



# Addressing nonresponse bias in forest inventory change estimation using response homogeneity classifications

James A. Westfall<sup>a,\*</sup>, Mark D. Nelson<sup>b</sup>

<sup>a</sup> U.S. Forest Service, Northern Research Station, York, PA, USA

<sup>b</sup> U.S. Forest Service, Northern Research Station, St. Paul, MN, USA



## ARTICLE INFO

### Keywords:

Disturbance

Post-stratification

Land use conversion

Sample bias

## ABSTRACT

Estimating amounts of change in forest resources over time is a key function of most national forest inventories (NFI). As this information is used broadly for many management and policy purposes, it is imperative that accurate estimations are made from the survey sample. Robust sampling designs are often used to help ensure representation of the population, but often the full sample is unrealized due to hazardous conditions or possibly lack of land access permission. Potentially, bias may be imparted to the sample if the nonresponse is nonrandom with respect to forest characteristics, which becomes more difficult to assess for change estimation methods that require measurements of the same sample plots at two points in time, i.e., remeasurement. To examine potential nonresponse bias in change estimates, two synthetic populations were constructed: 1) a typical NFI population consisting of both forest and nonforest plots, and 2) a population that mimics a large catastrophic disturbance event within a forested population. Comparisons of estimates under various nonresponse scenarios were made using a standard implementation of post-stratified estimation as well as an alternative approach that groups plots having similar response probabilities (response homogeneity). When using the post-stratified estimators, the amount of change was overestimated for the NFI population and was underestimated for the disturbance population, whereas the response homogeneity approach produced nearly unbiased estimates under the assumption of equal response probability within groups. These outcomes suggest that formal strategies may be needed to obtain accurate change estimates in the presence of nonrandom nonresponse.

## 1. Introduction

Although assessments of changes in forest resources based on time series of remotely sensed data are becoming more technically advanced and commonplace (Cohen et al., 2010; Nguyen et al., 2019; Schleeweis et al., 2020; Woodcock et al., 2020), estimates from forest inventory field plots still often serve as the primary information source for official statistics and many analytical purposes (Zhao et al., 2019; Oswalt et al., 2019). A key aspect of many national forest inventory (NFI) programs is the ability to monitor change in numerous forest resource attributes. For example, change in total forest area (possibly by forest type or stand size class) as well as biomass and carbon attributes often rely on NFI data (Charru et al., 2010; Gschwantner et al., 2016). Accurate estimation of change statistics is often vital for making effective management and policy decisions affecting forest resources. Further, establishment of sound change accounting practices is necessary for participation in programs aimed at reducing deforestation and forest degradation (Park

et al., 2013). As NFI often provide the basis for these information needs, statistical rigor in the sampling design is paramount for achieving accurate statistical estimates. Fortunately, there exists considerable literature in this regard due to longstanding implementation of forest inventories in various countries (Bechtold and Patterson, 2005; Tomppo et al., 2011; Zeng et al., 2015). However, it is nearly impossible to avoid nonresponse in sample based NFIs, where no information is obtained for some of the selected sample units primarily due to hazardous conditions or lack of access permission by private landowners (Patterson et al., 2012). In the U.S., there is some evidence for increasing incidence of denied access by private landowners (Westfall et al., 2022a, b), with similar increasing nonresponse by landowners to natural resource-related mail surveys (Connelly and Brown, 2003). Although some research of nonresponse mitigation in forest inventory has been conducted (Fischer et al., 2011; Corona et al., 2014; Fattorini et al., 2013), there is often little operational effort to explicitly address potential biases in the sample due to nonresponse and consequently some bias

\* Corresponding author.

E-mail address: [james.westfall@usda.gov](mailto:james.westfall@usda.gov) (J.A. Westfall).

<https://doi.org/10.1016/j.fecs.2023.100099>

Received 11 October 2022; Received in revised form 1 February 2023; Accepted 13 February 2023

2197-5620/© 2023 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

in the estimates may be present (Goeking and Patterson, 2013; Mag-nussen et al., 2017). Rather, nonresponse is often assumed to have minor effects on estimation bias due to the infrequent occurrence but increasing rates of nonresponse may negate that assumption.

A key area of interest in change estimation is to quantify effects of disturbance on forest resources. Disturbance events occur across a wide range of intensity and spatial scale (Stueve et al., 2011; Wilson et al., 2019; Fahey et al., 2020). While all disturbances have the potential to alter forest ecosystem structure and function, it is often that relatively large catastrophic events are of most concern due to the impact potential (Edgar et al., 2019; Senf and Seidl, 2021). In a forest inventory context, these severe disturbances can be problematic in that resulting hazardous conditions can preclude field crew visits to plots and thus nonresponse is more prevalent in the disturbance zone (Nelson et al., 2009). As most forest inventories employ equal probability sampling designs, the result is that information from the disturbance zone is underrepresented in the sample outcome. Failure to acknowledge and address this phenomenon and proceeding to apply standard estimation procedures will likely produce inaccurate population estimates of the disturbance effect (Nelson et al., 2009).

In this study, the effect of nonresponse outcomes on estimates of change is examined in the context of a typical NFI post-stratified estimation method and an alternative approach that explicitly incorporates knowledge of factors leading to nonresponse. The analysis accounts for two scenarios: 1) typical NFI nonresponse outcomes in a general population, and 2) a specific population in which a large disturbance occurs. Thus, the objectives are to: 1) increase awareness of nonresponse issues in the context of repeated plot measurements, 2) assess the potential magnitude of nonresponse bias that may be occurring in change estimates derived from forest inventory data, and 3) present a potential framework for using response homogeneity classes in a change estimation context when nonresponse does not occur uniformly at randomly in the population.

