HBP D11 2023 Sorting Subsample Experiment

Courtney Meier

2023-09-20

Goal

To determine the efficacy of subsampling various proportions of HBP samples in D11 to reduce long sort times associated with removing OSD from current-year clipped biomass. **Subsampling is only evaluated in the context of clip harvests that do not require sorting to functional group** (i.e., non-peak biomass clips), because the subsample to total mass ratios will not apply to individual herbGroups.

Experimental Setup and Analyses

- Select n=10 plots (all plots for OAES), resulting in a maximum of n=20 `clipID` to test, due to both `exclosure = Y and N` for each plot. Random spatially-balanced plot locations, and locations of clipIDs within plots, will provide an unbiased estimate of biomass throughout the Tower airshed.
- For each `clipID` harvested in the field, test subsampling efficiency at various levels of sorting by creating subsamples (current-year + OSD) with the following percentages of the total `freshMass`:
  - 10%
  - 15%
  - 25%
  - 50%

  The sum of all the subsamples = 100%; that is, the fresh mass of the entire clip strip.
- When subsampling is employed, calculate dryMass as follows: `dM = fM * (ssDM/ssFM)`, where:
  - dM = dryMass of current-year biomass in the clipID (no OSD)
  - fM = total freshMass in the clipID (current-year + OSD)
  - $ssDM$ = subsampleDryMass of current-year biomass in the subsample (no OSD)
  - $ssFM$ = fresh mass of all biomass in the subsample (current-year + OSD)
- Compare `dryMass` results calculated via subsampling with `dryMass` obtained with no subsampling, and use mixed effects models to analyze results.

Procedure

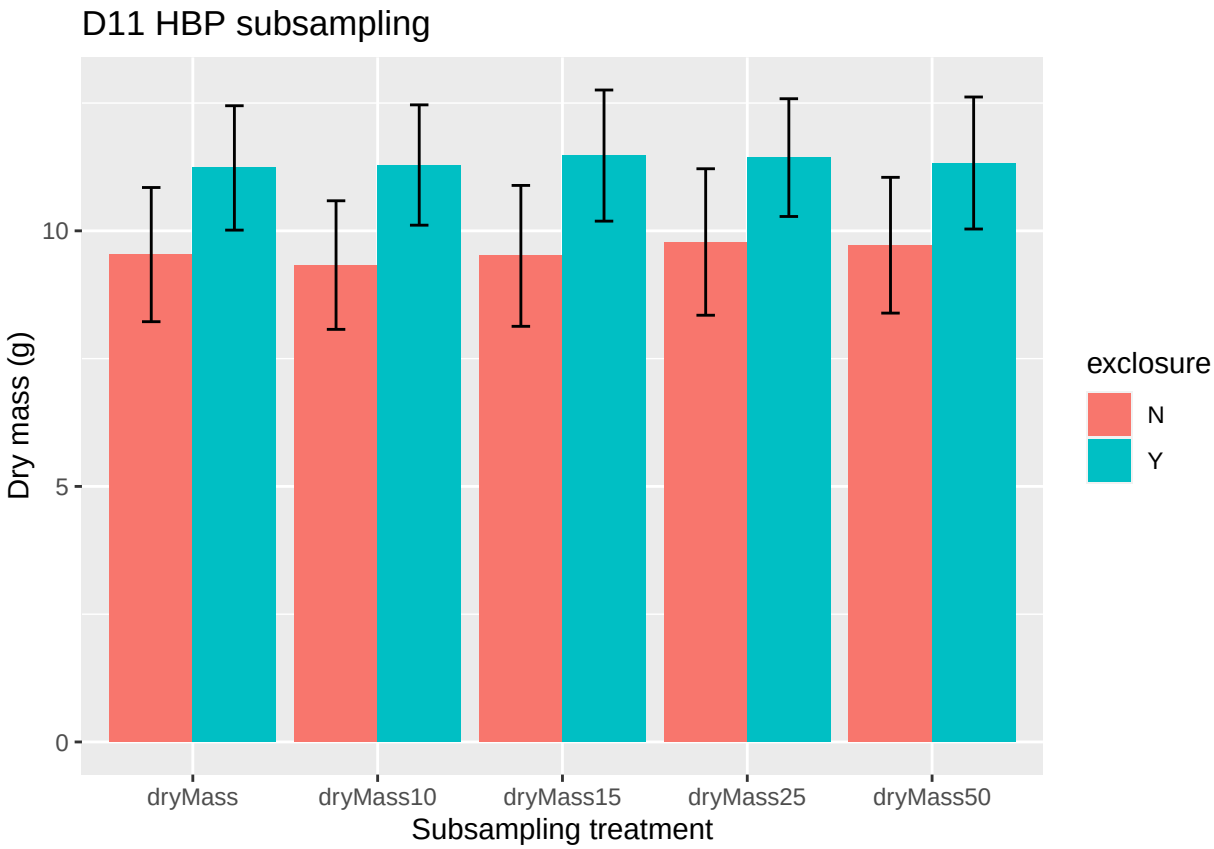
1. Perform clip harvest in the field as normal, and bring clipped biomass back to the laboratory in cold storage as normal.
2. Identify up to n=20 clipIDs (n=10 for `exclosure = Y` and n=10 for `exclosure = N`), originating from each of the 10 plotIDs.
3. Thoroughly mix biomass from each `clipID` to homogenize as thoroughly as possible.
   a. For large amounts of biomass, and when there is more than one bag of biomass for a given `clipID`, use a large bag, box, tray or equivalent vessel to mix the biomass.
4. For each clipID, weigh and record to 0.01 g:

