Minnesota Forest Land Area Estimation Using National Forest Inventory Data

by

Christopher Edgar, Michael Carson, and John Young

Staff Paper Series No. 256

Department of Forest Resources

May 2019

College of Food, Agricultural, and Natural Resource Sciences
University of Minnesota
St. Paul, MN

For more information about the Department of Forest Resources and its teaching, research, and outreach programs, contact the department at:

Department of Forest Resources
University of Minnesota
115 Green Hall
1530 Cleveland Avenue North
St. Paul, MN 55108-6112
Phone: (612) 624-3400
Fax: (612) 625-5212
Email: frweb@umn.edu
Webpage: https://www.forestry.umn.edu/our-department/publications/staff-paper-series

Table of Contents

List of Tables

List of Figures

Minnesota Forest Land Area Estimation Using National Forest Inventory Data

Christopher Edgar, Michael Carson, and John Young[1]

Abstract

The United States conducts a national forest inventory (NFI) for the purpose of providing information on the forest resources of the nation. The United States Forest Service, through the Forest Inventory and Analysis (FIA) program, administers the NFI in cooperation with state forestry agencies. NFI data and information are used by states to support forest management and policy decision-making. Here we review the sample design and estimation procedures of the NFI, with special emphasis on those aspects related to forest land area. Using the 2017 inventory data for the state of Minnesota, we compute an estimate of 17.7 million acres of forest land with a sampling error of 0.52%. Our estimate matches that published by FIA. This report documents the process of obtaining and organizing the data and application of formulae to compute the estimate and associated estimate of sampling error. All data and calculations are assembled in an MS Excel worksheet as a resource for anyone seeking to develop a fuller understanding of this important data set and appropriate estimation procedures[2]. We then extend the estimation to that of ash forest land area, information needed for assessment of potential impacts of the emerald ash borer. For 2017, we estimate that ash (black, green, and white) is a component (at least some live volume) on 4.33 million acres, is 25% or more live volume on 1.78 million acres, and is 50% or more of live volume on 1.10 million acres of forest land in the state.

Keywords: Minnesota, national forest inventory, estimation, area, forestland, sample design.

---

[1] Christopher Edgar is Research Assistant Professor, Michael Carson is Research Forester, and John Young is Research Assistant, Department of Forest Resources, and University of Minnesota. This research was supported by the Department of Forest Resources, the Interagency Information Cooperative, the Minnesota Agricultural Experiment Station Project MIN-42-078, and the USDA NIFA McIntire-Stennis Cooperative Forest Research Program.

2 Contact the lead author at cedgar@umn.edu to obtain links to the MS Excel workbook and other files used in the analysis.

1. Introduction and Objective

Forests are an important part of the natural heritage of Minnesota, contributing to the unique character and quality of life of the state. Forests are not static, but rather are subject to a variety of forces that can alter their extent and condition. As forests change, the socioeconomic and ecological benefits they provide may also change. Regular monitoring of the resource is essential to ensuring that forests are sustained and the benefits derived from them continue into the foreseeable future. The United States (U.S.), like many other countries, conducts a national forest inventory (NFI) for the purpose of monitoring the nation's forests. The NFI is administered by the U.S. Forest Service Forest Inventory and Analysis (FIA) program in cooperation with state forestry agencies. States make extensive use of the information for management and policy decision-making purposes. The state of Minnesota provides additional funding to intensify data collection, a sign of the importance officials place on the inventory and the information produced.

NFI data and information are available in a number of different formats. State reports have been a mainstay since the earliest years of the inventory. New state estimates are published in the form of annual online resource updates (e.g. Miles et al., 2017). More detailed analysis and assessment of a state's forests conditions and trends is conducted and reported every five years (e.g. Miles et al., 2016). New formats for reporting that can be placed in the category of digital engagement have been and continue to be developed (e.g. U.S. Forest Service, 2019a). Online analytical tools (e.g. EVALIDator, DATIM) provide users the ability to produce custom estimates through a convenient user interface (Vogt and Smith, 2017). The capabilities of these tools has increased in recent years and complex estimation and analyses are possible. For those seeking to perform their own estimation and analyses, the FIA Database (FIADB) including tree, condition, plot, strata, and population data records is publicly available (U.S. Forest Service, 2019c).

This study is motivated by our interest in understanding the estimation procedures used in the NFI. We set out to reproduce a published estimate using FIADB and the estimation procedures documented in several technical FIA reports. Measures of volume and growth are historically associated with forest inventory (Spurr, 1952; Avery and Burkhart, 2002). In broad-scale forest inventories, such as an NFI, forest land may not be delineated and thus its estimation is an

objective of the inventory. We were also interested in knowing the area of forest land susceptible to emerald ash borer. Our initial review of existing reports suggested that an answer was not straightforward and thus we saw this as an application worth exploring.

We reviewed the sample design and estimation procedures with a focus on information relevant and sufficient for area estimation. Primary references on sample design and estimation procedures were Bechtold and Patterson (2005) and a MS Excel workbook (U.S. Forest Service, 2019b) obtained from the U.S. Forest Service website. The workbook contained a number of estimation examples using a small, artificial data set. Pugh et al. (2019) contains a number of estimation examples using FIADB with structured query language (SQL). We desired an estimation example using actual FIA data in MS Excel, believing this would be an instructive format accessible to most forestry students and foresters having had forest measurements instruction at the level of Avery and Burkhart (2002) or Husch et al. (2002). In this report we separated the data assembly step from the estimation step. We used simple SQL statements sufficient to assemble the necessary data from FIADB and then conducted the computation of the estimate and its standard error in MS Excel. MS Excel workbook and data files can be obtained through an email to the lead author (cedgar@umn.edu).

2. Sample Design and Estimation

Knowledge of the process that generates sample data is essential in estimation. In this section we provide an overview of the sample design of the NFI with information specific to forest land area in Minnesota. The information presented is not exhaustive and users are directed to the primary reference (Bechtold and Patterson, 2005) on FIA sample design and estimation procedures for more detailed information. A second reference, intermediate in detail between what is presented in this report and that of Bechtold and Patterson (2005), is section 1.2 of the FIADB user manual (Burrill et al. 2018).

At the core of FIA is a network of permanent plots linked to a systematic-unaligned grid having a base spatial intensity of approximately one plot per 6,000 ac (Reams et al., 2005). A number of states contribute funding to increase the spatial intensity of the plot network. Minnesota's contribution doubles the spatial intensity to approximately one plot per 3,000 ac and there are more than 18,000 plots (base + intensified) across Minnesota's 54 million ac. In Minnesota the

plot network is systematically divided into five panels. One panel is measured each year and the measurement cycle, the length it takes to measure all the plots, is thus 5 years (Reams et al., 2005).

The FIA plot is a cluster of four points with one point in the center and the other three points 120 ft distance at azimuths of 0, 120, and 240 degrees from the center point (Figure 1). Land conditions observed on four subplots are the basis for area estimation. (The larger macroplot is not used in Minnesota.) Subplots are formed by extending a 24.0 ft radius around each point. The total area of all four subplots is 7,238 ft^2 or approximately $1/6^{th}$ ac. Forest conditions are identified and mapped on each subplot (Bechtold and Scott, 2005). FIA defines forest land as an area having, or had, at least 10% canopy cover of live trees, being at least 1 ac in size, and at least 120 ft wide (Burrill et al. 2018).

