



# ASReml workshop

## 1.0 Course CD – Installation

Arthur Gilmour



NSW DEPARTMENT OF  
**PRIMARY INDUSTRIES**



# asrwin110.exe

- Run this first. It uses InstallShield to set up the ASReml directory structure and file associations.
- Installs ASReml 1.10 as  
`C:\Program Files\ASReml\bin\asrem1.e`
- I suggest you rename `asrem1.exe` to `asrem1110a.exe`



# bin directory

---

- contains `asrem1110h.exe` and `asrem1162zb.exe`
- Copy the directory contents to `C:\Program Files\ASReml\bin`
- This is the latest builds of 1.10 and and 1.62.
- Copy `asrem1162zb.exe` as `asrem1.exe` so that it is the default version.



# Documentation

- There are two ASReml documents. The User Guide relates to version 1.00. The Reference Manual has been updated to reflect 1.62 but is not as easy to use.
- Place these in the `Doc` directory
- Copy the workshop notes into a `Workshop` directory
- Copy the Exercises into an `Exercises` directory.



# WinEdt or ConText

---

- ConText is simpler to set up and free
- WinEdt allows reviewing Postscript graph files if GSVIEW is also installed.



# Context

---

- <http://www.context.cx/>
- Run ContextSetup.exe from the CD  
Copy ASRem1.ch1 to  
C:\Program Files\Context\Highlighters

# Context: Attach ASReml

- <Options> <Environment Options>  
<Execute keys>

<Add> file extentions .as, .asr

<F9>Execute

```
"C:\Program files\ASReml\bin\asrem1.exe"
```

```
Start In %p, Parameters %f, Hint Asrem1
```

<Add> file extentions .pin

<F9>Execute

```
"C:\Program files\ASReml\bin\asrem1.exe"
```

```
Start In %p, Parameters -p %f, Hint
```

```
PIN file
```



# Attach UserGuide

- <F12>Execute <ADOBE ACROBAT PATH>  
Start In %p, Parameters  
"C:\Program files\ASReml\doc\UserGuide.pdf"
- Find <ADOBE ACROBAT PATH> by right clicking the ADOBE READER icon and selecting <Properties>





