A TOOL FOR TEACHING, BUILDING AND COMPARING
INCOMPLETE BLOCK DESIGNS

Johannes van LEEUWEN\(^1\)
Victor NETO\(^2\)

- ABSTRACT: Options for the choice of an experimental design have drastically increased; with an appropriate computer program, today any design can be analyzed. Nevertheless, the incomplete block design (IBD) continues to be avoided, resulting in a less efficient use of research resources. One of the plausible reasons is the fact that the IBD’s treatment/block composition is not well understood. We introduce the concurrence table (CT), a simple tool which clarifies the IBD’s structure, showing for all treatment pairs how often the two treatments of a pair occur together in a block. We show the use of the CT: to reveal the features of the balanced and the partially balanced incomplete block design; to check a design for connectedness; and to compare IBDs. The CT is helpful in building block designs with a less regular pattern.

- KEYWORDS: Design of experiments; understanding incomplete blocks; concurrence matrix; connectedness; weighted concurrence.

1 Introduction

To obtain a low experimental error, researchers group experimental units (EUs) into homogeneous blocks. The most popular block design is Fisher(1926)’s randomized complete block design (RCBD), where each block contains one complete replication of the treatments. Experiments often compare many treatments. In these cases, an incomplete block design (IBD), with a well-chosen block size, gives a lower experimental error than the RCBD, as the RCBD will have large and consequently heterogeneous blocks. Nevertheless, the RCBD continues to be by far the most employed experimental design (NETO, 1988; HINKELMANN and KEMPTHORNE, 2008), utilized in perhaps 85% of the cases (MEAD, 1990a, p.579).

The very frequent use of the RCBD suggests that researchers choose this design too often, in detriment to resource optimization. The limited use of the IBD can have several reasons. Potential users may fear that they will not be capable of analyzing the data obtained from an IBD. Today that needs no longer be a problem, since computer programs can do the analysis, although data manipulation may be necessary. The IBD will also lack

\(^1\)Instituto Nacional de Pesquisas da Amazônia - INPA, Coordenação de Tecnologia e Inovação, Postal Box 2223, 69.080-971, Manaus, AM, Brasil. E-mail: johannes.leeuwen@gmail.com

\(^2\)Agro Biotechnology Institute, MARDI campus, 43400, Serdang, Selangor, Malaysia. E-mail: emirwait@gmail.com
popularity, as its block structure is not well understood. On the other hand, the structure of the RCBD is straightforward and intuitive: all blocks are equal.

Not understanding the IBD’s treatment/block composition makes it difficult to compare alternative IBDs. To solve these problems and to contribute to a more frequent employment of the IBD, we present the concurrence table (CT), a simple tool showing the IBD’s logical structure. We demonstrate the CT’s use: to reveal the characteristics of the balanced IBD and the partially balanced IBD; to demonstrate the concept of design connectivity; and to compare IBDs. The CT can help building the IBDs with a less regular pattern, advocated by Mead (1990a, 1990b). Our paper shows that conferring a central role to the CT is very useful in explaining IBDs to those not at ease with these unusual designs.

2 Concurrence table

A balanced incomplete block design (BIBD) has \(t\) treatments, \(r\) replications and \(b\) blocks of \(k\) EUs, with \(k < t\) and \(b > r\). The design in Figure 1 (numbers indicate blocks; capital letters treatments) is a BIBD with \(t = 7\), \(r = 4\), \(b = 7\), \(k = 4\) (COCHRAN and COX, 1957, plan 11.8). It is balanced as each pair of treatments appears the same number of times \((\lambda)\) in a block, implying that all two-treatment comparisons have the same variance.

The design of Figure 1 (left) is a collection of seven groups (blocks) of four treatments. Why are these blocks constituted the way they are? To clarify the design’s structure, an inventory is made of the occurrences of treatment pairs in the blocks. This is done with the help of the upper right triangle of a treatment-by-treatment table (Figure 1, right). For each treatment pair, the triangle has one cell. The six treatment pairs of block 1 are AB, AC, AF, BC, BF and CF. For each treatment pair a dot is placed in the corresponding cell. The same is done for the pairs of block 2 and so on. This leads to the table (Figure 1, right), which we call concurrence table (CT), as it shows how often the two treatments of a pair of treatments concur in a block.

