

CONSTRUCTING AUGMENTED EXPERIMENT DESIGNS WITH GENDEX

By

Walter T. Federer, Cornell University

And

Nam-Ky Nguyen, Design Computing

ABSTRACT

It is shown how to construct augmented row-column, augmented lattice square, and augmented complete and incomplete block experiment designs. Seven examples are used to demonstrate the procedure. The first example is for $s = 3$ standards or checks and $n = 24$ new treatments with $r = 8$ replications of the checks arranged in an eight row by six column plan. The second example uses an optimal input design for Example 1. Example 3 is constructed for $s = 4$ checks and $n = 120$ treatments arranged in a 15-row by 12-column plan. There are $r = 15$ replications for the check treatments. The fourth example uses an optimal input plan for Example 3. The fifth example is an augmented lattice square experiment design with $s = 2k = 10$ standards and $n = 60$ new treatments with $r = 4$ replications of the check treatments. In Example 6 it is shown how to construct an augmented randomized complete block experiment design for $s = 3$ check treatments, $r = 5$ replications of the checks, and $n = 30$ new treatments. The last example involves the construction of an augmented incomplete block experiment design for $s = 8$ checks in incomplete blocks of size $k = 4$, $r = 3$ replications of the checks, and $n = 60$ new treatments.

Key words: Augmented row-column experiment design, augmented lattice square experiment design, Latin square, Optimal design, Augmented randomized complete block experiment design, Augmented incomplete block experiment design, Randomization.

BU-1610-M in the Technical Report Series of the Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, New York 14853

INTRODUCTION

Augmented experiment designs are constructed by selecting *any* experiment design for s check treatments with r replications of the check treatments. Then the block sizes or the number of rows and columns are expanded so that n new treatments may be accommodated. The new treatments are treatments that have been untried or untested, and they are to be screened for a given response, say yield. The new treatments represent a random selection of new entities, e. g. genotypes, and are therefore random effects. Augmented designs are used world-wide in screening new genotypes and other untried and untested material.

It is shown how to construct augmented row-column experiment designs using the BIB module of the toolkit Gendex (Nguyen, 2001) and by editing the output from the BIB module to obtain an optimal design. A similar procedure is used to construct augmented lattice square

experiment designs (Federer, 2002). Augmented block experiment designs are obtained by using the BIB module to create a blocked design for s checks with r replications of the checks. Then the block sizes $k \leq s$ are enlarged to accommodate the n new treatments. Examples are used to illustrate the procedure.

AUGMENTED ROW-COLUMN EXPERIMENT DESIGNS

To obtain an augmented row-column experiment design using Gendex, we proceed as follows. First, select the number of rows, r , and the number of columns, c , to be used. Then set the number of treatments equal to the smaller of the number of rows or columns. Using the BIB module of Gendex (See Federer *et al.*, 2001a), letting $r > c$, and letting $c = v$ the number of treatments, construct a randomized complete block experiment design with r replicates but select the non-resolvable option. Copy this design from BIB.HTM and paste it in notepad. Edit out all non-data lines and spaces. Use this design and the RCD module to construct an r -row by c -column plan with v treatments. Let $s < v$ be the number of standard or check treatments. Then delete all numbers greater than s . This will leave $r(v - s) = n$ blank places. Fill in these n spaces with n new treatments each in once. Below we give two examples of an augmented experiment design. The first example is an 8-row by 6-column design with $s = 3$ standards. The second example is a 15-row by 12-column design with $s = 4$ standards.

Example 1--The output for $v = 6$, $k = 6$, and $r = 8$ from the BIB module was obtained, copied in notepad, and all non-design items and spaces were edited out. This material was saved as rc6.txt in the Gendex directory. Then using the RCD module, the following output was obtained.

RCD 2.2: Construct a non-resolvable row-column design of size (r, k, b)
(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: Non-resolvable row-column design for $v=6$, $r=8$, $k=8$ and $b=6$.

```
try #          3
seed          96691959
# of iterations 10
f             3564.94
concurrences  14(1) 15(1) 16(2) 17(2) 18(3) 19(3) 20(2) 21(1)
E(column)     0.9115
E(row)        1
E             0.9115
E/U           0.9352
```

Plan:

0	1	3	5	2	4
0	2	4	5	3	1
4	2	1	5	0	3
1	5	2	0	4	3
4	5	0	2	3	1
3	0	5	1	4	2
5	3	2	4	1	0
5	2	3	0	4	1

Note: RCD used 0.11 seconds.

