TOS DESIGN OPTIMIZATION: LITTERFALL AND FINE WOODY DEBRIS

PREPARED BY	ORGANIZATION	DATE
Sam Simkin	SCI	7/23/2019
Katie Jones	SCI	7/23/2019
Courtney Meier	SCI	7/23/2019

APPROVALS	ORGANIZATION	APPROVAL DATE
Kate Thibault	SCI	12/17/2024

RELEASED BY	ORGANIZATION	RELEASE DATE
Courtney Meier	SCI	12/17/2024

The National Ecological Observatory Network is a project solely funded by the National Science Foundation and managed under cooperative agreement by Battelle. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

1 RELATED DOCUMENTS AND ACRONYMS 1
1.1 Reference Documents . . . 1
1.2 Acronyms . . . 1

2 BACKGROUND AND GOALS 1
2.1 Description of sample design and available data . . . 1
2.1.1 2018 Sampling reduction . . . 2
2.1.2 Optimization analyses . . . 2
2.2 Analytical Goals . . . 5

3 METHODS AND RESULTS 7
3.1 Data Cleaning Methods . . . 7
3.2 Data Imputation Methods . . . 7
3.3 Litterfall Variance components analysis . . . 8
3.3.1 Response variable, test specifics . . . 8
3.3.2 Methods . . . 8
3.3.3 Results . . . 8
3.4 Litterfall Power analysis . . . 9
3.4.1 Response variable, test specifics . . . 9
3.4.2 Methods . . . 9
3.4.3 Results . . . 10
3.5 Fine woody debris variance components analysis . . . 13
3.5.1 Response variable and test specifics . . . 13
3.5.2 Methods . . . 14
3.5.3 Results . . . 14
3.6 Fine woody debris Power analysis . . . 15
3.6.1 Response variable, test specifics . . . 15
3.6.2 Methods . . . 15
3.6.3 Results . . . 16
3.7 Litterfall functional group sorting subset analysis . . . 17
3.7.1 Response variable, test specifics . . . 17
3.7.2 Methods . . . 17
3.7.3 Results . . . 18

4 DISCUSSION 20

5 RECOMMENDATIONS 21
5.1 Recommendations applicable to all sites: . . . 21
5.2 Per site recommendations . . . 22

6 ACKNOWLEDGEMENTS 25

7 REFERENCES 25

LIST OF TABLES AND FIGURES

Table 1 Summary of data included in Litterfall optimization analyses. 3
Table 2 Summary of sites and data included in Litterfall and Fine Woody Debris optimization analyses. Tower Plot Size = dimensions of established Tower Plots; Traps Per Plot = number of herbaceous clip harvests performed per plot per bout; Tower NLCD Class = NLCD class of vegetation in Tower footprint, * = multiple NLCD types present and dominant type is listed. 4
Table 3 The minimum number of plots per site that should be sampled. The criterion is the power to detect a 20 percent inter-annual change in total litter production in 80 percent or more of 500 simulated datasets. A value of ”> 30” for the minimum number of traps indicates that there was less than 80 percent power to detect a 20 percent inter-annual change even with the most intensive sampling design (30 plots per site) examined in the simulations. 12
Table 4 The minimum number of traps per site for which sorting and weighing is required, for the litter functional group at each site with the highest proportion of total litter mass. The criterion is the percentage of 500 random draws in which the mean of the estimated litter functional group mass fell within +/- 10 percent of the mean of the observed litter functional group mass. When the ’Trap subsample size’ column entry is ’Sort All’, sorting subsets of any size resulted in less than 80% of random draws with a mean estimated litter functional group mass within +/- 10% of the mean of the observed mass. 19
Table 5 Tentative per site plot number recommendations for elevated litter trap sampling based on analysis of 1 - 3 years of litterfall productivity data. 22
Table 6 At sites where optimization analyses supported reducing the number of plots, and where there were 3 years of data available, the proposed reduced set of plot IDs to sample elevated litter traps, prioritized by Morton order. 25

Figure 1 Proportion of the total variance explained (y-axis) by each component in the mixed model (different colors, representing random factors in the mixed model) that predicts mass for each litter functional group (facet panels) at each site (x-axis). 9
Figure 2 Heat map indicating which site (x-axis) by litter functional group (y-axis) combinations had data sufficient to fit a linear mixed effect model (above) to detect annual variation in litter production. Year was treated as a fixed effect, and red cells indicate a site/functional group combination for which an effect of year could be detected at a significance level of p < 0.05. 10
Figure 3 Power to detect a 20% change in total annual litter mass (for all design scenarios and all functional groups) between consecutive years, using a critical value of p < 0.05, at each site. 11
Figure 4 Black cells indicate sampling efforts (number of plots on y-axis and number of traps on x-axis) at each site (also on x-axis) that have sufficient power to detect a 20% change in total annual litter mass between years. 12
Figure 5 Proportion of the total variance explained (y-axis) by each random factor in the mixed model that predicts mass for each fine woody debris functional group (facet panels) at each site (x-axis). Colors represent different random factors in the mixed model. 15

Figure 6 Heat map indicating which site (x-axis) by fine woody debris functional group (y-axis) combinations had data sufficient to fit a linear mixed effect model (above) to predict variation in annual fine woody debris production. The 'year' was treated as a fixed effect, and red cells indicate a site/functional group combination for which an effect of 'year' could be detected at a significance level of p < 0.05. 16
Figure 7 Black cells indicate sampling efforts (number of plots on y-axis and number of traps on x-axis) at each site (also on x-axis) that have sufficient power to detect a 60% change in total annual fine woody debris production between years. 17
Figure 8 Effect of trap subsample size on the ability to estimate mean functional group abundance within +/- 10% of observed functional group abundance. Analyses were performed for the five litter functional groups that averaged greater than or equal to 10% of total litter production. The x-axis represents the number of sorted traps used to estimate mean functional group abundance. The y-axis represents the percentage of 500 random samples in which the mean of the estimated functional group mass fell within +/- 10 % of the mean observed functional group mass. Data point shapes represent different years in which litter was collected. The mean percentage of total litter mass from a given functional group at a site is displayed in the lower left corner of each figure facet. Functional groups with more than 10% of total litter mass are depicted in black, while those with <= 10% are depicted in gray. 19

1 RELATED DOCUMENTS AND ACRONYMS

1.1 Reference Documents

RD[01]	NEON.DOC.000001	NEON Observatory Design
RD[02]	NEON.DOC.002652	NEON Level 1, Level 2, and Level 3 Data Products Catalog
RD[03]	NEON.DOC.000914	TOS Science Design for Plant Biomass and Productivity
RD[04]	NEON.DOC.001710	TOS Protocol and Procedure: Litterfall and Fine Woody Debris

1.2 Acronyms

Acronym	Definition
ANPP	Aboveground Net Primary Productivity
AOS	Aquatic Observation System
NLCD	National Land Cover Database
TOS	Terrestrial Observation System
TWG	Technical Working Group

2 BACKGROUND AND GOALS

2.1 Description of sample design and available data

Estimates of litterfall and fine woody debris production contribute to estimates of annual Aboveground Net Primary Productivity (ANPP) at plot, site and continental scales, and provide essential data for understanding vegetative C fluxes over time. Sampling occurs in elevated and ground trap pairs located within NEON Tower plots. Tower plots are established with 20m x 20m dimensions at short-stature vegetation sites (n=30 per site), or with 40m x 40m dimensions at tall-stature vegetation sites (n=20 per site). Two trap pairs are deployed in each 40m x 40m plot; one trap pair is deployed in 20m x 20m plots. Where the overstory is continuous, trap placement is random within the plot. In non-continuous (i.e. patchy) woody vegetation, litterfall and fine woody debris sampling is targeted to areas of the plot where woody vegetation is present. The selected sampling strategy, random or targeted, is used at all plots within a site. This protocol is not implemented at sites where average % cover of individuals > 2 m height is less than 10% (RD[04]).

