**3Next-generation mapping for regional smoke management and emissions inventories: incorporating underlying uncertainty in wildland fuel characterization**

Prichard, S.J., Kennedy, M.C., Andreu, A., Eagle, P.E., and French, N.

**Introduction**

Raster maps of vegetation and biomass are increasingly used in wildfire hazard assessments (Rollins 2009, Scott et al. 2013), emissions inventories (e.g., US EPA 2017), carbon mapping (Blackard et al. 2007, Pan et al. 2011) and local to regional wildland fire and smoke management planning. Traditionally, single biomass values have been assigned to mapped pixels, often based on broadly classified vegetation type and assignment using look-up tables or nearest neighbor imputation methods (e.g., Rollins et al. 2004, Pierce et al. 2009, Keane et al. 2013, Riley et al. 2016). Ideally, maps of vegetation and combustible biomass (wildland fuels) would be based on highly replicated and accurate field measurements to calibrate relationships with remotely sensed imagery and include estimated uncertainty. However, wildland fuels are highly dynamic and variable across time and space (Keane et al. 2012), and their inherent variability generally makes it impractical to collect sufficient measurements to represent the broad geographic and structural diversity of wildland vegetation and fuels.

Fuel mapping therefore generally relies on classifications of fuels based on mapped vegetation or interpretation of remotely sensed imagery rather than measured values (Keane et al. 2015). For example, LANDFIRE maps surface and canopy fuel characteristics across the United States, assigning fuelbeds from the Fuel Characteristics Classification System (Ottmar et al. 2007), fuel loading models (Lutes et al. 2009), fire behavior fuel models (Anderson 1982) and canopy characteristics based on remotely sensed imagery and other data layers (Rollins et al. 2009). Mapped fuels are often represented as estimates of biomass per area by major vegetation type, and some representations also include estimates of biomass by fuel stratum category (e.g., canopy, shrub, herbaceous, downed wood by size class, litter and organic soil or duff). For biomass or emissions inventories (e.g., US EPA 2017) fuel characteristics are summarized at relatively coarse scales (1-km pixels) and aggregate finer-scale variability in vegetation.

Advances are being made in biomass mapping that rely on statistical relationships between field measures of biomass and remotely sensed datasets. Currently, satellite imagery such as Landsat Thematic Mapper (Landsat TM) and repeat-pass interferometric synthetic aperture radar such as Moderate Resolution Imaging Spectroradiometer (MODIS) data layers can be used to map biomass and carbon (e.g., Thurner et al. 2014). Combinations of space-borne and airborne light detection and ranging (LiDAR) offer promising advances in higher-resolution vegetation mapping because LiDAR methods can penetrate upper canopy layers and do not have oversaturation errors as with other satellite imagery associated with high above-ground biomass (Boudreau et al. 2008, Hu et al. 2016). Remotely sensed datasets are generally useful for characterizing upper canopy fuels of dense forests and shrublands but are not suitable for surface fuel characterization. Because surface fuels are highly variable and are not easily predicted from canopy infromation, there are generally high misclassification rates when canopy fuels data are used to predict surface fuels (REFS). In addition, the majority of combustible biomass of surface fuels is within the first 10-cm of the fuelbed and often composed of intermixed live and dead vegetation including shrubs, grasses, downed wood and litter that is difficult to quantify from remotely sensed imagery.

Uncertainty in biomass estimates underlies any existing fuel classification and is generally not acknowledged, much less quantified (Congalton et al. 2014; Urbanski et al. 2011). Whether empirically or modeled, each mapped pixel has associated error (e.g., the standard error of the mean). Given that point estimates are used in regional assessments and to inform management decisions, critical questions include: 1) how close to the population value do we expect the point estimate to be, 2) how much error are we incurring with a single estimate of fuel loading, and 3) how does that error propogate within our downstream models or summaries? In general it is not particularly informative to validate individual pixels in a continental-scale fuel map using plot-level data that may not represent the full pixel, and are themselves subject to sampling variability – such a validation will inevitably fail (Keane et al. 2013). Nor is it defensible to represent all instances of a fuel type by the same set of fuel loadings, as fuels are highly variable and generally vary at multiple spatial scales and over time (Keane et al. 2015).

Despite the acknowledged variability of fuels at multiple spatial scales (Keane et al. 2012), there are currently no products that incorporate uncertainty in estimating the biomass of wildland fuels in North America. There are multiple sources that contribute to the total uncertainty in model prediction, and these can be classified broadly into four groups, including model structure uncertainty, parameter estimation uncertainty, data input uncertainty, and natural variability/stochasticity (O’Neill and Gardner, 1979; Beck, 1987; Turley and Ford, 2009). For model applications that require input fuel loadings, it is important to understand how uncertainty in fuel loading values propagate to uncertainty in model predictions (Hanna 1988). Characterizing a pixel as a single point estimate masks the underlying uncertainty, resulting in pseudoprecision in model predictions. A credible estimate of the uncertainty (or variability) in fuel loading associated with a given pixel or location requires knowledge of the likely distribution of the input data.

The goal of this study was provide the capacity to characterize the distributions underlying mapped fuel inputs in order to facilitate uncertainty analysis in modeling applications. To do this, we developed a geospatial database of quantified fuel loading values to characterize the inherent variability of fuels within and across major vegetation types of the United States and Canada and to identify gaps in fuels observations. For vegetation types that had sufficient quantification of fuels by major category (e.g., canopy, shrub, herbaceous, downed wood, litter and duff), we developed empirical distribution estimates of observed fuel biomass, hereafter referred to as fuel loads, by major category. Published distributions will be useful for informing the first-generation fuels mapping that incorporates uncertainty estimates by major fuel category. Results of this study also will help inform future sampling needs to better represent the biomass of wildland fuels. Because fuel loads are a common input in fire and smoke models, uncertainty intervals informed by the fuel loading database can be used to better understand uncertainty in predictions of wildland fuel consumption and emissions and in regional to national mapping applications for biomass, carbon and emissions inventories.

**Methods**

*Fuel loading database*

The U.S. Fuel Loading Database was created as part of a JFSP-funded project (15-1-01-1 Mapping Fuels for Regional Smoke Management and Emissions Inventories). The database stores existing dry-weight biomass observations by major fuel category (i.e., tree crowns, snags, shrubs, herb, downed wood by size class, litter and duff) across the United States. Our team began by compiling existing databases and importing fuel loadings in a standard unit of measure (Mg/ha). Existing databases, including the source data for fuel loading models (Lutes et al. 2009) and LANDFIRE public source reference database (https://www.landfire.gov/lfrdb.php) were compilations of published literature and plot data. Table 1 lists the databases and provides a brief description and a source reference. We next conducted a literature review of biomass, fuel characterization and fuel consumption literature and added over 150 individual references. Minimum standards for including observations in the database were that they: 1) contained a source reference such as FIA inventory plot and sample year or journal article citation, 2) had an identifiable vegetation type, and 3) relied on field measurements as opposed to photo monitoring sites or other visual estimations.

To group fuel loading observations by vegetation type, a standard mapping classification was needed. Because LANDFIRE is a widely used mapping source of geospatial fuels and vegetation, we chose to use Existing Vegetation Type Group (www.landfire.gov/NationalProductDescriptions21.php). There are 640 existing vegetation types within LANDFIRE and a total of 207 EVT Groups. Given that the objective of the database was to quantify the distribution of fuel loads within vegetation types, we opted to use a more generic vegetation classification (EVT Group), which is provided within the LANDFIRE EVT layer, to ensure greater numbers of records within each vegetation group. It also reduced uncertainty in assigning vegetation type to each record. Most records within the database had either a general description of vegetation, a listing of major species, a Society of American Foresters or Society of Rangeland Management cover type, or a more general Forest Type (e.g., FIA plots). We developed crosswalk tables to convert cover and forest types to EVT Groups. For records that only had a general vegetation description, we individually assigned a vegetation type.

