

# LANDFIRE EVT AutoKey Input Data

Karen C. Short

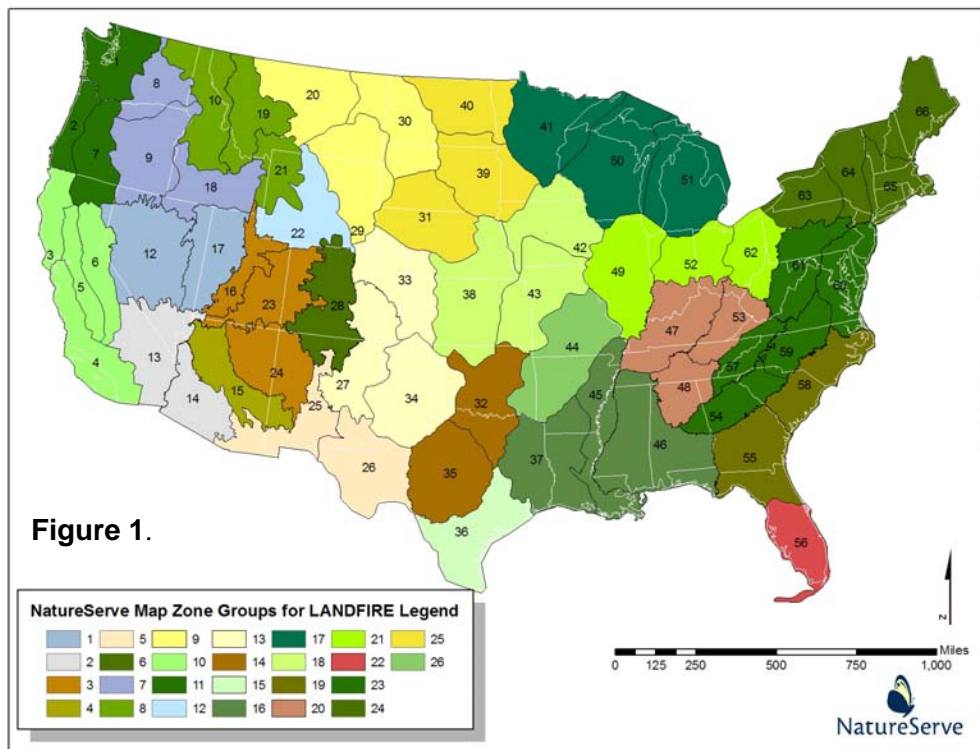
24 October 2006

## OVERVIEW

Three input tables are needed for the automated processing of field-sampled plot data for LANDFIRE Existing Vegetation Type (EVT) mapping. These three tables are: 1) *EVT*, 2) *Plots*, and 3) *Species*. The *EVT* table comprises the ruleset (a.k.a. “key”) used to assign any given plot to one of the candidate EVT’s. The *Plots* table includes plot-level floristic and topo-geographic data from field-sampled units. The *Species* table includes species-level data from the same field-sampled units reported in the *Plots* table.

Two additional tables included in the Key Database contain ancillary plot and species information needed to generate the full suite of output tables from the AutoKey (a.k.a. “key classifier”) program and to run the diagnostic queries. These two tables are: 1) *Plots Descriptive* and 2) *Spp List*. The *Plots Descriptive* table includes plot-level project and additional topo-geographic information as well as links to any digital plot photos. The *Spp List* reports the lifeform (i.e., whether tree, shrub, graminoid, or forb) of each plant included in the *Species* table.

These five input tables are populated through the coordinated efforts of NatureServe ecologists and the LANDFIRE Reference Database (LFRDB) team working for Systems for Environmental Management in Missoula, Montana. The plot data (i.e., in the *Plots*, *Species*, and *Plots Descriptive* tables) are compiled by the LFRDB team on a LANDFIRE mapzone-by-mapzone basis. The *EVT* table is created and the *Spp List* table is informed by NatureServe ecologists on a “mapzone group” basis (e.g., see Figure 1), although refinement of the *EVT* table at the mapzone level may be ultimately necessary.