## 2. Methods

### 2.1. Data

Initial testing of methods was conducted using a synthetic population where the values for a generic forest attribute response variable (mimicking tree biomass in tonnes per hectare) were known for each population element at two points in time. For each population element at each time point, classifications were assigned as either being forest or nonforest land use status. To facilitate use of post-stratified estimation methods used to improve the precision of estimates, three post-strata were created and each population element was assigned to a post-stratum. In practice, maps constructed from classified remote sensing data are often used to create the post-strata and determine the post-stratum weights (McRoberts et al., 2002). Post-stratum weights for the

synthetic population were determined from the proportions of population elements in each post-stratum. The response variable for each forested population element at T1 was generated based on the post-stratum assignment ( $h$ ) and a  $N(\mu_h, \sigma_h)$  normal distribution where the distributions were 1)  $\mu_1 = 25$  and  $\sigma_1 = 5$ , 2)  $\mu_2 = 100$  and  $\sigma_2 = 12.5$ , and 3)  $\mu_3 = 200$  and  $\sigma_3 = 25$  for post-strata  $h = 1$  to 3 (Table 1). The population elements having no forest area had the response value set to zero.

Population element response values at T2 were generated using the following methods regardless of post-stratum: 1) values of the response variable at T2 for forested elements that were nonforest at T1 were obtained randomly from a  $N(50, 10)$  distribution as areas reverting to forest over the measurement interval would likely have relatively small values, 2) a 'growth' factor approach was used for elements that were forest at both T1 and T2, where the T1 value was increased by a randomly selected value between 5% and 10% to obtain the T2 amount, and 3) nonforest elements were always zero-valued by definition (Table 1). The elements forested at T1 that became nonforest at T2 had a negative change equal to the T1 value. As such, these elements that changed status between T1 and T2 tended to have relatively large negative change amounts (Tables 1 and 2). Overall, there was approximately a 2% reduction in the population total between T1 and T2.

A sample of the population elements is considered to be analogous to plots that may be selected in a forest inventory sample. Once the sample is drawn, complete remeasurement sampling is assumed where each

**Table 2**

Synthetic NFI population post-stratum RHG weights ( $W_{hg}$ ), number of population elements ( $N_{hg}$ ), total change ( $T2 - T1$ ), and percentage of office plots (NF) at T1 and T2. Overall values for the post-strata are also given for weight ( $W_h$ ) and population size ( $N_h$ ), total change ( $T2 - T1$ ), and percentage of office plots (NF) at T1 and T2.

| Post-stratum | RHG       | $W_{hg}^a$ | $N_{hg}$ | $T2 - T1$  | Office plots (%) |       |
|--------------|-----------|------------|----------|------------|------------------|-------|
|              |           |            |          |            | T1               | T2    |
| 1            | NF-NF     | 0.665      | 46,561   | 0          | 100.0            | 100.0 |
| 1            | NF-F/F-NF | 0.066      | 4603     | -21,946.8  | 53.8             | 46.2  |
| 1            | F-F       | 0.269      | 18,836   | 35,370.3   | 0                | 0     |
| 1            |           | 0.350      | 70,000   | 13,423.6   | 70.1             | 69.6  |
| 2            | NF-NF     | 0.120      | 11,990   | 0          | 100.0            | 100.0 |
| 2            | NF-F/F-NF | 0.160      | 15,988   | -701,744.9 | 49.7             | 50.3  |
| 2            | F-F       | 0.720      | 72,022   | 540,503.1  | 0                | 0     |
| 2            |           | 0.500      | 100,000  | -161,241.8 | 19.9             | 20.0  |
| 3            | NF-NF     | 0.289      | 8657     | 0          | 100.0            | 100.0 |
| 3            | NF-F/F-NF | 0.138      | 4136     | -368,979.1 | 52.2             | 47.8  |
| 3            | F-F       | 0.574      | 17,207   | 258,423.6  | 0                | 0     |
| 3            |           | 0.150      | 30,000   | -110,555.5 | 36.1             | 35.4  |

<sup>a</sup> The RHG weights shown here reflect the proportion of observations in the stratum population. In practice, the actual RHG weights within strata are estimated from the selected sample.

**Table 1**

Description of synthetic NFI population construction by post-stratum and forest/nonforest population element status at both points in time.

| Post-stratum | Time 1             |                |                          | Time 2             |                |                          |
|--------------|--------------------|----------------|--------------------------|--------------------|----------------|--------------------------|
|              | Population element | Population (%) | Observation distribution | Population element | Population (%) | Observation distribution |
| 1            | Forest             | 10.5           | $N(25, 5)$               | Forest             | 9.4            | $T1 + (5\%-10\%)$        |
|              |                    |                |                          | Nonforest          | 1.1            | 0                        |
| 1            | Nonforest          | 24.5           | 0                        | Forest             | 1.2            | $N(50, 10)$              |
|              |                    |                |                          | Nonforest          | 23.3           | 0                        |
| 2            | Forest             | 40.0           | $N(100, 12.5)$           | Forest             | 36.0           | $T1 + (5\%-10\%)$        |
|              |                    |                |                          | Nonforest          | 4.0            | 0                        |
| 2            | Nonforest          | 10.0           | 0                        | Forest             | 4.0            | $N(50, 10)$              |
|              |                    |                |                          | Nonforest          | 6.0            | 0                        |
| 3            | Forest             | 9.6            | $N(200, 25)$             | Forest             | 8.6            | $T1 + (5\%-10\%)$        |
|              |                    |                |                          | Nonforest          | 1.0            | 0                        |
| 3            | Nonforest          | 5.4            | 0                        | Forest             | 1.1            | $N(50, 10)$              |
|              |                    |                |                          | Nonforest          | 4.3            | 0                        |

sample plot is remeasured periodically (Schreuder et al., 2004). Selected plots were designated for a field visit when forestland was present or are considered as an ‘office’ plot where no field visit was required due to no forest area being present. This corresponds to sampling methodology of the United States NFI, wherein sample plots having zero probability of forest condition (e.g., open water, cropland, developed impervious surface) as determined by office staff via photointerpretation of high resolution aerial imagery, are not assigned for field measurement. As office-determined nonforest plots are not subject to nonresponse, the only plots used in estimation that had zero probability of nonresponse were those that were nonforest office plots at both time points (Note that occasionally a field plot may be nonforest, but it is still subject to potentially being nonresponse). Thus, only the field plots were subject to either partial or complete nonresponse.

