

Diseños de muestreo de las Evaluaciones Forestales Nacionales

Ronald E. McRoberts¹, Erkki O. Tomppo² y Raymond L. Czaplewski ³

EN ESTE CAPÍTULO, APRENDEREMOS LO SIGUIENTE:

- Cómo definir una población con miras a realizar un muestreo
- Cómo seleccionar el tamaño y la forma de una configuración de parcelas
- Cómo distinguir entre diseños de muestreo aleatorizado simple, sistemático, estratificado y de muestreo en grupo
- Métodos para elaborar un diseño de muestreo
- Cómo calcular medias y varianzas de poblaciones
- Cómo estimar los errores de muestreo
- Consideraciones específicas para los inventarios forestales en áreas tropicales

Resumen

Las Evaluaciones Forestales Nacionales (EFN) se realizan con una mayor corrección si se dispone de cálculos acerca de los atributos forestales suficientemente precisos y defendibles desde un punto de vista científico. Este apartado aborda el diseño estadístico del plan de muestreo de un inventario forestal, incluyendo el proceso utilizado a la hora de definir la población que se va a muestrear así como la selección de muestra que satisfará los requisitos de precisión de la EFN. Por este motivo, un estadista con experiencia debería formar parte de cualquier equipo encargado del diseño de un inventario forestal nacional. Sin embargo, si esto no fuera posible, este apartado proporcionará líneas de actuación y recomendaciones para elaborar diseños de muestreo relativamente simples que reduzcan el riesgo a la vez que aumenten las posibilidades de éxito.

1. Introducción

El diseño de muestreo que va a ser la base del

programa técnico de una Evaluación Forestal Nacional (EFN) requiere de un planteamiento teórico que debe implementarse sobre el terreno (véase *Implantación de una EFN*). Entender los conceptos básicos relacionados con el diseño estadístico y los métodos de estimación es uno de los componentes del proceso general de la Gestión de la información y registro de datos de las Evaluaciones Forestales Nacionales

OBJETIVOS

El propósito es calcular el estado de los bosques de un país utilizando los datos recopilados a partir de una muestra de parcelas de campo. Los objetivos básicos de una EFN son cuatro: (1) obtener estimaciones o cálculos nacionales del área forestal total, subdividida según las principales categorías de condiciones y tipos de bosques; cantidades y distribuciones del arbolado por especies y categorías de tamaño; datos de volumen de madera según las características de los árboles; productos forestales no madereros; estimaciones del cambio en estos atributos de los bosques e indicadores de biodiversidad; (2) obtener

¹ Northern Research Station, EE. UU. Forest Service, 1992 Folwell Avenue, Saint Paul, Minnesota 55108

EE. UU.

² The Finnish Forest Research Institute, PO Box 18. FI-01301 Vantaa, Finlandia

1

³ Northern Research Station, EE. UU. Forest Service, 1992 Folwell Avenue, Saint Paul, Minnesota 55108
EE. UU.

estimaciones lo suficientemente precisas para regiones geográficas seleccionadas, como la nación, áreas subnacionales, provincias o regiones y municipios; (3) recopilar suficientes tipos y cantidades de información para satisfacer los requisitos internacionales sobre informes; y finalmente, (4) lograr una solución de compromiso aceptable entre costes, precisión y resolución geográfica de las estimaciones. Véase el apartado Variables que, generalmente, intervienen en los Inventarios Forestales Nacionales.

SUPOSICIONES Y SIMPLIFICACIÓN DE LIMITACIONES

Varias son las suposiciones que sustentan el tema que se trata a continuación. En primer lugar, se asume que no es posible contar con estadistas expertos con conocimientos del diseño de inventarios de recursos naturales y del análisis de datos. En segundo lugar, se asume que es posible contar con datos auxiliares, como mapas que muestran información de regiones ecológicas, cubiertas de terreno, tipos de suelo, elevación, fronteras políticas y administrativas y sistemas de transporte. En tercer lugar, se asume que es posible contar con modelos para la predicción de atributos, como los volúmenes de árboles individuales obtenidos a partir de mediciones de árboles básicas. Incluso teniendo en cuenta estas suposiciones, el debate acerca de todas las posibilidades que ofrece el diseño de muestreo de una EFN va más allá del objetivo de este apartado. Por lo tanto, se establecen tres puntos que delimitan dicha discusión. Por un lado, el debate se limita a un tema relativamente simple, los diseños polivalentes que pueden emplearse de forma fiable únicamente si se cuenta con la colaboración de un experto local. Por otro, el debate se limita a aquellos diseños que son flexibles y que, por lo tanto, reducen el riesgo de sesgo así como de pérdida de credibilidad. Y en tercer lugar y último, el debate se limita a los diseños que cuentan con muestras de igual probabilidad o, en el caso de los diseños estratificados, muestras de igual probabilidad

por estratos.

¿POR QUÉ REALIZAR EL MUESTREO?

La descripción más precisa de una población es el resultado de mediciones detalladas de todos y cada uno de sus miembros, lo que comúnmente se conoce como censo. Sin embargo, realizar un censo de este tipo suele ser imposible debido a su alto coste y a los problemas logísticos que lleva asociados. Imagine tratar de medir todos y cada uno de los árboles de una extensión forestal de un millón de hectáreas. Una muestra mide una parte de la población que, en ingeniería forestal, suele ser bastante pequeña. Los cálculos basados en los datos recopilados a partir de una muestra medida se extrapolan al conjunto de la población, cuya mayor parte no se ha analizado.

Se piensa en ello como una “suposición” o un “cálculo” del estado de una población basado en la muestra de unos pocos miembros de dicha población. Si la muestra es representativa del total de la población, la estimación será precisa y será menos probable que se desvíe del valor real de la población. De lo contrario, los cálculos serán imprecisos y engañosos, esta falta de precisión no podrá detectarse a simple vista y la verdadera precisión del cálculo no se sabrá porque el estado real del conjunto de la población no se conoce. Lo más adecuado en estos casos es aumentar las posibilidades de medir una muestra representativa. Esto se lleva a cabo haciendo uso de reglas, defendibles desde un punto de vista científico, para seleccionar la muestra, maximizar el número de unidades de muestra observadas o medidas y minimizar los errores de medición de cada una de ellas (véase Calidad de los datos). No resulta difícil obtener los datos. Es mucho más complicado conseguir datos precisos y fiables que se utilizarán a la hora de tomar decisiones importantes.

DEFINIR LA POBLACIÓN

El cálculo, defendible desde un punto de vista científico, de los atributos de una población se basa en un conjunto de teorías matemáticas que se deben respetar si se va a utilizar para

defender la precisión de los cálculos basados en muestras. La cuidadosa selección de un marco de muestreo, de la configuración de la parcela y del diseño de muestreo que se van a emplear son pasos cruciales del proceso que no deben llevarse a cabo de forma independiente ya que cada una de las decisiones influye en las otras. La teoría matemática comienza con una definición precisa de la población cuyos atributos se van a medir. Por ejemplo, para una localidad que cuenta con cinco millones de hectáreas, de las cuales un millón de ellas es zona forestal, la población estadística debería describirse de diferentes formas, aunque siempre de una manera lógica:

- Cientos de formaciones forestales y polígonos no forestales
- Decenas de millones de parcelas de muestreo potenciales de 0,1 hectáreas
- Diez millones de píxeles de 30 x 30 m detectados de forma remota
- Miles de millones de árboles
- Número infinito de puntos

Véase el apartado Unidades de observación para obtener más información.

En los inventarios forestales, no existe una definición de población perfecta. El problema clave en las aplicaciones básicas de un muestreo forestal es definir de una forma precisa las fronteras geográficas de la población objetivo, como pueden ser los terrenos, ya sean forestales o no forestales, de un país que se encuentra fuera de las fronteras geopolíticas de las zonas urbanas. Es frecuente encontrarse con el hecho de que determinadas partes de la población objetivo no se pueden medir. Por ejemplo, aquellas zonas remotas e inaccesibles o en las que el acceso no es seguro. Estas áreas deberían identificarse claramente en un mapa, incluso si las fronteras reales no están claramente definidas, y excluirse de la población que se va a medir. Los cálculos defendibles desde un punto de vista científico deben limitarse exclusivamente a la población medida.

ELEGIR UN MARCO DE MUESTREO

Se distinguen tres términos: marco de

muestreo, diseño de muestreo y configuración de parcela. Marco de muestreo: conjunto de todas las unidades de muestra posibles; diseño de muestreo: subconjunto de unidades de muestra que representan a la población; configuración de parcela: tamaño, forma y componentes de la parcela de campo.

El marco de muestreo ofrece otras ventajas, ya que considera un bosque como una población de puntos infinitos. Un enfoque que se emplea en este marco de muestreo es la popular parcela Bitterlich, que resulta muy eficaz a la hora de calcular las variables relacionadas con el tamaño de los árboles. Las configuraciones de parcela alternativas basadas en puntos miden una determinada zona e imputan sus atributos a un punto concreto. Cuando un punto se encuentra cerca de un límite o de una valla, es más fácil asignarlo a un lado o a otro, mientras que las parcelas con diferentes diseños pueden extenderse más allá de vallas o límites. Se recomienda considerar la población forestal como un conjunto infinito de puntos, así como utilizar las mediciones físicas de una determinada zona para describir el estado en un punto de la muestra.