## DEVELOPING THE EVT AUTOKEY INPUTS

The steps in development of any given set of input tables are as follows:

- 1) **Species list review** – A list (spreadsheet) of plants reported from all plots compiled for a mapzone group is provided by the LFRDB team to the NatureServe ecologist(s) who will create the *EVT* table for that cluster. Included in the spreadsheet are the lifeform (i.e., tree, shrub, subshrub, vine, forb, graminoid) and, if herbaceous, duration (i.e., annual, biennial, perennial) of each plant as per the NRCS PLANTS database. The PLANTS database often indicates that a given plant can be considered one of several lifeforms and/or life-cycle types. However, for the purposes of auto-keying plots to LANDFIRE EVTs, each plant in the list must be assigned a *single* lifeform (tree, shrub, forb, or graminoid) and, if herbaceous, a *single* duration (annual or perennial) *for the mapzone group at hand*. The LFRDB team relies on the NatureServe ecologist(s) to identify the lifeform and, if relevant, life cycle that best characterizes each plant in the mapzone-group list. (If particular taxa cannot be “pigeonholed” into a single lifeform or duration category, that is acceptable, but see caveats under 3 and 4 below.)
- 2) **Creation of the *Spp List* table** – The *Spp List* table, which reports the lifeform of all plants that may be included in the *Species* table and is needed to run the diagnostic queries in the Key Database, is populated from the reviewed species list for the mapzone group.
- 3) **Creation of the *Species* table** – This table includes the plant-composition data for each plot. Because it includes relative cover estimates (by lifeform), it cannot be populated until the appropriate lifeform for each species has been identified. In other words, the relative cover of Gambel oak, for example, cannot be determined until after it is known whether Gambel oak is to be considered a shrub or a tree in the mapzone group at hand. Unknown species and taxa that cannot be pigeonholed into a single lifeform during the species list review will necessarily lack relative cover values in the *Species* table.
- 4) **Creation of the *Plots* table** – This table affords plot-level estimates of total vascular plant cover and cover by lifeform or lifeform-life cycle (i.e., perennial herbs), and is thus dependent on the lifeform and life cycle assignments in the reviewed species list for the mapzone group. Unknown species will contribute to the total vegetation cover estimate only, and taxa that cannot be pigeonholed into a single lifeform (or life cycle if herbaceous) cannot be factored into their relevant total cover estimates.
- 5) **Creation of the *Plots Descriptive* table**
- 6) **Creation of the *EVT* table** – Once the four other inputs are established, the rest of the inaugural Key Database is developed using a “generic” version of the *EVT* table. The AutoKey program and diagnostic queries are run repeatedly in the process of developing the final, full-blown *EVT* table(s) for the mapzone group. The required format of the *EVT* table is defined in Appendix A.

## SPECIAL CONSIDERATIONS

The following should be borne in mind when exploring the data in the input tables:

- 1) ***Spp List*** – The processing of field-sampled vegetation data for LANDFIRE is a multi-year endeavor. Part of this data processing is to map each of the plant names provided in each original database to the appropriate Symbol Key and botanical name in the NRCS PLANTS database. The PLANTS database, however, is a dynamic entity. It is updated periodically to reflect changes in taxonomy, nomenclature, and accepted characterizations of distribution, growth habit, duration, and nativity. For consistency’s sake, the LFRDB team maps to a “snapshot” of the NRCS PLANTS database taken in January 2005. It is this “frozen” list of plant names and symbols that we use to standardize the nomenclature in the Key Database and the LFRDB. The only

exception made at this point is to add any true *species nova* (e.g., *Sporobolus pinetorum*) as encountered in the plot data.

- 2) **Species** – 1) Only data reported for *live* plants are included. 2) The design of this table reflects our very early consideration of using basal area to “key” plots when tree-cover estimates were unavailable (e.g., as is the case with much Forest Inventory and Analysis data). However, to do so seemed to necessitate development of separate EVT and Plots tables for basal area and cover data (and introduced other data-processing complications). We have instead opted to derive percentage canopy cover from stem diameters of individual trees as informed by Bechtold (2003, 2004) and Crookston and Stage (1999), on an as-needed basis.
- 3) **Plots** – The percentage cover values provided in this table have been calculated by summing the *Species* cover values across relevant lifeform/life cycle categories within each plot (i.e., shbcov = the summed cover of all shrubs). Such derivation is necessary to ensure consistency with the lifeform/life cycle assignments in the mapzone-group species list. The drawback of these estimates is that, because we must sum across species to get them and we’ve no way of adjusting for potential overlap among species’ canopies with any confidence, they invariably overestimate total cover in multistoried stands (and can exceed values of 100%). (The definition of that “20” code in those method fields is essentially as above: the data were derived by summing the *Species* cover values across relevant lifeform/life cycle categories within each plot.)
- 4) **FIA** – Data collected for the Forest Inventory and Analysis (FIA) Program and certain other efforts must be “dissolved” from a cluster of “subplots” into a single plot record for our purposes, as only a single geo-reference is provided for the cluster. In these cases, we filter for “homogeneous” (i.e., single condition) clusters before amalgamating the component data into a single record.

