ANOVA & REML

A GUIDE TO LINEAR MIXED MODELS IN AN EXPERIMENTAL DESIGN CONTEXT

Mick O'Neill

STatistical Advisory & Training Service Pty Ltd

Last updated August 2010

Introduction

In recent years a general algorithm, Restricted Maximum Likelihood (REML) has been developed for estimating variance parameters in linear mixed models (LMM).

This manual will review classic statistical techniques (ANOVA & REGRESSION) and demonstrate how LMM (REML) can be used to analyse normally distributed data from virtually any situation. For balanced data, REML reproduces the statistics familiar to those who use ANOVA, but the algorithm is *not* dependent on balance. It allows for spatial and/or temporal correlations, so can be used for repeated measures or field-correlated data. Unlike ANOVA, REML allows for changing variances, so can be used in experiments where some treatments (for example different spacings, crops growing over time, treatments that include a control) have a changing variance structure. The statistical package GenStat is used throughout. The current version is 13, although the analyses can generally be performed using the Discovery Edition released in 2010.

We have not separated the LMM (REML) section from ANOVA in this manual. The reason is clear. ANOVA is an appropriate analysis for a model

$$Yield = mean + fixed\ effects + random\ effects$$

where the random error terms are normal, independent, each with constant variance. This model includes simple random sampling (there are no random effects), regression, t tests and analysis of variance F tests.

LMM (REML) is also appropriate analysis for a model

$$Yield = mean + fixed\ effects + random\ effects$$

where the random error terms are normal, possibly correlated, with possibly unequal variances. The algorithm does not insist on balanced data, unlike ANOVA.

In general, data from two familiar text books will be used as examples. The editions we used are the following.

Snedecor, G.W. and Cochran, W.G. (1980). Statistical Methods. Seventh Edition. Ames Iowa: The Iowa State University Press.

Steel, R.G.D. and Torrie, J.H. (1980). Principles and Procedures of Statistics: a Biometrical Approach. Second Edition. New York: McGraw-Hill Kogakusha.

Several examples were kindly supplied by Curt Lee (Agro-Tech, Inc., Velva, North Dakota, USA). Other sources for data include:

Cochran, W. and Cox, G. (1957). Experimental Designs. Second Edition. Wiley 1957.

Diggle, P.J. (1983). Statistical Analysis of Spatial Point Patterns. London: Academic Press.

McConway, K. (1950). Statistical modelling using GENSTAT / K.J. McConway and M.C. Jones, P.C. Taylor. London : Arnold in association with the Open University.

Mead, R. and Curnow, R.N. (1990). Statistical methods in agricultural and experimental biology. Chapman and Hall, London.

Pearce, S.C. (1976). Field experimentation with fruit trees and other perennial plants. Second Edition. Farnham Royal: Commonwealth Agricultural Bureaux.

Reynolds, P.S. (1994). Time-series analyses of beaver body temperatures. In *Case Studies in Biometry*. N. Lange, L. Ryan, L. Billard, D. Brillinger, L. Conquest and J. Greenhouse (editors), 211–228. New York: John Wiley.

Schabenberger, O. and Pierce, F.J. (2001). Contemporary statistical models for the plant and soil sciences.

Sokal, R.R. and Rohlf, F.J. (1995). Biometry. The Principles and Practice of Statistics in Biological Research. Third Edition. New York: W.H Freeman and Company.

The training manual was prepared by Mick O'Neill from the **Statistical Advisory & Training Service Pty Ltd**. Contact details are as follows.

Mick O'Neill mick@stats.net.au
Statistical Advisory & Training Service Pty Ltd
www.stats.net.au

Table of Contents

Estimation and modelling

Whenever we conduct an experiment, no matter how complex, the analysis we perform always relates to way we set up the experiment: if we vary our methods, we vary the type of analysis we perform.

Moreover, the analysis we perform is always associated with an underlying model that involves any factors in the experiment and includes any random terms (like experimental error).

In this manual we will demonstrate these concepts starting from the most simple random sampling, and show that linear mixed models (LMM) with a residual maximum likelihood (REML) algorithm is a general model with an associated analysis that includes regression, time series and analysis of variance (ANOVA) as special cases.

Random samples from a single treatment or group

Example 1 Coefficients of digestibility of dry matter, fed corn silage, in percent (Steel and Torrie, page 93) fed to randomly selected sheep

Sheep	57.8	56.2	61.9	54.4	53.6	56.4	53.2

We are clearly interested in estimating the *mean* coefficient of digestibility for sheep, μ, hoping that these $n = 7$ randomly chosen sheep are representative of the entire population. We are also interested in estimating the *variation* in coefficients of digestibility, expressed say as a variance, σ^2.

Assume now that the coefficient of digestibility, Y, is normally distributed, ie $Y \sim N(\mu, \sigma^2)$. Then the simple model is that for each randomly chosen sheep, its coefficient of digestibility will differ from the mean value μ only by a random amount, which is what we call the error. The errors for the 7 sheep are all assumed independent,

The model for this random strategy is simply

$$Y = \textit{coefficient of digestibility} = \mu + \textit{Error}$$

where $\textit{Error} \sim N(0, \sigma^2)$. The parameter μ is a fixed parameter, and the parameter σ^2 is the only parameter in the random part of the model.

Immediately we have a special case of a general model

$$Y = \textit{fixed parameters} + \textit{random effects}$$

where the only fixed parameter is μ. Alternatively, we can pull μ out and express the model as

$$Y = \mu + \textit{fixed effects} + \textit{random effects}$$

where in this case there are no additional fixed effects (like possible breed effects which make the mean coefficient of digestibility different across breeds).

Maximum likelihood (ML)

Parameters of distributions are often estimated using the technique of *maximum likelihood* (ML) *estimation*. This technique maximizes what is known as the likelihood, though it is equivalent, and often easier, to maximize the log-likelihood. For the normal population, the likelihood of a random sample of size n is simply the product of the density function of the normal distribution evaluated at each of the data points. The log-likelihood is therefore

$$\log L = -\frac{n}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{Y_i - \mu}{\sigma}\right)^2 .$$

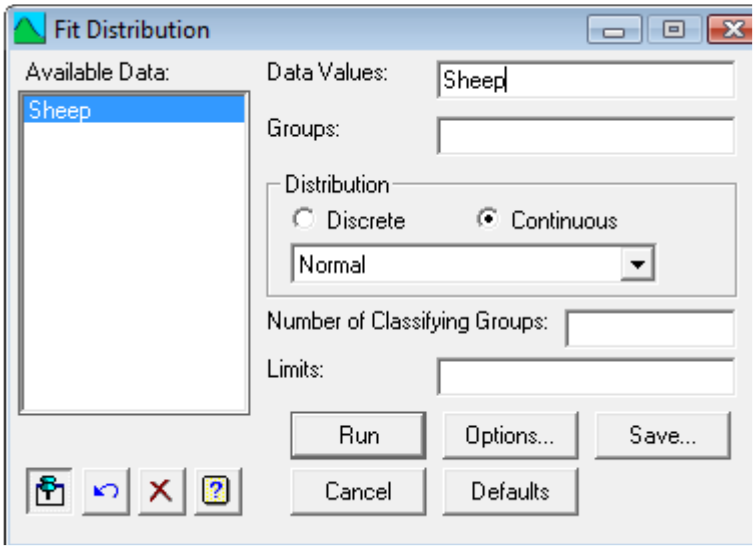
It is straightforward (mathematically) to show that the ML estimators of μ and σ^2 are

$$\hat{\mu} = \bar{y}\,, \qquad \hat{\sigma}^2_{ML} = \frac{\sum_{i=1}^{n}(Y_i - \bar{y})^2}{n} = s_n^2 .$$

Maximum likelihood estimators do not necessarily have optimal small-sample properties. It is true that the ML estimate of σ^2 is biased, in the sense that the mean over repeated sampling settles down on the value $(n\text{-}1)/n \times \sigma^2$ rather than on σ^2 itself.

For these data, the ML estimates are $\hat{\mu} = 56.214$, $s_n^2 = 7.727$, $s_n = 2.780$.

Early monographs such as Steel and Torrie and Snedecor and Cochran introduced the idea of estimating parameters like the mean μ and standard deviation σ of a normal population without reference to the concept of maximum likelihood. They used n as a divisor of the variance estimate rather than $(n\text{-}1)$. To justify this, they talk about bias or sampling with and without replacement. Some authors talk about using n as the divisor when calculating the *population* variance and $(n\text{-}1)$ when calculating the *sample* variance. Indeed, scientific calculators have σ_n and σ_{n-1} buttons. Excel has VARP and VAR formulae for the two sorts of variances (which we label s_n^2 and s_{n-1}^2 respectively), and STDEVP and STDEV for the equivalent standard deviations.



GenStat has a menu (Stats > Distributions > Fit Distributions…) that allows various distributions to be fitted to data. Maximum likelihood estimation is used in this menu to fit the parameters of these distributions. As can be seen, one simply indicates the data to be used and selects the distribution to be fitted. The number of classifying groups and the limits are optional (for controlling the number and positions of cut-points).

Fit continuous distribution

Sample statistics

Sample Size	7
Mean	56.21
Variance	9.01
Skewness	0.84
Kurtosis	-0.56

Quartiles:

25%	50%	75%
53.6	55.4	54.0

Summary of analysis

Observations: Sheep
Parameter estimates from individual data values
Distribution: Normal (Gaussian)
X distributed as Normal(m,s**2)
Deviance: 0.21 on 0 d.f.

ML estimate of μ

ML estimate of σ

Estimates of parameters

	estimate	s.e.	correlations	
m	**56.2143**	1.0510	1.0000	
s	**2.7798**	0.7435	0.0000	1.0000

Residual maximum likelihood (REML)

The idea of residual maximum likelihood (REML) is only a couple of decades old. The idea is this:

We take the likelihood and partition it into two components. The first component is a likelihood of one or more statistics and involves all *fixed* parameters like μ (and may involve variance parameters as well). The second component is a *residual* likelihood and involves only the variance parameters of the *random* effects. *We then maximize each component separately.* The estimates of the variance parameters are known as REML estimates.

For samples from a normal population, the first component turns out to be the likelihood for the sample mean $\bar{y}$, the second likelihood is that of variates associated with the sample variance. Specifically,

$$\log L = \left[-\frac{1}{2}\ln(2\pi\sigma^2/n) - \frac{1}{2}\left(\frac{\bar{y}-\mu}{\sigma/\sqrt{n}}\right)^2\right] + \left[-\frac{n-1}{2}\ln(2\pi\sigma^2) - \frac{1}{2}\ln(n) - \frac{1}{2}\sum\left(\frac{Y_i-\bar{y}}{\sigma}\right)^2\right]$$

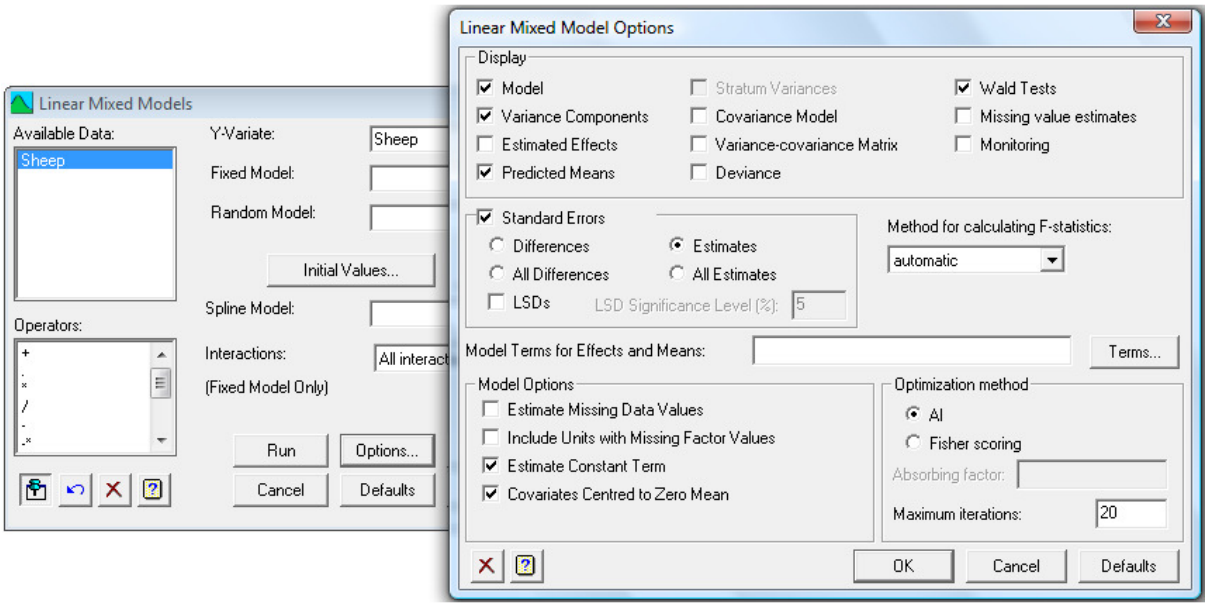
involves μ (and, unimportantly, σ) — involves σ only (not μ)

The separate solutions are

$$\hat{\sigma}^2_{REML} = \frac{\sum_{i=1}^{n}(Y_i - \bar{y})^2}{n-1} = s^2_{n-1}, \qquad \hat{\mu} = \bar{y}.$$

Thus, the familiar estimate for σ^2 is actually a REML estimate, $s^2_{n-1} = 9.015$, and this estimate is unbiased. For more complex models, the REML estimate is less biased than the ML estimate.

For the sheep data, REML estimates are available using the menu Stats > Mixed Models (REML) > Linear Mixed Models... In this menu GenStat will always fit a constant term (μ) and, if you do not include an error term, it will add one for you. Simply enter the coefficient of digestibility column as the **Y-variate** and leave the **Fixed Model** and **Random Model** blank. We need to click **Predicted Means** in **Options**, and as a general rule, click **Deviance** as well.



REML variance components analysis

Response variate: Sheep
Fixed model: Constant
Number of units: 7

Residual term has been added to model

Sparse algorithm with AI optimisation

REML estimate of σ^2

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
Residual		Identity	Sigma2	9.015	5.205

Table of predicted means for Constant

ML/REML estimate of μ

56.21 Standard error: 1.135

se of mean = s/√*n* - uses REML estimate of σ

Notice in the output that a "Residual term has been added to model". We can deliberately put an error term if we wish (for example, if we decide to include a correlation into our model). For a sample of size *n* there are *n* error terms, each being independent with the same distribution, $N(0, \sigma^2)$. We therefore need to set up a factor that contains *n* levels corresponding to the *n* data values. In this case we would set up a factor column with levels 1, …, 7 called say Replicate and use Replicate as the **Random Model**. Alternatively, GenStat has an in-built device to do this: simply type **'*Units*'** in the **Random Model**.

Deviance

Selecting the option **Deviance** produces this additional information:

Deviance: -2*Log-Likelihood

Deviance	d.f.
21.14	5

Note: deviance omits constants which depend on fixed model fitted.

Deviance plays the role that the Residual SS plays in ANOVA. The deviance that GenStat prints out is proportional to -2×LogL, where LogL is the log-likelihood of the variance components. (The actual definition actually has the constant 2π removed):

Deviance really is only used to compare models where the null hypothesis involves the variance parameter of a random effect. Asymptotically, a *change in deviance* for one (nested) model compared to a larger model follows a χ^2 distribution, and the degrees of freedom to use are the *change in df*. The nested model arises by replacing in the larger model the new parameters that are given in the null hypothesis.

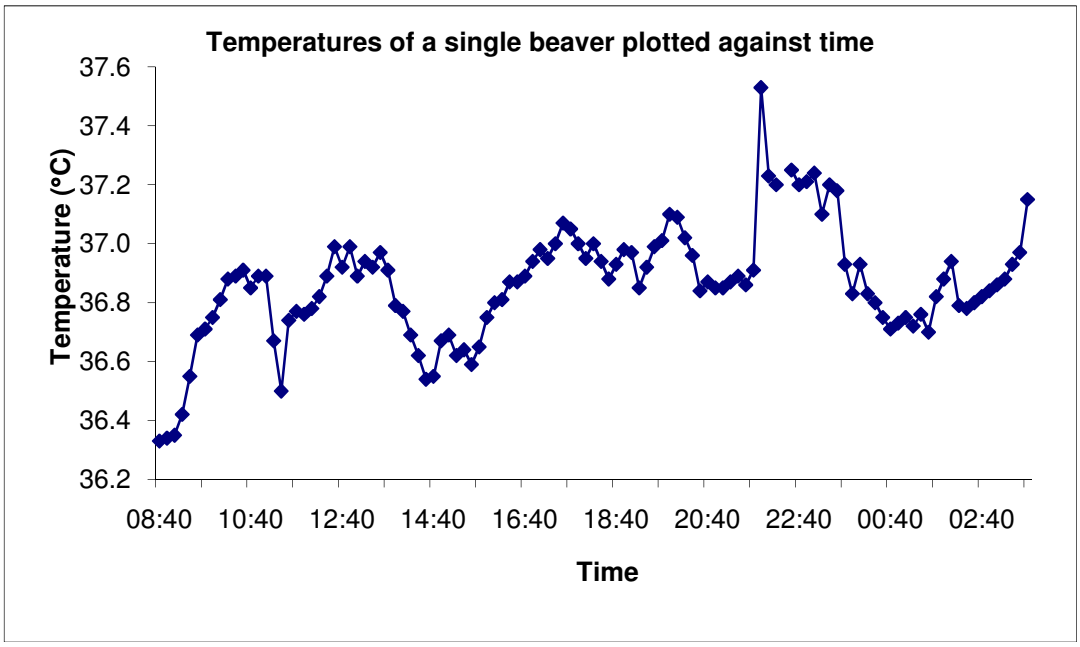
Correlated samples

Using a REML algorithm in experiments involving fixed effects and random effects is not restricted to independent data, or to data with the same variance in any one stratum. It is an extremely flexible estimating tool, and has become the standard way of analyzing data from agricultural trials.

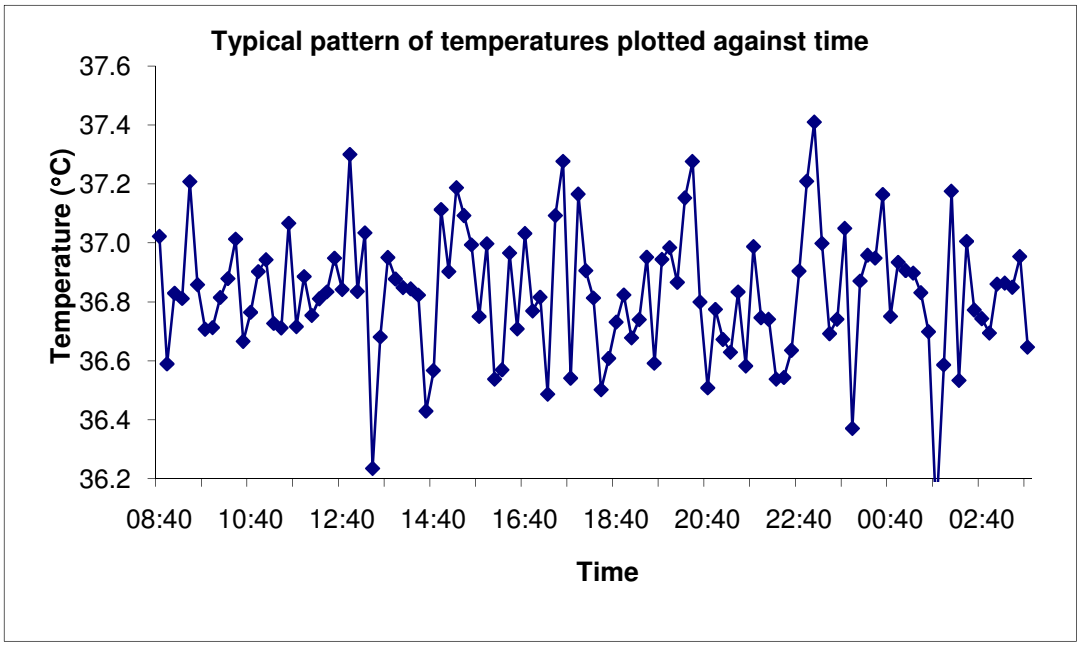
This manual is not a place to describe in great detail the concepts of correlated data over time. At this point all we want to do is demonstrate that very often we need to analyze data that is serially correlated.

A good example to illustrate serially correlated data is the famous beaver body temperatures taken every 10 minutes, taken from *Case Studies in Biometry* (Lange *et al*. 1994). A plot of these temperatures for a single animal is shown on the left hand page, and for comparison, a plot of notional temperatures randomly sampled from a normal distribution at each time with the same mean and variance as the overall beaver temperatures had. It is clear that there is an essential difference between the two plots.

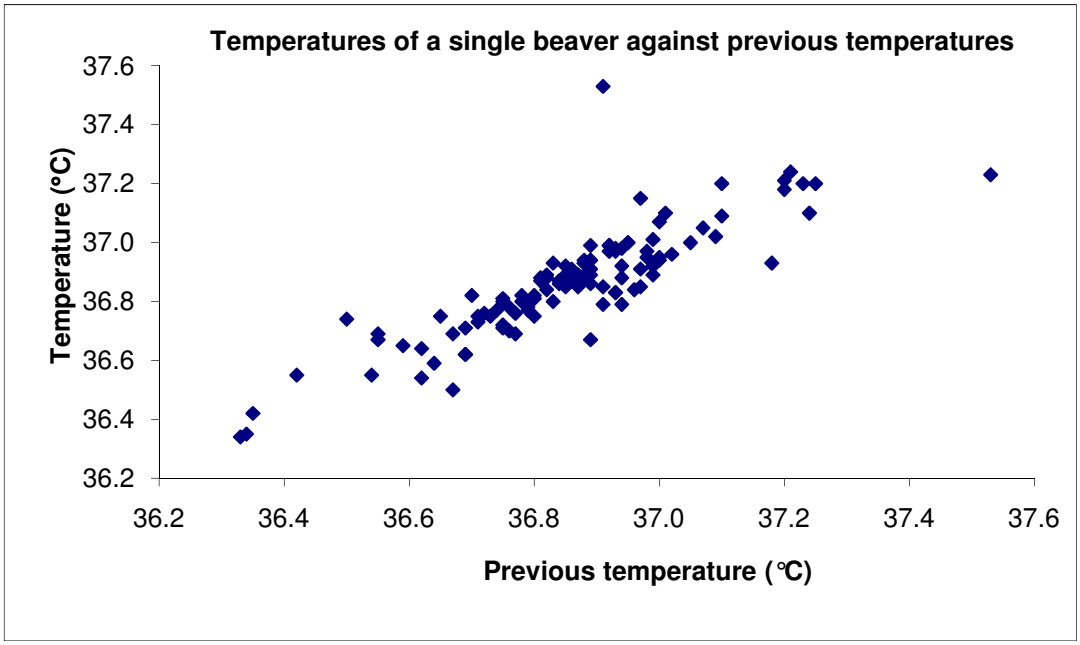
Plot of temperatures of a single beaver every ten minutes

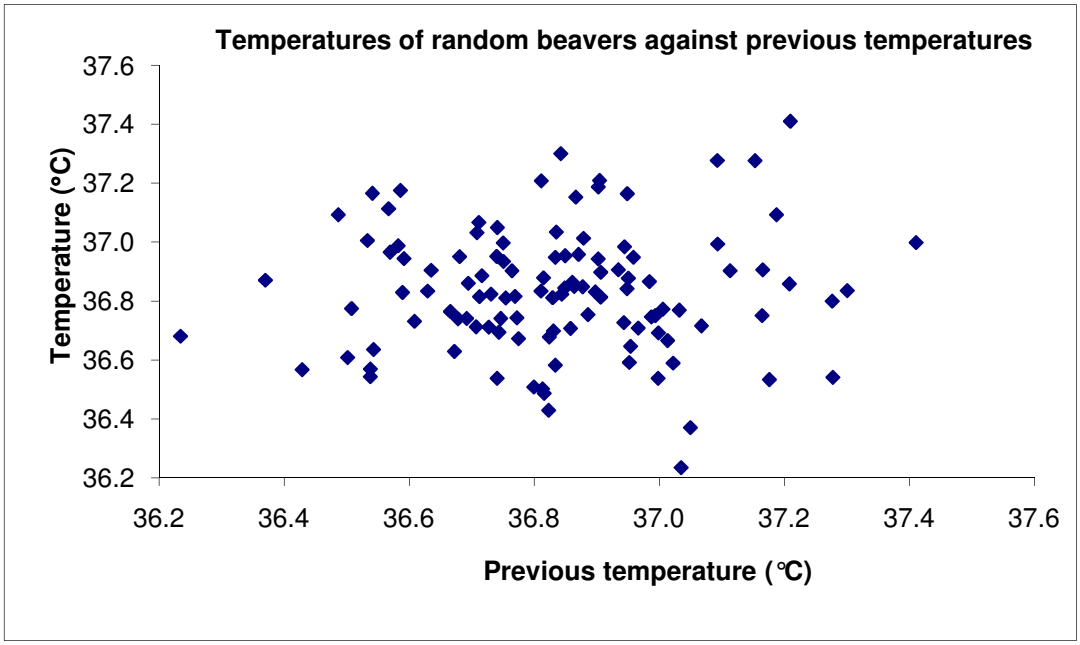


Notional plot of temperatures of beavers randomly selected every ten minutes



To emphasize the difference even more strongly, here are plots of the temperatures at time *t* plotted against the temperatures at time *t*-1.



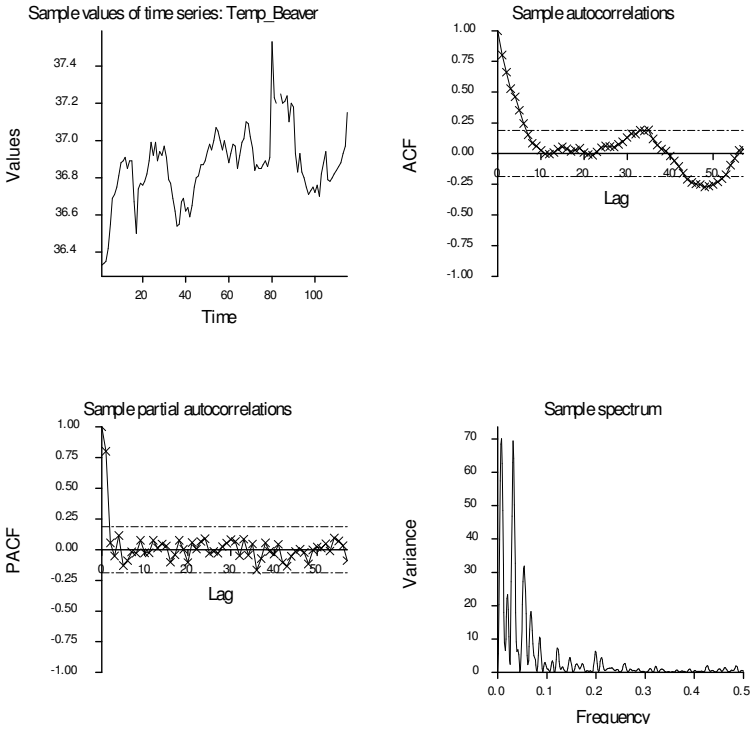


The temperatures of a single beaver are clearly correlated in time: we call this a *serial correlation*. The model is the same as the previous model for coefficients of digestibility, only the assumptions underlying the model are different:

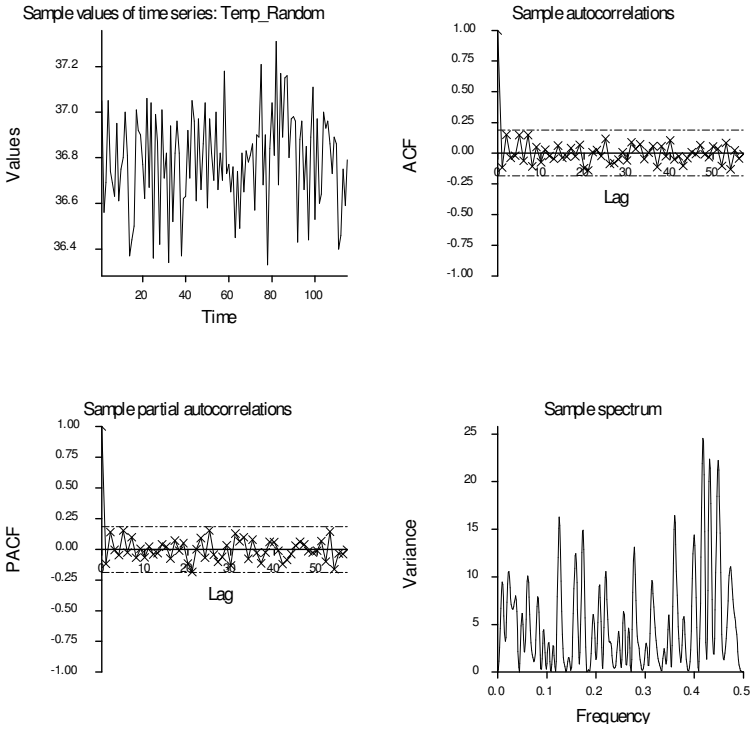
$$Y = \textit{Temperature of a beaver} = \mu + \textit{Error}$$

where *Error* ~ N(0, σ^2), however some correlation structure exists among the individual error terms. This is the subject of *time series analysis*.

Time series plots for beaver data



Time series plots for random data with same mean and standard deviation



Example 2 Temperatures of a single beaver taken every 10 minutes (left to right)

36.33	36.34	36.35	36.42	36.55	36.69	36.71	36.75	36.81	36.88
36.89	36.91	36.85	36.89	36.89	36.67	36.50	36.74	36.77	36.76
36.78	36.82	36.89	36.99	36.92	36.99	36.89	36.94	36.92	36.97
36.91	36.79	36.77	36.69	36.62	36.54	36.55	36.67	36.69	36.62
36.64	36.59	36.65	36.75	36.80	36.81	36.87	36.87	36.89	36.94
36.98	36.95	37.00	37.07	37.05	37.00	36.95	37.00	36.94	36.88
36.93	36.98	36.97	36.85	36.92	36.99	37.01	37.10	37.09	37.02
36.96	36.84	36.87	36.85	36.85	36.87	36.89	36.86	36.91	37.53
37.23	37.20	*	37.25	37.20	37.21	37.24	37.10	37.20	37.18
36.93	36.83	36.93	36.83	36.80	36.75	36.71	36.73	36.75	36.72
36.76	36.70	36.82	36.88	36.94	36.79	36.78	36.80	36.82	36.84
36.86	36.88	36.93	36.97	37.15					

There are various ways that we can model this correlation structure. In time series literature, they define autoregressive (AR) models, moving average (MA) models, combinations of these known as ARMA models for data, or ARIMA models for differences in data values.

It is not always easy to identify which structure to use for a given data set. Two types of correlations are helpful in deciding on a particular structure. The set of these is known as the *autocorrelation function* (ACF) and *partial autocorrelation function* (PACF).

The *autocorrelation* $\boldsymbol{r_1}$ is the sample correlation between successive pairs of data, $\{Y_t, Y_{t-1}\}$, lagged by one time period.

The *autocorrelation* $\boldsymbol{r_2}$ is the sample correlation between successive pairs of data, $\{Y_t, Y_{t-2}\}$, lagged by two time periods, … and so on for other autocorrelations.

The *partial autocorrelation* $\boldsymbol{r_{2.1}}$ is the sample correlation between successive pairs of data, $\{Y_t, Y_{t-2}\}$, *adjusted for the effect of* Y_{t-1}. It is like performing a regression of Y_t on Y_{t-1}, saving the residuals and calculating a correlation of these with Y_{t-2}. This is extended to higher-order lags as well. As a starting point it is conventional to define $\boldsymbol{r_{1.0}}$ as $\boldsymbol{r_1}$, the first autocorrelation.

Both AC and PAC functions have specific forms for the different types of correlation structures.

Use **Stats > Time Series > Data Exploration**

	Beaver	Random	Beaver	Random
Unit	ACF	ACF	PACF	PACF
1	1	1	1	1
2	**0.802**	-0.117	**0.802**	-0.117
3	0.663	0.151	0.055	0.139
4	0.527	-0.036	-0.053	-0.004
5	0.463	-0.021	0.115	-0.047
6	0.353	0.149	-0.130	0.153
7	0.245	-0.063	-0.089	-0.026
8	0.153	0.148	-0.017	0.099
9	0.085	-0.107	-0.030	-0.068
10	0.061	0.050	0.077	0.005
11	0.027	-0.074	-0.024	-0.066
12	-0.004	0.029	-0.026	0.024
13	-0.004	-0.023	0.075	-0.042
14	0.009	-0.046	0.013	-0.031
15	0.036	0.061	0.046	0.039
16	0.056	-0.037	0.030	0.021
17	0.039	-0.029	-0.103	-0.074
18	0.015	0.041	-0.042	0.071
19	0.029	-0.025	0.076	-0.011
20	0.044	0.068	0.002	0.051

For the beaver data and the random temperature data, the ACF and PACF values are obtained as follows. Select **Time Series > Data Exploration** and the data to be investigated. In Options,

choose **Partial Autocorrelation Functions** if these are required. The default should include ACF and PACF plots.

ACF and PACF plots for beaver temperatures and random temperatures are given on the left hand page for the first twenty lags. The horizontal lines on each plot are confidence bands around zero values.

There is clearly a difference. For the beaver data, the ACF declines steadily while the PACF values are basically zero (note that, by definition, lag-1 correlations are unity). For the random data, both ACF and PACF functions are zero.

In this manual we will mention three correlation structures that are commonly used in biological sciences.

a) Uniform correlation model

This model says that the correlation between two data values is the same irrespective of the time or distance between them.

The uniform correlation matrix looks like $\begin{pmatrix} 1 & \rho & \cdots & \rho & \rho \\ \rho & 1 & & \rho & \rho \\ \vdots & & \ddots & \vdots & \\ \rho & \rho & & 1 & \rho \\ \rho & \rho & \cdots & \rho & 1 \end{pmatrix}$.

A uniform correlation structure applies, for example, whenever blocks are assumed random in a randomized block design. This means that the yields in a block are all uniformly correlated – which often is less than satisfactory. More likely, plots closer together are more highly correlated than plots far apart.

It is the only correlation structure that allows a split-plot ANOVA to be used validly for units in an experiment that are repeatedly measured in time.

b) AR1 or power model

This model says that the correlation between two data values declines exponentially with the time or distance between them. When time intervals or distances between plots are equal, the model is described as an AR1 model with correlations ρ, ρ^2, ρ^3, ρ^4, …. The power model is more general, with a correlation of ρ^s between observations s units apart – the units can be unequally spaced.

Data that follow an AR1 model are basically made up as follows.

The observation at time t is linearly related to that at time t-1 –this is a lag 1 process

Mathematically: $Y_t = \mu + \phi_1 (Y_{t-1} - \mu) + \textit{independent error}$,

where in this model $\rho = \phi_1$.

The AR1 correlation matrix looks like $\begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \\ \rho & 1 & \rho & \rho^2 & \rho^3 & \\ \rho^2 & \rho & 1 & \rho & \rho^2 & \cdots \\ \rho^3 & \rho^2 & \rho & 1 & \rho & \\ \rho^4 & \rho^3 & \rho^2 & \rho & 1 & \\ & & \vdots & & & \ddots \quad \vdots \\ & & & & & \cdots \end{pmatrix}$

The beaver data appears to follow an AR1 process, since the pattern of autocorrelations is (approximately) 0.8, 0.8^2=0.64, 0.8^3=0.51, 0.8^4=0.41, 0.8^5=0.33, 0.8^2=0. 26, …. The actual pattern is 0.8, 0.66, 0.53, 0.46, 0.35, 0.25, ….

c) AR2 or lag 2 model

For this process the dependent error depends only on the previous *two* dependent errors:

The observation at time *t* depends only on the previous *two* observations, those at time *t*-1 and at time *t*-2.

Mathematically: $Y_t = \mu + \phi_1 (Y_{t-1} - \mu) + \phi_2 (Y_{t-2} - \mu) + \textit{independent error}$,

where in this model the correlations are $\rho_1 = \phi_1/(1-\phi_2)$, $\rho_2 = \phi_2 + \phi_1^2/(1-\phi_2)$, …

The formulae for the higher-lag correlations in the AR2 correlation matrix become more complex. Suffice to say that the AR2 sequence ρ, ρ_2, ρ_3, ρ_4, … declines somewhat faster than the AR1 sequence ρ, ρ^2, ρ^3, ρ^4, ….

Deciding on a correlation structure

Generally we do not have a long run of correlated data, so time series devices that assist us to choose the most appropriate correlation model are unavailable.

Since correlations are some of the parameters of the random effects, we can use *change in deviance* to test whether some are zero or not.

In the AR2 model, setting $\phi_2 = 0$ produces an AR1 model.

In the AR1 model, setting $\phi_1 = 0$ produces an independent model.

We cannot compare uniform and AR1 models, since no value of ρ in the AR1 structure leads to a uniform correlation matrix. However, since a minimum deviance is associated with a maximum likelihood, the model having the smaller deviance is worth exploring. Generally, we support the choice by an investigation of the residuals: if the chosen model is appropriate, there should be no remaining trend in the residuals.

Time Series analysis of beaver data

Output series: Temperature Noise model: _erp

Residual deviance = 1.087
Innovation variance = 0.009569

Estimate of the *independent* error component of the model

Number of units present = 115
Residual degrees of freedom = 112

Estimate of the correlation between two temperatures 10 minutes apart

Summary of models

Model	Orders: Type	Delay B	AR P	Diff D	MA Q	Seas S
_erp	ARIMA	-	1	0	0	1

Parameter estimates

Model	Seas. Period	Diff. Order	Delay	Parameter	Lag	Ref	Estimate	s.e.	t
Noise	1	0	-	Constant	-	1	36.8489	0.0826	446.26
				Phi (AR)	1	2	**0.8968**	0.0473	18.96

REML analysis of beaver data

Assume an AR1 stationary model for temperature. We can use change in deviance to test this model, namely

$Temperature_t = \mu + \varepsilon_t$ independent model for the errors

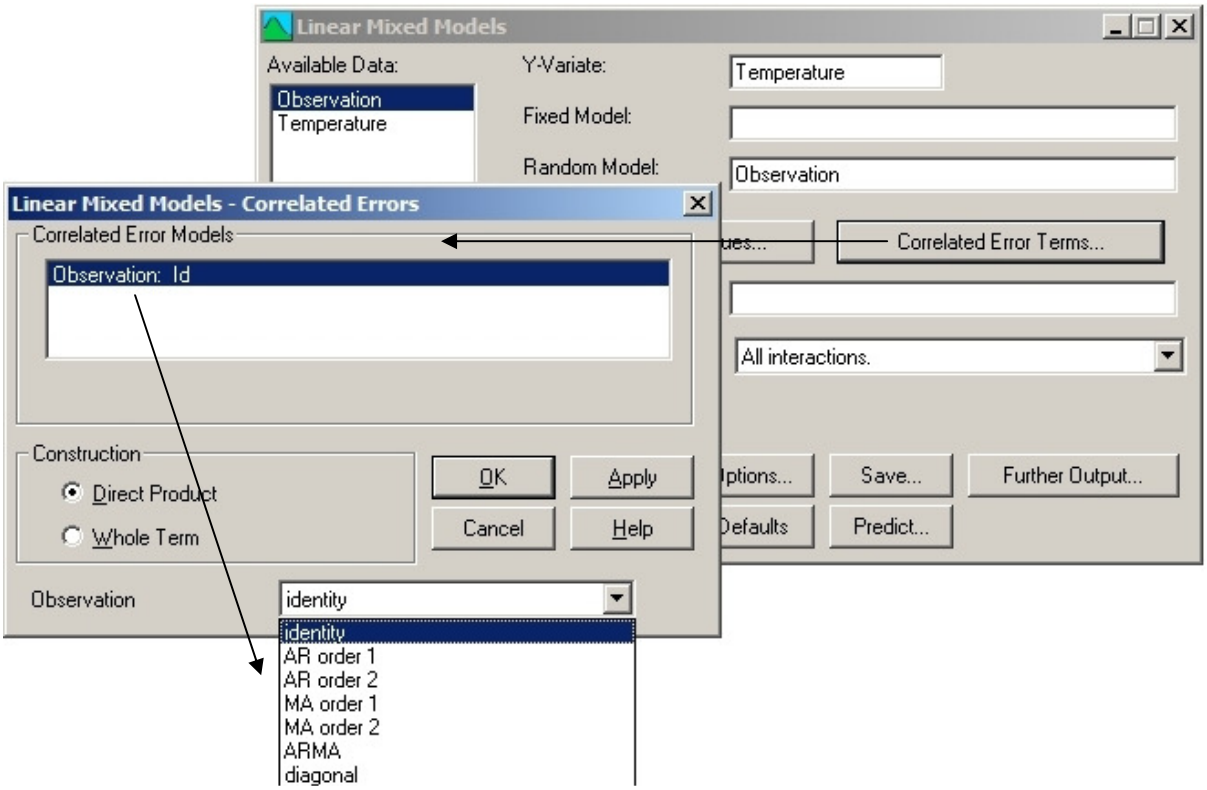
against the AR1-correlated model

$Temperature_t = \mu + \phi_1 \varepsilon^*_{t-1} + \varepsilon_t$ AR1-correlated model for the errors

Note that the estimates will be slightly different than those obtained using GenStat's Time Series menu. LMM (REML) used REML rather than ML to estimate the variance parameters.

For the *independent* model, we leave the **Fixed Model** blank (there is no predictor variate, just an overall mean which GenStat adds automatically). The **Random Model** consists of a factor to identify the n units, so we could set up our own Observation factor (with n = 115 levels), or just use the in-built '*Units*', or just leave it blank (since GenStat will add an independent error term for us). However, in order to set up a correlation structure later, we will add Observation at this stage.

For the *dependent* model, we again leave the **Fixed Model** blank (there is still no predictor variate). The **Random Model** consists of a factor to identify the *dependent* units ε^*_{t-1}; we use the factor Observation and declare an AR1 structure for this. Note that we could also set an AR2 structure (which assumes that the temperature at time t depends directly on the previous *two* temperatures) and test whether this more complex model is statistically better than the AR1 model. Unfortunately for this example the mathematical algorithm does not converge for the AR2 model.



The deviances for the two models are as follows. Clearly the AR1 model is superior to the independent error model.

Model	deviance	d.f.	change in deviance	change in d.f.	*P*-value
Identity	-253.56	112			
AR1	-411.23	110	157.67	2	<0.001

To maximize the explanation in GenStat's output we also use click **Covariance Model** in the LMM (REML) **Options**.

REML variance components analysis

Response variate: Temp_Beaver
Fixed model: Constant
Random model: Observation + '*units*'
Number of units: 114 (1 units excluded due to zero weights or missing values)

'*units*' used as residual term

Covariance structures defined for random model

Covariance structures defined within terms:

Term	Factor	Model	Order	No. rows
Observation	Observation	Auto-regressive (+ scalar)	1	115

Estimated parameters for covariance models

Random term(s)	Factor	Model(order)	Parameter	Estimate	s.e.
Observation	Observation	AR(1)	**phi_1**	**0.9337**	0.0472
			Scalar	113.4	218.2

Note: the covariance matrix for each term is calculated as G or R where
var(y) = Sigma2(ZGZ'+R), i.e. relative to the residual variance, Sigma2.

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
	'*units*'	Identity	Sigma2	0.000580	0.0010881

Estimated covariance models

Variance of data estimated in form:
V(y) = Sigma2(gZGZ' + I)

where: V(y) is variance matrix of data
Sigma2 is the residual variance
g is a gamma for the random term — G is the covariance matrix for the random term
Z is the incidence matrix for the random term — I is the residual (identity) covariance matrix

Note: a gamma is the ratio of a variance component to the residual (Sigma2)
Random Term: Observation

G is a single matrix
Scalar Sigma2*g: 0.06575

Factor: Observation
Model : Auto-regressive

Covariance matrix (first 10 rows only):

1	1.000									
2	0.934	1.000								
3	0.872	0.934	1.000							
4	0.814	0.872	0.934	1.000						
5	0.760	0.814	0.872	0.934	1.000					
6	0.710	0.760	0.814	0.872	0.934	1.000				
7	0.663	0.710	0.760	0.814	0.872	0.934	1.000			
8	0.619	0.663	0.710	0.760	0.814	0.872	0.934	1.000		
9	0.578	0.619	0.663	0.710	0.760	0.814	0.872	0.934	1.000	
10	0.539	0.578	0.619	0.663	0.710	0.760	0.814	0.872	0.934	1.000
	1	2	3	4	5	6	7	8	9	10

Residual term: '*units*'
Sigma2: 0.0005800

I is an identity matrix (114 rows)

Deviance: -2*Log-Likelihood

Deviance	d.f.
-411.23	110

Table of predicted means for Constant

36.87

Interpretation of the analysis

- The REML estimate of ρ (or ϕ_1 – labeled phi_1 in the output) is 0.9337; the ML time series estimate was 0.8968. Thus, the AR1 model assumes that the correlations between the temperatures are $(0.9337)^2 = 0.872$ for two units of time apart, $(0.9337)^3 = 0.814$ for three units of time apart, $(0.9337)^4 = 0.760$ for four units of time apart, $(0.9337)^5 = 0.710$

for five units of time apart, and so on. These values form the covariance matrix printed above.

- The scalar 113.4 is multiplied by the "variance estimate" 0.000580 giving 0.066 as the REML estimate of the variance of any temperature at a particular time point. This is confirmed in the output (Scalar Sigma2*g: 0.06575). This is the variance of the *dependent* error term in the model.

  In the time series output, this needs to be reconstructed from the properties of the time series. For the assumptions to work, the "innovative variance", i.e. the variance of the independent error component, turns out to be:

  variance(independent error) = $(1 - \rho^2)$ variance(temperature at time t)

  Hence

  variance(temperature at time t) = variance(independent error) / $(1 - \rho^2)$

  which is estimated as $0.009569/(1-0.8968^2) = 0.049$. Remember this is a ML estimate.

- The estimated REML model is

$$\begin{aligned} Temperature_t &= 36.87 + 0.9337\,\varepsilon^*_{t-1} + \varepsilon_t \\ &= 36.87(1-0.9337) + 0.9337\times Temperature_{t-1} + \varepsilon_{t-1} \\ &= 2.444 + 0.9337\times Temperature_{t-1} + \varepsilon_{t-1} \end{aligned}$$

  Thus, the temperature at time t is approximately 2.444°C + 0.9337 *times* the temperature at time t-1.

Simple linear regression

Example 3 Yields of potatoes receiving various amounts of fertilizer (Snedecor and Cochran, page 150).

Amount	0	4	8	12	*mean fertiliser* = 6.000
Yield	8.34	8.89	9.16	9.50	*mean yield* = 8.973

The linear regression model can be expressed either as

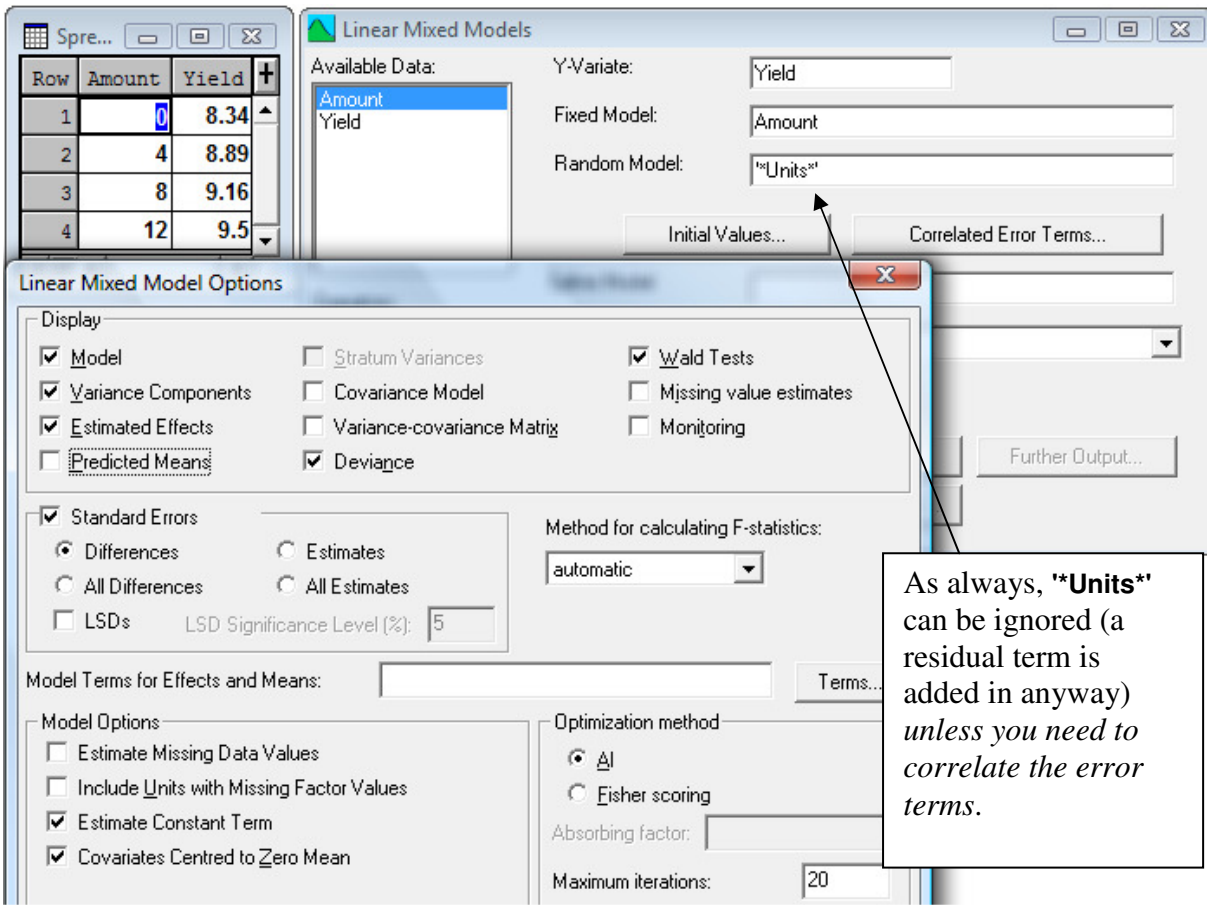
$$Yield = intercept + slope \times Fertiliser + Error$$

or as

$$Yield = mean\ yield + slope\ (Fertiliser - mean\ fertiliser) + Error$$

Notice that this model is in the form *mean + fixed effect + random effect*. The assumptions made when using a regression ANOVA (independent normally distributed errors with constant variance) fit within a LMM (REML) framework, and hence the analyses should be identical.

It is the second form of the model that GenStat has as the default in its LMM (REML) menu. To obtain the first form, go into **Options** and untick **Covariates Centred to Zero Mean**. You should also click **Deviance** and, for regression, the **Estimated Effects** (that is, mean Y and slope, or intercept and slope respectively).



Regression analysis

Response variate: Yield
Fitted terms: Constant, Amount

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	1	0.70312	0.703125	**82.00**	0.012
Residual	2	0.01715	0.008575		
Total	3	0.72028	0.240092		

Percentage variance accounted for 96.4
Standard error of observations is estimated to be 0.0926.

Estimates of parameters

Parameter	estimate	s.e.	t(2)	t pr.
Constant	8.4100	0.0775	108.55	<.001
Amount	0.0938	0.0104	9.06	0.012

REML variance components analysis

Response variate: Yield
Fixed model: Constant + Amount
Random model: '*units*'
Number of units: 4

'*units*' used as residual term

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
'	*units*'	Identity	Sigma2	0.00858	0.008575

Deviance: -2*Log-Likelihood

Deviance	d.f.
-1.75	1

Wald tests for fixed effects

Fixed term	Wald statistic	n.d.f.	**F statistic**	d.d.f.	F pr
Amount	82.00	1	**82.00**	2.0	0.012

and, for the default **Covariates Centred to Zero Mean**:

Table of effects for Constant

8.973 Standard error: 0.0463

Table of effects for Amount

0.09375 Standard error: 0.010353

If **Covariates Centred to Zero Mean** is unticked:

Table of effects for Constant

8.410 Standard error: 0.0775

Table of effects for Amount

0.09375 Standard error: 0.010353

So LMM (REML):

- produces the same F statistic (82.00) as regression produces for the ANOVA(called v.r. in that analysis);

- produces the same line of best fit

$$Yield = 8.410 + 0.09375\ Fertiliser$$

or equivalently

$$Yield = 8.973 + 0.09375\ (Fertiliser - 6.0)$$

The mean amount of fertilizer (6.0) is not part of the REML output, it needs to be calculated separately.

Unpaired *t* test – special case of a one-way treatment design (no blocking)

Example 4 Coefficients of digestibility of dry matter, of sheep and steers fed corn silage, in percent (Steel and Torrie, page 93)

	Sheep	Steers
	57.8	64.2
	56.2	58.7
	61.9	63.1
	54.4	62.5
	53.6	59.8
	56.4	59.2
	53.2	
mean	56.21	61.25
sd	3.00	2.83

The first decision to make is whether you are prepared to believe that the two population variances are equal. There is a variance ratio test for this, *but this test relies very heavily on the data being normally distributed*, so use it with care. Unless you change the default in **Options**, GenStat does the *F* test for you.

To test H_0: $\sigma_1^2 = \sigma_2^2$ for normally distributed data:

$F_{obs} = \frac{s_1^2}{s_2^2} \sim F$ variable with $(n_1\text{-}1)$ and $(n_1\text{-}1)$ *df*

If the test does not fail, then the unpaired *t* test is used to test the means, with $sed = \sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$ and $df = (n_1-1)+(n_2-1)$. Here, s_p^2 is a weighted average of the two treatment variances (see Appendix).

If the test does fail, then an *approximate t* test is used to test the means, with $sed = \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$. The degrees of freedom are calculated from the formula alongside; if the two sample variances are close, the *approximate df* are close to $(n_1\text{-}1)+(n_2\text{-}1)$. When the two sample variances are different, the *approximate df* will be closer to the *df* associated with the larger variance.

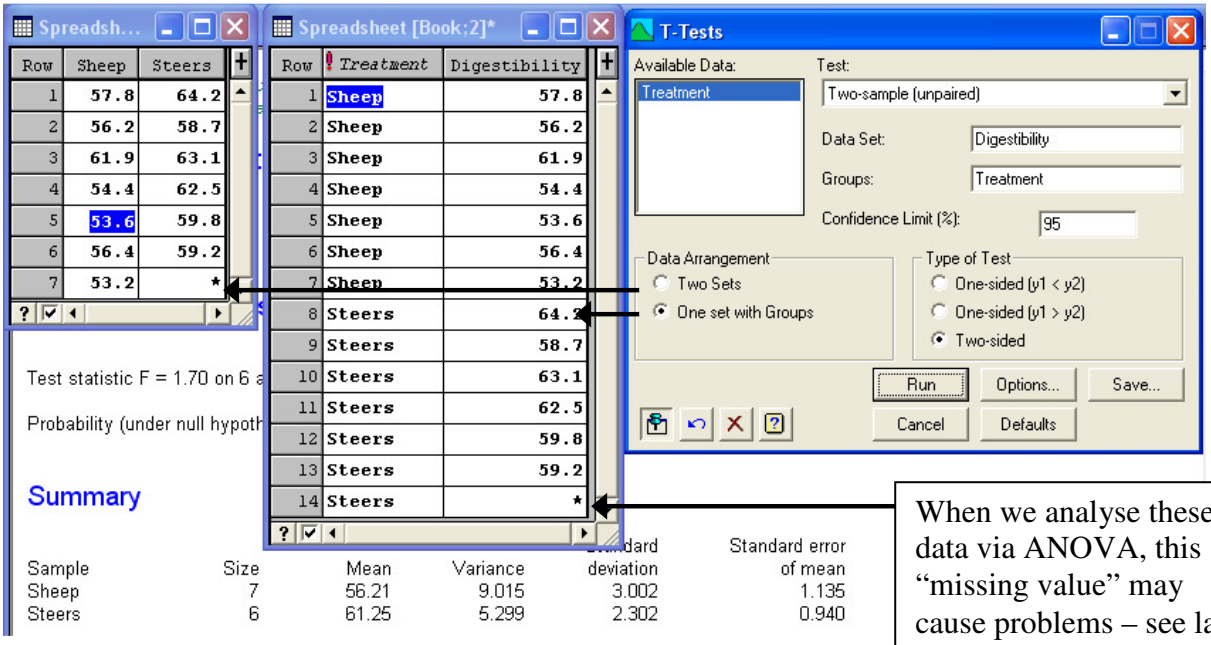
$$df = \left[\frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1}+\frac{\left(s_2^2/n_2\right)^2}{n_2-1}}\right]$$

To analyse the data, use **Stats > Statistical Tests > One- and two-sample t-tests…**. GenStat allows the data to be organized either in separate columns for the separate treatments, or in one combined data column *plus* a factor column to identify which observation each treatment belongs to. Since this is a special case of a more general design, we chose to illustrate the latter approach, see the output on the left hand page.

