Journal of Forestry, 2019, 1–6
doi:10.1093/jofore/fvz050
Field Notes - silviculture
Received July 27, 2019; Accepted August 14, 2019
Advance Access publication October 07, 2019

Field Notes - silviculture

Field Note: A New Conversion of Forest Inventory and Analysis Data for Use in the Forest Vegetation Simulator

John D. Shaw and Aaron Gagnon

John D. Shaw (john.d.shaw@usda.gov), USDA Forest Service, Rocky Mountain Research Station, Ogden, UT 84401. Aaron Gagnon (aaron.gagnon@usda.gov), USDA Forest Service, Forest Management Service Center, Fort Collins, CO 80526.

Forest Inventory and Analysis (FIA) data are useful for a wide range of applications (Shaw 2017), including research, National Forest planning, and private consultation. For some of these applications, FIA data are the primary source of input for projections done using the Forest Vegetation Simulator (FVS; Dixon 2002). For the past several years, there has been a barrier to full use of FIA data in FVS, because of rapid changes in the FIA database (FIADB) and difficulty maintaining a translation process that correctly and comprehensively produces FVS-ready tables from FIADB. In addition, the use of Microsoft Access as a data-distribution format was discontinued by the FIA program, because of the fact that FIADB contents for several individual states had exceeded the 2 gigabyte capacity of Access. The need for an updated translation process and new data distribution approach warranted a new approach to both issues.

The FIA program first provided FVS-ready data through its Mapmaker program in the early to mid-2000s, resulting in a large number of state-level retrievals by users (Miles 2008). However, changes to Forest Service information technology infrastructure led to the retirement of Mapmaker at version 2.1. Given the lack of a translation tool, Shaw (2009) described a "roadmap" that could be used to develop a next-generation translation process, should the investment of programming resources be made.

The need for a replacement translation process was met by the Forest Service Forest Management Service Center, which is responsible for development, maintenance, training, and user support for FVS. The new approach to translation was a stand-alone application called FIA2FVS (hereafter FIA2FVS 1.0; Vandendriesche 2012), which could use data from FIADB as ASCII (text) tables or in the form of a Microsoft Access database, and produce FVS-ready data in the form of an Access database formatted for use with the FVS Database Extension (Crookston et al. 2003). Because of rapid changes to FIADB and the need to support FIA2FVS as an end-user application, the maintenance burden became unmanageable, and FIA2FVS 1.0 support was dropped. However, discontinuation of FIA2FVS 1.0 was based on agreement between the Service Center and FIA that FIA would take the lead on a replacement process.

Translating FIA Data to FVS-Ready Format

The new effort to convert data from FIADB into FVS-ready format continued under the FIA2FVS label (hereafter FIA2FVS 2.0), using FIA2FVS 1.0 code and logic as a starting-point. However, several development goals made FIA2FVS 2.0 a substantially improved process as compared to the original version:

- The translation process would correct any data issues known to exist. Most notable of these were mismatches between site index base age used in FIADB and the base ages expected by FVS, which could result in over- or under-prediction of individual tree growth.
- FIA plot design information would be preserved in the translated data when possible.

Management and Policy Implications

The Forest Inventory and Analysis database (FIADB) and the Forest Vegetation Simulator (FVS) comprise a powerful toolset that can be used for a wide variety of applications. Both are available to the public and cover all forest types and ownerships in the US states and territories. This allows managers and researchers to assess the status and trends of a wide variety of forest types over large areas, and evaluate the outcomes of possible future management scenarios. As the body of Forest Inventory and Analysis (FIA) data grows under the annualized inventory system, plot remeasurements eventually become validation sources for projected forest conditions, in both managed and unmanaged stands. Current managers and planners can use FIA data collected over the past 20 years to develop plans for future management direction, and future managers and planners will have the capability to assess the results of those plans in combination with unplanned influences, such as fire and insect outbreaks. The increased integration of FIADB and FVS will include development of a series of user guides, which have not been available with previous data-delivery efforts, to help managers and planners use FIA data from past plot visits to inform the development of FVS simulations.

