Supplementary File 4: *Tutorial for WOMBAT*

Software notes

WOMBAT is a slightly less flexible but free alternative to ASReml, written by Karin Meyer (2006). It can be downloaded via http://agbu.une.edu.au/~kmeyer/wombat.html. 32-bit and 64-bit Linux executables, capable of analysing complex models and large amounts of data, as well as a less powerful and efficient Windows version, are available. As is the case for ASReml, users will require a simple text editor (e.g. TextPad or ConTEXT) to generate the input files. A manual covering all the basics, as well as a number of more advanced options can be downloaded at http://agbu.une.edu.au/~kmeyer/download.php?file=WombatManual.pdf. Additional information related to running WOMBAT and REML estimation in general can also be found at the WOMBAT Wiki (http://agbu.une.edu.au/~kmeyer/dokuwiki/doku.php).

Reference: Meyer, K. (2006). WOMBAT – Digging deep for quantitative genetic analyses by restricted maximum likelihood. Proc. 8th World Congr. Genet. Appl. Livest. Prod., Communication. 27: 4.

Getting started

Installing and executing WOMBAT is straightforward and is covered in detail in Chapter 3 of the manual. While we strongly encourage you to read the manual before starting this tutorial, we here provide a short overview of how to get up and running with the Windows version of WOMBAT (also see section 3.1.2 of the manual).

Download the zipped program file ("wombatwin2.zip") from the WOMBAT homepage. Unzip the file and copy "wombat.exe" to a folder of your choice (e.g. *C:\Documents and Settings\MyName\My Documents\Wombat*). Copy the pedigree file ("gryphon.ped_WOMABT") and the four data files ("gryphon_uni.dat", "gryphon_bi.dat", "gryphonRM.dat" and "gryphonRM_pe.dat") for this tutorial into the same folder. The pedigree file provided is so named to ensure the correct version is downloaded. Once you have it in the appropriate folder it should be renamed as simply "gryphon.ped".

In this tutorial we will assume that wombat.exe is located in the same folder as your pedigree and data files. Alternatively you may choose to have WOMBAT in its own folder, for example in C:\Program Files\wombat, and keep it separate from any data files. In this case you need to add this directory to the list of folders that Windows automatically searches for executable files (listed in the so-called PATH environment variable). How to do this can be found in the WOMBAT manual or in the Windows documentation.

WOMBAT is run from the command prompt, which is something some users may be unfamiliar with. To open a command line window, click on Start → Run and type cmd. Alternatively, you may try to locate the shortcut to the Command Prompt, which is probably hiding under Start → Programs → Accessories. You will now see a black window, with some text indicating in which folder you are located (e.g. *C:\Documents*). You should now navigate to the folder containing wombat.exe and your input-files. You can see the contents of the current folder by entering dir (short for directory). To move one folder down in the hierarchy, enter cd (short for change directory) and the name of the folder, so in our case cd My Documents, followed by cd Wombat. Enter cd.. if you need to go one folder up in the hierarchy. If you need to go to a different disk drive (for example a USB-stick that got assigned the letter E), simply enter E:. If

you are located in the folder containing all the necessary files (check with `dir`), you are ready to go!

Tutorial 1 (WOMBAT) - Estimating the heritability of birth weight

This tutorial will demonstrate how to run a univariate animal model using the software WOMBAT and the example data files provided.

1) Scenario
In a population of gryphons there is strong positive selection on birth weight, with heavier born individuals having on average higher fitness. To find out whether increased birth weight will evolve in response to this selection, and if so, how quickly, we want to estimate the heritability of birth weight.

2) Data files
Open `gryphon.ped` and `gryphon_uni.dat` in your text editor. The structure and contents of these files are fairly self-explanatory. The pedigree file `gryphon.ped` contains three columns containing unique numerical IDs that correspond to each animal, its father, and its mother. Note that this is a multigenerational pedigree, with the earliest generation (for which parentage information is necessarily missing) at the beginning of the file. For later born individuals maternal identities are all known, but paternity information is incomplete (a common situation in real world applications). Unknown parents are coded as 0. An important requirement for the WOMBAT pedigree file is that only numerical identities are allowed and that offspring must have a higher number than their parents.

The phenotype data, as well as additional factors and covariates that you may wish to include in your model are contained in `gryphon_uni.dat`. Columns correspond to individual identity (ANIMAL), maternal identity (MOTHER), year of birth (BYEAR), SEX (1: female, 2: male) and birth weight (BWT). Unlike in ASReml, in WOMBAT each row of the data file contains a record for only one trait. When data for multiple traits is available, each trait comes on a new line (see Tutorial 2). As is the case for the pedigree file, only numbers are allowed. Missing values are not allowed in the data file in WOMBAT.

