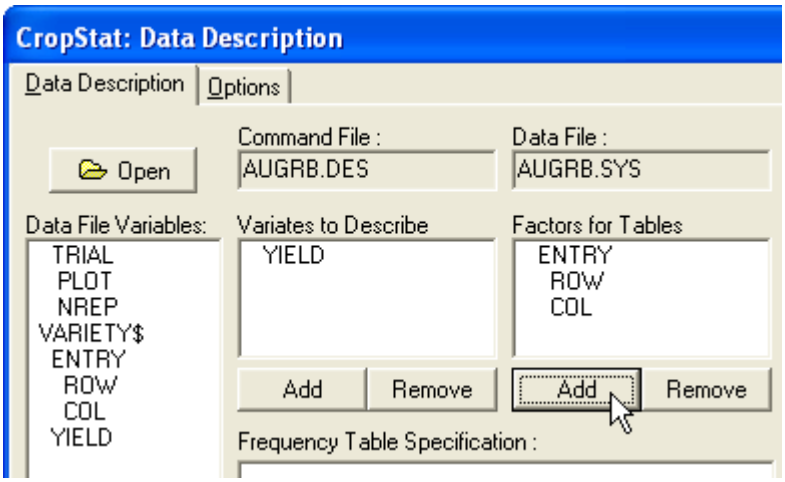




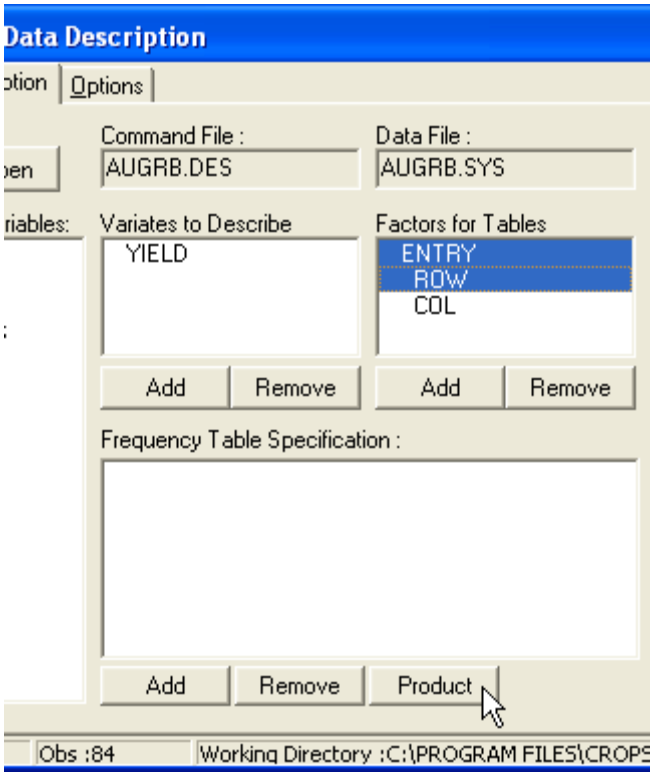
- Enter *YIELD* to the **Variates to Describe** box by selecting it from the **Data File Variables** list and clicking the **Add** button.



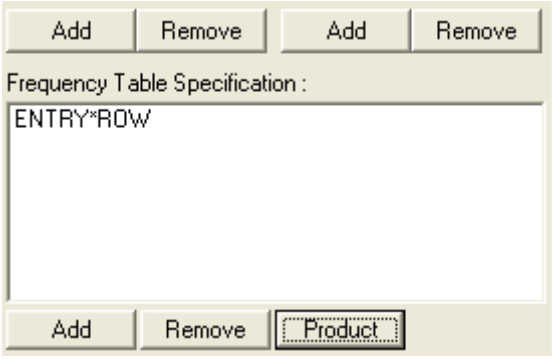
- Highlight *ENTRY*, *ROW* and *COL* and click **Add** under the **Factors for Tables** box.



- In the **Factors for Tables** box, highlight *ENTRY* and *ROW* then click **Product** below the **Frequency Table Specification** box.



ENTRY and *ROW* selected in the **Factors for Tables** box

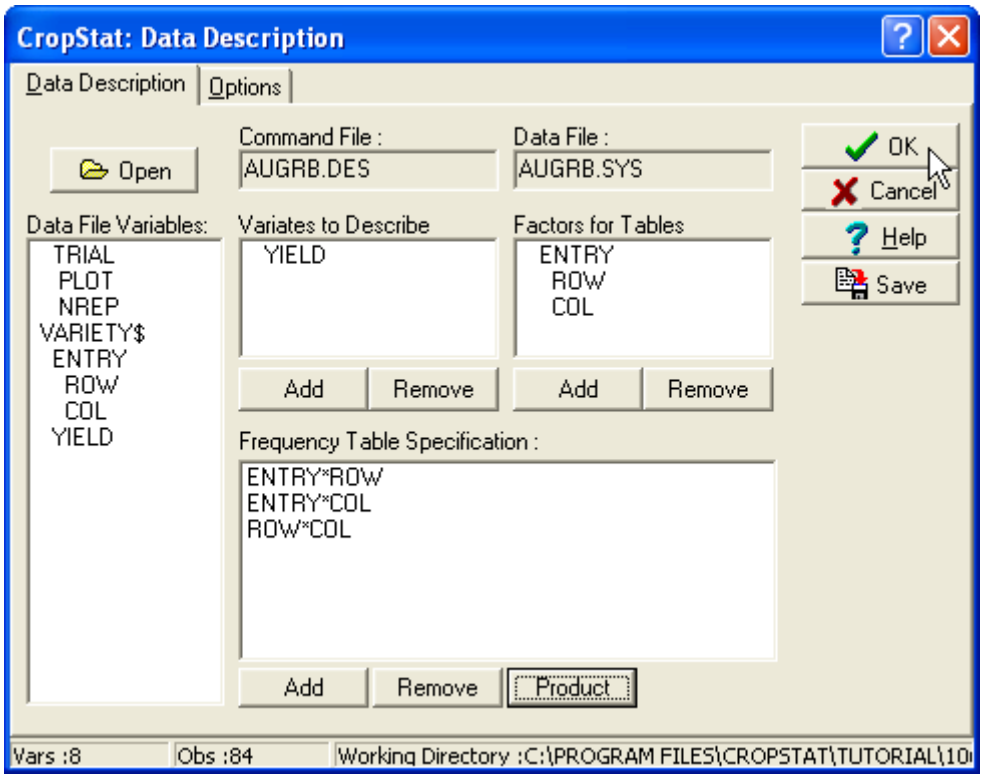


*ENTRY*ROW* added in the **Frequency Table Specification** box after clicking the **Product** button

- Highlight *ENTRY* and *COL* (Click *ENTRY* then hold Ctrl-key and click *COL*) and then click **Product** under the **Frequency Table Specification** box.

- Repeat to enter *ROW*COL*.

- Click **OK** to describe the data.



III. Sample Output

Summary Statistics: Count, count of missing values, maximum, record of maximum, minimum, record of minimum, mean, standard deviation, median, median absolute deviation, and a robust standard error are printed for each variate in the **Variates to Describe** list.

```
VARIATE DESCRIPTION PROGRAM  FILE AUGRB   30/ 9/ 4 13:58
 ------------------------------------------------------------------- :PAGE   1

  VARIATE       YIELD
                 V  8
  NUMBER           84
  MISSING           0

  MAXIMUM       8.250
  REC OF MAX       66
  MINIMUM       1.050
  REC OF MIN       70

  MEAN          4.332
  STD DEV       1.377

  MEDIAN        4.275
  MED ABS DEV  0.8750
  ROBUST SE     1.296
```

This is followed by a vertical display of a Box-Plot showing the median, +, inner hinges, I, outer hinges, -, extreme values, * and count of outliers which exceed the plot limits. The Anderson-Darling A^2 test for normality is also shown.

```
  PLOT LIMIT    8.250
  NO > LIMIT     0
                  *
                  -
                  -
                  -
                  -
                  -
                  I

                  +
                  I
                  -
                  -
                  -
                  -
                  -
  NO < LIMIT     0
  PLOT LIMIT    1.050
  EDF A**2(3)   0.38
```

For each two-way frequency table specified, the counts in each cell are given as well as marginal counts. In this example, test entries must only appear once over all rows and all columns and check entries must appear exactly four times. Partial listing of the two-way frequency table is presented below.

```
VARIATE DESCRIPTION PROGRAM  FILE AUGRB   30/ 9/ 4 13:58
------------------------------------------------------------------- :PAGE   2

               COUNTS FOR ENTRY X ROW                        SECTION  1

    CLASSES      1    2    3    4 TOTA LS
-----------------------------------------
          1     1    0    0    0    1
          2     0    0    1    0    1
          3     0    0    1    0    1
          4     0    0    0    1    1
          5     0    0    0    1    1
          .     .    .    .    .    .
          .     .    .    .    .    .
          .     .    .    .    .    .
         61     1    1    1    1    4
         62     1    1    1    1    4
         63     1    1    1    1    4
         64     1    1    1    1    4
         65     1    1    1    1    4
         66     1    1    1    1    4
-----------------------------------------
  TOTALS       21   21   21   21   84
-----------------------------------------
CHI-SQUARE (195 DF) = 180.0     UPPER TAIL P-VALUE = 0.772
** WARNING ** 100% OF CELLS HAD EXPECTED COUNTS LESS THAN FIVE
```

BALANCED ANALYSIS OF VARIANCE

At the end of the tutorial, the user should be able to

- perform analysis of variance from experimental designs with single error (e.g., RCB) with least square difference and test of residuals
- perform analysis of variance from experimental designs with 2 or more errors (e.g., Split Plot)
- perform partitioning of sum of squares of main and interaction effects
- perform analysis of variance with single and multiple data selection

I. Sample Problem: Fertilizer Experiment

Using the datafile *NRATE.sys* that was created in the tutorial on Data and File Management, perform an analysis of variance on grain yield to compare

1. the control with the applied treatments
2. sources (prilled urea, SCU, USG, USG/urea)
3. levels of nitrogen (low, high)
4. effect of source at different levels of application.

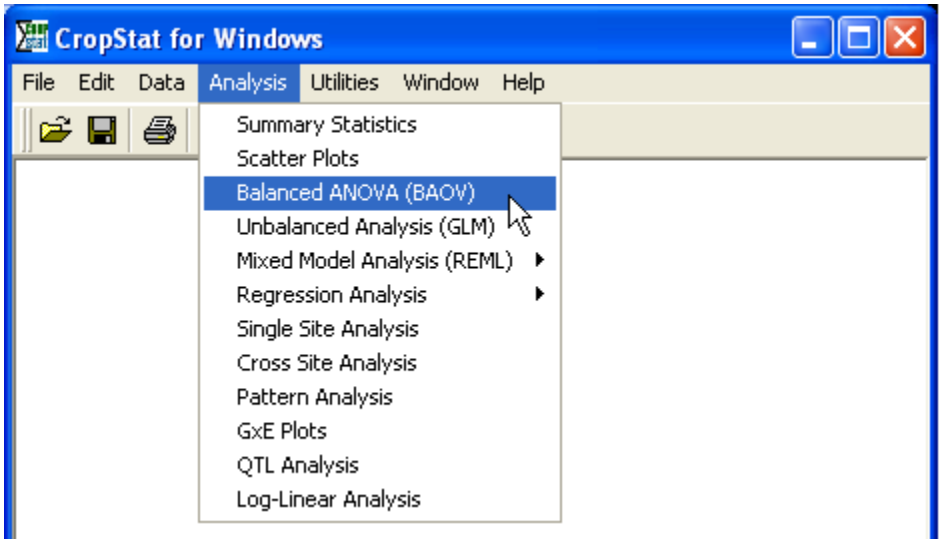
The outline of the ANOVA for this problem is given below.

Source	df	
Rep	2	
Treatments	8	
Control vs. treated		1
Source (S)		3
Level (L)		1
S × L		3
Error	16	
Total	26	

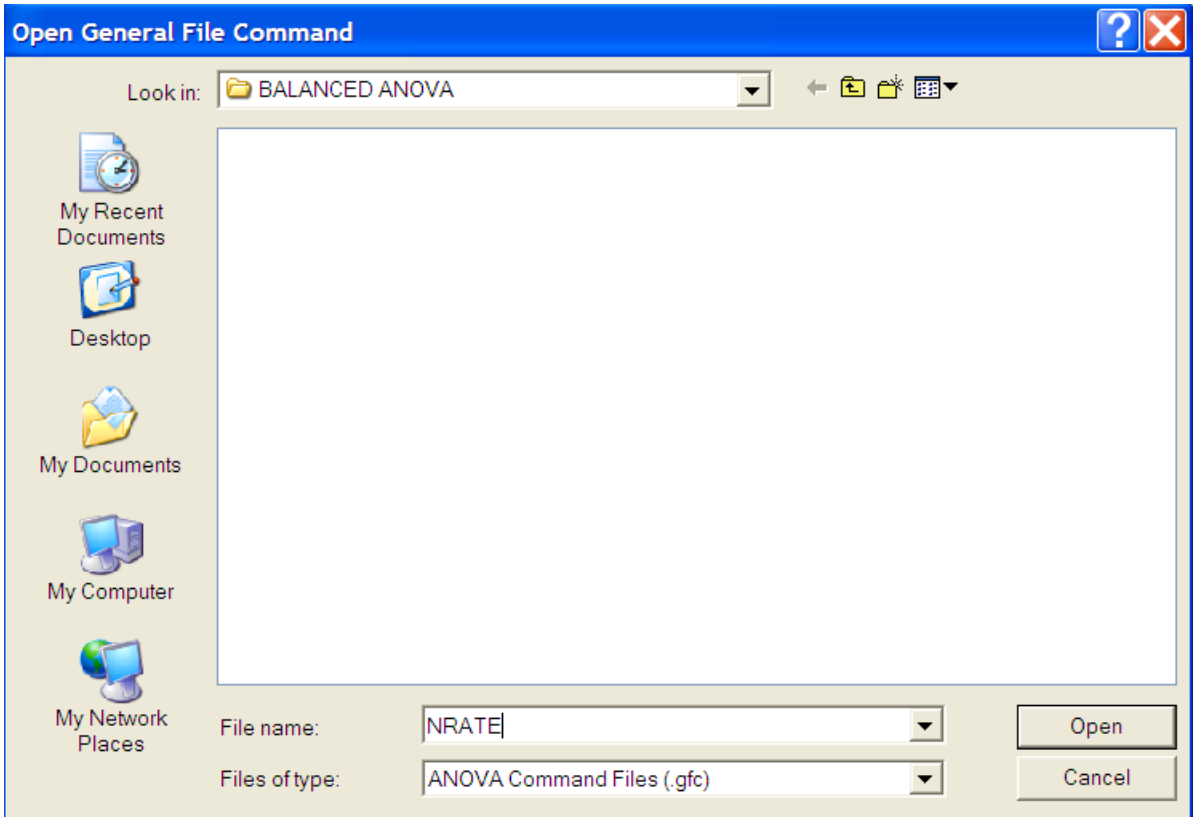
II. Analysis of the Fertilizer Experiment

- Re-open NRATE.SYS from the DATA AND FILE MANGEMENT subfolder. Clik on **File ⇒ Save as.** Click the **Save in** box and go inside the directory *C:\MY CROPSTAT* create a subfolder BALANCED ANOVA and save NRATE.SYS.

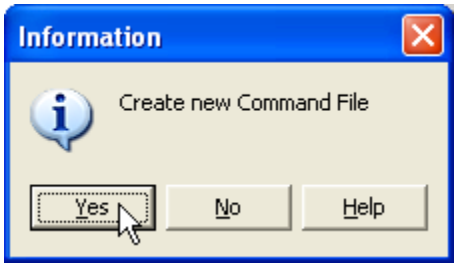
- Select **Analysis|Balanced ANOVA** from the Main Window.



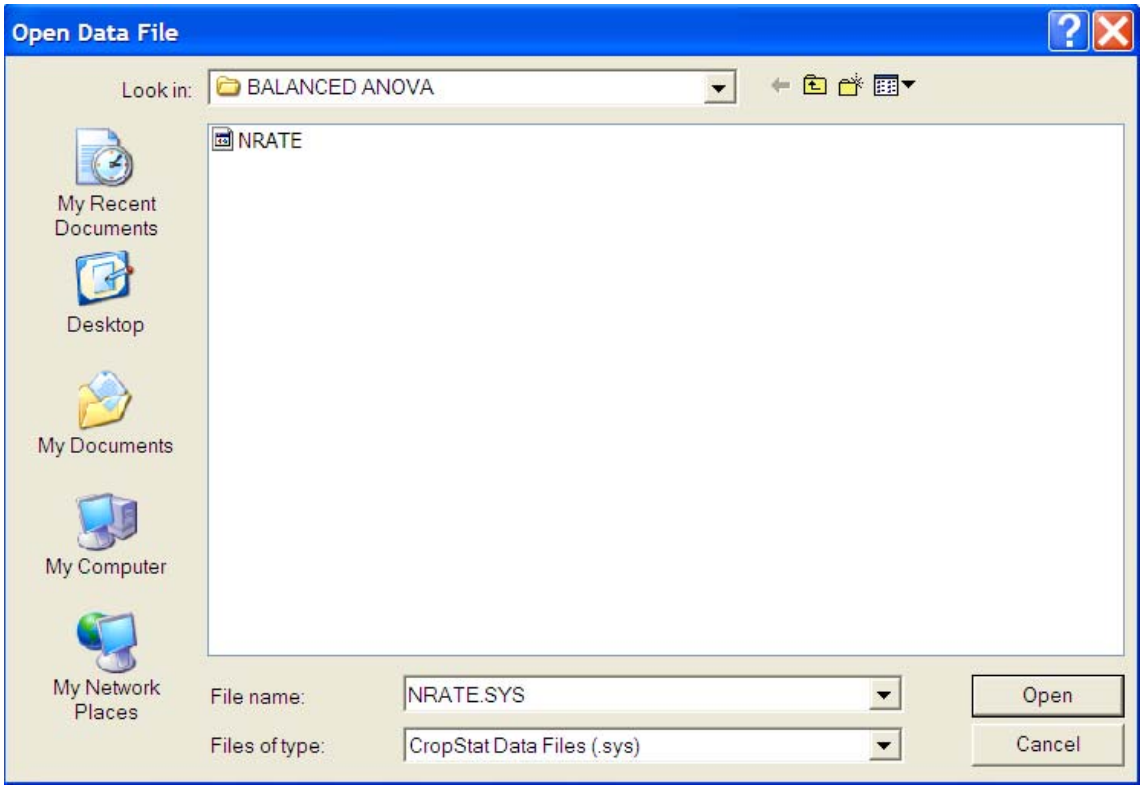
- The **Open** dialog box will prompt you to enter a name for the command file. Click the **Look In** box to go to your working drive *C:\MY CROPSTAT\BALANCED ANOVA*.



- Enter *NRATE* in the **File name** box. Click the **Open** button. Since *NRATE.GFC* does not exist, a message box will appear confirming if you want to create the file. Click **Yes**.

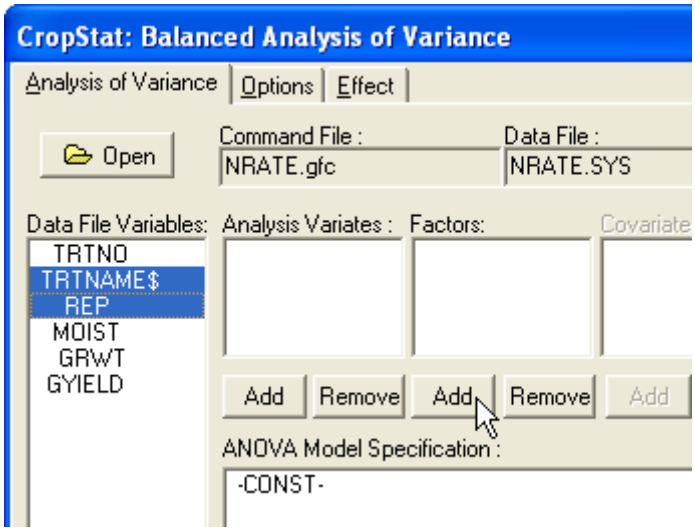


- Enter the name of the data file to be used. Enter *NRATE.SYS* in the **File name** box.

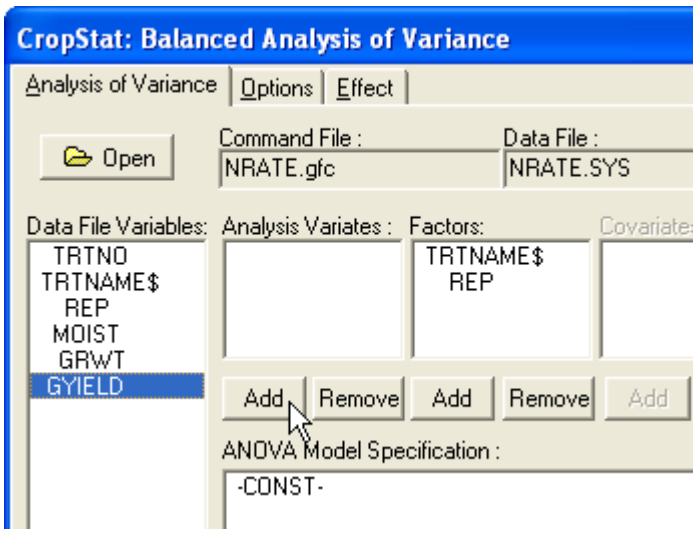


- Click **Open**. The **Analysis of Variance** dialog box will appear to help you in specifying the details of the analysis.

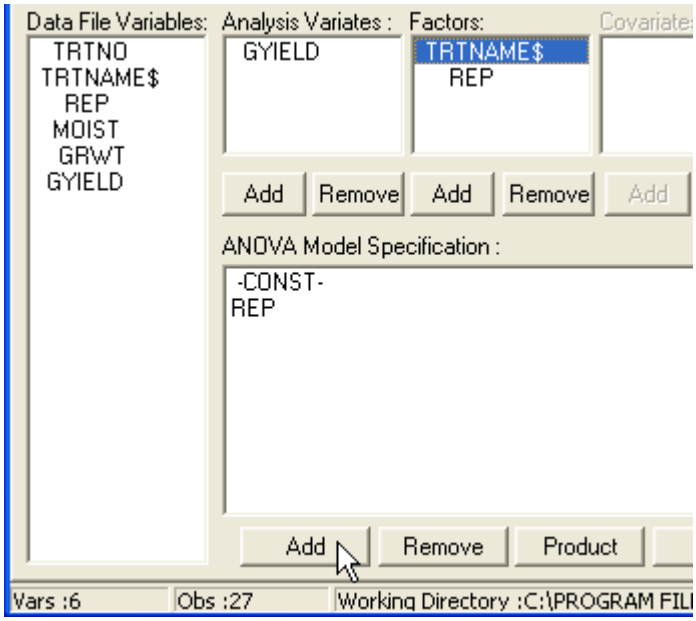
- To specify variates to be used as factors, select the variates *REP* and *TRTNAME$* from **Data File Variables** list and click the **Add** button under the **Factors** list.



- To specify variates to be analyzed, select these variates from **Data File Variables** list and click the **Add** button under the **Analysis Variates** list. For this example, select *GYIELD*.



- To specify the ANOVA model, select *REP* and *TRTNAME$* from the **Factors** list. Click **Add** under the **ANOVA Model Specification** list. The constant effect is automatically entered into the model.



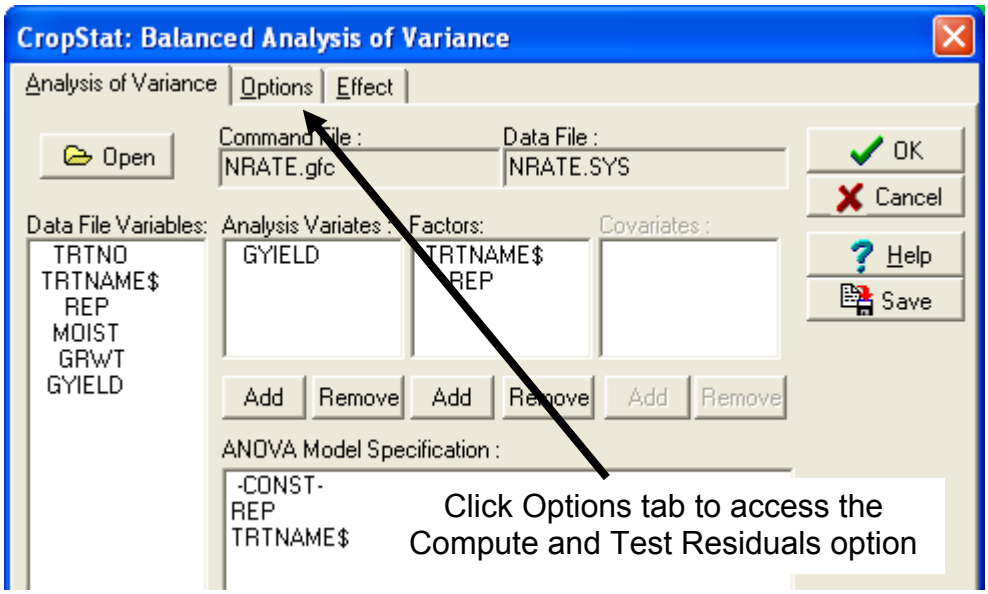
- To add the main effects, select each factor from the **Factors** list. Click the **Add** button under the **ANOVA Model Specification** list. To remove effect(s), highlight the effects(s) and click the Remove button.

Note: To define interaction effects, click the **Product** button. You must enter main effects and lower order interactions before higher order interactions with any group of factors.

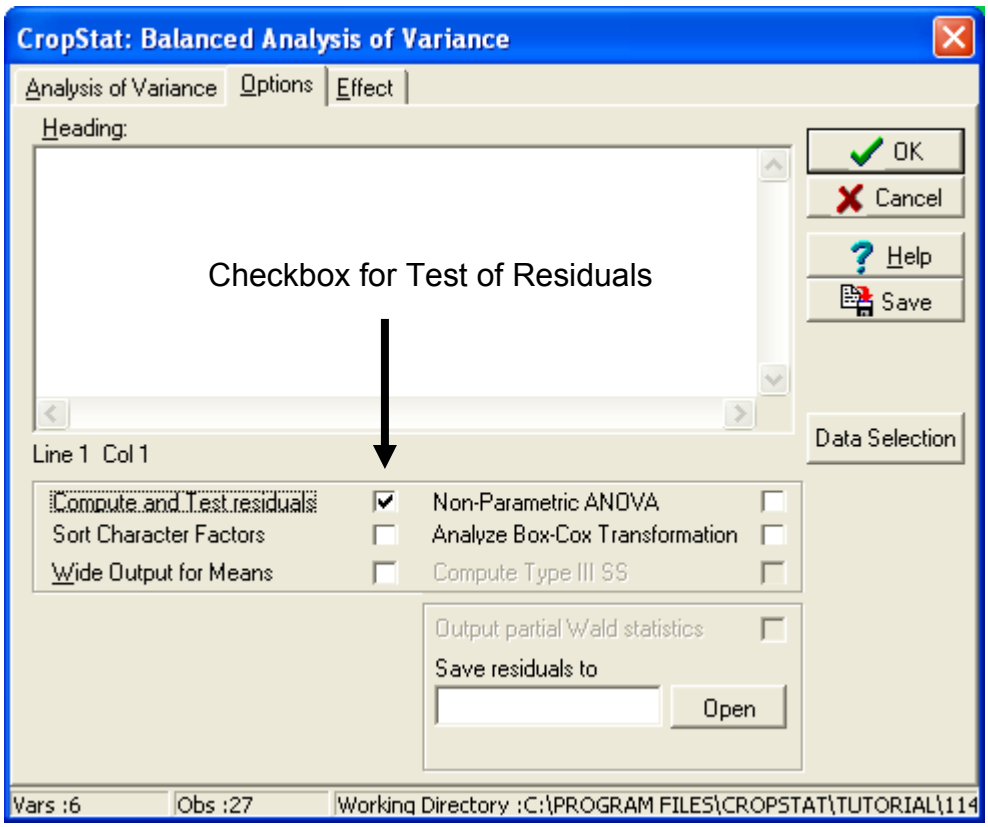
Example:	A	main effect
	B	main effect
	A*B	interaction

To generate all possible combinations of selected factors, select the factors from the **Factors** list and click the **Cross** button under the **ANOVA Model Specification** list.

- To request for Test of Residuals, click the **Options** tab.



- Click the options checkbox for **Compute and Test Residuals**. Click **OK** to run the analysis.



III. Sample Output

The following text will appear in the Text Editor with filename *NRATE.OUT*.

1. ANOVA table

```
 BALANCED ANOVA FOR VARIATE   GYIELD  FILE NRATE   27/ 1/ 5 13:46
------------------------------------------------------------------ :PAGE   1
VARIATE V006 GYIELD

LN   SOURCE OF VARIATION            DF  SUMS OF     MEAN     F RATIO PROB  ER
                                        SQUARES     SQUARES                LN
=============================================================================
1 REP                                2 .343195     .171597       0.56 0.586  3
2 TRTNAME$                           8 19.8485     2.48106       8.13 0.000  3
* RESIDUAL                          16 4.88432     .305270
-----------------------------------------------------------------------------
* TOTAL (CORRECTED)                 26 25.0760     .964461
-----------------------------------------------------------------------------
```

2. Test of assumptions

```
TUKEY'S TEST FOR TRANSFORMABLE NON-ADDITIVITY
SS= 0.328493E-01, F(1, 15)=   0.10, P= 0.752
REGRESSION SLOPE=-0.18042     SUGGESTED POWER TRANSFORMATION=  2.9815

MEAN ORTHOGONAL RESIDUAL = -0.8493E-01, P-VALUE= 0.554


PLOT OF LS RESIDUALS AGAINST FITTED VALUES
          .............................................................
          :      *                                                    :
          :                                                           :
          :                                                           :
          :                                                           :
     0.9 -:                                                           :
          :                            *                              :
          :                                                           :
          :                                                           :
          :                                        *                  :
     0.3 -:                           *    *                         :
          :                                    **   *              *  :
          :                                         *           *  *  :
          :                                  *          *  *  *       :
          :                          *             *    * *  *        :
    -0.3 -:                                                 *      *  :
          :                                                           :
          :   *                                                       :
          :                            * *                            :
          :        *                                                  :
    -0.9 -:                                                           :
          :...........................................................:
           :         :         :         :         :         :
          3.3       3.9       4.5       5.1       5.7       6.3
THE FOLLOWING RECORDS HAVE LARGE RESIDUALS, (>2.5 SES)

 TRTNAME$         REP      OBS. VALUE  FITTED VALUE  RESIDUAL
CONTROL     2               5.006        3.677        1.329

BOX PLOT OF STUDENTIZED RESIDUALS FROM LPLT= -1.791    TO ULPT=  3.124
NO.<LPLT                                                    NO.>UPLT
       0 ----------------I    +     I----------------              *   0

MEDIAN= -0.1558E-01 ANDERSON-DARLING STATISTIC=  0.497
```

3. Table of means

```
TABLE OF MEANS FOR FACTORIAL EFFECTS  FILE NRATE   27/ 1/ 5 13:46
------------------------------------------------------------------- :PAGE   2
MEANS FOR EFFECT REP
--------------------------------------------------------------------------------

     REP          NOS       GYIELD
1                   9        5.34136
2                   9        5.51992
3                   9        5.61308

SE(N=   9)                  0.184171
5%LSD  16DF                 0.552147
--------------------------------------------------------------------------------

MEANS FOR EFFECT TRTNAME$
--------------------------------------------------------------------------------

 TRTNAME$         NOS       GYIELD
CONTROL             3        3.64863
L-N UREA            3        4.83205
L-N SCU             3        4.95429
L-N USG             3        6.02263
L-N USG/UREA        3        5.40650
H-N UREA            3        5.69666
H-N SCU             3        6.35465
H-N USG             3        6.61284
H-N USG/UREA        3        5.89482

SE(N=   3)                  0.318993
5%LSD  16DF                 0.956347
--------------------------------------------------------------------------------
```

4. Summary table

```
ANALYSIS OF VARIANCE SUMMARY TABLE  FILE NRATE   27/ 1/ 5 13:46
------------------------------------------------------------------- :PAGE   3

 F-PROBABLIITY VALUES FOR EACH EFFECT IN THE MODEL. SECTION - 1

 VARIATE       GRAND MEAN  STANDARD   DEVIATION  C OF V |REP     |TRTNAME$|
               (N=   27)   ------------------- SD/MEAN |        |        |
              NO.          BASED ON   BASED ON    %    |        |        |
              OBS.         TOTAL SS   RESID SS         |        |        |
  GYIELD       27  5.4915    0.98207    0.55251     10.1 0.5856   0.0002
```

IV. Testing Contrast Hypotheses

As noted in section I, interest in TREATMENT centers on several specific hypotheses:

1. Compare the control with the applied treatments.
2. Among the applied treatments,
   - compare the 4 sources of nitrogen (prilled urea, SCU, USG, USG/urea);
   - compare the 2 levels of nitrogen (low, high); and
   - compare the interaction between sources and levels of nitrogen.

These hypotheses can be formulated in terms of contrasts among the treatment means. Using hypothesis 1, if CONTROL is no different to APPLIED TREATMENTS, the mean for the CONTROL should be about the same as the average of the means for the other treatments, i. e.,

CONTROL - (low-N-rate urea + low-N-rate SCU + low-N-rate USG + low-N-rate USG/urea + high-N-rate urea + high-N-rate SCU + high-N-rate USG + High-N-rate USG/urea)/8

should differ from zero only due to error. The statistical aspect of the analysis is to test whether the value of the contrast is significantly different from zero or not.

To specify this contrast in CropStat, we must specify the coefficients for each mean:

1 -0.125 -0.125 -0.125 -0.125 -0.125 -0.125 -0.125 -0.125

The order of the coefficients is critical. They must be entered in the order of the means that appear in the table of treatment means in the previous analysis output.

Note that the effect SOURCE has 3 degrees of freedom. CropStat reads only single-df contrasts so we have to specify all three single-df contrasts for effect SOURCE. Any orthogonal set of contrasts which compares sources will do, but the simplest ones are those that compare

Prilled urea with other sources	-	0 3 -1 -1 -1 3 -1 -1 -1
SCU with USG and USG/UREA	-	0 0 2 -1 -1 0 2 -1 -1
USG vs. USG/UREA	-	0 0 0 1 -1 0 0 1 -1

To compare low and high levels of nitrogen, the contrast will be

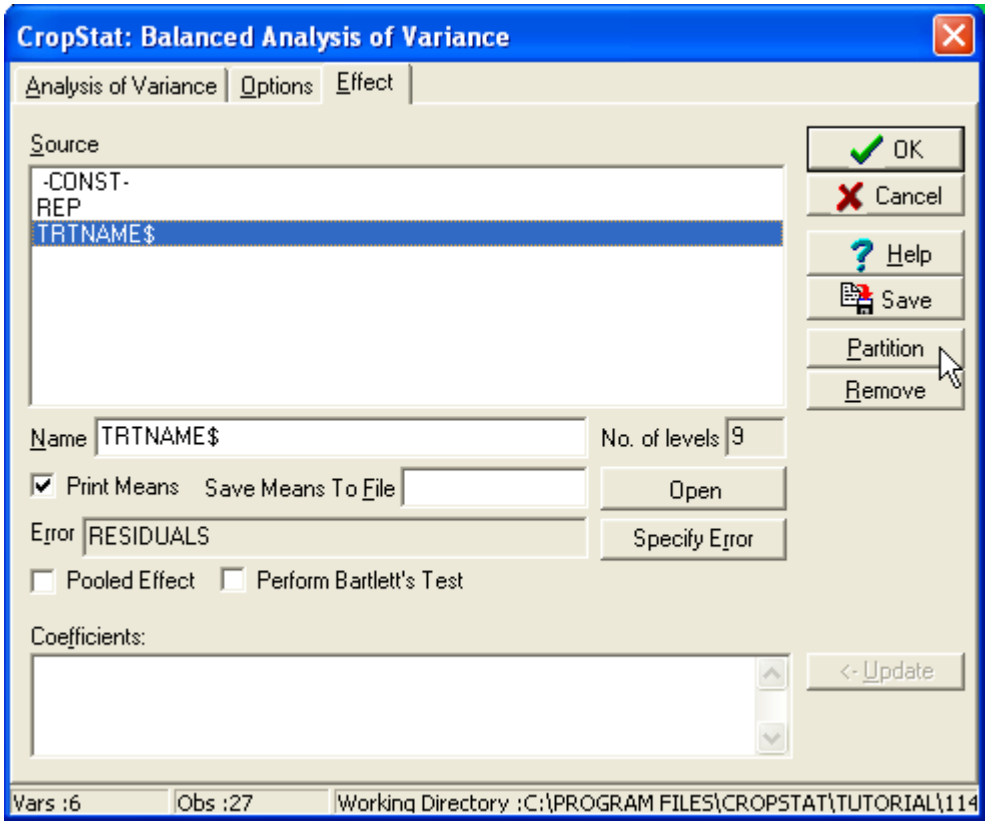
0 1 1 1 1 -1 -1 -1 -1

For the interaction between source and level, the coefficients are the products of the coefficients for source and level: (urea with other sources) × level, (SCU with USG) × level, (USG with urea) × level.

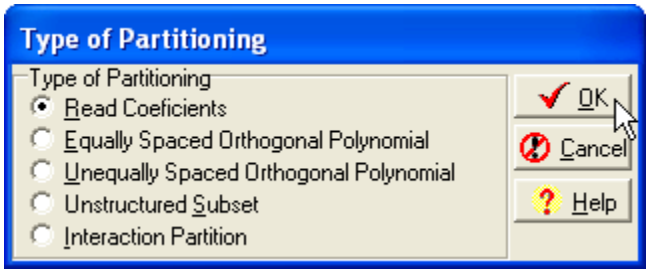
Partitioning of Effects

- Close the Text Editor and go back to the **ANOVA** dialog box by selecting **Analysis ⇒ Balanced ANOVA**. In the **Open ANOVA Command File** dialog, click the **Look in** box and go inside *C:\MY CROPSTAT\BALANCED ANOVA* folder and open *NRATE.GFC*.

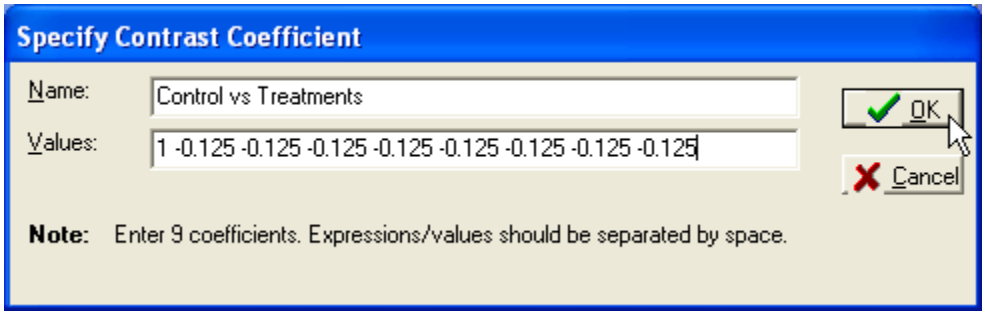
- Click the **Effect** tab. Select the effect you want to partition from the **Source** list, in this case *TRTNAME$*. Click the **Partition** button.



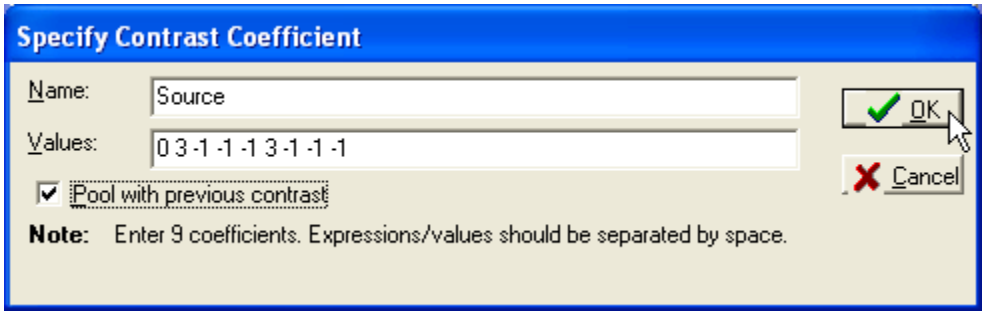
- Select **Read Coefficients** as the type of partitioning. Click **OK.**



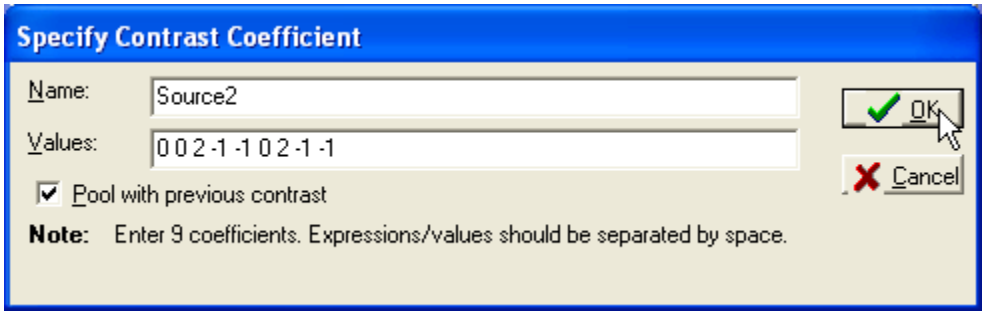
- In the **Specify the contrast coefficient** box, type the contrast name (*Control vs Treatments*) and the corresponding contrast coefficients. Enter the values separated by a space. Click **OK** when you are finished.



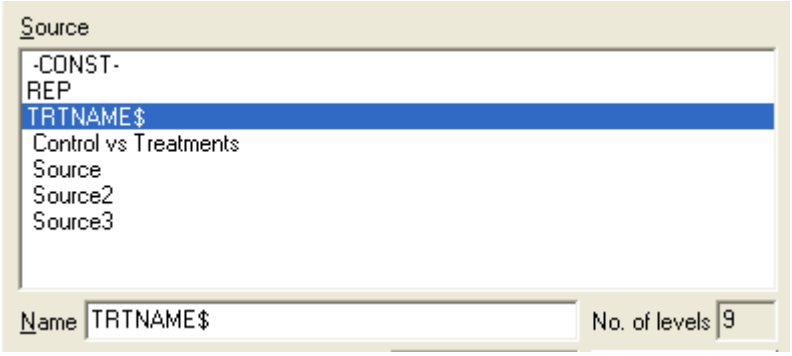
- To specify the first partition for *SOURCE*, highlight *TRTNAME$* from the **Source** list. Click the **Partition** button. Choose **Read Coefficients** as the type of partitioning. Enter the contrast coefficients for the hypothesis Urea vs. Others and *Source* as the contrast name. Since this is the second contrast of the analysis, the **Pool with previous contrast** check box is added on the **Contrast Coefficient** box. If you want to add/pool the Sum of Squares of this contrast to the previous contrast, click on the box and a check (✓) will appear on the box.



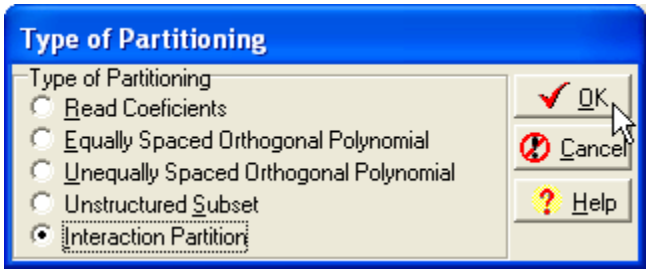
- To specify the second partition for source, highlight *TRTNAME$* from the **Source** list. Click the **Partition** button. Choose **Read Coefficients** as the type of partitioning. Enter the contrast coefficients for the hypothesis SCU vs. USG (0 0 2 -1 -1 0 2 -1 -1). Enter *SOURCE2* as the contrast name and click on the **Pool with previous contrast** box.



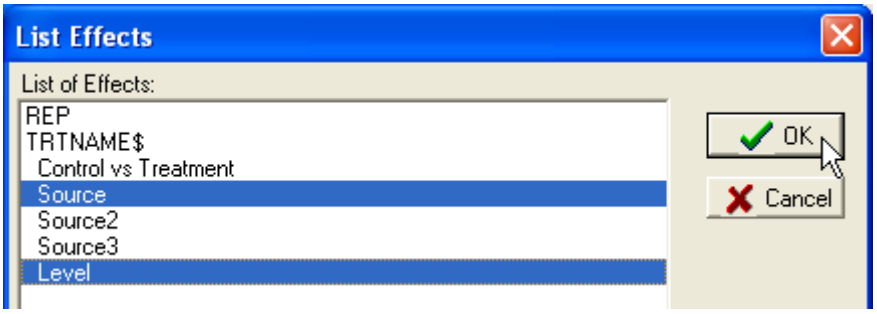
- Enter the third partition for source (0 0 0 1 -1 0 0 1 -1). Enter *SOURCE3* as the contrast name and pool it with the previous two.



