**Miro 2021 BLUP notes**

**10/18/2021**

1. **Introduction**
   1. This is a short explanation of the analyses and how to use the data. Please note that it is not a manuscript-quality document, it is just an outline in bullet-point format just do record what was done, and give the primary results.
2. **Overview**
   1. BLUP analyses of four Miro datasets
      * E. camaldulensis progeny test (1)
      * E. urophylla progeny tests (2)
      * E. pellita progeny test (1)
      * E. pellita clonal tests (3)
   2. Progeny tests
      * Progeny test analyses were done in similar manners, but not exactly the same due to differences in number of tests and availability of information from other Camcore trials.
      * For each progeny dataset, there are two sheets in the .xlsx file
        + A list of prov and family BLUPs
        + A list of candidate trees for inspection and selection
          - Selected trees can go into further clonal testing, be grafted into clone banks or orchards, or serve as primary seed collection trees in a seedling seed orchard.
      * The primary analysis variable is Volume.
        + I convert all data to standardized volume (stVOL).
          - This is a useful step to remove differences in variances due to scale effects (different sizes of trees, or slight differences in measurement age).
          - This also makes data interpretation easy, as stVOL has a mean = 100%, and genetic value predictions are expressed as units of % above and below that mean.
      * Generally there are four kinds of variables predicted in the BLUP analyses
        + Phat = provenance effect
        + GCA = family within provenance effect
        + FamGain = Phat + GCA
        + Gw = within-family genetic gain (the tree compared to the family mean, i.e., FamGain)
3. **Camaldulensis Progeny**
   1. Data Available
      * One test (727701J1, age 5, good survival 89%
      * There were four controls: Family = 999 uro, 998 pell, 711 uro, 301 pell
      * 14 families from 3 provenances

Prov Provname

1 Laura River

2 Morehead River

3 Kennedy River

* 1. Results
     + Single-site analysis only
     + large provenance differences (p2b = 0.08, phat = -11% to +13%)
     + heritability a bit low (h2b = 0.09), but OK
     + This was a small test, so the selection list contains 55 candidates from 578 surviving trees.

1. **Urophylla Progeny**
   1. Data Available
      * Two tests (497701J1, 497701J2), age 4, decent survival 72% to 74%
        + Both tests in Ghana
      * 54 families from 19 provenances
   2. Other Data
      * Many other Camcore tests (approx. 190) from around the world.
   3. Approach
      * First analyze the two Miro Ghana tests together
        + Calculate Phats + GCA + gainw
          - FamGain = Phat + GCA
        + Compare Phats and GCA with other results
          - Modify FamGain or use average for Ghana
        + Generate Selection List
   4. Results
      * Important provenance variation (p2 = 0.12, phat = -35% to +39%)
      * Moderate family within prov variation (h2 = 0.12, GCA = -15% to +23%)
      * FamGain variation is substantial (FamGain = -42% to +62%)
      * Comparison of these Miro-only BLUPs with BLUPs for other countries showed best correlation with data from Mexico and Venezuela.
        + Corr(FamGainMiro, FamGainMex) = 0.85
        + Corr(FamGainMiro, FamGainVen) = 0.84
      * With two Miro tests, the Miro FamGain values are reasonably precise, and with the high correlation with Mexico and Venezuela, the family ranks would not show large changes if we did some kind of combined analysis.
      * Select list contains 145 candidates from 1993 surviving trees.
2. **Pellita Progeny**
   1. Data Available
      * one test (647701J1), age 5, decent survival 73%
      * 45 families from 6 provs
      * ProvList

Prov Provname

1 Bupul

2 Jagebob

3 Kweel

5 Okaba

6 Caruk

* 1. Other Data
     + There are a number of other tests from other Camcore members. Preliminary analysis showed Mexico and Venezuela to have some value.
  2. Analysis Notes
     + With only one Miro test, there would be more value to add in some data from the other regions.
     + Analyze the Miro test alone.
     + Consider how to do a combined-test analysis.
  3. Results
     + Miro stand-alone
       - Very little provenance variance, phats = ±3%
       - GCA variance is moderate, GCA = ±11%
     + Combined Mex-Ven-Miro analyses
       - Included two tests from Venezuela with age 3 data, and five tests from Mexico with ages 2, 3, or 5-year data.
         * In general, ages 3 and 5 have very similar genetic parameters, and can be considered the same trait once the scale effects are dealt with through standardization.
         * The age 2 test had very good growth, similar to other age 3 datasets, and so was considered as age 3.
       - Therefore the analysis trait was stVOL35.
     + The first model was a homogeneous variance model
       - This effectively considers growth in Miro, Mexico and Venezuela to be exactly the same trait, and weights data from all tests equally.
         * Little provenance variance, phat = -2 to +3%
         * GCA variance is substantial, GCA = -20 to +22%
         * FamGain = -18% to +21%
     + The second model used a Country-Corr model
       - This considers growth in Miro, Mexico and Venezuela to be different traits, but uses data from the other regions to help with the prediction according to how correlated it appears to be.
       - Technical details
         * Use structured G matrix UNR to allow different correlation of Mex and Ven with Miro.
         * Use Fam(prov) as a combined random effect for model simplification. This will predict FamGain directly, and will not predict Phat and GCA as separate components.
       - Results
         * The two approaches gave very similar results, I prefer the second Country-Corr model.
         * Corr(FamGainmex, FamGainMiro) = 0.98
         * Corr(FamGainven, FamGainMiro) = 0.78
         * FamGain = -16% to +18%
     + The stand-alone Miro BLUPs and the MexVenMiro BLUPs were similar, with a Corr(FamGainMiro, FamGainMexVenMiro) = 0.77.
       - I have more confidence in the values from the MexVenMiro analysis, and used those to
       - generate the selection list.
     + Select list contained 106 candidates from 927 surviving trees.

