TEACHING MANUAL

Management of Plant Genetic Resources

Compiled and Edited by

Sherry R Jacob
Neeta Singh
Kalyani Srinivasan
Veena Gupta
J Radhamani
Anjali Kak
Chitra Pandey
Sushil Pandey
J Aravind
IS Bisht
RK Tyagi

भाकृअनुप
ICAR

रा.पा.आ.सं.ब्यूरो
NBPGR

ICAR - NATIONAL BUREAU OF PLANT GENETIC RESOURCES
New Delhi

Teaching Manual
"Management of Plant Genetic Resources"

Compiled and edited by

Sherry R Jacob
Neeta Singh
Kalyani Srinivasan
Veena Gupta
J Radhamani
Anjali Kak
Chitra Pandey
Sushil Pandey
J Aravind
IS Bisht
RK Tyagi

ICAR-National Bureau of Plant Genetic Resources
New Delhi

Citation: Jacob Sherry R, N Singh, K Srinivasan, V Gupta, J Radhamani, A Kak, C Pandey, S Pandey, J Aravind, IS Bisht and RK Tyagi (2015) Management of Plant Genetic Resources, National Bureau of Plant Genetic Resources, New Delhi, 323 p.

Published in 2015

ICAR-National Bureau of Plant Genetic Resources
New Delhi 110012, India

Copies can be obtained from:
The Director
ICAR-National Bureau of Plant Genetic Resources
New Delhi 110012, India
E-mail: director@nbpgr.ernet.in
Website: http//nbpgr.ernet.in

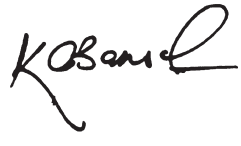
Contents

Foreword

The National Bureau of Plant Genetic Resources (NBPGR) has been serving the nation since 1976, as the nodal organization for management of plant genetic resources for food and agriculture (PGRFA). It is primarily a service oriented organization and caters to the requirements of the national crop improvement programmes by facilitating germplasm management.

The institute has a very strong cadre of scientists having best expertise in various components of PGR management. In order to harness this expertise for developing effective human resources for the future of PGRFA, the Post Graduate School of Indian Agricultural Research Institute (IARI), New Delhi, incepted the PGR discipline in 1996, as a part of its post-graduate programme. This was followed by initiation of the Ph.D. programme in 2004. Till date, 46 M.Sc. and 13 Ph.D. students have graduated in the PGR discipline and are successfully pursuing their careers in respective fields.

The teaching programme has been strengthened over the years by the addition of new faculty members and introduction of topics on recent advances in each course. This manual on **Management of Plant Genetic Resources** is yet another effort in this direction, whereby a comprehensive presentation of the major relevant aspects of PGR management has been attempted. The manual would serve as a ready-reckoner for all students of PGR and I hope students would make effective use of this compilation for upgrading their knowledge.



(KC Bansal)

Acknowledgements

The team of editors takes this opportunity to express their sincere gratitude to Dr KC Bansal, Director, ICAR-National Bureau of Plant Genetic Resources (NBPGR), New Delhi, for encouraging and extending his support for bringing out this teaching manual.

We also place on record, our gratitude to the Joint Director (Education) and Dean, Post Graduate School, Indian Agricultural Research Institute (IARI), New Delhi, for his support.

The core information used in this compilation is a reproduction of the lectures delivered by respective authors during the training entitled "Management of Plant Genetic Resources" held at NBPGR on 16-25 September 2013. We thank each and every subject matter expert for their meticulous contribution.

Financial support provided by the Division of Agricultural Education, Indian Council of Agricultural Research and Post Graduate School, IARI, New Delhi, under "Strengthening of PG Education Programme (Deemed University)", is gratefully acknowledged.

Editors

Designing of Experiments and Analysis of Data in Plant Genetic Resource Management

Rajender Parsad
Indian Agricultural Statistics Research Institute, Library Avenue, Pusa, New Delhi

In genetic resources environment, which is a field in the forefront of biological research, an essential activity is to test or evaluate the new germplasm/provenances/superior selections (new treatments, henceforth called test treatments) with the existing provenances or released varieties (checks, henceforth called control treatments). A problem in these evaluation studies is that the quantity of the genetic material collected from the exploration trips is very limited or cannot be made available since a part of this is to be deposited in genebank. The available quantity of seed is often not sufficient for replicated trials. Moreover, the number of new germplasm or provenances to be tested is very high (usually about 1000-2000 and sometimes upto 3000 accessions). A problem of interest is to design such experiments for making comparisons among the test treatments, among the control treatments and test treatments vs control treatments. The basic interest in these trials is to identify the promising germplasms. The promising germplasms identified are then subjected to more rigorous experimentation by allowing replications (generally two to three) of the test treatments (the promising germplasms or entries) along with the controls. In order to test the adaptability of these germplasms or entries in different environmental conditions, the trials are generally conducted over different environments (locations or years) to identify the promising lines. The purpose of the present note is to make an attempt to highlight some considerations in designing of experiments and analysis of experimental data. No claim is made about it being exhaustive.