To account for different nonresponse probabilities, response homogeneity groups (RHG) were created within each post-stratum (Table 2). In the context of change estimation, plots can be classified in four ways based on their forest/nonforest status at each time point: 1) forested at T1 and nonforest at T2 (F–NF), 2) nonforest at T1 and forest at T2 (NF–F), 3) forest at both times (F–F), and 4) nonforest at both times (NF–NF). Only plots in the latter classification (NF–NF) are assured to have no nonresponse, as these plots would have been nonforest office plots for both measurements where both T1 and T2 forest attribute values are zero and thus the change is also zero. The other three categories would have been subject to nonresponse at either T1 (F–NF), T2 (NF–F), or both (F–F), with the latter being considered the most likely to incur nonresponse due to access requirements at both measurements. Thus, three RHGs are created to account for differing nonresponse probabilities (Särndal et al., 1992; section 15.6.2): 1) no nonresponse possible (NF–NF), 2) nonresponse possible at only one of two measurements (F–NF, NF–F), and 3) potential nonresponse at both measurements (F–F).

A second synthetic population was created to illustrate a specific case where forest disturbance affects nonresponse. In this situation, it is considered that the disturbance creates hazardous conditions such that plots within the disturbance zone are more prone to nonresponse than those elsewhere. In this case, consider a smaller population with 50,000 elements that is composed entirely of forest land at both T1 and T2, e.g., a large publicly owned forest where access to field plots is always granted. Three post-strata are created and the generic response values for population elements at T1 are generated from  $N(\mu_h, \sigma_h)$  distributions where the distributions were  $N(75, 7.5)$ ,  $N(125, 10)$ , and  $N(200, 25)$  for post-strata  $h = 1$  to 3, respectively. As done previously, T2 values were obtained from the T1 values using a randomly selected increase between 5% and 10%. Additionally, an extreme disturbance occurring between T1 and T2 caused losses in the generic response variable on 10,000 population elements that define the disturbance zone. It was assumed the T2 value of the generic response variable was randomly reduced by 80%–100% (i.e., 0–20% remaining) for each population element subject to the disturbance (Table 3), simulating severe damage levels typical of high intensity disturbances such as wildfires, derechos, or hurricanes (Boose et al., 1994). Overall, this structure results in an approximately 12% reduction in the population total from T1 to T2. Under this scenario, RHGs within post-strata are created to differentiate between population elements within and external to the disturbance zone (Table 4). In

**Table 4**

Disturbance population post-stratum RHG weights ( $W_{hg}$ ), number of population elements ( $N_{hg}$ ), and total change (T2 – T1) values. Overall values for the post-strata are also given for weight ( $N_h$ ) and population size ( $N_{hg}$ ), and total change (T2 – T1).

| Post-stratum | RHG    | $W_{hg}^a$ | $N_{hg}$ | T2 – T1  |
|--------------|--------|------------|----------|----------|
| 1            | DIST   | 0.203      | 2026     | –135,728 |
| 1            | UNDIST | 0.797      | 7974     | 44,878   |
| 1            |        | 0.200      | 10,000   | –90,850  |
| 2            | DIST   | 0.120      | 2988     | –334,203 |
| 2            | UNDIST | 0.720      | 12,012   | 112,775  |
| 2            |        | 0.500      | 15,000   | –221,428 |
| 3            | DIST   | 0.289      | 4986     | –885,119 |
| 3            | UNDIST | 0.574      | 20,014   | 299,910  |
| 3            |        | 0.150      | 25,000   | –585,209 |

<sup>a</sup> The RHG weights shown here reflect the proportion of observations in the stratum population. In practice, the actual RHG weights within strata are estimated from the selected sample.

practice, this differentiation would likely be accomplished via map information that delineates the disturbance area.

## 2.2. Analysis

Analytical results were obtained for a typical implementation of post-stratified estimation (Scott et al., 2005) as well as for implementation of RHGs within the post-strata (Goeking and Patterson, 2013; Westfall, 2022). First, the standard FIA post-stratified procedures are outlined, with subsequent extension to accommodate the RHG approach. Initially, plot values to be used in estimation are obtained by adjusting the observed value to account for plots having partial nonresponse within each post-stratum  $h$ .

$$\bar{p}_h = \frac{\sum_{i=1}^{m_h} a_{hi}}{m_h} \quad (1)$$

$$y'_{hi} = \frac{y_{hi}}{\bar{p}_h} \quad (2)$$

where  $a_{hi}$  is the proportion of measured area for plot  $i$  in post-stratum  $h$ ,  $m_h$  is the number of plots (excluding completely nonsampled plots) in post-stratum  $h$ ,  $\bar{p}_h$  is the post-stratum  $h$  adjustment factor,  $y_{hi}$  is the original observed value for plot  $i$  in post-stratum  $h$ , and  $y'_{hi}$  is the adjusted value for plot  $i$  in post-stratum  $h$ . The mean ( $\bar{y}'_h$ ) and estimated variance of the mean ( $v(\bar{y}'_h)$ ) for each post-stratum  $h$  are obtained from simple random sample estimators:

$$\bar{y}'_h = \frac{\sum_{i=1}^{m_h} y'_{hi}}{m_h} \quad (3)$$

$$v(\bar{y}'_h) = \frac{\sum_{i=1}^{m_h} (y'_{hi} - \bar{y}'_h)^2}{m_h(m_h - 1)} \quad (4)$$

In this implementation of post-stratification, the mean of plots that

**Table 3**

Description of synthetic disturbance population construction by post-stratum and disturbance zone at both points in time.

| Time 1       |             |                |                  | Time 2      |                |                           |
|--------------|-------------|----------------|------------------|-------------|----------------|---------------------------|
| Post-stratum | Zone        | Population (%) | Plot observation | Zone        | Population (%) | Plot observation          |
| 1            | Undisturbed | 20.0           | $N(75, 7.5)$     | Disturbed   | 4.1            | T1 + (5%–10%)             |
|              |             |                |                  | Undisturbed | 15.9           | [T1 + (5%–10%)] × (0–20%) |
| 2            | Undisturbed | 30.0           | $N(125, 10)$     | Disturbed   | 6.0            | T1 + (5%–10%)             |
|              |             |                |                  | Undisturbed | 24.0           | [T1 + (5%–10%)] × (0–20%) |
| 3            | Undisturbed | 50.0           | $N(200, 25)$     | Disturbed   | 10.0           | T1 + (5%–10%)             |
|              |             |                |                  | Undisturbed | 40.0           | [T1 + (5%–10%)] × (0–20%) |

are completely nonsampled are, in effect, assumed to be equal to the post-stratum mean calculated from the observed plots. Finally, estimates of the population total and corresponding variance are calculated by combining across post-strata. Equation (5) is used to obtain the estimated population total ( $\hat{y}'$ ) using post-stratum mean  $\bar{y}'_h$  values, post-stratum weights ( $W_h$ ), and the population size ( $N$ ). The estimated variance of the population total ( $v(\hat{y}')$ ) includes both sampling uncertainty and the variability due to random post-stratum sample sizes (Scott et al., 2005).