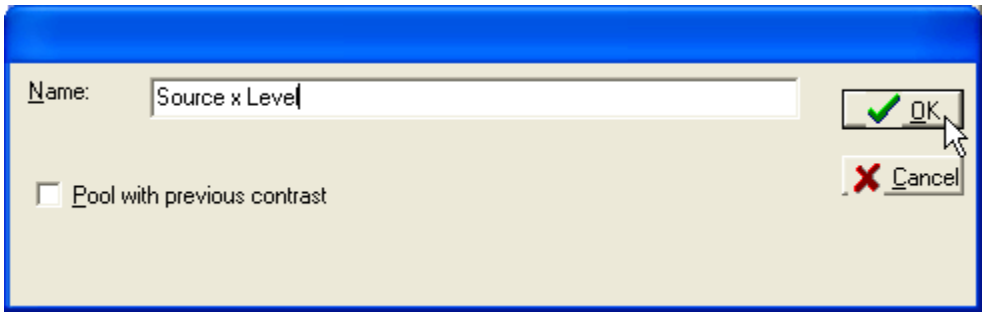
- To compare high and low levels of nitrogen, highlight *TRTNAME$* from the **Source** list. Click the **Partition** button and choose **Read Coefficients** as the type of partitioning. Enter the contrast coefficients for this hypothesis (0 1 1 1 1 -1 -1 -1 -1) and enter *LEVEL* as the contrast name. Do not click on the **Pool with previous contrast** box.
- To define the interaction contrast, we must specify and pool three contrasts: Source × Level, Source2 × Level, and Source3 × Level. Highlight *TRTNAME$* from the **Source** list. Click the **Partition** button and choose **Interaction Partition** as the type of partitioning.



- A list showing all the effects that have been created appears. Click *SOURCE*, then holding the <CTRL key>, click *LEVEL*. Click **OK**.

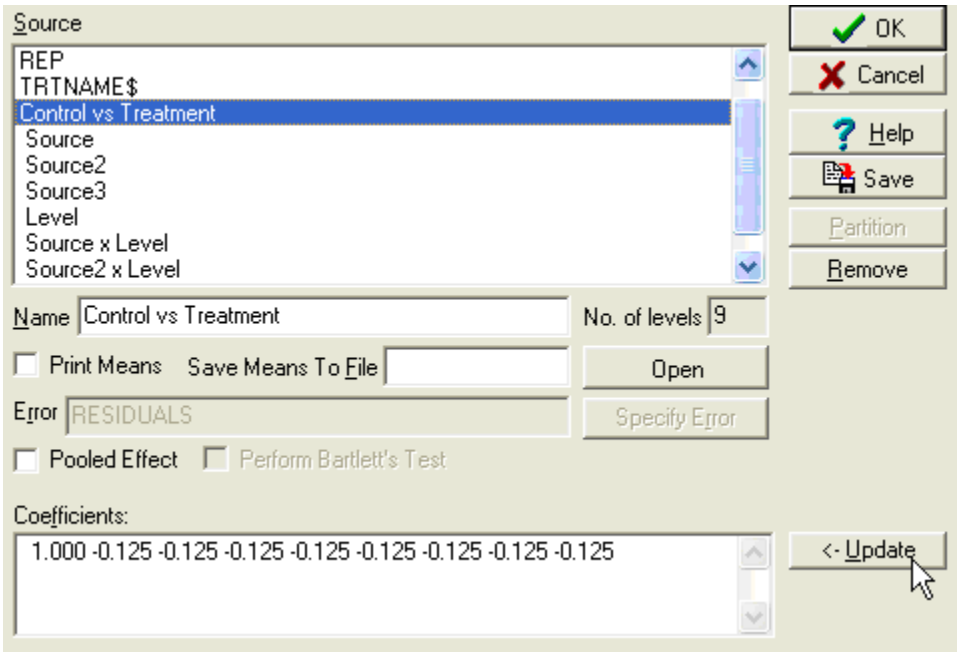


- Enter the contrast name *Source × Level*. Since you do not want to add the Sum of Squares of this contrast to the previous contrast, do not click on the **Pool with previous contrast** box.



- Do the same procedure to enter the contrasts for *SOURCE2* × *LEVEL* and *SOURCE3* × *LEVEL*. Pool these effects with the previous effect. Click **OK** to run the analysis.

- To edit the contrast coefficient, click the contrast name and edit the coefficients through the Update box and click **Update.**



- To change the Contrast Name, go to the **Name** box and type in the Contrast Name.

V. Sample Output

The following text saved in *NRATE.OUT* will appear in the Text Editor.

1. ANOVA table

```
 BALANCED ANOVA FOR VARIATE   GYIELD  FILE NRATE    3/ 2/ 4 11: 0
 ------------------------------------------------------------------ :PAGE   1
 VARIATE V005 GYIELD

LN   SOURCE OF VARIATION            DF  SUMS OF     MEAN     F RATIO PROB  ER
                                       SQUARES     SQUARES                 LN
 ============================================================================
 1 REP                               2 .343195     .171597       0.56 0.586 11
 2 TRTNAME$                          8 19.8485     2.48106       8.13 0.000 11
 3 Control vs. Treatment             1 11.4615     11.4615      37.55 0.000 11
 4 Source                            3 3.44391     1.14797       3.76 0.032 11
 5 Level                             1 4.19211     4.19211      13.73 0.002 11
 6 Source x Level                    3 .750921     .250307       0.82 0.504 11
 * RESIDUAL                         16 4.88432     .305270
 ----------------------------------------------------------------------------
 * TOTAL (CORRECTED)                26 25.0760     .964461
 ----------------------------------------------------------------------------
```

2. Test of assumptions for ANOVA

```
 TUKEY'S TEST FOR TRANSFORMABLE NON-ADDITIVITY
 SS= 0.328482E-01, F(1, 15)=   0.10, P= 0.752
 REGRESSION SLOPE=-0.18042     SUGGESTED POWER TRANSFORMATION=  2.9815

 MEAN ORTHOGONAL RESIDUAL = -0.8493E-01, P-VALUE= 0.554

 PLOT OF LS RESIDUALS AGAINST FITTED VALUES
           .............................................................
           :      *                                                    :
           :                                                           :
           :                                                           :
           :                                                           :
      0.9 -:                                                           :
           :                          *                                :
           :                                                           :
           :                                                           :
           :                                      *                    :
      0.3 -:                          *    *                          :
           :                                    **   *              *  :
           :                                         *           *  *  :
           :                                  *           *  *  *      :
           :                           *              *    * *  *      :
     -0.3 -:                                                 *      *  :
           :                                                           :
           :   *                                                       :
           :                           * *                             :
           :        *                                                  :
     -0.9 -:                                                           :
           :...........................................................:
            :         :         :         :         :         :
           3.3       3.9       4.5       5.1       5.7       6.3
 THE FOLLOWING RECORDS HAVE LARGE RESIDUALS, (>2.5 SES)

 TRTNAME$         REP      OBS. VALUE  FITTED VALUE  RESIDUAL
 CONTROL     2               5.006        3.677        1.329

 BOX PLOT OF STUDENTIZED RESIDUALS FROM LPLT= -1.791    TO ULPT=  3.124
 NO.<LPLT                                                     NO.>UPLT
        0 ----------------I    +     I----------------              *   0

 MEDIAN= -0.1558E-01 ANDERSON-DARLING STATISTIC=  0.497
```

3. Table of means

```
TABLE OF MEANS FOR FACTORIAL EFFECTS  FILE NRATE   30/ 9/ 4 15:54
------------------------------------------------------------------- :PAGE   2
MEANS FOR EFFECT REP
--------------------------------------------------------------------------------

     REP         NOS       GYIELD
1                  9        5.34136
2                  9        5.51992
3                  9        5.61308

SE(N=   9)                 0.184171
5%LSD  16DF                0.552147
--------------------------------------------------------------------------------

MEANS FOR EFFECT TRTNAME$
--------------------------------------------------------------------------------

 TRTNAME$        NOS       GYIELD
CONTROL            3        3.64863
L-N UREA           3        4.83205
L-N SCU            3        4.95429
L-N USG            3        6.02263
L-N USG/UREA       3        5.40650
H-N UREA           3        5.69666
H-N SCU            3        6.35465
H-N USG            3        6.61284
H-N USG/UREA       3        5.89482

SE(N=   3)                 0.318993
5%LSD  16DF                0.956347
--------------------------------------------------------------------------------
```

4. ANOVA summary table

```
ANALYSIS OF VARIANCE SUMMARY TABLE  FILE NRATE   30/ 9/ 4 15:54
------------------------------------------------------------------- :PAGE   3

 F-PROBABLIITY VALUES FOR EACH EFFECT IN THE MODEL. SECTION - 1

 VARIATE       GRAND MEAN  STANDARD   DEVIATION  C OF V |REP     |TRTNAME$|Control |Source3 |Level   |Source3 |
               (N=   27)   ------------------- SD/MEAN |        |        |vs. Treat|        |        |x Level |
              NO.          BASED ON   BASED ON     %    |        |        |ment    |        |        |        |
              OBS.         TOTAL SS   RESID SS          |        |        |        |        |        |        |
  GYIELD       27  5.4915    0.98207    0.55251     10.1 0.5856   0.0002   0.0000   0.0320   0.0020   0.5040
```

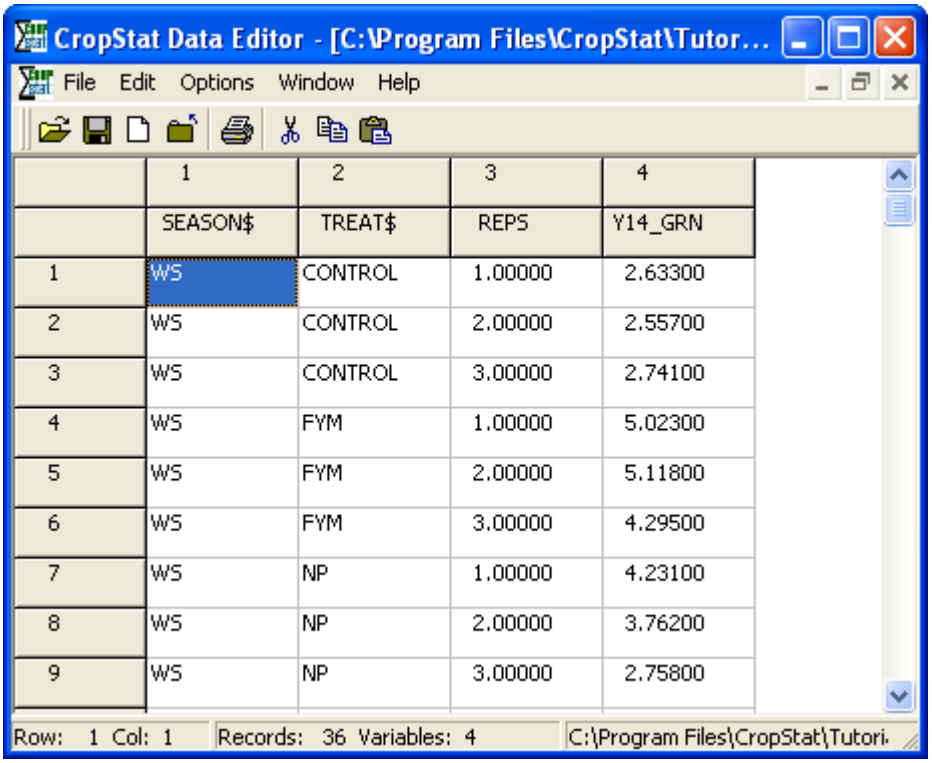
VI. Data Subsets

A. Single Selection Set

Single Selection is used for data selection for only one set of data. The selection can be specified by the variates in the data set. Up to four variates can be used for the selection.

The data were taken from a fertilizer experiment in RCBD in 1995 with two seasons (Dry and Wet Seasons). The experiment has 6 treatments and 3 replicates stored in *FERT1.SYS*.

The researcher wants to have an ANOVA for the fertilizer treatments (excluding control) in the wet season (WS) only. (***Note:*** this means that you want to have a subset with Season = WS and all the treatments except the Control). Use the **Data Selection Single Selection Set** to analyze the wet season data only.



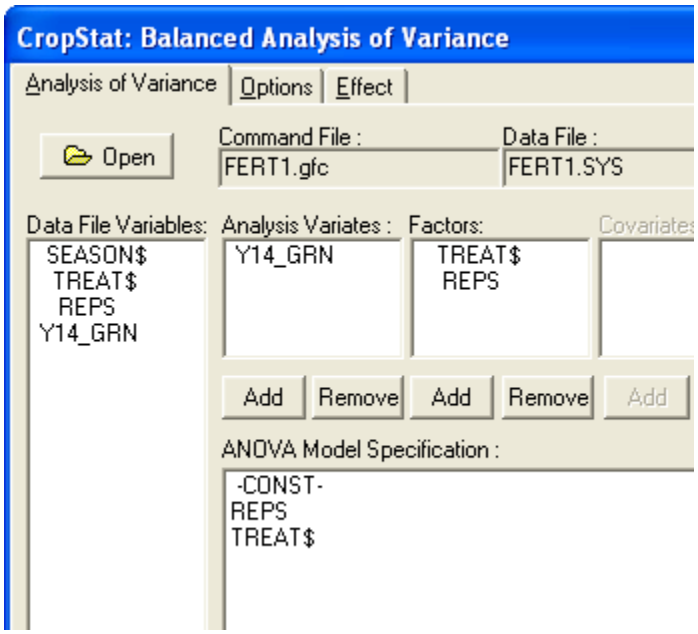
	1	2	3	4
	SEASON$	TREAT$	REPS	Y14_GRN
1	WS	CONTROL	1.00000	2.63300
2	WS	CONTROL	2.00000	2.55700
3	WS	CONTROL	3.00000	2.74100
4	WS	FYM	1.00000	5.02300
5	WS	FYM	2.00000	5.11800
6	WS	FYM	3.00000	4.29500
7	WS	NP	1.00000	4.23100
8	WS	NP	2.00000	3.76200
9	WS	NP	3.00000	2.75800

- Select **File ⇒ Open.** Go to *C:\PROGRAM FILES\ CROPSTAT7.2\TUTORIAL\ \TUTORIAL DATASETS* and open Fert1.SYS.

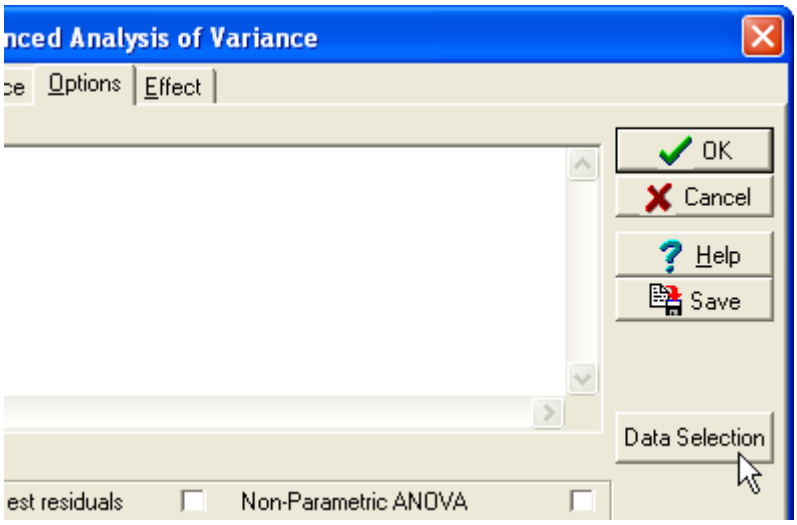
- Select on **File ⇒ Save as.** Click the **Save in** box and go inside the directory C:\MY CROPSTAT\ BALANCED ANOVA and save FERT1.SYS.

- Select **Analysis|Balanced ANOVA** from the Main Window.

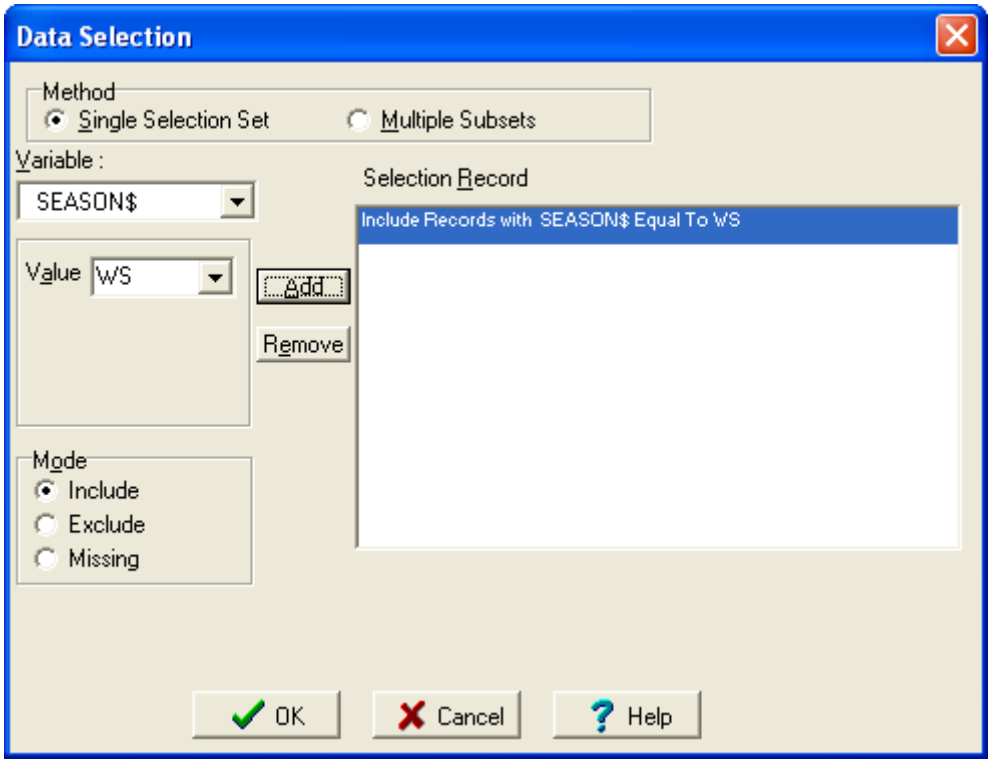
- The **Open** dialog box will prompt you to enter a name for the command file. Click the **Look In** box to go to your working drive *C:\MY CROPSTAT\BALANCED ANOVA.*
- Enter *FERT1* in the **File name** box. Click **Open** button. Since *FERT1.GFC* does not exist, a message box will appear confirming if you want to create one. Click **Yes**.
- Enter the name of the data file to be used. Enter *FERT1*.SYS in the **File name** box.
- Click **Open**. The **Analysis of Variance** dialog box will appear. Specify *Y14_GRN* as the Analysis Variates; *TREAT$* and *REPS* as the Factors. Note*:* *SEASON$* is not included as a factor.
- From the Factors list box, add *REPS* and *TREAT$* in the ANOVA Model Specification box.



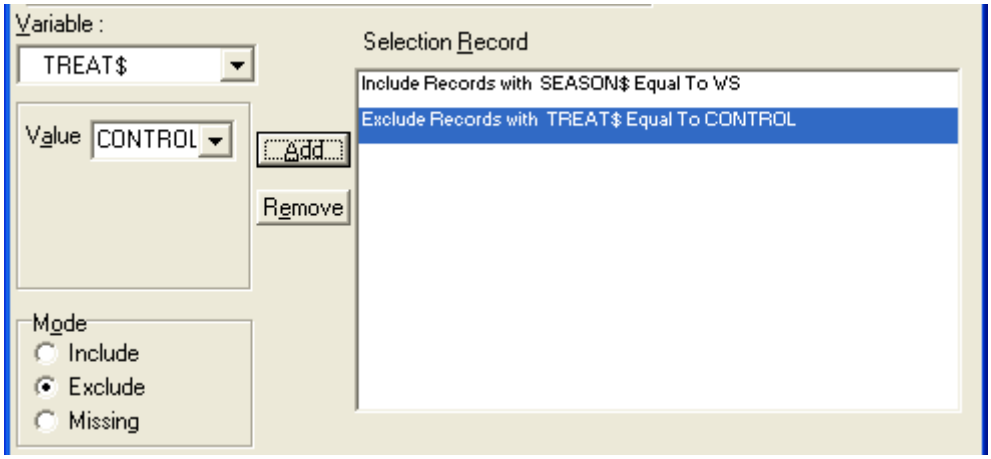
- Got o the **Options** tab then click the **Data Selection** button.


nced Analysis of Variance
ce
Options
Effect
OK
Cancel
Help
Save
Data Selection
est residuals
Non-Parametric ANOVA

- Cick on **Single Selection** select *SEASON$* from the **Variable** list; select *WS* from the **Value** list; select **Include** from the **Mode;** click **Add.** This selection means that the subset includes all values with Season equal to *WS*.



- To exclude the *CONTROL* treatments from the first subset; select *TREAT$* from the **Variable** list; select *CONTROL* from the **Value** list; select **Exclude** from the **Mode**; and click **Add**. This subset contains all the data with Season = WS and all treatments except the Control. Then click **OK** to close the Data Selection window.



- Click **OK** in the Option Window.

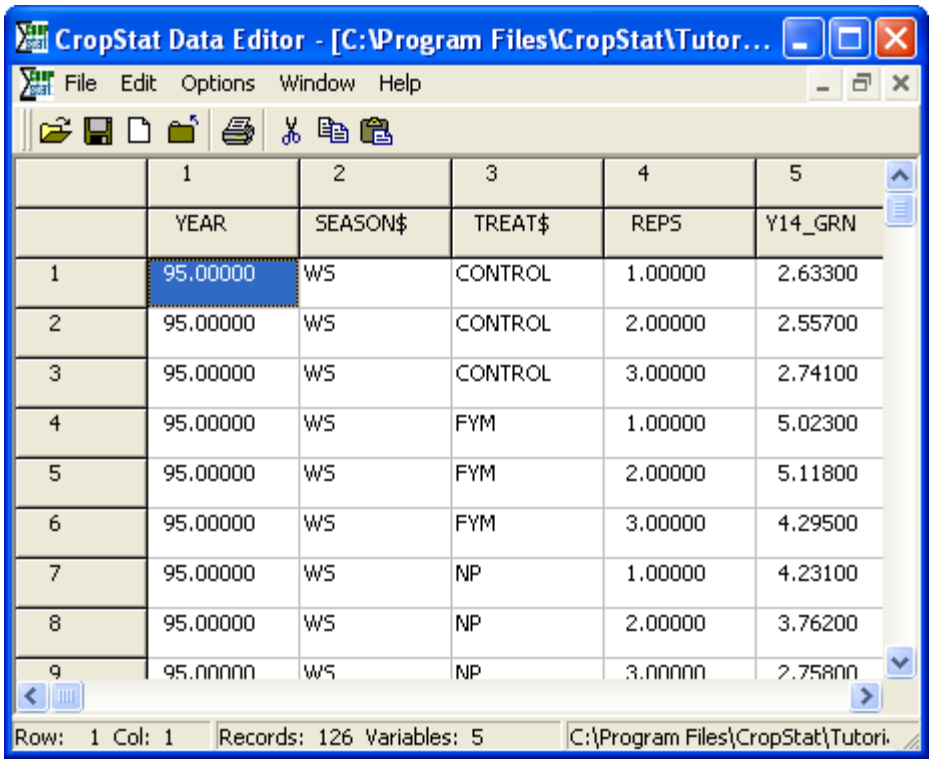
- Sample output

```
 BALANCED ANOVA FOR VARIATE  Y14_GRN  FILE FERT1    1/10/ 4  8:47
------------------------------------------------------------------ :PAGE   1
VARIATE V004 Y14_GRN
15 DATA RECORDS SELECTED FROM FILE FERT1
INCLUDE RECORDS WITH   SEASON$   (  1) EQUAL TO WS
EXCLUDE RECORDS WITH    TREAT$   (  2) EQUAL TO CONTROL

LN   SOURCE OF VARIATION            DF  SUMS OF     MEAN     F RATIO PROB  ER
                                        SQUARES     SQUARES                LN
=============================================================================
1 REPS                               2 .245583     .122791       0.57 0.589  3
2 TREAT$                             4 7.54391     1.88598       8.80 0.005  3
* RESIDUAL                           8 1.71489     .214361
-----------------------------------------------------------------------------
* TOTAL (CORRECTED)                 14 9.50438     .678885
-----------------------------------------------------------------------------
TABLE OF MEANS FOR FACTORIAL EFFECTS  FILE FERT1    1/10/ 4  8:47
------------------------------------------------------------------ :PAGE   2
15 DATA RECORDS SELECTED FROM FILE FERT1
INCLUDE RECORDS WITH   SEASON$   (  1) EQUAL TO WS
EXCLUDE RECORDS WITH    TREAT$   (  2) EQUAL TO CONTROL
MEANS FOR EFFECT REPS
-------------------------------------------------------------------------------

    REPS         NOS      Y14_GRN
1                  5        4.71440
2                  5        4.74540
3                  5        4.45980

SE(N=   5)                 0.207056
5%LSD   8DF                0.675189

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

MEANS FOR EFFECT TREAT$
-------------------------------------------------------------------------------

   TREAT$        NOS      Y14_GRN
FYM                3        4.81200
NP                 3        3.58367
NP+FYM             3        5.60067
NPK                3        4.13067
NPK+FYM            3        5.07233

SE(N=   3)                 0.267308
5%LSD   8DF                0.871665
-------------------------------------------------------------------------------

ANALYSIS OF VARIANCE SUMMARY TABLE  FILE FERT1    1/10/ 4  8:47
------------------------------------------------------------------ :PAGE   3
15 DATA RECORDS SELECTED FROM FILE FERT1
INCLUDE RECORDS WITH   SEASON$   (  1) EQUAL TO WS
EXCLUDE RECORDS WITH    TREAT$   (  2) EQUAL TO CONTROL

 F-PROBABLIITY VALUES FOR EACH EFFECT IN THE MODEL. SECTION - 1

 VARIATE       GRAND MEAN  STANDARD   DEVIATION  C OF V |REPS    |TREAT$  |
               (N=   15)   -------------------- SD/MEAN |        |        |
              NO.           BASED ON   BASED ON    %    |        |        |
              OBS.          TOTAL SS   RESID SS         |        |        |
 Y14_GRN       15  4.6399    0.82394    0.46299     10.0 0.5893   0.0054
```

B. Multiple Subsets

Multiple Subsets is used to define multiple selections on a single SYS file. For example if a trial is repeated at several sites or over several seasons, the data for all repetitions could be entered in a single file, each identified by one or more categorical variables. Data subsets are defined by specifying at most 4 categorical variables. The analysis is repeated for each subset defined.

In this example, two variables such as *Year* and *Season* will be used for creating a subset. *FERT2.SYS* contains the data of a fertilizer experiment from 1995 to1998 for dry and wet seasons. The experiment has 6 treatments and 3 replicates. An Analysis of Variance will be done for each Year × Season subset.



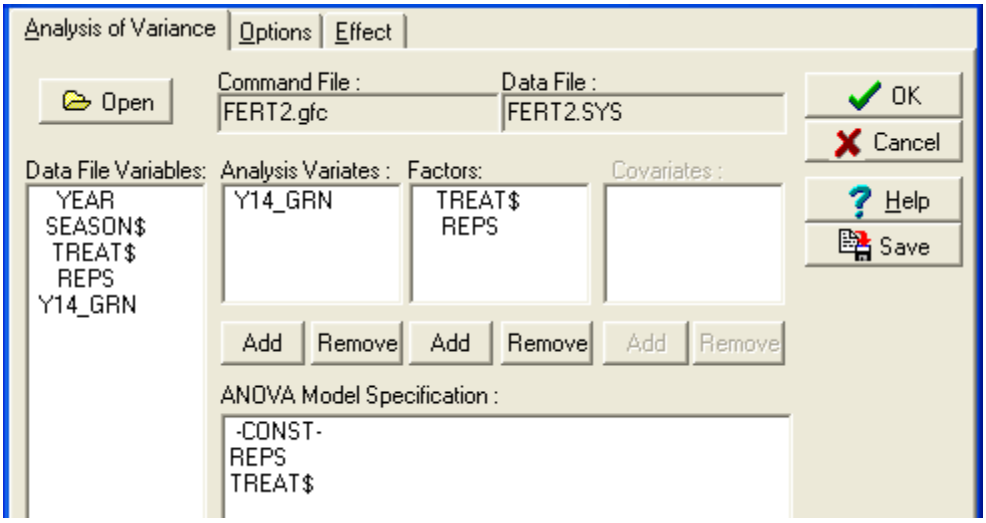
	1	2	3	4	5
	YEAR	SEASON$	TREAT$	REPS	Y14_GRN
1	95.00000	WS	CONTROL	1.00000	2.63300
2	95.00000	WS	CONTROL	2.00000	2.55700
3	95.00000	WS	CONTROL	3.00000	2.74100
4	95.00000	WS	FYM	1.00000	5.02300
5	95.00000	WS	FYM	2.00000	5.11800
6	95.00000	WS	FYM	3.00000	4.29500
7	95.00000	WS	NP	1.00000	4.23100
8	95.00000	WS	NP	2.00000	3.76200
9	95.00000	WS	NP	3.00000	2.75800

- Select **File ⇒ Open.** Go to *C:\PROGRAM FILES\ CROPSTAT7.2\TUTORIAL\ \TUTORIAL DATASETS* and open Fert2.SYS.

- Select on **File ⇒ Save as.** Click the **Save in** box and go inside the directory C:\MY CROPSTAT\ BALANCED ANOVA and save FERT2.SYS.

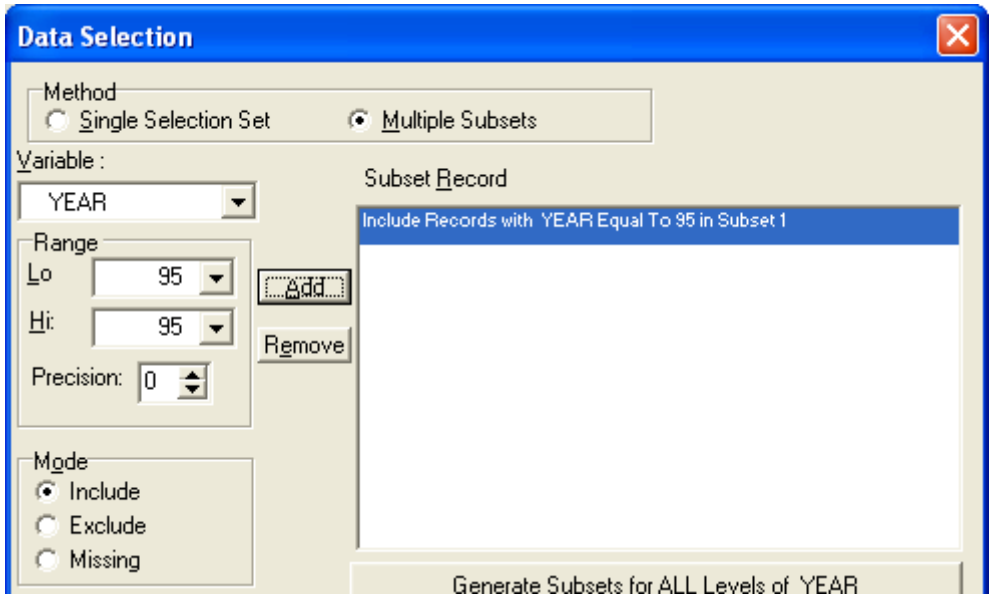
- Select **Analysis|Balanced ANOVA** from the Main Window.
- The **Open** dialog box will prompt you to enter a name for the command file. Click the **Look In** box to go to your working drive *C:\MY CROPSTAT\BALANCED ANOVA*.
- Enter *FERT2* in the **File name** box. Click **Open** button. Since *FERT2.gfc* does not exist, a message box will appear confirming if you want to create one. Click **Yes**.

- Enter the name of the data file to be used. Enter *FERT2.SYS* in the **File name** box.
- Click **Open**. The **Analysis of Variance** dialog box will appear. Specify *Y14_GRN* as the Analysis Variates; *TREAT$* and *REPS* as the Factors. Note: *YEAR* and *SEASON$* are not included as factors.
- From the Factors list box, add *REPS* and *TREAT$* in the ANOVA Model Specification box.



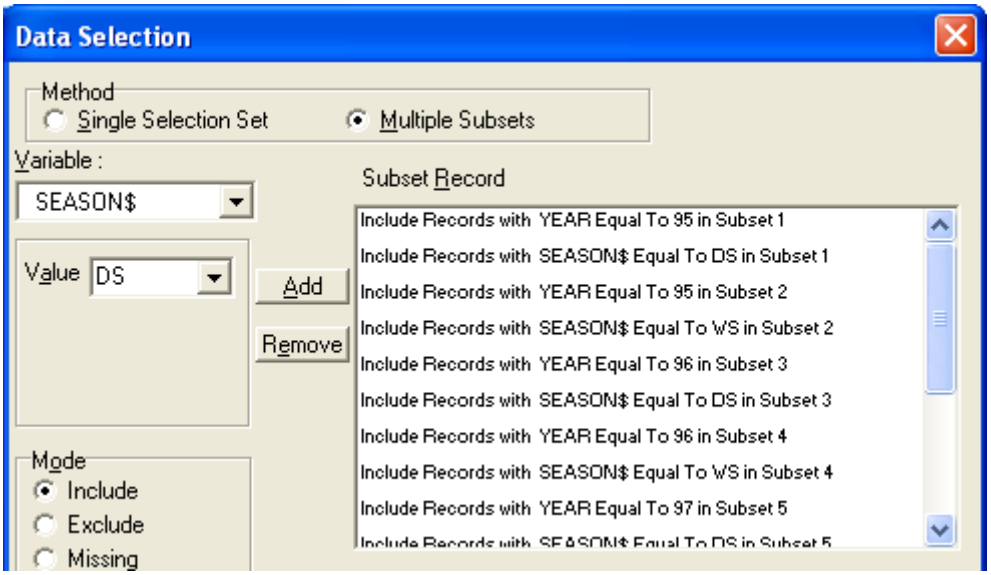
- Click **Options|Data Selection.**

- Click on **Multiple Subsets;** select *YEAR* from the **Variable** list; select lower bound of the Year **(click the Lo and select** *95***)**; select **Include** from the **Mode;** click **Add.** Note: the first level of the first subset variable should be selected to generate the levels. If the first level is not selected, the desired subsets will be individually produced using the **Subset** counter box.


Data Selection
Method
Single Selection Set
Multiple Subsets
Variable :
YEAR
Range
Lo
95
Hi:
95
Precision:
0
Add
Remove
Subset Record
Include Records with YEAR Equal To 95 in Subset 1
Mode
Include
Exclude
Missing
Generate Subsets for ALL Levels of YEAR

- To select the season, click *SEASON$* under the **Variable** list window and click **Add**.

- Then click on **Generate Subsets for ALL Levels of YEAR SEASON$.** The **Data Selection** produced 8 subsets which stands for Year × Season combinations.



- Click **OK** in the **Option** window to run the analysis.

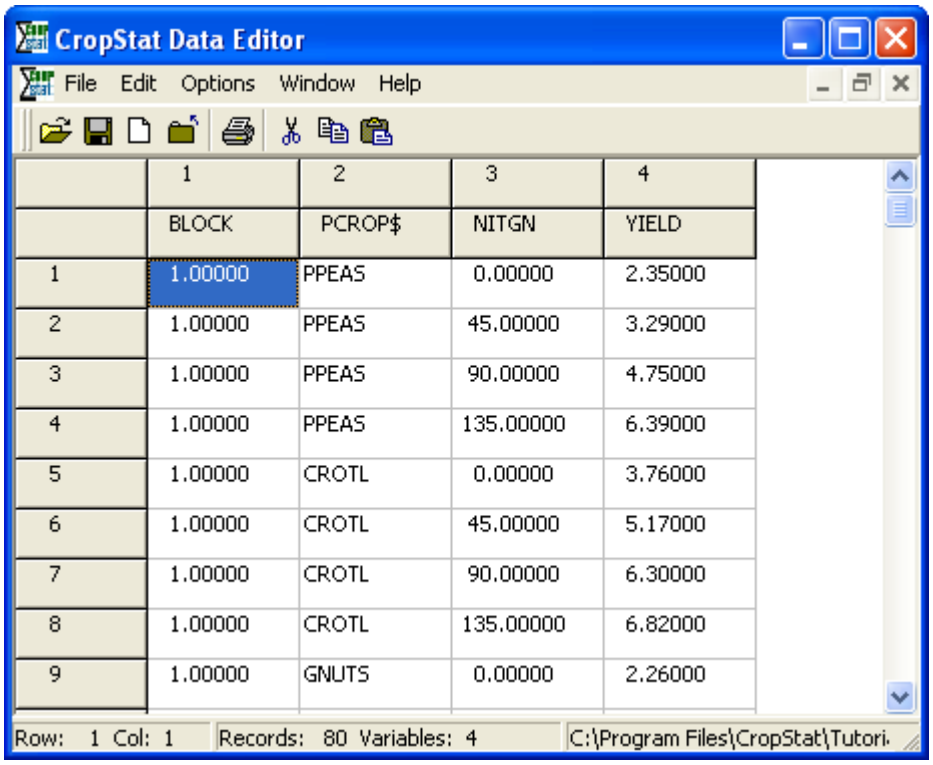
- Sample output (partial)

```
BALANCED ANOVA FOR VARIATE  Y14_GRN  FILE FERT2    1/10/ 4  9:55
------------------------------------------------------------------ :PAGE   1
VARIATE V005 Y14_GRN
18 DATA RECORDS SELECTED FROM FILE FERT2
INCLUDE RECORDS WITH     YEAR    (  1) BETWEEN 95.00     AND 95.00
INCLUDE RECORDS WITH   SEASON$   (  2) EQUAL TO DS

LN   SOURCE OF VARIATION            DF  SUMS OF     MEAN     F RATIO PROB  ER
                                        SQUARES     SQUARES               LN
=============================================================================
1 REPS                               2 .798162E-01 .399081E-01   0.29 0.757  3
2 TREAT$                             5 17.4011     3.48022      25.33 0.000  3
* RESIDUAL                          10 1.37407     .137407
-----------------------------------------------------------------------------
* TOTAL (CORRECTED)                 17 18.8550     1.10912
-----------------------------------------------------------------------------
TABLE OF MEANS FOR FACTORIAL EFFECTS  FILE FERT2    1/10/ 4  9:55
------------------------------------------------------------------ :PAGE   2
18 DATA RECORDS SELECTED FROM FILE FERT2
INCLUDE RECORDS WITH     YEAR    (  1) BETWEEN 95.00     AND 95.00
INCLUDE RECORDS WITH   SEASON$   (  2) EQUAL TO DS
MEANS FOR EFFECT REPS
-----------------------------------------------------------------------------

    REPS         NOS      Y14_GRN
1                  6        2.93883
2                  6        3.10167
3                  6        3.02850

SE(N=   6)                 0.151331
5%LSD  10DF                0.476851
-----------------------------------------------------------------------------

MEANS FOR EFFECT TREAT$
-----------------------------------------------------------------------------

   TREAT$        NOS      Y14_GRN
CONTROL            3        1.45800
FYM                3        1.95067
NP                 3        3.68167
NP+FYM             3        4.05467
NPK                3        3.13967
CR                 3        3.85333

SE(N=   3)                 0.214015
5%LSD  10DF                0.674369
-----------------------------------------------------------------------------

ANALYSIS OF VARIANCE SUMMARY TABLE  FILE FERT2    1/10/ 4  9:55
------------------------------------------------------------------ :PAGE   3
18 DATA RECORDS SELECTED FROM FILE FERT2
INCLUDE RECORDS WITH     YEAR    (  1) BETWEEN 95.00     AND 95.00
INCLUDE RECORDS WITH   SEASON$   (  2) EQUAL TO DS

 F-PROBABLIITY VALUES FOR EACH EFFECT IN THE MODEL. SECTION - 1

 VARIATE       GRAND MEAN  STANDARD   DEVIATION  C OF V |REPS    |TREAT$  |
               (N=   18)   ------------------- SD/MEAN |        |        |
              NO.          BASED ON   BASED ON    %    |        |        |
              OBS.         TOTAL SS   RESID SS         |        |        |
 Y14_GRN       18  3.0230     1.0531    0.37068     12.3 0.7568   0.0000
```

VII. Analysis of Variance for Split-plot design

A. Data

The data were taken from a trial carried out by the TLU at Maroua (Talleyrand, H. 1991) to examine the effect of previous crop on response to nitrogen in a following crop of maize. The ASC file *SPLIT.TXT* in the *CROPSTAT7.2\TUTORIAL\TUTORIAL DATASETS* contains the data for this trial. Importing the file using the Data Editor and save inside the MY CROPSTAT\BALANACED ANOVA, a partial listing of the file of the *SPLIT.SYS* is given below.



	1	2	3	4
	BLOCK	PCROP$	NITGN	YIELD
1	1.00000	PPEAS	0.00000	2.35000
2	1.00000	PPEAS	45.00000	3.29000
3	1.00000	PPEAS	90.00000	4.75000
4	1.00000	PPEAS	135.00000	6.39000
5	1.00000	CROTL	0.00000	3.76000
6	1.00000	CROTL	45.00000	5.17000
7	1.00000	CROTL	90.00000	6.30000
8	1.00000	CROTL	135.00000	6.82000
9	1.00000	GNUTS	0.00000	2.26000

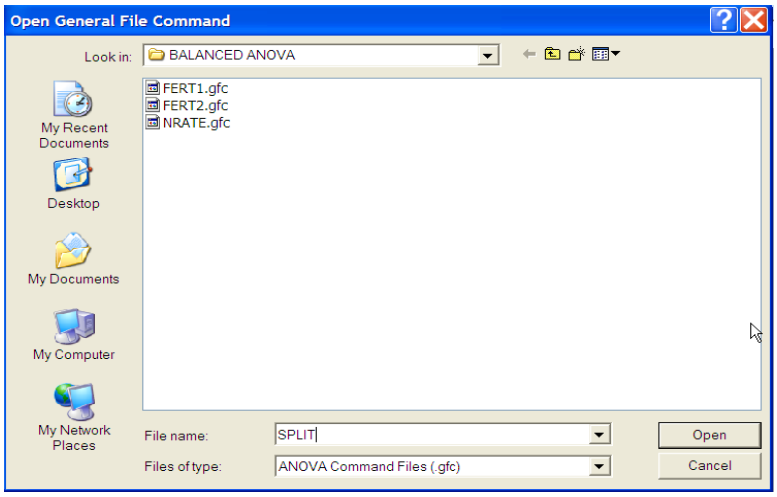
B. Analysis in CropStat

For a split-plot design, the outline of the ANOVA table is given below.

SOURCE	**DF**
BLOCK	3
PCROP$	4
Error A	12
NITGN	3
PCROP$ * NITGN	12
Error B	45
TOTAL	**79**

The effect of PCROP$ must be tested against Error A, the third line in the ANOVA table, and the other treatment effects against Error B. It is important to note that for a split-plot design in randomized complete blocks, Error A is formally equivalent to the BLOCK * PCROP$ interaction.

