International Journal of Wildland Fire 2020, 29(9), 793-806 © IAWF 2020
doi: 10.1071/WF19160_AC

Supplementary material

Quantifying how sources of uncertainty in combustible biomass propagate to prediction of wildland fire emissions

Maureen C. Kennedy[A,D], *Susan J. Prichard*[B], *Donald McKenzie*[B] *and Nancy H. F. French*[C]

[A]School of Interdisciplinary Arts and Sciences, Division of Sciences and Mathematics, University of Washington, 1900 Commerce Street, Tacoma, WA 98402, USA.

[B]School of Environmental and Forest Sciences, University of Washington, Box 352100, Seattle, WA 98195-2100, USA.

[C]Michigan Tech Research Institute, Michigan Technological University, 3600 Green Court, Suite 100, Ann Arbor, MI 48105, USA.

[D]Corresponding author. Email: mkenn@uw.edu

Supplementary Methods

Distribution estimation

To characterize the distribution of fuel loading for a given vegetation group, Prichard *et al.* used a hurdle estimation procedure because of the large proportion of zeroes for some of the fuel loading components. Hereafter, we let x_j be the values of fuel loading across all entries for the j^{th} fuel component (e.g., 1000-hr fuel loading). The first step in the model is to estimate the probability that the fuel load is zero (that fuel is not present). This is estimated as the simple proportion of observations that have zero loading for that fuel component. A continuous distribution function is then estimated for the set of observations that have greater than zero fuel loading (that have crossed the "hurdle"). The density function for the j^{th} fuel component in the k^{th} EVG ($f_{kj}(x)$) can be written as (Lachenbruch 2002):

$$f_{kj}(x,d) = \pi_{kj}^{1-d}((1-\pi_{kj})h_{kj}(x))^d, \qquad (1)$$

where *h(x)* is the estimated continuous distribution function for x>0, *d* = 1 if x non-zero, 0 if x 0, and π is the probability of observing a zero.

For the continuous portion of each fuel component in each EVT group, (Prichard *et al.* 2019) chose either a lognormal or a gamma distribution (Table S1). The lognormal probability distribution function, with parameters μ, σ, is written as:

$$h(x) = \frac{1}{\sigma x\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, \qquad (2)$$

where σ_x is the standard deviation of ln(x) and μ is the mean of ln(x).

The gamma probability distribution function, with parameters α, β, is written as:

$$h(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}. \qquad (3)$$

For our sensitivity and uncertainty analysis we sampled from the empirical hurdle distributions (Table S1) using a 2-part procedure. We let *N* be the total number of loadings that are sampled. First we took a random draw from a binomial distribution where the parameter *p* is the estimated proportion of zero-valued entries for the given fuel component and EVG and *n* is the number of randomly drawn zero entries. We then sampled *N*-*n* = *m* random draws from the estimated continuous portion of the distribution (equation 2 or 3, above). All random samples were drawn using the R Statistical Package (R Core Team 2017). This method provided samples from the independent marginal distributions of each fuel loading component.

Correlation matrices

For each EVG and set of target fuel types we estimated a Spearman's rank correlation matrix using only complete cases (only cases for which all variables of interest are entered). We let *K* be

the number of fuel types that are targeted for the analysis (e.g., $K = 6$ for flaming fuel types). The matrix X (N x K) was computed by taking N independent samples for each fuel component from the best fitting marginal empirical distribution from the fuels database for each EVG. We used the matrix method of (Iman and Conover 1982) to approximate a given rank correlation structure for sampled input data, resulting in an N x K sampled input matrix X* with correlation structure similar to the empirical correlation structure (see supplementary material for details). This method prevents combinations of fuels in individual model runs that are not likely to occur in the real world.

We adapt the method of (Iman and Conover 1982) by first estimating Spearman's rank correlation structure on the group of fuel types for the test case EVT groups, using only complete cases to estimate the correlation matrix (only cases for which all variables of interest are entered). Let C be the empirical correlation matrix (KxK). P' is the Cholesky factorization of C (generated by the R chol function), and P is the transpose of the resulting matrix.

Define the scores a(i) (i=1,…,N):

$$a(i) = \Phi^{-1}\left(\frac{i}{N+1}\right),$$

Where Φ^{-1} is the inverse of the standard normal cumulative distribution function. Then we create a new NxK matrix R, where each column is an independent random sample (without replacement) from the vector of scores A (comprised of a(i) above). We then generate the matrix R* as RP'. The matrix R* will have a rank correlation structure similar to C.

