

The background of the cover is an aerial photograph of a large agricultural field, likely a cornfield, with rows of crops stretching towards the horizon. The lighting suggests a sunrise or sunset, with a warm glow on the left side. In the bottom right corner, there is a black and white portrait of a man with a prominent mustache, wearing a dark suit jacket, a white shirt, and a patterned tie. The overall composition is professional and academic.

NELDER STATISTICS

The Rothamsted approach to ANOVA

FORWARD

Nelder statistics is the Rothamsted approach to the analysis of variance. It differs from other approaches as it does not stress linear models but rather symmetry models and hypotheses for structured data. The covariance matrix, null analysis of variance and expectation of mean squares are all derived from the initial block structure describing the symmetries and pattern in the design. This structural approach evolved around John Nelder's theory of general balance (1965) and was implemented in the Wilkinson (1970) algorithm. This idea behind general balance, unlike other theories of experimental design, requires that the structure of treatments be specified separately from the dispersion structure (Payne et al, 1992). This approach is completed by a least squares procedure that consists of a sequence of sweeps as determined by the linear factorial model. This model is specified by way of the block and treatment structure. An extension of this procedure produces a multi-stratum analysis that preserves the distinction between the terms describing the underlying variability structure and those indicating the treatments applied. This is of special interest to designs with one or several error terms as commonly encountered in agricultural field experiments.

PURPOSE

