Environ Monit Assess (2026) 198:239
https://doi.org/10.1007/s10661-026-15060-3

RESEARCH

Methods for estimating status and change in presence and foliar cover area of invasive plants using national forest inventory data

James A. Westfall · Andrew J. Lister · Scott A. Pugh · Tonya W. Lister · Cassandra M. Kurtz

Received: 26 June 2025 / Accepted: 1 February 2026

Abstract The influence of invasive plant species (IPS) on forest ecosystems has been of sufficient concern to warrant data acquisition on their presence and abundance. These data are needed at various spatial scales for both land management and policy purposes. One potential source of data are national forest inventories (NFIs) that often record IPS observations on sample plots that span relatively large spatial and temporal extents, thus providing data users with a wide range of analytical opportunities to meet desired information needs. In the NFI of the USA, which is conducted by the USA Forest Service's Forest Inventory and Analysis (FIA) program, IPS presence and foliar cover area are visually observed and recorded nationally on a spatially balanced subset of inventory plots measured each year. As with any sample-based assessment, the implementation of statistically appropriate methods is imperative for credibility of the final estimates. In this paper, statistical methods for calculating population estimates and associated measures of uncertainty for current status and change in IPS attributes are presented. The estimators address both presence and cover area estimation, as well as estimation of area proportions relative to total area domains. Also included are data and various numeric examples that assist in understanding the estimation steps and provide for verification of accuracy in results. The methods should be applicable to most IPS data collected in other forest inventories with little or no modification.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10661-026-15060-3.

J. A. Westfall (✉) · A. J. Lister · T. W. Lister
U.S. Forest Service, Northern Research Station, York, PA, USA
e-mail: james.westfall@usda.gov

A. J. Lister
e-mail: andrew.lister@usda.gov

T. W. Lister
e-mail: tonya.lister@usda.gov

S. A. Pugh
U.S. Forest Service, Northern Research Station, Houghton, MI, USA
e-mail: scott.a.pugh@usda.gov

C. M. Kurtz
U.S. Forest Service, Northern Research Station, St. Paul, MN, USA
e-mail: cassandra.m.kurtz@usda.gov

Keywords Species occupancy · Understory vegetation · Change estimation · Post-stratification · Ratio estimate

Introduction

Invasive plant species (IPS) are nonnative species that can cause deleterious impacts by spreading aggressively and outcompeting surrounding vegetation. Within forest ecosystems, IPS can alter nutrient

Published online: 21 February 2026

availability, supplant desirable species, impact aesthetics, and cause negative economic consequences (Kaufman & Kaufman, 2023; Senator & Rozenberg, 2017). Not only are these plants economically burdensome within forests, but they are also costly to manage in farmland and urban areas (Arrington, 2021; Bajwa et al., 2019). Many of these invasive species use efficient propagation methods such as rooting from small fragments and prolific seed production (Gioria et al., 2023; Simberloff, 2009). There has also been recognition of positive aspects of IPS presence (Tassin & Kull, 2015), and beneficial uses such as for ecological site remediation, medicinal, and culinary purposes have been documented (Scott, 2010; Shen et al., 2023). However, their negative impacts pose challenges in the context of maintaining ecologically robust forest ecosystems (Blossey, 1999).

Due to the potential ramifications of IPS on forest resources, information regarding their presence and abundance is critical for land managers and policy makers. IPS data are collected at various spatial scales depending on the analytical need and funding availability, but often holistic landscape scale assessments rely on large regional- or national-scale data such as those collected in national forest inventories (NFI). For example, NFI data in Germany were used to assess invasion risk of two non-native tree species (Bindewald et al., 2021), analyses of invasive herbaceous and tree species have been conducted using NFI data from Spain (Alberdi et al., 2017), and non-native plant species richness and abundance were evaluated from NFI data across New Zealand (Jo et al., 2024). In the NFI of the USA, which is conducted by the USA Forest Service's Forest Inventory and Analysis (FIA) program, IPS presence and foliar cover are visually observed and recorded for a subset of plots that are designated for additional data collection to address data needs related to forest health indicators (Kurtz, 2013; Morin, 2020) . Numerous IPS assessments have been made using FIA data (Christensen et al., 2016; Kurtz, 2023; Oswalt et al., 2015); however, reporting of results is often accompanied by little documentation of the analytical methods and lack estimation of statistical uncertainty. Thus, there is a need to elucidate a formal estimation procedure for IPS analyses using FIA data with possible extensions to other NFI attributes and programs as well. The estimation processes used by the FIA program have been published (Scott et al., 2005), so a guiding principle of this work is that IPS analyses should generally conform to the existing methods. Consequently, the objectives of this manuscript are (1) to present estimators pertaining to estimating the status of IPS presence and cover (including sampling variance), (2) to extend these methods to include estimation of temporal change in IPS status, and (3) to provide numeric examples to aid in the reproducibility of estimates for data users.

Methods

Data

The methods described below are in the context of the sample design (Reams et al., 2005), plot configuration (Bechtold & Scott, 2005), and field measurement protocols (U.S. Forest Service, 2024) implemented by FIA. The sample design is quasi-systematic with post-stratification employed to increase the precision of estimates. IPS measurements are taken on a 1/8 (12.5%) subset of FIA plots and thus the sampling intensity for IPS is approximately 1 plot per 30,351 ha (75,000 ac). The plot configuration is composed of four circular subplots, each having a 7.32 m (24 ft) radius with one subplot located at plot center and the remaining three subplots centered at azimuths 120, 240, and 360 degrees and distance of 36.58 m (120 ft) from plot center (Fig. 1). A mapping protocol is used within each subplot to record any boundaries between differing conditions, such as between forest and nonforest land use or within forest areas that may diverge in characteristics such as forest type, stand size class, or ownership (U.S. Forest Service, 2024). These mapped conditions allow for the estimation of subplot area proportions for each delineated condition. The measurements for IPS are collected in forested conditions within each subplot and include species identification and percent foliar cover for each IPS present. Cover is visually estimated by species for each measured condition on the subplot separately and recorded as a percent of the entire subplot. Recording the condition-level values as a percent of subplot area allows the total proportion for the entire subplot area to be estimated as the sum of the condition values. Area of presence for a given IPS is the condition area within the subplot where the species was observed. Nonforest conditions are assumed to

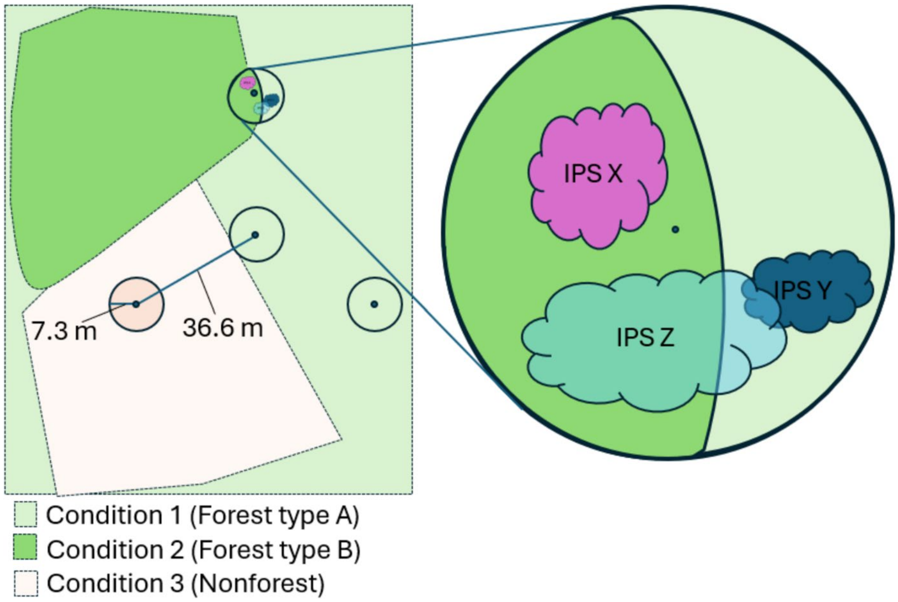

Fig. 1 The FIA cluster plot for IPS data collection consists of an array of four subplots (7.3 m radius) with three satellite subplots separated from a center subplot by 36.6 m at azimuths of 0, 120, and 240 degrees. In this example, there are three conditions on the plot, delineated using FIA's condition-defining rules, which are based on differences in land use, forest type, or certain other circumstances. The foliar cover proportion for each IPS is recorded for each condition on the subplot. This example depicts estimated cover area percent for IPS species X (10%), species Y (5%), and species Z (17%). No IPS measurements are collected on nonforest portions of plots

have no extant IPS. Plants qualifying for inclusion in IPS observation protocols are defined by a state-specific species list (found in the FIA database table REF_INVASIVE_SPECIES), which is accessible as part of the FIA reference table set available at: https://apps.fs.usda.gov/fia/datamart/CSV/FIADB_REFERENCE.zip. Documentation of table contents is provided by Burrill et al. (2024).

