

BBC 2023 Alaska Subsample Experiment

Courtney Meier

2024-06-26

Goal

To determine the efficacy of subsampling BBC core samples at the D18 TOOL and D19 DEJU sites to reduce time required to sort roots into size categories.

Experimental Setup and Analyses

- Select $n=10$ cores for each plot site. Cores are sampled from random spatially-balanced plot locations, and locations of clipIDs within plots are also random. The sampling design provides an unbiased estimate of root biomass throughout the Tower airshed.
- For each `sampleID` collected in the field, test subsampling efficiency at various levels of sorting by creating subsamples with the following percentages of the total core mass:
 - 10%
 - 15%
 - 25%
 - 50%

The sum of all the subsamples = 100%; that is, the mass of the entire core sample.

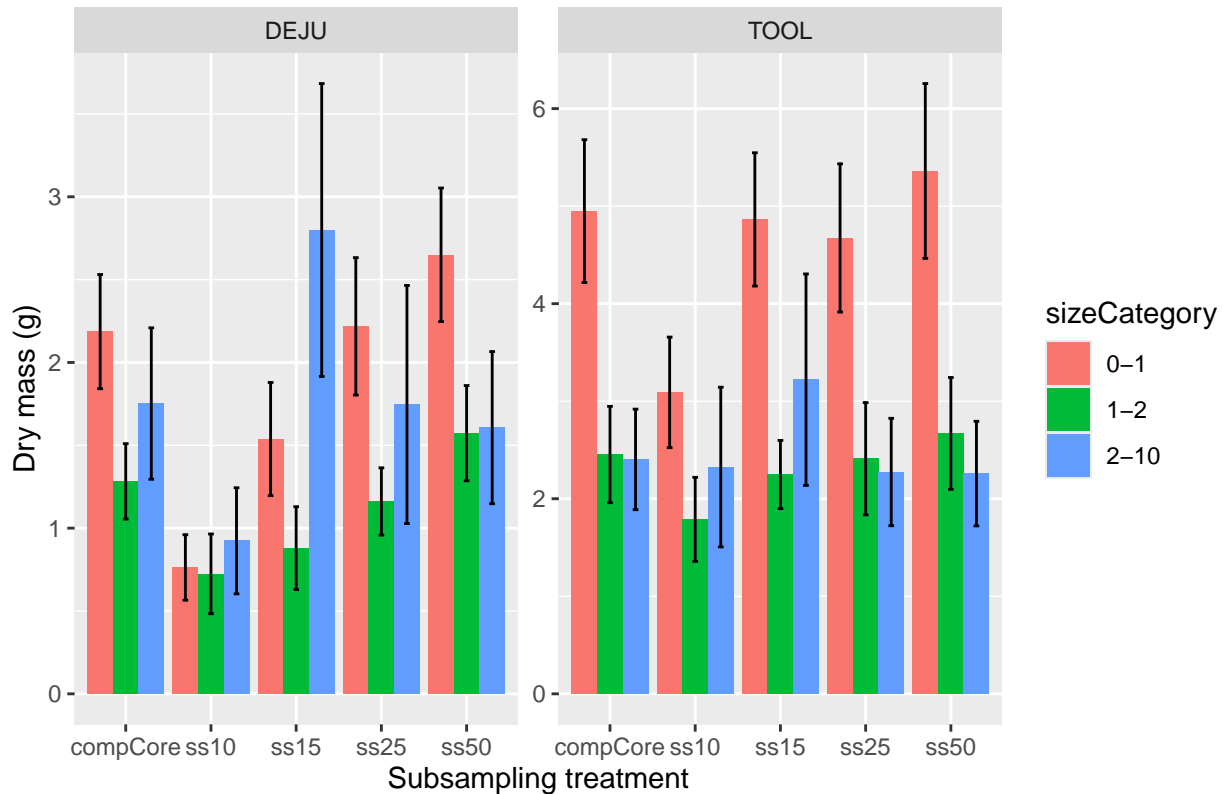
- When subsampling is employed, estimated `dryMass` is calculated as follows: $dM = ssDM/X$, where:
 - dM = estimated root `dryMass` for a given level of subsampling
 - X = the proportion of the total core sample represented by the subsample (ranging from 0-1)
- Compare `dryMass` results calculated via subsampling with `dryMass` obtained with no subsampling, and use mixed effects models to analyze results.