Designing the experiments with unreplicated tests

We begin by describing the experimentation as it was carried out in the initial stages. The check plots are extensively used for controlling spatial heterogeneity. Here, the new selections are planted in long and narrow plots and checks are planted at regular intervals. Indices in terms of functions of check yields are developed for comparing the yields of the new selections. The four commonly used indices are (i) the yield of the nearest check, (ii) the mean yields of the two nearest checks, one on either side of the test plot, (iii) the weighted mean of these two checks, where weights are inversely related to the distance from the test plot, and (iv) the mean of all the checks in the range. For more details on these indices one may refer to Kempton and Fox (1997). All comparisons using these indices are subjective in nature. Therefore, there is a need to develop objective methods of making comparisons of the checks and the new selections. One possible way of performing objective comparisons is based on a measure of experimental error computed from the variation among the observations of the check yields. A minimum significant difference is then obtained

to compare the yield of the new selection with the mean of the check under the assumption that the variability of the new selections is similar to that of the checks. If the yield of a new selection is more than the sum of the mean of the check yields and the minimum significant difference, then this new selection is declared as significantly different from the check mean.

Augmented designs

A major drawback of the approach just described is that the checks are systematically placed and the variation in the check yields is not a true representative of the total field variability. To take care of such problems, one should conduct the experiments using Augmented (Hoonuiaku) Designs given by Federer (1956). These designs were introduced to fill a need arising in screening new strains of sugarcane at Experimental Station of Hawaiin Sugarcane Planters Association, Hawaii on the basis of agronomic characters other than yield.

For designing the experiments for preliminary screening, the experimenter cannot afford to replicate the test treatments. The control treatments, however, can be replicated. Since the designs will have a single replication of the test treatments, it is feared that one may not be able to get an estimate of the experimental error. This is not true and the experimental error is provided by the replications of the control treatments. It is also feared that with single replication of the test treatments, one may not be able to make all the possible paired comparisons among the test treatments, the control treatments and between the test treatments and control treatments. A design that allows making all the possible paired comparisons is called a connected design. In terms of this concept, the design with single replication of the test treatments may seemingly be a disconnected design. To overcome this problem, the control treatments are replicated in all the blocks. This ensures that the design in control treatments is a randomized complete block (RCB) design. If it is not possible to replicate each of the control treatments in all the blocks, then one may take a connected incomplete block design (like balanced incomplete block design, partially balanced incomplete block design, Lattice design, etc.) in control treatments. The augmented design is then obtained by adding the test treatments in the blocks of the basic design in control treatments.

To make the exposition general, **an augmented experimental design is any standard experimental design in controls augmented with additional treatments in the complete block, incomplete block, the row, the column, etc.** The additional treatments require enlargement of the complete blocks, incomplete blocks in block designs and rows and columns in row-column designs, etc. The groupings in an augmented design may be of unequal sizes in block designs.

Augmented designs are available for *0*-way, *1*-way, *2*-way, *etc*. heterogeneity settings. Augmented designs eliminating heterogeneity in one direction are called augmented block designs and augmented designs eliminating heterogeneity in two directions are called augmented row-column designs. Federer (1956, 1961) gave the analysis, randomization

procedure and construction of these designs by adding the new treatments to the blocks of a RCB Design and balanced lattice design. Federer (1963) gave procedures and designs useful for screening material inspection and allocation with a bibliography. Federer and Raghavarao (1975) who obtained augmented designs using RCB design and linked block designs for one-way heterogeneity setting gave a general theory of augmented designs. They also gave a method of construction of augmented row - column designs using a Youden Square design and also provided formulae for standard errors of estimable treatment contrasts. Federer, Nair and Raghavarao (1975) gave systematic methods of construction of augmented row column designs. They have also given a procedure for the analysis of the data generated from an augmented row-column design. The estimable contrasts in such designs may be (i) among new varieties (test treatments), (ii) among check varieties (control treatments), and (iii) among all check and new varieties simultaneously. Indeed it may be possible to estimate the contrasts between check and new varieties. We shall concentrate on augmented designs for 1-way elimination of heterogeneity settings. In general, the randomization procedure for an augmented block design is:

1. Follow the standard randomization procedure for the known design in control treatments or check varieties.

2. Test treatments or new varieties are randomly allotted to the remaining experimental units, as one does in unblocked designs.

3. If a new treatment appears more than once, assign the different entries of the treatment to a block at random with the provision that no treatment appears more than once in a block until that treatment appears once in each of the blocks.

The analysis of variance of the data generated from an augmented block design with $v = u + w$ treatments comprising of w tests and u controls arranged in b blocks having k_1 plots in block *1*, k_2 plots in block *2*, and so on, and k_b plots in block b, such that $k_1 + k_2 + \cdots + k_b = n$, the total number of plots in the design, is sketched below:

ANOVA

Source	**D.F.**	**S.S.**	**M.S.**	**F-Calculated**
Blocks (eliminating treatments)	*b -1*	***ASSB***	*MSSB*	*MSSB/ MSE=F(B)*
Treatments (eliminating blocks)	*v -1*	***ASST***		
Among Tests	*w - 1*	*SST*	***MSST***	***MSST/MSE=F(T)***
Among Controls	*u - 1*	*SSC*	***MSSC***	***MSSC/MSE=F(C)***
Tests *vs* Controls	*1*	*SSTC*	***MSSTC***	***MSSTC/MSE=F(TC)***
Error	***n – v – b + 1***	*SSE*	*MSE*	*MSE*
Total	**n -1**	TSS		