Provision of example calculations relied on data from the state of Maryland (MD), USA. Estimates of current status were obtained from data collected between 2013 and 2019 on 109 plots (hereafter identified as 2019 data). Lacking an appropriate set of actual remeasurement data, estimation of change was based on a subjective perturbation of the 2019 data to create a synthetic dataset that represented change over time. Modifications to the original 2019 data included removal of three IPS of rare occurrence and minor perturbations to plot IPS cover percent and forest area values. Thus, the change estimates do not reflect actual change in IPS presence, IPS foliar cover, or forest land area. For all data, IPS cover percent data were converted to proportions to align with forest area proportion data and facilitate consistency in estimations. The example calculations provided assume that the 2019 data were the observations for time 1 (T1) and the synthetic data represented conditions at time 2 (T2), presumably associated with the next remeasurement cycle.

The field data are used to construct population estimates following the post-stratified estimation procedures described by (Scott et al., 2005). For a given state, the total area is partitioned into $g=1$ to G populations referred to as estimation units (Westfall et al., 2022). Within each estimation unit g, canopy cover classes predicted from classified remotely sensed imagery are used to apportion the area into $h=1$ to H post-strata, the areas of which determine the stratum weights (W_{gh}) (Gormanson et al., 2023). Typically, a minimum of 10 sample units per post-stratum is strived for (Westfall et al., 2011); however, fewer sample units are considered acceptable in some cases, e.g., estimation unit 1, which represents sample units in water which will all have observations of zero forest land and IPS attributes. Table 1 provides a summary of the data by estimation unit and post-stratum.

Table 1 Summary of 2019 and synthetic data sets across all IPS by estimation unit g and stratum h where A_g is the estimation unit area, W_{gh} is the stratum weight, n_{gh} is the stratum sample size, $\overline{p}_{gh}$ is the partial nonresponse adjustment factor, A_{ghid} is the plot i area proportion in the forest domain d adjusted by $\overline{p}_{gh}$, S_{ghi} is the plot i number of species, and C^*_{ghi} is the plot i foliar cover proportion adjusted by $\overline{p}_{gh}$. For detailed descriptions of these variables and their context within the estimation framework, consult the Analysis section below

							Minimum			Mean			Maximum		
Data	Estimation unit	Area (ha)	Stratum	Weight	Sample size	Adjustment factor	Forest area proportion	Species count	Cover proportion	Forest area proportion	Species count	Cover proportion	Forest area proportion	Species count	Cover proportion
	g	A_g	h	W_{gh}	n_{gh}	$\overline{p}_{gh}$	A_{ghid}	S_{ghi}	C^*_{ghi}	A_{ghid}	S_{ghi}	C^*_{ghi}	A_{ghid}	S_{ghi}	C^*_{ghi}
T1 (2013–2019)	1	175,997.1	1	1.0000	3	1.000	0.000	0.00	0.0000	0.0000	0.00	0.0000	0.000	0.00	0.0000
	2	2,202,465.3	1	0.4927	46	0.995	0.000	0.00	0.0000	0.0146	0.07	0.0007	0.565	3.00	0.0327
	2	2,202,465.3	2	0.5073	42	0.964	0.000	0.00	0.0000	0.6079	4.02	0.0973	1.037	23.00	0.7006
	3	329,145.5	1	1.0000	18	0.940	0.000	0.00	0.0000	0.7984	4.67	0.1121	1.064	24.00	0.8643
T2 (Synthetic)	1	175,997.1	1	1.0000	3	1.000	0.000	0.00	0.0000	0.0000	0.00	0.0000	0.000	0.00	0.0000
	2	2,202,465.3	1	0.4927	46	0.995	0.000	0.00	0.0000	0.0109	0.07	0.0007	0.424	3.00	0.0343
	2	2,202,465.3	2	0.5073	42	0.964	0.000	0.00	0.0000	0.6010	3.98	0.1009	1.037	23.00	0.7194
	3	329,145.5	1	1.0000	18	0.940	0.000	0.00	0.0000	0.7797	4.56	0.1173	1.064	24.00	0.9064

Analysis

The estimation methods described below generally follow those documented for the FIA program (Scott et al., 2005; Westfall et al., 2022). However, the referenced FIA documentation does not specifically address estimation of species presence and cover area for special observation protocols such as those for IPS. A key element in defining the specifications for the desired estimate is the use of a binary (0/1) domain indicator variable δ. The domain d is a user specified set of criteria and can include a wide range of data elements collected at the plot-, condition-, subplot-, or plant-level. In the context of this document, key domain components are forest land use and presence/absence status of an IPS of interest. Further refinements can be made such as specifying any combination of location (e.g., county), condition (e.g., forest type), and/or plant (e.g., minimum cover present) domain elements. For example, consider a domain d specified by forest type B (e.g., Fig. 1) having presence of IPS species X. In this case, $\delta = 1$ only if the IPS species X is present, the land use is forest, and the forest type is classified as B; otherwise, $\delta = 0$. The domain-based estimation approach allows users to customize estimates to meet their specific information needs. As will be seen below, the domain indicator variable plays a key role in determining the observed value for each plot.

Estimation of IPS presence area and proportion of a land domain

To estimate the total area where IPS are present, each subplot is evaluated for presence/absence in the context of the specified domain d (e.g., d is forest land containing a selected IPS species of interest). The area proportion of IPS presence on each plot is estimated by summing the area proportion of forest conditions in each subplot where IPS associated with domain d are present; i.e., the condition area is considered the presence area where the domain of interest is satisfied. The following equation accounts for the fact that some subplots may have more than one condition k that is forested:

$$A\_IPS_{ghid} = \sum_{j=1}^{4}\sum_{k=1}^{K}\frac{A_{ghijk}\delta_{ghijkd}}{\overline{p}_{gh}} \quad (1)$$

where within stratum h of estimation unit g, A_IPS_{ghid} is the area proportion of IPS presence on plot i in domain d, A_{ghijk} is the proportion of plot-level area on plot i in subplot j having condition k, δ_{ghijkd} is the indicator variable for domain d on plot i in subplot j having condition k, and $\overline{p}_{gh}$ is an adjustment factor to account for plots only partially accessed (i.e., that have partial nonresponse). The adjustment factor is calculated as:

$$\overline{p}_{gh} = \frac{\sum_{i=1}^{n_{gh}} a_{ghi}}{n_{gh}} \quad (2)$$

where, in stratum h of estimation unit g, a_{ghi} is the accessible area proportion of plot i and n_{gh} is the number of plots at least partially accessible. The application of this adjustment factor has the same effect on the estimate as imputing the stratum average value for the attribute to any missing portions of the plots. To proceed with estimating the total area of IPS presence in domain d, estimate the mean area proportion $\overline{A\_IPS}_{ghd}$ and its variance $v(\overline{A\_IPS}_{ghd})$ for each stratum h in estimation unit g:

$$\overline{A\_IPS}_{ghd} = \frac{\sum_{i=1}^{n_{gh}} A\_IPS_{ghid}}{n_{gh}} \quad (3)$$

$$v(\overline{A\_IPS}_{ghd}) = \frac{\sum_{i=1}^{n_{gh}}\left(A\_IPS_{ghid} - \overline{A\_IPS}_{ghd}\right)^2}{n_{gh}(n_{gh}-1)} \quad (4)$$