Procedure

1. Soil cores (6.65 cm diameter x 30 cm max depth) were collected in the field and frozen in plastic bags for processing at a later date. To generate subsamples, thaw frozen core sample and homogenize soil/root mixture in a plastic bag or tub.
2. Weigh the whole core and create subsamples at desired percentages (above) based on the fresh mass of the total core (roots + soil).
3. Sort subsamples separately into root sizeCategory (0-1, 1-2, 2-10 mm) following the normal procedure in the BBC protocol, and dry and weigh sorted subsamples separately for each sizeCategory.
4. Record the `dryMass` of each sorted subsample by sizeCategory for a given core sample, then recombine all subsamples within a sizeCategory and weigh to obtain the “whole core” `dryMass`.
 - a. Enter the “whole core” `dryMass` for each sizeCategory into Fulcrum for ingest as normal.

Data Summary

Alaska BBC subsampling



Results: Mixed-effects model analysis

Approach: For each site, construct a model with **sizeCategory** as the sole fixed effect where the mean is allowed to vary randomly by **sampleID**. Next, construct a model with an additional **treatment** fixed effect, with the mean similarly allowed to vary randomly by **sampleID**. Theoretically, if subsampling has no effect on estimated dryMass the model with **treatment** will not be significantly better according to AIC or p-value than the simpler model that only has **sizeCategory**.

DEJU results

```
# DEJU simplest mixed-effects model
m1DEJU <- lmerTest::lmer(formula = estDryMass ~ sizeCategory + (1|sampleID),
  data = df %>% dplyr::filter(siteID == "DEJU"),
  REML = FALSE)

# DEJU mixed-effects model with subsampling 'treatment'
m2DEJU <- lmerTest::lmer(formula = estDryMass ~ treatment + sizeCategory + (1|sampleID),
  data = df %>% dplyr::filter(siteID == "DEJU"),
  REML = FALSE)

# DEJU mixed-effects model with treatment*sizeCategory interaction
m3DEJU <- lmerTest::lmer(formula = estDryMass ~ treatment*sizeCategory + (1|sampleID),
  data = df %>% dplyr::filter(siteID == "DEJU"),
  REML = FALSE)
```

```

# Model comparison
anova(m1DEJU, m2DEJU, m3DEJU)

## Data: df %>% dplyr::filter(siteID == "DEJU")
## Models:
## m1DEJU: estDryMass ~ sizeCategory + (1 | sampleID)
## m2DEJU: estDryMass ~ treatment + sizeCategory + (1 | sampleID)
## m3DEJU: estDryMass ~ treatment * sizeCategory + (1 | sampleID)
##      npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## m1DEJU     5 508.62 523.67 -249.31   498.62
## m2DEJU     9 498.85 525.94 -240.43   480.85 17.768  4    0.00137 **
## m3DEJU    17 498.05 549.23 -232.02   464.05 16.804  8    0.03222 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

anova(m2DEJU, m3DEJU)

## Data: df %>% dplyr::filter(siteID == "DEJU")
## Models:
## m2DEJU: estDryMass ~ treatment + sizeCategory + (1 | sampleID)
## m3DEJU: estDryMass ~ treatment * sizeCategory + (1 | sampleID)
##      npar    AIC    BIC logLik deviance Chisq Df Pr(>Chisq)
## m2DEJU     9 498.85 525.94 -240.43   480.85
## m3DEJU    17 498.05 549.23 -232.02   464.05 16.804  8    0.03222 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



- The addition of treatment to the m2DEJU model adds significant explanatory value compared to sizeCategory alone (lower AIC and  $p < 0.005$  compared to m1DEJU).
- Takehome from m2DEJU: With sizeCategory accounted for, the estimated dryMass is not the same across all subsampling treatments.
- Overall: The m3DEJU model is significantly better than the m2DEJU model → effects of subsampling treatment are not the same across sizeCategory.
- Need to determine which treatments are not significantly different from the dryMass for the complete core.



## Type III Analysis of Variance Table with Satterthwaite's method
##              Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
## treatment          24.061   6.0153     4    140  5.3402 0.0004949 ***
## sizeCategory        16.342   8.1709     2    140  7.2538 0.0010057 **
## treatment:sizeCategory 20.111   2.5138     8    140  2.2317 0.0285053 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
## method [lmerModLmerTest]
## Formula: estDryMass ~ treatment * sizeCategory + (1 | sampleID)
## Data: df %>% dplyr::filter(siteID == "DEJU")
##
##      AIC      BIC    logLik deviance df.resid
##    498.0    549.2   -232.0    464.0      133
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.1584 -0.4473 -0.0688  0.3135  4.0605
##