- The translation burden would be removed from the end-user; users would need only to download data for the geographic area of interest.
- Data would be translated in a multiformat configuration, in response to user requests.
- FVS-ready data would be provided in SQLite database format as a solution to the 2 gigabyte limit of Access databases.
- Because of the storage capability afforded by SQLite, FVS-ready tables and all of the publicly available FIADB tables would be included in one database, opening up capabilities that have not previously been available to users.
- FIA2FVS 2.0 will be supported by a substantial body of documentation that not only helps users understand how the translation process was done, but provides guidance on how to leverage FIA data that are not used directly by FVS, and how to use FIA data and FVS together to accomplish certain tasks.

Translation specifications were based on Shaw (2009), the "blank" Access database provided on the FVS website, and FIADB documentation. There have been many changes since Shaw (2009) was published, and the last version of FIA2FVS 1.0 was released, such as the addition of condition-level, per-acre weights for downed woody materials in FIADB, and the expansion of fuel variables in the FVS Fire and Fuels Extension (Rebain 2010) to more piece size classes and separate variables for hard and soft pieces.

For some situations, the translation of FIA variables into FVS-compatible format is not straightforward, so translation required compromise or the establishment of rules for which there is not necessarily one "correct" approach. This is best understood using a comparison of FIA sampling and stand data typically prepared for FVS, and the resulting data structures.

In a typical forest inventory, stands are commonly delineated prior to sampling based on age, composition, density, minimum size, and other stratifying features. Plots are then arranged within each stand according to the inventory design (e.g., Figure 1A). In FIA inventories, stand boundaries are not identified prior to the time of plot visit. Rather, there is a pre-determined plot center location, to which the crew navigates using aerial photography, GPS, and other reference information. Stand boundaries are then determined in relation to the location of subplots used in the applicable plot design. It should be noted here that FIA uses the term "condition" rather than stand. In most cases, the terms are synonymous, but FIA delineates forested conditions based on six criteria: reserved status (available for or prohibited from harvest), owner group, forest type, stand size class, regeneration status (natural versus artificial), and tree density. As a result, it is possible for an otherwise homogeneous stand to be divided based on ownership or, for example, a wilderness area boundary.

For most FIA periodic inventories with data included in FIADB (prior to ca 2000), the plot footprint was intended to sample only one condition, and that condition was determined by the plot center. If the subplot arrangement straddled condition boundaries, all outside subplots would be moved into the sampled condition according to a prescribed procedure (Figure 1b). Under the current four-subplot, mapped-plot design (Hahn et al. 1995; Bechtold and Scott 2005), which has been in use since the late 1990s, for a majority of plots sampling forest conditions all subplots sample a single condition (Figure 1C). However, unlike most periodic inventory designs, subplots are not moved when the plot footprint samples more than one condition. Instead, condition boundaries are mapped (Figure 1D).

In typical FVS input data, plots sampled within a stand (i.e., Figure 1A) are linked to the stand with a unique identifier (Stand_ID), and projections can be run for all plots within a stand or by individual plot. During most FIA periodic inventories, FIA plots typically sampled a single condition, with the subplots arranged within the condition boundary. In this comparison, there is effectively equivalence between the area sampled by