To estimate total IPS presence area and the associated variance in domain d for estimation unit g, combine strata means and variances using the stratum weights W_{gh} from the post-stratification and the known area of the estimation unit A_g:

$$\widehat{A\_IPS}_{gd} = A_g \sum_{h=1}^{H} W_{gh}\overline{A\_IPS}_{ghd} \quad (5)$$

$$v\left(\widehat{A\_IPS}_{gd}\right) = \frac{A_g^2}{n_g}\left[\sum_{h=1}^{H} W_{gh}n_{gh}v\left(\overline{A\_IPS}_{ghd}\right) + \sum_{h=1}^{H}(1-W_{gh})\frac{n_{gh}}{n_g}v\left(\overline{A\_IPS}_{ghd}\right)\right] \quad (6)$$

Having the estimates of total presence area and associated variances for each estimation unit, calculation of state-level estimates is accomplished by summation across the G estimation units in the state:

$$\widehat{A\_IPS}_d = \sum_{g=1}^{G} \widehat{A\_IPS}_{gd} \tag{7}$$

$$v(\widehat{A\_IPS}_d) = \sum_{g=1}^{G} v(\widehat{A\_IPS}_{gd}) \tag{8}$$

Practitioners may also be interested in evaluating $\widehat{A\_IPS}_d$ in the context of a specified land area domain, e.g., $\widehat{A\_IPS}_d$ as a proportion of total land area $\hat{A}_{d\prime}$, where *d'* is a secondary domain that could be defined as all forest, a selected forest type, or other land type domain of interest. It is important to emphasize here that *d'* is commonly an area-based domain that has some relationship to the domain *d*. For example, domain *d* may be specified as presence of multiflora rose (*Rosa multiflora* Thunb.) in oak/hickory forest type and the associated domain *d'* is oak/hickory forest type. Similar calculations to those performed above are implemented to construct the estimate. First obtain the area proportion in domain *d'* for each plot:

$$A_{ghid'} = \sum_{j=1}^{4} \sum_{k=1}^{K} \frac{A_{ghijk}\delta'_{ghijkd'}}{\bar{p}_{gh}} \tag{9}$$

where $A_{ghid'}$ is the area proportion in domain *d'* on plot *i* in stratum *h* of estimation unit *g*. $\delta'_{ghijkd'}$ equals 1 if the presence of domain *d'* is indicated and equals 0 otherwise. The next step is to estimate the mean area proportion $\bar{A}_{ghd'}$ and its variance $v(\bar{A}_{ghd'})$ for each stratum h in estimation unit g:

$$\bar{A}_{ghd'} = \frac{\sum_{i=1}^{n_{gh}} A_{ghid'}}{n_{gh}} \tag{10}$$

$$v(\bar{A}_{ghd'}) = \frac{\sum_{i=1}^{n_{gh}} \left(A_{ghid'} - \bar{A}_{ghd'}\right)^2}{n_{gh}(n_{gh}-1)} \tag{11}$$

Now the estimated area in domain *d'* for the estimation unit *g* and its estimated variance are calculated by combining stratum means and variances across strata:

$$\hat{A}_{gd'} = A_g \sum_{h=1}^{H} W_{gh}\bar{A}_{ghd'} \tag{12}$$

$$v\left(\hat{A}_{gd'}\right) = \frac{A_g^2}{n_g}\left[\sum_{h=1}^{H} W_{gh} n_{gh} v\left(\bar{A}_{ghd'}\right) + \sum_{h=1}^{H} (1 - W_{gh)} \frac{n_{gh}}{n_g} v\left(\bar{A}_{ghd'}\right)\right] \tag{13}$$

Estimations analogous to Eqs. (7) and (8) are performed to obtain state-level estimates $\hat{A}_{d'}$ and $v(\hat{A}_{d'})$. The desired estimate of area proportion of IPS presence within the land area associated with *d'* is calculated as a ratio estimate:

$$\widehat{A\_P}_{dd'} = \frac{\widehat{A\_IPS}_d}{\hat{A}_{d'}} \tag{14}$$

with its associated variance having the form of:

$$v(\widehat{A\_P}_{dd'}) = \frac{1}{\hat{A}_{d'}^2}\left[v(\widehat{A\_IPS}_d) + \widehat{A\_P}_{dd'}^2 v(\hat{A}_{d'}) - 2\widehat{A\_P}_{dd'} cov(\hat{A}_{d'}, \widehat{A\_IPS}_d)\right] \tag{15}$$

where $\hat{A}_{d'}$, $v(\hat{A}_{d'})$, $\widehat{A\_P}_{dd'}$, and $v(\widehat{A\_IPS}_d)$ are as defined before. The remaining statistic to be estimated to complete the variance estimation at the state level is the covariance between the estimates of the numerator and denominator $cov(\hat{A}_{d'}, \widehat{A\_IPS}_d)$. The estimations for each stratum *h* within estimation unit *g* are given by:

$$cov(\bar{A}_{ghd'}, \overline{A\_IPS}_{ghd}) = \frac{\sum_{i=1}^{n_{gh}} \left(A\_IPS_{ghid} - \overline{A\_IPS}_{ghd}\right)\left(A_{ghid'} - \bar{A}_{ghd'}\right)}{n_{gh}(n_{gh}-1)} \tag{16}$$

In like manner to the steps taken for variances, the covariances are combined across the *H* strata to estimate the covariance for the estimation unit *g*:

$$cov\left(\hat{A}_{gd'}, \widehat{A\_IPS}_{gd}\right) = \frac{A_g^2}{n_g}\left[\sum_{h=1}^{H} W_{gh} n_{gh} cov(\bar{A}_{ghd'}, A\_\bar{I}PS_{ghd} + \sum_{h=1}^{H} (1 - W_{gh}) \frac{n_{gh}}{n_g} cov(\bar{A}_{ghd'}, A\_\bar{I}PS_{ghd})\right] \tag{17}$$

Finally, the estimated covariance at the state level is obtained by summing the covariances of the estimates across the *G* estimation units:

$$cov\left(\hat{A}_{d\prime}, \widehat{A\_IPS}_d\right) = \sum_{g=1}^{G} cov\left(\hat{A}_{gd\prime}, \widehat{A\_IPS}_{gd}\right) \tag{18}$$

Estimated IPS cover area

To estimate the total cover area, the cover area proportion on each subplot C^*_{ghijd} is estimated in the context of the specified domain d:

$$C^*_{ghijd} = \sum_{k=1}^{K} C_{ghijk}\delta_{ghijkd} \tag{19}$$

where within plot i of stratum h in estimation unit g, C_{ghijk} is the cover proportion for plot i, subplot j, in condition k. The overall cover area proportion C^*_{ghid} for plot i is estimated by summing over the four subplots and dividing by the aforementioned partial plot adjustment factor given in Eq. (2):

$$C^*_{ghid} = \sum_{j=1}^{4} \frac{C^*_{ghijd}}{\overline{p}_{gh}} \tag{20}$$

The sample mean and its variance in stratum h of estimation unit g are estimated as:

$$\overline{C}^*_{ghd} = \frac{\sum_{i=1}^{n_{gh}} C^*_{ghid}}{n_{gh}} \tag{21}$$

$$v(\overline{C}^*_{ghd}) = \frac{\sum_{i=1}^{n_{gh}} \left(C^*_{ghid} - \overline{C}^*_{ghd}\right)^2}{n_{gh}(n_{gh} - 1)} \tag{22}$$