For the coefficients of digestibility of dry matter,

- there is no evidence (P=0.580) that the population variances are not equal
- there is strong evidence (P=0.007) that the population means are different. Steers have coefficients of digestibility that are, on average, 5.0% higher than for sheep. We are 95% confident that the true difference is between 1.7% and 8.4%.

GenStat's unpaired t test procedure



When we analyse these data via ANOVA, this "missing value" may cause problems – see later

Two-sample t-test

Variate: Digestibility
Group factor: Treatment

Test for equality of sample variances

Test statistic F = 1.70 on 6 and 5 d.f.
Probability (under null hypothesis of equal variances) = 0.58

Step 1. GenStat tests
H_0: $\sigma_1^2 = \sigma_2^2$ using $F = s_1^2 / s_2^2$.
Here there is *no evidence* that the population variances are not equal (P=0.580).

Summary

Sample	Size	Mean	Variance	Standard deviation	Standard error of mean
Sheep	7	56.21	9.015	3.002	1.135
Steers	6	61.25	5.299	2.302	0.940

Difference of means: -5.036
Standard error of difference: 1.506

95% confidence interval for difference in means: (-8.350, -1.721)

Test of null hypothesis that mean of Digestibility with Treatment = Sheep is equal to mean with Treatment = Steers

Test statistic t = -3.34 on 11 d.f.
Probability = 0.007

One-way (no Blocking) Model

Apart from individual random errors, the only possible differences in the data can come from individual treatment effects, leading to a model

Yield = mean + treatment effect + error

With *t* treatments, there can only be *t*-1 treatment effects in a model that contains an overall mean: the effects measure how far a particular treatment is from the overall mean. Note that the general regression model allows factors as explanatory variates. ANOVA is therefore just a special case of multiple linear regression. However, the model is also a special case of a LMM, and hence the t-test can be performed using ANOVA, regression or LMM (REML).

Regression output

Here is GenStat's output from **Stats > Regression Analysis > Linear Models** and choosing **General Linear Regression** from the drop down selection. The model is referenced to level 1 (Sheep), hence Constant is the estimate of the Sheep mean. The coefficient Treatment Steers is what you add to the Constant to obtain the mean for the second level (Steers) and hence is the difference in means (Steers-Sheep).

Regression analysis

Response variate: Digestibility
Fitted terms: Constant, Treatment

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	1	81.93	81.927	11.18	0.007
Residual	11	80.58	7.326		
Total	12	162.51	13.543		

Percentage variance accounted for 45.9
Standard error of observations is estimated to be 2.71.

Message: the following units have large standardized residuals.

Unit	Response	Residual
3	61.90	2.27

Estimates of parameters

Parameter	estimate	s.e.	t(11)	t pr.
Constant	56.21	1.02	54.95	<.001
Treatment Steers	5.04	1.51	3.34	0.007

Parameters for factors are differences compared with the reference level:

Factor	Reference level
Treatment	Sheep

- Same *P*-value as that for *t* test of means
- v.r. = t^2 $11.18 = 3.34^2$
- 7.326 is the pooled estimate of variance
- Constant is mean for level 1 Sheep
- *Difference in means* is 5.04
- *sem* = 1.02 for Sheep
- *sed* = 1.51

Analysis of Variance output

Use **Stats** > **Analysis of Variance**. There is a special menu item for this design, but we prefer to use the **General** analysis of variance. We have also gone into **Options** and selected **l.s.d.**s. Without changing the stacked spreadsheet, the output is as follows.

Analysis of variance

Variate: Digestibility

Source of variation	d.f.	(m.v.)	s.s.	m.s.	v.r.	F pr.
Treatment	1		88.754	88.754	12.12	0.005
Residual	11	(1)	80.584	7.326		
Total	12	(1)	162.511			

Message: the following units have large residuals.

units 3 5.69 s.e. 2.40

Tables of means

Grand mean 58.73

Treatment	Sheep	Steers
	56.21	61.25

Standard errors of differences of means

Table	Treatment
rep.	7
d.f.	11
s.e.d.	1.447

(Not adjusted for missing values)

Least significant differences of means (5% level)

Table	Treatment
rep.	7
d.f.	11
l.s.d.	3.184

(Not adjusted for missing values)

This is not exactly the same analysis, because with unequally replicated treatments, if you leave a row in with an asterisk (*) to signify a missing value, GenStat assumes you want to estimate the missing value. This is rather an old fashioned approach. It *over-estimates* the Treatment SS and the resulting variance ratio is therefore too large.

If you really do have missing values, there is an **Unbalanced Treatment Structure** you can use in this case. (Basically, GenStat analyses the data via regression for you.)

If this is a case of a deliberate choice of sample size (for example, these are the only steers you could get hold of), then a correct analysis is obtained after deleting the row with the *.

Here are both analyses. The similarities are obvious.

Unbalanced Treatment Structure output

(i) Including the row with the missing value, choosing Unbalanced Treatment Structure

Analysis of an unbalanced design using GenStat regression

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ Treatment	1	81.927	81.927	**11.18**	**0.007**
Residual	11	80.584	7.326		
Total	12	162.511	13.543		

Predictions from regression model

Treatment	Prediction
Sheep	56.21
Steers	61.25

Standard error of differences between predicted means	**1.506**
Least significant difference (at 5.0%) for predicted means	**3.314**

(ii) Deleting the row with the non-observed value, choosing General Analysis of Variance

Analysis of variance

Variate: Digestibility

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Treatment	1	81.927	81.927	**11.18**	**0.007**
Residual	11	80.584	7.326		
Total	12	162.511			

Message: the following units have large residuals.

units 3	5.69	approx. s.e.	2.49

Tables of means

Grand mean 58.54

Treatment	Sheep	Steers
	56.21	61.25
rep.	7	6

Standard errors of differences of means

Table	Treatment
rep.	unequal
d.f.	11
s.e.d.	1.506

Least significant differences of means (5% level)

Table	Treatment
rep.	unequal
d.f.	11
l.s.d.	3.314

LMM (REML) analysis of one-way design (no blocking)

The **Fixed Model** is again Treatment. Since there is only one random error term we can ignore the **Random Model**, since as always GenStat allows us to omit the error in the final stratum – it adds it in for us. Tick to obtain deviances and predicted means. From Version 11 l.s.d. values can be selected as well. Missing values are ignored, as in regression, so the * that may be in the stacked dataset is simply ignored.

REML variance components analysis

Response variate: Coefficient
Fixed model: Constant + Treatment
Number of units: 13 (1 units excluded due to zero weights or missing values)

Residual term has been added to model

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
Residual		Identity	Sigma2	7.326	3.124

Deviance: -2*Log-Likelihood

Deviance	d.f.
36.64	10

Note: deviance omits constants which depend on fixed model fitted.

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Treatment	11.18	1	11.18	11.0	0.007

Message: denominator degrees of freedom for approximate F-tests are calculated using algebraic derivatives ignoring fixed/boundary/singular variance parameters.

Standard error of differences: 1.506

Table of predicted means for Constant

58.73 Standard error: 0.753

Table of predicted means for Treatment

Treatment	Sheep	Steers
	56.21	61.25

Standard error of differences: 1.506

Approximate least significant differences (5% level) of REML means Treatment

Treatment Sheep	1	*	
Treatment Steers	2	3.314	*
		1	2

Notice that regression, LMM (REML) and ANOVA (except with the missing unit retained) analyses give virtually the same information as the *t* test did. We obtained:

- the equivalent test statistic (*F* instead of t^2);
- the same *P*-value for testing the difference between the two means (0.007);
- the same estimate of variance (7.326) and hence the same s.e.d. value (1.506);
- the same means and l.s.d. values

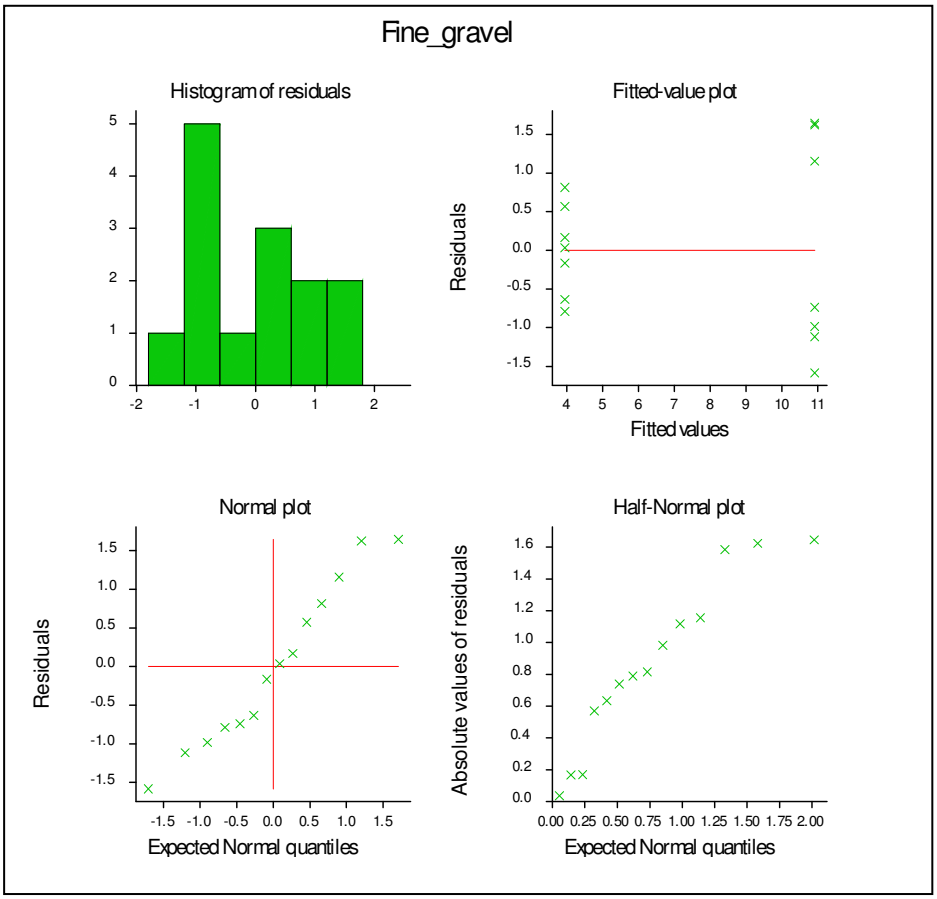
An advantage to the *t* test is the calculation of the confidence interval for treatment mean difference (μ_{steers}-μ_{sheep}). With the other approaches you need to add and subtract the l.s.d. value (3.314) to the mean difference (61.25-56.21) to obtain the confidence interval. Another advantage is the default automatic check on equality of treatment variances, which is a very important assumption underlying ANOVA. We will demonstrate how to do this in LMM (REML) with the next example.

An advantage to the ANOVA approach is that unusual values (ie standardized residuals outside the range (-2, +2)) are flagged. It is also important to routinely examine (standardized) residual plots.

Unpaired *t* test – example of unequal variances – Satterthwaite's approximate *t* test

Example 5 Fine gravel in soil, in percent (Steel and Torrie, page 107)

	Good soil	Poor soil
	5.9	7.6
	3.8	0.4
	6.5	1.1
	18.3	3.2
	18.2	6.5
	16.1	4.1
	7.6	4.7
mean	10.91	3.94
variance	40.12	6.95



Both means and variances in the two samples *appear* to be different. What statistical evidence is there that the mean percentage of fine gravel in the soil differs in the two soil types?

We first analysed the data via a one-way (no blocking) analysis of variance, and examined the residual plot. It is clear that the soil with the higher fitted value (obviously the good soil) has a larger visual scatter of residuals compared to that for the poor soil. This is a reflection of the different variances in the two samples.

An analysis in GenStat via a *t* test results in strong statistical evidence ($P = 0.020$) that the mean percentages of fine gravel differ. However, the test of equal variances is marginal. GenStat actually proceeds to use the standard unpaired *t* test because technically the *F* test does not fail ($P = 0.05$ to two decimals; it is actually 0.0509). We make three points.

- The *F* test depends heavily on normally distributed data, and percentages are unlikely to be normally distributed, so the *P*-value is somewhat unreliable.
- Failure to reject in this case is most likely to be caused by the low level of replication.
- We often make decisions about homogeneity of variance in more complex analyses of variance from an inspection of the standardized residual plot, rather than a formal test.

As mentioned previously, the default in GenStat for this test is to allow it to decide automatically what test to use for the means. To illustrate the approximate procedure, we over-rode GenStat by going into the **Options** menu, as shown. The change for an equally replicated experiment is only in the *df* of the *t* test (and hence in the *P*-value). Remember, it is not an exact t test. Here, the *df* used are obtained from the Satterthaite formula and are closer to 6 than to 12, since the variances are quite different in the sample.

GenStat output for the automatic *t* test of the fine gravel data

Two-sample t-test

Variate: Fine_gravel
Group factor: Soil

Test for equality of sample variances

Test statistic F = 5.77 on 6 and 6 d.f.
Probability (under null hypothesis of equal variances) = 0.05

Step 1. Test for equality of variances

Summary

Sample	Size	Mean	Variance	Standard deviation	Standard error of mean
good	7	10.914	40.12	6.334	2.394
poor	7	3.943	6.95	2.636	0.996

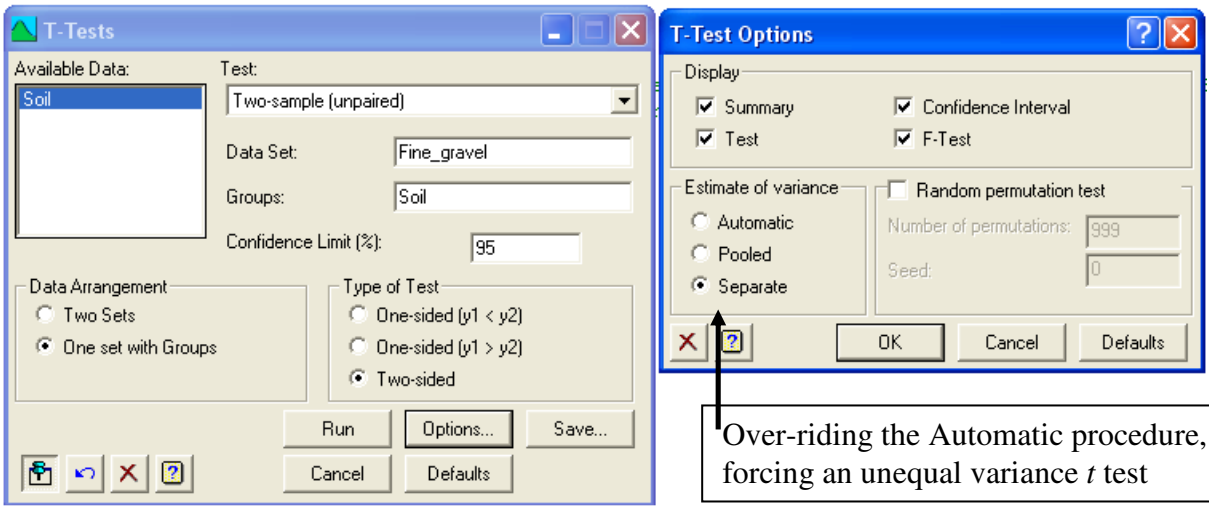
Difference of means: 6.971
Standard error of difference: 2.593

95% confidence interval for difference in means: (1.321, 12.62)

Test of null hypothesis that mean of Fine_gravel with Soil = good is equal to mean with Soil = poor

Test statistic t = 2.69 on 12 d.f.
Probability = 0.020

Step 2. Test for equality of means



Over-riding the Automatic procedure, forcing an unequal variance *t* test

Difference of means: 6.971
Standard error of difference: 2.593

95% confidence interval for difference in means: (0.9937, 12.95)

Test of null hypothesis that mean of Fine_gravel with Soil = good is equal to mean with Soil = poor

Test statistic t = 2.69 on approximately 8.02 d.f.
Probability = 0.028

Change to Step 2. Calculates approximate *df* for *t* test (8 instead of 12) and gives new *P*-value

LMM (REML) output for two sample t test (unequal variances)

The model for this dataset is as follows.

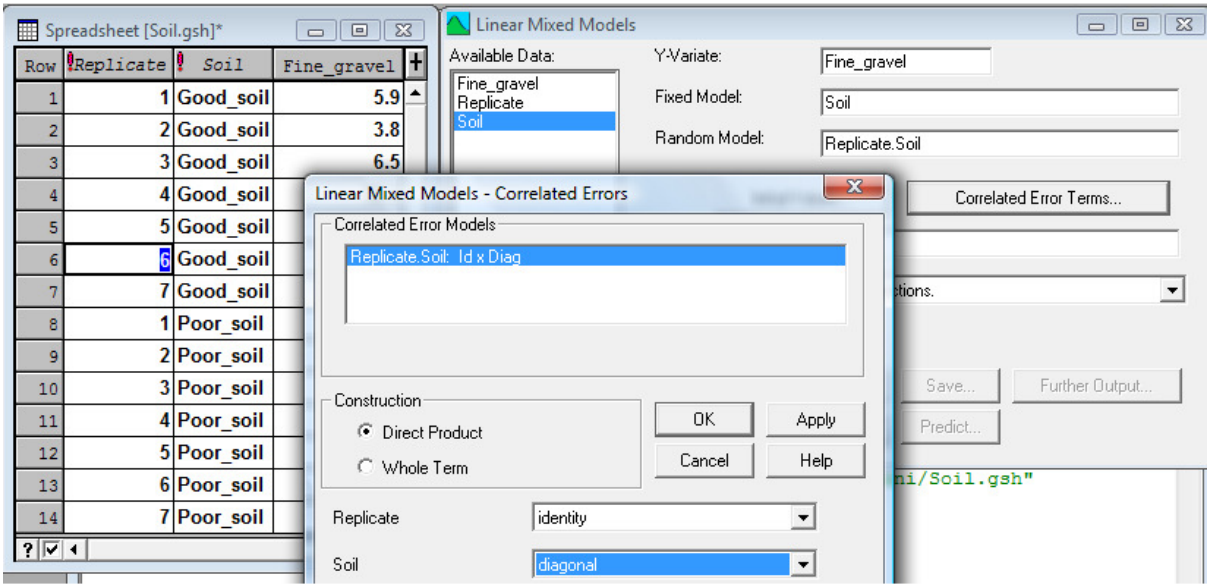
Fine gravel percentage = mean + soil effect + error.

There are two competing hypotheses as far as variances are concerned. The first is that the variance of the good soil is equal to that of the poor soil. The alternative is that they are different. Since these are parameters in the random part of the model, we test equality by change in deviance.

Equality of variances is represented in the **Correlated Error Terms** sub-menu as an Identity variance matrix. For this matrix, the off-diagonal elements are all zero, reflecting the absence of any correlation in the data; the diagonal elements are all unity, reflecting the equality of variances. The variance matrix is actually σ^2 times the identity matrix.

Inequality of variances is represented in the **Correlated Error Terms** sub-menu as a Diagonal variance matrix. For this matrix, the off-diagonal elements are again zero, reflecting the absence of any correlation; the diagonal elements are different multipliers, reflecting the equality of variances. The different variances are obtained by multiplying σ^2 by the diagonal elements of the variance matrix.

In order to actually access the **Correlated Error Terms** sub-menu, we need to enter the residual term ourselves. As always, the residual term must be a factor that indexes over all the data, in such a way as the factor Soil is present. Then we can set the levels of that factor to have a Diagonal variance matrix. We therefore need to set up a Replicate factor to index over the 7 replicates of each of good and poor soil:



We run the analysis twice, once with Identity and once with Diagonal and record the deviance information:

Model	Estimates of parameters in model	Deviance	d.f.	P
unequal variances	σ^2_{good} = 40.1 (6 *df*), σ^2_{poor} =6.9 (6 *df*)	49.68	10	
equal variances	σ^2 = weighted average = 7.326	53.79	11	
change in deviance		4.11	1	0.043

Here, the change in deviance is based on an asymptotic χ^2 distribution, not the F distribution. Since we have significance at 5%, we use the unequal variance output.

REML variance components analysis

Response variate: Fine_gravel
Fixed model: Constant + Soil
Random model: Replicate.Soil
Number of units: 14

Replicate.Soil used as residual term with covariance structure as below

Covariance structures defined for random model

Covariance structures defined within terms:

Term	Factor	Model	Order	No. rows
Replicate.Soil	Replicate	Identity	0	7
	Soil	Diagonal	2	2

Residual variance model

Term	Factor	Model(order)	Parameter	**Estimate**	s.e.
Replicate.Soil	Sigma2	1.000	fixed		
	Replicate	Identity	-	**-**	-
	Soil	Diagonal	d_1	**40.12**	23.17
			d_2	**6.950**	4.012

d_1 and d_2 are the diagonal elements and represent the two soil variances

Deviance: -2*Log-Likelihood

Deviance	d.f.
49.68	10

Note: deviance omits constants which depend on fixed model fitted.

Tests for fixed effects

The F statistic is identical to the square of the Satterthwaite t test obtained earlier:
Test statistic t = 2.69 on approximately 8.02 d.f.

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Soil	7.23	1	7.23	8.0	0.028

Message: denominator degrees of freedom for approximate F-tests are calculated using algebraic derivatives ignoring fixed/boundary/singular variance parameters.

Table of predicted means for Constant

7.429 Standard error: 1.2966

Table of predicted means for Soil

Soil	Good_soil	Poor_soil
	10.914	3.943

Standard error of differences: 2.593

Approximate least significant differences (5% level) of REML means

Soil

Soil Good_soil	1	*		
Soil Poor_soil	2	**5.980**	*	Appropriate l.s.d. value
		1	2	

Note. If GenStat produces a Sigma2 value that is not unity, then d_1 will be 1.000 and d_2 a multiplier different to 1.000. These are GenStat's gamma (multiplier) values. The Sigma parameterization is easily obtained by capturing the `REML` line, copying it to a new `Input` window and modifying the `PARAMETERIZATION` option:

```
REML [PRINT=model,components,means,deviance,waldTests; PSE=differences;\
PARAMETERIZATION=sigmas;MVINCLUDE=*; METHOD=ai; MAXCYCLE=20000] Fine_gravel
```

Paired *t* test – special case of a one-way treatment design (in randomised blocks)

Example 6 Sugar concentrations of nectar in half heads of red clover kept at different vapor pressures for eight hours (from Steel and Torrie, page 103)

Head	4.4 mm Hg	9.9 mm Hg	difference
1	62.5	51.7	10.8
2	65.2	54.2	11.0
3	67.6	53.3	14.3
4	69.9	57.0	12.9
5	69.4	56.4	13.0
6	70.1	61.5	8.6
7	67.8	57.2	10.6
8	67.0	56.2	10.8
9	68.5	58.2	10.3
10	62.4	55.8	6.6
mean	67.04	56.15	10.89
sd	2.82	2.72	2.22

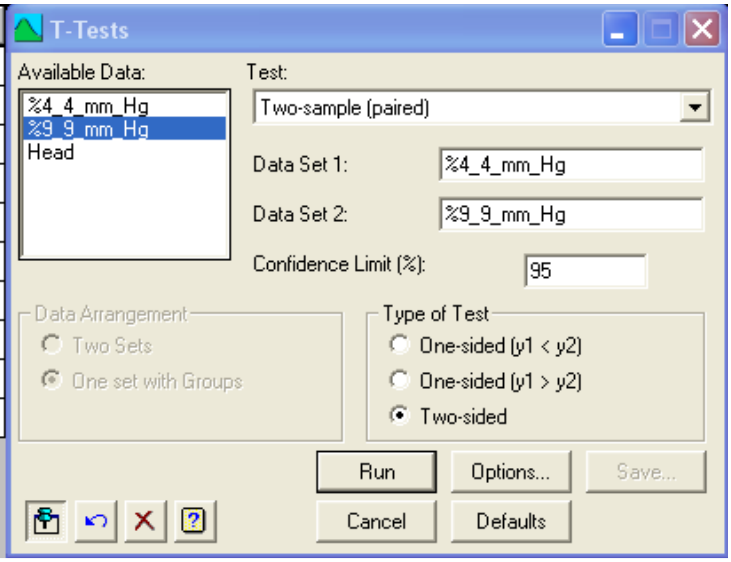
This example is quite different to the previous two examples. In this case, we cannot place the 10 concentrations in any order in each column: they are *paired*. The heads of red clover are divided into half heads; one is randomly subjected to a vapor pressure of 4.4 mm Hg, the other to a vapor pressure of 9.9 mm Hg. Each head of clover is likely to vary in its sugar concentration, and the only way to remove this variation is to take differences, and analyse these in a one sample *t* test.

When we have more than two treatments in an experiment that is blocked in some way, then we need to analyse the data using an ANOVA *F* test, setting up a "block" factor as well as a "treatment" factor.

Firstly, in GenStat, paired t test data must be set up in separate columns for separate treatments.

As a paired t test

Row	Head	%4_4_mm_Hg	%9_9_mm_Hg
1	1	62.5	51.7
2	2	65.2	54.2
3	3	67.6	53.3
4	4	69.9	57
5	5	69.4	56.4
6	6	70.1	61.5
7	7	67.8	57.2
8	8	67	56.2
9	9	68.5	58.2
10	10	62.4	55.8



One-sample t-test

Variate: Y[1].

Sample	Size	Mean	Variance	Standard deviation	Standard error of mean
VP_4_4-VP_9_9	10	10.89	4.914	2.217	0.7010

95% confidence interval for mean: (9.304, 12.48)

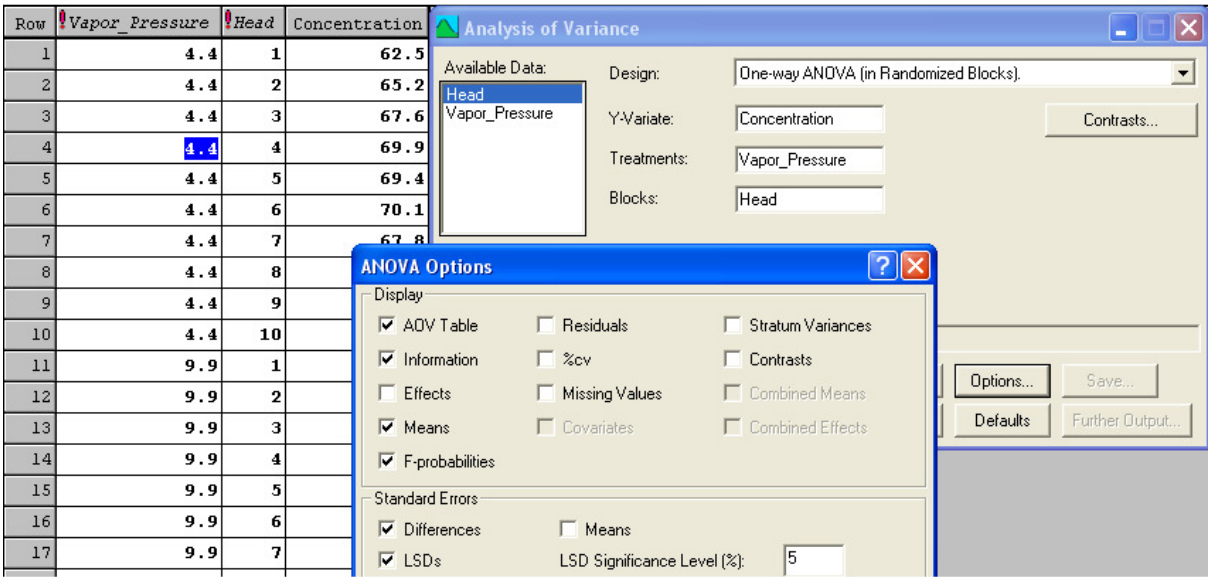
Test of null hypothesis that mean of VP_4_4-VP_9_9 is equal to 0

Test statistic t = 15.53 on 9 d.f.
Probability < 0.001

There is strong statistical evidence (P<0.001) that the mean sugar concentration of nectar differs in heads of red clover kept at different vapor pressures for eight hours. The best estimated mean difference is 10.89%, and we are 95% confident that the true difference lies between 9.30% and 12.48%.

To analyse the data via ANOVA or regression, we must stack the data, and provide a factor column to identify the various head (acting as blocks).

Paired t test as a one-way treatment design (in randomized blocks)



Notice in the output that GenStat organizes the ANOVA into the two strata for this experiment. Individual heads form the top stratum, and since these are not replicated (there is no other “head 1” or “head 2” etc), there is no *P*-value for this variance ratio. The second stratum is the “Heads.Units” stratum, that is, the half head put into one of two vapor pressure treatments (at random). These are replicated in a balanced way (each treatment occurs once in each block).

Thus, the actual block structure is Head + Head.Vapor_Pressure or Head.Vapor_Pressure (see GenStat’s syntax rules in the Appendix). The final error term has been dropped from the **Blocks** structure, as GenStat always allows this final stratum to be ignored (it adds it for us).

Analysis of variance

Variate: Concentration

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Head stratum	9	116.114	12.902	5.25	
Head.*Units* stratum					
Vapor_Pressure	1	592.960	592.960	**241.32**	<.001
Residual	9	22.115	2.457		
Total	19	731.189			

Message: the following units have large residuals.

Head 10 *units* 1	-2.14	s.e.	1.05
Head 10 *units* 2	2.14	s.e.	1.05

Tables of means

Variate: Concentration

Grand mean 61.60

Vapor_Pressure	4.4	9.9
	67.04	56.15

Standard errors of differences of means

Table	Vapor_Pressure
rep.	10
d.f.	9
s.e.d.	0.701

Least significant differences of means (5% level)

Table	Vapor_Pressure
rep.	10
d.f.	9
l.s.d.	1.586

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Head	12.902	9.000	**5.222**
Head.*Units*	2.457	9.000	**2.457**

Again, notice

- the relationship between the *t*-value of 15.53, and the *F*-value of 241.32 ($15.53^2 = 241.32$);
- the same *P*-value (*P*<0.001, though it is hard to see the similarity, *P* is so small;
- the mean difference is 67.04 - 56.15 = 10.89 ± 1.586, giving rise to the same confidence interval.

Regression output

Remember that a *t* test is just a special case of regression. There are two models to consider when testing whether the vapor pressure treatment effect is zero.

Maximal model
Sugar concentration = *overall mean* + *Head* effect + *Vapor pressure* effect + *Error*

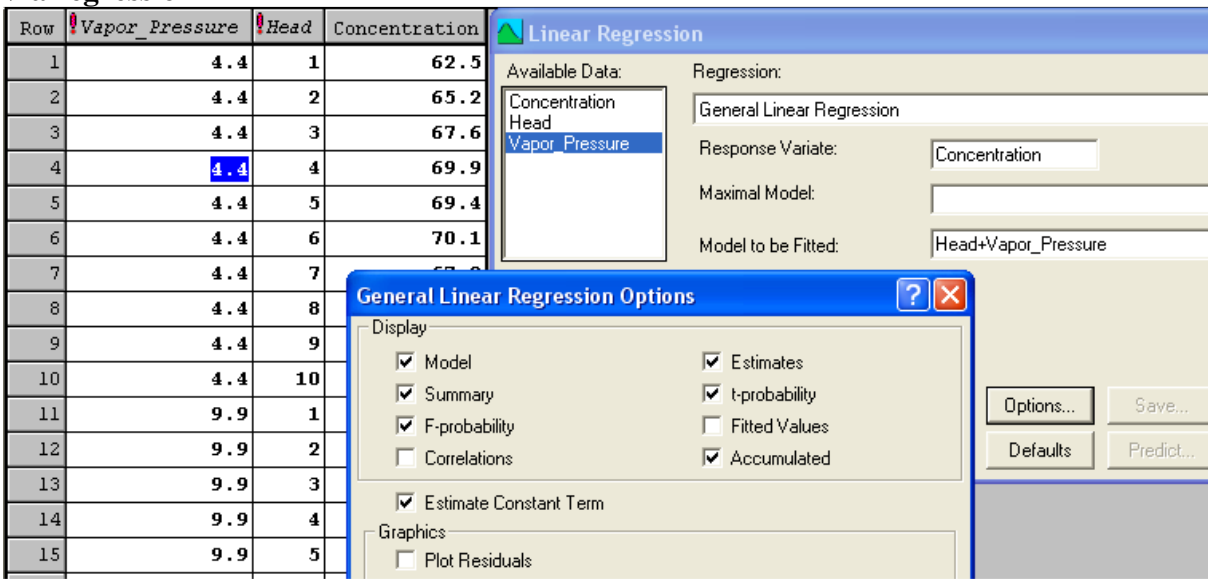
dropped

Reduced model
Sugar concentration = *overall mean* + *Head* effect + *Error*

Technically you need to run both models. The best estimate of error variance is obtained as the Residual MS from the ANOVA of the maximal model. The effect of treatments over and above that of blocks is obtained by subtracting the residual sums of squares from the two ANOVAs; divide this by the change in degrees of freedom to obtain the Treatment MS. The variance ratio is constructed as the ratio of the Treatment MS and Residual MS from the maximal model.

In GenStat's **General Linear Regression Option** menu, the effect of blocks (*Heads*) and treatments (*vapor pressure*) can be assessed by turning on **Accumulated**.

Via regression



Regression analysis

Response variate: Concentration
Fitted terms: Constant + Head + Vapor_Pressure

Summary of analysis

Source	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	10	709.08	70.908	28.86	<.001
Residual	9	22.11	2.457		
Total	19	731.19	38.484		

Percentage variance accounted for 93.6

Standard error of observations is estimated to be 1.57.

Message: the following units have large standardized residuals.

Unit	Response	Residual
10	62.40	-2.04
20	55.80	2.04

Estimates of parameters

Parameter	estimate	s.e.	t(9)	t pr.
Constant	**62.55**	**1.16**	**53.80**	**<.001**
Head 2	2.60	1.57	1.66	0.132
Head 3	3.35	1.57	2.14	0.061
Head 4	6.35	1.57	4.05	0.003
Head 5	5.80	1.57	3.70	0.005
Head 6	8.70	1.57	5.55	<.001
Head 7	5.40	1.57	3.44	0.007
Head 8	4.50	1.57	2.87	0.018
Head 9	6.25	1.57	3.99	0.003
Head 10	2.00	1.57	1.28	0.234
Vapor_Pressure 9.900	**-10.890**	**0.701**	**-15.53**	**<.001**

Parameters for factors are differences compared with the reference level:

Factor	Reference level
Head	1
Vapor_Pressure	4.400

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ Head	9	116.115	12.902	5.25	0.011
+ Vapor_Pressure	**1**	**592.961**	**592.961**	**241.32**	**<.001**
Residual	9	22.114	2.457		
Total	19	731.190	38.484		

The default model produces a **Constant** (the mean for vapor pressure 4.4) and a mean difference of -10.890, labeled **Vapor_Pressure 9.900**. This is highly significant, with a *t*-value of -15.53, the same (apart from sign) as was produced by the paired *t* test. The **Accumulated** analysis is the RCBD ANOVA, though it is an application of the general technique for comparing a maximal and reduced model.

Notice also that 1.16 is actually the s.e.m. and 0.701 the s.e.d..

LMM (REML) analysis of one-way treatment design in randomized blocks

Blocks in a field experiment are almost always treated as random factors, although it makes no difference to the test of treatment means whether it is treated as fixed or random – we will demonstrate this property later.

In this case, the factor Head is almost certainly a random factor: heads were chosen from a large number of heads, at random. GenStat assumed it to be random in the ANOVA output, producing variance components for the **Head** stratum as well as the **Heads.Units** stratum:

Estimated stratum variances

Stratum	variance	effective d.f.	**variance component**
Head	12.902	9.000	**5.222**
Head.*Units*	2.457	9.000	**2.457**

Hence, for linear mixed models, we have:

Fixed Model: Vapor_Pressure.
Random Model Head + Head.Vapor_Pressure
(or Head/Vapor_Pressure, or for simplicity Head since GenStat adds an error term for the lowest stratum if we omit it).

REML variance components analysis

Response variate: Concentration
Fixed model: Constant + Vapor_Pressure
Random model: Head + Head.Vapor_Pressure
Number of units: 20

Head.Vapor_Pressure used as residual term

Estimated variance components

Random term	component	s.e.
Head	**5.222**	3.096

identical to the Head stratum variance of the ANOVA

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
	Head.Vapor_Pressure	Identity	Sigma2	**2.457**	1.158

identical to the Residual MS of the ANOVA

Deviance: -2*Log-Likelihood

Deviance	d.f.
53.71	16

The F statistic is identical to the variance ratio in the ANOVA, as are df.

Wald tests for fixed effects

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Vapor Vapour_pressure	241.32	**1**	**241.32**	**9.0**	<0.001

Table of predicted means for Constant

61.59 Standard error: 0.803

Table of predicted means for Vapor_Pressure

Vapor_Pressure	1	2
	67.04	56.15

Means, s.e.d. and l.s.d. values are identical to those from the ANOVA.

Standard error of differences: 0.7010

Approximate least significant differences (5% level) of REML means

Vapour_pressure

Vapour_pressure %4_4_mm_Hg	1	*	
Vapour_pressure %9_9_mm_Hg	2	1.586	
		1	2

Completely randomized design (CRD), or one-way design (no blocking)

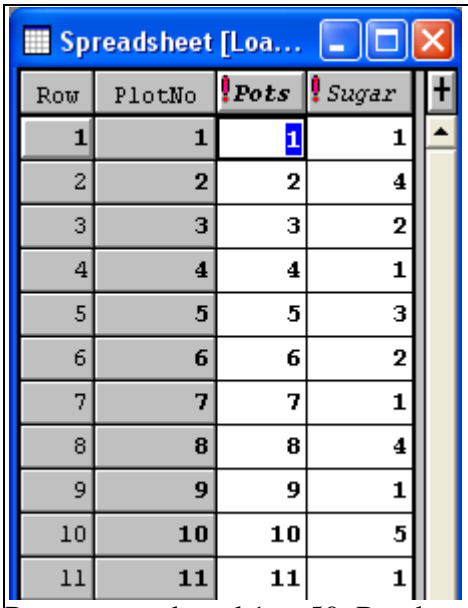
The data are from an experiment in plant physiology. Lengths of pea sections grown in tissue culture with auxin present were recorded. The purpose of the experiment was to test the effects of various sugar media on growth as measured by length.

Treatment structure: Single factor with 5 levels: sugar treatments (including a control)
Block Structure: None: 10 replicates for all treatments

Example 7 The effect of different sugars on length, in ocular units (× 0.114 = mm), of pea sections grown in tissue culture with auxin present (Sokal & Rohlf 3rd Ed. page 218)

Replicate	*Control*	*2% glucose added*	*2% fructose added*	*1% glucose + 1% fructose added*	*2% sucrose added*
1	75	57	58	58	62
2	67	58	61	59	66
3	70	60	56	58	65
4	75	59	58	61	63
5	65	62	57	57	64
6	71	60	56	56	62
7	67	60	61	58	65
8	67	57	60	57	65
9	76	59	57	57	62
10	68	61	58	59	67

In this experiment we have 50 pots (labelled 1 to 50) with no blocking required. The pots are placed in a growth chamber, and the treatments randomized to the pots (eg using GenStat's **Design** menu; notice that GenStat creates a factor column Pots, with levels 1 to 50):

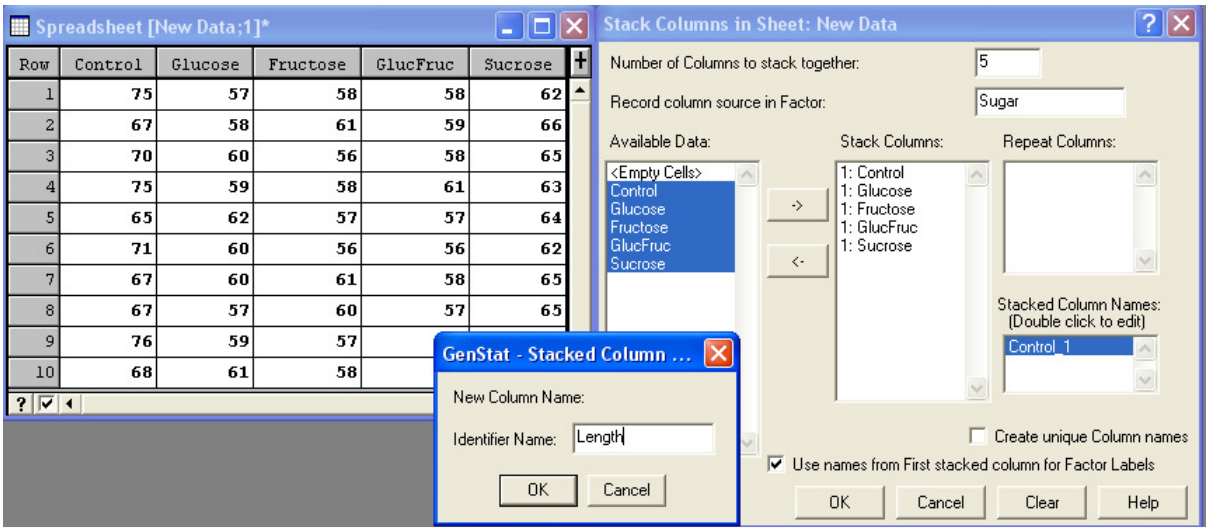


Pots are numbered 1 to 50. Random allocation of the *Control* treatment is shown

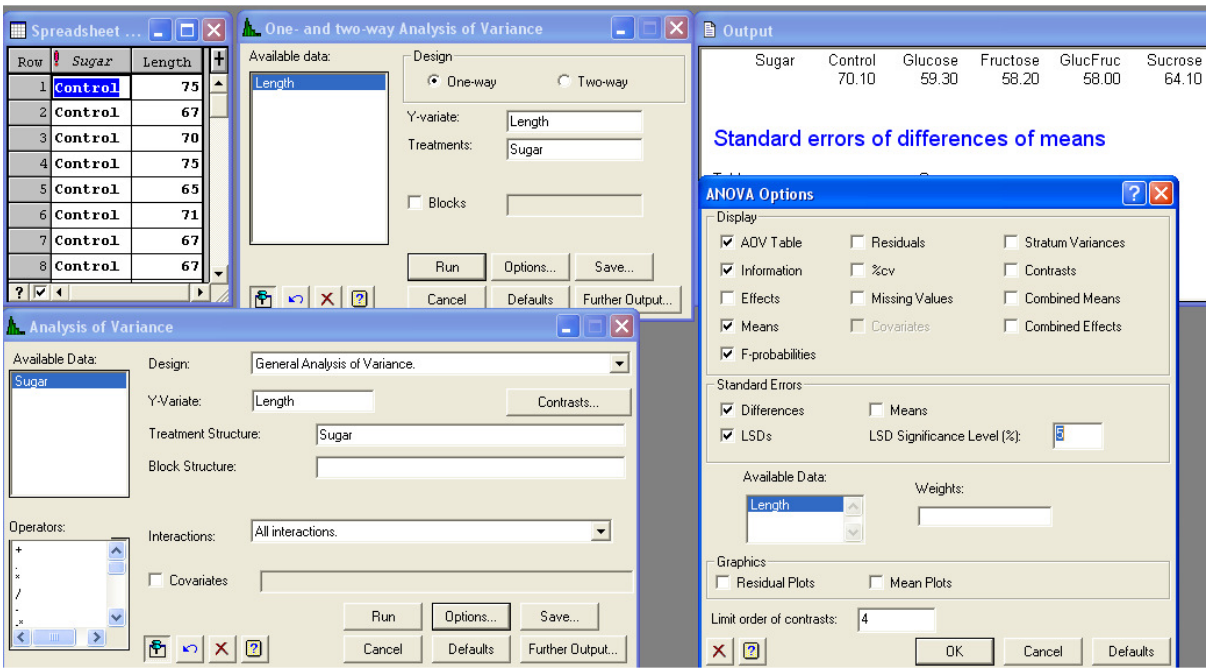
1 Control	2	3	4 Control	5
6	7 Control	8	9 Control	10
11 Control	12	13	14	15
16	17	18	19 Control	20 Control
21	22	23 Control	24	25
26 Control	27	28	29	30
31	32	33	34	35
36	37	38	39	40
41	42	43	44 Control	45
46	47	48	49	50

Data and analysis in GenStat

We firstly stack the data into a *variate* labelled `Length`, and create an identifier *factor* for the `Sugar` treatments. It is much more sensible to use treatment **labels** or treatment **levels** where possible. (Note that this can be done while stacking the data.) GenStat will always use labels or levels in its output. You can see that GenStat replaces the identifying numbers with actual labels.



Choose **One- and Two-way** to obtain the basic CRD ANOVA; alternatively, choose **General Analysis of Variance** and use Pots as the **Block Structure**. Note that GenStat allows the final stratum to be omitted, so you can, for this design, leave the **Block Structure** blank. Notice that we selected to output the 5% l.s.d. values. The s.e.(difference) is set as the default output; we could also have chosen to obtain the s.e.(mean). The (standardised) residual plot can be drawn once the analysis is obtained: return to the **Analysis of Variance** window, select **Further Output**, **Residual Plots** and **Standardized**.



Analysis of variance

Variate: Length

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Sugar	4	1077.320	269.330	49.37	<.001
Residual	45	245.500	5.456		
Total	49	1322.820			

Message: the following units have large residuals.

units 5	-5.10	s.e.	2.22
units 9	5.90	s.e.	2.22

Tables of means

Variate: Length

Grand mean 61.94

Sugar	Control	Glucose	Fructose	GlucFruc	Sucrose
	70.10	59.30	58.20	58.00	64.10

Standard errors of means

Table	Sugar
rep.	10
d.f.	45
e.s.e.	0.739

Standard errors of differences of means

Table	Sugar
rep.	10
d.f.	45
s.e.d.	1.045

Least significant differences of means (5% level)

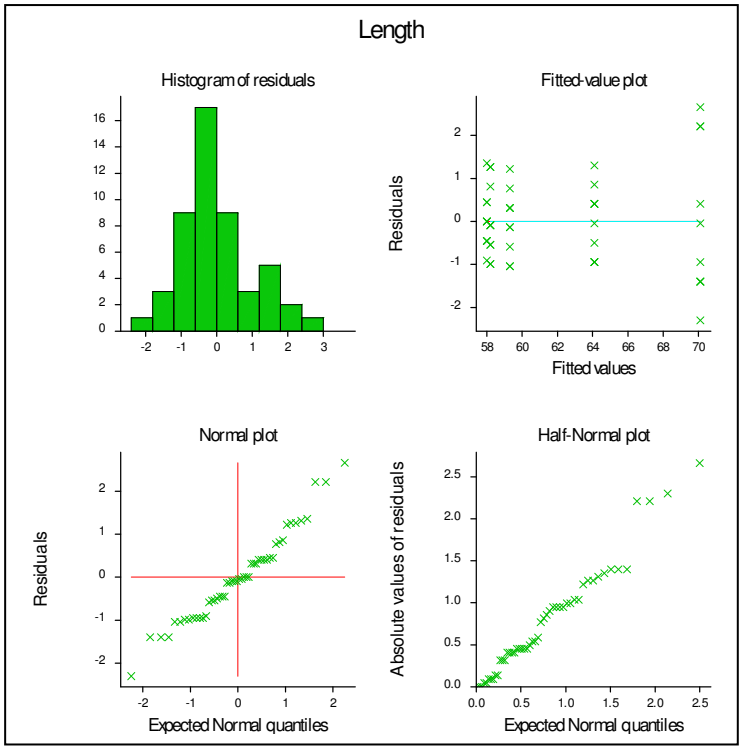
Table	Sugar
rep.	10
d.f.	45
l.s.d.	2.104

Notice:

- 5.456 is the average of the sample variances 15.878, 2.678, 3.511, 2.000, 3.211, each with (10-1) = 9 *df*.
- 269.33 is the weighted sample variance of the sugar means 70.1, 59.3, 58.2, 58.0, 64.1. Since an unweighted variance would (if the population treatment means were all equal) estimate $\sigma^2/10$, the Sugar MS is 10 × sample variance.

Before discussing the analysis in any more detail, we should inspect the (standardized) residual plot.

There are problems with this analysis. The standardised residual plot uncovers a large variance for the data in the treatment with the largest fitted value, which on inspection is the `Control` treatment. This is common in agricultural trials, and leads to special ways of analysing the data.



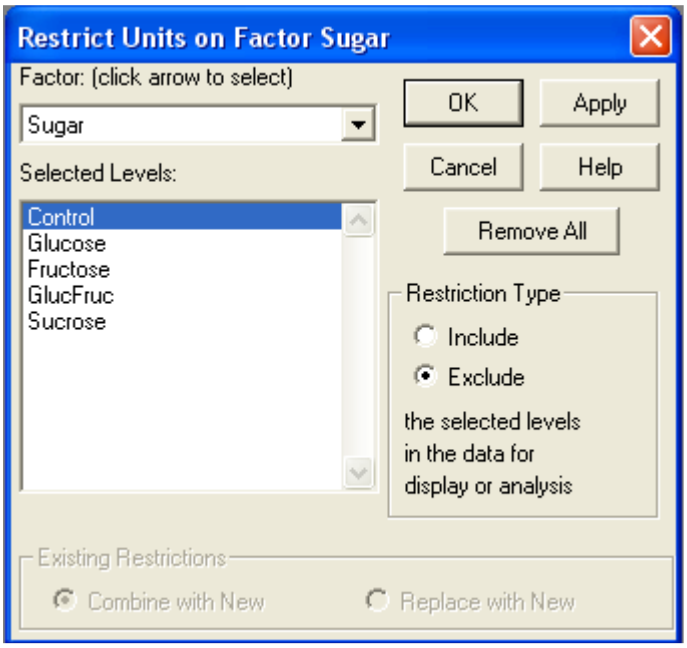
Sometimes it is possible to find a transformation that overcomes the problem, especially if the problem is one of fanning. Fanning often indicates log-normal (rather than normal) data, or data for which the variance increases as mean^2.

In this case, untreated data simply behave differently to treated data in terms of variability. One possibility is to separate out the treated and control data, and analyse these sets of data separately. The variance for the untreated data is very large (15.878 with 9 *df*) compared to the variances for the treated data (whose average is 2.850 with $4 \times 9 = 36$ *df*). Keeping the treated data allows fair comparisons among the four sugar treatments. If one really wanted to compare the control mean with one of the four sugar means, a variation of Satterthwaite's approximate *t* test (see page 39) can be used.

Alternatively, a Linear Mixed Model can be used that allows two variances, one for untreated data and another for treated data. Both tests (tests of equality of the four sugar treatment means, test of the mean of the untreated data versus the mean of the treated data) are done in the one analysis.

Restricting the analysis to a subset of treatments

There are several ways to do this, but the easiest is click inside the spreadsheet, then select **Spread** > **Restrict/Filter** > **To Groups (factor levels)**, select the Control treatment and **Exclude**.



Now click back into the **Analysis of Variance** box and click on **OK** to re-run the analysis. The sugar means are the same (as they must be) but the Control mean is left blank. The Residual MS is now only 2.850 instead of the earlier 5.456, representing a much fairer variance estimate for comparing the 4 sugar means (resulting in a reduced l.s.d. value of 1.531 instead of the earlier 2.104).

Analysis of variance

Variate: Length

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Sugar	3	245.000	81.667	28.65	<.001
Residual	36	102.600	2.850		
Total	39	347.600			

Tables of means

Grand mean 59.90

Sugar	Control	Glucose	Fructose	GlucFruc	Sucrose
		59.30	58.20	58.00	64.10

Standard errors of differences of means

Table	Sugar
rep.	10
d.f.	36
s.e.d.	0.755

Least significant differences of means (5% level)

Table	Sugar
rep.	10
d.f.	36
l.s.d. 1.531	

To compare the Control mean (which has an estimated standard deviation of $s_1 = 3.985$ with 9 *df*) with one of the 4 sugar means (which has an estimated standard deviation of $s_p = \sqrt{2.850} = 1.688$ with 36 *df*) is achieved by an extension of Satterthwaite's test.

Approximate t test of $\mu_{untreated} = \mu_{sucrose}$

Difference in means = 70.1 – 64.1 = 6.0. $sed = \sqrt{\frac{s_1^2}{n_1} + \frac{s_p^2}{n_2}} = \sqrt{1.873} = 1.368$. Hence,

$$t_{obs} = 6.0/1.368 = 4.38.$$

The degrees of freedom are calculated from a formula modified using the formula on page 34, with n_2=10 and n_2=40.

$$df = \left[\frac{\left(\frac{s_1^2}{n_1} + \frac{s_p^2}{n_2} \right)^2}{\frac{\left(s_1^2/n_1 \right)^2}{n_1 - 1} + \frac{\left(s_p^2/n_2 \right)^2}{df \text{ of } s_p^2}} \right] = 12.42.$$

There is strong statistical evidence (P<0.001) that the control and sucrose means are different. The modified df for comparing the control mean against the mean of all 4 sugar treatments (i.e for n_2=40) is 9.82.

LMM (REML) analysis of CRD (unequal variances)

Firstly, the treatment variances (each with 9 *df*) fall into two groups. The variance for the untreated pots (15.878) appears quite different to that for the treated pots. The average variance for treated pots is 2.850.

Treatment variances

Control	glucose 2%	fructose 2%	gluc_fruct 1%	sucrose 2%
15.878	2.678	3.511	2.000	3.211

As before, the **Fixed Model** is the Sugar factor with 5 levels.

The **Random Model** is Pots (a factor with levels 1 to 50). However, this model assumes that the variance is constant (Identity). We are interested in allowing the variance to change depending on the treatment.

The worst case is when every treatment has a different variance. What is believed is that only the Control treatment has a different variance.

Another way of extracting the tests of interest is

- to compare treated and untreated pots;
- for the treated pots, to compare among the four sugar treatments.

The spreadsheet can be set up with a factor (called say Control_Rest) to identify control and treated pots. We will use the label “control” to identify a control pot and a label “treated” to identify a treated pot.