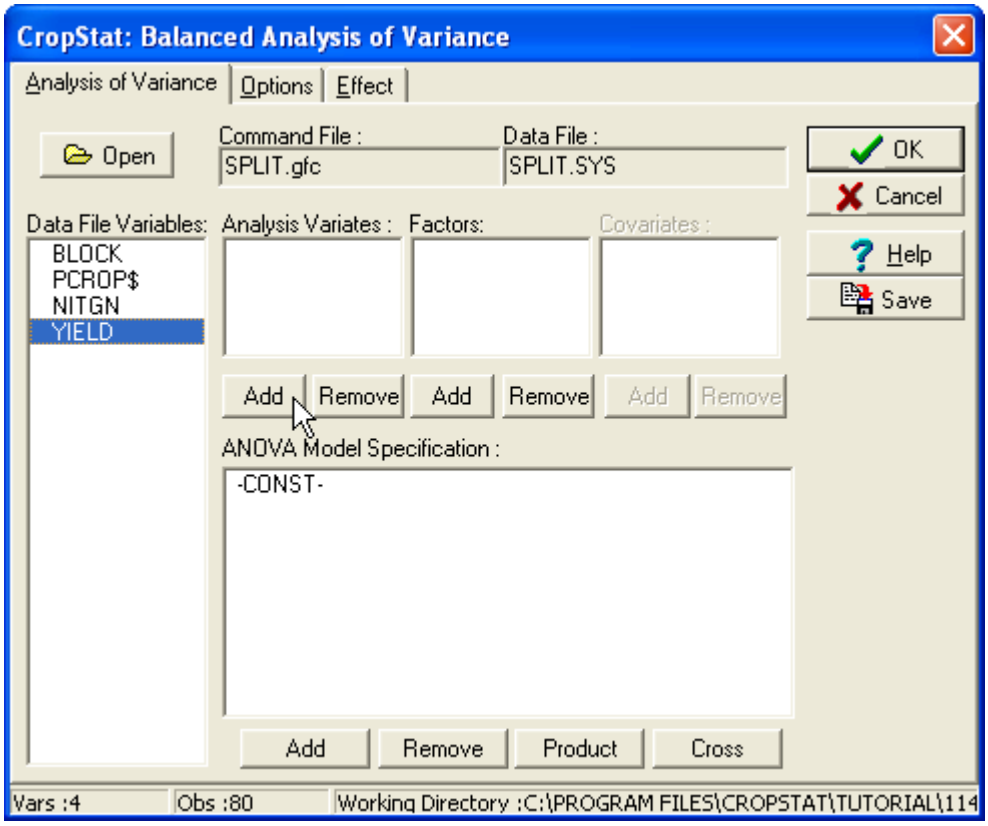
- Select **Analysis|Balanced ANOVA** from the Main Window.
- In the **Open** dialog, click the **Look In** box and go to drive *C*. Double-click MY *CROPSTAT\BALANCED ANOVA*. In the **File name** edit box, type *SPLIT*.



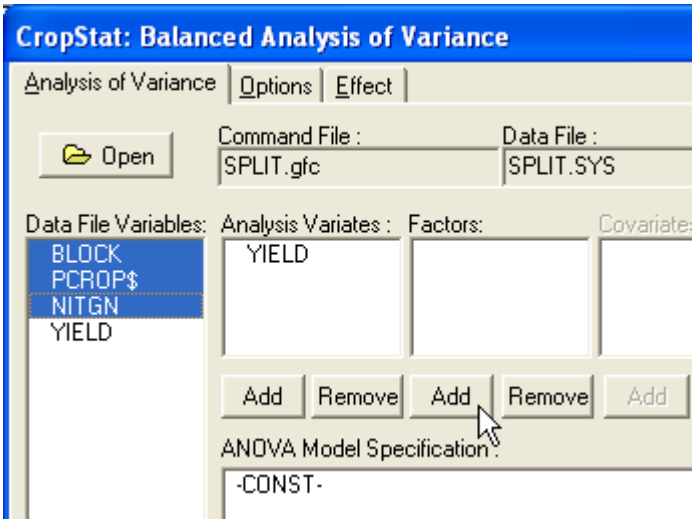
- Since no command file *SPLIT.GFC* exists, CropStat will ask if the user wants to create that command file. Click the **Yes** button.

- In the **Open** dialog, select *SPLIT.SYS*. Click **Open**.

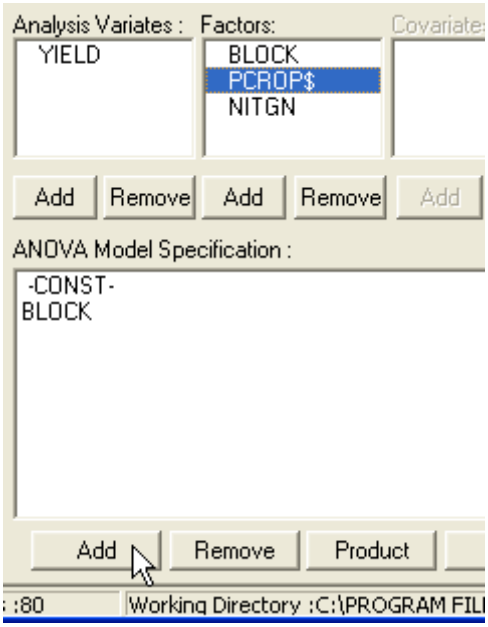
- In the **Analysis of Variance** dialog, select *YIELD* from **Data File Variables** list. Click **Add** under **Analysis Variates** list.



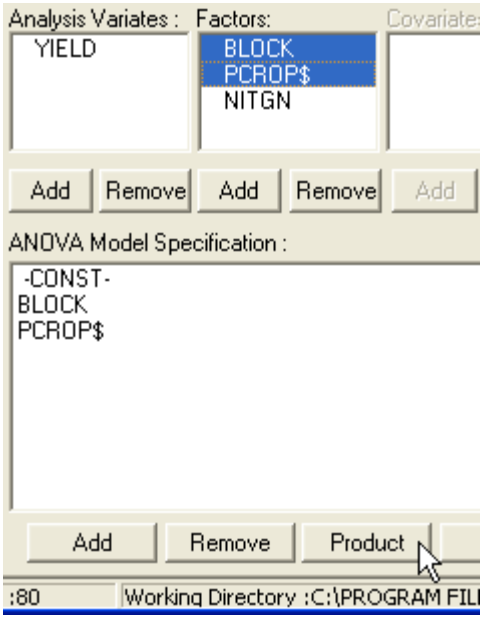
- Select *BLOCK*, *PCROP,* AND *NITGN* from **Data File Variables** list. Click **Add** under **Factors** list.



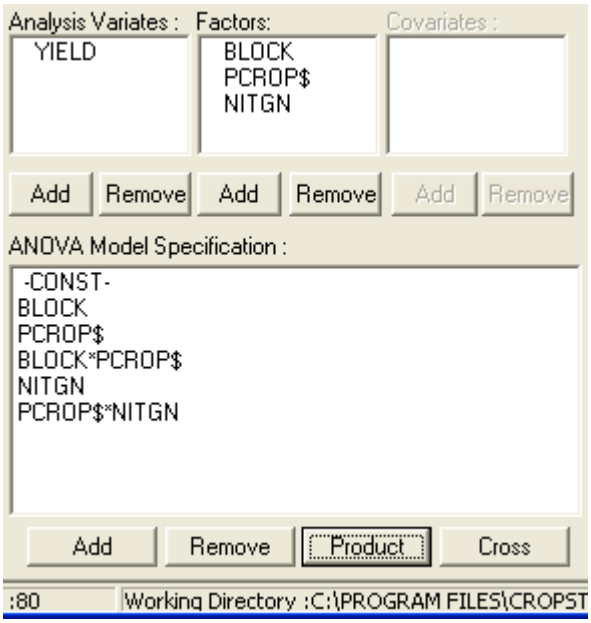
- Select *BLOCK* from **Factors** list and click **Add** under the **ANOVA Model Specification** list. Do the same to add *PCROP* to the model.



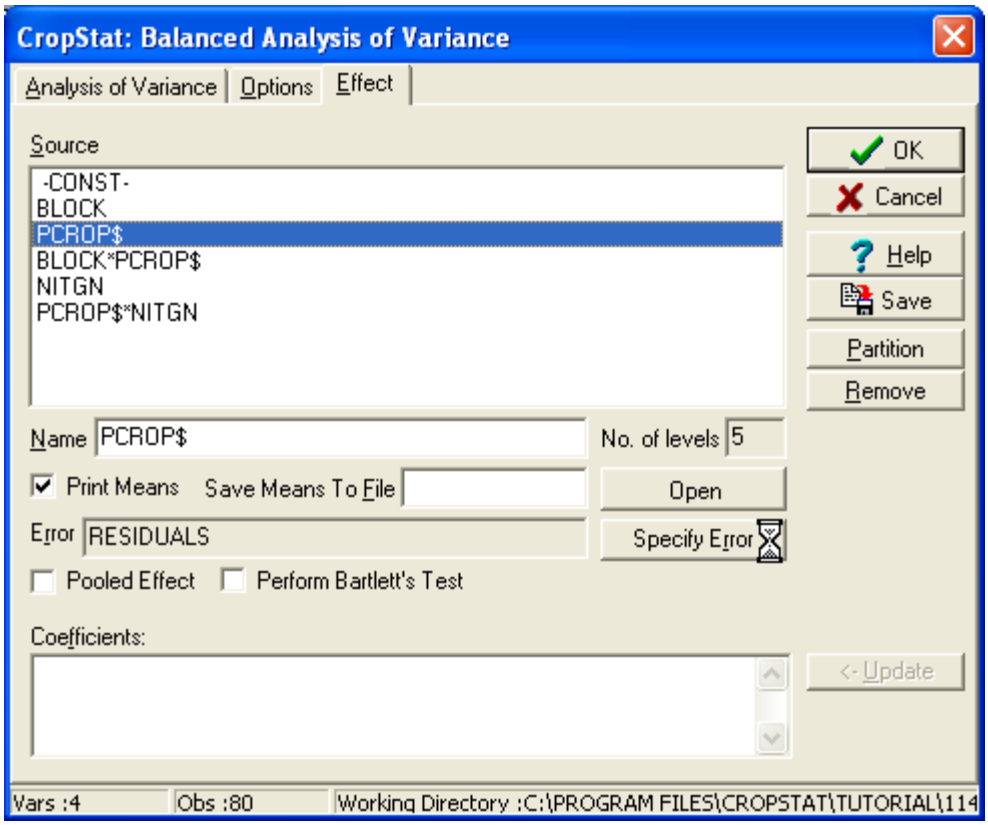
- Select *BLOCK* and *PCROP* from the **Factors** list and click the **Product** button under the **ANOVA Model Specification** list.



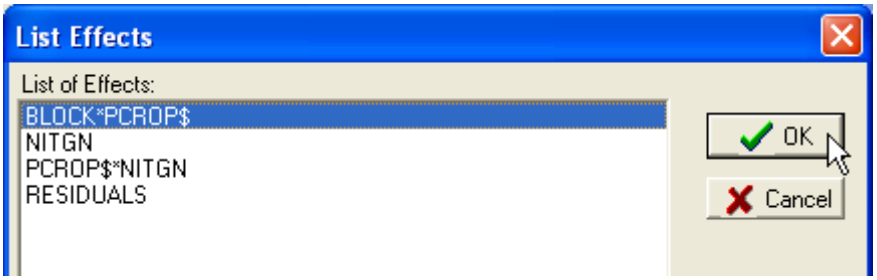
- Repeat the previous steps to enter *NITGN* and *NITGN*PCROP* in the model.



- By default, CropStat tests *PCROP* against Error B or the Residuals Error. To change the error term for *PCROP*, click the **Effect** tab. Select *PCROP* from the **Source** list. Click the **Specify Error** button.

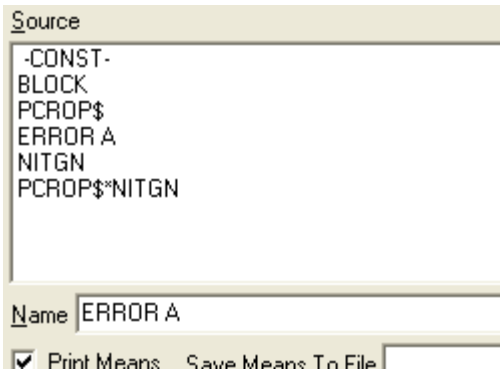


- Select *BLOCK*PCROP$* from the **List of Effects**. Click **OK** to proceed.

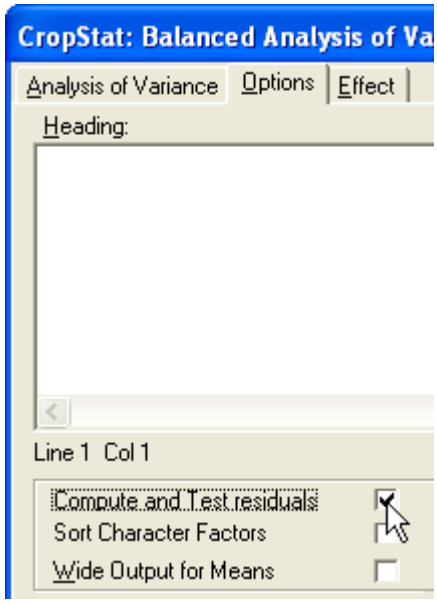


- You may want to change the name of *BLOCK*PCROP$* effect to *Error A*. To do so, select *BLOCK*PCROP$* from the **Source** list.

- Click the **Name** edit box and type *Error A*.



- To request for Test of Residuals, click the **Options** tab. Click the **Compute and Test Residuals** check box and click **OK** to run the analysis.



- Sample Output (Partial)

  - ANOVA Table

```
 BALANCED ANOVA FOR VARIATE    YIELD  FILE SPLIT    1/10/ 4 10:56
------------------------------------------------------------------- :PAGE   1
VARIATE V004 YIELD

LN   SOURCE OF VARIATION            DF  SUMS OF     MEAN     F RATIO PROB  ER
                                        SQUARES     SQUARES               LN
=============================================================================
1 BLOCK                              3 .694944     .231648       0.45 0.724  3
2 PCROP$                             4 30.7157     7.67892      14.97 0.000  3
3 Error A                           12 6.15679     .513066       3.55 0.001  6
4 NITGN                              3 93.7898     31.2633     216.10 0.000  6
5 PCROP$*NITGN                      12 5.87911     .489926       3.39 0.001  6
* RESIDUAL                          45 6.51013     .144670
-----------------------------------------------------------------------------
* TOTAL (CORRECTED)                 79 143.746     1.81958
-----------------------------------------------------------------------------
BLOCK NITGN                           9   2.0074       1.78 0.105
BLOCK PCROP$ NITGN                   36   4.5027       0.56 0.895

TUKEY'S TEST FOR TRANSFORMABLE NON-ADDITIVITY
SS= 0.268378E-01, F(1, 44)=   0.18, P= 0.675
REGRESSION SLOPE=-0.31749E-01 SUGGESTED POWER TRANSFORMATION=  1.3030

MEAN ORTHOGONAL RESIDUAL = -0.1129    , P-VALUE= 0.050
```

  - Scatter plot

```
PLOT OF LS RESIDUALS AGAINST FITTED VALUES
          .............................................................
          :                                                           :
          :                                          *                :
          :                                                           :
          :                                                           :
     0.6 -:                                  *                        :
          :                  *                                        :
          :          *    *         *     *                   *       :
          :          *       *           *            *   *           :
          :                                           *    *          :
     0.2 -:    *               *        *     *     * 2               :
          :          *                     * ** *    *     * *  *  *   :
          :                                *    *   2     *     **     :
          :           *      * *    *         2     *   * *  2         :
          :                         *         *              *        :
    -0.2 -:           *                           *  *   *       *     :
          :                *          **   *   **    *2    *          *  :
          :            **      *  *   *                  2              :
          :                                 *          *                :
          :                             *                             :
    -0.6 -:     *                                                       :
          :...........................................................:
           :         :         :         :         :         :
          1.5       2.5       3.5       4.5       5.5       6.5
THE FOLLOWING RECORDS HAVE LARGE RESIDUALS, (>2.5 SES)

   BLOCK        PCROP$       NITGN      OBS. VALUE  FITTED VALUE  RESIDUAL
1           PPEAS       135              6.390        5.554       0.8362

BOX PLOT OF STUDENTIZED RESIDUALS FROM LPLT= -1.983    TO ULPT=  2.931
NO.<LPLT                                                      NO.>UPLT
       0 -------------I          +      I------------------         *   0

MEDIAN= -0.1205E-01 ANDERSON-DARLING STATISTIC=  0.457
```

Exercise 3

3.1 A CropStat file (NUTRIENT) has the following file description:

FACTORS:

ID	NAME	LEVEL
R	REPLICATION	6
N	NUTRIENT LEVEL	4

NAMES OF FACTOR LEVELS:

NUTRIENT LEVEL (N):

1 25 mg
2 75 mg
3 125 mg
4 175 mg

VARIABLE: Tiller Number
Plant Height

Perform an appropriate analysis of variance on Tiller Number and Plant Height.

3.2 A CropStat file (**INFEST**) has the following file description:

FACTORS:

ID	NAME	LEVEL
R	REP	4
L	INFESTATION LEVEL	5
V	VARIETY	4

NAMES OF FACTOR LEVELS:

INFESTATION LEVEL (L):

1 50 LARVAE/HILL
2 100 LARVAE/HILL
3 200 LARVAE/HILL
4 400 LARVAE/HILL
5 800 LARVAE/HILL

VARIETY (V)

1 IR1820
2 IR36
3 IR40
4 IR54

VARIABLE: YIELD LOSS (G/M2)

Perform an analysis of variance and appropriate mean comparison. The analysis of variance should have the following format:

Source of Variation	**df**	
REP	3	
VARIETY (V)	3	
GROUP1: IR1820,IR54		1
GROUP2: IR36, IR40		1
GROUP1 VS GROUP2		1
ERROR (a)	9	
INFESTATION LEVEL (L)	4	
LINEAR		1
QUADRATIC		1
CUBIC		1
RESIDUAL		1
V × L	12	
GROUP1 × LINEAR		1
GROUP1 × QUADRATIC		1
GROUP1 × CUBIC		1
GROUP2 × LINEAR		1
GROUP2 × QUADRATIC		1
GROUP2 × CUBIC		1
(GROUP1 VS GROUP2) × LINEAR		
(GROUP1 VS GROUP2) × QUADRATIC		1
(GROUP1 VS GROUP2) × CUBIC		1
RESIDUAL		3
ERROR (b)	48	

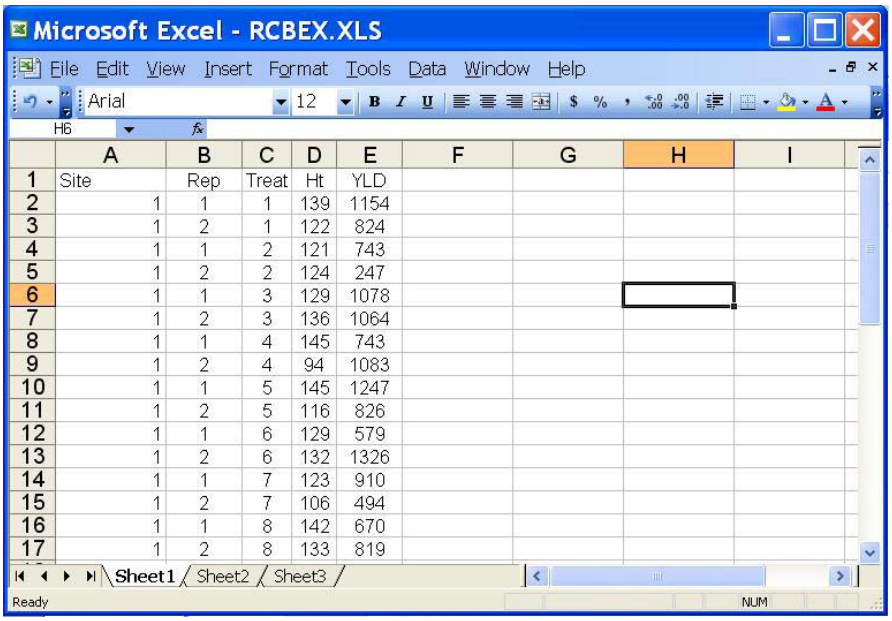
LINEAR MIXED MODELS

At the end of the tutorial, the user should be able to

- perform mixed model analysis on experimental data with random and fixed effects
- perform analysis with multiple data selection

I. Analysis of Randomized Complete Block Design

- For our example, we want to analyze the following Excel file, RCBEX.XLS, stored in *CROPSTAT\TUTORIAL\TUTORIAL DATASETS* folder.

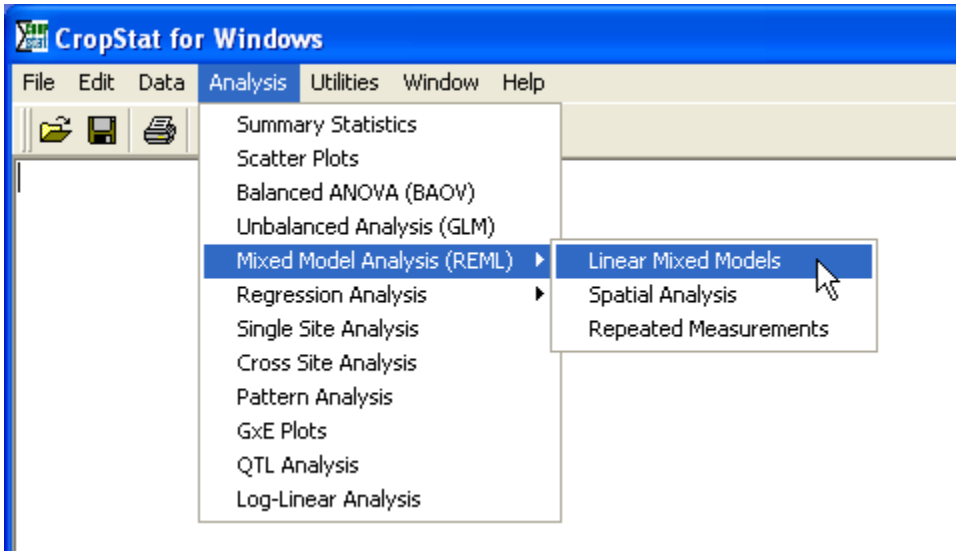


	A	B	C	D	E
1	Site	Rep	Treat	Ht	YLD
2	1	1	1	139	1154
3	1	2	1	122	824
4	1	1	2	121	743
5	1	2	2	124	247
6	1	1	3	129	1078
7	1	2	3	136	1064
8	1	1	4	145	743
9	1	2	4	94	1083
10	1	1	5	145	1247
11	1	2	5	116	826
12	1	1	6	129	579
13	1	2	6	132	1326
14	1	1	7	123	910
15	1	2	7	106	494
16	1	1	8	142	670
17	1	2	8	133	819

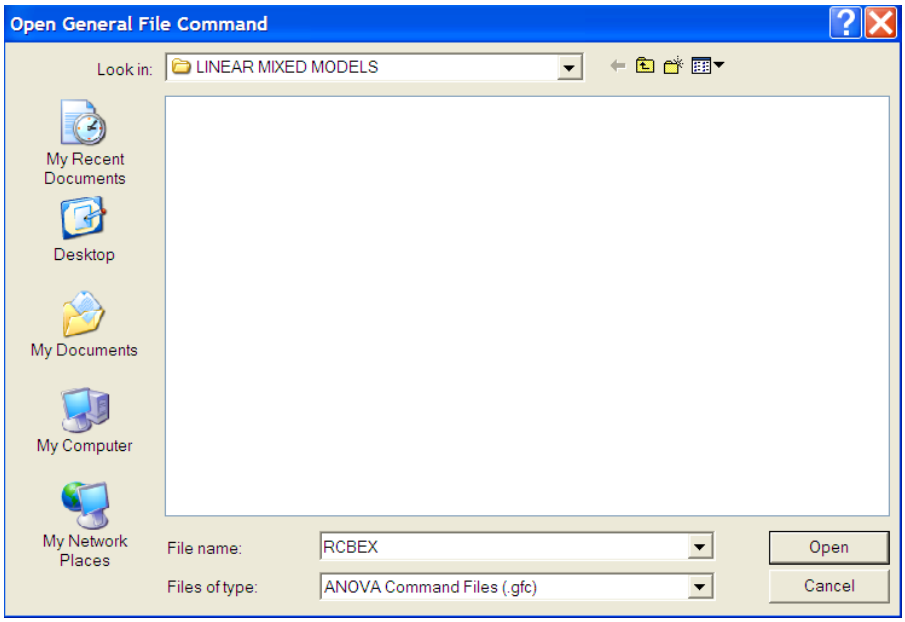
- Convert RCBEX.XLS to .SYS file, before this data can be analyzed in CropStat. To import an Excel file into Cropstat follow the procedure given in the Data and File management module.

- Create a subfolder LINEAR MIXED MODEL inside working folder C:\MY CROPSTAT. Save the imported file RCBEX.SYS inside this created folder.

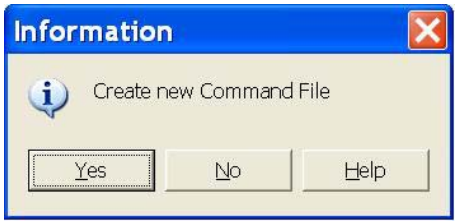
- Go to **Mixed Model Analysis|Linear Mixed Models** from the Analysis menu.



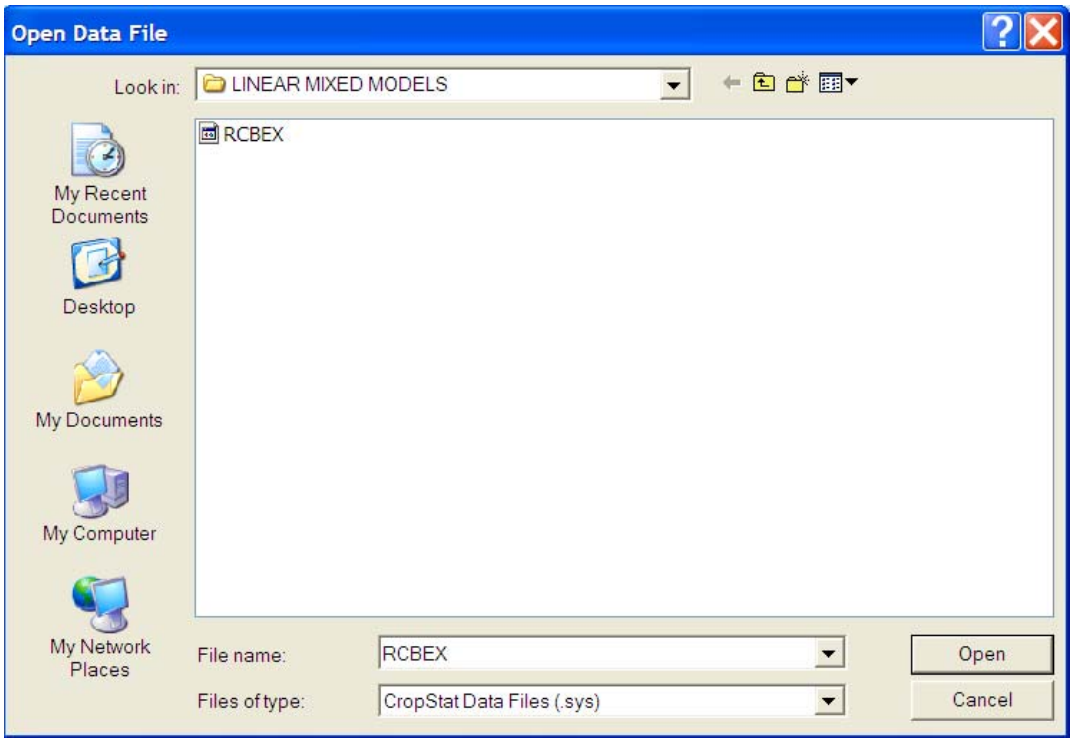
- Click the **Look in** box and go inside your working folder *C:\MY CROPSTAT\LINEAR MIXED MODEL*. Specify a filename for your general file command (GFC) as RCBEX. Cropstat uses .gfc as the default file type.



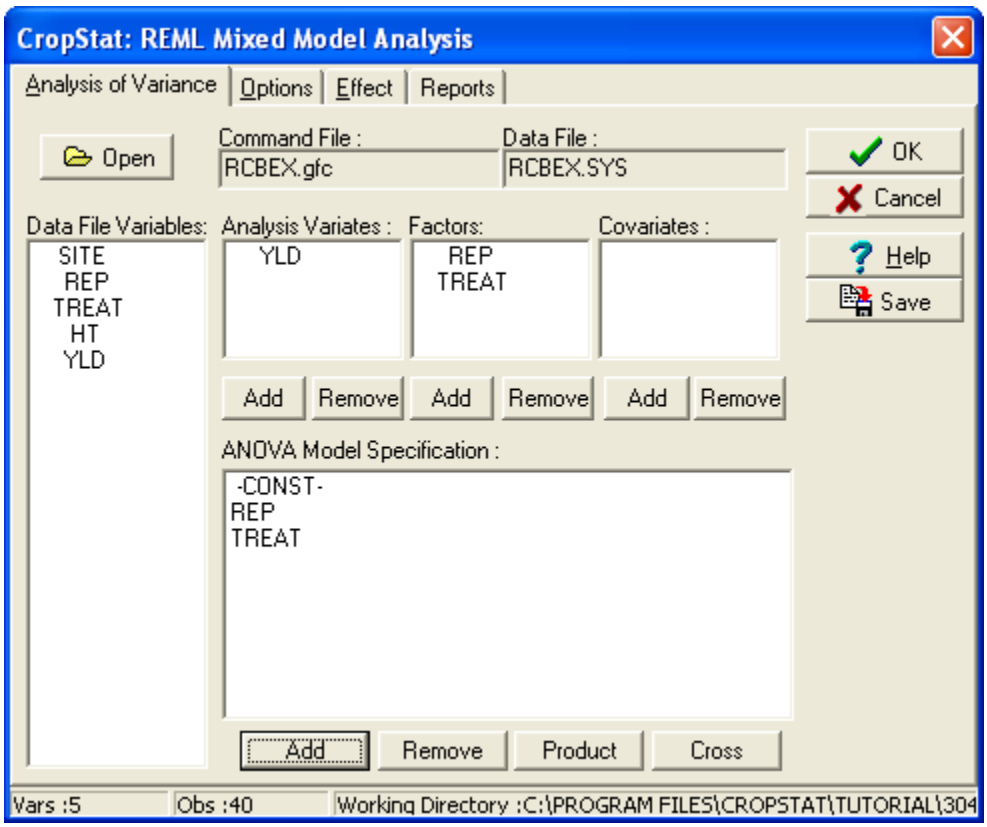
- Click **Yes** to create the file.



- Open the SYS file to be used for the analysis.

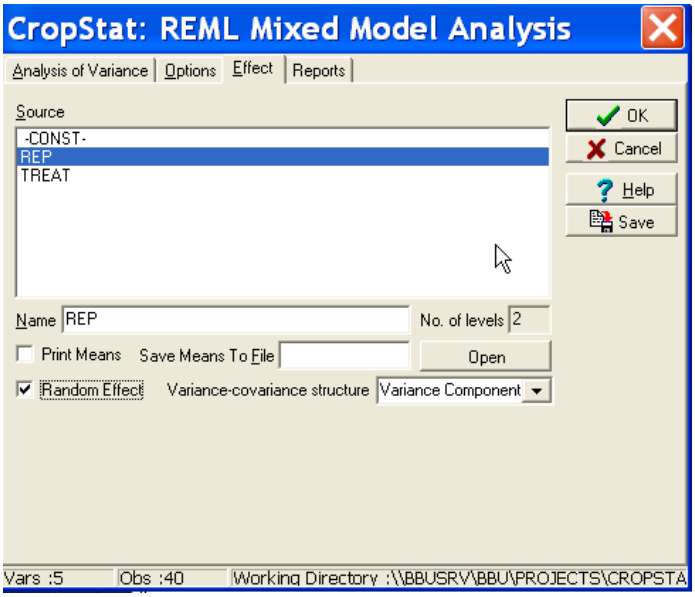


- The **REML Mixed Model Analysis** dialog box will appear. Fill-in the details needed for the analysis.



- From the **Data File Variables** list; highlight all variables to be analyzed then **Add** to the **Analysis Variates** box; and highlight the treatment and block variables then **Add** to the **Factors** box.
- From the **Factors** box, highlight the treatment and block variables then **Add** to the **ANOVA Model Specification** box.
- **Click Effect** to go to the Effect window.

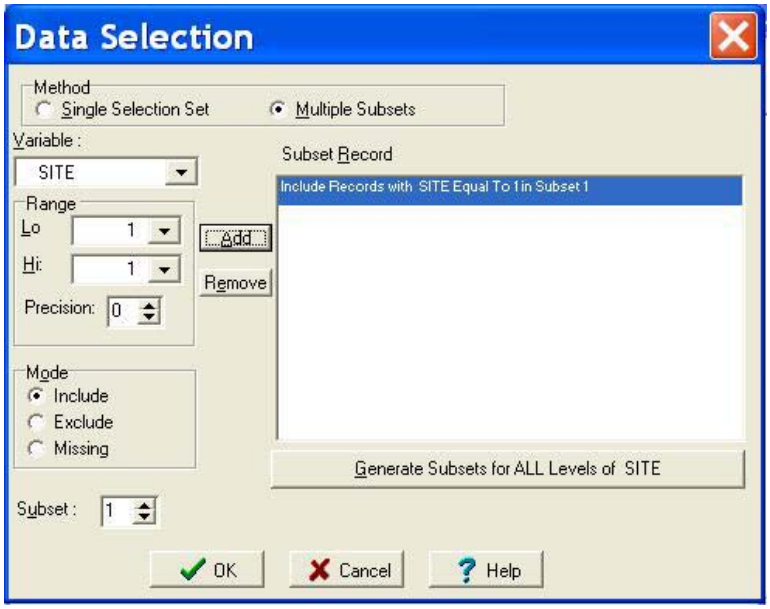
- Specify which factors are random effects. Highlight the **Block** factor then click the **Random Effect** check box to indicate **Block** as a random effect.



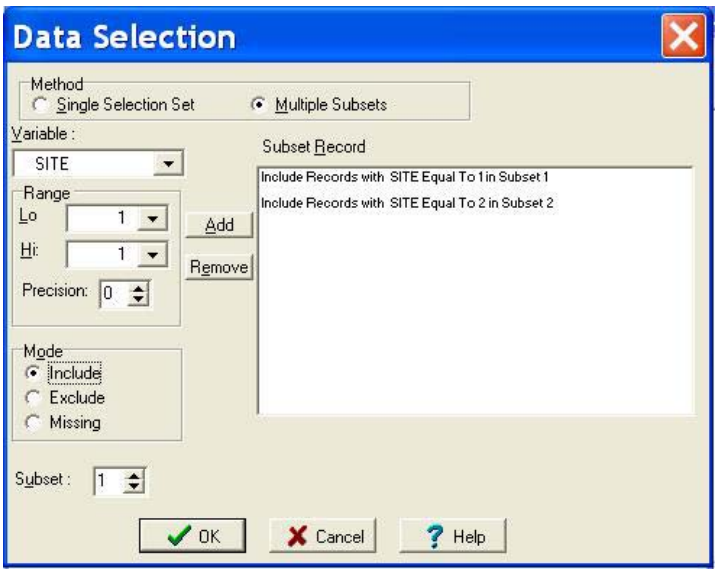
- Click the Print Means check box if you want to suppress the printing of the block means. The default is to print the means of all effects in the model.

- To do analysis for each site, click **Options** to go the Options window. Create subsets of the data to perform separate analyses for each site.

- Click on **Multiple Subsets,** select Site from the **Variable** list and click add; Then click on **General Subsets for ALL Levels of Site.** The **Data Selection** produced 2 subsets for Site**.**



- Click **Ok** and you will be back in the **Options** Window. Click **Ok** and the output will appear in the Text Editor.



II. Sample Output

```
IRREML(V1.0) - REML ANALYSIS  FILE RCBEX    8/ 2/ 5 15: 5
------------------------------------------------------------------ :PAGE   1
THE IRREML PROGRAM WAS WRITTEN BY DOUGLAS CLARKSON OF SCIENCEOPS FOR IRRI
DATA RECORDS SELECTED FROM FILE RCBEX
INCLUDE RECORDS WITH     SITE    (  1) BETWEEN 1.000     AND 1.000

Command File: C:\MY CROPSTAT\LINEAR MIXED MODELS\RCBEX.gfc   Data File: RCBEX


Data File: RCBEX

Number of Records:    20 Non missing observations:    20

Number of Columns in the Fixed Effects Model:    10

Number of columns in the random effects model:     2

Variables in Data Set: REP  TREAT  YLD

Classification Variables: REP  TREAT

Levels of the classification variables

   2 CODES:(Number  Label) for Variable: REP
(  1            1)(  2            2)(

  10 CODES:(Number  Label) for Variable: TREAT
(  1            1)(  2            2)(  3            3)(  4            4)(  5
5)
(  6            6)(  7            7)(  8            8)(  9            9)( 10
10)

Model Specification

Intercept in model: Yes

The Fixed Effects Model
    YLD = Intercept + TREAT
```

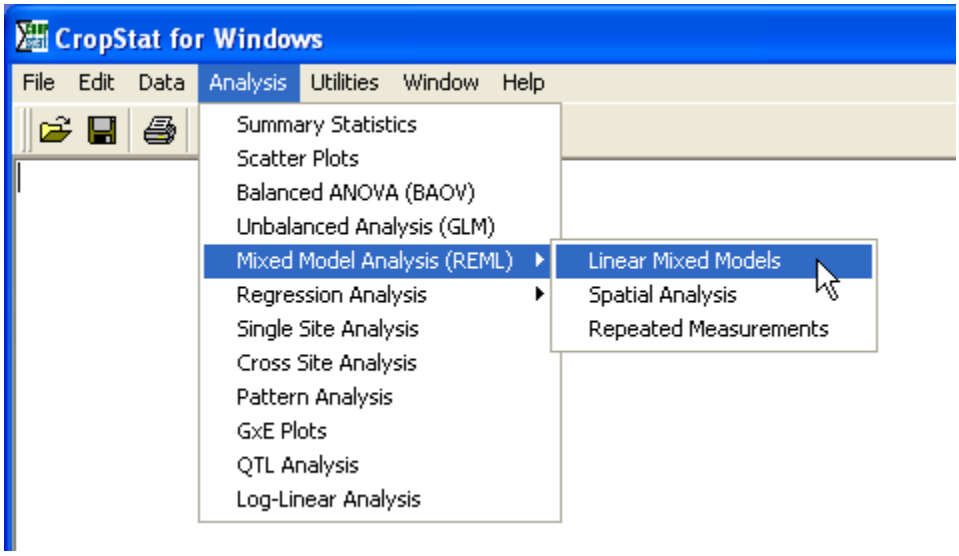
```
 The Random Effects Terms
     REP


 RANDOM EFFECT COVARIANCE MODEL.  1 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 REP                            1- 1  diagonal

 RESIDUAL EFFECT COVARIANCE MODEL.  0 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 RESIDUAL                                 sigmasq(1)xI
 Message: Relative function convergence

 Final REML criterion:      -65.710490060239749

 Variance/Covariance component parameters
 Dep Name       Coef    Std. Error      Z       Pr > |Z|   Component
  1 REP(1) ..  0.1000E-05  0.1576      0.6346E-05   1.000      0.9380E-01

 The scale parameters
 Dep.  Sigma_Squared Std. Error      Z         Pr > |Z|
Dep(1) .......    0.9380E+05  0.4394E+05   2.135      0.3276E-01

 Asymptotic Covariance Matrix of the Variance/Covariance Components
                                  1            2
           1   1 REP(1)..   0.248E-01  -0.206E+04
           2 Dep(1)......  -0.206E+04   0.193E+10

       Fixed Effect ANOVA Table - Partial Wald Tests
       Denominator Degrees of Freedom: Residual DF
 Dep Effect    DFNum   DFDen   F - Statistic   P > |F|
  1    TREAT      9  10.00        1.090       0.4440

       ANOVA Table for Sequentially Deleted Fixed Effects
       Denominator Degrees of Freedom: Residual DF
 Dep Effect    DFNum   DFDen   F - Statistic   P > |F|
  1    TREAT      9  10.00        1.090       0.4440

                 Balanced Least Squares Means Fixed
 Dep Level                LSMean   Std. Error
  1  TREAT            1    989.1       216.6
  1  TREAT            2    495.1       216.6
  1  TREAT            3    1071.       216.6
  1  TREAT            4    913.2       216.6
  1  TREAT            5    1036.       216.6
  1  TREAT            6    952.1       216.6
  1  TREAT            7    702.2       216.6
  1  TREAT            8    744.3       216.6
  1  TREAT            9    915.2       216.6
  1  TREAT           10    1319.       216.6

 Standard Errors of Differences
  Minimum     Mean      Maximum
   306.3       306.3       306.3
```

```
IRREML(V1.0) - REML ANALYSIS  FILE RCBEX    8/ 2/ 5 15: 5
 ------------------------------------------------------------------ :PAGE   2
 THE IRREML PROGRAM WAS WRITTEN BY DOUGLAS CLARKSON OF SCIENCEOPS FOR IRRI
 DATA RECORDS SELECTED FROM FILE RCBEX
 INCLUDE RECORDS WITH     SITE    (  1) BETWEEN 2.000     AND 2.000

 Command File: C:\MY CROPSTAT\LINEAR MIXED MODELS\RCBEX.gfc   Data File: RCBEX


 Data File: RCBEX

 Number of Records:    20 Non missing observations:    20

 Number of Columns in the Fixed Effects Model:    10

 Number of columns in the random effects model:     2

 Variables in Data Set: REP  TREAT  YLD

 Classification Variables: REP  TREAT

 Levels of the classification variables

    2 CODES:(Number  Label) for Variable: REP
(  1            1)(  2            2)(

   10 CODES:(Number  Label) for Variable: TREAT
(  1            1)(  2            2)(  3            3)(  4        4)(  5     5)
(  6            6)(  7            7)(  8            8)(  9        9)( 10    10)

 Model Specification

 Intercept in model: Yes

 The Fixed Effects Model
     YLD = Intercept + TREAT

 The Random Effects Terms
     REP


 RANDOM EFFECT COVARIANCE MODEL.  1 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 REP                           1- 1  diagonal

 RESIDUAL EFFECT COVARIANCE MODEL.  0 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 RESIDUAL                             sigmasq(1)xI
 Message: Relative function convergence

 Final REML criterion:      -70.791176560268866

 Variance/Covariance component parameters
 Dep Name       Coef    Std. Error      Z       Pr > |Z|   Component
  1 REP(1) ..  0.1000E-05  0.1509      0.6626E-05   1.000      0.2591

 The scale parameters
 Dep.  Sigma_Squared Std. Error      Z         Pr > |Z|
Dep(1) .......    0.2591E+06  0.1220E+06   2.125      0.3362E-01

 Asymptotic Covariance Matrix of the Variance/Covariance Components
                                1           2
          1   1 REP(1)..   0.228E-01  -0.574E+04
          2 Dep(1)......  -0.574E+04   0.149E+11

       Fixed Effect ANOVA Table - Partial Wald Tests
       Denominator Degrees of Freedom: Residual DF
```

```
Dep Effect    DFNum   DFDen   F - Statistic   P > |F|
 1    TREAT      9  10.00       0.8615       0.5837

      ANOVA Table for Sequentially Deleted Fixed Effects
      Denominator Degrees of Freedom: Residual DF
Dep Effect    DFNum   DFDen   F - Statistic   P > |F|
 1    TREAT      9  10.00       0.8615       0.5837

                Balanced Least Squares Means Fixed
Dep Level               LSMean   Std. Error
 1  TREAT            1    784.6       360.0
 1  TREAT            2    370.5       360.0
 1  TREAT            3    780.6       360.0
 1  TREAT            4    1027.       360.0
 1  TREAT            5    1395.       360.0
 1  TREAT            6    370.3       360.0
 1  TREAT            7    949.6       360.0
 1  TREAT            8    408.0       360.0
 1  TREAT            9    1030.       360.0
 1  TREAT           10    828.0       360.0

Standard Errors of Differences
 Minimum     Mean      Maximum
  509.0       509.0       509.0
```

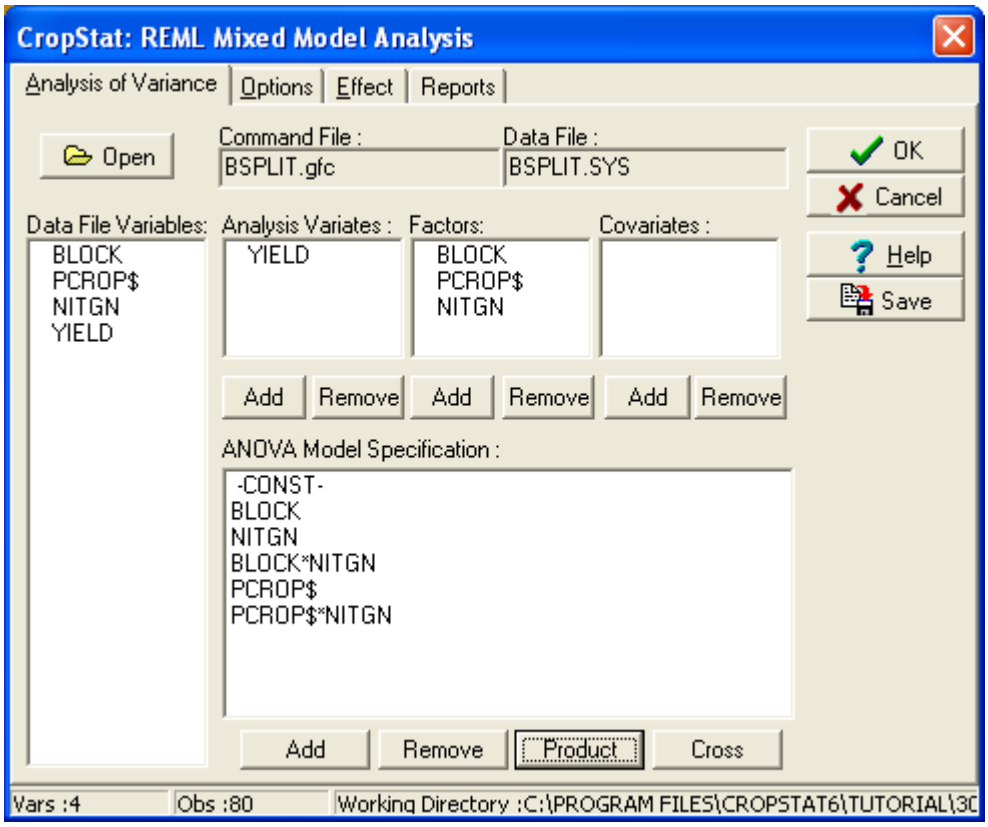
III. Analysis of Split-plot Design

- For our example, we'll use the same data in the Analysis of Variance chapter. Data is stored in BSPLIT.SYS.
- Select **File ⇒ Open.** Go to *C:\PROGRAM FILES\ CROPSTAT7.2\TUTORIAL\ \TUTORIAL DATASETS* and open BSSPLIT.SYS.
- Select on **File ⇒ Save as.** Click the **Save in** box and go inside the directory *C:\MY CROPSTAT\ LINEAR MIXED MODEL* and save BSPLIT.SYS.
- From the **Main Window** go to **Analysis|Mixed Model Analysis|Linear Mixed Models**.