In the above table the total of the sum of squares due to among tests, among controls and tests Vs controls need not be equal to adjusted treatment sum of squares. If $F_{(w-1),(n-v-b+1),\alpha}(tabulated) < F(T)$ then the effects of test treatments differ significantly. If > then there is no reason to believe that test treatments effects are different. Similar inferences can be drawn using *F (C)* and *F(TC).* The test treatment yields are then adjusted for block effects. These adjusted yields are used for multiple comparison procedures *viz.* least significant difference method, Scheffe's method of multiple comparisons, Tukey's method of all pairwise comparisons, etc. It may be noted here that the least significant difference should be applied only after the *F*-test in the analysis of variance is significant at desired level of significance. However, if one is interested only in the test treatments vs control treatments comparisons, then the Dunnett's multiple comparison procedure may be used. For a more detailed discussion on multiple comparison procedures, one may refer to Dean and Voss (1999). The statistical analysis of the data can be carried out using the PROC GLM in SAS. The steps involved alongwith several options are given below:

```
data augment;

input block treat yield;

cards;

: : :

: : :

;

proc glm;

class block treat;

model yield = block treat/ss2;

lsmeans treat/stderr pdiff;

contrast 'Among tests' treat 1 –1, treat 1 1 –2, treat 1 1 1 –3, …., treat 1 1 1 1 …. 1 –(w-1);

contrast 'among controls' treat 0 0 0 …. 0 1 –1, treat 0 0 0 …. 0 1 1 –2; …, treat 0 0 0 …0 1 1 … 1 -(u-1);

contrast 'tests vs controls' treat u u … u -w -w … -w;

run;
```

The analysis can also be carried out using the following steps of SPSS.

From the menu choose: Analyze → General Linear Models → Univariate → Select dependent variable (character to be analysed) and fixed factors (blocks and treatments) → Model → Custom (with block and treatment effect in the model). For obtaining the sum of squares of among tests, among controls and tests vs controls, we can paste the above steps in the syntax window and use the option /LMATRIX for contrast analysis. For example the sum of squares among tests can be obtained using

/LMATRIX = "among tests"

treat 1 –1 0 0 … 0;

treat 1 1 –2 0 0 …0;

treat 1 1 1 –3 0 0 …0;

…;

…;

…;

treat 1 1 1 1 …. 1 –(w-1)

Analysis can also be carried out using MS-EXCEL, the detailed steps can be taken from the authors. The online analysis module for analysis of data generated from augmented block designs is also available on Design Resources Server (www.iasri.res.in/design) at http://www.iasri.res.in/design/SpadWeb/Default.aspx.

Augmented Randomized Complete Block Designs

In National Agricultural Research System (NARS), most of such experiments are conducted using an augmented randomized complete block design. Therefore, for simplicity and ease of understanding, we shall describe the Augmented Randomized Complete Block Designs. Let us consider the experimental situation where *w* test treatments are to be compared with *u* control treatments via *n* experimental units arranged in *b* blocks such that j^{th} block is of size k_j *(>u); j=1(1)b* . For an augmented randomized complete block design, each of the control treatments is replicated *b* times and occur once in every block and test treatments occur once in one of the blocks. Therefore, it can easily be seen that in the j^{th} block there are k_j - *u* = n_j test treatments *j=1(1)b.* The randomization procedure is same as for an augmented block design discussed above. A more specified randomization procedure for an augmented randomized complete block design is

1. Randomly allot *u* controls to *u* of the k_j experimental units in each block.

2. Randomly allot the *w* test treatments to the remaining experimental units.

3. If a new treatment appears more than once, assign the different entries of the treatment to a complete block at random with the provision that no treatment appears more than once in a complete block until that treatment occurs once in each of the complete blocks.

For the analysis of data one may follow the steps of augmented block designs as discussed above. It may, however, be of importance to mention here that if one does not want to go for analysis for variance for whole data, he/she can only perform ANOVA on the control data thinking that test treatments are not present. The error mean square (MSE) will be same for comparing various treatment means. The error degrees of freedom in the above analysis shall be *(b-1)(u-1).* The minimum number of blocks (replications of the controls), to have at least *e (≥10)* degrees of freedom for error, is . For example, if there are four checks, then the minimum number of blocks should be .

The adjusted yields for test treatments can be obtained by subtracting the difference of block mean based on yield of checks and grand mean of check yields from the corresponding test yield. These adjusted test yields and check mean yields can then be subjected to all possible paired comparisons. These comparisons can broadly be classified into four categories. The different categories of comparisons along with their corresponding standard errors are

(i) Between two control treatment means

$$SE(1) = \sqrt{\frac{2MSE}{b}}$$

(ii) Between two test treatments in the same block

$$SE(2) = \sqrt{2MSE}$$

(iii) Between two test treatments not in the same block

$$SE(3) = \sqrt{2MSE(1+1/u)}$$

(iv) Between a test treatment and a control treatment

$$SE(4) = \sqrt{MSE(1+1/b+1/u-1/bu)}$$

These standard errors can be used to obtain the minimum significant difference for different categories of comparisons. A test treatment whose adjusted yield is more than the

sum of the mean yield of a check and the minimum significant difference is said to be better than the check. The test treatments with yield levels upto the satisfaction of breeder can be selected for further national level trials.

Keeping in view the importance of this design and for the ease of Biological Research Workers Agarwal and Sapra (1995) have developed a user friendly program AUGMENT1 at the Documentation Unit of National Bureau of Plant Genetic Resources, New Delhi, to analyze the data of Augmented RCB design. It is interesting to note that the augmented RCB design is variance balanced with respect to tests *vs* controls comparisons. As mentioned earlier, the augmented designs can be obtained by augmenting test treatments in the blocks of incomplete block designs, row-column designs, etc. in control treatments.