Using this output, the following is the augmented experiment design

0	1	3	4	2	5
0	2	6	7	8	1
9	2	1	10	0	11
1	12	2	0	13	14
15	16	0	2	17	1
18	0	19	1	20	2
21	22	2	23	1	0
24	2	25	0	26	1

where the checks are numbered 0, 1, and 2, and the $n = 24$ new treatments are numbered from 3 to 26.

Example 2--The above design is not balanced in columns as much as is possible. For the input plan in notepad, edit the BIB design in notepad to obtain the following 6×6 Latin square design with the first two rows repeated to make an 8×6 Latin rectangle design:

1	2	3	4	5	0
0	1	2	3	4	5
5	0	1	2	3	4
4	5	0	1	2	3
3	4	5	0	1	2
2	3	4	5	0	1
1	2	3	4	5	0
0	1	2	3	4	5

The output for a row-column design using the above and the RCD module is:

RCD 2.2: Construct a non-resolvable row-column design of size (r, k, b)
 (C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: Non-resolvable row-column design for $v=6$, $r=8$, $k=8$ and $b=6$.

```
try #          2
seed          166110494
# of iterations 10
f             3746.63
concurrences  18(9) 19(6)
E(column)     0.9746
E(row)        1
E             0.9746
E/U           1
```

Plan:

0	2	3	4	5	1
3	1	2	5	0	4
1	2	4	0	3	5
1	0	3	2	4	5
5	3	2	1	4	0
4	1	5	3	2	0
2	5	0	4	1	3
0	4	1	3	5	2

Note: RCD used 0.11 seconds.

Note that the $E = 0.9746$ and $E/U = 1$ for the above but $E = 0.9115$ and $E/U = 0.9352$ for the design in Example 1. Hence this is a much better design. Note that in order to obtain the above, it was necessary to edit the input from a BIB output to obtain the desired form. The check treatment numbers would need to be randomly assigned.

To obtain the augmented row-column experiment design, one may delete numbers 3, 5, and 0 or numbers 1, 3, and 5. Then the missing cells would be allocated to the 24 new genotypes. The former deletion produces an augmented experiment design wherein row, column, and check treatment effects are estimable. However many plant breeders would desire the latter deletion plan as the checks would be uniformly distributed among the new genotypes. All row, column, and check effects do not have solutions for this design. An analysis is possible using functions of row and column effects such as linear, quadratic, cubic, etc. trends of row and columns totals on row and column positions. Federer (1998) and Federer and Wolfinger (1998, 2002) have given computer codes for this situation and for treating the new genotypes as random effects.

If the numbers 2, 4, and 0 are the check treatments for a numerical example, a part of the SAS/GLM analysis of variance, ANOVA, output is

Source	DF	Type I SS
Row	7	26.0000
Column	4	17.6667
Treatment	2	0.0000

Source	DF	Type III SS
Row	6	23.3333
Column	4	17.3333
Treatment	2	0.0000

For the Type I (nested) ANOVA, one column degree of freedom for columns is confounded with row effects. This shows up in both Type I ANOVA as one column degree of freedom is lost and in Type III ANOVA as one degree of freedom is lost for both rows and columns.

If treatment numbers 0, 1, and 3 are used as the check treatments, the ANOVAs are

Source	DF	Type I SS
Row	7	26.0000
Column	5	9.5782
Treatment	2	0.0186

Source	DF	Type III SS
Row	7	22.9773
Column	5	9.3265
Treatment	2	0.0186

For this augmented row-column design, the ANOVA for the check treatments contains the correct degrees of freedom for all three variables.

Example 3--In order to construct an augmented experiment design for $s = 4$ checks and $n = 120$ new treatments in a 15-row by 12-column design, we first use the BIB module to construct a randomized complete block experiment design for $v = 12$ treatments and $r = 15$ replicates. This plan is pasted in notepad, edited, and saved as rc15.txt, say. Then using the RCD module with rs15.txt, the following out put was obtained:

RCD 2.2: Construct a non-resolvable row-column design of size (r, k, b)
(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: Non-resolvable row-column design for $v=12$, $r=15$, $k=15$ and $b=12$.

```
try #      83
seed      -255160763
# of iterations 38
f         56068.88
concurrences 21(1) 23(1) 25(1) 26(2) 27(2) 28(4) 29(4) 30(1) 31(6) 32(9)
33(7) 34(12) 35(2) 36(5) 37(3) 38(1) 39(3) 40(1) 41(1)
E(column)  0.9262
E(row)     1
E          0.9262
E/U        0.9365
```

Plan:

4	0	10	1	8	11	7	9	2	6	3	5
0	7	11	8	6	3	9	4	2	1	10	5
5	10	11	2	3	7	9	1	4	0	6	8
7	2	11	0	10	5	8	1	6	4	3	9
3	2	9	6	0	10	11	1	4	5	8	7
5	7	8	6	11	3	10	9	1	4	2	0
1	10	2	11	5	7	6	9	0	8	3	4
9	2	11	7	3	1	4	6	10	8	5	0
8	2	3	7	10	5	0	9	11	1	4	6
3	4	2	1	8	6	10	0	7	9	5	11
8	3	6	1	10	7	5	4	2	9	0	11
10	0	5	6	3	9	2	8	11	1	7	4
7	5	0	8	4	10	9	3	2	11	6	1
5	4	3	1	8	0	10	6	2	7	9	11
1	9	4	7	11	10	8	0	6	3	5	2

Note: RCD used 0.93 seconds.

The following 15-row by 12-column augmented experiment design with $s = 4$ checks and $n = 120$ new treatments, numbers 4 to 123, is obtained:

4	0	5	1	6	7	8	9	2	10	3	11
0	12	13	14	15	3	16	17	2	1	18	19
20	21	22	2	3	23	24	1	25	0	26	27
28	2	29	0	30	31	32	1	33	34	3	35
3	2	36	37	0	38	39	1	40	41	42	43
44	45	46	47	48	3	49	50	1	51	2	0
1	52	2	53	54	55	56	57	0	58	3	59
60	2	61	62	3	1	63	64	65	66	67	0
68	2	3	69	70	71	0	72	73	1	74	75
3	76	2	1	77	78	79	0	80	81	82	83
84	3	85	1	86	87	88	89	2	90	0	91
92	0	93	94	3	95	2	96	97	1	98	99
100	101	0	102	103	104	105	3	2	106	107	1
108	109	3	1	110	0	111	112	2	113	114	115
1	116	117	118	119	120	121	0	122	3	123	2

where the checks are numbered 0, 1, 2, and 3 and the $n = 120$ new treatments are numbered from 4 to 123.

Example 4--A better experiment design may be obtained by editing the above design, rc15.txt, in notepad to make it a 12×12 cyclical Latin square with rows 1, 2, and 4 duplicated to make 15 rows. Doing this, the output from the RCD module is:

RCD 2.2: Construct a non-resolvable row-column design of size (r, k, b)
(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: Non-resolvable row-column design for $v=12$, $r=15$, $k=15$ and $b=12$.

```
try #          17
seed          1510801413
# of iterations 43
f             58764.24
concurrences  33(30) 34(36)
E(column)     0.989
E(row)        1
E             0.989
E/U           1
```

Plan:

9	2	11	0	7	3	1	5	10	6
	4	8							
2	4	0	5	11	9	3	10	6	7
	8	1							
0	8	10	3	2	4	5	9	7	6
	1	11							
0	5	3	1	4	11	6	2	7	10
	9	8							
7	0	6	2	3	8	11	1	9	4
	10	5							

3	0	11	2	1	4	5	6	8	9
	7	10							
1	6	8	10	0	7	9	4	3	5
	11	2							
9	3	2	7	4	5	8	6	1	11
	0	10							
10	11	1	8	9	5	0	3	4	2
	7	6							
11	10	1	9	6	2	7	4	0	8
	5	3							
11	9	5	4	10	1	3	7	2	8
	6	0							
6	10	2	11	1	0	4	8	5	3
	9	7							
5	7	9	0	3	6	10	11	8	1
	2	4							
8	1	4	6	5	10	2	7	11	0
	3	9							
4	1	7	3	8	2	6	0	5	9
	10	11							

Note: RCD used 0.38 seconds.

The row-column experiment design obtained above has an E column of 0.989 and an E/U = 1 whereas the design of Example 3 only has an E column of 0.9262 and an E/U = 0.9365. There are only two types of occurrences for this example as opposed to several for Example 3. Hence more optimal row-column designs of this nature can be obtained by the above method than by using the method of Example 3.

