



# LANDFIRE

## Quality Assurance Plan

September 30, 2006

## **Document HISTORY**

<b>Tasks</b>	<b>Date</b>	<b>Responsibility</b>
Refine QA/QC procedures, and develop standardized check lists or forms	June 30, 2005	<b>Complete</b>
Fuzzy set rules for accuracy assessment	June 30, 2006	<b>Complete</b>
Refine accuracy assessment protocols	July 31, 2006	<b>Complete</b>
Reporting of LANDFIRE accuracy	Per progress	<b>Complete</b>
Case studies at the TNC application projects	Next 2 years	<b>Complete</b>
Guidelines and recommendations regarding scale of applications, minimum evaluation area, and uncertainties in data	September 30, 2006	<b>Complete</b>
Submitted to LANDFIRE Business Leads and LANDFIRE Executive Oversight Committee	September 30, 2006	<b>Complete</b>
Submit Product Quality Control Plan Implementation Plan	December 1, 2006	<b>Complete</b>

# LANDFIRE Product Quality Control and Assessment Plan

## 2006 Team Membership

Tom Bobbe	USDA Forest Service, Remote Sensing Applications Center <a href="mailto:tbobbe@fs.fed.us">tbobbe@fs.fed.us</a>
Greg Dillon	USDA Forest Service, Fire Sciences Laboratory, RMRS <a href="mailto:gdillon@fs.fed.us">gdillon@fs.fed.us</a>
Don Ohlen	USGS EROS Data Center <a href="mailto:ohlen@usgs.gov">ohlen@usgs.gov</a>
Stephen Prisley	Department of Forestry, Virginia Tech <a href="mailto:prisley@vt.edu">prisley@vt.edu</a>
Matt Reeves	USDA Forest Service, Fire Sciences Laboratory, RMRS <a href="mailto:mreeves@fs.fed.us">mreeves@fs.fed.us</a>
James L. Smith	The Nature Conservancy [ <b>2006 Chair</b> ] <a href="mailto:jim_smith@tnc.org">jim_smith@tnc.org</a>
Stephen Stehman	SUNY-Environmental Science and Forestry <a href="mailto:svstehma@syr.edu">svstehma@syr.edu</a>
Zhiliang Zhu	USGS EROS Data Center <a href="mailto:zhu@usgs.gov">zhu@usgs.gov</a>

# LANDFIRE Product Quality Control and Assessment Plan

## Executive Summary

The interagency LANDFIRE project is required to provide a Product Quality Control and Assessment Plan (PQCA Plan). The LANDFIRE Product Quality Working Team (PQWT) developed this draft work plan for review and adoption for implementation. We propose that LANDFIRE PQCA Plan include the following five components:

- 1) Quality control procedures built into LF production procedures to minimize errors and uncertainties in input data and methods.
- 2) Quantitative accuracy assessments that spatially and statistically assesses the accuracy of certain LANDFIRE map products.
- 3) Qualitative assessments and peer review of LANDFIRE products that are not amenable to quantitative assessment.
- 4) TNC LANDFIRE Application Projects that test the utility of LANDFIRE deliverables, and
- 5) Development of general guidelines for applying LANDFIRE products.

All PQCA Plan methods are in accordance with best practices in the geospatial and statistical fields.

The philosophy of the PQCA Plan is one of full and open disclosure of all information pertaining to the quality of LANDFIRE products. This information will support appropriate application of LANDFIRE products, and allow users to conduct additional product assessments beyond what is included in LANDFIRE reports. The objective of the PQCA plan, therefore, is to outline how such information will be generated and reported.

This PQCA Plan is a robust set of procedures that will require significant time and resources to complete. The PQCA Plan will be reviewed after the first complete cycle of assessments to determine if adjustments to the methods are necessary to make implementable with existing resources. The PQCA Plan will also be reviewed and revised as needed, and is owned by the PQWT.

## **Introduction**

This document represents a description of the initial LANDFIRE Product Quality Control and Assessment Plan (PQCA), and is the responsibility of the LANDFIRE Product Quality Working Team. This is a living document, and will be reviewed and revised as needed.

LANDFIRE technical teams will assess the quality of deliverable layers in a variety of ways, depending on the fundamental characteristics of each layer. Quantitative accuracy assessments will be performed on layers that are directly mapped from imagery or field plots because reference data exists for the direct comparison, specifically Environmental Site Potential (ESP), Existing Vegetation Type (EVT) and canopy characteristics layers. Other LANDFIRE products will be assessed qualitatively because they are derived through simulation or modeling, such as vegetation models, fire regimes, and fuels. In both cases, technical teams will develop and publish reports summarizing the quality of each LANDFIRE deliverable layer. The PQWT will review reports and submit them to the LANDFIRE Business Team. The philosophy of the LANDFIRE Product Quality Team is to provide as much information about the characteristics of LANDFIRE products as possible to support appropriate application of the products, and to allow for additional product assessments beyond what is included in LANDFIRE reports.

The plan is composed of five sections:

1. Quality Control Procedures,
2. Quantitative Accuracy Assessment,
3. Qualitative Assessment and Peer Review of Deliverables,
4. Application Projects, and
5. Guidelines for Applications Scale.

Following these sections are numerous appendices with additional details, forms or examples.

Questions about this plan should be directed to the LANDFIRE helpdesk ([helpdesk@landfire.gov](mailto:helpdesk@landfire.gov)).

## **Quality Control Procedures**

The LANDFIRE Team consists of nine individual production teams:

- 1) LANDFIRE Reference Data Base (MFSL/SEM)
- 2) Vegetation Modeling (TNC)
- 3) Biophysical Gradients (MFSL/EDC)
- 4) Potential Vegetation (MFSL)
- 5) Vegetation Rectification (MFSL)

- 6) Existing Vegetation (EDC)
- 7) Fuels (MFSL)
- 8) Fire Regime (MFSL)
- 9) Data Library (MFSL)

Existing QC procedures were developed by individual production teams to ensure that inherent errors in input data and methods are minimized, and the highest quality results possible are produced using the current LANDFIRE methodology. QC procedures are developed and implemented separately by each of the LANDFIRE production teams. Current QC procedures are described in an *Appendix A*. QC procedures were implemented at the beginning of the production process, and are under constant review. The information contained in Appendix A is current, but it will be updated in the spring of 2007. It is beyond the scope of this plan to detail all the production steps in the LANDFIRE process, but all LANDFIRE procedural steps are documented by the production teams in Procedure Tables that are available separately.

The PQWT will work with LANDFIRE Technical Team and individual production teams to ensure that the QC procedures are followed consistently and information about QC are provided as specified. Towards these goals, we recommend that task area teams consider the following two actions:

- Provide full documentation of all logic and data consistency checks, data integrity, and QC that occur throughout the LANDFIRE production process. This QC documentation will be provided to the PQWT, included in this plan, and updated as required.
- Standardize all QC steps so that QC processes and philosophy are consistently applied. Documents or tools that ensure this consistency will be developed and utilized by each production team, such as lists, diagrams, or checklists

### **Quantitative Accuracy Assessment Plan**

Some type of quantitative accuracy assessment will be conducted on the following LANDFIRE Products:

1. Environmental Site Potential (ESP)
2. Existing Vegetation Type (EVT)
3. Existing Vegetation Height (EVH)
4. Existing Vegetation Cover (EVC)
5. Canopy Bulk Density (CBD)
6. Canopy Base Height (CBH)

These products are amenable to quantitative assessment because field reference information exists for comparison, or in the case of CBD and CBH, because other reference data are available to assess model predictive ability.

## **Metrics**

LANDFIRE will utilize standard map accuracy assessment metrics for Environmental Site Potential, and the three Existing Vegetation layers. Contingency tables (*aka* Error Matrices) are cross-tabulations of the mapped category compared to the actual category for reference features. Reference features will vary depending on the specific “feature” of the accuracy assessment (See *Feature* below). From each error matrix, the Percent Agreement, Producer’s Agreement and User’s Agreement will be computed and reported. Users may compute additional metrics from the error matrices as they deem necessary.

Canopy characteristics (Bulk Density and Base Height) are not amenable to all the map accuracy assessment procedures described below. In lieu of these procedures, standard statistics will be computed to help users understand the characteristics of these products, namely Pearson’s Correlation Coefficient, Mean Absolute Error (MAE), and bias.

In addition, the team will compute and report other statistics that could provide insight about the map products to the LANDFIRE user. For ESP and EVT in each Map Zone and Super Zone, the percent of area by class will be reported, along with percent of the area in each class based upon the field reference data.

## **Area**

A two-tiered geographic approach will be used to report quantitative accuracy assessment results. The first tier will provide accuracy assessment results for individual mapping zones, while the second tier will provide accuracy results for aggregated mapping zones called Super Zones (Figure 1). Super Zones will be used with Holdout samples because of low holdout sample sizes for individual mapped categories in single Map Zones. Super Zones will be identical to Model Zones that were defined for the LANDFIRE Rapid Assessment phase of the project. There are 11 Super Zones for the lower 48 states.

**Figure 1.** LANDFIRE Super Zones (in red) for the conterminous United States



### **Procedures**

Two different quantitative accuracy assessment procedures will be utilized to provide more information to the user of LANDFIRE products cross-validation and holdout samples. Both procedures are common in the mapping literature, and have been widely investigated and applied by numerous agencies and organizations. Specific citations can be found in the References section. Both procedures will result in error matrices.

Cross-Validation: Cross validation is a statistical technique that indicates how well the mapping model “fits” the data used to construct it, and is computed as a part of the model development process. There is generally a relationship between model fit and model accuracy, and most users find cross-validation results to be informative. In LANDFIRE, a 10-fold cross-validation will be performed using protocols developed by the production team, and results will be reported as error matrices for vegetation products. Cross validation of canopy fuel predictions will not result in error matrices but will produce estimates of MAE, r and bias.

Holdout Sample: As an independent measure of product accuracy, a Holdout Sample approach will also be utilized. The LANDFIRE holdout sample will be a 2% probability sample of all available reference plot data. The plan calls for selecting a 2% holdout sample using a pre-defined spatial systematic framework. *Appendix B* describes considerations leading to this choice of designs for the sampling of holdout field plots. A Holdout Sample is a direct measure of the accuracy of the final map, although low sample sizes can limit our ability to make precise statistical estimates from the error matrices since some categories may have small sample sizes (or even no holdout samples). Because of low plot numbers in the Canopy Base Height and Canopy Bulk Density mapping processes, there will be no holdout sample analysis conducted for these two products.

### Pseudo Plots

LANDFIRE products are derived from field plots stored in the LANDFIRE Reference Data Base. In some cases, insufficient plots exist for some categories to support the existing vegetation type mapping process. In these cases, so-called “pseudo plots” are generated from the existing plots for use in the modeling process. However, since these are not “real” plots, pseudo-plots are not used in the Cross Validation or Holdout Sample assessment procedures.

### Lifeform/Lifezone Stratification

To improve the quality of LANDFIRE products, the production teams use a stratification process for model development. Instead of creating a single mapping model for the entire zone, the zone is stratified by lifeform (tree, shrub or herbaceous) or lifezone (topographic position) or a combination of lifezone and lifeform. For instance, two strata have been used thus far when mapping Environmental Site Potential, Upland and Riparian. Several strata are used in the mapping of Existing Vegetation, and they could vary from map zone to map zone. Cross-validation assessments will be performed separately for each stratum.

It should be recognized that the existing set of reference data plots does not represent a valid probability sample of the landscape. The LANDFIRE reference database forms the foundation for nearly all LANDFIRE deliverables. It is used for developing training sites for imagery classification, validating and testing simulation models, developing vegetation classifications, creating empirical models, determining and archiving data layer attributes, and assessing the accuracy of maps and models (Caratti, 2006). The reference database contains all relevant plot level information and provides the means to generate, test, and validate predictive models and LANDFIRE deliverables. Once each plot is converted to a common format, it is keyed to an existing vegetation type (EVT) and environmental site potential (ESP) using sequence table classifiers based solely on floristic composition. Many reference data plots were not selected by a probability sampling design, but rather were chosen purposefully to meet the objectives of a particular study, because of convenient access, or because of special interest (i.e., “exemplar” of a particular vegetation type). Consequently, while the holdout sample is a probability sample from the “population” of existing reference data plots, it is not a probability sample of the landscape, and training data used in the cross-validation assessments are also not necessarily a representative sample of the landscape. For this reason, subjective judgment is needed to determine how representative the accuracy results are to the full landscape.