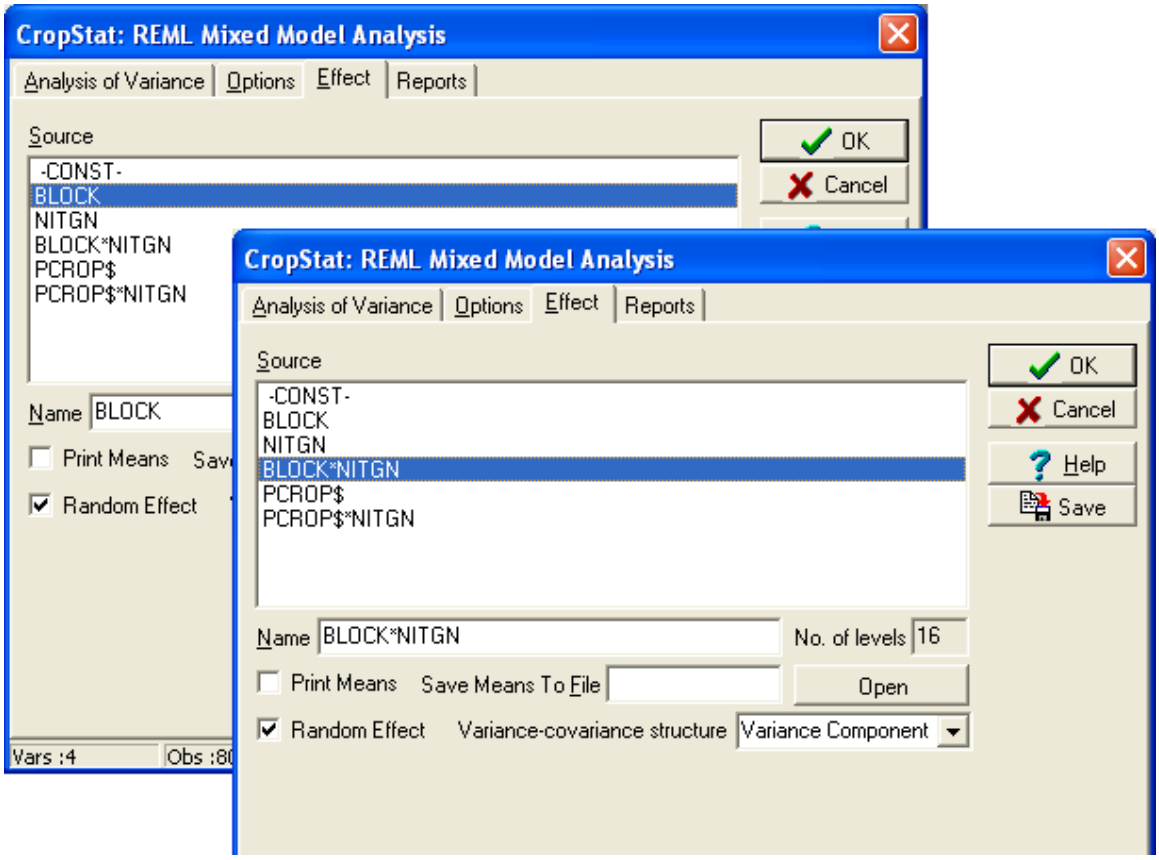


- Click the **Look in** box and go inside *C:\MY CROPSTAT\LINEAR MIXED MODEL*. Specify the command filename BSPLIT and select data file BSPLIT.SYS to use in the analysis.

In the **REML Mixed Model Analysis** window, specify the variates, factors and model in the ANOVA tab.



- In the **Effect** tab, specify *BLOCK* and *BLOCK*NITGN* as Random Effects.



- Click OK to run the analysis.

IV. Sample Output

```
 IRREML 2.0.7: REML ANALYSIS FOR COMMAND SET BSPLIT  FILE BSPLIT    1/ 4/ 9 11: 7
 ------------------------------------------------------------------------ :PAGE   1
 THE IRREML PROGRAM WAS WRITTEN BY DOUGLAS CLARKSON OF SCIENCEOPS FOR IRRI

 Command File: C:\MY CROPSTAT\LINEAR MIXED MODELS\BSPLIT.gfc   Data File: BSPLIT

 Number of Records:    80

 Variables in Data Set: BLOCK  PCROP$  NITGN  YIELD

 SUMMARY STATISTICS FOR NUMERIC VARIATES
 VARIATE       NOBS   MINIMUM     MAXIMUM     MEAN     STD. DEV.
 YIELD          80.   1.410       7.140       4.772       1.349

 Classification Variables: BLOCK  PCROP$  NITGN

 Levels of the classification variables

    4 CODES:(Number  Label) for Variable: BLOCK
(  1            1)(  2            2)(  3            3)(  4            4)(

    5 CODES:(Number  Label) for Variable: PCROP$
(  1 CROTL       )(  2 CWPEA       )(  3 GNUTS       )(  4 MAIZE       )(  5 PPEAS       )

    4 CODES:(Number  Label) for Variable: NITGN
(  1            0)(  2           45)(  3           90)(  4          135)(
IRREML: REML ANALYSIS FOR VARIATE YIELD  FILE BSPLIT    1/ 4/ 9 11: 7
 ------------------------------------------------------------------------ :PAGE   2
 Number of non-missing dependent observations:           80
 Check estimability of effect means: T

 Model Specification
 Intercept in model: Yes
 The Fixed Effects Model
     YIELD = Intercept + NITGN + PCROP$ + PCROP$:NITGN
 The Random Effects Terms
     BLOCK + BLOCK:NITGN

 RANDOM EFFECT COVARIANCE MODEL.  2 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE SCALE SAME NBLOCK GROUPING VARS
 ---------------------------------------------------------------------------
 BLOCK                          1-  1  diagonal   1     1      1
 BLOCK:NITGN                    2-  2  diagonal   1     1      1

 RESIDUAL EFFECT COVARIANCE MODEL.  0 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE SCALE SAME NBLOCK GROUPING VARS
 ---------------------------------------------------------------------------
 RESIDUAL                           sigmasq(1)xI

 Number of columns in the fixed effects model:     20
 Number of columns in the random effects model:    20

 Message: Relative function convergence

 Final REML criterion:           1.196508369884569
 Likelihood value -2LogL:      107.879607220230852

 Variance/Covariance component parameters
 Dep Name               Gamma Coef. Std. Error    Z       Pr > |Z|    Scaled Gamma  Std.
Error
  1 BLOCK(1) ........  0.1936E-02  0.4872E-01  0.3974E-01  0.9683      0.4300E-03  0.1082E-01
  1 BLOCK:NITGN(1) ..  0.8745E-03  0.1032      0.8474E-02  0.9932      0.1942E-03  0.2290E-01
```

```
 The scale parameters
 Dep.  Sigma_Squared Std. Error      Z          Pr > |Z|
Dep(1) .......    0.1447      0.3050E-01   4.743      0.2101E-05

 Asymptotic Covariance Matrix of the Gamma Estimates
                                    1            2            3
  1   1 BLOCK(1).........    0.365E-01  -0.185E-01  -0.643E-08
  2   1 BLOCK:PCROP$(1)..   -0.185E-01   0.120      -0.508E-02
  3 Dep(1)...............   -0.643E-08  -0.508E-02   0.930E-03

        ANOVA Table for Sequentially Deleted Fixed Effects
        Denominator Degrees of Freedom: Containment method
 Dep Effect       DFNum   DFDen   F - Statistic   P > |F|
  1 PCROP$:NITGN     12  60.00        3.387       0.8204E-03
  1        NITGN      3  60.00        216.1       0.4128E-31
  1       PCROP$      4  12.00        16.81       0.7345E-04
□IRREML:  PREDICTIONS  FILE BSPLIT   14/ 3/ 8 13:31
 ------------------------------------------------------------------- :PAGE   3

              Balanced Least Squares Means Fixed
 Dep Level          LSMean   Std. Error
  1  PCROP$ CROTL    5.656      0.1690
  1  PCROP$ CWPEA    5.034      0.1690
  1  PCROP$ GNUTS    4.845      0.1690
  1  PCROP$ MAIZE    3.762      0.1690
  1  PCROP$ PPEAS    4.561      0.1690

 Standard Errors of Differences
  Minimum      Mean      Maximum
  0.2390      0.2390      0.2390

                 Balanced Least Squares Means Fixed
 Dep Level                LSMean   Std. Error
  1  NITGN           0     3.171      0.1055
  1  NITGN          45     4.471      0.1055
  1  NITGN          90     5.385      0.1055
  1  NITGN         135     6.061      0.1055

 Standard Errors of Differences
  Minimum      Mean      Maximum
  0.1203      0.1203      0.1203

                        Balanced Least Squares Means Fixed
 Dep Level                                LSMean   Std. Error
  1  PCROP$:NITGN CROTL            0     4.243      0.2360
  1  PCROP$:NITGN CWPEA            0     3.900      0.2360
  1  PCROP$:NITGN GNUTS            0     2.457      0.2360
  1  PCROP$:NITGN MAIZE            0     2.187      0.2360
  1  PCROP$:NITGN PPEAS            0     3.067      0.2360
  1  PCROP$:NITGN CROTL           45     5.500      0.2360
  1  PCROP$:NITGN CWPEA           45     4.665      0.2360
  1  PCROP$:NITGN GNUTS           45     4.948      0.2360
  1  PCROP$:NITGN MAIZE           45     3.200      0.2360
  1  PCROP$:NITGN PPEAS           45     4.040      0.2360
  1  PCROP$:NITGN CROTL           90     6.065      0.2360
  1  PCROP$:NITGN CWPEA           90     5.393      0.2360
  1  PCROP$:NITGN GNUTS           90     5.770      0.2360
  1  PCROP$:NITGN MAIZE           90     4.477      0.2360
  1  PCROP$:NITGN PPEAS           90     5.217      0.2360
  1  PCROP$:NITGN CROTL          135     6.815      0.2360
  1  PCROP$:NITGN CWPEA          135     6.180      0.2360
  1  PCROP$:NITGN GNUTS          135     6.205      0.2360
  1  PCROP$:NITGN MAIZE          135     5.183      0.2360
  1  PCROP$:NITGN PPEAS          135     5.920      0.2360

 Standard Errors of Differences
  Minimum      Mean      Maximum
  0.2690      0.3235      0.3337
```

COMBINED ANALYSIS OF VARIANCE

At the end of the tutorial, the user should be able to

- perform combined analysis of variance for series of experiments for balanced and unbalanced data
- specify error term for a specific source of variation

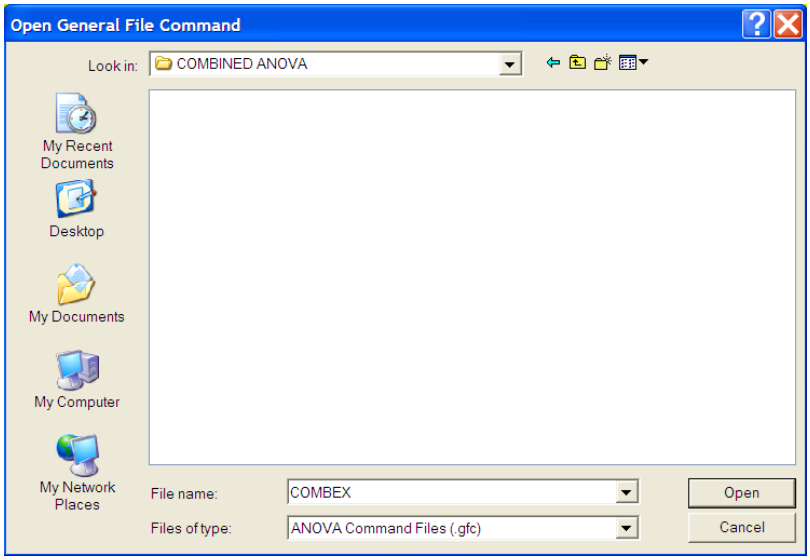
I. Mixed Model Analysis of Variance for a RCB design

- An experiment was conducted in two sites using ten treatment levels with two replicates. Data is stored in *RCBEX.SYS* inside the *CROPSTAT7.2\TUTORIAL\ TUTORIAL DATASETS* folder, open this file by selecting **File ⇒ Open.**

- Select **File ⇒ Save-as**. Click the **Save in** box and go inside your working folder *C:\MY CROPSTAT*. Create a subfolder COMBINED ANOVA then click **Save**.

- Choose **Mixed Model Analysis|Linear Mixed Models** from the **Analysis Menu**.

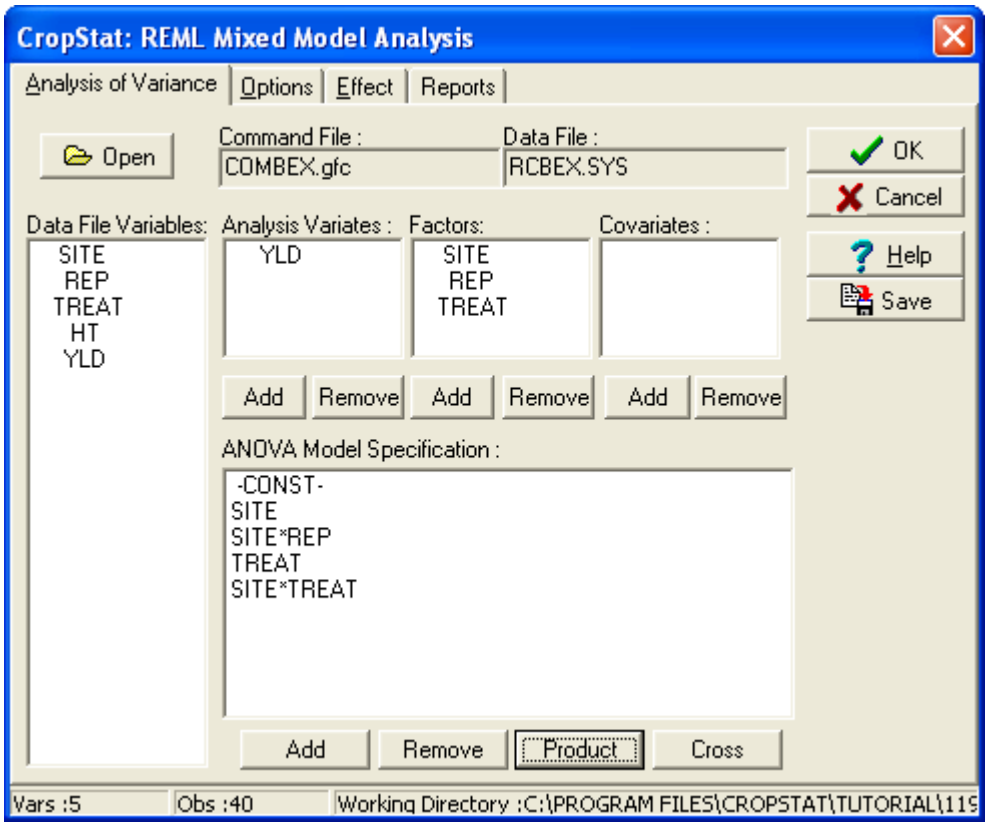
- Open the *MY CROPSTAT\COMBINED ANOVA*. Specify name of the General File Command. Check if you are in the right directory. Enter filename "**COMBEX**".



- Click **Open**. When asked to create a new command file click **Yes.**

- Specify the data file you will use in the analysis.

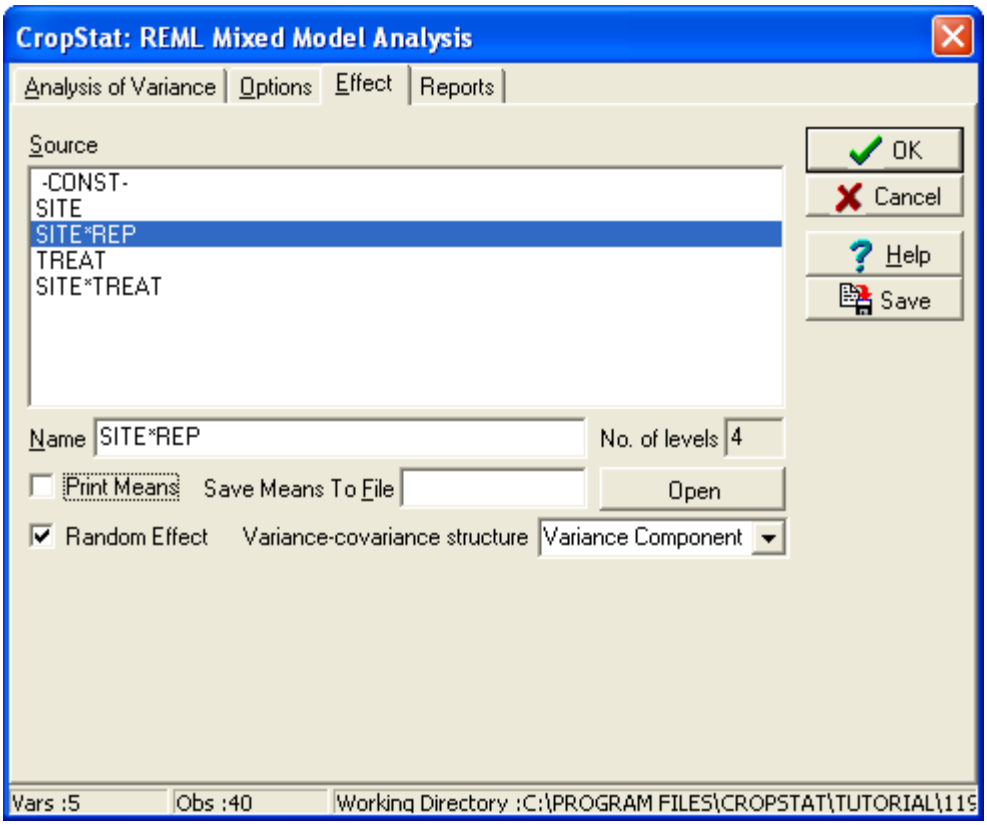
- The **REML Mixed Model Analysis** dialog box will open. Fill-in the details of the analysis.



- From the **Data File Variables** list, highlight all variables to be analyzed then **Add** to the **Analysis Variates** box; and highlight the *TREAT*, *BLOCK/REP* and *SITE* variables then **Add** to the **Factors** box.

- From the **Factors** box, highlight the *SITE* factor then **Add** to the **ANOVA Specification** box. Highlight *SITE* and *REP* factors simultaneously (using CTRL-click for the second) then click **Product**. Highlight the *TREAT* factor, click Add. Highlight the *TREAT* and *SITE* variables then click **Product**.

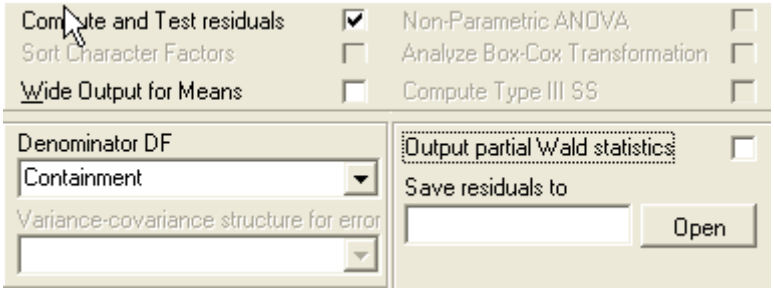
- Click **Effect** to go to the Effect window.



- Decide whether *REP* is random or fix. If random, highlight SITE*REP then click **Random Effect** check box.

- To suppress printing of SITE*REP means click **Print Means** check box.

- To request a plot of residuals against fitted values, tick the **Compute and Test residuals** option in the **Options** tab.

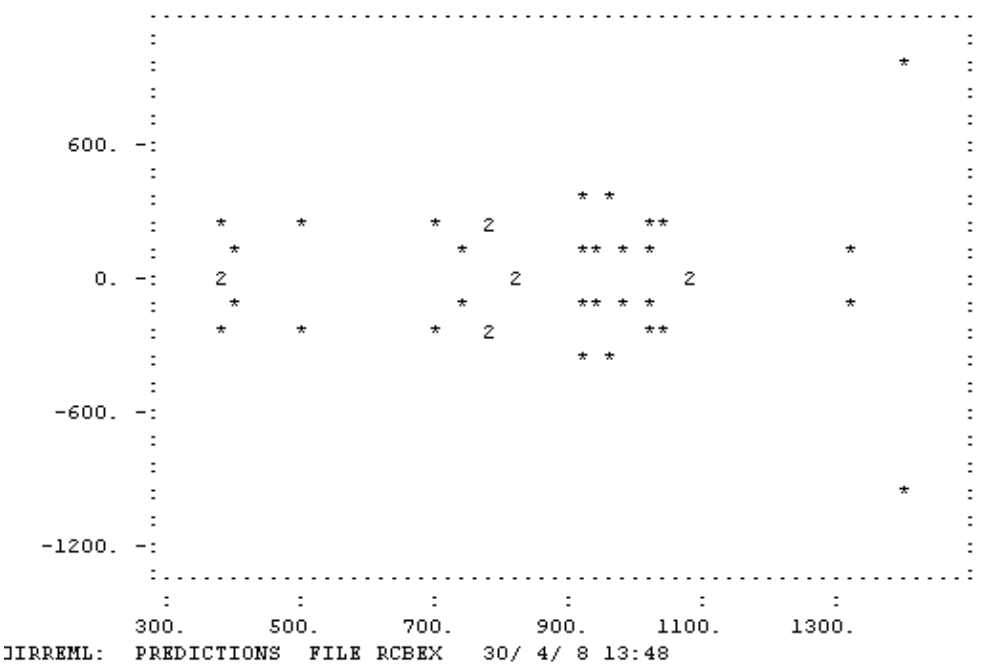


- Click **OK** and the output will appear in the **Text Editor**.

II. Sample Output

```
IRREML(V1.0) - REML ANALYSIS  FILE COMBEX    3/ 2/ 5 16:37
 ------------------------------------------------------------------ :PAGE   1
 THE IRREML PROGRAM WAS WRITTEN BY DOUGLAS CLARKSON OF SCIENCEOPS FOR IRRI

 Command File: C:\MY CROPSTAT\COMBINED ANOVA\COMBEX.gfc   Data File: RCBEX

 Data File: COMBEX

 Number of Records:    40 Non missing observations:    40

 Number of Columns in the Fixed Effects Model:    20

 Number of columns in the random effects model:     4

 Variables in Data Set: SITE  REP  TREAT  YLD

 Classification Variables: SITE  REP  TREAT

 Levels of the classification variables

    2 CODES:(Number  Label) for Variable: SITE
(  1            1)(  2            2)(

    2 CODES:(Number  Label) for Variable: REP
(  1            1)(  2            2)(

   10 CODES:(Number  Label) for Variable: TREAT
(  1            1)(  2            2)(  3            3)(  4            4)(  5
5)
(  6            6)(  7            7)(  8            8)(  9            9)( 10
10)

 Model Specification

 Intercept in model: Yes

 The Fixed Effects Model
     YLD = Intercept + SITE + TREAT + SITE:TREAT

 The Random Effects Terms
     REP(SITE)


 RANDOM EFFECT COVARIANCE MODEL.  1 SPECIFIED STRUCTURES
 TERM             PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 REP(SITE)                    1- 1  diagonal

 RESIDUAL EFFECT COVARIANCE MODEL.  0 SPECIFIED STRUCTURES
 TERM             PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 RESIDUAL                           sigmasq(1)xI
 Message: Relative function convergence

 Final REML criterion:     -137.740353669744138

 Variance/Covariance component parameters
 Dep Name             Coef    Std. Error      Z       Pr > |Z|   Component
  1 REP(SITE)(1) ..  0.1000E-05  0.1079      0.9268E-05   1.000      0.1765

 The scale parameters
 Dep.  Sigma_Squared Std. Error      Z         Pr > |Z|
Dep(1) .......    0.1765E+06  0.5866E+05   3.008      0.2625E-02

 Asymptotic Covariance Matrix of the Variance/Covariance Components
                                 1           2
     1   1 REP(SITE)(1)..   0.116E-01  -0.195E+04
     2 Dep(1).............  -0.195E+04   0.344E+10
```

```
       ANOVA Table for Sequentially Deleted Fixed Effects
       Denominator Degrees of Freedom: Containment method
Dep Effect      DFNum   DFDen   F - Statistic   P > |F|
 1 SITE:TREAT       9  20.00        0.5716       0.8044
 1       TREAT      9  20.00         1.273       0.3100
 1        SITE      1   2.00        0.8084       0.4635

Plot of residuals against predicted values
```




```
JIRREML:  PREDICTIONS  FILE RCBEX   30/ 4/ 8 13:48
------------------------------------------------------------------ :PAGE   3

               Balanced Least Squares Means Fixed
Dep Level              LSMean   Std. Error
 1  SITE            1    913.8       93.93
 1  SITE            2    794.4       93.93

Standard Errors of Differences
 Minimum     Mean      Maximum
  132.8      132.8       132.8

                Balanced Least Squares Means Fixed
Dep Level              LSMean   Std. Error
 1  TREAT           1    886.8       210.0
 1  TREAT           2    432.8       210.0
 1  TREAT           3    926.0       210.0
 1  TREAT           4    970.1       210.0
 1  TREAT           5    1216.       210.0
 1  TREAT           6    661.2       210.0
 1  TREAT           7    825.9       210.0
 1  TREAT           8    576.2       210.0
 1  TREAT           9    972.6       210.0
 1  TREAT          10    1074.       210.0

Standard Errors of Differences
 Minimum     Mean      Maximum
  297.0      297.0       297.0
```

Balanced Least Squares Means Fixed

Dep	Level			LSMean	Std. Error
1	SITE:TREAT	1	1	989.1	297.0
1	SITE:TREAT	2	1	784.6	297.0
1	SITE:TREAT	1	2	495.1	297.0
1	SITE:TREAT	2	2	370.5	297.0
1	SITE:TREAT	1	3	1071.	297.0
1	SITE:TREAT	2	3	780.6	297.0
1	SITE:TREAT	1	4	913.2	297.0
1	SITE:TREAT	2	4	1027.	297.0
1	SITE:TREAT	1	5	1036.	297.0
1	SITE:TREAT	2	5	1395.	297.0
1	SITE:TREAT	1	6	952.1	297.0
1	SITE:TREAT	2	6	370.3	297.0
1	SITE:TREAT	1	7	702.2	297.0
1	SITE:TREAT	2	7	949.6	297.0
1	SITE:TREAT	1	8	744.3	297.0
1	SITE:TREAT	2	8	408.0	297.0
1	SITE:TREAT	1	9	915.2	297.0
1	SITE:TREAT	2	9	1030.	297.0
1	SITE:TREAT	1	10	1319.	297.0
1	SITE:TREAT	2	10	828.0	297.0

Standard Errors of Differences

Minimum	Mean	Maximum
420.1	420.1	420.1

III. Combined Analysis of Variance for an Alpha Design

An experiment recorded the grain yield of 44 varieties conducted in three trials using alpha designs. Data has missing values in one trial. Interaction between varieties and sites is of interest. The data is stored in *UPLAND.XLS* in the *CROPSTAT7.2\TUTORIAL* TUTORIAL DATASETS folder.

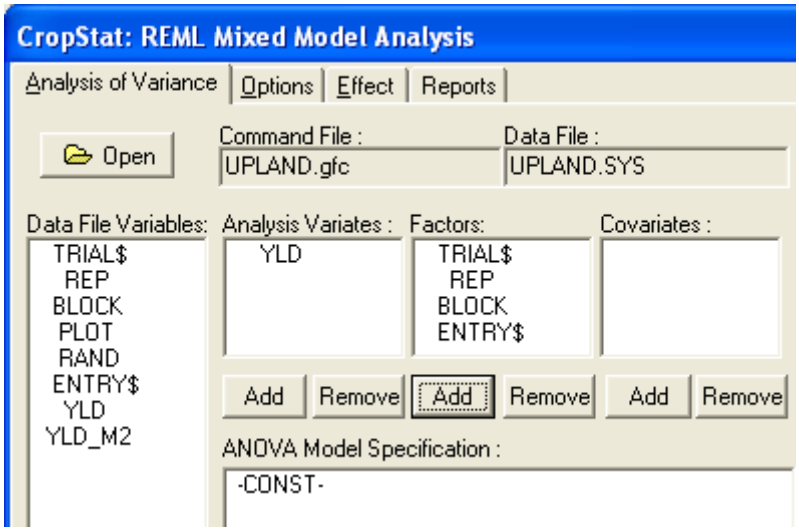
- Import *UPLAND.XLS* using the Data Editor. Save the data as *UPLAND.SYS* inside your working directory *C:\MY CROPSTAT\COMBINED ANOVA* folder. (See instructions on how to import data in an EXCEL file in the Data and File Management tutorial module.)

- From the **Main Window** go to **Analysis Mixed Model Analysis|Linear Mixed Models**.

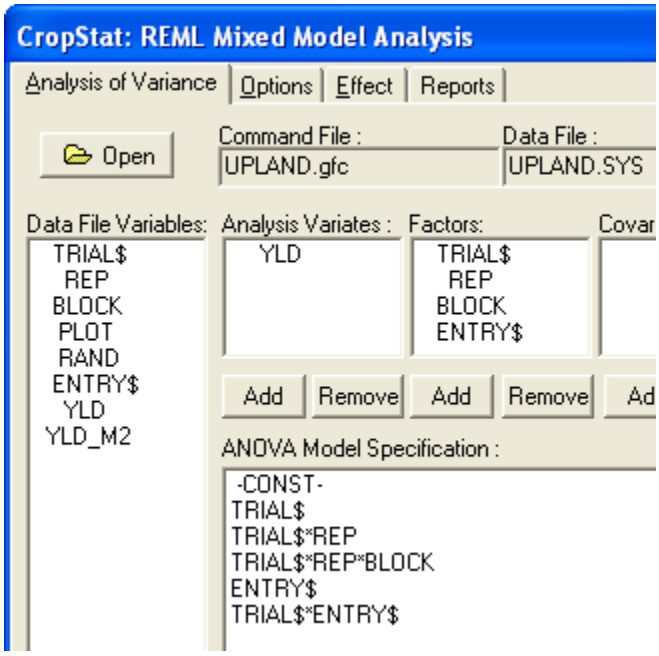
- In the **Open** dialog box**,** click the **Look in** box to switch to your working drive, *C:\MY CROPSTAT\COMBINED ANOVA***.**

- Specify the general file command *UPLAND.GFC.* In the same folder select the data file, *UPLAND.SYS* in the **Open** dialog box.

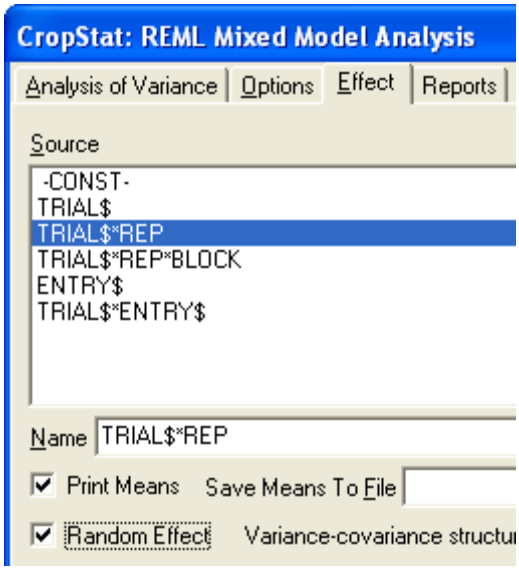
- The **REML Linear Mixed Model Analysis** dialog box will appear. Specify *YLD* as the variate and *TRIAL$, REP, BLOCK* and *ENTRY$* as factors.



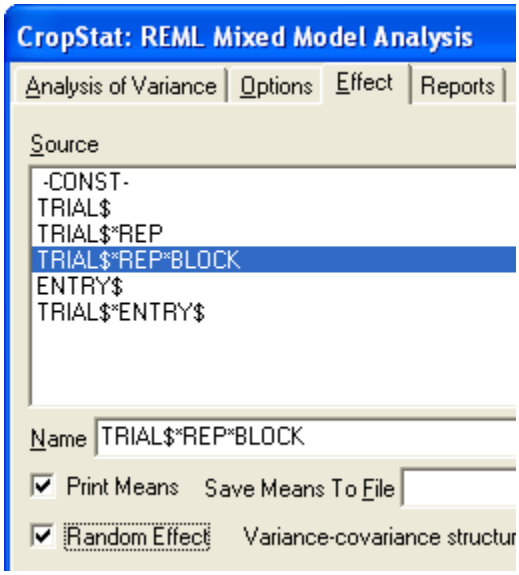
- Specify the ANOVA Model with the following sources of variation: *TRIAL$, TRIAL$*REP, TRIAL$*REP*BLOCK, ENTRY$* and *TRIAL$*ENTRY$*. See the Analysis of Variance module for a detailed step by step instruction on how to specify the ANOVA model.



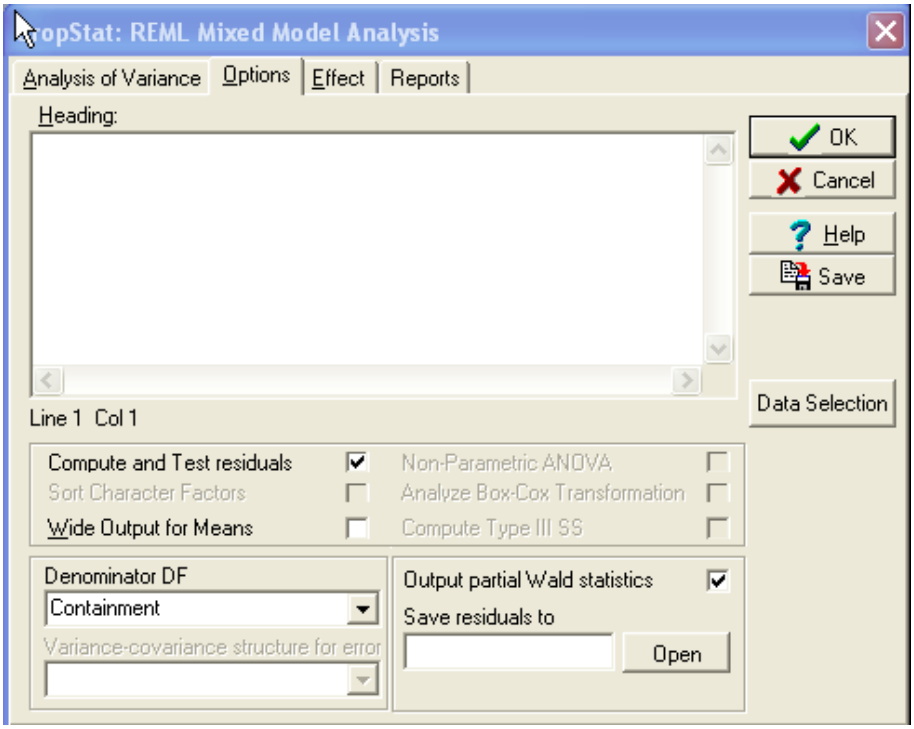
- In the **Effect** tab specify *TRIAL$*REP* as a random factor.



- Specify *TRIAL$*REP*BLOCK* as random factor too.



- To plot residuals against fitted values, tick the **Compute and Test residuals** option in the **Options** tab. Also, to output the Partial Wald Tests, tick and **Output partial Wald Statistics** checkbox.

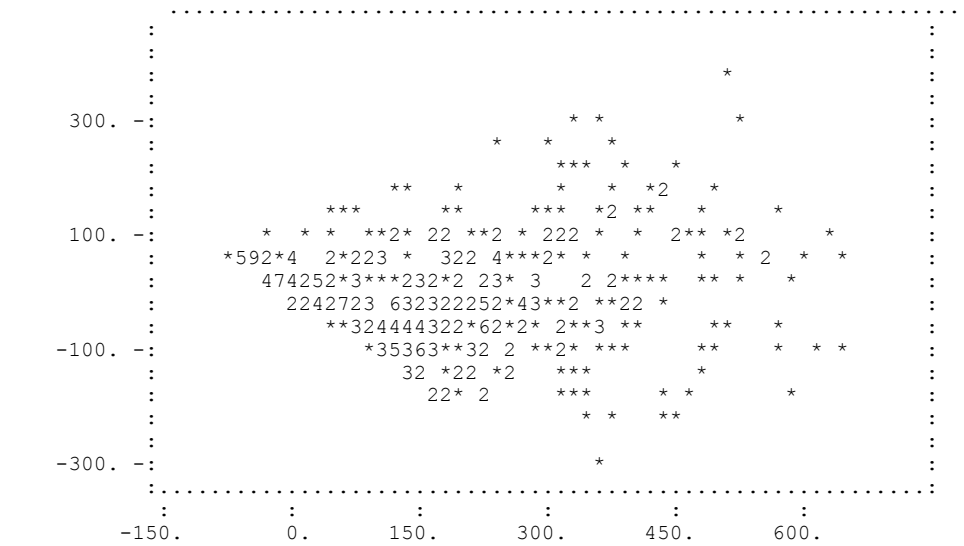


- Click the **OK** button to run the analysis.

