**Next-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**

Mapping vegetation and biomass is increasingly relied upon to inform 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 wildland fire, fuels and smoke planning at regional to local scales. Traditionally, single biomass values have been assigned to mapped pixels and used as best-estimate values, often based on broadly classified vegetation type and assignment based on 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). In reality, wildland fuels are highly dynamic, with high variability across time and space (Keane et al. 2012). Given their spatial variability, it would be untenable to map fuels over an entire continent at the characteristic scales at which they vary.

Fuel mapping therefore generally relies on remotely sensed imagery and classifications of fuels based on mapped vegetation or other remotely sensed interpretations. 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. Mapped fuels are often represented as a discrete estimates of biomass by fuel type, commonly referred to as fuel loadings (e.g., shrub, herbaceous, downed wood by size class, litter and organic soil or duff). For biomass or emissions inventories (e.g., US EPA 2017) these maps summarize fuel characteristics at relatively coarse scales (1-km pixels) and aggregate finer-scale variability.

Advances are being made in biomass mapping that relies on statistically-derived relationships between field measures of biomass and remotely sensed datasets. Currently, satellite imagery such as Landsat Thematic Mapper (30m resolution) 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 spaceborne and airborne light detection and ranging (LiDAR) offer promising advances in higher-resolution vegetation mapping because LiDAR methods can penetrate canopy layers and do not have oversaturation errors as with other satellite imagery associated with higher above-ground biomass (Boudreau et al. 2008, Hu et al. 2016). All remote sensing techniques rely on field-based estimates of surface and canopy fuels and assessments of uncertainty in relationships and classifications.

Underlying fuel classifications, however, is uncertainty in fuel estimates that is generally not acknowledged, much less quantified (other ref? Congalton et al. 2014). For example, 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 – 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 these vary at multiple spatial scales and are generally independent of each other.

The best practice for mapping data with inherent spatial variability is to represent the underlying uncertainty in the base fuels layer. This measure of uncertainty then can be used to understand the reliability of the fuel loading estimates and also to evaluate how uncertainty propagates to variability in modeled response variables such as predicted wildland fire emissions which are generally highly dependent on available fuel and consumption (Larkin et al. 2014). 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.

For many modeling studies of wildland fire and vegetation, the importance of incorporating variability 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). 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.

In this study, we developed a geospatial database of measured fuel loading values to characterize the inherent variability of fuels 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 probability distributions of observed fuel biomass, hereafter referred to as fuel loads, by major category. Published probability 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.

**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 measurements by major fuel category across the United States. Our team started 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.

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 look up site locations based on site descriptions. Many records (n = 2470) had geospatial location but no assigned 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 assignment errors incurred by spatial assignment, we tagged each of these records as having auto-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. herbaceous 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).

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.

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. *Should we calculate any error rate?*

As the database was compiled, supported fuel loading fields were expanded to accommodate various studies and approaches. Table 2 presents the fuel loading fields 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.

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.

**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 - REDUCED |  | 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. All analyses were conducted in the R statistical program (R ref, version). To identify candidate distributions to fit to individual fuel loading categories an exploratory data analysis (EDA) was conducted on select EVT Groups with substantial representation (> 1000 entries) in a preliminary version of the database. Histograms, boxplots, and normal QQ plots were used to understand prominent distribution shapes and to identify any possibly outlying entries (which were then checked in QA/QC efforts).

  After the EDA we identified four candidate distributions to be estimated for fuel loading categories: normal, lognormal, gamma, and Weibull. The lognormal, gamma, and Weibull distributions are right skewed, a prominent pattern observed in the EDA. All candidate distributions have two parameters. The normal and lognormal distributions have a parameter describing the mean value and standard deviation. The lognormal distribution follows a normal distribution when in the log of the values of the random variable. The gamma and Weibull distributions are described by a shape parameter and a scale parameter. The gamma distribution and Weibull distributions in particular are known to have more flexible shape than the normal and lognormal distributions (find refs on all these). Density functions for each distribution are given (below, in supplementary, Table \_?)

The fitdistrplus (Delignette-Muller and Dutang 2015) R statistical package was used to estimate parameters for each candidate distribution for each combination of fuel loading category and EVT. A minimum threshold of (100, 200, based on references for expected error in distribution fitting relative to sample size) was enforced for distribution fitting. Distributions were not estimated for any EVT fuel loading category with fewer than N entries. The lognormal, gamma and Weibull distributions require values > 0. We excluded any values of 0 in the database or we increased all values by this small amount.

  All four candidate distributions were estimated for all EVT fuel loading categories that achieved greater than the minimum number of entries. Log-likelihood values were recorded for each distribution estimate, where (refs, details on log-likelihood as necessary). The candidate distribution that achieves the maximum log-likelihood value was chosen for the EVT fuel loading category. (where candidate distributions had relatively equivalent likelihoods (threshold?) the log-normal was preferred for (simplicity, potential to estimate joint distributions).

  Across all EVTS results were summarized by tallying the number of times a candidate distribution was chosen for each fuel loading category.

(Equivalence testing?). Or some other check on the quality of the fit. Uncertainty in the distribution fitting was quantified using (equivalence test, or a bootstrap on the parameters for the chosen distribution).  I like bootstrap confidence intervals for the parameters of the distribution.

**RESULTS**

Results

* Summary of EVTGroups and number of observations
* Identified data gaps
* Probability distributions (discuss with Maureen)
* Sensitivity analysis for wildland fire emissions by region
  1. Tally summaries
  2. Sample distributions and uncertainty
  3. Maps/visualizations

**DISCUSSION**

* 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
* Identified data gaps – future research needs

**CONCLUSIONS**

1. Future development (maintenance, expansion to North American fuels)

**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]