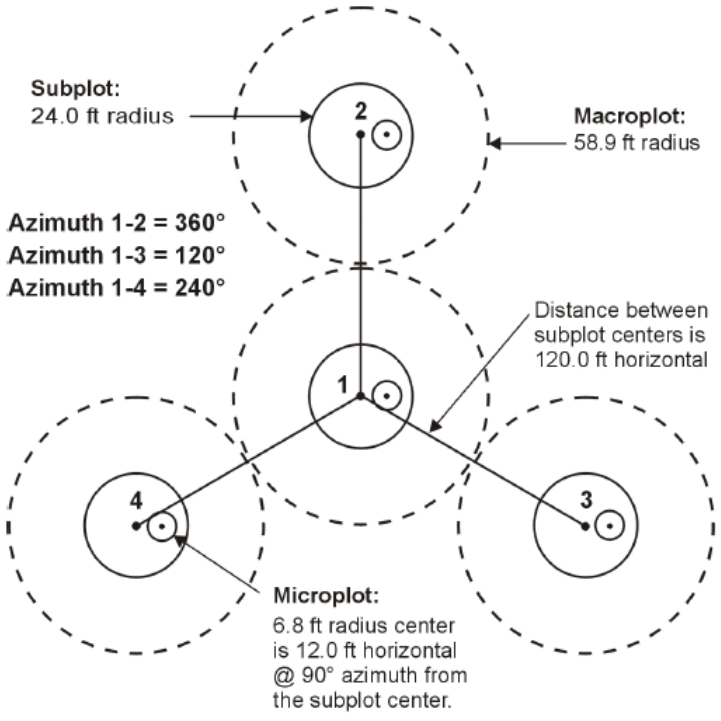

Figure 1. Design of the FIA plot (Source: Burrill et al., 2018).

FIA uses a design-based, post-stratified approach to estimation that is documented in Scott et al. (2005). The term estimation unit is used by FIA to denote the area where estimates of population parameters are desired. Post-stratification is used for the purpose of variance reduction. An estimation unit is stratified into non-overlapping subdivisions that together fill the estimation unit. Early in the annual inventory, strata in Minnesota were based on land cover classes (forest/nonforest, edge/interior). In recent years strata have been based on tree canopy cover from the National Land Cover Database.

In producing population estimates, FIA combines all the panels of data. Using this approach, sample sizes are increased and variance reduction is achieved (Scott et al., 2005), albeit at the cost of reduced temporal resolution of the estimate (e.g. Edgar et al., 2019). The specific approach used by FIA to combine panels is known as the temporally indifferent method, which entails combining all panels into the equivalent of a large periodic survey with the same stratification applied (Patterson and Reams, 2005).

For area estimation, the plot observation is the proportion of the plot in the domain of interest. Each plot is assigned to one and only one stratum. Strata means and variances are calculated. Population (i.e. estimation unit) total is estimated. Statewide estimates are computed by adding the estimates and variances of all estimation units in the state (Scott et al. 2005).

3. Data

The inventory data are available in a large, relational database referred to as FIA Database (FIADB). Many of the measurements or observations made on trees (e.g. height), conditions (e.g. ownership), and plots (e.g. approximate latitude and longitude) plus values derived or computed from measurements (e.g. tree volume, site quality class) are in FIADB. Information to identify populations, panels, and stratifications needed for estimation are also included in FIADB.

The US Forest Service makes FIADB publicly-available at the FIA DataMart (US Forest Service 2019c). When working with FIA it is helpful to note when data were loaded to the DataMart. FIA data are regularly updated and knowing the date establishes the version of the data, which is important when data or estimates are compared. At the time of this analysis, the most recent

Minnesota data were loaded to the DataMart on 06-18-2018. Eight files of Minnesota data were downloaded (Table 1).

Table 1. FIADB files used in the analysis.

File Name	Records (Number)	Size[1] (KB)
MN_POP_EVAL_GRP	17	4
MN_POP_EVAL_TYP	122	13
MN_POP_EVAL	64	15
MN_POP_ESTN_UNIT	1,807	310
MN_POP_STRATUM	8470	1,336
MN_POP_PLOT_STRATUM_ASSGN	1,053,904	121,217
MN_PLOT	147,532	35,258
MN_COND	159,699	51,403

[1]File sizes may vary between computers.

The first 6 tables are population tables that provide information about the population evaluation group, population evaluation type, population evaluation, population estimation units, population strata, and the population assignment of plots to strata, respectively. The last 2 tables are records of plots and conditions, respectively. A schematic of the tables and their relationships is presented in Figure 2. All of the tables in FIADB are fully documented in the FIADB User Manual (Burrill et al., 2018) and the reader is referred to that publication if more information is desired.

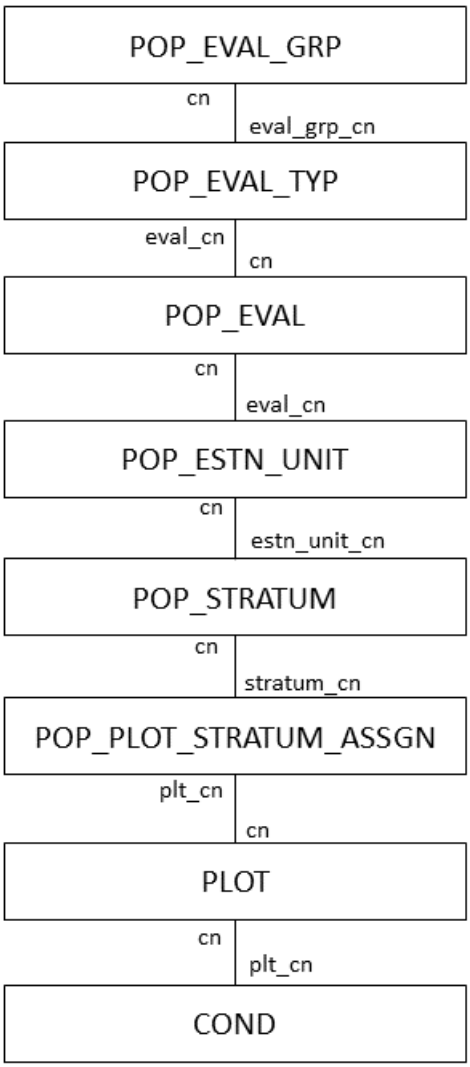

Figure 2. FIADB tables used in area estimation and their relationships. Boxes represent tables in FIADB and lines are used to demonstrate relationships with field names upon which joins are made displayed outside the boxes.

Data access is facilitated through use of a database program. Data files are provided in comma separated value format and can be imported into different programs. Selection of a specific program is primarily a matter of personal choice. The database was provided in MS Access until recently when database size exceeded MS Access limits. For many applications only part of FIADB is needed and therefore MS Access may still be a suitable option. FIA is currently providing online resources to assist users in building Postgres versions of FIADB. For this report SAS was utilized.

3.1 Evaluation

Central to how FIA organizes plots and stratifications for population estimation is the *evaluation identifier* (or *evalid*). *Evaluation identifier* is a 6-digit code that follows the pattern of 2-digit state code, 2-digit year code, and 2-digit evaluation type. Rather than specifying the *evaluation identifier* directly, we have found it easier to specify the *evaluation group* (or *eval_grp*) and *evaluation type* (or eval_typ*)*, which together define the *evaluation identifier*. Evaluation group is a 6-digit code with the pattern 2-digit state code and 4-digit year. The state code for Minnesota is 27 and thus the *evaluation group* for estimation of Minnesota 2017 conditions is 272017. Plots may be appropriate for one type of estimate but not another. For example, change estimates (e.g. growth, removals, mortality) are computed from plots measured on two occasions. Certain measurements may only be made on certain plots, such as down woody material. It is sufficient for our purposes to note that EXPCURR is the evaluation type used to produce current area estimates. We refer the reader to the FIADB documentation (Burrill et al., 2018) for a listing and description of the available evaluation types and for what attributes they apply.