As the database was assembled, we performed a series of quality assurance and control measures. We first screened any records that did not have geospatial location. For each of these records, we attempted to assign a geospatial location and standardized existing location data into latitude and longitude (decimal degrees). In some cases, it was necessary to assign site locations based on site descriptions. Many records (n = 2470) had geospatial location but no associated vegetation type or information. For these, we overlaid record locations with the EVT Groups layer in ArcGIS and assigned a likely EVT Group based on spatial location. Due to the potential errors incurred by spatial assignment, we tagged each of these records as having spatially-assigned vegetation types. Fuel loading values were summarized into fields defined in Table 2. In many instances, simple summations were required to create summary inputs (e.g. herb load was calculated as the sum of forb and graminoid loadings and total CWD is the sum of all sound and rotten coarse wood classes).

The fuel loading database includes data from 292 sources from existing fuel loading databases and scientific literature. Entries from existing databases were presumed to be quality checked by the source agency and were not rechecked. A random selection of 15-20% of the literature sources were checked for data entry errors, including errors in unit conversions, standardizations of woody fuel size classes, and site descriptors. Due to the extent of data entry errors an additional 10% of the literature sources were checked to ensure a higher level of accuracy. *Should we report any error rate?*

For every record that had a published source reference, we obtained the source reference and included a full citation. For quality assurance and quality control, we subsampled 30% of all source references and confirmed that entered data was accurate. Most identified errors were simple rounding errors and were corrected. In a few cases, some fuel categories were missing from the inputs and were added from the published source. In other cases, fuel categories were inaccurate and corrected within the database entries.

As the database was compiled, supported fuel loading fields were expanded to accommodate various studies and approaches. Table 2 presents the fuel categories and definitions within the database. Many categories are sparsely populated but are included because they are important within particular EVT Groups. For example, moss and ground lichen are important in many boreal and subboreal vegetation types but are relatively rare in other ecosystems and associated EVT Groups.

**Table 1**: Major source databases within the Fuel Loading database.

|  |  |  |  |
| --- | --- | --- | --- |
| **Database** | **Number of records** | **Years** | **Source** |
| FFS | 128 |  | Fire and fire surrogates (McIvor) |
| FLM database | 8555 |  | Source data for the fuel loading model development (Keane) |
| FOFEM fuels | 1095 |  | Old database compiled to inform FOFEM fuel loading profiles (Reinhardt, Lutes) |
| Forest Inventory and Analysis Program | 15,061 | 2015 | David Chojnacky, University of Vermont– downloaded from - <http://web.gis.vt.edu/forestry/dwm/index.php> |
| LFRDB | 18,012 |  | LFRDB\_Public\_20100122.mdb |
| Natural Fuels Photo Series | 550 | 1998-2016 | <https://www.fs.fed.us/pnw/fera/research/fuels/photo_series> |

**Table 2**: Fuel loading database fields and definitions. To date, the database contains nearly 40,000 records and was designed to accommodate additional records as they become available.

|  |  |  |
| --- | --- | --- |
| **Field** | **Definition** | **Sample entry** |
| LFEVTGroupID | Unique ID for each EVT Group number | 693 |
| LFEVTGroup | EVT Group Name | Spruce-Fir-Hardwood Forest |
| sourceID | Unique ID for each source reference | 571 |
| Source | Source reference | Natural Fuels Photo Series Volume Iia, PMS 836 |
| studyPointID | Unique study point ID | 48753 |
| Plotname | Plot name if provided | AKHD 15 |
| State | State name | AK |
| inventoryYear | Inventory or sampling year | 2007 |
| veg\_notes | Vegetation description | Closed spruce-paper birch forest |
| us\_loading: Mg/ha | Understory crown loading (check) | 1.52 |
| ms\_loading: Mg/ha | Midstory crown loading (check) | 22.88 |
| os\_loading: Mg/ha | Overstory crown loading (check) | 91.32 |
| tree\_crown\_loading: Mg/ha | Total tree crown loading - sum of understory, midstory and overstory |  |
| tree\_loading: Mg/ha | Total aboveground tree biomass, including boles |  |
| snag\_loading: Mg/ha | Total aboveground biomass of dead trees, all decay classes | 13.56 |
| shrub\_loading: Mg/ha | Total aboveground biomass of shrubs | 3.43 |
| herb\_loading: Mg/ha | Total aboveground biomass of herbaceous plants including grasses and other nonwoody plants | 0.06 |
| 1hr\_loading: Mg/ha | 0-1/4 inch or 0.67 cm diameter wood | 0.9 |
| 10hr\_loading: Mg/ha | 1/4 to 1 inch or 0.67 to 2.54 cm diameter wood | 1.34 |
| 100hr\_loading: Mg/ha | 1-3 inch or 2.54 to 7.6 cm diameter wood | 2.46 |
| fwd\_loading: Mg/ha | Sum of fine wood (1, 10, 100-hr) wood |  |
| 1KhrS\_loading: Mg/ha | Sound wood 3 to 9 inches or 7.62 to 22.86 cm diameter (S1000hr wood) | 0.22 |
| 1KhrR\_loading: Mg/ha | Rotten wood 3 to 9 inches or 7.62 to 22.86 cm diameter (R1000hr wood) | 0 |
| 1Khr\_loading: Mg/ha | Sum of 1000hr wood |  |
| 10KhrS\_loading: Mg/ha | Sound wood 9 to 20 inches or 22.86 to 50.8 cm diameter (S10,000hr wood) | 0 |
| 10KhrR\_loading: Mg/ha | Rotten wood 9 to 20 inches or 22.86 to 50.8 cm diameter (R10,000hr wood) | 0 |
| 10Khr\_loading: Mg/ha | Sum of 10,000hr wood |  |
| GT10KhrS\_loading: Mg/ha | Sound wood > 20 inches or 50.8 cm diameter (S >10,000hr wood) |  |
| GT10KhrR\_loading: Mg/ha | Rotten wood > 20 inches or 50.8 cm diameter (R >10,000hr wood) |  |
| GT10Khr\_loading: Mg/ha | Sum of >10,000hr wood |  |
| cwd\_sound\_loading: Mg/ha | Sum of sound coarse wood (1000, 10,000, and >10,000hr wood) |  |
| cwd\_rotten\_loading: Mg/ha | Sum of rotten coarse wood (1000, 10,000, and >10,000hr wood) |  |
| cwd\_loading: Mg/ha | Sum of coarse wood (1000, 10,000, and >10,000hr wood) |  |
| moss\_loading: Mg/ha | Biomass of surface fuel cryptograms (arboreal moss not included) | 1.48 |
| lichen\_loading: Mg/ha | Biomass of ground lichens (arboreal lichens not included) | 0 |
| litter\_depth: cm | Depth of the litter layer (Oi soil layer) is included because many sources record this instead of loading. A generic bulk density value can be used to estimate biomass from this. |  |
| litter\_loading: Mg/ha | Litter biomass (Oi soil layer) | 4.68 |
| duff\_depth: cm | Depth of the duff layer (Oe and Oa soil layers) is included because many sources record this instead of loading. A generic bulk density value can be used to estimate biomass from this. |  |
| duff\_loading: Mg/ha | Duff biomass (combined upper and lower duff layers) |  |

*Fuel loading distributions*

Database values were clustered by LANDFIRE EVT Groups for estimation of loading distributions. All analyses were conducted in the R statistical program (version 3.4.1; (R Core Team 2017)), and distributions estimated using the R fitdistr package (Delignette-Muller and Dutang 2015). To identify candidate distributions for individual fuel loading categories, an exploratory data analysis (EDA) was conducted on select EVT Groups with substantial representation (> 1000 entries). Histograms, boxplots, and normal QQ plots were used to understand prominent distribution shapes and to assist in QA/QC of the database. This exploratory data analysis showed that many of the fuel types had a high proportion of values that were zero, and the fuel loading distributions tended to be right-skewed rather than symmetric. Due to these features, we chose a hurdle estimation procedure, described in the next section.