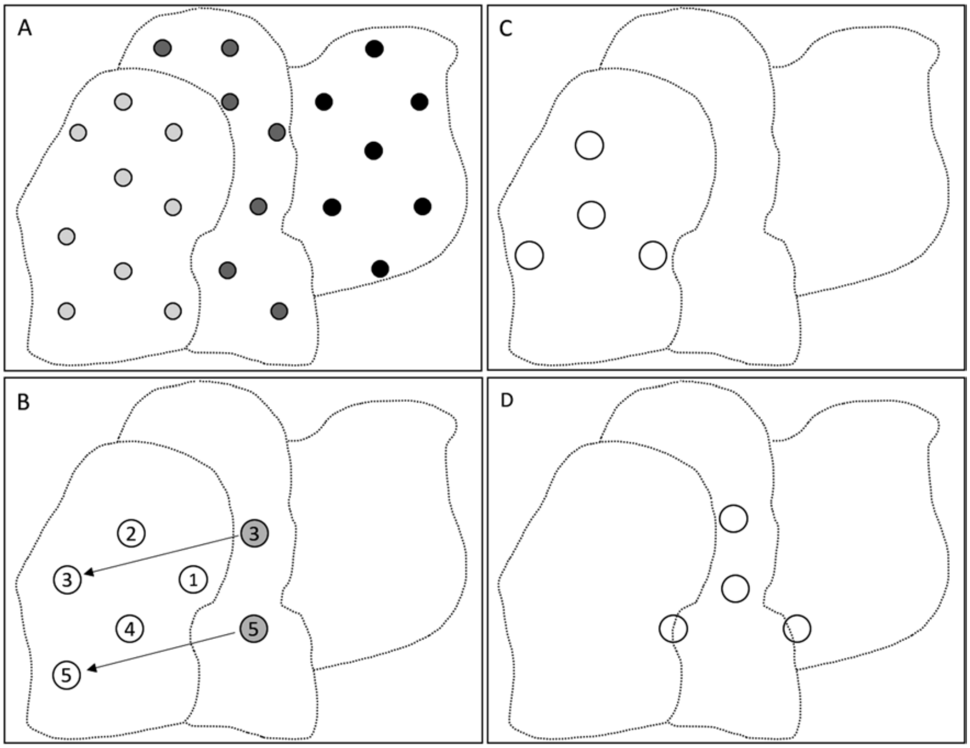

Figure 1. Representation of stands and plots in FIA and FVS. (A) Common forest inventory scenario with three delineated stands sampled by varying numbers of plots. These are consistent with FVS terminology of stands and plots. (B) Representative sample design from FIA periodic inventory, where the plot design consists of five subplots. In this case, the plot footprint, as determined by the pre-established location of plot center (subplot 1), straddles a stand boundary, and subplots 3 and 5 are systematically relocated to sample the stand occupied by plot center. This is effectively the same sampling scenario as (A) for one stand, but with different terminology: the stand is an FIA condition, which can be delineated by ownership or reserved status, the FIA "plot" consists of all five subplots, and the FIA subplots are the equivalent of plots in (A). (C) Current FIA mapped-plot design, consisting of four subplots, where the plot footprint samples a single condition. (D) Same plot design as (C), but with the plot footprint straddling multiple conditions. The trees and variables associated with each condition are apportioned to the whole or partial subplots they occupy.

an FIA plot and what is typically represented in FVS as a stand, and FIA subplots are equivalent to plots in FVS. However, other plot design details preclude the projection of FIA subplots as separate units. In these cases, data are represented on a per-acre basis.

The complicating issue in the case of multicondition plots is that most of the variables that are attributed to stands in other kinds of inventories, such as forest type, site index, or stand age, are recorded at the condition level in FIADB, i.e., the condition is a basic analysis element in FIA data compilation. This situation poses some challenges and limitations for data translation. Because conditions can divide a plot in an infinite number of ways, it is not possible to code the plot design into FVS input. As a result, in any case where the condition proportion of a plot is <1, all trees and seedlings must be converted to a per-acre basis.

Ignoring condition boundaries and translating a multicondition plot as a single unit raises a related issue. If, for example, there are two forested conditions present on the plot, then one condition will possibly have a different site index, forest type, and/or stand age than the other. When coding the data for a whole plot, only one set of these values can be used by FVS. The preferred solution, as expressed by users who would project whole parts regardless of the number of conditions present, was to determine the largest forested condition and use the condition-level values associated with it. Although this obviously represents a loss of available data, this approach may be acceptable if the projection results desired by users are not particularly sensitive to this compromise.

Implications of Design Coding on FVS Projection

Internal FVS growth models can contain absolute and ranked variables (Table 1). Therefore, selection of condition, whole plot, or subplot as the projection unit can influence how trees are grown by FVS. As an example,

Table 1. Components of a large-tree-diameter growth model in FVS.