The *evaluation identifier* can be determined through a select query of the population evaluation group, population evaluation type, and population evaluation tables.

```
select
      peg.eval_grp,
      pet.eval_typ,
      pev.evalid,
      pev.eval_descr
from
      fs_fiadb.pop_eval_grp   as peg,
      fs_fiadb.pop_eval_typ   as pet,
      fs_fiadb.pop_eval       as pev
where
      peg.cn=pet.eval_grp_cn
  and pet.eval_cn=pev.cn
  and peg.eval_grp=272017
  and pet.eval_typ='EXPCURR';
```

The *evaluation identifier* for *evaluation group* 272017 and *evaluation type* EXPCURR is 271701 (Table 2). *Evaluation identifier* 271701 is constructed from the 2-digit state code 27, 2-digit year code 17, and 2-digit evaluation type 01. *Evaluation description (or eval_descr)* confirms the evaluation is Minnesota 2017. The description notes that the plots were measured 2013-2017,

reflecting the fact that Minnesota is on a 5-year inventory cycle. The final information in the evaluation description confirms the evaluation is appropriate for current area estimates.

Table 2. Evaluation identifier for Minnesota 2017 current area and current volume estimates.

eval_grp	eval_typ	evalid	eval_descr
272017	EXPCURR	271701	MINNESOTA 2017: 2013-2017: CURRENT AREA, CURRENT VOLUME

3.2 Estimation Units

A select query of the population evaluation group, population evaluation type, population evaluation, and population estimation unit tables provides the estimation unit information for the evaluation.

```
select
      peu.estn_unit,
      peu.estn_unit_descr,
      peu.area_used,
      peu.p1pntcnt_eu
from
      fs_fiadb.pop_eval_grp   as peg,
      fs_fiadb.pop_eval_typ   as pet,
      fs_fiadb.pop_eval       as pev,
      fs_fiadb.pop_estn_unit  as peu
where
      peg.cn=pet.eval_grp_cn
  and pet.eval_cn=pev.cn
  and pev.cn=peu.eval_cn
  and peg.eval_grp=272017
  and pet.eval_typ='EXPCURR'
order by
      peu.estn_unit;
```

There are 16 estimation units in the evaluation (Table 3). *Area used to calculate all expansion factors* (*area_used*) and *phase 1 point count for estimation unit* (*p1pntcnt_eu*) are provided for each estimation unit. Estimation units 1, 2, 15 and 16 correspond to large federal land holdings (i.e. Boundary Waters Canoe Area, Chippewa National Forest, Superior National Forest, and Voyageurs National Park). Additional estimation units were constructed from FIA survey units (1 = Aspen-Birch unit, 2 = Northern Pine unit, 3 = Central Hardwood unit, and 4 = Prairie unit)

with additional subdivision by census water status and major ownership group. Plot distribution for each estimation unit is displayed in Figure 3.

Table 3. Estimation units for the evaluation.

estn_unit	estn_unit_descr	area_used	p1pntcnt_eu
1	BWCA	766334.5	3445829
2	Chippewa NF	641491.1	2884470
3	Inland Census Water Unit 1	599085	2693790
4	Inland Census Water Unit 2	1485605	6680035
5	Inland Census Water Unit 3	625405.2	2812139
6	Inland Census Water Unit 4	350376.5	1575470
7	Private Unit 1	3002133	13499114
8	Private Unit 2	6458263	29039625
9	Private Unit 3	11160398	50182807
10	Private Unit 4	18599019	83630619
11	Public Unit 1	3573391	16067777
12	Public Unit 2	3960072	17806490
13	Public Unit 3	767144.1	3449469
14	Public Unit 4	570994.8	2567485
15	Superior NF	1328214	5972322
16	Voyagers NP	121087.5	544471

Figure 3. Approximate locations of plots in the 16 estimation units used in the evaluation. The four survey unit boundaries are displayed. Aspen-Birch unit covers northeast Minnesota. Northern Pine unit covers north central Minnesota. Central Hardwood unit cover central and southeast Minnesota. Prairie unit covers the western and south central parts of the state.

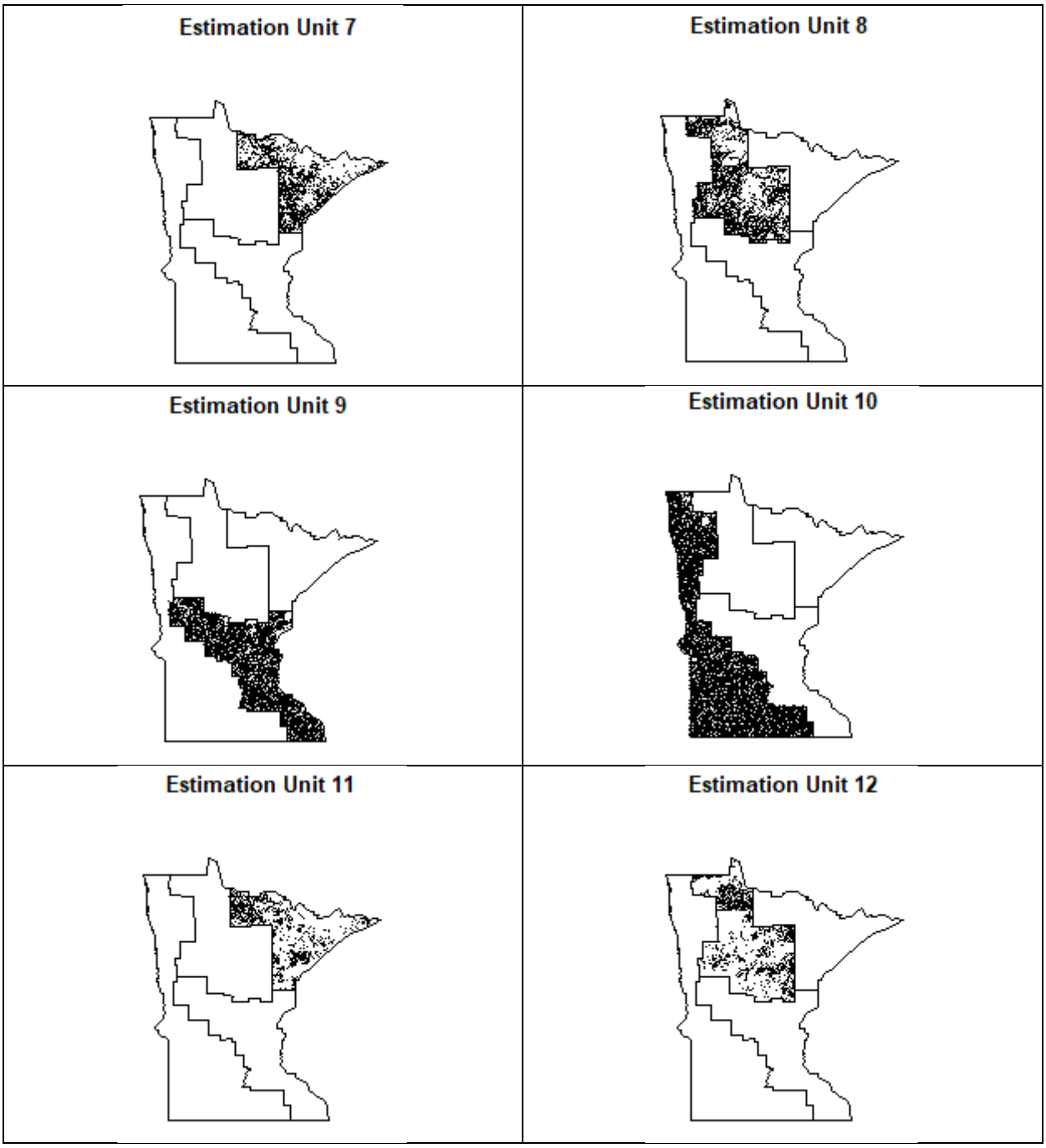

Figure 3 (continued). Approximate locations of plots in the 16 estimation units used in the evaluation. The four survey unit boundaries are displayed. Aspen-Birch unit covers northeast Minnesota. Northern Pine unit covers north central Minnesota. Central Hardwood unit cover central and southeast Minnesota. Prairie unit covers the western and south central parts of the state.