Elevated traps are 0.5 m^2 (70.7 cm x 70.7 cm) and are intended to collect materials < 50 cm in length and < 2 cm butt-end diameter. Ground traps are 1.5 m^2 (3 m x 0.5 m) and target material > 50 cm in length and < 2 cm butt-end diameter, as well as large fronds that are not reliably intercepted by elevated traps.

Frequency of litter collection from elevated traps varies based on the dominant vegetation at the site. Sites dominated by deciduous species are sampled once in the spring to capture winter production then every two weeks during fall senescence. Sites dominated by evergreen species are sampled once a month throughout the year. Mixed forests with both deciduous and evergreen species are sampled according to a hybrid approach, once a month with increased fall frequency. Ground traps are sampled once a year regardless of the dominant vegetation.

Samples are collected in the field and sorted into eight functional groups:

- Leaves
- Needles
- Twigs/branches
- Woody material
- Seeds
- Flowers
- Other (lichen, mosses, frass, etc.)
- Mixed

Sorted samples are dried to constant mass and weighed. Dry mass measurements are reported in grams with a minimum precision of 0.01 g.

2.1.1 2018 Sampling reduction

Prior to the 2018 sampling season, all collections were sorted into the functional groups listed above prior to drying and measurement of dry mass. For the 2018 sampling season, sorting of litterfall to functional group was reduced to traps from a spatially balanced subsample of five plots per bout. The reduced 2018 sorting effort corresponds to sorting either 5 traps (for 20m x 20m plots) or 10 traps (for 40m x 40m plots). All other traps are collected and measured as a mixed sample.

2.1.2 Optimization analyses

Optimization analyses utilized L0 data from 2016 and 2017 from all sites that implemented the Litterfall and Fine Woody Debris protocol for the complete sampling season within a year (Table 1).

Table 1: Summary of data included in Litterfall optimization analyses.

Domain	Site	Years
1	BART	2016, 2017, 2018
1	HARV	2016, 2017, 2018
2	BLAN	2016, 2017, 2018
2	SCBI	2016, 2017, 2018
2	SERC	2016, 2017, 2018
3	JERC	2016, 2017, 2018
3	OSBS	2016, 2017, 2018
4	GUAN	2016, 2017, 2018
5	STEI	2016, 2017, 2018
5	TREE	2016, 2017, 2018
5	UNDE	2016, 2017, 2018
6	KONZ	2016, 2017, 2018
6	UKFS	2017, 2018
7	GRSM	2016, 2017, 2018
7	MLBS	2018
7	ORNL	2016, 2017, 2018
8	DELA	2016, 2017, 2018
8	LENO	2017, 2018
8	TALL	2016, 2017, 2018
10	RMNP	2018
11	CLBJ	2017, 2018
13	NIWO	2016, 2017, 2018
14	SRER	2017, 2018
16	ABBY	2017, 2018
17	SJER	2017, 2018
19	BONA	2018
19	DEJU	2017, 2018
19	HEAL	2016, 2017, 2018

A summary of the sites with data included in the Litterfall and Fine Woody Debris optimization analyses is provided in Table 2.

Table 2: Summary of sites and data included in Litterfall and Fine Woody Debris optimization analyses. Tower Plot Size = dimensions of established Tower Plots; Traps Per Plot = number of herbaceous clip harvests performed per plot per bout; Tower NLCD Class = NLCD class of vegetation in Tower footprint, * = multiple NLCD types present and dominant type is listed.

Domain	Site	Site Name	State	Tower Plot Size	Traps Per Plot	Tower NLCD Class
D01	BART	Bartlett Experimental Forest	NH	40m x 40m	2	Deciduous forest*
D01	HARV	Harvard Forest	MA	40m x 40m	2	Mixed forest*
D02	BLAN	Blandy Experimental Farm	VA	20m x 20m	1	Cultivated crops*
D02	SCBI	Smithsonian Conservation Biology Institute	VA	40m x 40m	2	Deciduous forest
D02	SERC	Smithsonian Environmental Research Center	MD	40m x 40m	2	Deciduous forest
D03	JERC	Jones Ecological Research Center	GA	40m x 40m	2	Evergreen forest*
D03	OSBS	Ordway-Swisher Biological Station	FL	40m x 40m	2	Evergreen forest
D04	GUAN	Guanica Forest	PR	40m x 40m	2	Evergreen forest
D05	STEI	Steigerwaldt Land Services	WI	40m x 40m	2	Deciduous forest
D05	TREE	Treehaven	WI	40m x 40m	2	Mixed forest*
D05	UNDE	University of Notre Dame Environmental Research Center	MI	40m x 40m	2	Deciduous forest*
D06	KONZ	Konza Prairie Biological Station	KS	20m x 20m	1	Grassland herbaceous*
D06	UKFS	University of Kansas Field Station	KS	40m x 40m	2	Deciduous forest*
D07	GRSM	Great Smoky Mountain National Park	TN	40m x 40m	2	Deciduous forest*
D07	MLBS	Mountain Lake Biological Station	VA	40m x 40m	2	Deciduous forest

Domain	Site	Site Name	State	Tower Plot Size	Traps Per Plot	Tower NLCD Class
D07	ORNL	Oak Ridge National Laboratory	TN	40m x 40m	2	Deciduous forest
D08	DELA	Dead Lake	AL	40m x 40m	2	Deciduous forest*
D08	LENO	Lenoir Landing	AL	40m x 40m	2	Woody wetlands*
D08	TALL	Talladega National Forest	AL	40m x 40m	2	Deciduous forest*
D10	RMNP	Rocky Mountain National Park	CO	40m x 40m	2	Evergreen forest*
D11	CLBJ	Caddo LBJ National Grassland	TX	20m x 20m	1	Deciduous forest*
D13	NIWO	Niwot Ridge LTER	CO	20m x 20m	1	Grassland herbaceous*
D14	SRER	Santa Rita Experimental Range	AZ	40m x 40m	2	Shrub scrub
D16	ABBY	Abby Road	WA	40m x 40m	2	Evergreen forest*
D17	SJER	San Joaquin Experimental Range	CA	40m x 40m	2	Evergreen forest*
D19	BONA	Caribou Creek / Bonanza Creek LTER	AK	40m x 40m	2	Deciduous forest*
D19	DEJU	Delta Junction	AK	40m x 40m	2	Evergreen forest*
D19	HEAL	Healy	AK	20m x 20m	1	Dwarf scrub*

2.2 Analytical Goals

Statistically rigorous analyses are needed to assess the capacity of NEON data to address Observatory goals and to guide sampling design optimization efforts. Initial spatial and temporal sampling designs for the NEON Terrestrial Observation System (TOS) and Aquatic Observation System (AOS) were developed in collaboration with Technical Working Groups (TWGs) comprised of community experts, and were captured in Science Design documents (RD[01], RD[03]). The initial designs relied on analysis of published datasets (where relevant), analysis of NEON prototype data collection efforts, and subject matter expertise.

Analysis and evaluation of the data provides a feedback loop that enables assessment of the Designs relative to Observatory goals. Moreover, results of these analyses allow the NEON TOS and AOS to effectively prioritize sampling costs in the face of uncertain future funding levels. As the NEON Observatory matures and moves into full operations, it is critical for TOS and AOS initial design assumptions to be tested with multiple years of data collected from NEON sites. At a high-level, analyses must:

- Describe where the data support reduction of TOS and AOS sampling effort without compromising NEON's ability to meet design requirements (e.g., target sites for reduction in sampling effort, spatial scales where replication is not necessary, etc.).

- Evaluate effects of design modifications on uncertainty and ability to detect year-to-year changes for key response variables (e.g., effects of reduced spatial and/or temporal sampling effort).

Specification of analyses suitable for evaluating the efficacy of the NEON TOS and AOS sampling designs required development of specific questions. Power analyses and sampling simulations were then used to determine the degree to which design modifications affect the ability of the data to address those questions. Questions and analytical approaches were developed in consultation with NEON Technical Working Groups comprised of subject matter experts selected from the ecological community. Key questions to address for Litterfall and Fine Woody Debris sampling are:

1. What aspects of the sampling design are associated with the most variability in the data? Specifically, what proportion of the variance at each site can be attributed to year, NLCD Class, plot, or trap within plot?