*Hurdle distribution fitting*

It is common in empirical studies of biomass (or, more commonly, abundance) for there to be excessive density at zero (Welsh et al. 1996, Lecomte et al. 2013) relative to the density functions commonly estimated for such data. Often the non-zero distribution is skewed to the right, implying a distribution such as the log-normal or the gamma distribution is more appropriate than the normal distribution (Lecomte et al. 2013). One method to contend with excessive density at zero is to estimate two models for the data, one that predicts the probability of observing a zero, and a second that models the distribution of non-zero values (Welsh et al. 1996, Lachenbruch 2002). It can be shown that the maximum likelihood estimate for the two-part model can be obtained by finding maximum likelihood estimates for each part individually (Duan et al. 1983, Welsh et al. 1996). Such a two-stage (two-part) estimation procedure has been called by many names, but we will use the nomenclature of a "hurdle model." Qualitatively, the hurdle to be crossed is having a non-zero fuel loading, and once that hurdle is crossed (x>0) a continuous distribution is estimated for the data. The density function for the jth fuel type in the kth EVT group (fkj(x)) can be written as [@Lachenbruch2002]:

fkj(x,d)= πkj1-d ((1-πkj)hkj(x))d, (1)

where h(x) is the estimated continuous distribution function (in this case, gamma or lognormal) for x>0, d = 1 if x non-zero, 0 if x 0, and π is the probability of observing a zero. For this distribution, the expected value is:

E(x) = (1-π)E(h(x)) (2)

For the continuous portion of each fuel type (in each EVT group) we estimated and compared lognormal and gamma distributions, where the lognormal probability distribution function, with parameters μ, σ, is written as:

x > 0 (3),

where σ is the standard deviation of ln (x) and μ is the mean of ln(x). The expected value of the lognormal distribution is:

(4)

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

x>0 (5)

With expected value:

E(x) = αβ (6)

Estimation of the hurdle distribution occurs in two steps. Let nkj be the total number of entries in the database for a particular fuel type (j) in a particular EVT group (k), and xkji be the ith fuel loading value for fuel type j in EVT group k. Then:

1. Estimate where I is an indicator function that takes a value of 1 if the entry has a value of 0, 0 otherwise and is the estimation probability of zero loading.

2. For the remaining non-zero entries (x), use the fitdistr function in the R fitdistrplus package to find the maximum likelihood estimates of distribution parameters for the lognormal and gamma distributions.

For initial distribution fitting we decided on a minimum of 30 *non-zero entries* required for a distribution to be estimated. This balanced our ability to estimate more distributions with the uncertainty in estimating distributions for small sample sizes (where with 95% confidence n = 30 is expected to obtain an estimated distribution with cumulative distribution function at most 0.25 away from the true cumulative distribution (Massart 1990). (See supplementary material for error analysis for distribution fitting)

Assessing distribution estimates

There are 30 total fuel types, and a total of 134 EVT groups in the current database. In general it is best practice to inspect distribution fits graphically as part of an assessment of the distribution fit, but this is untenable with so many individual distributions to be estimated in our dataset. Instead, we use several goodness of fit quantities to evaluate the distribution fits (see below; or possibly describe in supplementary material).

Kolmogorov-Smirnov test

The Kolmogorov-Smirnov (KS) test is used for the null hypothesis that a given data set follows a specified theoretical distribution. In general it is designed for situations where the full theoretical distribution is specified *a priori* and performs poorly if distribution parameter values estimated from the data are used to specify the distribution for the KS test (Lilliefors 1967). We use a Monte-Carlo (MC) procedure to estimate the p-value for the estimated distribution against the data, where a smaller p-value indicates that the observed data is statistically different than the estimated distribution (following (Lilliefors 1967)).

In the MC procedure, we calculate KS statistic for observed distribution relative to "theoretical" distribution at estimated parameter values. Then for 5000 MC replicates we take n (n=number of observed values in original distribution fit) random draws from the "theoretical" distribution at estimated parameter values. For each of these, we estimate the same theoretical distribution, then perform KS test of random to theoretical distribution at estimated parameter values. This generates 5000 KS values when the null hypothesis is true, thus a "null" distribution. The p-value is then calculated as:

(7)

where nmc is the number of MC replicates in the null distribution (5000), di is the value of the KS statistic for MC replicate i, and dobs is the observed value of the KS statistic. I is an indicator function that takes a value of 1 if the observed statistic is greater than the simulated, 0 otherwise. The sum tallies the number of simulated statistics are smaller than the observed statistic. We divide by nmc+1 because we have nmc+1total statistics (including dobs). We can then evaluate, against some α level of significance, which distributions are "fail to reject" (ftr).

For an application like this, interpretation of the KS test suffers from two issues related to sample size. At low sample sizes the test has insufficient statistical power to reject the null hypothesis; in these instances, result (ftr) does not necessarily provide evidence in favor of the estimated distribution (the null hypothesis). A large sample size presents the opposite problem: as sample size increases, the effect size necessary to reject the null hypothesis decreases. At large sample sizes this means that although the observed data are statistically different than the estimated distribution, the difference may not be of practical significance. For these reasons, we use equivalence tests to aid our interpretation of the goodness of fit between observed data and estimated distributions.

Equivalence tests

Robinson and Froese (2004) recommend an equivalence test to compare empirical data to model predictions using a two-one-sided t-test (TOST). In equivalence testing a maximum allowable error (or error tolerance) is defined, and the null hypothesis is that the observed distribution is outside of the error tolerance relative to a theoretical distribution. If the observed distribution is seen to be within the maximum error (or error tolerance), then the null hypothesis is rejected and the observed data is judged to be "equivalent" to the theoretical distribution (within the error tolerance). Here we use TOST to assess adequate matching between our observed empirical cumulative distribution of fuel type and the theoretical cdf associated with each candidate distribution. Let x(i) be the ith quantile of the empirical data distribution, and be the ith quantile of the theoretical distribution. Then the difference between the observed and theoretical cumulative distributions (xdi) is:

We then calculate as the mean distance between observed and theoretical cumulative distributions and use TOST to determine statistically if the observed and theoretical distributions differ by more than a specified error tolerance ε. This requires an error tolerance to be specified, which for our application would be a relatively arbitrarily defined threshold.

Prichard et al. (2014) use a similar equivalence procedure to evaluate the uncertainty of the fits of observed fuel consumption relative to those predicted by empirical consumption equations. For their analysis, rather than choosing a single arbitrary error threshold, they repeated the equivalence test with increasing ε until the first epsilon at which the equivalence test null hypothesis was rejected. This then defined the bound of uncertainty for that fuel type. We adapt their approach here, repeating the equivalence test for increasing error thresholds between observed and theoretical distributions for distributions estimated both with zeroes (and an offset), and distributions estimated for only values > 0. We then compare the minimum ε that rejects the null hypothesis to assess the uncertainty in the distribution estimates.

For assessing distribution estimates, the best fits would be fuel types with a KS p-value > α, and a small ε value for the equivalence procedure outlined above. We assigned broad goodness of fit classifications based on these two goodness of fit metrics (Table 3) in combination with the sample size. A fit was considered excellent if it was based on ≥ 100 entries, associated with a non-significant KS MC p-value, and had an ε value ≤ 0.05. A fit was considered good if it was based on ≥ 30 entries, had a non-significant KS p-value, and 0.05 < ε ≤ 0.15; alternatively, a fit was considered good if it has > 30 entries, and a significant KS MC p-value associated with an ε ≤ 0.05. A fit was considered poor if it has ≥ 30 entries, associated with a significant KS p-value and a large (> 0.15) ε value. The distribution was not estimated for any fuel type X EVT group combination with < 30 entries, and assigned an NA here.