```

```

## Random effects:
## Groups Name Variance Std.Dev.
## sampleID (Intercept) 0.5086 0.7132
## Residual 1.1264 1.0613
## Number of obs: 150, groups: sampleID, 10
##
## Fixed effects:
## Estimate Std. Error df t value Pr(>|t|)
## (Intercept) 2.18627 0.40436 63.69951 5.407 1.03e-06
## treatmentss10 -1.42347 0.47464 140.00000 -2.999 0.00321
## treatmentss15 -0.64860 0.47464 140.00000 -1.367 0.17397
## treatmentss25 0.03169 0.47464 140.00000 0.067 0.94686
## treatmentss50 0.46343 0.47464 140.00000 0.976 0.33056
## sizeCategory1-2 -0.90400 0.47464 140.00000 -1.905 0.05888
## sizeCategory2-10 -0.43413 0.47464 140.00000 -0.915 0.36195
## treatmentss10:sizeCategory1-2 0.86530 0.67124 140.00000 1.289 0.19949
## treatmentss15:sizeCategory1-2 0.24571 0.67124 140.00000 0.366 0.71488
## treatmentss25:sizeCategory1-2 -0.15296 0.67124 140.00000 -0.228 0.82008
## treatmentss50:sizeCategory1-2 -0.17230 0.67124 140.00000 -0.257 0.79780
## treatmentss10:sizeCategory2-10 0.59493 0.67124 140.00000 0.886 0.37697
## treatmentss15:sizeCategory2-10 1.69592 0.67124 140.00000 2.527 0.01263
## treatmentss25:sizeCategory2-10 -0.03743 0.67124 140.00000 -0.056 0.95561
## treatmentss50:sizeCategory2-10 -0.60905 0.67124 140.00000 -0.907 0.36578
##
## (Intercept) ***
## treatmentss10 **
## treatmentss15
## treatmentss25
## treatmentss50
## sizeCategory1-2 .
## sizeCategory2-10
## treatmentss10:sizeCategory1-2
## treatmentss15:sizeCategory1-2
## treatmentss25:sizeCategory1-2
## treatmentss50:sizeCategory1-2
## treatmentss10:sizeCategory2-10
## treatmentss15:sizeCategory2-10 *
## treatmentss25:sizeCategory2-10
## treatmentss50:sizeCategory2-10
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation matrix not shown by default, as p = 15 > 12.
## Use print(x, correlation=TRUE) or
## vcov(x) if you need it

```

- Across **sizeCategory**, the **ss10** treatment generates significantly lower estimated **dryMass** compared to the complete core.
- Visually, the 2-10 mm **sizeCategory** for the 15% subsample is much more variable and has higher average mass than the 2-10 mm roots in the complete core. However, across all **sizeCategories** the 15% subsample is not significantly different from the complete core.
- Conservatively, subsampling at 25% is supported by the data and appears most similar to the complete core.