### **Assessment Feature**

Much of the LANDFIRE reference data has been collected on plots that can putatively be associated with single 30 m pixels of the maps. However, because of the complexity of the mapping process, we will not limit the assessment of

LANDFIRE accuracy to an individual pixel assessment. To assist potential users as much as possible, the PQWT determined that three different accuracy assessments would be performed, each unique in some way in their interpretation and content: Pixel Assessment (PX), Fuzzy Category Assessment Units (FC), and Spatially Filtered Assessment Units (SF).

Pixel Assessment: Since the resolution of LANDFIRE spatial products is 30m, the most basic assessment unit is the pixel. However, users should regard pixel based accuracy assessments with caution. First, pixel assessments assume that field plots are contained within an individual pixel, field plots are located within the correct pixel, and that the vegetation within the pixel is homogeneous and mutually exclusive from other vegetation classes. Pixel assessments should be regarded as relatively conservative estimates of product accuracy.

Fuzzy Category Assessment: The assessment will be extended to include methods that account for overlap of map categories. LANDFIRE map units are based upon ecological systems which have strong biophysical and floristic overlapping features between classes (Comer et al 2003). Thus, it is informative to go beyond the determination of individual, crisp class accuracy and extend the analysis to allow for partial agreement between closely related classes. The PQWT will work with NatureServe to develop the criteria for defining the fuzziness of mapping units, similar to procedures being used in the Southwest ReGAP and Southeast ReGAP projects. We will utilize procedures from these projects whenever feasible. Fuzzy Category assessments will not be performed on the quantitative products (Existing Vegetation Height and Existing Vegetation Cover).

Spatially Filtered Assessment: The purpose of this assessment unit is to explore the impact of two confounding spatial factors on accuracy assessment results. First, there are potential geo-registration errors between the imagery and the reference data. Care is taken during the QC process to adjust plots for gross location errors, but it is possible if not likely that field plots do not definitively fall in a single pixel without question. In addition, using a spatial filter also more accurately represents the thematic resolution of the LANDFIRE products, which is far greater than 30 meters. Using methods described or utilized by literally dozens of authors (Hagen 2003 and Hoppus, et al 2006 for instance), accuracy results for a 5x5 spatial filter will be computed and reported. Spatial filtering will only be applied to analyses with the holdout sample.

Table 1 contains a summary of the standard quantitative accuracy assessment products that will be produced by the LANDFIRE team for each product.

**Table 1.** Standard LANDFIRE quantitative accuracy assessment products

			LANDFIRE Product					
Area	Procedure	Unit			Existing Vegetation Height	Existing Vegetation Cover	Canopy Base Height	Canopy Bulk Density
Map Zone	Cross Validation (CV)	PX	X <i>(2 models)</i>	X <i>(~6 models)</i>	X	X	X	X
		FC	X <i>(2 models)</i>	X <i>(~6 models)</i>				
		SF						
	Holdout Sample (HO)	PX	X	X				
		FC	X	X				
		SF	X	X				
Super Zone	Holdout Sample (HO)	PX	X	X				
		FC	X	X				
		SF	X	X				

**PX** – Pixel by Pixel Error Matrix; *Percent Accuracy; Producer’s Accuracy; User’s Accuracy*  
**FC** – Fuzzy Category Error Matrix; *Percent Accuracy; Producer’s Accuracy; User’s Accuracy*  
**SF** – 5x5 Spatial Window Error Matrix; *Percent Accuracy; Producer’s Accuracy; User’s Accuracy*

## **Qualitative Assessment and Peer Review of LANDFIRE Deliverables**

Qualitative evaluations will be conducted where possible on products that cannot have a formal accuracy assessment conducted on them because of the lack of true reference data. Reference data for these products do not exist for some products for several reasons, such as the impossibility of obtaining extensive spatial vegetation data that represents the reference time period (Pre-European Settlement), and the expense of performing extensive FRCC designations in the field. To compensate for this lack of reference data, qualitative methods will be used to evaluate the products. These products fall into three groups for qualitative assessment, 1) Non-spatial Vegetation Models, 2) Fuel Layers, and 3) Fire Regime Layers.

### **Non-spatial Vegetation Reference Condition Models**

Non-spatial vegetation models are created by non-LANDFIRE personnel at local expert workshops. These models undergo a rigorous internal and external model review process prior to delivery to the Missoula Fire Sciences Lab. Model contention is resolved through a majority evidence approach. MSFL personnel who utilize the models during the LANDFIRE production process also examine the models and provide feedback. See *Appendices C, D, E*, for additional details on the model review process.

### **Fuels Layers**

No extensive spatially referenced fuels information exists to use as reference data for a formal accuracy assessment. However, the Fuels team developed a set of protocols or processes for evaluating LANDFIRE fuels products, depending on product. Anderson Fuel Models and Scott/Burgan Fuel Models products will be evaluated and calibrated by local experts at Fuels Calibration Workshops in each LANDFIRE Mapzone. Fuel model as a % of the landscape will also be computed. The Fuels Team will conduct FLAMMAP (Finney, 2006) runs to identify aberrant fire behavior predictions induced by LANDFIRE data. In addition, LANDFIRE personnel will be meeting with FBAN and LTAN personnel after the fire season to evaluate the efficacy of the fuels layers during the 2006 fire season.

### **Fire Regime**

LANDFIRE Fire Regime layers (FRCC, FRCC Departure, Fire Regime Groups, Historical Simulated Fire Return Interval, Percent Non-lethal fire, Percent Mixed Severity Fire, Percent Stand Replacement Fire, Succession Classes) are developed using complex simulation models that incorporate multiple diverse input products. In particular, the Reference Condition Vegetation Models (see above), WX-BGC (estimate abiotic site factors such as weather, solar radiation, etc.) and LANDSUM (simulated fire history) drive the fire regime mapping process. There are no generally accepted analytical methods for estimating

(quantitatively) the accuracy from a series of model inputs. Each of the models, however, has supporting information in this plan or in separate publications. Reference Condition Model QC and review processes are outlined in various plan appendices. Keane, Holsinger and Pratt (2006) discuss the characteristics of the LANDSUM model, and Keane and Holsinger describe the WX-BGC model.

As a final evaluation of the products, fire regime data layers will be directly compared to local data in LANDFIRE Application Projects (see below).

## **LANDFIRE Application Projects**

The Nature Conservancy will establish a set of LANDFIRE Application Projects (LAP) around the United States. To date, LAPs exist in Washington, Oregon, Nevada, Michigan, New Mexico, Florida/Georgia, North Carolina, Iowa and Arkansas, ranging in size from 115,000 acres to approximately 2,000,000 acres. Additional projects are currently being pursued in Alaska and other parts of the country. The goal of the LAPs is to compare LANDFIRE Rapid Assessment and National Implementation data to local data. Although the primary focus will be on evaluating FRCC and related data, other LANDFIRE products and derivatives will be utilized and compared in most projects.

The LAPs are a key component of the data assessment strategy. Each LAP will be conducted on a multi-partner, working landscape, thus the data evaluations will be in a true application setting. Where possible, we will encourage local participants to utilize LANDFIRE data in simulated decision-making situations, and to compare the outcomes to decisions made using local data. While we may not be able to extrapolate the results of LAPs widely in a geographic sense, they will provide a unique picture of how well LANDFIRE data is working in the “real world”, and a basic understanding of the quality and characteristics of LANDFIRE data from the perspective of a data user.

## **Guidelines for Application Scale**

### **Scale of Applications**

The LANDFIRE Charter stated that the project deliverables are intended for regional and national applications, such as regional risk evaluations and as input into Fire Planning and Analysis (FPA) resource allocations. A region may be envisioned as a large watershed of tens or hundreds of thousands square kilometers, or expressed as scale of 1:250,000. Even though LANDFIRE has a nominal raster data resolution of 30 meters, the true information content as a function of all input data with a variety of original resolutions, is far coarser than 30 meters. Together with inherent errors within input and output data layers, an appropriate scale of applications should be multiple aggregations of LANDFIRE map pixels.

As an example, many appropriate application scales may include risk evaluations for communities (approximately hundreds of square kilometers), watershed-wide fuel planning, FARSITE simulation of wildland fires sized thousands of hectares, insect & disease modeling of national forests, and carbon and biomass studies of square kilometer blocks.

While LANDFIRE map products will be used for a broad range of applications, it is conceivable that some of the applications may even use LANDFIRE data in a manner that far exceeds an appropriate scale of use in terms of spatial resolution, thematic information content, and inherent uncertainties about Product Quality and reliability. The PQWT will work with LANDFIRE technical task teams to ensure that facts about LANDFIRE data resolution and appropriate scale of applications be publicized and published in various venues to ensure that the user community understands how to appropriately use LANDFIRE map products.

### **LANDFIRE National Application Scale Definition**

Application scale is defined as the geographic extent at which LANDFIRE National products are of sufficient quality [structure, completeness, accuracy] to be applicable and useful.

### **LANDFIRE National Product Application Scale**

Although LANDFIRE National data are delivered as 30 meter pixels, the data should NOT be utilized at the individual pixel level, or on small groups of pixels. LANDFIRE National products were designed to support 1) national (*all states*) strategic planning, 2) regional (*single large states or groups of smaller states*), and 3) strategic/tactical planning for large sub-regional landscapes and Fire Management Units (*e.g., significant portions of states, multiple federal administrative entities*).

The applicability of LANDFIRE National data to support fire planning on less extensive geographies will vary by product, location and specific use. Further investigation by local experts and by LANDFIRE team partners during the project and over time will help inform this decision about local applicability. However, it is the responsibility of the local user, using appropriate LANDFIRE National metadata and local expertise and knowledge, to determine if/how LANDFIRE National data can be used for their particular area of interest.

## **Plan References**

Anderson, J.A., J.D. Hanson, and R.H. Haas. 1993. Evaluating Landsat Thematic Mapper derived vegetation indices for estimating above - ground biomass on semiarid rangelands. *Remote Sensing of Environment* 45:165-175.

Bechtold, W. A. 2004. Largest-crown-width prediction models for 53 species in the western United States. *Western Journal of Applied Forestry* 19:245-251.

Brohman, R., and L. Bryant. 2005. Existing Vegetation Classification and Mapping Technical Guide. General Technical Report WO-67. U.S. Department of Agriculture Forest Service, Ecosystem Management Coordination Staff., Washington, DC

Comer, P.; Faber-Langendoen, D.; Evans, R.; Gawler, S.; Josse, C.; Kittel, G.; Menard, S.; Pyne, M.; Reid, M.; Schulz, K.; Snow, K.; Teague, J. 2003. Ecological systems of the United States: A working classification of U.S. terrestrial systems. Arlington, VA: NatureServe.

Caratti, J.F. 2006. The LANDFIRE Reference Database. In: The LANDFIRE Prototype Project: nationally consistent and locally relevant geospatial data for wildland fire management. General Technical Report RMRS-GTR-175. Fort Collins, CO: Missoula Fire Sciences Laboratory, Missoula, MT.

Finney, M.A. 2006. FlamMap. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory, Missoula, MT.

Friedl, M. A., Woodcock, C., Gopal, S., Muchoney, D., Strahler, A. H., and Barker-Schaaf, C. (2000), A note on procedures used for accuracy assessment in land cover maps derived from AVHRR data, International Journal of Remote Sensing 21: 1073-1077.