ELEGIR UNA CONFIGURACIÓN DE PARCELA

La configuración de la parcela viene definida por el tamaño y la forma y determina las variables que se van a medir en cada ubicación de la parcela de muestra. En la elección de configuraciones de parcela se incluyen las parcelas de área variable, las parcelas de área fija, las subdivisiones de parcelas en subparcelas y las parcelas agrupadas, cuyos tamaños y formas deben tenerse en consideración. Las parcelas de área variable que utilizan el muestreo Bitterlich son particularmente efectivas a la hora de obtener cálculos precisos de los atributos forestales relacionados con el tamaño de los árboles. Las parcelas de área fija, aunque no tienen por qué ser las más adecuadas para medir determinados atributos forestales, son un elemento importante cuando el objetivo del muestreo es obtener cálculos acerca de una gran variedad de atributos forestales y

tienden a ser más compatibles con los datos auxiliares. El muestreo agrupado reduce el desplazamiento entre parcelas a la vez que ofrece un número de parcelas suficiente. El tamaño y la forma óptimos se pueden determinar utilizando una simulación de muestreo y antes de obtener la información, aunque las parcelas circulares a menudo se emplean en los inventarios forestales.

Los problemas relacionados con la selección de una determinada configuración de parcela se describen en los apartados Unidades de observación y Optimización del diseño de parcela.

MEDIR LAS PARCELAS DE MUESTRA

El apartado Observación y medición en Evaluaciones Forestales Nacionales resume las consideraciones más importantes a tener en cuenta a la hora de medir las parcelas de muestra. Para obtener información más detallada, véase la referencia en línea Statistical Techniques for Sampling and Monitoring Natural Resource (Técnicas estadísticas para muestrear y supervisar los recursos naturales) (Schreuder y otros, 2004). En este apartado se resaltan dos aspectos de este problema: el uso de datos detectados de forma remota para medir parcelas y las parcelas temporales frente a las parcelas permanentes.

En primer lugar, los datos detectados de forma remota gracias a satélites de resolución media y fotografías aéreas a gran altitud (escalas de 1:24.000 a 1:60.000) ofrecen mediciones más económicas acerca de los indicadores generales del estado forestal, en su mayoría cambios en el área forestal. Sin embargo, la mayoría de las mediciones realizadas de condiciones forestales detalladas son imposibles de llevar a cabo con estos sensores (véase Datos obtenidos de forma remota para las Evaluaciones Forestales Nacionales). Gracias a fotografías aéreas realizadas a baja altitud y sensores como LIDAR se pueden obtener mediciones detalladas del estado forestal. Todos estos sensores suelen ser bastante caros y las imágenes que ofrecen son muy limitadas en

cuanto a extensión, por lo que no son capaces de cubrir, de extremo a extremo, el conjunto del terreno de un país. Sin embargo, en principio, estos sensores podrían utilizarse para medir una muestra de las ubicaciones durante la realización de estudios nacionales. Por ejemplo, puede resultar más económico medir inicialmente una parcela con los datos obtenidos a partir de un sensor remoto para determinar si ésta cuenta con una cubierta de terreno forestal accesible o un uso del terreno forestal. En caso contrario, las visitas del personal de campo a dichas ubicaciones no podrían garantizarse.

En segundo lugar, el cálculo de los cambios y las tendencias de los bosques de un país a menudo constituye una parte importante de una EFN. Si las ubicaciones de las parcelas de muestra están lo suficientemente documentadas, éstas se pueden volver medir más adelante para obtener cálculos más precisos del cambio forestal, tales como el crecimiento arbóreo, la mortalidad, el cultivo, la regeneración y los cambios de las condiciones forestales de las diferentes áreas así como aquellos sufridos en las categorías de uso del terreno. (Véanse los apartados Observaciones temporales frente a observaciones permanentes, Observaciones y mediciones y Evaluación de los cambios). Volver a medir las parcelas aumenta la eficacia del cálculo y ayuda a comprender mejor los componentes del cambio. Sin embargo, si se utilizan parcelas permanentes, sus ubicaciones deben estar documentadas de una forma muy precisa. Se puede hacer introduciendo y anclando un barrote en el suelo, en el centro o en una de las esquinas de una parcela, para marcar la posición y documentando cuidadosamente cómo localizarlo a partir de una ubicación de partida fácil de identificar y cómoda, que puede estar a unos kilómetros de distancia. La marca no debería ser visible, de modo que la parcela sea realmente representativa de cientos de hectáreas que nunca se van a medir. Una parcela de muestra no será representativa si recibe un tratamiento especial, como que se proteja frente a la

explotación u otras alteraciones. Una marca obvia en el suelo podría influir en como otras personas tratan dicha ubicación.

Aunque volver a medir los mismos árboles puede proporcionar cálculos más precisos del cambio, este enfoque es más costoso ya que los mismos centros de las parcelas y árboles se deben reubicar en el momento de cada medición. Entre las alternativas que existen para estimar el cambio a partir de parcelas temporales se incluye el cálculo del crecimiento arbóreo según el incremento de las perforaciones y el cálculo a grandes rasgos del área forestal y el cambio en el volumen mediante la comparación de cálculos independientes obtenidos de mediciones de diferentes conjuntos de parcelas temporales realizadas en diferentes momentos. Sin embargo, la explotación, la mortalidad y la regeneración son elementos difíciles de calcular utilizando datos obtenidos a partir de parcelas temporales. Por lo tanto, siempre que sea posible, se recomienda utilizar parcelas permanentes o una combinación de éstas y parcelas temporales (por ejemplo, Ranneby y otros, 1987).

2. Diseño de muestreo

Muestreo subjetivo o dirigido y muestreo probabilístico. El muestreo subjetivo trata de emplear un juicio profesional para seleccionar aquellas unidades de muestra que se consideran representativas del conjunto de la población. A menudo es más cómodo medir estas unidades, lo que reduce el coste. Aunque los datos que se recopilan así describen de una forma precisa el estado de las ubicaciones muestreadas, no son representativos del conjunto de la población. Los defensores del muestreo subjetivo confían en la capacidad de los expertos para seleccionar una muestra representativa y argumentan que este enfoque es suficientemente bueno para fines prácticos. En determinadas situaciones sencillas, esto puede ser cierto. Sin embargo, ¿qué pasa si el usuario de los datos no tiene la misma confianza en los expertos? Los datos más

costosos pueden dejar de tener valor, ya que el diseño de muestreo no es defendible desde un punto de vista científico crítico. Del mismo modo, las ubicaciones de muestreo más cómodas a menudo se encuentran cerca de carreteras, las cuales están frecuentemente asociadas con accidentes geográficos, usos del terreno, históricas de gestión y patrones de paisaje. ¿Son estas ubicaciones realmente representativas del conjunto de la población? La respuesta es discutible. Es mucho más sencillo desacreditar la precisión de los cálculos de la población obtenidos a partir de una muestra subjetiva que probarlos de otra forma.

El muestreo probabilístico sustituye los juicios subjetivos por reglas objetivas basadas en probabilidades conocidas de la selección de cada miembro de una población. Por ejemplo, supongamos que un bosque de un millón de hectáreas comprende una población de parcelas de 10 m x 10 m. Dicha población podría estar compuesta de 100 millones de estas parcelas. Si una de estas parcelas se eligiera de forma aleatoria, su probabilidad de selección sería de 1/100.000.000. Si una muestra aleatoria simple de 1.000 parcelas se eligiera para calcular las condiciones del conjunto de la población de 1 millón de hectáreas, cada miembro de dicha población tendría una probabilidad de selección de aproximadamente $1.000/100.000.000 = 1/100.000$ y cada parcela medida de la muestra podría considerarse representativa de otras 99.999 parcelas no medidas. El resultado más importante es que el muestreo probabilístico es un método objetivo si se emplean reglas precisas y una base matemática para calcular los atributos de la población en función de una muestra. La probabilidad de que un experto seleccione cualquiera de las parcelas de muestra potenciales es desconocida y las matemáticas que fundamentan el muestreo subjetivo no se pueden aplicar de una forma que se pueda defender desde un punto de vista científico. Por lo tanto, se recomienda el uso del muestreo probabilístico frente al muestreo subjetivo así como el uso de un

muestreo de igual probabilidad, en el que las posibles ubicaciones de las unidades de muestra tienen las mismas probabilidades de ser seleccionadas.

SELECCIONAR UN DISEÑO DE MUESTREO PROBABILÍSTICO

Muchas de las dificultades asociadas con la selección de un diseño de muestreo surgen a raíz de dos factores: por un lado, las unidades de muestra se distribuyen en un determinado espacio y sus observaciones pueden estar correlacionadas de forma espacial y, por otro lado, los diferentes diseños de muestreo tienen costes diferentes. La correlación espacial entre las observaciones de variables de interés ejerce una poderosa influencia sobre la elección de los diseños de muestreo. Factores ecológicos, climáticos y del terreno así como las prácticas de gestión forestal producen observaciones a partir de parcelas cercanas entre sí y serán, de media, más similares que las observaciones a partir de otras que se encuentran alejadas unas de otras. El resultado es que, en sentido estricto, el desarrollo de un diseño de muestreo 100% óptimo es una tarea imposible, ya que las diferentes EFN evaluadas y sus variables derivadas difieren significativamente en el espacio. Por este motivo, puesto que los diseños de muestreo óptimos pueden ser diferentes según las variables, la optimización puede requerir la minimización del error estándar de una única pero importante variable, como puede ser el volumen de madera o, en el caso de una función ponderada, de los errores estándar de un pequeño número de variables. Una solución parcial es minimizar los efectos de la correlación espacial estableciendo ubicaciones de muestreo lo más alejadas posible. Esto también tiene en consideración el hecho de que las observaciones de parcelas de muestra que están más separadas unas de otras, proporcionan más información a la muestra. En muestreos forestales, esto a menudo sugiere diseños hexagonales. Los primeros costes del muestreo se atribuyen a los desplazamientos que se hacen desde y hasta la ubicación de las unidades de muestra y a

las mediciones que se realizan de ellas. Estos costes dependen de la estructura del paisaje y de los bosques, del tipo de medición que se va a tomar y de las condiciones topográficas, económicas y de transporte.