$$\hat{y}' = N \sum_{h=1}^H W_h \bar{y}'_h \quad (5)$$

$$v(\hat{y}') = \frac{N^2}{m} \left[ \sum_{h=1}^H W_h m_h v(\bar{y}'_h) + \sum_{h=1}^H (1 - W_h) \frac{m_h}{m} v(\bar{y}'_h) \right] \quad (6)$$

where  $m = \sum_{h=1}^H m_h$  is the total number of at least partially observed plots.

Implementation of an RHG estimation framework is accomplished by using the RHGs created within each post-stratum, wherein the nonresponse occurrence is considered to be random within each RHG. From an estimation perspective, Westfall et al. (2022a,b) suggest the use of ratio-to-size estimators within each RHG to avoid potential bias in estimates of variance due to the use of the adjustment factor method described above:

$$\bar{y}_{hg} = \frac{\sum_i^{m_{hg}} y_{hgi}}{\sum_i^{m_{hg}} a_{hgi}} \quad (7)$$

$$s_{hg}^2 = \frac{m_{hg}}{m_{hg} - 1} \frac{\sum_i^{m_{hg}} y_{hgi}^2 - 2\bar{y}_{hg} \sum_i^{m_{hg}} y_{hgi} a_{hgi} + \bar{y}_{hg}^2 \sum_i^{m_{hg}} a_{hgi}^2}{(\sum_i^{m_{hg}} a_{hgi})^2} \quad (8)$$

where  $\bar{y}_{hg}$  is the mean response for RHG  $g$  in post-stratum  $h$ ,  $y_{hgi}$  is the observed value from the  $i$ th plot in RHG  $g$  for post-stratum  $h$ ,  $a_{hgi}$  is the proportion of the plot that was accessible, and  $s_{hg}^2$  is the sample variance in RHG  $g$  for post-stratum  $h$ . Subsequently, RHG need to be combined to estimate means and variances for each post-stratum. This requires weights to be determined for each RHG, which are based on the proportion of plots selected into the sample for each RHG relative to the number of plots selected for post-stratum  $h$  (Goeking and Patterson, 2013).

$$w_{hg} = n_{hg} / n_h \quad (9)$$

where  $w_{hg}$  is the estimated weight for RHG  $g$  in post-stratum  $h$ ,  $n_{hg}$  is the number of plots in RHG  $g$  within post-stratum  $h$ , and  $n_h$  is the number of plots in post-stratum  $h$ . It is important here to recognize that  $n$  refers to the sample sizes that also include completely nonresponse plots; whereas  $m$  notation used earlier excludes plots that were entirely inaccessible. In the context of change estimation, some assumptions must be made in practice to assign nonresponse plots to an RHG for calculation of  $w_{hg}$  - acknowledging that nonresponse occurs when the plot is unobserved at either T1, T2, or both. In the case of the U.S. NFI and office/field plot considerations, it may be reasonable to assume that plots observed at one of the two time periods had the same status when unobserved such that these plots would be classified as either F-F or NF-NF accordingly. When a plot is nonresponse at both T1 and T2, it was probably a field plot both times and is likely to be F-F. Of course, exceptions will occur and it must be accepted that RHG assignments will not always be entirely accurate.

Additionally, unlike  $W_h$ , the  $w_{hg}$  are estimated from the sample and thus RHG weights are an additional source of variability. Estimation of the  $w_{hg}$  can be considered as a special case of double sampling for post-stratification where the first phase sample consists only of points associated with the population elements selected into the sample. Further,

Särndal et al. (1992) propose that variation in nonresponse that would occur among different samples should be included in estimates of uncertainty. In consideration of these factors, it is proposed the estimators presented in Westfall (2022) be used to obtain post-stratum means and variances:

$$\bar{y}_h = \sum_{g=1}^G w_{hg} \bar{y}_{hg} \quad (10)$$

$$v(\bar{y}_h) = \sum_{g=1}^G \left( \frac{n_{hg} - 1}{n_h - 1} \right) \frac{s_{hg}^2}{m_h} + \sum_{g=1}^G \left( \frac{n_{hg} - 1}{n_h - 1} \right) \frac{(1 - w_{hg}) s_{hg}^2}{m_h^2 w_{hg}} + \frac{1}{n_h - 1} \sum_{g=1}^G \frac{n_{hg}}{n_h} (\bar{y}_{hg} - \bar{y}_h)^2 + \sum_{g=1}^G \left( \frac{n_{hg}}{n_h} \right)^2 \frac{1 - \sum_i^{m_{hg}} a_{hgi} / n_{hg}}{\sum_i^{m_{hg}} a_{hgi}} s_{hg}^2 \quad (11)$$

Hereafter, the remainder of the estimation process proceeds as described above in equations (5) and (6) by substituting  $\bar{y}_h$  and  $v(\bar{y}_h)$  in place of  $\bar{y}'_h$  and  $v(\bar{y}'_h)$ , respectively, to obtain RHG-based estimates of the population total  $\hat{y}$  and its variance  $v(\hat{y})$ . Construction of 95% confidence intervals was accomplished using  $\hat{y} \pm 1.96 \sqrt{v(\hat{y})}$  for the typical post-stratified estimates and  $\hat{y} \pm 1.96 \sqrt{v(\hat{y})}$  for RHG estimates.