Gopal, S., and C.E. Woodcock, 1994. Theory and Methods for Accuracy Assessment of Thematic Maps Using Fuzzy Sets, Photogrammetric Engineering and Remote Sensing, 60(2):181-188.

Hagen, A. (2003), Fuzzy set approach to assessing similarity of categorical maps, International Journal of Geographical Information Science 17: 235-249.

Hoppus, M. S. Arner, and A. Lister. 2006. Stratifying FIA Ground Plots Using a 3-year Old MRLC Forest Cover Map and Current TM Derived Variables Selected by "Decision Tree" Classification.

Lennartz, S. 2005. Presentation on mapping Ecological Systems in western Washington for GAP at Pecora 16.

Keane, R.E., L.M. Holsinger, and S.D. Pratt. 2006. Simulating Historical Landscape Dynamics Using the Landscape Fire Succession Model LANDSUM version 4.0. General Technical Report RMRS-GTR-171CD. United States Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO

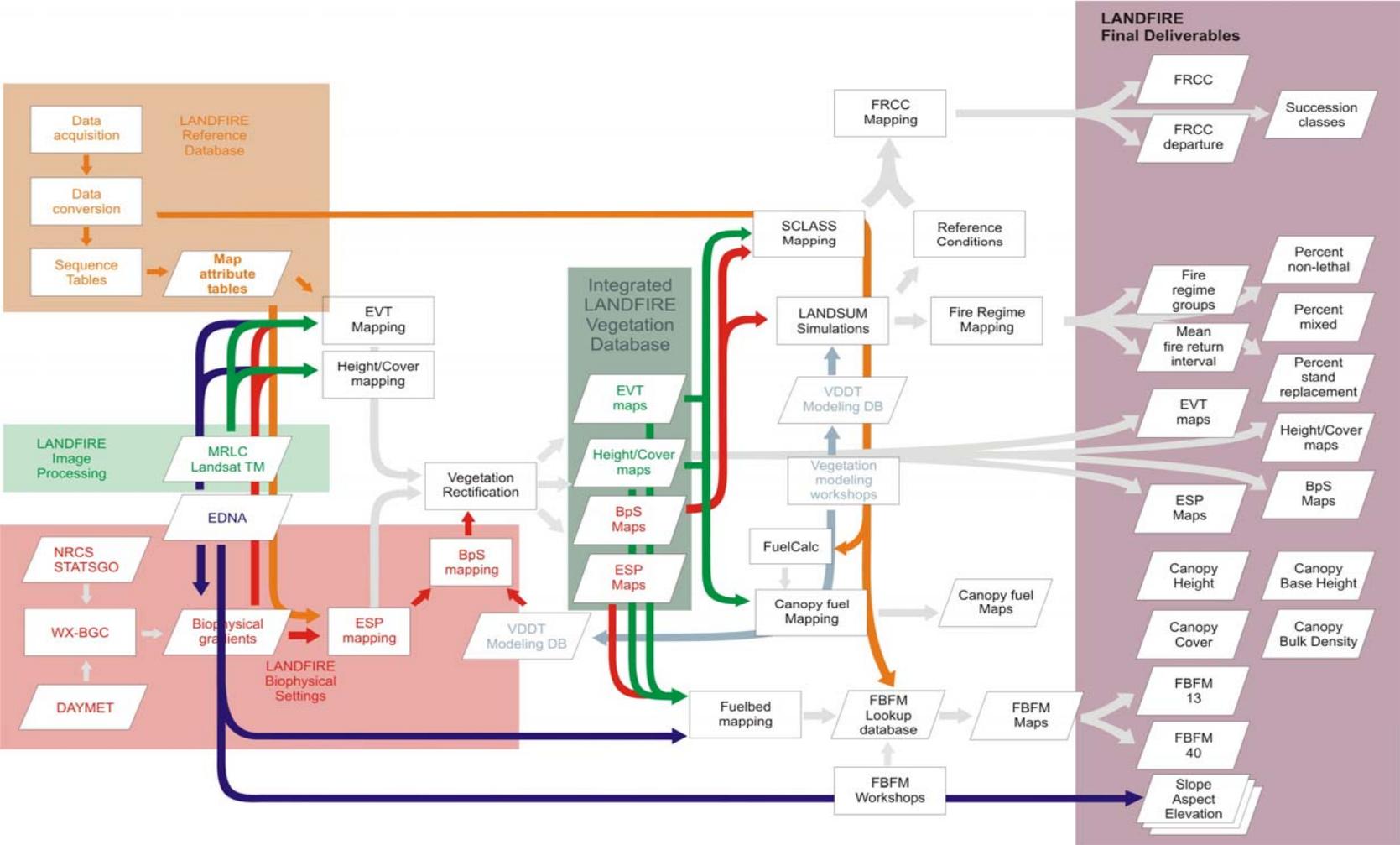
Keane, R.E., and L.M. Holsinger. 2006. Simulating Biophysical Environment for Gradient Modeling and Ecosystem Mapping Using the WXFIRE Program: Model Documentation and Application. General Technical Report RMRS-GTR-168DVD.

United States Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Lowry, J. 2005. Presentation on SW ReGap at Pecora 16.

Steele, B. M. (2005), Maximum posterior probability estimators of map accuracy, Remote Sensing of Environment 99: 254-270.

# LANDFIRE Production Process Flow Chart



## **Appendix A**

### **Existing LANDFIRE Quality Control Procedures**

LANDFIRE has implemented a set of systematic QC measures that assure a high level of confidence in maintaining high Product Quality standards. Specific QC procedures are integrated into this project to test completeness of spatial and attribute information.

#### **LANDFIRE Reference Database**

QC procedures are integrated into this project to help identify potential errors that may exist with the field reference plot data. The collection of reference data can include plot data from various projects collected at various scales ranging from individual ranger districts to national programs, collected for various objectives, and/or have been collected over a variety of dates ranging back to the late 1980's. The quality of these data sets for the most part has been good, but errors in the data do occur and therefore steps are used to minimize the occurrence of errors. The LFRDB Team spend the bulk of their collective time ensuring that data are accurately represented as valid and useful components of viable source datasets in the LFRDB. To this end, each dataset and each data element therein must be carefully evaluated throughout the conversion and compilation process. The level of the evaluation/conversion processes can differ widely from dataset to dataset. Each set invariably has its own "issues."

The QC process includes these steps:

1. Conduct cursory QC on data to eliminate any data with irreconcilable geospatial or information content errors.
  - Map plots from the coordinate information in the source database and look for plots outside the stated domain of the sampling effort. Plots clearly outside the expected sampling area are removed from the pool. This step does not apply to FIA data, as no FIA coordinate information is available to us.
  - Remove plot records with key data missing (e.g., FIA grid plots with no tree data).
2. Check for and correct or remove records with information errors or otherwise unsuitable data.
  - Select the most recent record(s) for a plot (e.g., max of cycle for FIA plots) or, in the cases of element occurrence data (e.g., exotic plant records), select the most recent record with the most species reported for a given site and remove all others.
  - If cluster plot design (e.g., FIA, CVS), remove records containing data from subplots installed across multiple "conditions" (e.g., cover types), as defined and indicated by the data source.
  - Map and update all species codes for plants within study domain to symbol keys from the NRCS PLANTS database snapshot taken January 2005. If unable to match, seek additional information (i.e., a

valid/comprehensive decoder) from the data source. If still unable to reconcile, update the item label to “unknown.”

- Derive percentage canopy cover for species from tree data (i.e., stem diameter and TPA), as necessary (e.g., for PNW FIA). Derivation for western species is based on relationships reported in Bechtold (2004).
- Check for information errors in cover fields, such as missing data or cover values out of range. If all species on plot are missing cover values, discard the plot. If one or a few species on a plot are missing cover values, assign a default cover of 0.5 (trace) to those items.
- Check for other information errors such codes that don't make sense or are not in a look-up table (e.g., crown class codes or decay class codes). Make correction or if unable to reconcile, update value to null.
- Check for other information errors/inconsistencies (e.g., in the event that we have live trees with percent live crown indicated as zero in the source dataset, we will update zero value to null).

3. Conduct final pre-distribution QC for erroneous or suspicious data in the

#### Map Attribute Table

- Must have same number of records as the Plots table in the Key DB.
- Check Date column for faulty dates or dates out of range
- MaterID: All records should have Master ID.
- Albers\_x and Albers\_y: All non-FIA data should have Albers coordinates. Check for same coordinates with different plot ID's
- LF\_Zone: All records should be labeled with the appropriate zone.
- Orig\_Eco\_System, Orig\_Alliance, and Orig\_CoverType: Make sure all available cover type information has been placed in the appropriate field. Double check all datasets.
- LF\_EVT and LF\_EVTCode: Update fields with EVT Key data outputs. Make sure all records have an EVT assignment. If EVT is blank update LF\_EVT to none and LF\_EVTCode to 0.
- EVT\_Lifeform: Use established EVT/Lifeform look up table to update. Make sure all records have an EVT\_Lifeform the exception being records with LF\_EVT or none.
- DomSp: Find the species on a plot within the EVT lifeform that has the greatest cover. Use the Species table from the KeyDB for calculations. If two species have the equal cover sort species ascending and use 1<sup>st</sup> species as DomSp. Check to see how many records have a dominant species. For records that have an EVT assignment but no dominant species take note of their EVT.
- DomSpC or DomSpBA: Place the dominant species absolute cover or absolute basal area in the appropriate field. Make sure all records that have a DomSp also have a DomSpC or DomSpBA.
- CoDomSp: Find the species on a plot within the EVT lifeform that has the second greatest cover. Use the Species table from the KeyDB for calculations. If two species have the equal cover sort species ascending

and use 2<sup>nd</sup> species as CoDomSp. Some records may not have a CoDomSp.

- CoDomSpC or CoDomSpBA: Place the co-dominant species absolute cover or absolute basal area in the appropriate field. Make sure all records that have a CoDomSp also have a CoDomSpC or CoDomSpBA.
- Original Lifeform Cover: Update with QAQC'd Life-form cover calls from the PD. Herb cover- add the PD forb and gram cover.
- Original Lifeform Heights: Ensure life-form height calls have been brought through from the original data. Double check datasets.
- LF\_TreeC and LF\_ShruhC: update columns with treecov and shbcov fields in the Plots table from the Key DB.
- LF\_HerbC: calculate herb cover from the Species table in the Key DB by updating species lifeform of F and G to H.
- LF\_TreeHt, LF\_ShruhHt, and LFHerbHt: calculate life-form height from the Species table in the Key DB. Make sure to calculate weighted heights using the life-form RC or RBA. Then add weighted heights for Trees, Shrubs and Herbs (forb and gram). There should be no 0's, use null to represent no data.
- Orig\_BpS: Make sure all available BpS/PVT information has been placed in the appropriate field. Double check all datasets.
- LF\_BpS and LF\_BpS Code: Update fields with BpS Key data outputs. Make sure all records have a BpS assignment. If BpS is blank update LF\_BpS to none and LF\_BpSCode to 0.
- Photo\_1 and Photo\_2: Make sure all available photos are listed in these two columns. Photos will need to be updated in the MAT to run the photo extraction program.
- Reference: Ensure any applicable reference materials such as PDF's or scientific papers are listed in this column.
- Loc\_Meth: All records should have either a G for GPS or an X for unknown. If blank or X double check data sets and supporting information.
- Loc\_Acc: If location accuracy information is available make sure it populates this field
- Ancillary Data: DistToRoad, NLCD\_Code, NDVI\_Diff, and Withhold\_for\_Validation: All records should have values in these fields. Use the intial zone ancillary draw to update fields.
- NLCD\_Des: All records should have a description. Make sure codes and descriptions match. Use NLCD code/description look up table to populate this field.
- Orig\_LifeformCover\_QAQC: All records that have an original life-form cover call should also have a QAQC code. Ensure all codes make sense and each dataset has only one code.
- LF\_LifeformCover\_QAQC: All records should have a QAQC code of 20.
- SppCover\_QAQC: All records should have the appropriate QAQC code. Make sure there is only one code per a dataset.

