Sampling tree breeding trials

Luis Apiolaza & Elena Moltchanova
University of Canterbury
Private Bag 4800, Christchurch, New Zealand

This presentation is based on work for the New Zealand Dryland Forests Initiative, and connects with projects and discussions with Clemens Altaner, Nicholas Davies, John Walker, Ruth McConnochie, Paul Millen and Shaf van Ballekom.
http://www.nzdfi.org.nz
Motivation

• We are domesticating *Eucalyptus bosistoana* for the production of durable and high performance timber.

• Pretty much any tree breeding program involves quantity & quality of wood + adaptation traits.

• Some traits cheap and easy to assess, while the rest are very expensive -> sampling.
Measuring longitudinal growth strain in standing trees

Strain is measured with a resistance or a CIRAD tool

We need 8-10 measures/tree to get a proper description of strain.
From (very slowly) assessing trees to (slowly) assessing 1-2 year old plants
Another example: Heartwood variability

Stained disks of *E. bosistoana* showing heartwood percentage
In the old days: truncation sampling

<table>
<thead>
<tr>
<th>cheap to assess</th>
<th>expensive to assess</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**h² of cheap trait**: no bias, increasing precision with larger samples

**h² of expensive trait**: bias, increasing precision with larger samples

**r_g between traits**: Large bias, poor precision
Better: random sampling

- **cheap to assess**
  - No bias, increasing precision with larger samples

- **expensive to assess**
  - Decreasing bias and increasing precision with larger samples
  - Decreasing bias and increasing precision with larger samples
Sometimes random is too random: Rated Set Sampling

<table>
<thead>
<tr>
<th>cheap to assess</th>
<th>expensive to assess</th>
</tr>
</thead>
</table>

Using additional info (cheap trait) we can improve representativeness of sample, increasing precision

1. Choose multiple ‘sets’ of observations

2. Within each set rank observations based on cheap trait

3. Choose smallest unit in first set, second smallest in second set, etc.

4. You have a sample.

Goes back to 1950s, current revival in environmental monitoring
Example of coverage

Follows performance of random sampling, but more precise
And on top we have spatial trends
Balanced Acceptance Sampling

• Based on Halton sequences, which are deterministic but appear random for many purposes, generating well-spread positions.
• In general, evenly spatially balanced designs are more precise.
• This can be adapted to consider additional information from multiple covariates (e.g. cheap trait) and groupings (e.g. families)
Example spatial sampling

Random sample

Balanced Acceptance Sampling
In summary

• Sampling is a necessity in tree breeding programs
• A poor sampling scheme will deliver misleading, poor or unusable data
• Good sampling schemes will increase precision (or maintain it for smaller samples)
• Explicit spatial constraints to sampling are needed to make the most of our trials