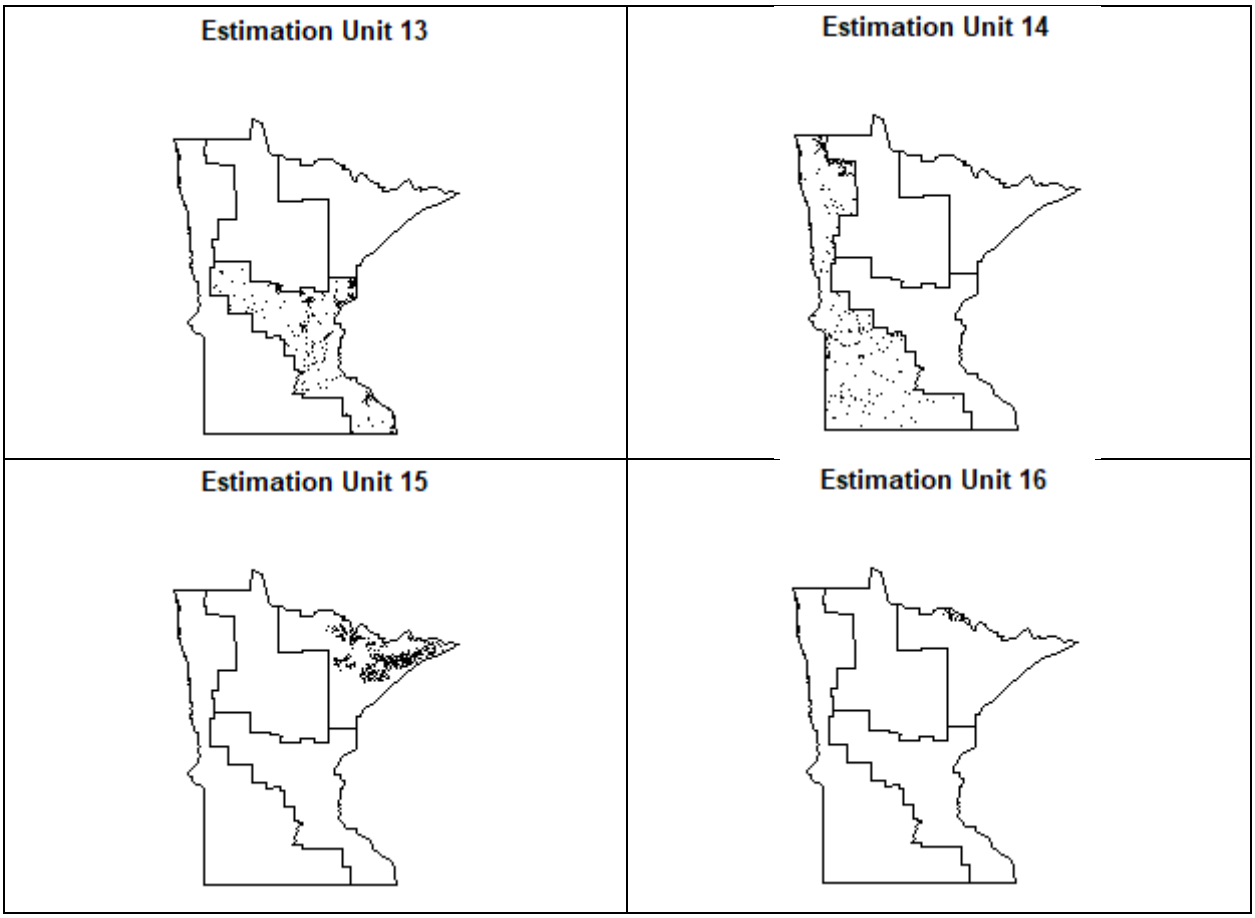

Figure 3 (continued). Approximate locations of plots in the 16 estimation units used in the evaluation. The four survey unit boundaries are displayed. Aspen-Birch unit covers northeast Minnesota. Northern Pine unit covers north central Minnesota. Central Hardwood unit cover central and southeast Minnesota. Prairie unit covers the western and south central parts of the state.

3.3 Strata

A select query of the population evaluation group, population evaluation type, population evaluation, population estimation unit, and population stratum tables provides the strata information for the evaluation.

```
select
      peu.estn_unit,
      psm.stratumcd,
      psm.stratum_descr,
      psm.p1pointcnt,
      psm.p2pointcnt,
      psm.adj_factor_subp
from
      fs_fiadb.pop_eval_grp         as peg,
      fs_fiadb.pop_eval_typ         as pet,
```

```
      fs_fiadb.pop_eval             as pev,
      fs_fiadb.pop_estn_unit        as peu,
      fs_fiadb.pop_stratum          as psm
where
      peg.cn=pet.eval_grp_cn
  and pet.eval_cn=pev.cn
  and pev.cn=peu.eval_cn
  and peu.cn=psm.estn_unit_cn
  and peg.eval_grp=272017
  and pet.eval_typ='EXPCURR'
order by
      peu.estn_unit,
      psm.stratumcd;
```

The query retrieves the list of strata with *estimation unit* (*estn_unit*), *stratum code* (*stratumcd*), *stratum description* (*stratum_descr*), *phase 1 point count* (*p1pointcnt*), *phase 2 point count* (*p2pointcnt*), and *adjustment factor for the subplot* (*adj_factor_subplot*). There are 50 strata in the evaluation. The first 10 records are displayed in Table 4. The number of strata per estimation unit ranges between 1 and 5 with a mode of 4. In this evaluation, canopy cover classes formed from 2011 National Land Cover Data tree canopy cover data are used for strata. Stratum codes 1, 2, 3, 4, and 5 correspond to canopy cover of 0-5%, 6-50%, 51-65%, 66-80%, and 81-100%, respectively. For example, *estimation unit* 1 has 3 strata. The first two canopy cover classes, 0-5% and 6-50%, form separate strata. Canopy cover classes 3, 4, and 5 are collapsed to form one stratum given, which was given the code 345. Canopy cover classes are collapsed to in order to meet minimum number of plots per strata requirements (Westfall et al., 2011).

Table 4. First 10 stratum (of 50 total) for the evaluation.

estn_unit	stratumcd	stratum_descr	p1pointcnt	p2pointcnt	adj_factor_subp
1	1	Canopy cover 0 - 5	942687	41	1.014993
1	2	Canopy cover 6 - 50	1377288	54	1.005651
1	345	Canopy cover 51 - 100	1125854	39	1
2	1	Canopy cover 0 - 5	539682	35	1
2	2	Canopy cover 6 - 50	1298382	97	1
2	34	Canopy cover 51 - 80	1046406	73	1
3	1	Canopy cover 0 - 5	2547056	179	1.001399
3	234	Canopy cover 6 - 80	146734	11	1
4	12345	Canopy cover 0 - 100	6680035	493	1
5	12345	Canopy cover 0 - 100	2812139	212	1.001674

Phase 1 point count is used to determine stratum weight. *Phase 2 point count* is the number of field plots in the stratum. *Adjustment factor for the subplot* is an adjustment that will be applied in the estimation to account for partially nonsampled plots (see section 4.3.1.2 of Scott et al., 2005).