Comparisons between estimations approaches were based on Monte Carlo simulations. For the generalized change estimation scenario, a sample of 2000 population elements was drawn randomly at each iteration without regard to post-stratum or RHG assignment. This approach produces samples consistent with a post-stratified design where sample sizes within post-strata are random and proportions of sample plots within post-strata often differ from post-strata weights ( $W_h$ ). For the primary change estimation population, mean nonresponse rates for completely inaccessible sample units were 10% (F-F), 5% (F-NF, NF-F), and 0 (NF-NF). Similarly, partial nonresponse was applied on average to 4% of F-F plots, 2% of F-NF/NF-F, and 0 if NF-NF plots where the nonresponse proportion for each plot was randomly determined from a  $U(0, 1)$  uniform distribution. This nonresponse scenario produces roughly 6% of plots selected into the sample being complete nonresponse and 2% of plots having partial nonresponse. Sample units subject these nonresponse rates were chosen randomly at each iteration. The mean of  $R = 20,000$  population total estimates for the typical post-stratified estimates and those where RHGs were employed were compared to the known population total change of  $Y = -258,373.7$ . Similarly, the means of the estimated variances for both estimation approaches ( $v(\hat{y})$  and  $v(\hat{y}')$ ) were assessed against the variance of the  $R = 20,000$  population total estimates.

For the disturbance change scenario, a sample of 500 population elements was obtained at each iteration in the same manner as described above. It was assumed that all sample units outside the disturbance zone were accessible. Four nonresponse outcomes were examined within the disturbance zone, where complete/partial nonresponse rates of 20/2, 40/3, 60/4 and 80/4 percent were examined. Sample units subject to either complete or partial nonresponse were selected randomly at each iteration. The nonresponse is applied to the sample units in the disturbance zone to reflect the hazardous conditions that directly result in nonresponse that otherwise would not be present. As before, comparisons were made between mean estimates of change from the  $R = 20,000$  iterations and the known population change of  $Y = -897,487.0$ . As with the NFI scenario, formula-based variances were compared with the simulation results.

### 3. Results

In the standard NFI change estimation scenario, implementation of the typical post-stratified estimator that assumes nonresponse sample unit observations are on average equal to the post-stratum mean

**Table 5**

Estimates, percent bias, and 95% confidence interval coverage for RHG and standard post-stratified estimation methods for the NFI and disturbance populations.

| Population (CNR/PNR) | RHG method |              |           |              |                     | Standard post-stratification |               |            |               |                     |
|----------------------|------------|--------------|-----------|--------------|---------------------|------------------------------|---------------|------------|---------------|---------------------|
|                      | Estimate   |              | Bias (%)  |              | 95% CI coverage (%) | Estimate                     |               | Bias (%)   |               | 95% CI coverage (%) |
|                      | $\bar{y}$  | $v(\bar{y})$ | $\bar{y}$ | $v(\bar{y})$ |                     | $\bar{y}'$                   | $v(\bar{y}')$ | $\bar{y}'$ | $v(\bar{y}')$ |                     |
| Typical NFI (6/2)    | -257,912.4 | 1.72E + 10   | 0.18      | -0.16        | 94.6                | -321,555.9                   | 1.91E + 10    | -24.45     | -0.55         | 93.7                |
| Disturbance (20/2)   | -896,952.7 | 1.97E + 10   | 0.06      | 0.36         | 94.9                | -639,875.3                   | 1.75E + 10    | 28.70      | 0.27          | 49.9                |
| Disturbance (40/3)   | -896,777.0 | 1.99E + 10   | 0.08      | 1.30         | 94.9                | -365,360.2                   | 1.47E + 10    | 59.29      | 0.97          | 1.6                 |
| Disturbance (60/4)   | -898,699.9 | 2.04E + 10   | -0.14     | 0.81         | 94.9                | -65,387.3                    | 1.09E + 10    | 92.71      | 1.70          | 0                   |
| Disturbance (80/4)   | -894,251.2 | 2.17E + 10   | 0.36      | 0.03         | 94.7                | 255,548.7                    | 5.85E + 09    | 128.47     | -2.13         | 0                   |

produced an estimated change that was -63182.2 (-24.4%) different than the known population change (Table 5). In comparison, the RHG estimator resulted in an estimate that differed from the known value by only -461.3 (0.18%). The formula-based estimates of variance were consistent with the variance of the simulations estimates for both estimators.

For change estimation under a large disturbance scenario where nonresponse is caused by the disturbance, standard post-stratified estimation produced biases in estimates that were dependent on the amounts of nonresponse present. When 20% of disturbance plots were complete nonresponse and 2% of disturbance plots had partial nonresponse (denoted 20/2), the estimated change was 28.7% smaller than the known population total (Table 5). Under the 40/3 nonresponse scenario, the estimate was 59.3% too small, the 60/4 nonresponse rate produced an underestimation of 92.7%, and 80/4 nonresponse was 128.5% too small. In contrast, the differences between RHG estimates and known population totals were approximately 0.1% for nonresponse scenarios 20/2, 40/3, and 60/4, respectively, and nearly 0.4% for the 80/4 nonresponse (Table 5).

From a more practical perspective, users often rely on a confidence interval to provide an indication of estimate reliability. Typical interpretation of a 95% confidence interval would be that in repeated sampling of the population, 95% of the confidence intervals would contain the true population value. However, biased estimates can cause disruption to the theoretical construct due to the shift in confidence interval central location. As the estimates from the RHG approach for the disturbance population were largely unbiased, the confidence interval coverage of the true value was consistently near 95%. However, post-stratified estimates of the disturbance suffered poor coverage, with only about 50% coverage for the 20/2 nonresponse situation and near zero coverage for more intense nonresponse outcomes (Table 5).

#### 4. Discussion

The results for the NFI population indicated the RHG method produced estimates of change in the response variable that were essentially unbiased. However, there was a tendency for nonresponse to cause overestimation (more change was estimated than occurred in the population) of change in the NFI population when using post-stratified estimation. In this case, the preponderance of nonresponse on F/F plots that generally have positive change values shifts the results more towards NF/NF (zero change) and the group of status change (NF/F and F/NF) plots which had an overall tendency to negative change (Table 2). Other population change scenarios may provide different results. For example, if the population change was positively valued and status change plots had smaller change than the F/F plots then the change estimate would be too small (underestimation). Outcomes would also be dependent on the proportions of plot types within the RHG and on both the post-stratum and RHG weights. Temporal correlations in plot nonresponse may also interact with population change dynamics to affect estimation results. Thus, the population used here may be considered one of numerous potential NFI scenarios and inventory practitioners are encouraged to make evaluations in the context of their specific inventory circumstances. Nonetheless, potentially adverse effects of nonrandom nonresponse on

population change estimates using the typical post-stratified estimator were demonstrated. This outcome is troubling as inaccurate estimates of change that are too large are more likely to be judged as being of consequential practical/statistical importance with subsequent implementation or continuation of activities and policies that may actually not be achieving desired outcomes. Similarly, underestimation of change may provide indications of status quo in the population and thus no responses to alter trajectories are needed (Field et al., 2004; Legg and Nagy, 2006).