# WinEdt

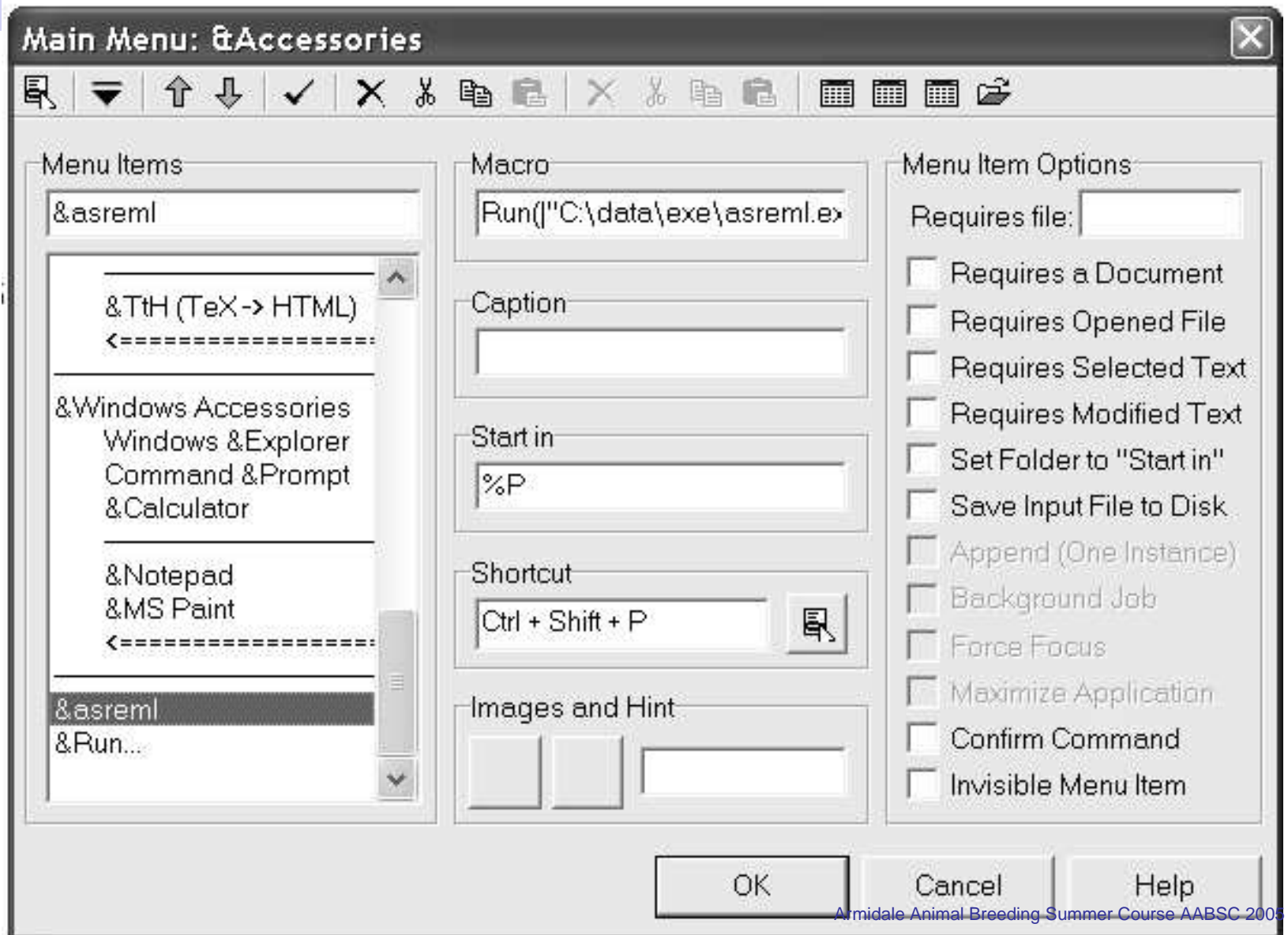
- WinEdt is a Third Party Shareware Editor available from <http://www.winedt.com> especially suited to Windows XP.
- 30 day Demo license; thence USD40
- Teach it to run ASReml:  
<Options> <Menu Setup>  
DoubleClick <&Accessories> in 'popups' list  
Cursor down to Run and Left Click  
RightClick to **insert a new Macro**



# Fill in fields

- Menu Item: &ASReml
- Macro: `Run( | "C:\Program Files\ASReml\  
bin\ASReml.exe" "%N".as, |, "%P" );`
- Start In: %P
- ShortCut: 'Ctrl+Shft+A' (or P or Z)
- Allocate a button and add ASReml to the toolbar.
- Build a submenu to invoke ASReml different ways

# Example





# ASReml workshop

---

## 1.1 Getting Started with ASReml

Arthur Gilmour



NSW DEPARTMENT OF  
**PRIMARY INDUSTRIES**



# Why start? Why am I here?

- You understand the principles of linear mixed models
- You want a flexible and comprehensive program to fit them.
- You are paying. What do you expect?



# The focus of this workshop

---

- Principles for using ASReml.
- Animal breeding applications?
- Exploring variance modelling issues?



# Today

---

- Mechanics of running ASReml
- Basic principles



# Rest of Week

---

- Animal models
- Multivariate Modelling
- Repeated measures
- Prediction and Testing





# Getting ASReml

- <http://www.asreml.co.uk>
- Windows version has automatic 30day demo license.
- Other versions: license available on request.  
<mailto:asreml@asreml.co.uk>
- Current release version 1.10h
- Workshop version 1.62zb from  
<http://www.asreml.co.uk/download-beta>
- 1.62 is a beta version: Release 2 due in 2005.



# Licensing

- VSN-International
- Support sold separately



# Documentation

---

- User Guide (1.00)
- Reference manual (1.62)
- Examples
- Cookbook



# Installation (Windows)

---

- InstallShield
- C:\Program Files\ASReml  
bin contains program  
doc contains pdf manual  
examples contains examples



# A batch process

---

- ASReml is not a typical Windows application
- Run from Windows Explorer or Command Prompt (DOS Box) or customized WinEdt/ConText
- There is a basic Menu mode for simple jobs.