IV. Sample Output

```
IRREML(V1.0) - REML ANALYSIS  FILE UPLAND    3/ 2/ 5 11:35
 ------------------------------------------------------------------ :PAGE   1
 THE IRREML PROGRAM WAS WRITTEN BY DOUGLAS CLARKSON OF SCIENCEOPS FOR IRRI

 Command File: C:\MY CROPSTAT\COMBINED ANOVA\UPLAND.gfc

 Data File: UPLAND

 Number of Records:   396 Non missing observations:   391

 Number of Columns in the Fixed Effects Model:   132

 Number of columns in the random effects model:    45

 Variables in Data Set: TRIAL$  REP  BLOCK  ENTRY$  YLD

 Classification Variables: TRIAL$  REP  BLOCK  ENTRY$

 Levels of the classification variables

    3 CODES:(Number  Label) for Variable: TRIAL$
(  1 UDS02-6     )(  2 UDS02-7     )(  3 UDS02-14    )(

    3 CODES:(Number  Label) for Variable: REP
(  1            1)(  2            2)(  3            3)(

    4 CODES:(Number  Label) for Variable: BLOCK
(  1            1)(  2            2)(  3            3)(  4            4)(

   44 CODES:(Number  Label) for Variable: ENTRY$
(  1 C22         )(  2 IR64        )(  3 IR72        )(  4 AUS196      )(  5 AZUCENA
)
(  6 IRAT170     )(  7 IRAT177     )(  8 IRAT212     )(  9 IRAT216     )( 10 PALAWAN
)
( 11 PSBRC80     )( 12 PSBRC82     )( 13 UPLRI-5     )( 14 UPLRI-7     )( 15 VANDANA
)
( 16 DINORADO    )( 17 WAB638-1    )( 18 WAYRAREM    )( 19 MARAVILHA   )( 20
PRIMAVERA   )
( 21 WAB181-18   )( 22 WAB56-125   )( 23 WAB96-1-1   )( 24 IR55419-04  )( 25 IR55423-
01  )
( 26 IR60080-46A )( 27 B6144F-MR-6-)( 28 CT13370-12-2)( 29 CT13377-4-2-)( 30 CT13382-
8-3-)
( 31 CT6510-24-1-)( 32 CT6516-24-3-)( 33 IR47686-30-3)( 34 IR65261-09-1)( 35 IR65907-
116-)
( 36 IR66417-18-1)( 37 IR66421-062-)( 38 IR66424-1-2-)( 39 IR68702-072-)( 40 IR70358-
84-1)
( 41 IR70360-38-1)( 42 IR71524-44-1)( 43 IR71525-19-1)( 44 IR72768-15-1)(

 Model Specification

 Intercept in model: Yes

 The Fixed Effects Model
     YLD = Intercept + TRIAL$ + ENTRY$ + TRIAL$:ENTRY$

 The Random Effects Terms
     REP(TRIAL$) + BLOCK(TRIAL$,REP)


 RANDOM EFFECT COVARIANCE MODEL.  2 SPECIFIED STRUCTURES
 TERM             PARAMETER INDICES  STRUCTURE SCALE SAME NBLOCK GROUPING VARS
 ------------------------------------------------------------------------------
 REP(TRIAL$)                   1-  1  diagonal   1     1      1
 BLOCK(TRIAL$,REP)             2-  2  diagonal   1     1      1

 RESIDUAL EFFECT COVARIANCE MODEL.  0 SPECIFIED STRUCTURES
 TERM             PARAMETER INDICES  STRUCTURE SCALE SAME NBLOCK GROUPING VARS
 ------------------------------------------------------------------------------
 RESIDUAL                          sigmasq(1)xI

 Number of columns in the fixed effects model:    132
 Number of columns in the random effects model:    45
```

```
 Message: Relative function convergence

 Final REML criterion:       -1463.444035106248521
 Likelihood value -2LogL:     3402.898230306497226

 Variance/Covariance component parameters
 Dep Name                     Gamma Coef. Std. Error    Z       Pr > |Z|    Scaled Gamma  Std.
Error
  1 REP(TRIAL$)(1) ........  0.4977      0.3183       1.564      0.1179       7502.       4748.
  1 BLOCK(TRIAL$,REP)(1) ..  0.9547E-01  0.6787E-01   1.407      0.1595       1439.       987.1

 The scale parameters
 Dep.  Sigma_Squared Std. Error      Z          Pr > |Z|
Dep(1) .......    0.1507E+05   1418.       10.63      0.2186E-25

 Asymptotic Covariance Matrix of the Gamma Estimates
                              1           2           3
  1   1 REP(TRIAL$)(1)...   0.101      -0.308E-04   -64.7
  2   1 BLOCK(TRIAL$,REP)  -0.308E-04   0.461E-02   -31.5
  3 Dep(1)...............   -64.7       -31.5       0.201E+07

        Fixed Effect ANOVA Table - Partial Wald Tests
        Denominator Degrees of Freedom: Containment method
 Dep Effect      DFNum   DFDen   F - Statistic   P > |F|
  1        TRIAL$      2   6.00       0.5876       0.5847
  1        ENTRY$     43 259.00        4.275       0.1188E-12
  1 TRIAL$:ENTRY$     86 259.00        1.643       0.1559E-02

        ANOVA Table for Partially Deleted Fixed Effects
        Denominator Degrees of Freedom: Containment method
 Dep Effect      DFNum   DFDen   F - Statistic   P > |F|
  1 TRIAL$:ENTRY$     86 259.00        1.644       0.1540E-02
  1        ENTRY$     43 259.00        4.277       0.1164E-12
  1        TRIAL$      2   6.00       0.6328       0.5632

  ANOVA Table for Sequentially Deleted Fixed Effects
        Denominator Degrees of Freedom: Containment method
 Dep Effect      DFNum   DFDen   F - Statistic   P > |F|
  1 TRIAL$:ENTRY$     86 259.00        1.644       0.1540E-02
  1        ENTRY$     43 259.00        7.576       0.1017E-25
  1        TRIAL$      2   6.00       0.7389       0.5166

Plot of residuals against predicted values
```




```
 Balanced Least Squares Means Fixed
 Dep Level              LSMean   Std. Error
  1  TRIAL$ UDS02-6      160.1       52.30
  1  TRIAL$ UDS02-7      247.4       52.30
  1  TRIAL$ UDS02-14     219.7       52.37

 Standard Errors of Differences
  Minimum     Mean      Maximum
   73.96       73.99       74.01
```

Balanced Least Squares Means Random

Dep	Level			LSMean	Std. Error
1	REP(TRIAL$)	1	UDS02-6	115.2	18.25
1	REP(TRIAL$)	2	UDS02-6	87.84	18.25
1	REP(TRIAL$)	3	UDS02-6	277.4	18.25
1	REP(TRIAL$)	1	UDS02-7	226.9	18.25
1	REP(TRIAL$)	2	UDS02-7	189.9	18.25
1	REP(TRIAL$)	3	UDS02-7	325.5	18.25
1	REP(TRIAL$)	1	UDS02-14	131.1	20.02
1	REP(TRIAL$)	2	UDS02-14	228.2	18.25
1	REP(TRIAL$)	3	UDS02-14	299.9	18.25

Standard Errors of Differences

Minimum	Mean	Maximum
25.62	26.05	27.09

Balanced Least Squares Means Random

Dep	Level				LSMean	Std. Error
1	BLOCK(TRIAL$,REP)	1	UDS02-6	1	100.3	31.12
1	BLOCK(TRIAL$,REP)	2	UDS02-6	1	126.3	31.12
1	BLOCK(TRIAL$,REP)	3	UDS02-6	1	107.0	31.12
1	BLOCK(TRIAL$,REP)	4	UDS02-6	1	127.1	31.12
1	BLOCK(TRIAL$,REP)	1	UDS02-7	1	273.0	31.12
1	BLOCK(TRIAL$,REP)	2	UDS02-7	1	182.2	31.12
1	BLOCK(TRIAL$,REP)	3	UDS02-7	1	211.2	31.12
1	BLOCK(TRIAL$,REP)	4	UDS02-7	1	241.3	31.12
1	BLOCK(TRIAL$,REP)	1	UDS02-14	1	141.6	31.48
1	BLOCK(TRIAL$,REP)	2	UDS02-14	1	117.5	31.48
1	BLOCK(TRIAL$,REP)	3	UDS02-14	1	151.5	33.38
1	BLOCK(TRIAL$,REP)	4	UDS02-14	1	114.0	34.41
1	BLOCK(TRIAL$,REP)	1	UDS02-6	2	89.84	31.12
1	BLOCK(TRIAL$,REP)	2	UDS02-6	2	93.51	31.12
1	BLOCK(TRIAL$,REP)	3	UDS02-6	2	93.39	31.12
1	BLOCK(TRIAL$,REP)	4	UDS02-6	2	74.61	31.12
1	BLOCK(TRIAL$,REP)	1	UDS02-7	2	193.2	31.12
1	BLOCK(TRIAL$,REP)	2	UDS02-7	2	215.7	31.12
1	BLOCK(TRIAL$,REP)	3	UDS02-7	2	185.0	31.12
1	BLOCK(TRIAL$,REP)	4	UDS02-7	2	165.7	31.12
1	BLOCK(TRIAL$,REP)	1	UDS02-14	2	217.4	31.17
1	BLOCK(TRIAL$,REP)	2	UDS02-14	2	196.6	31.22
1	BLOCK(TRIAL$,REP)	3	UDS02-14	2	220.0	31.44
1	BLOCK(TRIAL$,REP)	4	UDS02-14	2	278.8	31.18
1	BLOCK(TRIAL$,REP)	1	UDS02-6	3	273.0	31.12
1	BLOCK(TRIAL$,REP)	2	UDS02-6	3	275.8	31.12
1	BLOCK(TRIAL$,REP)	3	UDS02-6	3	292.3	31.12
1	BLOCK(TRIAL$,REP)	4	UDS02-6	3	268.5	31.12
1	BLOCK(TRIAL$,REP)	1	UDS02-7	3	293.5	31.12
1	BLOCK(TRIAL$,REP)	2	UDS02-7	3	348.3	31.12
1	BLOCK(TRIAL$,REP)	3	UDS02-7	3	334.3	31.12
1	BLOCK(TRIAL$,REP)	4	UDS02-7	3	326.0	31.12
1	BLOCK(TRIAL$,REP)	1	UDS02-14	3	303.3	31.29
1	BLOCK(TRIAL$,REP)	2	UDS02-14	3	319.7	31.18
1	BLOCK(TRIAL$,REP)	3	UDS02-14	3	317.9	31.29
1	BLOCK(TRIAL$,REP)	4	UDS02-14	3	258.6	31.22

Standard Errors of Differences

Minimum	Mean	Maximum
41.16	44.02	47.29

Balanced Least Squares Means Fixed

Dep	Level	LSMean	Std. Error
1	ENTRY$ C22	72.87	54.04
1	ENTRY$ IR64	206.2	54.03
1	ENTRY$ IR72	223.5	51.19
1	ENTRY$ AUS196	315.1	51.18
1	ENTRY$ AZUCENA	103.6	51.19
1	ENTRY$ IRAT170	107.5	51.19
1	ENTRY$ IRAT177	136.0	51.18
1	ENTRY$ IRAT212	326.8	51.18
1	ENTRY$ IRAT216	133.8	51.19
1	ENTRY$ PALAWAN	196.2	51.18
1	ENTRY$ PSBRC80	198.3	51.18
1	ENTRY$ PSBRC82	217.4	51.18
1	ENTRY$ UPLRI-5	148.5	51.18
1	ENTRY$ UPLRI-7	258.4	51.18
1	ENTRY$ VANDANA	500.8	51.18
1	ENTRY$ DINORADO	49.92	54.02
1	ENTRY$ WAB638-1	18.31	51.18
1	ENTRY$ WAYRAREM	98.36	51.18
1	ENTRY$ MARAVILHA	158.0	51.18
1	ENTRY$ PRIMAVERA	409.9	51.19
1	ENTRY$ WAB181-18	243.3	51.18
1	ENTRY$ WAB56-125	415.1	51.19
1	ENTRY$ WAB96-1-1	61.01	51.18
1	ENTRY$ IR55419-04	265.4	51.19
1	ENTRY$ IR55423-01	303.9	51.18
1	ENTRY$ IR60080-46A	224.6	51.19
1	ENTRY$ B6144F-MR-6-	253.3	51.18
1	ENTRY$ CT13370-12-2	116.5	51.18
1	ENTRY$ CT13377-4-2-	136.2	51.18
1	ENTRY$ CT13382-8-3-	95.60	51.19
1	ENTRY$ CT6510-24-1-	286.1	51.18
1	ENTRY$ CT6516-24-3-	158.4	51.18
1	ENTRY$ IR47686-30-3	53.50	51.18
1	ENTRY$ IR65261-09-1	117.2	51.17
1	ENTRY$ IR65907-116-	280.0	54.04
1	ENTRY$ IR66417-18-1	148.6	54.03
1	ENTRY$ IR66421-062-	128.6	51.19
1	ENTRY$ IR66424-1-2-	244.6	51.19
1	ENTRY$ IR68702-072-	208.7	51.18
1	ENTRY$ IR70358-84-1	439.8	51.18
1	ENTRY$ IR70360-38-1	162.3	51.19
1	ENTRY$ IR71524-44-1	272.7	51.17
1	ENTRY$ IR71525-19-1	283.2	51.20
1	ENTRY$ IR72768-15-1	422.4	51.19

Standard Errors of Differences

Minimum	Mean	Maximum
58.42	59.69	64.29

Balanced Least Squares Means Fixed

Dep	Level	LSMean	Std. Error
1	TRIAL$:ENTRY$ UDS02-6 C22	9.104	88.66
1	TRIAL$:ENTRY$ UDS02-7 C22	68.72	88.64
1	TRIAL$:ENTRY$ UDS02-14 C22	140.8	102.8
1	TRIAL$:ENTRY$ UDS02-6 IR64	120.4	88.66
1	TRIAL$:ENTRY$ UDS02-7 IR64	382.9	88.64
1	TRIAL$:ENTRY$ UDS02-14 IR64	115.5	102.7
1	TRIAL$:ENTRY$ UDS02-6 IR72	185.5	88.64
1	TRIAL$:ENTRY$ UDS02-7 IR72	175.4	88.66
1	TRIAL$:ENTRY$ UDS02-14 IR72	309.6	88.69
1	TRIAL$:ENTRY$ UDS02-6 AUS196	162.5	88.66
1	TRIAL$:ENTRY$ UDS02-7 AUS196	269.8	88.66
1	TRIAL$:ENTRY$ UDS02-14 AUS196	513.0	88.64
1	TRIAL$:ENTRY$ UDS02-6 AZUCENA	13.41	88.64
1	TRIAL$:ENTRY$ UDS02-7 AZUCENA	167.3	88.66
1	TRIAL$:ENTRY$ UDS02-14 AZUCENA	130.2	88.70
1	TRIAL$:ENTRY$ UDS02-6 IRAT170	26.36	88.66
1	TRIAL$:ENTRY$ UDS02-7 IRAT170	193.4	88.66
1	TRIAL$:ENTRY$ UDS02-14 IRAT170	102.7	88.68
1	TRIAL$:ENTRY$ UDS02-6 IRAT177	46.43	88.66
1	TRIAL$:ENTRY$ UDS02-7 IRAT177	206.5	88.64

1	TRIAL$:ENTRY$ UDS02-14 IRAT177	155.0	88.65
1	TRIAL$:ENTRY$ UDS02-6 IRAT212	212.2	88.62
1	TRIAL$:ENTRY$ UDS02-7 IRAT212	347.3	88.66
1	TRIAL$:ENTRY$ UDS02-14 IRAT212	421.1	88.68
1	TRIAL$:ENTRY$ UDS02-6 IRAT216	123.4	88.66
1	TRIAL$:ENTRY$ UDS02-7 IRAT216	159.5	88.66
1	TRIAL$:ENTRY$ UDS02-14 IRAT216	118.6	88.69
1	TRIAL$:ENTRY$ UDS02-6 PALAWAN	58.55	88.66
1	TRIAL$:ENTRY$ UDS02-7 PALAWAN	401.2	88.62
1	TRIAL$:ENTRY$ UDS02-14 PALAWAN	128.9	88.69
1	TRIAL$:ENTRY$ UDS02-6 PSBRC80	204.1	88.62
1	TRIAL$:ENTRY$ UDS02-7 PSBRC80	153.6	88.66
1	TRIAL$:ENTRY$ UDS02-14 PSBRC80	237.2	88.66
1	TRIAL$:ENTRY$ UDS02-6 PSBRC82	239.2	88.62
1	TRIAL$:ENTRY$ UDS02-7 PSBRC82	316.7	88.66
1	TRIAL$:ENTRY$ UDS02-14 PSBRC82	96.19	88.69
1	TRIAL$:ENTRY$ UDS02-6 UPLRI-5	12.02	88.66
1	TRIAL$:ENTRY$ UDS02-7 UPLRI-5	181.8	88.66
1	TRIAL$:ENTRY$ UDS02-14 UPLRI-5	251.8	88.63
1	TRIAL$:ENTRY$ UDS02-6 UPLRI-7	171.3	88.66
1	TRIAL$:ENTRY$ UDS02-7 UPLRI-7	412.6	88.62
1	TRIAL$:ENTRY$ UDS02-14 UPLRI-7	191.3	88.68
1	TRIAL$:ENTRY$ UDS02-6 VANDANA	343.8	88.64
1	TRIAL$:ENTRY$ UDS02-7 VANDANA	563.2	88.64
1	TRIAL$:ENTRY$ UDS02-14 VANDANA	595.5	88.67
1	TRIAL$:ENTRY$ UDS02-6 DINORADO	26.73	88.66
1	TRIAL$:ENTRY$ UDS02-7 DINORADO	30.54	88.66
1	TRIAL$:ENTRY$ UDS02-14 DINORADO	92.49	102.7
1	TRIAL$:ENTRY$ UDS02-6 WAB638-1	3.911	88.66
1	TRIAL$:ENTRY$ UDS02-7 WAB638-1	15.74	88.66
1	TRIAL$:ENTRY$ UDS02-14 WAB638-1	35.27	88.65
1	TRIAL$:ENTRY$ UDS02-6 WAYRAREM	64.06	88.64
1	TRIAL$:ENTRY$ UDS02-7 WAYRAREM	160.7	88.64
1	TRIAL$:ENTRY$ UDS02-14 WAYRAREM	70.35	88.66
1	TRIAL$:ENTRY$ UDS02-6 MARAVILHA	41.92	88.64
1	TRIAL$:ENTRY$ UDS02-7 MARAVILHA	296.2	88.62
1	TRIAL$:ENTRY$ UDS02-14 MARAVILHA	136.0	88.71
1	TRIAL$:ENTRY$ UDS02-6 PRIMAVERA	471.9	88.66
1	TRIAL$:ENTRY$ UDS02-7 PRIMAVERA	507.0	88.66
1	TRIAL$:ENTRY$ UDS02-14 PRIMAVERA	250.8	88.67
1	TRIAL$:ENTRY$ UDS02-6 WAB181-18	109.2	88.64
1	TRIAL$:ENTRY$ UDS02-7 WAB181-18	239.3	88.66
1	TRIAL$:ENTRY$ UDS02-14 WAB181-18	381.5	88.66
1	TRIAL$:ENTRY$ UDS02-6 WAB56-125	275.5	88.66
1	TRIAL$:ENTRY$ UDS02-7 WAB56-125	423.4	88.62
1	TRIAL$:ENTRY$ UDS02-14 WAB56-125	546.4	88.70
1	TRIAL$:ENTRY$ UDS02-6 WAB96-1-1	62.10	88.66
1	TRIAL$:ENTRY$ UDS02-7 WAB96-1-1	57.00	88.64
1	TRIAL$:ENTRY$ UDS02-14 WAB96-1-1	63.94	88.65
1	TRIAL$:ENTRY$ UDS02-6 IR55419-04	414.7	88.64
1	TRIAL$:ENTRY$ UDS02-7 IR55419-04	276.9	88.64
1	TRIAL$:ENTRY$ UDS02-14 IR55419-04	104.4	88.70
1	TRIAL$:ENTRY$ UDS02-6 IR55423-01	247.4	88.64
1	TRIAL$:ENTRY$ UDS02-7 IR55423-01	340.3	88.64
1	TRIAL$:ENTRY$ UDS02-14 IR55423-01	324.0	88.68
1	TRIAL$:ENTRY$ UDS02-6 IR60080-46A	113.3	88.66
1	TRIAL$:ENTRY$ UDS02-7 IR60080-46A	288.5	88.66
1	TRIAL$:ENTRY$ UDS02-14 IR60080-46A	272.1	88.67
1	TRIAL$:ENTRY$ UDS02-6 B6144F-MR-6-	177.3	88.64
1	TRIAL$:ENTRY$ UDS02-7 B6144F-MR-6-	390.4	88.66
1	TRIAL$:ENTRY$ UDS02-14 B6144F-MR-6-	192.2	88.67
1	TRIAL$:ENTRY$ UDS02-6 CT13370-12-2	107.1	88.64
1	TRIAL$:ENTRY$ UDS02-7 CT13370-12-2	115.1	88.64
1	TRIAL$:ENTRY$ UDS02-14 CT13370-12-2	127.4	88.67
1	TRIAL$:ENTRY$ UDS02-6 CT13377-4-2-	227.4	88.64
1	TRIAL$:ENTRY$ UDS02-7 CT13377-4-2-	164.5	88.66
1	TRIAL$:ENTRY$ UDS02-14 CT13377-4-2-	16.79	88.66
1	TRIAL$:ENTRY$ UDS02-6 CT13382-8-3-	120.3	88.66
1	TRIAL$:ENTRY$ UDS02-7 CT13382-8-3-	52.35	88.66
1	TRIAL$:ENTRY$ UDS02-14 CT13382-8-3-	114.1	88.69
1	TRIAL$:ENTRY$ UDS02-6 CT6510-24-1-	149.2	88.64
1	TRIAL$:ENTRY$ UDS02-7 CT6510-24-1-	454.8	88.64
1	TRIAL$:ENTRY$ UDS02-14 CT6510-24-1-	254.4	88.68
1	TRIAL$:ENTRY$ UDS02-6 CT6516-24-3-	170.3	88.66
1	TRIAL$:ENTRY$ UDS02-7 CT6516-24-3-	203.1	88.66

1	TRIAL$:ENTRY$ UDS02-14 CT6516-24-3-	101.8	88.64
1	TRIAL$:ENTRY$ UDS02-6 IR47686-30-3	19.23	88.64
1	TRIAL$:ENTRY$ UDS02-7 IR47686-30-3	15.07	88.64
1	TRIAL$:ENTRY$ UDS02-14 IR47686-30-3	126.2	88.67
1	TRIAL$:ENTRY$ UDS02-6 IR65261-09-1	35.67	88.64
1	TRIAL$:ENTRY$ UDS02-7 IR65261-09-1	135.9	88.64
1	TRIAL$:ENTRY$ UDS02-14 IR65261-09-1	180.1	88.64
1	TRIAL$:ENTRY$ UDS02-6 IR65907-116-	186.2	88.66
1	TRIAL$:ENTRY$ UDS02-7 IR65907-116-	425.6	88.66
1	TRIAL$:ENTRY$ UDS02-14 IR65907-116-	228.1	102.8
1	TRIAL$:ENTRY$ UDS02-6 IR66417-18-1	35.89	88.66
1	TRIAL$:ENTRY$ UDS02-7 IR66417-18-1	169.2	88.66
1	TRIAL$:ENTRY$ UDS02-14 IR66417-18-1	240.8	102.7
1	TRIAL$:ENTRY$ UDS02-6 IR66421-062-	123.7	88.66
1	TRIAL$:ENTRY$ UDS02-7 IR66421-062-	177.0	88.64
1	TRIAL$:ENTRY$ UDS02-14 IR66421-062-	85.21	88.69
1	TRIAL$:ENTRY$ UDS02-6 IR66424-1-2-	261.5	88.66
1	TRIAL$:ENTRY$ UDS02-7 IR66424-1-2-	86.02	88.66
1	TRIAL$:ENTRY$ UDS02-14 IR66424-1-2-	386.4	88.69
1	TRIAL$:ENTRY$ UDS02-6 IR68702-072-	329.2	88.64
1	TRIAL$:ENTRY$ UDS02-7 IR68702-072-	218.2	88.66
1	TRIAL$:ENTRY$ UDS02-14 IR68702-072-	78.56	88.67
1	TRIAL$:ENTRY$ UDS02-6 IR70358-84-1	401.4	88.62
1	TRIAL$:ENTRY$ UDS02-7 IR70358-84-1	387.2	88.66
1	TRIAL$:ENTRY$ UDS02-14 IR70358-84-1	530.9	88.67
1	TRIAL$:ENTRY$ UDS02-6 IR70360-38-1	114.7	88.66
1	TRIAL$:ENTRY$ UDS02-7 IR70360-38-1	223.7	88.66
1	TRIAL$:ENTRY$ UDS02-14 IR70360-38-1	148.4	88.70
1	TRIAL$:ENTRY$ UDS02-6 IR71524-44-1	244.9	88.64
1	TRIAL$:ENTRY$ UDS02-7 IR71524-44-1	333.9	88.64
1	TRIAL$:ENTRY$ UDS02-14 IR71524-44-1	239.3	88.64
1	TRIAL$:ENTRY$ UDS02-6 IR71525-19-1	224.1	88.66
1	TRIAL$:ENTRY$ UDS02-7 IR71525-19-1	286.7	88.64
1	TRIAL$:ENTRY$ UDS02-14 IR71525-19-1	338.9	88.73
1	TRIAL$:ENTRY$ UDS02-6 IR72768-15-1	359.3	88.66
1	TRIAL$:ENTRY$ UDS02-7 IR72768-15-1	413.6	88.64
1	TRIAL$:ENTRY$ UDS02-14 IR72768-15-1	494.4	88.71

Standard Errors of Differences

Minimum	Mean	Maximum
101.2	118.7	135.7

Exercise 4

EXPERIMENT ON NITROGEN SOURCE CONDUCTED IN 3 SITES

PROBLEM	The source of nitrogen being used by farmers in a given area is very expensive.
OBJECTIVES	To investigate if *sesbania* could be used as an alternative source of nitrogen.
SITES	3 farmer's fields chosen at random in Nueva Ecija, Philippines
DESIGN	Randomized complete block in each site
	Treatment - 3 sources of nitrogen (no nitrogen, *Sesbania* (Basal) + 60 kg N, 120 kg N
	Replications – 3
FILENAME	COMBEXER

Perform a combined analysis of variance and interpret results.

ANALYSIS OF REPEATED MEASURES

At the end of the tutorial, the user should be able to

- generate orthogonal polynomial coefficients
- create new variables using these coefficients
- perform analysis of repeated measures using these new variables
- convert data from parallel to serial
- perform analysis of repeated measures using mixed models

I. Introduction

Analysis of repeated measures involves studies wherein repeated measurements are taken on each experimental unit over a period of time or space. The experimental unit may be a plot, animal or plant. For example:

- height measured at different growth stages of the plant
- weekly gain in weight of an animal
- daily soil moisture content of a plot
- weekly yield of a fruit crop
- number of faults along a bolt of cloth
- root density at different soil depth
- yearly yield of a long term fertility experiment
- soil samples taken at different depths
- grain yield data taken from a long term experiment

Analysis could be carried out for each of the ***r*** times, but with these analyses the change of treatment effect over time cannot be evaluated. Hence, what is needed is an analysis which provides a test for the interaction between treatment and time. In assuming a split-plot design where time or interval between measurements represents the different subplots, the interaction between time and treatment can be measured. One problem with this analysis is that subplot treatments are supposed to be applied at random to the different subplots. This is obviously not possible with time or interval of measurements. Consequently, neighboring plots or measurements will tend to be more highly correlated than distant plots or measurements. The result of this is that estimates of measurement error tend to be too small and too many significant differences will be declared.

One way to avoid the problem of correlated measurements is to choose one or more measures to summarize the trend over repetitions for each sampling unit, and then base the analysis on those measures. This reduces the problem to univariate analysis

of different measures – one for each experimental unit, standard methods, like analysis of variance can be used. The key is to choose good summary measures. Some common choices are:

- Average over time
- Slope of regression over time (or higher order polynomial coefficients)
- Total increase (last point minus first point)
- Area under a time curve
- Maximum or minimum point

The use of orthogonal polynomials for multiple summary measures has the advantage that the separate measures are independent for normal samples and hence the multiple tests are valid. This approach will be illustrated in section III below.

With the right choice of summary measures this type of analysis can be very useful. However, the disadvantage of this method is that information may be lost by reducing each trend to single measures.

Another approach is to explicitly model the covariance structure of the repeated measures using a mixed model with special parametric structure on the covariance matrices. This model must pay special attention to the covariance structure due to the sequential nature of the time or interval measurements. This method will be illustrated in section IV below.

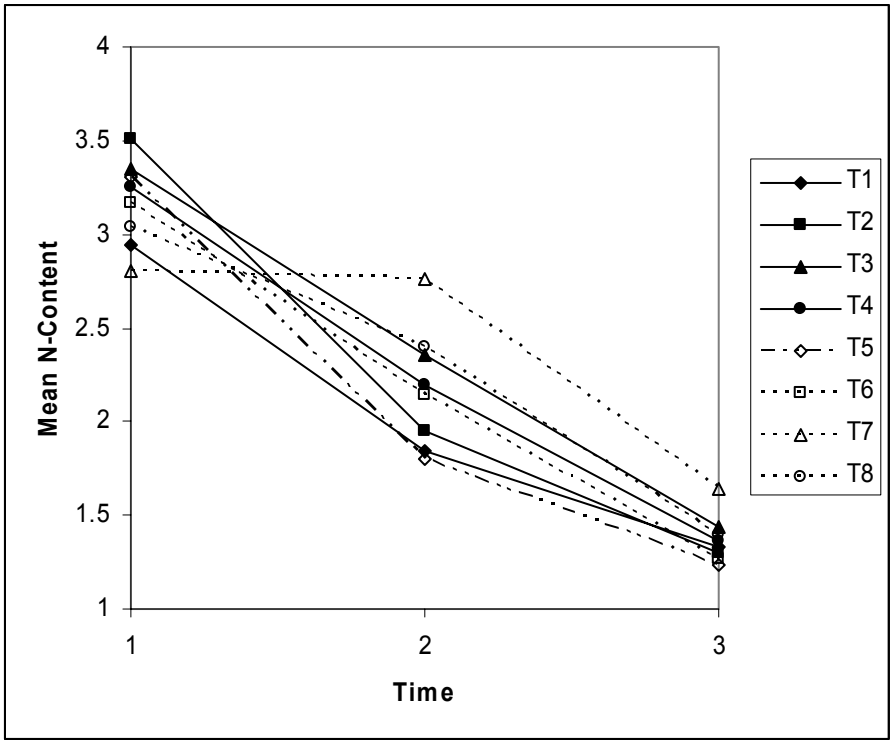
II. Sample problem: Nutrient Content of rice plant at different growth stages

To illustrate how to perform repeated measures analysis in CropStat using both the analysis of summary measures and mixed model analysis, we use the following nutrient content data collected at three different growth stages of rice plants.

Treatment	Time 1	Time 2	Time 3
		Rep 1	
1	3.26	1.88	1.40
2	3.84	2.36	1.53
3	3.50	2.20	1.33
4	3.43	2.32	1.61
5	3.43	1.98	1.11
6	3.68	2.01	1.26
7	2.97	2.66	1.87
8	3.11	2.53	1.76
		Rep 2	
1	2.98	1.74	1.24
2	3.74	2.14	1.21
3	3.49	2.28	1.54
4	3.45	2.33	1.33
5	3.24	1.70	1.25
6	3.24	2.33	1.44
7	2.90	2.74	1.81
8	3.04	2.22	1.28
		Rep 3	
1	2.78	1.76	1.44
2	3.09	1.75	1.28
3	3.03	2.48	1.46
4	2.81	2.16	1.40
5	3.45	1.78	1.39
6	2.84	2.22	1.12
7	2.92	2.67	1.31
8	3.20	2.61	1.23
		Rep 4	
1	2.77	2.00	1.25
2	3.36	1.57	1.17
3	3.36	2.47	1.41
4	3.32	1.99	1.12
5	3.09	1.74	1.20
6	2.91	2.00	1.24
7	2.42	2.98	1.56
8	2.81	2.22	1.29

III. Analysis of Repeated Measures using Summary Statistics

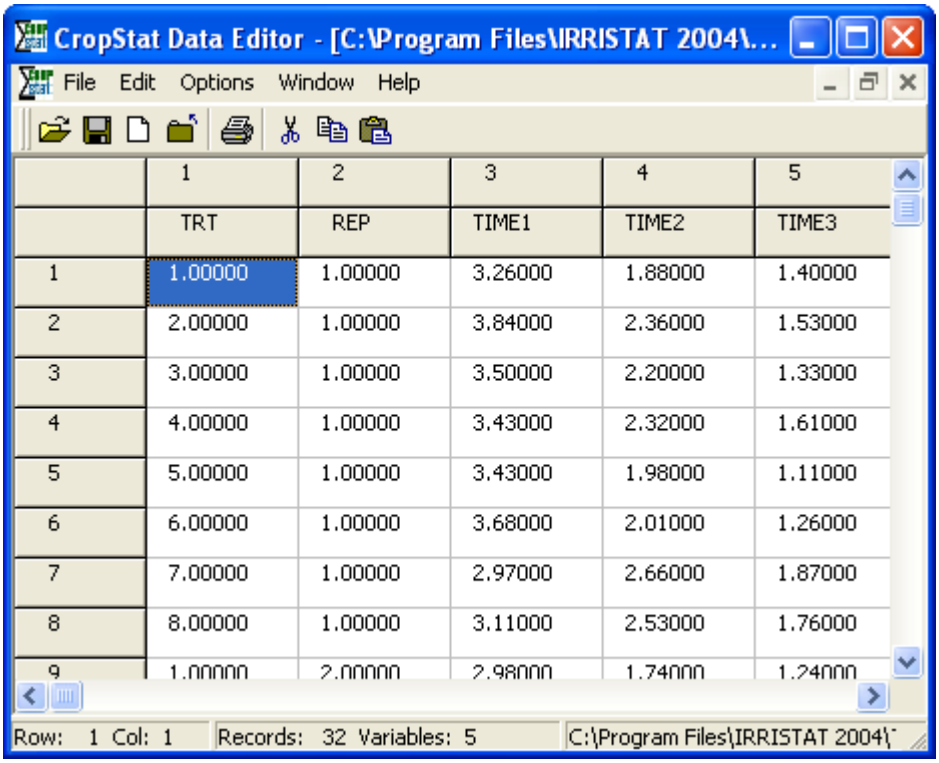
Plotting the data should be the first step in the analysis of repeated measures. Quite often the main conclusions from the analysis can already be seen from a good plot of the data. The figure below shows the mean nitrogen content for each treatment across time.



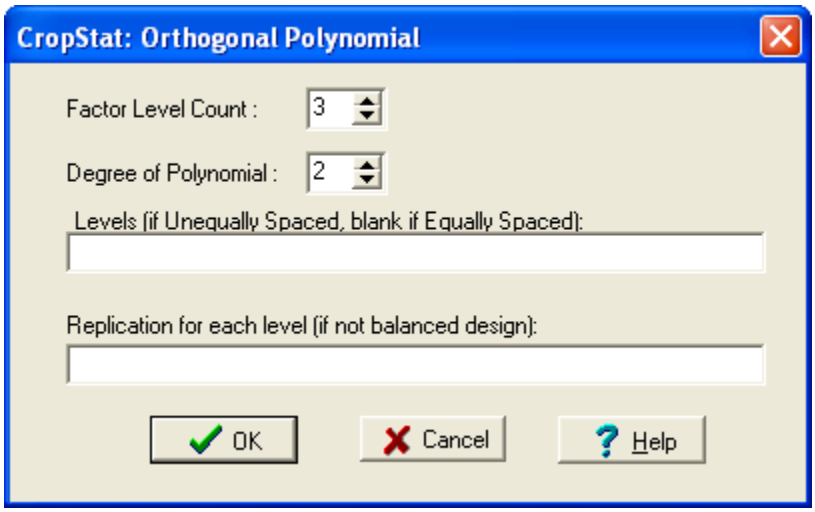
From the figure it seems that nitrogen content is decreasing from time 1 to time 3 and perhaps there are small differences among the treatments. It also seems that the slope is similar for all the treatments except treatment 7. To calculate this summary measure, follow these steps:

- Create a subfolder REPEATED MEASURES inside working folder C:\MY CROPSTAT.

- Using the Data editor import *Repmeas1.XLS* which is stored inside *CROPSTAT7.2\TUTORIAL* TUTORIAL DATASETS folder. Save the data as *Repmeas1.SYS* inside the MY CROPSTAT\REPEATED MEASURES folder.



- Generate orthogonal polynomial contrasts for the linear and quadratic trends for the 3 levels of time. Select **Utilities|Orthogonal Polynomial** from the **Main Window**. The **Orthogonal Polynomial** dialog box will appear.



- Enter the number of levels of the independent variable in the **Factor Level Count** spin box. For this example, the number of levels of TIME is 3. Set the **Factor Level Count** spin box to 3.

- Enter the degree of polynomials required. The maximum number of polynomials that can be requested is one less the number of levels but not exceeding 5. For this example, since we wish to analyze linear and quadratic trends then the number of polynomials to be requested is 2. Enter 2 in the **Degree of Polynomial** spin box.

- If the levels are not equally spaced, enter the level values separated by spaces or commas on the **Levels** edit box. Otherwise, there is no need to enter the level values. For this example the levels of TIME are equally spaced so there is no need to enter the values.

- Click **OK**.

- From the **Main Window**, CropStat will give the following output. Note: the orthogonal polynomial coefficients in standardized units should be used.

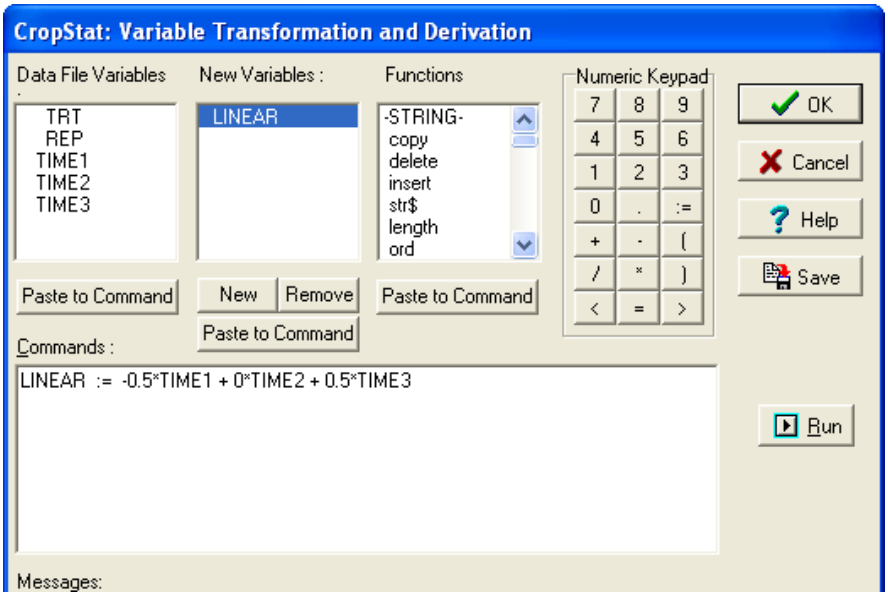
```
Standardized Units
  SUM [P(I)*Y(I)^K] = 1 for each Polynomial K=1..NP

P(1) :
    -0.500000     0.000000     0.500000
P(2) :
  0.500000    -1.000000     0.500000
```

- Add a new variable for the orthogonal contrast to *REPMEAS1.SYS*. Select **Window|Data Editor** from the **Main Window**. Open *REPMEAS1.SYS*.

- Select **Options|Recode** from the **Data Editor**.

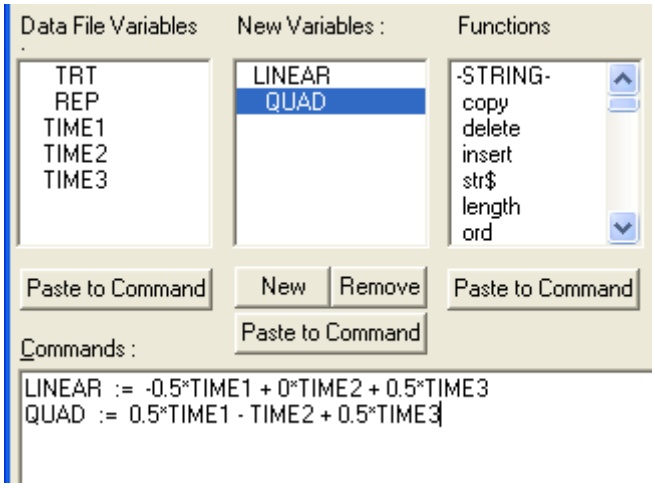
- Click the **New** button of the dialog box. Type *LINEAR* then click **OK**.

- Enter the formula *LINEAR=-0.5*TIME1 + 0*TIME2 + 0.5*TIME3*. Select *LINEAR* and click **Paste** under the **New Variables** list. Click *:=* operator from the numeric keypad. Enter the rest of the formula using the numeric keypad and **Paste to Command** key.



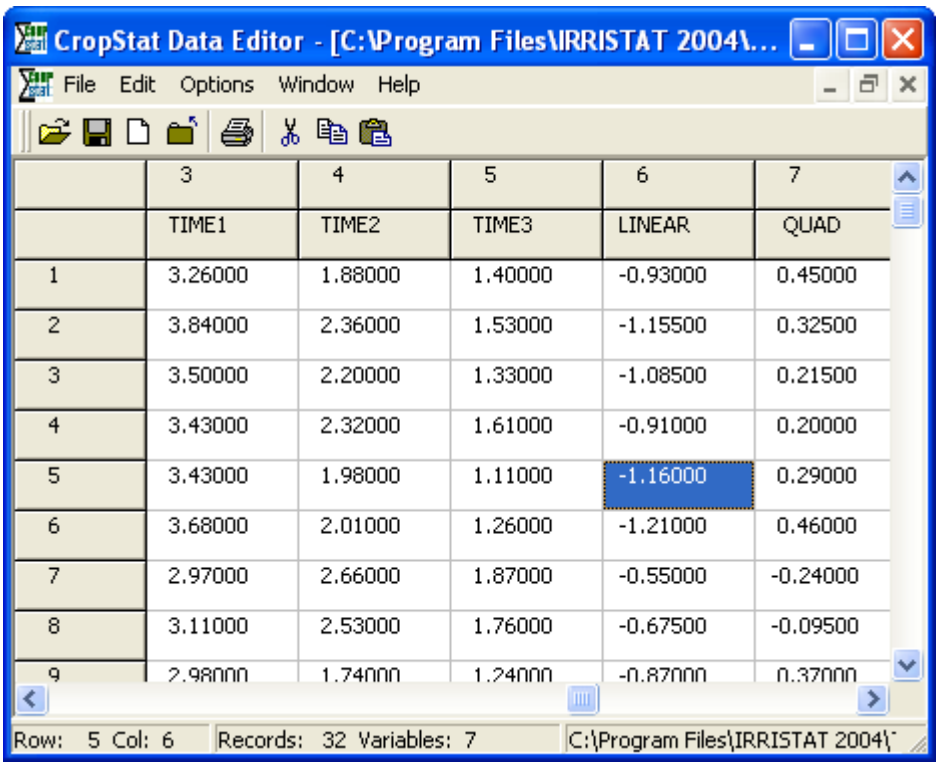
- Click the **New** button of the dialog box. Type *QUAD* then click **OK**.

- Enter the formula *QUAD=0.5*TIME1 - TIME2 + 0.5*TIME3*. Select *QUAD* and click **Paste** under the **New Variables** list. Click *:=* operator from the numeric keypad. Enter the rest of the formula using the numeric keypad and **Paste to Command** key.



- Click **Save** button to save the formula. CropStat will write it to the heading of your data file. Click **Run** to generate the values for the new variable.

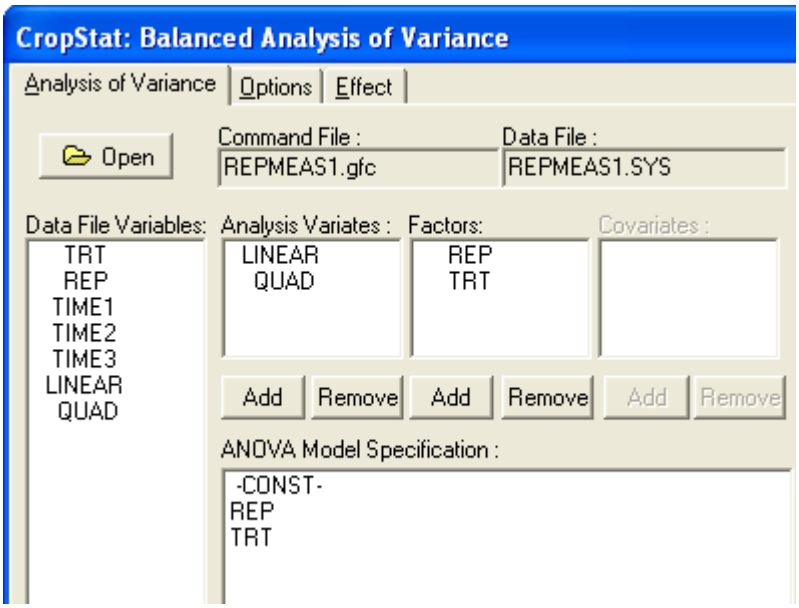
- Click **OK** to close the **Transformation** dialog.



	3	4	5	6	7
	TIME1	TIME2	TIME3	LINEAR	QUAD
1	3.26000	1.88000	1.40000	-0.93000	0.45000
2	3.84000	2.36000	1.53000	-1.15500	0.32500
3	3.50000	2.20000	1.33000	-1.08500	0.21500
4	3.43000	2.32000	1.61000	-0.91000	0.20000
5	3.43000	1.98000	1.11000	-1.16000	0.29000
6	3.68000	2.01000	1.26000	-1.21000	0.46000
7	2.97000	2.66000	1.87000	-0.55000	-0.24000
8	3.11000	2.53000	1.76000	-0.67500	-0.09500
9	2.98000	1.74000	1.24000	-0.87000	0.37000

- Save the file *REPMEAS1.SYS* and exit the **Data Editor**.

- After creating the summary variable, perform an ANOVA on the newly created variables as you would for any variable of the experiment. That is, if the experiment was conducted in a randomized complete block (RCB) design, then ANOVA using RCB should be performed on the summary variable. The dialog box for the ANOVA should be as follows:



IV. Sample output

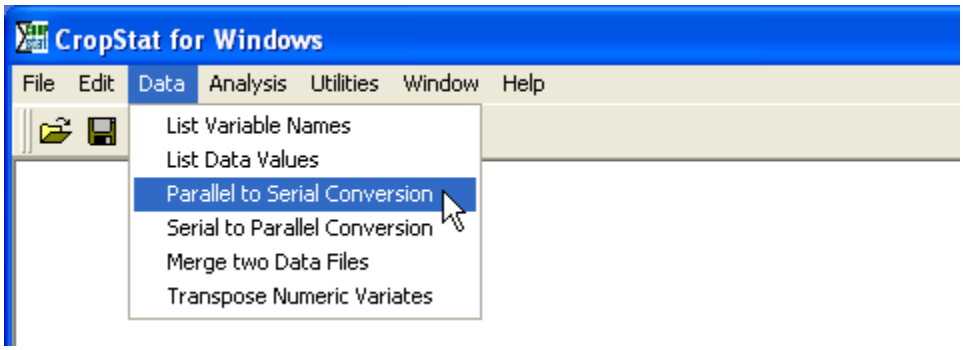
```
BALANCED ANOVA FOR VARIATE   LINEAR  FILE REPMEAS1   15/ 3/ 5 10:36
------------------------------------------------------------------- :PAGE   1
VARIATE V006 LINEAR

LN   SOURCE OF VARIATION            DF  SUMS OF     MEAN     F RATIO PROB  ER
                                        SQUARES     SQUARES                LN
=============================================================================
1 REP                                3 .758407E-01 .252802E-01   1.26 0.315  3
2 TRT                                7 .728534     .104076       5.17 0.002  3
* RESIDUAL                          21 .422422     .201153E-01
-----------------------------------------------------------------------------
* TOTAL (CORRECTED)                 31 1.22680     .395741E-01
-----------------------------------------------------------------------------

BALANCED ANOVA FOR VARIATE     QUAD  FILE REPMEAS1   15/ 3/ 5 10:36
------------------------------------------------------------------- :PAGE   2
VARIATE V007 QUAD

LN   SOURCE OF VARIATION            DF  SUMS OF     MEAN     F RATIO PROB  ER
                                        SQUARES     SQUARES                LN
=============================================================================
1 REP                                3 .229728     .765761E-01   1.80 0.177  3
2 TRT                                7 3.16406     .452009      10.63 0.000  3
* RESIDUAL                          21 .892634     .425064E-01
-----------------------------------------------------------------------------
* TOTAL (CORRECTED)                 31 4.28642     .138272
-----------------------------------------------------------------------------

TABLE OF MEANS FOR FACTORIAL EFFECTS  FILE REPMEAS1   15/ 3/ 5 10:36
------------------------------------------------------------------- :PAGE   3
MEANS FOR EFFECT REP
-----------------------------------------------------------------------------

     REP         NOS       LINEAR         QUAD
1                  8       -.959375     0.200625
2                  8       -.936250     0.138750
3                  8       -.843125     -.687502E-02
4                  8       -.862500     0.212499E-01

SE(N=   8)                 0.501440E-01 0.728924E-01
5%LSD  21DF                0.147472     0.214375
-----------------------------------------------------------------------------

MEANS FOR EFFECT TRT
-----------------------------------------------------------------------------

     TRT         NOS       LINEAR         QUAD
1                  4       -.807500     0.295000
2                  4       -1.10500     0.447500
3                  4       -.955000     0.325000E-01
4                  4       -.943750     0.108750
5                  4       -1.03250     0.470000
6                  4       -.951250     0.762500E-01
7                  4       -.582500     -.542500
8                  4       -.825000     -.180000

SE(N=   4)                 0.709143E-01 0.103085
5%LSD  21DF                0.208557     0.303172
-----------------------------------------------------------------------------

ANALYSIS OF VARIANCE SUMMARY TABLE  FILE REPMEAS1   15/ 3/ 5 10:36
------------------------------------------------------------------- :PAGE   4

 F-PROBABLIITY VALUES FOR EACH EFFECT IN THE MODEL. SECTION - 1

 VARIATE       GRAND MEAN  STANDARD   DEVIATION  C OF V |REP     |TRT     |
               (N=   32)   -------------------- SD/MEAN |        |        |
              NO.           BASED ON   BASED ON    %    |        |        |
              OBS.          TOTAL SS   RESID SS         |        |        |
  LINEAR       32 -.90031    0.19893    0.14183     15.8 0.3147   0.0016
    QUAD       32 0.88438E-010.37185    0.20617    233.1 0.1767   0.0000
```

V. REML Analysis of Repeated Measurements

- Create a SYS file for the data set used above. There should be only one variable containing all times of measurement.