a. **freshMass** = total fresh mass in the clipID (current-year + OSD).
b. It is important that **freshMass** for the entire **clipID** and subsample fresh masses (below) are collected for a given **clipID** as close to each other in time as possible. That is, avoid weighing **freshMass** for a given **clipID** hours apart from the subsample fresh masses as water loss will affect the experimental results.

5. Based on the **freshMass**, calculate the desired subsample fresh masses for testing. For example, assuming **freshMass** = 100 g, the target subsample fresh masses are:
   a. 10% subsample –> 10 g
   b. 15% subsample –> 15 g
   c. 25% subsample –> 25 g
   d. 50% subsample –> 50 g
6. Label a coin envelope for each subsample above with the information below.
   a. **subsampleTest**: 10%, 15%, 25% or 50%
   b. **clipID**
   c. **collectDate**
   d. **exclosure**: Y/N
7. Weigh each fresh subsample created above (current-year + OSD), and record the information below.
   a. **subsampleTest**: as above
   b. **clipID**
   c. **collectDate**
   d. **exclosure**: Y/N
   e. **subsampleFreshMass**: To the nearest 0.01 g; for subsamples < 0.5 g total mass, weigh to the nearest 0.0001 g
8. Sort current-year biomass from OSD for each subsample, and place sorted, current-year biomass into the corresponding labeled coin envelope.
   a. Sorted OSD may be discarded at this point.
9. Dry subsamples until dry; minimum of 48 h @ 65 °C, track drying progress as normal.
10. Remove dry samples from the oven one at a time, and immediately weigh and record:
    a. **subsampleDryMass**: To the nearest 0.01 g; for masses < 0.5 g, weigh to the nearest 0.0001 g.

Data summary

```
## `summarise()` has grouped output by 'treatment'. You can override using the
## `.groups` argument.
```



Results: Mixed-Effects model analysis

1. Models accounting for subsampling effect on **dryMass** and using **clipID** as a random effect (no need to use **exclosure** as a fixed effect since random effect accounts for variation across exclosure treatment).