#### LITERATURE CITED

- Bailey, R.G. 1994. Ecoregions of the United States. USDA Forest Service, Washington DC.
- Bechtold, W.A. 2003. Crown-diameter prediction models for 87 species of stand-grown trees in the eastern United States. *Southern Journal of Applied Forestry* 27:269-278.
- 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.
- Crookston, N.L., and A.R. Stage. 1999. Percent canopy cover and stand structure statistics from the Forest Vegetation Simulator. USDA Forest Service, Rocky Mountain Research Station, General Technical Report RMRS-GTR-24.
- McNab, W.H., D.T. Cleland, J.A. Freeouf, J.E. Keys Jr., G.J. Nowacki, and C.A. Carpenter. 2005. Description of ecological subregions: sections of the conterminous United States. USDA Forest Service, Washington DC.
- Omernik, J.M. 1995. Ecoregions: a framework for managing ecosystems. *George Wright Forum* 12:35-50.

## APPENDIX A

The EVT sequence table is the source of the rules used to assign a given plot to a given LANDFIRE EVT. This table must conform as follows for the plot-assignment program to perform as intended:

Field Name	Data Type	Description	Format Rules
Row_No	Number (currently Double)	Defines sequence of rules against which to evaluate plots	Unique for each row
EVT_Code	Number	Code for the LANDFIRE EVT to which this rule applies	If blank, row will never be used to test plots
ESLF_Label	Number	Code for corresponding Ecological System from NatureServe's Systems database	
MapZone	Text	LANDFIRE mapping zone(s) for which this rule applies	May be blank or a series of comma-delimited numbers
EVT_Name	Text	Name of the LANDFIRE EVT to which this rule applies	
totcov	Text	Criterion: Total canopy cover (%) of all vegetation on the plot	May be blank or contain a condition (<X%, >X%, =X%, X-Y%, present, absent)
wdycov	Text	Criterion: Total canopy cover (%) of all woody vegetation on the plot	May be blank or contain a condition (<X%, >X%, =X%, X-Y%, present, absent)
treecov	Text	Criterion: Total canopy cover (%) of all trees on the plot	May be blank or contain a condition (<X%, >X%, =X%, X-Y%, present, absent)
conifertotcov	Text	Criterion: Total canopy cover (%) of all conifer trees on the plot	May be blank or contain a condition (<X%, >X%, =X%, X-Y%, present, absent)
shbcov	Text	Criterion: Total canopy cover (%) of all shrubs on the plot	May be blank or contain a condition (<X%, >X%, =X%, X-Y%, present, absent)
pherbcov	Text	Criterion: Total canopy cover (%) of all perennial herbaceous plants on the plot.	May be blank or contain a condition (<X%, >X%, =X%, X-Y%, present, absent)
grasscov	Text	Criterion: Total canopy cover (%) of all graminoids on the plot.	May be blank or contain a condition (<X%, >X%, =X%, X-Y%, present, absent)

Field Name	Data Type	Description	Format Rules
spp1	Memo	Criterion: First set of indicator species.	May be null or contain comma-delimited list of species names. Valid elements of the latter are: Genus, Genus species, Genus “spp.”, Genus species “ssp.” subspecies, Genus species “var.” variety, Genus species “ssp.” subspecies “var.” variety
relcov1	Text	Criterion: Relative cover (%) threshold or presence/absence criterion for first set of indicator species.	If species group is defined, must afford a condition (<X%, >X%, =X%, X-Y%, present, absent)
spp2	Memo	“	“
relcov2	Text	“	“
spp3	Memo	“	“
relcov3	Text	“	“
spp5	Memo	“	“
relcov5	Text	“	“
spp6	Memo	“	“
relcov6	Text	“	“
spp7	Memo	“	“
relcov7	Text	“	“
spp8	Memo	“	“
relcov8	Text	“	“
spp9	Memo	“	“
relcov9	Text	“	“
spp10	Memo	“	“
relcov10	Text	“	“
Subsection	Text	Criterion: Ecological subsection, as described in McNab et al. (2005).	May be blank or a series of sections or subsections delimited by commas
Elevation	Number	Criterion: Range or threshold of elevation (meters) above mean sea level (extracted from the DEM).	May be blank or contain a condition (<X, >X, =X, X-Y)
EPA_Ecoregion	Text	Criterion: Ecoregion (level III or IV) as described in Omernik (1995).	May be blank or a series of ecoregions (level III or level IV)
TNC_Ecoregion	Text	Criterion: Ecoregion based on Bailey (1994) as modified for The Nature Conservancy’s planning purposes.	May be blank or a series of ecoregions

<b>Field Name</b>	<b>Data Type</b>	<b>Description</b>	<b>Format Rules</b>
planted	Number	Criterion: The original data source indicates that the plot is in a planted stand: 1 = yes.	Enter 1 if you wish to restrict plots that key out in a given line to those identified as planted, or plantations, in the original datasets. Otherwise leave blank.