Among the treated pots, the four sugar treatments can be compared using GenStat’s nested shortcut. In other words, the treatment structure is:

Fixed Model: Control_Rest/Sugar

The following choices set up difference variance structures among the treatments

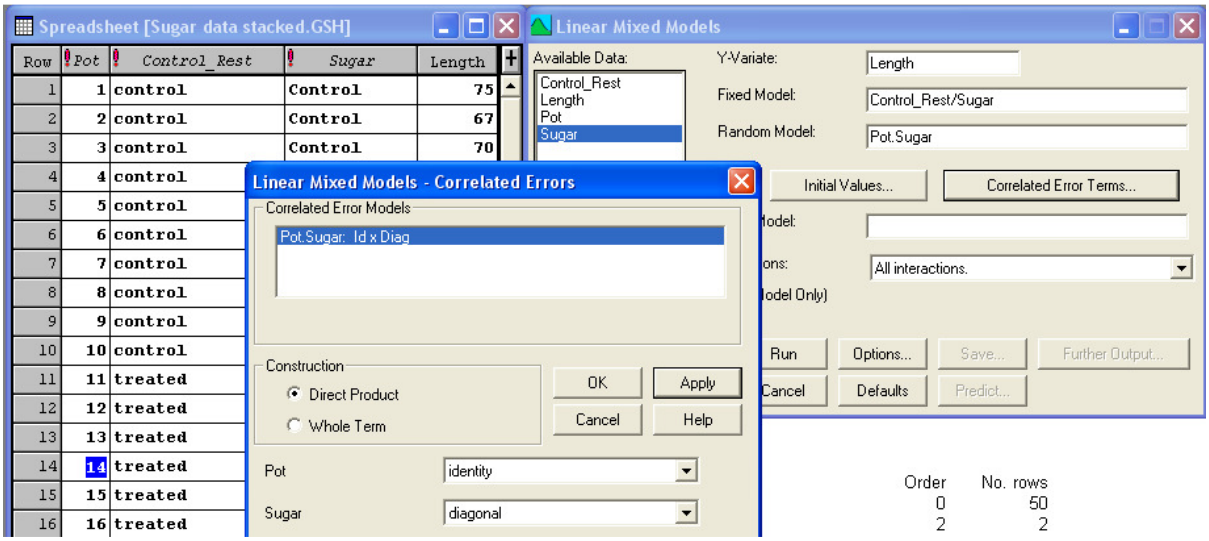
Random Model: Pots.Sugar	allows a different variance for all 5 sugar treatments by selecting **Diagonal** for Sugar in **Correlated Error Terms**
Random Model: Pots.Control_Rest	sets up one variance for the control treatment, and a separate variance for the other 4 sugar treatments, by selecting **Diagonal** for Control_Rest in **Correlated Error Terms**;
Random Model: Pots	sets up a constant variance for all 5 treatments by selecting **Identity** for Pot in **Correlated Error Terms**.

The models can be compared by change in deviance as usual.

Note that the default in GenStat is to produce multipliers rather than actual variances when selecting a **Diagonal** variance structure. To have GenStat print out the different variance estimates instead, use the

```
PARAMETERIZATION=sigmas
```

option of REML. You will need to run the default model, copy the three lines from the **Input** window, add the option and re-run the window.



The deviances for each of the three models are as follows.

Model	Random Model	Deviance	d.f.	Change in deviance	Change in d.f.	*P* value
All 5 treatment variances different	Pots.Sugar	118.3	40	0.80	3	0.849
Control variance different	Pots.Control_Rest	119.1	43	13.76	1	<0.001
Common variance	Pots	132.86	44			

Clearly allowing the control treatment to have a different variance is a better assumption than one with all variances equal ($P<0.001$); it appears unnecessary to allow all five treatments variances to be different ($P=0.849$).

Having the **Fixed Model** as Control_Rest/Sugar allows the comparison of the control treatment with the remaining sugar treatments to be equivalent to a *t* test with unequal variances. The apparent interaction Control_Rest.Sugar is actually a main effect, testing the differences among the four sugar treatments.

The full analysis is as follows (using the sigmas parameterization)..

REML variance components analysis

Response variate: Length
Fixed model: Constant + Control_Sugar_F + Control_Sugar_F.Sugar
Random model: Pots.Control_Sugar_F
Number of units: 50

Residual term has been added to model

Covariance structures defined for random model

Covariance structures defined within terms:

Term	Factor	Model	Order	No. rows
Pots.Control_Sugar_F	Pots	Identity	0	50
	Control_Sugar_F	Diagonal	2	2

Estimated parameters for covariance models

Random term(s)	Factor	Model(order)	Parameter	Estimate	s.e.
Pots.Control_Sugar_F					
	Pots	Identity	-	-	-
	Control_Sugar_F	Diagonal	d_1	14.88	7.48
			d_2	1.850	0.672

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
Residual		Identity	Sigma2	1.000	aliased

For this parameterization, individual variances are estimated to be

var(yield) = 1.000 × (14.88+1.000) = 15.88 for control data, and
= 1.000 × (1.850+1.000) = 2.85 for treated data.

Notice that 15.877 is actually the sample variance of the control data, whereas 2.850 is the average of the four sugar variances, each with 9 *df*. Hence the variance estimate for the control data has 9 *df*, while the average sugar variance has 36 *df*.

Deviance: -2*Log-Likelihood

Deviance	d.f.
119.10	43

Note: deviance omits constants which depend on fixed model fitted.

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Control_Sugar_F	62.71	1	62.71	9.8	<0.001
Control_Sugar_F.Sugar	85.96	3	28.65	36.0	<0.001

Table of predicted means for Control_Rest.Sugar

Sugar:	Control	gluc_2%	fruc_2%	gluc_fruc_1%	gluc_fruc_1%
Control_Sugar_F					
control	70.10	*	*	*	*
treated	*	59.30	58.20	58.00	64.10

Since the means have one of two estimated variances, the s.e.d. values will differ depending on whether a control mean is involved (1.37), or not (0.75). Use the **Standard Errors All Differences** option to obtain a complete set of s.e.d and l.s.d. values.

Notice the following.

- The Wald F statistic and d.f. for the (nested) component Control_Rest.Sugar are the same as those from th ANOVA of just the treated data:

Analysis of variance

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Sugar	3	245.000	81.667	**28.65**	<.001
Residual	36	102.600	2.850		
Total	39	347.600			

- The Wald F statistic and d.f. for the component Control_Sugar_F.Sugar are the same those from the Satterthwaite approximate *t* test of the control mean versus the mean of all treated pots:

Control mean = 70.1 (based on 10 observations), var = 15.878, $df = 9$
Sugar mean = 59.9 (based on 40 observations), var = 2.850, $df = 36$

Difference in means = 10.2, s.e.d. = $\sqrt{\frac{15.878}{10}+\frac{2.850}{40}} = 1.288$

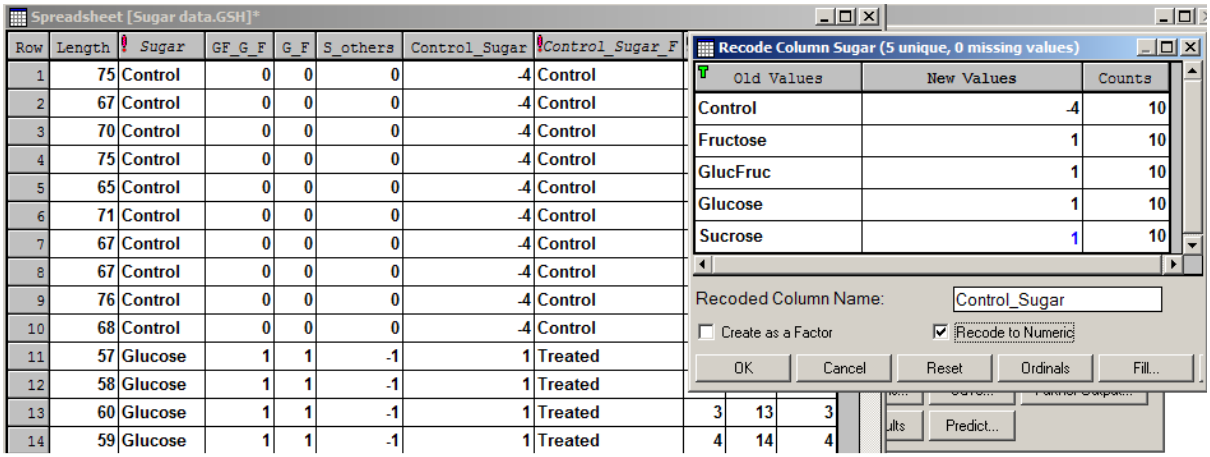
$t = 10.2/1.288 = 7.919$, or $F = t^2 = 7.919^2 = 62.711$ (d.f. calculation shown earlier).

Using contrasts in REML

There is an FCONTRASTS procedure (from version 12) that allows you to fit contrasts in REML by commands. However, we have done this directly in the spreadsheet menu choosing, by way of illustration:

(i) control vs overall sugar,
(ii) sucrose vs other sugar treatments,
(iii) glucose vs fructose, and
(iv) the mean of glucose and fructose vs the combination glucose/fructose treatment

To set these variates up, each time click in the Sugar factor column and use Spread > Factor > Recode. We need a variate and hence untick Create as a Factor and tick Recode to Numeric. Define the new values and name the contrast appropriately, as shown in the following screen capture:



Simply replace the **Fixed Model** `Control_Sugar_F/Sugar` with

```
Control_Sugar+GF_G_F+G_F+S_others
```

The output is the same as before, with individual Wald F statistics for each of the 4 contrasts instead. The design is balanced, hence test of the sequential terms and dropping each term last are the same.

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Control_Sugar	62.71	1	62.71	9.8	<0.001
GF_G_F	1.32	1	1.32	36.0	0.259
G_F	2.12	1	2.12	36.0	0.154
S_others	82.53	1	82.53	36.0	<0.001

Dropping individual terms from full fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Control_Sugar	62.71	1	62.71	9.8	<0.001
GF_G_F	1.32	1	1.32	36.0	0.259
G_F	2.12	1	2.12	36.0	0.154
S_others	82.53	1	82.53	36.0	<0.001

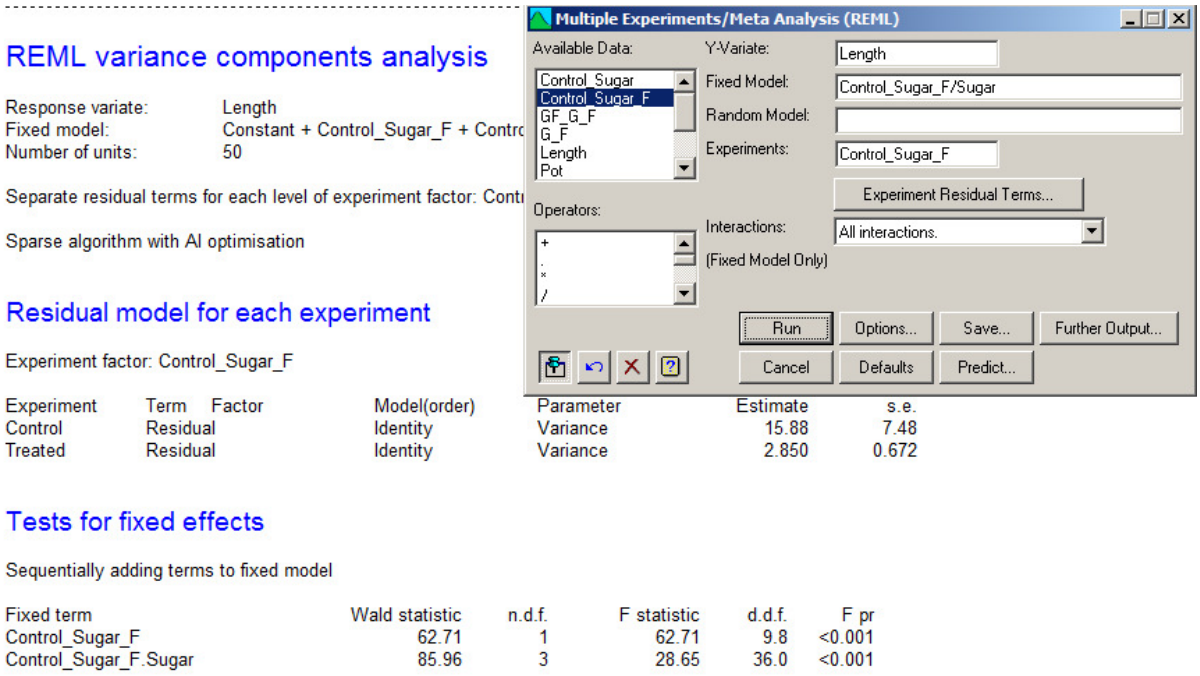
Meta Analysis - REML of Multiple Experiments menu

Prof Roger Payne kindly pointed out a more simple method of obtaining the analysis where the variance changes across (part of) one or more factors. This menu allows you to specify a changing variance across different experiments. In this case, we imagine that the control pots come from a separate experiment than the treated pots.

The **Fixed Model** is either Control_Rest/Sugar or `Control_Sugar+GF_G_F+G_F+S_others` as before.

The **Random Model** is Pots, since the changing variance is declared in the next line. Pots can be omitted, as is usual for a simple CRD (since GenStat adds an error term if one is not provided).

In this case, on the **Experiments** line simply indicate the factor `Control_Sugar_F` that contains the information to identify how the variance changes.



The output is the same as before with the exception of a more simple presentation of the variance estimates:

Residual model for each experiment

Experiment factor: Control_Sugar_F

Experiment	Term	Factor	Model(order)	Parameter	Estimate	s.e.
Control	Residual		Identity	Variance	15.88	7.48
Treated	Residual		Identity	Variance	2.850	0.672

Two-way design (no blocking) with subsamples

Mint plants were assigned at random to pots, 4 plants per pot, 18 pots in all and grown in a nutrient solution. Three pots were randomly assigned to one of six treatment combinations, as follows. All pots were randomly located during the time spent at either 8, 12 or 16 hours of daylight. Each group of pots was completely randomized within low- or high-temperature greenhouses during the time spent in darkness. Individual plants stem lengths were measured after one week.

Example 8 One week stem lengths (cm, Steel and Torrie pages 153-9)

Temperature																		
High										Low								
Hours of Daylight										Hours of Daylight								
8			12			16				8			12			16		
pot			pot			pot				pot			Pot			pot		
1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	
3.5	2.5	3.0	5.0	3.5	4.5	5.0	5.5	5.5	8.5	6.5	7.0	6.0	6.0	6.5	7.0	6.0	11.0	
4.0	4.5	3.0	5.5	3.5	4.0	4.5	6.0	4.5	6.0	7.0	7.0	5.5	8.5	6.5	9.0	7.0	7.0	
3.0	5.5	2.5	4.0	3.0	4.0	5.0	5.0	6.5	9.0	8.0	7.0	3.5	4.5	8.5	8.5	7.0	9.0	
4.5	5.0	3.0	3.5	4.0	5.0	4.5	5.0	5.5	8.5	6.5	7.0	7.0	7.5	7.5	8.5	7.0	8.0	

This design is slightly complex, in that half the pots have a restricted randomization for the time spent in one of the two greenhouses, each set at a different temperature. Ignoring that problem, it is clear that pots form replicates for the six treatment combinations: a pot containing 4 plants is moved to a random daylight position and a random position in a greenhouse; the 4 plants form sampling units.

Treatment Structure

You need to supply two factor columns, properly labeled, to identify the six **Temperature** and **Light** treatment combinations applied to each pot. The **Treatment Structure** is then **Temperature** + **Light** + **Temperature.Light**. By the Rule 2 simplifies to **Temperature***Light**.

Block Structure

Choice 1

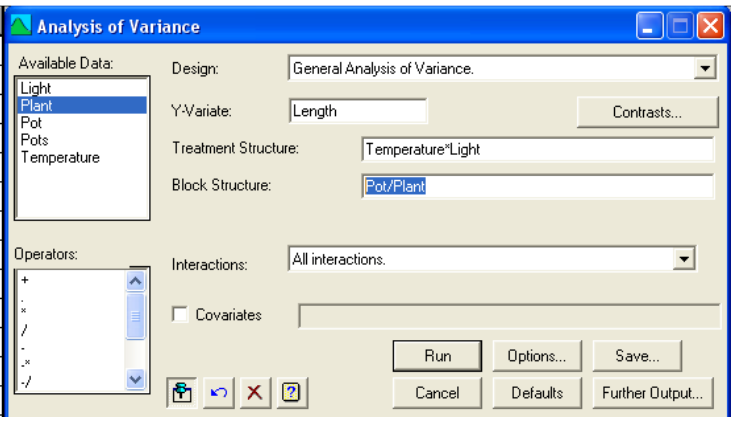
Generally we recommend that the replicates be numbered from 1 to the total number of replicates, across all treatments . There are 18 pot replicates, and in our spreadsheet we called this column **Pots**. Plants in pots are samples. There are two strata, and hence the **Block Structure** is **Pots+Pots.Plant**. By Rule 3 this simplifies to **Pots/Plant**. GenStat also allows the final error term to be omitted, so **Pots** is also permissible.

Choice 2 (not recommended)

Steel and Torrie, however, used 1, 2, 3 for each treatment combination, so we differentiate this factor as **Pot**. If you decide to use this numbering system, then the **Block Structure** *cannot* be **Pot/Plant**: as mentioned, this expands to **Pot+Pot.Plant**, and GenStat will assume that Pot #1 in every treatment is a block. Rather, you need to use **Pot.Treatment/Plant**, which expands to **Pot.Treatment + Pot.Treatment.Plant**. Here, **Treatment** is a factor that enumerates all six treatments and **Pot** has levels 1, 2, 3. We don't have such a treatment factor column, so you would need to **Insert** a new column and **Fill** this column from 1 to 6, each number repeated nine times. The analysis is identical to that obtained in *Choice 1*.

Analysis of Two-way Design (no Blocking) with subsamples

Row	Pots	Plant	Light	Temperature	Length
1	1	1	8	High	3.5
2	1	2	8	High	4
3	1	3	8	High	3
4	1	4	8	High	4.5
5	2	1	8	High	2.5
6	2	2	8	High	4.5
7	2	3	8	High	5.5
8	2	4	8	High	5
9	3	1	8	High	3
10	3	2	8	High	3
11	3	3	8	High	2.5
12	3	4	8	High	3
13	4	1	12	High	5



Analysis of variance

Variate: Length

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Pots stratum					
Temperature	1	151.6701	151.6701	70.45	<.001
Light	2	22.2986	11.1493	5.18	0.024
Temperature.Light	**2**	**5.6736**	**2.8368**	**1.32**	**0.304**
Residual	12	25.8333	2.1528	2.30	
Pots.Plant stratum	54	50.4375	0.9340		
Total	71	255.9132			

…

Tests these means

Tables of means

Variate: Length

Grand mean 5.78

Temperature	High	Low
	4.33	7.24

Light	8.	12.	16.
	5.50	5.29	6.56

Temperature	Light	8.	12.	16.
High		3.67	4.12	5.21
Low		7.33	6.46	7.92

Standard errors of differences of means

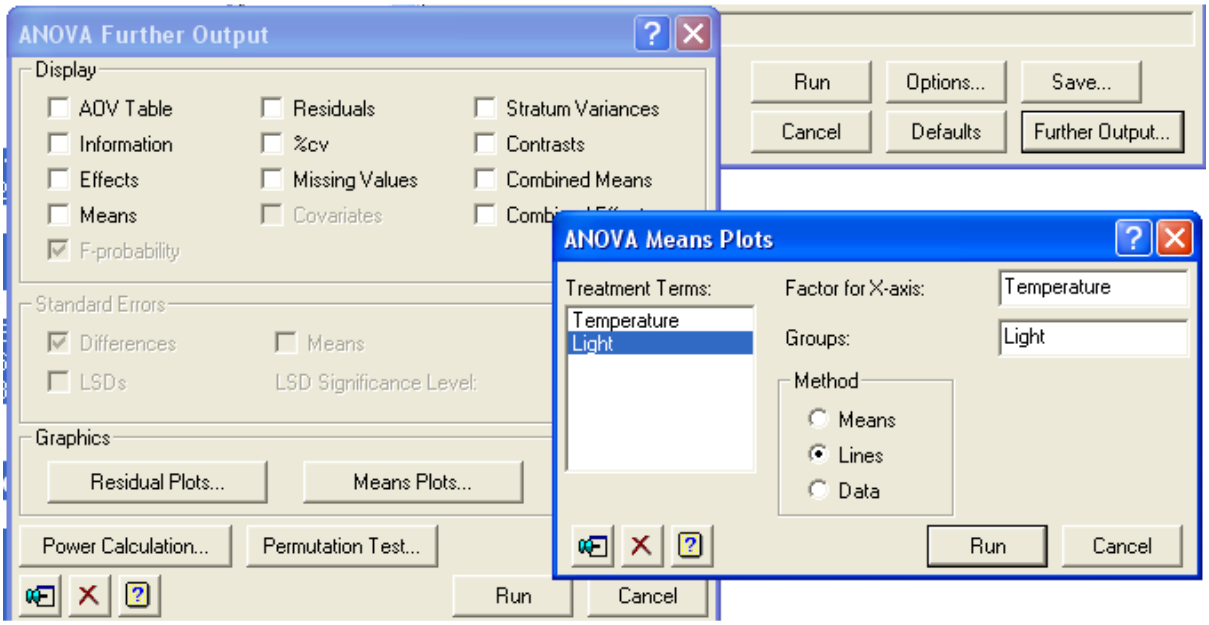
Table	Temperature	Light	Temperature Light
rep.	36	24	12
d.f.	12	12	12
s.e.d.	0.346	0.424	0.599

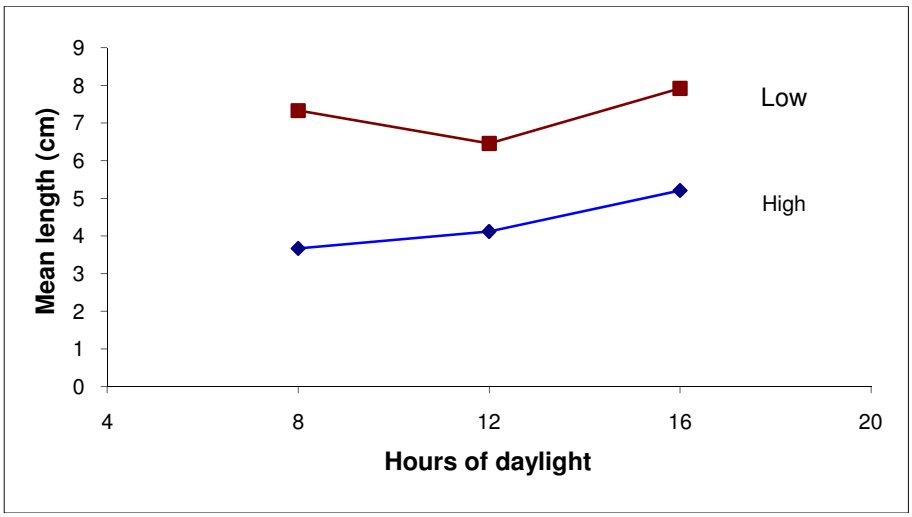
Least significant differences of means (5% level)

Table	Temperature	Light	Temperature Light
rep.	36	24	12
d.f.	12	12	12
l.s.d.	0.753	0.923	1.305

When interpreting this analysis, *it is important to interpret the interaction first (for more complex designs, from highest-order interaction backwards).* A two-way interaction tests whether any change in the response of the plant to temperature is consistent for both high and low temperatures. Thus, it examines the response to temperature in the following table. The response is best plotted (**Further Output > Means Plot**).

	Hours of light		
Temperature	8	12	16
High	3.67	4.12	5.21
Low	7.33	6.46	7.92





The responses are parallel within statistical variation ($P = 0.304$). Hence, attention can focus on the average effect of temperature, as well as the average effect of light. These are known as *main effects*. Both are strongly significant – see the ANOVA table.

Interest focuses on how much variation is there from plant to plant (the sampling variance) as opposed to pot to pot variation. Note that each of 6 treatments provides (3-1) = 12 residual df for estimating σ^2.

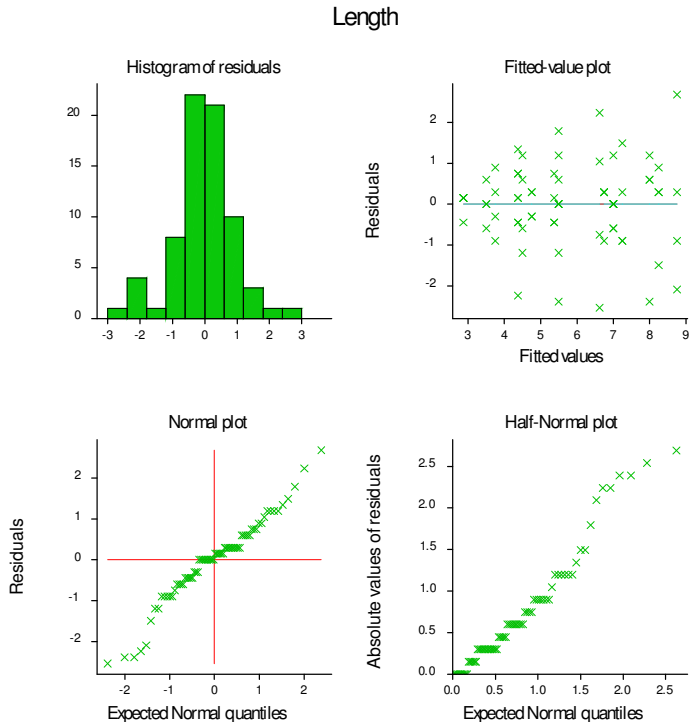
Estimates of the sampling and experimental variances are obtained by clicking on **Stratum Variances** in **Options** prior to running the analysis. The output is the following. There is three times more variation between plants in a pot than between pots.

Estimated stratum variances

Variate: Length

Stratum	variance	effective d.f.	variance component	
Pots	2.153	12.000	0.305	variance among pots
Pots.Plant	0.934	54.000	0.934	variance among plants in a pot

Finally, below is the standardised residual plot. You can make up your own mind whether the variation across all sampling units is constant.



LMM (REML) analysis

The **Treatment Structure** is Temperature*Light and the **Block Structure** is Pots.Plants.

Here is the LMM (REML) analysis. The means are as before and are suppressed in this output.

REML variance components analysis

Response variate: Length
Fixed model: Constant + Light + Temperature + Light.Temperature
Random model: Pot + Pot.Plant
Number of units: 72

Pot.Plant used as residual term

Estimated variance components

Random term	component	s.e.
Pot	0.3047	0.2243

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
	Pot.Plant	Identity	Sigma2	0.934	0.1798

Approximate stratum variances

Use Fisher scoring to obtain this

Stratum	variance	effective d.f.
Pot	**2.1528**	**12.00**
Pot.Plant	**0.9340**	**54.00**

Wald tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Light	10.36	2	5.18	12.0	0.024
Temperature	70.45	1	70.45	12.0	<0.001
Light.Temperature	2.64	2	1.32	12.0	0.304

Notice:

- The variance estimates (and *df*) are the same as obtained from ANOVA;
- The F statistics and P values are the same as those from the ANOVA.

Two-way design (in randomized blocks)

Snedecor and Cochran present the yields of cowpea hay (pounds per 1/100 Morgen plot) from 3 varieties, each grown with 3 row spacings (4", 8" and 12" apart).

Firstly, let's use GenStat's **Design** menu to generate a field plan (the monograph does not give us a field layout). One random design is the following:

Spreadsheet [Book;3]*

Row	PlotNo	Block	Plots	Variety	Spacing
1	11	1	1	1	4
2	12	1	2	1	12
3	13	1	3	3	4
4	14	1	4	1	8
5	15	1	5	2	12
6	16	1	6	3	8
7	17	1	7	3	12
8	18	1	8	2	4
9	19	1	9	2	8
10	21	2	1	3	8
11	22	2	2	3	4
12	23	2	3	1	4
13	24	2	4	2	8
14	25	2	5	1	8
15	26	2	6	3	12
16	27	2	7	2	4
17	28	2	8	2	12
18	29	2	9	1	12
19	31	3	1	1	4
20	32	3	2	3	4
21	33	3	3	2	8
22	34	3	4	3	8
23	35	3	5	2	4
24	36	3	6	1	8
25	37	3	7	2	12
26	38	3	8	3	12
27	39	3	9	1	12

	BLOCK 1	BLOCK 2	BLOCK 3
1	Variety 1 **Spaced 4"**	Variety 3 **Spaced 8"**	Variety 1 **Spaced 4"**
2	Variety 1 **Spaced 12"**	Variety 3 **Spaced 4"**	Variety 3 **Spaced 4"**
3	Variety 3 **Spaced 4"**	Variety 1 **Spaced 4"**	Variety 2 **Spaced 8"**
4	Variety 1 **Spaced 8"**	Variety 2 **Spaced 8"**	Variety 3 **Spaced 8"**
5	Variety 2 **Spaced 12"**	Variety 1 **Spaced 8"**	Variety 2 **Spaced 4"**
6	Variety 3 **Spaced 8"**	Variety 3 **Spaced 12"**	Variety 1 **Spaced 8"**
7	Variety 3 **Spaced 12"**	Variety 2 **Spaced 4"**	Variety 2 **Spaced 12"**
8	Variety 2 **Spaced 4"**	Variety 2 **Spaced 12"**	Variety 3 **Spaced 12"**
9	Variety 2 **Spaced 8"**	Variety 1 **Spaced 12"**	Variety 1 **Spaced 12"**

Note that spacing experiments, by definition, are unlikely to produce plot mean (or plot total) yields *whose variances are constant*. Why is that?

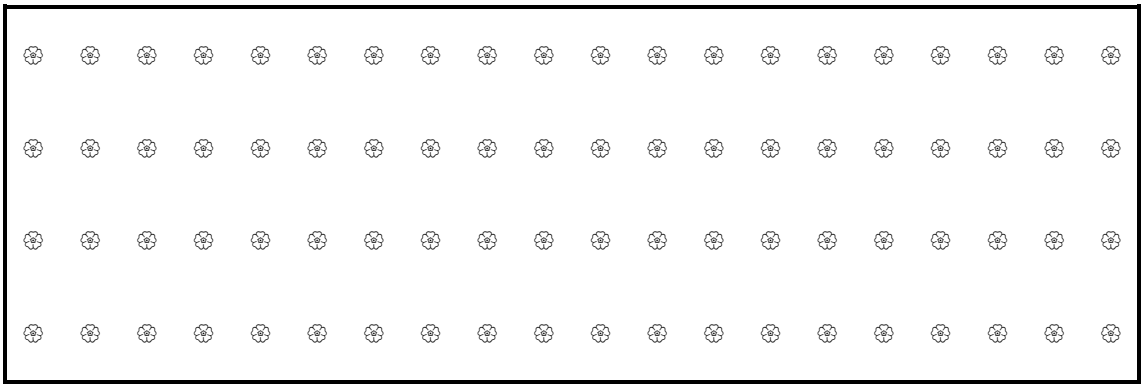
From statistical theory, if you add independent variates whose individual variances are the same, the variance of the sum is the sum of the individual variances. Let σ^2 be the variance *on a per plant basis*. Then, for independently growing plants,

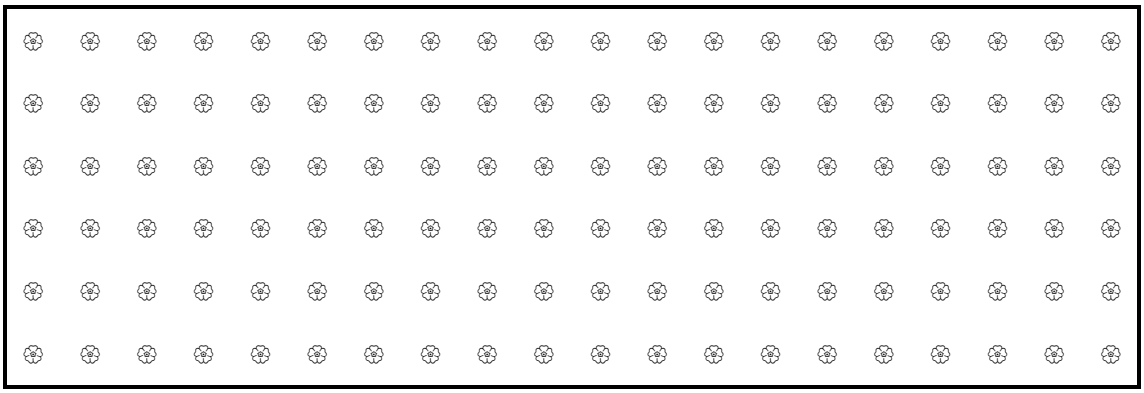
$$\text{var}(\textit{Total yield}) = \text{var}(Y_1 + \ldots + Y_n) = n\,\sigma^2$$

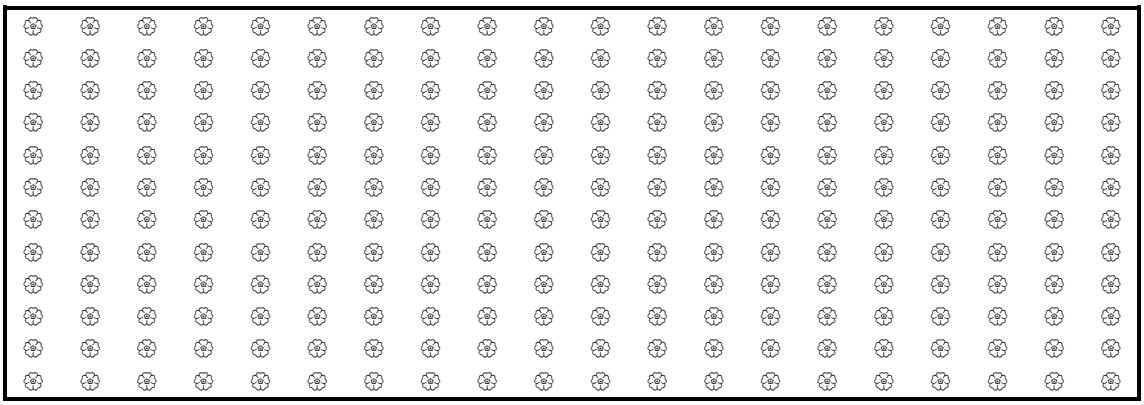
and hence

$$\text{var}(\textit{Mean yield}) = \text{var}(\bar{y}) = \sigma^2 / n$$

Now put that in the context of this spacing experiment. The plot area is 0.01 Morgen which is about 86 m^2. Spacings are about 10, 20, 30 cm. The number of rows of varying shapes depends on the shape of the plot. We'll assume for illustration that we have multiples of 1.2m areas for rows. The 12" spacing is equivalent to 30cm row spacing, so 4 rows are used at that spacing, 6 rows at 20cm spacing and 12 rows at 10cm spacing.







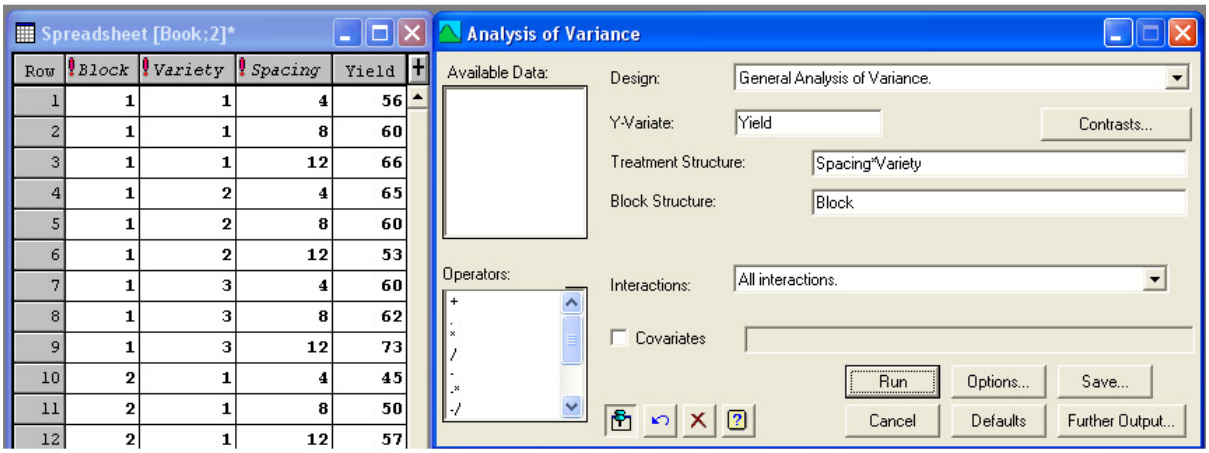
Plots like these (with row spacings 12", 8", 4") consist of varying numbers of plants (in the ratio 4:6:12). Other combinations are possible. The point is, total yield (or mean yield) obtained from plots with varying numbers of plants will have changing variance *if the plants grow independently*.

With plant competition, the variance of total yield could well even out across all shaped plots. Plant competition means that the yields become spatially correlated. We will ignore this problem for the moment. Changing variance and correlated yield models are available in **Linear Mixed Models (REML)**.

Example 9 Yields (pounds) of cowpea hay from Snedecor and Cochran, page 309.

Variety	Spacing	Block 1	Block 2	Block 3	Block 4
I	4	56	45	43	46
	8	60	50	45	48
	12	66	57	50	50
II	4	65	61	60	63
	8	60	58	56	60
	12	53	53	48	55
III	4	60	61	50	53
	8	62	68	67	60
	12	73	77	77	65

There are two strata in this experiment, **Block** and **Block.Plot**. The **Block Structure** is therefore **Block** + **Block.Plot**, or simply **Block/Plot**. Since the smallest stratum can be omitted, **Block** is sufficient.



The full analysis of the data, including L.S.D. values and stratum variances, is as follows.

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	3	255.64	85.21	4.82	
Block.*Units* stratum					
Variety	2	1027.39	513.69	29.07	<.001
Spacing	2	155.06	77.53	4.39	0.024
Variety.Spacing	4	765.44	191.36	10.83	<.001
Residual	24	424.11	17.67		
Total	35	2627.64			

Tables of means

Variate: Yield

Grand mean 57.81

Variety	1	2	3
	51.33	57.67	64.42

Spacing	4.	8.	12.
	55.25	57.83	60.33

Variety	Spacing	4.	8.	12.
1		47.50	50.75	55.75
2		62.25	58.50	52.25
3		56.00	64.25	73.00

Standard errors of differences of means

Table	Variety	Spacing	Variety Spacing
rep.	12	12	4
d.f.	24	24	24
s.e.d.	1.716	1.716	2.972

Least significant differences of means (5% level)

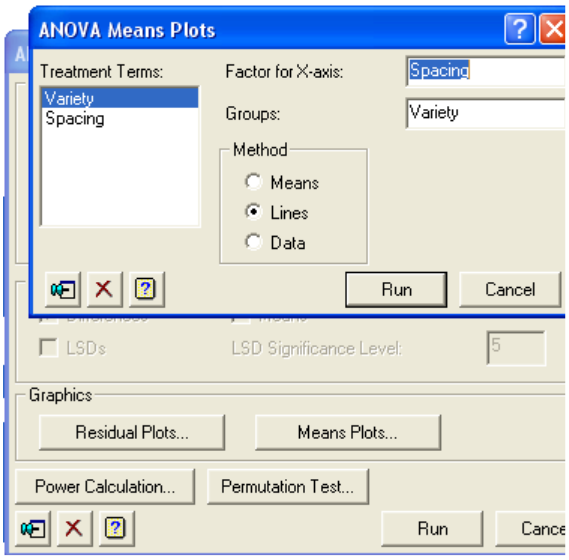
Table	Variety	Spacing	Variety Spacing
rep.	12	12	4
d.f.	24	24	24
l.s.d.	3.542	3.542	6.135

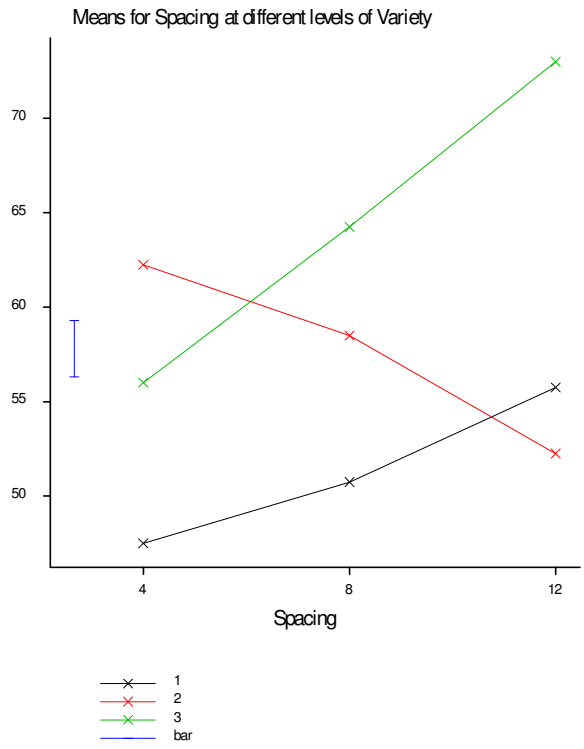
Estimated stratum variances

Variate: Yield

Stratum	variance	effective d.f.	variance component
Block	85.213	3.000	7.505
Block.*Units*	17.671	24.000	17.671

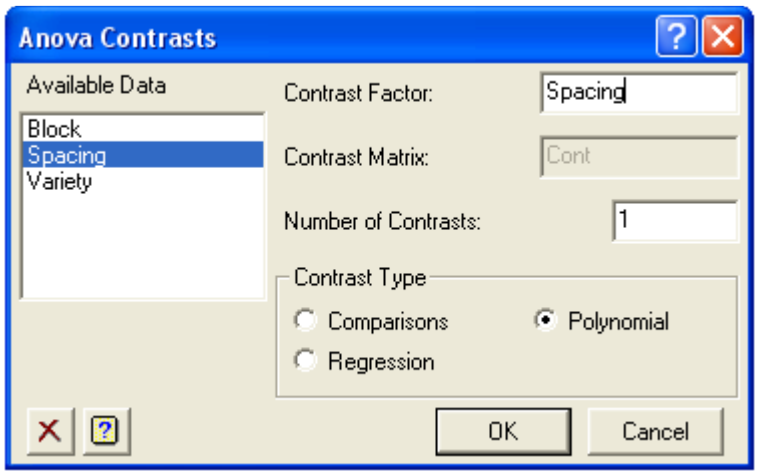
There is strong statistical evidence ($P<0.001$) that the change in mean yield at different row spacings is not the same for all three varieties. A means plot illuminates the differences:





There is a strong linear trend in mean yield, but the means for variety 2 *decrease* with increasing spacing. Varieties 1 and 3 must have heavy vegetative growth that requires at least 12" to approach optimal yield.
These linear trends can be incorporated into the ANOVA, using the **Contrast** button on the ANOVA table.

Using the Contrast Matrix
Firstly, for the factor **Spacing** we are interested in a linear trend: this is a situation where **POL** (polynomial regression/contrast) can be used.



Click on the **Contrast** button, select the **Spacing** factor and nominate **Polynomial**. The degree of the polynomial you wish to fit is the **Number of Contrasts**. In this case leave this as 1 and click **OK**. GenStat replaces **Spacing** in the treatment structure with **POL(Spacing;1)**.

We are also interested in sub-hypotheses for the **Variety** factor. In this case, two are more natural than other choices:

- H_0: Variety 1 and Variety 3 means are equal: we wish to assess $\mu_3-\mu_1$.
- H_0: Variety 2 mean and the *average mean* of Variety 1 and Variety 3 are equal: we wish to assess $(\mu_3+\mu_1)/2-\mu_2$.

Contrasts are simply the coefficients of the means in the questions asked. For any contrast, the coefficients will add to zero. GenStat allows two types of questions, labelled **Comparisons** and **Regression**.

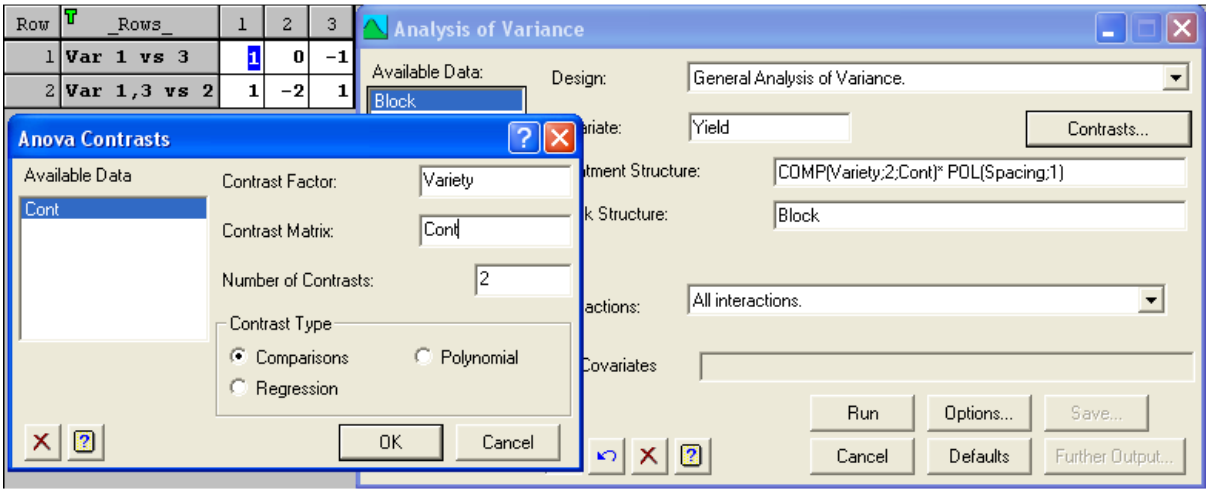
Comparisons allows any number of questions to be asked, with no restrictions on the questions asked. Their component sums of squares will not add to the Variety SS.

For *t* treatments, **Regression** allows up to (*t*-1) questions, with restrictions on the questions asked. The questions must be *orthogonal*, that is, balanced in a special way. The component sums of squares for all (*t*-1) contrasts will add to the Treatment SS. Even if the contrasts are orthogonal, the **Comparisons** choice can be used. The only difference is that GenStat does not report deviations when **Comparisons** is selected.

- Variety 1 vs 3: $\mu_3-\mu_1$ is equivalent to (-1, 0 ,1) multipliers of (μ_1, μ_2, μ_3) respectively

- Variety 1&3 vs 2: $(\mu_3+\mu_1)/2-\mu_2$ is equivalent to (½, -1, ½) multipliers of (μ_1, μ_2, μ_3). It is preferable to enter integers rather than fractions, so multiplier by a constant (in this case 2) to remove fractions. The contrast is then (1, -2, 1)

Click on the **Contrast** button, select the **Variety** factor and nominate **Regression** and enter the **Number of Contrasts** you wish to make (here 2). GenStat opens up a table (which is names, by default, **Cont**, or **Cont_1** if **Cont** exists) with (here) 2 rows (questions) and 3 columns (levels). Names of the levels are placed above the columns. Enter the contrast coefficients, and double click on the grey areas of the rows, where the names of each contrast can be set up. Then return to the **ANOVA Contrasts** menu and click **OK**. GenStat replaces **Variety** in the treatment structure with **REG(Variety;2;Cont)** or **COMP(Variety;2;Cont)** if you chose **Comparisons**.



The new ANOVA table is as follows.

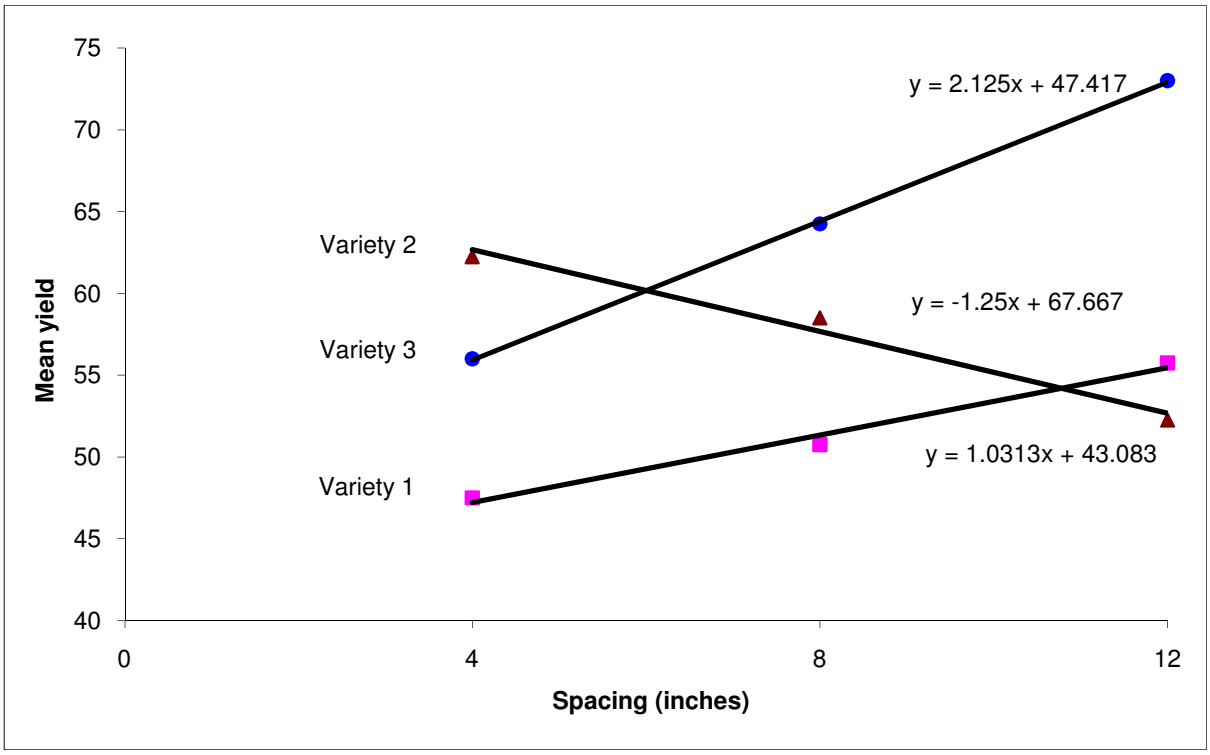
Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	3	255.64	85.21	4.82	
Block.*Units* stratum					
Variety	2	1027.39	513.69	29.07	<.001
Var 1 vs 3	1	1027.04	1027.04	58.12	<.001
Var 1,3 vs 2	1	0.35	0.35	0.02	0.890
Spacing	2	155.06	77.53	4.39	0.024
Lin	1	155.04	155.04	8.77	0.007
Deviations	1	0.01	0.01	0.00	0.978
Variety.Spacing	4	765.44	191.36	10.83	<.001
Var 1 vs 3.Lin	1	76.56	76.56	4.33	0.048
Var 1,3 vs 2.Lin	1	682.52	682.52	38.62	<.001
Var 1 vs 3.Dev	1	0.52	0.52	0.03	0.865
Var 1,3 vs 2. Dev	1	5.84	5.84	0.33	0.571
Residual	24	424.11	17.67		
Total	35	2627.64			

Note that with 3 spacing levels, Dev is identical to the quadratic term. With 4 spacing levels and a linear model requested, Dev will be the combined quadratic and cubic components: it's what is left after the requested polynomial is fitted. This table adds the following to what we knew already. The slope in the regression of the means of varieties 1 and 3 are marginally different (P=0.048), whereas the slope for variety 2 in comparison is strikingly different (P<0.001) to an average slope for variety 1 and 3 means.

Here are trend lines added in Excel:



If we just wish to estimate the fitted regressions using GenStat, it is easier to use a general regression ignoring blocks (because the design is orthogonal). The factor column **Spacing** needs to be converted to a variate instead (simply point to the column, right click and select **Convert to Variate**). The **Model to be fitted** is **Variety*Spacing**. *We are using this model simply to obtain the linear equations, not to test hypotheses.*

Estimates of parameters

Parameter	estimate	s.e.	t(30)	t pr.
Constant	43.08	3.65	11.80	<.001
Spacing	1.031	0.423	2.44	0.021
Variety 2	24.58	5.17	4.76	<.001
Variety 3	4.33	5.17	0.84	0.408
Spacing.Variety 2	-2.281	0.598	-3.82	<.001
Spacing.Variety 3	1.094	0.598	1.83	0.077

Parameters for factors are differences compared with the reference level:

Factor	Reference level
Variety	1

The model for the reference Variety 1 comes out immediately:
Mean yield = 43.08 + 1.031 *Spacing*

For variety 2 we add 24.58 to the intercept and -2.281 to the slope:
Mean yield = 67.66 - 1.250 *Spacing*

For variety 3 we add 4.33 to the intercept and 1.094 to the slope:
Mean yield = 47.41 + 2.125 *Spacing*

LMM (REML) analysis

The **Treatment Structure** is Spacing*Variety and the **Block Structure** is Block/Plot. In the earlier discussion, there was consideration about whether the variance was constant, proportional to the number of plants in a plot, or somewhere in between. We explore these issues using change in deviance.

The estimates of the stratum variances were:

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Block	85.213	3.000	7.505
Block.*Units*	17.671	24.000	17.671

In order to allow a changing variance model for different spacings, we need to ensure that Spacing appears in the **Block Structure** so we can use **Correlated Error Terms**. We can change Block/Plot for an expression in which the Plot part is replaced by a factor expression which ranges over the same set of values. Plot goes from 1 to 9 in each block. These track which combination of variety and spacing is used in each plot. Hence an equivalent expression for the **Block Structure** is Block.Spacing.Variety. The deviances for common variance (Identity) and variances changing over Spacing levels (Diagonal) are as follows:

	deviance	d.f.	Change in deviance	Change in d.f.	P value
Identity	121.74	25			
Diagonal	120.37	23	1.37	2	0.504

For this experiment, there is no evidence that a changing variance model is necessary (P=0.504). The rest of the analysis gives the same variance estimates and equivalent test values as for ANOVA.

REML variance components analysis

Response variate: Yield
Fixed model: Constant + Variety + Spacing + Variety.Spacing
Random model: Block + Block.Variety.Spacing
Number of units: 36

Block.Variety.Spacing used as residual term

Estimated variance components

Random term	component	s.e.
Block	**7.50**	7.75

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
	Block.Variety.Spacing	Identity	Sigma2	17.67	5.10

Wald tests for fixed effects

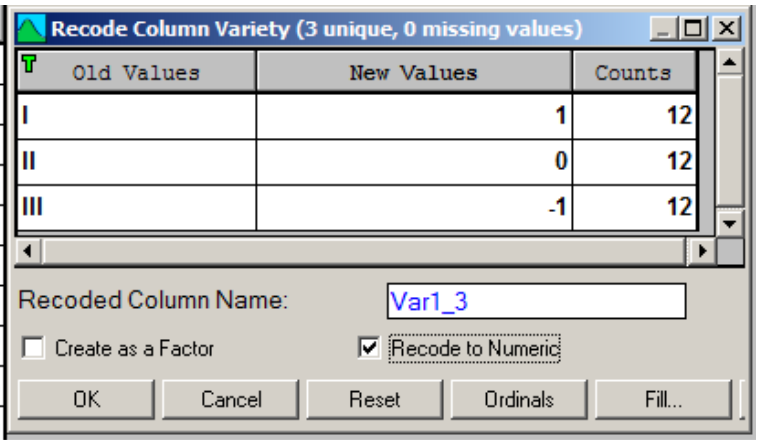
Fixed term	Wald statistic	d.f.	Wald/d.f.	chi pr
Variety	58.14	2	29.07	<0.001
Spacing	8.77	2	4.39	0.012
Variety.Spacing	43.32	4	10.83	<0.001

Using contrasts in REML

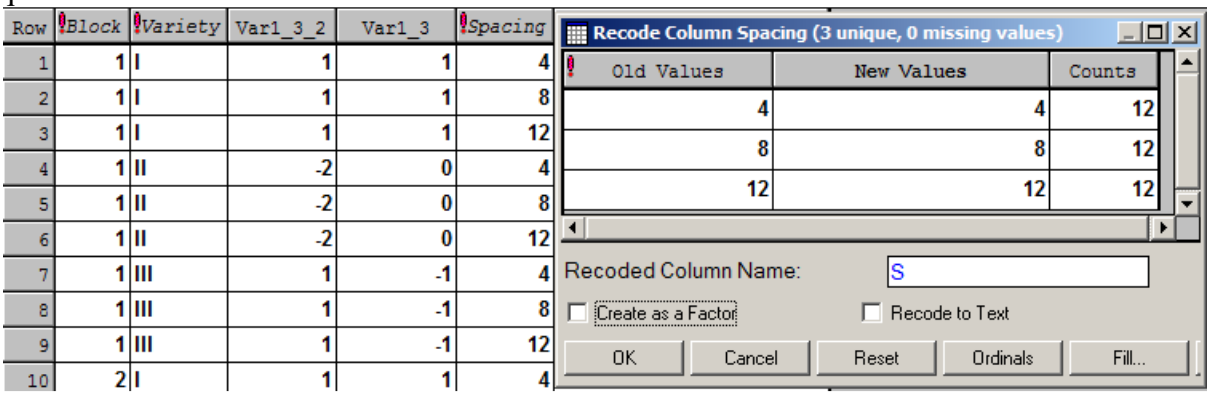
We will do this directly by replacing the two factors with variates that represent the contrasts and trends.