Further, it may be worthwhile for those desiring to estimate disturbance impacts to consider the effects disturbance can have on nonresponse patterns and amounts (Nelson et al., 2009; Smith and Gray, 2021). As with the NFI population, estimates based on the RHG approach for the disturbance population were essentially unbiased at all levels of nonresponse studied (Table 5). In standard post-stratified estimation, differences in response probabilities between disturbed and undisturbed areas produced substantial underestimation of disturbance amounts. The results pertaining to the population studied here suggest that the disturbance effect is underestimated by nearly 30% under a 20/2 nonresponse scenario, with even further underestimation associated with higher levels of nonresponse. In extreme cases such as 80/4 nonresponse occurrence, the bias was severe enough to produce a positive estimate of change when the true population value was negative. Thus, catastrophic disturbances that cause increased nonresponse in the disturbed area can produce very misleading results under a typical implementation of post-stratified estimation. This suggests management and policy reactions to disturbance outcomes may be inadequate as the severity of impacts is underestimated in proportion to the amount of nonresponse. Failure to recognize the full extent of disturbance impacts may have long-term detrimental socio-economic consequences as well as inhibiting effective land management practices (De Grandpré et al., 2018; Montagné-Huck and Brunette, 2018).

The primary issue that affects the standard post-stratified estimation approach is a combination of post-stratum heterogeneity and the assumption that completely inaccessible sample units are reflective of the mean based on observed plots within the post-stratum. For example, in post-stratum 1 under a typical NFI change scenario (Table 2), there are nearly 90% of population elements that are NF-NF and therefore: 1) have a change value of zero, and 2) will never be subjected to nonresponse. Thus, the post-stratum total (or mean) change will tend to be relatively small due to the large number of zeros. In contrast, other sample elements in the post-stratum may have relatively large change values, e.g., F-NF, and also are potential nonresponse candidates. When the post-stratum mean is small but sample units with relatively large change values are nonresponse, it is incongruous that the post-stratum mean and nonresponse sample unit values will be of similar magnitude. For estimates of current status, it is often the case that post-stratum means are underestimated in the presence of nonresponse (Westfall, 2022). In the context of change estimation, the phenomenon is more complex due to the presence of both positive and negative values which may result in either overestimation or underestimation of post-stratum means depending on how nonresponse transpires in a given sample. Nonetheless, bias in the sample often arises from nonresponse and mitigation via RHG or other methods (Magnussen et al., 2017; Westfall and Wilson, 2022) can reduce

negative impacts on population change estimates.

Estimated variances for both estimation approaches performed well with respect to the simulation variance, however the RHG approach generally had smaller differences than the standard post-stratified method. Astute readers may have noticed the estimated variance was larger using the post-stratified estimators in comparison to RHG estimation for the NFI population but was smaller in the analysis of the disturbance population. Examination of the results using the NFI population suggested that creation of post-stratum variances from RHG variances (Eq. (11)) produced smaller values than directly estimating post-stratum variances (Eq. (4)). Conceptually, this is intuitive as RHG essentially post-stratify the original post-strata into more homogenous assemblages – which is the underlying premise of variance reduction via (post-) stratified estimation (Cochran, 1977). This effect translated into overall variances being smaller for RHG estimation when combining variances across post-strata (Eq. (6)).

Conversely, for the disturbance population the means of the DIST and UNDIST RHG (Table 4) are substantially different than the overall post-stratum mean. When the post-stratum variance is calculated from RHG variances, the third term of Eq. (11) becomes relatively large and contributes to larger values than are realized using the standard post-stratified approach (Eq. (4)). In addition, the 4th term of Eq. (11) makes a larger contribution as the amount of nonresponse increases, e.g., the 4th term contributes an additional 8% to the total variance under the 80/4 nonresponse scenario. Another phenomenon that occurs with the disturbance population is that as nonresponse amounts increase in the disturbed zone, the sample becomes more homogeneous (mostly undisturbed plots) and this situation favors smaller variances using post-stratified estimators; whereas the RHG approach maintains the DIST/UNDIST distinction mentioned previously that produces higher variances. Särndal and Lundström (2005) provide a useful perspective on the situation by stating that ‘... if an estimator is greatly biased, it is poor consolation that its variance is low’.

Within each post-stratum, various factors (e.g., ownership) can cause nonrandom nonresponse among sample units. The purpose of the RHG method is to further partition sample units such that the assumption of nonresponse occurring randomly and with equal probability is better satisfied. The simulations applied nonresponse under this assumption and thus robust agreement was obtained between population estimates and known values. It should be acknowledged that, while the RHGs used in this study are intuitive with regards to basic nonresponse tenets, in practice the construction of RHGs for nonresponse in forest inventories such as NFIs may include various and possibly interrelated factors that are difficult to identify. Thus, formulation of RHG where plots therein truly have (approximately) equal response probabilities may be a challenging endeavor. First, identification of the underlying nonresponse mechanisms may prove elusive without further focused investigation (Gao et al., 2020). For some NFIs, further refinements that account for different nonresponse probabilities between public and private land ownership or other known accessibility factors may be appropriate if not already included in development of the post-strata. Also, methodological differences in forest (F) and nonforest (NF) plot status delineation and plot vs. stratum definition (e.g., land use vs. land cover) may alter the distribution of plots requiring a field visit or introduce class mismatches between plots and their assigned RHG. Another factor not explicitly addressed within this study is RHG construction to minimize bias that may occur if nonresponse plots tend to cluster spatially within a certain areal subdomain of the population. Subsequently, a large number of combinations of relevant factors may preclude full accommodation into an RHG structure due to small sample sizes in certain assemblages and may result in corresponding sub-optimal remediation of nonresponse bias effects on population estimates.