Note: A survey of the literature reveals that generally the experiments described above are conducted using an augmented randomized complete block design. However, the experimenters would often like to know how many times the control treatments be replicated in each of the blocks so as to maximize the efficiency per observation for making test treatments *vs* control treatments(s) comparisons? Parsad and Gupta (2000) have answered this question. Let us assume that there are w test treatments which occur only once in the design and each of the u controls occurs in each of the b blocks. Then to maximize the efficiency per observation the number of times each control appears in each of the blocks is

$$a = \frac{\sqrt{u+b-1}\,\sqrt{w}}{ub}$$

provided $u = 1, b \leq w$, For example, consider the problem of obtaining the optimum number of replications of the controls in an experiment with $w = 24$, $u = 3$, $b = 4$. We have Similarly, for $w=98$, $u=2$, $b=7$, we have

$a = \frac{1}{3}\sqrt{\frac{6\times 24}{4\times 4}} = 1.$ Similarly, for $w=98$, $u=2$, $b=7$, we have $a = \frac{1}{2}\sqrt{\frac{8\times 98}{7\times 7}} = 2$

Remark: For a single control situation, *i.e.* $u = 1$, the above expression reduces to $a = \sqrt{\frac{w}{b}}$ and it can easily be seen that for $u = 1, b \leq w$, which is always true.

There may, however, arise many combinations of w, u and b for which the above expression of a does not yield a positive integer value of a. In such situations, a question that arises is as to what integer value of a should be taken? To answer this question, the efficiency per observation has been calculated for $w \leq 100$, $b \leq 25$ and $u \leq 10$ such that $b + u - 1 \leq w$ and a has been taken as $a^* = int\,(a)$ and $int\,(a) + 1$ besides taking $a=1$. A close scrutiny reveals that if value of $a > \#.\ 42$ then take $a^* = int\,(a) + 1$ and for values of a smaller than or equal to $\#.\ 42$ take $a^* = int\,(a)$ for $u \geq 2$. For $u = 1$, the same rule applies but the value of a is taken as $\#.\ 45$ instead of $\#.\ 42$.

Parsad and Gupta (2000) also gave the steps of the analysis of experimental data generated from the augmented completed block designs using *a* as above. As mentioned earlier, the error mean square (MSE) can be obtained using the ANOVA on control treatments only. The error degrees of freedom in this case will be *uab - u - b+1*. The minimum number of blocks (replications of the controls), to have at least *e (≥10)* degrees of freedom for error, is $\frac{e+u-1}{ua-1}$.

Statistical Package for Augmented Designs

A user friendly, menu driven, graphic user interface (GUI) based Statistical Package called STATISTICAL PACKAGE FOR AUGMENTED DESIGN (SPAD) has been developed at IASRI, New Delhi. The package generates randomized layout of augmented designs and performs the analysis of data generated. For given number of test treatments, number of control treatments and number of blocks, it computes the optimum replication number of each control treatment in every block of the design such that the efficiency per observation of the test treatments *vs* control treatment(s) comparisons is maximum. The package also provides flexibility in choosing the replication number of each control treatment in every block. The user can define the replication of each control treatment in every block. Once the user defines the number of test treatments, number of control treatments, and number of blocks in the design, the randomized layout of the design is generated. The package also provides the analysis of the data generated from augmented designs. A null hypothesis on any user-defined contrast can also be tested.

The package is very useful for classroom teaching as well as for the researchers in statistics with interest in experimental designs. The package has been developed using Microsoft Visual C++ 6.0. Software is completely stand-alone and can be installed on any hardware platform with 32 Bit Microsoft Windows Operating System. Software can be executed with minimum specification of RAM for host Operating System. Installation of SPAD takes 2 MB of hard disk space and at least 1 MB free space for its working. Software is menu driven and is very user friendly. It has a rich edit control for text editor and supports cut, copy, paste, undo, find and find-replace facilities. A Context Sensitive Help with Contents, Index and Search facilities is also available. The software is designed to assist experimenters in planning and analysing augmented designs.

Generation of augmented design

We begin with the generation of randomized layout of augmented complete block design with each control replicated *a* () times in each block. When *a* = 1, it reduces to usual augmented randomized complete block design and when *a* = *r*, the number of replications of control treatments per block that maximize the efficiency per observation, then we get the

randomized layout of the augmented complete block design that maximizes the efficiency per observation with respect to test treatments *vs* control treatment(s) contrasts. One can select the option *Augmented Designs* from the menu and then select the sub-option *Generate Design*. On selecting the sub-option *Generate Design*, a form for entering the design parameters is displayed. For generation of randomized layout of augmented design, the input in terms of number of control treatments, number of test treatments and blocks available with experimenter is required. Once the user enters the design parameters, the replication of control treatment(s) that maximizes the efficiency per observation is automatically computed and suggested to the experimenter. There is flexibility for user to change the replication number of the control treatments. To change replication of control treatments, one has to check on the "*Change Replication of Control*" check box. This will enable an edit box for replication of control treatments, where desired number of replication for control treatments can be given.

Once the desired number of replications of control treatment(s) is entered, the box for entering replication of test treatments and block sizes get activated. Software also displays the total number of experimental units required. The block sizes are to be entered by the user. The package accepts blocks with unequal sizes also.

Analysis of data generated from augmented design using SPAD

The data pertaining to an augmented block design is analysed as per procedure of analysis of general block designs. The treatment sum of squares is partitioned into different components of interest *viz.* (i) among test treatments, (ii) among control treatments and (iii) among test treatments and control treatments. The pair wise treatment comparisons can be simplified for an augmented complete block design in which each of the control treatments appear in each block '*a*' times.