1. **Pellita Clones**
   1. Miro Data Available
      * Three tests, age 3 HT & DBH, good survival 90%
        + B50 = Sierra Leone, age 3
        + F06 = Sierra Leone, age3
        + K83 = Ghana, age 2
      * 92 clones
      * Controls: 5 pellita families, 7 pellita control bulks, 2 uro fams, 2 GU fams, 2 unknown hybrid fams or clones

|  |  |
| --- | --- |
|  | Clone |
| **E.pellita** | Fam P88 |
| **E.pellita** | Fam P95 |
| **E.pellita** | Fam P92 |
| **E.pellita** | Fam P83 |
| **E.pellita** | Fam P77 |
| **E.pellitaCont** | IBSSO |
| **E.pellitaCont** | IBSSO0113 |
| **E.pellitaCont** | SSB |
| **E.pellitaCont** | SSB15160 |
| **E.pellitaCont** | SSB17033 |
| **E.pellitaCont** | VAFS |
| **E.pellitaCont** | VAFS00408 |
| **EGRAxURO** | Fam GH7 |
| **EGRAxURO** | Fam GH32 |
| **EURO** | Fam EU8 |
| **EURO** | Fam EU5 |
| **Hybrid** | Fam B6 |
| **Hybrid** | Fam B53 |

* 1. Pell Clone Analysis
     + Trait = stVOL23, mean = 100%
     + Single-site parms
       - Some variation in single-site H2
         * B50K H2= 0.10
         * F06 H2= 0.41
         * K83 H2= 0.24
     + Multiple-site BLUP, three tests,
       - Moderate level of GxE
         * Overall clonal rBg = 0.58
       - Run the three paired-site analyses to check if GH is very different
         * SL\_B50K & SL\_F06 rBg = 0.53
         * SL\_B50K & GH\_K83 rBg = 0.49
         * SL\_F06 & GH\_K83 rBg = 0.72
         * No indication that the Ghana test is very different, so use a 3-sites combined
     + Combined site analysis variance structure
       - Can use the statement repeated / group=testid;
         * Estimates different residual variance for the tests.
       - Using one common standardization CV = 15 18 44 gives very similar residual variances
         * and therefore similar H2 and rBg for the three tests.
       - Using site-specific CV gives different residual variances for the three sites, which look a bit more like the single-site parms.
       - Fairly similar BLUPs from either approach, will use the heterogeneous test residual variance model to calculate the final BLUPs.
     + Ghat = overall clonal blup, expressed in ± percent above or below the pellita clonal mean
       - Range from -25% to +32%

1. **Select Lists from Progeny Test Analyses**
   1. General procedure
      * The same basic steps apply to all three Select lists of candidates.
      * The lists contain candidates for selection, based on the family BLUP and the within-family gain prediction (gw).
        + There are more candidates from the better tests.
        + From poor families, there may be only one candidate, if there was a progeny with an outstanding phenotype.
      * All candidates should be inspected in the field to confirm data accuracy, form traits, general health, etc.
        + Typically, we include brief comments on the data sheet for each tree to indicate any problems, and if the tree is a definite Yes, a definite No, or if intermediate in quality, any small defects which should be noted.
        + I use a five category system: YES, OK+, OK, OK-, and NO.
        + If the tree is OK or OK-, I will usually record comments like
          - Slightly crooked
          - Slight sweep
          - Small fork at top
          - Edge tree, no border trees
          - Etc.
      * Please inspect trees and send comments back to us for final selection and assignment of CamNums for our database.
      * If you find trees that like very much that are not on the list, send us the information and we can compare the measurement data and gain prediction.
   2. What to do with selected trees?
      * This is mostly a Miro decision, but you can use the selections for clonal trials, or collect seed from these trees as an elite seed crop, or graft them into clone banks or orchards for future breeding.
      * For thinning the progeny tests into seed orchards, we recommend marking the selections as Primary Seed Trees, thinning the rest of the stand using silvicultural and spacing and the FamGain BLUPs to guide the decisions as to what trees to retain and what to thin.
        + You may want to do the thinning in phases to reduce the chance of windthrow.
        + Final target might be around 200 to 250 stems per hectare.
2. **Questions**
   1. If you have questions or want to discuss something about these lists, please let us know and we can set up a time.