Un punto de inicio común a la hora de seleccionar un diseño de muestreo es conocer los límites superiores aceptables de los errores estándar de los cálculos, así como del coste. Optimizar el diseño de muestreo, dado el marco de muestreo y la configuración de la parcela, implica seleccionar un procedimiento de distribución espacial de las ubicaciones de las unidades de muestra de forma que los errores estándar se minimicen sin sobrepasar los costes totales permitidos. En ocasiones, esto no es posible y puede ser necesario tratar de llegar a un acuerdo.

MUESTREO ALEATORIO SIMPLE

Un muestreo aleatorio simple sitúa las parcelas de muestra de forma aleatoria dentro de la población muestreada (Figura 1a). De manera casual, pueden existir agrupaciones espaciales y terrenos vacíos en la distribución de parcelas; a pesar de ello, continúa siendo un muestreo probabilístico válido. Las coordenadas geográficas de cada una de las parcelas de muestra incluidas en una muestra aleatoria se pueden seleccionar gracias a un generador de números aleatorios que utilice las coordenadas permitidas y limitadas a la población muestreada. Aparte de esto, no se tienen en consideración ni la seguridad, ni la dificultad de medición de las parcelas ni los desplazamientos desde y hasta las ubicaciones de las parcelas. Se trata del diseño de muestreo de igual probabilidad menos arriesgado, pero también es el menos eficaz con respecto al coste y a la precisión de los cálculos, en parte debido a la correlación espacial entre las observaciones.

MUESTREO SISTEMÁTICO

Un muestreo sistemático utiliza una cuadrícula fija, o una matriz, para asignar parcelas en un patrón regular (Figura 1b). La ventaja del muestreo sistemático es que maximiza la distancia media entre parcelas y,

por lo tanto, minimiza la correlación espacial entre observaciones e incrementa la eficacia estadística. Además, el muestreo sistemático, el cual se considera representativo en cierto sentido, puede resultar muy convincente para los responsables de la toma de decisiones que no tienen experiencia con muestreos. Los muestreos sistemáticos se pueden basar en cuadrículas rectangulares o en matrices hexagonales. Por ejemplo, una parcela de muestra podría establecerse en las intersecciones de una cuadrícula de 2 x 2 km. Para seleccionar el punto inicial y la orientación de dicha cuadrícula se utiliza un número aleatorio, pero no es necesaria ninguna otra información de este tipo. Este diseño de muestreo es el más común en ingeniería forestal. El mayor riesgo es que la orientación de la cuadrícula pueda, por algún casual, coincidir o ser paralela a las características naturales o introducidas por el hombre, como carreteras o caminos de grava formados como resultado del deshielo de glaciares. En el caso

de grandes áreas geográficas, se debe evitar la orientación de las líneas de la cuadrícula a lo largo de las líneas de longitud. A elevadas latitudes, la naturaleza convergente de estas líneas de cuadrícula norte-sur puede hacer que las ubicaciones de las parcelas de muestra estén más cerca unas de otras que a latitudes menores. Los diseños de muestreo basados en matrices hexagonales mitigan este problema (White y otros, 1992).

Los diseños de muestreo sistemáticos no alineados combinan las características de los diseños de muestreo aleatorios simples y las de los diseños de muestreo sistemáticos. En estos diseños, una parcela de muestra se asigna a una ubicación de una cuadrícula o una matriz de células seleccionada de forma aleatoria (Figura 1c).

Muestreo agrupado

Por motivos prácticos, como el aumento de la eficacia del coste y la reducción de los desplazamientos del personal de campo, las parcelas de muestra pueden organizarse

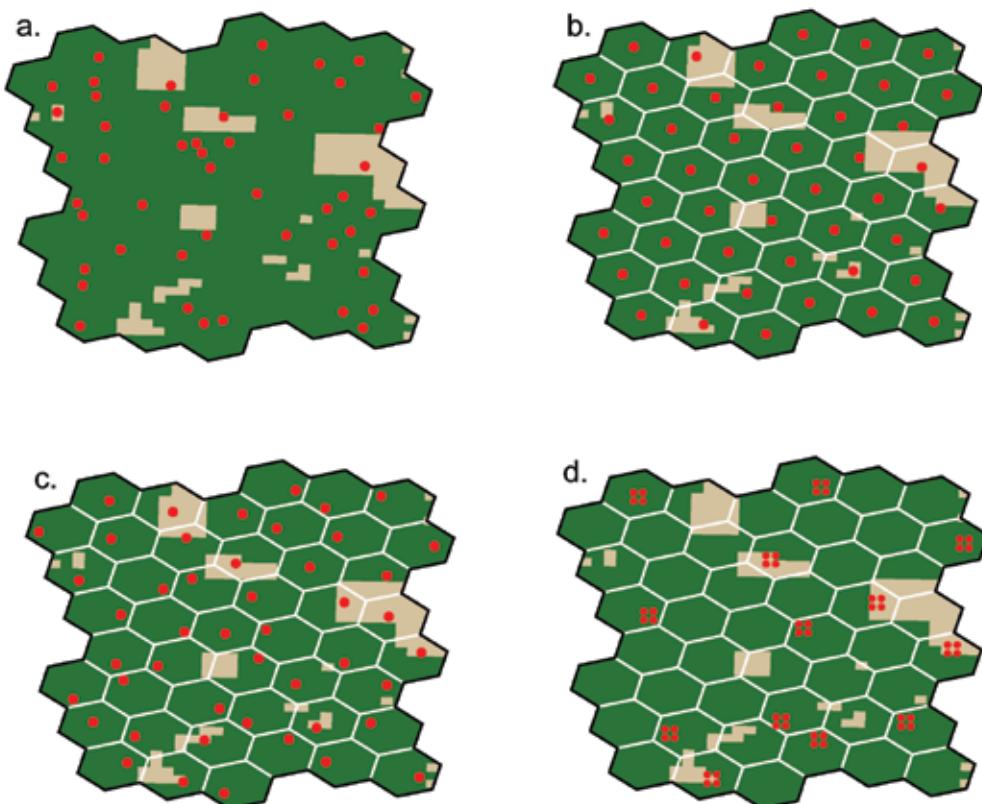


Figura 1. (a) diseño de muestreo aleatorio simple, (b) diseño de muestreo sistemático alineado, (c) diseño de muestreo sistemático no alineado, (d) diseño de muestreo sistemático, agrupado, no alineado con el mismo número de parcelas pero agrupadas.

en grupos, lo que nos lleva a un muestreo sistemático agrupado y a un muestreo sistemático agrupado estratificado. En el muestreo sistemático agrupado, los grupos se distribuyen en la población utilizando cuadrículas o polígonos, como por ejemplo hexágonos.

Varias preguntas son relevantes a la hora de planificar un diseño de muestreo basado en grupo: (1) ¿cuál es la distancia espacial entre los grupos? (2) ¿cuál es la forma del grupo? (3) ¿cuál es el número de parcelas por grupo? y (4) ¿cuál es la configuración de parcela de muestra? Para responder a estas preguntas, es necesaria la información preliminar de la distribución espacial y la correlación de las variables de interés. La correlación, como función de distancia entre las parcelas de campo, calculada mediante variogramas, se puede emplear para comparar la eficacia de los diferentes diseños de muestreo.

2.1 MUESTREO ESTRATIFICADO

El muestreo estratificado conlleva, en primer lugar, dividir la población en subpoblaciones que no se solapen, denominadas estratos y que, de forma conjunta, incluyen el conjunto de la población y, a continuación, dibujar una muestra independiente a partir de cada estrato. Si la muestra de cada estrato es una muestra aleatoria simple, la totalidad del procedimiento se describe como muestreo aleatorio estratificado. Varias razones pueden justificar el muestreo estratificado (Cochran, 1977; Schreuder y otros, 1993). En primer lugar, la estratificación se emplea para aumentar la precisión de las estimaciones de la población. Para comprender la obtención de ganancia potencial en la precisión que puede alcanzarse gracias a la estratificación, son necesarias algunas anotaciones y fórmulas. Con un muestreo aleatorio simple (SRS, por sus siglas en inglés), el cálculo de la media de la población es

$$\bar{y}_{SRS} = \frac{1}{n} \sum_{i=1}^n y_i \quad , \quad [1]$$

y el cálculo de la varianza de la media es

$$Var(\bar{y}_{SRS}) = \frac{s^2}{n} \quad [2]$$

donde n es el tamaño de la muestra, y_i es una observación y

$$s^2 = \frac{1}{n-1} \sum (y_i - \bar{y})^2 \quad , \quad [3]$$

es el cálculo de la muestra de la varianza de la población. Cochran (1977) proporciona fórmulas básicas para el cálculo estratificado. Al ignorar los factores de corrección de población finitos y los errores de cálculo en estratos ponderados, un estimador no sesgado de la media de la población y de la varianza es

$$\bar{y}_{Str} = \sum_{h=1}^L W_h \bar{y}_h \quad , \quad [4]$$

y

$$Var(\bar{y}_{Str}) = \sum_{h=1}^L W_h^2 \frac{s_h^2}{n_h} \quad , \quad [5]$$

donde

$$\bar{y}_h = \frac{1}{n_h} \sum_{j=1}^{n_h} y_{hj} \quad , \quad [6]$$

$$s_h^2 = \frac{1}{n_h - 1} \sum_{j=1}^{n_h} (y_{hj} - \bar{y}_h)^2 \quad , \quad [7]$$