The estimated total and associated variance for estimation unit g are:

$$\widehat{C}^*_{gd} = A_g \sum_{h=1}^{H} W_{gh}\overline{C}^*_{ghd} \tag{23}$$

$$v\left(\widehat{C}^*_{gd}\right) = \frac{A_g^2}{n_g}\left[\sum_{h=1}^{H} W_{gh}n_{gh}v\left(\widehat{C}^*_{gd}\right) + \sum_{h=1}^{H}(1 - W_{gh)}\frac{n_{gh}}{n_g}v\left(\widehat{C}^*_{gd}\right)\right] \tag{24}$$

As shown previously in Eqs. (7) and (8), summation across the G estimation units provides the appropriate estimates at the state level:

$$\widehat{C}^*_{d} = \sum_{g=1}^{G} \widehat{C}^*_{gd} \tag{25}$$

$$v(\widehat{C}^*_{d}) = \sum_{g=1}^{G} v(\widehat{C}^*_{gd}) \tag{26}$$

Estimated mean cover proportion of a land domain

It may also be of interest to estimate the IPS cover area proportion in the context of a specified land domain. For example, a common land domain of interest is land on which a given IPS is present (i.e., land containing the IPS is designated as $d\prime$). The proportion estimate therefore is defined as the proportion cover of an IPS on land where that IPS is present. Alternatively, other land domains could be specified such as all forest land. Regardless of how d' is specified, estimation proceeds identically as in the previous section using Eqs. (1–18) to obtain estimates of $\widehat{C}^*_d$ instead of $\widehat{A\_IPS}_d$. Ultimately, the estimated ratio $\widehat{C^*\_P}_{dd'}$ and its variance are defined as:

$$\widehat{C^*\_P}_{dd'} = \frac{\widehat{C}^*_d}{\widehat{A}_{d'}} \tag{27}$$

$$v\left(\widehat{C^*\_P}_{dd'}\right) = \frac{1}{\widehat{A}^2_{d'}}\left[v\left(\widehat{C}^*_d\right) + \widehat{C^*\_P}^2_{dd'}v\left(\widehat{A}_{d'}\right) - 2\widehat{C^*\_P}_{dd'}cov(\widehat{C}^*_d, \widehat{A}_{d'})\right] \tag{28}$$

Combining covers of multiple species

The methods described above provide the basis for estimating plot-level values based on domain d for a single IPS. However, there may be cases where cover information is desired for several species in combination. Using FIA data collection protocols, the cover observations are made independently by species with no accounting of any overlapping cover among species. This situation thus requires further consideration as assumptions other than no overlapping cover or entirely overlapping cover may be preferred. For example, the principle of inclusion–exclusion for calculating the probability of union for multiple events (Kahn et al., 1996) may be invoked to approximate the total combined cover of multiple species that may partially overlap under the assumption of random assortment. In this framework, the percent cover of a species is treated as a probability of occurrence at any given point in the subplot under the assumption of random spatial distribution. Therefore, estimations to obtain the requisite plot observations in the context of domain d that combine cover measurements from

$s = 1$ to S species may be approximated via (Jennings et al., 2009):

$$C^*_{ghijd} = \sum_{k=1}^{K}\left[1 - \prod_{s=1}^{S}\left(1 - C_{ghijks}\delta_{ghijkd}\right)\right] \quad (29)$$

where within plot i of stratum h in estimation unit g, C^*_{ghijd} is the proportion of plot area in subplot j that is covered by all species in the specified domain d, δ_{ghijkd} is a domain indicator that includes S species, and C_{ghijks} is the cover proportion of species s for subplot j in condition k. The overall cover area proportion C^*_{ghid} for plot i is obtained by summing over the four subplots:

$$C^*_{ghid} = \sum_{j=1}^{4}\frac{C^*_{ghijd}}{\overline{p}_{gh}} \quad (30)$$

Having established C^*_{ghid} for S > 1 species, calculations can proceed as described earlier, e.g., beginning at Eq. (21) for combined cover estimates of species defined in domain d.

Estimating change in area

This section will focus on two primary analytical needs: (1) estimation of change in the area for presence and cover and (2) estimation of change in the proportion for presence and cover area. The first need addresses whether the area of occurrence is changing, whereas the second frames the analysis in the context of forest area change. Estimating change requires an initial measurement at time T1 and a subsequent measurement at time T2 such that the change is determined by the periodic difference. For the purposes of this document, the estimations of current status outlined above are considered to be generated from the T1 measurement, e.g., $\widehat{A\_IPS}_d$ and $\widehat{A}_{d'}$ using Eqs. (1)–(13). The corresponding estimates from the T2 measurement also arise from Eqs. (1) to (13) and will be denoted with a subscript, e.g., $\widehat{A\_IPS}_{(T2)d}$ and $\widehat{A}_{(T2)d'}$. As it is commonplace to use a unified post-stratification scheme for change estimation, it was assumed the same post-stratification sample plot assignments and stratum weights from T1 are also applied to T2 estimates. Area of change in presence $\widehat{\Delta A\_IPS}_d$ is calculated as:

$$\widehat{\Delta A\_IPS}_d = \widehat{A\_IPS}_{(T2)d} - \widehat{A\_IPS}_d \quad (31)$$

where $\widehat{A\_IPS}_d$ and $\widehat{A\_IPS}_{(T2)d}$ are the estimated IPS area in domain d at T1 and T2, respectively. As $\widehat{\Delta A\_IPS}_d$ is the difference between two random variables; the variance $v\left(\widehat{\Delta A\_IPS}_d\right)$ takes the form of:

$$v\left(\widehat{\Delta A\_IPS}_d\right) = v(\widehat{A\_IPS}_d) + v(\widehat{A\_IPS}_{(T2)d}) - 2cov(\widehat{A\_IPS}_d, \widehat{A\_IPS}_{(T2)d}) \quad (32)$$

The estimation procedures described above for $\widehat{A\_IPS}_d$ can be followed to obtain the estimated values for $\widehat{A\_IPS}_d$, $\widehat{A\_IPS}_{(T2)d}$, $v(\widehat{A\_IPS}_d)$, and $v(\widehat{A\_IPS}_{(T2)d})$. Estimation of the covariance proceeds generally as indicated previously (i.e., Eq. (16)) except that observations from two points in time are used instead of two attributes at a given point in time. Specifically, begin calculations for each stratum h within estimation unit g:

$$cov(\overline{A\_IPS}_{ghd}, \overline{A\_IPS}_{(T2)ghd}) = \frac{\sum_{i=1}^{n_{gh}}\left(A\_IPS_{ghid} - \overline{A\_IPS}_{ghd}\right)\left(A\_IPS_{(T2)ghid} - \overline{A\_IPS}_{(T2)ghd}\right)}{n_{gh}(n_{gh} - 1)} \quad (33)$$

Subsequently, combine the estimated covariances across strata to estimate the covariance between the attribute totals for the estimation unit g under post-stratification:

$$cov\left(\widehat{A\_IPS}_{gd}, \widehat{A\_IPS}_{(T2)gd}\right) = \frac{A_g^2}{n_g}\left[\sum_{h=1}^{H} W_{gh}n_{gh}cov(\overline{A\_IPS}_{ghd}, \overline{A\_IPS}_{(T2)ghd}) + \sum_{h=1}^{H}(1 - W_{gh})\frac{n_{gh}}{n_g}cov(\overline{A\_IPS}_{ghd}, \overline{A\_IPS}_{(T2)ghd})\right] \quad (34)$$

Obtain the estimated covariance at the state level from:

$$cov(\widehat{A\_IPS}_d, \widehat{A\_IPS}_{(T2)d}) = \sum_{g=1}^{G} cov\left(\widehat{A\_IPS}_{gd}, \widehat{A\_IPS}_{(T2)gd}\right) \tag{35}$$

Estimation of change in cover area proceeds identically as described above in Eqs. (31)–(35) for presence area change except that calculations have their basis in C^*_{ghid} and $C^*_{(T2)ghid}$ instead of A_IPS_{ghid} and $A\_IPS_{(T2)ghid}$. The corresponding notation for the cover area change estimate and associated variance at the state level would be $\Delta\widehat{C}^*_d$ and $v(\Delta\widehat{C}^*_d)$, respectively.