3.4 Plots

A select query of the population evaluation group, population evaluation type, population evaluation, population estimation unit, population stratum, population plot stratum assignment, and plot tables provides the plot information for the evaluation.

```
select
      p.cn                                  as plt_cn,
      p.invyr,
      p.statecd,
      p.unitcd,
      p.countycd,
      p.plot,
      psm.estn_unit,
      psm.stratumcd
from
      fs_fiadb.pop_eval_grp                 as peg,
      fs_fiadb.pop_eval_typ                 as pet,
      fs_fiadb.pop_eval                     as pev,
      fs_fiadb.pop_estn_unit                as peu,
```

```
      fs_fiadb.pop_stratum                as psm,
      fs_fiadb.pop_plot_stratum_assgn     as ppsa,
      fs_fiadb.plot                       as p
where
      peg.cn=pet.eval_grp_cn
  and pet.eval_cn=pev.cn
  and pev.cn=peu.eval_cn
  and peu.cn=psm.estn_unit_cn
  and psm.cn=ppsa.stratum_cn
  and ppsa.plt_cn=p.cn
  and peg.eval_grp=272017
  and pet.eval_typ='EXPCURR'
order by
      p.cn;
```

The evaluation includes 17,676 plots. The query returns the p*lot sequence number* (*plt_cn*), *inventory year* (*invyr*), *state code* (*statecd*), *survey unit code* (*unitcd*), *county code* (*countycd*), *plot number* (*plot*), *estimation unit* (*estn_unit*), and *stratum code* (*stratumcd*) (Table 5). Plot sequence number is a unique sequence number that identifies the plot record. As previously noted, one panel of plots is measured every year. Inventory year denotes the year when most of the plots in the panel were measured. (The actual date of measurement can be determined from MEASYEAR, MEASMON, and MEASDAY fields in the plot table.) Names of counties and survey units can be found in the FIADB user manual (Burrill et al., 2018).

Table 5. First 15 plots (of 17,676 total) for the evaluation.

plt_cn	invyr	statecd	unitcd	countycd	plot	estn_unit	stratumcd
15738820020004	2013	27	1	71	22247	11	2
15738821020004	2013	27	2	61	29172	8	3
15738822020004	2013	27	2	1	20037	12	3
15738823020004	2013	27	3	141	20125	5	12345
15738824020004	2013	27	3	25	20186	9	1
15738825020004	2013	27	3	163	20053	9	2
15738826020004	2013	27	1	17	20365	7	2
15738827020004	2013	27	3	25	20069	9	1
15738828020004	2013	27	4	47	20219	10	1
15738829020004	2013	27	4	47	20114	10	1
15738830020004	2013	27	4	47	20025	10	1
15738831020004	2013	27	2	29	20140	8	2
15738832020004	2013	27	2	29	20358	8	3
15738833020004	2013	27	4	51	20100	10	1
15738834020004	2013	27	4	151	20211	10	1

Noted in the section Sample Design and Estimation is the FIA approach of combining all the panels of data when producing estimates. In Minnesota, where the inventory cycle is 5 years, evaluations will consist of plots spanning 5 consecutive inventory years ending with the year that was used in forming the evaluation group (i.e. 2017). Plots will be approximately evenly distributed across the 5 inventory years. A barplot of plot counts by inventory year for *evaluation identifier* 271701 illustrate this pattern (Figure 4).

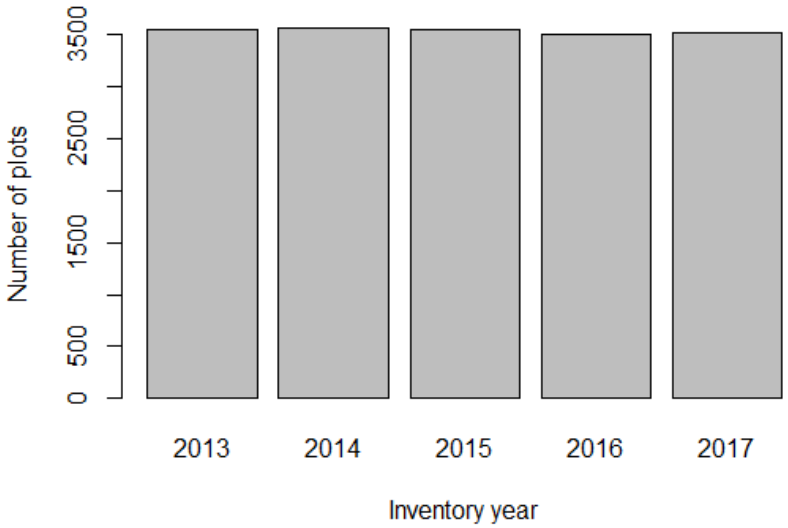

Figure 4. Counts by inventory year of the 17,676 plots in the evaluation.

3.5 Conditions

A select query of the population evaluation group, population evaluation type, population evaluation, population estimation unit, population stratum, population plot stratum assignment, plot, and condition tables provides the condition information for the evaluation.

```
select
      c.cn                                  as cnd_cn,
      c.plt_cn,
      c.invyr,
      c.statecd,
      c.unitcd,
      c.countycd,
      c.plot,
      c.condid,
      c.cond_status_cd,
      c.condprop_unadj
from
      fs_fiadb.pop_eval_grp                 as peg,
      fs_fiadb.pop_eval_typ                 as pet,
      fs_fiadb.pop_eval                     as pev,
      fs_fiadb.pop_estn_unit                as peu,
      fs_fiadb.pop_stratum                  as psm,
      fs_fiadb.pop_plot_stratum_assgn       as ppsa,
      fs_fiadb.plot                         as p,
      fs_fiadb.cond                         as c
where
```

```
      peg.cn=pet.eval_grp_cn
  and pet.eval_cn=pev.cn
  and pev.cn=peu.eval_cn
  and peu.cn=psm.estn_unit_cn
  and psm.cn=ppsa.stratum_cn
  and ppsa.plt_cn=p.cn
  and p.cn=c.plt_cn
  and peg.eval_grp=272017
  and pet.eval_typ='EXPCURR'
order by
      c.cn;
```

There are 21,488 records in the evaluation. For each condition the query retrieves the *condition sequence number* (*cnd_cn*), *plot sequence number* (*plt_cn*), *inventory year* (*invyr*), *state code* (*statecd*), *survey unit code* (*unitcd*), *survey unit code* (*unitcd*), *county code* (*countycd*), *plot number* (*plot*), *condition class number* (*condid*), *condition status code* (*cond_status_cd*), and *condition proportion unadjusted* (*condprop_unadj*) (Table 6).

As previously noted, conditions on FIA plots are identified and mapped. Differences in *condition status code* (i.e. accessible forest land, nonforest land, noncensus water, census water, and nonsampled with possibility of forest land) will trigger a change in *condition class number*. On accessible forest land, differences in reserved status, owner group, forest type, stand-size class, regeneration status, and stand density trigger changes in *condition class number*. The proportion of the plot in the condition is recorded in *condition proportion unadjusted* (Burrill et al., 2018). The first 10 conditions in the evaluation are displayed in Table 6.

Table 6. First 10 conditions (of 21,448 total) for the evaluation.

cnd_cn	plt_cn	invyr	statecd	unitcd	countycd	plot	condid	cond_status_cd	condprop_unadj
234663901020004	15738820020004	2013	27	1	71	22247	1	1	0.75
234663902020004	15738820020004	2013	27	1	71	22247	2	1	0.25
234663968020004	15738821020004	2013	27	2	61	29172	1	1	1
234664025020004	15738822020004	2013	27	2	1	20037	1	1	1
234664128020004	15738831020004	2013	27	2	29	20140	1	2	0.477289
234664129020004	15738831020004	2013	27	2	29	20140	2	1	0.522711
234664163020004	15738832020004	2013	27	2	29	20358	1	1	1
234664252020004	15738836020004	2013	27	1	75	29138	1	1	0.75
234664253020004	15738836020004	2013	27	1	75	29138	2	1	0.25
234664290020004	15738841020004	2013	27	2	21	29166	1	1	0.965516

3.6 Excel Workbook

All of the records from the five queries were assembled into one MS Excel workbook named FIA-Data.xlsx. Query records were saved in individual worksheets (Table 7). All of the data used in producing an estimate of forest land area for Minnesota 2017 are contained in the Excel workbook.