For Variety contrasts, click in the Variety column and use Spread > Factor > Recode. We need a variate and hence untick Create as a Factor and tick Recode to Numeric. Use the same contrasts as for ANOVA:

Row	Block	Variety	Spacing	Yield
1	1	I	4	56
2	1	I	8	60
3	1	I	12	66
4	1	II	4	65
5	1	II	8	60
6	1	II	12	53
7	1	III	4	60
8	1	III	8	62
9	1	III	12	73
10	2	I	4	45

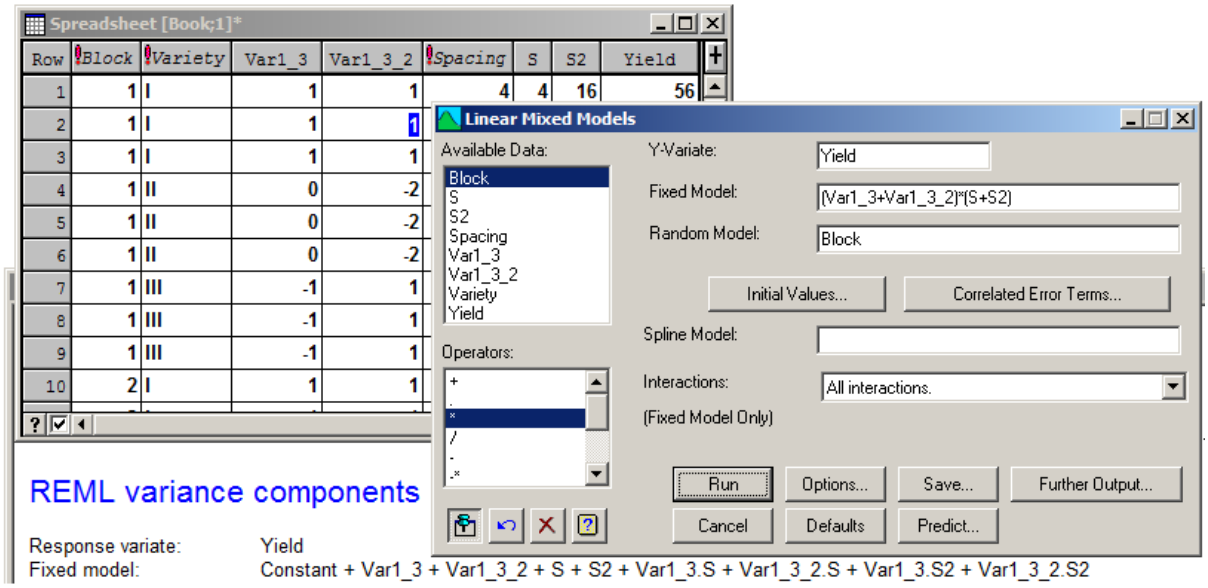


For Spacing trends, click in the Spacing column and use Spread > Factor > Recode. There are spacing levels already defined, so simply untick Create as a Factor and name the new column S (say). Repeat and use squared spacing levels for a column named S2 (say) representing the quadratic trend.



Row	Block	Variety	Var1_3_2	Var1_3	Spacing
1	1	I	1	1	4
2	1	I	1	1	8
3	1	I	1	1	12
4	1	II	-2	0	4
5	1	II	-2	0	8
6	1	II	-2	0	12
7	1	III	1	-1	4
8	1	III	1	-1	8
9	1	III	1	-1	12
10	2	I	1	1	4

Here we are not using orthogonal polynomials for Spacing, and so we need to examine the Wald statistics *sequentially* – i.e. we ignore the P Wald statistics in Dropping individual terms from full fixed model. Each factor in the fixed model Variety*Spacing is replaced by the two variate contrasts/polynomials, so (Var1_3+Var1_3_2)*(S+S2):



REML variance components analysis

Response variate: Yield
Fixed model: Constant + Var1_3 + Var1_3_2 + S + S2 + Var1_3.S + Var1_3_2.S + Var1_3.S2 + Var1_3_2.S2
Random model: Block
Number of units: 36

Residual term has been added to model

Sparse algorithm with AI optimisation
All covariates centred

Estimated variance components

Random term	component	s.e.
Block	7.50	7.75

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
Residual		Identity	Sigma2	17.67	5.10

Deviance: -2*Log-Likelihood

Deviance	d.f.
161.60	25

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Var1_3	58.12	1	58.12	24.0	<0.001
Var1_3_2	0.02	1	0.02	24.0	0.890
S	8.77	1	8.77	24.0	0.007
S2	0.00	1	0.00	24.0	0.978
Var1_3.S	4.33	1	4.33	24.0	0.048
Var1_3_2.S	38.62	1	38.62	24.0	<0.001
Var1_3.S2	0.03	1	0.03	24.0	0.865
Var1_3_2.S2	0.33	1	0.33	24.0	0.571

- These P values are the same as those in the ANOVA.

Illustration that assuming *blocks are random* does not affect the test of fixed treatments

The tests of fixed effects from a REML analysis with Block a *random* component are:

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Variety	58.14	2	**29.07**	24.0	**<0.001**
Spacing	8.77	2	**4.39**	24.0	**0.024**
Variety.Spacing	43.32	4	**10.83**	24.0	**<0.001**

With Block a *fixed* component we obtain:

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Block	14.47	3	4.82	24.0	0.009
Variety	58.14	2	**29.07**	24.0	**<0.001**
Spacing	8.77	2	**4.39**	24.0	**0.024**
Variety.Spacing	43.32	4	**10.83**	24.0	**<0.001**

F statistics and P values for the two main effects and the interaction are unchanged. In the second analysis there is an additional test of the fixed block effects. For the first analysis there is a variance component instead for the random block term.

The means are also unchanged. However, the standard errors of individual means will be larger for the random block model, since the treatment means all involve an additional random block term. Standard errors of differences, however, are unchanged, since this block term cancels out in the difference (assuming a balanced design). Hence decisions based on comparing means are also unaffected by the assumption about blocks.

For example, the standard error of a varietal mean is 1.214 when blocks are assumed fixed, but 1.830 when they are random; the standard error of a difference is 1.716 in both cases.

Illustration that assuming *blocks are random* is equivalent to a *uniform correlated error structure*

Take any two plots (say plot *j* and plot *k*) in block *i*. The simple RCBD model with fixed treatments implies

$Y_{ij} = mean + Block_i + Treatment_j + Error_{ij}$
and
$Y_{ik} = mean + Block_i + Treatment_k + Error_{ik}$

Since $Block_i \sim N\left(0,\sigma^2_{Block}\right)$ independently of $Error_{ij} \sim N\left(0,\sigma^2\right)$ we obtain

$$var\left(Y_{ij}\right) = var\left(Y_{ik}\right) = \sigma^2_{Block} + \sigma^2$$
and
$$covar\left(Y_{ij},Y_{ik}\right) = \sigma^2_{Block}$$

giving the following correlation between the two plots:

$$corr\left(Y_{ij},Y_{ik}\right) = \frac{\sigma^2_{Block}}{\sigma^2_{Block}+\sigma^2} = \theta \text{ say.}$$

The estimated stratum variances from the ANOVA are $\hat{\sigma}^2_{Block}$ = 7.505 and $\hat{\sigma}^2$ = 17.671. This implies that the yields in any two plots in each block are uniformly correlated, the estimated correlation being 7.505/(7.505+17.671) 0.298.

When you wish to use a correlated error structure in LMM (REML) you need to drop Block from the **Random Model**, and use just Block.Plot, since the correlation model supercedes the two random components model. (This is more fully described on page 656 in GenStat's *Statistics Guide* via the **Help** screen.)

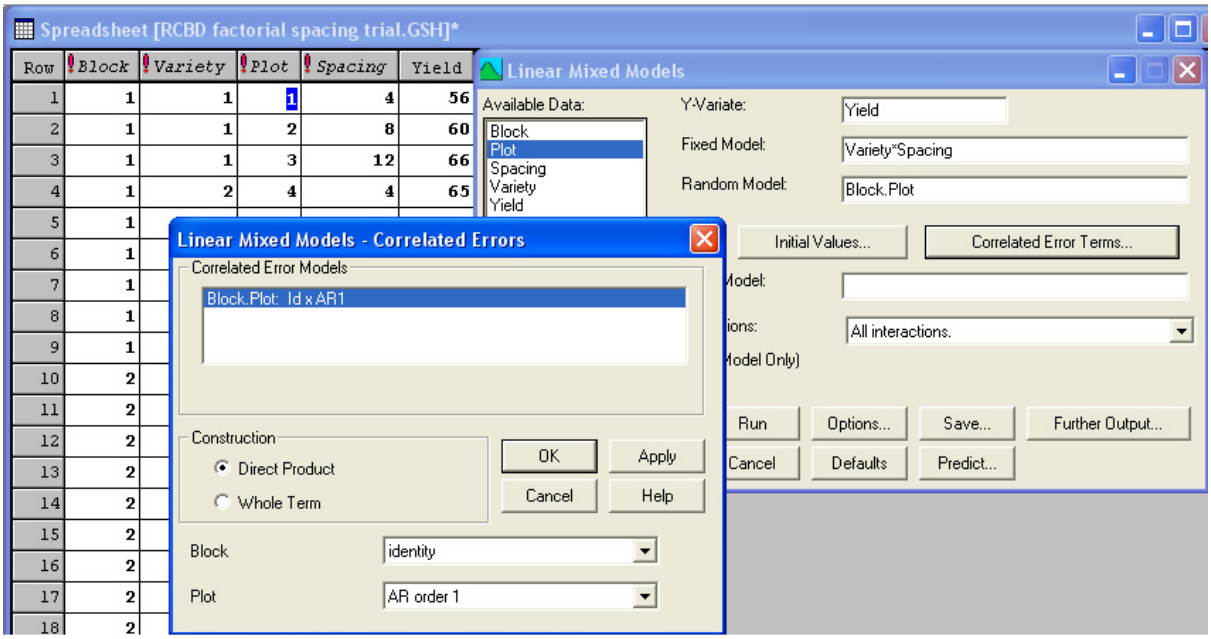
Unfortunately, **Uniform** is not currently listed in the menu's available **Correlated Error Terms**, but it is an option in the actual procedure. The way around this is to run a different correlation structure, copy the appropriate lines of code to a new **Input Window**, modify the line and re-run the window of code. Here we chose AR1:

Copy from GenStat's Input window:

```
VCOMPONENTS [FIXED=Block+Variety*Spacing; FACTORIAL=9; CADJUST=none]
RANDOM=Block.Plot; INITIAL=1; CONSTRAINTS=none
VSTRUCTURE [TERMS=Block.Plot; FORMATION=direct] MODEL=identity,ar1;
ORDER=*,1; FACTOR=Block,Plot
REML [PRINT=model,components,deviance,waldTests; PSE=differences;
MVINCLUDE=*; METHOD=AI; MAXCYCLE=20] Yield
```

Change to **uniform**, then use Run > Submit Window to re-run the analysis with a uniform correlation structure

To illustrate this, we need to supply an error term that indexes over the 4 blocks and 9 plots in each block. We will first add a factor column Plot with 9 levels (corresponding to the 3 varieties × 3 spacings used in each block). We then select an AR1 correlated error term from the menu, copy the input, change AR to uniform and rerun the analysis.



REML variance components analysis

Response variate: Yield
Fixed model: Constant + Variety + Spacing + Variety.Spacing
Random model: Block.Plot
Number of units: 36

Block.Plot used as residual term with covariance structure as below

Covariance structures defined for random model

Term	Factor	Model	Order	No. rows
Block.Plot	Block	Identity	1	4
	Plot	**Uniform**	1	9

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
		Block.Plot	Sigma2	**25.18**	8.96
	Block	Identity	-	-	-
	Plot	Uniform	theta1	**0.2981**	0.2286

Deviance: -2*Log-Likelihood

Deviance	d.f.
121.74	25

These F statistics and P values are identical to when we had a random model Block + Block.Plot, and are identical to those from the ANOVA.

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Variety	58.14	**2**	**29.07**	**24.0**	<0.001
Spacing	8.77	**2**	**4.39**	**24.0**	0.024
Variety.Spacing	43.32	**4**	**10.83**	**24.0**	<0.001

There is no random block term in the model, but the presence of a uniform correlation structure within blocks implies such a term. We can work the formula for the uniform correlation backwards to calculate the block variance component:

The estimate 25.18 is actually the *combined* estimate ($\hat{\sigma}^2_{Block} + \hat{\sigma}^2$). The uniform correlation is $0.2981 = \hat{\sigma}^2_{Block} / (\hat{\sigma}^2_{Block} + \hat{\sigma}^2) = \hat{\sigma}^2_{Block} / 25.18$, so that $\hat{\sigma}^2_{Block} = 0.2981 \times 25.18 = 7.506$ (as was obtained earlier).

In field trials, it is unlikely that a uniform correlation applies spatially or temporally. It is more likely that plots closer together (in time or space) are more strongly correlated than plots further apart. Hence, AR models are commonly used in the modern analyses of field trials. The example above does not have a known field plan, so we illustrate this with the eelworm data later on.

GenStat's examples in their on-line Statistics guide go even further. Once you start imposing complex correlation structures on the spatial design, there remains the possibility of including other sources of variation (measurement error, sampling error etc). Again, we will illustrate this with the eelworm data.

Three-way design (in randomized blocks) – missing values

Consider the following factorial treatment structure with two varieties, V, (labelled A, B), two levels of witchweed, W, (infested, I, or not infested, U) and 4 fertilisers, F, (0 = none, 1 = super only, 2 = super + manure and 4 = super + N + K). Two randomized blocks were used. The yields, Y, and the field plan are as follows:

Example 10 Maize RCBD experiment with 2 varieties × 2 witchweed infestations × 4 fertilisers, from SC Pearce, P132.

Block	V	W	F	Y	V	W	F	Y	V	W	F	Y	V	W	F	Y
1	B	I	F3	13.5	B	U	F1	12.8	A	I	F3	15.8	B	I	F4	11.6
	A	I	F1	10.4	B	U	F4	17.1	A	I	F2	12.5	A	U	F1	14.8
	B	I	F2	11.8	B	U	F2	16.9	B	I	F1	9.5	A	I	F4	11.3
	B	U	F3	22.3	A	U	F3	24.9	A	U	F4	19.9	A	U	F2	19.7
2	B	U	F2	16.0	A	I	F1	10.0	B	I	F2	9.5	A	U	F4	19.2
	A	U	F2	18.0	B	U	F1	13.0	B	I	F1	9.6	A	U	F3	22.0
	B	I	F3	13.4	A	I	F4	11.4	B	U	F4	16.6	B	U	F3	20.0
	A	I	F2	10.1	B	I	F4	9.2	A	U	F1	14.0	A	I	F3	13.6

This is a straightforward 3-way factorial treatment design. Ignoring any potential problems with the assumptions, the ANOVA is as follows:

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	1	11.5200	11.5200	19.61	
Block.*Units* stratum					
Variety	1	19.2200	19.2200	32.72	<.001
Fertiliser	3	167.7450	55.9150	95.20	<.001
Witchweed	1	338.0000	338.0000	575.48	<.001
Variety.Fertiliser	**3**	**0.7050**	**0.2350**	**0.40**	**0.755**
Variety.Witchweed	**1**	**3.6450**	**3.6450**	**6.21**	**0.025**
Fertiliser.Witchweed	**3**	**22.2250**	**7.4083**	**12.61**	**<.001**
Variety.Fertiliser.Witchweed	3	0.3300	0.1100	0.19	0.903
Residual	15	8.8100	0.5873		
Total	31	572.2000			

As usual with factorial experiments, interpret highest-order interactions downwards. If the 3-factor interaction is significant, that means that the pattern in a 2-way table of means differs across the levels of the third factor. For example, had Variety.Fertiliser.Witchweed been significant, we would conclude that for plots infested with witchweed, the change in response to the four fertilisers for varieties A and B is different to plots not infested with witchweed.

In this case, the 3-factor interaction is not significant so we can turn our attention to 2-factor interactions. Since the design is balanced, the order of the three 2-way interactions is

irrelevant. Below are the P values for a different order (Variety, Witchweed, Fertiliser). You can see that the variance ratios and P values for the three 2-way interactions are unchanged:

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	1	11.5200	11.5200	19.61	
Block.*Units* stratum					
Variety	1	19.2200	19.2200	32.72	<.001
Witchweed	1	338.0000	338.0000	575.48	<.001
Fertiliser	3	167.7450	55.9150	95.20	<.001
Variety.Witchweed	**1**	**3.6450**	**3.6450**	**6.21**	**0.025**
Variety.Fertiliser	**3**	**0.7050**	**0.2350**	**0.40**	**0.755**
Witchweed.Fertiliser	**3**	**22.2250**	**7.4083**	**12.61**	**<.001**
Variety.Witchweed.Fertiliser	3	0.3300	0.1100	0.19	0.903
Residual	15	8.8100	0.5873		

Now suppose that the bottom right hand corner plot was damaged due to rain. The plot yield, 13.6, is missing. The treatment involved was in a lower yielding block (block 2), the higher yielding variety A, the highest yielding fertiliser regime and the plot was infested with witchweed resulting in much lower yields.

We saw with example 1 that using a missing value code in ANOVA had a completely different outcome than omitting the row completely. With an * in lieu of a data value, a missing value formula is used to replace the yield, resulting in an apparent balanced data set (albeit with an adjustment to the residual degrees of freedom). While that may be approximately OK (treatment F values are somewhat inflated) it could become misleading. Omitting the entire row and using the unbalanced treatment structure ANOVA produces just one possible order of the factors and interactions.

In an unbalanced design, it is important to look at the P values for an interaction (or main effect) *adjusted for all other interactions (or main effects) of the same order*.

Occasionally the numbers of replicates in the treatment combinations may be unbalanced *simply because of the design limitations*. For example, an animal trial may involve brred and sex, and an equal number of male, female and neuter horses may not be available for all breeds of horses. Or in a sample survey an unequal number of males and females are canvassed across another category such as profession. ANOVA will not work for such unbalanced treatment structures. Suppose we omit the final row of the current data set and re-run the ANOVA. You will see the following error message:

```
Fault 8, code AN 1, statement 1 on line 411

Command: ANOVA [PRINT=aovtable,information,means,stratumvariance; FACT=32; CONTR
Design unbalanced - cannot be analysed by ANOVA.
Model term Fertiliser (non-orthogonal to term Block) is unbalanced.
```

Switching to unbalanced treatment structure ANOVA gives P values for the order of the factors and interactions in the fixed model:

Analysis of an unbalanced design using GenStat regression

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ Block	1	10.5376	10.5376	17.82	<.001
+ Variety	1	20.7471	20.7471	35.09	<.001
+ Fertiliser	3	193.8668	64.6223	109.31	<.001
+ Witchweed	1	313.0126	313.0126	529.46	<.001
+ Variety.Fertiliser	3	1.6177	0.5392	0.91	0.460
+ Variety.Witchweed	**1**	**2.3300**	**2.3300**	**3.94**	**0.067**
+ Fertiliser.Witchweed	3	20.0202	6.6734	11.29	<.001
+ Variety.Fertiliser.Witchweed	3	0.5422	0.1807	0.31	0.821
Residual	14	8.2767	0.5912		
Total	30	570.9510	19.0317		

whereas putting Variety last gives:

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.
+ Block	1	10.5376	10.5376	17.82	<.001
+ Fertiliser	3	186.3230	62.1077	105.06	<.001
+ Witchweed	1	319.8248	319.8248	540.98	<.001
+ Variety	1	21.4788	21.4788	36.33	<.001
+ Fertiliser.Witchweed	3	19.9113	6.6371	11.23	<.001
+ Fertiliser.Variety	3	1.0697	0.3566	0.60	0.624
+ Witchweed.Variety	**1**	**2.9869**	**2.9869**	**5.05**	**0.041**
+ Fertiliser.Witchweed.Variety	3	0.5422	0.1807	0.31	0.821
Residual	14	8.2767	0.5912		
Total	30	570.9510	19.0317		

You can see the dilemma: do we trust the 0.041 P value for Variety.Witchweed, or the 0.067 P value? The answer is *we should use the P value for* Variety.Witchweed *when it is the* ***last*** *2-factor interaction entered in the model*. The reason is that we need to **adjust** for the behaviour of maize across all four fertiliser regimes and both varieties before we can decide whether the response to infestation of witchweed is the same for the two varieties.

So, since all 2-factor interactions need to be entered last, that means we need to run at least three different unbalanced treatment structure ANOVAs.

Before looking at how REML handles this, we note the following. Since the 3-factor interaction is not significant, the corresponding Mean Square must be statistically similar to the Residual Mean Square (for the variance ratio to be not significantly larger than 1). We can therefore omit the three-factor interaction from the treatment structure. The repercussion is to move this interaction into the residual term, thus increasing the precision of the estimate of variance and increasing the power of the remaining tests.

To remove the three-factor interaction, either use the GenStat shortcut A*B*C-A.B.C, or else simply enumerate the remaining model: A+B+C+A.B+A.C+B.C (or A*B+A*C+B*C since repeated terms in the expansion of this model are simply ignored). We will do this in the next section.

LMM (REML) analysis

Remember that REML uses only the data present and hence it makes no difference whether an * is used or the row deleted entirely.

The **Fixed Model** is Variety*Witchweed*Fertiliser and the **Random Model** Block as with ANOVA:

REML variance components analysis

Response variate: Yield
Fixed model: Constant + Variety + Fertiliser + Witchweed + Variety.Fertiliser + Variety.Witchweed + Fertiliser.Witchweed + Variety.Fertiliser.Witchweed
Random model: Block
Number of units: 31 (1 units excluded due to zero weights or missing values)

Residual term has been added to model

Sparse algorithm with AI optimisation

Estimated variance components

Random term	component	s.e.
Block	0.6028	0.9084

Residual variance model

Term	FactorModel(order)	Parameter	Estimate	s.e.
Residual	Identity	Sigma2	0.591	0.2234

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Variety	35.19	1	35.19	14.0	<0.001
Fertiliser	328.35	3	109.45	14.0	<0.001
Witchweed	529.10	1	529.10	14.0	<0.001
Variety.Fertiliser	2.78	3	0.93	14.0	0.453
Variety.Witchweed	3.90	1	3.90	14.0	0.068
Fertiliser.Witchweed	33.73	3	11.24	14.0	<0.001
Variety.Fertiliser.Witchweed	0.95	3	0.32	14.0	0.813

Dropping individual terms from full fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Variety.Witchweed.Fertiliser	0.95	3	0.32	14.0	0.813

Notice

- GenStat has two sections of tests of fixed effects. The **Sequentially adding terms to fixed model** section is equivalent to the order produced by the unbalanced treatment structure ANOVA, except that with the latter a Block term is included, thereby affecting slightly the subsequent F values.

  The **Dropping individual terms from full fixed model** section is what should be used with

unbalanced data, since this is where the Wald statistics are placed for each term adjusted for all other terms of the same order.

- In this case, the 3-factor interaction can dropped (P=0.813). When we actually drop this from the model and re-run the analysis with:
  Fixed Model: Variety*Fertiliser*Witchweed-Variety.Fertiliser.Witchweed
  we obtain:

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Variety	40.09	1	40.09	17.0	<0.001
Fertiliser	374.12	3	124.71	17.0	<0.001
Witchweed	603.03	1	603.03	17.0	<0.001
Variety.Fertiliser	3.16	3	1.05	17.0	0.394
Variety.Witchweed	4.45	1	4.45	17.0	0.050
Fertiliser.Witchweed	38.47	3	12.82	17.0	<0.001

Dropping individual terms from full fixed model

Fixed term	**Wald statistic**	**n.d.f.**	**F statistic**	**d.d.f.**	**F pr**
Variety.Fertiliser	**1.72**	**3**	**0.57**	**17.0**	**0.640**
Variety.Witchweed	**5.70**	**1**	**5.70**	**17.0**	**0.029**
Fertiliser.Witchweed	**38.47**	**3**	**12.82**	**17.0**	**<0.001**

The P values for each of the 2-factor interactions is obtained adjusted for the other 2-factor interactions, so it is as if GenStat is running three models for us. Had these interactions all been not significant we could drop them from the model, leaving main effects only; the Wald statistics in the Dropping individual terms from full fixed model section are all adjusted.

Any significant interaction that needs to be included in a model should have the main effects and lower-order interactions included as well.

Thus, we conclude that the final model for this example involves three main effects (variety, fertliser and witchweed) and two significant interactions Variety.Witchweed (P=0.029) and Fertliser.Witchweed (P<0.001).

Table of predicted means for Variety.Witchweed

Witchweed Variety	I	U
A	11.95	19.06
B	11.01	16.84

The yield for Variety A is relatively lower than Variety B in plots infested with witchweed than uninfested.

Table of predicted means for Witchweed.Fertiliser

Fertiliser Witchweed	F1	F2	F3	F4
I	9.88	10.97	14.20	10.87
U	13.65	17.65	22.30	18.20

See Appendix 6 for an example showing the reliability of REML means for missing values.

Three-way design (in randomized blocks) – changing variance

McConway *et al.* (1999) reported the results of an experiment which had a randomised block design, in more or less the following words. There were 64 plots, arranged in four blocks each of size sixteen. Each block was a rectangular piece of land, measuring 3m × 32m. Each block was divided into sixteen plots by splitting the long side of the block into sixteen 2m pieces. So, each plot was a 3m× 2m rectangle of land. The River Thames runs along one edge of the field used in this experiment, and usually floods part of the field each year. The blocks were designed so that the long side of each block was parallel to the river-bank. The blocks were different distances from the river-bank.

The experiment was about growing turnips for fodder. The turnips would not normally be harvested because they are grown to provide food for farm animals in winter; the farmer simply releases animals into the field and the animals graze on the turnips. The turnips are not even the main crop in the field during the growing season; the turnips are sown after the main crop is removed.

There were sixteen treatments in this experiment. The combinations are formed from: two different varieties – Barkant or Marco; two different sowing dates – one as soon as possible after the main crop has been harvested, the other a week later; and four different sowing densities – 1, 2, 4 or 8 kg ha^{-1}. Treatment combinations were allocated to plots within blocks at random.

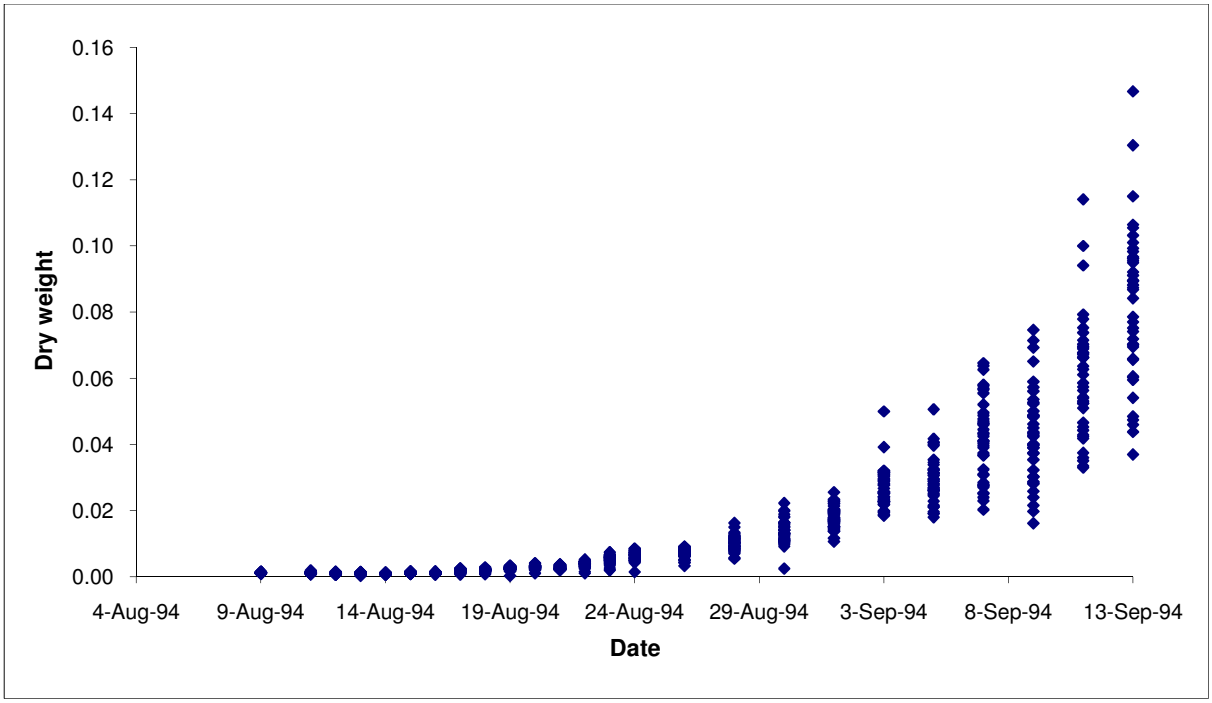
Example 11 Yield of turnips (kg), from McConway *et al.* (1999)

variety	sowing date	sowing density (kg ha^{-1})	Block 1	Block 2	Block 3	Block 4
Barkant	21/08/1990	1	2.7	1.4	1.2	3.8
		2	7.3	3.8	3.0	1.2
		4	6.5	4.6	4.7	0.8
		8	8.2	4.0	6.0	2.5
	28/08/1990	1	4.4	0.4	6.5	3.1
		2	2.6	7.1	7.0	3.2
		4	24.0	14.9	14.6	2.6
		8	12.2	18.9	15.6	9.9
Marco	21/08/1990	1	1.2	1.3	1.5	1.0
		2	2.2	2.0	2.1	2.5
		4	2.2	6.2	5.7	0.6
		8	4.0	2.8	10.8	3.1
	28/08/1990	1	2.5	1.6	1.3	0.3
		2	5.5	1.2	2.0	0.9
		4	4.7	13.2	9.0	2.9
		8	14.9	13.3	9.3	3.6

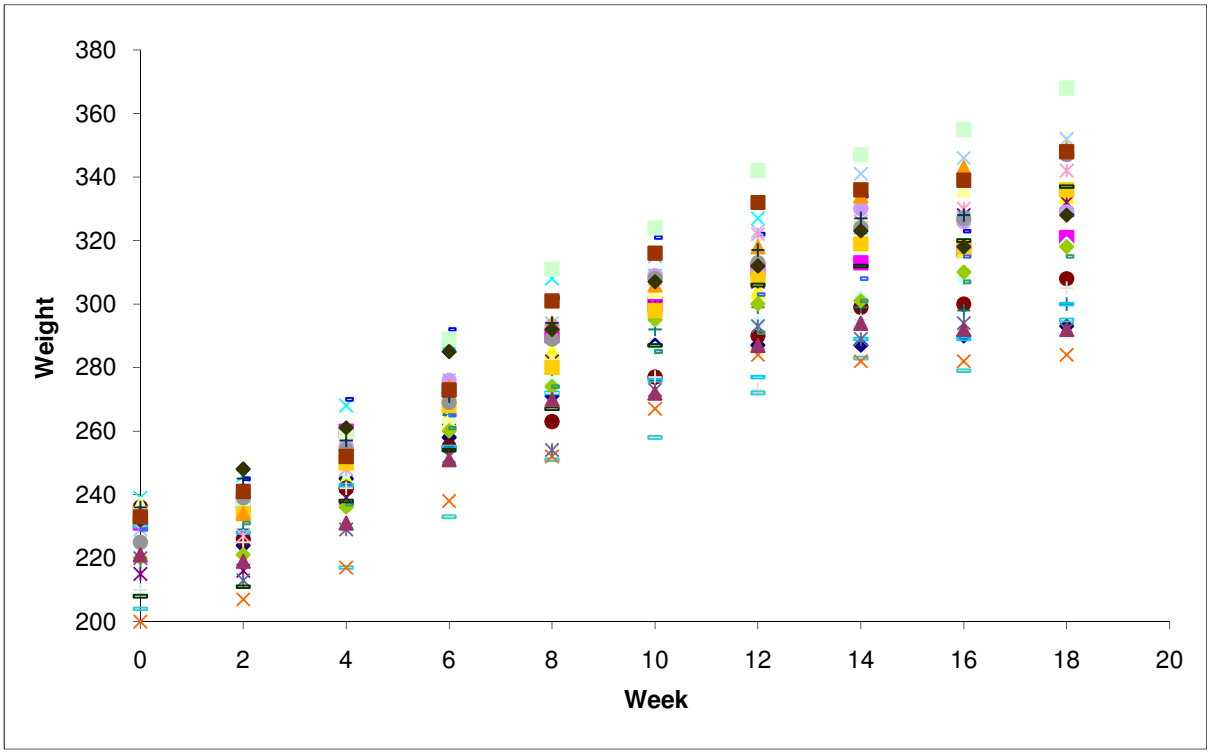
Again, this is a density trial, and hence the variance *may* change over different planting densities.

The plants are also grown for two different time periods. It is almost always the case that the variance of plant yield increases over time. The following is an example of this.

An experiment was conducted by a former student at The University of Sydney (Jason Moodie) on lettuce growth for the first 30 days after transplanting seedlings. Dry weights, fresh weights and leaf areas were measured every day or every second day. It is clear that the variance increases over time.



A second example is calf weight for the first nineteen weeks after birth which we consider again later:



Again, the variance appears to increase as the calves grow. The means and variances over time for these thirty calves are as follows.

Week	0	2	4	6	8	10	12	14	16	18	19
Mean	226.20	230.33	246.87	265.63	281.17	294.87	304.73	312.87	315.13	324.07	325.47
Variance	105.54	155.13	165.22	184.86	242.97	283.77	306.55	340.67	389.15	470.06	444.60

The points are

- we should expect the variance to change when plants are grown for different lengths of time
- we should expect the variance to change with density (it may not, depending on the extent of plant competition).

Firstly, here is the standard ANOVA assuming constant variance:

Analysis of variance

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	3	163.737	54.579	5.69	
Block.*Units* stratum					
Density	3	470.378	156.793	16.35	<.001
Sowing	1	233.708	233.708	24.37	<.001
Variety	1	83.951	83.951	8.75	0.005
Density.Sowing	3	154.793	51.598	5.38	0.003
Density.Variety	3	8.647	2.882	0.30	0.825
Sowing.Variety	1	36.451	36.451	3.80	0.057
Density.Sowing.Variety	3	17.999	6.000	0.63	0.602
Residual	45	431.611	9.591		
Total	63	1601.275			

Tables of means

Variate: Yield

Grand mean 5.38

Density	1	2	4	8
	2.14	3.35	7.33	8.69

Sowing	21-Aug-90	28-Aug-90
	3.47	7.29

Variety	Barkant	Marco
	6.52	4.23

Density	Sowing	21-Aug-90	28-Aug-90
1		1.76	2.51
2		3.01	3.69
4		3.91	10.74
8		5.18	12.21

Density	Variety	Barkant	Marco
1		2.94	1.34
2		4.40	2.30
4		9.09	5.56
8		9.66	7.73

Sowing	Variety	Barkant	Marco
21-Aug-90		3.86	3.08
28-Aug-90		9.19	5.39

	Sowing	21-Aug-90		28-Aug-90	
Density	Variety	Barkant	Marco	Barkant	Marco
1		2.28	1.25	3.60	1.43
2		3.83	2.20	4.98	2.40
4		4.15	3.68	14.03	7.45
8		5.18	5.18	14.15	10.28

Standard errors of differences of means

Table	Density	Sowing	Variety	**Density Sowing**
rep.	16	32	32	**8**
d.f.	45	45	45	**45**
s.e.d.	1.095	0.774	0.774	**1.548**

Table	Density Variety	Sowing Variety	Density Sowing Variety
rep.	8	16	4
d.f.	45	45	45
s.e.d.	1.548	1.095	2.190

Estimated stratum variances

Stratum	variance	effective d.f.	**variance component**
Block	54.579	3.000	**2.812**
Block.*Units*	9.591	45.000	**9.591**

LMM (REML) analysis

For this experiment, the **Fixed Model** is Variety*Date*Density and the **Random Model** is Block/Plot. As before, plots are completely described by the combination of Variety*Date*Density, leading to Block+ Block.Variety*Date*Density as the **Random Model**. That allows use to investigate Diagonal structures for Date and/or Density.

						Change in:		
Block	**Variety**	**Sowing date**	**Density**	**deviance**	**d.f.**	**deviance**	**d.f.**	**P value**
Identity	Identity	Diagonal	Diagonal	162.05	42			
Identity	Identity	Identity	Diagonal	168.10	43	6.05	1	0.014
Identity	Identity	Diagonal	Diagonal	162.05	42			
Identity	Identity	Diagonal	Identity	175.71	45	13.66	3	0.003
Identity	Identity	Identity	Identity	183.92	46			

If we start assuming that the variance changes over *time* as well as over *densities*, we can then test whether an adequate model has only a changing variance over densities ($P = 0.014$), or a

changing variance over time ($P = 0.003$). We clearly should allow the variance to change over both factors.

REML variance components analysis

Response variate: weight
Fixed model: Constant + density + sowing + variety + density.sowing + density.variety + sowing.variety + density.sowing.variety
Random model: block + block.density.sowing.variety
Number of units: 64

block.density.sowing.variety used as residual term with covariance structure as below

Covariance structures defined for random model

Term	Factor	Model	Order	No. rows
block.density.sowing.variety				
	block	Identity	0	4
	density	**Diagonal**	**4**	**4**
	sowing	**Diagonal**	**2**	**2**
	variety	Identity	0	2

Estimated variance components

Random term	component	s.e.
block	0.160	0.328

Output using `PARAMETERIZATION=sigmas`

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
block.density.sowing.variety		Sigma2	1.000	fixed	
	block	Identity	-	-	-
	density	Diagonal	d_1	1.000	fixed
			d_2	2.195	1.358
			d_3	10.48	6.35
			d_4	7.682	4.661
	sowing	Diagonal	d_1	1.030	0.507
			d_2	3.143	1.481
	variety	Identity	-	-	-

Estimated covariance models

Variance of data estimated in form:

V(y) = sZZ' + Sigma2.R

where: V(y) is variance matrix of data
s is the variance component for the random term
Z is the incidence matrix for the random term
Sigma2 is the residual variance
R is the residual covariance matrix

Random Term: block
Scalar s: 0.1604

Residual term: block.density.sowing.variety
Sigma2: 1.000
R uses direct product construction

To assist in understanding this output, we turned on the option **Covariance Model**. GenStat has scaled σ^2 to 1. The information on variance estimates is then obtained in the diagonal covariance matrices of the factors making up the residual term. To take one block and one variety, the variance of Y is obtained by evaluating the direct product of the two diagonal covariance matrices:

$$\begin{pmatrix} 1.000 & 0 & 0 & 0 \\ 0 & 2.195 & 0 & 0 \\ 0 & 0 & 10.48 & 0 \\ 0 & 0 & 0 & 7.682 \end{pmatrix} \otimes \begin{pmatrix} 1.030 & 0 \\ 0 & 3.143 \end{pmatrix}$$

The matrix in the text book is a direct product of a 4×4 and a 2×2, giving an 8×8 matrix with elements obtained by element-by-element multiplication of the separate matrices:

1.000×1.030	0	0	0	0	0	0	0
0	1.000×3.143	0	0	0	0	0	0
0	0	2.195×1.030	0	0	0	0	0
0	0	0	2.195×3.143	0	0	0	0
0	0	0	0	10.48×1.030	0	0	0
0	0	0	0	0	10.48×3.143	0	0
0	0	0	0	0	0	7.682×1.030	0
0	0	0	0	0	0	0	7.682×3.143

This calculates as:

	Density	1		2		4		8	
Density	Sowing date	21	28	21	28	21	28	21	28
1	21-Aug-90	1.030	0	0	0	0	0	0	0
	28-Aug-90	0	3.143	0	0	0	0	0	0
2	21-Aug-90	0	0	2.261	0	0	0	0	0
	28-Aug-90	0	0	0	6.899	0	0	0	0
4	21-Aug-90	0	0	0	0	10.794	0	0	0
	28-Aug-90	0	0	0	0	0	32.939	0	0
8	21-Aug-90	0	0	0	0	0	0	7.912	0
	28-Aug-90	0	0	0	0	0	0	0	24.145

Thus, the variance of an observation for any block and variety, whose density is 1 kg ha^{-1} and sown on 21/08/1990 is estimated to be 0.1604 (= block variance) + 1.030 = 1.190. For a similar combination but sown a week later, it is 0.1604 + 3.143 = 3.301.

The same variances are obtained using `PARAMETERIZATION=gammas`. GenStat estimates σ^2 to be 1.030 and scales the leading diagonal element of the covariance matrix for sowing date:

	sowing	Diagonal	d_1	1.000	fixed
			d_2	3.053	1.328

Deviance: -2*Log-Likelihood

Deviance	d.f.
162.05	42

Wald tests for fixed effects

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Variety	9.35	1	9.35	22.2	0.006
Density	38.47	3	12.09	18.9	<0.001
Sowing	7.86	1	7.86	21.9	0.010
Variety.Density	0.40	3	0.13	18.9	0.944
Variety.Sowing	2.03	1	2.03	21.9	0.168
Density.Sowing	14.50	3	4.56	18.8	0.015
Variety.Density.Sowing	1.44	3	0.45	18.8	0.719

Next we present just the two-way means for density and sowing for illustration. Since there are changing variances over the levels of some factors, we should turn on the option **Standard**

Errors All Differences so that individual differences can be compared or estimated with the correct precision.

For example, to compare the two variety means at a density of 4 kg ha^{-1}, we select treatments numbered 5 and 6 from the Standard errors of differences between pairs table for the in the output. We then read the value where the row marked

density 4.sowing 28/8/90	**6**

intersects with the column marked **5**. The mean difference is 10.737-3.912 = 6.825 ± 2.338. Note from the Wald statistic that the df are 18.8, so for assessing the significance of this difference we would use 18.8 or 19 df. The t value is 6.825/2.338 = 2.92, and this is highly significant (P=0.009). The 95% confidence interval for the true varietal difference at 4 kg ha^{-1} is (1.93, 11.72) kg ha^{-1}.

The ANOVA had a P value for Sowing.Variety of 0.057 and a constant s.e.d. of 1.548 for comparing two means and. Using this average-type value leads to difficulty in comparing means with appropriate precision.

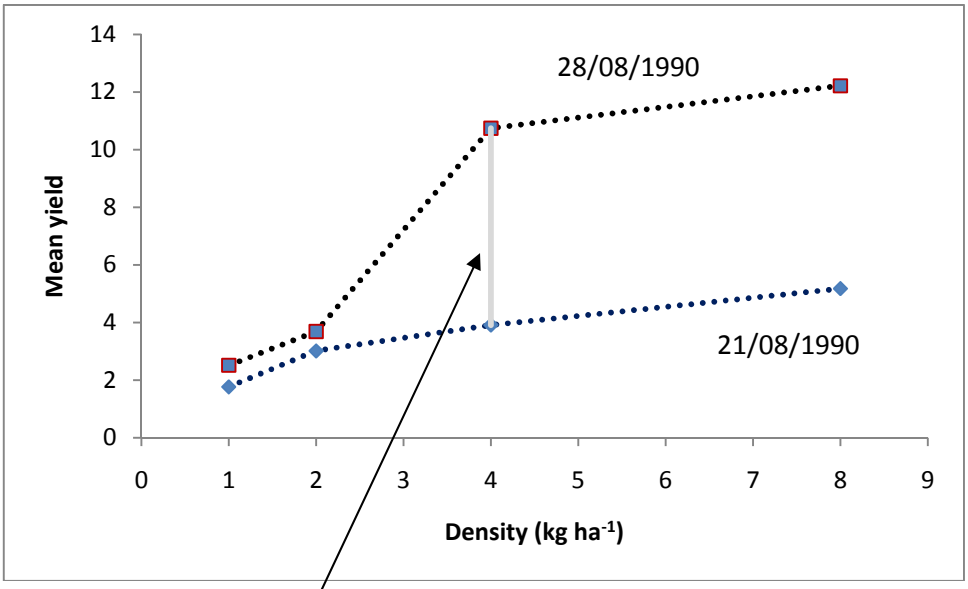


Table of predicted means for density.sowing

sowing	21/8/90	28/8/90
density		
1kg/Ha	1.763	2.512
2kg/Ha	3.013	3.688
4kg/Ha	3.912	10.737
8kg/Ha	5.175	12.212

To compare these means, use

Standard errors of differences between pairs

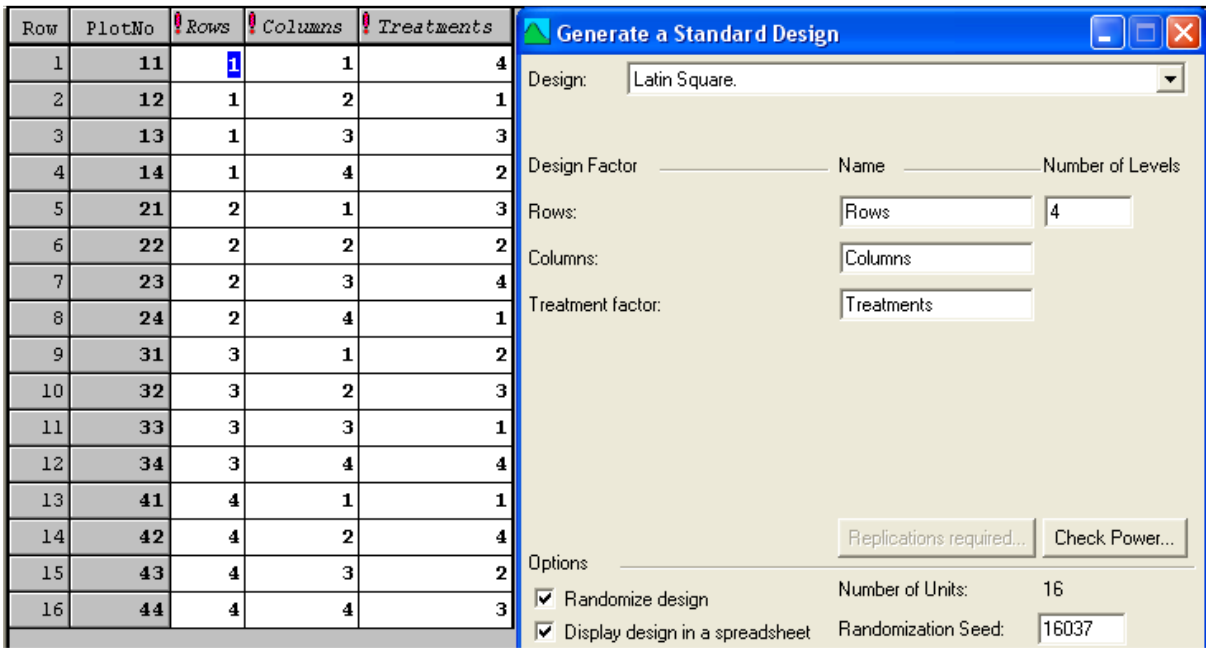
density 1.sowing 21/8/90	**1**	*							
density 1.sowing 28/8/90	**2**	0.722	*						
density 2.sowing 21/8/90	**3**	0.641	0.822	*					
density 2.sowing 28/8/90	**4**	0.996	1.120	1.070	*				
density 4.sowing 21/8/90	**5**	1.216	1.320	1.277	1.487	*			
density 4.sowing 28/8/90	**6**	2.061	2.124	2.098	2.232	2.338	*		
density 8.sowing 21/8/90	**7**	1.057	1.175	1.127	1.361	1.529	2.260	*	
density 8.sowing 28/8/90	**8**	1.774	1.847	1.817	1.970	2.090	2.671	2.002	*
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**

Latin Square design

Occasionally we need to block in two directions in the field (especially in animal trials, where individual animals form one block, and the experiment is repeated over time, time forming a second block).

For a Latin Square design, we need to have as many blocks in both directions as we have treatments. We then balance the allocation of treatments so that each occurs just once in each row and once in each column.

Here is GenStat's **Design** menu for generating a random 4×4 design:



Row	PlotNo	Rows	Columns	Treatments
1	11	1	1	4
2	12	1	2	1
3	13	1	3	3
4	14	1	4	2
5	21	2	1	3
6	22	2	2	2
7	23	2	3	4
8	24	2	4	1
9	31	3	1	2
10	32	3	2	3
11	33	3	3	1
12	34	3	4	4
13	41	4	1	1
14	42	4	2	4
15	43	4	3	2
16	44	4	4	3

Treatment allocation for this random design:

	Column block			
Row block	1	2	3	5
1	**4**	**1**	**3**	**2**
2	**3**	**2**	**4**	**1**
3	**2**	**3**	**1**	**4**
4	**1**	**4**	**2**	**3**

We have marked a typical row block, a typical column block, and a typical plot (the intersection of a row block and a column block). Thus, there are three strata, and hence the **Block Structure** is

Row + Column + Row.Column

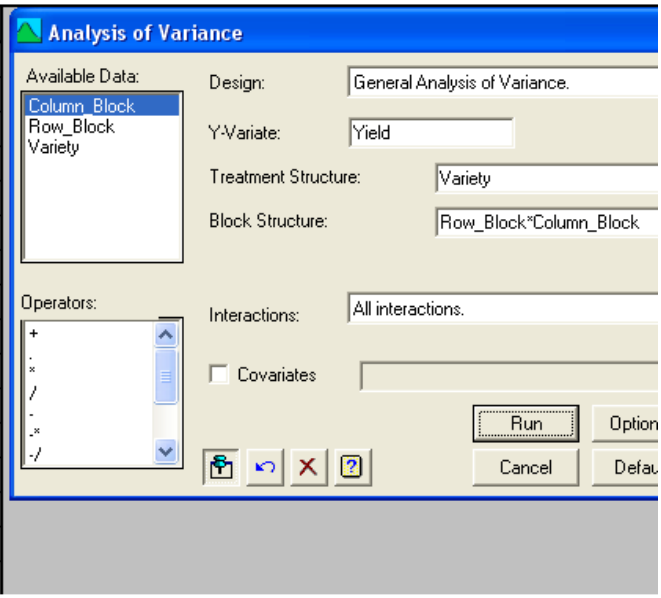
which can be shortened to Row*Column, or, since the final stratum can always be omitted, Row + Column.

Example 12 Wheat yields (kg per plot) from Steel and Torrie, page 224.

	Column block			
Row block	1	2	3	4
1	C	D	B	A
2	B	A	C	D
3	D	C	A	B
4	A	B	D	C

Column block			
1	2	3	4
10.5	7.7	12.0	13.2
11.1	12.0	10.3	7.5
5.8	12.2	11.2	13.7
11.6	12.3	5.9	10.2

Row	Row_Block	Column_Block	Variety	Yield
1	1	1	C	10.5
2	2	1	B	11.1
3	3	1	D	5.8
4	4	1	A	11.6
5	1	2	D	7.7
6	2	2	A	12
7	3	2	C	12.2
8	4	2	B	12.3
9	1	3	B	12
10	2	3	C	10.3
11	3	3	A	11.2
12	4	3	D	5.9
13	1	4	A	13.2
14	2	4	D	7.5
15	3	4	B	13.7
16	4	4	C	10.2



Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Row_Block stratum	3	1.9550	0.6517	1.44	
Column_Block stratum	3	6.8000	2.2667	5.00	
Row_Block.Column_Block stratum					
Variety	**3**	**78.9250**	**26.3083**	**58.03**	**<.001**
Residual	6	2.7200	0.4533		
Total	15	90.4000			

Message: the following units have large residuals.

Row_Block 4 Column_Block 4 -0.85 s.e. 0.41

Tables of means

Variate: Yield

Grand mean 10.45

Variety	A	B	C	D
	12.00	12.27	10.80	6.72

Standard errors of differences of means

Table	Variety
rep.	4
d.f.	6
s.e.d.	0.476

Least significant differences of means (5% level)

Table	Variety
rep.	4
d.f.	6
l.s.d.	1.165

Estimated stratum variances

Variate: Yield

Stratum	variance	effective d.f.	variance component
Row_Block	0.652	3.000	0.050
Column_Block	2.267	3.000	0.453
Row_Block.Column_Block	0.453	6.000	0.453

From the stratum variances, columns show more variability than rows.

LMM (REML) analysis

For this design there are three variance estimates coming from the three strata – rows, columns and plots. As before, the **Fixed Model** contains the one factor, Variety, while the **Random Model** is Row_Block + Column_Block + Row_Block.Column_Block, or simply Row_Block*Column_Block.

REML variance components analysis

Response variate:	Yield
Fixed model:	Constant + Variety
Random model:	Row_block + Column_block + Row_block.Column_block
Number of units:	16

Row_block.Column_block used as residual term

Sparse algorithm with AI optimisation

Estimated variance components

Random term	component	s.e.
Row_block	0.0496	0.1482
Column_block	0.4533	0.4673

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
Row_block.Column_block		Identity	Sigma2	0.453	0.2617

Deviance: -2*Log-Likelihood

Deviance	d.f.
13.97	9

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Variety	174.10	3	58.03	6.0	<0.001

Table of predicted means for Constant

10.45 Standard error: 0.393

Table of predicted means for Variety

Variety	A	B	C	D
	12.00	12.27	10.80	6.72

Standard error of differences: 0.4761

Notice, as usual:

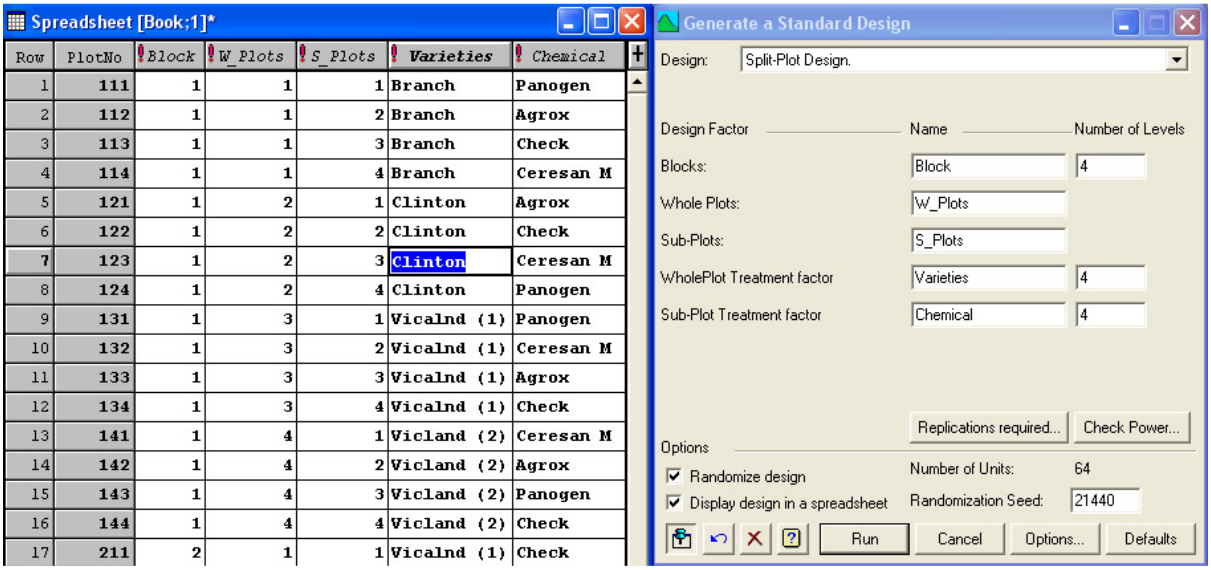
Variety

Variety A	1	*	
Variety B	2	1.165	*

- The estimates of variance are the same as the stratum variances given in the ANOVA.
- The F statistic is the same as the variance ratio of the ANOVA.
- The means and s.e.d. values are the same as from ANOVA. REML also gives 1.165 as the common least significant difference (5% level) of means (in a complete matrix of values).