3) Running the model
To run an analysis, write the parameter file below using your text editor. We suggest saving it as `gryphon1.par` (WOMBAT requires the extension `.par` for the parameter file). Make sure that `gryphon1.par` is located in the same folder as `wombat.exe`, `gryphon.ped` and `gryphon_uni.dat`. Note also that any text after # (which has to be in column 1 of the line!) is ignored and we have taken advantage of this to annotate the example below and highlight some key features.

A WOMBAT parameter file has the following sections: comment line, analysis type, name (and location) of the pedigree file, name (and location) of the data file, data specification, animal model specification and the starting values.

```
# A single optional comment line of up to 74 characters
COMMENT WOMBAT analysis of Gryphon birth weight

# The type of analysis to be performed. In this case a UNIvariate analysis
ANALYSIS UNI

# Name the pedigree file, assuming it is in the same folder as the parameter file
PEDS gryphon.ped

# Name the data file, assuming it is in the same folder as the parameter file
DATA gryphon_uni.dat

# Description of the data file
# Each variable to be fitted in the model needs to be followed by
# the maximum number of levels
  ANIMAL 900
  MOTHER 400
  BYEAR 35
  SEX 2
  BWT
END DATA

# Model specification
# Type of effect (FIXed, COVariate, RANdom) and variable name
# NRM indicates that a pedigree is available for ANIMAL
MODEL
  RAN ANIMAL NRM
# The trait to be analysed
  TRAIT BWT
END

# Specify starting values and the number of rows and columns in the matrix
# These are not very important here
VAR residual 1
1.0
VAR ANIMAL 1
1.0
```

To run this analysis you now have to switch to the command prompt and navigate to the folder containing `gryphon1.par` (see *Getting started*). This fits a simple univariate animal model with a single fixed effect (the mean) and a single random effect (the additive genetic effect). To run the model as specified in the parameter file, type `wombat`, followed by the name of the parameter file. So in this case you enter `wombat gryphon1.par`. This will produce a large number of files, but in this tutorial we are mainly interested in `SumEstimates.out` which contains the most important results. Open this file in your text editor:

```
======= Version 17-03-2009 ======================================== **KM** ====

        Program WOMBAT : Estimates of covariance components
================================================================================

 WOMBAT analysis of Gryphon birth weight

 Analysis type        :    "UNI"
 Data file            :    "gryphon_uni.dat"
 Pedigree file        :    "gryphon.ped"
 Parameter file       :    "gryphon1.par"


 No. of traits        =       1                   BWT
 No. of records       =     854              854
 No. of parameters    =       2
 Maximum log L        =          -1246.462
 -1/2 AIC & AICC      =          -1248.462            -1248.469
 -1/2 BIC             =          -1253.212     "Penalty factor" =   3.375

 Parameter estimates with approx. sampling erors
```

```
    1    CHOL Z 1 1                  1.99765              0.129761
    2    CHOL A 1 1                  1.79052              0.173784


 Convergence criteria for last 3 iterates
 Change in log likelihood      =    12.609370     0.279775     0.000180
 Change in parameter vector    =     0.094751     0.017635     0.000462
 Norm of gradient vector       =       95.7971      11.9300       0.2945
 Newton decrement              =      -23.8956      -0.5535      -0.0004

***** Estimates of residual covariances   *********************************
      Order of fit           =         1
 Covariance matrix
    1     3.9906
 Matrix of correlations and variance ratios
    1      0.5545
 Covariances & correlations with approximate sampling errors
    1  COVS Z 1 1            3.99061         0.518435        vrat     0.555    0.076

***** Estimates for RE  1    "ANIMAL"    *********************************
      No. of levels          =      1084
      Covariance structure =      NRM
      Order of fit           =         1
 Covariance matrix
    1     3.2059
 Matrix of correlations and variance ratios
    1      0.4455
 Covariances & correlations with approximate sampling errors
    2  COVS A 1 1            3.20595         0.622327        vrat     0.445    0.076

***** Estimates of phenotypic covariances   *********************************
 Covariance matrix
    1     7.1966
 Covariances & correlations with approximate sampling errors
    3  COVS T 1 1            7.19655         0.368869
```

The first half of `SumEstimates.out` contains some information on the type of model you fitted and the names of the different input files, as well as some statistical information on the fit of the model. This information is followed by the estimates of the different variance components.
In this model we fitted a single random effect, so we have partitioned the phenotypic variance into two components, a residual (lines 31-38) and an ANIMAL (lines 40-49) component. The above output first lists the residual variance (V_R) as 3.9906, and then it lists ANIMAL random effect (V_A) which is estimated as 3.2059. Finally, it also lists the phenotypic variance (V_P) (lines 51-55), which is simply the sum of above components.

`SumEstimates.out` lists the approximate standard errors for each of these components. Given that the ratio of V_A to its standard error (0.622327) is considerable larger than 2 (i.e. the parameter estimate is more than 2 SEs from zero) V_A is probably significantly greater than zero (but see 7. Testing significance of random effects).