- SppHeight\_QAQC: Bad tree, shrub, and herb heights need to be calculated from the following FIEMON tables: SCCover, TDTree, and TDSap. Bad tree height is >150 ft. Bad shrub height is >40 ft. Bad Herb height is >6 ft. Place appropriate code in this field if the plot contains any one or combination of bad life-form heights.
- The MAT table may need the following metric conversions, check to see if source data had already been converted to metric:
  - DomSpBA: square feet per acre to square meters per hectare (multiply by 0.229568411).
  - CoDomSpBA: square feet per acre to square meters per hectare (multiply by 0.229568411).
  - Orig\_TreeHt: feet to meters (multiply by 0.3048).
  - Orig\_ShrubHt: feet to meters (multiply by 0.3048).
  - Orig\_HerbHt: feet to meters (multiply by 0.3048).
  - LF\_TreeHt: feet to meters (multiply by 0.3048).
  - LF\_ShrubHt: feet to meters (multiply by 0.3048).
  - LF\_HerbHt: feet to meters (multiply by 0.3048).
  - Loc\_Acc: must be in meters.

#### **Species Comp Table:**

- MasterID: Make sure all records have master ID.
- Symbol Key and Scientific Name: All records should have a Symbol Key and Scientific Name.
- AC or ABA: make sure all records have a cover value and there are no 0 or -1. Records with a symbol key like NOTREES can have null cover values.
- RC or RBA: all records that have a lifeform of T, S, F, or G should have a value in this column.
- Cov/BA Code: Make sure all records are labeled with the appropriate code.
- Height: make sure there are no 0 or -1.
- Life\_Form: Update this column with the appropriate life form found in the zones species list.
- Duration: Update this column with appropriate duration found in zones species list. Only grass and forb species should have a duration assignment.
- The Species Comp table may need the following metric conversions, data should already be in metric but double check:
- Height: feet to meters (multiply by 0.3048).

#### **Constancy Queries and Tables:**

- Constancy Queries: The MAT and Species Comp table must be filled and QAQC'd to run these queries. Once that is finished begin by checking the design view then running the Plot CT and Species CT queries. If everything looks ok with those queries then check the design view of the constancy queries to make sure you only include species with a constancy of greater than 5. If everything is hooked up ok and run the query. Look over query results to make sure it looks accurate

- Constancy tables: Once all looks good with the constancy queries run the append queries to populate the constancy tables. Double check tables.

### **Ancillary Tables:**

- Z#\_Geo: Should have same records as the Map Attribute table. Make sure all columns for all records are filled with the appropriate ancillary data.
- Z#\_Gradients: Should have same records as the Map Attribute table. Make sure all columns for all records are filled with the appropriate ancillary data.
- Z#\_Imagery: Should have same records as the Map Attribute table. Make sure all columns for all records are filled with the appropriate ancillary data.

### **Disturbance History**

- MasterID: Make sure all records have a Master ID and that all plots in the MAT are included.
- Sample\_Year: Make sure all records have a sample year and check the range. Update Sample year using date column and following formula: Right ([Date], 4)
- Dirsturbance1, Dirsturbance2, Dirsturbance3 : Make sure all codes have been deciphered into descriptions and descriptions are short enough to fit in the text field. It is ok to mix treatment info with disturbance info if original data mixed the 2 in one column. Make sure there are no 0, -1, 999, unavailable, etc. If there is no disturbance information for a plot leave field blank.
- Disturb1\_Sev, Disturb2\_Sev, Disturb3\_Sev: Make sure field includes severity description (not code) and all descriptions are sensible and fit in text field. Make sure there are no 0, -1, 999, unavailable, etc. If there is no disturbance severity information for a plot leave field blank.
- Disturb1\_Year, Disturb2\_Year, Disturb2\_Year: Make sure field includes only year, check year against sample year, and check year range. Make sure there are no 0, -1, 999, unavailable, etc. If there is no disturbance year information for a plot leave field blank.
- Treatment1, Treatment2, Treatment3: Make sure all codes have been deciphered into descriptions and descriptions are short enough to fit in the text field. This column should contain only treatment information. Make sure there are no 0, -1, 999, unavailable, etc. If there is no treatment information for a plot leave field blank.
- Trtmnt1\_Year, Trtmnt2\_Year, Trtmnt3\_Year: Make sure field includes only year, check year against sample year, and check year range. Make sure there are no 0, -1, 999, unavailable, etc. If there is no treatment year information for a plot leave field blank.
- Source\_Comments: If disturbance information is in a general comments field in the raw data bring information into this field. Make sure the full length of the comment placed in this field. QAQC field and weed out non-disturbance information.

### **MAP Attribute QAQC**

- MasterID: Make sure all records have a Master ID and that all plots in the MAT are included.
- Field\_Comments: Place unedited source data comments from disturbance history table in this field. Make sure full length of comment is placed in field.
- LFRDB\_QAQC\_Flag: Flag plots with erroneous and suspect information.
- Flag plots with bad dates.
- \_\_\_\_\_
- \_\_\_\_\_
- \_\_\_\_\_
- LFRDB\_QAQC\_Comment: List reason for flagging plots. If bad date give date range of source data.
- Update the following fields with information supplied by the appropriate Mapping team:
  - BpS\_Discard
  - BpS\_Discard\_Comment
  - EVT\_Discard
  - EVT\_Discard
  - CanFuel\_Discard
  - CanFuel\_Discard\_Comment
  - SurFuel\_Discard
  - SurFuel\_Discard\_Comment

### **LANDFIRE Biophysical Gradients**

The following QC procedures are followed during the development of biophysical gradient layers:

1. Create Topographic and Soil Inputs to Simulation Unit development
  - Check output grids from soils\_slope.aml – extents should match z#\_dem, data are checked for reasonable values and there should be no internal NO DATA values
  - Check output grids from simunit\_reclass.aml for extents, be sure all pixels fall into a class.
2. Create Simulation Units for WX-BGC Runs.
  - Check the extent of z#\_mmap in ArcMap, make sure it matches input grids/bils for the zone.
  - Run checkloc.exe – checks that every pixel has a simulation unit.
3. Process and Distribute Biophysical Gradient Layers
  - Run mapimg.bat – scales the raw BILs (averages) from 0 to 65535 and flags pixels with data outside the expected range (65535 values)
  - Correct “problem” pixels in BILs with data outside the expected range.
  - Document the calculation of integer BILs - Use the .stats files of the scaled BILs from mapimg.bat to create a spreadsheet of each output variable’s absolute min/max, scaled min/max, gain, offset. Note variables which had to be fixed (had simunits with data outside the expected range). Post

this spreadsheet on the MFSL network drive with BILs and images, so user's (e.g. BpS modeler) know which variables had to be fixed.

- Check extent/values for all the images.
- Create metadata for each image

### **Potential Vegetation (ESP and BpS) Mapping**

The potential vegetation mapping process uses plot information from the LFRDB and geospatial information from the LANDFIRE biophysical gradient layers as inputs. The main phases of potential vegetation mapping are: 1) assignment of plots to ESP map units using sequence tables, 2) QC of plot data prior to ESP mapping, 3) ESP mapping, and 4) BpS mapping. The following QC steps are taken to ensure the highest quality products, within the project's methodological and time constraints.

#### 1. Assigning plots in the LFRDB to ESP map units

- Develop draft ESP sequence table, based on EVT sequence table. Use reference information in draft Model Tracker Database, NatureServe system descriptions, and any other publications or on-line information that can be found about the ecology and floristics of the given map zone.
- Run bootstrap analysis to identify which systems are closely related or easily confused with each other. Use bootstrap results to guide adjustments to the ESP sequence table.
- Create draft maps based on ESP assignments in draft sequence table. Use spatial output to guide adjustments to the ESP sequence table.
- Check that assignments of ESP are logically consistent, in most cases, with EVT plot assignments.

#### 2. QC of plot data prior to ESP mapping

- Assemble plot data from the LFRDB (including ESP assignments) and the Key Database into the BpS training sites database.
- Evaluate histograms and descriptive stats for predictor variables; note any variables that should be dropped due to obvious problems.
- Screen out plots with bad gradient values (flagged with 65535) and plots not classified by sequence table (ESP code < 1000). If an individual gradient has a large number of plots flagged as bad (65535), consider dropping the gradient from further modeling instead of dropping the plots.
- Run bootstrap analysis. Add data from bootstrap analysis to the BpS training sites database.
- Run a draft classification with See5 and an associated map and cross-validation error matrix. This serves as the baseline for subsequent comparisons.
- Within the training sites database, run standard QC queries to drop plots with common problems. These include: 1) plots with 0% correct classification in the bootstrap analysis, 2) plots predicted to another class >60% of the time in the bootstrap analysis, and 3) plots where indicator species have very low (e.g., < 5%) absolute cover.

- Within the training sites database, construct “targeted queries” to drop additional plots whose ESP assignments seem incorrect. Use reference information in draft Model Tracker Database, NatureServe system descriptions, and any other publications or on-line information that can be found about the ecology and floristics of the given map zone.
- Evaluate the results of dropping plots by running a draft classification with See5 and an associated map and cross-validation error matrix. Evaluate the producer’s accuracy by class and the overall accuracy. Evaluate spatial distributions of each system in the draft map. If classes have been lost due to dropping plots, if accuracies have decreased, or if spatial patterns are problematic, consider evaluating individual plots that were dropped and possibly adding them back in.
- Iterate the above steps until accuracies and spatial patterns are acceptable, within time limits.

### 3. ESP Mapping

- Check plots from adjacent zones that will be used as training sites in mapping. If a system is mapping only due to plots in adjacent zones, evaluate whether it makes sense to map in current zone. If not, drop the plots for that system.
- Produce predicted lifezone classifications and associated maps and cross-validation matrices. Evaluate each and decide if one will be used in upland ESP classification.
- Produce upland ESP classification with See5 and an associated map and 10-fold cross-validation error matrix. Evaluate error matrix for potential problems. Evaluate spatial patterns in the map. If problems are easily correctable, then correct them and rerun classification. If problems cannot be corrected easily and quickly, document them and possibly make recommendations for how they can be dealt with in rectification.
- Produce riparian ESP classification with See5 and an associated map and 10-fold cross-validation error matrix. Evaluate error matrix for potential problems. Evaluate spatial patterns in the map. If problems are easily correctable, then correct them and rerun classification. If problems cannot be corrected easily and quickly, document them and possibly make recommendations for how they can be dealt with in rectification.

### 4. BpS Mapping

- Review final Model Tracker Database to determine which systems from the ESP layer need to be split or refined for BpS. Develop a list of potential splits/refinements, and a strategy for implementing each.
- Send questions and proposed strategies to TNC contact (regional modeling lead and/or national staff) to solicit feedback from appropriate individuals (modelers, local ecologists, etc.). Alternately, send questions and proposed strategy directly to appropriate individuals for advice or confirmation on specific splits/refinements.
- Incorporate any feedback into splitting/refining strategies.
- For splits requiring See5 modeling, produce a classification and associated map. Also run 5 iterations of 10-fold cross-validation (to

account for high variability created by the often skewed plot distributions in BpS splits). Review individual cross-validation error matrices and average overall errors. Review spatial patterns in the map. If errors are high and/or spatial patterns are not achieving the intention of the split/refinement consider modifying the strategy and running again or possibly eliminating the given split or refinement.

- For splits not requiring See5 modeling, evaluate spatial patterns and revise strategy if needed until the intention of the given split/refinement is achieved.
- In final steps, multiple pieces (various “split” grids and original ESP grid) are brought together to make the BpS layer. Scripts are used to automate this process. Numerous checks are incorporated into these scripts to ensure: a) consistency from map zone to map zone, and b) that spatial and logical errors are not introduced.
- Visually check final output grid to ensure that all splits/refinements were incorporated as intended.