<table>
<thead>
<tr>
<th>Block</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
</tr>
<tr>
<td>5</td>
<td>E</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
</tr>
<tr>
<td>7</td>
<td>G</td>
</tr>
</tbody>
</table>

![Figure 1](image)

Figure 1 - Balanced incomplete block design (BIBD) (left), corresponding concurrence table (CT) (right). Each occurrence of a treatment pair in a block corresponds to a dot in the corresponding cell of the CT. As the CT is symmetric in the diagonal, only the upper right triangle is shown. There are no elements on the diagonal: the cells of the diagonal do not match treatment pairs.
All treatment pairs got two dots. Hence the design of Figure 1 is balanced, with $\lambda = 2$. The CT shows a design’s structure and indicates the relative precision of the treatment comparisons (a larger $\lambda$ corresponds to a more precise comparison). In our experience, constructing this table in the classroom is a handy way to demonstrate why an IBD is structured the way it is. As an exercise, students are asked to randomize an IBD from a textbook and check the outcome for balance by building its CT.

The CT can also be obtained from the incidence matrix. The incidence matrix $N$ of a design with $t$ treatments and $b$ blocks has $t$ rows and $b$ columns and consists of elements $n_{ij}$ (the number of times treatment $i$ occurs in block $j$). By multiplying the incidence matrix $N$ with its transpose $N'$, the concurrence matrix $NN'$ is obtained. This square, $t \times t$, symmetric matrix, with off-diagonal elements $\lambda$, is important in the theory of IBDs (John, 1971, p.220). The CT corresponds to the non-diagonal elements of $NN'$. Figure 2 shows $N$ and $NN'$ of the design in Figure 1.

For an IBD with many blocks and treatments, it is tedious to obtain the CT by hand. In such a case, $N$ is written out block by block (column by column) in a computer worksheet, after which CT is obtained from $NN'$ by retaining only the non-diagonal elements.

### 3 Partially balanced incomplete block design

In a partially balanced incomplete block design (PBIBD), $\lambda$ has more than one value, mostly two. Figure 3 shows a PBIBD with design parameters $t = 9$, $r = 3$, $b = 9$, $k = 3$ (Cochran and Cox, 1957, p.453) and its CT. In this case the complete CT is given, as that is helpful in what follows.

To obtain a CT with a more regular pattern, which better shows the design’s properties, the CT of Figure 3 is reordered (permutated symmetrically), which leads to the CT of Figure 4. Reordering is done by placing similar columns side by side, while rows are maintained in the same order as columns (done effortlessly using a computer spreadsheet). [For a discussion of reordering, see Liiv (2010)].

---

**Figure 2** - Incidence matrix $N$ of the design in Figure 1 (left), corresponding concurrence matrix $NN'$ (right).

<table>
<thead>
<tr>
<th>Block</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>E</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>F</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>G</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>
Partial balanced incomplete block design (PBIBD) (left), corresponding CT (right). The blocks are grouped in 3 replications, numbered I, II and III.

The CT has empty cells ($\lambda = 0$) and cells with one dot ($\lambda = 1$). Treatment comparisons with $\lambda = 1$ (e.g.: comparing A with B) are estimated more precisely (have a lower variance) than those with $\lambda = 0$ (e.g.: comparing A with F). B is called a “first associate” to A, and F a “second associate” to A. Comparing the rows of the first associates A and B, it can be seen that they have three first associates in common (C, E, G), while A has two second associates (F, H) which are first associates to B, and B has two second associates (D, I) which are first associates to A. It is easily verified in the CT of Figure 4, that all pairs of first associates follow this pattern.