Model: Data from eventID = "HBP.2023.CLBJ.18" only

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: estDryMass ~ treatment + (1 | clipID)
##    Data: longDF %>% dplyr::filter(eventID == "HBP.2023.CLBJ.18")
##
##      AIC      BIC   logLik deviance df.resid
##    370.5    387.6   -178.3    356.5       78
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.5230 -0.3012 -0.0772  0.3370  2.3320
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  clipID   (Intercept) 32.466   5.698
##  Residual              1.523   1.234
## Number of obs: 85, groups:  clipID, 17
##
## Fixed effects:
##                    Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)         9.32412    1.41398 18.28119   6.594 3.16e-06 ***
```

```
## treatmentdryMass10 -0.57647    0.42328 68.00000  -1.362    0.178
## treatmentdryMass15 -0.05882    0.42328 68.00000  -0.139    0.890
## treatmentdryMass25  0.24118    0.42328 68.00000   0.570    0.571
## treatmentdryMass50  0.19824    0.42328 68.00000   0.468    0.641
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) trtM10 trtM15 trtM25
## trtmntdrM10 -0.150
## trtmntdrM15 -0.150  0.500
## trtmntdrM25 -0.150  0.500  0.500
## trtmntdrM50 -0.150  0.500  0.500  0.500
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## treatment 7.2347  1.8087     4    68  1.1877 0.3241
```

Model: Data from eventID = "HBP.2023.CLBJ.34" only

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: estDryMass ~ treatment + (1 | clipID)
##    Data: longDF %>% dplyr::filter(eventID == "HBP.2023.CLBJ.34")
##
##      AIC      BIC   logLik deviance df.resid
##    262.0    279.1   -124.0    248.0       78
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.95982 -0.46198  0.00238  0.46201  2.41570
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  clipID   (Intercept) 18.2862  4.2762
##  Residual              0.3568  0.5973
## Number of obs: 85, groups:  clipID, 17
##
## Fixed effects:
##                    Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)        11.34176    1.04721 17.53165  10.830 3.42e-09 ***
## treatmentdryMass10  0.41294    0.20487 68.00000   2.016   0.0478 *
## treatmentdryMass15  0.26000    0.20487 68.00000   1.269   0.2087
## treatmentdryMass25  0.21000    0.20487 68.00000   1.025   0.3090
## treatmentdryMass50  0.08647    0.20487 68.00000   0.422   0.6743
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) trtM10 trtM15 trtM25
## trtmntdrM10 -0.098
## trtmntdrM15 -0.098  0.500
## trtmntdrM25 -0.098  0.500  0.500
## trtmntdrM50 -0.098  0.500  0.500  0.500
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
```

```
##           Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## treatment 1.7297 0.43242     4    68  1.2121 0.3138
```

Model: Data from both eventIDs

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: estDryMass ~ treatment + (1 | clipID)
##    Data: longDF
##
##      AIC      BIC   logLik deviance df.resid
##    661.8    683.8   -323.9    647.8      163
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.7672 -0.2862 -0.0286  0.2627  3.0004
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  clipID   (Intercept) 26.6322  5.1606
##  Residual              0.9912  0.9956
## Number of obs: 170, groups:  clipID, 34
##
## Fixed effects:
##                     Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)         10.33294    0.90136  36.03158  11.464 1.41e-13 ***
## treatmentdryMass10  -0.08176    0.24147 136.00000  -0.339    0.735
## treatmentdryMass15   0.10059    0.24147 136.00000   0.417    0.678
## treatmentdryMass25   0.22559    0.24147 136.00000   0.934    0.352
## treatmentdryMass50   0.14235    0.24147 136.00000   0.590    0.556
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) trtM10 trtM15 trtM25
## trtmntdrM10 -0.134
## trtmntdrM15 -0.134  0.500
## trtmntdrM25 -0.134  0.500  0.500
## trtmntdrM50 -0.134  0.500  0.500  0.500
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## treatment 1.9734 0.49334     4   136  0.4977 0.7374
```

2. Null model, and using **clipID** as a random effect. Theoretically, if all subsampling results in the same dryMass as the full sort, the null model should be no different than the "treatment" model.