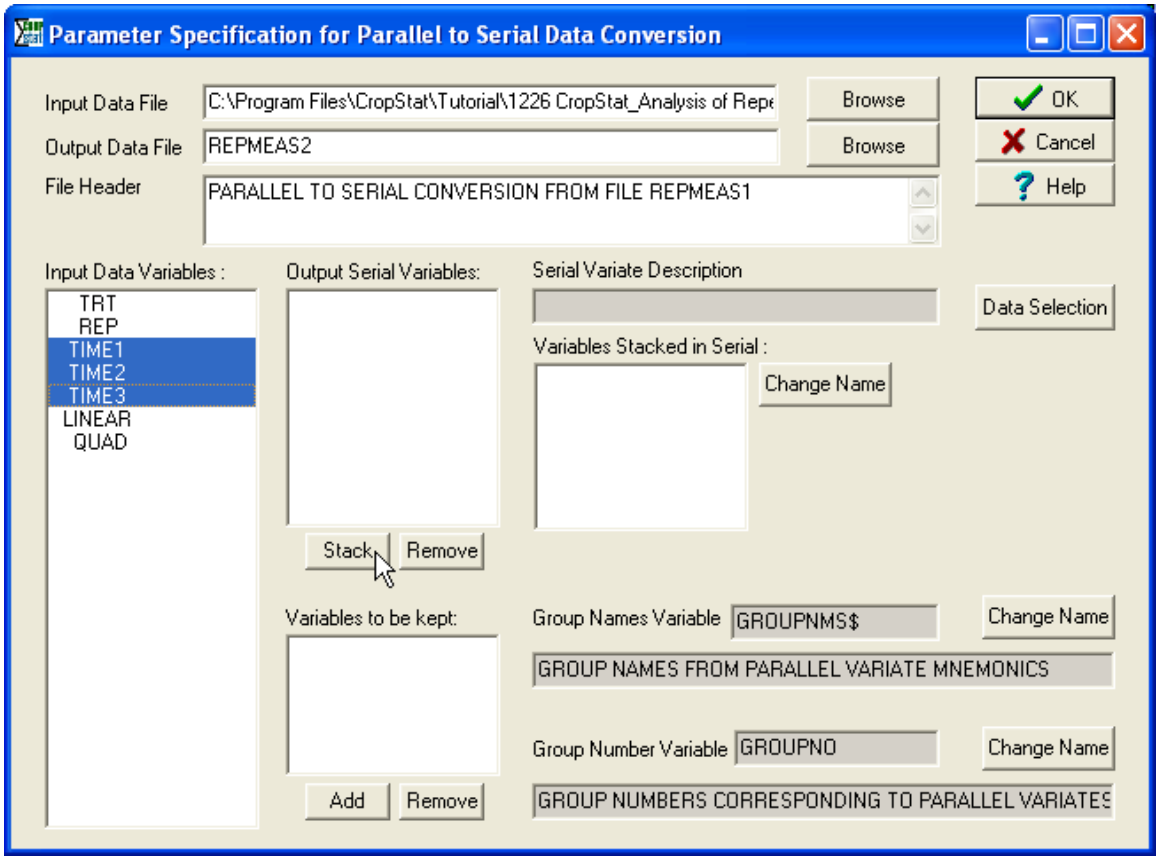
TIME	NITCONT	REP	TRT
1	3.26	1	1
1	3.84	1	2
1	3.5	1	3
1	3.43	1	4
1	3.43	1	5
1	3.68	1	6
1	2.97	1	7
1	3.11	1	8
1	2.98	2	1
1	3.74	2	2
…	…	…	…
2	1.88	1	1
2	2.36	1	2
2	2.2	1	3
2	2.32	1	4
…	…	…	…

- To save the data file in the format shown above, from the **Main Window**, go to **Data|Parallel to Serial Conversion**.

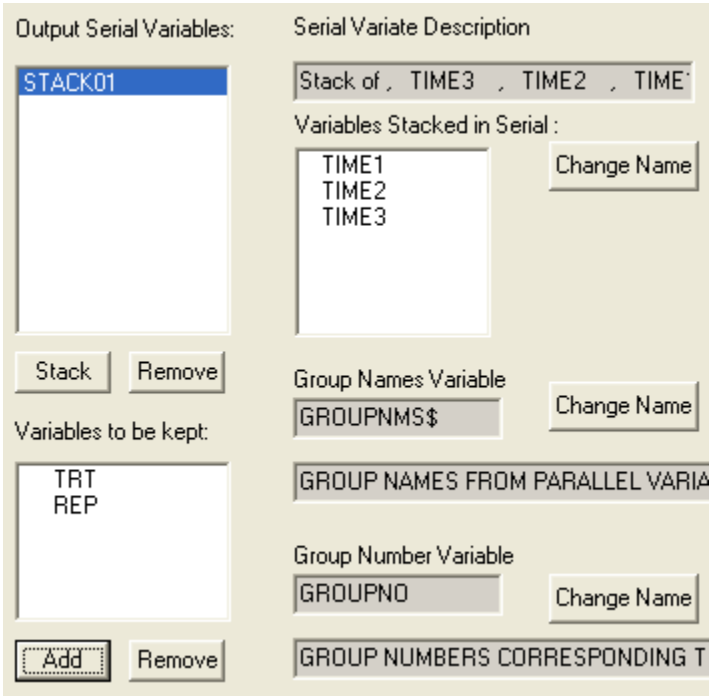


- Open *REPMEAS1.SYS* .In the **Save As** dialog box type a different filename, say, *REPMEAS2.SYS*.

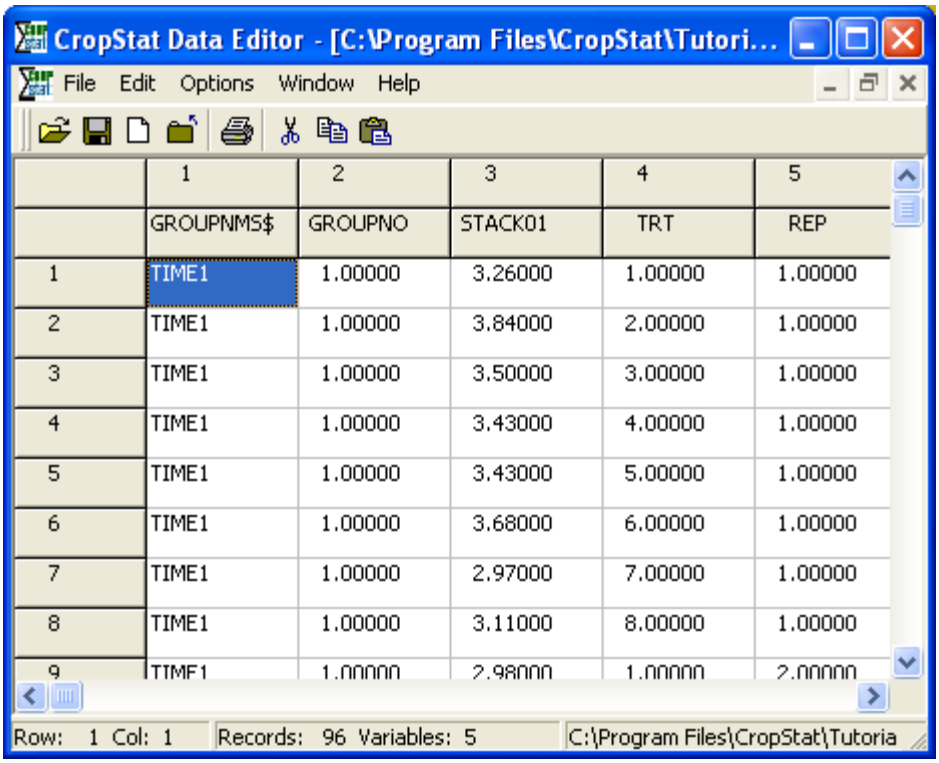
- In the **Parameter Specification for Parallel to Serial Data Conversion** box select *TIME1, TIME2* and *TIME3* from the **Input Data Variables** list and click the **Stack** button below the **Output Serial Variables** box.



- Select *REP* and *TRT* from the **Input Data Variables** list and click the **Add** button below the **Variables to be kept** box.

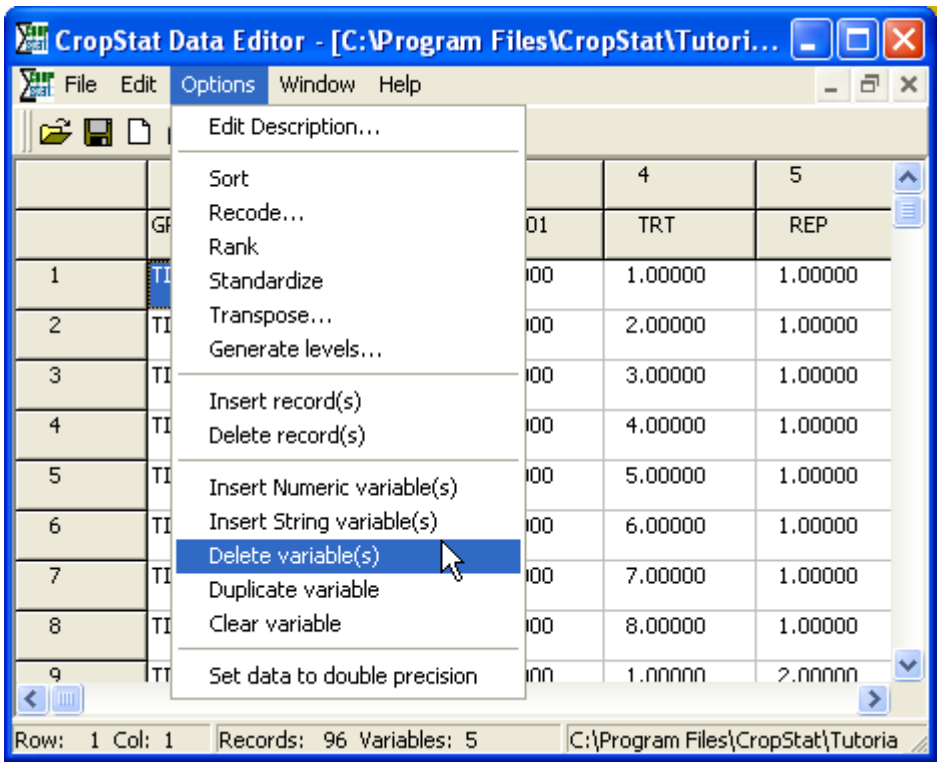


- Click the **OK** button.
- The newly created data file *REPMEAS2.SYS* will automatically open in the **Data Editor**.

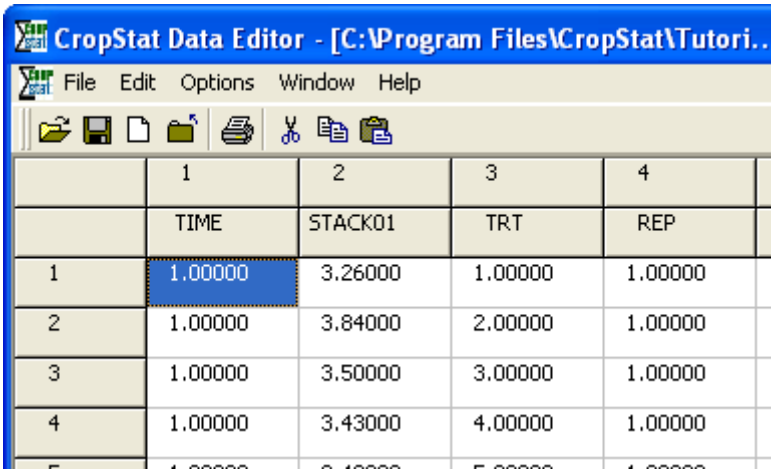


	1	2	3	4	5
	GROUPNMS$	GROUPNO	STACK01	TRT	REP
1	TIME1	1.00000	3.26000	1.00000	1.00000
2	TIME1	1.00000	3.84000	2.00000	1.00000
3	TIME1	1.00000	3.50000	3.00000	1.00000
4	TIME1	1.00000	3.43000	4.00000	1.00000
5	TIME1	1.00000	3.43000	5.00000	1.00000
6	TIME1	1.00000	3.68000	6.00000	1.00000
7	TIME1	1.00000	2.97000	7.00000	1.00000
8	TIME1	1.00000	3.11000	8.00000	1.00000
9	TIME1	1.00000	2.98000	1.00000	2.00000

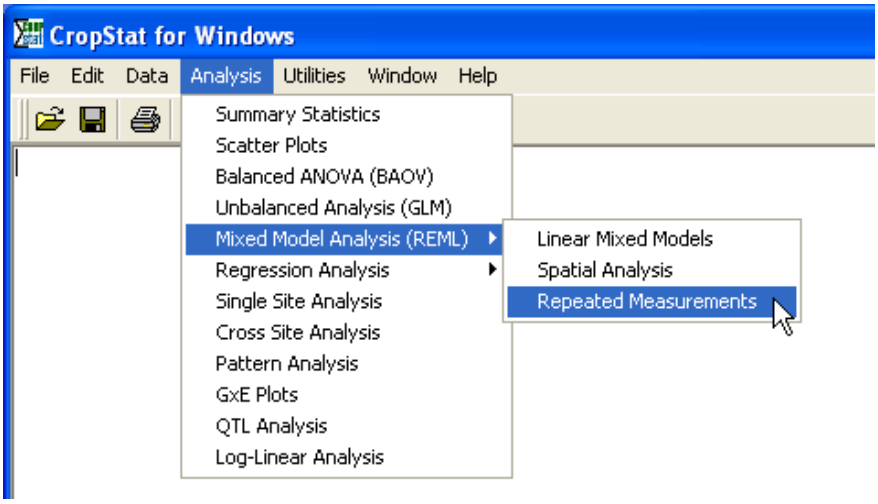
- The variables *GROUPNMS$* and *GROUPNO* both refers to the time the observations were taken, hence, we can delete either one of the variables. In this example we'll delete *GROUPNMS$*. Click any cell under the *GROUPNMS$* variable. From the menu go to **Options** then select **Delete variable(s)**.



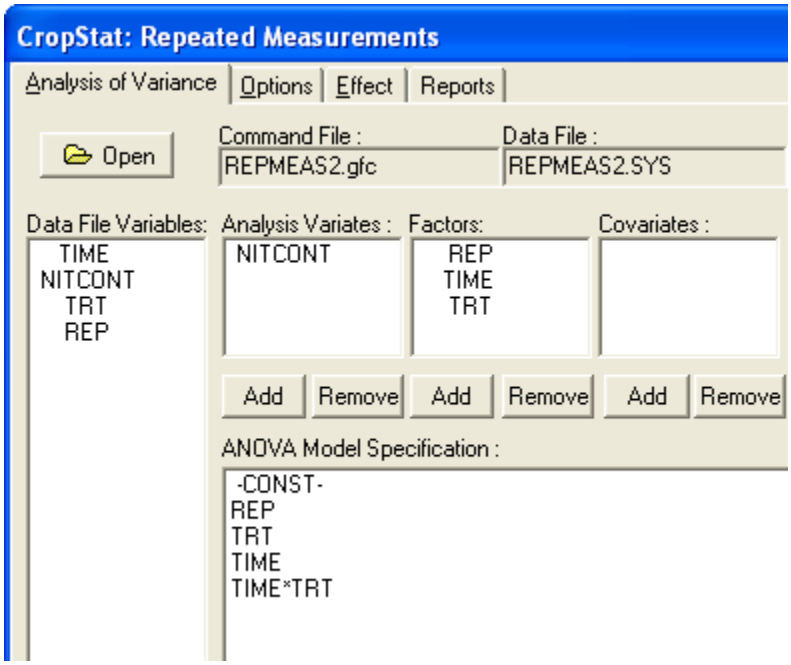
- To rename *GROUPNO* to *TIME*, click on the column header *GROUPNO*. In the **Variable name** box replace *GROUPNO* with *TIME*. Do the same for *STACK01* by renaming it to *NITCONT*.



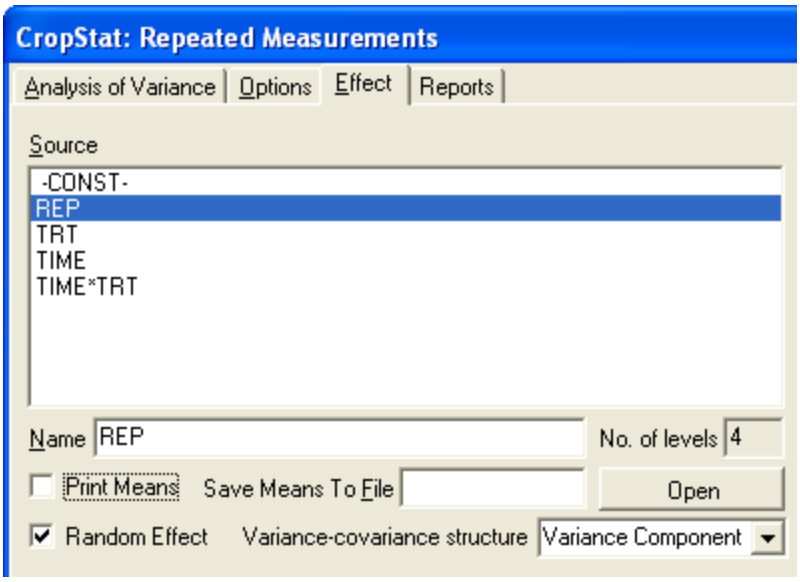
- Save then close the file.
- Choose **Mixed Model Analysis|Repeated Measurements** from the Analysis Menu.



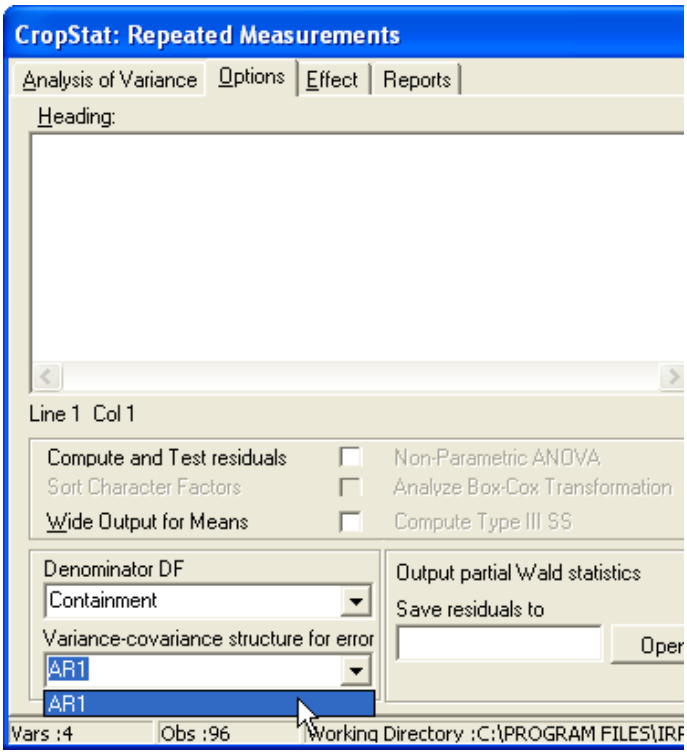
- In the **Open** dialog box, click the **Look In** box and open the *MY CROPSTAT\ REPEATED MEASURES* folder.
- In the **File name** box, type *REPMEAS2* as the name of the command file.
- Since no command file *REPMEAS2.GFC* exists, click the **Yes** button to create new a command file.
- In the next **Open** dialog box, type *REPMEAS2.SYS* as the data to be analyzed.
- The **Repeated Measurements** dialog box will open. Fill-in the information needed for the analysis.
- Specify *NITCONT* as the **variate** and *REP, TIME TRT* as **factors**. Include *TIME* and all its interaction with the other factors except *REP* in the model



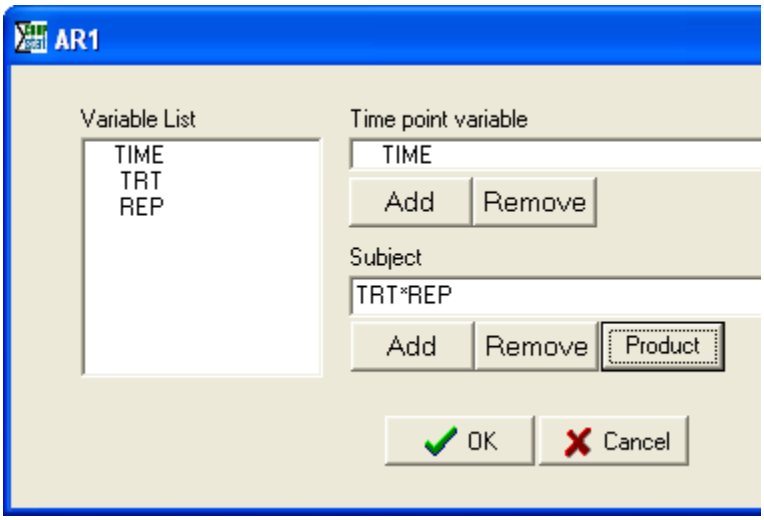
- To specify the random effect click the **Effect** tab.
- From the **Source** box select *REP* and click on the **Random Effect** checkbox. Leave the default structure 'Variance Component' for the *REP* effect.



- In the **Options** tab, specify the variance structure for the error by clicking on the arrow in the **Variance-Covariance Structure for Error** list box.



- Click AR1 in the **Variance-Covariance Structure for Error** combo box. A pop-up window will appear, the **AR1** dialog box for repeated measures. Specify *TIME* as the **Time point variable** and *TRT*REP* as **Subject**.



- Click the **OK** button to return to the **Options** tab. Click the **Ok** button to run the analysis.

VI. Sample Output

```
IRREML 1.0.2: REML ANALYSIS end  FILE REPMEAS2   12/ 5/ 5 15: 9
 ------------------------------------------------------------------ :PAGE   1
 THE IRREML PROGRAM WAS WRITTEN BY DOUGLAS CLARKSON OF SCIENCEOPS FOR IRRI

 Command File: C:\MY CROPSTAT\REPEATED MEASURES\REPMEAS2.gfc   Data File: REPMEAS2

 Number of Records:    96

 Variables in Data Set: TIME  REP  TRT  NITCONT

 SUMMARY STATISTICS FOR NUMERIC VARIATES
 VARIATE      NOBS   MINIMUM     MAXIMUM     MEAN     STD. DEV.
 NITCONT        96   1.110       3.840       2.241      0.7963

 Classification Variables: TIME  REP  TRT

 Levels of the classification variables

    3 CODES:(Number  Label) for Variable: TIME
(  1            1)(  2            2)(  3            3)(

    4 CODES:(Number  Label) for Variable: REP
(  1            1)(  2            2)(  3            3)(  4            4)(

    8 CODES:(Number  Label) for Variable: TRT
(  1            1)(  2            2)(  3            3)(  4            4)(  5            5)
(  6            6)(  7            7)(  8            8)(


IRREML: REML ANALYSIS FOR VARIATE NITCONT  FILE REPMEAS2   12/ 5/ 5 15: 9
 ------------------------------------------------------------------ :PAGE   2
 Number of non-missing dependent observations:           96

 Model Specification
 Intercept in model: Yes
 The Fixed Effects Model
     NITCONT = Intercept + TRT + TIME + TIME:TRT
 The Random Effects Terms
     REP

 RANDOM EFFECT COVARIANCE MODEL.  1 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 REP                          1- 1  diagonal

 RESIDUAL EFFECT COVARIANCE MODEL.  1 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 RESIDUAL                     1- 1  rAR1(TIME)xI(REP))xI(TRT))x

 Number of columns in the fixed effects model:     24
 Number of columns in the random effects model:     4

 Message: Relative function convergence

 Final REML criterion:       63.718304862062666

 Variance/Covariance component parameters
 Dep Name        Gamma Coef.  Std. Error    Z       Pr > |Z|   Var. Component  Std. Error
  1 rAR1(1) ..  0.8884E-01   0.1540      0.5768      0.5641
  1 REP(1) ...   0.2857      0.2775       1.030      0.3032      0.1045E-01  0.9930E-02
```

Correlation coefficient across time.

Variance component of *REP*

```
 The scale parameters
 Dep.          Sigma_Squared Std. Error      Z         Pr > |Z|
Dep(1) .......   0.3657E-01  0.6304E-02   5.801      0.6587E-08

 Asymptotic Covariance Matrix of the Variance/Covariance Components
                                1              2              3
         1   1 rAR1(1)..   0.237E-01  -0.197E-02   0.152E-03
         2   1 REP(1)...  -0.197E-02   0.770E-01  -0.365E-03
         3 Dep(1).......   0.152E-03  -0.365E-03   0.397E-04

 Warning: Denominator degrees of freedom estimates do not account
 for measurement error parameters.

      ANOVA Table for Sequentially Deleted Fixed Effects
      Denominator Degrees of Freedom: Residual DF
 Dep Effect    DFNum   DFDen   F - Statistic   P > |F|
  1 TIME:TRT     14  72.00        7.528       0.2065E-08
  1     TIME      2  72.00        717.5       0.2834E-47
  1      TRT      7  72.00        4.077       0.8087E-03

IRREML:  PREDICTIONS  FILE REPMEAS2   12/ 5/ 5 15: 9
 ------------------------------------------------------------------- :PAGE   3

                Balanced Least Squares Means Fixed
 Dep Level              LSMean   Std. Error
  1  TRT            1    2.042      0.7769E-01
  1  TRT            2    2.253      0.7769E-01
  1  TRT            3    2.379      0.7769E-01
  1  TRT            4    2.272      0.7769E-01
  1  TRT            5    2.113      0.7769E-01
  1  TRT            6    2.191      0.7769E-01
  1  TRT            7    2.401      0.7769E-01
  1  TRT            8    2.275      0.7769E-01

 Standard Errors of Differences
  Minimum     Mean      Maximum
  0.8276E-01  0.8276E-01  0.8276E-01

                Balanced Least Squares Means Fixed
 Dep Level              LSMean   Std. Error
  1  TIME           1    3.171      0.6128E-01
  1  TIME           2    2.182      0.6128E-01
  1  TIME           3    1.370      0.6128E-01
```

Error variance

Test for fixed effect.

```
Standard Errors of Differences
 Minimum     Mean      Maximum
 0.4564E-01  0.4630E-01  0.4762E-01

                        Balanced Least Squares Means Fixed
Dep Level                                  LSMean   Std. Error
 1  TIME:TRT            1            1    2.947      0.1084
 1  TIME:TRT            2            1    1.845      0.1084
 1  TIME:TRT            3            1    1.333      0.1084
 1  TIME:TRT            1            2    3.507      0.1084
 1  TIME:TRT            2            2    1.955      0.1084
 1  TIME:TRT            3            2    1.297      0.1084
 1  TIME:TRT            1            3    3.345      0.1084
 1  TIME:TRT            2            3    2.358      0.1084
 1  TIME:TRT            3            3    1.435      0.1084
 1  TIME:TRT            1            4    3.252      0.1084
 1  TIME:TRT            2            4    2.200      0.1084
 1  TIME:TRT            3            4    1.365      0.1084
 1  TIME:TRT            1            5    3.303      0.1084
 1  TIME:TRT            2            5    1.800      0.1084
 1  TIME:TRT            3            5    1.238      0.1084
 1  TIME:TRT            1            6    3.168      0.1084
 1  TIME:TRT            2            6    2.140      0.1084
 1  TIME:TRT            3            6    1.265      0.1084
 1  TIME:TRT            1            7    2.803      0.1084
 1  TIME:TRT            2            7    2.763      0.1084
 1  TIME:TRT            3            7    1.637      0.1084
 1  TIME:TRT            1            8    3.040      0.1084
 1  TIME:TRT            2            8    2.395      0.1084
 1  TIME:TRT            3            8    1.390      0.1084

Standard Errors of Differences
 Minimum     Mean      Maximum
 0.1291      0.1349      0.1352
```

Exercise 5

An experiment was conducted to compare the effect of three fertilizer treatments. A SPAD meter was used to measure the chlorophyll content of the leaves. Measurement was done every five days. Perform an analysis of repeated measures using orthogonal contrasts and mixed model. Interpret the results. The experimental design used was a randomized complete block design.

Data for SPAD reading taken 6 times every five days

Fert	Rep	SPAD1	SPAD2	SPAD3	SPAD4	SPAD5	SPAD6
1	1	32.96	34	34.2	35.83	35.6	35.63
2	1	34.56	37.7	38.26	36.8	38.63	36.86
3	1	36.4	37.36	37.13	35.5	36	39.4
1	2	34.1	36.3	37.13	32.66	35.06	35.7
2	2	35.63	37.9	33.96	35.93	37.16	40.7
3	2	34.86	37.36	37.1	37.36	38.1	40.4
1	3	34.56	35.73	33.36	33.36	31.7	35.16
2	3	35.53	37.9	36.16	36.16	38.43	36.9
3	3	32.23	36.6	38.36	38.36	39.93	40.7

Data are encoded and stored in an EXCEL file SPAD.XLS.

REGRESSION AND CORRELATION

At the end of the tutorial, the user should be able to

- perform simple and multiple regression analysis and test of residuals
- perform correlation analysis and test of residuals
- create scatterplots

I. Sample Problem

Data on grain yield (kg/ha) and percentage of nutrients N, P, and K in the grain were collected from 48 experimental plots. The data set is shown below. The researcher wants to investigate the relationship between the dependent variable (grain yield) and the three independent variables (N, P, and K). This translates into a multiple regression problem, which will be carried out using CROPSTAT. This data set is saved in the file *REG1.SYS* in the *CROPSTAT7.2\TUTORIAL\TUTORIAL DATASETS* folder.

Grain yield and percentage nutrient data*

PLOT	GY14	N%	P%	K%
01	1678	0.9849	0.0901	0.3987
02	4265	1.1714	0.0926	0.3814
03	2431	1.0756	0.0886	0.4134
04	2431	1.0435	0.0807	0.4027
05	4461	1.2101	0.0852	0.3851
06	3110	1.2084	0.0845	0.3746
07	4469	1.1643	0.0972	0.3794
08	4194	1.2369	0.0850	0.4373
09	3379	1.0247	0.2901	0.3108
10	7132	1.4624	0.2467	0.2545
11	4359	0.9954	0.2872	0.3095
12	3646	1.0404	0.2487	0.3137
13	6917	1.5309	0.2868	0.3071
14	6692	1.5001	0.2776	0.2506
15	7028	1.6112	0.2852	0.3091
16	7196	1.6009	0.3104	0.3168
17	4220	0.9392	0.2558	0.3039
18	7250	1.6738	0.1985	0.2474
19	4915	0.9525	0.2712	0.3116
20	4185	0.8468	0.2474	0.3051
21	7463	1.6020	0.3213	0.3144
22	7367	1.6678	0.1793	0.2295
23	7860	1.5521	0.3058	0.3296
24	7624	1.5707	0.2969	0.3078

PLOT	GY14	N%	P%	K%
25	2830	0.8732	0.2582	0.3204
26	3705	1.0743	0.2396	0.2390
27	3280	1.0124	0.2665	0.3016
28	2906	0.9288	0.2705	0.3078
29	4041	1.0279	0.2492	0.2925
30	3295	1.1412	0.2430	0.2475
31	4147	1.1594	0.2867	0.2924
32	4102	1.0733	0.2624	0.2736
33	3509	0.8955	0.2395	0.3815
34	5087	1.2949	0.2548	0.2978
35	4353	1.0541	0.2604	0.3597
36	3915	0.9689	0.2469	0.3806
37	5122	1.4547	0.3239	0.3195
38	4660	1.2749	0.2454	0.2556
39	5150	1.3373	0.3057	0.3158
40	5730	1.2832	0.3059	0.3104
41	3744	0.9977	0.2103	0.2943
42	5363	1.7496	0.1617	0.2353
43	4628	1.0091	0.2208	0.2993
44	4212	0.8938	0.2008	0.2919
45	5063	1.6833	0.2609	0.2983
46	5015	1.7700	0.1462	0.2204
47	6244	1.6487	0.2495	0.3163
48	5483	1.6799	0.2417	0.2923

*GY14 refers to grain yield (kg/ha) at 14% moisture content; N%, P%, and K% refer to percentage of nitrogen, phosphorus, and potassium in the grain, respectively.

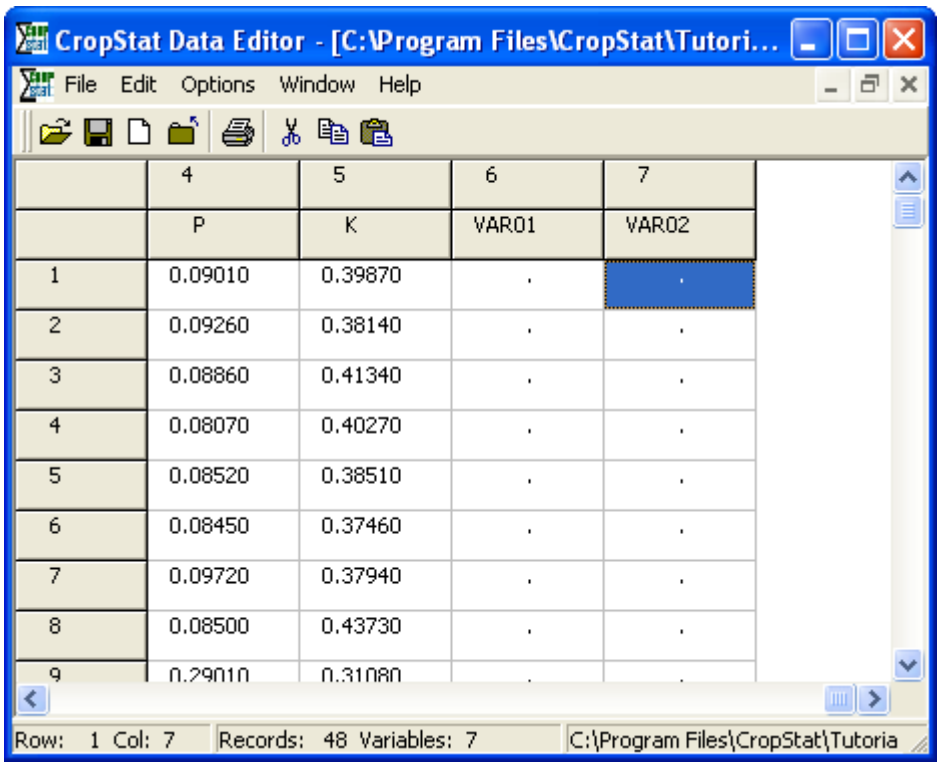
II. Regression and correlation analysis

- Open the data file *REG1.SYS* from the *CROPSTAT7.2\TUTORIAL\TUTORIAL DATASETS* folder.

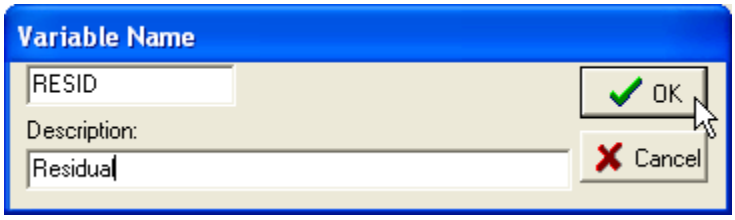
- Select **File ⇒ Save-as**. Click the **Save in** box and go inside your working folder *C:\MY CROPSTAT*. Create a subfolder REGRESSION AND CORRELATION then click **Save**

- Select **Window|Data Editor** from the Main Window. Enter *REG1.SYS* as the name of the file to open.

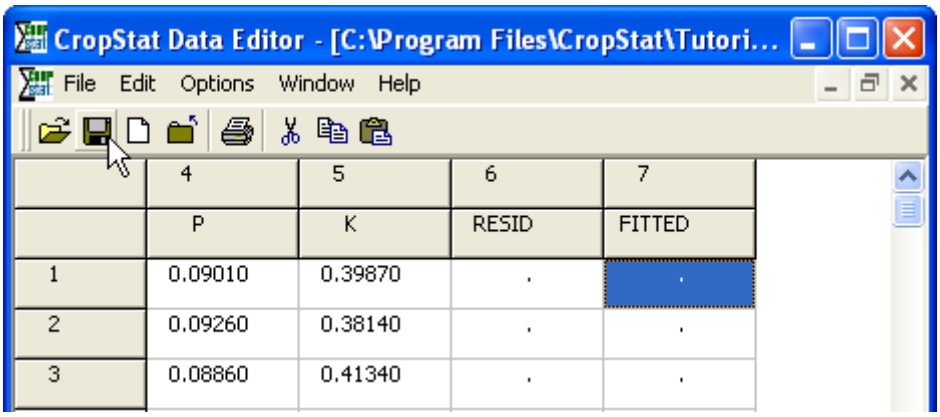
- Add two more variates in the data. These two additional variates will be used to hold residuals and fitted values after fitting the regression equation. To add new variates, use the right arrow key (→) to activate the column. The activated column has point (.) on its cell.



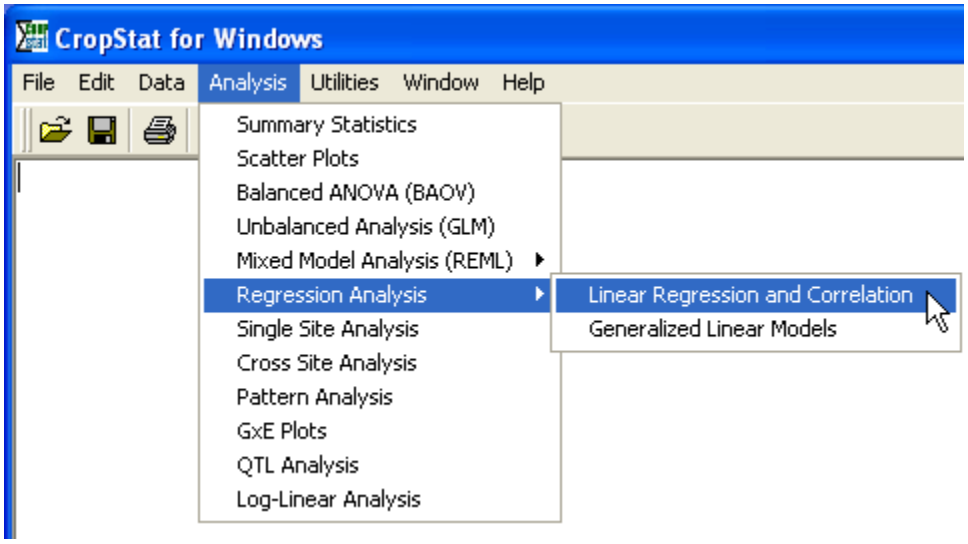
- To replace the variable name *VAR01* to *RESID*, click on *VAR01* (left button) and a dialog box will appear. Delete *VAR01* and type *RESID* then click **OK**. To change *VAR02* to *FITTED* , follow the same procedure.



- Click the **SAVE** icon to save the changes made on the file and exit from the **Data Editor.**



- Select **Analysis|Regression Analysis|Linear Regression and Correlation** from the **Main Window**.



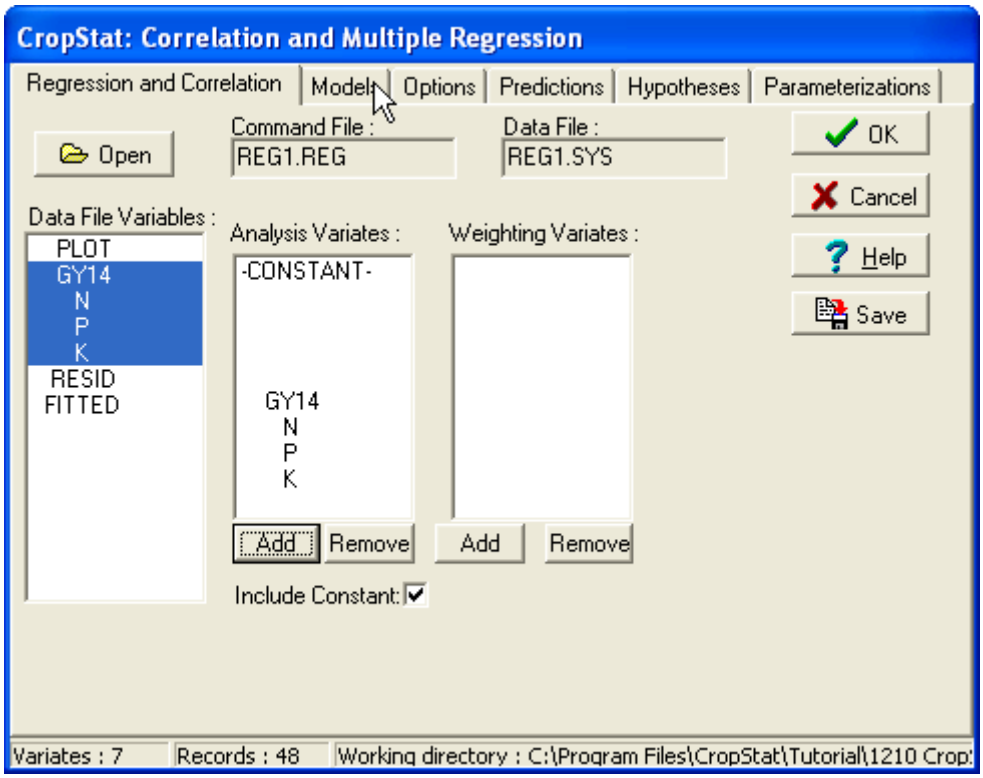
- In the **Open** dialog box, click the **Look in** box to switch to your working drive. *C:\MY CROPSTAT\REGRESSION AND CORRELATION*.

- To specify a name for the regression command file, type *REG1.REG* in the **File name** box. Click **Open**. Since no command file called *REG1.REG* exists, click **Yes** in the message box that appears. This will instruct CropStat to create a new one.

- In the **Open Data File** dialog box, specify a name for the data file by entering *REG1.SYS* in the **File name** box. Click **Open**.

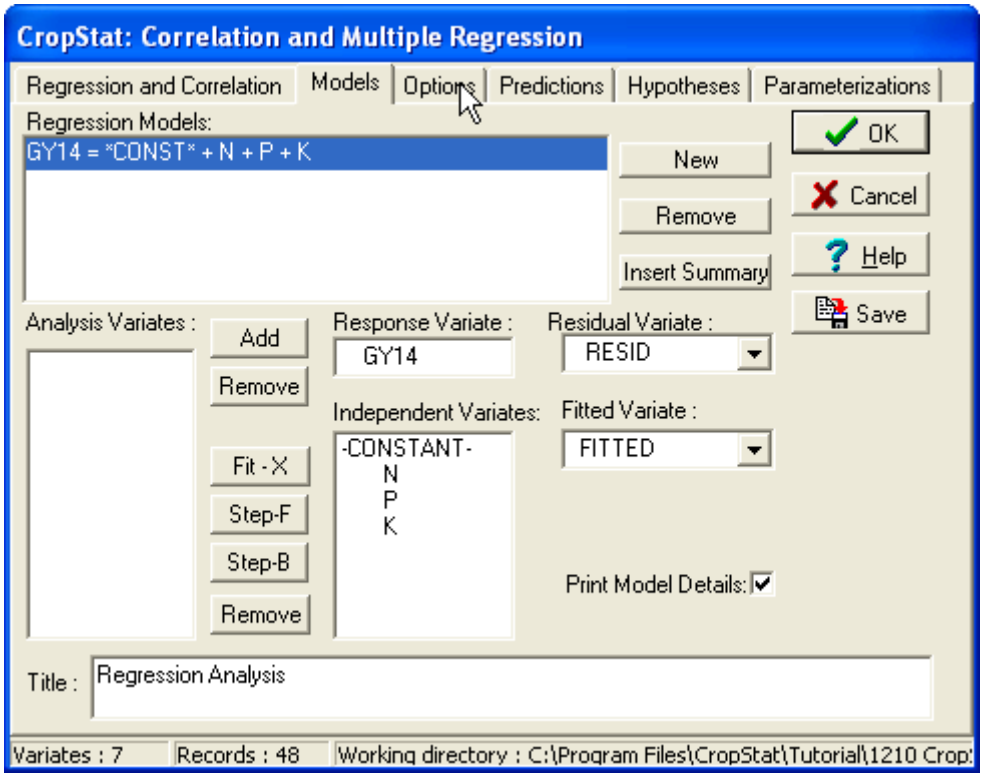
- Specify variates to be included in the model. Select *GY14*, *N*, *P,* and *K* from the **Data File Variables** list by clicking *GY14* and dragging the mouse down to *K*. (***Note***: Another way to select these variables is to click GY14, then click K while holding the <Shift key>.)