# How it works.

---

- Identify problem
- Collect and organise data: save as ASCII (.txt, .csv, .asd) file
- Prepare .as job file (Notepad, WinEdt, ConText, TextEdit)
- Run, Review, Revise, Rerun cycle
- Report



# Zinc example

---

- View Data
- Fit oneway analysis of variance to SeedZn
- View running
- View output



# Zinc data

---

- Source SeedZn LeafZn

1	61	24.1
1	63	23.8
2	51	16.0
2	64	19.0
6	69	22.6





# Zinc.as file

---

- Zinc concentration study

# Zinc.as file

- Zinc concentration study  
Source \* SeedZn LeafZn

- cf

	Source	SeedZn	LeafZn
--	--------	--------	--------

1	61	24.1
---	----	------

1	63	23.8
---	----	------

2	51	16.0
---	----	------

2	64	19.0
---	----	------

6	69	22.6
---	----	------



# Zinc.as file

---

- Zinc concentration study  
Source \* SeedZn LeafZn  
ZINC.DAT !Skip 1



# Zinc.as file

---

- Zinc concentration study  
Source \* SeedZn LeafZn  
ZINC.DAT !Skip 1  
SeedZn ~ mu Source



# Zinc.as file

- Zinc concentration study

```
Source * SeedZn LeafZn
```

```
ZINC.DAT !Skip 1
```