### **Existing Vegetation and Structure Mapping Procedure**

Existing vegetation types and structure are a required input, along with other data products, for modeling and mapping of fire fuels and fire regime condition class under the LANDFIRE project. The vegetation maps also represent an important geospatial deliverable to user communities for use in various environmental applications, which implies a certain degree of user expectations for quality and consistency of the vegetation data layers to be mapped.

The following steps are followed to insure QC with the mapping of existing vegetation type and structure:

#### 1.0 Spatial QC of field plot data.

- QC non-FIA data using NDVI difference data.
- Flag plots determined to have changed in Map Attribute Table and do not use in future analyses.
- Overlay points onto imagery stratified by Ecological System class groupings and NLCD data.
- Identify and flag points on roads or other similar types of locations (e.g., urban, agriculture) that should not be used for training.
- Identify plots typed as forest located in relatively intact non-forest locations (and vice versa).
- Identify plots typed as conifer located in relatively intact deciduous forest (and vice versa)
- Flag questionable plots in Map Attribute Table and do not use in future analyses.
- Develop a modified Map Attribute Table storing only field plots that pass the above QAQC process. Send the new MAT to Missoula.
- QC FIA data (same general process as in 1.1 and 1.2, only needs to be done by FIA analyst)

2. Conduct preprocessing steps on the data used in this mapping - Conduct visual quality check on the data layers, make sure no obvious seam lines, dropped pixels, or other quality problems.
3. Life form-specific cover type mapping
  - Generate life-form specific cross-validation error matrices.
  - Check for any visual and information content problems by examining cover type maps and interpreting error matrices.
  - Determine if there are any rare classes (field plots less than 30 and scattered), and decide how to treat such rare classes.
  - Determine if other major mapping errors exist, and correct by altering input parameters (if possible) as well as field reference data.
4. Life form-specific canopy height mapping.
  - Generate life-form specific cross-validation error matrices for height classes.
  - Check for errors in the three life form-specific height maps make sure ranges of height values logically make sense for their corresponding cover types
5. Life form-specific canopy cover mapping.
  - 5.1 Tree canopy cover.
    - Generate cross-validation error matrices, evaluate error and R2 values, determine effectiveness of the regression tree models.
  - 5.2 Shrub and herbaceous canopy cover – option 1.
    - Generate life-form specific error assessments based on cross-validation analysis.
    - Determine effectiveness of the regression tree models based on error analysis and whether changes need to be made to both field data and independent spatial data layers.
  - 5.3 Shrub and herbaceous canopy cover – option 2.
    - Generate life-form specific error assessments based on cross-validation analysis.
    - Determine effectiveness of the regression tree models based on error analysis and whether changes need to be made to both field data and independent spatial data layers.
6. Generate merged vegetation cover type and structure maps- Revisit, and revise if necessary, the merged vegetation cover type map (3.8) by using forest, shrub, and herbaceous canopy percent cover as references. Make sure cover types match with life-form canopy cover maps. For example, shrub cover types should not exist where forest canopy cover is relatively high. Also, transitions between forest, shrub, and herbaceous cover types generally match with the “10 percent canopy cover” rule, i.e. forest cover types have corresponding canopy cover at greater than 10 percent, else if shrub canopy cover greater than 10 percent, it would be shrub cover types, else herbaceous cover types.
  1. Conduct cross-validation and hold out accuracy assessments.
    - For individual mapping zones summarize all cross validation results: error matrix, overall accuracy, class-specific accuracy (user’s, producer’s accuracies). Holdout sample overall accuracy (sample size

too small to support class-specific accuracy) .Distribution of map land cover (i.e., percent of area by class for the entire map). Distribution of map land cover in the reference (ground) sample plots (percent of area by class for the entire “sample”)

- Extract the final existing vegetation cover type, canopy cover, and height class values and labels using withheld plot locations. Provide information to the Accuracy Working Team.
- For multiple mapping zones summarize all holdout sample results: error matrix, overall accuracy, class-specific accuracy (user’s, producer’s accuracies). Cross validation for map zones aggregated to superzone: overall, user’s, producer’s accuracies. “Fuzzy” classification scheme and “window” (e.g., 3x3 pixel, 5x5 pixel) based analysis.

### **Vegetation Models of Reference Conditions**

The LANDFIRE Vegetation Model QC is a living process, and will be continually revisited as model zones are completed, and as the rectification, BpS mapping, and fire regime simulation processes are refined to ensure that the process is complete and relevant. This QC process builds on what was utilized in the LANDFIRE Rapid Assessment. Some details varied due to the differences in model resolution and mapping methodology.

#### Initial Model QC

1. To ensure that the modeling workshops and resulting models reflect the widest possible range of local expert opinion, TNC developed a database of experts for each region. Each of these experts is contacted directly and invited to the workshop, and encouraged to invite any other experts not included in the original list. In addition, the workshop is advertised via the LANDFIRE website, and TNC also offers limited travel scholarships to noted, non-federal experts to encourage their attendance.
2. Following the modeling workshop, the Regional Lead accumulates all the models, and reviews them to ensure that all modelers followed the established rules, and that the model information is complete. To standardize this process, Regional Leads are required to utilize the LANDFIRE Model Q-C Checklist developed by the TNC LANDFIRE National Team (See Appendix A).

#### Peer Review

3. The Regional Lead compiles the models (VDDT models and model description PDF documents) and uploads them to a website where they can be downloaded by potential reviewers. All individuals who attended the workshop, all experts who were invited, and additional experts identified by workshop participants as potential peer reviewers are contacted and asked to conduct a peer review of any or all models. The

opportunity to review the models is publicized on the LANDFIRE and TNC websites. The goal is to engage as many peer reviewers as possible in the process. Peer reviewers can download and evaluate the actual VDDT model and/or the PDF model description. Peer reviewers complete a form (Appendix B) to standardize the review process.

4. The Regional Modeling Lead compiles and reads all peer reviews. Simple changes (spelling, missing fields, etc.) are immediately made to the Model Tracker Data Base description and/or the VDDT model by the Regional Modeling Lead. If the peer reviewer suggests more major revisions, the original modeler is contacted to discuss the recommendations. If the peer reviewer suggests that a model be added, the Regional Lead must identify and engage an individual to build the model. It is the responsibility of the Regional Lead to determine what changes are incorporated into the models, which models are added, or deleted. A protocol for handling the peer review process is adhered to by Regional Leads (Appendix C).
5. The final peer reviewed models are delivered to the TNC LANDFIRE National Team.

#### Final Model Review and Delivery

6. A member of the LANDFIRE National Team (not the Regional Lead) reviews the set of models for each mapping zone, using the model QC checklist (Appendix A).
7. Just prior to final delivery, the model set for each mapping zone undergoes a final check using the Veg Model Delivery Checklist (Appendix D). Once all items are verified, the model set is delivered to the Fire Science Lab modeling liaison with a Read Me file (e.g., Appendix E: Read Me 2005\_10\_15) and metadata.
8. The TNC National Team in conjunction with the MFSL project leadership developed a process for extremely contentious models, whereby a decision can be made about models when modelers or reviewers cannot agree nor compromise. The decision will be based upon “majority evidence”, as determined by an appropriate, knowledgeable individual at MFSL.

#### Final Model Release

9. A member of the LANDFIRE National Team reviews the entire set of models for each mapping zone, and compares them to models previously delivered for adjacent mapping zones. That individual determines if any

models are identical to previous models from adjacent mapping zones and modifies descriptions accordingly.

10. Model descriptions are reviewed and edited one final time by a member of the LANDFIRE National Team, or in some cases by the national LANDFIRE editor. Files (model descriptions as PDF files and VDDT databases) are prepared for web release along with accompanying metadata.

## Fuels

### Layers involved:

Fire Behavior Fuel Model 13 (FBFM 13)  
Fire Behavior Fuel Model 40 (FBFM 40)  
Canopy Height (CH)  
Canopy Cover (CC)  
Canopy Bulk Density (CBD)  
Canopy Base Height (CBH)

- 1) **INPUT - Plot Level** Quality Control (QC)
  - a. LANDFIRE Reference Database (LFRDB) screening
    - i. LFRDB QC flag = = 0
    - ii. Existing Vegetation Type (EVT) <> 201 (Unclassified Forest)
    - iii. Distance to road > 45 m
    - iv. Environmental Site Potential (ESP) discard = = 0
    - v. EVT discard = = 0
    - vi. Normalized Difference Vegetation Index (NDVI) difference < 20% change
    - vii. Discard values of Diameter at Breast Height (DBH) < 1 or DBH >= 120 inches
    - viii. Tree Height (HGT) < 1 or HGT >= 300
    - ix. Height to live tree crown (C-HBC) < 0 or C-HBC >= 300 or C-HBC >= HGT
    - x. Trees Per Acre (TPA) <= 0 or TPA >= 3000
  - b. Imagery Screening
    - i. Cloud and snow masked
    - ii. Visual inspection for gross radiometric differences
- 2) **OUTPUT – Plot level** (CBD and CBH estimates) QC
  - a. Discard CBD values > 0.75 or CBD < 0.012
  - b. Available Canopy Fuel (ACF) >= Total Canopy Weight (TCW) or < 0
  - c. TCW <= ACF
  - d. CBH < 0 or CBH > SH
  - e. Stand Height (SH) < 1 or SH > 300
  - f. CC < 1 or CC > 100
  - g. Basil Area (BA) < 1

- 3) **OUTPUT – Image level** (CBD and CBH estimates) QC
  - a. Visual inspection (Find strange patterns and fix them)
  - b. Fill cloud and snow areas with regional averages.
- 4) **OUTPUT – Logic Checks**
  - a. Ensure FBFM13 and 40 are calibrated in workshops.
  - b. Ensure Canopy Fuels are assigned in calibration workshops to work in conjunction with FBFM13 and 40.
  - c. Ensure that entire raster matrix defined by the zone mask is populated with a valid value for every fuel product.
  - d. CBD values are continuous from 0 to 0.40 kg m<sup>-3</sup> (0.025 lb ft<sup>-3</sup>) (to the nearest 0.01 kg m<sup>-3</sup>), whereas all values > 0.40 kg m<sup>-3</sup> are binned into a single thematic class of 45 (0.45 kg m<sup>-3</sup> \* 100).
  - e. CBH data are continuous from 0 to 99 (9.9 meters \* 10) (to the nearest 0.1 meter), whereas all values > 99 (9.9 meters \* 10) are binned into a single thematic class of 100 (10 meters \* 10).
  - f. Every hardwood (deciduous) EVT gets CBH of 10 meters and CBD of 0.01 kg M<sup>3</sup>.
  - g. Shrub spotting condition pixels are assigned CC of 5%, CBH of 0.2m, CBD of 0.1 kg/m<sup>3</sup>, and CH of 9.1m, typically applies to California chaparral types.
  - h. Ensure that every EVT dominated by Juniperus spp receives a CBH value of 0.5m.
  - i. Ensure that every FBFM that requires canopy characteristics gets a valid value for CBH, CH, CC, and CBD.
- 5) **OUTPUT – Data Checks**
  - a. Run FLAMMAP on 3 test areas. Be sure output is realistic.
- 6) **Output Quantitative Accuracy Assessment**
  - a. 10 fold cross validation accuracy
  - b. Values of accuracy include Mean Absolute Error, Bias, and Pearson's Correlation Coefficient (r).

### **LANDFIRE Data Library Team**

The data library team at the MFSL assists in managing and processing all LANDFIRE data at various stages. As such, they perform several checks for data quality, integrity, and consistency throughout the process. These checks are outlined below.

**Processing step:** Spatial overlays to associate predictor variables with LFRDB plot locations

Primary personnel: C. Toney

Description: Plot lists and coordinates for non-FIA plots are received from the LFRDB team. Coordinates are attached to FIA plot ids. Plots are

overlaid on raster datasets to obtain pixel values from the following predictor layers:

9 DEM derivatives  
4 soil layers  
69 biophysical gradient layers  
30 bands of Landsat imagery  
6 intermediate LANDFIRE vegetation layers

The derived data are sent to the LFRDB team for incorporation into the database.