   Understanding how variance is partitioned across design components provides insight for modifying sampling effort by providing answers to questions such as:

   - At what spatial scale does the sampling design capture the most variation?
   - How does variance partitioning differ among sites? Are there changes to the design that can be applied across all sites? Of primary interest: Do the data universally support removing multiple traps within plots for 40m x 40m plots?
   - Does a site's NLCD Classification influence the proportion of the variance explained by different components of the design?

   Sites with greater variation among years might require a prioritization of temporal sampling. Alternatively, there would be a need to promote spatial replication at sites with smaller temporal variance components and larger variance associated with spatial components of the design.

2. Is it possible to detect meaningful interannual variation in key response variables?

   Understanding the power to detect change between years provides a direct indication of the sample sizes required to detect meaningful ecological change. In essence this analysis enables understanding minimum sample sizes needed to detect change through time; the analysis provides one means to bound the lower end of the required sample size.

3. Can specified changes in litter production be detected at each site? What sample size and frequency of sampling are needed to detect a 20% change in total production from year to year? For example, is it possible to detect 20% year-to-year changes in litter production after reducing sampling from 2 traps to 1 trap per plot, or from 20 plots per site to 10 plots per site?

   The ability to detect change from year-to-year is central to the Observatory's mission to enable understanding and forecasting change.

4. Is sorting of all litterfall samples necessary to understand trends in productivity of functional groups? Specifically, at the site scale, to what extent does our estimate of production by functional

group from a subset of traps match our estimate when all traps are sorted? What size subset of plots per site should be sorted?

This analysis provides a means to assess whether a reduced effort in sorting samples may result in the same estimate of production.

3 METHODS AND RESULTS

3.1 Data Cleaning Methods

All data cleaning has been implemented in the NEON database resulting in a higher quality dataset for use by the external community. The L0 edits included removal of duplicates, identifying and either excluding or correcting consecutive sampling bouts with overlapping start and end dates (allowing for an accurate bout trapping day variable), and addition of new data fields (boutNumber, yearBoutBegan, and eventID) to better align this data product with other NEON data products. QA records were excluded from optimization analyses.

3.2 Data Imputation Methods

To optimize the number of traps within plots and the number of plots within sites, it is ideal to analyze total production data per trap for a standardized annual period. That is, for each trap we summed production across all bouts and calculated total annual litter production on a per trap basis as inputs for the analysis. However, many sites experience regular disturbance of elevated traps such that material cannot be collected from all bouts in a given year. To address this problem, missing data were imputed using the bout mean of other traps from the same site to avoid having to exclude traps that were missing a limited number of collection events. More specifically, the data imputation process began by identifying the subset of litter trap IDs with the most complete data (between 299 and 430 days per year, with >= 90% of annual days being undisturbed). Note that the number of days per year could deviate from 365 since collection dates could have occurred a month or more either direction from December 31. For the complete traps identified above, we summed production across all bouts per year to calculate annual production (relativizing to 365 days where applicable), merged annual production with bout-level data for each trap ID by year by functional group combination, and calculated the proportion of annual production contributed by each bout for each trap ID by functional group combination. For each year by bout by site ID by functional group combo, we calculated the average annual productivity and bout proportion, merged those averages with the raw bout data by trap ID, year, bout, and functional group, and then calculated the average bout mean by multiplying the average annual productivity by the average proportion of the year represented by that bout. If the mass for a trap ID by year by bout by site by functionalGroup combo was missing, we replaced it with the calculated mean for that trap ID by year by bout by site by functional group combo. This data imputation process increased the number of traps that could be used for litter analyses from 784 to 1032. If after this data-filling exercise there were still not between 299 and 430 trapping days per year, the data for that trap by year combination were excluded from further analysis. Litter production data from traps meeting the completeness criteria above were scaled to 365 days. The 28 sites with litter traps containing at least one year's worth of data were included in these analyses. The ground traps are collected during just one bout per year, so the data imputation process described above for litter was not applicable to ground traps.

3.3 Litterfall Variance components analysis

For each site with sufficient qualifying data, a variance partitioning analysis was used to reveal the proportion of the variance in the data associated with spatial (trap, plot, and NLCD vegetation type) and temporal (year) components of the design.

3.3.1 Response variable, test specifics

The response variable is dry mass of total litter production per trap over one year's worth of collection bouts.

3.3.2 Methods

We used linear mixed effects models (lme4 package for R) to partition variance in observed annual litter production among spatial (nlcdClass, plotID, trapID) and temporal (year) random grouping factors.

Variance partitioning was carried out at each site for each litter functional group, as well as for all litter functional groups combined. The analysis was based on the following model:

$$Y \sim 1 + (1|nlcdClass) + (1|plotID : nlcdClass) + (1|trapID : plotID : nlcdClass) + (1|year)$$

Where:

- `Y` = Annual litter production (grams)
- `nlcdClass`, `plotID`, `trapID`, and `year` are all random factors among which variance was partitioned

3.3.3 Results

The dominant source of variation in litter production differed by site (Figure 1).

- ***NLCD class***: At sites with data from > 1 NLCD vegetation type litter functional group mass predictably varied according to vegetation type. For example, 'needles' functional group mass was associated with NLCD evergreen forest cover classes that include conifers, and was not abundant in the shrub/scrub cover type. NLCD class was the primary contributor to variation in per trap total litter production at Domain 06 KONZ and Domain 07 ORNL.
- ***Plot***: Plot was the primary contributor to variation in total litter production at 12 sites (BLAN, BONA, CLBJ, GRSM, HARV, HEAL, NIWO, RMNP, SCBI, SERC, SJER, and UNDE). The amount of variance explained by the `plotID` variable depended on litter functional group.
- ***Trap***: Trap was the primary contributor to variation in total litter production at 11 sites (ABBY, BART, DEJU, DELA, JERC, OSBS, SRER, STEI, TALL, TREE, and UKFS).
- ***Year***: Only a subset of sites had more than one year of L0 data currently available for analysis. For sites with 2-3 years of data, `year` was the primary contributor to variation in total litter production only at Domain 04 GUAN and Domain 08 LENO. For functional groups flowers and leaves, `year` was a primary contributor to variation in 13 sites.