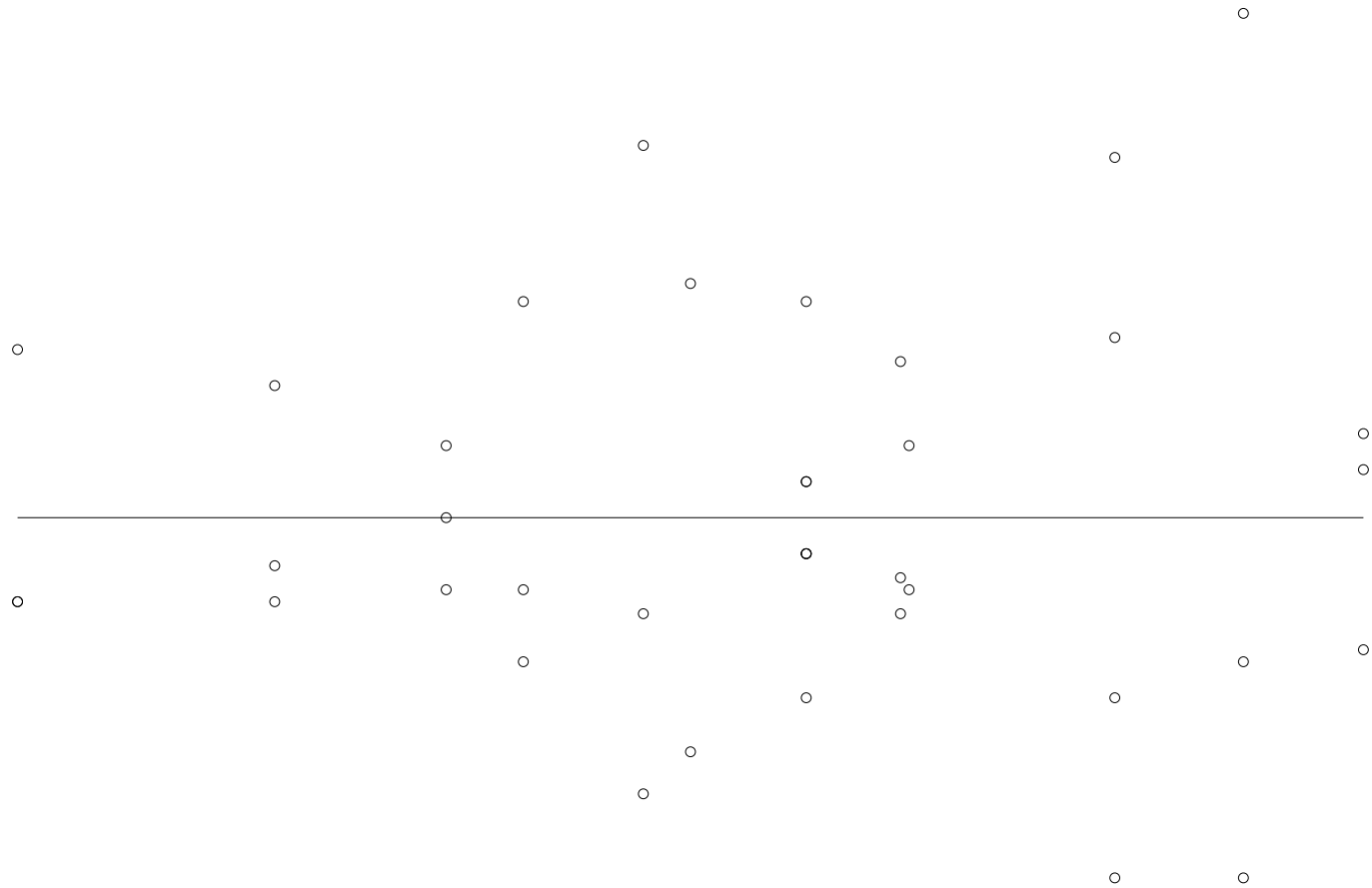
```
SeedZn ~ mu Source
```

- Run the job from command prompt, Explorer or WinEdt.

```
"C:\Program Files\ASReml\bin\ASReml.exe" ZINC
```

# Automatic plot of residuals

Zinc concentration study    Residuals vs Fitted values  
Residuals (Y)-10.00:14.00    Fitted values (X)    31.33:    83.67



# View zinc.asr



```
ASReml 1.63i [07 Jan 2005] Zinc concentration  
17 Jan 2005 12:23:06.591 32.00 Mbyte Windows
```

```
This Beta version is valid for 5 months from
```

```
Please send comments/queries to asrem
```

```
Licensed to: Arthur Gilmour
```

```
*****
```

```
* SYNTAX change: A/B now means A A.B
```

```
*
```

```
* Contact asreml@vsn-intl.com for licensing and
```

```
*****
```

# Data summary

■ Folder: C:\data\Prosper\Armidale2005

QUALIFIERS: !SKIP 1

Reading ZINC.DAT FREE FORMAT skipping 1 lines

Univariate analysis of SeedZn

Using 39 records of 39 read

Model term	Size	#miss	#zero	MinNon0	Mean	MaxNo
1 Source	24	0	0	1	11.9487	24
2 SeedZn	Variate	0	0	29.00	60.10	93.00
3 LeafZn		0	0	10.20	19.84	35.40
4 mu	1					



# Iteration sequence

- Forming 25 equations: 25 dense.

Initial updates will be shrunk by factor 0.316

NOTICE: 10 singularities detected.

1	LogL=-64.4770	S2=	44.160	24	df	1.000
---	---------------	-----	--------	----	----	-------

2	LogL=-64.4770	S2=	44.160	24	df	1.000
---	---------------	-----	--------	----	----	-------

Final parameter values						1.000
------------------------	--	--	--	--	--	-------

# Result summary

- Degrees of Freedom and Stratum Variances

	24.00	44.1597		1.0	
Source	Model terms	Gamma	Component	Comp/SE	%
Variance	39	241.00000	44.1597	3.46	0
Analysis of Var	NumDF	DenDF	F-incr	Prob	
4 mu	1	24.0	3190.25	<.001	
1 Source	14	24.0	12.94	<.001	

WARNING: The DenDF values are calculated ignoring  
variance parameters and may change

SLOPES FOR LOG(ABS(RES)) on LOG(PV) for Section  
0.50

Finished: 17 Jan 2005 12:23:08.854 LogL Conve



# ASReml workshop

---

## 1.2 Overview

Arthur Gilmour



NSW DEPARTMENT OF  
**PRIMARY INDUSTRIES**



# A Structured .as file

- First part defines the data
- Second part defines the analysis
- A minimal job has 4 lines
- Many jobs have over 20 lines
- File is built up in stages



# Definition part

---

- [ Job qualifiers ] (command line options)
- Job Title
- Data Definition
- [ Pedigree and GIV Files ]
- Data file name and qualifiers  
Zinc data analysis  
Source \* SeedZn LeafZn  
ZINC.DAT !SKIP 1