QA checks:

1. View plot locations spatially and verify all plots are inside the unbuffered map zone boundary (manual).
2. Visually inspect plot locations with satellite imagery and check for anomalous spatial distributions (manual).
3. Verify that the raster extent and cell size of each predictor layer match the reference layer for the map zone (automated).
4. Check derived output data for no-data values (automated)
5. Obtain the pixel row/column numbers for each plot coordinate (automated), which are used to detect potential duplication of plot data obtained from multiple sources, or identify areas with potential oversampling.

**Processing step:** Assemble and process data layers in preparation for Rectification procedures.

Primary personnel: J. Bramel, C. Toney

Description: Rectification is the internal LANDFIRE process that ensures logical consistency between the EVT and ESP map products. In order to compare these maps they must share a common data format, extent, and projection. The following layers are prepared for the rectification process:

Environmental Site Potential (ESP)  
Existing Vegetation (EVT)  
Existing Vegetation Cover (EVC)  
Existing Vegetation Height (EVH)  
Biophysical Settings (BPS)

Pre-rectification QC for ESP, EVT, EVC, EVH, BPS:

1. Input layer extents are checked for consistency against the in-house 3km mask (manual)

2. Projection information is validated (automated)
3. Visual inspection of each layer (manual)
4. NLCD and sparse vegetation pixel counts are checked for consistency across layers (manual)
5. Legends are checked for consistency and typos (manual)
6. Oftentimes there is a minor inconsistency between the masks used for the EVT-based maps and the ESP products. This inconsistency generally results in a consistent edge mismatch of 1-2 pixels. To correct this problem map edges are either clipped or filled using an ArcGIS Model. (automated)

Additional QA has been added for the EVT map:

1. EVT map is visually compared with imagery and NLCD92 (manual)
2. EVT map is generalized to lifeform and overlaid with the NLCD92 for comparison (automated)

**Processing step:** Incorporate rectification changes and produce final rectified vegetation layers.

Primary personnel: J. Bramel, C. Toney

Description: Tabular changes from the Rectification team are incorporated back into the vegetation grids. These grids are then passed on to the Fuels and Fire Regime teams and used as inputs to their processes.

Post-rectification QC:

1. NLCD pixel counts are checked for consistency across layers (automated)
2. Visual inspection of each layer (manual)
3. Change detection is run on pre and post-rectified versions of the maps (automated)

**Processing step:** Final deliverables prepared before sending to EROS for distribution on the National Map.

Primary personnel: H. Kreilick

Description: Final deliverables are received from data team leads by map zone. Spatial extents and no-data values are checked for consistency before sending to EROS. This ensures that there are no internal no-data values.

QA checks:

1. Visually inspect layer colormap and value attribute table (manual).
2. Ensure spatial extents match 3km buffered zone mask (manual).
3. Resolve spatial extent inconsistencies (automated).
4. Ensure no-data values are consistent with 3km grid (manual).
5. Notify team lead of any no-data issues and troubleshoot causes (manual).

**Processing step:** Verify that data distributed from the National Map are consistent with final deliverable layers archived in the LANDFIRE Data Library at MFSL.

Primary personnel: H. Kreilick

Description: Upon availability, a portion of each zone is downloaded from the National Map and compared to data that was sent to EROS. This ensures that data on the National Map are consistent with final deliverables archived at MFSL.

QA checks:

1. A portion of each zone for each layer is downloaded from the National Map, using the same extent for each layer per zone (manual).
2. Analysis is performed to verify that the downloaded data are equal to the sent data (automated).

## **Appendix B**

### **Accuracy Assessment “Holdout” Sampling Design**

The objective of the holdout sample is to provide a quantitative assessment of accuracy to supplement the cross-validation results. Withholding a sample of the available data from the classifier development process avoids the potential optimistic bias when accuracy is evaluated using the same data that was used to develop the classifier. Additionally, the holdout sample will be spatially separate from the training and cross validation data. That is, we will hold out sample plots within 3 km by 3 km blocks. The motivation for defining spatial units in the holdout sample derives from Friedl et al. (2000), who demonstrated that when training and accuracy assessment points came from within the same spatial units (clusters), the accuracy results were more optimistic than when the accuracy assessment points came from different spatial units from those used in classifier development. Friedl et al. (2000) explained this finding as resulting from the positive spatial correlation of the points within the clusters. An accuracy assessment point within the same cluster as training data would be more likely to be classified correctly because the training data in that same cluster carried relevant information to predict the accuracy point well.

#### **Proposed design**

Two main decisions are the size of the spatial units used to separate the holdout evaluation sample plots from the classifier training plots the sample size to withhold. A 3 km by 3 km block was chosen based on the subjective decision criteria that this would be a manageable block size to work with, and this size would provide reasonable assurance that most of the evaluation reference plots could be considered spatially independent of the training data. A systematic sampling design for selecting the 3 km by 3 km blocks was chosen because it would be simple to implement and it would provide good geographic coverage. We implemented a systematic design of every 7<sup>th</sup> block in both directions of a square grid. This design yields a sampling intensity of 2% (1 block out of every 49) and a projected sample size of approximately 100-200 accuracy assessment plots per LANDFIRE mapping zone in the west. All sample plots within the chosen sample blocks will be withheld from the classification process.

A systematic sample of this type is an equal probability sampling design, meaning that all plots would have the same probability of being withheld (0.02). Further, all subgroups of the plots would be sampled in the same proportion as their representation in the population. That is, if 20% of all available plots are condition A, on average, 20% of the holdout sample plots will be condition A plots. Stratified random sampling with proportional allocation has the same property, the difference being that it ensures that any stratified sample collected will have exactly the proportions of each stratum (subgroup) as are found in the whole region. But even if stratified sampling with proportional allocation is used, rare classes will still be represented by only a small number of plots in the total

sample, so the systematic design does not have a significant disadvantage in this regard.

The holdout sample is anticipated to provide adequate sample sizes for estimating accuracy for a combination of several LANDFIRE mapping zones. The primary reporting format will be at the support of the LANDFIRE Rapid Assessment Zones. The accuracy data and results will be available for each mapping zone, but the estimates derived for a mapping zone must be viewed with the recognition that the estimates will not be precise.

An alternative design to the holdout sample is to select a stratified random sample of plots to set aside, stratifying by the known class of the plot (e.g., vegetation type). An advantage of this stratified sampling design is that it ensures a specified sample size in each map class, including rare classes. The systematic sample of holdout “blocks” is an equal probability sampling design, and consequently the sample size for a rare class will be small given the overall small sample of holdout plots. Conversely, the stratified random sample would not create the spatial separation of plots achieved by the systematic holdout sample, so the two designs represent trade-offs of the advantage of larger sample sizes for the rare classes with spatial separation of the accuracy assessment evaluation plots from the classifier training reference data plots.

#### Spatial framework for conducting and reporting quantitative accuracy assessments

As mentioned above, the holdout sample of approximately 2% is designed for estimating accuracy for an area of about 4-5 mapping zones. These areas, called superzones, should be a) large enough to include an adequate number of holdout points; and b) ecologically similar, such that they will contain relatively similar vegetation map units. The PQWT proposes that Rapid Assessment (RA) Model Zones be used as a starting point for these superzones (see figure XX). RA Model Zones should be evaluated by technical teams at EROS, MFSL, and TNC to determine if any boundary lines should be shifted for the purposes of accuracy assessment superzones. RA Model Zones are spatially contiguous and ecologically similar aggregations of LANDFIRE mapping zones. They were established to facilitate development of vegetation mapping and modeling rules for the RA project. The PQWT also considered the use of polygons in the USDA Forest Service ECOMAP hierarchy (citation?) as superzones. While such areas would have ecological significance, the PQWT decided that it is preferable to have superzones that are clean aggregations of LANDFIRE mapping zones. Therefore, any adjustments made to RA Model Zone boundaries by technical teams must keep superzone boundaries coincident with mapping zone boundaries.

#### Proposed methods for conducting quantitative accuracy assessments

The PQWT will develop methods for conducting quantitative accuracy assessments of LANDFIRE vegetation layers (BpS, ESP, EVT, EVC, EVH).

Once methods are established, LANDFIRE technical teams will be responsible for conducting the assessments and preparing summary reports. Assessments will primarily be conducted on an entire superzone and utilize the 2% holdout sample. At the scale of a superzone, the 2% holdout sample will provide enough plots to make reasonable statements about the accuracy of most map units, with the exception of relatively rare or undersampled classes. At this time, the PQWT proposes several methods for assessing accuracy at a variety of spatial scales, as follows:

1. *Comparison of mapped pixel values to reference values at exact pixel locations of holdout points.* This method is an assessment of pixel-level accuracy, and provides estimates of overall accuracy, as well as an error matrix and estimates of producer's and user's accuracy for all map units represented in the holdout sample.
2. *Comparison of mapped pixel values to reference values within a 5x5-pixel area surrounding holdout point locations.* While the pixel-size of LANDFIRE deliverables is 30m, the effective thematic resolution of most layers is far coarser than this. There is also inherent error in the location of holdout points. With FIA plots in particular, the layout of the sample plot corresponds more appropriately to a 5x5-pixel area than to a single pixel (citation?). Therefore, by "growing" the spatial footprint of holdout points to a 5x5-pixel area (150m x 150m) we can possibly make statements about accuracy of LANDFIRE layers that are more in line with the data's thematic resolution. Following methods used by NLCD92 (citation?), mapped pixel values within the 5x5-pixel area may be compared to the reference value in a few different ways (e.g., agreement can be defined as any pixel, or alternately the majority of pixels, within the 5x5 area agreeing with the reference value). This type of assessment will also provide estimates of overall accuracy, as well as an error matrix and estimates producer's and user's accuracy for all map units represented in the holdout sample
3. *Test methods 1 and 2 above using only FIA holdout points.* The PQWT suggests that these tests be attempted, based on a) the spatially systematic design of FIA plots and b) the documented correspondence of FIA plot design to a 5x5-pixel footprint on the ground.
4. *Apply a fuzzy set assessment to the results of methods 1, 2, and 3.* Fuzzy set assessment recognizes that map units are not mutually exclusive and accounts for similarity between certain map units by adjusting the error matrix and recalculating overall accuracy, kappa, and producer's and user's accuracies accordingly (Gopal and Woodcock 1994, Green and Congalton 2004). Other recent mapping efforts by the GAP Analysis Program have also applied fuzzy set assessment concepts to maps using Ecological Systems as map units (Lennartz 2005, Lowry 2005, <http://earth.gis.usu.edu/swgap/mapquality.html>). Results of fuzzy set

assessments would not take the places of the “raw” results, but would be presented as additional information.

5. *Comparison of the composition of mapped pixels in a given area (superzone, mapping zone, watershed) to plots in the same area.* While not a true accuracy assessment, this type of compositional comparison could provide information about the appropriate scale of use for certain LANDFIRE layers. The idea behind this type of analysis is that if a classification unit (i.e., vegetation type) is represented in the plot database by a certain proportion of all plots in an area, a similar proportion of pixels should be mapped to this unit in the same area. At this time, the PQWT recommends performing such an analysis using a variety of spatial units, ranging from the entire superzone down to a small watershed (6<sup>th</sup> HUC?). The smallest spatial scale at which there is reasonable agreement between the plot vs. pixel composition could indicate an appropriate scale at which to use the data. One important element of a compositional comparison is a relatively even spatial distribution of plots. Therefore, the PQWT recommends first attempting this type of analysis with only FIA plots. If this yields too few plots for comparison, then additional plots should be selected using a spatially systematic approach (e.g., divide the landscape into 10km x 10km blocks and pull one plot from each block).

Finally, in addition to the above methods for estimating accuracy across a superzone, the PQWT recommends that some measures of accuracy also be reported by individual mapping zones. For maps created from See5 or Cubist models, cross-validation accuracies should be reported along with any error matrices and other appropriate summary statistics generated from the cross-validation analysis. Accuracy assessments using the holdout plots, such as those outlined in methods 1, 2, 3, and 4 above, may also be performed at the mapping zone scale, although there may be too few holdout points to yield meaningful results.