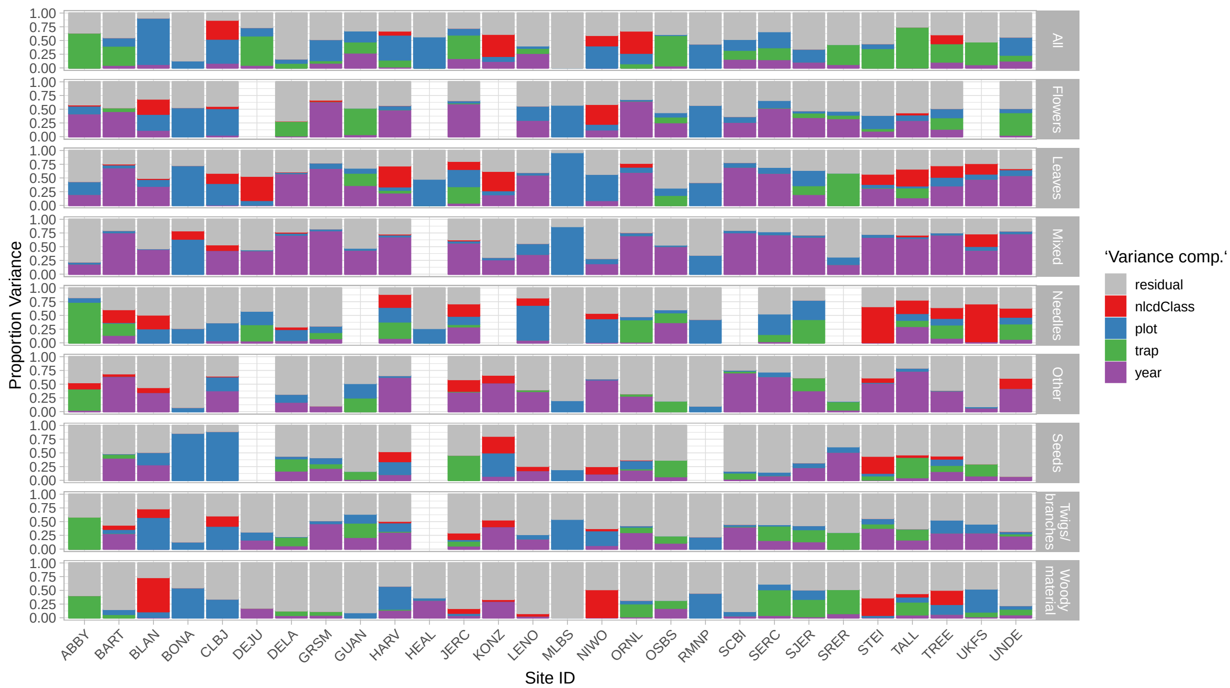

Figure 1: Proportion of the total variance explained (y-axis) by each component in the mixed model (different colors, representing random factors in the mixed model) that predicts mass for each litter functional group (facet panels) at each site (x-axis).

3.4 Litterfall Power analysis

3.4.1 Response variable, test specifics

The response variable for this test is the effect of year on dry mass of each functional group and for all functional groups combined.

3.4.2 Methods

We assessed whether we could detect a difference in litter mass by fitting the following mixed effect model to the data for each site by litter functional group combination:

$$AnnualLitterMass \sim 1 + year + (1|nlcdClass) + (1|plotID : nlcdClass) + (1|trapID : plotID : nlcdClass)$$

We then conducted a power analysis to assess whether a 20% change in year-to-year litter production could be detected. We used the parameterized mixed models to simulate 500 replicate data sets for alternative sampling designs for each site by litter functional group combination. Power for each design scenario was calculated as the proportion of replicate simulations in which we could detect the imposed 20% change in litter production between years using a critical value cutoff of t = 2 (approximately p < 0.05).

For example, a power of 0.90 indicates that 90% of the replicate simulations produced a data set in which a mixed model detected a `year` fixed effect with a p value < 0.05.

3.4.3 Results

An effect of `year` could be detected for total litter production at seventeen sites (BART, HARV, BLAN, SCBI, SERC, JERC, OSBS, GUAN, TREE, UNDE, UKFS, GRSM, LENO, CLBJ, SRER, SJER, and DEJU) (Figure 2). At eight other sites with sufficient data (STEI, KONZ, ORNL, DELA, TALL, NIWO, ABBY, and HEAL) a year effect could not be detected. The remaining three sites did not have sufficiently complete data from at least 2 years, even after filling in missing data for some trapIDs using bout means. At sites where a `year` effect was significant, these results indicate that the current sampling design is capable of detecting extant year-to-year variation in litter production.

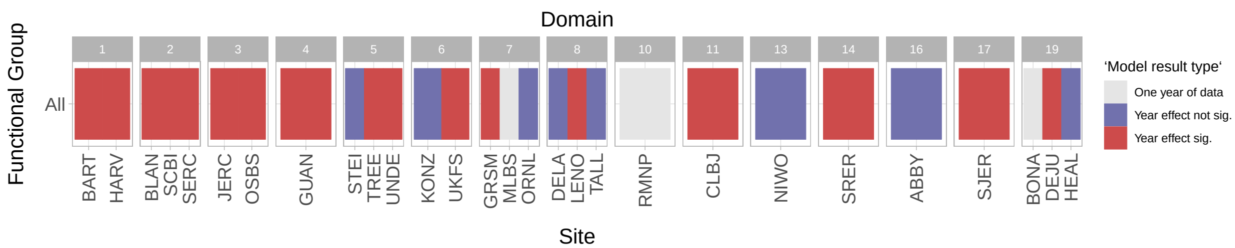

Figure 2: Heat map indicating which site (x-axis) by litter functional group (y-axis) combinations had data sufficient to fit a linear mixed effect model (above) to detect annual variation in litter production. Year was treated as a fixed effect, and red cells indicate a site/functional group combination for which an effect of year could be detected at a significance level of p < 0.05.

The power to detect an effect of `year` varied considerably across sites (Figure 3). More specifically, the power analysis indicates that out of 28 sites, 16 sites currently have sufficient power to detect a 20% year-to-year change in total litter production with the full sampling design of 20 plots with 2 traps per site. For sites with tall-stature vegetation and 40m x 40m Tower plots, reducing the number of plots per site to 10 or reducing the number of traps per plot to 1 reduces the number of sites with sufficient power from 16 to 13 (Figure 4 and Table 3). Conversely, boosting the number of plots from 20 to 30 while keeping the number of traps per plot fixed at 2 increases the power of an additional 4 sites to a sufficient level. Overall, if the number of traps per plot was kept at 2, these results indicate that:

- For 13 sites, sampling could be reduced from 20 plots per site to 10 plots per site
- For 14 sites, sampling could be reduced from 20 plots per site to 15 plots per site
- For 3 sites, sampling should be increased from 20 plots per site to 30 plots

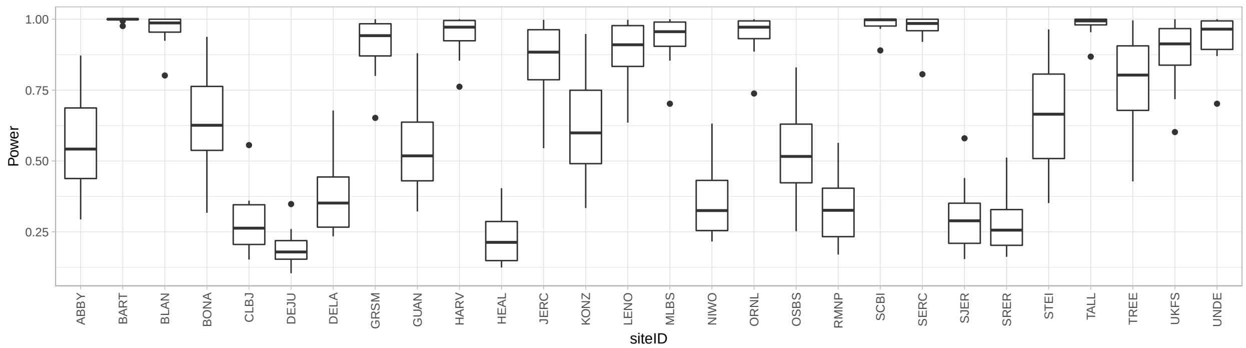

Figure 3: Power to detect a 20% change in total annual litter mass (for all design scenarios and all functional groups) between consecutive years, using a critical value of $p < 0.05$, at each site.

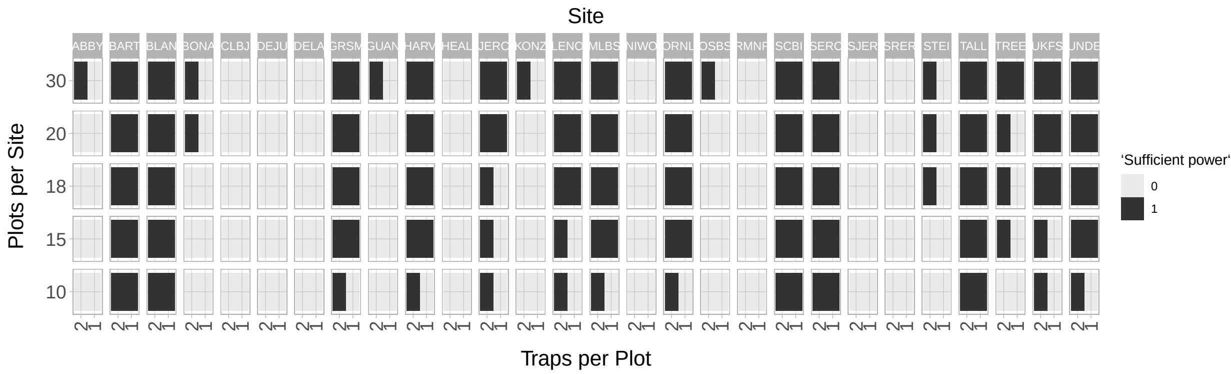

Figure 4: Black cells indicate sampling efforts (number of plots on y-axis and number of traps on x-axis) at each site (also on x-axis) that have sufficient power to detect a 20% change in total annual litter mass between years.