Estimating change in area proportion

In previous sections, options for estimating area proportions relative to a specified land domain were presented. Given that the methods for estimating change in area proportions build upon the previous calculations, the below presentation will use the same notation presented earlier in the section on presence area proportion. Recall this proportion was given earlier in Eq. (14) as $\widehat{A\_P}_{dd'} = \frac{\widehat{A\_IPS}_d}{\widehat{A}_{d'}}$. In the context of change estimation, the two estimates corresponding with different points in time ($\widehat{A\_P}_{dd'}$ = initial measurement, $\widehat{A\_P}_{(T2)dd'}$= subsequent measurement) and their difference $\Delta\widehat{P}_{dd'}$ are:

$$\widehat{A\_P}_{dd'} = \frac{\widehat{A\_IPS}_d}{\widehat{A}_{d'}} \tag{36}$$

$$\widehat{A\_P}_{(T2)dd'} = \frac{\widehat{A\_IPS}_{(T2)d}}{\widehat{A}_{(T2)d'}} \tag{37}$$

$$\Delta\widehat{A\_P}_{dd'} = \widehat{A\_P}_{(T2)dd'} - \widehat{A\_P}_{dd'} \tag{38}$$

It is clear from this formulation that $\Delta\widehat{A\_P}_{dd'}$ is the difference between two ratio estimates. Nonetheless, the general form for the estimated variance of $\Delta\widehat{A\_P}_{dd'}$ conforms to typical construction:

$$v\left(\Delta\widehat{A\_P}_{dd'}\right) = v(\widehat{A\_P}_{dd'}) + v(\widehat{A\_P}_{(T2)dd'}) - 2cov(\widehat{A\_P}_{dd'}, \widehat{A\_P}_{(T2)dd'}) \tag{39}$$

Methods for estimating $v(\widehat{A\_P}_{dd'})$ and $v(\widehat{A\_P}_{(T2)dd'})$ have been presented in previous sections, i.e., Eqs. (15)–(18). Thus, the remaining challenge is to estimate $cov(\widehat{A\_P}_{dd'}, \widehat{A\_P}_{(T2)dd'})$. Cochran (1977) prescribes a general form for estimating the covariance between two ratios. Adapting that to the situation of FIA IPS estimation suggests the specification for stratum h in estimation unit g is:

$$cov\left(\widehat{A\_P}_{ghdd'}, \widehat{A\_P}_{(T2)ghdd'}\right) = \frac{1}{n_{gh}(n_{gh}-1)\widehat{A}_{ghd'}\widehat{A}_{(T2)ghd'}} \sum_{i=1}^{n_{gh}} \Big(A\_IPS_{ghid} A\_IPS_{(T2)ghid} - \widehat{A\_P}_{ghdd'} A\_IPS_{(T2)ghid} A_{ghid'} - \widehat{A\_P}_{(T2)ghdd'} A\_IPS_{ghid} A_{(T2)ghid'} + \widehat{A\_P}_{ghdd'} \widehat{A\_P}_{(T2)ghdd'} A_{ghid'} A_{(T2)ghid'} \Big) \tag{40}$$

where $\widehat{A}_{ghd'} = \overline{A}_{ghd'} W_{gh} A_g$ and $\widehat{A}_{(T2)ghd'} = \overline{A}_{(T2)ghd'} W_{gh} A_g$ are the estimated total land area in domain d' for stratum h and estimation unit g for T1 and T2, respectively. However, there are instances where $\widehat{A}_{ghd'} = 0$ and/or $\widehat{A}_{(T2)ghd'} = 0$ which produces an undefined result when Eq. (40) is used. An alternative approach is presented that uses the state-level calculations of land area such that division by zero is more easily avoided. For this method, estimation of the covariance between the two ratio estimates begins by obtaining all the relevant pairwise covariances within the $h = 1$ to H strata in each of the $g = 1$ to G estimation units. Using the same formulation as Eq. (33), these estimates are:

$$cov(\overline{A\_IPS}_{ghd}, \overline{A\_IPS}_{(T2)ghd}) = \frac{\sum_{i=1}^{n_{gh}} \left(A\_IPS_{ghid} - \overline{A\_IPS}_{ghd}\right)\left(A\_IPS_{(T2)ghid} - \overline{A\_IPS}_{(T2)ghd}\right)}{n_{gh}(n_{gh}-1)} \tag{41}$$

$$cov(\overline{A}_{ghd'},\overline{A}_{(T2)ghd'})=\frac{\sum_{i=1}^{n_{gh}}\left(A_{ghid'}-\overline{A}_{ghd'}\right)\left(A_{(T2)ghid'}-\overline{A}_{(T2)ghd'}\right)}{n_{gh}(n_{gh}-1)} \tag{42}$$

$$cov(\overline{A\_IPS}_{ghd},\overline{A}_{(T2)ghd'})=\frac{\sum_{i=1}^{n_{gh}}\left(A\_IPS_{ghid}-\overline{A\_IPS}_{ghd}\right)\left(A_{(T2)ghid'}-\overline{A}_{(T2)ghd'}\right)}{n_{gh}(n_{gh}-1)} \tag{43}$$

$$cov(\overline{A}_{ghd'},\overline{A\_IPS}_{(T2)ghd})=\frac{\sum_{i=1}^{n_{gh}}\left(A_{ghid'}-\overline{A}_{ghd'}\right)\left(A\_IPS_{(T2)ghid}-\overline{A\_IPS}_{(T2)ghd}\right)}{n_{gh}(n_{gh}-1)} \tag{44}$$

Now obtain each estimated covariance for estimation unit g following the standard approach, e.g., for $cov\left(\widehat{A\_IPS}_{gd},\widehat{A\_IPS}_{(T2)gd}\right)$:

$$cov\left(\widehat{A\_IPS}_{gd},\widehat{A\_IPS}_{(T2)gd}\right)=\frac{A_g^2}{n_g}\left[\sum_{h=1}^{H}W_{gh}n_{gh}cov\left(\overline{A\_IPS}_{ghd},\overline{A\_IPS}_{(T2)ghd}\right)+\sum_{h=1}^{H}(1-W_{gh})\frac{n_{gh}}{n_g}cov\left(\overline{A\_IPS}_{ghd},\overline{A\_IPS}_{(T2)ghd}\right)\right] \tag{45}$$

Perform analogous calculations to estimate $cov\left(\widehat{A}_{gd'},\widehat{A}_{(T2)gd'}\right)$, $cov\left(\widehat{A\_IPS}_{gd},\widehat{A}_{(T2)gd'}\right)$, and $cov\left(\widehat{A}_{gd'},\widehat{A\_IPS}_{(T2)gd'}\right)$. In typical fashion, obtain the estimated state-level statistics by summing over the G estimation units:

$$cov\left(\widehat{A\_IPS}_{d},\widehat{A\_IPS}_{(T2)d}\right)=\sum_{g=1}^{G}cov\left(\widehat{A\_IPS}_{gd},\widehat{A\_IPS}_{(T2)gd}\right) \tag{46}$$

Obtain the values for $cov\left(\widehat{A}_{d'},\widehat{A}_{(T2)d'}\right)$, $cov\left(\widehat{A\_IPS}_{d},\widehat{A}_{(T2)d'}\right)$, and $cov\left(\widehat{A}_{d'},\widehat{A\_IPS}_{(T2)d}\right)$ similarly and estimate the overall covariance between the ratios $\widehat{A\_P}_{dd'}$ and $\widehat{A\_P}_{(T2)dd'}$ from:

$$\begin{aligned}cov\left(\widehat{A\_P}_{dd'},\widehat{A\_P}_{(T2)dd'}\right)==\frac{1}{\widehat{A}_{d'}\widehat{A}_{(T2)d'}}&\left[cov\left(\widehat{A\_IPS}_{d},\widehat{A\_IPS}_{(T2)d}\right)-\widehat{A\_P}_{dd'}cov\left(\widehat{A}_{d'},\widehat{A\_IPS}_{(T2)d}\right)\right.\\&\left.-\widehat{A\_P}_{(T2)dd'}cov\left(\widehat{A\_IPS}_{d},\widehat{A}_{(T2)d'}\right)+\widehat{A\_P}_{dd'}\widehat{A\_P}_{(T2)dd'}cov\left(\widehat{A}_{d'},\widehat{A}_{(T2)d'}\right)\right]\end{aligned} \tag{47}$$

The key difference between Eqs. (47) and (40) is the term $\frac{1}{\widehat{A}_{d'}\widehat{A}_{(T2)d'}}$, which is only undefined when the estimated area in domain *d'* is zero at the state level. This circumstance should be rare relative to zero estimated area in domain *d'* at either the post-stratum h or estimation unit g levels.

