

7 List of Terms

Bold terms in the explanations are included in this list. Font style of quoted text has not been altered.

Accuracy

A measure of “correctness”; in the context of this document, accuracy expresses the degree to which the map agrees with reality; an estimate of accuracy is the degree to which the map agrees with a sample of **reference observations**.

Accuracy assessment

The process of estimating a measure of map accuracy using a **sample of reference observations**.

Bias

A property of an **estimator**; we say that the bias of an estimator $\hat{\mu}$ of a population parameter μ is the difference between μ and the expected value, $\hat{\mu}$ over all possible samples; that is, $\text{Bias}(\hat{\mu}) = E(\hat{\mu}) - \mu$ (Casella & Berger, 2002, p. 330). Note that “...because an estimate is a number, it has no variance and no bias (Särndal et al., 1992, p. 41). The term *biased* [or unbiased] *estimate* is not recommended, although it appears occasionally in the literature.

Cluster

A “sampling unit [that] consists of group or *cluster* of smaller units” (Cochran, 1977, p. 233).

Commission error

Commission error is the proportion or percentage of the area mapped as the category of interest that is erroneously predicted based on comparison to the **reference classification**. Commission error is the complement of **user’s accuracy** (Olofsson et al., 2013).

Confidence interval

A 95% confidence interval for a population parameter, μ , expresses uncertainty in the parameter estimate, $\hat{\mu}$, and is calculated using the sample data. Confidence intervals are often, but not necessarily, in the form $\hat{\mu} \pm a \cdot \text{SE}(\hat{\mu})$ where $\text{SE}(\hat{\mu})$ is the **standard error** of the estimate and a is a statistic related to the desired confidence level (see **z-score**). Among the aggregate set of confidence intervals constructed using all samples that could be realized using the sampling design, 95% of such intervals are expected to include the true value of the population parameter μ , although which intervals do and which do not include μ is generally unknown. The IPCC Good Practice Guidelines (IPCC, 2006, Section 3.1.3) recommend the use of 95% confidence intervals in greenhouse gas inventories.

Design-based inference

The process of drawing an **inference** for a **population parameter** by analysing a probability **sample** selected from the population. With design-based inference, the observation for a population unit is a constant value, apart from negligible measurement error. Design-based properties of estimators such as bias and variance are determined from the sampling distribution

constructed from the aggregate set of population parameter estimates obtained from all possible samples that could be realized from the sampling design. See also **Inference**.

Estimate

The value obtained from the **estimator** when applied to a specific **sample**. An estimate of μ is denoted $\hat{\mu}$. Note that in the literature, the estimator is usually denoted in the same way as the estimate (e.g. Cochran, 1977, p. 11; Särndal et al., 1992, p. 40) because there will not likely be confusion regarding which use we intend: for example, an estimator of μ is denoted $\hat{\mu}$. Note that because variance and bias are properties of estimators, and “because an estimate is a number, it has no variance and no bias. The term *biased* [or *unbiased*] *estimate* is [not recommended but is] nevertheless used occasionally.” (Särndal et al., 1992, p. 41).

Estimator

“The rule by which an estimate of some population characteristic [i.e. parameter] μ is calculated from the sample results” (Cochran, 1977, p. 11). Note that that an estimator is not the same thing as an estimate – instead “An estimator is a function of the sample, while an estimate is the realized value of an estimator (that is, a number) that is obtained when a sample is actually taken” (Casella & Berger, 2002, p. 312).

Expected value

“The expected value, or expectation, of a random variable is merely its average value, where we speak of ‘average’ value as one that is weighted according to the probability distribution. [...] The *expected value* or *mean* of a random variable X is denoted by $E(X)$ ” (Casella & Berger, 2002, p. 55).

Inference

In a sampling framework, an inference expresses the relationship between the **population parameter**, μ , and its estimate, $\hat{\mu}$, in probabilistic terms, typically in the form of either of a **confidence interval** or a test of hypothesis (Dawid, 1983).

Margin of error

A relative measure of the **uncertainty** in an **estimate**. Note that the definitions of margin of error are not all the same. Typically, it is calculated as the ratio of the half width of a 95% confidence interval to an estimate.

Model-assisted estimator

An **estimator** used in design-based inference that incorporates auxiliary information to increase precision by comparing a model’s predictions, often in the form of map unit values, to a probability **sample** of **reference** observations (Särndal et al., 1992, p. 219). Model-assisted estimators can be particularly effective when the response variable is continuous (e.g., proportion forest or biomass) rather than categorical (e.g., forest/non-forest or forest change class).