For an augmented incomplete block design, the significance of all possible pair wise treatment comparisons can be tested by automatically generating all the possible elementary treatment contrasts.

For performing the analysis of data generated through an augmented block design, an ASCII data file in a specified format is required. The existing ASCII data file can be opened in the SPAD window using File-Open options. A new data file can also be created in the SPAD window using File-New option. One can also copy and paste data into SPAD editor from any windows based software like Excel or which supports clipboard operations. For creation of data file in a specified format, the treatments are renumbered as 1, 2, ..., u, u + 1, ..., $u + w$. Here first u treatments are the control treatments and u+1, ..., $u + w$ are the test treatments. Data file contains at least three columns; first column represents block number, second column represents treatment number and third column consists of observed value of character. If there is more than one character to be analyzed, then the characters

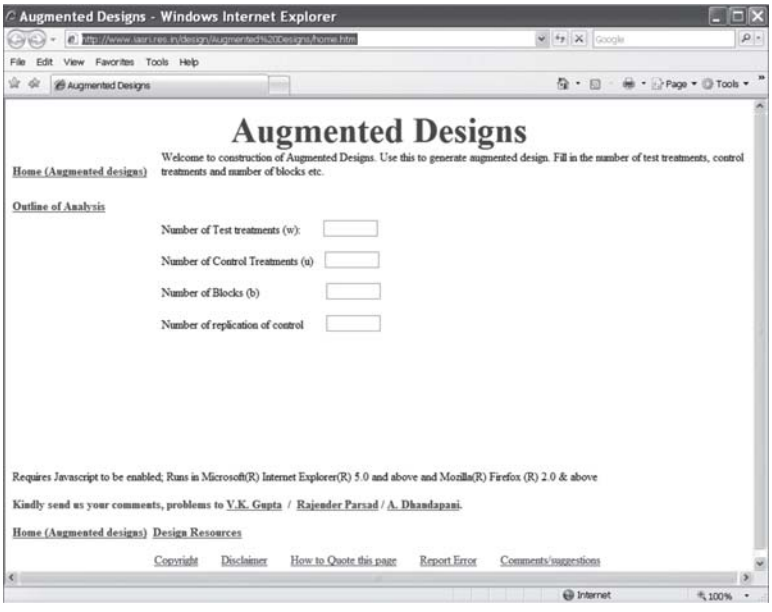
can be entered from fourth column onwards. There is no limitation on the number of characters present in the file. All these data values must be separated by a SPACE or a TAB.

For performing the analysis of the data generated through an augmented block design, one can select the sub-option *Analyze Block Design* from Option *Augmented Design* in the menu. A click on sub-option *Analyze Block Design* displays a dialog box. In this dialog box user must specify the character to analyze this time. This box will only appear if data file has more then one character. Once a character is selected for the analysis, complete analysis with two ANOVA tables; one for testing the equality of treatment effects and another for testing the equality of block effects, R^2, Coefficient of Variation, Root Mean Square Error (RMSE), General Mean and adjusted treatment means is generated. For partitioning the treatment sum of squares into components of interest *viz.* (i) among test treatments, (ii) among control treatments and (iii) among test treatments *vs* control treatments, one can select the sub-option *Contrast Analysis*. There are three options within the contrast analysis viz. (i) Augmented CB design, (ii) GBD for Tests *vs* Control(s) and (iii) User Defined Contrasts. Here *Tests* is used for test treatments and *Controls* for control treatments. If the data is generated from an augmented design in which each control treatment appears equally often in all the blocks, then the option *Augmented CB* design can be used for obtaining partitioned sum of squares and critical differences for performing all possible pairwise treatment comparisons. If the data is generated from an augmented incomplete block design, then the option *GBD for Tests vs Controls* may be used. In this option, the exact probability levels of significance of all possible pairwise treatment comparisons are given in a $(u + w) \times (u + w)$ matrix. A null hypothesis on any other contrast of interest can be tested using *User Defined Contrasts*.

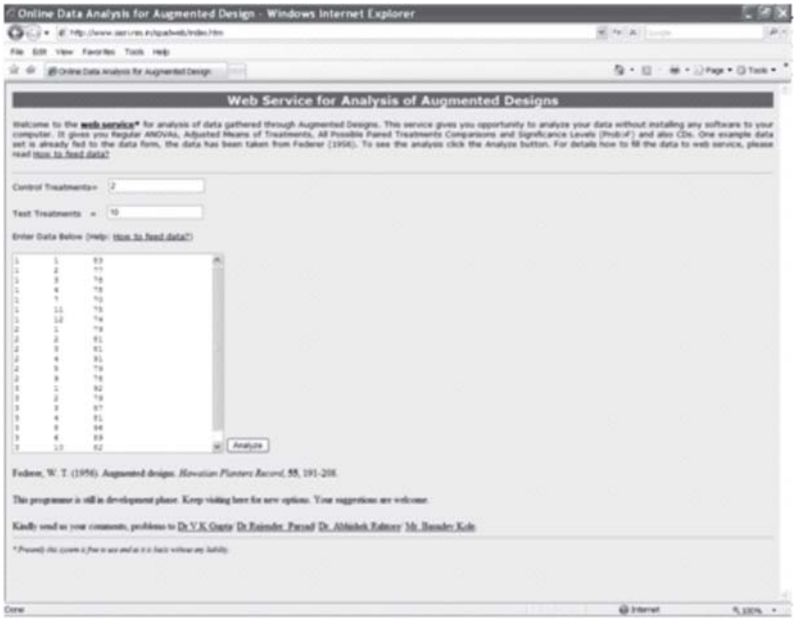
This software is an effort to help the experimenters who are conducting experiments using augmented block designs. This software provides a complete solution in the sense that it is capable of generating the augmented complete block design that maximizes the efficiency per observation and analysing the data generated from any augmented block design. In fact the data from any general block design can be analysed using SPAD. Software is very easy to use and can be operated without any help or training. However, there is still a need to incorporate the features of analysis of covariance, stability analysis and variance components estimation from the data generated from an augmented block design.