Growth variable	Independent variable	Description
ln(dds)*=	b_0	Intercept
	$+ b_0 \cdot \ln \text{dbh}$	Log of diameter at breast height at beginning of estimation period
	$+ b_0 \cdot \text{dbh}^2$	Squared diameter at breast height
	$+ b_0 \cdot \ln \text{crwn}$	Log of percent crown ratio
	$+ b_0 \cdot \text{hrel}$	Relative height
	$+ b_0 \cdot \text{SI}$	Site index for the species
	$+ b_0 \cdot \text{plttba}$	Plot basal area
	$+ b_0 \cdot \text{pntbal}$	Plot basal area in trees larger than the subject tree
	$+ b_0 \cdot \tan \text{slp}$	Tangent of slope in degrees
	$+ b_0 \cdot f \cos$	Tangent of slope, cosine of aspect
	$+ b_0 \cdot f \sin$	Tangent of slope, sine of aspect
	$+ b_0 \cdot \text{fortype}$	Categorical variable for forest type group
	$+ b_0 \cdot \text{ecounit}$	Categorical variable for ecological unit group
	$+ b_0 \cdot \text{plant}$	Categorical variable for planted stands

Note: The two variables that are based on rank of the modeled tree—hrel and pntbal—can vary based on user section of projecting by FIA condition, whole plot, or subplot.
*dds = (diameter inside bark at time 0 + periodic diameter growth)^2 – diameter inside bark^2 (Wykoff et al. 1982).

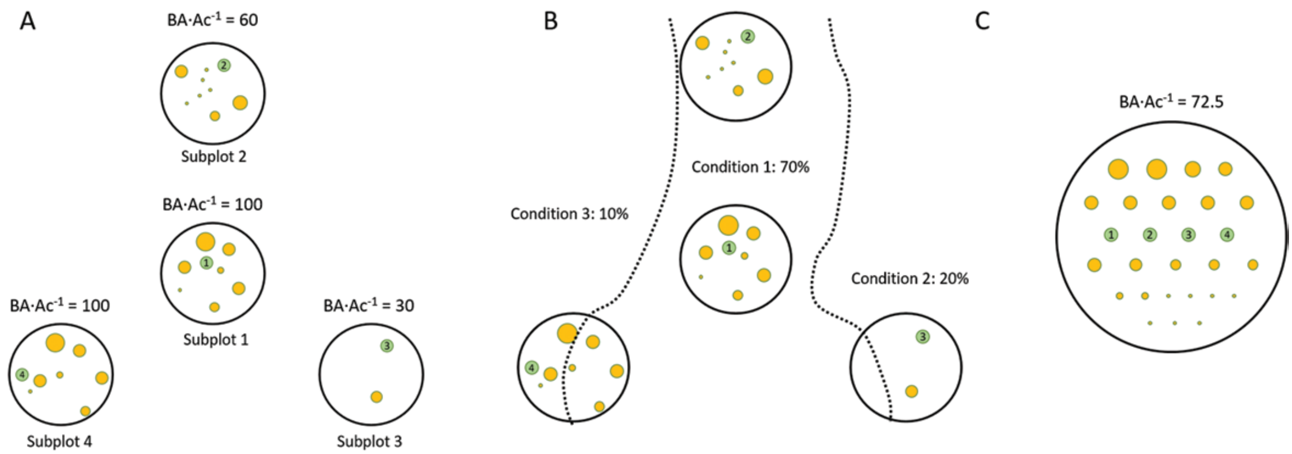

Figure 2. Plot design implications for projection of individual tree growth. The four numbered example trees are identical with respect to species, size, and site quality. The three representations are for (a) FIA annual plot design sampling a single forest condition, (b) the same plot design sampling multiple forest conditions, and (c) any FIA plot design that cannot be coded into FVS in its native form and is therefore coded on a per-acre basis.

consider a case where the current four-subplot design samples a single condition (Figure 2A). This example is mocked up such that the numbered trees on each subplot are identical, and coded as such in the data. By observation, it is apparent that the surrounding stand conditions are not all equal, and it can be generally agreed that the competitive status of each numbered tree is different—trees 1 and 4 occur on the densest plots, but tree 1 experiences more intense competition because of the proximity of other large trees. Their competitive rank in the forest, from highest to lowest competitive status, would be: 1 > 4 > 2 > 3. Therefore, a true-to-life projection should grow each of the numbered trees differently.