4) Estimating heritability
We can calculate the h^2 of birth weight from the components above since $h^2 = V_A/V_P = V_A/(V_A+V_R)$. Thus according to this model, $h^2 = 3.20595 / (3.20595 + 3.99061) = 0.445$. Conveniently, WOMBAT estimates not only the absolute variance components, but also the proportion of the phenotypic variation accounted for by the different random effects, again along with their sampling error. These proportions are reported following 'vrat' (as 0.445 ± 0.076 SE) in the above model output, after the variance component and its estimated standard error. For the animal effect this proportion is usually equal to the heritability (but see Tutorial 3).

5) Adding fixed effects

To add effects to a univariate model, simply modify the model specification in `gryphon1.par`. For example we might know (or suspect) that birth weight is a sexually dimorphic trait and therefore include the sex of the individual in the model as a fixed effect.

```
MODEL
  FIX SEX
  RAN ANIMAL NRM
  TRAIT BWT
END
```

Save these changes and run `gryphon1.par` as before. Note that this will overwrite the previous output files, so don't forget to rename or move any files you might want to keep.

Unlike ASReml, WOMBAT will not provide the significance of the fixed effects in your model. However, generalised least-squares estimates for all fixed effects fitted, together with means and the number of observations for each group, can be found in `FixSolutions.out`.

It is worth noting that the variance components in `SumEstimates.out` have changed slightly. In fact, since SEX effects were previously contributing to the residual variance of the model, our estimate of V_R (denoted "residual" in the output) is now slightly lower than before. This has an important consequence for estimating heritability since if we calculate V_P as V_A+V_R, then as we include fixed effects we will soak up more of the residual variance driving V_P. Assuming that V_A is more or less unaffected by the fixed effects fitted, then as V_P goes down we expect our estimate of h^2 to go up.

```
***** Estimates of residual covariances   ************************************
     Order of fit          =        1
 Covariance matrix
   1    2.9384
 Matrix of correlations and variance ratios
   1      0.4898
 Covariances & correlations with approximate sampling errors
   1 COVS Z 1 1         2.93843        0.416148      vrat     0.490   0.074

***** Estimates for RE  1    "ANIMAL"   ************************************
     No. of levels        =     1084
     Covariance structure =    NRM
     Order of fit          =        1
 Covariance matrix
   1    3.0604
 Matrix of correlations and variance ratios
   1      0.5102
 Covariances & correlations with approximate sampling errors
   2 COVS A 1 1         3.06042        0.524357      vrat     0.510   0.074
```

Indeed, after fitting SEX, h^2 has increased slightly from 0.45 to 0.51. Which is the better estimate? It depends on what your question is. The first is an estimate of the proportion of variance in birth weight explained by additive effects, whereas the latter is an estimate of the proportion of variance in birth weight that is explained by additive genetic effects *after conditioning on sex*.

6) Adding additional random effects
This is done by simply modifying the model statement in the same way as before:

```
MODEL
  FIX SEX
  RAN ANIMAL NRM
  RAN BYEAR
  TRAIT BWT
END
```

As we are fitting an additional random effect, we will also need to provide an additional starting value for this random effect:

```
VAR residual 1
1.0
VAR ANIMAL 1
1.0
VAR BYEAR 1
1.0
```

This results in an additional variance component reported in `SumEstimates.out`

```
***** Estimates of residual covariances   ***********************************
     Order of fit         =       1
 Covariance matrix
   1    2.3093
 Matrix of correlations and variance ratios
   1      0.3912
 Covariances & correlations with approximate sampling errors
   1 COVS Z 1 1        2.30926        0.345100     vrat     0.391    0.064

***** Estimates for RE  1   "ANIMAL"   **************************************
     No. of levels       =    1084
     Covariance structure =    NRM
     Order of fit         =       1
 Covariance matrix
   1    2.7068
 Matrix of correlations and variance ratios
   1      0.4586
 Covariances & correlations with approximate sampling errors
   2 COVS A 1 1        2.70685        0.442218     vrat     0.459    0.067

***** Estimates for RE  2   "BYEAR"   **************************************
     No. of levels       =      34
     Covariance structure =    IDE
     Order of fit         =       1
 Covariance matrix
   1    0.88626
 Matrix of correlations and variance ratios
   1      0.1502
 Covariances & correlations with approximate sampling errors
   3 COVS B 1 1        0.886260       0.269593     vrat     0.150    0.040
```

Here the variance in BWT explained by birth year is 0.886, and based on its approximate sampling error it appears to be significant (but see 7. Testing significance of random effects). Thus we would conclude that year to year variation (e.g., in climate, resource abundance) contributes to V_P. Note that although V_A has changed somewhat, most of what is now partitioned as a birth year effect was previously included in V_R. Thus what we have really done here is to partition environmental effects into differences arising from year to year differences versus everything else, and thus we do not expect much change in h^2 (as now $h^2 = V_A/ (V_A+V_{BY}+V_R)$).