Table 3: The minimum number of plots per site that should be sampled. The criterion is the power to detect a 20 percent inter-annual change in total litter production in 80 percent or more of 500 simulated datasets. A value of "> 30" for the minimum number of traps indicates that there was less than 80 percent power to detect a 20 percent inter-annual change even with the most intensive sampling design (30 plots per site) examined in the simulations.

Site	Minimum plots per site with 2 traps per plot	Minimum plots per site with 1 trap per plot
ABBY	30	> 30
BART	10	10
BLAN	10	10
BONA	30	> 30
CLBJ	> 30	> 30
DEJU	> 30	> 30
DELA	> 30	> 30
GRSM	10	18
GUAN	30	> 30
HARV	10	15
HEAL	> 30	> 30
JERC	10	20
KONZ	30	> 30
LENO	10	18
MLBS	10	15
NIWO	> 30	> 30
ORNL	10	15

Site	Minimum plots per site with 2 traps per plot	Minimum plots per site with 1 trap per plot
OSBS	30	> 30
RMNP	> 30	> 30
SCBI	10	10
SERC	10	10
SJER	> 30	> 30
SRER	> 30	> 30
STEI	18	> 30
TALL	10	10
TREE	15	30
UKFS	10	20
UNDE	10	15

In conclusion, the analyses outlined here represent a framework that allows assessment of sampling sufficiency to detect a 20% change in year-to-year litter production. We recommend that sampling optimization decisions be made with a minimum of 3 years worth of data that show a consistent trend at a given site, a threshold that has been reached in 18 sites to date (Table 1).

Should sampling effort decisions become necessary based on just 2 years (7 sites) or 1 year (3 sites) of data collection at a site, analyses presented here are a basis for decision-making. However, it must be noted that the inter-annual variability observed in just two years of sampling (2016 and 2017) may not be representative of long-term variability. Considering this important caveat, the analysis reveals that:

- For 13 of the 28 sites with sufficient data for at least one year, a reduction in the number of plots per site from 20 to 10 is supported by the power analysis. At these sites, a sample size of n=10 plots still enables detecting a 20% change in annual litter production in 80% of the simulated datasets.
- At an additional 2 sites, reducing sampling effort from n=20 plots to n=18 or n=15 plots is supported by this analysis.
- There are 12 sites which have insufficient power to detect inter-annual change even with the full complement of 20 plots per site and 2 traps per plot.

3.5 Fine woody debris variance components analysis

For each site with sufficient qualifying data, this analysis partitions the variance of fine woody debris production and illuminates patterns of variability across both spatial and temporal design factors.

3.5.1 Response variable and test specifics

The response variable is dry mass of total fine woody debris production per trap over one year from 1.5m^2 ground traps. Data from smaller, 0.5m^2, elevated traps used for quantifying litterfall production are not analyzed here.

3.5.2 Methods

The same methodology that was used for the litterfall analysis was used for the fine woody debris analysis. To recapitulate, we used linear mixed effects models (lme4 package for R) to partition variance in observed annual fine woody debris mass among spatial (`nlcdClass`, `plotID`, `trapID`) and temporal (`year`) grouping factors.

Variance partitioning was carried out at each site for each fine woody debris functional group, as well as for all fine woody debris functional groups combined. The analysis was based on the following model:

$$Y \sim 1 + (1|nlcdClass) + (1|plotID : nlcdClass) + (1|trapID : plotID : nlcdClass) + (1|year)$$

Where:

- `Y` = Annual fine woody debris production (grams)
- `nlcdClass`, `plotID`, `trapID`, and `year` are all random factors among which variance was partitioned

3.5.3 Results

The residual variation in total fine woody debris was high across most sites, with `nlcdClass` and `year` having dominant importance at a subset of sites (Figure 5).

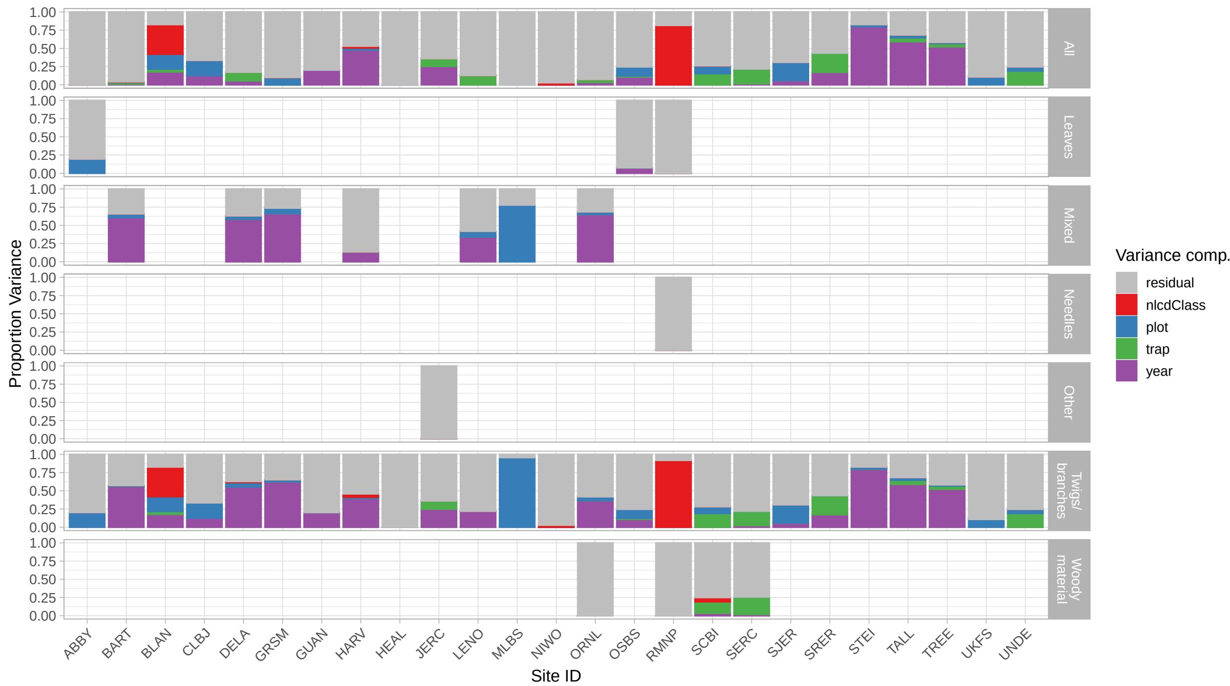

Figure 5: Proportion of the total variance explained (y-axis) by each random factor in the mixed model that predicts mass for each fine woody debris functional group (facet panels) at each site (x-axis). Colors represent different random factors in the mixed model.

3.6 Fine woody debris Power analysis

3.6.1 Response variable, test specifics

The response variable for this test is the effect of year on dry mass of each fine woody debris functional group and for all fine woody debris functional groups combined.