Web Resources on Augmented Designs

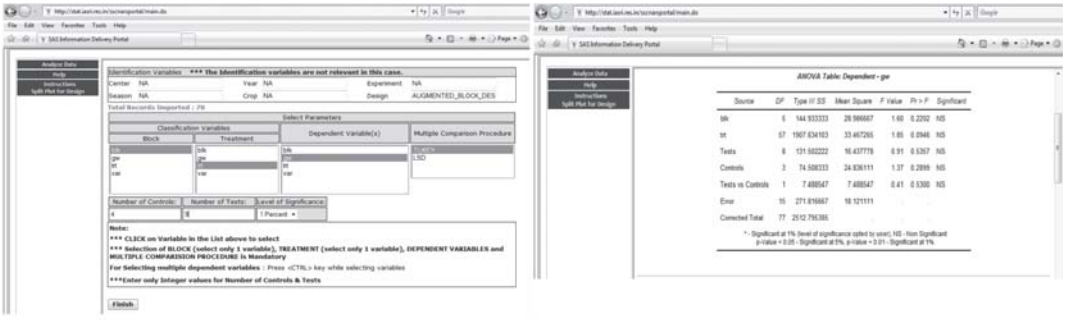
Online software for generation of randomized layout of an augmented randomized complete block design for given number of test treatments, control treatments and number of blocks with given block sizes, not necessarily equal, is developed and is available at www.iasri.res.in/design/Augmented Designs/home.htm.The design can be generated with optimum replication of control treatments in each block so as to maximize efficiency per observation. A screenshot is as shown below:



An experimenter can perform online analysis of data generated from augmented randomized block designs. This is available at **www.iasri.res.in/spadweb/index.htm**.



Analysis of Data generated from Augmented Block Designs can also be performed from IP Authenticated Indian NARS Statistical Computing Portal (http://stat.iasri.res.in/sscnarsportal) using the link augmented block designs. The analysis report includes ANOVA Table, adjusted



treatment means and minimum significant differences for control treatment means, test treatments in same block, test treatments in different blocks, test versus control treatment.

Designing experiments with some replications of the tests

In the experimental situations discussed above, the material on test treatments is scarce and it is not possible to replicate the tests, though the controls are replicated and it is the replications of the controls alone that provide us the experimental error. Such experiments are conducted to identify the promising genotypes or lines or test treatments. Once these are identified, then a further experimentation is carried out with the tests identified as promising. In these experiments, it is possible to replicate the tests also besides replicating the controls. However, it is taken care of that replication number for new treatments is less than the replication number for control treatments. For the earlier situations, the designs have been obtained by augmentation of standard experimental design in control treatments while for the later situation, the standard experimental design is taken in new treatments and augmented with suitable number of replication of control treatments and some new blocks in new and/or control treatments may also be added. The efficient designs for comparing the test (new varieties) treatments with that of control (check varieties) treatments have been developed under different names *viz*. Reinforced incomplete block designs, supplemented balanced block designs, inter and intra group balanced block designs with varying replications, balanced bipartite block designs and balanced treatment incomplete block designs. An excellent comparison of these designs and analytical procedures has been comprehended in the '*Monograph of supplemented block designs*' by Nigam, Gupta and Narain (1979).

For the last two decades, attempts have been made by various research workers to investigate the optimality and construction aspects of designs used for making tests *vs* controls comparisons. For a critical and excellent review on the subject, reference may be made to Hedayat, Jacroux and Majumdar (1988), Majumdar (1996) and Gupta and Parsad

(1999). Methods of construction and the catalogues of these designs along with their efficiencies can be used by experimenters for choosing an experimental design. A considerable amount of work has also been done at the Division of Design of Experiments, IASRI, New Delhi for obtaining efficient designs for single control as well as for more than one control treatments under both proper and non-proper block design settings. The catalogues of efficient designs have been prepared.

Most of these experiments are still being conducted using randomized complete block designs. In these designs the tests as well as controls are equally replicated and hence all possible paired comparisons are estimated with same variance. The main objective of these experiments is, however, to make the test treatments-vs- control treatments comparisons with more precision than the tests vs controls comparisons. Therefore, to meet this objective of these experiments, an orthogonal block design in which the control treatments should be replicated times that of the test replications in each of the *b* blocks should be used. However, with the large number of treatments, the blocks of the RCB design become large and it may not be possible to maintain homogeneity among the experimental units within blocks. As such the primary purpose of forming blocks to have homogeneous experimental units within a block gets defeated. A direct consequence of laying out an experiment in a RCB design or an orthogonal block design with large number of treatments is that the coefficient of variation (CV) of the design may become large. This amounts to saying that the error sum of squares is large compared to the sum of squares attributable to the model. Hence, the small treatment differences may not be detected as significant. It also leads to a poor precision of comparisons of the two treatment effects. The experimental error through designing can be controlled by the use of incomplete block designs for making test treatments-control treatments comparisons. In these designs, the number of experimental units in a block is generally smaller than the total number of treatments. Consequently, the per plot variance is small leading to a considerable reduction in error mean squares as compared to that of a RCB design or an orthogonal block design in which the control treatments requires to be replicated times that of the test replications in each of the *b* blocks and where the blocks are large resulting thereby in a large per plot variance and hence a large error mean square. The precision of treatment comparisons of the treatment effects through incomplete block designs is also high if we use the incomplete blocks designs for making test treatments-control treatments comparisons. Despite the merit of incomplete block designs, these designs have not found favour of the experimenters. The possible reasons for that includes

1. Lack of awareness about the designs and the randomized layout of the design and
2. The fear of the analysis of the data.