están comprendidos en las medias y varianzas del estrato, respectivamente; $h=1, 2, \dots, L$ indica el estrato; j indica las observaciones en el estrato; n_h indica el número de observaciones de la muestra en el estrato h -ésimo con $n_1+n_2+\dots+n_L=n$; y W_h es la ponderación del estrato que representa la proporción de la población en el estrato h -ésimo. Los efectos de

la estratificación y del cálculo estratificado en la precisión a menudo se evalúan utilizando la eficacia relativa (RE, por sus siglas en inglés), definida como

$$RE = \frac{Var(\bar{y}_{SRS})}{Var(\bar{y}_{Str})}, \quad [8]$$

donde $RE > 1$ indica un efecto beneficioso. La eficacia relativa se puede interpretar como el incremento del tamaño de la muestra general necesario para alcanzar la misma precisión utilizando un cálculo basado en el muestreo aleatorio simple obtenido mediante la estratificación y el cálculo estratificado. Desde un punto de vista cuantitativo, las ganancias en la precisión ocurren cuando las varianzas de las medias de los estratos calculados son sustancialmente inferiores a la varianza de la media general calculada (por

ejemplo, $\frac{s_h^2}{n_h} < \frac{s^2}{n}$) o cuando el estrato con

mayor $\frac{s_h^2}{n_h}$ representa pequeñas porciones de la población (por ejemplo, cuando Wh es pequeño). Desde un punto de vista cualitativo, las ganancias en precisión ocurren cuando las poblaciones heterogéneas se dividen en varias subpoblaciones homogéneas. Por norma general, esto significa que los estratos tienen varianzas, medias, o ambas, sustancialmente diferentes.

Un segundo motivo para optar por la estratificación es que puede contribuir a evitar el sesgo del cálculo, en función del estimador elegido. Por ejemplo, al personal de campo de la EFN generalmente se le otorga acceso a las ubicaciones de las parcelas que se encuentran en terrenos que son propiedad pública. Sin embargo, si se requiere el permiso de los propietarios privados para medir parcelas de muestra que se encuentran en sus tierras, de forma inevitable algunos de ellos denegarán el acceso. En casos extremos, el ratio de las parcelas privadas y públicas en la porción inaccesible de la muestra debe tener menos

relevancia que el ratio de los terrenos forestales privados y públicos en la misma. Si las composiciones de las especies o las prácticas de gestión presentan diferencias sustanciales entre los terrenos forestales de propiedad pública y los de propiedad privada, puede darse el sesgo. Una solución consiste en estratificar las tierras según su propiedad, lo que conlleva cálculos de muestra independientes para los dos estratos de propiedad (McRoberts, 2003).

Un tercer motivo para la estratificación es acomodar los diferentes protocolos de muestreo o diferentes procedimientos de cálculo de las distintas subpoblaciones. Por ejemplo, un porcentaje importante de los costes del muestreo se puede atribuir a los desplazamientos desde y hacia las ubicaciones de las parcelas. Si los datos obtenidos a partir de sensores remotos se pueden emplear para determinar que algunas de ellas están ubicadas en terreno no forestal, los costes derivados de los desplazamientos pueden verse significativamente reducidos al no enviar al personal de campo a dichas ubicaciones. Sin embargo, como resultado de una técnica de medición diferente puede ser necesario un estimador diferente para estos estratos.

Las ventajas más importantes del cálculo estratificado se obtienen cuando la población se estratifica y los tamaños de muestra de los estratos se determinan antes de llevar a cabo el muestreo. El proceso de determinar los tamaños de muestra de los estratos o, de forma equivalente, asignar muestras a estratos, puede realizarse de diferentes maneras y con diferentes objetivos. A menudo, las muestras se asignan a los estratos según determinados atributos de estos. Un enfoque de fácil implantación consiste en asignar parcelas de muestra a los estratos según los tamaños de estos últimos. Si en el estrato se emplea un muestreo aleatorio simple o uno sistemático, este produce muestras de igual probabilidad dentro del estrato, lo que puede simplificar el cálculo. Sin embargo, con este enfoque, las varianzas de las medias de los estratos pueden diferir de forma significativa. Si se desea obtener cálculos precisos y comparables de las

medias de los estratos, las muestras se pueden asignar a los estratos según las varianzas de estos. Una desventaja potencial de este enfoque es que se necesitan cálculos correctos de las varianzas de los estratos antes de asignar las muestras al estrato. Por último, puede que los cálculos de las medias de algunos estratos sean más importantes que otras. En este caso, las muestras se pueden asignar a los estratos según una evaluación subjetiva de su importancia.

A menudo los objetivos del muestreo no permiten realizar el muestreo aleatorio estratificado. Por ejemplo, un diseño de muestreo sistemático se puede utilizar como medio para optimizar la precisión de los cálculos de diferentes variables a la vez. Aunque los resultados más importantes de la estratificación pueden no alcanzarse en el caso de determinadas variables, los efectos beneficiosos de aumentar la precisión y excluir el sesgo del cálculo pueden continuar garantizando tanto la estratificación post-muestreo como el cálculo estratificado. Por lo tanto, incluso si no se emplea el muestreo estratificado, se recomienda considerar el cálculo estratificado post-muestreo, ya que el aumento de la precisión a menudo puede alcanzarse con poco coste o esfuerzo adicional.

A la hora de crear estratos, se puede emplear casi cualquier fuente de datos siempre y cuando dos tareas puedan llevarse a cabo de forma consistente. En primer lugar, se debe determinar la ponderación del estrato, calculada como la proporción de la población representada por cada estrato. Por otra parte, cada parcela debe asignarse a un único estrato. La creciente disponibilidad de las diferentes capas de datos temáticos digitales proporciona grandes posibilidades para las fuentes de datos que se pueden emplear para crear estratos. Además, la creciente disponibilidad de sistemas de información geográfica (SIG) simplifica en gran medida el cumplimiento de ambas tareas. Una elección común de estratificación de datos son las clasificaciones de las cubiertas del terreno, a partir de las cuales se pueden crear categorías forestales

y no forestales agregadas y utilizarlas como estratos (McRoberts, 2002). Utilizar un SIG con esta capa simplifica de forma significativa las dos tareas de estratificación. En los SIG, cada una de las unidades de terreno de la clasificación de la cubierta del terreno se asigna a un estrato basado en la categoría asignada a dicha unidad. El cálculo de las ponderaciones de los estratos se simplifica gracias a la funcionalidad de los SIG para determinar el área total de las áreas de terreno asignadas al mismo estrato y dividiéndolas por el área total de la población muestreada. Una parcela se asigna al estrato de la unidad cartográfica que contiene sus centros. Se pueden utilizar otras elecciones de capas de datos digitales para crear estratos como, entre otros, mapas de suelo, mapas de división de clima, provincias ecológicas, límites administrativos, mapas de propiedad y unidades de gestión de terreno.

2.2 Estimadores de ratio y estimadores de error de Matérn

Aunque el tema de los estimadores estadísticos se trata en otro apartado o puede consultarse en el libro en línea Statistical Techniques for Sampling and Monitoring Natural Resources (Schreuder y otros, 2004), en este documento nos centramos en la importancia de una selección de estimadores consistente en el diseño de muestreo para poder obtener cálculos de varianza válidos. En los diseños de muestreo sistemáticos y basados en grupos, es especialmente importante que los estimadores tengan en cuenta, de una forma adecuada, la posible correlación espacial entre observaciones. Puesto que su utilidad en los diseños de muestreo debe tener en cuenta la correlación espacial, a continuación se ofrece un pequeño comentario acerca de los estimadores de Matérn (Matérn, 1960).

Puesto que los cálculos de los inventarios forestales en ocasiones son o bien medias o bien totales, ya sea para un área o volumen, las variables derivadas relevantes en los inventarios forestales a menudo tienen la siguiente forma:

$$M = \frac{X}{Y}, \quad [9]$$

donde X e Y son expectativas de variables aleatorias, x e y. Por ejemplo, supongamos el cálculo de un área forestal media por estrato de utilización de terreno para parcelas de muestra que pueden interseccionar con varios estratos, todos dentro de la categoría de terreno forestal. Un método para tener en consideración este fenómeno, particularmente importante en el muestreo por puntos, es utilizar la información únicamente del punto central. Si $x_i=1$ cuando el punto central de la parcela pertenece al estrato en cuestión, en caso contrario $x_i=0$; $y_i=1$ cuando el punto central se encuentra en terreno forestal, en caso contrario $y_i=0$. Entonces, el estimador de ratio para el área media es:

$$m = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n y_i} = \frac{\bar{X}}{\bar{Y}} \quad [10]$$

donde n es el número de unidades de la muestra. Si $E(\cdot)$ indica la expectativa estadística; entonces,

$$E(m) \approx \frac{E\bar{x}}{\bar{Y}} = \frac{\bar{X}}{\bar{Y}} = M, \quad [11]$$

significa que m prácticamente no tiene sesgo cuando n es grande.