The work of John Nelder, one of the most influential statisticians of his generation, is seldom mentioned in applied statistical courses in the United States. The concepts of **General Balance**, **Strata**, and the **Multi-Stratum ANOVA** are often unfamiliar to research agronomists. This is unfortunate as these ideas are fundamental to understanding experimental design and data analysis. This tradition is built upon methods developed and taught at Rothamsted Research. This philosophy is implemented in the GenStat ANOVA algorithm. The purpose of this document is a brief introduction to this subject matter. Those unfamiliar with this approach may gain an appreciation for the method and use it to become more efficient in the design and analysis of crop experiments. I contend that more practical skill is acquired in a few short trainings using Nelder's approach than that gained from in an entire semester of traditional course work. I was introduced to this method by the teaching methods and consulting of Mick O'Neill, technical assistance from Roger Payne, and the writings of Stephen Senn, Sue Welham and Emlyn Williams.

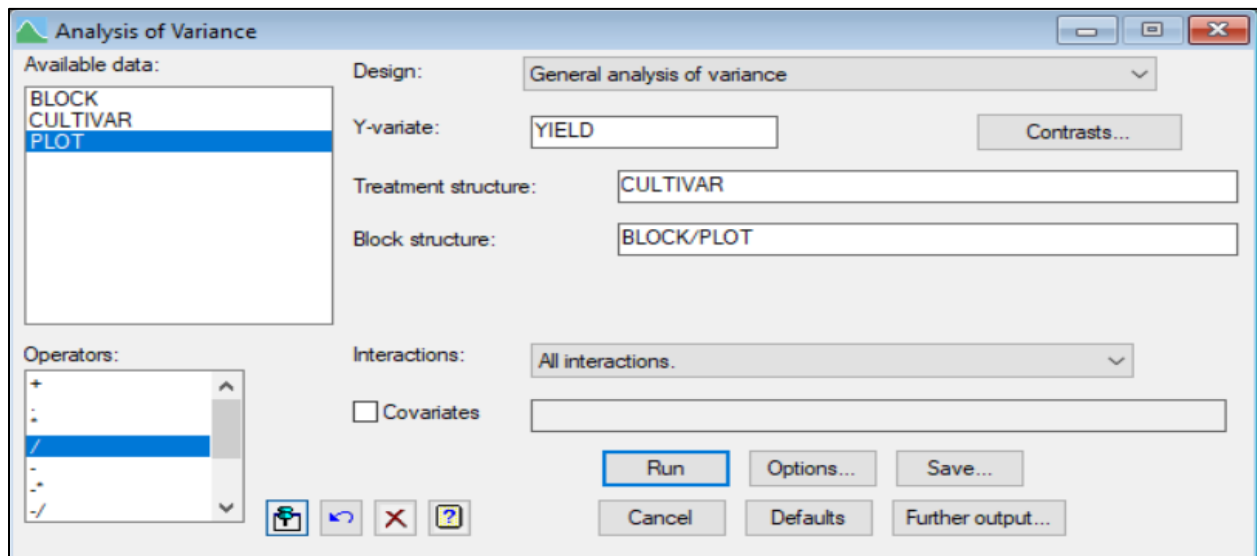
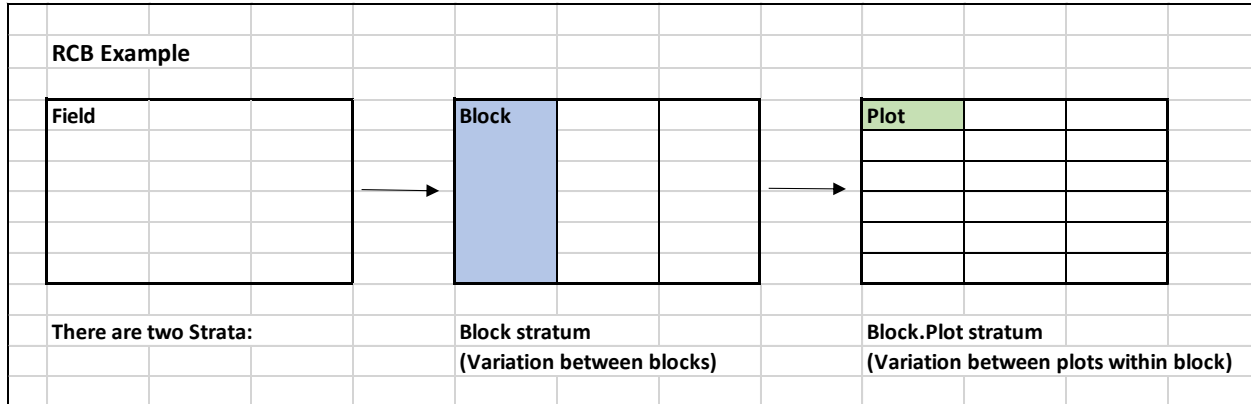
Nelder's concept of general balance combines the works of Fisher, Yates, Finney and others into a single framework so that any design which exhibits the properties of **general balance** can be described in terms of two formula. The theory enables the error structure of the experiment to be derived from the block structure. The treatment model, how treatments are applied to the experimental units, is specified by the treatment structure. The properties of general balance are; block terms are mutually orthogonal, treatment terms are mutually orthogonal, contrast of each term have equal efficiency factors in each stratum they are estimated. This idea covers a wide range of experimental designs (Payne, 1999).