Model-based Inference

Inference based on the perspective that an observation for a **population unit** is a **random variable** and that the model used to make predictions for all units in the **population** has been adequately

specified in the sense that there is no systematic lack of the model fit to data that represent the entire population. See also **Inference**.

Monte Carlo method

A technique in which a large quantity of randomly generated numbers are studied using a probabilistic model to find an approximate solution to a numerical problem that would be difficult to solve by other methods. For example, suppose a sample s is a realization of set-valued random variable S specified by the sampling design. The **expected value** and **variance** of a statistic $Q(S)$ are unknown but $Q(s)$ can be calculated once the sample data has been collected. If we select many **samples** (let's say 10,000) using the same design, each of size n , and compute Q for each sample, the average and variance of the values $Q(s)$ will approximate the expectation and variance of $Q(S)$. "This method of approximating the value of quantities that may be hard to calculate by analytic means is known as a Monte Carlo simulation" (Särndal et al., 1992, p. 35).

Omission error

Omission error is the proportion or percentage of area with the **reference classification** of the category of interest that is erroneously predicted (mapped) to be in other categories. Omission error is the complement of **producer's accuracy** (Olofsson et al., 2013).

Overall accuracy

The overall accuracy is the "overall proportion of area correctly classified" (Stehman, 1997, p. 79).

Parameter

See **population parameter**.

Population

"The aggregate [that we want to obtain information about] from which the sample is chosen" (Cochran, 1977, p. 5). (Example: all likely voters in the next U.S. presidential election; an example in the context of this document, is all Landsat pixels of a study area).

Population parameter

"Numerical characteristics, or parameters, of the population [with constant values that will be estimated from a sample]" (Rice, 1995, p. 186). (Examples: the area deforested in Brazil 2010-2015 or a fixed value in a model.)

Population unit

An individual member of the set of elements that make up the **population**. Population units are referred to as *elements* in Särndal et al. (1992, p. 9).

Post-stratified estimation

Post-stratification refers to a **stratification** of the study area that is independent of the selection of the **sample** and the stratification is applied subsequent to sampling in the estimation of parameters. For example, if a systematic sample of plots exists and we want to estimate the area of forest, the use of a forest/non-forest map to stratify the area is likely to increase **precision** in the

area **estimate**, even if stratified random sampling is not used. In this case, a **stratified estimator** applied to the sample data is referred to as a **post-stratified estimator**.

Precision

In the context of estimation, Cochran (1977, p. 16) states that because “of the difficulty of ensuring that no unsuspected bias enters into estimates [sic], we will usually speak of the *precision* of an estimate instead of its *accuracy*. Accuracy refers to the size of deviations from the true mean μ , whereas precision refers to the size of deviations from the mean m obtained by repeated application of the sampling procedure.” In the context of this document, we often characterize the **precision** of an **estimate** with a 95% **confidence interval** – the larger the interval the less the precision (and greater the **uncertainty**).

Probability distribution function (PDF)

From IPCC (2006, Volume 1, Chapter 3, Section 3.1.3): The PDF describes the range and relative likelihood of possible values. The PDF can be used to describe **uncertainty** in the **estimate** of a quantity that is a fixed constant whose value is not exactly known, or it can be used to describe inherent variability.

Probability sample

A **sample** drawn from a **population** using a randomization mechanism such that “the inclusion probability for each element of the sample is known, and the inclusion probabilities are non-zero for all elements of the population.” (Stehman, 1999).

Producer’s accuracy

From Stehman (1997, p. 79): “Producer’s accuracy for [category] j [is] the conditional probability that an area classified as category j by the reference data is classified as category j by the map.” When expressed in terms of area, producer’s accuracy is the proportion of area that has the **reference classification** of the category of interest that is correctly predicted (mapped) as that category. For a simple random sample of **reference observations**, each of which represents an equal area, producer’s accuracy for a category is estimated by dividing the number of correctly classified map units in the category by the number of reference units in the category (Lillesand et al., 2008, p. 586). For stratified random sampling of reference data, producer’s accuracy is estimated in two steps: (1) the number of map units in to each map category (or stratum) is multiplied by the stratum weight; and (2) the product for the category of interest is divided by the sum of the products. Producer’s accuracy is the complement of **omission error** (Olofsson et al., 2013)

Random variable

A variable whose value depends on possible outcomes of some process (a coin toss for example). More formally, it is defined as (Gut 2009): “A random variable is a (measurable) function from a probability space to the real numbers”. Also called stochastic variable.