El cálculo de los errores estándar es complicado debido a la correlación espacial que puede surgir a partir de cambios en la tendencia de las variables, así como debido al muestreo utilizado, sistemático o agrupado. Matérn (1947, 1960) sugirió la varianza de error, σ^2 , como medida de la fiabilidad del estimador y también propuso un estimador de varianza. Si i indica las parcelas de campo, r indica los grupos de las parcelas de campo

y considera los residuos del grupo $z_r = x_r - my_r$ donde

$$x_r = \sum_{i \in r} x_i \quad y_r = \sum_{i \in r} y_i$$

Asumamos que los residuos componen la realización de un proceso estocástico estacionario (débilmente estacionario) de segundo orden. La varianza del proceso se puede calcular mediante fórmulas cuadráticas

$$T = \sum_r \sum_s C_{rs} z_r z_s, \quad \text{donde } C_{rs} = C_s,$$

$$\sum_r \sum_s C_{rs} = 0 \quad \sum_r C_{rr} = 1 \quad \text{donde } r$$

y s se refieren a los grupos de las parcelas de campo. Los estimadores de esta fórmula son insesgados si el proceso z no está correlacionado de forma espacial y conservadores si el proceso está correlacionado de forma positiva (Matérn, 1960). Este enfoque se ha utilizado en inventarios de Suecia y Finlandia (Ranneby, 1981; véase también Tomppo y otros, 1997; y Heikkinen, 2006) y se aplica a los estratos de muestreo como sigue. En cada estrato, el grupo g de cuatro grupos de parcelas de campo (r_1, r_2, r_3, r_4)

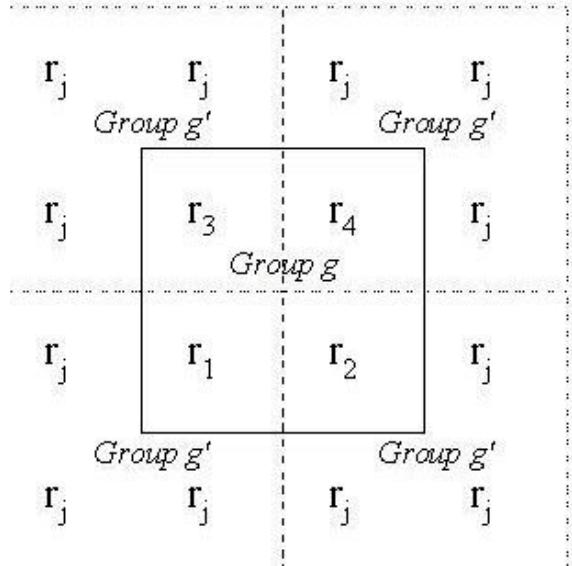


Figura 2. Conjunto de grupos y grupos de parcelas de muestra.

está formado de modo que cada grupo pertenece a cuatro grupos diferentes (Figura 2). La desviación de la media del grupo, \bar{y}_r , a partir de la media del estrato \bar{y} se tiene en cuenta para cada grupo r . Indica $z_r = (\bar{y}_r - \bar{y})n_r$, donde n_r es el número de puntos relevantes de la muestra del grupo r (en este ejemplo, $n_r = 4$). A menudo se utilizan las ponderaciones $c_{r1} = c_{r2} = c_{r3} = c_{r4} = \frac{1}{2}$. Las fórmulas cuadráticas se pueden expresar como $T_g = (z_{r1} - z_{r2} - z_{r3} - z_{r4})^2/4$ y los estimadores de error estándar resultantes para cada estrato son

$$s = \sqrt{\frac{k \sum_g T_g}{\sum_i y_i}}, \quad [12]$$

donde g indica un conjunto de grupos en el estrato, i indica las parcelas del estrato y k es el número de grupos incluidos en la agrupación (en este caso, $k = 1$). Los estimadores de error estándar del conjunto del área de estudio se pueden obtener al combinar los estimadores específicos a nivel de estrato con la fórmula habitual del muestreo estratificado (ecuaciones [4] y [5]). Este procedimiento es relevante para aquellos estratos que cuentan con un gran número de parcelas de campo, preferiblemente y como mínimo, cien.

3. Tamaño de la muestra

Determinar el tamaño de la muestra es uno de los pasos más importantes a la hora de desarrollar un diseño de muestra. Si la muestra es demasiado pequeña, el nivel de incertidumbre será grande y si la muestra es demasiado grande, el coste será innecesariamente elevado. Es posible cuantificar la confianza prevista en los cálculos futuros obtenidos a partir de una muestra probabilísticamente válida. A medida que el número de muestras aumenta, la varianza del error de cálculo disminuye, la precisión de los cálculos aumenta y es posible otorgar más confianza a la estimación. Generalmente, se conoce el valor exacto del cálculo pero no el del estado real del bosque. Gracias

a los muestreos probabilísticos, es posible determinar la probabilidad de que un cálculo se encuentre en una distancia especificada a partir de un valor real. Éstas son las funciones del “intervalo de confianza”, un intervalo estimado de proporciones que probablemente incluya valores reales, pero desconocidos, de la proporción de bosque; y el “coeficiente de confianza”, la probabilidad de que intervalos de confianza similares creados con diferentes muestras contengan una proporción real de bosque.

El escenario más simple es el cálculo de proporciones utilizando una muestra aleatoria simple, por ejemplo, el cálculo del terreno forestal de un país. Supongamos que una EFN abarca una población muestreada de cinco millones de hectáreas y que en una muestra aleatoria simple en la que $n=1.000$ parcelas, 400 de ellas son terreno forestal. La proporción estimada de bosque es del 40%; sin embargo, ¿cuál es el nivel de confianza de este cálculo? Supongamos que un coeficiente de confianza del 80% es aceptable. Esto significa que para 80 parcelas de muestra, el porcentaje real, pero desconocido, de bosque se encuentra dentro del intervalo de confianza. A partir de las tablas y cifras disponibles (Czaplewski, 2003), con $n=1.000$ y un cálculo del 40% de terreno forestal, el intervalo de confianza va del 38 al 42%. Otro ejemplo, supongamos que en la población existe un tipo raro de bosque, pero cuya cantidad exacta no se conoce. Sin embargo, en la muestra aleatoria simple de $n=1.000$ parcelas, esto no se observó y el porcentaje estimado de esta extraña condición forestal es del 0% para el conjunto del país. Para el mismo 80% del coeficiente de confianza, el intervalo de confianza de este cálculo va del 0,0% al 0,2%. Por lo tanto, el cálculo del área de este extraño tipo de bosque en un país de cinco millones de hectáreas va de las 0 a las 10.000 hectáreas. El último ejemplo es una localidad de 100.000 hectáreas para la cual la medición de una muestra de $n=20$ de las 1.000 parcelas, reveló que 18 de ellas son forestales. El cálculo para esta localidad es de un 90% de cubierta forestal, con un intervalo

de confianza que va del 75,5 al 97,3% o de las 75.500 a las 97.300 hectáreas. Es posible realizar otros cálculos de tamaños de muestra mediante el uso de “calculadores de tamaño de muestra” interactivos que se encuentran disponibles en Internet. Estos ejemplos demuestran que es posible realizar cálculos nacionales precisos para tipos de cubierta forestal comunes utilizando relativamente pocas parcelas de muestra. Sin embargo, a menudo son necesarios tamaños de muestra mayores si la EFN requiere información de tipos de bosque raros o de pequeñas zonas del país. Es el tamaño de la muestra lo que determina la precisión de los cálculos en una EFN y no el tamaño del conjunto de la población muestreada.

Para determinar el tamaño de la muestra que se necesita es necesario un cálculo de la desviación estándar de las diferencias entre los valores individuales a nivel de parcela y sus valores medios. Esta desviación estándar puede calcularse realizando un estudio piloto o bien un inventario que mida una pequeña muestra de parcelas forestales y que determine la variabilidad que existe entre ellas. Por ejemplo, supongamos que el inventario piloto incluye 60 parcelas y el volumen de madera se mide en cada una de ellas. Además, supongamos que el volumen medio es =100 m³/ha, la varianza entre parcelas es =2.500 m⁶/ha²y la desviación estándar es =50 m³/ha. Si las observaciones de las parcelas piloto se distribuyen normalmente, cerca de un sexto de ellas tendrá (100-50)=50 m³/ha o menos y otro sexto, 100+50=150 m³/ha o más. Asumamos que el requisito de precisión de la EFN es calcular el volumen de madera medio por hectárea con un ±5% de “tolerancia” o una “diferencia permitida máxima” (Dmáx.=0,05) con un coeficiente de confianza del 66%. El tamaño necesario de la muestra n es de aproximadamente 100 parcelas de muestra.

$$n = \left(\frac{\hat{s}}{\hat{x}D_{\max}} \right)^2 = \left(\frac{50}{100 \times 0.05} \right)^2 = 100 \quad , [15]$$

Si este requisito de precisión de la EFN es

para la totalidad del país, 100 parcelas de muestra son suficientes. Si la precisión de la EFN es para cada 10 unidades subnacionales, es necesario un total de 1.000 parcelas de muestra. Los tamaños de la muestra aumentan de forma significativa a medida que la tolerancia aceptable disminuye. Una tolerancia del ±1% requeriría que el tamaño de muestra se aumentara de n=100 a n=2.500 parcelas de muestra (ecuación 15) en este ejemplo. El tamaño de muestra requerido aumenta si los coeficientes de confianza son mayores. Por ejemplo, se requieren cuatro veces más parcelas de muestra para mejorar la precisión de un coeficiente de confianza del 66% hasta un nivel del 95%. Es posible realizar cálculos más exactos y detallados acerca de los tamaños de muestra utilizando los “calculadores de tamaño de muestra” disponibles en Internet.