Changes in estimates of IPS cover proportion over time essentially follow the same estimation process where notationally A_IPS would be replaced by C^* throughout using Eqs. (36)–(47). For example, the covariance in Eq. (47) would take the form of:

$$cov\left(\widehat{C^*\_P}_{dd'},\widehat{C^*\_P}_{(T2)dd'}\right) == \frac{1}{\widehat{A}_{d'}\widehat{A}_{(T2)d'}}\Big[cov\left(\widehat{C}^*_{d},\widehat{C}^*_{(T2)d}\right)-\widehat{C^*\_P}_{dd'}cov\left(\widehat{A}_{d'},\widehat{C}^*_{(T2)d}\right) -\widehat{C^*\_P}_{(T2)dd'}cov\left(\widehat{C}^*_{d},\widehat{A}_{(T2)d'}\right)+\widehat{C^*\_P}_{dd'}\widehat{C^*\_P}_{(T2)dd'}cov\left(\widehat{A}_{d'},\widehat{A}_{(T2)d'}\right)\Big] \tag{48}$$

Once the calculations have been completed based on C^* data, the final state-level proportion change estimates are defined as:

$$\Delta\widehat{C^*\_P}_{dd'}=\widehat{C^*\_P}_{(T2)dd'}-\widehat{C^*\_P}_{dd'} \tag{49}$$

$$v\left(\Delta\widehat{C^*\_P}_{dd'}\right)=v(\widehat{C^*\_P}_{dd'})+v(\widehat{C^*\_P}_{(T2)dd'})-2cov\left(\widehat{C^*\_P}_{dd'},\widehat{C^*\_P}_{(T2)dd'}\right) \tag{50}$$

For all estimations, the standard error (se) of the estimate (generically indicated as $\hat{\cdot}$) is obtained from the square root of the variance:

$$se(\hat{\cdot})=\sqrt{v(\hat{\cdot})} \tag{51}$$

The percent sampling error is calculated as:

$$\%se(\hat{\cdot})=\frac{se(\hat{\cdot})}{|(\hat{\cdot})|}100 \tag{52}$$

Approximate 95% confidence interval bounds for a given estimate were $\pm z_{1-\alpha/2}se(\hat{\cdot})$ where $z_{1-\alpha/2}=1.96$ under an assumed normal distribution.

Application to urban FIA

The Urban FIA program uses a different plot configuration than that of the standard NFI. In this case, the plot consists of a single circular plot having a radius of 14.63 m (48 ft). Further, the IPS data collection protocols are applied to the entire plot area, including nonforest land and water. Given that, the land domain of interest needs to be carefully considered since classes such as water are included and the user might prefer to exclude these areas in the domain specification. There is also the opportunity to customize the estimation realm to nonforest land or subsets thereof, e.g., residential land use class. In this example, the estimate would be the proportion of residential land area that has a given IPS present.

Essentially, the calculations described above are applicable to estimates from Urban FIA except there are no longer four subplots. Thus, the primary modification to the relevant equations is to exclude the summation over the j subplots. Respecifying Eq. (1) for Urban FIA would therefore be:

$$A\_IPS_{ghid}=\sum_{k=1}^{K}\frac{A_{ghik}\delta_{ghikd}}{\overline{p}_{gh}} \tag{53}$$

where A_{ghik} is the land area in domain d of condition k within plot i of stratum h in estimation unit g. Users desiring estimates based on certain land use types can do so when specifying the domain d. When making area proportion estimates, Eq. (9) would be modified in the same manner to accommodate the urban data collection protocols. Equations (19) and (20) can be consolidated because there are no subplot calculations:

$$C^*_{ghid}=\sum_{k=1}^{K}\frac{C_{ghik}\delta_{ghikd}}{\overline{p}_{gh}} \tag{54}$$

Once the desired plot-level observations have been generated, the NFI equations described previously can be used to obtain the desired estimates for urban populations.

Example calculations

Example calculations were developed for a selected subset of the estimation procedures described above. The MD data described previously were used with a domain d specification of all IPS on forest land; i.e., the condition is in the domain if the condition is forested and has any IPS present. For area and cover

Table 2 Estimates, standard errors $se(\cdot)$, percent sampling errors $\%se(\cdot)$, and 95% confidence intervals (lower bound CI_L, upper bound CI_U) for example calculations. Estimates, standard errors, and confidence interval bounds are in hectares for $\widehat{A\_IPS}_d$, $\Delta\widehat{A\_IPS}_d$, $\widehat{C}_d^*$, and $\Delta\widehat{C}_d^*$; units of proportion otherwise

Attribute	Estimate	$se(\cdot)$	$\%se(\cdot)$	95% CI_L	95% CI_U
$\widehat{A\_IPS}_d$	484,919.2	78,171.9	16.12	331,702.3	638,136.1
$\Delta\widehat{A\_IPS}_d$	−14,267.0	5125.4	35.93	−24,312.9	−4221.2
$\widehat{A\_P}_{dd'}$	0.506	0.075	14.91	0.36	0.65
$\Delta\widehat{A\_P}_{dd'}$	−0.006	0.108	1924.37	−0.22	0.21
$\widehat{C}_d^*$	146,422.6	36,203.5	24.73	75,463.8	217,381.5
$\Delta\widehat{C}_d^*$	5713.3	1374.6	24.06	3019.1	8407.6
$\widehat{C^*\_P}_{dd'}$	0.153	0.085	55.34	−0.01	0.32
$\Delta\widehat{C^*\_P}_{dd'}$	0.009	0.121	1346.74	−0.23	0.25

proportion estimates, the domain *d'* is forest land. The plot-level values created under this construct and used in calculations are provided as Supp. 1. The relevant summary statistics based on these data are given at the post-stratum h, estimation unit g, and state of MD levels of aggregation in Supp. 2. Specifically, the examples include estimations for IPS presence area ($\widehat{A\_IPS}_d$), IPS presence area change ($\Delta\widehat{A\_IPS}_d$), IPS presence area proportion ($\widehat{A\_P}_{dd'}$), IPS presence area proportion change ($\Delta\widehat{A\_P}_{dd'}$), IPS cover area ($\widehat{C}_d^*$), IPS cover area change ($\Delta\widehat{C}_d^*$), IPS cover area proportion ($\widehat{C^*\_P}_{dd'}$), and IPS cover area proportion change ($\Delta\widehat{C^*\_P}_{dd'}$). A summary of the results is given in Table 2.