Nelder proposed that based on this concept, a general computer program could be written which if given these two structures would extrapolate the correct analysis of variance, contrasts, and errors for any experiment exhibiting the properties of general balance. Wilkinson completed the implementation of this concept into a general balance algorithm (Senn, 2019).

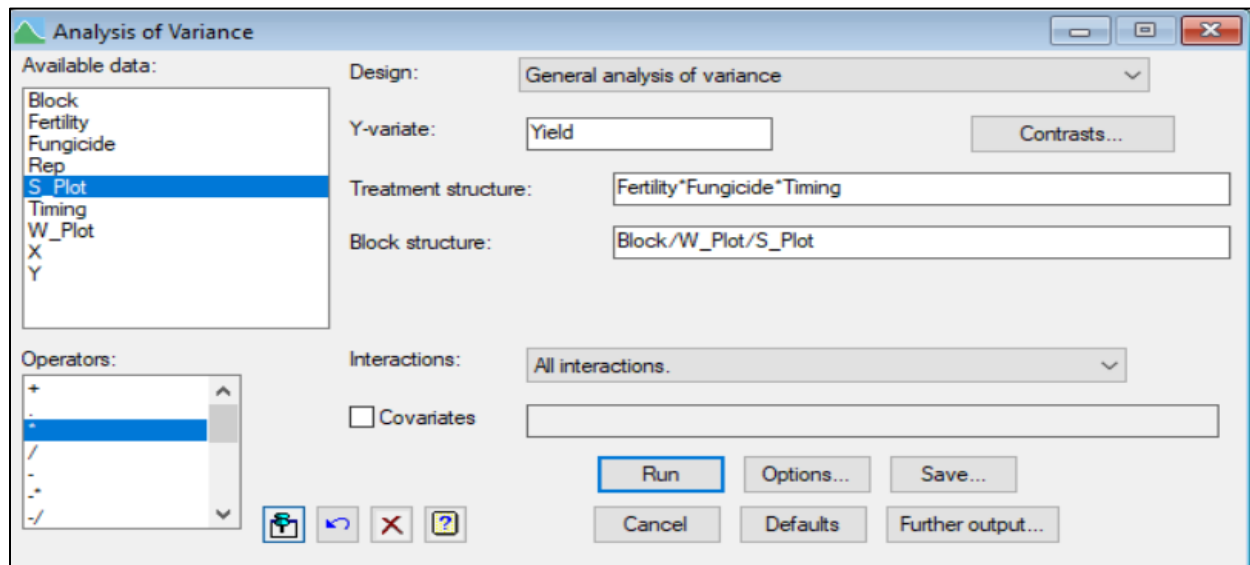
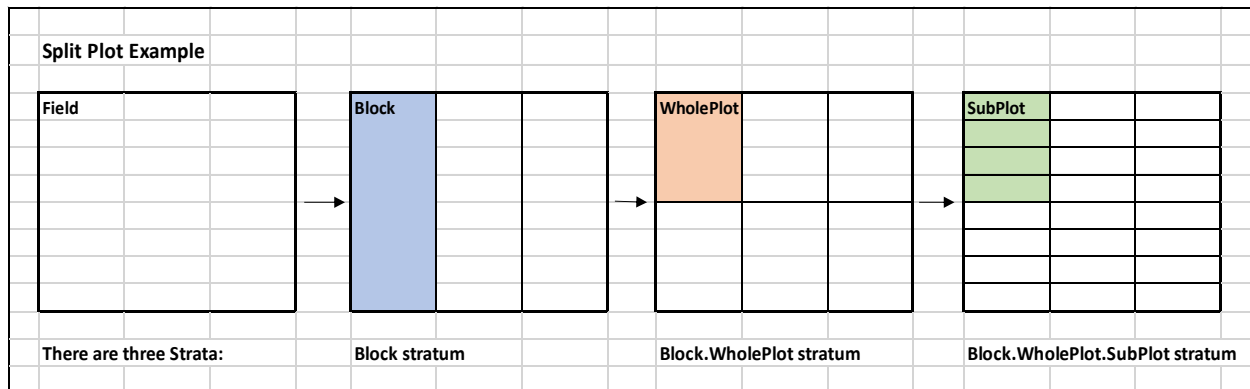
To complete the analysis the user, supplies the data for analysis, associated **indexing** information, and specification of the model. The model is specified symbolically with operators. This is a factor approach to the analysis of data which uses a **sweeping** method in contrast to the regression model method used by most programs. In a sweep method the sum squares and effects are calculated on residuals, not directly on the original yields (Mead et al, 2012, Wilkinson, 1970 and Payne et al, 1992).

The Wilkinson algorithm exploits the properties of balance in experimental designs to provide a uniquely efficient analysis that allows for different sources of random variation or error **strata** (Payne, 2018). The error strata are printed to an intuitive **multi-stratum ANOVA** table. The advantage of this is the recognition between the block and treatment structure so that treatment effects are always allocated to the correct strata so appropriate variances are calculated (Welham et al, 2015). Consequently, if you specify the block structure correctly the treatment terms get tested at the correct level of structure.

This approach is incorporated into the GenStat ANOVA. Some will argue it is the superior or correct approach to data analysis for stratified designs and will prevent many errors by both the professional and novice.