4. Comparación entre diseños de muestreo

Una forma efectiva de comparar diseños de muestras es a través de la simulación, si hay disponible un modelo de área forestal. El modelo puede obtenerse a partir de inventarios previos o a partir de cálculos de variables de interés obtenidos mediante imágenes por satélite. En la Figura 3 se muestra un ejemplo de los errores estándar obtenidos a partir de diseños de muestras y utilizados para calcular la media del volumen de masa arbórea en crecimiento. La ubicación de la muestra se encuentra en el norte de Finlandia y tiene una extensión de 6,47 millones de hectáreas, un área de terreno forestal de 4,19 millones de hectáreas y un volumen medio de terreno forestal de 52,7 m³/ha.

A nivel de píxel, se ha recreado un mapa forestal que se extiende de un límite a otro utilizando los datos de campo obtenidos a partir de inventarios anteriores, imágenes por satélite y datos de mapas digitales (Tomppo & Halme, 2004; Tomppo, 2004). Imágenes por satélite de diferentes resoluciones ofrecen una fuente de información, además de los

mapas existentes. Del mismo modo, se puede emplear un inventario piloto para recopilar información con el objetivo de desarrollar el diseño de muestra final. Es posible seleccionar subáreas representativas dentro de la población, en las que realizar los inventarios piloto. Sin embargo, estos inventarios piloto deben reconocerse y aceptarse como lejos de ser óptimos. Además, se pueden crear nuevos diseños de muestra utilizando la información obtenida a partir de inventarios previos, tal y como ha ocurrido en países donde los inventarios forestales se han llevado a cabo desde los años 20 y 30 (p. ej., Ilvessalo, 1927).

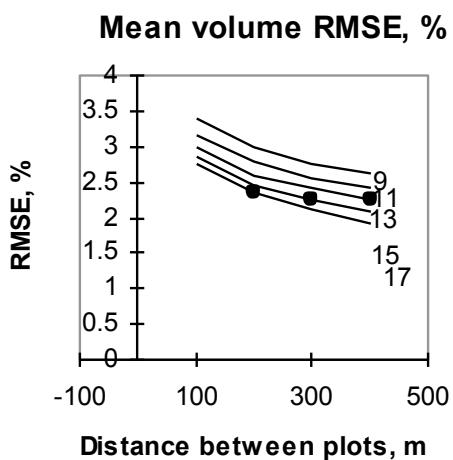


Figura 2. Errores estándar basados en simulaciones de muestreo con distancias diferentes entre parcelas de campo y con grupos compuestos por 9-17 parcelas. La distancia entre grupos es de 10 km. (Figura de Tomppo, et al. 2001; elaboración por Helena Henttonen)

5. Consideraciones de muestreo a tener en cuenta en los inventarios forestales tropicales

En los últimos años, la preocupación por los efectos del cambio climático, así como las acciones para paliar dichos efectos, han motivado un intenso interés en la realización de inventarios forestales en países tropicales

con el objetivo de calcular los niveles de dióxido de carbono y su cambio. Dichos inventarios, a menudo denominados sistemas de supervisión, producción de informes y verificación (MRV, por sus siglas en inglés) y en los casos en los que están orientados a los objetivos de la UNFCCC REDD (reducción de las emisiones producidas por la deforestación y la degradación forestal), son similares a los inventarios forestales nacionales (IFN) aunque el objetivo de los MRV puede limitarse a las variables relacionadas con la biomasa, y la población de interés del MRV puede limitarse a aquellos terrenos expuestos a las emisiones de gases de efecto invernadero producidos por el hombre. Sin embargo, debido a las semejanzas entre los MRV y los IFN, las naciones tropicales en vías de desarrollo a menudo diseñan sus propios IFN de modo que puedan utilizarse como MRV o diseñan sus propios MRV de forma que puedan completar un IFN. Por lo tanto, las directrices que se detallan a continuación son válidas tanto para los MRV como para los IFN.

Por definición, un programa de supervisión se centra en el cambio y en las tendencias. Además, en los últimos años, también los IFN han demostrado un creciente interés en los mismos temas. Por este motivo, la selección de las configuraciones de parcela, los diseños de muestreo y los esquemas de estratificación está provocado, en parte, por el deseo de calcular el cambio.

Configuración de parcela

La selección de una configuración de parcela se fundamenta en varios principios generales, muchos de los cuales son iguales en los inventarios boreales, templados y tropicales, aunque otros también son diferentes. Se sabe que realizar un cálculo preciso del cambio es más complicado que obtener un cálculo preciso de las condiciones actuales, en especial cuando el cambio en cuestión únicamente afecta a una zona pequeña. Sin embargo, la precisión del cálculo del cambio se puede aumentar volviendo a medir las mismas parcelas en momentos posteriores.

Además, el área de terreno de interés podría estratificarse para reducir la varianza gracias a una variable relacionada con la probabilidad del cambio. Por lo tanto, el énfasis en el cálculo del cambio de los inventarios tropicales defiende la utilización de una proporción relativamente grande de parcelas permanentes que, a su vez, están a favor del mercado y establecimiento de ubicaciones de árboles de modo que se puedan localizar en inventarios sucesivos. Aunque el proceso de establecimiento y medición de una parcela temporal es menos costoso que el de una permanente, el proceso de establecimiento y medición de diferentes parcelas temporales en dos momentos del tiempo no resulta menos costoso que el establecimiento, la medición y la nueva medición de una única parcela permanente.

Aunque no existe un consenso en cuanto a la forma de la parcela, a menudo las parcelas circulares son más recomendables ya que requieren puntos de control únicos: los centros de la parcela. Las parcelas rectangulares requieren cuatro puntos de control, uno en cada esquina. Además, para un área de parcela determinada, una parcela circular tiene un perímetro menor, lo que significa que serán necesarias menos decisiones según determinados árboles pertenezcan o no a la parcela. Del mismo modo, determinar las coordenadas de árboles individuales, lo cual puede ser necesario para evaluar su cambio, puede ser más sencillo en las parcelas circulares que sólo cuentan con un único punto de control, en comparación con las parcelas rectangulares que tienen cuatro. Sin embargo, si las densidades de los árboles son bastante grandes, las parcelas rectangulares largas y estrechas pueden ser la mejor alternativa.

Por motivos de eficacia logística, los programas de supervisión e inventario generalmente configuran las parcelas en grupos. Debido a los esperados problemas de acceso, configurar las parcelas en grupos puede ser una tarea mucho más importante en los programas tropicales. Por lo tanto, el

tamaño de la parcela individual y el número de parcelas incluidas en los grupos están sujetos a varias consideraciones importantes, las cuales están, por norma general, relacionadas con consideraciones logísticas, económicas y de precisión (Tomppo y otros, 2010a, 2011; Scott, 1993). En primer lugar, el tamaño y el número de las parcelas debe ser lo suficientemente reducido como para que el personal de campo pueda realizar la medición de la totalidad del grupo en un día. La mayor parte del porcentaje del coste originado por la medición de una parcela en bosques boreales y templados se deriva de los desplazamientos desde y hacia la ubicación de la parcela, y este porcentaje tiende a ser mayor en los bosques tropicales en los que muchas regiones están muy alejadas o son prácticamente inaccesibles. Por lo tanto, el nivel de eficacia más alto se alcanza si el personal de campo no tiene que volver de nuevo a la misma ubicación de la parcela en días posteriores. En segundo lugar, las características de la parcela, como el radio de las parcelas circulares o la longitud de las rectangulares, deben medirse en un plano horizontal y no en terreno irregular. Debido a que las mediciones realizadas sobre plano horizontal son más difíciles de realizar cuando las parcelas son de gran tamaño, particularmente en terrenos ondulados y montañosos, se recomienda utilizar parcelas más pequeñas. En tercer lugar, el establecimiento de parcelas permanentes frente a aquellas temporales facilita el cálculo del cambio, que generalmente requiere el marcado o el establecimiento de las coordenadas de árboles individuales. El último enfoque es más complicado de llevar a cabo en grandes parcelas de bosques tropicales densos, ya que un número mayor de árboles se interpondrá entre aquellos de interés y los puntos de control. Un argumento a favor de utilizar parcelas más grandes en los inventarios tropicales es que los bosques tropicales son generalmente más diversos que los boreales o templados, lo que implica que el área total que se observa en cada ubicación de muestreo debería ser mayor

para poder así registrar una mayor diversidad. Sin embargo, este mayor tamaño también podría alcanzarse aumentando el número de parcelas pequeñas dentro del mismo grupo de parcelas. Este enfoque resulta más económico cuando las correlaciones espaciales entre las observaciones de las variables de interés son grandes, pero disminuyen al aumentar la distancia.

También se puede lograr una mayor eficacia del muestreo al utilizar en la medición pequeñas parcelas que contengan árboles de menor diámetro. En el caso de las parcelas circulares, las subparcelas a menudo se anidan, es decir, tienen forma de círculos concéntricos, todas con el mismo centro. Los particulares tamaños de las subparcelas y los umbrales de diámetro correspondientes a éstas deberían basarse en el número de árboles que se prevé encontrar en ellas, las semejanzas entre árboles y el tiempo de desplazamiento entre las subparcelas pertenecientes a la misma parcela o entre parcelas pertenecientes al mismo grupo.

Por último, la naturaleza remota e inaccesible de muchos bosques tropicales implica que los inventarios pueden tener que confiar en una combinación de datos de parcela y datos obtenidos gracias a la detección remota. Por lo tanto, las consideraciones de la detección remota pueden ser necesarias a la hora de seleccionar una configuración de parcela. Por ejemplo, una parcela debería ser lo suficientemente grande como para constituir una muestra adecuada de los árboles del elemento terrestre correspondiente a aquel que se ha detectado de forma remota, como puede ser el caso de un píxel de imagen satélite o huella de LIDAR que muestra el centro del elemento terrestre. Además, el deseo de alinear diferentes parcelas en un mismo grupo mediante diferentes elementos detectados de forma remota puede requerir que las distancias entre parcelas sean, al menos, tan grandes como las dimensiones de los elementos detectados.