Split-plot design (in randomized blocks)

Firstly, we will use GenStat's **Design** menu to generate a field plan to correspond to Steel and Torrie's oats experiment (page 383) with four varieties randomised to whole plots and four chemical seed treatments (one of which is a control) to split plots. Appropriate factor labels have replaced numbers.



Row	PlotNo	Block	W_Plots	S_Plots	Varieties	Chemical
1	111	1	1	1	Branch	Panogen
2	112	1	1	2	Branch	Agrox
3	113	1	1	3	Branch	Check
4	114	1	1	4	Branch	Ceresan M
5	121	1	2	1	Clinton	Agrox
6	122	1	2	2	Clinton	Check
7	123	1	2	3	Clinton	Ceresan M
8	124	1	2	4	Clinton	Panogen
9	131	1	3	1	Vicalnd (1)	Panogen
10	132	1	3	2	Vicalnd (1)	Ceresan M
11	133	1	3	3	Vicalnd (1)	Agrox
12	134	1	3	4	Vicalnd (1)	Check
13	141	1	4	1	Vicland (2)	Ceresan M
14	142	1	4	2	Vicland (2)	Agrox
15	143	1	4	3	Vicland (2)	Panogen
16	144	1	4	4	Vicland (2)	Check
17	211	2	1	1	Vicalnd (1)	Check

Notice that GenStat creates three factor columns (Block, W_Plot and S_Plot), one for each of the three strata in this experiment. The field plan is also printed in the **Output** window.

Treatment combinations on each unit of the design

	S_Plots	1	2	3	4
Block	W_Plots				
1	1	4 3	4 4	4 1	4 2
	2	3 4	3 1	3 2	3 3
	3	1 3	1 2	1 4	1 1
	4	2 2	2 4	2 3	2 1
2	1	1 1	1 4	1 3	1 2
	2	2 4	2 1	2 3	2 2
	3	3 1	3 4	3 2	3 3
	4	4 2	4 4	4 1	4 3
3	1	4 4	4 1	4 2	4 3
	2	1 3	1 1	1 2	1 4
	3	3 2	3 3	3 1	3 4
	4	2 2	2 4	2 3	2 1
4	1	1 2	1 3	1 1	1 4
	2	3 3	3 4	3 1	3 2
	3	2 4	2 1	2 3	2 2
	4	4 4	4 3	4 1	4 2

Treatment factors are listed in the order: Varieties, Chemical.

This field plan is reproduced graphically with labels:

Block 1	Panogen	Agrox	Check	Ceresan M	Branch
	Agrox	Check	Ceresan M	Panogen	Clinton
	Panogen	Ceresan M	Agrox	Check	Vicland (1)
	Ceresan M	Agrox	Panogen	Check	Vicland (2)
Block 2	Check	Agrox	Panogen	Ceresan M	Vicland (1)
	Agrox	Check	Panogen	Ceresan M	Vicland (2)
	Check	Agrox	Ceresan M	Panogen	Clinton
	Ceresan M	Agrox	Check	Panogen	Branch
Block 3	Agrox	Check	Ceresan M	Panogen	Branch
	Panogen	Check	Ceresan M	Agrox	Vicland (1)
	Ceresan M	Panogen	Check	Agrox	Clinton
	Ceresan M	Agrox	Panogen	Check	Vicland (2)
Block 4	Ceresan M	Panogen	Check	Agrox	Vicland (1)
	Panogen	Agrox	Check	Ceresan M	Clinton
	Agrox	Check	Panogen	Ceresan M	Vicland (2)
	Agrox	Panogen	Check	Ceresan M	Branch

There are clearly three strata here: blocks, the ¼ block strips (the whole-plots) that the varieties are randomised to, and the ¼ whole-plot shapes (the split-plots) that the seed protectants were assigned to at random. The **Block Structure** is therefore

Block + Block.Whole_Plot + Block.Whole_Plot.Split_plot

with the shortcut

Block/Whole_Plot/Split_plot

which describes the way the units were formed in the field: whole-plots were formed as large units within blocks, and split-plots were formed as smaller units within whole-plots.

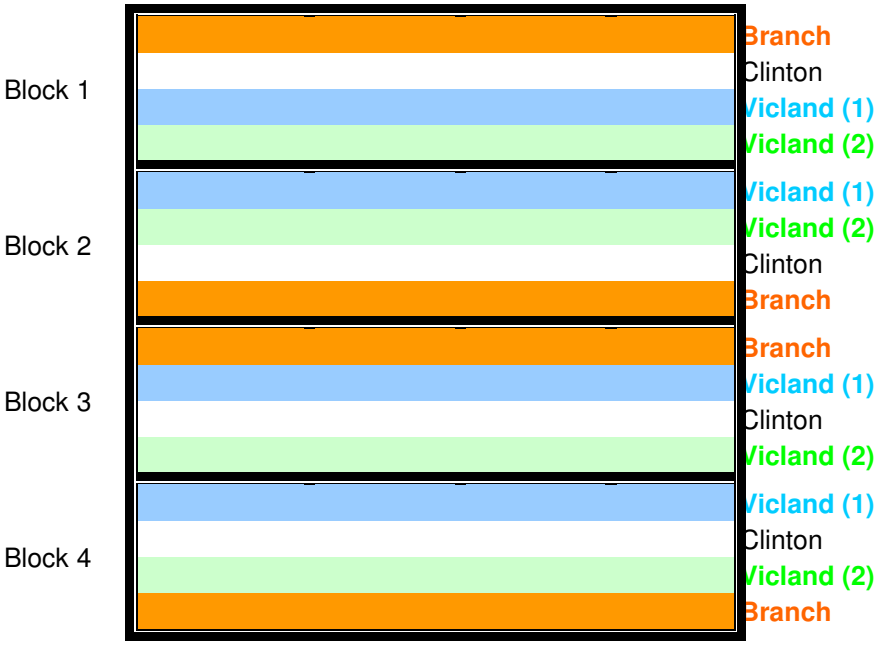
Providing you set up these three factors, this structure is what you would use irrespective of the complexity of the whole-plot treatment and the split-plot treatment structures. For example, the treatments applied to whole-plots could have a 3×4 factorial structure, while those applied to the split-plots a (2×2+1) incomplete factorial structure.

For this example, there were simple structures for both whole-plot and split-plot treatment structures. Hence the following **Block Structure** can be used instead:

Block + Block.Variety + Block.Variety.Chemical

In fact this design can be thought of in two ways.

1. RCBD with varieties as treatments.

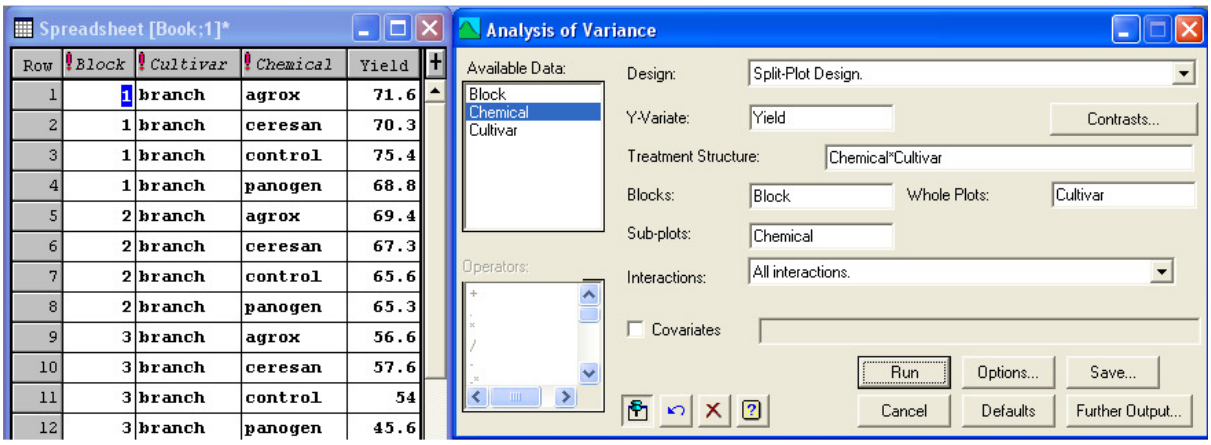


This, in fact, forms the whole-plot part of the combined split-plot ANOVA.

2. Four separate RCBDs, one per variety, with seed chemical protectants as treatments. This is one such layout, for Branch.

Block 1	Panogen	Agrox	Check	Ceresan M	Branch
Block 2	Ceresan M	Agrox	Check	Panogen	Branch
Block 3	Agrox	Check	Ceresan M	Panogen	Branch
Block 4	Agrox	Panogen	Check	Ceresan M	Branch

In fact, this is an important concept in checking the assumptions at the split-plot level. This ANOVA produces 9 *df* for the Residual MS. There are four such residuals to check for "homogeneity"; their average is, in fact, the split-plot Residual MS in the combined analysis. The combined analysis is feasible only when these individual variance components are commensurable.



Example 13 From Snedecor and Cochran page 384

		Seed chemical protectant			
Cultivar	Block	Control	Ceresan M	Panogen	Agrox
Vicland (1)	1	42.9	53.8	49.5	44.4
	2	41.6	58.5	53.8	41.8
	3	28.9	43.9	40.7	28.3
	4	30.8	46.3	39.4	34.7
Vicland (2)	1	53.3	57.6	59.8	64.1
	2	69.6	69.6	65.8	57.4
	3	45.4	42.4	41.4	44.1
	4	35.1	51.9	45.4	51.6
Clinton	1	62.3	63.4	64.5	63.6
	2	58.5	50.4	46.1	56.1
	3	44.6	45.0	62.6	52.7
	4	50.3	46.7	50.3	51.8
Branch	1	75.4	70.3	68.8	71.6
	2	65.6	67.3	65.3	69.4
	3	54.0	57.6	45.6	56.6
	4	52.7	58.5	51	47.4

First, the standard split-plot ANOVA is obtained (using the specific split-plot menu).

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	3	2842.87	947.62	13.79	
Block.Cultivar stratum					
Cultivar	3	2848.02	949.34	13.82	0.001
Residual	9	618.29	68.70	3.38	
Block.Cultivar.Chemical stratum					
Chemical	3	170.54	56.85	2.80	0.054
Cultivar.Chemical	9	586.47	65.16	3.21	0.006
Residual	36	731.20	20.31		
Total	63	7797.39			

Message: the following units have large residuals.

Block 2 Cultivar clinton	-7.27	s.e. 3.11
Block 2 Cultivar vicland2	6.45	s.e. 3.11
Block 2 Cultivar clinton Chemical panogen	-8.24	s.e. 3.38
Block 2 Cultivar vicland2 Chemical agrox	-9.09	s.e. 3.38
Block 3 Cultivar clinton Chemical panogen	9.81	s.e. 3.38
Block 4 Cultivar vicland2 Chemical control	-8.34	s.e. 3.38

Tables of means

Variate: Yield

Grand mean 52.81

Cultivar	branch	clinton	vicland1	vicland2
	61.07	54.31	42.46	53.41

Chemical	agrox	ceresan	control	panogen
	52.23	55.20	50.69	53.13

Cultivar	Chemical	agrox	ceresan	control	panogen
branch		61.25	63.43	61.93	57.68
clinton		56.05	51.38	53.93	55.88
vicland1		37.30	50.63	36.05	45.85
vicland2		54.30	55.38	50.85	53.10

Standard errors of differences of means

Table	Cultivar	Chemical	Cultivar Chemical
rep.	16	16	4
s.e.d.	2.930	1.593	4.025
d.f.	9	36	26.78
Except when comparing means with the same level(s) of			
Cultivar			3.187
d.f.			36

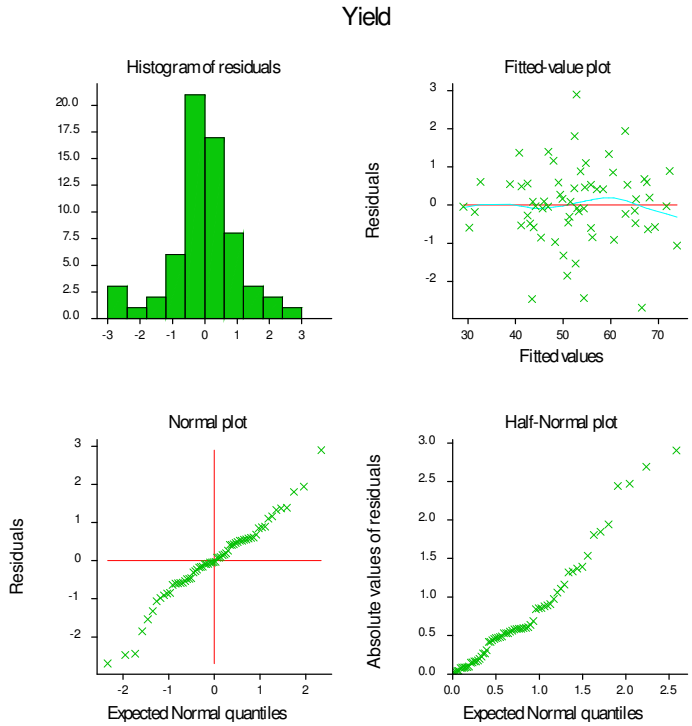
Least significant differences of means (5% level)

Table	Cultivar	Chemical	Cultivar Chemical
rep.	16	16	4
l.s.d.	6.629	3.232	8.263
d.f.	9	36	26.78
Except when comparing means with the same level(s) of			
Cultivar			6.463
d.f.			36

GenStat organizes the analysis into three strata corresponding to what was done in the field. Notice the following.

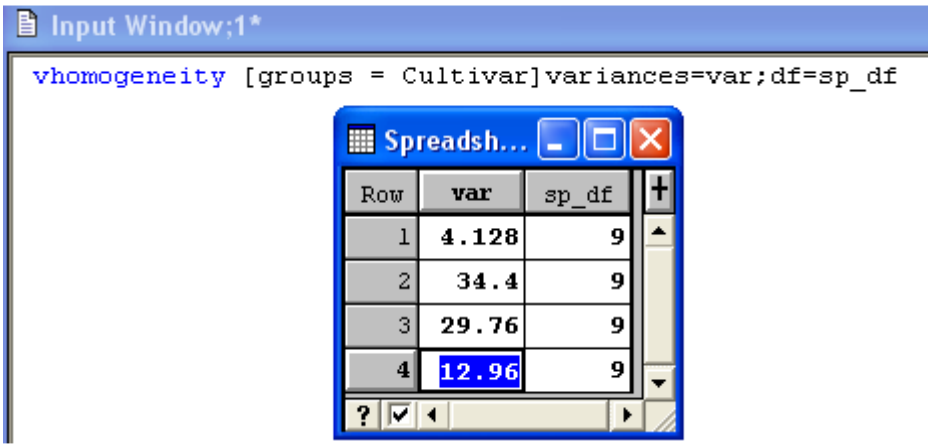
- Cultivar is tested in the whole-plot stratum, since whole-plots are the replicates for this treatment factor.
- Chemical and Cultivar.Chemical are tested in the split-plot stratum, since split-plots are the replicates for this treatment/interaction.
- There are several s.e.d. and l.s.d. values. Each is used for an appropriate treatment mean comparison. Not all comparisons lead to exact *t* tests. Performing a two stage randomization in the field has made the subsequent analysis slightly more complex than a one stage randomization.

Before interpreting the analysis, we should check the residual plot. Maybe there is some fanning, but nothing jumps out as a major problem.



Before interpreting the analysis, the components that form the split-plot error should be checked.

We do this in GenStat by clicking in the spreadsheet, then **Restrict/Filter > To Groups (factor levels)**. Select Cultivar and, one by one, each of the levels to perform a simple RCBD ANOVA. The Residual MS values (each with 9 *df*) are 4.128(Vicland (1)), 34.40 (Vicland (2)), 29.76 (Clinton), 12.96 (Branch). These appear quite different. Their average is 20.312, which is the split-plot Residual MS, with 4×9 = 36 *df*. In fact, performing a Bartlett test of homogeneity of variances on these indicates significance at *P*=0.021.

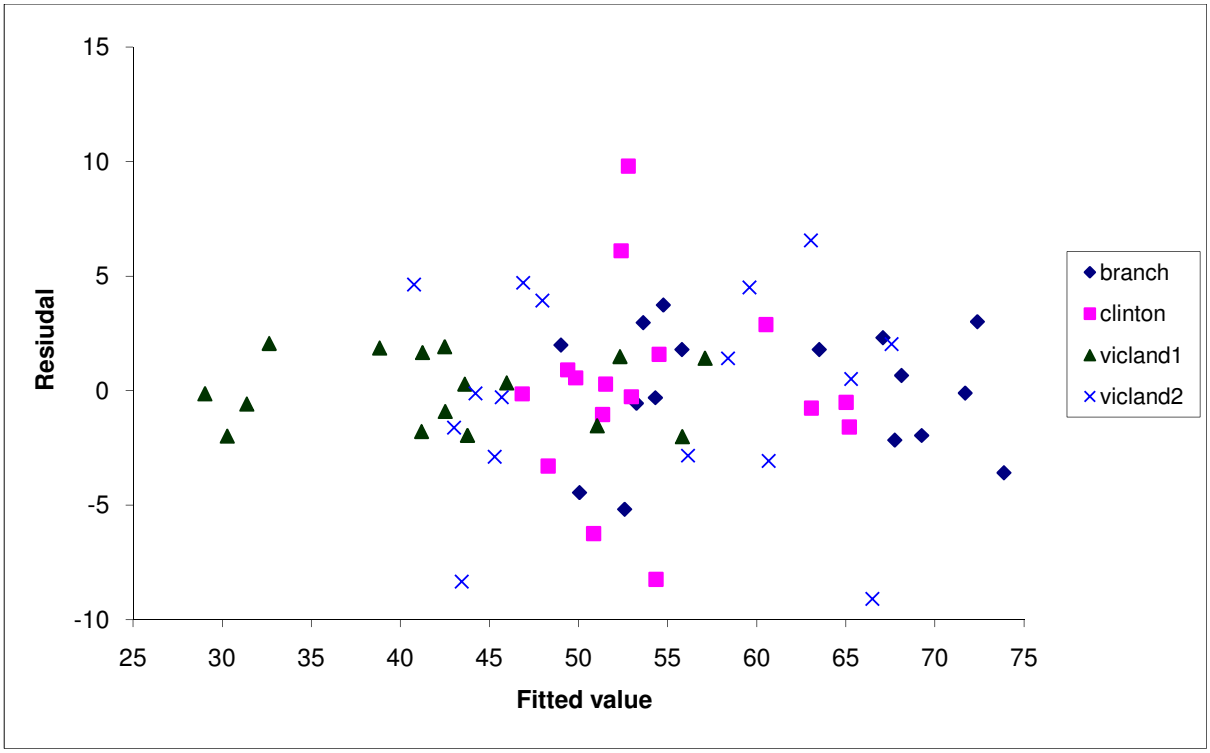


Bartlett's test for homogeneity of variances

Chi-square 9.75 on 3 degrees of freedom: probability 0.021

Steel and Torrie give further information about these varieties. Vicland (1) is a variety infected with *H. victoriae*, Vicland (2) is the same variety but is not infected. Clinton and Branch are varieties resistant to *H. victoriae*. The variation in the Vicland (1) data appears smaller than for the other varieties. It is possible that the actual levels of this factor are associated with different variances: one level is expected to have consistently smaller yields, since these seeds have been infected. Linear Mixed Models (REML) allows us to model this.

Has the combined analysis overlooked this problem? If we **Save** the fitted values and residuals, we can obtain a residual plot with different colours for the different varieties.



In this plot, the residuals from Vicland (1) appear less varied than the other varieties (corresponding to the significantly smaller variance in the yields of this variety). It would appear that the combined split-plot analysis is inappropriate for these data.

LMM (REML) analysis of split-plot design (in randomized blocks)

For this split-plot there are three strata: blocks, whole-plots and split-plots. Hence, the **Random Model** is Block/W_Plot/S_Plot. In order to allow a changing variance across cultivars, we need to mention them in the **Random Model**. Cultivars were allocated at random to the whole plots, so we can express the **Random Model** as Block/Cultivar/S_Plot, Block/Cultivar/Chemical, or simply as Block/Cultivar since the final stratum can be omitted. The stratum variances were estimated in ANOVA as follows:

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Block	947.624	3.000	54.933
Block.Cultivar	68.699	9.000	12.097
Block.Cultivar.Chemical	20.311	36.000	20.311

Standard split-plot analysis via LMM (REML)

REML variance components analysis

Response variate: Yield
Fixed model: Constant + Cultivar + Chemical + Cultivar.Chemical
Random model: Block + Block.Cultivar + Block.Cultivar.Chemical
Number of units: 64

Block.Cultivar.Chemical used as residual term

Estimated variance components

Random term	component	s.e.
Block	**54.93**	48.40
Block.Cultivar	**12.10**	8.18

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
	Block.Cultivar.Chemical	Identity	Sigma2	**20.31**	4.79

Deviance: -2*Log-Likelihood

Deviance	d.f.
237.21	45

Wald tests for fixed effects

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Cultivar	41.46	3	13.82	9.0	0.001
Chemical	8.40	3	2.80	36.0	0.054
Cultivar.Chemical	28.87	9	3.21	36.0	0.006

Table of predicted means for Constant

52.81 Standard error: 3.848

Table of predicted means for Cultivar

Cultivar	Branch	Clinton	Vicland (1)	Vicland (2)
	61.07	54.31	42.46	53.41

Standard error of differences: 2.930

Table of predicted means for Chemical

Chemical	Control	Ceresan	Panogen	Agrox
	50.69	55.20	53.12	52.22

Standard error of differences: 1.593

Table of predicted means for Cultivar.Chemical

Chemical Cultivar	Control	Ceresan	Panogen	Agrox
Branch	61.92	63.42	57.67	61.25
Clinton	53.93	51.38	55.88	56.05
Vicland (1)	36.05	50.63	45.85	37.30
Vicland (2)	50.85	55.38	53.10	54.30

Standard errors

Chemical Cultivar	Agrox	Ceresan	Control	Panogen
Branch	4.67	4.67	4.67	4.67
Clinton	4.67	4.67	4.67	4.67
Vicland (1)	4.67	4.67	4.67	4.67
Vicland (2)	4.67	4.67	4.67	4.67

LMM (REML) gives the same means, s.e.m., s.e.d. and l.s.d. values as ANOVA, but in full matrix form.

Next, we demonstrate how to check for changing variance across cultivars. Given the nature of the cultivars and seed chemical protectants, we might expect this variance to change only at the split-plot level. The following change in deviance table explores various models for Cultivar in firstly the split-plot error term (Block.Cultivar.Chemical) and then in the whole-plot error term (Block.Cultivar).

Model for Cultivar in Block.Cultivar	Model for Cultivar in Block.Cultivar.Chemical	deviance	d.f.	change in deviance	change in d.f.	*P* value
Identity	Identity	237.21	45			
Identity	Diagonal	225.78	42	**11.43**	**3**	**0.010**
Diagonal	Diagonal	223.69	39	2.09	3	0.554

The analysis allowing for a changing variance at the split-plot level is as follows. Use **Save** if you want to take the s.e.d. values into Excel or Word most efficiently.

REML variance components analysis

Response variate:	Yield
Fixed model:	Constant + Cultivar + Chemical + Cultivar.Chemical
Random model:	Block + Block.Cultivar + Block.Cultivar.Chemical
Number of units:	64

Block.Cultivar.Chemical used as residual term with covariance structure as below

Sparse algorithm with AI optimisation

Covariance structures defined for random model

Covariance structures defined within terms:

Term	Factor	Model	Order	No. rows
Block.Cultivar.Chemical	Block	Identity	0	4
	Cultivar	**Diagonal**	**4**	**4**
	Chemical	Identity	0	4

Allowing the variance to change across cultivars

Estimated variance components

Random term	component	s.e.
Block	55.842	48.137
Block.Cultivar	7.728	6.384

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
	Block.Cultivar.Chemical	Sigma2	1.000	fixed	
	Block	Identity	-	-	-
	Cultivar	Diagonal	d_1	12.81	5.98
			d_2	33.19	16.03
			d_3	4.060	1.898
			d_4	37.03	17.41
	Chemical	Identity	-	-	-

These were the four individual Residual MS from separate RCBD analyses, one for each cultivar.

Deviance: -2*Log-Likelihood

Deviance	d.f.
225.78	42

Wald tests for fixed effects

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Cultivar	73.54	3	24.04	7.5	<0.001
Chemical	93.98	3	31.33	19.4	<0.001
Cultivar.Chemical	58.23	9	5.90	18.8	<0.001

Dropping individual terms from full fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Cultivar.Chemical	58.23	9	5.90	18.8	<0.001

This section is not important in this analysis, but, without the interaction in the model, P values for both main effects *when entered last* would be available here. Important for unbalanced designs.

Table of predicted means for Constant

52.81 Standard error: 3.845

Table of predicted means for Cultivar

Cultivar	Branch	Clinton	Vicland (1)	Vicland (2)
	61.07	54.31	42.46	53.41

Standard errors of differences between pairs

Cultivar Branch	1	*			
Cultivar Clinton	2	2.60	*		
Cultivar Vicland (1)	3	2.22	2.49	*	
Cultivar Vicland (2)	4	2.64	2.87	2.54	*
		1	2	3	4

`

Table of predicted means for Chemical

Chemical	Control	Ceresan	Panogen	Agrox
	50.69	55.20	53.12	52.22

Standard errors of differences between pairs

Chemical Control	1	*			
Chemical Ceresan	2	1.65	*		
Chemical Panogen	3	1.65	1.65	*	
Chemical Agrox	4	1.65	1.65	1.65	*
		1	2	3	4

Table of predicted means for Cultivar.Chemical

Chemical Cultivar	Control	Ceresan	Panogen	Agrox
Branch	61.92	63.42	57.67	61.25
Clinton	53.93	51.38	55.88	56.05
Vicland (1)	36.05	50.63	45.85	37.30
Vicland (2)	50.85	55.38	53.10	54.30

Standard errors

Chemical Cultivar	Agrox	Ceresan	Control	Panogen
Branch	4.37	4.37	4.37	4.37
Clinton	4.92	4.92	4.92	4.92
Vicland (1)	4.11	4.11	4.11	4.11
Vicland (2)	5.01	5.01	5.01	5.01

Standard errors of differences between pairs

Cultivar Branch.Chemical Control	1	*			
Cultivar Branch.Chemical Ceresan	2	2.53	*		
Cultivar Branch.Chemical Panogen	3	2.53	2.53	*	
…	…	…	…	…	…
Cultivar Vicland (2).Chemical Agrox	16	4.04	4.04	4.04	4.04
		1	2	3	4

etc.

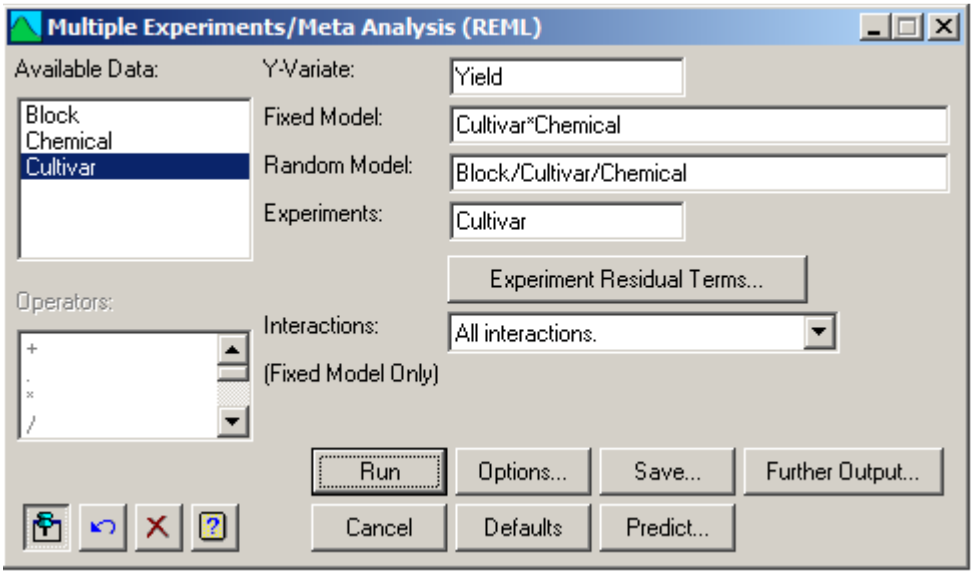
Notice that the s.e.m. values are all higher than those obtained from the split-plot ANOVA, which were given as 2.846 (the first of the two possibilities). For the ANOVA, the block effect sums to 0 for each mean so the block effect is not part of the calculation.

Standard errors of means

Table	Cultivar	Chemical	**Cultivar Chemical**
rep.	16	16	**4**
e.s.e.	2.072	1.127	**2.846**
d.f.	9	36	**26.78**
Except when comparing means with the same level(s) of			
Cultivar			**2.253**
d.f.			**36**

Meta Analysis (REML) analysis

Since the variance appears to change across a single factor (Cultivar), the analysis is simply performed using Stats > Meta Analysis > REML of Multiple Experiments. The fixed and random models are those from ANOVA or LMM; we simply declare Cultivar as the "notional" factor over which the residual changes across "Experiments":



REML variance components analysis

Response variate: Yield
Fixed model: Constant + Cultivar + Chemical + Cultivar.Chemical
Random model: Block + Block.Cultivar + Block.Cultivar.Chemical
Number of units: 64

Separate residual terms for each level of experiment factor: Cultivar

Sparse algorithm with AI optimisation

Estimated variance components

Random term	component	s.e.
Block	55.84	48.14
Block.Cultivar	7.73	6.38
Block.Cultivar.Chemical	1.00	aliased

Residual model for each experiment

Experiment factor: Cultivar

Experiment	Term	Factor	Model(order)	Parameter	Estimate	s.e.
Branch		Residual	Identity	Variance	11.81	5.98
Clinton		Residual	Identity	Variance	32.19	16.03
Vicland (1)		Residual	Identity	Variance	3.060	1.898
Vicland (2)		Residual	Identity	Variance	36.03	17.41

Deviance: -2*Log-Likelihood

Deviance	d.f.
225.78	42

Note: deviance omits constants which depend on fixed model fitted.

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Cultivar	73.54	3	24.04	7.5	<0.001
Chemical	93.98	3	31.33	19.4	<0.001
Cultivar.Chemical	58.23	9	5.90	18.8	<0.001

Notice that the parameterization for the variances is slightly different here. A random error term is added with a variance σ^2 whose estimate (1.00) is shown as "aliased":

Random term	component	s.e.
Block.Cultivar.Chemical	1.00	aliased

This value needs to be added to the separate estimates of variances for the four cultivars (e.g. for Branch, the estimate of variance is 11.81+1.00 = 12.81, which was the the Residual MS from the RCB analysis of the Branch data in the four blocks with the four chemical treatments).

With the more realistic modeling of changing variances across cultivars in the split-plot experiment, sem and sed values all change. Selecting to show Standard Errors of All Estimates in the options shows the effect of this change. With a constant variance model, the sem value is 4.67. With a changing variance model, it varies from a low 4.11 to a high 5.01. Unlike the ANOVA, the calculation of the s.e.m. value involves the block variance, the whole-plot variance and the split-plot variance.

Table of predicted means for Cultivar.Chemical

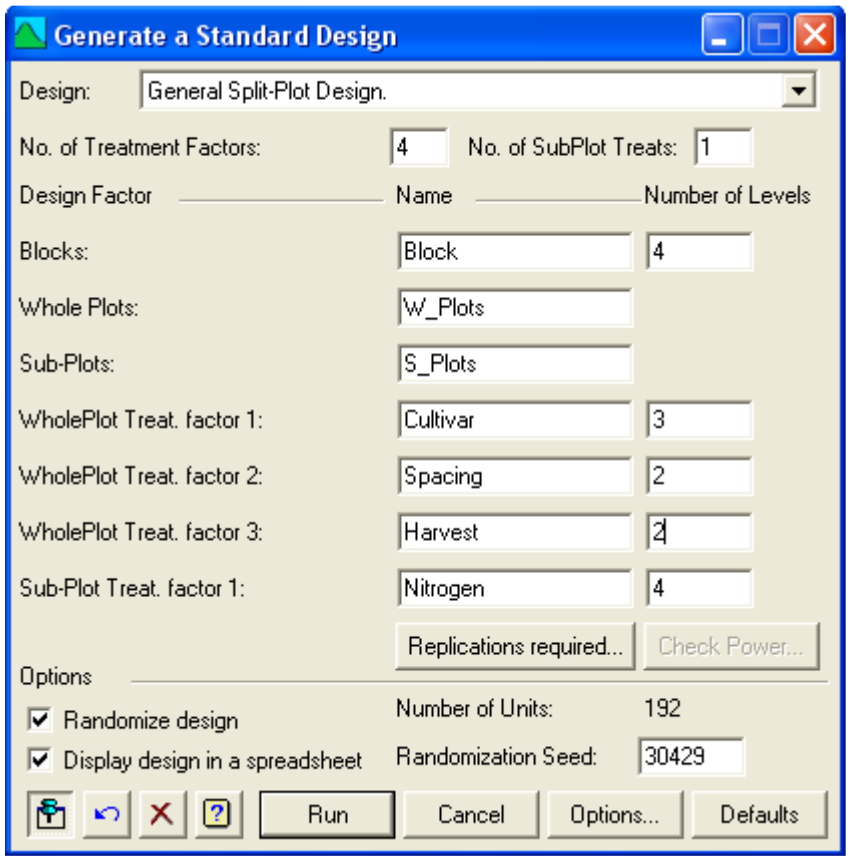
Chemical Cultivar	Agrox	Ceresan	Control	Panogen
Branch	61.25	63.43	61.93	57.67
Clinton	56.05	51.38	53.93	55.88
Vicland (1)	37.30	50.62	36.05	45.85
Vicland (2)	54.30	55.38	50.85	53.10

Standard errors

Chemical Cultivar	Agrox	Ceresan	Control	Panogen
Branch	4.37	4.37	4.37	4.37
Clinton	4.92	4.92	4.92	4.92
Vicland (1)	4.11	4.11	4.11	4.11
Vicland (2)	5.01	5.01	5.01	5.01

General split-plot design

The split-plot design in the previous section had just one treatment factor applied to whole-plots and to split-plots. There is no restriction on the treatment structure in either stratum. GenStat's Design menu allows for a general split-plot design. You simply indicate how many treatment factors there are altogether, and how many of these are allocated to split-units. The following example produces a random design with cultivar × spacing × harvest treatments (3×2×4 = 24 combinations) allocated to whole-plots, and four levels of nitrogen allocated to split-plots within each whole-plot.



GenStat creates, as before, a Block stratum, a W_Plot stratum and a S_Plot stratum. This time, there are three factors required to fully define the whole-plots. Nevertheless, the Block Structure remains as Block/W_Plot/S_Plot.

PlotNo	Block!	W_Plots!	S_Plots!	Cultivar!	Spacing!	Harvest!	Nitrogen!
1101	1	1	1	2	2	2	1
1102	1	1	2	2	1	2	2
1103	1	1	3	2	2	2	4
1104	1	1	4	2	1	1	3
1105	1	1	5	2	1	2	3
1106	1	1	6	2	1	2	1
1107	1	1	7	2	1	1	2
1108	1	1	8	2	2	2	3
1109	1	1	9	2	2	1	4
1110	1	1	10	2	2	2	2
1111	1	1	11	2	1	1	1
1112	1	1	12	2	2	1	3
1113	1	1	13	2	1	1	4
1114	1	1	14	2	2	1	2
1115	1	1	15	2	2	1	1
1116	1	1	16	2	1	2	4
1201	1	2	1	3	2	1	2

etc …

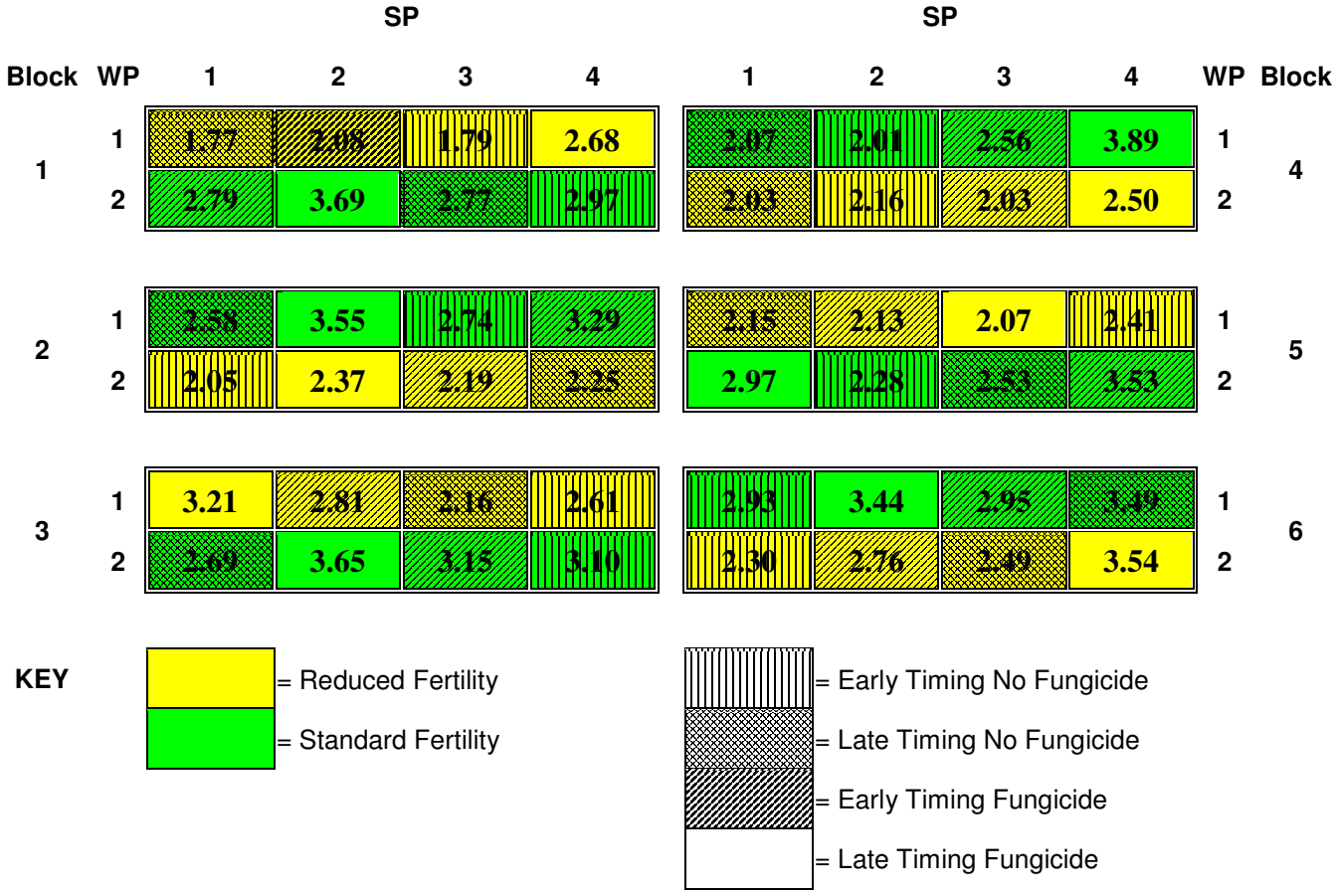
Split-plot design with a two-way factorial split treatment structure

Curt Lee (Agro-Tech, Inc., Velva, North Dakota, USA) kindly supplied data from the following experiment on wheat.

Six blocks were set up and each divided into two whole-plots (WP). One whole-plot was randomly fertilized with a full recommended rate of nitrogen fertilizer (Standard), the other not fertilized (Reduced). The final applied-N plus residual-N was 100 lbs for the standard fertility and 50 lbs for the reduced fertility plots.

Each whole-plot was divided into four split-plots (SP). The four treatments allocated randomly to these plots were a fungicide treatment (or a blank treatment), and an early (at the tillering stage) or a late (at the flag leaf stage) application of the fungicide and the blank.

Example 14 Wheat split-plot experiment with a factorial split-plot treatment structure



The blank plots were sprayed with the treatments that contained all the carrier material (water, solvents, etc), except the active ingredient. Thus, since a treatment was actually applied to the blank plots, the split-plot treatments can be thought of as a 2×2 factorial combination.

Alternatively, you can think of the split-plot treatments as a simple set of four treatments, and extract three **contrasts** to estimate the following characteristics.

a) Estimate the effect of the fungicide versus no fungicide, by comparing the mean yields from fungicide (early and late) plots to no fungicide (early and late) plots.

b) Estimate the effect of different timing by comparing the mean yields from fungicide early plots to fungicide late plots.

c) Estimate the effect of the two check treatments by comparing the mean yields from no fungicide early plots to no fungicide late plots. (They should yield the same, unless they are getting something out of the carrier materials.)

With the split-plot treatment as a 2 × 2 factorial

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	5	2.1795	0.4359	4.18	
Block.W_Plot stratum					
Fert	1	4.7376	4.7376	45.38	0.001
Residual	5	0.5220	0.1044	0.97	
Block.W_Plot.S_Plot stratum					
Fung	1	2.7552	2.7552	25.51	<.001
Timing	1	0.5043	0.5043	4.67	0.039
Fert.Fung	1	0.2002	0.2002	1.85	0.183
Fert.Timing	1	0.0261	0.0261	0.24	0.626
Fung.Timing	1	0.6674	0.6674	6.18	0.019
Fert.Fung.Timing	1	0.0000	0.0000	0.00	0.993
Residual	30	3.2398	0.1080		
Total	47	14.8322			

Message: the following units have large residuals.

Block 4 W_Plot 1 S_Plot 4	**0.710**	**s.e.**	**0.260**
Block 5 W_Plot 2 S_Plot 4	**0.642**	**s.e.**	**0.260**
Block 6 W_Plot 1 S_Plot 4	**0.583**	**s.e.**	**0.260**

Tables of means

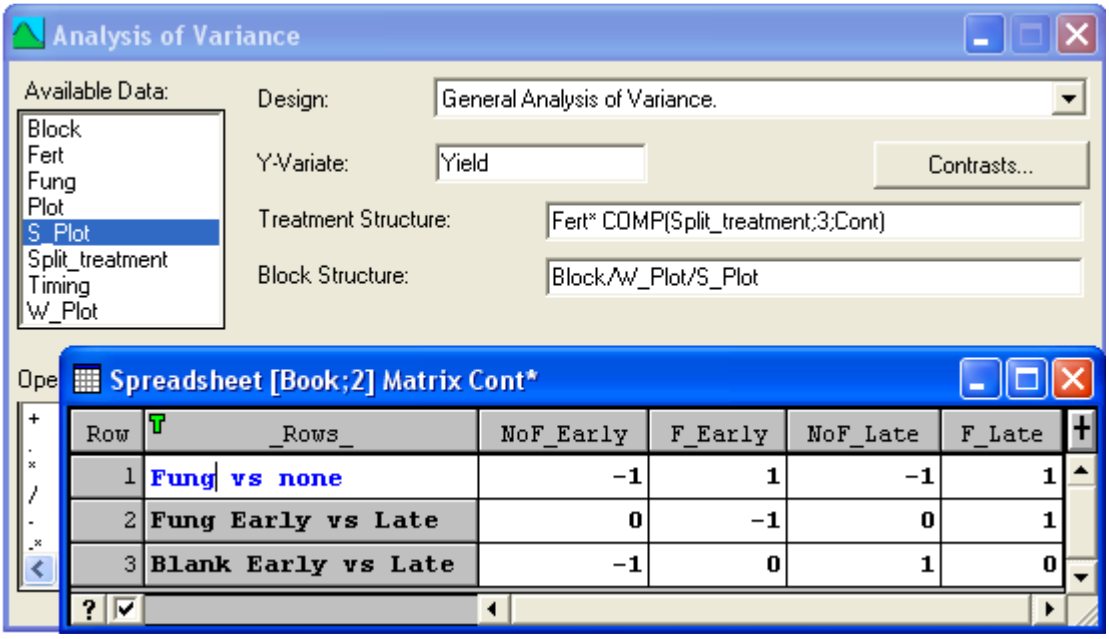
etc…

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Block	0.4359	5.000	0.0414
Block.W_Plot	0.1044	5.000	-0.0009
Block.W_Plot.S_Plot	0.1080	30.000	0.1080

The three residuals were all from edge plots in blocks 4, 5 and 6. On checking, the research company discovered that these plots had not been trimmed to equal length. For their analysis they went back, measured each plot and corrected the yield based on actual harvested plot length. We will not do that here.

With the split-plot treatment as 4 simple treatments with structure



Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	5	2.1795	0.4359	4.18	
Block.W_Plot stratum					
Fert	1	4.7376	4.7376	45.38	0.001
Residual	5	0.5220	0.1044	0.97	
Block.W_Plot.S_Plot stratum					
Split_treatment	3	3.9269	1.3090	12.12	<.001
Fung vs none	**1**	**2.7552**	**2.7552**	**25.51**	**<.001**
Fung Early vs Late	**1**	**1.1660**	**1.1660**	**10.80**	**0.003**
Blank Early vs Late	1	0.0057	0.0057	0.05	0.820
Fert.Split_treatment	3	0.2263	0.0754	0.70	0.560
Fert.Fung vs none	1	0.2002	0.2002	1.85	0.183
Fert.Fung Early vs Late	1	0.0126	0.0126	0.12	0.735
Fert.Blank Early vs Late	1	0.0135	0.0135	0.13	0.726
Residual	30	3.2398	0.1080		
Total	47	14.8322			

Message: the following units have large residuals.

Block 4 W_Plot 1 S_Plot 4	0.710	s.e.	0.260
Block 5 W_Plot 2 S_Plot 4	0.642	s.e.	0.260
Block 6 W_Plot 1 S_Plot 4	0.583	s.e.	0.260

Tables of effects and contrasts

Block.W_Plot.S_Plot stratum

Split_treatment contrasts

Fung vs none 0.96, s.e. 0.190, ss.div. 3.00

Fung Early vs Late 0.44, s.e. 0.134, ss.div. 6.00

Blank Early vs Late -0.03, s.e. 0.134, ss.div. 6.00

Fert.Split_treatment contrasts

Fert.Fung vs none, e.s.e. 0.268, ss.div. 1.50

Fert	Reduced	Standard
	-0.26	0.26

Fert.Fung Early vs Late, e.s.e. 0.190, ss.div. 3.00

Fert	Reduced	Standard
	-0.05	0.05

Fert.Blank Early vs Late, e.s.e. 0.190, ss.div. 3.00

Fert	Reduced	Standard
	-0.05	0.05

From the ANOVA, we see that:

- applying the fungicide late, at the flag leaf stage, gives significantly better yields ($P = 0.003$). The difference in means (for which see below) is 0.44 ($\pm$ 0.134) kg/plot.
- Using fungicide has a yield advantage, on average, of $\frac{1}{2}(0.96 \pm 0.190) = 0.48 \pm 0.095$ kg/plot ($P < 0.001$). The ½ arises because the contrast we want is $\frac{1}{2}(\mu_2+\mu_4)-\frac{1}{2}(\mu_1+\mu_3)$ and we currently have $(\mu_2+\mu_4)-(\mu_1+\mu_3)$ which is essentially row 1 of the contrast matrix in the screen capture above.

Tables of means

Grand mean 2.670

Fert	Reduced	Standard
	2.356	2.984

Split_treatment	NoF_Early	F_Early	NoF_Late	F_Late
	2.446	2.689	2.415	3.130

Fert	Split_treatment	NoF_Early	F_Early	NoF_Late	F_Late
Reduced		2.220	2.333	2.142	2.728
Standard		2.672	3.045	2.688	3.532

Standard errors of differences of means

Table	Fert	Split_treatment	Fert Split_treatment
rep.	24	12	6
s.e.d.	0.0933	0.1342	0.1889
d.f.	5	30	32.32

Except when comparing means with the same level(s) of

Fert	0.1897
d.f.	30

This design is straightforward and will not be repeated in LMM (REML).

Possible field layout for split-split-plot experiment

Block 1

1 2 3	13000
3 2 1	16000
3 1 2	10000
3 1 2	16000
2 3 1	13000
2 3 1	10000

Block 2

3 2 1	10000
2 3 1	16000
2 1 3	13000
3 1 2	13000
1 2 3	16000
1 3 2	10000

Block 3

2 3 1	16000
3 2 1	10000
3 1 2	13000
1 2 3	16000
1 3 2	10000
3 2 1	13000

Block 4

1 3 2	10000
1 3 2	16000
1 2 3	13000
2 1 3	10000
3 1 2	16000
3 2 1	13000

Key to fertilizer:

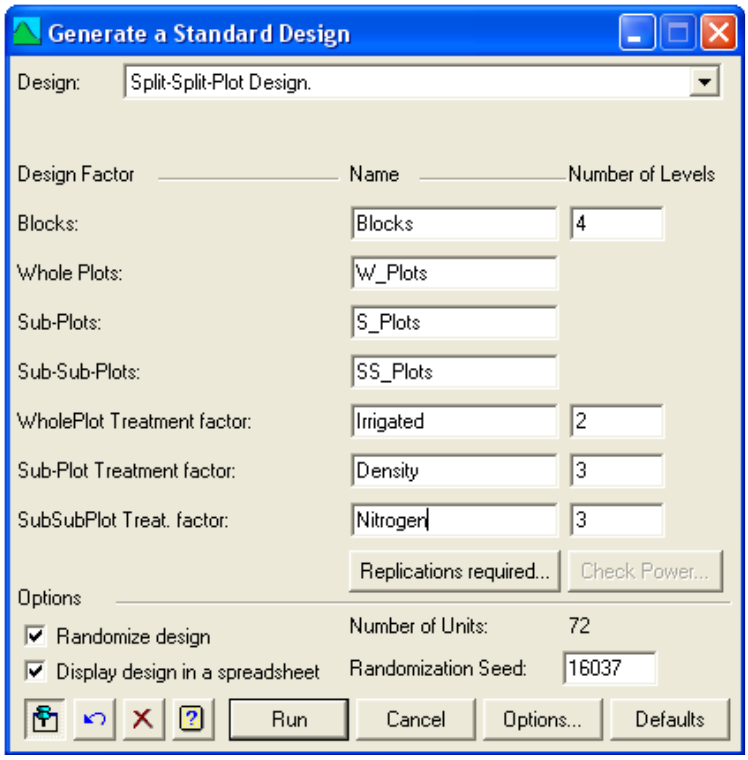
1= 60 lb nitrogen
2=120 lb nitrogen
3=180 lb nitrogen

Key to irrigation:

Irrigated	
Non-irrigated	

Key to Spacing:

Spacing	13000	
Spacing	16000	
Spacing	10000	



Split-split-plot design (in randomized blocks)

An experiment was conducted to determine that effects of irrigation, planting density (or stand), and fertilizer level on the yield of corn. The smallest area that could be irrigated was half a block – or one whole-plot. The two irrigation treatments were randomly allocated to the whole-plots in each of four blocks. Each whole-plot was divided into three split-plots, and with three planting densities (rates of 10,000, 13,000 and 16,000 plants acre^{-1}) randomly allocated to each. Finally, each split-plot was divided into three split-split-plots, with three fertilisers (60, 120 and 180 lb of nitrogen) randomly allocated to each.

This is quite a different layout compared to a simple RCBD in which all 18 treatment combinations could occur in any plot of each block. In this case, practical limitations dictated the layout; the penalty is a more complex analysis. The **Block Structure** comes about as follows.

- Blocks were identified in the field, so Block forms the first stratum.
- Half block areas were prepared and one of these in each block was (randomly) irrigated, forming a Block.Irrigated stratum. Irrigated and non-irrigated plot means are compared within this stratum, which is basically an RCBD with 4 blocks and 2 treatments.
- Each half-block was split into three areas and one of three spacings used (randomly) in each. Thus, we have a third stratum, Block.Irrigated.Spacing, and these units are used in constructing Spacing and Spacing.Irrigated *F*-tests.
- Each spacing strip was split into three even smaller areas and one of three fertilisers applied (randomly) in each. This gives rise to a fourth and final stratum, Block.Irrigated.Spacing.Fertiliser, and these units are used in constructing *F*-tests for the Fertiliser main effect and any interaction involving this factor.

To summarise, the **Block Structure** is

Block + Block.Irrigated + Block.Irrigated.Spacing + Block.Irrigated.Spacing.Fertiliser

which simplifies to Block/Irrigated/Spacing/Fertiliser.

Example 15 Yields of corn (bushels acre^{-1}) from Snedecor & Cochran page 328

		Non-irrigated blocks				Irrigated blocks			
Stand	Fertilizer	1	2	3	4	1	2	3	4
10,000	60	90	83	85	86	80	102	60	73
	120	95	80	88	78	87	109	104	114
	180	107	95	88	89	100	105	114	114
13,000	60	92	98	112	79	121	99	90	109
	120	89	98	104	86	110	94	118	131
	180	92	106	91	87	119	123	113	126
16,000	60	81	74	82	85	78	136	119	116
	120	92	81	78	89	98	133	122	136
	180	93	74	94	83	122	132	136	133

Analysis of variance

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	3	194.44	64.81	0.14	
Block.Irrigated stratum					
Irrigated	1	8277.56	8277.56	17.59	0.025
Residual	3	1411.78	470.59	2.03	
Block.Irrigated.Stand stratum					
Stand	2	1758.36	879.18	3.78	0.053
Irrigated.Stand	2	2747.03	1373.51	5.91	0.016
Residual	12	2787.94	232.33	2.69	
Block.Irrigated.Stand.Fertilizer stratum					
Fertilizer	2	1977.44	988.72	11.45	<.001
Irrigated.Fertilizer	2	953.44	476.72	5.52	0.008
Stand.Fertilizer	4	304.89	76.22	0.88	0.484
Irrigated.Stand.Fertilizer	4	234.72	58.68	0.68	0.611
Residual	36	3108.83	86.36		
Total	71	23756.44			

Message: the following units have large residuals.