Although the analysis of the data generated from these designs can be carried out as per procedure of augmented block designs using PROC GLM of SAS, still there is need to develop statistical softwares that can generate the randomized layout of the design and carry out the analysis of data. Moreover, close interactions between the experimenters and

the statisticians through training/extension programmes or one to one discussions are necessary.

These designs are also useful in on farm experiments in which the farmers' practice is to be compared with that of treatments identified from research stations. What treatment is to be taken as farmers' practice is a problem as this varies from farmer to farmer? One possible solution of this problem is that we take as many controls as there are farmers in the trial and add the farmers' practice to each of the blocks of one replication of resolvable incomplete block design. The problem still needs attention. Some work in this direction has been done by Nigam, Parsad and Gupta (2006).

For further details regarding methods of construction, applications and bibliography on designs for making test treatments-control treatment comparisons a reference may be made to Design Resources Server available at **www.iasri.res.in/design**. We now give the analysis of the data generated from an experiment laid out as a reinforced block design that is efficient for making test treatments-control treatment(s) comparisons.

Sampling in Field Experiments

In the experiments related to germplasm evaluation, generally, the data on multiple quantitative characteristics are recorded. Most of these characters are recorded on a random sample of plants selected from the plot. The number of plants selected is generally arbitrary and subjective in nature. It has also been observed that the analysis of data generated from these characters is performed based on mean of the observations taken from the plants pertaining to the same plot rather than utilizing the individual observations. This amounts to losing some information. Therefore, it is suggested that this data should be analysed as per procedure of the analysis of data on sampling in field experiments. This analysis helps in obtaining an estimate of the sampling variance. This estimate of sampling variance then can be used to obtain the optimum sample size. This optimum sample size may be considered for the trials at the second stage for the future first stage trials on the same crop. Sometimes, the optimum sample size may differ from character to character. In such situations, the largest optimum sample size shall be considered. For details on the analysis of sampled experimental data one may refer to Nigam and Gupta (1979) and Gomez and Gomez (1984).

Multivariate Analytical Techniques

The data on multiple quantitative characteristics (e.g. plant height, grain yield, flowering date, growth habit, seedling vigour, etc) may also be used to perform multivariate analysis of variance. The data on qualitative characters like grain colour, agronomic scale, disease reaction, etc. is also recorded. The data on quantitative and qualitative multiple characteristics may also be used to carry out the cluster analysis for studying the genetic diversity of the accessions in gene banks by grouping accessions into sub-populations or clusters according to attributes of agronomic performance, disease resistance, and so on. For various clustering

procedures based on categorical and continuos variables, one may refer to Franco et al (1998, 1999).

The data on multiple characteristics may also be used for evaluating the association of trait expression by computing the correlation coefficients between all possible pairs of the traits. A relationship between yield and other biometrical characters may also be established using the linear or curvilinear regression.

Mult–location/Season Trials

In any plant improvement programme or the germplasm evaluation trial, the experimenter would like to test the adaptability of promising germplasms at different environmental conditions. This is achieved by conducting the multi-locational trials. Let there be entries and checks to be tested at *L* locations. The entries in these trials may be unreplicated at each of the locations (if the seed for the test treatments is enough for just one replication at each of the locations) or may have replications. The experiment is conducted using a block design at each of the locations. Let there be blocks at the t^{th} location. The combined analysis of the data can be carried out using the following steps.

Step 1: Perform the ANOVA individually for each location. We get mean square errors for each of the *L* locations.

Step 2: Use the Bartlett's test for testing the equality of error variances.

Step 3: If the error variances are homogeneous, *i.e.*, they do not differ significantly, then use the ANOVA on the original data as given in Step 5.

Step 4: If the error variances are significantly different, then divide each observation of a location by its corresponding root mean square error (Aitken's transformation) and use the ANOVA on the transformed data.

Step 5: In the model take the treatment effects, the environment effects or the location effects, the block (environments) effects and treatments*environments interaction. The outline of the ANOVA is given as follows:

ANOVA

Source	**D.F.**
Environments	*L-1*
Blocks (Environments)	$\sum_{t=1}^{L}(b_t - 1)$
Treatments (eliminating blocks)	*v -1*
Among Tests	*w - 1*
Among Controls	*u - 1*
Tests *vs* Controls	*1*
Treatment*Environment	*(L-1)(v -1)*
Error ***By subtraction***	
Total	**L(w+ub)-1**

This ANOVA enables the experimenter to test the treatment * environment interaction. If, however, the treatment * environment interaction is not significant, then another analysis of variance can be performed by dropping the interaction term from the model. In these trials, the checks generally represent the national level checks (established varieties in terms of yield, disease resistant, etc.) for overall comparisons and local checks for testing the adaptability for a specific environment. The local checks are generally different at each location. Some of the tests may also be different at each of the locations. Therefore, only some of the treatments are common to environments. The analysis of the data generated from such trials can also be analyzed by making certain modifications in the above procedure. For details, please see Searle (1971).