Diseño de muestreo

La selección de un diseño de muestreo para su utilización en un inventario forestal tropical también implica la consideración de varios factores. Por un lado, el equilibrio espacial es una característica muy valorada en estos casos, lo que significa que se podrán observar grandes regiones geográficas de población. El equilibrio espacial a menudo se alcanza incorporando un componente sistemático en el diseño de muestreo. Este componente sistemático puede ser o bien una red de líneas de cuadrícula perpendiculares o bien el teselado de una población en polígonos regulares. Los diseños alineados espacialmente establecen parcelas en las intersecciones de la cuadrícula o en los centros de los polígonos, mientras que los diseños no alineados espacialmente establecen las parcelas en ubicaciones aleatorias y en el interior de los rectángulos que forman las líneas de las cuadrículas o los polígonos regulares.

La teledetección también se puede tener en cuenta a la hora de seleccionar un diseño de muestreo. Por ejemplo, los bosques tropicales a menudo se caracterizan por tener un número de días limitado en los que el cielo está despejado. Por este motivo, puede ser complicado para sensores basados en satélites, como Landsat o SPOT, obtener fotos sin nubes. Los datos de LIDAR, que se obtienen a través de plataformas aéreas y utilizan técnicas láser, a menudo se ofrecen como alternativa. Además, el láser es capaz de atravesar la cubierta forestal y recopilar información relevante para el cálculo del volumen, la biomasa y el contenido de dióxido de carbono de los árboles. Si las parcelas se ubican en las intersecciones de las cuadrículas perpendiculares, se facilita la adquisición de datos de LIDAR a través de plataformas aéreas en franjas, ya que se pueden emplear líneas de vuelo rectas.

Finalmente, al crear las redes de cuadrícula y teselado, el uso de proyecciones de área iguales también debe tenerse en cuenta. En caso contrario, las parcelas ubicadas a

mayores distancias del ecuador representarán un área de población menor que aquellas que se encuentran más cercanas a él. Aunque los esquemas de ponderación se pueden utilizar en proyecciones de área desiguales, a menudo su realización es excesivamente compleja y requiere mucho tiempo.

Tal y como se ha mencionado anteriormente, el porcentaje mayor del coste derivado de la medición de una parcela se corresponde con el desplazamiento desde y hacia ella; este porcentaje puede ser mucho mayor en los bosques tropicales que cuentan con regiones remotas e inaccesibles (Tomppo y otros, 2010a, 2011). Por lo tanto, la eficacia económica indica que las parcelas deben establecerse en grupos en lugar de hacerse de forma individual. Existen varios enfoques acerca del muestreo agrupado. Un enfoque es configurar una parcela como varias subparcelas en un patrón regular y cerca unas de otras (McRoberts y otros, 2005). En este enfoque, los datos de todas las subparcelas se pueden agregar y atribuir al centro de la misma. Un segundo enfoque consiste en establecer parcelas en grupos configurados como rectángulos, mitades de rectángulos u otras formas geográficas (Tomppo, 2006). Un tercer enfoque es el muestreo agrupado en dos fases. Con este enfoque, las unidades de muestreo primarias, como polígonos con forma de grandes rectángulos, se seleccionan de forma aleatoria en primer lugar y, a continuación, las unidades de muestreo secundarias en forma de parcela se establecen dentro de los polígonos seleccionados en ubicaciones escogidas de forma aleatoria. Al utilizar el muestreo agrupado, se debe tener en cuenta la correlación espacial que existe entre las observaciones de las parcelas dentro de un mismo grupo. Si la distancia entre pares de parcelas es inferior al intervalo de correlación espacial, las observaciones tenderán a ser similares y el muestreo será menos eficaz.

Estratificación

Los enfoques de muestreo estratificados se utilizan por varios motivos pero,

principalmente, para variar las intensidades del muestreo de forma que se ajusten a los criterios seleccionados. Por ejemplo, en el caso de un MRV enfocado a regiones geográficas sujetas a emisiones de dióxido de carbono provocadas por el hombre, podrían ser aceptables intensidades de muestreo menores, en el caso de aquellas zonas remotas o inaccesibles en las que su desarrollo o cultivo es menos probable. Además, el coste asociado con intensidades de muestreo mayores en regiones remotas puede ser excesivo. Sin embargo, el muestreo, incluso con intensidades menores, puede realizarse en estas regiones con el objetivo de alcanzar el equilibrio espacial.

Varios principios también guían los enfoques estratificados de muestreo. En primer lugar, son recomendables los estratos con límites estables. Por otra parte, los cambios en los límites de los estratos con diferentes intensidades de muestreo conllevan distintas probabilidades de inclusión en el muestreo, así como complicaciones a la hora del cálculo. Además, los cálculos estratificados requieren que una parcela se asigne a un único estrato. Si el estrato al cual la parcela está asignado cambia entre mediciones, surgen dificultades en el estrato al que la observación de cambio de la parcela debería asignarse. Por este motivo, los estratos definidos según la topografía, las zonas climáticas, los biomas o las fronteras políticas son más recomendables que aquellos definidos por atributos forestales como la densidad o el tipo de bosque.

El muestreo estratificado se implementa con más frecuencia utilizando uno de los tres esquemas de asignación de parcela. Gracias a la asignación equitativa, el mismo número de parcelas se asigna a todos los estratos, independientemente de sus tamaños. Este esquema es recomendable si el objetivo es calcular estratos individuales. Con la asignación óptima, las intensidades de muestreo seleccionadas para los estratos se basan en criterios de optimización, como los costes de medición y la variación a nivel de estrato de las observaciones de variables

de interés, como pueden ser el volumen o la biomasa, u otros posibles cambios. Para aquellos estratos con una mayor variación o menores costes de medición, se seleccionan intensidades de muestreo mayores. Con la asignación proporcional, las intensidades de muestreo seleccionadas para los estratos son proporcionales a su tamaño. Cochran (1977) ofrece un extenso debate acerca de las diferentes estrategias alternativas. En el caso de los países tropicales con regiones remotas y prácticamente inaccesibles, en ocasiones será necesaria alguna fórmula de asignación óptima para mitigar los excesivos costes asociados con el muestreo de estas regiones. Una asignación proporcional y óptima se puede implementar de una forma sencilla utilizando diseños de muestreo basados en redes de líneas de cuadrícula perpendiculares. En el caso de las asignaciones proporcionales, las parcelas o parcelas agrupadas se establecen en las intersecciones de la cuadrícula independientemente del estrato asociado con la intersección de la cuadrícula. Con la asignación óptima, las intensidades de muestreo se pueden aumentar o disminuir en función del estrato al seleccionar las intersecciones de la plantilla en las que las parcelas están establecidas. Por ejemplo, si la intensidad de muestreo se va a reducir unas cuatro veces, las parcelas pueden establecerse en las intersecciones únicamente de las segundas líneas de la cuadrícula en cada dirección.

Caso práctico: Tanzania

En un diseño de muestreo llevado a cabo en Tanzania, Tomppo y otros (2010a) utilizaron un muestreo doble para realizar la estratificación y la asignación óptima de parcelas por estratos. La primera fase del muestreo consistió en la evaluación previa de una gran variedad de parcelas de campo para así asignar las categorías de volumen y coste. Según el resultado de esta evaluación, los estratos se definieron utilizando los volúmenes medios previstos de las masas arbóreas en crecimiento a nivel

de agrupación, así como el coste estimado de la medición de un grupo de parcelas. Las categorías de volumen se basaron en las predicciones de volumen obtenidas a través de imágenes por satélite, observaciones de parcelas de campo fuera de Tanzania y sólidos modelos cuyas predicciones se calibraron utilizando los cálculos de volumen aéreo del país. La asignación Neyman (Cochran, 1977) se empleó para seleccionar los límites de las categorías de volumen, para así maximizar la precisión del cálculo de volumen general, asumiendo un tamaño de muestra fijo. Las categorías de coste se basaron en el análisis de los SIG y en la opinión de expertos locales según el número de días (uno, dos o más de dos) necesarios para medir un grupo de parcelas. La selección de los intervalos de las categorías, que afecta a la ganancia que se puede obtener empleando la estratificación, requiere una labor de investigación mayor. La segunda fase consistió en la medición de campo de las parcelas, donde las intensidades de muestreo entre estratos se seleccionaron gracias a una asignación óptima (Cochran, 1977). Gracias a ella, las intensidades de muestreo son proporcionales a la cantidad donde oh es la desviación estándar entre estratos de las observaciones de la variable de interés (volumen de la masa arbórea en crecimiento) y ch es el coste medio en términos de tiempo de medición necesario para un grupo de parcelas en el estrato h . Se puede encontrar más información acerca del diseño de muestreo en Tomppo y otros (2010a).

En los trópicos, a menudo es necesario hacer uso de los mapas de vegetación disponibles para separar el terreno forestal del no forestal. Sin embargo, si los grupos de parcelas no se delimitan correctamente o se consideran como terreno no forestal, de la misma manera que se hace con el terreno forestal, los errores presentes en los mapas pueden desvirtuar los resultados, ya que el terreno forestal considerado como no forestal no se puede medir. Sin embargo, los costes asociados con el muestreo de terreno considerado no forestal pueden disminuir al asignar intensidades de

muestreo menores a estos terrenos. Además, la medición de campo de los grupos de parcelas que se realiza fuera del bosque y sin masa arbórea en crecimiento a menudo se puede evitar evaluando estas agrupaciones utilizando la información de terreno obtenida a partir otras fuentes de confianza, tal y como se sugirió en el caso de Brasil (Tomppo, 2009).