However, we get a somewhat different result if we also add a random effect for MOTHER to test for maternal effects:

```
MODEL
  FIX SEX
  RAN ANIMAL NRM
  RAN BYEAR
  RAN MOTHER
  TRAIT BWT
END

VAR residual 1
1.0
VAR ANIMAL 1
1.0
```

```
VAR BYEAR 1
1.0
VAR MOTHER 1
1.0
```

This gives estimated variance components of

```
***** Estimates of residual covariances   *********************************
      Order of fit          =        1
 Covariance matrix
   1     1.6290
 Matrix of correlations and variance ratios
   1      0.2748
 Covariances & correlations with approximate sampling errors
   1  COVS Z 1 1         1.62899        0.371418     vrat    0.275    0.068

***** Estimates for RE  1    "ANIMAL"   *********************************
      No. of levels         =     1084
      Covariance structure =    NRM
      Order of fit          =        1
 Covariance matrix
   1     2.2985
 Matrix of correlations and variance ratios
   1      0.3877
 Covariances & correlations with approximate sampling errors
   2  COVS A 1 1         2.29851        0.496268     vrat    0.388    0.078

***** Estimates for RE   2    "BYEAR"   *********************************
      No. of levels         =       34
      Covariance structure =    IDE
      Order of fit          =        1
 Covariance matrix
   1    0.88204
 Matrix of correlations and variance ratios
   1      0.1488
 Covariances & correlations with approximate sampling errors
   3  COVS B 1 1        0.882044        0.263251     vrat    0.149    0.039

***** Estimates for RE   3    "MOTHER"   *********************************
      No. of levels         =      394
      Covariance structure =    IDE
      Order of fit          =        1
 Covariance matrix
   1     1.1185
 Matrix of correlations and variance ratios
   1      0.1887
 Covariances & correlations with approximate sampling errors
   4  COVS C 1 1         1.11851        0.238627     vrat    0.189    0.038
```

Here the partitioning of significant maternal variance has resulted in a further decrease in V_R, but also in a decrease in V_A. The latter is because maternal effects of the sort we simulated to create these data files (consistent differences between mothers) will increase the similarity among maternal siblings. Consequently they can look very much like additive genetic effects and, if present but unmodelled, represent a type of "common environment effect" that can - and will - cause upward bias in V_A and h^2.

7) Testing significance of random effects
A final point to note in this tutorial is that while the relative size of the sampling error reported in SumEstimates.out can be used as an indicator of likely statistical significance, these approximate sampling errors are not recommended for formal hypothesis testing. A better approach is to use likelihood ratio tests.

For example, to test the significance of maternal effects we should run models with and without maternal identity as a random effect, and compare the final log-likelihoods (`Maximum log L`) as reported in `SumEstimates.out`.

Fitting a model with MOTHER using

```
MODEL
  FIX SEX
  RAN ANIMAL NRM
  RAN BYEAR
  RAN MOTHER
  TRAIT BWT
END
```

gives `maximum log L = -1094.81`, whereas a model without MOTHER using

```
MODEL
  FIX SEX
  RAN ANIMAL NRM
  RAN BYEAR
  TRAIT BWT
END
```

gives `maximum log L = -1113.77`

A test statistic equal to twice the difference in these log-likelihoods is assumed to be distributed as chisquare with one degree of freedom. So in this case we would conclude that the maternal effects are highly significant since:

$\chi^2_{1DF} = 2\,(-1094.81 - -1113.77) = 37.92$, which corresponds to P<0.0001

If we take $\alpha = 0.05$, we would thus conclude that the inclusion of MOTHER as an additional random effect significantly improves the model.

Tutorial 2 (WOMBAT) – A bivariate animal model

This tutorial will demonstrate how to run a multivariate animal model using the software WOMBAT and the example data files provided.

1) Scenario

Since natural selection rarely acts on single traits, to understand how birth weight might evolve in our population of gryphons, we may want to test for possible correlations with other traits. If tarsus length at fledging is also under positive selection, what implications does this have for birth weight and vice versa? If the two traits are positively genetically correlated then this will facilitate the evolution of larger size (since response of one trait will induce a positively correlated response in the other). If there is negative genetic covariance then this could act an evolutionary constraint.

Multivariate models allow us to estimate parameters for each trait alone (i.e. V_A, h^2 etc), but also yield estimates of covariance components between traits. These include the (additive) genetic covariance COV_A among two traits, which is often rescaled to give the genetic correlation r_G. However, correlations among traits can also arise through other random effects (e.g. maternal effects) and these sources can be explicitly modelled in a bivariate analysis as well.

