**Master Sample Methods**

**Paul van Dam-Bates - pbates@doc.govt.nz**

**Sample generation:**

1. Choose Island (eg. North Island, South Island, Stewart Island)
2. Select random seed for Halton sequence that falls within island.
3. Use BAS (Halton Sequence) to generate points.

*Provide to user:*

Since it’s trivial computationally to generate the points we can provide the code that will automatically select the sample. We will build a package to select the master sample based on specific parameters and island locations. We could then host this on GitHub. We could also invest in a sampling interface that does this outside of R (or in a ShinyApp). The algorithm is very simple to write for generating the points and is not required to be in R.

**Tier 1 Sample:**

The Tier 1 sample is for our own (DOC) use and will typically not be sampled in the frequency or intensity needed for individual user projects. It will be used in the planning process to contribute to sample size etc. However, we do not want extra visits to Tier 1 sites in order to reduce bias created by the impact of sampling a plot. Regional councils who are extending the Tier 1 grid will also follow this example. To incorporate it in the analysis stage it will work as a spatially balanced equal probability sample laid on top of BAS.

**Intensification:**

If a particular group would like to extend the monitoring in an area that we are already doing, we can easily provide points in their region to visit that extends both our knowledge and theirs. This will be done by selecting the next set of points from the Master Sample that falls within that region.

**Stratification:**

To stratify select the first n1 points that fall in strata 1 and the first n2 points that fall within strata 2 etc. from the Master Sample. This is essentially intensification.

**Panel Designs:**

Select N + k points from the Master Sample where N is the total number of points needed for all panels and k is the oversample to replace points through time.

Each panel *i* is made up of the first *ni* points that occur in the sample.

As plots are set up; if extra locations are needed, use the next point in order and add one from the k oversample to the final panel. If a full rotation of plots has been completed and new points are needed, use points in the k oversample to replace the missing points. This method ensures that the overall sample is a pure BAS sample while we maintain decent spatial balance between each panel. Panels combined are still spatially balanced and overall maintain an equal probability of inclusion.

*Example of Panel Design*: [1-n, 1-4] Design

Sampling intention:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Panel/Year | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | … |
| 1 | X | X | X | X | X | X | X | X | X |
| 2 | X |  |  |  |  | X |  |  |  |
| 3 |  | X |  |  |  |  | X |  |  |
| 4 |  |  | X |  |  |  |  | X |  |
| 5 |  |  |  | X |  |  |  |  | X |
| 6 |  |  |  |  | X |  |  |  |  |

3 points in panel 2 are not actually valid. Each following panel then lends to panel 2 and panel 6 replaces its points from the oversample.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Panel/Year | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | … |
| 1 | X | X | X | X | X | X | X | X | X |
| 2 | X |  |  |  |  | X |  |  |  |
| 3 |  | X |  |  |  |  | X |  |  |
| 4 |  |  | X |  |  |  |  | X |  |
| 5 |  |  |  | X |  |  |  |  | X |
| 6 |  |  |  |  | X |  |  |  |  |

k oversample points

Realised overall BAS sample goes from {xi}1Nto{xi}1N+3

Now, we have run our sample for 5 years and a landslide removes a point from Panel 1 in year 7. Add a point from oversample.

This could also be used if funding was increased and we wanted to increase sample size in a particular panel.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Panel/Year | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | … |
| 1 | X | X | X | X | X | X | X | X | X |
| 2 | X |  |  |  |  | X |  |  |  |
| 3 |  | X |  |  |  |  | X |  |  |
| 4 |  |  | X |  |  |  |  | X |  |
| 5 |  |  |  | X |  |  |  |  | X |
| 6 |  |  |  |  | X |  |  |  |  |

k oversample points

Realised overall BAS sample goes from {xi}1N+3to{xi}1N+4