



# **Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio**

**INFC**

## **IL DISEGNO DI CAMPIONAMENTO**

documento preparato dall'Istituto Sperimentale  
per l'Assestamento Forestale e per l'Alpicoltura



per il Ministero per le Politiche Agricole e Forestali,  
Ispettorato Generale del Corpo Forestale dello Stato



documento predisposto dal prof. Lorenzo Fattorini del Dipartimento Metodi quantitativi dell'Università degli studi di Siena, con contributi di Giovanni Tabacchi

versione n. 1.0 del 2 febbraio 2004

documento registrato in data 16 febbraio 2004 ai sensi della legge 02/02/1939 n. 374

## **SOMMARIO**

### **1 Premessa**

### **2 Impostazione del problema**

### **3 Piano di campionamento**

### **4 Formulazione delle stime**

4.1 Stima delle coperture

4.2 Stima di somme di coperture

4.3 Stima di rapporti di coperture

4.4 Stima delle biomasse

4.5 Stima di somme di biomasse

4.6 Stima di densità di biomasse

### **5 Problemi inerenti l'accorpamento di alcune classi e tipi di uso del suolo**

### **Bibliografia**

## 1 Premessa

Il disegno di campionamento adottato per l'inventario forestale nazionale prevede due modalità di suddivisione del territorio nazionale, i distretti territoriali e le forme di uso del suolo.

La prima impiega un criterio amministrativo, per il quale le Regioni e le Province Autonome sono assunte come unità di disaggregazione territoriale; esse rappresentano le unità minime di evidenziamento dei risultati inventariali e nel documento sono citate come *distretti territoriali*.

La seconda modalità di ripartizione del territorio si basa sulle diverse forme del suo utilizzo e si articola essenzialmente in due livelli: la *classe di uso del suolo* e il *tipo di uso del suolo*, essendo quest'ultimo una ripartizione gerarchica della classe.

Posta in relazione alla nomenclatura effettivamente impiegata nel progetto di inventario, ampiamente illustrata nei documenti *Manuale di fotointerpretazione per la classificazione di prima fase*, *Guida alla classificazione della vegetazione forestale* e *Istruzioni per il rilievo degli attributi di seconda fase*, la generica dizione di *classe di uso del suolo* assumerà il significato di **Sottoclasse di uso del suolo** (ad es. *Formazioni forestali*, *Praterie pascoli e incolti*, *Altre superfici artificiali*, ecc.), che rappresenta nella fattispecie l'effettivo livello nomenclaturale osservabile - su ortofoto carta - con le unità campionarie di prima fase.

Analogamente, la generica dizione di *tipo di uso del suolo* assumerà il significato di **Categoria forestale** (ad es. *Boschi di abete rosso*, *Pinete di pini mediterranei*, *Faggete*, *Castagneti*, ecc.), che in questo caso rappresenta l'effettivo livello nomenclaturale osservabile – al suolo - con le unità campionarie di seconda fase.

Gli altri livelli della nomenclatura adottata per l'inventario forestale nazionale ( le *Classi di uso del suolo*, le *Categorie inventariali* e le *Sottocategorie forestali*) rappresentano utili momenti di aggregazione o disaggregazione delle classificazioni che seguono o che precedono; ovviamente anche queste classificazioni avvengono in una precisa fase inventariale, nella prima per le Classi di uso del suolo, nella seconda per le Categorie inventariali e per le Sottocategorie forestali, senza esserne però la principale connotazione.

Un chiarimento è necessario anche per i termini *copertura* e *biomassa*. Il primo termine assume qui il significato di *peso* di uno strato e con la dizione *stima della copertura* si intende la stima dell'estensione relativa (proporzione sul totale), e quindi assoluta

(superficie in ettari), di uno strato forestale rispetto all'insieme della superficie territoriale. Il termine *biomassa* assume invece il significato generale di attributo quantitativo, cioè questa parola diventa sinonimo di una qualsiasi variabile quantitativa, ad es. il volume commerciale, il numero di alberi, l'incremento corrente legnoso, il peso fresco della chioma, il peso secco dei residui legnosi presenti sul terreno, osservata con le unità di campionamento di terza fase

## 2 Impostazione del problema

Si indichi con  $\mathcal{A}$  l'area dell'intera nazione, pari a  $A = 30\,127\,761$  ettari. Nel territorio nazionale vengono individuati diversi usi del suolo e diversi distretti territoriali. Sia allora  $\mathcal{A}(l,k)$  l'area relativa alla classe di uso del suolo  $l$  nel distretto territoriale  $k$  e sia  $A(l,k)$  la sua superficie, da cui

$$\pi(l,k) = \frac{A(l,k)}{A}$$

rappresenta la copertura di tale area. Inoltre, in corrispondenza di alcune classi di uso del suolo che implicano la presenza di alberi, si indichi con  $T(l,k)$  la biomassa relativa all'area  $\mathcal{A}(l,k)$ .

