**Sampling with replacement with unequal inclusion probabilities**

**Estimating area**

For estimating total area, we apply a stratified estimator that sums the estimated area from each stratum,

(1)

where H=number of strata. is the estimated total area of the variable of interest in stratum *h*,

(2)

where the summation is over *nh* pixels sampled in stratum *h*, *yh,u* = area of sample pixel *u* in stratum *h* if the pixel has the characteristic being estimated (e.g., forest loss due to fire) and *yh,u* = 0 otherwise. For the with-replacement, unequal probability sampling design implemented, the selection probability for pixel *u* in stratum *h* at each independent draw of the sampling protocol is

(3)

where *ah,u* is the area of pixel *u* in stratum *h* and *Ah* is the total area of all pixels in stratum *h*. In effect, for each sample pixel *yh,u*/*ph,u* provides an estimate of the population total, and equation (2) represents the mean of these *nh* estimates. The variance estimator uses the sum of the estimated variances for the H strata (Särndal et al. 1992, p.52)

(4)

where

(5)

For computational purposes, equation 2 can be re-expressed as:

(6)

where *zh,u* = *Ah* if the pixel has the characteristic being estimated (e.g., forest loss due to fire) and *zh,u*=0 otherwise (*Ah* is the total area of stratum *h*). Note that if the pixel has the characteristic of interest, , and because , then .

Substituting and simplifies equation (5) to:

(7)

where is the sample variance for the *nh* values of *zh,u* in stratum *h*.

**Estimating User’s and Producer’s Accuracies**

User’s and producer’s accuracies are estimated by a ratio estimator of the form

(8)

where

(9)

is the estimated total area of variable *y* and

(10)

is the estimated total of variable *z*, *n* is the total sample size across all strata, and is the inclusion probability for pixel *u* in stratum *h*,

(11)

with *nh* = the sample size from stratum *h*. The definitions of *y* and *z* determine the parameter estimated by the ratio. For both user’s accuracy and producer’s accuracy, *yu* = *ah,u* (the area of pixel *u)* if pixel *u* is mapped as class C and has reference class C (i.e., the map and reference class labels agree for class C), otherwise *yu* = 0. For user’s accuracy, *zu* = *ah,u* if pixel *u* is mapped as class C, otherwise *zu* = 0. For producer’s accuracy, *zu* = area of pixel *u* if pixel *u* has reference class C, otherwise *zu* = 0.

The inclusion probability is the probability that pixel *u* would be included in the sample taking into account the full selection of the *nh* sample units in each stratum, whereas the selection probability *ph,u* is the draw-by-draw probability for each sample unit. The same estimates and standard errors will result whether the estimator formulas are expressed using inclusion or selection probabilities. Whereas the standard error of estimated area is easy to compute in terms of selection probabilities, the standard error of the ratio estimator is easier to express using inclusion probabilities. The estimated variance for the ratio estimator is (Särndal et al. 1992, p. 178, Equation 5.6.7)

(12)

where the double summation is over all possible pairs of sample pixels. The pairwise inclusion probability , defined as the probability that both pixels *u* and *v* will be jointly included in the sample, is the product of the inclusion probabilities, because of the with replacement selection protocol. For the case of *u*=*v* (same pixel), . Given the pairwise inclusion probabilities of the sampling design, for all pairs *u*≠*v* and for *u*=v. Then equation (12) simplifies to

(13)

where the summation is over all units in the sample.

**References**

Särndal, C. E., Swensson, B., and Wretman, J. (1992), *Model-Assisted Survey Sampling*. Springer-Verlag, New York.