Block 1 Irrigated Irrigated Stand 13,000	12.7	s.e.	6.2
Block 1 Irrigated Irrigated Stand 16,000	-13.6	s.e.	6.2
Block 2 Irrigated Irrigated Stand 10,000 Fertilizer 60.	14.7	s.e.	6.6
Block 3 Irrigated Irrigated Stand 10,000 Fertilizer 60.	-14.6	s.e.	6.6

Tables of means

Grand mean 99.7

Irrigated	Non-irrigated	Irrigated
	89.0	110.4

Stand	10,000	13,000	16,000
	92.8	103.6	102.8

Fertilizer	60.	120.	180.
	92.9	100.6	105.7

Irrigated	Stand	10,000	13,000	16,000
Non-irrigated		88.7	94.5	83.8
Irrigated		96.8	112.7	121.8

Irrigated	Fertilizer	60.	120.	180.
Non-irrigated		87.3	88.2	91.6
Irrigated		98.6	113.0	119.8

Stand	Fertilizer	60.	120.	180.
10,000		82.4	94.4	101.5
13,000		100.0	103.8	107.1
16,000		96.4	103.6	108.4

Irrigated	Stand	Fertilizer	60.	120.	180.
Non-irrigated	10,000		86.0	85.3	94.8
	13,000		95.2	94.2	94.0
	16,000		80.5	85.0	86.0
Irrigated	10,000		78.8	103.5	108.3
	13,000		104.7	113.2	120.2
	16,000		112.2	122.3	130.7

Standard errors of differences of means

Table	Irrigated	Stand	Fertilizer	Irrigated Stand
rep.	36	24	24	12
s.e.d.	5.11	4.40	2.68	7.21
d.f.	3	12	36	9.53
Except when comparing means with the same level(s) of				
Irrigated				6.22
d.f.				12

Table	Irrigated Fertilizer	Stand Fertilizer	Irrigated Stand Fertilizer
rep.	12	8	4
s.e.d.	5.98	5.81	8.99
d.f.	5.54	30.80	21.28
Except when comparing means with the same level(s) of			
Irrigated	3.79		8.22
d.f.	36		30.80
Stand		4.65	
d.f.		36	
Irrigated.Stand			6.57
d.f.			36
Irrigated.Fertilizer		8.22	
d.f.			30.80

Least significant differences of means (5% level)

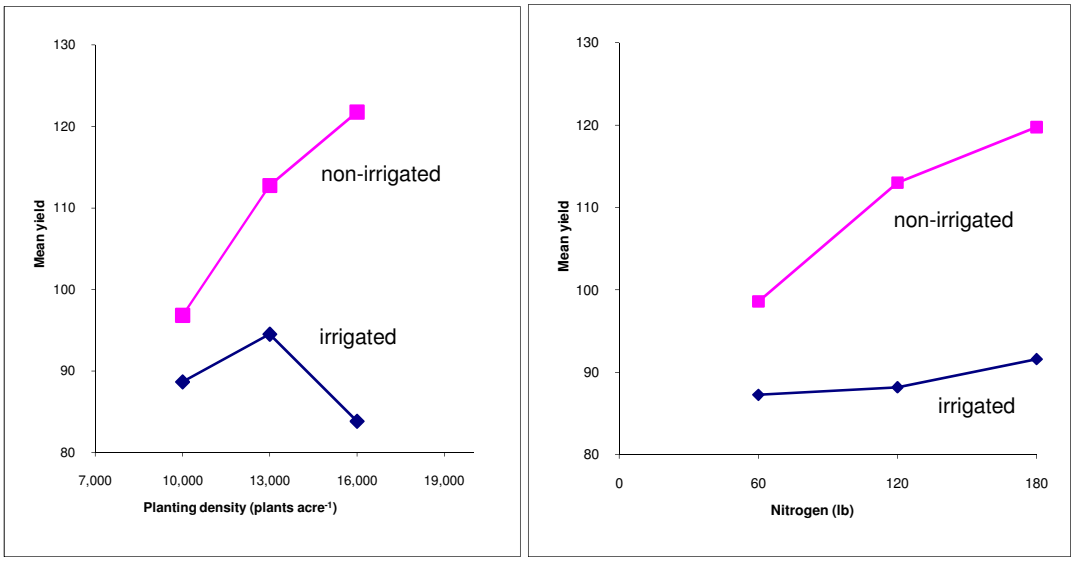
Table	Irrigated	Stand	Fertilizer	Irrigated Stand
rep.	36	24	24	12
l.s.d.	16.27	9.59	5.44	16.17
d.f.	3	12	36	9.53
Except when comparing means with the same level(s) of				
Irrigated				13.56
d.f.				12

Table	Irrigated Fertilizer	Stand Fertilizer	Irrigated Stand Fertilizer
rep.	12	8	4
l.s.d.	14.92	11.85	18.67
d.f.	5.54	30.80	21.28
Except when comparing means with the same level(s) of			
Irrigated	7.69		16.76
d.f.	36		30.80
Stand		9.42	
d.f.		36	
Irrigated.Stand			13.33
d.f.			36
Irrigated.Fertilizer			16.76
d.f.			30.80

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Block	64.81	3.000	-22.54
Block.Irrigated	470.59	3.000	26.47
Block.Irrigated.Stand	232.33	12.000	48.66
Block.Irrigated.Stand.Fertilizer	86.36	36.000	86.36

Comparing 2-way and 3-way means is now a complex procedure. Note, however, that comparing two densities (/two fertilizers) both of which were irrigated (or non-irrigated) is straightforward (the l.s.d. values are 13.56/7.69), and so on. The differences in means come down to two significant interactions, and the following plots make these differences clear:

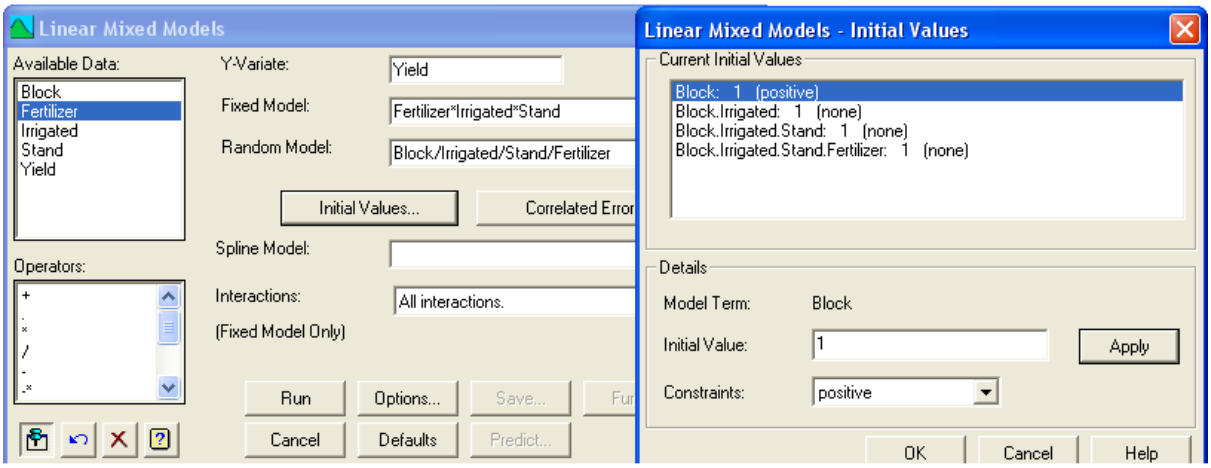


Note that the Block MS is smaller than the highest stratum Residual MS, which is unusual. When analysing via REML we would be advised to force variance components to be positive. In the analysis above, we also ignored the potential variance problem we discussed previously brought about by having varying planting densities.

LMM (REML) analysis

This experiment illustrates the occasional need to restrict the variance estimates to be positive. In the ANOVA, the variance of the block stratum was estimated as -22.54 simply because the Block MS was smaller than the Residual MS in the whole-plot analysis. This indicates the absence of any block effect.

For a split-split-plot design there are four strata, the **Fixed Model** being the same as the **Treatment Structure** of ANOVA (Fertilizer*Irrigated*Stand) and the **Random Model** being the same as the **Block Structure** (Block/Irrigated/Stand/Fertilizer). To ensure that all stratum variances are positive, you need to click **Initial Values**, choose Block and select **positive** for **Constraints**.



REML variance components analysis

Response variate: Yield
Fixed model: Constant + Irrigated + Stand + Fertilizer + Irrigated.Stand + Irrigated.Fertilizer + Stand.Fertilizer + Irrigated.Stand.Fertilizer
Random model: Block + Block.Irrigated + Block.Irrigated.Stand + Block.Irrigated.Stand.Fertilizer
Number of units: 72

Block.Irrigated.Stand.Fertilizer used as residual term

Estimated variance components

Random term	component	s.e.
Block	**0.00**	**bound**
Block.Irrigated	3.93	20.15
Block.Irrigated.Stand	48.66	32.34

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
	Block.Irrigated.Stand.Fertilizer	Identity	Sigma2	86.36	20.35

Deviance: -2*Log-Likelihood

Deviance	d.f.
338.38	50

Wald tests for fixed effects

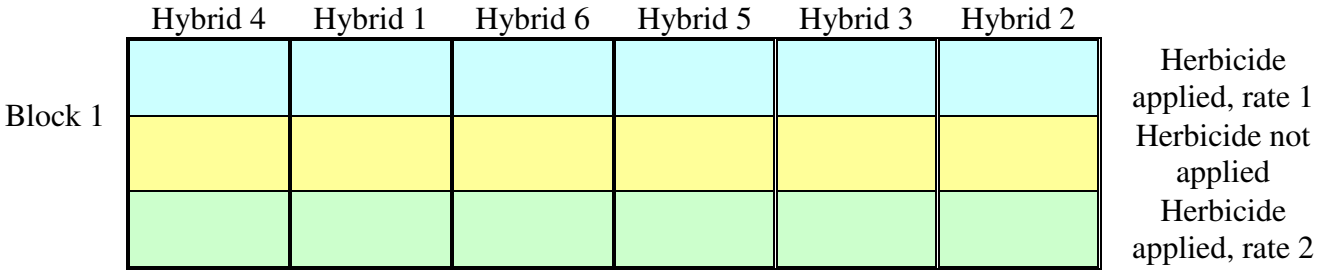
Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Irrigated	30.92	1	30.92	6.0	0.001
Stand	7.57	2	3.78	12.0	0.053
Fertilizer	22.90	2	11.45	36.0	<0.001
Irrigated.Stand	11.82	2	5.91	12.0	0.016
Irrigated.Fertilizer	11.04	2	5.52	36.0	0.008
Stand.Fertilizer	3.53	4	0.88	36.0	0.484
Irrigated.Stand.Fertilizer	2.72	4	0.68	36.0	0.611

Dropping individual terms from full fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Irrigated.Stand.Fertilizer	2.72	4	0.68	36.0	0.611

Criss-cross/split-block/strip-plot design

This design has various names in the literature, but the essential difference is that a second (possibly factorially structured) treatment is randomly applied across large areas of each block, generally at right angles to the first treatment. For example, this is one block from a factorial trial in which hybrids are allocated to four plots in the block, and a herbicide treatment (absent, or one of two rates) is applied to one-three block areas stripped across the plots.



A corresponding split-plot design has the herbicide treatment applied at random to the three small plots within each whole-plot. This more complex arrangement is often the only practical way of running the experiment, but comes at the cost of greater complexity in treatment comparisons.

The levels of the herbicide treatment are also applied to large areas in each block. Thus, there are two types of whole-plots. There are now four strata: Block, Block.Hybrid, Block.Herbicide, and Block.Hybrid.Herbicide (an individual plots whose yields are measured).

Example 16 Curt Lee (Agro-Tech, Inc., Velva, North Dakota, USA) kindly supplied data from the following experiment on sunflower (yield in lb/acre). Hybrid number shown in each block (V1 to V7).

Block	Herbicide	V1	V2	V3	V4	V5	V6	V7
1	check	810.6	1369.7	1830.8	1335.8	1563.6	1419.5	726.8
	rate 1	776.8	1115.4	1497.0	1610.8	1637.0	1236.2	679.4
	rate 2	595.2	1175.9	1260.0	1204.3	1465.2	1172.2	669.8
		V6	V5	V4	V7	V2	V1	V3
2	rate 1	1429.4	1152.8	1150.4	744.1	1099.0	735.2	1413.9
	check	1517.5	1971.4	1737.6	643.4	916.2	608.3	1747.6
	rate 2	1696.2	1467.0	1456.1	662.8	906.7	562.5	1417.2
		V4	V2	V6	V7	V3	V1	V5
3	rate 2	1383.6	1328.2	1301.4	671.6	1805.0	709.7	1536.6
	check	1638.7	1250.8	1411.5	762.6	1827.9	601.4	1685.0
	rate 1	1727.8	1201.4	1576.8	748.2	1340.2	670.8	2193.3
		V4	V1	V7	V2	V5	V6	V3
4	rate 1	1414.4	562.3	833.6	1085.4	1480.6	1323.9	1683.9
	rate 2	1329.2	845.3	884.5	1069.9	1822.1	1277.1	1734.2
	check	1318.4	760.4	842.6	1147.4	1729.5	1212.6	1450.5

It is common practice to place treatments for dose response experiments in sequential order (not randomized) in the first block of a field trial. This is used to accommodate farmer tours so they may walk through the trial and see the expected differences. There is a debate as to whether the demonstration block should be used as part of the research data, but we will do so here.

Using ANOVA, the **Treatment Structure** is clearly Hybrid*Herbicide.

The **Block Structure** is slightly more complex to formulate with a shortcut. The four strata mentioned above technically is all that is needed to set up the block structure, so:

Block + Block.Hybrid + Block.Herbicide + Block.Hybrid.Herbicide

which by the rules is abbreviated to Block/(Hybrid*Herbicide).

Analysis of variance

Variate: Yield

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	3	165289.	55096.		
Block.Hybrid stratum					
Hybrid	6	10890886.	1815148.	65.92	<.001
Residual	18	495649.	27536.	0.83	
Block.Herbicide stratum					
Herbicide	2	111333.	55666.	1.22	0.360
Residual	6	274730.	45788.	1.38	
Block.Hybrid.Herbicide stratum					
Hybrid.Herbicide	12	130254.	10855.	0.33	0.979
Residual	36	1192168.	33116.		
Total	83	13260309.			

Message: the following units have large residuals.

Block 2 Hybrid V6	205.	s.e. 77.
Block 2 Hybrid V5 Herbicide Check	281.	s.e. 119.
Block 2 Hybrid V5 Herbicide H1	-275.	s.e. 119.
Block 3 Hybrid V3 Herbicide H1	-277.	s.e. 119.
Block 3 Hybrid V5 Herbicide H1	355.	s.e. 119.

Tables of means

Grand mean 1231.

Hybrid	V1	V2	V3	V4	V5	V6	V7
	687.	1139.	1584.	1442.	1642.	1381.	739.

Herbicide	Check	H1	H2
	1280.	1219.	1193.

Hybrid	Herbicide	Check	H1	H2
V1		695.	686.	678.
V2		1171.	1125.	1120.
V3		1714.	1484.	1554.
V4		1508.	1476.	1343.
V5		1737.	1616.	1573.
V6		1390.	1392.	1362.
V7		744.	751.	722.

Standard errors of differences of means

Table	Hybrid	Herbicide	Hybrid Herbicide
rep.	12	28	4
s.e.d.	67.7	57.2	128.6
d.f.	18	6	54.21
Except when comparing means with the same level(s) of			
Hybrid			132.1
d.f.			41.33
Herbicide			125.0
d.f.			53.62

Least significant differences of means (5% level)

Table	Hybrid	Herbicide	Hybrid Herbicide
rep.	12	28	4
l.s.d.	142.3	139.9	257.8
d.f.	18	6	54.21
Except when comparing means with the same level(s) of			
Hybrid			266.8
d.f.			41.33
Herbicide			250.7
d.f.			53.62

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Block	55096.4	3.000	708.9
Block.Hybrid	27536.0	18.000	-1859.9
Block.Herbicide	45788.3	6.000	1810.4
Block.Hybrid.Herbicide	33115.8	36.000	33115.8

There are strongly significant differences ($P < 0.001$) among hybrids, but no interaction or herbicide effect. The interpretation is therefore straightforward. In the presence of a significant interaction, individual means will have to be compared using one of three l.s.d. values, none of which leads to a strict t test (notice the non-integer degrees of freedom).

Notice also the negative Block.Hybrid stratum variance. When using LMM (REML) we would set that to be non-negative. The analysis is straightforward using the fixed and random models described above.

More complex field designs: a split-strip plot experiment

This experiment was used by Schabenberger and Pierce (2001), page 599, to illustrate a REML analysis in SAS. Four soybean cultivars were used as whole-plots in each of four replicate blocks. Two row spacings (9", 18") were used, each applied at random to half of each whole-plot in a vertical direction. In addition, five target plant populations (60, 120, …, 300 thousand per acre) were used, each applied at random to one-fifth of each whole-plot in a horizontal direction. The field plan therefore appears as follows.

Example 17 Soybean example, from Schabenberger and Pierce (2001), page 599

Block 1

AG4601		AG4701		AG3701		AG3601	
120	120	300	300	60	60	300	300
300	300	240	240	240	240	60	60
180	180	60	60	300	300	180	180
240	240	120	120	180	180	120	120
60	60	180	180	120	120	240	240
9	18	9	18	9	18	9	18

Block 2

AG4601		AG3701		AG3601		AG4701	
180	180	180	180	240	240	120	120
60	60	240	240	60	60	300	300
240	240	120	120	120	120	60	60
300	300	60	60	300	300	180	180
120	120	300	300	180	180	240	240
9	18	9	18	18	9	18	9

Block 3

AG3701		AG4701		AG3601		AG4601	
60	60	60	60	120	120	120	120
180	180	180	180	240	240	60	60
240	240	300	300	180	180	180	180
300	300	120	120	60	60	300	300
120	120	240	240	300	300	240	240
18	9	18	9	18	9	18	9

Block 4

AG3701		AG4601		AG3601		AG4701	
60	60	120	120	60	60	120	120
300	300	240	240	180	180	300	300
240	240	300	300	120	120	180	180
120	120	180	180	300	300	60	60
180	180	60	60	240	240	240	240
18	9	9	18	18	9	9	18

There are five strata in this experiment, and the block structure is the sum of these terms:

1. Block stratum
2. Block.Cultivar stratum
3. Block.Cultivar.Row stratum
4. Block.Cultivar.Plant stratum
5. Block.Cultivar.Row.Plant stratum

The yields for the corresponding treatments are as follows.

	Column							
Row	1	2	3	4	5	6	7	8
1	19.5	26.2	26.4	32.5	23.4	21.3	29.4	32.0
2	23.9	23.3	25.7	24.2	24.0	25.9	25.2	26.1
3	22.0	21.9	19.0	16.3	27.6	28.1	31.5	29.1
4	19.4	20.0	22.9	21.7	21.8	21.9	26.6	25.0
5	19.0	15.8	26.0	27.9	25.9	22.0		
6	23.4	22.4			26.0	32.9	21.9	23.9
7	20.6	19.7	26.9	25.9		27.9	31.4	26.5
8	28.2	27.9	25.6	24.8	32.1	34.2	24.5	21.4
9	25.9	28.5	23.0	23.3	26.5	40.2	28.9	30.5
10	22.0	30.3	28.8	30.4	25.1	35.9	28.0	23.3
11	17.8	22.3	16.5	19.3	22.0	28.9	23.6	21.6
12	20.9	23.3	23.3	26.6	27.9	36.9	17.2	20.8
13	26.5	26.2	28.0	30.4	27.0	32.1	24.9	24.6
14	25.9	24.2	24.2	30.1	23.2	26.9	33.0	35.3
15	22.8	19.0	22.0	26.9	26.9	34.5	30.7	25.3
16	16.2	13.0	20.4	23.6	21.4	17.6	25.2	21.1
17	26.5	25.4	21.0	24.4	23.3	26.9	26.7	26.1
18	27.5	21.9	23.2	26.2	16.0	23.2	25.5	23.5
19		17.9	24.4	21.7	21.3	27.1	14.7	15.6
20	19.8	22.2	15.6	17.7	26.2	32.4	26.0	26.4

There are six missing yields. GenStat will analyse the data via **General Analysis of Variance**. However, missing values are inserted and therefore F tests are inflated upwards. In addition, there may well be a change in variance across both row spacings and plant populations, and there may well be a better spatially correlated model to use, so it is preferable to use LMM (REML).

Treatment Structure: Cultivar*RowSpacing*PlantPop

Block Structure:

Block+Block.Cultivar+Block.Cultivar.Row+Block.Cultivar.Plant+Block.Cultivar.Row.Plant

Here is part of the ANOVA output.

Analysis of variance

Variate: Yield

Source of variation	d.f.	(m.v.)	s.s.	m.s.	v.r.	F pr.
Block stratum	3		419.420	139.807	7.24	

Block.Cultivar stratum						
Cultivar	3		531.821	177.274	9.18	0.004
Residual	9		173.723	19.303		
Block.Cultivar.PlantPop stratum						
PlantPop	4		1173.923	293.481	33.72	<.001
Cultivar.PlantPop	12		139.997	11.666	1.34	0.230
Residual	46	(2)	400.406	8.704	2.20	
Block.Cultivar.RowsSpacing stratum						
RowsSpacing	1		38.125	38.125	3.47	0.087
Cultivar.RowsSpacing	3		185.301	61.767	5.63	0.012
Residual	12		131.682	10.974	2.77	
Block.Cultivar.PlantPop.RowsSpacing stratum						
PlantPop.RowsSpacing	4		18.891	4.723	1.19	0.327
Cultivar.PlantPop.RowsSpacing	12		122.997	10.250	2.59	0.011
Residual	44	(4)	174.146	3.958		
Total	153	(6)	3388.902			

Message: the following units have large residuals.

Block 2 Cultivar AG3601 PlantPop 120.	4.51	s.e. 1.58
Block 2 Cultivar AG3601 PlantPop 240.	-3.89	s.e. 1.58
Block 3 Cultivar AG4601 PlantPop 300.	4.41	s.e. 1.58
Block 1 Cultivar AG3601 RowsSpacing 9.	-2.02	s.e. 0.91
Block 1 Cultivar AG3601 RowsSpacing 18.	2.02	s.e. 0.91
Block 2 Cultivar AG3601 PlantPop 120. RowsSpacing 9.	-2.60	s.e. 1.04
Block 2 Cultivar AG3601 PlantPop 120. RowsSpacing 18.	2.60	s.e. 1.04

May well be due to a changing variance in the field. ANOVA assumes constant variance

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Block	139.807	3.000	3.013
Block.Cultivar	19.303	9.000	0.358
Block.Cultivar.PlantPop	8.704	46.000	2.373
Block.Cultivar.RowsSpacing	10.974	12.000	1.403
Block.Cultivar.PlantPop.RowsSpacing	3.958	44.000	3.958

LMM (REML) analysis

There are four blocks, three fixed factors (4 cultivars × 2 row spacings × 4 target plant populations) in a five stratum layout. To obtain a better analysis than ANOVA, we use LMM (REML) with the following models:

Fixed Model: Cultivar*PlantPop*RowSpace
Random Model:
Block+Block.Cultivar+Block.Cultivar.Row+Block.Cultivar.Plant+Block.Cultivar.Row.Plant

REML variance components analysis

Response variate: Yield
Fixed model: Constant + Cultivar + PlantPop + RowsSpacing + Cultivar.PlantPop + Cultivar.RowsSpacing + PlantPop.RowsSpacing + Cultivar.PlantPop.RowsSpacing
Random model: Block + Block.Cultivar + Block.Cultivar.PlantPop + Block.Cultivar.RowsSpacing + Block.Cultivar.PlantPop.RowsSpacing
Number of units: 154 (6 units excluded due to zero weights or missing values)

Block.Cultivar.PlantPop.RowsSpacing used as residual term

Estimated variance components

Random term	component	s.e.
Block	3.037	2.896
Block.Cultivar	0.452	1.054
Block.Cultivar.PlantPop	2.421	1.017
Block.Cultivar.RowsSpacing	1.245	0.868

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
Block.Cultivar.PlantPop.RowsSpacing		Identity	Sigma2	3.927	0.835

Approximate stratum variances

Stratum	variance	effective d.f.
Block	133.551	3.00
Block.Cultivar	18.841	8.99
Block.Cultivar.PlantPop	8.688	45.83
Block.Cultivar.RowsSpacing	9.834	11.90
Block.Cultivar.PlantPop.RowsSpacing	3.927	44.29

Deviance: -2*Log-Likelihood

Deviance	d.f.
395.74	109

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Cultivar	24.38	3	8.13	9.0	0.006
RowSpacing	3.17	1	3.17	11.5	0.102
PlantPop	131.80	4	32.95	46.2	<0.001
Cultivar.RowSpacing	19.08	3	6.36	11.5	0.009
Cultivar.PlantPop	14.47	12	1.21	46.3	0.308
RowSpacing.PlantPop	4.19	4	1.05	45.3	0.393
Cultivar.RowSpacing.PlantPop	31.20	12	2.60	45.5	0.010

The similarities are clear, with the differences between the two analyses (apart from *P* values) due to the fact that REML uses just the data and ignores missing values.

However, we should investigate whether the variance changes with changing row spacing and changing plant population. Unfortunately, GenStat's analysis failed to converge when we tried this. To make headway, we tried the following.

The Block.Cultivar variance component is very small (0.452) and in fact can be deleted (the change in deviance is 395.96 - 395.74 = 0.22 with 1 d.f.). This is a simpler analysis which, apart from round-off error due to iteration with many parameters, produces the same variance components and close *P* values, with the exception that the individual Block and Block.Cultivar variance components of the first analysis (3.037 and 0.452) are replaced by a combined variance component of 3.465. This analysis is equivalent to treating the $b \times c$ plots (b blocks $\times$ c cultivars) as strips in the field into which the other factors are randomised (in two different ways). The analysis with changing variances for these factors did converge.

Random Model: Strip+ Strip.PlantPop+ Strip*RowSpace+ Strip.PlantPop.RowSpace, or simply Strip/(PlantPop*RowSpace)

Correlated Error Terms: use Identity⊗Diagonal⊗Diagonal for Strip.PlantPop.RowSpace

It turns out that that this more complex model is unnecessary, with a change in deviance of 401.01 – 397.48 = 3.53 with 5 d.f. (3.53 would be not significant if there was just 1 d.f.). Statistically, the first LMM (REML) analysis is the one to use for decisions; biologically, the plants within plots are competing to the point that a common variance model appears adequate.

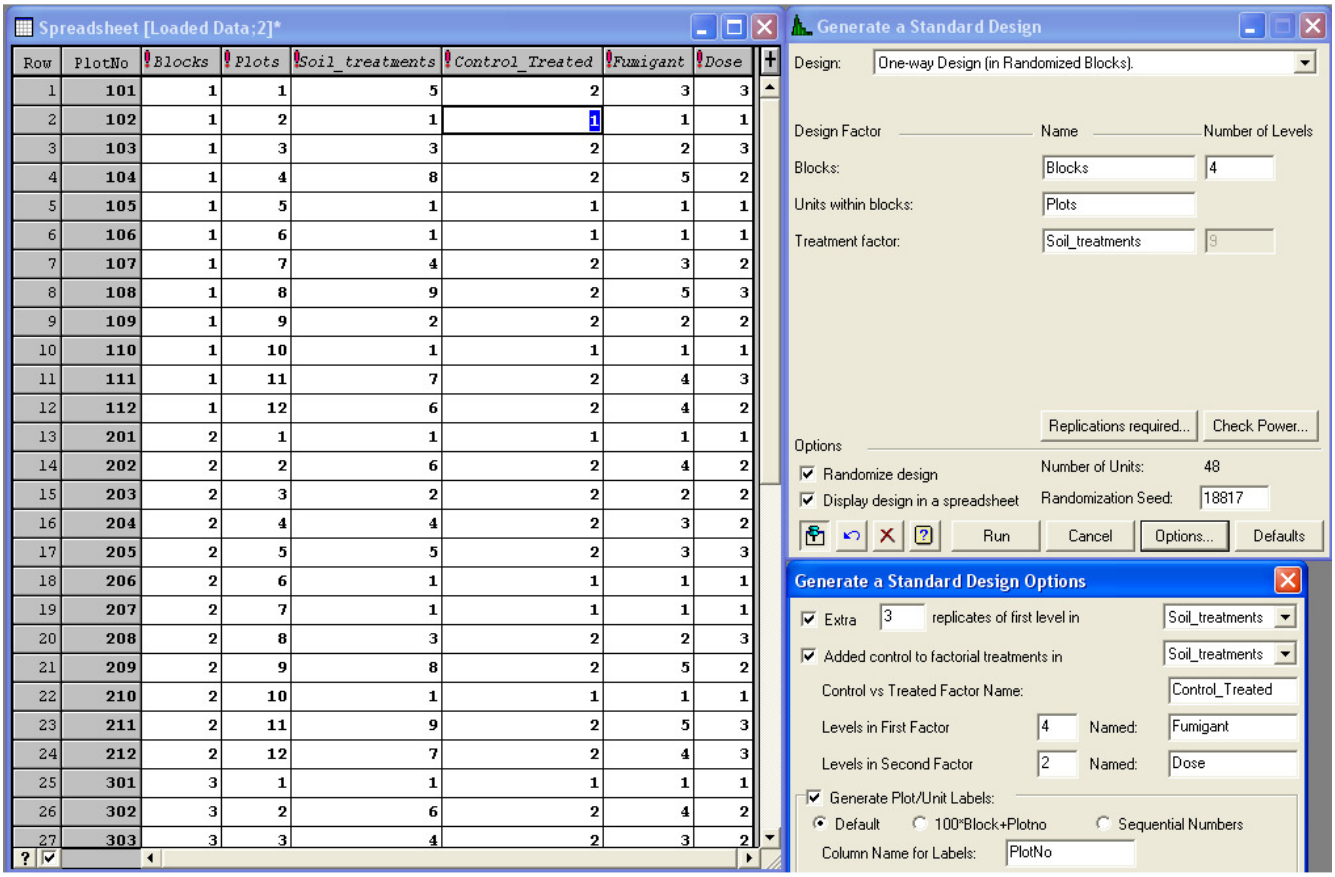
The only point to add is that the design is unbalanced (with 6 missing values) and hence the *P* values depend on the order the factors are added to the model. As usual with unbalanced data, the *P* value to use for a factor should be the one obtained from an analysis with that factor entered last.

Spatial model: two-way design (in randomized blocks) *plus* a control *plus* extra replication of the control *plus* a covariate

An experiment was laid out in four randomized blocks, designed to determine the effectiveness of four soil fumigants in keeping down the numbers of eelworms in the soil. The fumigants were chlorodinitrobenzene (CN), carbon disulphide jelly (CS) and two proprietary preparations, "Cymag" (CM) and "Seekay" (CK). Each fumigant was tested both in a single and double dose. There was a 9th treatment, viz a control (no fumigant): four plots in each block were left untreated. The purpose was to supply an accurate standard against which the performance of the fumigants was measured. The fumigants were ploughed in during spring, after which a crop of oats was sown. Before and after harvest, 400g of soil was taken from each plot and the number of eelworm cysts counted.

Generating a random design in GenStat prior to running the experiment

Although there is a 4 × 2 factorial structure (Fumigant × Dose), once the control treatment is added the treatment structure is a bit more complex. Since the control is "no fumigant", there is no way of having a single and double dose of "nothing". So initially, we need to think of this as a one-way treatment design with (4×2+1) levels. We have 9 treatments, 8 of which are factorially structured. So in the **Design** menu we select **One-way (in Randomized Blocks)**, set the number of treatments to 9, then go into **Options**. We set up a 1 *df* contrast for the treated versus untreated plots, and set up the 4 × 2 factorial structure in that menu. In addition, we can get GenStat to replicate the Control treatment 4 times (an additional 3 replicates per block):



Row	PlotNo	Blocks	Plots	Soil_treatments	Control_Treated	Fumigant	Dose
1	101	1	1	5	2	3	3
2	102	1	2	1	1	1	1
3	103	1	3	3	2	2	3
4	104	1	4	8	2	5	2
5	105	1	5	1	1	1	1
6	106	1	6	1	1	1	1
7	107	1	7	4	2	3	2
8	108	1	8	9	2	5	3
9	109	1	9	2	2	2	2
10	110	1	10	1	1	1	1
11	111	1	11	7	2	4	3
12	112	1	12	6	2	4	2
13	201	2	1	1	1	1	1
14	202	2	2	6	2	4	2
15	203	2	3	2	2	2	2
16	204	2	4	4	2	3	2
17	205	2	5	5	2	3	3
18	206	2	6	1	1	1	1
19	207	2	7	1	1	1	1
20	208	2	8	3	2	2	3
21	209	2	9	8	2	5	2
22	210	2	10	1	1	1	1
23	211	2	11	9	2	5	3
24	212	2	12	7	2	4	3
25	301	3	1	1	1	1	1
26	302	3	2	6	2	4	2
27	303	3	3	4	2	3	2

Notice that GenStat creates a factor (with 1s and 2s) to compare treated and untreated plots: a 1 represents an untreated plot (throughout the spreadsheet) and 2 a treated plot. Then, in the **Output** window, the **Treatment Structure** is shown as Control_Treated/(Fumigant*Dose). Remember that the / operator has a higher priority than the * operator, so the parentheses are important in this structure, to force the / operator on all three terms in the factorial structure. This might be clearer with the following explanation.

If you examine the other factor levels in the spreadsheet you will see that the combination of fumigant number (2, 3, 4, 5) and dose number (2 = single, say, and 3 = double) occurs only when the Control_Treated level is 2 (ie treated). Fumigant and dose treatments are "nested" inside the treated versus control contrast. The effect is that, in the ANOVA, apparent first-order interactions (like Control_Treated.Fumigant) are actually main effects and the apparent second-order interaction (Control_Treated.Fumigant.Dose) is first-order interaction

Analysis of variance

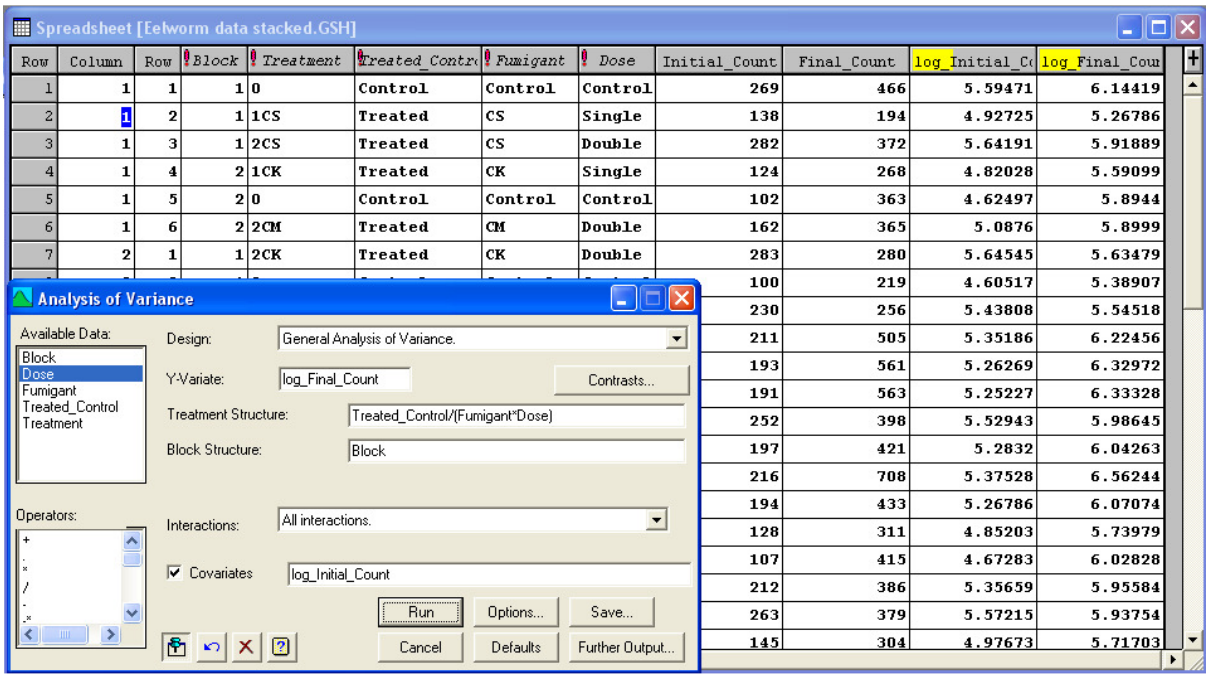
Source of variation	d.f.	**think of this component as:**
Blocks stratum	3	Blocks
Blocks.Plots stratum		
Control_Treated	1	**Control_Treated** contrast
Control_Treated.Fumigant	3	**Fumigant** main effect (for treated plots)
Control_Treated.Dose	1	**Dose** main effect (for treated plots)
Control_Treated.Fumigant.Dose	3	**Fumigant.Dose** interaction (for treated plots)
Residual	36	

Example 18 Dose (1 = single, 2 = double) and type of fumigant, and eelworm counts (initial above final) in field position, from Cochran and Cox page 46

0	2CK	1CN	1CM	2CM	2CS	2CK	0
269	283	252	212	95	127	80	134
466	280	398	386	199	166	142	590
1CS	0	0	2CM	1CK	1CN	1CM	0
138	100	197	263	107	89	41	74
194	219	421	379	236	332	176	137
2CS	1CK	0	2CN	0	0	2CN	1CS
282	230	216	145	88	25	42	62
372	256	708	304	356	212	308	221
1CK	0	1CS	2CK	2CK	0	1CK	1CM
124	211	194	222	193	209	109	153
268	505	433	408	292	352	132	454
0	2CN	2CS	1CN	0	2CN	2CS	0
102	193	128	42	29	9	17	19
363	561	311	222	254	92	28	106
2CM	0	1CM	0	1CS	1CN	0	2CM
162	191	107	67	23	19	44	48
365	563	415	338	80	114	268	298

Had we used GenStat to design the trial, we need only add the two data columns (final and initial counts) and **Run** the analysis via the **Spread** menu.

The analysis is performed in GenStat by initially setting up two factor columns: a Block factor with 4 levels and a soil Treatment factor with 9 levels. Then in Options, we set up a factor to identify treated and untreated plots, and two treatment factor columns, Dose (Single, Double) and Fumigant (CK=Seekay, CM=Cymag, CN=chlorodinitrobenzene, CS= carbon disulphide jelly). We have the added complication that the control is replicated 4 times in each block.



Row	Column	Row	Block	Treatment	Treated_Contr	Fumigant	Dose	Initial_Count	Final_Count	log_Initial_C	log_Final_Cou
1	1	1	1	0	Control	Control	Control	269	466	5.59471	6.14419
2	1	2	1	1CS	Treated	CS	Single	138	194	4.92725	5.26786
3	1	3	1	2CS	Treated	CS	Double	282	372	5.64191	5.91889
4	1	4	2	1CK	Treated	CK	Single	124	268	4.82028	5.59099
5	1	5	2	0	Control	Control	Control	102	363	4.62497	5.8944
6	1	6	2	2CM	Treated	CM	Double	162	365	5.0876	5.8999
7	2	1	1	2CK	Treated	CK	Double	283	280	5.64545	5.63479
								100	219	4.60517	5.38907
								230	256	5.43808	5.54518
								211	505	5.35186	6.22456
								193	561	5.26269	6.32972
								191	563	5.25227	6.33328
								252	398	5.52943	5.98645
								197	421	5.2832	6.04263
								216	708	5.37528	6.56244
								194	433	5.26786	6.07074
								128	311	4.85203	5.73979
								107	415	4.67283	6.02828
								212	386	5.35659	5.95584
								263	379	5.57215	5.93754
								145	304	4.97673	5.71703

There are some issues to sort out with data like these.

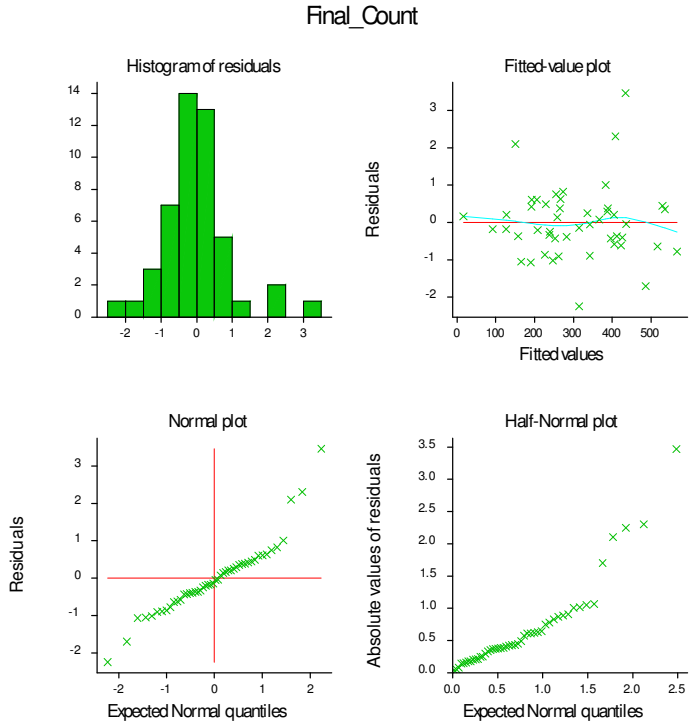
- The data are not normally distributed. It is possible that they are Poisson, in which case the variance is the same as the mean, and if the means change then so must the variances. Hence a logistic regression might be preferable to ANOVA. Alternatively, we could transform the data to achieve approximate constant variance. For Poisson data the square root transformation used to be recommended. With large counts, a log transformation may be better: differences in means are then more easily back-transformed and interpreted.

- The final counts may well depend on the initial worm counts: if the worms are not uniformly spread at the start of the experiment, then differences at the end may be misleading. We should incorporate initial counts as a covariate. If we log-transform final counts, then we should log-transform initial counts as well.

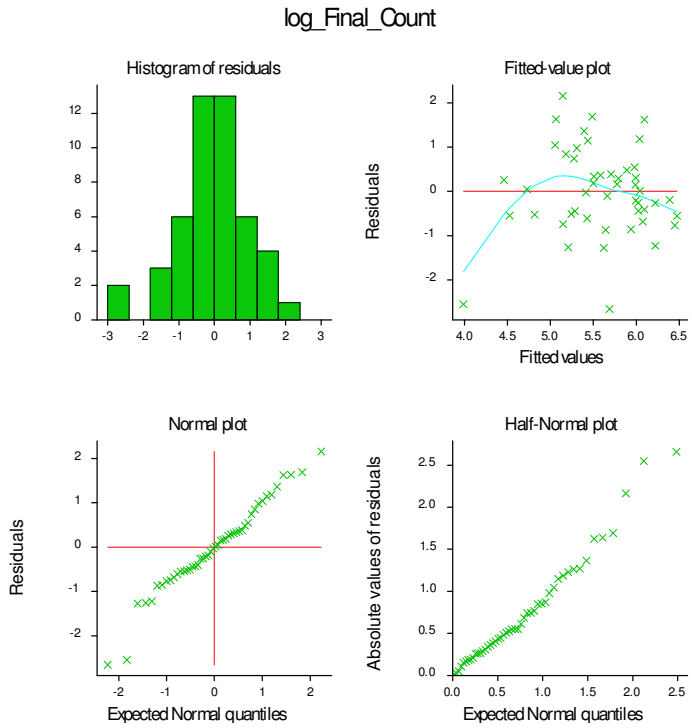
- The Poisson distribution tends to a normal distribution with increasing mean count. Thus, we could use LMM (REML) assuming an approximate normal distribution with a changing variance, and possibly a spatially correlated error structure. Notice that the four blocks are formed as a 2 × 2 layout in the field, and in each block the plots are arranged in a 3×4 grid. If there is a gradient left to right and top to bottom across blocks, we might expect a gradient left to right and/or top to bottom *within* the blocks. What has become

known as a Row-Column analysis might then remove a trend in the field more successfully than the 2 × 2 block layout.

We will look at some of these actions. Firstly, an analysis of final counts with initial counts as a covariate shows a distinct fanning in the standardised residuals:



We therefore analyse the data log-transformed:



Analysis of log(final counts), with log(initial counts) as a covariate

Analysis of variance (adjusted for covariate)

Variate: log_Final_Count
Covariate: log_Initial_Count

Source of variation	d.f.	s.s.	m.s.	v.r.	cov.ef.	F pr.
Block stratum						
Covariate	1	4.76145	4.76145	11.74		0.076
Residual	2	0.81127	0.40563	4.23	4.58	
Block.*Units* stratum						
Treated_Control	1	1.16420	1.16420	12.13	1.00	0.001
Treated_Control.Fumigant	3	2.08349	0.69450	7.24	0.92	<.001
Treated_Control.Dose	1	0.04506	0.04506	0.47	0.99	0.498
Treated_Control.Fumigant.Dose	3	0.31977	0.10659	1.11	1.00	0.358
Covariate	1	5.21084	5.21084	54.31		<.001
Residual	35	3.35793	0.09594		2.48	
Total	47	16.92526				

Message: the following units have large residuals.

Block 3 *units* 11	-0.770	approx. s.e.	0.264
Block 4 *units* 8	-0.654	approx. s.e.	0.264

Tables of means (adjusted for covariate)

Variate: log_Final_Count
Covariate: log_Initial_Count

Grand mean 5.582

Treated_Control	Control	Treated
	5.805	5.470
rep.	16	32

Treated_Control	Dose	Control	Double	Single
Control		5.805		
Treated			5.432	5.508

Treated_Control	Fumigant	Control	CK	CM	CN	CS
Control		5.805				
	rep.	16				
Treated			5.195	5.667	5.798	5.220
	rep.		8	8	8	8

Treated_Control	Dose	Fumigant	Control	CK	CM	CN	CS
Control	Control		5.805				
		rep.	16				
Treated	Double			5.216	5.589	5.882	5.041
		rep.		4	4	4	4
	Single			5.174	5.745	5.713	5.399
		rep.		4	4	4	4

Standard errors of differences of means

Table	Treated_Control	Treated_Control Dose	Treated_Control Fumigant	Treated_Control Dose Fumigant	
rep.	unequal	16	unequal	unequal	
d.f.	35	35	35	35	
s.e.d.			0.1596	0.2226	min.rep
	0.0949	0.1097	0.1382	0.1760	max-min
			0.1129X	0.1113X	max.rep

(No comparisons in categories where s.e.d. marked with an X)

Least significant differences of means (5% level)

Table	Treated_Control	Treated_Control Dose	Treated_Control Fumigant	Treated_Control Dose Fumigant	
rep.	unequal	16	unequal	unequal	
d.f.	35	35	35	35	
l.s.d.			0.3241	0.4520	min.rep
	0.1927	0.2227	0.2806	0.3573	max-min
			0.2291X	0.2260X	max.rep

(No comparisons in categories where l.s.d. marked with an X)

Estimated stratum variances (adjusted for covariate)

Variate: log_Final_Count
Covariate: log_Initial_Count

Stratum	variance	effective d.f.	variance component
Block	0.3029	2.746	0.0173
Block.*Units*	0.0953	35.254	0.0953

Clearly initial counts go a long way to explaining differences in final counts. Incorporating the initial counts as a covariate:

- is strongly significant ($P<0.001$);
- reduces the Residual MS from 0.2380 to less than *half* that value, 0.0959;
- more accurately tests whether treated plots have significantly lower eelworm cysts than control plots, taking initial counts into account ($P=0.001$);
- detects that the type of fumigant is very important ($P<0.001$).

A very important feature of interpreting means of log-transformed data should be mentioned.

- The back-transformed mean of log-transformed data is the *geometric mean* of the original data. For log-normal data, the geometric mean is a much better estimate of a "typical" value than the arithmetic mean, since the importance of very large values in the calculation is greatly reduced.

- The back-transformed difference in two means of log-transformed data is the *ratio of the two geometric means* of the original data. For example, for the carbon disulphide jelly (CS) fumigant, the effect of a single compared to a double dose is 5.399 – 5.041 = 0.358 on the log-scale. This back-transforms to 1.43. Thus, a plot with a single dose of carbon disulphide jelly applied typically has 43% more eelworms cysts than a similar plot with a double dose.

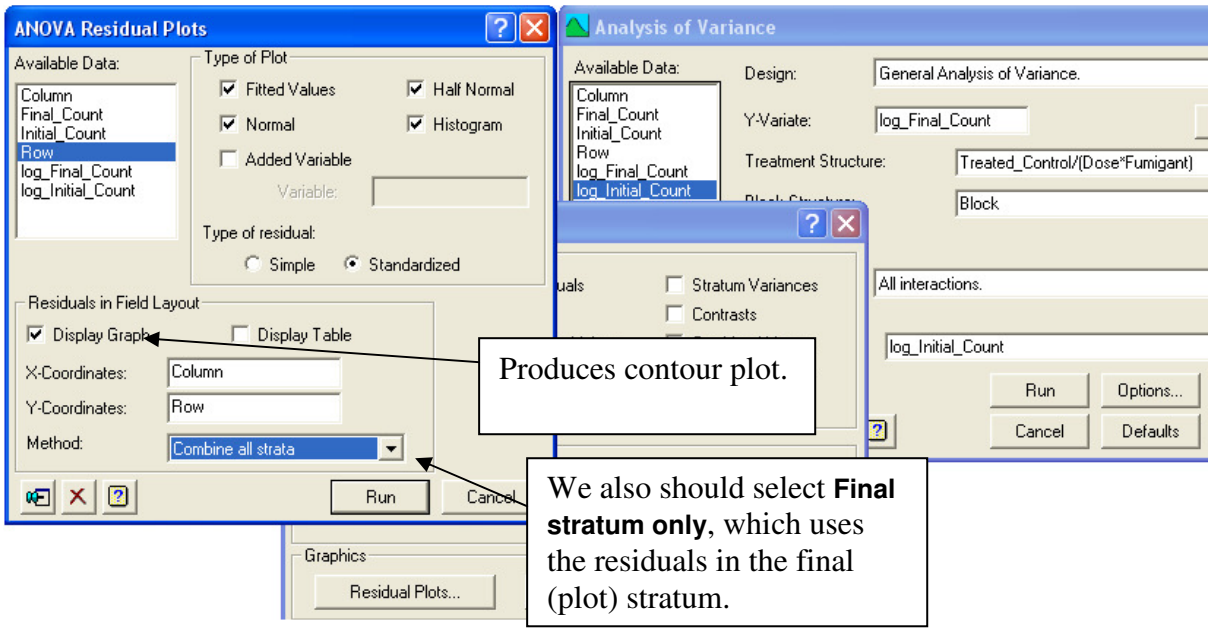
- The l.s.d. value for the comparison above is 0.4520 and this is based on 35 *df* for which *tcrit* is 2.030. The value to add and subtract to the difference in means above is 2.030×0.4520 = 0.918. The 95% confidence interval on the log-scale is (-0.560, 1.276). Back-transforming the end points gives a confidence interval for the ratio of (0.571, 3.581). Thus, while a plot with a single dose of carbon disulphide jelly applied typically has 43% more eelworms cysts than a similar plot with a double dose, we are only 95% confident that this ratio is between just over a half (0.571×), to a little more than three and a half times (3.581×). Other differences are treated similarly.

Residuals plotted in field position

There is still one other plot to check: a plot of the residuals *in field position*, with an accompanying contour plot. To obtain this plot, we need to supply *two variates*: the X-coordinate and the Y-coordinate of each plot in field position. Imagine an X-Y coordinate system overlaying the experimental site (consisting of plots in a 6×8 layout) with the origin in the bottom left hand corner of the site.

	0	2CK	1CN	1CM	2CM	2CS	2CK	0
Y=6	269	283	252	212	95	127	80	134
	466	280	398	386	199	166	142	590
	1CS	0	0	2CM	1CK	1CN	1CM	0
Y=5	138	100	197	263	107	89	41	74
	194	219	421	379	236	332	176	137
	2CS	1CK	0	2CN	0	0	2CN	1CS
Y=4	282	230	216	145	88	25	42	62
	372	256	708	304	356	212	308	221
	1CK	0	1CS	2CK	2CK	0	1CK	1CM
Y=3	124	211	194	222	193	209	109	153
	268	505	433	408	292	352	132	454
	0	2CN	2CS	1CN	0	2CN	2CS	0
Y=2	102	193	128	42	29	9	17	19
	363	561	311	222	254	92	28	106
	2CM	0	1CM	0	1CS	1CN	0	2CM
Y=1	162	191	107	67	23	19	44	48
	365	563	415	338	80	114	268	298
(0,0)	X=1	X=2	X=3	X=4	X=5	X=6	X=7	X=8

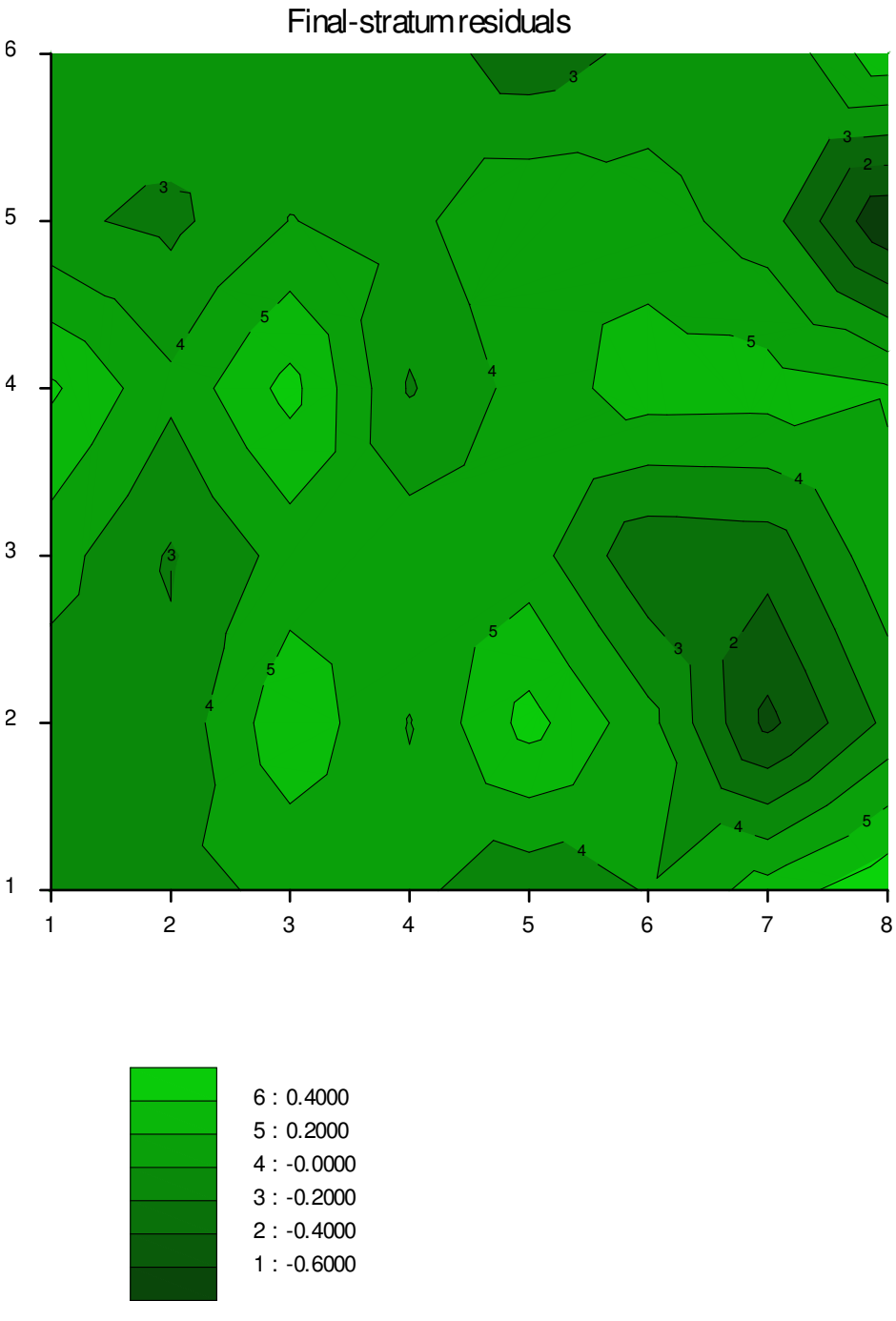
The data are ordered down column 1 first, so we need to set up Y as (6, 5, 4, 3, 2, 1, 6, 5, …) and X as (1, 1, 1, 1, 1, 1), …, (8, 8, 8, 8, 8, 8) by right-clicking on each column and selecting **Fill** (with the **Starting Value** for Y being 6, the **Ending Value** 1 and **Increment** -1).