Reference data

Data characterizing the most accurate available assessment of the true condition at the **sample** location (example: fine-resolution satellite imagery).

Reference classification

The most accurate available assessment of the true condition of a **population unit** (example: deforestation).

Reference observations

The **reference classification** applied to the collection of all **sample units**.

Sample

A subset of **population units** selected from the **population**.

Sampling frame

List of **population units** that can be selected for inclusion in a **sample**. The population units in the list are referred to as **sampling units**. In other words, a frame is a device that provides observational access to the population by associating the **population units** with the sampling units (Särndal et al., 1992, p. 9).

Sample unit

The **sampling units** drawn from the **sampling frame** for inclusion in a **sample**. I.e., a sample unit is different from sampling and **population units**.

Sampling unit

Entities that make up the **sampling frame** (Särndal et al., 1992, p.5). In the literature, sometimes there is no distinction made between **population units** and sampling units (e.g. Cochran, 1977, p. 6). However, population and samplings units are different entities because the sampling frame and the **population** are sometimes different entities (in many situations though, the sampling frame is equivalent to the population).

Simple random sampling

“A method for selecting n units out of the N such that every one of [the sets of n specified units] has an equal chance of being drawn.” (Cochran 1977, p. 18).

Spatial assessment unit

In the context of accuracy assessment a spatial assessment unit is a “unit for comparing the map and reference classifications” (Stehman & Wickham, 2011, p. 3044; typically a pixel, block of pixels or polygon).

Standard deviation

The standard deviation of X is the square root of the **variance**: $SD(X) = \sqrt{V(X)}$ and is often denoted by $S(X)$, $SD(X)$, $D(X)$, or σ_X . Because of the square root, the standard deviation has the same unit as the random variable as opposed to the variance. A standard deviation calculated from **sample** data

is sometimes referred to as the *sample standard deviation* to distinguish it from the population standard deviation. The standard deviation of an **estimate** is referred to as its **standard error**.

Standard Error

The standard error is the **standard deviation** (i.e. square root of the **variance**) of an **estimator** (Rice, 1995, p. 192). For example, consider the situation in the explanation of variance below: we want to estimate the mean \bar{Y} of a **population** of size N with variance σ^2 . To do this, we select a simple random sample of n units (y_1, \dots, y_n) ; \bar{y} is an estimate of \bar{Y} and the estimated variance of the sample mean is $\hat{V}(\bar{y}) = s^2 \div n$. The standard deviation of \bar{y} is referred to as its **standard error**. Because the sample variance is an **unbiased estimator** of the population variance, we can estimate the standard error using the sample variance: $SE(\bar{y}) = \sqrt{\hat{V}(\bar{y})} = s \div \sqrt{n}$. Note that because a standard error is a standard deviation of an estimator, all standard errors are also standard deviations but not all standard deviations are standard errors. This sometimes causes confusion even though the definitions of standard error and standard deviation are consistent. The confusion is exacerbated by the common use of the letter S to denote both standard errors and standard deviations.

Strata

Strata are “subpopulations that are non-overlapping, and together comprise the whole population” (Cochran, 1977, p. 89)

Stratified estimator

The stratified estimator is used in design-based inference to estimate of the mean of a **population** often applied to a sample selected by stratified random sampling (Cochran, 1977, p. 91). The estimator is expressed as the sum of the means of the simple random samples within strata weighted by stratum weights calculated as relative proportions of the population within strata. When applied to a simple random or simple systematic sample from the entire population (i.e., without using the strata in the sampling design), the stratified estimator is called a **post-stratified estimator**.

Stratified random sampling

When “simple random sampling is taken in each stratum” (Cochran 1977, p. 89).

Time series

“A time series is a sequence of observations taken sequentially in time. [...] An intrinsic feature of a time series is that, typically, adjacent observations are dependent.” (Box et al. 1994).

Time series analysis

“Time series analysis is concerned with techniques for the analysis of this dependency [of adjacent observations]”. (Box et al. 1994).

Unbiased estimator

“An estimator $\hat{\mu}$ of μ is unbiased if the mean value of $\hat{\mu}$, taken over all possible samples obtained using the [design], is equal to μ ” (Cochran, 1977, p. 11); or in other words, the **estimator** is

characterized as unbiased if it produces an “estimate [that] is correct ‘on the average’” (Rice, 1995, p. 192) over all possible samples. See also **bias**.

Uncertainty

The opposite of **precision**.