Table 3: Criteria to rate quality of distribution fits. Ratings are excellent, good, or poor. Criteria are based on the number of entries in the database (higher n is evidence for a quality fit), the Kolmogorov-Smirnov Monte Carlo p-value (p>0.05 is evidence for a quality fit), and error threshold for equivalence test (ε at which the equivalence test null hypothesis is first rejected; a lower ε is evidence for a quality fit). Note that there are two criteria combinations to achieve a “Good” rating depending on the criterion by which an Excellent fit is not obtained (e.g., either the KS p-value is too small, or the ε value is too large). A poor fit is obtained when both the criteria for an Excellent fit are not met. An Excellent fit also requires n≥100.

|  |  |  |  |
| --- | --- | --- | --- |
| n | KS | ε | Rating |
| ≥100 | > 0.05 | (0,0.05] | Excellent |
|  | > 0.05 | (0.05,0.15] | Good |
|  | < 0.05 | (0, 0.05] | Good |
|  | < 0.05 | > 0.05 | Poor |
| [30,100) | > 0.05 | ≤ 0.15 | Good |
|  | < 0.05 | ≤ 0.05 | Good |
|  | < 0.05 | > 0.05 | Poor |
|  |  |  |  |

Uncertainty in distribution estimates

Finally, we use a bootstrap procedure to estimate a standard deviation for estimated distribution parameter values and to generate a 95% confidence interval for each distribution parameter value. The bootstrap estimates are generated using the bootdist function in the fitdistrplus package in R. In general the observed data are resampled with replacement and the distribution parameters estimated for each resampling of the bootstrap. This is repeated 5000 times to generate a distribution of parameter values. From this distribution a standard deviation of each estimated parameter can be calculated, and a 95% confidence interval as the 0.025 and 0.975 quantiles of the distribution.

Outliers

On preliminary exploratory analysis, and as part of our quality assurance effort, we identified extreme outliers in the database as any value > Q3 + 4\*IQR, where Q3 is the third quartile for the empirical distribution and IQR is the interquartile range (Q3-Q1). First we determined if the outlier was due to an error in rounding, units or data entry. For those values that were not entered in error, we estimated distributions both with and without the value of the identified outlier. Below we give results for distributions estimated without outliers. All distribution estimates, both with and without outliers, are presented in the Supplementary Material.

Sample distributions

For the purposes of demonstrating comparisons of distributions among fuel types and EVT groups we present here distributions for EVT groups that represent eastern mixed hardwood forests (682 yellow birch-sugar maple, 655 beech-maple-basswood and 666 eastern floodplain) and conifer forests (683 peatland, 631 ponderosa pine, and 625 Douglas-fir, ponderosa pine and lodgepole pine). For these EVT groups we present distributions for total tree loading, coarse woody debris loading (downed wood > 7.6 cm diameter), duff loading, and litter loading. These were chosen because they represent major EVT groups across the continental US and because they also had sufficient representation of the chosen fuel types to estimate distributions.

To place the fuel loading distributions in the context of smoke management and emissions estimation, we made use of two existing datasets on measured pre-burn and post-burn fuel loadings for southeastern pine and western pine understory fuels. In each case, a relatively large dataset of fuel loading observations was available from a past study on fuel consumption (Prichard et al. 2017) and could be compared with the calculated distributions for the associated EVT groups (697 loblolly pine; 631 ponderosa pine) based on much larger collections of fuel loading observations. Because fuel loading estimation is of particular importance for emissions calculations, we compared emissions estimation using 1) field-based estimates of fuel loadings from Prichard et al. (2017) and 2) distribution-based estimates of fuel loadings from the fuel loading database.

**Results**

Database coverage

Of the 198 Landfire (2014) EVT groups, the database contains records for 134 EVT groups. Of those, 68 EVT groups had sufficient entries to estimate at least one fitted distribution of a fuel category (Table S1). Based on broad physiognomic or land use category, the majority of land area in the United States is forest and woodland (32%) followed by shrublands (19%), agriculture (17%) and non-vegetated pixels (16%). However, the percentage of EVT Groups adequate for distribution fitting is highly skewed toward forest and woodland EVT groups (70%) with only 22% and 13% of shrubland and grassland EVT groups represented by fitted distributions. Of the forest and woodland EVT Groups there was higher representation of fitted distributions for coniferous forests (78%) than broadleaf forests (68%) and mixed forests (63%).

**Table X**: Percentage of total US land area by physiognomic class or agriculture and the percentage representation of each class with at least one fitted fuel loading distribution.

|  |  |  |
| --- | --- | --- |
|  | **% Area** | **% Representation** |
| Agriculture | 17% | 0% |
| Barren/developed/water | 16% | 0% |
| Forest and woodland | 32% | 70% |
| Grassland and tundra | 13% | 13% |
| Shrubland | 19% | 22% |
| Wetland and marsh | 2% | 13% |

Empirical distribution estimates

The results of all fitted distributions that were estimated are given in supplementary material. Here we give example results for representative EVT groups and fuel types. For eastern hardwood forests common in the continental US, tree loading is best represented by a gamma distribution (with nearly no zeroes). Observed values range from near 0 to near 500 Mg ha-1, with variability depending on EVT group. The example fits given here are all classified as at least “good” or “excellent.”

A common characteristic of many distributions is the high proportion of zero values and high coefficients of variation (CV). The presence and amount of CWD and duff is particularly variable across records and leads to high CV in distribution estimates. These two categories also have a wide range in reported biomass values. For example, in the 3 sample conifer forest distributions (Figure x), CWD ranges from 0 to nearly 150 Mg/ha, and the proportion of records with 0 CWD ranges from 8 to 21%. Mean CWD is greatest in Douglas-fir/ponderosa pine/lodgepole pine (DF-PP-LP) forests and lowest in peat forests. In each case, standard deviations meet or exceed estimated means (Table X). Although CWD loading is generally lower in mixed hardwood forests, ranging from 0-50 Mg/ha, the proportion of zero values is still high and with 5.6% for mixed hardwood forests dominated by yellow birch and sugar maple (YB-SM) to 14% for beech-maple-basswood (B-M-B) forests and nearly half of all records for eastern floodplain forests (47%). Mean CWD loadings are quite similar across all mixed hardwood sites as are the shapes of the distributions (Fgirue X).

The presence of duff in our three sample conifer forest types is more uniform that CWD records with relatively low proportions of zero (4.4 to 7.3%). However, loading values range widely from 0 to 250 Mg/ha in peatland forests and 0 to >100 Mg/ha in the other sample conifer forests. Mean duff loading is markedly higher in peatland forests (73 Mg/ha) compared to the other conifer forests (16-20 Mg/ha), and as with CWD, standard deviations are wide. Presence of duff is more variable in mixed hardwood sites with the proportion of records without duff ranging from 4.5 to 23%. Mean duff loading ranges from 8 to 23 Mg/ha in mixed hardwood sites, and standard deviations are either exceed or are close to the estimated mean.

Presence of litter and total aboveground biomass (TAB) of trees is much more uniform across fitted distributions than CWD and duff fuel types with less than 5% of records with zero values in sample distributions. Tree TAB ranges from 0-500 Mg/ha for conifer forests with highest mean tree biomass in DF-PP-LP forests (134 Mg/ha) compared to 75 Mg/ha for the other forest types. Mean tree TAB is comparable for mixed hardwood sites, ranging from 91 Mg/ha in eastern floodplain forets to 135 Mg/ha in YB-SM forests and high standard deviation and CV values.