Discussion

Of particular relevance when interpreting the results of the example calculations are those resulting from the 2019 MD data that represent actual plot observations. Specifically, it was estimated that over 50% ($\widehat{A\_P}_{dd'} = 0.506$) of forest land in MD has an IPS present. However, this estimate is not highly precise as indicated by the 95% confidence interval being (0.36, 0.65). The interpretation is that if many independent inventories of IPS were conducted in MD using the same sample design and number of plots, it would be expected that 95% of those inventories would produce 95% confidence intervals that would contain the true proportion. Similarly, the area proportion of IPS foliar cover suggests about 15% ($\widehat{C^*\_P}_{dd'} = 0.153$) of forest land is affected and the 95% confidence interval is (−0.01, 0.32). The interpretation of the confidence interval remains as described above, but clearly a negative value of IPS cover is not possible. The inclusion of 0 in the confidence interval indicates the estimate of 0.153 is not statistically different from zero (at the 95% confidence level). Thus, it would be expected that some estimates of zero IPS foliar cover would occur if the sampling process and estimation were repeated a large number of times. Nonetheless, for this sample, it is clear that IPS cover is greater than zero in the population because IPS were found on some plots. From a statistical hypothesis testing perspective, where the null hypothesis is $C^*\_P_{dd'} = 0$, the results indicate there is inadequate evidence to reject the null hypothesis based on the sample size used in this study. However, clearly $C^*\_P_{dd'} > 0$ and thus a type II error has occurred (Fairweather, 1991).

The purpose of the synthetic change data was to illustrate change estimation calculations; therefore, results for $\Delta\widehat{A\_IPS}_d, \Delta\widehat{A\_P}_{dd'}, \Delta\widehat{C}_d^*$, and $\Delta\widehat{C^*\_P}_{dd'}$ cannot be taken as actual estimates of change in IPS area or cover in MD forests. The synthetic data were constructed with changes in both IPS and forest area to illustrate the importance of accounting for both when estimating proportional change ($\Delta\widehat{A\_P}_{dd'}$ and $\Delta\widehat{C^*\_P}_{dd'}$). Estimates of change in the proportion of area and cover were relatively small ($\Delta\widehat{A\_P}_{dd'} = -0.006$, $\Delta\widehat{C^*\_P}_{dd'} = 0.009$). Small change estimates are often accompanied by associated $\%se(\cdot)$ that are extraordinarily large; in this case, the standard errors are > 1000% of the estimates. This outcome highlights the difficulty in obtaining precise estimates of change proportions using IPS data when change is at very small levels. However, when the IPS proportional change is small, the estimate becomes less relevant to decision makers and large sampling errors may be inconsequential. If data consumers require precise estimates of change regardless of the magnitude of the estimate, possible remedies under the proposed estimation framework include using a more specific source of auxiliary data for post-stratification that is highly correlated with IPS change (Haakana et al., 2019) or dramatically increasing sample sizes. Alternatively, other types of estimators could be considered but investigation would be needed to understand any potential efficiency gains (Moisen et al., 2022).

The examples provided above comprise a small number of potential information needs specific to relevant managerial or policy questions. Alternative estimates, e.g., those pertaining to a single species, are easily constructed via definition of the domain *d*. In this context, there may be situations where a species of interest has not been found on any sample plots and thus all plots will be zero-valued, resulting in estimates and associated variances that are zero. For estimations of change, some estimated covariances may also be zero. Regardless, practitioners should not be concerned in these cases and continue with the estimation process as described to obtain the desired results.

Interpretation of NFI estimates under the above framework hinges partly on the meaning of "presence" in relation to the FIA plot design (Fig. 1). It is important to note that estimates of total IPS presence area are dependent on the size of the plot in a manner similar to other occupancy estimates used in ecology such as species-area curves (Condit et al., 1996). FIA uses a single stage cluster plot design with a triangular array of four 7.3-m radius circular subplots (Fig. 1). If there is an IPS occurrence on forest land within 7.3 m of the subplot center point, the forest condition area is labeled as "IPS present" and plot level presence area proportion values are generated through aggregation of subplot condition areas having IPS presence. In the IPS occupancy context, the areas within each of the four 7.3-m radius subplots having IPS presence provide the basis for observation. Therefore, it would be appropriate to consider the support area used to define presence as being forest land within 7.3 m of a given location in the study area. This differs from the interpretation of traditional FIA results, which are not directly tied to the plot size used.

Users preferring a different scale of interpretation, such as at the plot scale, should be able to modify the provided equations accordingly. In this case, the area being sampled by the plot would be considered to be the 43.9-m radius circular area that circumscribes the plot (Fig. 1). Plot values would be determined by assessing IPS presence area at the plot scale, i.e., simultaneously across all subplots combined. For example, in Fig. 1, the current approach would assign a plot-level value of 0.25 to IPS presence for each of the species shown; a plot-scale approach would assign a value of 1. This means that with the subplot approach, the furthest distance an IPS can be from any subplot center to be considered present is 7.3 m, whereas with the plot approach, the furthest distance from plot center would be 43.9 m. The choice of a plot-level vs. a subplot-level support region depends upon the desired information needs. For example, if one assumes that IPS impacts on some phenomenon of interest (e.g., tree regeneration) is limited to a small area around the location of that phenomenon, then a subplot support region may be appropriate. Alternatively, if the analyst believes that IPS within the general area of a phenomenon (e.g., IPS colonization risk), then a larger support might be more suitable. Given that the outcomes are directly dependent on the underlying support area for presence classification, data users should be cognizant of the convention being used and how that may affect the interpretation of the results.

Conclusion

Due to the importance of IPS information in land management decision making, it is essential that IPS data consumers are provided with scientifically credible estimates and associated measures of reliability. The documentation of estimation methods pertinent to IPS data collected by FIA provides a general framework that should accommodate information needs such as basic area estimates and those area estimates as proportions of underlying land areas of interest. Methods for estimation of change over time are also included, which may be equally or more important than current status assessments for evaluation of threat increase or persistence. These methods may be applicable to IPS data collected in a similar manner in other forest inventory efforts. There may also be extensions to other inventory data where presence and cover data are commonly collected such as understory vegetation.

Acknowledgements The authors are indebted to the associate editor and anonymous reviewers for their commitment of time and expertise that resulted in valuable suggestions for improvement.

Author contribution Conceptualization: JAW; methodology: JAW and AJL.; formal analysis: JAW; writing—original draft preparation: JAW and AJL in consultation with SAP, TWL, and CMK; writing—review and editing: All authors; All authors have read and agreed to the content of the manuscript.

Funding Not applicable.

Data availability The data used in numeric examples are provided with the manuscript. More generally, invasive plant data associated with the NFI in the USA can be accessed at https://research.fs.usda.gov/products/dataandtools/fia-datamart.

Declarations

Competing interests The authors declare no competing interests.

References

Alberdi, I., Vallejo, R., Álvarez-González, J. G., Condés, S., González-Ferreiro, E., Guerrero, S., Hernández, L., Martínez-Jauregui, M., Montes, F., Oliveira, N., & Pasalodos-Tato, M. (2017). The multi-objective Spanish National Forest Inventory. *Forest Systems, 26*(2), Article e04S. https://doi.org/10.5424/FS/2017262-10577

Arrington, A. (2021). Urban foraging of five non-native plants in NYC: Balancing ecosystem services and invasive species management. *Urban Forestry & Urban Greening, 58*, Article 126896. https://doi.org/10.1016/j.ufug.2020.126896

Bajwa, A. A., Farooq, M., Nawaz, A., Yadav, L., Chauhan, B. S., & Adkins, S. (2019). Impact of invasive plant species on the livelihoods of farming households: Evidence from *Parthenium hysterophorus* invasion in rural Punjab, Pakistan. *Biological Invasions, 21*(11), 3285–3304. https://doi.org/10.1007/s10530-019-02047-0

Bechtold, W. A., & Scott, C. T. (2005). The forest inventory and analysis plot design. In W. A. Bechtold, & P. L. Patterson (Eds.), The enhanced forest inventory and analysis program - national sampling design and estimation procedures. *U.S. Forest Service General Technical Report, SRS-80*, pp. 37–52. https://doi.org/10.2737/SRS-GTR-80

Bindewald, A., Miocic, S., Wedler, A., & Bauhus, J. (2021). Forest inventory-based assessments of the invasion risk of *Pseudotsuga menziesii* (Mirb.) Franco and *Quercus rubra* L. in Germany. *European Journal of Forest Research, 140*(4), 883–899. https://doi.org/10.1007/s10342-021-01373-0

Blossey, B. (1999). Before, during and after: The need for long-term monitoring in invasive plant species management. *Biological Invasions, 1*, 301–311. https://doi.org/10.1023/A:1010084724526

Burrill, E. A., DiTommaso, A. M., Turner, J. A., Pugh, S. A., Christensen, G., Kralicek, K. M., Perry, C. J., Lepine, L. C., Walker, D. M., & Conkling, B. L. (2024). *The Forest Inventory and Analysis Database, FIADB user guides, volume: Database description (version 9.3), nationwide forest inventory (NFI)* (p. 1026). U.S. Forest Service.