```
## Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's
##   method [lmerModLmerTest]
## Formula: estDryMass ~ 1 + (1 | clipID)
##    Data: longDF
##
##      AIC      BIC   logLik deviance df.resid
##    655.8    665.2   -324.9    649.8      167
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.8908 -0.2938 -0.0651  0.2928  3.1272
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  clipID   (Intercept) 26.629   5.160
##  Residual              1.006   1.003
## Number of obs: 170, groups:  clipID, 34
##
## Fixed effects:
##             Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)  10.4103     0.8883 34.0000   11.72 1.73e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Data: longDF
## Models:
## m4ME: estDryMass ~ 1 + (1 | clipID)
## m3ME: estDryMass ~ treatment + (1 | clipID)
##      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## m4ME    3 655.78 665.19 -324.89   649.78
## m3ME    7 661.81 683.76 -323.90   647.81 1.9764  4     0.7401
```

3. Linear model, using treatment*exclosure, and no random effect.

```
##
## Call:
## lm(formula = estDryMass ~ exclosure * treatment, data = longDF)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.7517 -4.4638 -0.3566  3.3773 13.3175
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    9.53500    1.25852   7.576 2.69e-12 ***
## exclosureY                     1.69562    1.83459   0.924    0.357
## treatmentdryMass10            -0.20500    1.77982  -0.115    0.908
## treatmentdryMass15            -0.02500    1.77982  -0.014    0.989
## treatmentdryMass25             0.24667    1.77982   0.139    0.890
## treatmentdryMass50             0.18333    1.77982   0.103    0.918
## exclosureY:treatmentdryMass10  0.26188    2.59451   0.101    0.920
## exclosureY:treatmentdryMass15  0.26688    2.59451   0.103    0.918
## exclosureY:treatmentdryMass25 -0.04479    2.59451  -0.017    0.986
## exclosureY:treatmentdryMass50 -0.08708    2.59451  -0.034    0.973
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.339 on 160 degrees of freedom
## Multiple R-squared:  0.02904,    Adjusted R-squared:  -0.02558
## F-statistic: 0.5316 on 9 and 160 DF,  p-value: 0.85
```

```
## Analysis of Variance Table
##
## Response: estDryMass
##                       Df Sum Sq Mean Sq F value  Pr(>F)
## exclosure              1  133.4 133.438  4.6804 0.03199 *
## treatment              4    2.0   0.493  0.0173 0.99941
## exclosure:treatment    4    1.0   0.250  0.0088 0.99985
## Residuals            160 4561.6  28.510
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Summary of subsampling results

Table 1: Difference in estimated dryMass to total dryMass for subsampling (%)

eventID	ss10	ss15	ss25	ss50
HBP.2023.CLBJ.18	-6.2	-0.6	2.6	2.10
HBP.2023.CLBJ.34	3.6	2.3	1.9	0.76

Conclusions

1. Random effect associated with **clipID** explains considerable variability in the data. See summary outputs for all mixed-effects models.
2. Summary output for Mixed Effects model indicates that D11 Clip Strips can be subsampled at the 15% level for all bouts with **herbGroup** = ALL.
   - Comparing results from two separate bouts, estimates of dryMass using a 10% subsample are more variable relative to total dryMass from the full sort compared to other subsampling levels.
   - Estimated dryMass from the 10% subsample is significantly different than the full sort for eventID = "HBP.2023.CLBJ.34" ($p = 0.05$)
   - When eventIDs are analyzed separately (m1ME, m2ME), other subsamples are not significantly different than the entire sorted subsample ($p > 0.22$) –> go with 15% subsample.
   - With both eventIDs analyzed together, there are no differences between the full sort dryMass and estimated dryMass from subsampling ($p = 0.74$).
3. Effect of `exclosure*treatment`, as evaluated in standard linear model (m1LM):
   - Effect of `exclosure` is significant ($p < 0.05$).
   - Effect of `treatment` not significant ($p = 0.99$).
   - No detection of `exclosure:treatment` interaction effect ($p = 0.99$).