3.6.2 Methods

The same methodology that was used for the litter analysis was used for the fine woody debris analysis. To recapitulate, we assessed whether we could detect a difference in fine woody debris mass by fitting the following mixed effect model to the data for each site by litter fine woody debris group combination:

$$AnnualFineWoodyDebris \sim 1 + year + (1|nlcdClass) + (1|plotID : nlcdClass) + (1|trapID : plotID : nlcdC$$

We then conducted a power analysis to assess how well we can detect a 20%, 60%, or 100% change in annual fine woody debris mass in consecutive years. We used the parameterized mixed models to simulate 500 replicate data sets for alternative sampling designs for each site by fine woody debris functional group combination. Power for each design scenario was calculated as the proportion of replicate simulations in which we could detect the imposed 20%, 60%, or 100% change in standing biomass between years using a critical value cutoff of t = 2 (approximately $p < 0.05$). For example, a power of 0.90 indicates that 90% of the replicate simulations produced a data set in which a mixed model detected a year fixed effect with a p value < 0.05.

3.6.3 Results

An effect of `year` could be detected for total fine woody debris production at 12 sites (HARV, BLAN, JERC, OSBS, GUAN, STEI, TREE, DELA, TALL, CLBJ, SRER, and SJER), but a `year` effect could not be detected at 9 sites (BART, SCBI, SERC, UNDE, GRSM, ORNL, LENO, NIWO, and HEAL)(Figure 6).

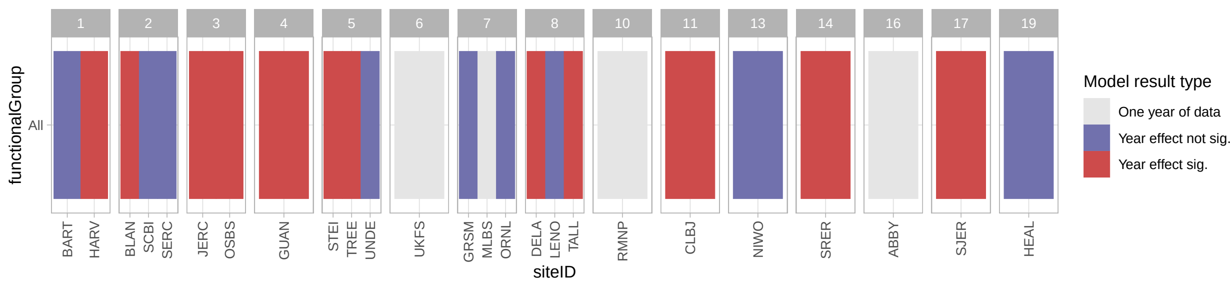

Figure 6: Heat map indicating which site (x-axis) by fine woody debris functional group (y-axis) combinations had data sufficient to fit a linear mixed effect model (above) to predict variation in annual fine woody debris production. The 'year' was treated as a fixed effect, and red cells indicate a site/functional group combination for which an effect of 'year' could be detected at a significance level of $p < 0.05$.

The power to detect a 20% interannual change in total fine woody debris production was universally low across all design scenarios and all sites, consistent with expectations of more stochastic production of twigs and branches within fine woody debris compared to production of functional groups like leaves and needles in litterfall. When we examined our power to detect a 60% interannual change in total fine woody debris production we found considerable differences among sites (Figure 7). Even with the full suite of 20 plots and 2 traps per plot we observed sufficient power to detect a 60% interannual change in total fine woody debris production in 14 out of 25 sites. If the number of sampled plots dropped from 20 per site to 10 per site then only 7 of 25 sites would have sufficient power. Overall, there is too much variability in fine woody debris production to discern more modest 20% interannual changes with the current sampling design, much less in a reduced intensity sampling design going forward. This is in contrast to litter production where potential decreases in sampling intensity may be feasible in some sites.

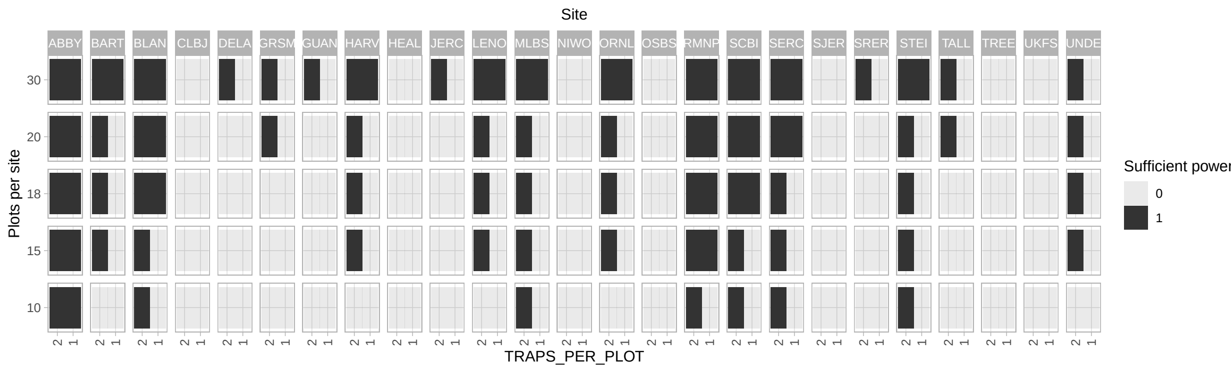

Figure 7: Black cells indicate sampling efforts (number of plots on y-axis and number of traps on x-axis) at each site (also on x-axis) that have sufficient power to detect a 60% change in total annual fine woody debris production between years.

3.7 Litterfall functional group sorting subset analysis

The Litterfall sampling design specifies that the abundance of litter functional groups should be quantified by sorting litter collected from each trap from all plots (functional groups are: leaves, needles, seeds, twigs/branches, and woody material). This analysis asks whether the relative abundance of litter functional groups observed across all plots may be estimated accurately by sorting litter from traps selected from only a subset of plots. Moreover, if sorting a subset of plots is supported, we analyze the number of traps that must be sorted in order to accurately estimate the relative abundance of the most dominant functional group across all plots.

3.7.1 Response variable, test specifics

The response variable is the estimated dry mass of litter in specific functional groups inferred from sorting and weighing functional group samples from a subset of traps.

3.7.2 Methods

We randomly sampled (with replacement) the functional group fractions of a subset of 5-40 traps from the observed trap functional group fractions, calculated the mean functional group fraction within that subset, applied that mean functional group fraction to the traps not randomly sampled, and then multiplied by the observed total mass per trap to get a mass per functional group for all trapIDs. We then calculated the mean functional group mass for each random draw. We then repeated for 500 draws and calculated what percentage of draws were within +/- 10% of the mean of the observed data that nominally had all 40 traps sorted and weighed for each functional group. We calculated for each site the mean proportion of total litter mass represented by each functional group, and discarded from further consideration functional groups that in all sites represented less than 10% of total litter mass (i.e., flowers), as well as functional groups subject to a protocol change in the early legacy data (i.e., 'Mixed' and 'Other'). For the remaining five litter functional groups (Leaves, Needles, Seeds, Twigs/branches, and Woody Material) we established a target threshold of 80% of random draws falling within +/- 10% of the mean of the observed data.

3.7.3 Results

Subsampling analyses support the following observations (Figure 8 and Table 4):

- For 15 of 19 sites with adequate data, the most abundant functional group in litter was leaves, while in the remaining four sites (ABBY, DEJU, JERC, and OSBS) the most abundant litter functional group was needles.
- Only two sites (DELA and SJER) had a dominant functional group mean mass that did not fall within the bounds of the mean observed data +/- 10% in at least 80% of 500 draws, regardless of the number of traps sorted and weighed.
- At nine sites (BART, GRSM, GUAN, LENO, ORNL, SCBI, SERC, STEI, and SRER), the estimated mass of the most abundant functional group met the 80% target threshold when sorting occurred for only 10 of 40 traps (the newly revised practice implemented in the 2018 field season).
- In the remaining 8 sites the number of traps required to produce an estimate for the dominant litter functional group that met the 80% threshold ranged from 14 to 38.
- The second-most abundant functional group of litter frequently did not meet the 80% target threshold, and the functional groups of lesser abundance rarely met the 80% target threshold.

For the most abundant functional group at a site, we can recommend a reduction in sorting effort to varying degrees for all sites. However, reducing sorting to 10 traps (as recommended in 2018) is supported in only 9 of 19 sites. For the less abundant functional groups (not included in the attached table), there is already considerable spatial variability in estimates even when all 40 traps are sorted (the current full design).