Christensen, G. A., Waddell, K. L., Stanton, S. M., & Kuegler, O. (2016). *California's forest resources: Forest Inventory and Analysis, 2001–2010*. U.S. Forest Service General Technical Report. https://doi.org/10.2737/PNW-GTR-913

Cochran, W. G. (1977). *Sampling techniques*. John Wiley & Sons.

Condit, R., Hubbell, S. P., Lafrankie, J. V., Sukumar, R., Manokaran, N., Foster, R. B., & Ashton, P. S. (1996). Species-area and species-individual relationships for tropical trees: A comparison of three 50-ha plots. *Journal of Ecology, 84*(4), 549–562. https://doi.org/10.2307/2261477

Fairweather, P. G. (1991). Statistical power and design requirements for environmental monitoring. *Australian Journal of Marine and Freshwater Research, 42*, 555–567. https://doi.org/10.1071/MF9910555

Gioria, M., Hulme, P. E., Richardson, D. M., & Pyšek, P. (2023). Why are invasive plants successful? *Annual Review of Plant Biology, 74*(1), 635–670. https://doi.org/10.1146/annurev-arplant-070522-071021

Gormanson, D. D., Pugh, S. A., Barnett, C. J., Morin, R. S., Sowers, P. A., & Westfall, J. A. (2023). *Statistics and quality assurance for the Northern Research Station Forest Inventory and Analysis Program, version 3*. U.S. Forest Service General Technical Report. https://doi.org/10.2737/NRS-GTR-210

Haakana, H., Heikkinen, J., Katila, M., & Kangas, A. (2019). Efficiency of post-stratification for a large-scale forest inventory—Case Finnish NFI. *Annals of Forest Science, 76*(1), 1–15. https://doi.org/10.1007/s13595-018-0795-6

Jennings, M. D., Faber-Langendoen, D., Loucks, O. L., Peet, R. K., & Roberts, D. (2009). Standards for associations and alliances of the U.S. National Vegetation Classification. *Ecological Monographs, 79*(2), 173–199. https://doi.org/10.1890/07-1804.1

Jo, I., Bellingham, P. J., Mason, N. W., McCarthy, J. K., Peltzer, D. A., Richardson, S. J., & Wright, E. F. (2024). Disturbance-mediated community characteristics and anthropogenic pressure intensify understory plant invasions in natural forests. *Journal of Ecology, 112*(8), 1856–1871. https://doi.org/10.1111/1365-2745.14367

Kahn, J., Linial, N., & Samorodnitsky, A. (1996). Inclusion-exclusion: Exact and approximate. *Combinatorica, 16*, 465–477. https://doi.org/10.1007/BF01271266

Kaufman, S. R., & Kaufman, W. (2023). *Invasive plants: Guide to identification and the impacts and control of common North American species*. Rowman & Littlefield.

Kurtz, C. M. (2013). *An assessment of invasive plant species monitored by the Northern Research Station Forest Inventory and Analysis Program, 2005 through 2010*. U.S. Forest Service General Technical Report. https://doi.org/10.2737/NRS-GTR-109

Kurtz, C. M. (2023). *An assessment of invasive plant species in northern U.S. forests*. U.S. Forest Service Research Note. https://doi.org/10.2737/NRS-RN-311

Moisen, G. G., Andersen, H. E., Bell, D. M., Coulston, J. W., Frescino, T. S, McConville, K. S., McRoberts, R. E., Patterson, P. L., Westfall, J. A., & Wilson, B. T. (2022). Emerging alternative estimators. In J. A. Westfall, J. W. Coulston, G. G. Moisen, & H. E. Andersen (Eds.), Sampling and estimation documentation for the Enhanced Forest Inventory and Analysis Program: 2022. *U.S. Forest Service General Technical Report, NRS-GTR-207*. pp. 76–101. https://doi.org/10.2737/NRS-GTR-207

Morin, R. S. (Comp.) (2020). Forest ecosystem health indicators. *U.S. Forest Service, FS-1151*. https://research.fs.usda.gov/sites/default/files/2024-07/wo-17827_forest_service_forest_ecosystem_health_indicators_508.pdf. Accessed 15 Dec 2025.

Oswalt, C. M., Fei, S., Guo, Q., Iannone, B. V., III., Oswalt, S. N., Pijanowski, B. C., & Potter, K. M. (2015). A subcontinental view of forest plant invasions. *NeoBiota, 24*, 49–54. https://doi.org/10.3897/neobiota.24.4526

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 W. A. Bechtold, & P. L. Patterson (Eds.), *The enhanced forest inventory and analysis program - national sampling design and estimation procedures. U.S. Forest Service General Technical Report, SRS-80*. pp. 21–36. https://doi.org/10.2737/SRS-GTR-80

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 W. A. Bechtold, & P. L. Patterson (Eds.), *The enhanced forest inventory and analysis program - national sampling design and estimation procedures. U.S. Forest Service General Technical Report, SRS-80*. pp. 53–77. https://doi.org/10.2737/SRS-GTR-80

Scott, T. L. (2010). *Invasive plant medicine: The ecological benefits and healing abilities of invasives*. Healing Arts Press.

Senator, S. A., & Rozenberg, A. G. (2017). Assessment of economic and environmental impact of invasive plant species. *Biology Bulletin Reviews, 7*, 273–278. https://doi.org/10.1134/S2079086417040089

Shen, X., Ge, M., Handel, S. N., Wang, W., Jin, Z., & Kirkwood, N. G. (2023). Advancing environmental design with phytoremediation of brownfield soils using spontaneous invasive plants. *Science of the Total Environment, 883*, Article 163635. https://doi.org/10.1016/j.scitotenv.2023.163635

Simberloff, D. (2009). The role of propagule pressure in biological invasions. *Annual Review of Ecology, Evolution, and Systematics, 40*, 81–102. https://doi.org/10.1146/annurev.ecolsys.110308.120304

Tassin, J., & Kull, C. A. (2015). Facing the broader dimensions of biological invasions. *Land Use Policy, 42*, 165–169. https://doi.org/10.1016/j.landusepol.2014.07.014

U.S. Forest Service. (2024). *Forest Inventory and Analysis national core field guide for the nationwide forest inventory. version 9.4*. https://research.fs.usda.gov/understory/nationwide-forest-inventory-field-guide. Accessed 15 Dec 2025.

Westfall, J. A., Coulston, J. W., Moisen, G. G., & Andersen, H. E. (2022). Sampling and estimation documentation for the Enhanced Forest Inventory and Analysis Program: 2022. *U.S. Forest Service General Technical Report, NRS-GTR-207*. https://doi.org/10.2737/NRS-GTR-207

Westfall, J. A., Patterson, P. L., & Coulston, J. W. (2011). Post-stratified estimation: Within-strata and total sample size recommendations. *Canadian Journal of Forest Research, 41*, 1130–1139. https://doi.org/10.1139/x11-031

All authors have read, understood, and have complied as applicable with the statement on "Ethical responsibilities of Authors" as found in the Instructions for Authors.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.