If all other factors are held constant, and FVS was insensitive to density, FVS should project identical growth for each. If we consider only local plot basal area and tree rank in this example (Table 1), the situation changes slightly because of the way the data are coded. FVS is often described as quasi-distance-independent, which means that while the subplot level densities come onto play, inter-tree distances within subplots do not. Therefore, whereas the stem mapping in Figure 2A would suggest a different competitive status between trees 1 and 4, FVS uses no data that

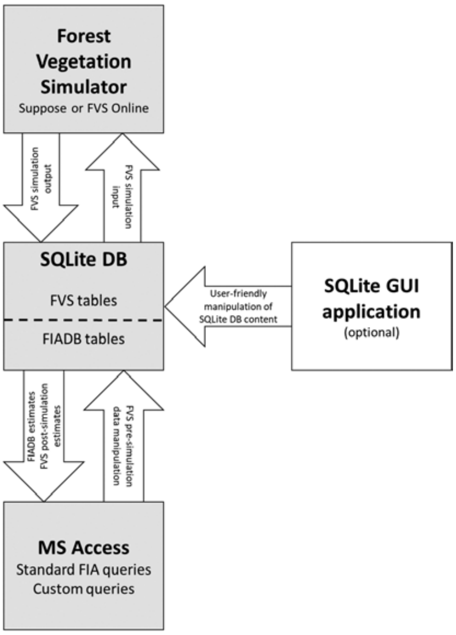

Figure 3. Schematic representation of one possible system of applications for FIA data analysis and FVS simulation. FVS-ready and FIADB tables are stored in the SQLite database, which can also be a repository for FVS output and the results of some queries in Access. Access can contain queries used to produce the FIA core table set and those developed by users, and can store results in Access format. An alternative approach would use a third-party application that is used to read, write, and modify the SQLite database directly.

can be used to make that distinction. As a result, the competitive status ranking within FVS is 1 = 4 > 2 > 3. This applies when projecting whole FIA plots or FIA subplots separately.

If the situation on the ground is that this is actually a multicondition plot (Figure 2B), the apparent competitive relations differ again. When plot data are coded as conditions, FVS does not recognize that trees located in different conditions may actually be located on the same subplot (e.g., subplot containing tree 4). Because FVS cannot preserve the design information of a multicondition plot for each condition at the subplot level, it is necessary to code each condition on a per-acre basis. Trees 1 and 2 are now included in the same growth computations, along with some trees from subplot 4. The factor used when scaling also includes the unoccupied portion of subplot 3. Trees 1 and 2 will now have adjacent rank within the condition, and therefore be grown similarly. Tree 3 will likely still be computed to have the lowest competitive status, but the per-acre density will be higher because of the sliver area of subplot 3 belonging to a different condition. The competitive ranking is now 1 = 2 <> 4 > 3.

Finally, some FIA plot designs used during periodic inventories simply cannot be represented in FVS in a way that preserves design details, such as having different breakpoint diameters for different species, having more than two subplot sizes, or recording different size classes on different numbers of subplots. In these cases, it is only possible to represent the data on a per-acre basis in FVS (Figure 2C), and the differences in competitive status among the four example trees are lost completely.

Whatever the effects of these design differences, they can be propagated through FVS simulations as computations of growth, mortality, volume, and other tree- and stand-level characteristics. The magnitude of these differences and whether they are offsetting or produce bias in large-scale simulations are mostly unknown. However, having the same source data coded in different ways in a single, easily accessible database should facilitate sensitivity analyses.

New Database Structure

The FVS Database Extension (Crookston et al. 2003) uses reserved table and variable names, but additional tables and variables can reside in the database without the need for special accommodation in FVS. We took advantage of this arrangement and the increased capacity of SQLite to deliver FIADB and FVS-ready data together (Figure 3). Suppose and the new FVS Online/Onlocal application have also been updated to work with the multiformat data arrangement. SQLite has no practical limit (current capacity 140 terabytes) for typical users, so large input databases and extensive FVS output can reside in the single database. This provides capabilities to users that have not been convenient, or even possible, before.

A major development consideration was giving the ability to leverage FIADB for the purpose of FVS simulation development. For example, FVS tables do not inherently include information on ownership, but FIADB does. Such information can be considered by the user when setting up simulations over large areas. By default, the FVS Groups variable contains such information as the plot measurement year, forest type, inventory type (e.g., first visit or remeasurement), but the list of categories has been kept to a moderate length in the default translation to avoid an excessive number of selection possibilities. If a user wishes to add a new category to Groups, such as ownership, it is simply a matter of creating a query that assembles the desired variables and rewrites the Groups column for each stand record. Then, in development of the simulation, groups of stands belonging to the desired categories can be selected separately and assigned a unique set of treatments. If there is future user demand for additional Groups variables, the translation process can be modified easily to update the default list.