To reduce the variability in the correlation in the sampled data input matrix, a new matrix correlation matrix (T) is generated as the Spearman rank correlation matrix for the matrix R*. We then calculate Q' as the Cholesky factorization of T, with Q the transpose of Q'. We solve

for the matrix $S = PQ^{-1}$, and $R_b^* = RS'$. We then rearrange each column in the sampled data input matrix X to match the rankings in the matrix R_b^*. For example, if the first row in R_b^* is the 5th ordered statistic, then we place the 5th ordered statistic for that column in X in the first row (call the resulting matrix X*). This preserves the overall marginal distribution for each column X*, while approximating the rank correlation structure across each row in X*.

Sensitivity indices

To calculate the Sobol sensitivity indices, we used the sobolEff function in the R sensitivity package (Pujol *et al.* 2017) where a higher value implies that the model is more sensitive to that input. To calculate the Sobol sensitivity index the sampled N x K matrix X* was divided into two new (N/2) x K matrices, X_1 and X_2, where X_1 is the first N/2 rows in X* and X_2 is the last N/2 rows in X*. The sobolEff function then rearranged the X_1 and X_2 to create a new data input matrix by exchanging the columns in X_1 and X_2. This matrix rearrangement is necessary to use sampled data to estimate the Sobol sensitivity indices (Saltelli *et al.* 2010). The first-order Sobol sensitivity index characterizes the main effect of each fuel input averaged over the variability in the other fuel inputs. The total sensitivity index ranks the contribution of each variable to the variability in model predictions, including all interactions among the variables. Here we calculated both first-order and total sensitivity indices.

After Sobol rearrangement, the rank correlation structure of the final data input matrix had a poorer approximation of the original correlation matrix (C) than the original matrix X^*, and variance partitioning may not provide valid partitions in the case of correlated inputs (Jacques *et al.* 2006). We ranked the inputs by their importance in variance partitioning.

Table S1. Distributions and associated parameters for each EVG in the analysis, from Prichard et al. (2019)

EVG	Douglas-Fir Ponderosa Pine Lodgepole Pine (625)	Ponderosa Pine Forest (631)	Beech Maple Basswood (655)	Lodgepole Pine Forest and Woodland (622)	Ponderosa Pine Forest (631)	Spruce-Fir Forest and Woodland (639)
1hr	gamma(1.2,1.55) 0.04	lnorm(-1.96,1.45) 0.08	gamma(1.96,14.8) 0.01	0.90	0.22	1.12
10hr	gamma(1.50,0.46) 0.01	gamma(1.33,0.55) 0.02	gamma(2.36,1.79) 0.02	4.93	3.36	2.24
100hr	gamma(1.74,0.28) 0.06	gamma(1.52,0.32) 0.12	lnorm(1.29,0.73) 0.12	6.27	3.36	4.48
Herb	gamma(1.36,3.69) 0.01	gamma(1.44,2.79) 0.01	gamma(1.54,2.78) 0.03	0.45	0.11	0.67
Litter	gamma(1.56,0.15) 0.01	gamma(1.55,0.18) 0.003	gamma(3.05,0.31) 0.02	1.67	3.41	0.73
Shrub	gamma(1.15,0.55) 0.02	gamma(1.99,0.93) 0.10	gamma(1.87,1.05) 0.12	0.00	0.00	6.42
SCWD	0.90	17.92	6.72	lnorm(1.43,1.50) 0.204	lnorm(1.72,1.26) 0.37	lnorm(2.06,1.13) 0.21
RCWD	0.90	6.72	1.12	gamma(0.91,0.067) 0.314	lnorm(1.53,1.33) 0.49	gamma(1.28,0.14) 0.39
Duff	16.76	22.85	16.13	gamma(1.46,0.069) 0.027	gamma(1.07,0.07) 0.04	gamma(1.51,0.06) 0.01

Table S2. 95% prediction intervals (Mg ha^{-1}) when flaming fuels were sampled.