- Click **Add** under the **Analysis Variates** list.
- By default, CropStat includes **y**-intercept in the model _CONSTANT_ .

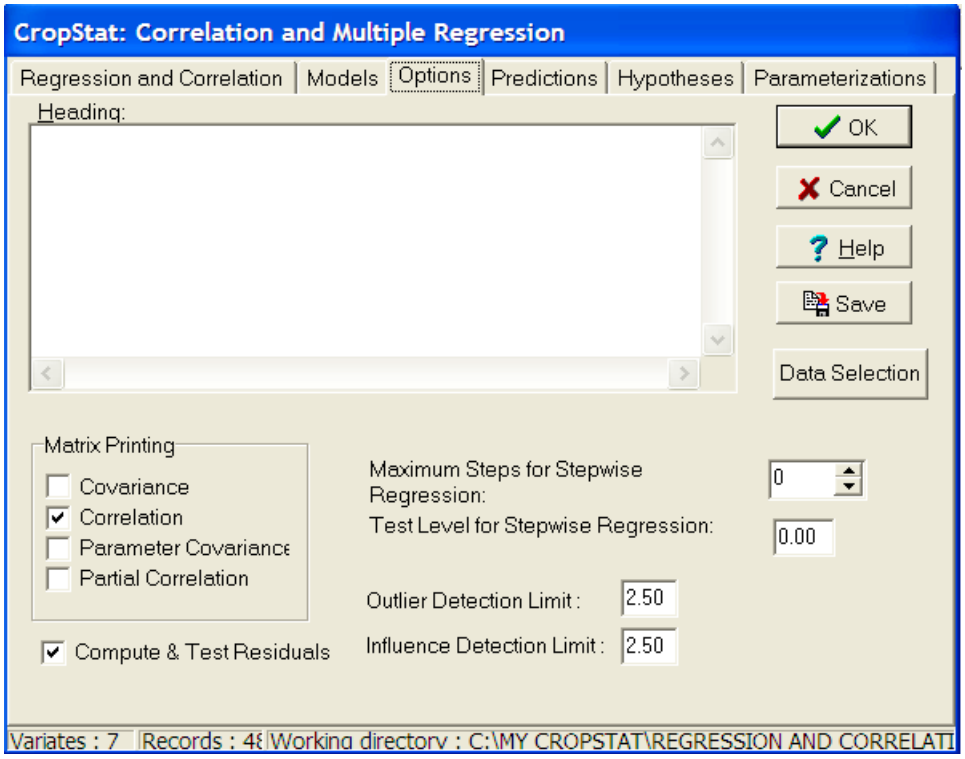


- To specify the regression model, click the **Models** tab.
- Click the **New** command button to enter a model in the **Regression Models** list.
- Select *GY14* from the **Analysis Variates** list. Click **Add** under the **Response Variate** box.
- Select *N*, *P*, and *K* from the **Analysis Variates** list. Click **Fit-X** on the **Independent Variates** list.
- To save residual values into the *REG1.SYS* file, specify the variate where you wish to write the residual values in the **Residual** edit box. Select *RESID* from the drop-down list.
- To save fitted values into the *REG1.SYS* file, specify the variate where you wish to write the fitted values in the **Fitted Variate** edit box. Select *FITTED* from the drop-down list. ***Note:*** *RESID* and *FITTED* are the columns we have just added to our data file. Any existing variates will do, but they will be overwritten if they already contain data.

- To add title, type *Regression Analysis* in the **Title** box.



- To instruct CropStat to print correlation, compute and test residuals, click the **Options** tab. Click the **Correlation** option in the **Matrix Printing** option group. Click the **Compute and Test Residuals** option. NOTE: If the **Compute and Test Residuals** option is not ticked the variables *RESID* and *FITTED* will not be filled-up with data. Click **OK** to run the regression analysis.



III. Sample Output

The following output saved in *REG1.OU*T appears on the Text Editor.

1. Correlation matrix

```
CORRELATION MATRIX AND STANDARD ERRORS  FILE REG1    1/10/ 4 16:13
------------------------------------------------------------------ :PAGE   1

  SECTION  1  CORR AND RESID. SES
            GY14           N              P              K
GY14          0.1552269E+04
N             0.7036688E+00  0.3306954E+00
P             0.4323169E+00  0.7691458E-01  0.7355986E-01
K            -0.4110870E+00 -0.2799415E+00 -0.4631045E+00  0.5090120E-01
```

2. Regression analysis

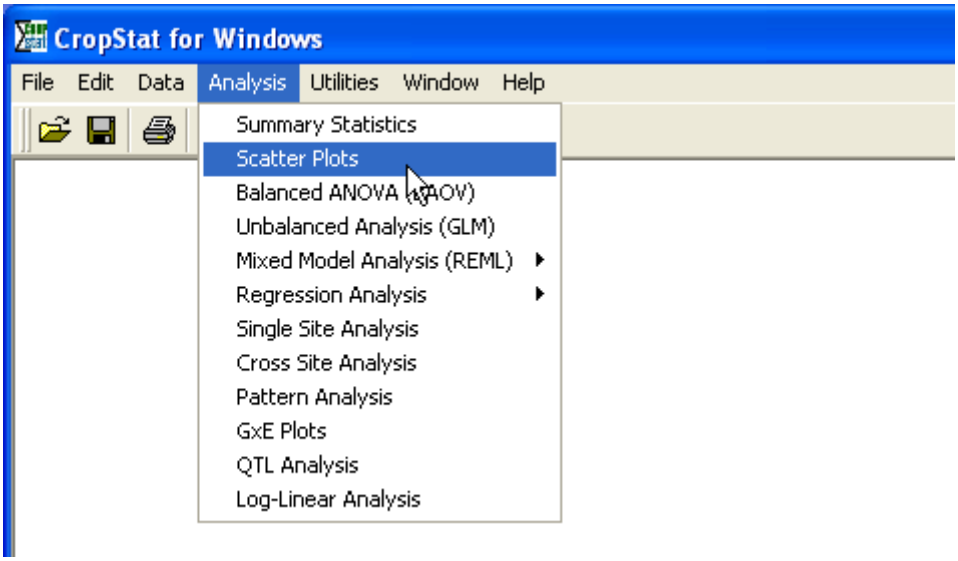
```
REGRESSION OF     GY14    (2  ) ON 4   VARIATES  FILE REG1    1/10/ 4 16:13
------------------------------------------------------------------ :PAGE   2

    TERM      NO  COEFFICIENT   STD. ERROR   F-VALUE  PRBF  TOLERANCE

CONSTANT       8  -99.63761      1541.00       0.004  0.947  0.8085E-02
       N       3   3092.318      441.910      48.967  0.000  0.9181
       P       4   7432.484      2151.88      11.930  0.001  0.7825
       K       5  -1938.088      3229.70       0.360  0.559  0.7255

    SOURCE         SS            DF    MS           F       PRBF

    REGRESSION   0.7270118E+08    3  0.2423373E+08  26.297  0.0000
    RESIDUAL     0.4054720E+08   44   921527.3
    TOTAL        0.1132484E+09   47

    R-SQUARED - 64.2%, (ADJUSTED FOR D.F.- 61.8%)
```

3. Residual analysis

```
.
.
.
RECORD  32 HAS INFLUENTIAL X-VAL. OBS=   4102.     RES=   2555.     FIT=   1547.
RECORD  32 HAS RESIDUAL=  3.3 SES OBS=   4102.     RES=   2555.     FIT=   1547.
RECORD  46 HAS INFLUENTIAL X-VAL. OBS=   5015.     RES=  -1018.     FIT=   6033.

BOX PLOT OF STUDENTIZED RESIDUALS FROM LPLT= -1.578    TO ULPT=  3.306
NO.<LPLT                                                      NO.>UPLT
       0 --------I         +          I------------------------------   0

MEDIAN= -0.1137E+00 ANDERSON-DARLING STATISTIC=  0.567
```

IV. Regression Scatterplot

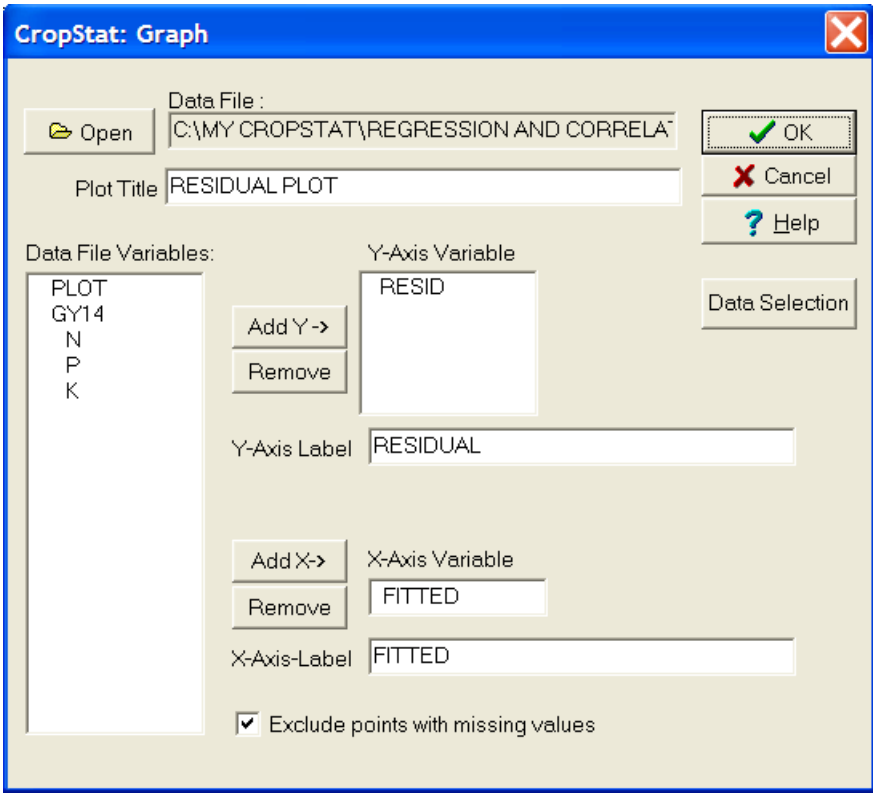
- Select **Analysis | Scatter Plots** from the Main Window.



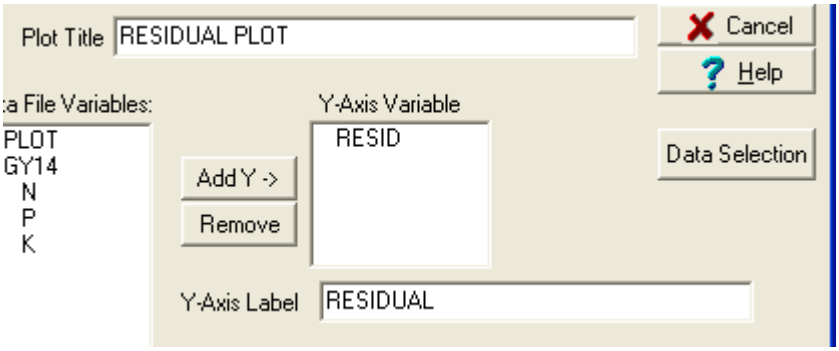
- In the **Open** dialog box, click the **Look In** box and select your working drive *C:\MY CROPSTAT\REGRESSION AND CORRELATION*.

- Enter *REG1.SYS* in the **File name** box. Click **Open**.

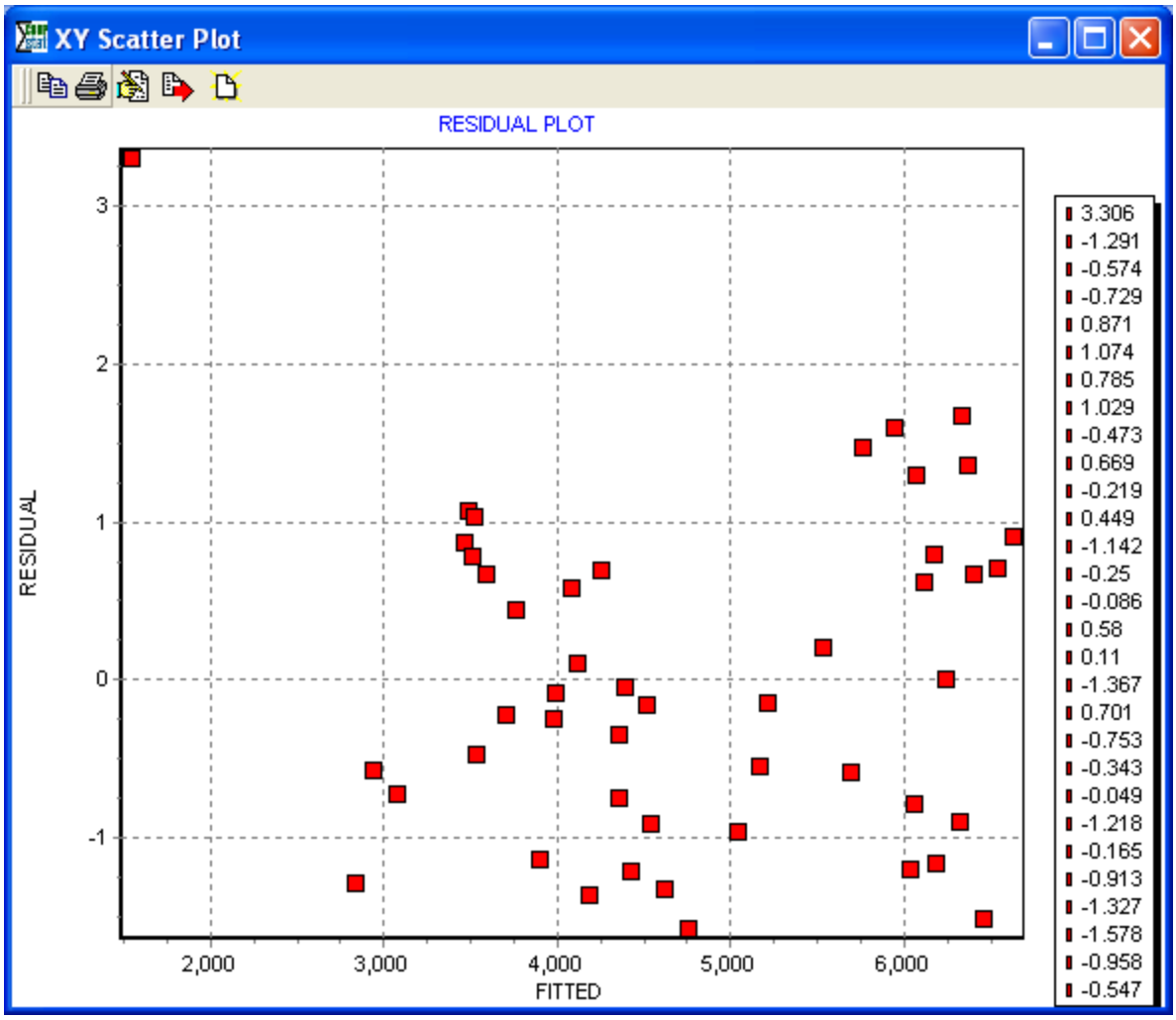
- In the **Graph Data** dialog box, specify the X axis variate by clicking *FITTED* from the **Data File Variables** list. Click **Add X** to enter the selected variate into the **X-axis variable** box. Type *FITTED* in the **X-Axis-Label** box.



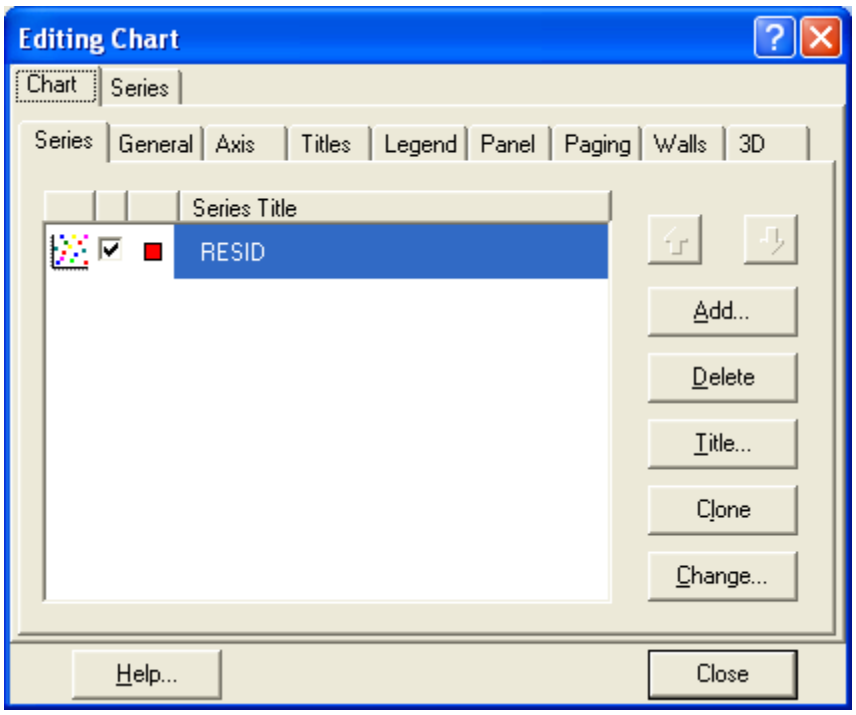
- Specify the **y axis variate**(s). Click *RESID* from the **Data File Variables** list. Click **Add Y** to enter the selected variate into **the Y-axis variable(s).** Type *RESIDUAL* in the **Y-Axis-Variable** box. Type *RESIDUAL PLOT* in the **Plot Title** box. Click **OK** to generate the graph.



- The output graph is shown below.



- If you would like to change the appearance of the plot, click  to edit chart. The following dialog will be available to facilitate chart editing.


Editing Chart
Chart
Series
Series
General
Axis
Titles
Legend
Panel
Paging
Walls
3D
Series Title
RESID
Add...
Delete
Title...
Clone
Change...
Help...
Close

Exercise 6

Given the following experimental data on rice yield (t/ha), plant height (cm) and tiller number, determine the relationships of these variables with each other using correlation and regression analysis. Obtain a model relating YIELD to the variables PLTHT and TILLER# and interpret results. Test for the significance of the parameter estimates and the regression equation. Evaluate the adequacy of the model obtained.

Observation	Yield (tha^{-1})	Plant Height (cm)	Tiller Number
1	5.75	110.5	14
2	5.94	105.4	16
3	6.01	118.1	15
4	6.54	104.5	18
5	6.73	93.6	15
6	6.75	84.1	18
7	6.90	77.8	18
8	7.86	75.6	19
9	6.56	96.2	17
10	6.40	92.6	14
11	7.92	76.4	19
12	5.60	112.1	13
13	5.81	109.5	14
14	6.33	89.8	17
15	6.95	78.3	18
16	7.25	75.9	19
17	5.50	111.8	15
18	5.88	108.9	14
19	6.86	92.5	16
20	7.46	78.9	19

SINGLE-SITE ANALYSIS FOR VARIETY TRIALS

At the end of the tutorial, the user should be able to

- perform analysis of variance for each site or season
- output treatment means to a SYS file

I. Sample Problem #1: Analysis of Randomized Complete Block (RCB) Design

To demonstrate single-site analysis, we will use genotype × environment (G×E) data from the upland rice ecosystem. This data set consists of 11 environments and 12 genotypes in a randomized complete block design with 4 replications. A partial listing of this data set is shown below.

```
GXE INTERACTIONS BETWEEN BREEDING SITES / 1994 RESULTS
COUNTRY$ SITE$ YEAR VARIETY$ VTYNO REP YIELD
 CIV     MB 94      AZU       1     1     0.6670
 CIV     MB 94      AZU       1     2     0.9580
 CIV     MB 94      AZU       1     3     0.3900
 CIV     MB 94      AZU       1     4     0.9360
 CIV     MB 94      BGORA     2     1     1.3290
 CIV     MB 94      BGORA     2     2     1.9600
 CIV     MB 94      BGORA     2     3     2.0870
 CIV     MB 94      BGORA     2     4     0.9350
 CIV     MB 94      GUAR      3     1     2.1730
 CIV     MB 94      GUAR      3     2     3.2250
 CIV     MB 94      GUAR      3     3     1.7650
 CIV     MB 94      GUAR      3     4     2.0970
 CIV     MB 94      IT146     4     1     2.6920
 CIV     MB 94      IT146     4     2     3.2940

...

 COL     CR 94      AZU       1     1     1.0930
 COL     CR 94      AZU       1     2     0.7600
 COL     CR 94      AZU       1     3     1.4380
 COL     CR 94      BGORA     2     1     1.3280
 COL     CR 94      BGORA     2     2     1.2150
```

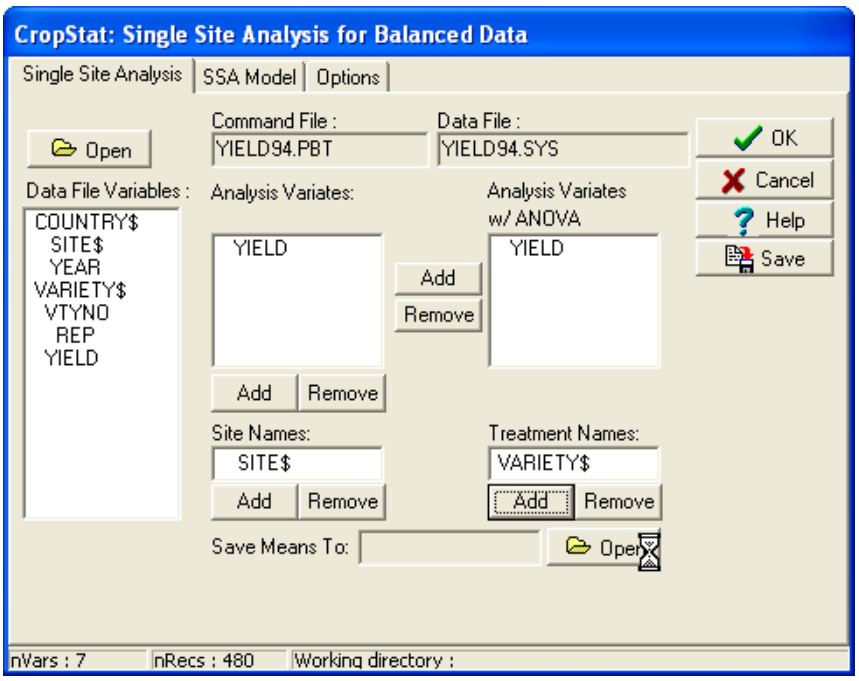
This data is already available in the file *YIELD94.SYS* in the *CROPSTAT7.2\TUTORIAL\TUTORIAL DATASETS* folder.

- Open the data file *YIELD94.SYS* from the *CROPSTAT7.2\TUTORIAL\TUTORIAL DATASETS* folder.

- Select **File ⇒ Save-as**. Click the **Save in** box and go go inside working folder *C:\MY CROPSTAT*. Create a subfolder SINGLE SITE ANALYSIS then click **Save.**

II. Steps for performing Single Site Analysis for Sample Problem #1

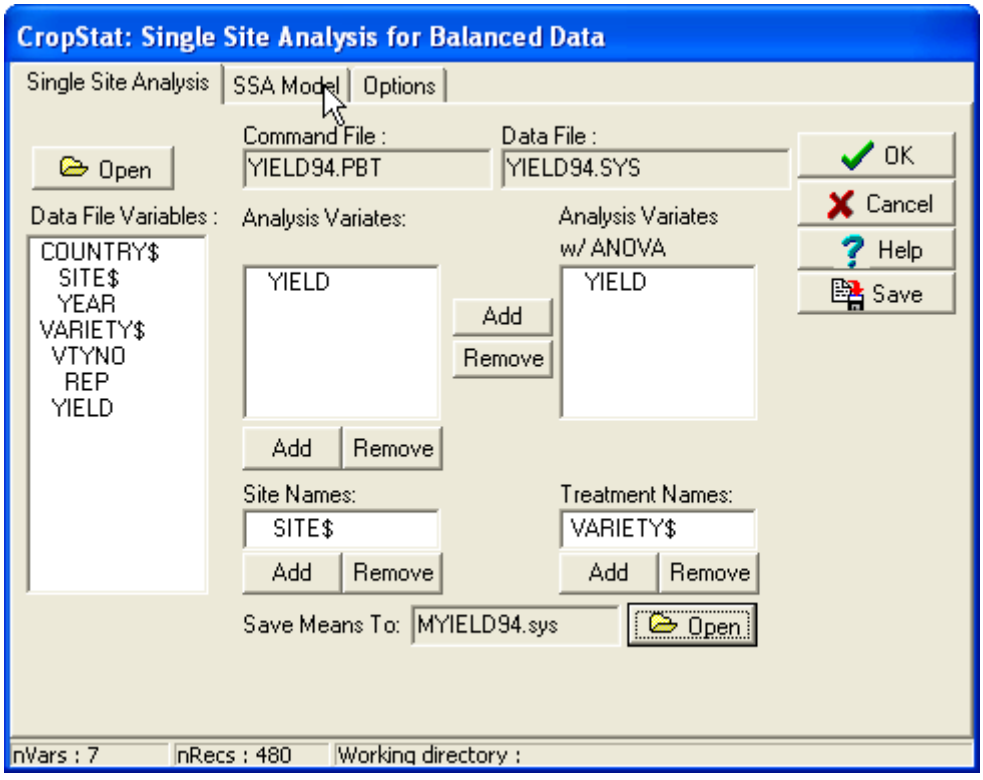
- Select **Analysis|Single Site Analysis** from the **Main Window**.
- Click the **Look in** box in the **Open** dialog box and go inside your working folder *C:\MY CROPSTAT\SINGLE SITE ANALYSIS*.
- In the **File name** box, type *YIELD94*. Click **Open** and **Yes** to create a new command file.
- In the **Open** dialog box, select *YIELD94.SYS* as the data file. Click **Open**. The dialog box for **Single Site Analysis** will appear.
- To specify variates to be analyzed, select *YIELD* from **Data File Variables** list. Click **Add** under the **Analysis Variates** list. (***Note:*** More than one variable can be specified for analysis but, for this example, a single variate will be used.)
- Variables added into the **Analysis Variates** list automatically appear in **Analysis Variates with ANOVA** list. So by default, CropStat does ANOVA for all the analysis variates. If you only want means presented for any variate, remove it from the list by clicking **Remove** button under the **Analysis Variates w/ ANOVA** list. In this example, we will ask for ANOVA for *YIELD*.
- To specify a variate containing site names, select SITE$ from **Data File Variables** list. Click **Add** under the **Site Names** edit box.
- To specify a variate containing extra treatment names, select VARIETY$ from **Data File Variables** list. Click **Add** under the **Treatment Names** edit box.


CropStat: Single Site Analysis for Balanced Data
Single Site Analysis
SSA Model
Options
Open
Command File :
YIELD94.PBT
Data File :
YIELD94.SYS
OK
Cancel
Help
Save
Data File Variables :
COUNTRY$
SITE$
YEAR
VARIETY$
VTYNO
REP
YIELD
Analysis Variates:
YIELD
Add
Remove
Analysis Variates w/ ANOVA
YIELD
Add
Remove
Site Names:
SITE$
Add
Remove
Treatment Names:
VARIETY$
Add
Remove
Save Means To:
Open
nVars : 7
nRecs : 480
Working directory :

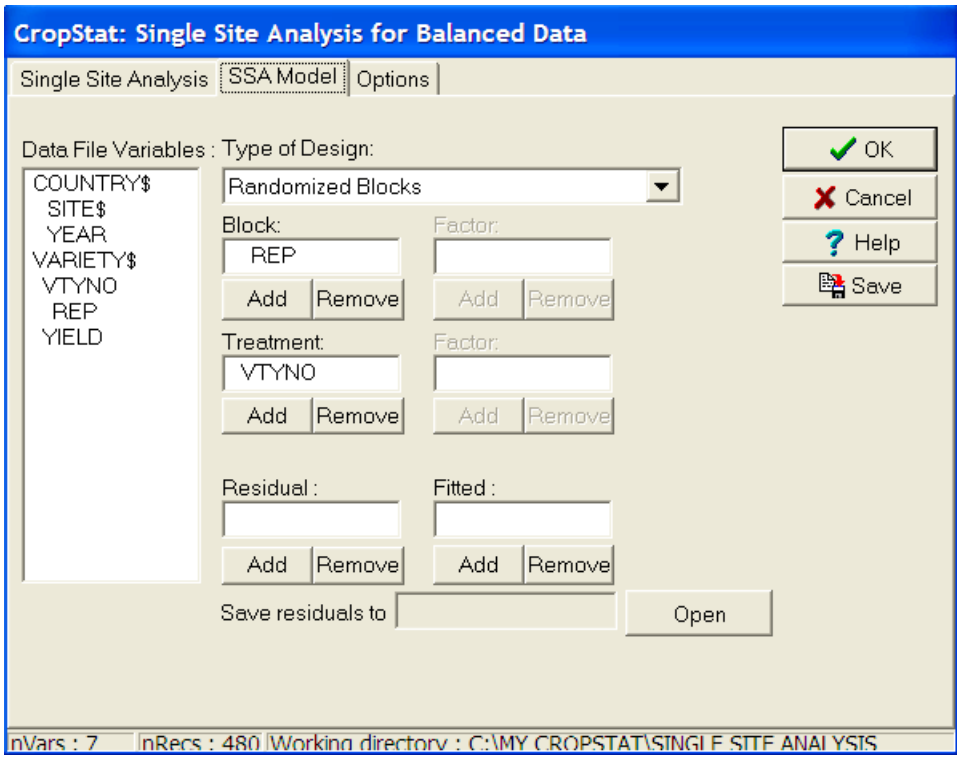
```
 MEANS FOR EACH VARIETY  - RANDOMIZED BLOCKS  FILE YIELD94   11/ 2/ 4  9:41
------------------------------------------------------------------- :PAGE   2
BH            - SECTION 1

       ENTRY                YIELD
1           AZU              1.344
2           BGORA           0.8480
3           GUAR             1.296
4           IT146           0.6880
5           OL5              1.524

6           OS6              1.388
7           UPL5             1.468
8           VAND             1.752
9           W181-18          1.576
10          W56-125          1.080

11          W56-50           1.568
12          W96-1-1          1.012
```

It is useful to include variate containing extra treatment names in addition to the treatment factor so that output is well annotated. These names are printed adjacent to the levels specified by the row factor in output tables.

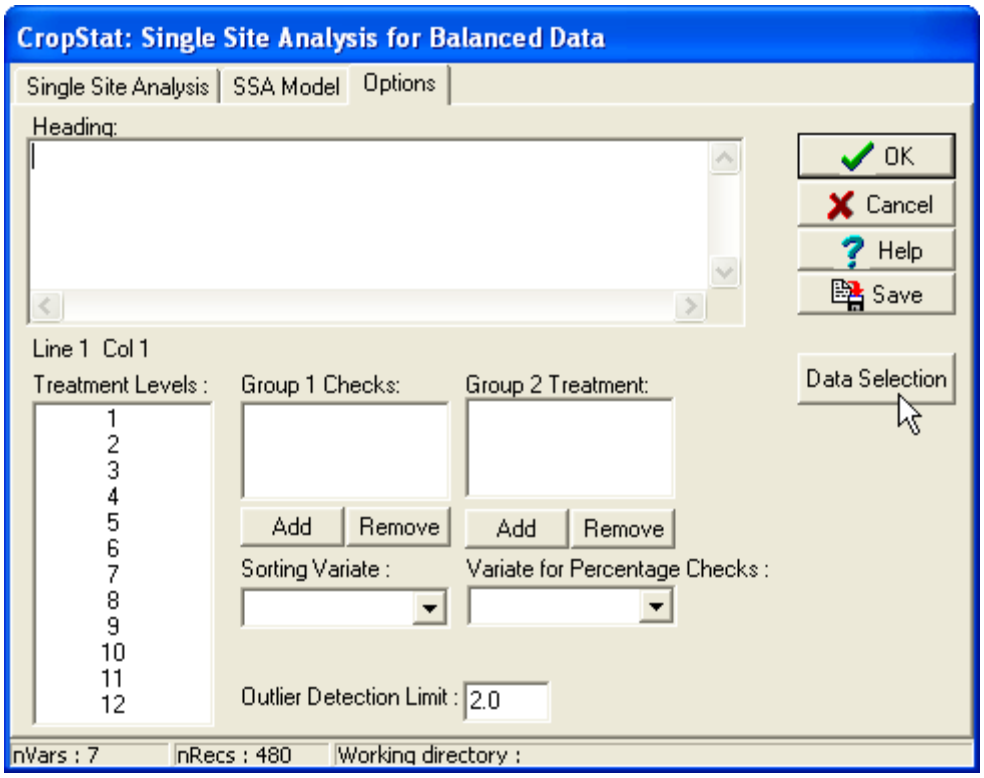
- To instruct CropStat to save the treatment means and error mean squares (EMS) to files, click the **Open** button next to the **Save Means to** box. Type *MYIELD94* in the **File name** box and click **Open**. The program will save the treatment means and EMS to *MYIELD94.SYS* and *MYIELD.EMS*, respectively.



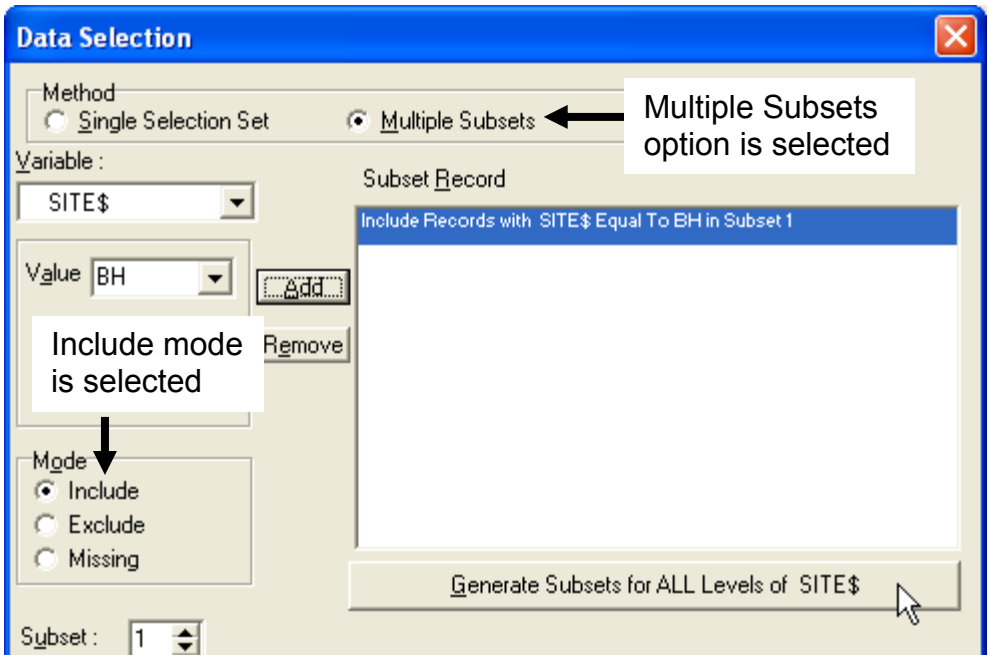
- To specify the model, click **SSA Model** tab.
- In the **SSA** page, click **Type of Design** box. Select *RANDOMIZED BLOCKS* from the drop-down list of designs.
- To specify block factor variate, Select *REP* from **Data File Variables** list. Click **Add** under the **Block** edit box.
- To specify the treatment variate, select *VTYNO* from **Data File Variables** list. Click **Add** under the **Treatment** edit box.



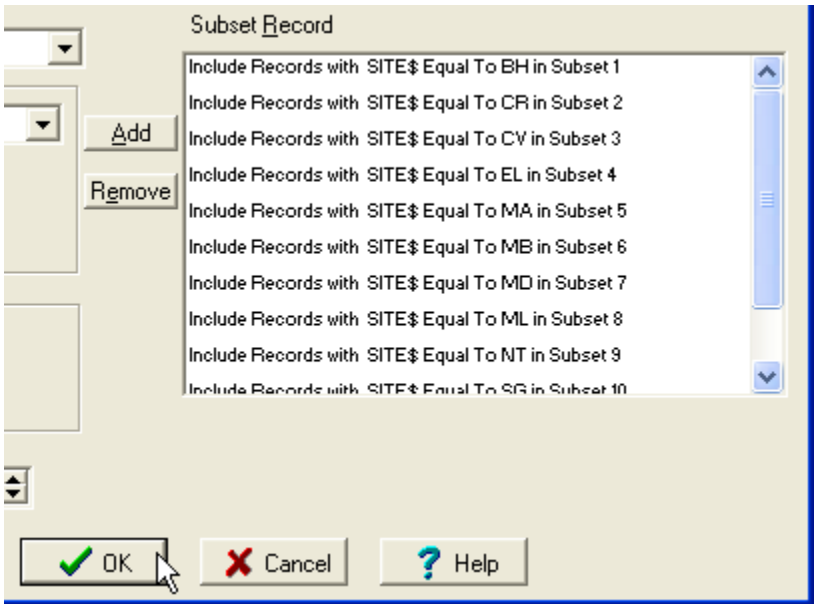
- To specify that this analysis is to be done repeatedly for the different sites, click the **Options** tab and click **Data Selection**.



- In the **Data Selection** page, click on **Multiple Subsets**, select *SITE$* on the **Variable** box; click **Add**.



- Click **Generate Subsets for ALL Levels of SITE$** button. This will request the analysis for each site.



- Click **OK** in the **Data Selection** page.
- Click **OK** in the **Options** page.

III. Sample Output for Sample Problem #1

The following is a partial listing of the output saved in *YIELD94.OUT*.

Results for site BH:

1. ANOVA table for single variates

```
ANOVA FOR SINGLE VARIATES - RANDOMIZED BLOCKS  FILE YIELD94   4/10/ 4 15:44
------------------------------------------------------------------------ :PAGE   1
DATA RECORDS SELECTED FROM FILE YIELD94
INCLUDE RECORDS WITH     SITE$   (  2) EQUAL TO BH
VARIATE V007 YIELD
-------------------------------------------------------------------------
SOURCE         D.F.        S.S.         M.S.          F       FPROB
-------------------------------------------------------------------------
BLOCKS            3     0.595051      0.198350
TREATMENT        11      4.60132      0.418302        5.86     0.000
ERROR M.S.       33      2.35612      0.713975E-01
TOTAL            47      7.55249
-------------------------------------------------------------------------
```

2. Residual analysis

```
VTYNO 4           , REP 1 HAS REDIDUAL -3.2 SEs:    -0.698
VTYNO 6           , REP 1 HAS REDIDUAL  2.9 SEs:     0.634
VTYNO 6           , REP 2 HAS REDIDUAL -2.3 SEs:    -0.506

BOX PLOT OF STUDENTIZED RESIDUALS FROM LPLT= -3.150    TO ULPT=  2.862
NO.<LPLT                                                      NO.>UPLT
       0 *        ----------------I    +      I-----------          *   0

MEDIAN= -0.5115E-01 ANDERSON-DARLING STATISTIC=  0.405
L.S.D. (5%)            0.38    COEFFICIENT OF VARIATION       20.63
```

3. Table of means

```
 MEANS FOR EACH VARIETY  - RANDOMIZED BLOCKS  FILE YIELD94   4/10/ 4 15:44
------------------------------------------------------------------- :PAGE   2
BH           - SECTION 1

       ENTRY               YIELD
1           AZU             1.344
2           BGORA          0.8480
3           GUAR            1.296
4           IT146          0.6880
5           OL5             1.524

6           OS6             1.388
7           UPL5            1.468
8           VAND            1.752
9           W181-18         1.576
10          W56-125         1.080

11          W56-50          1.568
12          W96-1-1         1.012

   MEANS                    1.295


OVERALL:
   MEANS                    1.295
   STD ERR                 0.1336
   5% LSD                  0.3844
   C.V.                      21.
   RES DF                    33.
```

Results for site CR:

1. ANOVA table for single variates

```
ANOVA FOR SINGLE VARIATES - RANDOMIZED BLOCKS  FILE YIELD94   4/10/ 4 15:44
------------------------------------------------------------------- :PAGE   3
DATA RECORDS SELECTED FROM FILE YIELD94
INCLUDE RECORDS WITH     SITE$   (  2) EQUAL TO CR
VARIATE V007 YIELD
-------------------------------------------------------------------------
SOURCE         D.F.        S.S.         M.S.          F       FPROB
-------------------------------------------------------------------------
BLOCKS            2     0.790917E-01  0.395459E-01
TREATMENT        11      20.3307       1.84824       11.03     0.000
ERROR M.S.       22      3.68566      0.167530
TOTAL            35      24.0954
-------------------------------------------------------------------------
```

2. Residual analysis

```
VTYNO 2           , REP 3 HAS REDIDUAL -2.8 SEs:    -0.902
VTYNO 9           , REP 3 HAS REDIDUAL  2.5 SEs:     0.794

BOX PLOT OF STUDENTIZED RESIDUALS FROM LPLT= -2.821    TO ULPT=  2.482
NO.<LPLT                                                      NO.>UPLT
       0 *            ----------I         +   I----------------------   0

MEDIAN=  0.1164E+00 ANDERSON-DARLING STATISTIC=  0.193
L.S.D. (5%)            0.69    COEFFICIENT OF VARIATION       19.50
```

3. Table of means

```
 MEANS FOR EACH VARIETY  - RANDOMIZED BLOCKS  FILE YIELD94   4/10/ 4 15:44
------------------------------------------------------------------------ :PAGE   4
CR           - SECTION 1

       ENTRY               YIELD
1           AZU             1.097
2           BGORA          0.8477
3           GUAR            2.999
4           IT146           2.079
5           OL5             1.422

6           OS6             2.995
7           UPL5            1.922
8           VAND            2.079
9           W181-18         2.281
10          W56-125         1.508

11          W56-50          2.687
12          W96-1-1         3.272

   MEANS                    2.099


OVERALL:
   MEANS                    2.099
   STD ERR                 0.2363
   5% LSD                  0.6931
   C.V.                      19.
   RES DF                    22.
```

IV. Alpha Designs

Peterson and Williams devised a new class of incomplete block designs called alpha designs. These designs are a generalization of Yates' original lattice designs. The main advantage of alpha designs is flexibility; they are available whenever the number of varieties ***v*** is a multiple of block size ***k***, and they can be easily adapted even when it is not.

Alpha designs fall under the category of resolvable designs. An incomplete block design is resolvable when the incomplete blocks can be arranged in complete replications. Non-resolvable designs exist but are used less frequently for field trials.

Within each replicate of an alpha design there are ***s*** blocks each of size ***k*** plots. There is thus a two-stage structure for the control of field variation. Replicates can allow an adjustment for large-scale variation and then within each replicate the blocks provide a second level of adjustment.

Efficiency Factor of an Alpha Design

Because treatments tested together in the same blocks are composed with high precision than those tested in different blocks, it is necessary to assign treatments to plots within blocks in such a way that comparisons have as similar a variance as possible.

For a particular set of ***r***, ***s***, and ***k***, many alpha designs (allocations of treatments to plots) are possible, so we must choose the most efficient alpha design. There are many ways of measuring the quality of incomplete block designs but the measure of most relevance where the interest is in comparing all pairs of varieties is the efficiency factor **E**. This is defined as the ratio of $2\sigma^2$ to V, where V is the mean variance of all differences between two varieties and σ^2 is the error variance in the within-block analysis of the incomplete block design. When choosing an alpha design, the aim is to choose one that has an efficiency factor **E** as large as possible. Designs with maximum **E** among all alpha designs are called alpha optimal.