Table 7. Evaluation data used to compute area estimate in FIA-Data.xlsx workbook.

Worksheet Name	Number of Records[1]	Number of Columns
Evaluation	1	4
Estimation unit	16	4
Stratum	50	6
Plot	17,676	8
Condition	21,488	10

[1] Number of rows in the worksheet is the number of records + 1 (header)

4. Estimation

4.1 Plot Attribute of Interest

In area estimation, the attribute of interest is the proportion of the plot in the domain of interest (equation 4.1 in Scott et al., 2005).

$$P_{hid} = \frac{\sum_{j}^{4} \sum_{k}^{K_{hij}} a_{mhijk} \delta_{hijkd}}{a_m \bar{p}_{mh}}$$

where

P_{hid} = proportion of plot *i* in the domain of interest *d*, for plots assigned to stratum *h*, adjusted for stratum *h* plots that overlap the population boundary.

a_{mhijk} = mapped area (acres) of subplot *j* covering condition *k* on plot *i* assigned to stratum *h*.

δ_{hijkd} = zero-one domain indicator function, which is 1 if condition *k* on subplot *j* of plot *i* assigned to stratum *h* belongs to the domain of interest *d.*

K_{hij} = the number of conditions that exist on subplot *j* of plot *i* assigned to stratum *h*

a_m = total area of the subplots

$\bar{p}_{mh}$ = mean proportion of stratum h mapped plot areas falling within the population

FIADB provides several calculated values that facilitate the calculation of the plot attribute of interest. The proportion of the plot in the condition is provided in the field *condition proportion unadjusted* (Condition worksheet) and thus it is not necessary to work with areas (a_{mhijk}, a_m) directly.

Every condition is evaluated to determine if the condition is in the domain of interest. A column (column K, label 'domain-indicator') has been added to the Condition worksheet. We desire an estimate of forest land area and thus the domain of interest in this example is forest land. Basic land classification information is recorded in *condition status code* (Condition worksheet) and a code value of 1 denotes accessible forest land. The aforementioned domain indicator (δ_{hijkd}) takes a value of 1 when condition status code value is 1. The domain indicator takes a value of 0 when the condition status code value is not 1.

A plot value is derived by summing *condition proportion unadjusted* across all conditions on the plot that are in the domain of interest. We demonstrate the approach for three plots:

The summed value for *plot sequence number* 157388<u>32</u>020004 (*unitcd* 2, *countycd* 29, *plot* 20358) is 1. There is one condition on the plot and it is accessible forest land with condprop_unadj value of 1.

The summed value for *plot sequence number* 157388<u>20</u>020004 (*unitcd 1, countycd* 71, *plot* 22247) is 1. There are two conditions on the plot and both are accessible forest land. The sum of c*ondition proportion unadjusted* values for the two conditions is 1 (i.e. 0.75 + 0.25).

The summed value for *plot sequence number* 157388<u>31</u>020004 (*unitcd* 2, *countycd* 29, *plot* 20140) is 0.522711. There are two conditions on the plot. The first condition is not accessible forest land and therefore it is not in the domain of interest. The second condition is accessible forest land and is in the domain of interest. The sum of c*ondition proportion unadjusted* values for the conditions in the domain of interest is 0.522711.

The final step in computing the attribute of interest for the plot is application of an adjustment for partially sampled plots (see section 4.3.1.2 of Scott et al., 2005). In FIADB the field *adjustment factor for the subplot* (Stratum worksheet) is the reciprocal of $\bar{p}_{mh}$in the formula above. Application of the *adjustment factor for the subplot* in computing the plot attribute of interest is demonstrated in Table 8.

Table 8. Adjustment factor application in plot attribute of interest.

Plot Sequence Number	Survey Unit Code	County Code	Plot Number	Sum of Condition Proportion Unadjusted for Conditions in the Domain of Interest	Adjustment Factor for the Subplot	Plot Attribute of Interest P_{hid}
157388<u>32</u>020004	2	29	20358	1	1.0231648795	1.0231648795
157388<u>20</u>020004	2	71	22247	1	1	1
157388<u>31</u>020004	2	29	20140	0.522711	1.0191526487	0.5327223002

The mean proportion of the plot in the domain of interest adjusted for partially nonsampled plots (P_{hid}) is reported for every plot in Plot worksheet.

4.2 Strata Means and Variances

Having computed the attribute of interest for every plot, we can now proceed with computing stratum mean and variance. The stratum mean is the sum of the plot values (P_{hid}) divided by the number of plots in the stratum. Scott et al. (2005) provides the following formula:

$$\overline{P_{hd}} = \frac{\sum_i^{n_h} P_{hid}}{n_h}$$

where

n_h is the number of plots in stratum *h*

and other variables as previously defined.

Scott et al. (2005) provides the following formula for the variance estimator:

$$v(\overline{P_{hd}}) = \frac{\sum_i^{n_h} P_{hid}^2 - n_h \overline{P_{hd}}^2}{n_h(n_h - 1)}$$

where

all variables as previously defined.

For an example, consider estimation unit 7 stratum 3. There are 248 plots (n_h) in the stratum (Stratum worksheet). The sum of the plot values ($\sum_i^{n_h} P_{hid}$) was 242.770 and the sum of the plot values squared ($\sum_i^{n_h} P_{hid}^2$) was 0.979 (Estimation worksheet). (Computation of the latter involves first squaring each individual plot value and then summing up the squared values. The column $P_{hid} \times P_{hid}$ was added to the Plot worksheet to facilitate the computation.)

$$\overline{P_{hd}} = \frac{242.770}{248} = 0.979$$

$$v(\overline{P_{hd}}) = \frac{240.553 - 248 \times 0.979^2}{248 \times (248 - 1)} = 0.000047$$

Strata weights are calculated as the *phase 1 point count* (Stratum worksheet) divided by the *phase 1 point count for estimation unit* (Estimation-Unit worksheet). Continuing with the example of estimation unit 7 stratum 3, *phase 1 point count* is 2,521,007 pixels and *phase 1 point count for estimation unit* is 13,499,114 pixels. The weight of stratum 3 is then computed as 2,521,007 divided by 13,499,114, which equals 0.187. Strata weights, means, and variances are reported for all 4 strata of estimation unit 7 in Table 9.

Table 9. Stratum statistics for estimation unit 7.

Stratum	Stratum Description	W_h	n_h	$\overline{P_{hd}}$	$v(\overline{P_{hd}})$
1	Canopy cover 0 - 5	0.4146890	564	0.528004	0.000381
2	Canopy cover 6 - 50	0.3614859	416	0.880266	0.000148
3	Canopy cover 51 - 65	0.1867535	248	0.978910	0.000047
45	Canopy cover 66 - 100	0.0370716	47	1.000000	0.000263

Estimation unit 7 covers more than 3.0 million acres (Estimation-Unit worksheet). There are 1,275 plots (Table 9) in the estimation unit. Stratum 1 is the largest stratum, accounting for approximately 41% of the estimation unit area and 44% of the plots. Stratum 45 is the smallest stratum, accounting for just under 4% of the area and 4% of the plots. Strata means do vary in a manner consistent with the canopy cover classification. The lowest mean proportion of plots in forest land was observed in the stratum created using the lowest canopy cover percentages (i.e. 0-5). Strata created from higher canopy percentages had higher mean proportions in forest land.

Strata weight, mean, and variance are reported by estimation unit for all 50 strata in the worksheet Estimate. With the strata information computed, we can proceed with estimation of population total and variance.