EVG	mod	Env	var	lower	upper
622	Consume	80	PM25F+S	0.26	0.78
622	Consume	97	PM25F+S	0.3	0.82
622	FOFEM	80	PM25F+S	0.25	0.98
622	FOFEM	97	PM25F+S	0.29	1.01
622	Consume	80	CO2F+S	29.22	86.28
622	Consume	97	CO2F+S	33.42	90.81
622	FOFEM	80	CO2F+S	28.05	108.67
622	FOFEM	97	CO2F+S	32.36	111.84
622	Consume	80	COF+S	1.8	5.32
622	Consume	97	COF+S	2.06	5.6
622	FOFEM	80	COF+S	1.73	6.7
622	FOFEM	97	COF+S	2	6.9
631	Consume	80	PM25F+S	0.43	0.86
631	Consume	97	PM25F+S	0.48	0.9
631	FOFEM	80	PM25F+S	0.52	1.17
631	FOFEM	97	PM25F+S	0.65	1.26
631	Consume	80	CO2F+S	47.95	94.97
631	Consume	97	CO2F+S	52.75	99.96
631	FOFEM	80	CO2F+S	58.09	130.21
631	FOFEM	97	CO2F+S	72.72	139.5
631	Consume	80	COF+S	2.96	5.86
631	Consume	97	COF+S	3.25	6.16
631	FOFEM	80	COF+S	3.58	8.03
631	FOFEM	97	COF+S	4.48	8.6
639	Consume	80	PM25F+S	0.23	0.48
639	Consume	97	PM25F+S	0.26	0.51
639	FOFEM	80	PM25F+S	0.23	0.53
639	FOFEM	97	PM25F+S	0.28	0.57
639	Consume	80	CO2F+S	26.08	54.13
639	Consume	97	CO2F+S	29.49	57.5
639	FOFEM	80	CO2F+S	25.42	58.94
639	FOFEM	97	CO2F+S	31.49	63.79
639	Consume	80	COF+S	1.33	2.76
639	Consume	97	COF+S	1.5	2.93
639	FOFEM	80	COF+S	1.3	3.01
639	FOFEM	97	COF+S	1.61	3.25

Table S3. 95% prediction intervals (Mg ha^{-1}) when smoldering fuels were sampled.

EVG	mod	Env	var	lower	upper
622	Consume	80	PM25F+S	0.16	1.01
622	Consume	97	PM25F+S	0.17	1.05
622	FOFEM	80	PM25F+S	0.21	1.68
622	FOFEM	97	PM25F+S	0.22	1.74
622	Consume	80	CO2F+S	17.9	112.02
622	Consume	97	CO2F+S	18.87	116.72
622	FOFEM	80	CO2F+S	23.21	186.45
622	FOFEM	97	CO2F+S	24.1	193.22
622	Consume	80	COF+S	1.1	6.91
622	Consume	97	COF+S	1.16	7.2
622	FOFEM	80	COF+S	1.43	11.5
622	FOFEM	97	COF+S	1.49	11.92
631	Consume	80	PM25F+S	0.11	0.81
631	Consume	97	PM25F+S	0.11	0.85
631	FOFEM	80	PM25F+S	0.13	1.39
631	FOFEM	97	PM25F+S	0.13	1.45
631	Consume	80	CO2F+S	12.14	90.34
631	Consume	97	CO2F+S	12.24	94.23
631	FOFEM	80	CO2F+S	14.56	154.27
631	FOFEM	97	CO2F+S	14.84	161.36
631	Consume	80	COF+S	0.75	5.57
631	Consume	97	COF+S	0.76	5.81
631	FOFEM	80	COF+S	0.9	9.51
631	FOFEM	97	COF+S	0.92	9.95
639	Consume	80	PM25F+S	0.19	1.37
639	Consume	97	PM25F+S	0.21	1.42
639	FOFEM	80	PM25F+S	0.2	2.37
639	FOFEM	97	PM25F+S	0.2	2.44
639	Consume	80	CO2F+S	20.79	152.18
639	Consume	97	CO2F+S	22.89	157.55
639	FOFEM	80	CO2F+S	21.75	263.56
639	FOFEM	97	CO2F+S	22.53	271.24
639	Consume	80	COF+S	1.28	9.39
639	Consume	97	COF+S	1.41	9.72
639	FOFEM	80	COF+S	1.34	16.25
639	FOFEM	97	COF+S	1.39	16.73