#### Proposed method for delivery of Product Quality information

For each superzone, LANDFIRE technical teams will produce a Product Quality report for each deliverable layer. For layers that will be assessed qualitatively through peer review, calibration workshops, or other methods, the Product Quality report will summarize all available information regarding the quality of the given layer for each map zone within the superzone, and possibly the superzone as a whole. For zones that will be assessed quantitatively with the methods outlined above, the Product Quality report will contain the following information:

- A summary of all accuracy assessment measures performed on the superzone, including a narrative, tables of overall accuracy, producer’s and user’s accuracy by class, and results of fuzzy set assessments;
- Full error matrices that would allow a user to conduct additional analyses if desired.

A summary of zone-specific cross-validation accuracies, and holdout accuracies if calculated, for all mapping zones in the superzone, including tables and error matrices where appropriate.

The actual format for delivery of LANDFIRE Product Quality reports remains to be determined. One possibility would be to make them available on-line via the [landfire.gov](http://landfire.gov) web site. The LANDFIRE PQWT will explore this issue before the assessments for the first superzone are completed. The PQWT and LANDFIRE technical teams will also consider which pieces of Product Quality information can be appropriately included in both the generic and spatially-explicit metadata prepared for each deliverable data layer.

## Appendix C LANDFIRE Model Q/C Checklist

Zone: \_\_\_\_\_

Date: \_\_\_\_\_

BPS \_\_\_\_\_

### Database

- BpS Name and Code match list provided exactly
- If BpS is split or lumped, those fields are checked and described
- General Information section complete
- Model evolution noted in comments field (if applicable)
- Peer review incorporated into comments field and reviewers listed (if applicable)
- Vegetation Classes section complete
- Structure data (height and % cover) mutually exclusive between classes
- Fire Regime section complete
- Appropriate Fire Regime Group selected
- References complete and alphabetized (double-check FEIS citation, if applicable)

### VDDT Model

- Disturbances (except Alt Succ) do not accelerate age
- Ages line up along main succession pathway
- Relative Age used only in A to A replacement disturbances
- If TSD used, probability of 1 and All Fire is TSD group
- Keep Rel is always false
- Run for 1000 years, 10 timesteps

	Database	Model
<b>A%</b>		
<b>B%</b>		
<b>C%</b>		
<b>D%</b>		
<b>E%</b>		
<b>AllFire</b>		
<b>Repl.</b>		
<b>Mixed</b>		
<b>Surf.</b>		

### Database-Model Matching

- Class labels (early#/mid#/late# and all/open/closed) match
- Percent in each class match
- Fire regime values match
- Non-fire disturbances and optional types checked appropriately.

## Appendix 4 Peer Review Form

**1. Save this form** with your last name and the 4-digit BPS ID number. Go to *File—Save As* and enter: *Name\_BPS#.doc*.

2. Basic Information			
Date of review:		Map Zone(s):	
BPS name:		BPS code (4-digit):	
Name:		Title:	
Affiliation:		Address:	
City:		State:	Zip:
Phone:		Email:	
Anonymity: (select one)	<input type="checkbox"/> <b>I would like my name listed as a reviewer.</b> (If you select this option, your feedback will be incorporated and your name will be listed on the BPS description as a reviewer.)		
	<input type="checkbox"/> <b>I would like to be an anonymous reviewer.</b> (If you select this option, your feedback will be incorporated and only the regional lead and national staff will know your name.)		

3. Rank your knowledge of this BPS.			
	Expert <sup>a</sup>	Knowledgeabl e <sup>b</sup>	Familiar <sup>c</sup>
a) How would you rate your understanding of the <b>fire regime</b> of this BPS throughout the entire model zone?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) How would you rate your understanding of the <b>succession processes</b> of this BPS throughout the entire model zone?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) How would you rate your understanding of the <b>composition and structure</b> of this BPS throughout the entire model zone?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) If your knowledge of this BPS varies considerably <b>geographically</b> , please indicate how it may vary within or between mapping zones (see map below). Leave blank if there is no substantial variation.			

<sup>a</sup>*Expert*: **In this BPS**, you have directed research or have at least 5 years of field experience, **and** feel confident in your understanding of the vast majority of related fire and/or other literature published in major professional journals.

<sup>b</sup>*Knowledgeable*: **In this BPS**, you have participated in research or have at least 3 years of field experience, **and** are familiar with some related fire and/or other literature published in major professional journals.

<sup>c</sup>*Familiar*. **In this BPS**, you have not directly participated in research and have less than 3 years field experience, but feel confident in your understanding of the majority of related fire and/or other literature published in major professional journals.

<b>4. Determine how you will perform your review.</b>		
Reviews can be performed via interactively modeling in VDDT (optional but preferred), or via review of BPS descriptions. If you choose to use VDDT software to review the models, ensure that you attribute the time definitions with at least 500 years (time steps), and use 10 Monte Carlo simulations.		
	<b>Review of BPS descriptions and the actual VDDT model</b>	<b>Review of BPS descriptions only</b>
<b>I performed this review via</b>	<input type="checkbox"/>	<input type="checkbox"/>

<b>5. Review the BPS description and model inputs, and answer the following questions.</b>
If you do not know the answer to any of the following questions, please enter “do not know”.
Assume that the reference fire regime and vegetation/fuels input and described for each BPS reflect historic conditions (i.e., pre-European settlement); and expected conditions if a natural fire regime were allowed to operate freely. Burning by Native Americans may or may not be considered part of the natural fire regime. Models are NOT intended to include states or processes that result from human-induced disturbances or management actions (except possibly Native American burning), and are constrained by the standardized model structure for this project (i.e., 3-5 classes (boxes) per model).

<b>5a. Rank this model overall.</b>				
Check one box for each row to classify your review overall. If you reject the model outright, please explain in further detail below.				
	<b>Accept as-is</b>	<b>Needs minor editing</b>	<b>Needs major editing</b>	<b>Reject outright</b>
<b>Model Description</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> <b>Model is redundant with another BPS (please specify):</b>  <input type="checkbox"/> <b>Model is not well-thought out or researched, for the reasons explained in the questions below.</b>  <input type="checkbox"/> <b>Other:</b>
<b>VDDT Model</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

<p><b>Introduction and Description</b></p> <p><i>The introductory BPS descriptions are intended to briefly describe the key factors that set this BPS apart from other BPSs. It should describe the geographic extent, biophysical site (e.g., major landform position, geologic substrate, elevation range), the vegetation, disturbance regimes, common adjacent BPSs, and information about scale.</i></p>		
5b.	Do the <b>introductory descriptions</b> adequately capture its distribution across the model zone? If not, what <b>specifically</b> should be added or removed from this description?	
<p><b>Mosaic of model classes A-E</b></p> <p><i>Model outputs summarizing the expected proportion of each of the reference model classes (A-E) in the BPS reflects the result of succession and disturbance processes operating concurrently over the long term.</i></p>		
5c.	Do the model <b>class descriptions</b> (A-E) appear to encompass the full spectrum of reference classes (including species composition, lifeform description, and canopy position) within the context of the standardized model structure (e.g., 3- to 5-box model)?	
5d.	Do the <b>structure data</b> described for each class (A-E) appear to be accurate, including percent cover and height of the upper layer lifeform, and tree size class (if applicable)?	
5e.	Do the <b>proportions of classes A-E</b> appear to reflect the landscape scale mosaic for this BPS ( $\pm 10\%$ for any one class) given a historic or reference fire regime?	<p>If No, please select one option:</p> <p><input type="checkbox"/> The proportions are inaccurate for the entire geographic area this model covers and the model should be rejected and remodeled. (Please complete additional questions above so that we know how to remodel this type.)</p> <p><input type="checkbox"/> The proportions are inaccurate for a subset of the geographic area this model covers, including these areas:</p>

**Disturbance Inputs**

Disturbance frequencies are translated to annual probabilities (1/ frequency in years) when entered into the VDDT model. Each disturbance can operate with different frequencies (i.e., different probabilities) and cause different transitions (i.e., have different effects) in different classes (i.e., boxes A-E). Fire disturbances are categorized in three severity classes (surface = <25% top-kill; mixed = 25-75% top-kill; replacement = >75% top-kill). Additional disturbances types may be modeled.

5f.	Does the <b>range in fire frequency</b> (fire return intervals) adequately capture the best available information for the BPS described?	
5g.	Are there <b>sources of published literature on fire frequency</b> that appear to be missing and which will change the range or central tendency of fire frequency used in the model if it were included? If so, provide the full citation of literature that should be considered.	
5h.	If you review the VDDT model, do the <b>differences in annual fire probabilities for each fire severity type by class</b> appear to capture the best available information on how fire frequency and type are distributed throughout this BPS? If not, specifically what should be changed within the model?	
5i.	Does the <b>distribution of fire severity</b> between stand replacement (>75% top-kill), mixed severity (25-75% top-kill) and surface (<25% top-kill) fire regimes adequately capture the best available information for the BPS described?	
5j.	Are there <b>sources of published literature on fire severity</b> that appear to be missing and which will change the distribution of fire severity used in the model if it were included? If so, provide the full citation of literature that should be considered.	

5k.	Are there any major <b>non-fire related disturbances</b> (e.g., hurricanes, insects) that have not been captured by the model? If so, what are they? For each, what would you estimate are their mean, minimum and maximum return intervals and severities (e.g., stand replacement, mosaic)? Which classes (A-E) do each operate in?	
-----	---	--

<b>Additional Feedback</b>		
5l.	Other comments, suggestions, or feedback.	

**When complete, please email this form to your regional lead.**

**THANK YOU FOR COMPLETING THE MODEL REVIEW!**

## **Appendix D**

### **Excerpts from LANDFIRE Vegetation Model Peer-Review Process**

#### **Summary for Regional Leads**

**May 2005**

#### **Introduction**

Soliciting and incorporating review are both explained below. Please consider the following general guidelines:

- Use PDF files whenever possible. This will reduce the file size, prevent confusion with review comments, and give you more control over how review comments are incorporated. If you do not have a PDF writer, see instructions below for downloading a free PDF-writer.
- Maintain an honest, open review process. Document feedback and comments by keeping folders of all communications and tracking feedback for each BPS in the ModelTracker Database.
- When in doubt, contact the national staff for clarification.

#### **Installing a PDF-Writer**

If you do not have access to a PDF-writer, you can download one for free at CutePDF (<http://www.cutepdf.com/Products/CutePDF/writer.asp>). You will need to download and install two programs (the CutePDF writer and a Ghostscript), both available at that address. When the CutePDF Writer is installed, you will be able to create PDF files by going to *File—Print* and selecting *CutePDF Writer* as your printer.

#### **Soliciting Review**

- 1) Complete the quality-control process on models and database entries.
- 2) Compile the list of reviewers. This should include:
  - a. All workshop participants.
  - b. Reviewers identified by modelers (in the ModelTracker Database).
  - c. People who were unable to attend the workshop, but expressed interest in providing review.
- 3) Create the summary list of BPSs using the *Summary for Review* report:
  - a. Open your master database and close the form that opens automatically.
  - b. Select *Reports* on the left menu bar, and open the *Summary for Review* report that will appear in the options at right.
  - c. When the report opens, turn it into a PDF file (see instructions above, if you don't have a PDF-writer).

- 4) Send an email to all of your reviewers soliciting their input. You may also want to send personalized messages or phone calls to key reviewers. An example invitation to review is in Appendix B. In the invitation, you should:
  - a. Invite them to tell you which models they can review. This will help you determine if you need to actively solicit reviewers for any BPSs.
  - b. Provide some background about the creation of the models.
  - c. Provide the list of BPSs modeled.
  - d. Attach the Summary for Review report (step 2 above) summarizing general information about each BPS.
  - e. Explain that model review can happen with or without VDDT and will require 30 minutes to 2 hours per model.
  - f. Offer the LANDFIRE website as a resource, including the modeling manual ([www.landfire.gov/workshop.html](http://www.landfire.gov/workshop.html)).
  - g. Note that if they express interest, you will forward more information.