# Analysis part

---

- [ Analysis qualifiers ]
- [ TABULATE ]
- Model line  
SeedZn ~ mu Source
- [ PREDICT ]
- [ Variance structures ]
- [ Component constraints ]



# Job qualifiers

- command line options

ASReml -<options> <jobname> <arguments>

- command line not easily modified under Windows

- First line of job

- Recognised by ! character

- !-<options> <arguments> or  
<qualifiers>



# Qualifier SYNTAX

---

- First character is !
- Three letters sufficient
- separate from arguments with a space
- Context specific

- Examples

!SKIP 1

!CONTINUE !EPS !WORKSPACE 512





# Common Job qualifiers

- !CONTINUE – Use parameter estimates from a previous run as starting values
- !FINAL – One more iteration
- !LOGFILE – write .asl file
- !EPS – graphics to .eps file
- !HARDCOPY – do not write graphics to screen
- !NOGRAPHICS – do not create graphics at all
- !WORKSPACE 512 – use 512 Mb workspace
- !REPEAT  $r$  – rerun job with arguments

# Job control continued

- M (Menu mode) and P (Pinfile mode) must be specified from the command line

```
ASReml -M zinc
```

```
ASReml -Pmyjob mypin
```

- Command line options and arguments override qualifiers and arguments on the job control line.

# Arguments

- A way of resetting options within a job
- Are inserted into the job where  $\$n$  appears:  
\$1 is replaced by the first argument  
\$2 is replaced by the second argument
- With `!REPEAT  $n$` , the first  $n$  arguments are built into the output filename, and, the job is run repeatedly after moving up any arguments after the  $n$ th

```
ASRem1 -r2 job alpha beta gamma
```

```
ASRem1 -r2 job alpha beta
```

```
ASRem1 -r2 job alpha gamma
```



# Title Line

---

- identifies the job
- must be present



# Comments

---

- On all lines, characters following # are stripped out
- Comment lines (a ! in column 1 followed by a space) are copied to the output file.
- Line length is 2000 characters in 1.62
- Reserved characters: #, !, \$



# Data definition

- Controls reading the data file and how the data fields are used in the analysis.
- definitions should appear in the order of the data in the file
- definition lines should be indented
- transformations may alter the fields; the label will apply to the transformed field.
- all data is held as real numbers.

# Basic definitions

- label (co)variate
- label \* simple factor coded 1 2 ...
- label !A [ $n$ ] alphabetically coded factor
- label !I [ $n$ ] numerically coded factor
- label !G  $n$  group of  $n$  variates
- label !P pedigree factor
- label !L list simple factor; levels named  
SEX !L male female 1 codes for male



# Zinc example

- `SeedZn` is interpreted as a (co)variate
- `Source *` generates 24 levels 1:2, 5:8, 11:14, 17:19 21, 24
- `Source 24` generates 24 levels 1:2, 5:8, 11:14, 17:19 21, 24





# Recoding levels

- Source !I generates 15 levels labelled 1:2, 5:8, 11:14, 17:19 21, 24
- Source !A generates 15 levels labelled 1:2, 5:8, 11:14, 17:19 21, 24
- Specify an indication of how many levels are expected after !I and !A if there are many (> 1000) levels



# Data qualifiers

- **!LL**  $n$  – sets character length of alphabetic labels
- **!SORT** – puts labels in alphabetic/numeric order (current term)
- **!SORTALL** – puts labels in alphabetic/numeric order (current and subsequent terms)
- **!SKIP**  $s$  – to skip  $s$  fields



# Transformations

---

- applied in order of definition.
- !D  $v$  – discards records with MV or  $v$  in current field
- !M  $v$  – converts values of  $v$  to missing values in current field
- yield !\*100

# ASReml workshop

## 1.3 Pedigree, giv and data files

Arthur Gilmour



NSW DEPARTMENT OF  
**PRIMARY INDUSTRIES**

# Pedigree file

- Expected if !P data qualifier specified
- contains ID SIREID DAMID
- in birth order (parents before progeny (see !SORT))
- !ALPHA !SKIP  $k$  !DIAG !GIV !INBRED !MGS  
!REPEAT !SELF !SORT
- e.g.  
`mydata.ped !skip 1 !diag`