TOOL results

```
# TOOL simplest mixed-effects model
m1TOOL <- lmerTest::lmer(formula = estDryMass ~ sizeCategory + (1|sampleID),
                        data = df %>% dplyr::filter(siteID == "TOOL"),
                        REML = FALSE)

# TOOL mixed-effects model with subsampling 'treatment'
m2TOOL <- lmerTest::lmer(formula = estDryMass ~ treatment + sizeCategory + (1|sampleID),
                        data = df %>% dplyr::filter(siteID == "TOOL"),
                        REML = FALSE)

# TOOL mixed-effects model with treatment*sizeCategory interaction
m3TOOL <- lmerTest::lmer(formula = estDryMass ~ treatment*sizeCategory + (1|sampleID),
                        data = df %>% dplyr::filter(siteID == "TOOL"),
                        REML = FALSE)

m4TOOL <- lmerTest::lmer(formula = estDryMass ~ treatment + sizeCategory -1 + (1|sampleID),
                        data = df %>% dplyr::filter(siteID == "TOOL"),
                        REML = FALSE)

# Model comparison
anova(m1TOOL, m2TOOL, m3TOOL)

## Data: df %>% dplyr::filter(siteID == "TOOL")
## Models:
## m1TOOL: estDryMass ~ sizeCategory + (1 | sampleID)
## m2TOOL: estDryMass ~ treatment + sizeCategory + (1 | sampleID)
## m3TOOL: estDryMass ~ treatment * sizeCategory + (1 | sampleID)
##      npar    AIC    BIC logLik deviance  Chisq Df Pr(>Chisq)
## m1TOOL     5 736.33 752.29 -363.16   726.33
## m2TOOL     9 734.48 763.22 -358.24   716.48 9.8437  4  0.04314 *
## m3TOOL    17 741.36 795.64 -353.68   707.36 9.1263  8  0.33175
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

anova(m2TOOL, m3TOOL)

## Data: df %>% dplyr::filter(siteID == "TOOL")
## Models:
## m2TOOL: estDryMass ~ treatment + sizeCategory + (1 | sampleID)
## m3TOOL: estDryMass ~ treatment * sizeCategory + (1 | sampleID)
##      npar    AIC    BIC logLik deviance  Chisq Df Pr(>Chisq)
## m2TOOL     9 734.48 763.22 -358.24   716.48
## m3TOOL    17 741.36 795.64 -353.68   707.36 9.1263  8  0.3318
```

- The addition of **treatment** to the m2TOOL model adds significant explanatory value compared to **sizeCategory** alone (lower AIC and $p < 0.05$ compared to m1TOOL).
- Takehome from m2TOOL: With **sizeCategory** accounted for, the estimated dryMass is not the same across all subsampling treatments.
- The interaction term is not significant in the m3TOOL model → subsampling has similar effects across all sizeCategories.
- Need to determine which subsampling treatments are not different from the complete core dryMass.

```
## Type III Analysis of Variance Table with Satterthwaite's method
##      Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
```

```

## treatment      26.593   6.648     4   168   2.5345   0.04211 *
## sizeCategory 191.532  95.766     2   168 36.5087 6.822e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
## method [lmerModLmerTest]
## Formula: estDryMass ~ treatment + sizeCategory + (1 | sampleID)
## Data: df %>% dplyr::filter(siteID == "T00L")
##
##      AIC      BIC    logLik deviance df.resid
##    734.5    763.2   -358.2    716.5      171
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.7919 -0.5479 -0.1481  0.3399  4.4578
##
## Random effects:
## Groups Name Variance Std.Dev.
## sampleID (Intercept) 2.359 1.536
## Residual 2.623 1.620
## Number of obs: 180, groups: sampleID, 12
##
## Fixed effects:
##              Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)    4.7237    0.5464  23.7067   8.645 8.62e-09 ***
## treatmentss10   -0.8673    0.3817 168.0000  -2.272  0.0244 *
## treatmentss15    0.1756    0.3817 168.0000   0.460  0.6460
## treatmentss25   -0.1491    0.3817 168.0000  -0.391  0.6965
## treatmentss50    0.1607    0.3817 168.0000   0.421  0.6743
## sizeCategory1-2 -2.2736    0.2957 168.0000  -7.689 1.19e-12 ***
## sizeCategory2-10 -2.0914    0.2957 168.0000  -7.073 3.90e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) trtm10 trtm15 trtm25 trtm50 szC1-2
## tretmntss10 -0.349
## tretmntss15 -0.349  0.500
## tretmntss25 -0.349  0.500  0.500
## tretmntss50 -0.349  0.500  0.500  0.500
## sizCtgry1-2 -0.271  0.000  0.000  0.000  0.000
## szCtgry2-10 -0.271  0.000  0.000  0.000  0.000  0.500

```

- The 10% subsample treatment has on average significantly lower dryMass compared to the complete core.
- Similar to DEJU, the 2-10 mm **sizeCategory** for the 15% subsample is much more variable and has higher average mass than the 2-10 mm roots in the complete core. Across all sizeCategories, the 15% subsample is not significantly different than the complete core.
- Conservatively, subsampling at 25% is supported by the data and appears most similar to the complete core.