4.3 Estimation Units Totals and Variances

In FIA, estimation units are the populations for which estimates of parameters are computed. Estimation unit total is computed with the following formula from Scott et al. (2005):

$$\hat{A}_d = A_T \sum_h^H W_h \overline{P_{hd}}$$

where

A_T = total area in the population (i.e. estimation unit) in acres

W_h = weight for stratum *h* (i.e. proportion of the population area, A_T, that is in stratum *h)*

and other variables as previously defined.

An estimate of the estimation unit variance is computed using the following formula of Scott et al. (2005):

$$v(\hat{A}_d) = \frac{{A_T}^2}{n}\left[\sum_h^H W_h n_h v(\overline{P_{hd}}) + \sum_h^H (1 - W_h)\frac{n_h}{n} v(\overline{P_{hd}})\right]$$

where

n is the total number of plots in the population (i.e. estimation unit)

and other variables as previously defined.

Calculation of the estimate is demonstrated with estimation unit 7. Total area in the estimation unit (A_T) is 3,002,133 acres (Estimation-Unit worksheet). Stratum mean and weight are reported in Table 9. The total forest land area estimate for estimation unit 7 is 2.27 million ac:

$$\hat{A}_d = 3002133 \times (0.415 \times 0.528 + 0.361 \times 0.880 + 0.187 \times 0.979 + 0.037 \times 1.000) \approx 2270000\ acres$$

The variance of the estimate is 807 million ac^2:

$$v(\hat{A}_d) = \frac{3002133^2}{1275} \times \left[(0.415 \times 564 \times 0.000381 + 0.361 \times 416 \times 0.000148 + 0.187 \times 248 \times 0.000047 + 0.037 \times 47 \times 0.000263) + \left((1 - 0.415) \times \frac{564}{1275} \times 0.000381 +\right.\right.$$

$$(1-0.361)\times\frac{416}{1275}\times 0.000148+(1-0.187)\times\frac{248}{1275}\times 0.000047+(1-0.037)\times\frac{47}{1275}\times 0.000263\Big)\Big]\approx 807000000\ acres^2$$

Estimates are reported for all 16 estimation units in the Estimate worksheet. With estimation unit calculations completed, we can proceed with computing a statewide estimate.

4.4 Statewide Total and Variance

Estimation units are considered independent populations and therefore a statewide estimate and variance is obtained by adding estimation unit estimates and variances. There are 16 estimation units in the state and the estimate and variance of each is summarized in Table 10 and in the Estimate worksheet. Total forest land area is estimated at 17,744,159 acres. The standard error of the estimate is 92,521 acres. The sampling error is 0.52%.

Table 10. Forest land area (ac) by estimation units in the evaluation.

Estimation Unit	Estimation Unit Description	Estimate	Variance
1	BWCA	711,051	178,198,329
2	Chippewa NF	564,459	115,102,752
3	Inland Census Water Unit 1	38,027	38,420,711
4	Inland Census Water Unit 2	13,851	21,738,811
5	Inland Census Water Unit 3	6,341	10,047,322
6	Inland Census Water Unit 4	1,748	1,140,431
7	Private Unit 1	2,272,758	807,024,516
8	Private Unit 2	3,021,335	1,915,994,419
9	Private Unit 3	2,318,794	1,715,688,185
10	Private Unit 4	654,878	709,480,480
11	Public Unit 1	3,222,195	1,062,536,723
12	Public Unit 2	2,967,366	1,570,561,722
13	Public Unit 3	456,914	187,986,758
14	Public Unit 4	125,700	71,728,059
15	Superior NF	1,265,082	118,508,778
16	Voyagers NP	103,659	36,056,344
Total		17,744,159	8,560,214,340

4.5 Comparison

Computations in the workbook returned an estimate of 17,744,159 acres of forest land in the Minnesota with a sampling error of 0.52%. Online analytical tool EVALIDator provides the opportunity to compare our estimate to what FIA computes and reports. EVALIDator reports an estimate of 17,744,159 acres of forest land in Minnesota with a sampling error of 0.52% (Figure 5). Agreement between the two sets of estimates provides confirmation of the validity of the computations performed in this study.

EVALIDator Version 1.8.0.00 - View report

Numerator attribute number and description: 0002 Area of forest land, in acres
FIADEF as the forest land definition.
Statecd/EVALID(s):
Minnesota 272017
Page variable=None (based on values from the Current inventory).
Row variable=State code (based on values from the Current inventory).
Column variable=EVALID (based on values from the Current inventory).
Filtering clause(s):

Estimate:

	EVALID	
State code	**Total**	**Minnesota 2017 rscd= 23 evalid= 271701**
Total	17,744,159	17,744,159
27 Minnesota	17,744,159	17,744,159

Sampling error percent (Confidence level 68%):

Note: for 95% confidence level multiply SE pct by 1.96

	EVALID	
State code	**Total**	**Minnesota 2017 rscd= 23 evalid= 271701**
Total	0.52	0.52
27 Minnesota	0.52	0.52

Figure 5. FIA report of area of forest land.

5. Application

Emerald ash borer is a wood-boring insect native to Asia. The first detection of the insect in the U.S. was in 2002 in the vicinity of Detroit, Michigan. Since then the insect has spread to 29 additional states (USDA APHIS, 2019). Extensive mortality of ash trees has been observed in areas where the insect has expanded (Pugh et al., 2011). The first detection of emerald ash borer in Minnesota occurred in St. Paul in 2009. Currently there are 19 counties under quarantine in the state. The largest areas of general infestation included much of the Twin Cities metropolitan area, portions of southeastern Minnesota adjacent near Wisconsin, and the city of Duluth. Black ash (*Fraxinus nigra* Marshall), green ash (*Fraxinus pennsylvanica* Marshall), and white ash (*Fraxinus americana* L.) occur in the state and all are susceptible to the emerald ash borer. As an initial assessment of potential impacts, we desired estimates of the amount of ash forest land area in the Minnesota.

5.1 Domain Specification

For every condition in the evaluation we computed the percentage net merchantable bole volume of live trees (at least 5 inches DBH) that was ash. This computation requires obtaining the tree table (MN_TREE.CSV) from the FIA DataMart. By condition, the volume of live ash trees was summed and divided by the sum of the volume of all live trees and the result multiplied by 100. The condition observations were then added as *ashvolper* column to the Condition worksheet in the workbook FIA-Data-Ash, which is a copy of the workbook FIA-Data. The *domain-indicator* column in the Condition worksheet was modified so that a value 1 was produced when a condition was forest land and had a specified percentage ash volume and a value 0 otherwise.

5.2 Estimates

Ash is present on an estimated 4.34 million of Minnesota's 17.7 million acres of forest land (Table 11). More restrictive criteria for what constitutes ash forest land reduces the estimate. There is an estimated 1.1 million acres of forest land where ash constitutes more than 50% or more of the stand volume. The FIA-Data-Ash workbook can be used to produce estimates for different thresholds.

Table 11. Forest land area (ac) estimates for varying levels of ash presence.

Stand Ash Volume	Estimate	Standard Error	Sampling Error
More than 0%	4,338,833	94,085	2.17%
At least 25%	1,775,875	64,822	3.65%
At least 50%	1,098,942	51,361	4.67%

We note the approach to specifying the percentage ash is not limited to using volume, but could be based on number, basal area, stocking, or biomass. The threshold could be amount of ash rather than percentage of the total. For example, the basal area per acre of ash could be computed for each condition and thresholds expressed in units square feet per acre used. We note that one needs to exercise care to properly account for the different plot sizes used to tally trees 1.0-4.9" and 5.0"+ when computing per acre estimates with diameter thresholds less than 5.0" inches. In computing per acre observations, condition proportion for plots with multiple conditions is often factored in the computation. We refer readers considering these options to the *balive* field in the Condition table. Successful computation of that value across a range of conditions would demonstrate correct accounting of different plot sizes and partial conditions.