I valori delle coperture e delle biomasse corrispondenti alle varie classi di uso del suolo ed ai vari distretti territoriali, insieme alle eventuali biomasse, costituiscono le quantità incognite da stimare.

Si noti che possono interessare anche le stesse quantità relative a classi di uso del suolo o a partizioni territoriali più ampie. Allora, se  $L$  indica un insieme di classi di uso del suolo, la sua superficie nel distretto territoriale  $k$  risulta ovviamente

$$A(L,k) = \sum_{l \in L} A(l,k)$$

da cui la copertura risulta

$$\pi(L,k) = \frac{A(L,k)}{A} = \sum_{l \in L} \pi(l,k)$$

così come la biomassa relativa risulta

$$T(L, k) = \sum_{l \in L} T(l, k)$$

In modo del tutto analogo, se  $K$  indica un insieme di distretti territoriali, la superficie della classe di uso del suolo  $l$  in tali distretti risulta

$$A(l, K) = \sum_{k \in K} A(l, k)$$

da cui la copertura risulta

$$\pi(l, K) = \frac{A(l, K)}{A} = \sum_{k \in K} \pi(l, k)$$

e la biomassa risulta

$$T(l, K) = \sum_{k \in K} T(l, k)$$

Infine, se  $L$  indica un insieme di classi di uso del suolo e  $K$  indica un insieme di distretti territoriali, la superficie complessiva di tale insieme in tali distretti risulta

$$A(L, K) = \sum_{l \in L} \sum_{k \in K} A(l, k)$$

da cui la copertura risulta

$$\pi(L, K) = \frac{A(L, K)}{A} = \sum_{l \in L} \sum_{k \in K} \pi(l, k)$$

e la biomassa risulta

$$T(L,K) = \sum_{l \in L} \sum_{k \in K} T(l,k)$$

In sostanza, da tali espressioni risulta che le coperture e le biomasse relative a classi di uso del suolo o a partizioni territoriali più ampie si possono ottenere sempre come somma delle coperture e delle biomasse relative a singole classi o distretti che le compongono. Di conseguenza, quando non è possibile stimare direttamente tali quantità aggregate, le stime si possono comunque ottenere come somma delle stime relative alle singole classi o ai singoli distretti.

### 3 Piano di campionamento

Nella **prima fase** di campionamento l'intera area di studio viene suddivisa in  $N = 301\,278^1$  quadrati di uguale superficie  $A/N = 100$  ettari. Successivamente, in ogni quadrato viene selezionato casualmente un punto  $\mathbf{p}_i$  ( $i = 1, 2, \dots, N$ ). Si indichi con  $U = \{1, 2, \dots, N\}$  la popolazione degli  $N$  punti selezionati nella prima fase. Tale popolazione di punti viene successivamente ripartita, utilizzando anche le ortofoto digitali, in 21 strati corrispondenti ad altrettanti distretti territoriali (le regioni e le due province autonome) in cui è diviso il territorio nazionale, elencati e identificati secondo una numerazione progressiva nella tabella 1, e in 17 strati corrispondenti ad altrettante classi di uso del suolo, elencate e identificate da una numerazione progressiva nella prima e seconda colonna della tabella 2. In uno strato identificato con l'indice 0 sono poi collocati tutti i punti non classificabili tramite ortofoto. Si indichi allora con  $U(l,k)$  ( $l = 0, 1, \dots, 17, k = 1, 2, \dots, 21$ ) lo strato dei punti di prima fase con classe di uso del suolo  $l$  nel distretto territoriale  $k$  e sia  $N(l,k)$  la dimensione di tale strato. Risulta quindi che  $w(l,k) = N(l,k)/N$  rappresenta il peso dello strato  $U(l,k)$  nella popolazione di punti  $U$ . Tale quantità può essere anche interpretata come la stima di prima fase della copertura della classe di uso del suolo  $l$  nel distretto territoriale  $k$  rispetto a tutta l'area di studio.

<sup>1</sup> Ovviamente, a causa della forma irregolare dei confini del territorio nazionale e della linea di costa, questa stima del numero di quadrati e quindi del numero di punti di campionamento ricadenti sul territorio potrà non coincidere con il numero di punti effettivamente posizionati su di esso secondo la procedura impiegata; quest'ultimo numero, infatti, a seguito delle attività di fotointerpretazione, è risultato pari a 301 329. Per un approfondimento sulle modalità pratiche di formulazione del campione di prima fase, si rimanda al documento di progetto *Procedure di definizione delle coordinate dei punti di campionamento*.