User’s accuracy

From Stehman (1997, p. 79): “User’s accuracy for [category] i [is] the conditional probability that an area classified as category i by the map is classified as category i by the reference data”. When expressed in terms of area, user’s accuracy is the proportion of the area that has the predicted class of the category of interest that is correctly classified as determined by comparison to the reference classification. For a simple random sample of reference observations, each of which represents an equal area, user’s accuracy is estimated by dividing the number of correctly classified map units in each category by the total number of units classified into that category (Lillesand et al., 2008, p. 586). User’s accuracy is the complement of **commission error** (Olofsson et al., 2013).

Variance

The formal definition of the variance of a random variable X with expected value $E(X)$, is $V(X) = E(X - E(X))^2$ and provides “a measure of the degree of spread of a distribution around its mean” (Casella & Berger, 2002, p. 59). But this definition is not very relevant in the context of this document. Instead, we are concerned with situations where a sample has been selected from a population (e.g. all pixels of a country) with the objective of estimating a certain population parameter, μ (e.g. the area of deforestation). For example, let’s say we have a population of N units (y_1, \dots, y_N) with mean \bar{Y} and variance σ^2 from which a sample of n units (y_1, \dots, y_n) has been selected by simple random sampling. We have collected reference observations for the n sample units. The sample variance is $s^2 = \sum (y_i - \bar{y})^2 \div (n - 1)$ (Cochran, 1977, p. 26). Again, this is also not very helpful as we are usually not interested in the sample variance but in the variance of an **estimate** (e.g. the area of deforestation). For simple random sampling designs, the sample mean \bar{y} is an unbiased estimator of the population mean \bar{Y} , and the sample variance s^2 is an **unbiased estimator** of the population variance σ^2 , such that $E(\bar{y}) = \bar{Y}$ and $E(s^2) = \sigma^2$ (Cochran, 1977, p. 22, 26). The variance of \bar{y} is $V(\bar{y}) = \sigma^2 \div n$ (assuming a small n relative N); the estimated variance of \bar{y} substitutes the sample variance s^2 for the population variance σ^2 to create the variance estimate of \bar{y} as $\hat{V}(\bar{y}) = s^2 \div n$. This is the variance that is of primary interest to us, and that allows us to calculate a **standard error** and a **confidence interval** of the estimated population parameter of interest.

Z-score

A z-score (also referred to as *standard score*), $z_{1-\alpha/2}$ where $1 - \alpha$ is the confidence level between 0 and 100%, is a constant such that the area under the standard normal density function in between $\pm z_{1-\alpha/2}$ is $1 - \alpha$ (Rice, 1995, p. 202). For example, a confidence interval at the 95% confidence level for an **estimate** $\hat{\mu}$ would be computed as $\hat{\mu} \pm z_{1-0.025} \times SE(\hat{\mu}) = \hat{\mu} \pm 1.96 \times SE(\hat{\mu})$. A condition for using z to compute a confidence interval is that the sampling distribution for $\hat{\mu}$ is approximately a normal distribution, which we can assume for large **sample** sizes (Särndal et al., 1992).

References

- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control* (third). Upper Saddle River, NJ: Prentice-Hall.
- Casella, G., & Berger, R. L. (2002). *Statistical inference* (2nd ed.). Pacific Grove, CA: Duxbury Press.
- Cochran, W. G. (1977). *Sampling Techniques*. New York, NY: Wiley.
- Dawid, A. P. (1983). *Inference, Statistical: I. Encyclopedia of Statistical Sciences Vol. 4*, S. Kotz, N. L. Johnson and C. B. Read (Eds.). New York, NY: Wiley.
- Gut, A. (2009). *An Intermediate Course in Probability*. New York, NY: Springer.
- IPCC. (2006). *2006 IPCC Guidelines for National Greenhouse Gas Inventories*. H. S. Eggleston, L. Buendia, K. Miwa, T. Ngara, & K. Tanabe (Eds.). Japan: IGES.
- Lillesand, T.M., Kiefer, R.W., Chipman, J.W. (2008). *Remote sensing and image interpretation*, 6th ed. New York, NY: Wiley.
- Olofsson, P., Foody, G. M., Stehman, S. V, & Woodcock, C. E. (2013). Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment*, 129, 122–131.
- Rice, J. A. (1995). *Mathematical statistics and data analysis* (2nd ed.). Belmont, CA: Duxbury Press.
- Särndal, C. E., Svensson, B. H., & Wretman, J. H. (1992). *Model assisted survey sampling*. New York, NY: Springer.
- Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62, 77–89.
- Stehman, S. V, & Wickham, J. D. (2011). Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. *Remote Sensing of Environment*, 115(12), 3044–3055.