# GIV files

- Generalized InVerse
- Has file extension `.giv` (`.grm` if not inverted)
- Lower triangle rowwise sparse format  
row column value
- `!SKIP  $s$`



# DataFile Line

- names the data file (enclosed in quotes if embedded blanks)
- Data file is typically an ASCII file  
TAB, SPACE or COMMA separated  
e.g. save from Excel as comma separated
- Missing values: \*, . and NA and empty fields in .csv file are taken as missing
- !SKIP *s* !FILTER *f* !SELECT *v* !SUMMARY
- zinc.dat !skip 1 !SUM

# DataFile Line Qualifiers

- Some 40 qualifiers are defined for this line or to immediately follow this line.

!MAXIT  $m$  !EPS !CONTINUE

!X  $x$  !Y  $y$  !JOIN !G  $g$

!CONTRAST  $t$   $f$  coefficients

!PVAL  $f$  points

!SPLINE  $t$  points



# ASReml workshop

## 1.4 Model line

Arthur Gilmour



NSW DEPARTMENT OF  
**PRIMARY INDUSTRIES**

# TABULATE directive

## Raw tables of means

- $y \sim$  factors
- Qualifiers
  - !COUNT – numbers in each cell
  - !SD – Standard deviation in each cell
  - !RANGE – of values in each cell
  - !STATS – same as !COUNT !RANGE !SD
- before model in 1.62, after in 1.10, 1.62
- `TABULATE Leaf Seed ~ Source !STATS`
- Multiple statements allowed



# Model line

- Univariate

$y \sim \langle \text{fixed dense} \rangle,$   
 $!r \langle \text{random sparse} \rangle,$   
 $!f \langle \text{fixed sparse} \rangle$

- $y$  is response variable

$\langle \text{fixed dense} \rangle$  terms appear in ANOVA table  
 $\langle \text{random sparse} \rangle$  and  $\langle \text{fixed sparse} \rangle$  are  
reordered to maximize sparsity during  
solution.



# Model terms

- Reserved terms
  - `mu` – constant term
  - `mv` – missing value estimates
  - `units` – extra residual
- Data terms e.g. `A B X Sex Treatment`
- Functions of terms
  - `at(Group,1)` `spl(X,10)` `fac(X)`
  - `log(X,1)` `forms log(X+1)`

# Model terms continued

- Combinations

`A.B Sex.sp1(X,5) at(site,3,5).row`

- Shorthand

`A*B – A B A.B`

`A/B – A A.B`

- Continuation of a model line is indicated by a trailing comma



# Random terms

---

- May be followed by an initial value for the variance component and a qualifier.
- Default initial value is 0.1
- `blocks 0.2 !GU`
  - !GP – force positive (default)
  - !GU – unrestricted
  - !GF – fixed



# PREDICT

---

- Multiple predict statements  
`predict A`
- More details later
- Variance structure lines
- See later



# Order of processing

---

- Read .as file down to model
- Read data file
- Produce data summaries
- Do Plots and tabulations
- Read Variance structure lines
- Fit model
- Report results





# GIGO

- Check ASReml has read the data correctly
  1. Number of records read/retained
  2. Mean and range of variables
  3. Distribution of data
    - !SUM summary
    - !X !Y plot
    - TABULATE



# ASReml workshop

## 1.5 Exercises

Arthur Gilmour



NSW DEPARTMENT OF  
**PRIMARY INDUSTRIES**



# Zinc data

---

- Use !SUM to explore data structure
- Define Source using !I
- Use !X !Y to plot SeedZn vs LeafZn
- Use sqrt() to transform to Square roots
- Fit Leaf ~ mu Seed !r Source



# Volts data

---

- User Guide 15.3
- Identify outliers - assess effect of dropping two.



# Oats data

---

- User Guide 15.1
- Split Plot design: Blocks|Variety|Nitrogen
- Use !CONTRAST to test for linear N trend



# Rats data

---

- User Guide 15.2



# Own data

- Prepare job to read and summarize the data