Although the list of variables added and modified in the FVS-ready tables and FIADB is too extensive to discuss in detail here, the changes generally fall into two categories: (1) those that were added to preserve the linkage between the tables used by FVS and the rest of FIADB; and (2) those that were added to provide user flexibility and better data transparency. Variables fitting the former case are primarily the control numbers that are unique record identifiers within their respective tables. Examples of the latter have various purposes. For example, the site index base age mismatch issue was complex and required extensive treatment. Although both FIA and FVS documentation could be used to determine when FIA site index values were appropriate for use in FVS, our experience was that many users overlooked such details and possibly produced biased stand projections. Our solution was to add FVS-compatible site index variables that paralleled the FIA values and to produce comprehensive documentation (Shaw and Gagnon, in preparation).

Conclusion

FIA2FVS 2.0 will provide better data quality and more user capabilities than previous translation programs and processes. Users will be able to use time more effectively when conducting FVS-related work, because the translation burden has been eliminated, and data quality will be ensured prior to delivery. In addition, user documentation should improve the user experience and promote greater consistency in results produced across the broad FVS user base.

Acknowledgments

We thank D. Jacobs and A. Zharkikh for their project management and programming skills, respectively. This paper was prepared by US Government employees on official time and is therefore in the public domain.

Literature Cited

Bechtold, W.A., and C.T. Scott. 2005. The forest inventory and analysis plot design. P. 27– 42 in *The enhanced forest inventory and analysis program - national sampling design and estimation procedures*, Gen. Tech. Rep. SRS-80, Bechtold, W.A.; Patterson, P.L. (eds.). USDA Forest Service, Southern Research Station, Asheville, NC.

Crookston, N.L., D.L. Gammel, S. Rebain, D. Robinson, C.E. Keyser, and C.A. Dahl. 2003. *Users guide to the database extension of the forest vegetation simulator version 2.0.* Internal Rep. USDA Forest Service, Forest Management Service Center, Fort Collins, CO. 60 p. (revised March 2015).

Dixon, G.E. (comp.). 2002. *Essential FVS: a user's guide to the forest vegetation simulator.* Internal Rep. USDA Forest Service, Forest Management Service Center, Fort Collins, CO. 226 p. (Revised: September 24, 2018).

Hahn, J.T., C.D. Maclean, S.L. Arner, and W.A. Bechtold. 1995. Procedures to handle inventory cluster plots that straddle two or more conditions. *For. Sci.* 41:12–25.

Miles, P.D. 2008. Forest inventory and analysis data for FVS modelers. P. 125–129 in *Third forest vegetation simulator conference; 2007 February 13–15; Fort Collins, CO. Proceedings RMRS-P-54*, Havis, R.N., and N.L. Crookston (comps.). USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Rebain, S.A. (comp.). 2010 (revised March 23, 2015). *The fire and fuels extension to the forest vegetation simulator: updated model documentation.* Internal Rep. USDA Forest Service, Forest Management Service Center, Fort Collins, CO. 403 p.

Shaw, J.D. 2009. Using FIA data in the Forest Vegetation Simulator. P. 16 in *Forest Inventory and Analysis (FIA) Symposium 2008, October 21-23, 2008, Park City, UT. Proc. RMRS-P-56CD*. McWilliams, W., Moisen, G., Czaplewski, R. (comps). USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Shaw, J.D. 2017. Introduction to the special section on forest inventory and analysis. *J. For.* 115:246–248.

Vandendriesche, D.A. 2012. *Advanced FVS tools for landscape planning*. USDA Forest Service, Forest Management Service Center, Fort Collins, CO. Interoffice publication. 191 p.

Wykoff, W.R., N.L. Crookston, and A.R Stage. 1982. *User's guide to the stand prognosis model*. Gen. Tech. Rep. INT-133. USDA Forest Service, Intermountain Forest and Range Experiment Station, Ogden, UT. 112 p.