.
The residuals in field position are:

	Final_stratum_residuals							
_['Column']	1	2	3	4	5	6	7	8
_['Row']								
6	**-0.077**	**-0.027**	**-0.104**	**-0.066**	**-0.326**	**-0.130**	**-0.191**	0.343
5	**-0.157**	**-0.253**	0.004	**-0.054**	0.190	0.099	**-0.114**	**-0.770**
4	0.434	0.047	0.470	**-0.219**	0.084	0.302	0.295	0.218
3	0.087	**-0.222**	0.079	0.124	0.093	**-0.356**	**-0.324**	0.141
2	**-0.127**	**-0.142**	0.350	**-0.007**	0.474	0.066	**-0.654**	**-0.152**
1	**-0.175**	**-0.055**	0.039	0.048	**-0.140**	0.012	0.284	0.555

These residuals should be random +/- across the field, since block effects are supposed to have dealt with any gradient in the field. Within each block the residuals will add to 0. Given that, deciding if the residuals are random in the field is fairly subjective. The accompanying contour plot smoothes over the individual residuals, but again, deciding if the light areas represent plots whose fitted counts are consistently larger than the observed counts is again subjective.


Final-stratum residuals
6
5
4
3
2
1
1
2
3
4
5
6
7
8
6 : 0.4000
5 : 0.2000
4 : -0.0000
3 : -0.2000
2 : -0.4000
1 : -0.6000

LMM (REML) analysis of the spatial data

Firstly, we reproduce the analysis of the eelworm Log(Final_Count) data. Recall that the ANOVA **Treatment Structure** is Treated_Control/(Dose*Fumigant) and in a separate box a covariate was defined. In LMM (REML), we move the covariate into the **Fixed Model**, which becomes Log_Initial_Count+Treated_Control/(Dose*Fumigant).

The **Random Model** is Block+Block.Plot, or simply Block. Neither formulation allows us to use a correlation structure spatially. We will discuss this issue after the basic REML analysis is completed:

REML variance components analysis

Response variate: log_Final_Count
Fixed model: Constant + log_Initial_Count + Treated_Control + Treated_Control.Fumigant + Treated_Control.Dose + Treated_Control.Fumigant.Dose
Random model: Block
Number of units: 48

All covariates centred

Estimated variance components

Random term	component	s.e.
Block	0.01730	0.02169

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
Residual		Identity	Sigma2	0.0953	0.02271

Deviance: -2*Log-Likelihood

Deviance	d.f.
-30.98	36

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Log_Initial_Count	61.54	1	61.54	25.4	<0.001
Treated_Control	12.25	1	12.25	35.2	0.001
Treated_Control.Dose	0.38	1	0.38	35.3	0.541
Treated_Control.Fumigant	22.01	3	7.33	36.0	<0.001
Treated_Control.Dose.Fumigant	3.33	3	1.11	35.3	0.358

Dropping individual terms from full fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Log_Initial_Count	70.30	1	70.30	25.4	<0.001
Treated_Control.Dose.Fumigant	3.33	3	1.11	35.3	0.358

REML estimates of the block and error variances are the same as the stratum variances. Once a covariate is added, the main effects depend on the order the factors are entered into the model (just as they would in the ANOVA). To illustrate this, we have removed the two factor interaction from the fixed model. The change to the last part of the analysis is:

Dropping individual terms from full fixed model

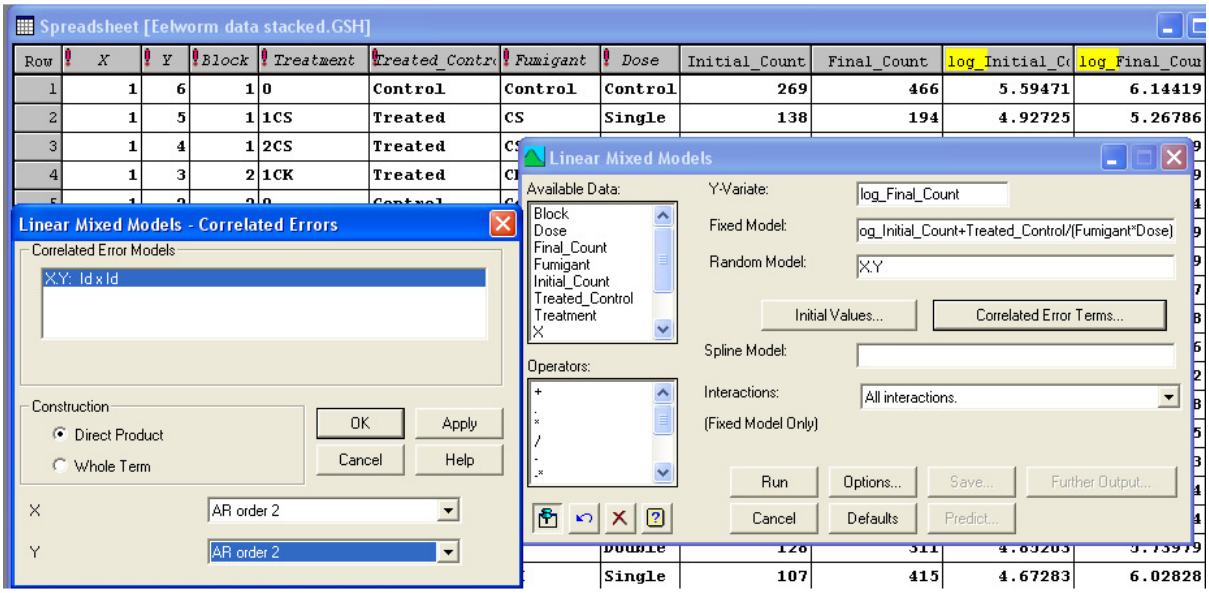
Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Log_Initial_Count	69.19	1	69.19	26.2	<0.001
Treated_Control.Dose	0.46	1	0.46	38.3	0.503
Treated_Control.Fumigant	21.81	3	7.27	39.1	<0.001

How do we incorporate a spatial correlation for this experiment?

Firstly, the field really consists of plots in a row by column layout. The original layout had four blocks in a 2×2 layout with each block consisting of 12 plots in a 3×4 layout. As hypothesized earlier, if there *is* a block effect, is it left to right across the field, or top to bottom, or both? If any of these, why is the gradient no reflected in the plots within a block?

To investigate these possibilities, we inserted a factor labelled Y with 6 levels, and a factor labelled X with 8 levels. The Y factor is filled from 6 down to 1 in order for the field layout to mimic the X-Y coordinate system with the original in the bottom left hand corner of the field.

The **Random Model** is then X.Y with at most an AR2 ⊗ AR2 spatially correlated model. We do not expect exactly the same scaled Wald statistics as before, since the assumed error structure is now different.



We can use change in deviance to check whether we a less complex model is adequate.

Model for X.Y	deviance	df	Change in deviance	Change in df	P-value
AR2.AR2	-34.32	33			
AR2.AR1	-34.17	34	0.15	1	0.699
AR2.Identity	-34.12	35	0.05	1	0.823
AR1.Identity	-33.76	36	0.36	1	0.549
Identity.Identity	-28.50	37	5.26	1	0.022

It would appear that an AR1 correlated model left to right is what is required in this case. The analysis is as follows.

REML variance components analysis

Response variate: log_Final_Count
Fixed model: Constant + log_Initial_Count + Treated_Control + Treated_Control.Fumigant + Treated_Control.Dose + Treated_Control.Fumigant.Dose
Random model: X.Y
Number of units: 48

X.Y used as residual term with covariance structure as below

Covariance structures defined for random model

Covariance structures defined within terms:

Term	Factor	Model	Order	No. rows
X.Y	X	Auto-regressive (+ scalar)	1	8
	Y	Identity	0	6

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
X.Y	Sigma2	0.113	0.0306		
	X	AR(1)	phi_1	0.4127	0.1753
	Y	Identity	-	-	-

Deviance: -2*Log-Likelihood

Deviance	d.f.
-33.76	36

Tests for fixed effects

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Log_Initial_Count	61.99	1	61.99	19.4	<0.001
Treated_Control	17.34	1	17.34	25.3	<0.001
Treated_Control.Fumigant	27.00	3	8.99	26.7	<0.001
Treated_Control.Dose	0.92	1	0.92	32.7	0.344
Treated_Control.Fumigant.Dose	3.29	3	1.10	29.4	0.367

Dropping individual terms from full fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Log_Initial_Count	64.38	1	64.38	19.4	<0.001
Treated_Control.Fumigant.Dose	3.29	3	1.10	29.4	0.367

Dropping the interaction between fumigants and dose:

Dropping individual terms from full fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Log_Initial_Count	63.95	1	63.95	20.7	<0.001
Treated_Control.Fumigant	27.27	3	9.08	30.5	<0.001
Treated_Control.Dose	0.90	1	0.90	36.7	0.349

Means, all s.e.d. and l.s.d. values are suppressed: they can be saved into an Excel file.

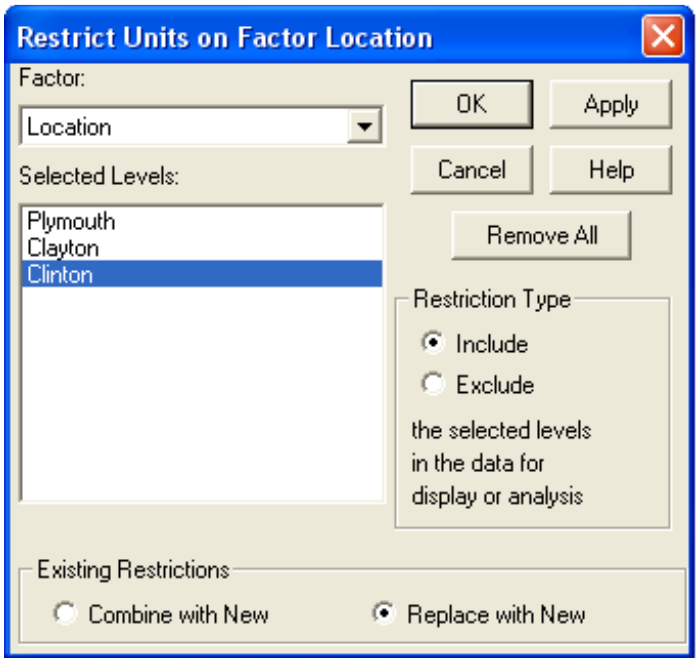
We could check whether an additional experimental error is necessary by adding '*Units*' to the residual. In this case, the change is deviance is negligible (0.15 on 1 df).

Multi-site experiments

Example 19 Twelve strains of soybeans were compared in separate randomized blocks at three locations in North Carolina. Data from Steel and Torrie page 399, 400

	Plymouth			Clayton			Clinton		
Variety	BL 1	BL 2	BL 3	BL 1	BL 2	BL 3	BL 1	BL 2	BL 3
Tracy	1307	1365	1542	1178	1089	960	1583	1841	1464
Centennial	1425	1475	1276	1187	1180	1235	1713	1684	1378
N72-137	1289	1671	1420	1451	1177	1723	1369	1608	1647
N72-3058	1250	1202	1407	1318	1012	990	1547	1647	1603
N72-3148	1546	1489	1724	1345	1335	1303	1622	1801	1929
R73-81	1344	1197	1319	1175	1064	1158	1800	1787	1520
D74-7741	1280	1260	1605	1111	1111	1099	1820	1521	1851
N73-693	1583	1503	1303	1388	1214	1222	1464	1607	1642
N73-877	1656	1371	1107	1254	1249	1135	1775	1513	1570
N73-882	1398	1497	1583	1179	1247	1096	1673	1507	1390
N73-1102	1586	1423	1524	1345	1265	1178	1894	1547	1751
R75-12	911	1202	1012	1136	1161	1004	1422	1393	1342

The first thing to decide is whether the variation at each site is consistent. Three separate RCBD analyses produced the following Residual MS estimates. These are obtained by clicking in the spreadsheet, selecting **Restrict/Filter > To Groups (factor levels)**. Select the **Location** factor and each level with **Replace with new**.



Location	df	Residual MS
Plymouth	22	24149
Clayton	22	12124
Clinton	22	22851
Average	**66**	**19708**

Do we have any right to combine the three estimates into a pooled estimate with 66 *df*? Since we assume normal data and independent experiments across locations, these can be tested by Bartlett's variance homogeneity test, (Chi-square 2.90 on 2 degrees of freedom: probability 0.234).

Next, locations are really included to make better breeding choices, so interest lies in interpreting the Strain.Location interaction. Technically, locations are *fixed* sites of interest and each site is unreplicated (as are blocks at each location). Hence, to place Location in a top-level stratum of its own (with no P value for Location) we place in the **Block Structure** rather

than in the **Treatment Structure**, simply as a device. (In the LMM (REML) section for this example we assume Strain and Strain.Location are both random factors.)

Next, block 1 at one location is not the same as block 1 at a different location. Hence we need to combine *blocks within locations*, thereby obtaining (3-1)×3 = 6 *df*.

The **Block Structure** we would recommend is then Location+Block.Location+Block.Location.Strain (GenStat allows the final stratum to be omitted)

This is the analysis that such a general ANOVA produces:

Analysis of variance

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Location stratum	2	3113626.	1556813.	134.87	
Location.Block stratum	6	69256.	11543.	0.59	
Location.Block.*Units* stratum					
Strain	**11**	**925090.**	**84099.**	**4.27**	**<.001**
Location.Strain	22	532900.	24223.	1.23	0.256
Residual	66	1300723.	19708.		
Total	107	5941596.			

...

Tables of means

Variate: Yield

Grand mean 1403.

Strain	Centennial	D74-7741	N72-137	N72-3058	N72-3148	N73-1102
	1395.	1406.	1484.	1331.	1566.	1501.

Strain	N73-693	N73-877	N73-882	R73-81	R75-12	Tracy
	1436.	1403.	1397.	1374.	1176.	1370.

Location	Strain	Centennial	D74-7741	N72-137	N72-3058	N72-3148
Plymouth		1405.	1395.	1473.	1299.	1599.
Clayton		1402.	1308.	1652.	1308.	1529.
Clinton		1378.	1517.	1327.	1385.	1570.

Location	Strain	N73-1102	N73-693	N73-877	N73-882	R73-81
Plymouth		1524.	1476.	1391.	1506.	1300.
Clayton		1464.	1476.	1414.	1375.	1334.
Clinton		1517.	1357.	1405.	1309.	1488.

Location	Strain	R75-12	Tracy
Plymouth		1055.	1418.
Clayton		1302.	1277.
Clinton		1172.	1415.

Standard errors of differences of means

Table	Strain	Location Strain
rep.	9	3
d.f.	66	66
s.e.d.	66.2	114.6

Least significant differences of means (5% level)

Table	Strain	Location Strain
rep.	9	3
d.f.	66	66
l.s.d.	132.1	228.9

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Location	1556813.0	2.000	42924.2
Location.Block	11542.7	6.000	-680.4
Location.Block.*Units*	19707.9	66.000	19707.9

Notice that the Location.Block MS (11543) is unexpectedly smaller than the Residual MS (19708) which gives rise to the negative variance component above. When then data are analysed using LMM (REML), it is advisable to force a zero bound for this variance component.

The Location MS is much larger than the Residual MS, indicating large variation in the overall mean yields over the three locations. Differences in means between the strains, however, are consistent across these locations (P=0.256).

LMM (REML) analysis assuming fixed locations and random strains

The block and treatment structures used in the ANOVA were:

Treatment Structure: Strain+Location+Location.Strain
Block Structure: Location.Block

Placing Location in the **Block Structure** was purely a device to prevent the unreplicated factor Location from having a P-value printed in the ANOVA. The same analysis is produced when Location is placed in the **Treatment Structure**, but no stratum variance is obtained then (GenStat treats factors in the **Treatment Structure** as fixed terms).

Generally, when a factor is regarded as random then any interaction involving that factor is also random. With the Steel and Torrie data it is unclear whether the three locations, or the twelve strains, were randomly chosen or were of specific interest. It is common that Strains, and hence Strains.Location, are random, and that is what we will assume (with Location fixed). What often occurs, moreover, is that the residual variances differ across locations. This was tested on page 102 via Bartlett's test of homogeneity of variance (and found to be not significant). Here we test it by change in deviance.

Firstly, we test whether the residual variances at each location are the same:

Fixed Model: Location
Random Model: Strain + Location.Strain + Location.Block + Location.Strain.Block
(Location.Block constrained to be positive)

Model	Deviance	d.f.	χ^2 P-value
Identity for Location in Location.Strain.Block	1172.04	101	
Diagonal for Location in Location.Strain.Block	1170.32	99	
Change	1.72	2	0.423

So, the simpler model with a constant residual variance at each location suffices ($P = 0.423$). The estimated variances (below) at each location are slightly different to those used in Bartlett's test in the design section, because in this analysis we constrained the Location.Block term to be non-negative:

Location	Variance from individual ANOVAs	Variance from combined REML
Plymouth	24149	21778 ± 5896
Clayton	12124	13591 ± 3882
Clinton	22851	21217 ± 5818

The output from the *constant variance* model is as follows.

REML variance components analysis

Response variate: Yield
Fixed model: Constant + Location
Random model: Strain + Strain.Location + Location.Block + Strain.Location.Block
Number of units: 108

Strain.Location.Block used as residual term

Estimated variance components

Random term	component	s.e.
Strain	6653.	4066.
Strain.Location	1732.	2654.
Location.Block	0.	bound

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
	Strain.Location.Block	Identity	Sigma2	19027.	3171.

Deviance: -2*Log-Likelihood

Deviance	d.f.
1172.04	101

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Location	128.54	2	64.27	22.0	<0.001

Clearly there are yield differences across locations ($P<0.001$), but this is neither surprising nor of interest. As a breeding trial, we are more interested in strain differences. However, we need to determine firstly whether there are genotype × environment interactions.

To test whether the random Location.Strain interaction is significant is equivalent to testing whether the Location.Strain variance is 0. The estimate from the analysis above is 1732 ± 2564. However, we can only test this hypothesis using change in deviance, with the new model omitting the random term to be tested.

Model	Deviance	d.f.	P-value
Including Location.Strain	1172.04	101	
Excluding Location.Strain	1172.55	102	
Change	0.51	1	0.475

This result indicates that strain differences are consistent across locations ($P = 0.475$).

Are there any differences among the strains themselves? Since Strain is also a random effect, we can only decide this by change in deviance. We take the no interaction model and drop Strain:

Model	Deviance	d.f.	χ^2 P-value
Including Strain	1172.55	102	
Excluding Strain	1186.87	103	
Change	14.32	1	<0.001

Strain differences are strongly significant ($P<0.001$). The final analysis we use excludes the Location.Strain interaction but includes the Block.Location random effect to emphasise the combined nature of the analysis.

REML variance components analysis

Response variate: Yield
Fixed model: Constant + Location
Random model: Strain + Block.Location
Number of units: 108

Estimated variance components

Random term	component	s.e.
Strain	7095.	3998.
Block.Location	0.	bound

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
Residual		Identity	Sigma2	20243.	2953.

Deviance: -2*Log-Likelihood

Deviance	d.f.
1172.55	102

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
Location	153.81	2	76.90	94.0	<0.001

Multiple Experiments/Meta Experiments (REML) menu

A combined analysis of separate experiments can be obtained using the meta analysis menu in one step. Note that this menu assumes you want separate variances for each experimental site:

Separate residual terms for each level of experiment factor: Location

Sparse algorithm with AI optimisation

Estimated variance components

Random term	component	s.e.
Strain	6272.	3941.
Strain.Location	2166.	2653.
Location.Block	0.	bound

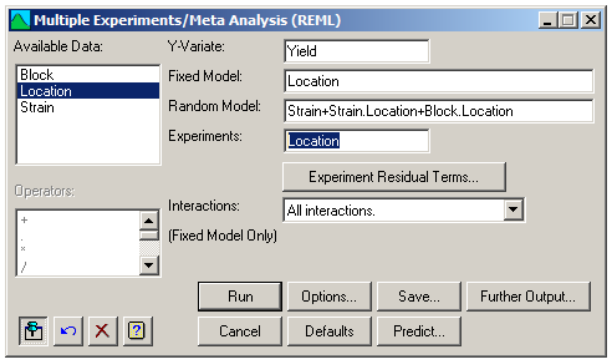
Residual model for each experiment

Experiment factor: Location

Experiment	Term	Factor	Model(order)	Parameter	Estimate	s.e.
Clayton	Residual		Identity	Variance	13591.	3882.
Clinton	Residual		Identity	Variance	21217.	5818.
Plymouth	Residual		Identity	Variance	21778.	5896.

Deviance: -2*Log-Likelihood

Deviance	d.f.
1170.32	99



BLUP estimates of strain means

The next question is how to estimate strain effects or strain means. GenStat provides Best Linear Unbiased Predictor (BLUP) means and/or effects for random terms using the **Save** menu. Before looking at these, what are they? For the following discussion we are indebted to Keith Boldman (Global Data Analysis Methods, Monsanto Company, Iowa).

A BLUP estimate applies to random effects only. The Strain effect technically has a mean of zero, and a variance of σ_s^2 say. However, we really wish to predict the genotype mean for each strain. Write the current model (omitting the random term Location.Block which has a zero variance and hence can be dropped from the model) as

$$Yield = \mu + stain\ effect + Error$$

At one extreme, we could use the i^{th} sample mean as an estimate of (μ + *stain effect*) for the i^{th} strain. This is appropriate when Strain is fixed, and is known as the Best Linear Unbiased Estimator (BLUE). This estimate is unbiased but may have a relatively large variance.

At the other extreme, with no genetic variance, the grand mean is the appropriate estimator for every strain. For our data, we have a genetic variance σ_s^2 which is significantly different to 0.

A BLUP mean is a compromise, or trade-off, between these two estimators. It is calculated by *shrinking* each sample strain mean somewhat toward the grand mean. The degree of shrinkage depends on the estimates of the genetic and environmental variance. The shrinkage ratio, h^2, is given by

$$h^2 = \frac{genetic\ var\ iance}{phenotypic\ var\ iance} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma^2 / r}$$

where r is the number of replicates of each strain and σ^2 is the residual variance. For our data, h^2 = 7095/(7095+20243/9) = 0.76. This ratio is applied to the *deviations* (differences between strain sample means and the grand mean). This reduces the various deviations, giving rise to BLUP effects and hence BLUP means. They are consequently "shrunk" toward the grand mean.

The BLUP effects and BLUP means were captured using Save in GenStat. Select to display the possible random terms. Double click on the random term whose BLUPS you wish to save (in this case Strain). The reduction in the following table is h^2×(deviation from grand mean): this reduction is added to the grand mean to produce the BLUP mean.

Strain	Sample mean	deviation from grand mean	h^2×deviation	BLUP Mean	ranking on sample mean	ranking on BLUP mean
Centennial	1395	-8.47	-6.43	1397	8	8
D74-7741	1406	3.19	2.43	1406	5	5
N72-137	1484	80.64	61.23	1464	3	3
N72-3058	1331	-72.58	-55.11	1348	11	11
N72-3148	1566	162.75	123.57	1527	1	1
N73-1102	1501	98.19	74.56	1478	2	2
N73-693	1436	32.97	25.04	1428	4	4
N73-877	1403	0.08	0.06	1403	6	6
N73-882	1397	-6.58	-5.00	1398	7	7
R73-81	1374	-29.47	-22.38	1381	9	9
R75-12	1176	-227.36	-172.63	1231	12	12
Tracy	1370	-33.36	-25.33	1378	10	10

In this example, no strain has a different ranking on the basis of sample and BLUP means.

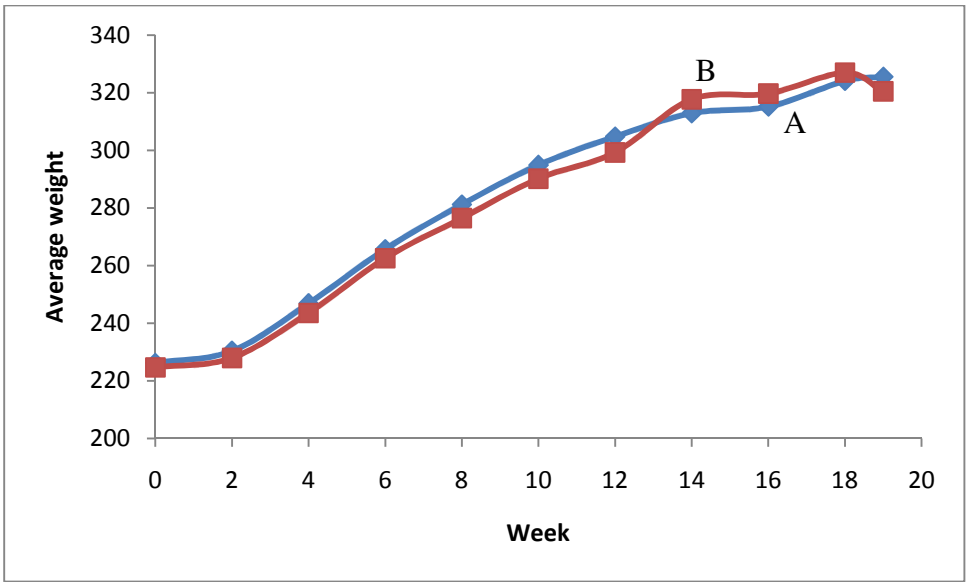
CRD Repeated Measures Example

Calves were randomly allocated to receive treatment A or B (30 calves per treatment). The weight of each calf was recorded 11 times (0, 2, 4, ..., 18, 19 wks). The first 3 calves in each treatment are as follows. Data are from Diggle (1983).

Example 20 Weights of calves from birth to 19 weeks

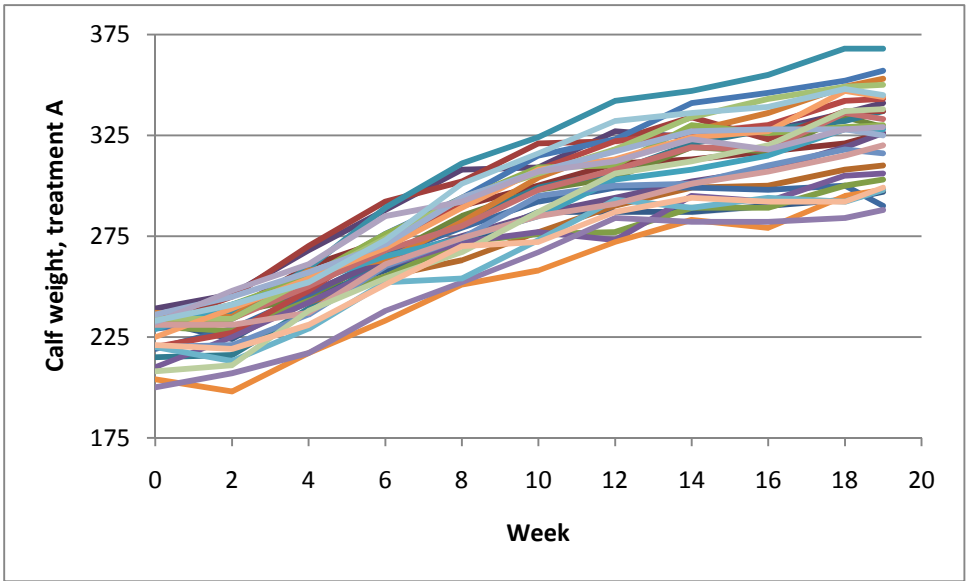
	Treatment A Calf:				Treatment B Calf:			
Week	1	2	3	...	1	2	3	...
0	233	231	232		210	230	226	
2	224	238	237		215	240	233	
4	245	260	245		230	258	248	
6	258	273	265		244	277	277	
8	271	290	285		259	277	297	
10	287	300	298		266	293	313	
12	287	311	304		277	300	322	
14	287	313	319		292	323	340	
16	290	317	317		292	327	354	
18	293	321	334		290	340	365	
19	297	326	329		264	343	362	

The trend in mean calf weights is similar for the two treatments, although mean calf weights for treatment B are consistently below those for treatment A until about week 13.



There is considerable variation in the weights at any week, and there is a suggestion that the variation increases over time (see the following plot for individual calf weights for treatment A). The means and variances over time are as follows. The variance at week 19 is four to six times larger than at birth.

	Week										
	0	2	4	6	8	10	12	14	16	18	19
Treatment	means										
A	226	230	247	266	281	295	305	313	315	324	325
B	225	228	244	263	276	290	299	318	320	327	320
	variances										
A	106	155	165	185	243	284	307	341	389	470	445
B	105	108	147	198	218	250	248	234	287	405	599



There are several ways you could analyse these data, but we will use the data to demonstrate various uses of REML for repeated measurements data.

Firstly, an old-fashioned ANOVA of the data would use time as a split-treatment in a split-plot experiment, with calves randomly assigned to one of two whole-plot treatments – thus, a CRD split-plot experiment. Of course this assumes constant variance over time (which appears an incorrect assumption). A split-plot also assumes that the split-units are also randomised, which for time is not possible. Since for each calf its weight at each time is in the same whole-plot, we have seen with a randomised block that this is equivalent to a *uniform correlation structure over time*.

Here is the split-plot output, ignoring any problems with the assumptions:

Analysis of variance

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Calf.Treatment stratum					
Treatment	1	455.01	455.01	0.20	0.658
Residual	58	133127.50	2295.30	35.37	
Calf.Treatment.Week stratum					
Week	10	846141.94	84614.19	1303.90	<.001
Treatment.Week	10	2264.16	226.42	3.49	<.001
Residual	580	37637.90	64.89		

Total	659	1019626.51		

...

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Calf.Treatment	2295.302	58.000	202.764
Calf.Treatment.Week	64.893	580.000	64.893

Before the advent of modern computers, statisticians developed tests of whether a uniform correlation structure (labelled "symmetry of the covariance matrix") is appropriate over time. When this assumption failed, an adjustment to the ANOVA is made by modifying the degrees of freedom in the split-plot part of the ANOVA. GenStat offers this in the Stats > Repeated Measurements > Analysis of Variance menu.

Box's tests for symmetry of the covariance matrix

Chi-square 599.67 on 64 degrees of freedom: probability <0.001

F-test 9.35 on 64 and 31776 degrees of freedom: probability <0.001

Greenhouse-Geisser epsilon

epsilon 0.2416

Analysis of variance

Variate: Week0,Week2,Week4,Week6,Week8,Week10,Week12,Week14,Week16,Week18,Week19

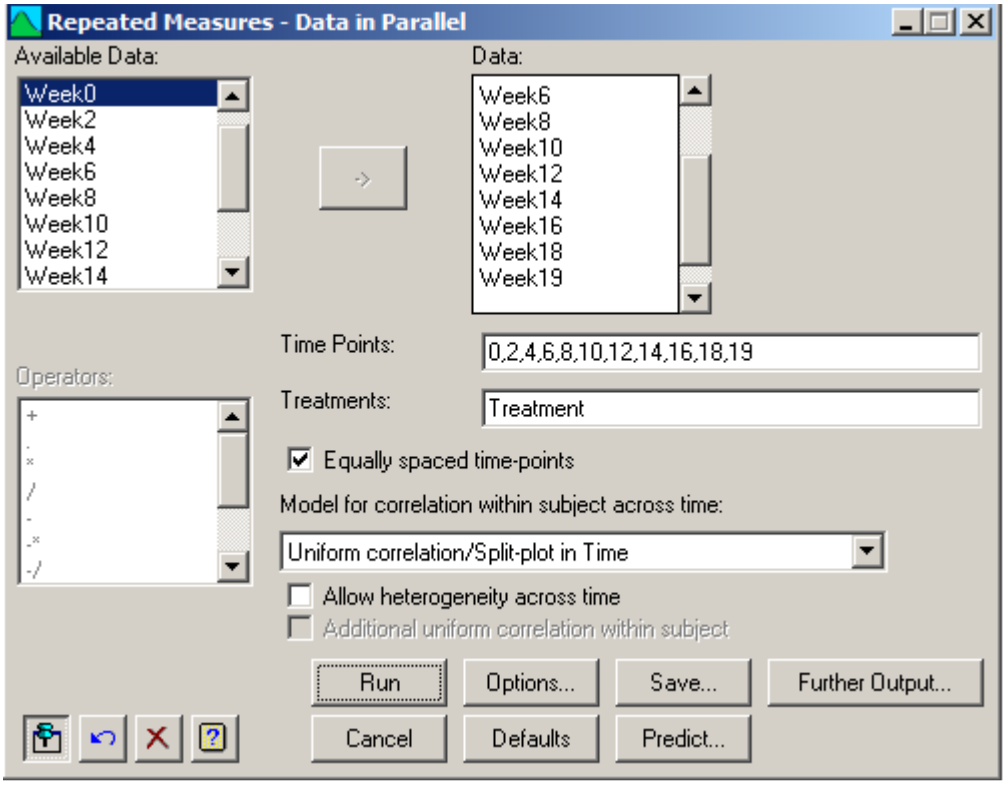
Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Subject stratum					
Treatment	1	455.01	455.01	0.20	0.658
Residual	58	133127.50	2295.30	35.37	
Subject.Time stratum					
d.f. correction factor 0.2416					
Time	10	846141.94	84614.19	1303.90	**<.001**
Time.Treatment	10	2264.16	226.42	3.49	**0.025**
Residual	580	37637.90	64.89		
Total	659	1019626.51			

(d.f. are multiplied by the correction factors before calculating F probabilities)

Again, this approach assumes constant variance, which for plants and animals growing over time is unlikely.

Repeated Measurements > Correlated Models by REML menu

There is a menu in GenStat which analyses CRD repeated measures data using REML. The data can be arranged in separate columns for separate times, or stacked.



Enter the columns of data (if unstacked). The Time Points are for labels in the output. The default correlation structure is uniform, which as we have seen is equivalent to a CRD split-plot with calf weights uniformly correlated over time. Therefore for this correlation structure it does not matter whether the time points are equally spaced or not.

REML variance components analysis

Response variate:	_Data
Fixed model:	Constant + %_Time + %_Treatment + %_Time.%_Treatment
Random model:	%_subject.%_Time
Number of units:	660

%_subject.%_Time used as residual term with covariance structure as below

Sparse algorithm with AI optimisation

Covariance structures defined for random model

Covariance structures defined within terms:

Term	Factor	Model	Order	No. rows
%_subject.%_Time	%_subject	Identity	1	60
	%_Time	Uniform	1	11

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
%_subject.%_Time			Sigma2	267.7	38.9
	%_subject	Identity	-	-	-
	%_Time	Uniform	theta1	0.7576	0.0368

Deviance: -2*Log-Likelihood

Deviance	d.f.
3581.85	636

Note: deviance omits constants which depend on fixed model fitted.

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
%_Time	13039.05	10	1303.90	580.0	<0.001
%_Treatment	0.20	1	0.20	58.0	0.658
%_Time.%_Treatment	34.89	10	3.49	580.0	<0.001

As can be seen:

- The Wald F statistics and df are the same as those from the CRD split-plot ANOVA.
- The estimate Sigma2 (267.7) is the total variance in the experiment. In the earlier ANOVA we selected to display stratum variances, of which there were two: Calf.Treatment (202.764) and Calf.Treatment.Week (64.893) so the total variance is 202.764+64.893 = 267.657.
- The whole-plot error variance can be reconstructed from the total variance and from the estimate of the uniform correlation (theta1), as we have seen before: 0.7576×267.657 = 202.8.

We saw that the variance was much larger at week 19 compared to at birth. REML allows the variance to change across time (Allow heterogeneity across time). The two models are compared using change in deviance:

	Deviance	d.f.	
Constant variance model	3581.85	636	
Changing variance model	3421.05	626	
Change	**160.8**	**10**	**<0.001**

Clearly the changing variance model is statistically better:

Covariance structures defined for random model

Covariance structures defined within terms:

Term	Factor	Model	Order	No. rows
%_subject.%_Time	%_subject	Identity	0	60
	%_Time	Uniform (het)	1	11

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
%_subject.%_Time	Sigma2	1.000	fixed		
	%_subject	Identity	-	-	-

	%_Time	Uniform het	theta1	0.7956	0.0357
			Scale row 1	**139.0**	29.9
			Scale row 2	**141.6**	28.5
			Scale row 3	**154.1**	29.8
			Scale row 4	**179.7**	34.9
			Scale row 5	**213.3**	41.4
			Scale row 6	**242.0**	46.5
			Scale row 7	**264.4**	52.6
			Scale row 8	**267.5**	52.3
			Scale row 9	**321.0**	62.9
			Scale row 10	**451.5**	91.4
			Scale row 11	**577.3**	119.9

Deviance: -2*Log-Likelihood

Deviance	d.f.
3421.05	626

Note: deviance omits constants which depend on fixed model fitted.

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
%_Time	8910.47	10	868.65	234.2	<0.001
%_Treatment	4.99	1	4.99	167.1	0.027
%_Time.%_Treatment	35.53	10	3.46	234.2	<0.001

These variances do increase with time, but they are not very close to the sample variances in all cases. By way of comparison, the average variances across treatments at each time are as follows:

0	2	4	6	8	10	12	14	16	18	19
105.4	131.8	156.2	191.7	230.3	267.1	277.4	287.4	338.1	437.4	521.6

The heterogeneity assumption says that the change in variance is consistent across treatments; possibly it should change with treatment. More probably, the uniform correlation assumption does not hold. Weights closer together are almost certainly more highly correlated than weights distant in time.

Unstructured, autoregressive/power and antedependence models

A simple model to explore is an AR1 structure (the autocorrelation model that applied to the beaver data). However, an AR1 model needs equally spaced time points. When you untick this option, AR1 and AR2 structures are no longer available. The available choices are:

- **Antedependence** order 1 or order 2. From GenStat's Statistics Guide: "Ante-dependence analysis can be regarded as a generalization of multivariate analysis of variance that allows for the patterns of covariances that typify repeated measurements. The variates observed at the successive times are said to have an antedependence structure of order *r* if each *i*th variate (*i*>*r*), given the preceding *r*, is independent of all further preceding variates (Gabriel 1961, 1962)." (See page 1051

for additional explanations.)

- **Power** model (City-block metric)
  If r is the correlation between weights two units of time apart, then r^t is the correlation between weights t units of time apart.

- **Unstructured**
  The whole variance-covariance matrix is estimated. It has no particular structure. It is equivalent to a multivariate CRD analysis with the weights at various times as the variates.

We commence with the unstructured model. For 11 time points there will be an 11 × 11 covariance matrix to print out. This involves 55 different parameter estimates. GenStat uses **v** (for **v**ariance or co**v**ariance) with the row number first and the column number last. So v_11 is the top corner element of the variance matrix (row 1, column 1) and is the *variance* at time 1; v_12 is the *covariance* between times 1 and 2; ... to v_1111 which is the bottom corner element of the variance matrix (row 11, column 11) and hence is the *variance* at time 11.

%_Time	Unstructured	**v_11**	**105.4**	19.6
		v_21	**98.77**	20.19
		v_22	**131.8**	24.5
etc to				
		v_1111	**521.6**	96.9

If you select the option **Covariance Model**, GenStat will rearrange these as a matrix, at least for the first 10 rows; we have added the final row below:

1	105.4										
2	98.8	131.8									
3	102.4	132.2	156.2								
4	95.2	136.8	160.3	191.7							
5	101.6	142.7	166.9	198.0	230.3						
6	104.6	147.0	175.1	210.5	237.7	267.1					
7	96.5	132.5	162.8	199.6	227.6	257.5	277.4				
8	100.0	141.1	169.2	204.4	231.9	261.4	265.4	287.4			
9	107.0	143.8	171.8	209.9	244.8	277.7	285.4	300.5	338.1		
10	102.2	147.0	178.8	218.3	250.4	288.1	287.9	309.0	348.0	437.4	
11	107.0	144.8	184.2	227.2	250.4	291.3	297.2	313.3	353.9	452.3	521.6
	1	2	3	4	5	6	7	8	9	10	11

You can confirm from the table on the previous page that the diagonal elements are simply the average variances across time for the points.

To convert these to a correlation matrix requires diving the covariances (the off-diagonal elements) by the appropriate two standard deviations. Thus, the correlation between the weights at weeks 0 and 2 is 98.8/SQRT(105.4×131.8) = 0.838. The full 11×11 unstructured correlation matrix for the weights over time is as follows:

Unstructured correlation matrix:

1	0.838	0.798	0.670	0.652	0.623	0.564	0.575	0.567	0.476	0.456
0.838	1	0.921	0.861	0.819	0.784	0.693	0.725	0.681	0.612	0.552
0.798	0.921	1	0.926	0.880	0.858	0.782	0.799	0.748	0.684	0.646
0.670	0.861	0.926	1	0.942	0.930	0.866	0.871	0.825	0.754	0.719
0.652	0.819	0.880	0.942	1	0.958	0.900	0.901	0.877	0.789	0.722
0.623	0.784	0.858	0.930	0.958	1	0.946	0.943	0.924	0.843	0.781
0.564	0.693	0.782	0.866	0.900	0.946	1	0.940	0.932	0.827	0.781
0.575	0.725	0.799	0.871	0.901	0.943	0.940	1	0.964	0.872	0.809
0.567	0.681	0.748	0.825	0.877	0.924	0.932	0.964	1	0.905	0.843
0.476	0.612	0.684	0.754	0.789	0.843	0.827	0.872	0.905	1	0.947
0.456	0.552	0.646	0.719	0.722	0.781	0.781	0.809	0.843	0.947	1

Would a power model be a good approximation to this? The correlations alongside 1 in the unstructured correlation matrix are the lag-1 correlations (i.e. the correlations between the weights at each time and the next time); they range from 0.838 to 0.964. Suppose that 0.9 is the overall lag-1 correlation. Then the lag-2 correlation would be $0.9^2 = 0.81$ under a power model, and so on. This is the pattern:

Lag	1	2	3	4	5	6	7	8	9	10
corr	0.90	0.81	0.73	0.66	0.59	0.53	0.48	0.43	0.39	0.35

The patterns are not too dissimilar, perhaps the individual lag-correlations in the matrix tend to be higher than the patterned power structure. The actual estimated power model (with no additional uniform correlation with subjects, but with changing variances over time) is as follows; phi_1 is the overall estimated lag-1 correlation:

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
%_subject.%_Time			Sigma2	1.000	fixed
	%_subject	Identity	-	-	-
	%_Time	Power(1) het	phi_1	0.9583	0.0061
			Scale row 1	**133.9**	23.6
			Scale row 2	**154.5**	26.5
			Scale row 3	**155.0**	25.7
			Scale row 4	**166.5**	27.2
			Scale row 5	**180.5**	28.9
			Scale row 6	**200.7**	32.0
			Scale row 7	**210.8**	34.0
			Scale row 8	**225.2**	36.1
			Scale row 9	**291.8**	47.2
			Scale row 10	**429.7**	70.4
			Scale row 11	**524.3**	86.3

The variance estimates (in bold) are not all close to the average sample variances (in order 105.4, 131.8, 156.2, 191.7, 230.3, 267.1, 277.4, 287.4, 338.1, 437.4, 521.6), so perhaps the model is not a good fit. Since the power structure is a special case of the unstructured model (the 55 individual correlations are replaced by (powers of) a single correlation, we use the change in deviance to determine the adequacy of fit. The df will be 55-1 = 54:

	Deviance	d.f.	P value
Power correlation model with changing variance	3043.48	626	
Unstructured correlation model	2938.73	572	
Change	104.75	54	<0.001

We conclude that the power structure is not an adequate fit.

The antedependence model is designed to be close to the unstructured model, and involves far fewer parameters. Firstly, we check whether order 1 or order 2 is necessary:

	Deviance	d.f.	P value
Antedependence order 1	3005.67	617	
Antedependence order 2	2977.86	608	
Change	27.81	9	0.001

The order 2 model is statistically better than the order 1 model. What do these look like?

The covariance matrix for the antedependence structure, $\mathbf{C}$ say, is defined as a function of a diagonal matrix $\mathbf{D}$ and a matrix $\mathbf{U}$ which has elements all zero apart from the diagonal elements (which are all 1) and, for the order 1 structure, one off diagonal element to the right alongside each diagonal element. For an order 2 structure, $\mathbf{U}$ has two off diagonal elements to the right alongside each diagonal element. Specifically, $\mathbf{C} = (\mathbf{U}\,\mathbf{D}^{-1}\,\mathbf{U}^{T})^{-1}$. GenStat produces the inverses of the diagonal elements of $\mathbf{D}$ (which are labelled dinv_1, dinv_2, ...) and the non-zero elements of $\mathbf{U}$.

Hence, for an order 1 structure over t time points, there are $t+(t-1) = 2t-1$ parameters to estimate (so 21 with 11 time points); for an order 2 structure over t time points, there are $t+(t-1)+(t-2) = 3(t-1)$ parameters to estimate (so 30 with 11 time points).

The antedependence structure is a special case of the unstructured model, for which there are $t(t+1)/2$ parameters to estimate (so 66 with 11 time points). The change in deviance for comparing an unstructured model with an antedependence order 2 structure will therefore have $(t-2)(t-3)/2$ df (so 36 for 11 time points):

	Deviance	d.f.	P value
Antedependence order 2	2977.86	608	
unstructured	2938.73	572	
Change	39.13	36	0.331

The antedependence order 2 model, with 36 fewer parameters, is not a significantly worse model than the unstructured model (P=0.331). However the power model (with variances changing across time) involves ever fewer parameters: 11 time variances and 1 correlation coefficient for a unit time difference. Since the power model is not a special case of the antedependence model, we cannot use change in deviance to compare them. GenStat offers as an option two coefficients that can be used in this situation.

Akaike's information criterion (AIC) and Schwartz information coefficient (SC)

These coefficients are both related to the deviance. As stated, they do not represent a formal test of two competing models, they are simply tools for model selection. The lower their value the less information is lost and the better the model is. GenStat offers these as options in the LMM (REML) menu.

The AIC and SC values for the power model with changing variances are 4243.89 and 4301.87; for the antedependence order 2 model they are 4320.50 and 4329.41, which are larger by 76.61 and 27.54 units respectively. The difference is largely because the power model involves fewer parameters, so is a trade off between the deviance and the number of parameters fitted. On the AIC and SC alone the power model appears the better choice. However, the change in deviance suggested the power model is not a good fit to the unstructured model, whereas the antedependence order 2 model is. The output for this model is:

REML variance components analysis

Response variate: _Data
Fixed model: Constant + %_Time + %_Treatment + %_Time.%_Treatment
Random model: %_subject.%_Time
Number of units: 660
%_subject.%_Time used as residual term with covariance structure as below

Covariance structures defined for random model

Covariance structures defined within terms:

Term	Factor	Model	Order	No. rows
%_subject.%_Time	%_subject	Identity	0	60
	%_Time	Antedependence	1	11

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
%_subject.%_Time			Sigma2	1.000	fixed
	%_subject	Identity	-	-	-
	%_Time	Antedependence(1)			
			dinv_1	0.009486	0.001778
			dinv_2	0.02549	0.00479
			dinv_3	0.04245	0.00790
			dinv_4	0.03680	0.00684
			dinv_5	0.03874	0.00723
			dinv_6	0.04578	0.00850
			dinv_7	0.03439	0.00643
			dinv_8	0.02994	0.00558
			dinv_9	0.04200	0.00780
			dinv_10	0.01263	0.00235
			dinv_11	0.01855	0.00344
			u_12	-0.9370	0.0809
			u_23	-1.003	0.056
			u_34	-1.026	0.056
			u_45	-1.033	0.049
			u_56	-1.032	0.041
			u_67	-0.9642	0.0446

u_78	-0.9569	0.0473
u_89	-1.046	0.039
u_910	-1.029	0.064
u_1011	-1.034	0.047

Estimated covariance models

Variance of data estimated in form:

V(y) = Sigma2.R

where: V(y) is variance matrix of data
Sigma2 is the residual variance
R is the residual covariance matrix
....

Factor: %_Time
Model: Antedependence

Covariance matrix (first 10 rows only):

1	105.4									
2	98.8	131.8								
3	99.1	132.2	156.2							
4	101.7	135.7	160.3	191.7						
5	105.0	140.1	165.6	198.0	230.3					
6	108.4	144.6	170.8	204.3	237.7	267.1				
7	104.5	139.4	164.7	197.0	229.2	257.5	277.4			
8	100.0	133.4	157.6	188.5	219.3	246.4	265.4	287.4		
9	104.6	139.5	164.8	197.1	229.3	257.7	277.6	300.5	338.1	
10	107.6	143.6	169.7	202.9	236.0	265.2	285.7	309.3	348.0	437.4
	1	2	3	4	5	6	7	8	9	10

Deviance: -2*Log-Likelihood

Deviance	d.f.
3005.67	617

Note: deviance omits constants which depend on fixed model fitted.

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
%_Time	3095.09	10	296.89	164.0	<0.001
%_Treatment	0.02	1	0.02	59.7	0.898
%_Time.%_Treatment	66.90	10	6.42	164.0	<0.001

One of the benefits of choosing an appropriate variance matrix over time is the appropriate precision for comparing treatment means at any time, or the difference in means for a particular treatment over time. A split-plot in time analysis assumes constant variance. For such an analysis, the same (inappropriate) sed value is used. Here are the means and sed values from the antedependence order 2 model:

Week	A	B	diff	sed
0	226.20	224.60	1.60	2.65
2	230.33	227.90	2.43	2.96
4	246.87	243.53	3.33	3.23
6	265.63	262.50	3.13	3.57
8	281.17	276.43	4.73	3.92
10	294.87	290.13	4.73	4.22
12	304.73	299.23	5.50	4.30
14	312.87	317.67	-4.80	4.38
16	315.13	319.67	-4.53	4.75
18	324.07	326.93	-2.87	5.40
19	325.47	320.47	5.00	5.90

Finally, we compare the variance matrix across time for the antedependence order 2 and unstructured models. You can see the variance estimates are the sample variances across time for both models. The covariances are identical to lag-2 (apart from the occasional round off error), and are not too different beyond lag-2.

antedependence order 2 variance matrix:

1	**105.4**										
2	98.8	**131.8**									
3	102.4	132.2	**156.1**								
4	105.8	136.8	160.3	**191.7**							
5	110.1	142.4	166.9	198.0	**230.3**						
6	116.8	151.1	177.1	210.5	237.7	**267.1**					
7	112.1	145.0	169.9	202.0	227.5	257.5	**277.3**				
8	114.1	147.5	172.9	205.6	231.9	261.4	265.4	**287.4**			
9	121.0	156.4	183.3	218.0	245.7	277.3	285.4	300.5	**338.0**		
10	124.4	160.9	188.6	224.2	252.8	285.2	293.6	309.0	348.0	**437.4**	
11	126.6	163.7	191.9	228.1	257.2	290.1	298.7	314.4	354.0	452.4	**521.9**
	1	2	3	4	5	6	7	8	9	10	11

unstructured variance matrix:

1	**105.4**										
2	98.8	**131.8**									
3	102.4	132.2	**156.2**								
4	95.2	136.8	160.3	**191.7**							
5	101.6	142.7	166.9	198.0	**230.3**						
6	104.6	147.0	175.1	210.5	237.7	**267.1**					
7	96.5	132.5	162.8	199.6	227.6	257.5	**277.4**				
8	100.0	141.1	169.2	204.4	231.9	261.4	265.4	**287.4**			
9	107.0	143.8	171.8	209.9	244.8	277.7	285.4	300.5	**338.1**		
10	102.2	147.0	178.8	218.3	250.4	288.1	287.9	309.0	348.0	**437.4**	
11	107.0	144.8	184.2	227.2	250.4	291.3	297.2	313.3	353.9	452.3	**521.6**
	1	2	3	4	5	6	7	8	9	10	11

RCBD repeated measures example - experiments repeated annually

Snedecor and Cochran presented an analysis of asparagus yields taken from an experiment in which planting occurred in 1929 and cuttings commenced in 1930. Data are available for four years from the same plots. This was a randomized block, with four plots in each block. The four plots corresponded to cuttings taken on June 1 each year, but for three of the plots additional cuttings were taken (but not analysed). The intent of the analysis was to detect if repeated cutting of asparagus affected plant vigour.

Example 21 Asparagus yields from four annual cuttings, from Snedecor and Cochran, page 330-2.

		Year			
Block	Cutting ceased	1930	1931	1932	1933
1	Jun-01	230	324	512	399
	Jun-15	212	415	584	386
	Jul-01	183	320	456	255
	Jul-15	148	246	304	144
2	Jun-01	216	317	448	361
	Jun-15	190	296	471	280
	Jul-01	186	295	387	187
	Jul-15	126	201	289	83
3	Jun-01	219	357	496	344
	Jun-15	151	278	399	254
	Jul-01	177	298	427	239
	Jul-15	107	192	271	90
4	Jun-01	200	362	540	381
	Jun-15	150	336	485	279
	Jul-01	209	328	462	244
	Jul-15	168	226	312	168

Clearly, the same plot is repeatedly measured, and hence yields for the same plot are most likely correlated across years.

Snedecor and Cochran overcame that problem by (a) an analysis of total annual yields, and (b) an analysis of the linear yield component over years (using multipliers -3, -1, 1, 3), which was (then) a way of overcoming the correlated nature of the data.

If you believe that the correlation structure over time was uniform, a split-plot RCBD would be appropriate (and would be the correct analysis if only two years were involved). This analysis is:

Analysis of variance

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	3	30169.6	10056.5	4.14	
Block.CuttingTime stratum					
CuttingTime	3	241376.6	80458.9	33.12	<.001
Residual	9	21860.8	2429.0	5.65	
Block.CuttingTime.Year stratum					
Year	3	518721.9	172907.3	401.94	<.001
CuttingTime.Year	9	51177.5	5686.4	13.22	<.001
Residual	36	15486.6	430.2		
Total	63	878793.0			

Estimated stratum variances

Stratum	variance	effective d.f.	variance component
Block	10056.54	3.000	476.72
Block.CuttingTime	2428.97	9.000	499.70
Block.CuttingTime.Year	430.18	36.000	430.18

LMM (REML) analysis

The analysis of the asparagus yields is an example of the need for a temporal correlation model for plots measured annually. Since the years were equally spaced, AR, antedependence and unstructured models are potential correlation models.

A split-plot in time analysis can be set up in REML as follows.

The random model for a general split-plot is
Block/Whole_Plot/Split_Plot
which expands to
Block + Block.Whole_Plot + Block.Whole_Plot.Split_Plot

Recall that for a randomised block with blocks random, the random model is
Block+Block.Plot

and this can be replaced by Block.Plot with a uniform correlation structure for the plots.

In the split-plot case, the split-plot treatment (Year) will be explored for an appropriate correlation structure. So by analogy with the RCB case, we work backwards and replace the last two random terms (Block.Whole_Plot + Block.Whole_Plot.Split_Plot) by a single term Block.Whole_Plot.Split_Plot with a uniform correlation structure on the split-plot units.

For the example, CuttingTime is the whole-plot treatment and Year the split-plot treatment, and we can use these factors in lieu of the unit names in the random model. Hence the split-plot ANOVA should be equivalent to a REML analysis with:

Fixed Model: Year*Cuttings
Random Model: Block+Block.Cuttings.Year with a uniform correlation structure on Year.

REML variance components analysis

Response variate: Yield
Fixed model: Constant + CuttingTime + Year + CuttingTime.Year
Random model: Block + Block.CuttingTime.Year
Number of units: 64

Block.CuttingTime.Year used as residual term with covariance structure as below

Sparse algorithm with AI optimisation

Covariance structures defined for random model

Covariance structures defined within terms:

Term	Factor	Model	Order	No. rows
Block.CuttingTime.Year	Block	Identity	1	4
	CuttingTime	Identity	0	4
	Year	**Uniform**	1	4

Estimated variance components

Random term	component	s.e.
Block	**476.7**	518.2

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
Block.CuttingTime.Year			**Sigma2**	**929.9**	296.2
	Block	Identity	-	-	-
	CuttingTime	Identity	-	-	-
	Year	Uniform	**theta1**	**0.5374**	0.1592

Deviance: -2*Log-Likelihood

Deviance	d.f.
386.30	45

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
CuttingTime	99.37	3	33.12	9.0	<0.001
Year	1205.81	3	401.94	36.0	<0.001
CuttingTime.Year	118.97	9	13.22	36.0	<0.001

You can see that

- the F statistics and df are identical to those from the ANOVA.