5.3 Comparison

Our estimates of ash forest land match estimates from EVALIDator. Producing these estimates in EVALIDator required addition of SQL. At the bottom of step 4 in EVALIDator is the option to input additional SQL that will be applied to the query. Figure 6 shows the output of the EVLAIDator retrieval of the estimate of forest land where ash is at least 25% of the stand volume. The SQL used is displayed in the line that starts with "Filtering clause(s)".

EVALIDator Version 1.8.0.00 - View report

Numerator attribute number and description: 0002 Area of forest land, in acres
FIADEF as the forest land definition.
Statecd/EVALID(s):
Minnesota 272017
Page variable=None (based on values from the Current inventory).
Row variable=State code (based on values from the Current inventory).
Column variable=EVALID (based on values from the Current inventory).
Filtering clause(s): and 0.25<=(select sum(volcfnet*decode(t.spcd,541,1,543,1,544,1,0))/sum(volcfnet) from fs_fiadb.tree t where cond.condid=t.condid and cond.plt_cn=t.plt_cn and t.statuscd=1)

Estimate:

	EVALID	
State code	**Total**	**Minnesota 2017 rscd= 23 evalid= 271701**
Total	1,775,875	1,775,875
27 Minnesota	1,775,875	1,775,875

Sampling error percent (Confidence level 68%):

Note: for 95% confidence level multiply SE pct by 1.96

	EVALID	
State code	**Total**	**Minnesota 2017 rscd= 23 evalid= 271701**
Total	3.65	3.65
27 Minnesota	3.65	3.65

Figure 6. FIA report of area of ash forest land (ash at least 25% of stand volume).

6. Conclusion

Using the latest data from the U.S. NFI and estimation methods documented in the literature, we compute an estimate of forest land area in Minnesota that matches the estimate published by the U.S. Forest Service. Users of NFI data who desire an understanding of the sample design and estimation procedures may find the step-by-step presentation and use of a spreadsheet a helpful learning tool. Our analysis focused on Minnesota. As a national program, key sample design and estimation procedures are consistent across the nation and thus the approach described here can be applied to other states and regions.

Acknowledgements

The authors wish to thank Dr. Scott Hillard of the Minnesota Department of Natural Resources for reviewing the staff paper and providing feedback that resulted in a more complete report. We think Pat Miles of the US Forest Service for assistance with the EVALIDator retrievals.

References

Avery, T.E.; Burkhart, H.E. 2002. Forest Measurements. 5^{h} ed. Waveland Press, Inc., Long Grove, IL.

Bechtold, W.A.; Patterson, P.L.; [Editors] 2005. The enhanced Forest Inventory and Analysis program – national sampling design and estimation procedures. General Technical Report SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Research Station. 85 p.

Bechtold, W.A.; Scott, C.T. 2005. The Forest Inventory and Analysis plot design. In: Bechtold, W.A., Patterson, P.L., eds. The enhanced Forest Inventory and Analysis program – national sampling design and estimation procedures. General Technical Report SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Research Station: 27-42.

Burrill, E.A.; Wilson, A.M.; Turner, J.A.; Pugh, S.A.; Menlove, J.; Christiansen, G.; Conkling, B.L.; David, W. 2018. The Forest Inventory and Analysis Database: database description and user guide version 8.0 for Phase 2. U.S. Department of Agriculture, Forest Service. 946 p. [Online]. Available at web address: http://www.fia.fs.fed.us/library/database-documentation/.

Edgar, C.B.; Westfall, J.A.; Klockow, P.A.; Vogel, J.G.; Moore, G.W. 2009. Interpreting effects of multiple, large-scale disturbances using national forest inventory data: A case study of standing dead trees in east Texas. Forest Ecology and Management 437: 27-40.

Husch, B.; Beers, T.W.; Kershaw, J.A. 2002. Forest Mensuration. John Wiley & Sons, New York.

Miles, P.D.; Crocker, S.J.; Walters, B.F.; Kepler, D. 2017. Forests of Minnesota, 2016. Resource Update FS-115. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 4 p. https://doi.org/10.2737/FS-RU-115.

Miles, P.D.; VanderSchaaf, C.L.; Barnett, C.; Butler, B.t J.; Crocker, S.J.; Gormanson, D.; Kurtz, C.M.; Lister, T.W.; McWilliams, W.H.; Morin, R.S.; Nelson, M.D.; Perry, C.H.; Riemann, R.I.; Smith, J.E.; Walters, B.F.; Westfall, J.A.; Woodall, C.W. 2016. Minnesota Forests 2013. Resource Bulletin NRS-104. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station. 134 p.

Minnesota Department of Agriculture. 2019. Emerald ash borer status. https://mnag.maps.arcgis.com/apps/webappviewer/index.html?id=63ebb977e2924d27b9ef0787ecedf6e9. Access May 2, 2019.

Patterson, P.L.; Reams, G.A. 2005. Combining panels for Forest Inventory and Analysis estimation. In: Bechtold, W.A.; Patterson, P.L., eds. The enhanced Forest Inventory and Analysis program – national sampling design and estimation procedures. General Technical Report SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Research Station: 69-73.

Pugh, S.A.; Liebhold, A.M.; Morin, R.S. 2011. Changes in ash tree demography associated with emerald ash borer invasion, indicated by regional forest inventory data from the Great Lakes States. Canadian Journal of Forest Research 41: 2165-2175.

Pugh, S.A.; Turner, J.A.; Burrill, E.A.; David, W. 2018. The Forest Inventory and Analysis Database: population estimation user guide (Edition: November, 2018). U.S. Department of Agriculture, Forest Service. 166 p. [Online]. Available at web address: http://www.fia.fs.fed.us/library/database-documentation/.

Reams, G.A.; Smith, W.D.; Hansen, M.H.; Bechtold, W.A.; Roesch, F.A.; Moisen, G.G. 2005. The Forest Inventory and Analysis sampling frame. In: Bechtold, W.A., Patterson, P.L., eds. The enhanced Forest Inventory and Analysis program – national sampling design and estimation procedures. General Technical Report SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Research Station: 11-26.

Scott, C.T.; Bechtold, W.A.; Reams, G.A.; Smith, W.D.; Westfall, J.A.; Hansen, M.H.; Moisen, G.G. 2005. Sample-based estimators used by the Forest Inventory and Analysis national information management system. In: Bechtold, W.A., Patterson, P.L., eds. The enhanced Forest Inventory and Analysis program – national sampling design and estimation procedures. General Technical Report SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Research Station: 43-67.

Spurr, S.H. 1952. Forest Inventory. The Ronald Press Company, New York.

USDA Animal and Plant Health Inspection Service. Emerald ash borer beetle. https://www.aphis.usda.gov/aphis/resources/pests-diseases/hungry-pests/the-threat/emerald-ash-borer/emerald-ash-borer-beetle. Accessed May 2, 2019.

U.S. Forest Service. 2019a. Engagement Portfolio. https://usfs.maps.arcgis.com/apps/PublicGallery/index.html?appid=e67afb7be2c8443ca64a2e8a8e473532. Accessed May 2, 2019.

U.S. Forest Service. 2019b. Examples of FIA point-in-time estimation procedures for several common cases. https://www.fia.fs.fed.us/library/sampling/docs/supplement5_121704.xls. Accessed May 2, 2019.

U.S. Forest Service. 2019c. FIA DataMart. https://apps.fs.usda.gov/fia/datamart/datamart.html. Accessed May 3, 2019.

Vogt, J.T.; Smith, W.B. 2017. Forest Inventory and Analysis: Fiscal year 2016 business report. FS-1075. Washington, D.C.: U.S. Department of Agriculture, Forest Service, 74 p.

Westfall, J.A.; Patterson, P.L.; Coulston, J.W., 2011. Post-stratified estimation: within-strata and total sample size recommendations. Can. J. For. Res. 41, 1130–1139.