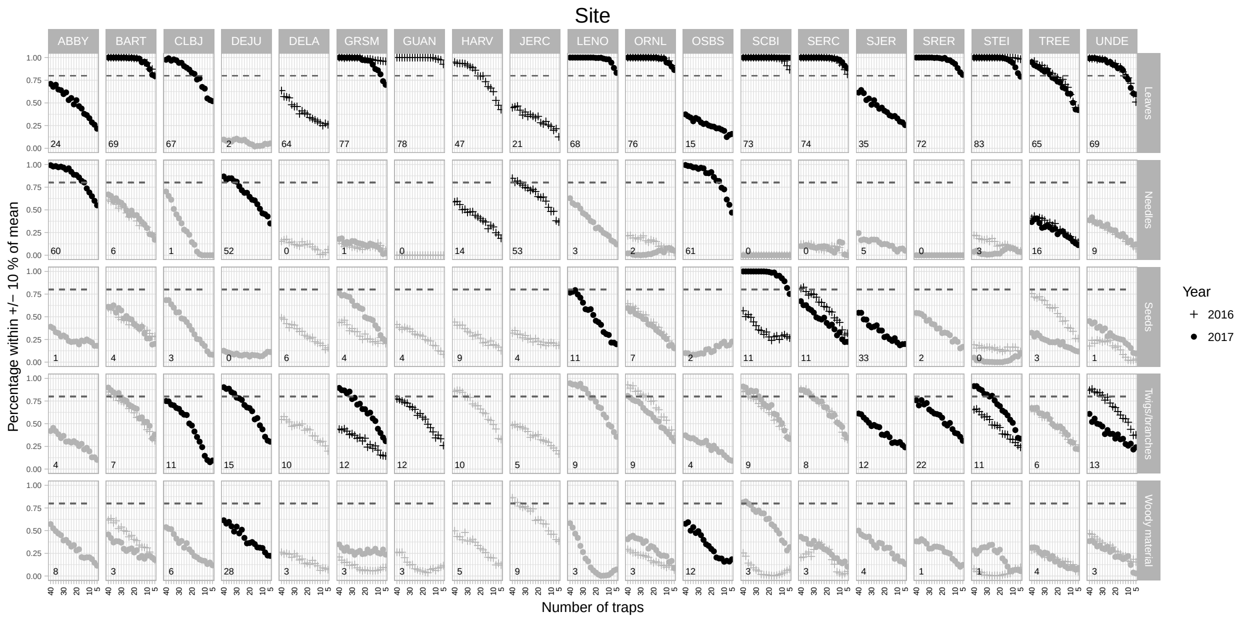

Figure 8: Effect of trap subsample size on the ability to estimate mean functional group abundance within +/- 10% of observed functional group abundance. Analyses were performed for the five litter functional groups that averaged greater than or equal to 10% of total litter production. The x-axis represents the number of sorted traps used to estimate mean functional group abundance. The y-axis represents the percentage of 500 random samples in which the mean of the estimated functional group mass fell within +/- 10 % of the mean observed functional group mass. Data point shapes represent different years in which litter was collected. The mean percentage of total litter mass from a given functional group at a site is displayed in the lower left corner of each figure facet. Functional groups with more than 10% of total litter mass are depicted in black, while those with <= 10% are depicted in gray.

Table 4: The minimum number of traps per site for which sorting and weighing is required, for the litter functional group at each site with the highest proportion of total litter mass. The criterion is the percentage of 500 random draws in which the mean of the estimated litter functional group mass fell within +/- 10 percent of the mean of the observed litter functional group mass. When the 'Trap subsample size' column entry is 'Sort All', sorting subsets of any size resulted in less than 80% of random draws with a mean estimated litter functional group mass within +/- 10% of the mean of the observed mass.

Domain	Site	Functional Group	Mean Functional Group Abundance (%)	Trap subsample size
D01	BART	Leaves	69	6
D01	HARV	Leaves	47	24
D02	SCBI	Leaves	73	5
D02	SERC	Leaves	74	5
D03	JERC	Needles	53	38

Domain	Site	Functional Group	Mean Functional Group Abundance (%)	Trap subsample size
D03	OSBS	Needles	61	14
D04	GUAN	Leaves	78	5
D05	STEI	Leaves	83	6
D05	TREE	Leaves	65	26
D05	UNDE	Leaves	69	14
D07	GRSM	Leaves	77	8
D07	ORNL	Leaves	76	5
D08	DELA	Leaves	64	Sort All
D08	LENO	Leaves	68	5
D11	CLBJ	Leaves	67	18
D14	SRER	Leaves	72	5
D16	ABBY	Needles	60	14
D17	SJER	Leaves	35	Sort All
D19	DEJU	Needles	52	30

4 DISCUSSION

The premise of these litter optimization analyses was to determine whether litter and fine woody debris measurements were sufficiently spatially homogeneous to reduce sampling or processing intensity, eliminate redundant sampling effort when possible, and maintain robust site-level litter production estimates and the ability to detect temporal trends. Our first question was whether it is possible to universally reduce sampling at all sites by eliminating collection of samples from multiple traps within each plot, and reduce litterfall sampling at all sites to 1 trap per plot. The results of the variance partitioning analysis show a high degree of spatially heterogeneous litter production within sites. The dominant components of the variance in litter production also varied considerably across sites. Most commonly the majority of the variance was associated with the spatial scale of the `plot` or `trap`. However, the `NLCD class` and `year` variables comprised the largest component of variance in total litter production at 2 sites each (Figure 1). Fine woody debris production was even more spatially heterogeneous than litter production (Figure 5), and variance partitioning results were similar to those for litter production in that the dominant component of the variance differed substantially among sites. Taken together, the variance partitioning results suggest that universally reducing sampling effort by collecting litter from only 1 trap per plot at all sites will lead to missing important sources of within-plot variation in litter production at some sites.

Given substantial spatial heterogeneity in litter production within sites, our next goal was to ascertain to what extent, if any, proposed or implemented reductions in sampling effort could save labor expenses without impacting site-level estimates or the ability to detect changes in litter production through time. The impact of each proposed or implemented change is discussed in detail below.

Based on an analysis of up to three years of data per site, reducing the number of plots sampled for litter production from 20 to 10 per site is supported for 13 of the 33 sites where litter is collected. Sampling effort may be reduced from 20 to 18 or 15 plots per site for an additional 2 sites. The data indicate that at the remaining sites there are: a) adverse impacts of reducing sampling intensity, or b) insufficient data to make analysis-informed recommendations. At 13 sites, results presented here show that an increase in sampling effort from 20 to 30 or more plots per site is necessary to detect a 20% inter-annual change in litter production. For sites where reductions in sampling effort were warranted (Figure 4), reducing the number of litter traps from 2 to 1 per plot had an impact on our ability to distinguish a 20% inter-annual change in litter production that was comparable to reducing the number of plots. Because we believe that site-specific adjustments of plot numbers present fewer complications to data end-users compared to eliminating within-plot replication on a per site basis, we provide per site plot number guidance in the 'Recommendations' section, and we maintain 2 traps per plot for all 40m x 40m plots. To reiterate:

- The plot-level sampling effort analysis revealed that the total number of plots required to detect a 20% year-to-year change in litter production varied considerably by site.
- The dataset we analyzed supports a reduced sampling effort at some sites; however, other sites require an increase in sampling effort in order to detect a 20% year-to-year change.
- All site-level sampling effort decisions would benefit from a minimum of 3 years of data before reductions in sampling effort are implemented (currently there are 18 sites with 3 years of data, 7 sites with only 2 years of data, and 3 sites with only 1 year of data).

Fine woody debris data are too variable within the dataset we analyzed to detect a 20% year-to-year change with the current sampling effort. Only if we were satisfied with detecting a 60-100% interannual change could we reduce sampling intensity from 20 to 10 plots per site at some sites.