The RHG methodology presented in the paper represents one potential way of addressing nonresponse bias in a forest inventory context. Other methods include ignoring missing values or employing different types of weighting schemes or imputation approaches, for which

Schreuder et al. (1993) provide helpful coverage in forest inventory applications. Schreuder et al. (1993) also reference the Kalton and Kasprzyk (1986) publication which provides an excellent overview of weighting and imputation methods. In the context of this paper, an approach that considers weighting based on plot nonresponse probabilities may be considered. However, a remaining research need is to assess these different approaches in practical forest inventory settings. It is likely that optimal solutions may vary due to a number of factors, including the ability to mitigate nonresponse bias, effects on the standard errors of estimates, operational cost and feasibility, and complexity of implementation. Nonetheless, a broader understanding of the advantages and disadvantages of various missing data compensation methods would provide practitioners a framework for identification and implementation of a suitable solution pertinent to their specific circumstances.

## 5. Conclusion

Accounting for potential nonresponse bias in estimates of change is more complex than for current status evaluations, primarily due to the need to assess nonresponse outcomes and underlying mechanisms at two points in time. In this study, nonresponse was based on plots requiring a field visit at either measurement time with commonly reported mechanisms of hazardous conditions or lack of access permission being assumed. Depending on the structure of change in the population and nonresponse severity, substantial differences between the post-stratified estimate and the true value can occur. For the forest inventory change scenarios presented, this study demonstrated that simply relying on post-stratum means to represent the average value of nonresponse plots is often an inadequate strategy for mitigating nonresponse bias. As an alternative, identification of nonresponse causal factors allows sample plots to be assigned to RHG that provide better representation of nonresponse plot values via group means from observed plots. In most cases, the plot information needed for placing plots into RHG based on factors such as office vs field plot and public vs private ownership already exists in the inventory database or can be otherwise easily obtained, e.g., via GIS map layers and plot location information. Accounting for more complex causal factors (e.g., ecological) may be considerably more difficult. In that context, nonresponse will remain nonrandom with respect to any underlying factors not specifically addressed in RHG formation. Nonetheless, the RHG change estimation framework described herein provides a potential pathway for forest inventory practitioners to reduce nonresponse effects on forest resource change estimates.

## Author contributions

James A. Westfall: Conceptualization, Formal Analysis, Writing-Original draft preparation. Mark D. Nelson: Conceptualization, Writing-Reviewing and Editing.

## Data availability

The data used in this study are available from the author upon request.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- Bechtold, W.A., Patterson, P.L., 2005. The enhanced forest inventory and analysis program - national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Forest Service, Southern Research Station, Asheville, NC.
- Boose, E.R., Foster, D.R., Fluet, M., 1994. Hurricane impacts to tropical and temperate forest landscapes. *Ecol. Monogr.* 64 (4), 369–400.