In Cochran and Cox’s (1957, secs.11.63a, 11.64a) calculation of the variances of a PBIBD, the terms “first associate” and “second associate” are interchanged. Consequently, the obtained variance for the comparison of first associates is larger than that for second associates, which is impossible. This never reported, twice occurring slip of the pen in the most popular handbook on experimental design (and in its translation into Spanish (COCHRAN and COX, 1965)) can be interpreted as a demonstration of the lack of popularity of IB designs.
The dots of the rows of the second associates A and F coincide, meaning that A and F have all their first associates (B, D, I, C, E, G) in common. The rows of A and F also have one empty cell in common, which corresponds to treatment H, second associate to A and F. Consequently the second associates A and F relate in the same way to the other (seven) treatments of the trial. In Figure 4, it is easily verified, that this is the case for every pair of second associates. The above-given presentation of the PBIBD’s features constitutes an attractive substitute for the textbook one given in Cochran and Cox (1957, p.453-455).

The CT of Figure 4 has three empty, disjoint, diagonal, 3x3 sub-tables. Each of them corresponds to a set of 3 second associates. The upper, empty 3x3 square is formed by the treatments A, F, H; the centre 3x3 square by B, D, I; and the lower 3x3 square by C, E, G. Adding the three blocks formed by these sets to the design of Figure 3, places dots in the three empty 3x3 sub-tables and transforms the PBIBD of Figure 3 into a BIBD of 12 blocks and 4 replications.

4 Connectedness

Treatments A and B are connected, if there is a sequence in which blocks and treatments alternate, which begins with A and ends with B. In such a sequence, the treatments, immediately preceding and following a block, occur both in that block (JOHN, 1971, p.233). In the design of Figure 3, the sequence A – 1 – C – 6 – F is one of the sequences connecting treatments A and F. If at least one pair of treatments is not connected, the design is said to be disconnected (HINKELMANN and KEMPTHORNE, 2005, p.14).

The concurrence matrix \( NN' \) of a disconnected design consists of disjoint, square, diagonal sub-matrices; the treatments within the sub-matrices being connected to each other (HINKELMANN and KEMPTHORNE, 2005, p.14). This property allows using the CT to check a design for connectedness. If a (reordered) CT cannot be subdivided into disjoint, diagonal sub-tables, the corresponding design is connected. This is demonstrated with the design of Figure 5.

Reordering the CT of Figure 5 results in the CT of Figure 6, which has two disjoint, diagonal sub-tables. Consequently, the design of Figure 5 is disconnected. The upper triangle of the reordered CT (Figure 6) represents a BIBD with \( \lambda = 2 \), formed by treatments A, C, E, H and blocks 1, 4, 7, 8; the lower triangle corresponds to a PBIBD (\( \lambda \) either 1 or 2) formed by the remaining blocks and treatments.

In plant breeding, an experiment of one replicate (\( r = 1 \)), laid out in incomplete blocks, is used to compare high numbers of progenies (sometimes even more than 1000). In this type of field trial, incomplete blocks with a number of plots close to 10 are commonly employed, while a control treatment is added to each block. The design of Figure 7 illustrates this setup for an experiment with control variety Z and 9 varieties (A, B, ..., I) to be tested in 3 blocks of 4 EUs. The CT (Figure 7, right) shows that the treatments of the design are connected by control treatment Z.
Figure 5 - Design \((t = 9, r = 3, b = 9, k = 3)\) (left); corresponding CT (right).

Figure 6 - Reordered CT of the design of Figure 5.

Figure 7 - Design with \(t = 10, r = 1\) (9 treatments), \(r = 3\) (control treatment Z), \(b = 3, k = 4\) (left); corresponding CT (right). As \(r\) has more than one value, a column for \(r\) has been added to the CT.
5 Indirect comparisons between treatments

When block effects are random and independent, block comparisons can give additional estimates of treatment differences, using the “second-order links” between treatments (JOHN, 1971, p.233-234). According to Mead (1990a, p.157) information on the relative importance of the second-order links is given by the non-diagonal elements of $(NN')(NN')$, a higher value indicating a stronger second-order link. As $NN'$ is symmetric, $(NN')(NN') = (NN')^2$, the square of concurrence matrix $NN'$.