2) Data files

The pedigree file used is the same as in Tutorial 1 (`gryphon.ped`). However, we now use the phenotypic data file `gryphon_bi.dat`. This datafile contains data for two traits, with one phenotypic value per line. The additional variable TRAITNO indicates which trait the variable TRAIT refers to. The datafile has to be sorted by ANIMAL and then by TRAITNO.

Note that the data for the other variables (ANIMAL, MOTHER, BYEAR and SEX) are identical for both traits, generating a substantial amount of duplicate data. Finally, it is worth pointing out that it is in fact relatively straightforward to perform the univariate analyses from Tutorial 1 using this dataset. For more details on how to do this, see the manual.

3) Running the model

Running multivariate analyses in WOMBAT requires only slightly more complex code than a univariate analysis (Tutorial 1). To run a bivariate analysis, write the following command file using your text editor. We suggest saving this file as `gryphon2.par` (extension `.par` denoting a WOMBAT parameter file). The main differences are in the parts describing the data structure, the model structure and the starting values.

```
# A single optional comment line of up to 74 characters
COMMENT WOMBAT analysis of Gryphon birth weight and tarsus length

# The type of analysis to be performed
# In this case a MUltiVariate analysis with 2 traits
ANALYSIS MUV 2

# Name the pedigree file, assuming it is in the same folder as the parameter file
PEDS gryphon.ped

# Name the data file, assuming it is in the same folder as the parameter file
# GRP indicates that we will be using the 'compact' data specification (see manual)
DATA gryphon_bi.dat GRP
# We have two traits, with the trait numbers (TRNOS) 1 and 2
  TRNOS 1 2
# Each variable to be fitted in the model needs to be followed by
# the maximum number of levels
  TRAITNO
  ANIMAL 900
  MOTHER 400
```

```
  BYEAR 35
  SEX 2
  NAMES BWT TARSUS
END

# Model specification
# Type of effect (FIXed, COVariate, RANdom, TRAIT) and variable names
# NRM indicates that a pedigree is available for ANIMAL
MODEL
  RAN ANIMAL NRM
  TRAIT BWT 1
  TRAIT TARSUS 2
END MODEL

# Specify starting values as upper triangle of variance-covariance matrix
# and the number of rows and columns in the matrix
# It doesn't matter if the values are next to each other,
# below each other, or something in between
VAR  residual 2
1.0
0.1
1.0
VAR  ANIMAL 2
1.0 0.1 1.0
```

Note that the starting values supplied here are arbitrary. If the model is difficult to fit then it can be because the starting values are too far from the best estimates. One way around this is to run single trait models first to get good starting values for the variances (but you still have to "guess" starting values for the covariances).

Run this model and look at the primary results file SumEstimates.out. Notice that there are now six (co)variance components reported.

```
***** Estimates of residual covariances   ************************************
     Order of fit           =        2
 Covariance matrix
   1    4.0020
   2    3.3260        17.747
 Eigenvalues of covariance matrix
Value       18.51         3.24
  (%)       85.11        14.89
 Matrix of correlations and variance ratios
   1     0.5563
   2     0.3947    0.5929
 Covariances & correlations with approximate sampling errors
   1  COVS Z 1 1        4.00198        0.519972      vrat    0.556    0.076
   2  COVS Z 1 2        3.32603        0.912341      corr    0.395    0.084
   3  COVS Z 2 2        17.7471        2.66419       vrat    0.593    0.093

***** Estimates for RE  1    "ANIMAL"   ************************************
     No. of levels        =     1084
     Covariance structure =     NRM
     Order of fit         =        2
 Covariance matrix
   1    3.1917
   2    2.4412        12.188
 Eigenvalues of covariance matrix
Value       12.81         2.57
  (%)       83.28        16.72
 Matrix of correlations and variance ratios
   1     0.4437
   2     0.3914    0.4071
 Covariances & correlations with approximate sampling errors
   4  COVS A 1 1        3.19172        0.623247      vrat    0.444    0.076
   5  COVS A 1 2        2.44117        1.06093       corr    0.391    0.132
   6  COVS A 2 2        12.1881        3.05787       vrat    0.407    0.093
```

The first part relates to the residual matrix and, since BWT is our first trait and TARSUS our second, these are $V_{R.BWT}$ (4.0020), COV_R (3.3260), $V_{R.TARSUS}$ (17.747). These are followed by the corresponding genetic terms $V_{A.BWT}$ (3.1917), COV_A (2.4412), $V_{A.TARUS}$ (12.188). Note that besides h^2_{BWT} (0.444 ± 0.076) and h^2_{TARSUS} (0.407 ± 0.093), we now also get an estimate of r_G (0.391 ± 0.132). Based on our quick and dirty check (is variance component / sampling error > 2?) all components are likely to be statistically significant.