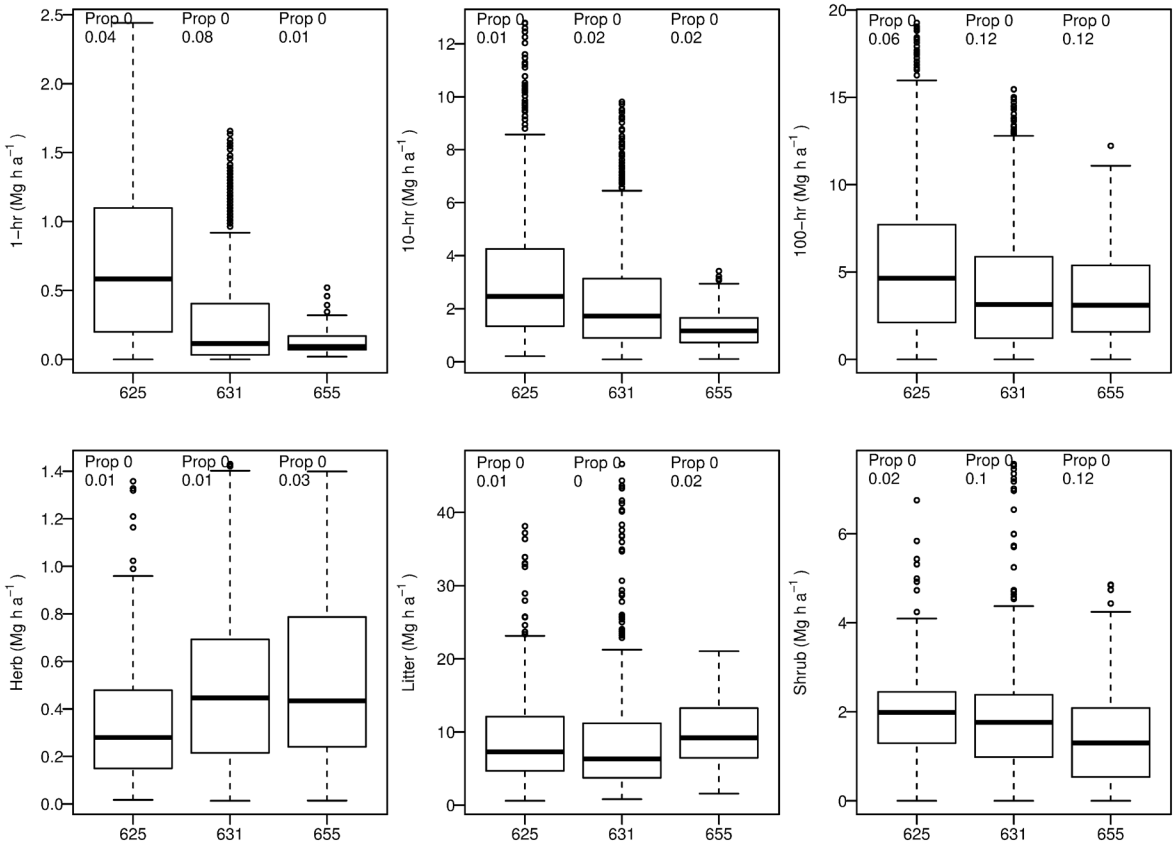

Figure S1. Empirical fuel loading distributions (middle 95%) for fuels assumed to be involved in the flaming phase of combustion. Prop 0 gives the proportion of entries with zero value. EVGs are Douglas-Fir Ponderosa Pine Lodgepole Pine (625), Ponderosa Pine Forest (631), Beech Maple Basswood (655).