Another feature of germplasm evaluation trials is that these trials are conducted over years. The list of entries may change from year to year because the new entries are included as they become available and those with poor performance are deleted from the further considerations. The analysis of such data can be analysed in the same way as described above with years replacing locations. However, Hill Jr. and Rosenberger (1985) proposed several procedures of combining total season yields from such trials conducted at the same location. They mainly concentrated on the problem of estimating the mean yields for the cultivars and experimental lines included in a series of trials that did not contain all entries in equal numbers. The methods of estimation included percent of checks, summation of differences between entries and checks, two-way analysis with trials and entries as factors. In the two-way classified, additive model considered by them the trial x genotype interaction

has not been included. They have presented five different versions of this model. Version 1 and 2 consider the fixed effects model. Version 1 is useful when the error variances for the different trials are homogeneous. Version can be used when the error variances for different trials are significantly different from each other. Versions 3 to 5 were the three different versions of Best Linear Unbiased Prediction and can be useful when the entries are considered as a random sample and their effect is taken as random.

For more details on the statistical tools in germplasm evaluation, a reference may be made to Dhillon *et al.,* (2004).

References

Agrawal RC and RL Sapra (1995) Augment 1: A microcomputer based program to analyze augmented randomized complete block design. *Indian J. Plant Genet. Resour.* **8 (1)**, 63-69.

Cox DR (1958) *Planning of Experiments.* John Wiley & Sons, New York.

Dean A and D Voss (1999) *Design and Analysis of Experiments*. Springer Texts in Statistics, Springer, New York, pp 67-97.

Dhillon BS, BS Singh, M Parsad, R Gupta VK and BB Singh (2004) Evaluation of plant genetic resources: issues, advances and opportunities. In: BS Dhillon, RK Tyagi, A Lal and S Saxena (eds). *Plant Genetic Resource Management*. Narosa Publishing House, New Delhi, pp 242-265.

Federer WT (1956) Augmented (Or Hoonuiaku) Designs. Hawaiian Planter's Record **55(2)**: 191-208.

Federer WT (1961) Augmented Designs with one way elimination of heterogeneity. *Biometrics.* **17**: 447-473.

Federer WT (1963) Procedures and designs useful for screening material in selection and allocation with a bibliography. *Biometrics* **19**: 553-587.

Federer WT and D Raghavarao (1975) On Augmented Designs. *Biometrics* **31**: 29-35.

Federer WT, Nair RC and D Raghavarao (1975) Some augmented row column designs. *Biometrics* **31**: 361-374.

Franco J, J Crossa , J Villasenor, S Taba and SA Eberhart (1998) Classifying genetic resources by categorical and continuous variables. *Crop Sci.* **38**: 1688-1696.

Franco J, J Crossa, J Villasenor, A Castillo, S Taba and SA Eberhart (1999) A two-stage, three-way method for classifying genetic resources in multiple environments. *Crop Sci.* **39**: 259-267.

Gomez KA and AA Gomez (1984) *Statistical Procedures for Agricultural Research, 2nd Edition*. John Wiley and Sons, New York.

Gupta VK and R Parsad (1999) Designs for making test treatments vs control treatment (s) comparisons – an overview. Special invited paper presented in the second annual conference of the society of Statistics, Computer and Applications held at St. Thomas College, Pala, Kerala organized during November 28 – December 1, 1999.

Gupta VK and R Parsad (2001) Block designs for comparing test treatments with control treatments - An overview. *Special issue of Statistics and Applications to felicitate the 80th Birthday of Dr. M.N.Das.* **3(1 & 2)**: 133-146.

Healey MJR (1956) The analysis of factorial experiments with additional treatments. *J. Agric. Soc.* **47**: 23-27.

Hill Jr RR and JL Rosenberger (1985) Methods for combining data from germplasm evaluation trials. *Crop Sci.* **25**: 467-470.

Kempton RA and PN Fox (1997) *Statistical Methods for Plant Variety Evaluation*. Chapman & Hall, London.

Majumdar D (1996) Optimal and efficient treatment-control designs. *Handbook of Statistics*, 13, Editors: S Ghosh and CR Rao, Elsevier Science B.V.

Nigam AK and VK Gupta (1979) *Handbook on Analysis of Agricultural Experiments*. IASRI, New Delhi.

Nigam AK, VK Gupta and P Narain (1979) *Monograph of Supplemented Block Designs*. IASRI, New Delhi.

Parsad R and VK Gupta (2000) A Note on augmented designs. *Indian J. Plant Genet. Resour.* **13(1):** 53-58.

Pearce SC (1960) Supplemented Balance. *Biometrika* **47**: 263-271.

Searle SR (1971) *Linear Models*. John Wiley and Sons, New York.

Sinha BK (1980) Optimal block designs. *Unpublished Seminar Notes*, Indian Statistical Institute, Calcutta.

Sinha BK (1982) On complete classes of Experiments for certain Invariant problems of Linear Inference. *J. Statist. Plann. Inf.* **7**: 171-180.

Ture TE (1994) Optimal row-column designs for multiple comparisons with a control: A complete catalog. *Technometrics*, **36**: 292-299.

ICAR-National Bureau of Plant Genetic Resources

Pusa Campus, New Delhi – 110 012

Phone: 01-011-25843697; Fax: 01-011-25842495

E-mail: director@nbpgr.ernet.in Website: https://www.nbpgr.ernet.in