4) Adding fixed and random effects

Fixed and random effects can be added just as in the univariate example. Given that our full model of BWT from Tutorial 1 had SEX as a fixed effect, as well as random effects of BYEAR and MOTHER, we could specify a bivariate formulation of this using:

```
MODEL
  FIX SEX
  RAN ANIMAL NRM
  RAN BYEAR
  RAN MOTHER
  TRAIT BWT 1
  TRAIT TARSUS 2
END MODEL

VAR  residual 2
1.0
0.1
1.0
VAR ANIMAL 2
1.0
0.1 1.0
VAR BYEAR 2
1.0 0.1
    1.0
VAR MOTHER 2
1.0 0.1 1.0
```

Note that there will now be twelve (co)variance components reported in `Sumestimates.out`.

```
***** Estimates of residual covariances   **********************************
     Order of fit        =      2
 Covariance matrix
   1     1.8443
   2     4.0141        12.484
 Eigenvalues of covariance matrix
Value        13.83          0.50
   (%)        96.51          3.49
 Matrix of correlations and variance ratios
   1      0.3098
   2      0.8366    0.3992
 Covariances & correlations with approximate sampling errors
   1  COVS Z 1 1        1.84430        0.344317     vrat    0.310    0.063
   2  COVS Z 1 2        4.01415        0.741260     corr    0.837    0.074
   3  COVS Z 2 2        12.4841        2.28937      vrat    0.399    0.079

***** Estimates for RE  1   "ANIMAL"   *************************************
     No. of levels       =    1084
     Covariance structure =    NRM
     Order of fit        =      2
 Covariance matrix
   1     1.9894
   2     3.3172        10.230
 Eigenvalues of covariance matrix
Value        11.40          0.82
   (%)        93.29          6.71
 Matrix of correlations and variance ratios
   1      0.3342
   2      0.7353    0.3271
 Covariances & correlations with approximate sampling errors
```

```
   4   COVS A 1 1          1.98936       0.441024       vrat     0.334    0.070
   5   COVS A 1 2          3.31717       0.903262       corr     0.735    0.109
   6   COVS A 2 2         10.2298        2.80786        vrat     0.327    0.085


***** Estimates for RE  2   "BYEAR"    *************************************
       No. of levels       =        34
       Covariance structure =     IDE
       Order of fit         =        2
 Covariance matrix
    1    0.97465
    2    0.16239        3.7382
 Eigenvalues of covariance matrix
 Value        3.75          0.97
   (%)        79.52         20.48
 Matrix of correlations and variance ratios
    1      0.1637
    2      0.0851    0.1195
 Covariances & correlations with approximate sampling errors
    7   COVS B 1 1         0.974645      0.282573       vrat     0.164    0.040
    8   COVS B 1 2         0.162391      0.418491       corr     0.085    0.213
    9   COVS B 2 2         3.73821       1.20651        vrat     0.120    0.035


***** Estimates for RE  3   "MOTHER"   *************************************
       No. of levels       =       394
       Covariance structure =     IDE
       Order of fit         =        2
 Covariance matrix
    1     1.1445
    2    -1.5567        4.8212
 Eigenvalues of covariance matrix
 Value        5.39          0.57
   (%)        90.38          9.62
 Matrix of correlations and variance ratios
    1      0.1923
    2     -0.6627    0.1542
 Covariances & correlations with approximate sampling errors
   10   COVS C 1 1         1.14450       0.230217       vrat     0.192    0.037
   11   COVS C 1 2        -1.55670       0.405193       corr    -0.663    0.249
   12   COVS C 2 2         4.82117       1.32019        vrat     0.154    0.040
```

5) Significance testing

Under the model above r_G is estimated at + 0.735 and the estimate of COV_A (3.3172) is greater than twice the sampling error (0.9033). We might therefore conclude there is evidence of a strong positive genetic correlation between the traits. However, to formally test the significance of COV_A we would have to compare the log-likelihood for this model to one in which we specify that $COV_A = 0$. Unfortunately, at this moment it is not possible to constrain individual covariance components in WOMBAT. This is currently one of its main limitations.

Tutorial 3 (WOMBAT) – A repeated measures animal model

This tutorial will demonstrate how to run a univariate animal model for a trait with repeated observations using the software WOMBAT and example data files provided.

1) Scenario

Since gryphons are iteroparous multiple observations of reproductive traits are available for some individuals. Here we have repeated measures of lay date (measured in days after Jan 1) for individual females of varying age from 2 (age of maturation) up until age 6. Not all females lay every year so the number of observations per female is variable. We want to know how repeatable the trait is, and (assuming it is repeatable) how heritable it is.

2) Data files

The pedigree file `gryphon.ped` is the same as we used in the preceding tutorials, but we now use the data file `gryphonRM.dat`. Open the latter in your text editor. Columns correspond to individual identity (ANIMAL), birth year (BYEAR), age in years (AGE), year of measurement (YEAR) and lay date (LAYDATE). Each row of the data file corresponds to a single phenotypic observation. Note that data are sorted by identity and then age. This does not only make the repeated observations on individuals readily apparent, it is also an important requirement for analysis.