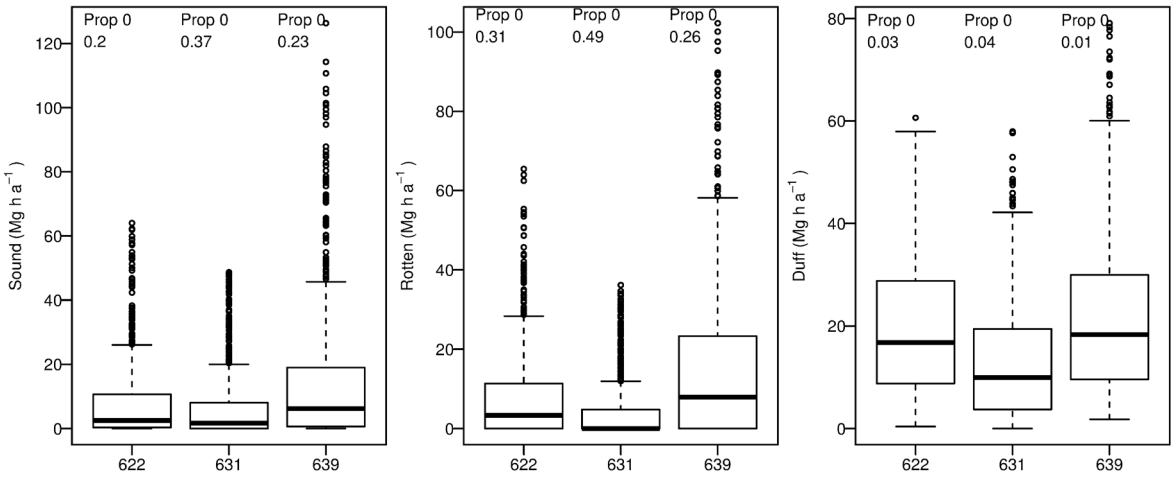

Figure S2. Empirical fuel loading distributions (middle 95%) for fuels assumed to be involved in the smoldering phase of combustion. Prop 0 gives the proportion of entries with zero value. EVGs are Lodgepole pine forest and woodland (622), Ponderosa pine forest (631), Spruce-fir forest and woodland (639).

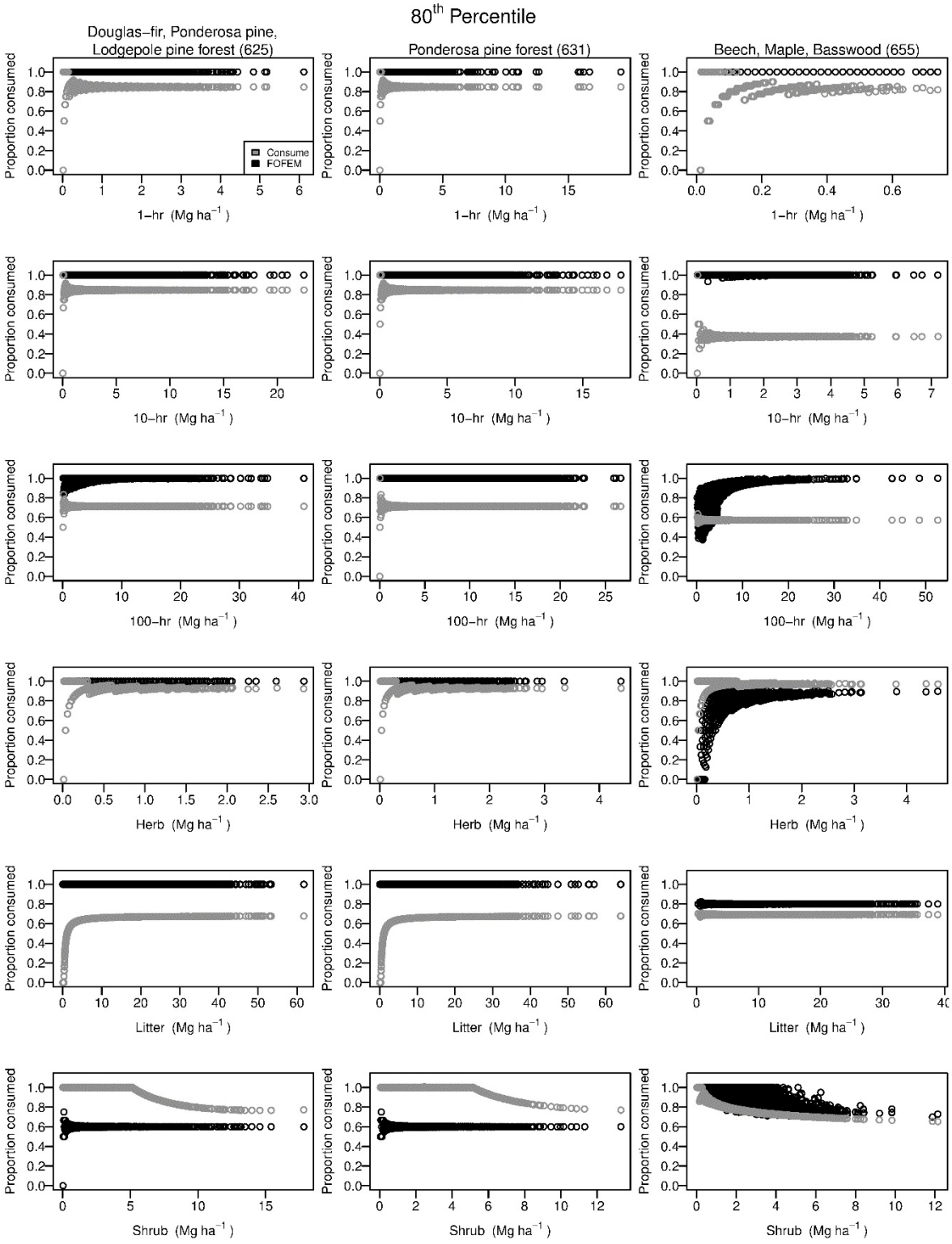

Figure S3. Proportion consumed with increased fuel loading, flaming fuels.