Considerer the PBIBD with $t = 6$, $r = 4$, $b = 6$, $k = 4$ (MEAD, 1990a, example 7.3). This design has $t(t - 1)/2 = 15$ different treatment pairs. In $b$ blocks of size $k$ occur $bk(k - 1)/2 = 36$ treatment pairs, allowing to repeat all 15 different pairs twice and 6 different pairs a third time. The two designs fitting these parameters are shown in Figure 8 and 9 and are compared in Table 1. Their CT’s have both six treatment pairs with $\lambda = 3$ and nine with $\lambda = 2$, but 12 of the 15 values of the upper triangle of their matrices $(NN')^2$ are different.

![Figure 8](image1.png)

**Figure 8** - Design I (left); corresponding CT (middle); upper triangle of $(NN')^2$ (right). Design I is a PBIBD with $t = 6$, $r = 4$, $b = 6$, $k = 4$. Treatment pairs AB, AC, BC, DE, DF, EF occur three times within a block ($\lambda = 3$), the remaining pairs two times ($\lambda = 2$).

![Figure 9](image2.png)

**Figure 9** - Design II (left); corresponding CT (middle); upper triangle of $(NN')^2$ (right). Design II has the same parameters as design I (Figure 8), but a different set of treatment pairs with $\lambda = 3$, resulting in different values in $(NN')^2$. 

When all treatment pairs have the same interest, those with a lower value for $\lambda$ should get a relatively larger value in $(NN')^2$ (MEAD, 1990a, p.157). In such a situation, design II (Figure 9) is the better choice (although the differences are very small); its more homogeneous variances confirm this.

As second example, consider the 16 possible PBIBDs with $t = 9$, $r = 4$, $b = 18$, $k = 2$ (MEAD, 1990a, example 7.14), enumerated and discussed in more detail in Mead (1990b, p.169-174). These designs have 36 different treatment pairs, of which 18 form the 18 blocks of the trial. For all designs, the $CT$’s upper right triangle contains 18 cells with $\lambda = 1$ and 18 cells with $\lambda = 0$. The $(NN')^2$, on the other hand, varies for the different designs. The average of the elements of $(NN')^2$ of the 18 treatments pairs with $\lambda = 0$ assumes 9 different values and shows a negative relation with the average of the variances of the 36 two-treatment differences. Furthermore, for the 16 possible designs, the number of different values for the variance of a two-treatment difference varies from 2 to 21, assuming 11 different values. An appropriate design can be chosen by comparing $CT$ and $(NN')^2$ of the 16 designs [not shown in Mead (1990a, 1990b)]. The choice will depend on whether all treatment comparisons have the same importance, or a particular set of treatment comparisons is more important. Even if three or more sets of treatment comparisons vary in their relative importance, it may still be possible to choose the most adequate design.

### 6 Irregular block design

The BIBD and the PBIBD (Figures 1, 3, 8, 9) have a very regular pattern: all treatments are repeated the same number of times and appear only once in a block or not at all ($n_i$ is either 0 or 1), while all blocks have the same size (JOHN, 1971, p.220, 251). These designs were developed in the forties of the last century, when these restrictions were necessary, as calculation was done by hand with the help of a mechanical calculator. Exceptions were limited to simple departures from these models [e.g.: Figure 7; Federer(1955)’s augmented design; Justesen and Keuls (1958)].

In the meanwhile, possibilities for data analysis have increased dramatically. Nowadays, employing an adequate statistical program as GenStat, the data of any experimental design can be analyzed (MEAD, 1990a, p.6, 103). This is fortunate, because,
in numerous situations, the lowest experimental error will be obtained by a block design with a less regular pattern. Natural groups (blocks) can have varying numbers of units (e.g.: litter size in animal experimentation, groups of individuals of a species or agricultural variety menaced with extinction). Greenhouses, cages and laboratory runs can impose block size. We propose to call a block design with a less regular pattern: irregular block design (IrBD). A name for this design, which importance can be expected to increase, will be helpful in teaching.

Introductions to experimental design typically give no attention to the IrBD. Mead (1990a) and Mead et al. (2012) do this radically different. In their discussion of blocking, the IrBD is presented as a perfectly acceptable option and many examples of this design are given. [For a discussion of the pros and contras of their approach, see Mead (1990b).]