- Charru, M., Seynave, I., Morneau, F., Bontemps, J.D., 2010. Recent changes in forest productivity: an analysis of national forest inventory data for common beech (*Fagus sylvatica* L.) in north-eastern France. *For. Ecol. Manag.* 260 (5), 864–874.
- Cochran, W.G., 1977. *Sampling Techniques*, 3rd edn. John Wiley & Sons, New York.
- Cohen, W.B., Yang, Z., Kennedy, R., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — tools for calibration and validation. *Remote Sens. Environ.* 114 (12), 2911–2924.
- Connelly, N., Brown, T., 2003. Factors affecting response rates to natural resource - focused mail surveys: empirical evidence of declining rates over time. *Soc. Nat. Resour.* 16, 541–549.
- Corona, P., Chirici, G., Franceschi, S., Maffei, D., Marcheselli, M., Pisani, C., Fattorini, L., 2014. Design-based treatment of missing data in forest inventories using canopy heights from aerial laser scanning. *Can. J. For. Res.* 44 (8), 892–902.
- De Grandpré, L., Waldron, K., Bouchard, M., Gauthier, S., Beaudet, M., Ruel, J.C., Hébert, C., Kneeshaw, D.D., 2018. Incorporating insect and wind disturbances in a natural disturbance-based management framework for the boreal forest. *Forests* 9 (8), 471.
- Edgar, C.B., Westfall, J.A., Klockow, P.A., Vogel, J.G., Moore, G.W., 2019. Interpreting effects of multiple, large-scale disturbances using national forest inventory data: a case study of standing dead trees in east Texas, USA. *For. Ecol. Manag.* 437, 27–40.
- Fahey, R.T., Atkins, J.W., Campbell, J.L., Rustad, L.E., Duffy, M., Driscoll, C.T., Fahey, T.J., Schaberg, P.G., 2020. Effects of an experimental ice storm on forest canopy structure. *Can. J. For. Res.* 50 (2), 136–145.
- Fattorini, L., Franceschi, S., Maffei, D., 2013. Design-based treatment of unit nonresponse in environmental surveys using calibration weighting. *Biomed. J.* 55 (6), 925–943.
- Field, S.A., Tyre, A.J., Jonzén, N., Rhodes, J.R., Possingham, H.P., 2004. Minimizing the cost of environmental management decisions by optimizing statistical thresholds. *Ecol. Lett.* 7 (8), 669–675.
- Fischer, C., Klein, C., Fehrmann, L., Fuchs, H., Panferov, O., 2011. A national level forest resource assessment for Burkina Faso—A field based forest inventory in a semiarid environment combining small sample size with large observation plots. *For. Ecol. Manag.* 262 (8), 1532–1540.
- Gao, Y., Huntsinger, L., White, E.M., Fried, J.S., 2020. Encouraging landowner participation in the inventory: barriers and possible solutions. In: Brandeis, T.J., comp (Eds.), *Proceedings of the 2019 Forest Inventory and Analysis (FIA) Science Stakeholder Meeting*, vol. 256. U.S. Forest Service, e-Gen. Tech. Rep. SRS-, p. 243.
- Goeking, S.A., Patterson, P.L., 2013. Stratifying to reduce bias caused by high nonresponse rates: a case study from New Mexico's Forest Inventory.. *Res. Note RMRS-RN-59* USDA, U.S. Forest Service, Rocky Mountain Research Station, Fort Collins, CO.
- Gschwantner, T., Lanz, A., Vidal, C., Bosela, M., Di Cosmo, L., Fridman, J., Gasparini, P., Kuliešis, A., Tomter, S., Schadauer, K., 2016. Comparison of methods used in European National Forest Inventories for the estimation of volume increment: towards harmonisation. *Ann. For. Sci.* 73 (4), 807–821.
- Kalton, G., Kasprzyk, D., 1986. The treatment of survey missing data. *Surv. Methodol.* 12 (1), 1–16.
- Legg, C.J., Nagy, L., 2006. Why most conservation monitoring is, but need not be, a waste of time. *J. Environ. Manag.* 78 (2), 194–199.
- Magnussen, S., Stinson, G., Boudewyn, P., 2017. Updating Canada's National Forest Inventory with multiple imputations of missing contemporary data. *For. Chron.* 93 (3), 213–225.
- McRoberts, R.E., Wendt, D.G., Nelson, M.D., Hansen, M.H., 2002. Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. *Remote Sens. Environ.* 81 (1), 36–44.
- Montagné-Huck, C., Brunette, M., 2018. Economic analysis of natural forest disturbances: a century of research. *J. For. Econ.* 32, 42–71.
- Nelson, M.D., Healey, S.P., Moser, W.K., Hansen, M.H., 2009. Combining satellite imagery with forest inventory data to assess damage severity following a major blowdown event in northern Minnesota, USA. *Int. J. Rem. Sens.* 30 (19), 5089–5108.
- Nguyen, T.H., Jones, S., Soto-Berelov, M., Haywood, A., Hislop, S., 2019. Landsat time-series for estimating forest aboveground biomass and its dynamics across space and time: a review. *Rem. Sens.* 12 (1), 98.
- Oswalt, S.N., Smith, W.B., Miles, P.D., Pugh, S.A., 2019. *Forest Resources of the United States, 2017: A Technical Document Supporting the Forest Service 2020 RPA Assessment*. Gen. Tech. Rep. WO-97. U.S. Forest Service, Washington Office, Washington, DC.
- Park, M.S., Choi, E.S., Youn, Y.C., 2013. REDD+ as an international cooperation strategy under the global climate change regime. *For. Sci. Tech.* 9 (4), 213–224.
- Patterson, P.L., Coulston, J.W., Roesch, F.A., Westfall, J.A., Hill, A.D., 2012. A primer for nonresponse in the US forest inventory and analysis program. *Environ. Monit. Assess.* 184 (3), 1423–1433.
- Särndal, C.-E., Swensson, B., Wretman, J., 1992. *Model Assisted Survey Sampling*. Springer-Verlag, New York.
- Särndal, C.E., Lundström, S., 2005. *Estimation in Surveys with Nonresponse*. John Wiley & Sons, Chichester.
- Schleeweis, K., Moisen, G., Schroeder, T., Toney, C., Freeman, E., Goward, S., Huang, C., Dungan, J., 2020. US national maps attributing forest change: 1986–2010. *Forests* 11, 653.
- Schreuder, H., Gregoire, T., Wood, G., 1993. *Sampling Methods for Multiresource Forest Inventory*. John Wiley & Sons, New York.
- Schreuder, H.T., Ernst, R., Ramirez-Maldonado, H., 2004. *Statistical Techniques for Sampling and Monitoring Natural Resources*. Gen. Tech. Rep. RMRS-GTR-126. U.S. Forest Service, Rocky Mountain Research Station, Fort Collins, CO.
- 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*. U.S. Forest Service, Southern Research Station, Gen. Tech. Rep. SRS-80, pp. 53–77.
- Senf, C., Seidl, R., 2021. Persistent impacts of the 2018 drought on forest disturbance regimes in Europe. *Biogeosciences* 18 (18), 5223–5230.
- Smith, R.J., Gray, A.N., 2021. Strategic monitoring informs wilderness management and socioecological benefits. *Conserv. Sci. Pract.* 3 (9), e482.
- Stueve, K.M., Perry, C.H., Nelson, M.D., Healey, S.P., Hill, A.D., Moisen, G.G., Cohen, W.B., Gormanson, D.D., Huang, C., 2011. Ecological importance of intermediate windstorms rivals large, infrequent disturbances in the northern Great Lakes. *Ecosphere* 2 (1), 1–21.
- Tomppo, E., Heikkinen, J., Henttonen, H.M., Ihalainen, A., Katila, M., Mäkelä, H., Tuomainen, T., Vainikainen, N., 2011. *Designing and Conducting a Forest Inventory-Case: 9th National Forest Inventory of Finland*, vol. 22. Springer Science & Business Media, Dordrecht.
- Westfall, J.A., 2022. An estimation method to reduce complete and partial nonresponse bias in forest inventory. *Eur. J. For. Res.* 141, 901–907.
- Westfall, J.A., Lister, A.J., Scott, C.T., 2022a. Evaluation of mapped-plot variance estimators across a range of partial nonresponse in a post-stratified national forest inventory. *Can. J. For. Res.* 52 (2), 280–285.
- Westfall, J.A., Wilson, B.T., 2022. Nonresponse bias in change estimation: a national forest inventory example. *Forestry* 95 (3), 301–311.
- Westfall, J.A., Schroeder, T.A., McCollum, J.M., Patterson, P.L., 2022b. A spatial and temporal assessment of nonresponse in the national forest inventory of the US. *Environ. Monit. Assess.* 194, 530.
- Wilson, D.C., Morin, R.S., Frelich, L.E., Ek, A.R., 2019. Monitoring disturbance intervals in forests: a case study of increasing forest disturbance in Minnesota. *Ann. For. Sci.* 76, 78.
- Woodcock, C.E., Loveland, T.R., Herold, M., Bauer, M.E., 2020. Transitioning from change detection to monitoring with remote sensing: a paradigm shift. *Remote Sens. Environ.* 238, 111558.
- Zeng, W., Tomppo, E., Healey, S.P., von Gadow, K., 2015. The national forest inventory in China: history-results-international context. *For. Ecosyst.* 2, 23.
- Zhao, M., Yang, J., Zhao, N., Liu, Y., Wang, Y., Wilson, J.P., Yue, T., 2019. Estimation of China's forest stand biomass carbon sequestration based on the continuous biomass expansion factor model and seven forest inventories from 1977 to 2013. *For. Ecol. Manag.* 448, 528–534.