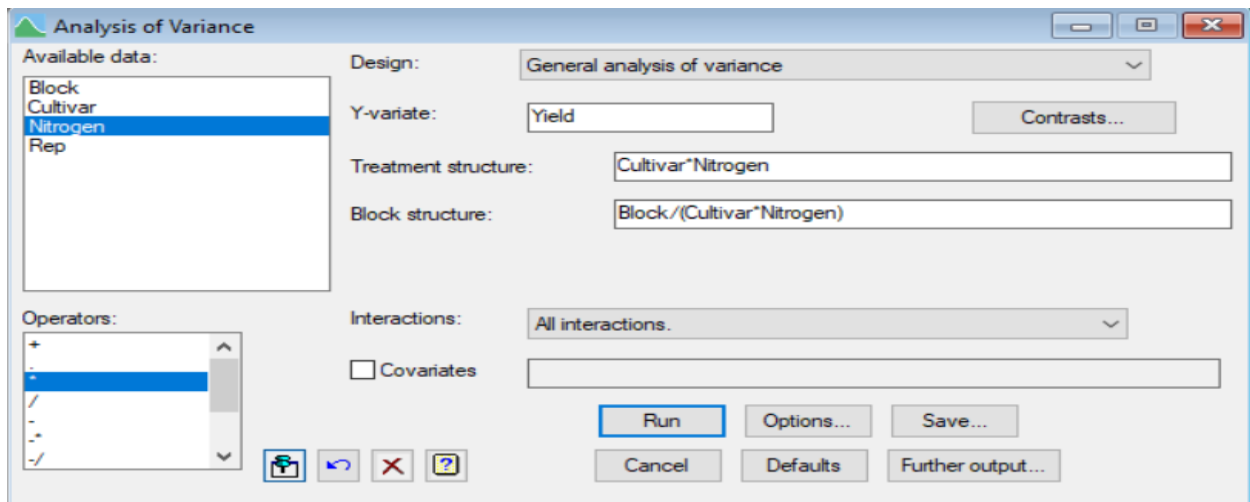
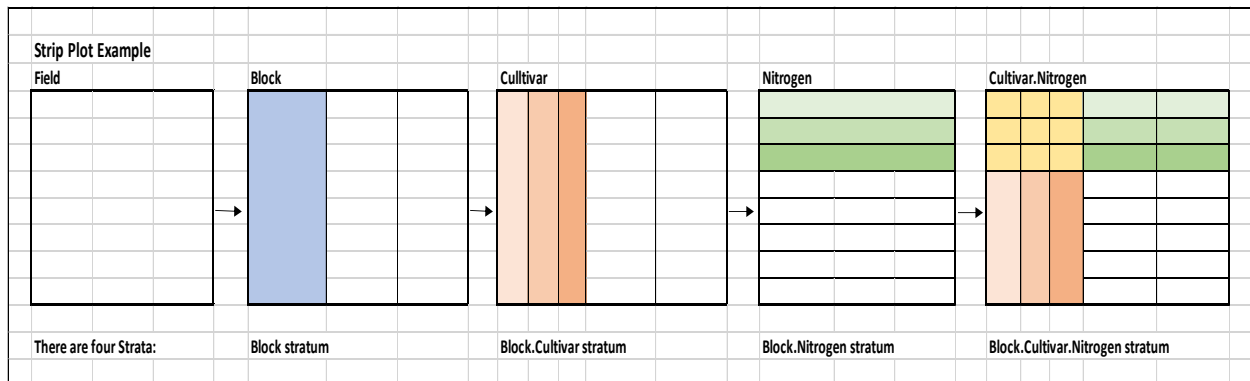
Analysis of variance					
Variate: YIELD					
Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
BLOCK stratum	2	410.47	205.23	10.77	
BLOCK.PLOT stratum					
CULTIVAR	20	2299.83	114.99	6.04	<.001
Residual	40	762.11	19.05		
Total	62	3472.41			



Analysis of variance


Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	5	659.77	131.95	4.26	
Block.W_Plot stratum					
Fertility	1	1328.21	1328.21	42.85	0.001
Residual	5	154.99	31.00	1.10	
Block.W_Plot.S_Plot stratum					
Fungicide	1	649.25	649.25	23.01	<.001
Timing	1	111.55	111.55	3.95	0.056
Fertility.Fungicide	1	47.05	47.05	1.67	0.206
Fertility.Timing	1	5.45	5.45	0.19	0.663
Fungicide.Timing	1	155.42	155.42	5.51	0.026
Fertility.Fungicide.Timing	1	0.50	0.50	0.02	0.895
Residual	30	846.49	28.22		
Total	47	3958.67			



Analysis of variance					
Variate: Yield					
Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	3	257.56	85.85		
Block.Cultivar stratum					
Cultivar	3	1282.19	427.40	76.64	<.001
Residual	9	50.19	5.58	0.72	
Block.Nitrogen stratum					
Nitrogen	3	761.31	253.77	33.07	<.001
Residual	9	69.06	7.67	1.00	
Block.Cultivar.Nitrogen stratum					
Cultivar.Nitrogen	9	237.94	26.44	3.44	0.006
Residual	27	207.69	7.69		
Total	63	2865.938			

Sweep

	The method of sweeping (Pearce et al, 1998)							
	A body of data is swept when each value is reduced by an appropriate mean.							
	Block	TRT	Data	TRT	Data	TRT	Data	<u>Block mean</u>
	1	A	22	B	17	C	21	20
2	C	26	A	17	B	23	22	
3	C	20	B	15	A	22	19	
1. Sweep by block means.								
<i>This gives the following quantities called deviations.</i>							(sum to zero)	
Block	TRT	Data	TRT	Data	TRT	Data	<u>Block Mean</u>	
1	A	2	B	-3	C	1	0	
2	C	4	A	-5	B	1	0	
3	C	1	B	-4	A	3	0	
The effect of block has now been removed.								
This value represents the effect of treatment and error .								
2. Calculate the new treatment means (after sweeping by block means)								
TRT	Block	Data	Block	Data	Block	Data	<u>TRT Mean</u>	
A	1	2	2	-5	3	3	0	
B	1	-3	2	1	3	-4	-2	
C	1	1	2	4	3	1	2	
3. Sweep by treatment means.								
<i>These quantities are called residuals.</i>							(sum to zero)	
Block	TRT	Data	TRT	Data	TRT	Data	<u>Block Mean</u>	
1	A	2	B	-1	C	-1	0	
2	C	2	A	-5	B	3	0	
3	C	-1	B	-2	A	3	0	
This value represents only the action of error.								