Finally, we aimed to assess whether the data support permanently adopting a change implemented for the 2018 field season that reduced the number of litter traps with full sorting and weighing of litter functional groups from 40 to 10 per site. Reducing the extent to which litter is sorted to functional group from 40 to 10 traps per site had no adverse impact in 9 of the 33 sites where the litterfall protocol is implemented. However, at 10 sites sorting litter from only 10 traps was insufficient to meet the goal of estimating litter functional group mass to within +/- 10% of the observed mean functional group mass. At 2 sites, the analysis showed that the full complement of 40 traps per site must be sorted. At the remaining 14 sites, data were insufficient to make a recommendation. Moreover, estimation of the mass of subdominant functional groups becomes increasingly uncertain when only a subsample of traps is sorted. With this result in mind, sorting a subsample of traps should only be implemented if that higher degree of uncertainty is acceptable and we focus our attention on generating accurate estimates of the dominant functional group (typically leaves or needles).

5 RECOMMENDATIONS

5.1 Recommendations applicable to all sites:

- A minimum of 3 years of data should be collected and analyzed for any given site prior to making changes to the current sampling design. Eighteen sites currently meet this criterion.
- Retain 2 traps per plot design in all 40 m x 40 m plots where litterfall sampling occurs. Changing the number of traps per plot on a per site basis to reflect variance partitioning results may lead

to end-user confusion with respect to increasing sampling design complexity, and is therefore not recommended.

- Ground traps and elevated traps need not remain paired and may be optimized individually for a given site.
- No changes should be made to the fine woody debris ground trap sampling effort at any sites. The current design is not capable of detecting a 20% change in year-to-year fine woody production at any sites.
- Reducing total number of plots sampled according to the results of the power analysis (Table 3) is recommended over sorting a subsample of all plots.
- The sorting subset analysis should be performed again once three years of data have been collected with the modified sampling effort.

5.2 Per site recommendations

Changes in litterfall sampling effort (elevated traps only) may be appropriate for some sites (Table 5). At sites with 3 years of data, the “Optimized Trap Number” obtained from optimization analyses is put forth as the “Proposed Trap Number” going forward. The estimated reduction in effort (“Proposed Effort Decrease”) is calculated as 100 x (“Current Trap Bouts” - “Proposed Trab Bouts”) / “Current Trap Bouts”. For the sites where reductions in the number of plots to be sampled are proposed, the plot IDs to be sampled, prioritized by Morton order, are provided (Table 6). Where there are currently only 1 or 2 years of data (10 of 28 sites) we provide the “Optimized Trap Number” for informational purposes but have proposed keeping the trap numbers at their current levels until 1 or more years of additional data are available. Likewise, where optimization analyses suggested that the “Optimized Trap Number” should be greater than the “Current Trap Number” we proposed keeping the trap numbers at their current levels due to pragmatic resource limitation concerns, but if those resource limitations were relaxed the “Optimized Trap Number” indicates where those additional resources could provide the greatest benefit.

Table 5: Tentative per site plot number recommendations for elevated litter trap sampling based on analysis of 1 - 3 years of litterfall productivity data.

Domain	Site	Bout Count	Current Trap Number	Current Trap Bouts	Optimized Trap Number	Proposed Trap Number	Proposed Trap Bouts	Proposed Effort Decrease	Remarks
D01	BART	8	40	320	20	20	160	50	
D01	HARV	10	41	410	20	20	200	51	
D02	BLAN	12	12	144	20	12	144	0	
D02	SCBI	12	41	492	20	20	240	51	
D02	SERC	12	40	480	20	20	240	50	
D03	JERC	12	41	492	20	20	240	51	

Domain	Site	Bout Count	Current Trap Number	Current Trap Bouts	Optimized Trap Number	Proposed Trap Number	Proposed Trap Bouts	Proposed Effort Decrease	Remarks
D03	OSBS	13	40	520	60	40	520	0	Implement litterfall sampling in at least 11 additional plots if resources permit. New plot establishment required.
D04	GUAN	13	40	520	60	40	520	0	Implement litterfall sampling in at least 11 additional plots if resources permit. New plot establishment required.
D05	STEI	6	34	204	36	34	204	0	Tower parcel lacks sufficient space to establish any new plots.
D05	TREE	6	40	240	30	30	180	25	
D05	UNDE	6	40	240	20	20	120	50	
D06	KONZ	7	4	28	60	4	28	0	
D06	UKFS	8	40	320	20	40	320	0	Only 2 years data.
D07	GRSM	7	39	273	20	20	140	49	
D07	MLBS	7	32	224	20	32	224	0	Only 1 year of data.
D07	ORNL	8	40	320	20	20	160	50	
D08	DELA	10	42	420	> 60	40	400	5	Implement litterfall sampling in at least 11 additional plots if resources permit. New plot establishment required.
D08	LENO	9	21	189	20	21	189	0	Only 2 years of data.
D08	TALL	13	40	520	20	20	260	50	
D10	RMNP	13	18	234	> 60	18	234	0	Only 1 year of data.

Domain	Site	Bout Count	Current Trap Number	Current Trap Bouts	Optimized Trap Number	Proposed Trap Number	Proposed Trap Bouts	Proposed Effort Decrease	Remarks
D11	CLBJ	11	26	286	> 60	26	286	0	Only 2 years of data. Implement litterfall sampling in at least 1 additional plot if resources permit. New plot establishment required.
D13	NIWO	7	12	84	> 60	12	84	0	Implement litterfall sampling in at least 1 additional plot if resources permit. New plot establishment required.
D14	SRER	13	36	468	> 60	36	468	0	Only 2 years of data. Tower parcel lacks sufficient space to establish any new plots.
D16	ABBY	12	26	312	60	26	312	0	Only 2 years of data. Tower parcel lacks sufficient space to establish any new plots.
D17	SJER	13	33	429	> 60	33	429	0	Only 2 years of data. Implement litterfall sampling in at least 11 additional plots if resources permit. New plot establishment required.
D19	BONA	8	20	160	40	20	160	0	Only 1 year of data.
D19	DEJU	8	40	320	> 60	40	320	0	Only 2 years of data. Implement litterfall sampling in at least 11 additional plots if resources permit. New plot establishment required.

Domain	Site	Bout Count	Current Trap Number	Current Trap Bouts	Optimized Trap Number	Proposed Trap Number	Proposed Trap Bouts	Proposed Effort Decrease	Remarks
D19	HEAL	8	7	56	> 60	7	56	0	

Table 6: At sites where optimization analyses supported reducing the number of plots, and where there were 3 years of data available, the proposed reduced set of plot IDs to sample elevated litter traps, prioritized by Morton order.

Site	Proposed Trap Number	Plot IDs to sample
BART	20	032, 070, 033, 034, 036, 037, 071, 039, 040, 041
HARV	20	033, 034, 035, 036, 037, 038, 039, 040, 041, 042
SCBI	20	067, 045, 046, 047, 048, 049, 050, 051, 052, 053
SERC	20	043, 044, 045, 047, 048, 049, 050, 067, 051, 052
JERC	20	047, 048, 049, 050, 051, 052, 053, 054, 055, 056
TREE	30	039, 040, 041, 042, 043, 044, 045, 046, 025, 047, 048, 049, 050, 051, 052
UNDE	20	065, 037, 038, 043, 044, 045, 047, 048, 049, 050
GRSM	20	047, 048, 049, 050, 051, 052, 053, 054, 055, 056
ORNL	20	036, 037, 043, 045, 046, 049, 050, 051, 052, 053
TALL	20	051, 044, 054, 047, 042, 056, 050, 061, 046, 043

6 ACKNOWLEDGEMENTS

Thank you to the TWG and the informal internal OS Optimization group for feedback, and to Andrew Landgraf and Eric Sokol for portions of the analysis code.

7 REFERENCES

Keller, M., D. S. Schimel, W. W. Hargrove, and F. M. Hoffman. 2008. A continental strategy for the National Ecological Observatory Network. Frontiers in Ecology and the Environment 6:282-284.