**Table X**: Empirical summaries for example fuel types and EVT groups. Prop 0 gives the proportion of entries with a value of zero loading for that fuel and EVT group. n gives the number of entries > 0 in the database for that fuel and EVT group. Quartiles (Q1, Q2, Q3; 25th, 50th, and 75th percentiles, respectively) are given in Mg ha-1 and represent those quantities for all entries > 0 for that fuel and EVT group. CV is the coefficient of variation, calculated as the standard deviation divided by the mean value for all entries > 0 for that fuel and EVT group. TAB = total aboveground biomass, CWD = coarse woody debris.

| **EVT**  **ID** | **EVT Group Name** | **Fuel type** | **Prop 0** | **n** | **Q1** | **Q2** | **Q3** | **CV** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 655 | Beech-maple-basswood forest | | | | | | | |
|  |  | Tree TAB | 0 | 159 | 76.1 | 121.69 | 169.55 | 0.51 |
|  |  | CWD | 0.142 | 139 | 2.45 | 6.34 | 12.89 | 0.98 |
|  |  | Duff | 0.148 | 138 | 2.3 | 5.69 | 17.01 | 1.21 |
|  |  | Litter | 0.018 | 161 | 6.47 | 9.32 | 13.28 | 0.52 |
| 682 | Yellow birch-sugar maple forest | | | | | | |  |
|  |  | Tree TAB | 0 | 387 | 87.23 | 127.16 | 176 | 0.52 |
|  |  | CWD | 0.056 | 370 | 3.53 | 7.42 | 13.44 | 1.05 |
|  |  | Duff | 0.045 | 378 | 6.22 | 17.93 | 33.69 | 0.9 |
|  |  | Litter | 0.005 | 400 | 5.85 | 8.85 | 12.7 | 0.53 |
| 666 | Eastern floodplain forest | | | | | | |  |
|  |  | Tree TAB | 0.003 | 1513 | 32.34 | 71.3 | 129.44 | 0.86 |
|  |  | CWD | 0.474 | 813 | 1.89 | 4.69 | 11.66 | 1.63 |
|  |  | Duff | 0.23 | 1164 | 2.3 | 3.45 | 7.75 | 1.7 |
|  |  | Litter | 0.039 | 1453 | 3.15 | 6.95 | 11.7 | 0.82 |
| 683 | Peatland forest | | | | | | | |
|  |  | Tree TAB | 0 | 68 | 40.62 | 75.15 | 104.48 | 0.59 |
|  |  | CWD | 0.162 | 57 | 1.57 | 4.26 | 7.84 | 1.09 |
|  |  | Duff | 0.049 | 58 | 25.56 | 58.72 | 86.26 | 0.87 |
|  |  | Litter | 0.016 | 60 | 3.55 | 7.11 | 11.58 | 0.89 |
| 631 | Ponderosa pine forest and woodland | | | | | | |  |
|  |  | Tree TAB | 0.013 | 301 | 31.59 | 66.85 | 105.4 | 0.75 |
|  |  | CWD | 0.211 | 946 | 3.58 | 9.95 | 19.38 | 1.38 |
|  |  | Duff | 0.044 | 475 | 4.55 | 10.65 | 21.45 | 1.07 |
|  |  | Litter | 0.003 | 578 | 3.67 | 6.34 | 11.83 | 1.41 |
| 625 | Douglas-fir, ponderosa pine and lodgepole pine forest | | | | | | |  |
|  |  | Tree TAB | 0.043 | 90 | 61.84 | 108.34 | 190.28 | 0.85 |
|  |  | CWD | 0.081 | 763 | 7.54 | 17.2 | 36.64 | 1.6 |
|  |  | Duff | 0.073 | 178 | 6.56 | 14.59 | 30.05 | 0.87 |
|  |  | Litter | 0.013 | 153 | 4.67 | 7.32 | 12.57 | 0.92 |

**DISCUSSION**

Fuels maps are used as data inputs to numerous modelling applications, including consumption and emissions models such as the First Order Fire Effects Model (REF) and Consume (REF), wildland fire behavior prediction tools such as FLAMMAP or FSSIM (REF) and smoke dispersion modeling tools and frameworks such as the Wildland Fire Emissions Information System (French) and BlueSky (Larkin et al. 2014). For many modeling studies of biomass and climate, the importance of incorporating uncertainty is the foundation of simulations. For example, coarse-scale dynamic vegetation models draw inputs from probability distributions in order to model stochastic processes of fire and climate (Quillet et al. 2010, Shankar et al. 2018). Bootstrapping is also used to understand the effect of sampling variability on model predictions; for example, Gregg and Hummel (2002) used bootstrapped tree lists to evaluate impacts on Forest Vegetation Simulator projections (Gregg and Hummel 2002). However, to date, probability distributions of mapped fuels have not been available for simulating fire behavior, effects and smoke production.

For many simulation studies of wildland fire and vegetation, the importance of incorporating uncertaintyis the foundation of modeling. For example, coarse-scale dynamic vegetation models draw inputs from probability distributions in order to model stochastic processes of fire and climate (Quillet et al. 2010, Shankar et al. 2018). Models for emissions inventories are becoming increasingly sophisticated and require corresponding complexity in input fuels datasets. Despite the acknowledged variability of fuels at multiple spatial scales (Keane et al. 2012), there are currently no products that incorporate uncertainty in estimating the biomass of wildland fuels in North America.

* Characterizing uncertainty in fuels
  + Variability is high by veg classification!
  + As data expands, might be able to classify more finely (e.g., by type rather than by group), possibly reducing variability and thereby uncertainty
* Demonstrating complex patterns in variability in fuel types
* Potential applications
  + Improved emissions inventories (adding error bars to estimates)
  + Simulation modeling (draw from known distributions of fuel loadings by category)
  + Global Climate Models
  + Carbon mapping

Inform SA/UA: If it is found that emissions estimates are particularly sensitive to certain fuel categories in a major vegetation type (e.g., forest floor in boreal forests or coarse wood in temperate mixed forests), this finding could help guide future field sampling efforts and for fire and fuels managers to provide finer-scale characterization of those fuel categories (Urbanski 2014, Peterson et al. 2018). If the estimated emissions in some fuel categories are insensitive to uncertainty, then a default representation (e.g., a mean value) is likely adequate.

* + Identified data gaps – future research needs

**CONCLUSIONS**

1. Future development (maintenance, expansion to North American fuels)
2. Importance of quantifying variability and uncertainty

**References**

Anderson, H.E., 1982. Aids to determining fuel models for estimating fire behavior. INT-GTR-122. US Forest Service Intermountain Forest and Range Experiment Station, Ogden, UT.

Blackard, J.A., Finco, M.V., Helmer, E.H. et al. 2007. Mapping US forest biomass using nationwide forest inventory data and moderate resolution information. Remote Sensing of Environment 112:1658-1677.

Boudreau, J., Nelson, R., Margolis, H., Beaudoin, A., Guindon, L. and Kimes, D. 2008. Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec. Remote Sensing of Environment, 112:3876–3890.

Congalton, R.G., Gu, J., Yadav, K., Thenkabail, P. and Ozdogan, M. 2014. Global land cover mapping: a review and uncertainty analysis. Remote Sensing 6:12070-12093.

Delignette-Muller, M.L. and Dutang, C. 2015. fitdistrplus: An R Package for Fitting Distributions. Journal of Statistical Software, 64:1-34. http://www.jstatsoft.org/v64/i04

Hu, T., Su, Y., Xue, B., Liu, J., Zhao, X., Fang, J. and Guo, Q. 2016. Mapping global forest aboveground biomass with spaceborne LiDAR, optical imagery, and forest inventory data. Remote Sensing 8:565.