Indexing (Randomized blocks example)

Generate a Standard Design

Design: One-way design (in randomized blocks)

Design factor: _____ Name: _____ Number of levels: _____

Blocks: Blocks 4

Units within blocks: Plots

Treatment factor: Cultivar 8

Replications required... Check power...

Options

Randomize design Number of units: 32

Display design in a spreadsheet Randomization seed: 1555

Run Cancel Options... Defaults

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

Indexing (General Split-plot example)

Generate a Standard Design

Design: General split-plot design

No. of treatment factors: 3 No. of sub-plot treats: 2

Design factor Name Number of levels

Blocks: Blocks 6

Whole plots: W_Plots

Sub-plots: S_Plots

Whole plot treat. factor 1: Fertility 2

Sub-plot treat. factor 1: Fungicide 2

Sub-plot treat. factor 2: Timing 2

Replications required... Check power...

Options

Randomize design Number of units: 48

Display design in a spreadsheet Randomization seed: 21537

Run Cancel Options... Defaults

Row	PlotNo	Block	W Plots	S Plots	Fertility	Fungicide	Timing
1	101	1	1	1	2	2	1
2	102	1	1	2	2	1	1
3	103	1	1	3	2	2	2
4	104	1	1	4	2	1	2
5	105	1	2	1	1	2	2
6	106	1	2	2	1	1	2
7	107	1	2	3	1	2	1
8	108	1	2	4	1	1	1
9	201	2	1	1	2	1	2
10	202	2	1	2	2	2	2
11	203	2	1	3	2	2	1
12	204	2	1	4	2	1	1
13	205	2	2	1	1	2	2
14	206	2	2	2	1	1	2
15	207	2	2	3	1	2	1
16	208	2	2	4	1	1	1
17	301	3	1	1	2	1	2
18	302	3	1	2	2	2	2
19	303	3	1	3	2	2	1
20	304	3	1	4	2	1	1
21	305	3	2	1	1	2	1
22	306	3	2	2	1	2	2
23	307	3	2	3	1	1	2
24	308	3	2	4	1	1	1
25	401	4	1	1	1	1	1
26	402	4	1	2	1	2	1
27	403	4	1	3	1	2	2
28	404	4	1	4	1	1	2
29	405	4	2	1	2	2	1
30	406	4	2	2	2	1	2
31	407	4	2	3	2	1	1
32	408	4	2	4	2	2	2

Indexing (Strip-plot example)

Generate a Standard Design

Design: Strip-plot (or criss-cross) design

No. of treatment factors: 2 No. of column treats: 1

Design factor Name Number of levels

Blocks: Blocks 4

Rows: Rows

Columns: Columns

Row treat. factor 1: Cultivar 3

Column treat. factor 1: Nitrogen 3

Replications required... Check power...

Options

Randomize design Number of units: 36

Display design in a spreadsheet Randomization seed: 25713

Run Cancel Options... Defaults

Row	PlotNo	Blocks	Rows	Columns	Cultivar	Nitrogen
1	1	1	1	1	1	3
2	2	1	1	2	1	2
3	3	1	1	3	1	1
4	4	1	2	1	3	3
5	5	1	2	2	3	2
6	6	1	2	3	3	1
7	7	1	3	1	2	3
8	8	1	3	2	2	2
9	9	1	3	3	2	1
10	10	2	1	1	2	3
11	11	2	1	2	2	1
12	12	2	1	3	2	2
13	13	2	2	1	1	3
14	14	2	2	2	1	1
15	15	2	2	3	1	2
16	16	2	3	1	3	3
17	17	2	3	2	3	1
18	18	2	3	3	3	2
19	19	3	1	1	1	1
20	20	3	1	2	1	3
21	21	3	1	3	1	2
22	22	3	2	1	2	1
23	23	3	2	2	2	3
24	24	3	2	3	2	2
25	25	3	3	1	3	1
26	26	3	3	2	3	3
27	27	3	3	3	3	2



AGRO-TECH
RESEARCH FARM

Curt Lee
Agro-Tech, Inc.
4489 Highway 41 N
Velva, North Dakota, USA
www.agrotechresearch.com