To construct an augmented row-column experiment design from the above, the experimenter may select the desired diagonals for check treatments, delete the remaining numbers, and fill in the deleted spaces with the new treatments (genotypes). The numbers to the check treatments should be randomized.

AUGMENTED LATTICE SQUARE EXPERIMENT DESIGN

To obtain an input design for a lattice square experiment design, the BIB module is used to obtain an incomplete block design for $v = k^2$ treatments in incomplete blocks of size k and r replicates. This design is saved as BIB.HTM. It is copied and pasted in notepad. Then, this plan is edited to obtain an appropriate form for constructing a lattice square plan that can be used for constructing an augmented lattice square design. Any two diagonals of the plan must contain the $s = 2k$ checks such that there are 2 checks in each row and in each column (See Federer, 2002). The remaining $n = rk(k - 2)$ empty spaces are filled with the n new treatments. The RRC module is used to obtain the plan.

Example 5--A one-resolvable incomplete block experiment design was obtained using the BIB module for $v = 25$, $k = 5$, and $r = 4$. This plan was edited to obtain the following plan that placed numbers 1 to 5 on the main right diagonal and numbers 6 to 10 on the diagonal above the main right diagonal. The plan was edited by rearranging the treatment numbers to conform with squares II and III plans given in Cochran and Cox (1957), page 498. The third and fourth

replicates were obtained by interchanging the row and column orientation. The resulting plan used as the input design for the RRC module is:

1	10	14	18	22
23	2	6	15	19
20	24	3	7	11
12	16	0	4	8
9	13	17	21	5
1	23	20	12	9
10	2	24	16	13
14	6	3	0	17
18	15	7	4	21
22	19	11	8	5
1	8	15	17	24
0	2	9	11	18
19	21	3	10	12
13	20	22	4	6
7	14	16	23	5
1	0	19	13	7
8	2	21	20	14

The output from the RRC module using the above input design is:

RRC 2.1: Construct a resolvable row-column design of size (r, k, s)
 (C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: Resolvable row-column design for $v=25$, $r=4$, $k=5$ and $s=5$.

```
try #           89
seed           1830982152
# of iterations 30
f             600
concurrences   1(200) 2(100)
E(column)     0.8182
E(row)        0.7863
E             0.6433
E/U           0.9292
```

Plan:

20	16	6	18	19
9	10	0	7	11
12	13	17	15	22
23	24	3	21	8
1	2	14	4	5
1	15	3	4	9
18	19	7	16	17
22	23	24	0	5
14	6	11	12	21

10	2	20	8	13
7	8	9	10	6
1	20	16	17	24
0	21	22	23	18
13	14	15	11	12
19	2	3	4	5
17	2	12	22	7
24	9	19	4	14
1	11	21	6	16
15	0	10	20	5
8	18	3	13	23

Note: RRC used 0.83 seconds.

The above procedure did not produce a lattice square design that contained numbers 0-4 in one diagonal and five other numbers, e. g. 5-9 or 10-14, in second diagonal. The following input design was also tried but again no lattice square design with the right conditions was obtained. Thus at this point it would appear that Gendex is not useful for constructing augmented lattice square designs.

0	21	17	13	9
5	1	22	18	14
10	6	2	23	19
15	11	7	3	24
20	16	12	8	4
0	22	16	10	5
6	1	23	17	11
12	7	2	21	18
19	13	8	3	20
24	15	14	9	4
0	22	19	12	6
7	1	23	15	13
14	8	2	24	16
17	10	9	3	20
21	18	11	5	4
0	23	18	14	7
8	1	24	19	10
11	9	2	20	15
16	12	5	3	21
22	17	13	6	4

AUGMENTED BLOCK EXPERIMENT DESIGNS

To obtain an augmented randomized complete block experiment design, the BIB module of Gendex is used to obtain a randomized plan for s checks or standards with r replications of

each check. Then the block sizes are enlarged to accommodate the n new treatments as shown in the following example.

Example 6-- The following output for $s = 3$ checks and $r = 5$ replicates was obtained using the BIB module of Gendex:

BIB 2.1: Construct incomplete block designs of size (v,k,r)
(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: Incomplete block design for $v=3$, $k=3$, $r=5$ and $b=5$.

```
try #           1
seed           1037721544650
# of iterations 0
E              1
E/U            1
concurrences   5(3)
```

Plan (Blocks are rows):

```
0      1      2
1      2      0
2      0      1
2      0      1
1      0      2
```

Note: BIB used 0.11 seconds.