The number of times a pair of varieties appears together within block of the design is called the concurrence of that pair of varieties. When the concurrence of pairs of varieties is high the pairwise variance is low, but if pairs of varieties do not appear together within blocks at all, the pairwise variance is high. The average pairwise variance will be minimized and **E** maximized if the concurrences of all pairs of varieties are as similar as possible. If the concurrences are either zero or one the design is known as an alpha (0,1)-design.

When the number of plots per block ***k*** is greater than the number of blocks per replicate ***s***, it is not possible to construct alpha (0,1)-designs, the next best thing is to restrict the number of concurrences to two, leading to alpha (0,1,2)-designs.

Availability of Alpha Designs

An advantage of alpha designs is that they are available for a wide range of combinations of parameters. There is no need to restrict resolvable incomplete block designs only to the square and rectangular lattice designs. Efficient designs can be constructed for all combinations of r, s, and k that would be required in practice.

For some number of varieties there may be a choice of block sizes, for example an alpha design for v=48 could have block sizes of 2, 4, 6, 8, 12, 16 or 24. As a general rule block size should be roughly equal to the square root of the number of varieties. So in the example above for v=48, we would consider designs with k=6 or 8. Often site considerations will dictate the block size; physical constraints on the layout of plots in the field might mean that blocks of six plots are much more natural than blocks of eight plots.

An apparent restriction on the availability of alpha designs is that the number of varieties is the product of the number of blocks per replicate and the block size, i.e. $v=sk$. So if we wanted a resolvable incomplete block design for 17 varieties it would appear difficult. A way around this problem is to allow designs with unequal block sizes. Provided the block sizes differ by no more than one, there is no need to review the model assumption that the pairwise variance between any two points within a block is the same. Alpha designs make it possible to produce efficient designs with block sizes differing by no more than one. This is done by simply deleting variety numbers from an alpha design for a larger number of varieties. So, in the previous example we could delete numbers 18, 19 and 20 to obtain a design for 17 varieties in blocks of size three or four. The largest variety numbers should be deleted before randomization to guarantee that block sizes differ by no more than one.

V. Sample Problem #2: Analysis of an Alpha Lattice Design

An experiment was conducted on 72 varieties in three replicates with nine blocks per replicate. Data is given below. Enclosed in parenthesis are the varieties.

		Plot							
Rep	Block	1	2	3	4	5	6	7	8
1	1	5.29(64)	2.56(47)	6.08(22)	5.30(41)	3.75(32)	1.51(13)	4.06(45)	3.93(42)
	2	2.69(57)	2.30(49)	3.05(29)	2.91(4)	3.38(70)	3.95(58)	2.39(48)	4.76(31)
	3	4.73(40)	3.32(6)	5.36(36)	5.54(67)	3.70(63)	3.45(46)	5.87(26)	2.14(66)
	4	5.08(72)	3.33(20)	1.70(7)	4.20(27)	4.10(24)	3.34(53)	3.66(33)	4.45(3)
	5	3.63(71)	4.76(37)	3.96(12)	3.58(65)	4.72(8)	5.72(25)	4.43(38)	4.69(15)
	6	1.57(16)	3.74(44)	3.48(14)	1.60(2)	2.61(17)	5.16(30)	5.73(69)	4.52(34)
	7	3.66(51)	3.16(18)	4.64(35)	6.31(21)	4.04(54)	5.22(11)	1.20(52)	4.20(10)
	8	4.24(62)	3.99(55)	3.99(60)	4.48(56)	4.63(19)	3.74(28)	4.37(61)	3.54(50)
	9	5.14(68)	5.36(39)	3.37(43)	2.70(1)	1.86(23)	2.78(5)	2.59(59)	4.79(9)
2	1	3.62(59)	2.98(7)	3.48(14)	2.50(13)	3.67(71)	2.69(16)	3.84(33)	4.75(9)
	2	4.32(35)	4.67(21)	4.49(68)	1.87(23)	4.99(30)	3.34(6)	4.85(40)	3.33(12)
	3	2.14(55)	4.51(36)	3.89(38)	3.24(50)	3.06(1)	3.14(65)	3.75(28)	3.68(64)
	4	5.28(67)	5.14(26)	4.58(39)	4.34(32)	4.87(22)	3.41(53)	2.75(20)	2.59(18)
	5	2.95(19)	3.79(58)	2.70(63)	3.10(45)	3.88(46)	3.17(48)	1.61(2)	4.19(3)
	6	3.94(27)	3.52(34)	3.74(62)	3.59(54)	3.64(60)	5.30(11)	3.90(51)	3.18(70)
	7	3.81(5)	3.30(15)	4.06(29)	2.57(66)	5.09(69)	4.36(41)	2.87(49)	4.25(42)
	8	1.20(52)	4.03(43)	4.08(37)	4.19(56)	4.36(61)	3.47(44)	2.69(17)	5.55(25)
	9	4.40(72)	3.68(24)	2.99(4)	2.60(47)	2.69(57)	3.71(31)	4.20(8)	2.08(10)
3	1	5.33(11)	4.43(19)	4.21(9)	3.79(5)	1.98(52)	4.83(21)	3.02(43)	1.62(66)
	2	5.34(30)	4.30(15)	5.06(67)	3.35(51)	2.02(23)	3.06(4)	4.50(68)	1.26(16)
	3	3.05(49)	5.31(41)	3.50(20)	2.99(55)	2.18(12)	2.87(50)	4.92(29)	4.69(3)
	4	3.40(24)	3.51(70)	4.76(40)	3.30(34)	2.63(65)	4.66(60)	3.57(28)	4.27(61)
	5	3.34(33)	2.35(7)	5.50(26)	2.70(63)	3.58(10)	2.87(18)	4.52(64)	4.47(38)
	6	3.96(71)	3.61(56)	4.20(27)	4.20(42)	3.19(32)	4.84(72)	5.72(22)	5.03(25)
	7	4.45(69)	4.20(54)	2.45(53)	3.63(35)	3.19(57)	2.72(31)	4.13(44)	0.79(2)
	8	3.94(14)	3.17(17)	4.18(46)	4.51(8)	3.59(59)	2.12(37)	2.58(47)	3.52(1)
	9	5.31(36)	0.88(13)	3.75(58)	3.88(45)	3.34(6)	3.73(62)	5.80(39)	1.89(48)

The data is in ALFAEX.SYS stored in the *CROPSTAT7.2\TUTORIAL\TUTORIAL DATASETS* folder.

- Open the data file ALFAEX.SYS from the *CROPSTAT7.2\TUTORIAL\TUTORIAL DATASETS* folder.

- Select **File →Save as.** Click the **Save in** box and go inside your working directory *C:\MY CROPSTAT\SINGLE SITE ANALYSIS* and click **Save**.

VI. Steps for performing Single Site Analysis for Sample Problem #2

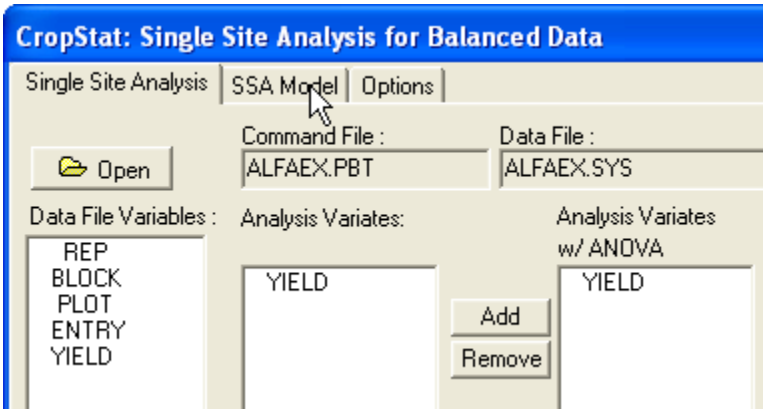
- Select **Single Site Analysis** from the **Analysis** menu of the Main CropStat Window.

- In the **Open** dialog box, click the **Look in** box and go inside your working folder MY *CROPSTAT\ SINGLE SITE*.

- Type *ALFAEX* in the **File name** edit box and click **Open**. Click **Yes** to create a new command file.

- Select *ALFAEX.SYS* data file and click **Open**. The Single Site Analysis for Balanced Data dialog appears.

- Select *YLD* from the **Data File Variables** list and click **Add** under the **Analysis Variates** box.

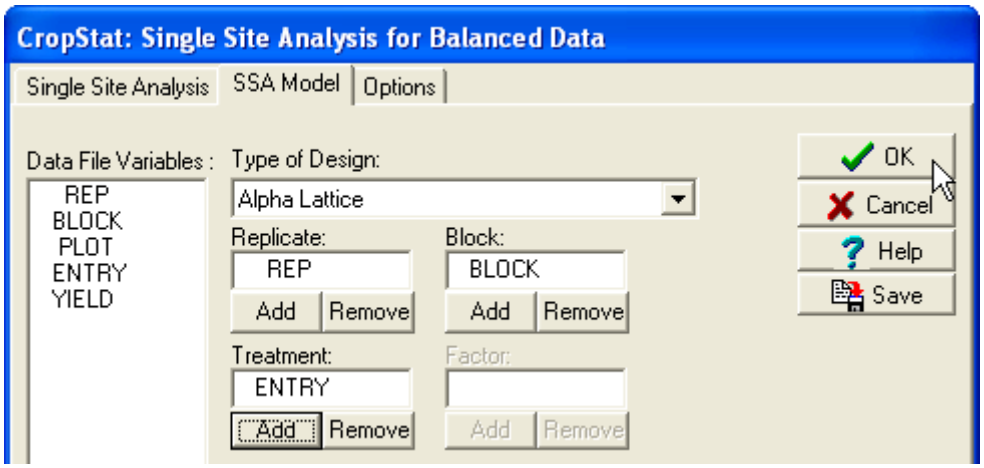


- Click the **SSA Model** tab to open the specification model window. In the **Type of Design** pull down list, select *Alpha Lattice*.

- Select *REP* from the **Data File Variables** list and click **Add** under the **Replicate** box.

- Highlight *BLOCK* from the **Data File Variables** and click **Add** under the **Block** box.

- Highlight *ENTRYNM1$* in the **Data File Variables** box and click **Add** under the **Treatment** box.



- Click **OK** to perform the analysis.

VII. Sample output for Sample Problem #2

a. Data description

```
---------------------------------------------------------------- :PAGE   1
THE IRREML PROGRAM WAS WRITTEN BY DOUGLAS CLARKSON OF SCIENCEOPS FOR IRRI

Command File: C:\MY CROPSTAT\SINGLE SITE ANALYSIS\ALFAEX.PBT

Data File: ALFAEX

Number of Records:   216 Non missing observations:   216

Number of Columns in the Fixed Effects Model:    72

Number of columns in the random effects model:    30

Variables in Data Set: REP  BLOCK  ENTRY  YIELD

Classification Variables: REP  BLOCK  ENTRY

Levels of the classification variables

    3 CODES:(Number  Label) for Variable: REP
(  1            1)(  2            2)(  3            3)(

    9 CODES:(Number  Label) for Variable: BLOCK
(  1            1)(  2            2)(  3            3)(  4            4)(  5
5)
(  6            6)(  7            7)(  8            8)(  9            9)(

   72 CODES:(Number  Label) for Variable: ENTRY
(  1            1)(  2            2)(  3            3)(  4            4)(  5
5)
(  6            6)(  7            7)(  8            8)(  9            9)( 10
10)
( 11           11)( 12           12)( 13           13)( 14           14)( 15
15)
( 16           16)( 17           17)( 18           18)( 19           19)( 20
20)
( 21           21)( 22           22)( 23           23)( 24           24)( 25
25)
( 26           26)( 27           27)( 28           28)( 29           29)( 30
30)
( 31           31)( 32           32)( 33           33)( 34           34)( 35
35)
( 36           36)( 37           37)( 38           38)( 39           39)( 40
40)
( 41           41)( 42           42)( 43           43)( 44           44)( 45
45)
( 46           46)( 47           47)( 48           48)( 49           49)( 50
50)
( 51           51)( 52           52)( 53           53)( 54           54)( 55
55)
( 56           56)( 57           57)( 58           58)( 59           59)( 60
60)
( 61           61)( 62           62)( 63           63)( 64           64)( 65
65)
( 66           66)( 67           67)( 68           68)( 69           69)( 70
70)
( 71           71)( 72           72)(
```

b. Model specification

```
 Model Specification

 Intercept in model: Yes

 The Fixed Effects Model
     YIELD = Intercept + ENTRY

 The Random Effects Terms
     REP + BLOCK(REP)
```

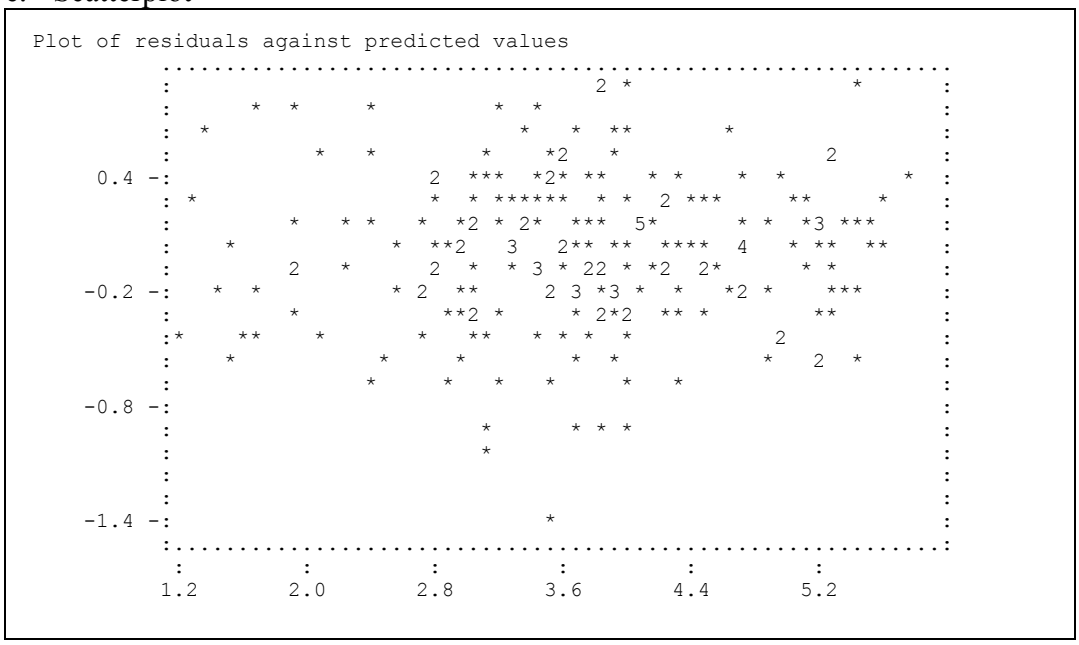
c. Variance/Covariance

```
 RANDOM EFFECT COVARIANCE MODEL.  2 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 REP                           1- 1  diagonal
 BLOCK(REP)                    2- 2  diagonal

 RESIDUAL EFFECT COVARIANCE MODEL.  0 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 RESIDUAL                              sigmasq(1)xI
 Message: Relative function convergence

 Final REML criterion:      -16.790117707187321

Variance/Covariance component parameters
 Dep Name              Gamma Coef. Std. Error    Z       Pr > |Z|    Var. Component  Std.
Error
  1 REP(1) .........  0.5132E-01  0.7604E-01  0.6749      0.4997      0.1259E-01  0.1861E-01
  1 BLOCK(REP)(1) ..  0.9074E-01  0.8994E-01   1.009      0.3130      0.2227E-01  0.2104E-01

 The scale parameters
 Dep.  Sigma_Squared Std. Error      Z         Pr > |Z|
Dep(1) .......    0.2454      0.3198E-01   7.673      0.1676E-13

 Asymptotic Covariance Matrix of the Variance/Covariance Components
                               1           2           3
    1   1 REP(1).........   0.578E-02  -0.534E-03  -0.182E-03
    2   1 BLOCK(REP)(1)..  -0.534E-03   0.809E-02  -0.119E-02
    3 Dep(1).............  -0.182E-03  -0.119E-02   0.102E-02
```

d. ANOVA table

```
       ANOVA Table for Sequentially Deleted Fixed Effects
       Denominator Degrees of Freedom: Residual DF
 Dep Effect    DFNum   DFDen   F - Statistic   P > |F|
  1    ENTRY     71 144.00        11.82       0.4469E-34
```

e. Scatterplot



f. Least Squares Means

Least Squares Means Fixed

Dep	Level		LSMean	Std. Error
1	ENTRY	1	3.155	0.3023
1	ENTRY	2	1.391	0.3024
1	ENTRY	3	4.444	0.3023
1	ENTRY	4	3.062	0.3024
1	ENTRY	5	3.495	0.3025
1	ENTRY	6	3.340	0.3024
1	ENTRY	7	2.273	0.3025
1	ENTRY	8	4.440	0.3024
1	ENTRY	9	4.553	0.3026
1	ENTRY	10	3.283	0.3023
1	ENTRY	11	5.269	0.3027
1	ENTRY	12	3.084	0.3022
1	ENTRY	13	1.552	0.3023
1	ENTRY	14	3.547	0.3026
1	ENTRY	15	4.046	0.3022
1	ENTRY	16	1.777	0.3024
1	ENTRY	17	2.782	0.3026
1	ENTRY	18	2.858	0.3024
1	ENTRY	19	3.985	0.3022
1	ENTRY	20	3.199	0.3024
1	ENTRY	21	5.244	0.3026
1	ENTRY	22	5.569	0.3027
1	ENTRY	23	1.968	0.3027
1	ENTRY	24	3.775	0.3023
1	ENTRY	25	5.360	0.3025
1	ENTRY	26	5.513	0.3026
1	ENTRY	27	4.141	0.3023
1	ENTRY	28	3.699	0.3028
1	ENTRY	29	4.019	0.3027
1	ENTRY	30	5.173	0.3026
1	ENTRY	31	3.842	0.3027
1	ENTRY	32	3.772	0.3027
1	ENTRY	33	3.543	0.3025
1	ENTRY	34	3.802	0.3024
1	ENTRY	35	4.215	0.3024
1	ENTRY	36	5.100	0.3023
1	ENTRY	37	3.559	0.3026
1	ENTRY	38	4.238	0.3024

```
 1  ENTRY           39    5.314      0.3022
 1  ENTRY           40    4.789      0.3023
 1  ENTRY           41    4.950      0.3026
 1  ENTRY           42    4.114      0.3025
 1  ENTRY           43    3.488      0.3026
 1  ENTRY           44    3.798      0.3025
 1  ENTRY           45    3.687      0.3026
 1  ENTRY           46    3.834      0.3023
 1  ENTRY           47    2.583      0.3023
 1  ENTRY           48    2.539      0.3027
 1  ENTRY           49    2.749      0.3027
 1  ENTRY           50    3.199      0.3027
 1  ENTRY           51    3.631      0.3024
 1  ENTRY           52    1.406      0.3026
 1  ENTRY           53    3.138      0.3023
 1  ENTRY           54    3.974      0.3026
 1  ENTRY           55    3.023      0.3027
 1  ENTRY           56    4.042      0.3024
 1  ENTRY           57    2.968      0.3027
 1  ENTRY           58    3.886      0.3027
 1  ENTRY           59    3.222      0.3026
 1  ENTRY           60    4.088      0.3028
 1  ENTRY           61    4.285      0.3026
 1  ENTRY           62    3.892      0.3023
 1  ENTRY           63    3.038      0.3024
 1  ENTRY           64    4.511      0.3023
 1  ENTRY           65    3.107      0.3024
 1  ENTRY           66    2.102      0.3023
 1  ENTRY           67    5.320      0.3023
 1  ENTRY           68    4.761      0.3027
 1  ENTRY           69    5.128      0.3024
 1  ENTRY           70    3.415      0.3024
 1  ENTRY           71    3.635      0.3023
 1  ENTRY           72    4.819      0.3024

Standard Errors of Differences
 Minimum     Mean      Maximum
 0.4045      0.4160      0.4180


                  Least Squares Means Fixed
Dep Level               LSMean   Std. Error
 1  BLOCK            1    3.516      0.9795E-01
 1  BLOCK            2    3.700      0.1021
 1  BLOCK            3    3.822      0.1002
 1  BLOCK            4    3.957      0.9984E-01
 1  BLOCK            5    3.647      0.9768E-01
 1  BLOCK            6    3.961      0.9798E-01
 1  BLOCK            7    3.706      0.1002
 1  BLOCK            8    3.586      0.1021
 1  BLOCK            9    3.668      0.9588E-01

Standard Errors of Differences
 Minimum     Mean      Maximum
 0.7751E-01  0.9111E-01  0.1074
```

Exercise 7

Regular controls in Latin Square Design

The file *PBAUGD.SYS* in the *CROPSTAT\TUTORIAL\2015 CROPSTAT_SINGLE SITE ANALYSIS* folder contains simulated data from a Regular controls in Latin Square design. There are 95 test entries numbered 1 to 95 and five replicated entries numbered 96 to 100. The 120 plots in the design are laid out in 10 rows of 12 plots each.

There are four simulated response variates, *Y1*, *Y2*, *Y3* and *Y4*. The value of *Y1* is computed as twice the entry number plus the row and column numbers plus a central normal random number with variance 16. *Y1* therefore has a consistent trend over rows and columns.

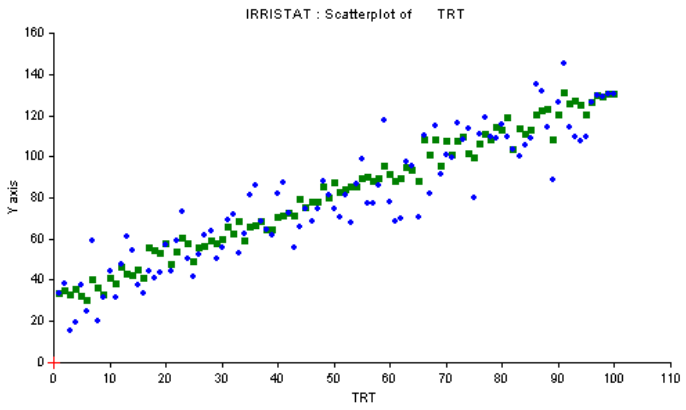
Y2 is computed as the entry number plus row number plus a normal deviate with variance 16. *Y2* has spatial trend across rows only. *Y3* and *Y4* are constructed from *Y1* and *Y2* with missing values in check and test entries.

Analyze *Y1* and *Y3*. Include *Y1* a second time in the analysis variates but exclude it from the Analysis with ANOVA box for the second instance. Include Row and Col as analysis variates but exclude them from the Analysis with ANOVA box. (Row and Col are also blocking variates on the Model tab). Specify a means save file.

In the analysis output you will notice two large residuals for replicated treatments 98 and 99 although these are not so large as to be definite outliers. The Row and Col effects are consistent with the model for *Y1*. The column *Y1* unadjusted for the second instance of *Y1* should contain the raw data for the test entries.

To see the effect of adjustment, open the saved means file in the Data Editor. Rename the second *Y1* variable to *Y1RAW* and save the file.

Select **Analysis|Scatter Plots** from the **Main Window**. Open the means data file you just saved. Plot *Y1* and *Y1RAW* against *TRT*. The adjusted values follow the true means more closely than the raw values.



SPATIAL ANALYSIS

At the end of the tutorial, the user should be able to

- understand what spatial variation is and how to control it
- perform spatial analysis using mixed models

I. Introduction

Spatial variability can be partly controlled by using an appropriate experimental design. Most variety trials use complete or incomplete block designs and are analyzed using the traditional analysis of variance. Block designs attempt a reduction of the experimental error by accounting for spatial heterogeneity among blocks. This approach does not consider the presence of spatial variability within blocks, and researchers face the problem of having to find blocks in the field that are homogeneous without knowing their most appropriate shape, dimension and orientation. When field variety trials are laid out in a rectangular array of *r* rows and *c* columns with replicates allocated contiguously, then spatial analysis can be performed with the aim of improving precision of estimates of variety effects and variety contrasts.

One useful measure for accounting for the heterogeneity patterns of the soil is to estimate the spatial autocorrelation of neighboring plots within rows or within columns. That is from the correlation between residuals at various distances apart. If there is no spatial pattern, all the correlations will be low. If there is pattern in the residuals, neighboring residuals will be more similar and so have higher correlation. Two-dimensional spatial analysis using an autoregressive model in the direction of the rows and columns has been found to be appropriate in many situations

Gilmour et al (1997) distinguished between global, natural and extraneous variation. For natural variation arising from unevenness of soil moisture, soil depth or other natural variation, they proposed using a separable autoregressive (AR) correlation structure, without differencing. Thus, they model the natural variation as the direct product of an AR correlation structure for columns and an AR correlation structure for rows, denoted by AR1xAR1. Extraneous variation includes effects introduced by the experimental operations. These operations are usually aligned with rows or columns and are usually modeled with random row and column effects. Global effects include any major (non-stationary) trends across the field. These can be fitted as linear trends, row and column contrasts and covariates.

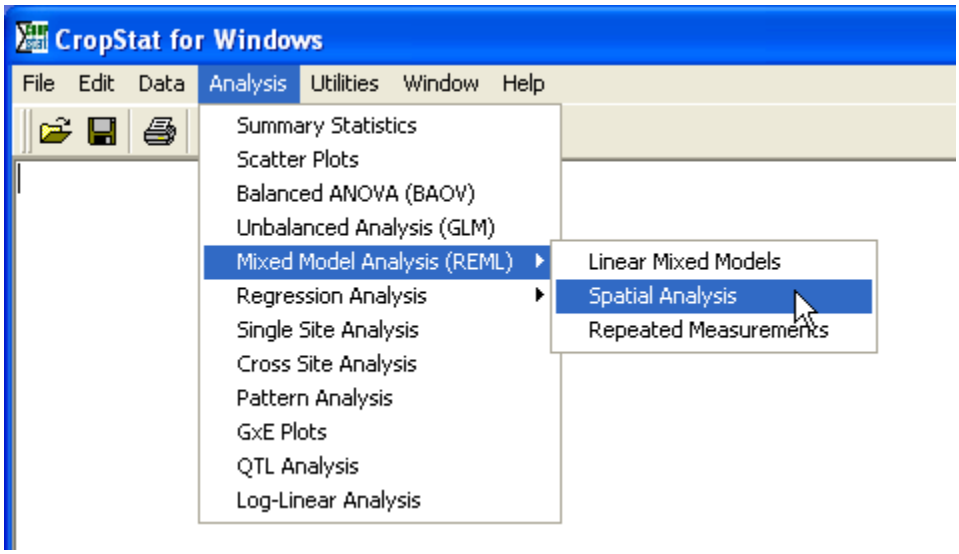
II. Analysis Using CropStat

To illustrate how to perform spatial analysis in CropStat, we use the sample data set we used in the analysis of the alpha design. Row and column variables have been added to the file called *AR1XAR1.SYS*. Note that for spatial analysis the replicates should be contiguous. Column or row numbers run through replicates.

- Open the data file *AR1XAR1.SYS* from the *CROPSTAT7.2\TUTORIAL\TUTORIAL DATASETS* folder.

- Select **File →Save as.** Click the **Save in** box and go inside your working folder *C:\MY CROPSTAT* and create a subfolder SPATIAL ANALYSIS then click **Save.**

- Choose **Mixed Model Analysis|Spatial Analysis** from the Analysis menu.



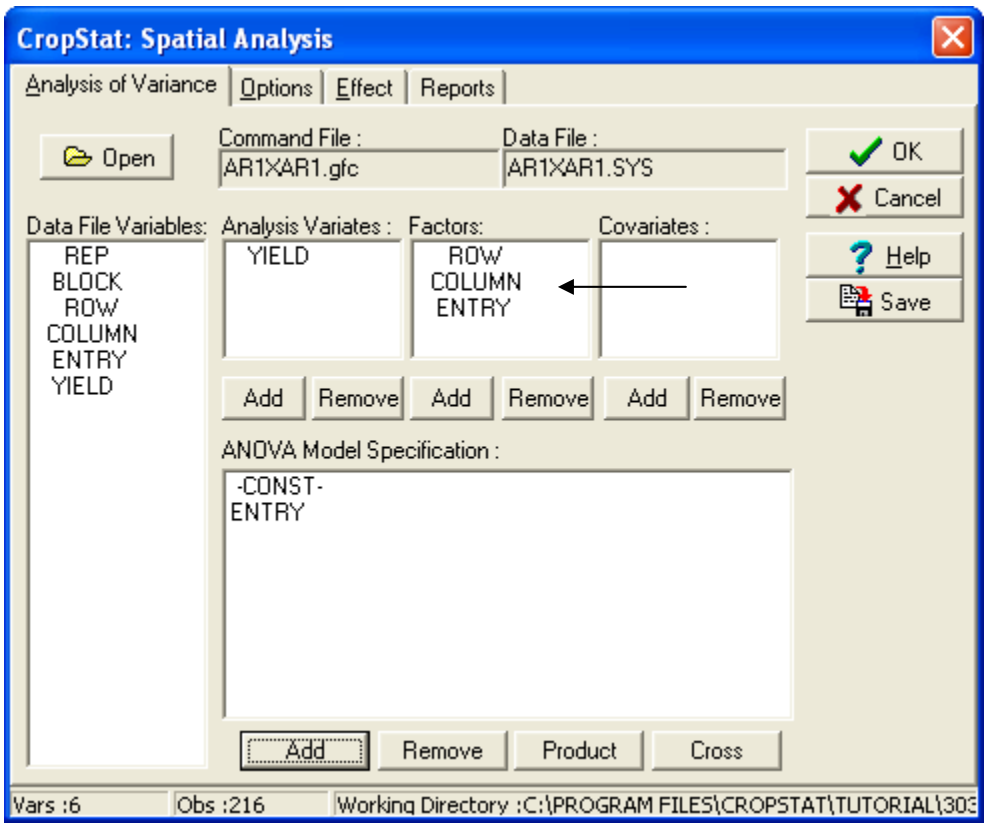
- The **Open** dialog box will prompt you to enter a name for the command file. Click the **Look In** box to go to your working drive *C:\MY CROPSTAT\SPATIAL ANALYSIS* folder.

- Enter AR1XAR1 in the **File name** box. Click **Open** button.

- Since *AR1XAR1.GFC* does not exist, a message box will appear confirming if you want to create the file. Click **Yes** to create new Command File.

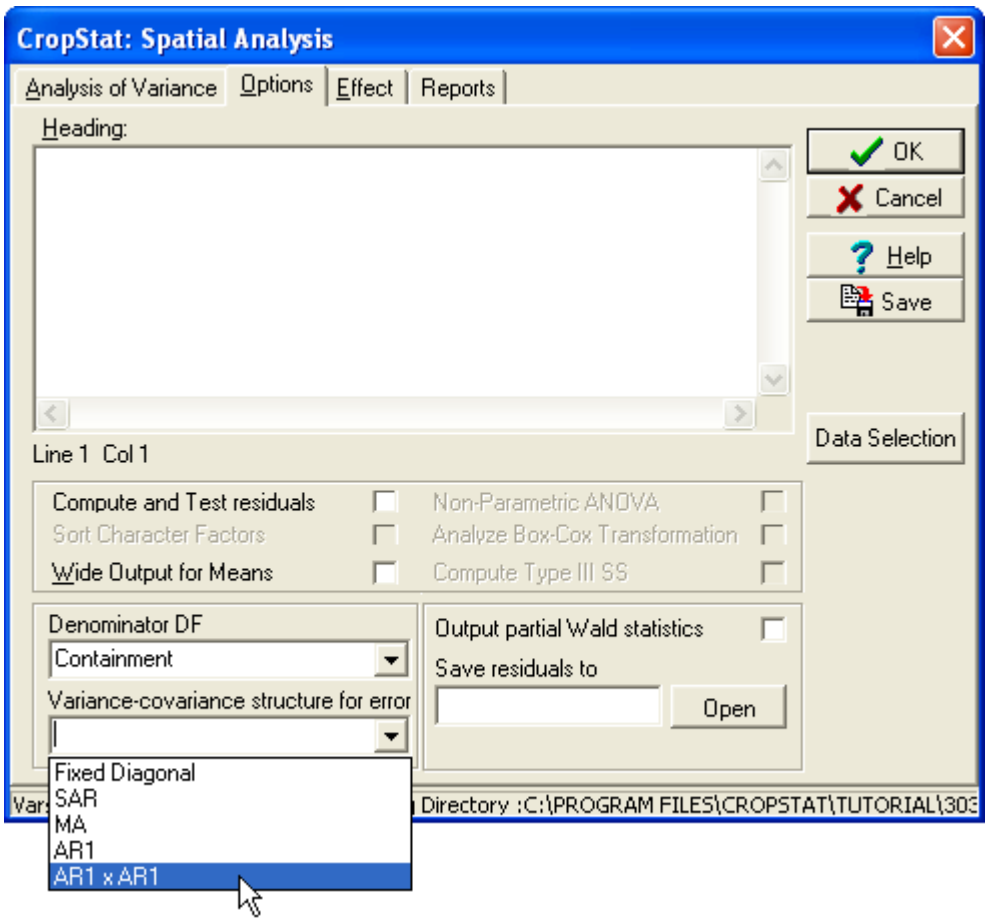
- Enter the name of the data file to be used. Enter *AR1XAR1.SYS* in the **File name** box.

- Click **Open**. The **Spatial Analysis** dialog box will appear for you to fill-in the details of the analysis.

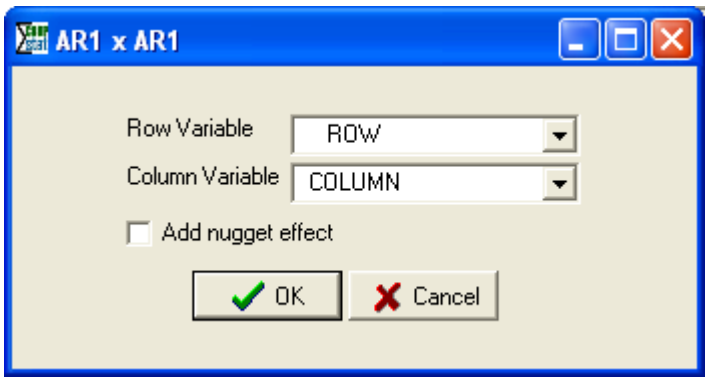


- From the **Data File Variable** list, highlight all variables to be analyzed then add to the **Analysis Variates** box; and highlight the *treatment, row, and column* variables then add to the **Factors** box.
- From the **Factors** box, highlight the treatment variable *ENTRY,* then add to the **ANOVA Model Specification** box. As for any ANOVA model see the Analysis of Variance module for a detailed step by step instruction on how to specify the ANOVA model.

- Click the **Options** tab. Specify the **Variance-Covariance Structure for Error**. Click the down-arrow key then choose AR1 X AR1.



- The AR1xAR1 pop-up window will appear.



- Click the **OK** button to return to the **Options** tab. Click the **OK** button to run the analysis.

III. Sample Output

```
IRREML 1.0.2: REML ANALYSIS end  FILE AR1XAR1   12/ 5/ 5 13: 2
 ------------------------------------------------------------------ :PAGE   1
 THE IRREML PROGRAM WAS WRITTEN BY DOUGLAS CLARKSON OF SCIENCEOPS FOR IRRI

 Command File: C:\MY CROPSTAT\SPATIAL ANALYSIS\AR1XAR1.gfc   Data File: AR1XAR1

 Number of Records:   216

 Variables in Data Set: ROW  COLUMN  ENTRY  YIELD

 SUMMARY STATISTICS FOR NUMERIC VARIATES
 VARIATE      NOBS   MINIMUM     MAXIMUM     MEAN     STD. DEV.
 YIELD         216  0.7900       6.310       3.729       1.085

 Classification Variables: ROW  COLUMN  ENTRY

 Levels of the classification variables

   27 CODES:(Number  Label) for Variable: ROW
(  1            1)(  2            2)(  3            3)(  4            4)(  5            5)
(  6            6)(  7            7)(  8            8)(  9            9)( 10           10)
( 11           11)( 12           12)( 13           13)( 14           14)( 15           15)
( 16           16)( 17           17)( 18           18)( 19           19)( 20           20)
( 21           21)( 22           22)( 23           23)( 24           24)( 25           25)
( 26           26)( 27           27)(

    8 CODES:(Number  Label) for Variable: COLUMN
(  1            1)(  2            2)(  3            3)(  4            4)(  5            5)
(  6            6)(  7            7)(  8            8)(

   72 CODES:(Number  Label) for Variable: ENTRY
(  1            1)(  2            2)(  3            3)(  4            4)(  5            5)
(  6            6)(  7            7)(  8            8)(  9            9)( 10           10)
( 11           11)( 12           12)( 13           13)( 14           14)( 15           15)
( 16           16)( 17           17)( 18           18)( 19           19)( 20           20)
( 21           21)( 22           22)( 23           23)( 24           24)( 25           25)
( 26           26)( 27           27)( 28           28)( 29           29)( 30           30)
( 31           31)( 32           32)( 33           33)( 34           34)( 35           35)
( 36           36)( 37           37)( 38           38)( 39           39)( 40           40)
( 41           41)( 42           42)( 43           43)( 44           44)( 45           45)
( 46           46)( 47           47)( 48           48)( 49           49)( 50           50)
( 51           51)( 52           52)( 53           53)( 54           54)( 55           55)
( 56           56)( 57           57)( 58           58)( 59           59)( 60           60)
( 61           61)( 62           62)( 63           63)( 64           64)( 65           65)
( 66           66)( 67           67)( 68           68)( 69           69)( 70           70)
( 71           71)( 72           72)(
```

```
IRREML: REML ANALYSIS FOR VARIATE YIELD  FILE AR1XAR1   12/ 5/ 5 13: 2
 ------------------------------------------------------------------ :PAGE   2
 Number of non-missing dependent observations:          216

 Model Specification
 Intercept in model: Yes
 The Fixed Effects Model
     YIELD = Intercept + ENTRY
 The Random Effects Terms
 None

 RANDOM EFFECT COVARIANCE MODEL.  0 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 None

 RESIDUAL EFFECT COVARIANCE MODEL.  1 SPECIFIED STRUCTURES
 TERM              PARAMETER INDICES  STRUCTURE
 ------------------------------------------------------------
 RESIDUAL                     1- 2  rAR1(ROW)xrAR1(COLUMN)

 Number of columns in the fixed effects model:     72
 Number of columns in the random effects model:     0

 Message: Relative function convergence

 Final REML criterion:      -19.051227980282228

 Variance/Covariance component parameters
 Dep Name           Gamma Coef. Std. Error     Z        Pr > |Z|    Var. Component  Std. Error
  1 AR1(ROW)(1)     0.6152E-01  0.8553E-01   0.6440      0.5196
  1 AR1(COLUMN)(2   0.1180      0.1044       1.130       0.2584
 The scale parameters
 Dep.  Sigma_Squared Std. Error       Z          Pr > |Z|
Dep(1) .......     0.2798      0.3330E-01   8.403      0.4338E-16

 Asymptotic Covariance Matrix of the Variance/Covariance Components
                                 1           2           3
        1   1 AR1(ROW)(1).    0.913E-02  -0.777E-03   0.192E-03
        2   1 AR1(COLUMN)(  -0.777E-03   0.109E-01   0.417E-03
        3   Dep(1).........   0.192E-03   0.417E-03   0.111E-02

 Warning: Denominator degrees of freedom estimates do not account
 for measurement error parameters.

      ANOVA Table for Sequentially Deleted Fixed Effects
      Denominator Degrees of Freedom: Residual DF
Dep Effect    DFNum   DFDen   F - Statistic   P > |F|
 1    ENTRY     71 144.00        11.07       0.1736E-32
```

Correlation coefficients along the rows and columns, respectively.

Test for fixed effect.

Balanced Least Squares Means Fixed

Dep	Level		LSMean	Std. Error
1	ENTRY	1	3.110	0.3023
1	ENTRY	2	1.286	0.3024
1	ENTRY	3	4.401	0.3033
1	ENTRY	4	3.036	0.3019
1	ENTRY	5	3.507	0.3025
1	ENTRY	6	3.305	0.3024
1	ENTRY	7	2.335	0.3019
1	ENTRY	8	4.485	0.3020
1	ENTRY	9	4.556	0.3090
1	ENTRY	10	3.329	0.3028
1	ENTRY	11	5.281	0.3024
1	ENTRY	12	3.140	0.3025
1	ENTRY	13	1.626	0.3023
1	ENTRY	14	3.546	0.3023
1	ENTRY	15	4.042	0.3024
1	ENTRY	16	1.868	0.3030
1	ENTRY	17	2.796	0.3022
1	ENTRY	18	2.851	0.3026
1	ENTRY	19	4.002	0.3023
1	ENTRY	20	3.169	0.3019
1	ENTRY	21	5.235	0.3023
1	ENTRY	22	5.532	0.3020
1	ENTRY	23	1.965	0.3018
1	ENTRY	24	3.694	0.3023
1	ENTRY	25	5.441	0.3030
1	ENTRY	26	5.567	0.3019
1	ENTRY	27	4.161	0.3023
1	ENTRY	28	3.673	0.3026
1	ENTRY	29	4.034	0.3018
1	ENTRY	30	5.133	0.3024
1	ENTRY	31	3.774	0.3030
1	ENTRY	32	3.791	0.3020
1	ENTRY	33	3.566	0.3023
1	ENTRY	34	3.776	0.3024
1	ENTRY	35	4.196	0.3027
1	ENTRY	36	5.173	0.3024
1	ENTRY	37	3.609	0.3018
1	ENTRY	38	4.263	0.3025
1	ENTRY	39	5.236	0.3020
1	ENTRY	40	4.769	0.3028
1	ENTRY	41	4.940	0.3024
1	ENTRY	42	4.133	0.3029
1	ENTRY	43	3.541	0.3019
1	ENTRY	44	3.835	0.3021
1	ENTRY	45	3.704	0.3022
1	ENTRY	46	3.811	0.3022
1	ENTRY	47	2.557	0.3020
1	ENTRY	48	2.376	0.3025
1	ENTRY	49	2.791	0.3028
1	ENTRY	50	3.181	0.3023
1	ENTRY	51	3.626	0.3023
1	ENTRY	52	1.407	0.3026
1	ENTRY	53	3.082	0.3020
1	ENTRY	54	3.974	0.3020
1	ENTRY	55	3.087	0.3023
1	ENTRY	56	4.009	0.3019
1	ENTRY	57	2.915	0.3027
1	ENTRY	58	3.923	0.3024
1	ENTRY	59	3.292	0.3023
1	ENTRY	60	4.080	0.3019
1	ENTRY	61	4.343	0.3028
1	ENTRY	62	3.897	0.3024
1	ENTRY	63	3.073	0.3019
1	ENTRY	64	4.486	0.3034
1	ENTRY	65	3.065	0.3021
1	ENTRY	66	2.138	0.3028
1	ENTRY	67	5.319	0.3023
1	ENTRY	68	4.743	0.3023
1	ENTRY	69	5.058	0.3023
1	ENTRY	70	3.346	0.3023
1	ENTRY	71	3.678	0.3030
1	ENTRY	72	4.817	0.3028

Standard Errors of Differences

Minimum	Mean	Maximum
0.4107	0.4265	0.4329

Exercise 8

A variety trial was conducted to evaluate the performance of 44 rice varieties. The design used was 11x4 alpha lattice. The layout is in Figure 1. Data are in DUALDATA.XLS.

a. Perform an analysis variance for an alpha design.
b. Perform an AR1xAR1 analysis adjusting for spatial variation.

Rep 1	Blk 1	21	10	13	25
	Blk 2	27	44	16	7
	Blk 3	18	36	22	8
	Blk 4	42	29	30	32
	Blk 5	14	6	43	35
	Blk 6	39	19	23	31
	Blk 7	4	24	5	33
	Blk 8	1	11	28	15
	Blk 9	26	20	2	40
	Blk 10	37	38	17	41
	Blk 11	12	3	34	9
Rep 2	Blk 1	34	23	27	30
	Blk 2	44	25	17	39
	Blk 3	7	10	8	26
	Blk 4	6	38	31	22
	Blk 5	21	14	18	33
	Blk 6	12	36	19	2
	Blk 7	11	40	24	3
	Blk 8	20	15	32	16
	Blk 9	5	42	9	35
	Blk 10	29	1	43	37
	Blk 11	13	41	28	4
Rep 3	Blk 1	36	32	24	25
	Blk 2	16	19	37	5
	Blk 3	26	33	39	1
	Blk 4	41	9	20	18
	Blk 5	38	34	10	15
	Blk 6	4	6	12	7
	Blk 7	3	21	29	31
	Blk 8	43	13	40	27
	Blk 9	35	23	11	8
	Blk 10	17	2	14	30
	Blk 11	42	28	22	44

Figure 1. Layout of 11x4 alpha lattice.