3) Estimating repeatability:
With repeated measures on individuals it is often of interest, prior to fitting a genetic model, to see how repeatable a trait is. We can estimate the repeatability of a trait as the proportion of phenotypic variance explained by individual identity using the parameter file below (we suggest saving this as `gryphon3.par`)

```
COMMENT WOMBAT analysis of Gryphon laydate

ANALYSIS UNI

# PEDS gryphon.ped
DATA gryphonRM.dat
  ANIMAL 900
  BYEAR 35
  AGE 5
  YEAR 35
  LAYDATE
END DATA

MODEL
  RAN ANIMAL
  TRAIT LAYDATE
END

VAR residual 1
1.0
VAR ANIMAL 1
1.0
```

Whereas before we have associated the term ANIMAL with the pedigree file (by adding NRM in the model specification), here we are interested in estimating the amount of variance explained by individual identity (rather than by additive effects). We do this by fitting ANIMAL without specifying NRM, and by removing or commenting out the section where we specify the pedigree file. The resulting partitioning of the phenotypic variance is seen in `SumEstimates.out` as

```
***** Estimates of residual covariances   **********************************
     Order of fit        =        1
 Covariance matrix
   1      21.296
```

```
Matrix of correlations and variance ratios
   1     0.6584
Covariances & correlations with approximate sampling errors
   1 COVS Z 1 1          21.2965        0.889605      vrat     0.658    0.028


***** Estimates for RE  1   "ANIMAL"   ***************************************
     No. of levels       =      469
     Covariance structure =     IDE
     Order of fit        =        1
Covariance matrix
   1     11.048
Matrix of correlations and variance ratios
   1     0.3416
Covariances & correlations with approximate sampling errors
   2 COVS A 1 1          11.0485        1.17572      vrat     0.342    0.028
```

Among-individual variance is given by the ANIMAL component, while the residual component represents within-individual variance. The repeatability of the trait (the proportion of V_P explained by ANIMAL) is $11.048/11.048 + 21.296 = 0.3416$. Given that we set up the simulation of the data such that mean lay date changes with age (initially increasing to age 5 before a late life decline) we might ask what the repeatability of lay date is after conditioning on age effects. This could be done by adding age into the model as a fixed effect, fitting a different mean lay date for each age class.

```
MODEL
  FIX AGE
  RAN ANIMAL
  TRAIT LAYDATE
END
```

Fitting the above gives

```
***** Estimates of residual covariances   ***************************************
     Order of fit        =        1
Covariance matrix
   1     16.380
Matrix of correlations and variance ratios
   1     0.5713
Covariances & correlations with approximate sampling errors
   1 COVS Z 1 1          16.3799        0.686632      vrat     0.571    0.027


***** Estimates for RE  1   "ANIMAL"   ***************************************
     No. of levels       =      469
     Covariance structure =     IDE
     Order of fit        =        1
Covariance matrix
   1     12.290
Matrix of correlations and variance ratios
   1     0.4287
Covariances & correlations with approximate sampling errors
   2 COVS A 1 1          12.2898        1.15610      vrat     0.429    0.027
```

The repeatability of lay date, after accounting for age effects, is now estimated as 0.429. So, just as we saw when estimating h^2 in Tutorial 1, the inclusion of fixed effects will alter the estimated effect size if we determine total phenotypic variance as the sum of the variance components. Thus, proper interpretation is vital.

Although here age was modelled as a 5 level factor, we could equally have fitted it as a continuous variable instead, in which case, given the late life decline, we should probably also include a quadratic term. We would do this by adding

```
MODEL
  COV AGE(2)
  RAN ANIMAL
  TRAIT LAYDATE
END
```

The (2) after AGE indicates that we include AGE as both a linear and a quadratic term. Although, as pointed out before, WOMBAT will not test for the significance of the effect of AGE, the parameter estimates for the linear and the quadratic term can be found in `FixSolutions.out`:

```
Covariables for trait no.      1    "LAYDATE"
    Covariable                     Reg.coeff    Solution
  1  AGE(2)                             1     0.958771
  1  AGE(2)                             2    -0.558404
```

4) Partitioning additive and permanent environment effects

While additive genetic effects will cause among-individual variation, so will other types of effect. Therefore we generally expect that the repeatability will set the upper limit for heritability. Non-additive contributions to consistent among-individual differences are normally referred to as "permanent environment effects". If a trait has repeated measures then it is necessary to model permanent environment effects in an animal model to prevent upward bias in V_A. To illustrate this, change `gryphon3.par` and fit the following simple animal model (don't forget to include a link to the pedigree file again):

```
PEDS gryphon.ped
DATA gryphonRM.dat
  ANIMAL 900
  BYEAR 35
  AGE 5
  YEAR 35
  LAYDATE
END DATA

MODEL
  RAN ANIMAL NRM
  TRAIT LAYDATE
END
```