### Making Model Data Available

- 1) Create a PDF document out of the database report for each BPS.
  - a. Open your master database. For each BPS (record/page) in the form, click on the *View Report* button. This will allow you to look at individual reports for each BPS. Create a PDF for each BPS.
  - b. Name each PDF file with the code for the BPS (e.g., Rapid Assessment: ROPIPO.pdf; LANDFIRE: 1125.pdf)
- 2) For each BPS, create a folder named for the BPS's code that contains the VDDT files and the database description:
  - a. CODE.pvt (VDDT file)
  - b. CODE.loc (VDDT file)
  - c. CODE.scn (VDDT file)
  - d. CODE.pdf (description document)
- 3) Upload your BPS data in a single zipped folder to the Model Reviewer Website.
  - a. Implement the following file structure. ***It is imperative that you use this file structure so that the website structure works and that all documentation for reviewers is consistent.*** Items in **bold** in the table below show text that must match exactly.

For LANDFIRE Models:	(Example)
<ul style="list-style-type: none"> <li>📁 LF_ModelZone_Date</li> <li> <ul style="list-style-type: none"> <li>📁 <b>Mapzone#</b></li> <li> <ul style="list-style-type: none"> <li>📁 MZCODE</li> <li>📁 MZCODE</li> <li>📁 MZCODE</li> <li>📁 (etc.)</li> </ul> </li> <li>📁 <b>Mapzone#</b></li> <li> <ul style="list-style-type: none"> <li>📁 MZCODE</li> </ul> </li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>📁 LF_GreatBasin_03_18_05</li> <li> <ul style="list-style-type: none"> <li>📁 Mapzone16</li> <li> <ul style="list-style-type: none"> <li>📁 161011</li> <li>📁 161012</li> <li>📁 161016</li> <li>📁 (etc.)</li> </ul> </li> <li>📁 Mapzone12</li> <li> <ul style="list-style-type: none"> <li>📁 121011</li> </ul> </li> </ul> </li> </ul>

📁 MZCODE	📁 121012
📁 MZCODE	📁 121016
📁 (etc.)	📁 (etc.)

- b. Create a zipped folder of the highest level folder above (e.g., LF\_ModelZone\_Date.zip). With most compression software, you can simply right-click on the folder in Windows Explorer and select “Create LF\_ModelZone\_Date.zip”.
  - c. Copy the single zipped folder to Jeannie Patton’s FTP site. Go to <ftp://ftp.tnc.org/>. Login as “partner” with the password “roanoke” under the *File—Login As* menu. Scroll to the *jpatton* folder and open it; open the *To Upload to Reviewers Website* folder. Paste your zipped folder there. Send Jeannie an email (at [jpatton@tnc.org](mailto:jpatton@tnc.org)) alerting her that your data is ready to be uploaded to the website. When she has uploaded the data, she will send you an email.
  - d. Already uploaded to the website are data files that are universal to the review process, including:
    - ReadMe Instructions for Review explaining the review process.
    - LFModelReview: a folder containing data and maps needed for model review process.
4. Send reviewers to the Model Review Website to download data using the following link: <http://www.tnc-ecomangement.org/Fire/Workshops2/workshops4.cfm>. This website is private and buried in the FLN website (i.e., there are no external links to it). It will soon be replaced by a site on nature.org, and we will inform you when that change occurs.

## **Incorporating Review**

### *General Guidelines:*

- Track all written feedback from reviewers in a systematic fashion to maintain a review process that is open, honest, and as robust as possible.
- Document all opinions and comments to the degree possible.
- If a reviewer chooses to be anonymous, protect their anonymity. List the number of anonymous reviewers in the Comments field in the database. (E.g., “Comments were provided by 2 anonymous reviewers.”)
- If a reviewer does not choose to be anonymous, please give them credit in the Reviewers section of the database. If there are more than three named reviewers, list additional reviewers in the Comments field in the database. (E.g., “Additional reviewers included Joe Bob ([joebob@tnc.org](mailto:joebob@tnc.org)) and Billy Jean ([bjean@tnc.org](mailto:bjean@tnc.org)).”)

### *Descriptive Changes*

In most cases, review will suggest descriptive changes only. Incorporate these as much as possible.

### *Quantitative Changes*

When reviewers suggest that changes should be made in the quantitative values for a model (e.g., percent in each class or frequency of fire), you may want to:

- 1) Make the changes in the model to see how significant the difference will be. If they are minor (e.g., within +/- 5% of a class, or within +/- 5 years in the fire frequency), you may choose to make the change.
- 2) Check with the modeler(s) to see how the requested changes differ from their understanding of the system.
- 3) Check with other reviewer(s) to determine if there is agreement about whether the changes should be made. Consider the knowledge self-ranking of each reviewer on the Model Review Form.
- 4) If there is not agreement among the modeler(s) and reviewer(s) document the dissenting opinions. You should document the differences in the Issues/Problems section of the database and any other section that the differences apply to. For example, if there is a debate about the frequency of fire, you may document it in the Issues/Problems and Disturbance Description sections.

### *Changes to the BPS list*

Changes cannot be made to the list of BPSs after the workshop. If additional types are requested, they can be modeled for regional needs by regional staff, but will not be included in the Rapid Assessment (unless, in some cases, you can provide mapping rules to Jim Menakis).

### *Widely Disparate Opinions*

In a small number of cases, there may be widely disparate opinions about a BPS. These will generally be BPSs with little data, high geographic variability, or contentious policy issues. In cases where disparate opinions limit your ability to incorporate review, or where you feel there is potential for conflict, please let the national RA staff know. For select BPSs, we will ask for additional review from a third party. In all cases, we will document all opinions.

### **Tracking Model Evolution**

Models will go through many iterations and may be used in multiple mapzones. When LANDFIRE is complete, we will try to consolidate duplicate models. For these reasons, it is imperative that we track model evolution carefully. Please let the following principles guide you when tracking model evolution:

- Be careful about changing authorship. The model author(s) (under *Modelers*) should generally be the person(s) who made the final adjustments to the VDDT model. Unless descriptive changes significantly alter the original meaning of the model, they do not change authorship.

- Track all contributors. If the model is based on another LANDFIRE model or a Rapid Assessment model, be sure list which model (see more below) and list contributors under *Modelers*, *Reviewers*, or *Comments*.
- Track mapzones. List only mapzones for which you know this exact model is used. If you are duplicating a model from another mapzone, list your mapzone and the previous mapzone.
- List duplicate models and previous iterations. Under *Comments*, track the model evolution:
  - *If this model is an exact duplicate of another LANDFIRE model*, type in the first sentence: “This model is identical to the model for the same BpS from mapzone XX [other mapzone]. It did/did not receive any peer review for mapzone YY [current mapzone].”
  - *If this model is similar to or based on another LANDFIRE model*, type in the first sentence: “This model is based on the LANDFIRE model for BpS ##### for mapzone ##. It did/did not receive any peer review for mapzone YY [current mapzone].”
  - *If this model is based on or is a duplicate of a Rapid Assessment model*, type in the first sentence: “This model is based on the Rapid Assessment model R#PNVG, authored by Joe Smith ([joe@email.com](mailto:joe@email.com)) and reviewed by Sue Smiley ([sue@email.com](mailto:sue@email.com)). Adjustments were made by [Reviewer or Regional Lead and email address] to make the model consistent with LANDFIRE BpS.”

### **Estimated Timelines**

In general, we will acquire from NatureServe (NS) the list of BPS by mapzone and the regional leads will have no less than two months to complete the models and review process before they deliver the models to national staff (i.e., Kelly or Darren). The national staff will have one month to compile and quality control models before delivering them to the Missoula Fire Sciences Lab (MFSL). All dates for LANDFIRE are hard and fast; there is very little flexibility in the schedule. Sequence and delivery dates for eastern U.S. and Alaska mapzones are not yet determined.

## Appendix E LANDFIRE Model Delivery Checklist

Mapzone: \_\_\_\_\_ Date: \_\_\_\_\_ Final  
QC: \_\_\_\_\_

### ***In MDTB.mdb:***

- Q/C all models (see separate checklist).
- Change the name of the database to be ##MTDB\_YYYY\_MM\_DD.mdb (where ## is the mapping zone number; use date of delivery).
- Update all BpS codes.
  - Add two digits to the beginning of the code to represent the mapping zone number.
  - Add a 0 (zero) to the end of the code, unless a model has been split, in which case number the splits sequentially starting with 1.
- Check all BpS names and make sure they are identical with the Ecological System Names.
- For any lumped models, make sure the entries are identical, that the lumped BpS ID is noted, and that the “Lump” field is checked.
- For any split models, make sure that the explanation of how to split the BpS is clear and that the “Split” field is checked.
- Double-check that the appropriate class names are used (names should be complete, not abbreviated). If necessary, do a Find/Replace to correct abbreviated names.
- Check that the correct lifeforms are listed (if time allows).
- Spell-check the *AllModels* table (if time allows).

### ***In VDDT.mdb***

- Change the name of the database to be ##VDDT\_YYYY\_MM\_DD.mdb (where ## is the mapping zone number; use date of delivery).
- Update all BpS codes and names to be identical with those in Model Tracker. (You can copy/paste the fields from one Access database to the other.)
- Check all BpS names and make sure they are identical with the Ecological System Names. (You can copy/paste the fields from one Access database to the other.)

### ***Cross-Check***

- Check the BpS in Model Tracker against the BpS plot list for the zone and against the VDDT database to ensure that we are delivering the correct BpS and we have a description and model for each.

### ***Prepare Data:***

- Create a PDF of all descriptions in the MTDB.mdb using the Report called *AllModels*. Name this ##Descriptions\_YYYY\_MM\_DD.pdf (where ## is the mapping zone number; use date of delivery).

- ❑ Create a zipped folder named ##Veg\_Models\_YYYY\_MM\_DD.zip (where ## is the mapping zone number; use date of delivery) containing:
  - ReadMe.pdf
  - ##Descriptions\_YYYY\_MM\_DD.pdf
  - ##MTDB\_YYYY\_MM\_DD.mdb
  - ##VDDT\_YYYY\_MM\_DD.mdb
  
- ❑ Deliver the zipped folder to Missoula Fire Sciences Laboratory via FTP ([ftp://ftp.tnc.org/kpohl/LANDFIRE\\_Model\\_Delivery](ftp://ftp.tnc.org/kpohl/LANDFIRE_Model_Delivery)) (username = partner; password = roanoke).
  
- ❑ Send an email to our colleagues telling them the data is uploaded. The following should be included in the email:

MFSL	Ann Wolf	<a href="mailto:atwolf@fs.fed.us">atwolf@fs.fed.us</a>
	Brendan Ward	<a href="mailto:bward@fs.fed.us">bward@fs.fed.us</a>
	Don Long	<a href="mailto:dlong01@fs.fed.us">dlong01@fs.fed.us</a>
	Greg Dillon	<a href="mailto:gdillon@fs.fed.us">gdillon@fs.fed.us</a>
	Sarah Pratt	<a href="mailto:sdpratt@fs.fed.us">sdpratt@fs.fed.us</a>
Tech Transfer	Steve Barrett	<a href="mailto:sbarrett@mtdiq.net">sbarrett@mtdiq.net</a>
TNC	Darren Johnson	<a href="mailto:darren_johnson@tnc.org">darren_johnson@tnc.org</a>
	Elena Contreras	<a href="mailto:econtreras@tnc.org">econtreras@tnc.org</a>
	Kori Buford	<a href="mailto:kbuford@tnc.org">kbuford@tnc.org</a>
	Randy Swaty	<a href="mailto:rswaty@tnc.org">rswaty@tnc.org</a>
	Jim Smith, PM	<a href="mailto:jim_smith@tnc.org">jim_smith@tnc.org</a>
	Regional Lead	<Regional Lead>