- The estimate of the variance of the random block effect (476.7) is the same as the Block stratum variance from the ANOVA.

- The estimate **Sigma2** (929.9) is the total variance of the two terms replaced in the REML with a uniform structure, ie the whole-plot error and the split-plot error, From the ANOVA, the two stratum variances were 499.70 and 430.18 respectively, and these add to 929.88.

- The estimate of the Block variance in an RCB was reconstructed by multiplying the uniform correlation by the total variance, so here the whole-plot error is simply 0.5374×929.88 = 499.71. This is the same as the whole-plot stratum variance from the ANOVA.

Years are equally spaced, and changing to an AR1 correlation structure over years (plus a random block effect) produces a similar size deviance (compared to uniform; we can't test the deviances for these two models as one is not a special case of the other). An AR2 structure is certainly unnecessary for these data (P=0.498). With an AR1 model, there also appears to be no need to have the variance change across years (P=0.440):

Correlation structure for Year	Deviance	d.f.	
Uniform	386.30	45	
AR1	382.77	45	
AR1 + changing variance (years)	380.07	42	Change=2.70, df=3, P=0.440
AR2	382.31	44	Change=0.46, df=1, P=0.498

When we try and fit an unstructured model over time the estimate of the block variance becomes negative; when constrained to be positive the deviance is 370.10 with 37 df. Hence, the AR1 model is a statistically acceptable model in comparison to the unstructured model (change in deviance = 12.67 on 8 df, P=0.124) and involves 8 (or 7 if Block is omitted) fewer parameters.

The antedependence order 2 model is not a significantly better model than the order 1 model (P=0.827) on the basis of the following change in deviance:

Correlation structure for Year	Deviance	d.f.
antedependence 1	374.61	40
antedependence 2	374.23	38
change	0.38	2

Finally, the antedependence order 1 model is also a statistically acceptable model in comparison to the unstructured model (change in deviance = 4.51 on 3 df, P=0.211). The AR1 model says that the asparagus yields are directly related to the previous year's yield, and indirectly related to the yields in earlier years. The antedependence order 1 model says that the yield is dependent on the previous year's yield, but given that yield, it is uncorrelated with the yields from previous years. It allows the variance to change across years as well.

Here is the full output from the antedependence model. The superiority of this model compared to the split-plot in time (uniform) model lies in the precision for comparing the

cutting time means within and across years. For the latter model, the sed for a comparison between a particular cutting time mean for any two years is 14.7; for comparing any two cutting time means in a particular year, or across years, is 21.6. For the antedependence order 1 model, the 14.7 common sed is replaced by a range of sed values whose minimum is 11.2 and whose maximum is 20.1; the 21.6 sed is replaced by a range of sed values whose minimum is 11.2 and whose maximum is 30.73. The maximum value applies to a comparison with 1932, a year in which both yields and the estimated variance were high.

REML variance components analysis

Response variate:	Yield
Fixed model:	Constant + CuttingTime + Year + CuttingTime.Year
Random model:	Block + Block.CuttingTime.Year
Number of units:	64

Block.CuttingTime.Year used as residual term with covariance structure as below

Covariance structures defined for random model

Covariance structures defined within terms:

Term	Factor	Model	Order	No. rows
Block.CuttingTime.Year	Block	Identity	0	4
	CuttingTime	Identity	0	4
	Year	Antedependence	1	4

Estimated variance components

Random term	component	s.e.
Block	86.000	155.907

Residual variance model

Term	Factor	Model(order)	Parameter	Estimate	s.e.
Block.CuttingTime.Year			Sigma2	1.000	fixed
	Block	Identity	-	-	-
	CuttingTime	Identity	-	-	-
	Year	Antedependence(1)			
			dinv_1	0.002333	0.001056
			dinv_2	0.001157	0.000480
			dinv_3	0.002082	0.000852
			dinv_4	0.002011	0.000830
			u_12	-0.7185	0.4525
			u_23	-1.139	0.196
			u_34	-0.6786	0.1554

Estimated covariance models

Variance of data estimated in form:

V(y) = sZZ' + Sigma2.R

where: V(y) is variance matrix of data
s is the variance component for the random term
Z is the incidence matrix for the random term

Sigma2 is the residual variance
R is the residual covariance matrix

Random Term: Block

Scalar s: 86.00

Residual term: Block.CuttingTime.Year
Sigma2: 1.000

R uses direct product construction

Factor: Block
Model: Identity (4 rows)

Factor: CuttingTime
Model: Identity (4 rows)

Factor: Year
Model: Antedependence

1	0.45	0.39	0.31
0.45	1	0.86	0.69
0.39	0.86	1	0.80
0.31	0.69	0.80	1

Covariance matrix:

1	428.7			
2	308.0	1085.2		
3	350.9	1236.3	1888.8	
4	238.1	839.0	1281.8	1367.2
	1	2	3	4

The *correlation* matrix among the 4 times is:

1	1	0.45	0.39	0.31
2	0.45	1	0.86	0.69
3	0.39	0.86	1	0.80
4	0.31	0.69	0.80	1
	1	2	3	4

Deviance: -2*Log-Likelihood

Deviance	d.f.
374.61	40

Note: deviance omits constants which depend on fixed model fitted.

Tests for fixed effects

Sequentially adding terms to fixed model

Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
CuttingTime	50.38	3	16.79	11.0	<0.001
Year	936.91	3	275.77	13.4	<0.001
CuttingTime.Year	79.34	9	7.51	16.9	<0.001

Table of predicted means for Constant

290.6 Standard error: 8.57

Table of predicted means for CuttingTime

CuttingTime	Jun_01	Jun_15	Jul_01	Jul_15
	356.6	322.9	290.8	192.2

Standard errors of differences between pairs

CuttingTime Jun_01	1	*			
CuttingTime Jun_15	2	20.4	*		
CuttingTime Jul_01	3	20.4	20.4	*	
CuttingTime Jul_15	4	20.4	20.4	20.4	*
		1	2	3	4

Standard error of differences: 20.37

Table of predicted means for Year

Year	1930	1931	1932	1933
	179.5	299.4	427.7	255.9

Standard errors of differences between pairs

Year 1930	1	*			
Year 1931	2	7.5	*		
Year 1932	3	10.0	5.6	*	
Year 1933	4	9.1	7.0	6.6	*
		1	2	3	4

Standard errors of differences

Average:	7.626
Maximum:	10.05
Minimum:	5.598

Average variance of differences: 60.43

Table of predicted means for CuttingTime.Year

Year	1930	1931	1932	1933
CuttingTime				
Jun_01	216.2	340.0	499.0	371.2
Jun_15	175.8	331.2	484.8	299.8
Jul_01	188.7	310.2	433.0	231.2
Jul_15	137.3	216.2	294.0	121.2

Standard errors of differences between pairs

CuttingTime Jun_01.Year 1930	1	*				
CuttingTime Jun_01.Year 1931	2	15.0	*			
CuttingTime Jun_01.Year 1932	3	20.1	11.2	*		
CuttingTime Jun_01.Year 1933	4	18.2	13.9	13.2	*	
CuttingTime Jun_15.Year 1930	5	14.6	19.5	24.1	21.2	*
CuttingTime Jun_15.Year 1931	6	19.5	23.3	27.3	24.8	15.0
CuttingTime Jun_15.Year 1932	7	24.1	27.3	30.7	28.5	20.1
CuttingTime Jun_15.Year 1933	8	21.2	24.8	28.5	26.1	18.2
CuttingTime Jul_01.Year 1930	9	14.6	19.5	24.1	21.2	14.6
CuttingTime Jul_01.Year 1931	10	19.5	23.3	27.3	24.8	19.5
CuttingTime Jul_01.Year 1932	11	24.1	27.3	30.7	28.5	24.1
CuttingTime Jul_01.Year 1933	12	21.2	24.8	28.5	26.1	21.2
CuttingTime Jul_15.Year 1930	13	14.6	19.5	24.1	21.2	14.6
CuttingTime Jul_15.Year 1931	14	19.5	23.3	27.3	24.8	19.5

CuttingTime Jul_15.Year 1932	15	24.1	27.3	30.7	28.5	24.1
CuttingTime Jul_15.Year 1933	16	21.2	24.8	28.5	26.1	21.2
		1	2	3	4	5

CuttingTime Jun_15.Year 1931	6	*				
CuttingTime Jun_15.Year 1932	7	11.2	*			
CuttingTime Jun_15.Year 1933	8	13.9	13.2	*		
CuttingTime Jul_01.Year 1930	9	19.5	24.1	21.2	*	
CuttingTime Jul_01.Year 1931	10	23.3	27.3	24.8	15.0	*
CuttingTime Jul_01.Year 1932	11	27.3	30.7	28.5	20.1	11.2
CuttingTime Jul_01.Year 1933	12	24.8	28.5	26.1	18.2	13.9
CuttingTime Jul_15.Year 1930	13	19.5	24.1	21.2	14.6	19.5
CuttingTime Jul_15.Year 1931	14	23.3	27.3	24.8	19.5	23.3
CuttingTime Jul_15.Year 1932	15	27.3	30.7	28.5	24.1	27.3
CuttingTime Jul_15.Year 1933	16	24.8	28.5	26.1	21.2	24.8
		6	7	8	9	10

CuttingTime Jul_01.Year 1932	11	*				
CuttingTime Jul_01.Year 1933	12	13.2	*			
CuttingTime Jul_15.Year 1930	13	24.1	21.2	*		
CuttingTime Jul_15.Year 1931	14	27.3	24.8	15.0	*	
CuttingTime Jul_15.Year 1932	15	30.7	28.5	20.1	11.2	*
CuttingTime Jul_15.Year 1933	16	28.5	26.1	18.2	13.9	13.2
		11	12	13	14	15

CuttingTime Jul_15.Year 1933	16	*
		16

Standard errors of differences

Average:	22.32
Maximum:	30.73
Minimum:	11.20

Average variance of differences: 525.3

Standard error of differences for same level of factor:

	CuttingTime	Year
Average:	15.25	23.70
Maximum:	20.10	30.73
Minimum:	11.20	14.64

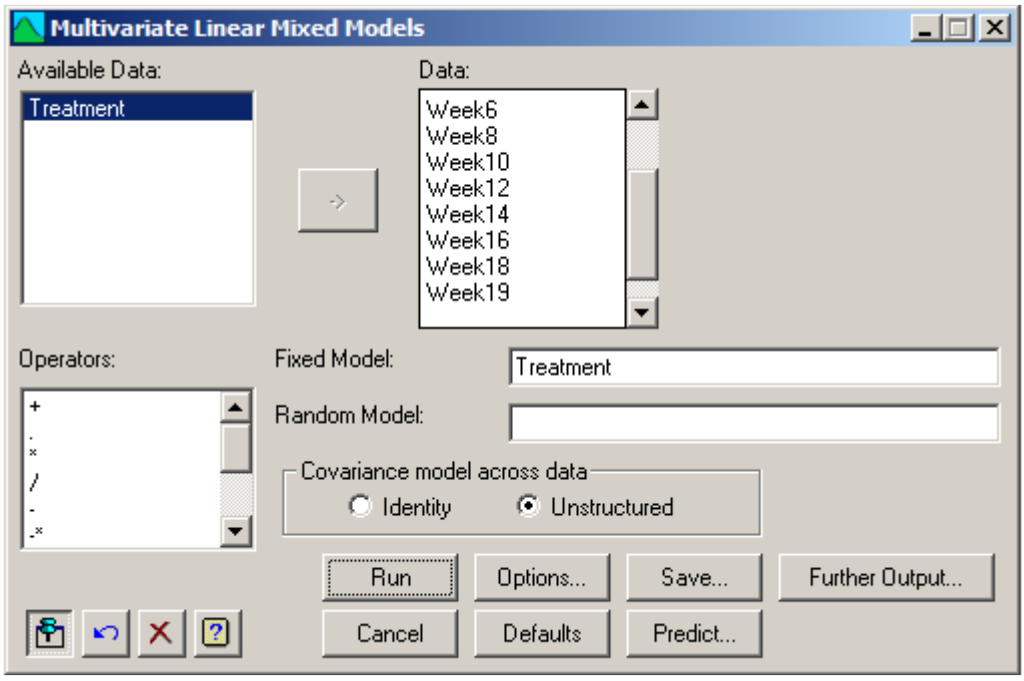
Average variance of differences:

241.7	596.2

Multivariate Linear Mixed Models for CRD

REML offers an alternative to multivariate analysis of variance (MANOVA) which becomes very useful for unbalanced data. To illustrate the two techniques we use the calf weights measured 11 times over the first 19 weeks from birth. We used these data previously to illustrate repeated measurements analysis when we assumed *unstructured* model (i.e. no particular structure for the variances and correlations) over time. This is essentially the method GenStat uses when selecting Stats > Mixed Models (REML) > Multivariate Linear Mixed Models. The data need to be unstacked for this menu. Basically, the test is comparing the *entire set of mean weights* across time for the two treatments is a single analysis.

There are two choices to make for the Covariance model across data. The first, Identity, simply assumes that the time variates are uncorrelated; a different variance will be fitted for each variate, hence the variance matrix fitted is Diagonal. The second will be shown to produce one of the MANOVA test statistics. As usual, we use change in deviance to decide between the two models.



Model	Deviance	d.f.	P value
Correlated times (Unstructured)	4211.4	627	
Uncorrelated times (Identity)	2938.7	572	
Change	1272.7	55	<0.001

There is overwhelming evidence that the data are correlated over time. The variances and covariances from this analysis were presented previously, as well as the reconstructed correlation matrix. (Remember that GenStat labels these v_11, v_12, v_22, ... in a long list in the output. Choose to show the Covariance Model to have them printed out in (lower triangular) matrix form, at least for up to 10 rows.

Full covariance matrix across the 11 time points:

1	105.4										
2	98.8	131.8									
3	102.4	132.2	156.2								
4	95.2	136.8	160.3	191.7							
5	101.6	142.7	166.9	198.0	230.3						
6	104.6	147.0	175.1	210.5	237.7	267.1					
7	96.5	132.5	162.8	199.6	227.6	257.5	277.4				
8	100.0	141.1	169.2	204.4	231.9	261.4	265.4	287.4			
9	107.0	143.8	171.8	209.9	244.8	277.7	285.4	300.5	338.1		
10	102.2	147.0	178.8	218.3	250.4	288.1	287.9	309.0	348.0	437.4	
11	107.0	144.8	184.2	227.2	250.4	291.3	297.2	313.3	353.9	452.3	521.6
	1	2	3	4	5	6	7	8	9	10	11

Tests for fixed effects

Sequentially adding terms to fixed model

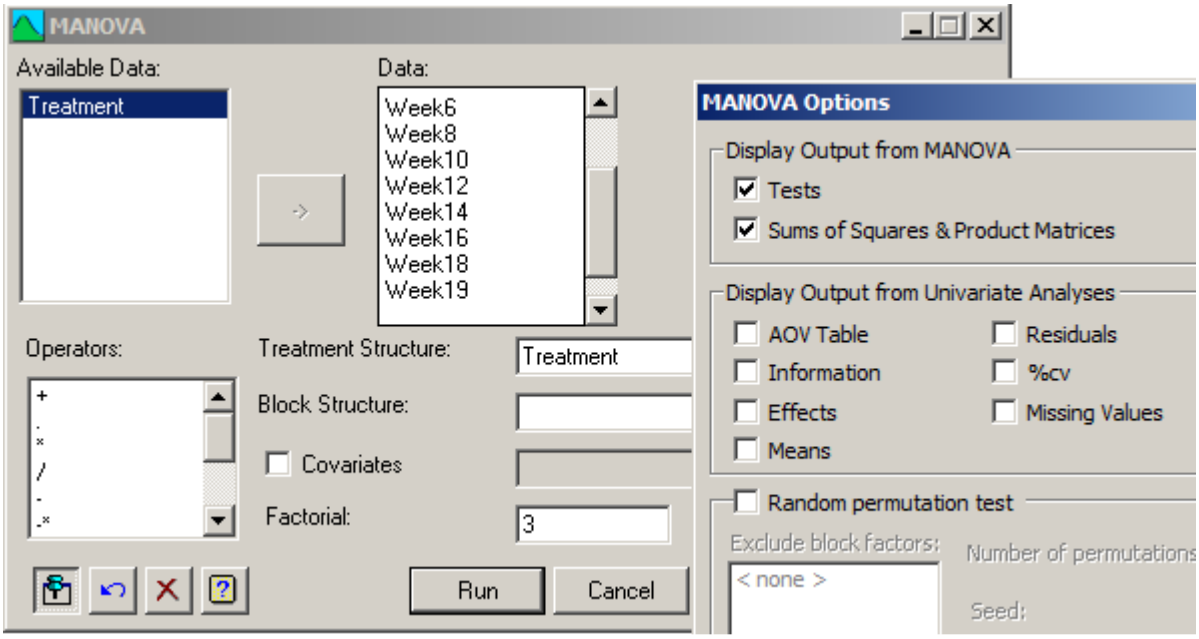
Fixed term	Wald statistic	n.d.f.	F statistic	d.d.f.	F pr
%_variable	36243.52	11	2726.79	48.0	<0.001
%_variable.%_Treatment	**86.07**	**11**	**6.48**	**48.0**	**<0.001**

There is a highly significant difference between the treatment A set of calf weight means and the treatment B set (P<0.001). The means, all s.e.d. and l.s.d. values are suppressed in this section.

Multivariate analysis of variance (MANOVA) for CRD

The MANOVA is obtained in Stats > Multivariate Analysis > MANOVA. In Options you can choose to have the sums of squares and products matrices printed out – these are the variance matrices for treatments and residual. You can also choose to have separate ANOVAs printed (AOV Table). This is appropriate if the data are uncorrelated over time, and essentially performs all the ANOVA in one step.

Firstly, a univariate ANOVA constructs an F statistic as the ratio (Treatment MS)/(Residual MS), or a scalar multiple of (Treatment SS)/(Residual SS). The problem confronting the early statisticians is how to generalize a ratio to MANOVA in which both Treatment SS and Residual SS are *matrices*: on the diagonal are sums of squares, off the diagonal are sums of products, so we re-label SS as SSP to reflect this. The denominator in the univariate F becomes an inverse of a matrix for a multivariate set of data, so the test is based on some aspect of (Treatment SSP)(Residual SSP)$^{-1}$. The MANOVA test statics are all named after statisticians who developed the different mathematical functions of this matrix expression. These tests are all based on some function of eigenvalues.



For the calf data the sums of squares and products matrices are as follows:

SSP matrices

Treatment

(Lower triangular part of each matrix is shown here, for times 0, 2, 4, …, 18, 19):

0	38.4										
2	58.4	88.8									
4	80	121.7	166.7								
6	75.2	114.4	156.7	147.3							
8	113.6	172.8	236.7	222.5	336.1						
10	113.6	172.8	236.7	222.5	336.1	336.1					
12	132	200.8	275	258.5	390.5	390.5	453.8				
14	-115.2	-175.2	-240	-225.6	-340.8	-340.8	-396	345.6			
16	-108.8	-165.5	-226.7	-213.1	-321.9	-321.9	-374	326.4	308.3		
18	-68.8	-104.6	-143.3	-134.7	-203.5	-203.5	-236.5	206.4	194.9	123.3	
19	120	182.5	250	235	355	355	412.5	-360	-340	-215	375
	0	2	4	6	8	10	12	14	16	18	19

Degree of freedom: 1

Residual

0	6114										
2	5729	7643									
4	5938	7667	9057								
6	5521	7933	9296	11116							
8	5891	8276	9681	11483	13360						
10	6065	8527	10157	12210	13785	15491					
12	5595	7686	9440	11575	13200	14936	16087				
14	5800	8182	9815	11856	13451	15161	15394	16668			
16	6205	8343	9967	12176	14200	16108	16554	17430	19608		
18	5929	8525	10372	12663	14524	16712	16697	17925	20183	25368	
19	6205	8400	10686	13180	14524	16897	17235	18169	20529	26232	30253
	0	2	4	6	8	10	12	14	16	18	19

Degree of freedom: 58

If you look at say the first ANOVA, you will see that the diagonal terms of the matrices are simply the Treatment SS (38.4) and Residual SS (6114).

Analysis of variance

Variate: Week0

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
units stratum					
Treatment	1	**38.4**	38.4	0.36	0.548
Residual	58	**6114.0**	105.4		
Total	59	6152.4			

Test statistics for MANOVA

Term	d.f.	Wilk's lambda	**Rao F**	**n.d.f.**	**d.d.f.**	**F prob.**
Treatment	1	0.4026	**6.48**	**11**	**48**	**0.000**

Term	d.f.	Pillai-Bartlett trace	Roy's maximum root test	Lawley-Hotelling trace
Treatment	1	0.5974	0.5974	1.484

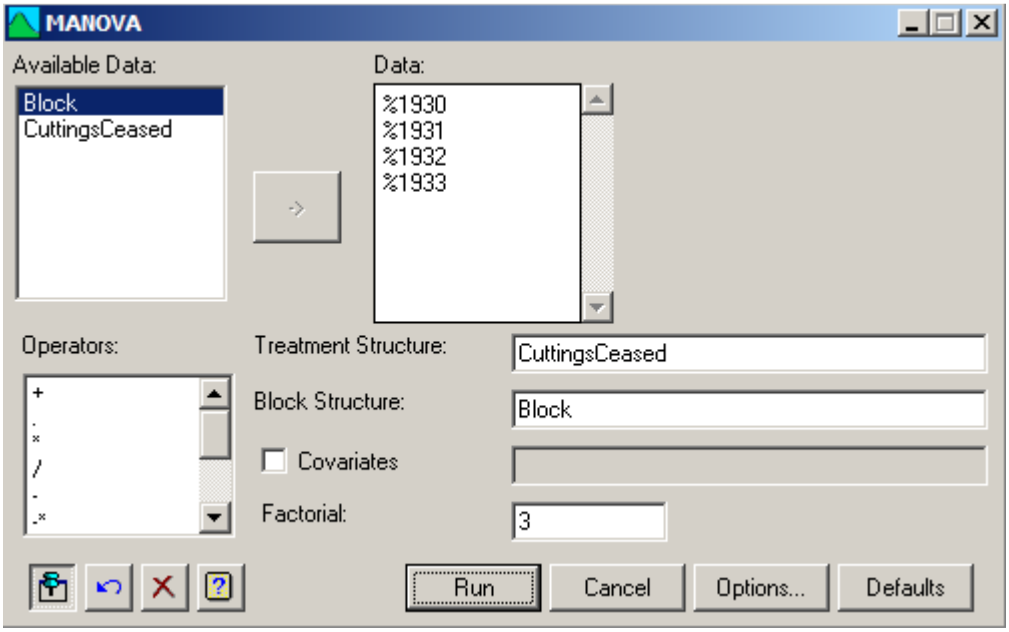
- Notice that the Rao F statistic of 6.48 is the same as the test of treatment means across variates in the Multivariate REML:

%_variable.%_Treatment	**86.07**	**11**	**6.48**	**48.0**	**<0.001**

The means and s.e.d. values are printed out as an option, but not l.s.d. values. MANOVA is also restricted to balanced data, so the REML approach has the advantage.

Multivariate analysis of variance (MANOVA) for a blocked design
We consider again the asparagus yields from four annual cuttings of plots treated with one of four cutting methods set out in four randomized blocks (Snedecor and Cochran, page 330-2).

The MANOVA is a simple extension of the CRD MANOVA – we simply set up the blocking structure using the unstacked data:



In Version 12 of GenStat there is a warning which we can ignore, as it does not affect tests or P values:

Multivariate analysis of variance

SSP matrices

Block stratum

Warning 11, code UF 2, statement 239 in procedure MANOVA

Residual SSP matrix for Block singular.

Residual

%1930	1800			
%1931	2761	6904		
%1932	4080	9801	14037	
%1933	3860	9212	12994	12520
	%1930	%1931	%1932	%1933

Degree of freedom: 3

Block._units_ stratum

CuttingsCeased

%1930	12941			
%1931	19944	38778		
%1932	32417	63546	104969	
%1933	38142	68034	114393	135867
	%1930	%1931	%1932	%1933

Degree of freedom: 3

Residual

%1930	4144			
%1931	2458	8363		
%1932	3020	8000	12316	
%1933	4012	4099	6004	7433
	%1930	%1931	%1932	%1933

Degree of freedom: 9

Test statistics

Block._units_ stratum

Term	d.f.	Wilk's lambda	**Rao F**	**n.d.f.**	**d.d.f.**	**F prob.**
CuttingsCeased	3	0.009994	**6.33**	**12**	**16**	**0.000**

Term	d.f.	Pillai-Bartlett trace	Roy's maximum root test	Lawley-Hotelling trace
CuttingsCeased	3	1.971	0.9586	25.24

Again, notice that the Rao F test is highly significant (P<0.001) – remember we never use 0.000 in a report. This variance ratio should be the same as the multivariate REML using an unstructured correlation matrix over time. Unfortunately, current versions of GenStat have a problem estimating the variance matrix - the default steps in the iteration routine are too large to lead to convergence - so we are unable to demonstrate the equivalence of the two analyses at this stage.

When setting up multivariate REML for an RCBD, use
Fixed Model: Time/Treatment
Random Model: Block.Time+Units.Time

and, if Time is unstructured for both random terms, the Rao F statistic of MANOVA will be the same as the Wald F test for Treatment.Time in the multivariate REML.

In the MANOVA output, the diagonal elements of the sum of squares and products matrices are simply the Block, Treatment and Residual sums of squares from the univariate ANOVAS. For example, here is the ANOVA for 1930. The three sums of squares are the leading element of the three matrices for Block, CuttingsCeased and Residual respectively:

Analysis of variance

Variate: %1930

Source of variation	d.f.	**s.s.**	m.s.	v.r.	F pr.
Block stratum	3	**1800.5**	600.2	1.30	
Block._units_ stratum					
CuttingsCeased	3	**12941.0**	4313.7	9.37	0.004
Residual	9	**4144.5**	460.5		
Total	15	18886.0			

The off-diagonal elements are the sum of products between the corresponding terms from pairs of ANOVAs. The Residual matrix provides the estimated correlations of the data among years. There are 9 df for each term in the matrix, so the variance matrix is:

460.4			
273.1	929.2		
335.6	888.9	1368.4	
445.8	455.4	667.1	825.9

and the correlation matrix from this is:

1			
0.418	1		
0.423	0.788	1	
0.723	0.520	0.628	1

The correlation matrix from the antedependence model (with no Block.Year random term) was similar, apart from the correlation between 1930 and 1933 data. (There are only 9 df for variances and covariances, so this discrepancy is not unsurprising.)

1			
0.45	1		
0.39	0.86	1	
0.31	0.69	0.80	1

Appendix 1 Revision of basic random sampling

Distribution of a **sample mean** of n data values from a normal distribution with mean μ and standard deviation σ	$\bar{y}$ is normally distributed with mean μ and standard deviation $\sqrt{\sigma^2/n}$
The **standard error of a mean** (sem)	sem = $\sqrt{\sigma^2/n}$ or $\sigma/\sqrt{n}$
Distribution of the **difference between two sample means** of n_1, n_2 data values (resp.) from normal distributions with means μ_1 and μ_2 and standard deviations σ_1 and σ_2	$\bar{y}_1 - \bar{y}_2$ is normally distributed with mean μ_1-μ_2 and standard deviation $\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$
The **standard error of a difference** between two means (sed)	sed $= \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$ $= \sqrt{\sigma^2(1/n_1 + 1/n_2)}$ when $\sigma_1 = \sigma_2$ $= \sqrt{2\sigma^2/n}$ when $\sigma_1 = \sigma_2$ and $n_1 = n_2$
The sample variance of Y_1, Y_2, … Y_n, defined as s^2, estimates σ^2	$s^2 = \dfrac{\sum_{i=1}^{n}(Y_i - \bar{y})^2}{n-1}$
The sample variance of $\bar{y}_1, \ldots, \bar{y}_t$ estimates σ^2/n	providing each mean comes from the same numbers of replicates from a common distribution

In experimental work, one almost never knows the true population variance σ^2, and hence it needs to be estimated. This affects the distribution used in analysing experimental data.

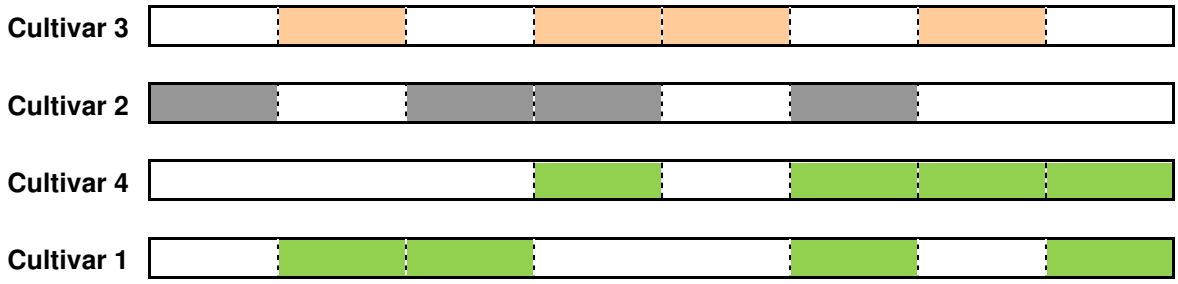
One-sample test statistic (we are usually interested in $\mu_1 = 0$)	$t = \dfrac{\bar{y}_1 - \mu_1}{\sqrt{s_1^2/n_1}} = \dfrac{\bar{y}_1 - \mu_1}{sem}$, $df = n\text{-}1$
Two-sample test statistics (we are usually interested in μ_1 - $\mu_2 = 0$). When we are happy to assume $\sigma_1^2 = \sigma_2^2$ we use a *pooled* estimate of variance obtained as a weighted variance with *df* as weights: $s_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1-1)+(n_2-1)}$	$t = \dfrac{(\bar{y}_1 - \bar{y}_2) - (\mu_1 - \mu_2)}{sed}$, where $sed = \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$ if $\sigma_1^2 \neq \sigma_2^2$, *df* complex $\sqrt{s_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$ if $\sigma_1^2 = \sigma_2^2$, $df = (n_1-1)+(n_2-1)$
95% confidence interval for μ	$\bar{y}_1 \pm t_{crit} sem$
95% confidence interval for μ_1 - μ_2	$(\bar{y}_1 - \bar{y}_2) \pm t_{crit} sed = (\bar{y}_1 - \bar{y}_2) \pm lsd$ where $lsd = t_{crit} sed$ is known as the "least significant difference"

For more complex analyses the estimate of variance used is based on the appropriate stratum variance (with appropriate degrees of freedom).

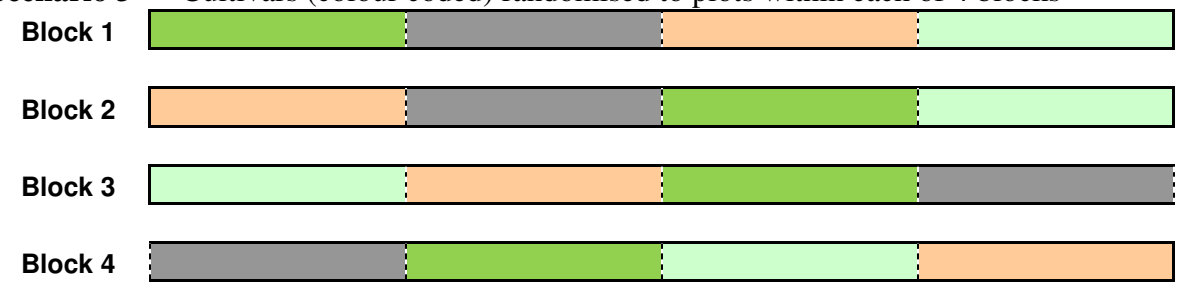
Various experimental scenarios

Scenario 1 Cultivars randomised to demonstration plots

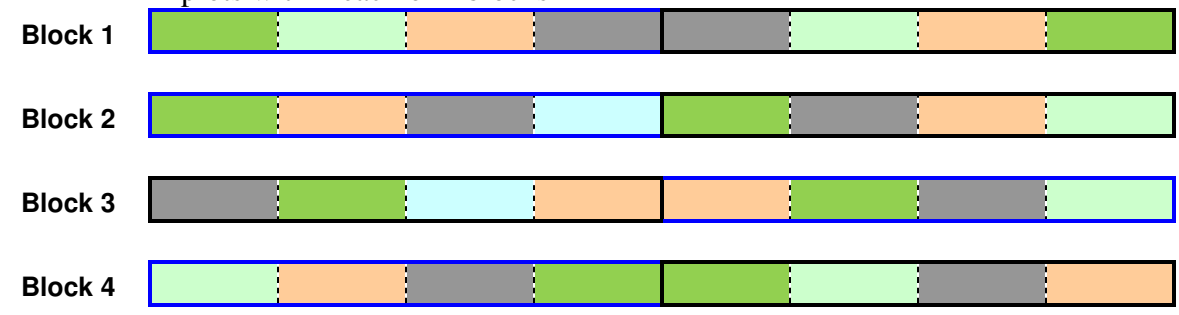
Scenario 2 Cultivars randomised to demonstration plots, 4 random grid samples taken in each



Scenario 3 Cultivars (colour coded) randomised to plots within each of 4 blocks



Scenario 4 A different method of **cultivation** (borders colour coded blue/black) is chosen at random to half of each block, then cultivars (colour coded) randomised to plots within each of 4 blocks



Appendix 2 Summary of basic experimental design concepts

Random sampling is important to remove bias and to allow the parameters (mean, standard deviation, and so on) of the distribution from which the sample is drawn to be estimated. The more **replicates** you can provide, the more accurate will be your estimates. How many replicates to provide is often the most difficult question to answer: as we will see, we need (a) some idea of the anticipated variation in our data, as well as (b) an understanding of how large a difference we are hoping to demonstrate, before a decision can be made. When it comes to designing an experiment, GenStat will always provide a random plan for the experiment: a "blueprint" that can be used in the field. The plan is a simple spreadsheet which we augment with the data available, and analyse by simple point and click.

Treatments can only be compared if they are properly replicated. Suppose you prepare four demonstration plots and sow out four cultivars, one in each plot (Scenario 1). You cannot then compare the yields from these plots, even if you obtain several sampling areas from each plot (Scenario 2). The cultivars are not replicated. Any differences in total yield could well be accidental location differences; there is no way of separating out the cultivar effects and the location effects.

Often you perform a number of randomisations in the field, leading to differently shaped experimental units. Treatments can only be compared using replicates of the same shape. We call these different shapes *strata*.

This leads to some basic principles.

i) An **experimental unit** is the smallest amount of experimental material that one treatment is randomised to.

ii) A **sampling unit** is the smallest amount of experimental material that is actually measured.

iii) Experimental units are used in forming tests of particular treatments. Sampling units just measure how "uniform" the experimental material is, and provide no degrees of freedom for these tests.

Basically, the way you design your experiment affects the way you analyse your data.

Scenario 3 is a properly replicated trial, with each cultivar sown out in different areas. Replicates are ¼ block shapes. Blocks form one stratum (and blocks are not replicated, so strictly cannot be tested) and plots in a block form a second stratum.

Scenario 4 is also properly replicated trial. However, the blocks (stratum 1) are first divided into two large areas (stratum 2) and different cultivation techniques applied to these two areas. Cultivars are applied to smaller plots (stratum 3) within these areas, thereby affecting the way we analyse the data, as we will see.

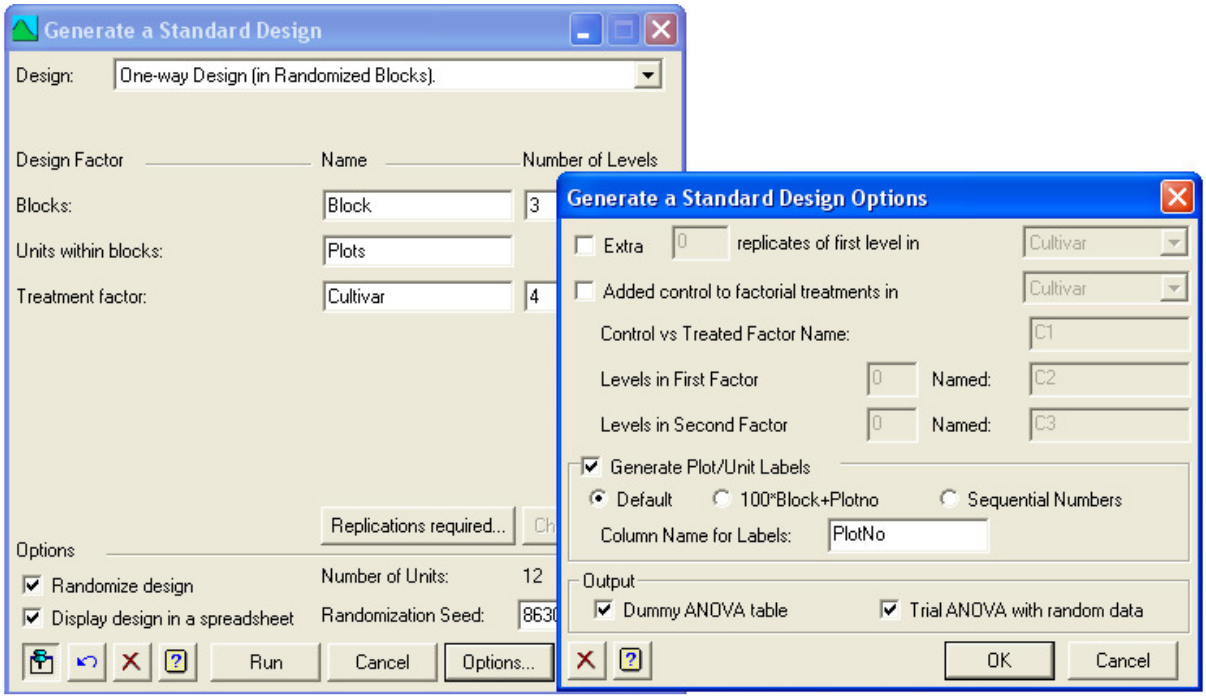
Appendix 3 GenStat's Design menu

GenStat has the ability to generate a random design for you. Most of the common designs are available, including incomplete factorial designs, and designs with additional replication for (say) a control treatment.

The design is a blueprint for conducting the experiment. It assigns the treatments to experimental units randomly. At the end of the experiment, add your data to the spreadsheet and, at least for normally or log-normally distributed data, all you need to do is point and click to have the analysis performed.

Firstly, let's illustrate the method with a simple one-way treatment design with four cultivars of oats (Vicland (1), Vicland (2), Clinton and Branch), set out in three randomized blocks in the field.

Use **Stats > Design > Generate a Standard Design**. Choose **One-way Design (in Randomized Blocks)**. Name the treatment factor and (optionally) the units to which the treatments are to be applied. Indicate the number of blocks and levels. In **Options**, you can **Trial ANOVA with random data**: this produces an analysis of random data, scaled so that the Residual MS is always 1.



GenStat creates a spreadsheet and outputs the analysis. Notice the following:

- The first column is a key to the plots in the field. The second integer is the block number, the first integer the plot number in that block. GenStat will use as many digits as required. Thus, for a design with 12 treatments in 3 blocks, the first two columns will indicate plots and the final column the block.

Analysis of variance

Variate: _Rand_

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Block stratum	2	13.973	6.986	6.99	
Block.Plots stratum					
Cultivar	3	6.448	2.149	2.15	0.195
Residual	6	6.000	**1.000**		
Total	11	26.421			

Tables of means

Variate: _Rand_

Grand mean 21.69

Cultivar	1	2	3	4
	22.94	21.05	21.36	21.41

Standard errors of means

Table	Cultivar
rep.	3
d.f.	6
e.s.e.	0.577

Least significant differences of means (5% level)

Table	Cultivar
rep.	3
d.f.	6
l.s.d.	1.998

Stratum standard errors and coefficients of variation

Variate: _Rand_

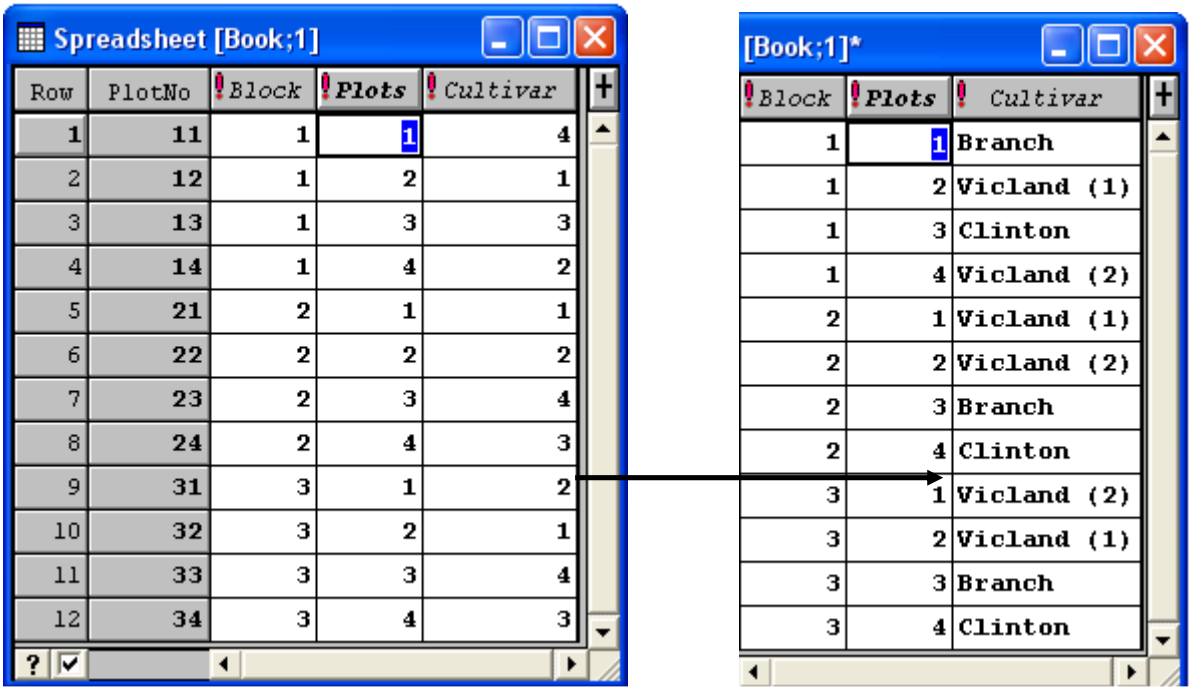
Stratum	d.f.	s.e.	cv%
Block	2	1.322	6.1
Block.Plots	6	1.000	4.6

Diagrammatic field plan

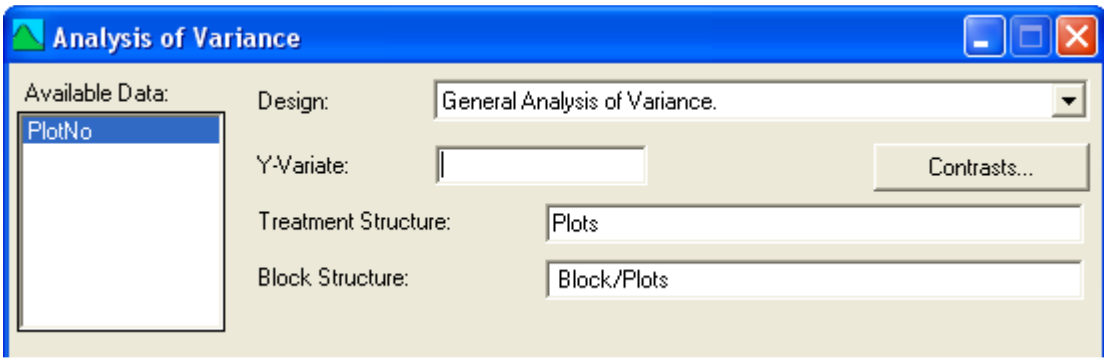
	Plot 1	**Plot 2**	**Plot 3**	**Plot 4**
Block 1	Branch	Vicland (1)	Clinton	Vicland (2)
Block 2	Vicland (1)	Vicland (2)	Branch	Clinton
Block 3	Vicland (2)	Vicland (1)	Branch	Clinton

GenStat will always generate a factor column for every stratum in the experiment. We have seen that for a block design, blocks, while unreplicated, form one stratum, and plots (which provide the replication for treatment comparisons) form the second stratum.

- The final column indicates which treatment to use in each plot in the field. This is the field plan. It is preferable at this stage to edit the column attributes (**F9** is the shortcut). In this case, change the 1, 2, 3, 4 for cultivars to their actual names. These names are then part of your statistical analysis once the data become available.



Having entered the experimental data into the spreadsheet, you can simply right click (in this example) on the PlotNo column in the spreadsheet, select **Analysis > Analysis of Variance**. The necessary structure is completed for you: your only task is to choose which variate you want analyzed this way.



The analysis will be like the one shown (which is for GenStat's random, scaled data).

Before proceeding to other designs, we need to discuss the shortcuts that GenStat uses for treatment and block structures.

Appendix 4 Overview of analysis of variance

Consider the analysis of variance for a one-way treatment design, firstly for the unblocked analysis and then for the randomized block analysis.

a) One-way treatment design, (no blocking)

ANOVA for one-way (no blocking)

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Treatment	1	81.927	81.927	11.18	0.007
Residual	11	80.584	7.326		
Total	12	162.511			

rep	7	6
mean	56.21	61.25
variance	9.015	5.299

Firstly, the sample variance of the 13 data values is 13.534. In the ANOVA table, this is the Total MS, and equals 162.511/12. GenStat does not complete this entry in the table (except in the regression menu).

The Residual SS (80.584) is the sum of squared residuals, (defined as observed – fitted). The Residual MS turns out to be the pooled variance estimate, that is, a weighted average of the individual treatment variances, with weights equal to the individual degreed of freedom of the sample variances:

$$7.326 = (6\times 9.015+5\times 5.299)/(6+5)$$

The Treatment MS is calculated as follows. Assuming common variances, if there *are* no treatment mean differences, the data from both treatments come from the same population. In that case, the i^{th} treatment mean is an estimate of σ^2/n_i. Accordingly, a weighted variance of these sample means, under the null hypothesis that the means are equal, will estimate σ^2. It also turns out that the Treatment MS and Residual MS are independent.

Thus, under the null hypothesis that the means are equal, the ratio

F = Treatment MS / Residual MS is

is distributed as an F variable with 1, 11 degrees of freedom.

For t treatments, the situation is no different. The mean squares are interpreted as follows.

To summarize:

ANOVA for one-way (no blocking)

Source of variation	**d.f.**	**m.s.**
Treatments	t-1	Weighted variance of treatment means
Residual	N-t	Pooled estimate of variance
Total	N-1	sample variance of the data

b) One-way treatment design, (in randomized blocks)

Analysis of variance

Variate: Concentration

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Head stratum	9	116.114	12.902	5.25	
Head.*Units* stratum					
Vapor_Pressure	1	592.960	592.960	241.32	<.001
Residual	9	22.115	2.457		
Total	19	731.189			

- The **Total MS** is still the sample variance of all the data. Thus, 731.189/19 = 38.484.
- The **Treatment MS** is a weighted variance of the treatment means, the weights being the number of blocks. The two vapor pressure means are 67.04 and 56.15. Each is based on 10 replicates. Thus, the **Treatment MS** is 10×sample variance of (67.04, 56.15) = 592.96.
- The **Block MS** is a weighted variance of the block means, the weights being the number of treatments. There are 10 block means, (57.1, …, 59.1) and each is based on two observations, one from each treatment. Thus, the **Block MS** is 2×sample variance of (57.1, …, 59.1) = 12.902.

The **Block SS** is still the sum of squares of the residuals. The **Block MS** is a Treatment × Block interaction: it measures the failure of the treatments to respond alike in each block.

To summarize:

ANOVA for one-way (in randomized blocks)

Source of variation	**d.f.**	**m.s.**
Blocks	b-1	Weighted variance of block means, with weights t
Treatments	t-1	Weighted variance of treatment means, with weights b
Residual	(b-1)(t-1)	Interaction between blocks and treatments
Total	bt-1	Sample variance of the data

More complex balanced designs have similar structures.

Appendix 5 Basic rules for expansion of formulae

The principle underlying a correct formulation of the blocking structure is to properly declare every type of experimental unit. For each stage of randomization a new experimental unit is created. Since the analysis exactly mimics the way the experiment is conducted in the field, a new stratum is created in the ANOVA table.

GenStat, however, allows you to omit the lowest level of randomization on the Block Structure line. If you omit the lowest level stratum in Linear Mixed Models (REML), GenStat (tells you that it) adds it to the model.

Block and treatment structures can be simplified using certain rules and operators.

Terms within parentheses are evaluated first. Otherwise, the order that GenStat uses to evaluate formulae which include operators is as follows (see GenStat Reference Manual):
1. .
2. //
3. /
4. *
5. + - -/ -*

Generally we use **. / * +** and **-**. Formulae involving a mixture of operators of rank (5) are computed left to right.

Let A, B, C … represent the names of factors and L and M a set of terms in a formula.

Rule 1	L.M	Sum of all pairwise combinations of terms in L with terms in M using the dot operator. For example: (A+B).(C+D.E) is the same as A.C + B.C + A.D.E + B.D.E
Rule 2	L*M	L+M+L.M. For example: A*C is the same as A + C + A.C (A+B)*C is the same as A + B + C + A.C + B.C
Rule 3	L/M	L+*L*.M where *L* is a term formed by combining all terms in L with the dot operator. For example: A/C is the same as A + A.C (A+B)/(C+D.E) is the same as A + B + A.B.C + A.B.D.E
Rule 4	L-M	L without any terms that appear in M. For example: (A+B)-(A+C) is the same as B A*B*C-A.B.C is the same as A+B+C+A.B+A.C+B.C

For an experiment with replication but no blocks, there should be a factor indexing the units that form replicates (plots, pots, animals, …). If there is sub-sampling within the replicate, provide an additional column to index those units. It is better to use Plot 1, 2, 3, … *p* rather than Treatment 1 (Plot 1, 2, 3), Treatment 2 (Plot 1, 2, 3) and so on. The **Block Structure** for this design can be left blank (as mentioned in paragraph 2 above), or written as **Plot** with the first method of indexing plots, or **Treatment.Plot** with the second. For the **Random Model:** in Linear Mixed Models (REML), there is an occasional advantage one way or another.

Appendix 6 REML means in the presence of one or more missing values

Suppose we have 8 participants randomized into two groups and tracked over 4 months.

Participant	Group	Time 0	Time 1	Time 2	Time 3
1	Control	8.8	8.5	8.7	8.5
2	Control	5.4	4.9	5	5.2
3	Control	2.4	2.5	2	2.2
4	Control	5.8	5.5	5.1	4.6
5	Treated	12.9	16.5	17.2	17.5
6	Treated	3.8	8.2	8.5	8.5
7	Treated	4.6	10.3	10.8	11.2
8	Treated	3.8	9.8	10.7	11.2
		Sample means			
	Control	5.60	5.35	5.20	5.13
	Treated	6.28	**11.20**	**11.80**	**12.10**

Next, suppose that Participant 7 dropped out of the trial after Time 0. This participant had an initial value of 4.6, only a little below the group average of 6.28. The treated group means at Times 1, 2 and 3 would not be expected to be very different from the ones above, provided that Participant 7 did not respond unexpectedly. That is, if the participant in question continued to have values just a little below the averages at these times, omitting these values at Times 1, 2 and 3 would (be expected to) *increase* the means just a little at those times.

Compare what happens when these three values are omitted:

Participant	Group	Time 0	Time 1	Time 2	Time 3
1	Control	8.8	8.5	8.7	8.5
2	Control	5.4	4.9	5	5.2
3	Control	2.4	2.5	2	2.2
4	Control	5.8	5.5	5.1	4.6
5	Treated	12.9	16.5	17.2	17.5
6	Treated	3.8	8.2	8.5	8.5
7	Treated	4.6			
8	Treated	3.8	9.8	10.7	11.2

		Sample means			
	Control	5.60	5.35	5.20	5.13
	Treated	6.28	**11.50**	**12.13**	**12.40**

This is a simple repeated measures analysis, with each participant having repeated measures at 4 times. We used a Linear Mixed Model (Residual Maximum Likelihood) analysis in GenStat - we refer to this analysis as LMM (REML). We allowed the variance to change over

time, and allowed for repeated data being correlated in an autoregressive order 1 (AR1) time series - a power model when the times are unequally spaced.

What does such a LMM (REML) analysis produce? Here are the sample and REML means with Participant 7 dropping out of the trial after Time 0:

Participant		Original sample means			
1	Control	5.60	5.35	5.20	5.13
2	Treated	6.28	**11.20**	**11.80**	**12.10**
3	Sample means with 3 missing values				
4	Control	5.60	5.35	5.20	5.13
5	Treated	6.28	**11.50**	**12.13**	**12.40**
6	LMM (REML) means				
7	Control	5.60	5.35	5.20	5.13
8	Treated	6.28	**11.02**	**11.64**	**11.91**

Means are slightly high in comparison to the original (known) sample means

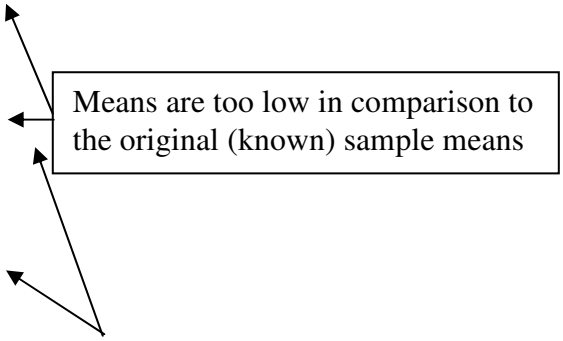
You can see that the sample means with 3 missing values *are adjusted downwards* for the treated group at times 1, 2 and 3, and are closer to what the original means were for the complete set of data.

Next suppose that Participant 5 dropped out of the trial after Time 0. This participant had an initial value of 12.9, a long way *above* the group average of 6.28. The treated group means at Times 1, 2 and 3 would therefore be expected to be very different from the original sample means, provided that Participant 5 did not respond unexpectedly. Since the participant had an initial pressure a long way above the average, omitting his values at Times 1, 2 and 3 would (be expected to) *lower* the means radically at those times. They would be very biased estimates of the true means, since the "worst" performing participant is excluded at those times.

Compare what happens when these three values are omitted, and what happens when we use a LMM (REML) analysis as described above:

Participant	Group	Time 0	Time 1	Time 2	Time 3
1	Control	8.8	8.5	8.7	8.5
2	Control	5.4	4.9	5	5.2
3	Control	2.4	2.5	2	2.2
4	Control	5.8	5.5	5.1	4.6
5	Treated	12.9			
6	Treated	3.8	8.2	8.5	8.5
7	Treated	4.6	10.3	10.8	11.2
8	Treated	3.8	9.8	10.7	11.2

Original sample means					
	Control	5.60	5.35	5.20	5.13
	Treated	6.28	**11.20**	**11.80**	**12.10**
Sample means with 3 missing values					
	Control	5.60	5.35	5.20	5.13
	Treated	6.28	**9.43**	**10.00**	**10.30**
LMM (REML) means					
	Control	5.60	5.35	5.20	5.13
	Treated	6.28	**11.54**	**12.26**	**12.47**



You can see that the sample means with 3 missing values *are adjusted upwards*, and by a long way, for the treated group at times 1, 2 and 3, and are closer to what the original means were for the complete set of data.