Keane, R.E., Burgan, R. and van Wagtendonk, J. 2001. Mapping wildland fuels for fire management across multiple scales: integrating remote sensing, GIS, and biophysical modeling. International Journal of Wildland Fire 10:301-319.

Keane, R.E. 2013. Describing wildland surface fuel loading for fire management: a review of approaches, methods and systems. International Journal of Wildland Fire 22:51-62.

Keane, R.E., Herynk, J.M., Toney, C., Urbanski, S.P., Lutes, D.C. and Ottmar, R.D. 2013. Evaluating the performance and mapping of three fuel classification system using Forest Inventory and Analysis surface fuel measurements. Forest Ecology and Management 305:248-263.

Krasnow, K., Schoennagel, T. and Veblen, T.T. 2009. Forest fuel mapping and evaluation of LANDFIRE fuel maps in Boulder County, Colorado, USA. Forest Ecology and Management 257:1603-1612.

Larkin, N.K., Raffuse, S.M. and Strand, T.M. 2014. Wildland fire emissions, carbon, and climate: US emissions inventories. Forest Ecology and Management 317:61-69.

Lutes, D.C., Keane, R.E. and Caratti, J. 2009. A surface fuel classification for estimating fire effects. International Journal of Wildland Fire 18:802-814.

Ottmar, R.D., Sandberg, D.V., Riccardi, C.L. and Prichard, S.J. 2007. An overview of the Fuel Characteristic Classification System – quantifying, classifying, and creating fuelbeds for resource planning. Canadian Journal of Forest Research 37:2383-2393.

Ottmar, R.D. 2014. Wildland fire emissions, carbon, and climate: modeling fuel consumption. Forest Ecology and Management 317:41-50.

Pierce, K.B. Jr., Ohmann, J.L., Wimberly, M.C., Gregory, M.J. and Fried, J.S. 2009. Mapping wildland fuels and forest structure for land management: a comparison of nearest neighbor imputation and other methods. Canadian Journal of Forest Research 39:1901-1916.

Peterson, J., Lahm, P., Fitch, M., George, M., Haddow, D., Melvin, M., Hyde, J. and Eberhardt, E. (editors) 2018. NWCG smoke management guide for prescribed fire. National Wildfire Coordinating Group PMS 420-2, NFES 001279. CHECK REF FORMAT

Reinhardt, E.D., Keane, R.E., and Brown, J.K. 1997. First order ﬁre effects model: FOFEM. GTR-INT-344.US Forest Service, Rocky Mountain Research Station, Ogden, Utah,

Riley, K.L., Grenfell, I.C. and Finney, M.A. 2016. Mapping forest vegetation for the western United States using modified random forests imputation of FIA forest plots. Ecosphere 7:e01472.

Rollins, M.G. 2009. LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. International Journal of Wildland Fire 18:235-249.

Rollins, M.G., Keane, R.E., and Parsons, R.A. 2004. Mapping fuels and fire regimes using remote sensing, ecosystem simulation, and gradient modeling. Ecological Applications 14:75-95.

Scott, J.H., Thompson, M.P., Calkin, D.E. 2013. A wildfire risk assessment framework for land and resource management. RMRS-GTR-315. US Forest Service Rocky Mountain Research Station. 83p.

Thurner, M.; Beer, C.; Santoro, M.; Carvalhais, N.; Wutzler, T.; Schepaschenko, D.; Shvidenko, A.; Kompter, E.; Ahrens, B.; Levick, S.R.; et al. Carbon stock and density of northern boreal and temperate forests. Global Ecology and Biogeography23:297–310.

Urbanski, S. 2014. Wildland fire emissions, carbon and climate: emissions factors. Forest Ecology and Management 317:51-60.

US EPA 2017 Profile of version 1 of the 2014 National Emissions Inventory.

<https://www.epa.gov/sites/production/files/2017-04/documents/2014neiv1_profile_final_april182017.pdf> [accessed June 21, 2018]

Beck, M. B. 1987. Water quality modeling: A review of the analysis of uncertainty. Water Resources Research 23:1393–1442.

Delignette-Muller, M., and C. Dutang. 2015. fitdistrplus: An R Package for Fitting Distributions. Journal of Statistical Software 64:1–34.

Duan, N., W. G. Manning, C. N. Morris, and J. P. Newhouse. 1983. A Comparison of Alternative Models for the Demand for Medical Care. Journal of Business & Economic Statistics 1:115–126.

Gregg, T. F., and S. Hummel. 2002. Assessing Sampling Uncertainty in FVS Projections Using a Bootstrap Resampling Method.

Hanna, S. R. 1988. Air quality model evaluation and uncertainty. JAPCA 38:406–412.

Hummel, S., and P. Cunningham. 2006. Estimating variation in a landscape simulation of forest structure. Forest Ecology and Management 228:135–144.

Lachenbruch, P. A. 2002. Analysis of data with excess zeros. Statistical Methods in Medical Research 11:297–302.

Lecomte, J. B., H. P. Benoît, S. Ancelet, M. P. Etienne, L. Bel, and E. Parent. 2013. Compound Poisson-gamma vs. delta-gamma to handle zero-inflated continuous data under a variable sampling volume. Methods in Ecology and Evolution 4:1159–1166.

Lilliefors, H. W. 1967. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. American Statistical Journal June:399–402.

Massart, P. 1990. The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. The Annals of Probability 18:1269–1283.

O’neill, R. V., and R. H. Gardner. 1979. Sources of Uncertainty in Ecological Models. Pages 447–463 in B. P. Zeigler, M. S. Elzas, G. J. Klir, and T. I. Oren, editors. Methodology in Systems Modelling and Simulation. North-Holland Publishing Company, Amsterdam. New York. Oxford, New York, US.