La ausencia de rutas de transporte, que no sean ríos, constituye un reto especial en los inventarios de bosques tropicales, como ocurre en el bioma amazónico. Por ejemplo, las carreteras pueden estar disponibles únicamente en determinados meses del año, seis aproximadamente en el caso del bioma amazónico. Además, algunos bosques pueden formar parte de proyectos de conservación medioambiental o destinarse para el uso exclusivo de las poblaciones indígenas. La estratificación basada en variables relevantes como la probabilidad de cambio y la medición de costes fomenta tanto la eficacia económica como la adherencia a los principios estadísticos del inventario.

6. Resumen

La creación de un diseño de muestreo adecuado para una EFN, un IFN o MRV es un paso crucial si los cálculos son lo suficientemente precisos y se pueden defender desde un punto de vista científico. Uno de los primeros pasos de este proceso es definir la población objetivo y seleccionar un marco de muestreo. Se recomienda un marco de muestreo de población infinita en el que las observaciones y mediciones del área de una parcela se atribuyen a un punto que se encuentra en el centro de ésta. Puesto que a menudo se espera que los inventarios reflejen estimaciones del cambio, se recomienda que el diseño de muestra incluya al menos algunas parcelas permanentes. El siguiente paso es distribuir las parcelas de campo a través de la población que se va a medir. Se ha proporcionado información y se ha debatido acerca de varios diseños de muestra de uso común: muestreo aleatorio simple, muestreo estratificado,

muestreo sistemático y muestreo agrupado. Si el diseño de muestreo incluye un componente sistemático, se recomienda precaución al usar cuadrículas rectangulares en las poblaciones objetivo con grandes componentes norte-sur. Aunque la elección de un determinado diseño de muestreo depende de una gran variedad de consideraciones, se recomienda, si no se va a emplear el muestreo estratificado, considerar la estratificación post-muestreo así como el cálculo estratificado. Por último, se puede encontrar información adicional acerca de estos y otros problemas de los diseños de muestreo más complejos en el material de referencia.

7. Referencias y recursos técnicos

- Cochran, W. G.** (1977). *Sampling techniques*. (3rd ed.). New York: Wiley.
- FAO, 2001.** Global Forest Resources Assessment 2000. Main report. FAO Forestry Paper 140. FAO Rome. 479 p.
- Cunia, T.** 1978. On the objectives and methodology of national forest inventories. In: Forest resource inventory. Joint Meeting of IUFRO Groups S4.02 and S4.04. 18-26 June 1978. Institutul de Cercetari si Amenajari Silvice, XI-XXIX. Bucharest.
- Czaplewski, Raymond L.** 2003. Chapter 5: Accuracy assessment of maps of forest condition: statistical design and methodological considerations, pp. 115-140. *Remote Sensing of Forest Environments: Concepts and Case Studies*. (Michael A. Wulder and Steven E. Franklin, Eds.) Kluwer Academic Publishers, Boston. 515p
- Heikkilä, J.** 2006. Assessment of uncertainty in spatially systematic sampling. In: Kangas, A. & Maltamo, M. (eds.). *Forest inventory. Methodology and applications*. Managing Forest Ecosystems. Vol 10. Springer, Dordrecht. p. 155-176.
- Ilvessalo, Y.** 1927. The forests of Suomi Finland. Results of the general survey of the forests of the country carried out during the years 1921--1924. In Finnish with English summary. *Communicationes ex Instituto Quaestionum Forestalium Finlandie* 11. Summary in English p. 321-395.
- Loetsch, F., & Haller, K. E.** 1973. *Forest Inventory*. Volume I. BLV Verlagsgesellschaft mbH, München.

439 pp.

Loetsch, F., Zöhrer, F. & Haller, K. E. 1973. Forest Inventory. Volume II. BLV Verlagsgesellschaft mbH, München. 469 pp.

Matérn, B. 1960. Spatial variation. *Meddelanden från statens skogsforskningsinstitut*. 49(5). 144 pp. Also appeared as *Lecture Notes in Statistics* 36. Springer-Verlag. 1986.

McRoberts, R.E. 2003. Compensating for missing plot observations in forest inventory estimation. *Canadian Journal of Forest Research* 33:1990-1997.

McRoberts, R.E., Wendt, D.G., Nelson, M.D., and Hansen, M.H. 2002. Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. *Remote Sensing of Environment* 81:36-44.

McRoberts, R.E., Bechtold, W.A.B., Patterson, P.L., Scott, C.T., & Reams, G.A. (2005). The enhanced Forest Inventory and Analysis program of the USDA Forest Service: historical perspective and announcement of statistical documentation. *Journal of Forestry* 103: 304-308.

Ranneby, B. 1981. Medelfelsformer till skattningar baserade på material från den 5:e riksskogstaxeringen. Abstract: Estimation of standard errors in systematic sampling. Swedish University of Agricultural Sciences, Section of Forest Biometry. Report 21. 19 s. (in Swedish).

Ranneby, B., Cruse, Th., Hägglund, B., Jonasson, H., Swärd, J. 1987. Designing a new national forest survey for Sweden. *Studia Forestalia Suecica* 177. 29 p.

Schreuder, H.T., Ernst, R., and Ramirez-Maldanado, H. 2004. Statistical techniques for sampling and monitoring natural resources. Gen. Tech. Rep. RMRS-GTR-126. USDA Forest Service, Rocky Mountain Research Station, Ft. Collins, CO. Available online at: <http://www.treesearch.fs.fed.us/pubs/viewpub.jsp?index=6287>.

Schreuder, H. T., Gregoire, T. G., Wood, G. B. 1993. Sampling Methods for Multiresource Forest Inventory. John Wiley & Sons, Inc, New York. 446 p. Scott, C.T. 1993. Optimal design of a plot cluster for monitoring. In: K. Rennolls, and G. Gertner (Eds.). *Proceedings of the IUFRO S.4.11 Conference. The optimal design of forest experiments and forest surveys.* pp. 233-242. Available by request from the author: ctscott@fs.fed.us.

Tomppo, E. 2004. Resource assessment: Forest resources. In: Burley, J., Evans, J. & Youngquist, J.A. (eds.). *Encyclopedia of forest sciences*. Elsevier, p. 965-973.

Tomppo, E. 2006. The Finnish National Forest Inventory. In: Kangas, A., & Maltamo, M. (Eds.). *Forest inventory, methodology and applications*. Springer. 179-194.

Tomppo, E. 2009. Establishing Methodological Basis and Building Partnership for Brazil's National Forest Inventory. A mission report for the Food and Agriculture Organization of the United Nations (FAO), Division FOMRD.

Tomppo E., & R.L. Czaplewski. 2002. Potential for a remote-sensing-aided forest resource survey for the whole globe. *Unasylva* 53(210):16-18.

Tomppo, E. & Halme, M. 2004. Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: a genetic algorithm approach. *Remote Sensing of Environment* 92:1-20.

Tomppo, E., Czaplewski, R. L., & Makisara, K. 2002. The role of remote sensing in global forest assessment: A remote sensing background paper for Kotka IV expert consultation. *Forest Resources Assessment (FRA) Working Paper 61*. Food and Agriculture Organization of the United Nations, Forest Resources Division, Forest Resources Assessment Programme, Rome, Italy.

Tomppo, E., Henttonen, H. & Tuomainen, T. 2001. Valtakunnan metsien 8. inventoinnin menetelmä ja tulokset metsäkeskuksittain Pohjois-Suomessa 1992-94 sekä tulokset Etelä-Suomessa 1986-92 ja koko maassa 1986-94. *Metsätieteen aikakauskirja* 1B/2001: 99-248. (in Finnish).

Tomppo, E., Gschwantner, T., Lawrence, M. & McRoberts, R.E. (eds.) 2010b. *National Forest Inventories - Pathways for common reporting*. Springer, 612 p. ISBN 978-90-481-3232-4

Tomppo, E., Heikkilä, J., Henttonen, H.M., Ihälainen, A., Katila, M., Mäkelä, H., Tuomainen, T. & Vainikainen, N. 2011. Designing and conducting a forest inventory - case: 9th National Forest Inventory of Finland. Springer, *Managing Forest Ecosystems* 21. 270 p. ISBN 978-94-007-1651-3

Tomppo, E., Varjo, J., Korhonen, K., Ahola, A., Ihälainen, A., Heikkilä, J., Hirvelä, H., Mikkelä, H., Mikkola, E., Salminen, S. & Tuomainen, T. 1997. Country report for Finland. In: *Study on European Forestry Information and Communication Systems. Reports on forestry inventory and survey systems*. Vol. 1. European Commission, p. 145-226.

Tomppo, E. Katila, M., Mäkisara, K., Peräsaari, J., Malimbwi, R., Chamuya, N., Otieno, J. Dalsgaard, S., & Leppänen, M. 2010a. A Report to the Food and Agriculture Organization of the United Nations (FAO) in support of Sampling Study for National Forestry Resources Monitoring and Assessment (NA-FORMA) in Tanzania. Available at <http://dl.dropbox.com>.

com/u/6700344/NAFORMA%20Sampling%20Study.pdf

White, D, Kimerling, J., and Overton, S.W. 1992.
Cartographic and geometrics components of a global
sampling design for environmental monitoring.
Cartography and Geographic Information Systems
19:5-22.