The estimate for the ANIMAL effect reported in `SumEstimates.out` is almost unchanged:

```
***** Estimates for RE  1    "ANIMAL"    *************************************
     No. of levels      =      615
     Covariance structure =    NRM
     Order of fit       =        1
 Covariance matrix
   1     11.673
 Matrix of correlations and variance ratios
   1      0.3476
 Covariances & correlations with approximate sampling errors
   2 COVS A 1 1          11.6733         1.39377         vrat     0.348    0.031
```

This suggests that almost all of the among-individual variance is – rightly or wrongly – being partitioned as V_A here. In fact, here we know that the partitioning is wrong as the simulated data included both additive genetic effects and additional effects that were not associated with the pedigree structure (i.e. permanent environment effects). To obtain an unbiased estimate of V_A we therefore have to fit ANIMAL twice, once with, and once without a pedigree attached. In WOMBAT this requires the duplication of the ANIMAL column in the datafile, which we will call PE (for Permanent Environment). We have done this for you in the `gryphonRM_pe.dat` datafile. The parameter file should now look like this:

```
COMMENT WOMBAT analysis of Gryphon laying date

ANALYSIS UNI

PEDS gryphon.ped

DATA gryphonRM_pe.dat
  ANIMAL 900
# A duplicate of ANIMAL
  PE 900
  BYEAR 35
  AGE 5
  YEAR 35
  LAYDATE
END DATA

MODEL
  RAN ANIMAL NRM
  RAN PE
  TRAIT LAYDATE
END

VAR residual 1
1.0
VAR ANIMAL 1
1.0
VAR PE 1
1.0
```

Which yields the following variance components estimates for ANIMAL and PE:

```
***** Estimates for RE  1   "ANIMAL"   ***************************************
     No. of levels      =      615
     Covariance structure =    NRM
     Order of fit       =        1
 Covariance matrix
   1     4.7569
 Matrix of correlations and variance ratios
   1       0.1473
 Covariances & correlations with approximate sampling errors
   2  COVS A 1 1         4.75691         1.81384       vrat     0.147    0.054

***** Estimates for RE  2   "PE"   ***************************************
     No. of levels      =      469
     Covariance structure =    IDE
     Order of fit       =        1
 Covariance matrix
   1     6.2365
 Matrix of correlations and variance ratios
   1       0.1931
 Covariances & correlations with approximate sampling errors
   3  COVS B 1 1         6.23646         1.76169       vrat     0.193    0.054
```

The estimate of V_A is now much lower (reduced from 11.67 to 4.76) since the additive genetic and permanent environment effects are now being separated properly. The heritability is again given in the standard output file (0.147). However, to obtain the repeatability (($V_A + V_{PE}$) / V_P) and its standard error, we have to tell WOMBAT explicitly how to calculate the repeatability. For this we first need to know how our random effects are numbered. In this case V_R is 1, V_A is 2 and V_{PE} is 3. WOMBAT automatically gives V_P the last number, in this case number 4. Knowing this we can first calculate $V_A + V_{PE}$ (which gets number 5), and then divide this sum by V_P (so 5 divided by 4). We do this by adding the following section to the end of our parameter file:

```
SE+USR
# Calculate 'VA+VPE', which is the SUM of 1 and 2
  SUM VA+VPE 2 3
# Calculate 'repeat', which is Variance component RAtio of 5 and 4
  VRA repeat 5 4
```

```
END
```

Parameters requested will be added to `SumEstimates.out`:

```
***** User defined functions of covariances*********************************
 Linear functions of covariances
   5  VA+VPE          10.9934          1.20027
 Variance ratios and correlations
   1  repeat          0.340318         0.281115E-01   vrat
```

Note that this provide a similar repeatability estimate as we got from the model including ANIMAL without the specification of a pedigree file (section 3), but with a slightly smaller standard error.

5) Adding additional effects and testing significance

Models of repeated measures can easily be extended to include other fixed or random effects. For example try including AGE, year of birth (BYEAR) and year of measurement (YEAR). Don't forget to provide additional starting values for the additional random effects. Also, if you still want WOMBAT to estimate the repeatability, V_P is now variance component 6, and $V_A + V_{PE}$ becomes 7.

```
MODEL
  FIX AGE
  RAN ANIMAL NRM
  RAN PE
  RAN YEAR
  RAN BYEAR
  TRAIT LAYDATE
END
```

This model will return additional variance components corresponding to year of measurement and birth year effects. The latter were not simulated, as should be apparent from the model fit (very low V_{BYEAR} and no change in log-likelihood if you would compare it to a reduced model with BYEAR excluded). However, as should be apparent from the variance component estimate and its sampling error reported in `SumEstimates.out`, YEAR effects were simulated. However, as pointed out in Tutorial 1 this should be formally confirmed by likelihood ratio tests.