Suppose that $n = 30$ new treatments are to be screened. Increase the block size to nine experimental units to include 6 new treatments in each of the five blocks. Using an X to designate where a new treatment is applied, the plan for the augmented randomized complete block experiment design is:

```
0 X X 1 X X 2 X X   X 1 X X 2 X X 0 X   X 2 X X 0 X X 1 X
X X 2 X X 0 X X 1   X 1 X X 0 X X 2 X
```

X denotes a new treatment and the checks are numbered 0, 1, and 2.

To obtain an augmented incomplete block experiment design, either of the BIB, ALPHA, or CIB modules of the Gendex toolkit may be used to obtain an incomplete block design for s treatments in incomplete blocks of size k with r replicates. This is demonstrated in the following example.

Example 7--Using the BIB module, obtain a randomized form of an incomplete block experiment design for $s = 8$, $k = 4$, and $r = 3$. A design obtained was:

BIB 2.1: Construct incomplete block designs of size (v,k,r)
(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: Incomplete block design for $v=8$, $k=4$, $r=3$ and $b=6$.

```

try #          1
seed          1037721720460
# of iterations 1
E             0.8235
E/U           1
concurrences  1(24) 3(4)

```

Plan (Blocks are columns):

```

1      5
3      7
4      2
0      6

```

```

4      5
7      0
3      6
2      1

```

```

6      1
5      7
3      0
4      2

```

Note: BIB used 0.05 seconds.

Enlarging the block size from $k = 4$ to 14 to accommodate 10 new treatments in each of the six incomplete blocks, we obtain the following plan where X denotes one of the 60 new treatments:

```

X X 1 X X 3 X X 4 X X 0 X X      X X 4 X X 7 X X 3 X X 2 X X
X 5 X X 7 X X 2 X X 6 X X X      X X 5 X X 0 X X 6 X X 1 X X

6 X X 5 X X 3 X X X 4 X X X
X X 1 X X 7 X X 0 X X 2 X X

```

COMMENT

The usefulness of Gendex to create augmented row-column and augmented block experiment designs is to obtain a randomized form of the design. For row-column designs the rows are randomized by the above procedure. The columns have been entered in a systematic fashion. The row randomization and random allocation of numbers to the treatments plus the random variation encountered in the lay-out of the experiment should suffice for statistical analyses. Gendex may be used to obtain the experiment design for the s check treatments but the construction of the augmented block design would then be accomplished manually. For the augmented block experiment designs, the block size is expanded to incorporate the n new treatments. Plant breeders would prefer to equally space the checks among the new treatments in the experiment for visual comparisons of a variety of characteristics.

LITERATURE CITED

Cochran, W. G. and G. M. Cox (1957). *Experimental Designs*. John Wiley & Sons, Inc., New York.

Federer, W. T. (1998). Recovery of interblock, intergradient, and intervariety information in incomplete block and lattice rectangle designed experiments. *Biometrics* 54(2):471-481.

Federer, W. T. (2002). Construction and analysis of an augmented lattice square design. *Biometrical Journal* 44(2):251-257.

Federer, W. T., B. Gross, N-K. Nguyen, and S. Nishinyabakobeje (2001a). GENDEX-Getting started and the BIB module. BU-1572-M in the Technical Report Series of the Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, New York 14853.

Federer, W. T., B. Gross, N-K. Nguyen, and S. Nishinyabakobeje (2001b). GENDEX-The ALPHA, CIB, RCD, and RRC modules. BU-1573-M in the Technical Report Series of the Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, New York 14853.

Federer, W. T. and R. D. Wolfinger (1998). SAS PROC GLM and PROC MIXED code for recovering inter-effect information. *Agronomy Journal* 90:545-551.

Federer, W. T. and R. D. Wolfinger (2002). PROC GLM and PROC MIXED codes for trend analyses for row-column designed experiments. In *Handbook of Formulas and Software for Plant Geneticists and Breeders* (Editor: M. S. Kang), Food Products Press, Binghamton, New York, pp (to appear).

Nguyen, Nam-Ky (2001). Gendex toolkit for generating design of experiments. <http://designcomputing.hypermart.net/Gendex>.