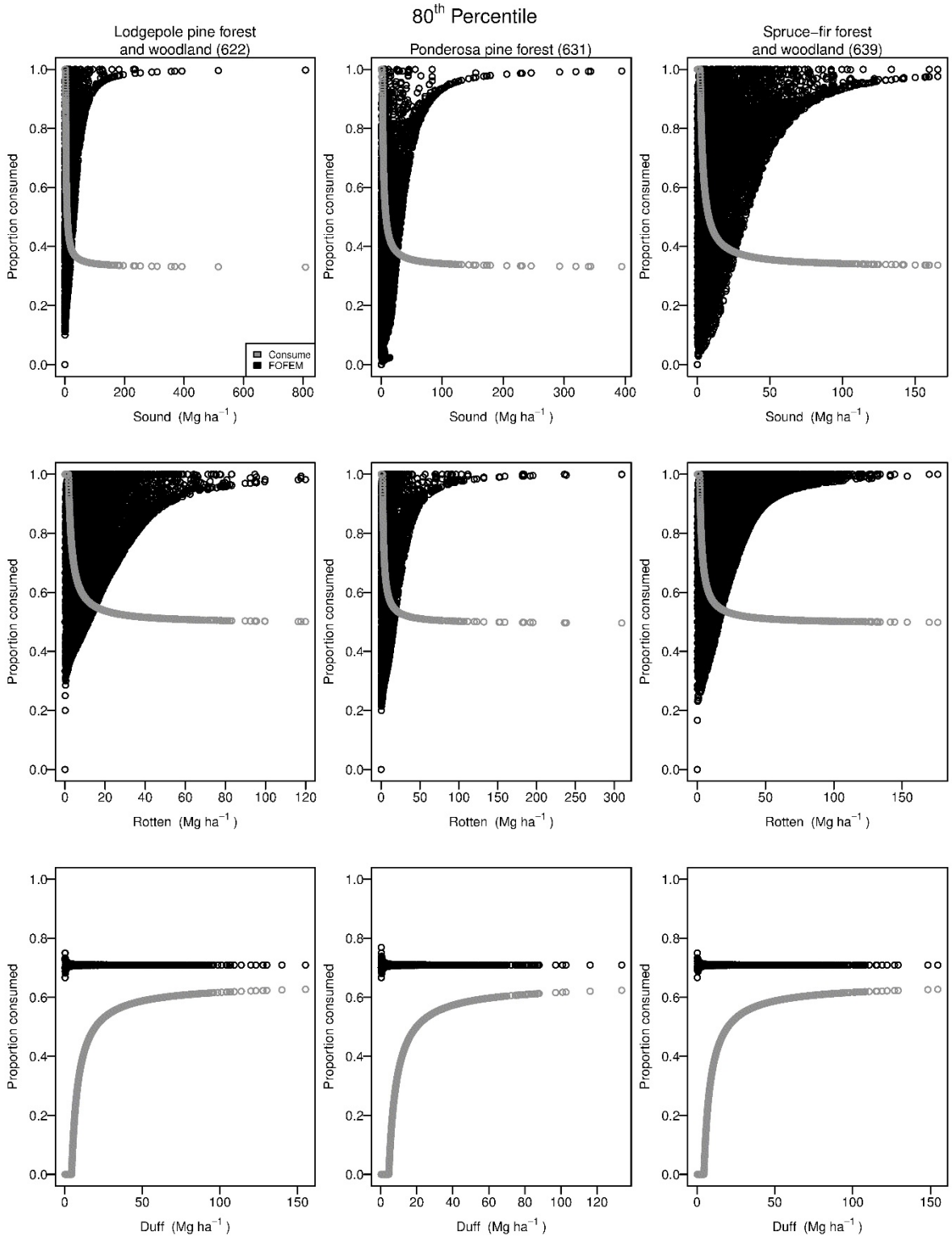

Figure S4. Proportion consumed with increased fuel loading, smoldering fuels.