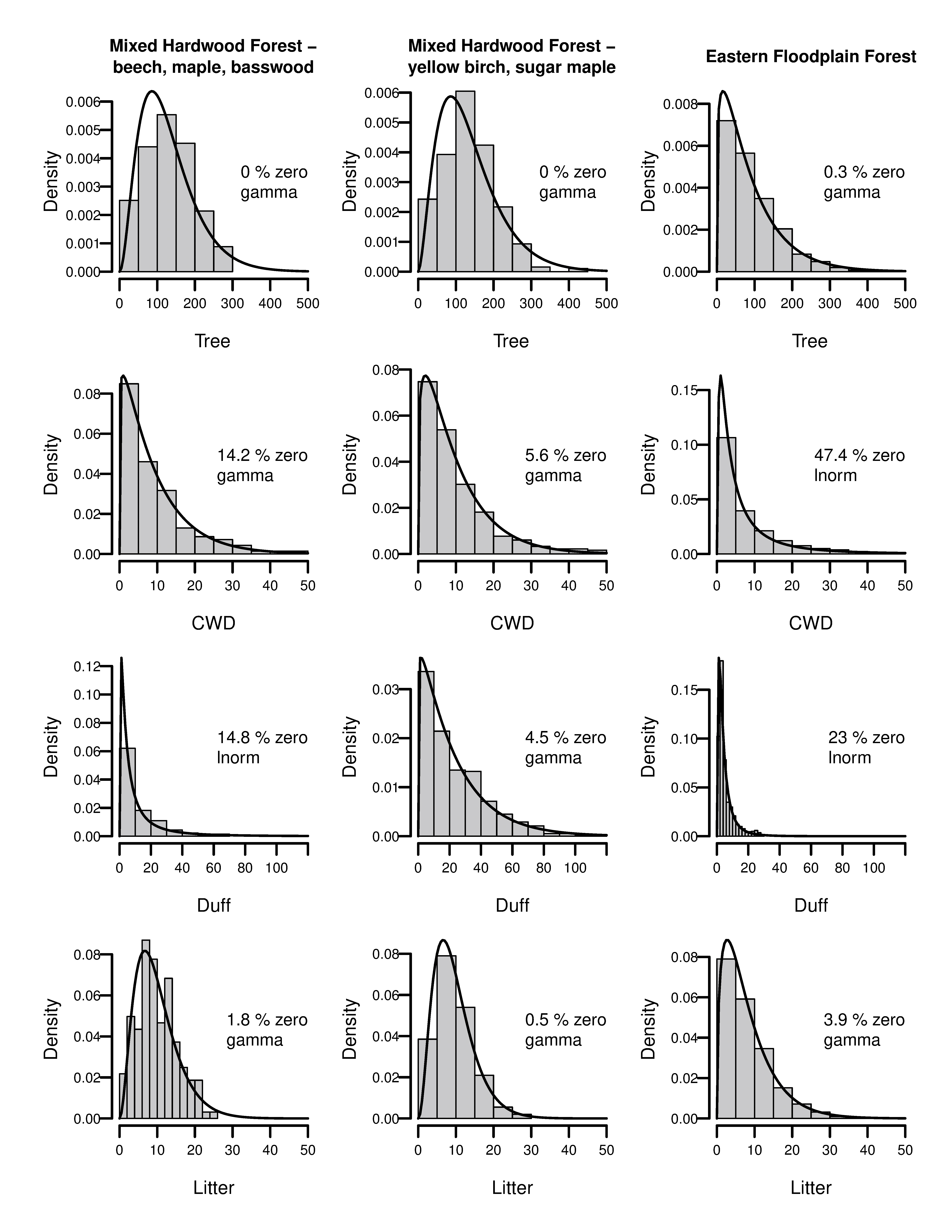
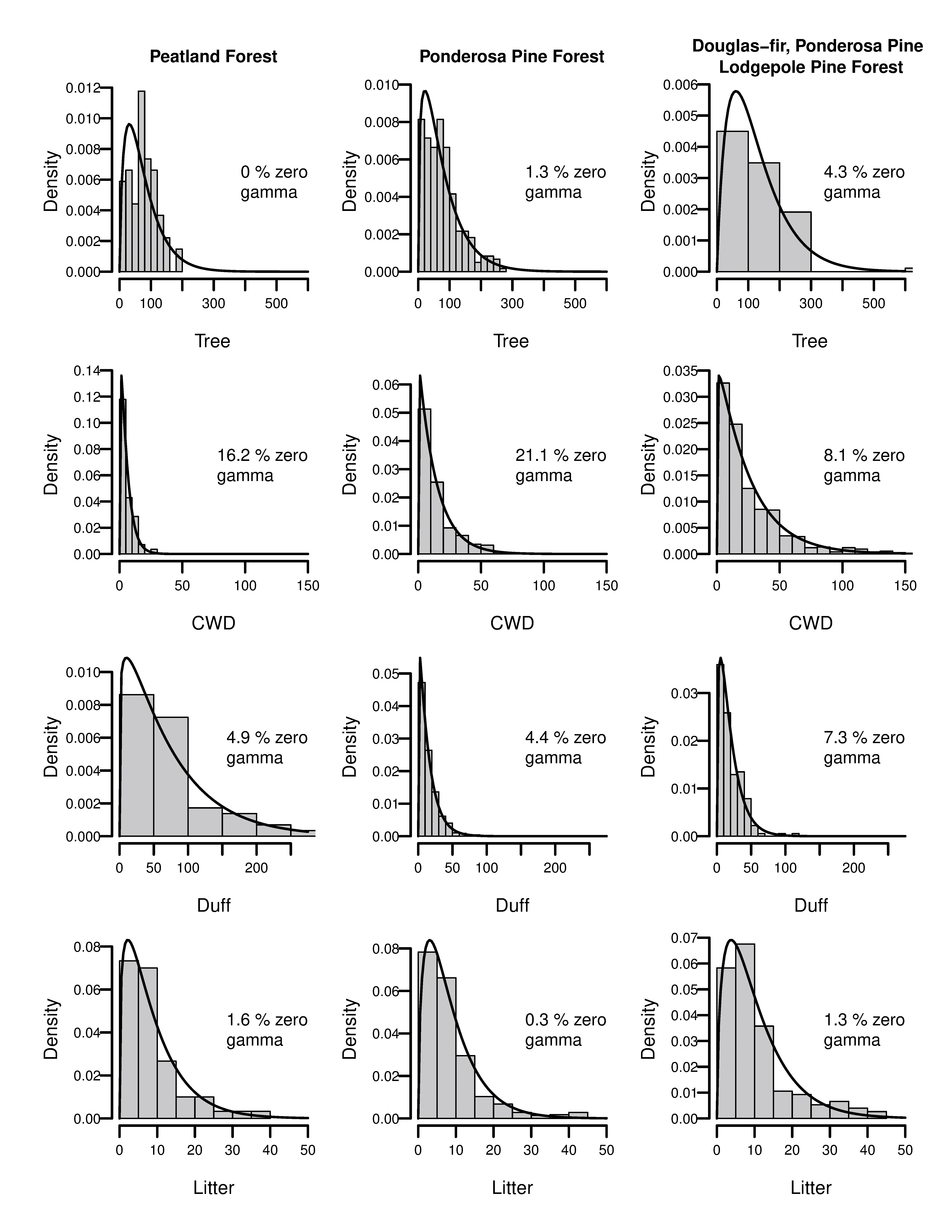
Prichard, S. J., E. C. Karau, R. D. Ottmar, M. C. Kennedy, J. B. Cronan, C. S. Wright, and R. E. Keane. 2014. Evaluation of the CONSUME and FOFEM fuel consumption models in pine and mixed hardwood forests of the eastern United States. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 44:784–795.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robinson, A. P., and R. E. Froese. 2004. Model validation using equivalence tests. Ecological Modelling 176:349–358.

Turley, M. C., and E. D. Ford. 2009. Definition and calculation of uncertainty in ecological process models. Ecological Modelling 220:1968–1983.

Welsh, A. H., R. B. Cunningham, C. F. Donnelly, and D. B. Lindenmayer. 1996. Modelling the abundance of rare species: statistical models for counts with extra zeros. Ecological Modelling 88:297–308.



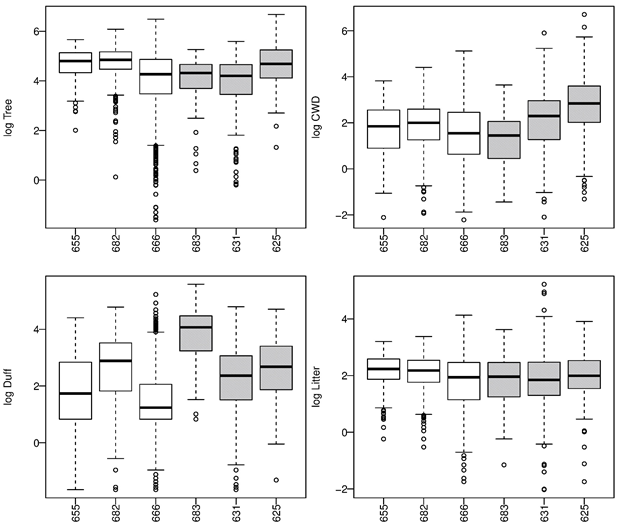


Figure: Boxplot comparison of example EVT groups and fuel loading. All values are log transformations of Mg ha-1, to aid comparisons. White boxes are hardwood, grey boxes conifer. Takeaways: for some fuels distributions are very similar (e.g., trees, litter), and for others there area some differences in distributions (CWD, duff). There is a lot of variability regardless of EVT group.