In their view it is no longer correct to adapt trials to a design found on lists of (P)BIBDs, as those given by Cochran and Cox (1957). If block size, number of treatments and number of replications do not automatically lead to a BIBD or PBIBD, a design is obtained by distributing the treatments over the available blocks. In such a case, the CT is useful in guiding the design’s construction. Reordering the CT might also be helpful.

In the case of the IrBD, the values of the CT no longer indicate in a completely accurate way, how “balanced” within-block treatment comparisons are. Correct values are given by the sums of weighted concurrences (MEAD, 1990a, p.190-191). The sum of weighted concurrences \( p_{ii'} \), of treatments \( i \) and \( i' \), is obtained from the following expression (MEAD, 1990a, p.142):

\[
p_{ii'} = \sum_j n_{ij}n_{ij'} / N_j
\]

The quantities \( n_{ij} \) and \( n_{ij'} \) indicate the number of times treatments \( i \) and \( i' \) occur in block \( j \). \( N_j \) is the total number of EUs of block \( j \).

The IrBD of Figure 10 illustrates this. Treatment pairs AB and BC occur both three times in a block, but the sum of weighted concurrences of AB is larger than that of BC. Consequently, comparison AB is estimated with more precision than comparison BC.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 10 - Irregular block design (IrBD) (left); corresponding CT (centre); table with sums of weighted concurrences (right).

7 Endnote

In this paper CT, \((NN')^2\) and sums of weighted concurrences are used for designs with one blocking system. They may also be employed for designs with two blocking systems (row and column designs). In such a case the tables \([CT, (NN')^2]\), sums of weighted
concurrences] for rows and columns can be added to get the full picture [as done in Mead (1990a, p.192) for the sums of weighted concurrences].

The devices discussed here allow building IBDs, comparing them and identifying the more appropriate candidates. To do this, it will be helpful to have a small computer program which, for a given design, produces $CT$, $(NN')^2$ and sums of weighted concurrences. [Users of programs as GenStat (MEAD, 1990b, p.190) and the free downloadable R can compare IBDs in a better way, employing the multiplicators of the error of the pairwise treatment comparisons.]

The $CT$ ignores an important structural design feature: resolvability. A design is resolvable, when its blocks can be arranged into replicates. The design of Figure 3 is resolvable; the one of Figure 1 is not. Resolvable block designs have two advantages. Block designs deal with two levels of random variation, resolvable block designs with three: (1) between EUs within blocks; (2) between blocks within replications; (3) between replications. If an IBD gets too messy (within-replicate allocation errors, many missing EUs), it can still be analyzed as RCBD.

Acknowledgements

The authors are grateful to an anonymous referee and to Professor Steven Gilmour, University of Southampton, and Pieter Vereijken, Biometris, Wageningen University & Research, who commented on an earlier version.


- RESUMO: As opções para a escolha de um delineamento experimental aumentaram drasticamente; hoje em dia dados de qualquer delineamento podem ser analisados, usando um programa de computação adequado. Mesmo assim, o delineamento em blocos incompletos (IBD) continua a ser evitado, o que resulta no uso menos eficiente dos recursos para a pesquisa. Uma das razões plausíveis é o não entendimento da distribuição dos tratamentos entre os blocos. Introduzimos o quadro da concorrência ($CT$), uma ferramenta simples que esclarece a estrutura do IBD, visualizando para todos os pares de tratamentos quantas vezes os dois tratamentos de um par ocorrem juntos num bloco. Mostramos o uso do $CT$: para conhecer as características do delineamento em blocos incompletos balanceados e em blocos incompletos parcialmente balanceados; para averiguar se um delineamento é conectado; e para comparar IBDs. O $CT$ ajuda na construção de delineamentos em blocos com um padrão menos regular.

- PALAVRAS-CHAVE: Delineamento experimental; compreender blocos incompletos; matriz de concorrência; conexidade; concorrência ponderada.

References


Received in 28.06.2014

Approved after revised in 24.09.2014