Table S1: Summary of EVT groups with sufficient records to support at least one fitted distribution. Percentage of land area is calculated as percentage of pixels represented by each EVT Group, based on the total pixels in the LANDFIRE EVT vegetation layer.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **EVT Group ID** | **EVT Group Name** | **Dominant species** | **% Land Area** | **NumStudyPoints** |
| 604 | Big Sagebrush Shrubland and Steppe | *Artemesia tridentata* | 4.428% | 144 |
| 646 | Mixedgrass Prairie |  | 3.140% | 46 |
| 613 | Desert Scrub |  | 2.048% | 53 |
| 630 | Pinyon-Juniper Woodland | *Pinus edulis*  *Juniperus* spp. | 1.736% | 626 |
| 682 | Yellow Birch-Sugar Maple Forest | *Betula alleghaniensis*  *Acer saccharum* | 1.639% | 415 |
| 660 | White Oak-Red Oak-Hickory Forest and Woodland | *Quercus alba*  *Q. rubra, Carya* spp. | 1.633% | 1161 |
| 655 | Beech-Maple-Basswood Forest | *Fagus grandifolia*  *Acer spp.*  *Tilia americana* | 1.619% | 167 |
| 617 | Grassland |  | 1.343% | 116 |
| 624 | Mesquite Woodland and Scrub | *Prosopis* spp. | 1.255% | 1309 |
| 683 | Peatland Forests | *Picea glauca*  *P. mariana* | 1.152% | 69 |
| 666 | Eastern Floodplain Forests |  | 1.072% | 1548 |
| 740 | Ruderal Forest |  | 1.064% | 347 |
| 758 | Black Spruce Forest and Woodland | Picea mariana | 0.961% | 131 |
| 631 | Ponderosa Pine Forest, Woodland and Savanna | *Pinus ponderosa* | 0.938% | 1426 |
| 661 | Chestnut Oak Forest and Woodland | *Quercus prinus* | 0.933% | 276 |
| 694 | Atlantic Swamp Forests |  | 0.883% | 235 |
| 664 | Chestnut Oak-Virginia Pine Forest and Woodland | *Quercus prinus*  *Pinus virginiana* | 0.864% | 51 |
| 685 | Pine-Hemlock-Hardwood Forest | *Pinus spp.*  *Tsuga canadensis*  *Acer* spp. | 0.825% | 145 |
| 625 | Douglas-fir-Ponderosa Pine-Lodgepole Pine Forest and Woodland | *Pseudotsuga menziesii*  *Pinus ponderosa*  *P. contorta* | 0.797% | 893 |
| 635 | Western Riparian Woodland and Shrubland |  | 0.760% | 136 |
| 639 | Spruce-Fir Forest and Woodland | *Picea glauca*  *Abies balsamea* | 0.706% | 995 |
| 662 | Post Oak Woodland and Savanna | *Quercus stellata* | 0.703% | 355 |
| 614 | Douglas-fir Forest and Woodland | *Pseudotsuga menziesii* | 0.684% | 990 |
| 693 | Spruce-Fir-Hardwood Forest | *Picea glauca*  *P. mariana*  *Abies balsamea* | 0.651% | 304 |
| 643 | Douglas-fir-Grand Fir-White Fir Forest and Woodland | *Pseudotsuga menziesii*  *Abies concolor*  *A. grandis* | 0.643% | 1056 |
| 756 | Birch-Aspen Forest | *Betula* spp.  Populus tremuloides | 0.569% | 48 |
| 615 | Douglas-fir-Western Hemlock Forest and Woodland | *Pseudotsuga menziesii*  *Tsuga heterophylla* | 0.507% | 1445 |
| 696 | Juniper-Oak | *Juniperus* spp.  *Quercus* spp. | 0.502% | 140 |
| 677 | Longleaf Pine Woodland | *Pinus palustris* | 0.458% | 114 |
| 668 | Eastern Small Stream Riparian Forests |  | 0.414% | 123 |
| 680 | Sweetgum-Water Oak Forest | *Liquidambar styraciflua*  *Quercus nigra* | 0.406% | 457 |
| 612 | Deciduous Shrubland |  | 0.392% | 93 |
| 684 | Pine Flatwoods | *Pinus palustris*  *P. taeda* | 0.381% | 1132 |
| 649 | Tallgrass Prairie | *Andropogon geradii*  *Schizachyrium scoparium* | 0.375% | 88 |
| 602 | Aspen Forest, Woodland, and Parkland | *Populus tremuloides* | 0.340% | 138 |
| 622 | Lodgepole Pine Forest and Woodland | *Pinus contorta* | 0.312% | 882 |
| 610 | Conifer-Oak Forest and Woodland |  | 0.247% | 187 |
| 690 | Shortleaf Pine-Oak Forest and Woodland | *Pinus echinata*  *Quercus* spp. | 0.245% | 137 |
| 689 | Shortleaf Pine Woodland | *Pinus echinata* | 0.225% | 125 |
| 626 | California Mixed Evergreen Forest and Woodland | *Abies spp.*  *Calocedrus decurrens*  *Pinus spp.* | 0.209% | 462 |
| 603 | Aspen-Mixed Conifer Forest and Woodland | *Populus tremuloides*  *Pinus spp.*  *Abies spp.* | 0.207% | 105 |
| 640 | Subalpine Woodland and Parkland |  | 0.200% | 281 |
| 675 | Inland Marshes and Prairies |  | 0.199% | 49 |
| 627 | Mountain Hemlock Forest and Woodland | *Tsuga mertensiana* | 0.196% | 350 |
| 638 | Sitka Spruce Forest | *Picea sitchensis* | 0.191% | 132 |
| 667 | White Oak-Beech Forest and Woodland | *Quercus alba*  *Fagus grandifolia* | 0.190% | 147 |
| 642 | Western Hemlock-Silver Fir Forest | *Tsuga heterophylla*  *Abies amabilis* | 0.182% | 1118 |
| 652 | Aspen-Birch Forest | *Populus tremuloides*  *Betula spp.* | 0.170% | 346 |
| 663 | Black Oak Woodland and Savanna | *Quercus kelloggii* | 0.136% | 203 |
| 676 | Jack Pine Forest | *Pinus banksiana* | 0.117% | 89 |
| 645 | Western Red-cedar-Western Hemlock Forest | *Thuja plicata*  *Tsuga heterophylla* | 0.117% | 368 |
| 633 | Red Fir Forest and Woodland | *Abies magnifica* | 0.109% | 169 |
| 632 | Red Alder Forest and Woodland | *Alnus rubra* | 0.103% | 90 |
| 629 | Western Oak Woodland and Savanna | *Quercus spp.* | 0.101% | 162 |
| 620 | Juniper Woodland and Savanna | *Juniperus spp.* | 0.098% | 617 |
| 665 | Red Pine-White Pine Forest and Woodland | *PInus resinosa*  *P. strobus* | 0.096% | 229 |
| 628 | Mountain Mahogany Woodland and Shrubland | *Cercocarpus spp.* | 0.089% | 69 |
| 657 | Cypress | *Taxodium distichum* | 0.053% | 55 |
| 695 | Virginia Pine Forest | *Pinus virginiana* | 0.049% | 84 |
| 659 | Bur Oak Woodland and Savanna | *Quercus macrocarpa* | 0.034% | 71 |
| 656 | Texas Live Oak | *Quercus fusiformis* | 0.034% | 236 |
| 686 | Pitch Pine Woodlands | *Pinus rigida* | 0.033% | 81 |
| 697 | Loblolly Pine Forest and Woodland | *Pinus taeda* | 0.024% | 1320 |
| 707 | Introduced Upland Vegetation-Treed |  | 0.021% | 61 |
| 679 | Maritime Forest |  | 0.019% | 79 |
| 691 | Southern Scrub Oak | *Quercus spp.* | 0.004% | 48 |
| 672 | Hammocks |  | 0.003% | 33 |
| 644 | Western Larch Forest and Woodland | *Larix occidentalis* | 0.002% | 135 |