Nella **seconda fase**, un campione  $S$  di  $n < N$  punti viene estratto dalla popolazione  $U$  secondo un campionamento stratificato. Si indichi allora con  $S(l,k)$  il campione di punti di seconda fase estratti dallo strato  $U(l,k)$  e sia  $n(l,k)$  la dimensione di tale campione. Differenti tassi di campionamento possono essere effettuati all'interno dei vari strati in relazione all'importanza delle varie classi di uso del suolo. In questo caso per ora si ipotizza che gli strati  $U(1,k)$  ( $k = 1,2,K,21$ ) relativi alle *formazioni forestali* siano campionati con tasso circa pari al 30%, così come gli strati  $U(0,k), U(2,k), U(8,k)$  ( $k = 1,2,K,21$ ) relativi rispettivamente ai *punti non classificabili da ortofoto*, alle *formazioni forestali rade* e all'*arboricoltura da legno*. Gli altri strati  $U(l,k)$  ( $l \neq 0,1,2,8, k = 1,2,K,21$ ) non vengono campionati in seconda fase.

Gli  $n$  punti campionati in seconda fase vengono poi visitati a terra e ulteriormente riclassificati sulla base di una tipologia di uso del suolo più articolata, secondo la classificazione ed i numeri di identificazione riportati nella terza e quarta colonna della tabella 2 e una classe ulteriore in cui sono raggruppati i punti che hanno subito un cambiamento di uso del suolo tra la prima e la seconda fase e i punti non raggiungibili. Si indichi allora con  $H_l$  il numero di tipi in cui viene ripartita la classe di uso del suolo  $l$  e con  $l.h$  un generico tipo di  $l$  ( $l = 1,2,8; h = 1,2,K, H_l$ ). In particolare si ha  $H_1 = 74$ ,  $H_2 = 20$  e  $H_8 = 4$ . Ovviamente, per gli strati non campionati, ovvero quelli relativi alle classi  $l \neq 0,1,2,8$ , non viene effettuata nessuna riclassificazione di seconda fase. Quindi, in questi casi,  $H_l = 1$  da cui l'unico tipo di  $l$  risulta  $l$  stesso. Infine i punti non classificabili in prima fase ( $l = 0$ ) vengono riclassificati a terra in uno dei 112 tipi possibili di seconda fase.

A questo proposito è opportuno sottolineare che alcune classi di uso del suolo possono essere osservate direttamente in prima fase o successivamente nella seconda fase. Per esempio la classe *zone aperte con vegetazione rada o assente* può essere osservata nella prima fase e quindi identificata tramite l'indice  $l = 7$ , oppure può essere osservata, molto meno frequentemente, nella seconda fase tra i punti classificati in prima fase nella classe  $l = 1$ , relativa alle *formazioni forestali*, o nella classe  $l = 2$ , relativa alle *formazioni forestali rade*. In questi casi essa sarà identificata, rispettivamente, tramite l'indice 1.73 e 2.19. Conviene tuttavia tenere distinte queste identificazioni, anche se inerenti alla stessa classe, ed eventualmente ottenere la copertura complessiva come somma delle singole coperture.



Inoltre, è estremamente importante tenere presente che *la classificazione di prima fase viene assunta come non soggetta ad errori*. In realtà, un punto classificato sulla base delle ortofoto nello strato  $U(l,k)$  potrebbe, una volta campionato in seconda fase e visitato a terra, risultare appartenente ad un altro strato. Data tuttavia l'approssimazione esistente nella localizzazione dei punti sul terreno, non si è in grado di discernere se questa diversa classificazione debba essere attribuita ad una errata classificazione di prima fase, ad un cambiamento nell'uso del suolo intervenuto nel frattempo o ad una poco accurata identificazione del punto a terra. Si preferisce allora attribuire tale eventuale discordanza alle due ultime cause e considerare queste particolari situazioni in modo tale che, come verrà meglio spiegato nel documento delle *Istruzioni per il rilievo degli attributi di seconda fase*, il punto in esame continui ad appartenere allo strato in cui era stato classificato in prima fase, assumendo come realtà quanto visto sull'ortofoto digitale.

Questo modo di procedere ha delle conseguenze rilevanti. In particolare questo implica che un punto di seconda fase classificato a terra nel tipo  $l.h$  ( $l = 1,2,8$ ) del distretto territoriale  $k$  può essere solo un punto di  $S(l,k)$  o un punto di  $S(0,k)$ . In pratica, al termine della seconda fase di campionamento il campione  $S(l,k)$  risulta suddiviso in un massimo di  $H_l$  campioni  $S(l.h,k)$ , ciascuno dei quali è costituito da  $n(l.h,k)$  punti classificati nel tipo  $l.h$  ( $l = 1,2,8, k = 1,2,K,21$ ), mentre il campione  $S(0,k)$  risulta suddiviso in un massimo di 112 campioni indicati con  $S_0(l.h,k)$ , corrispondenti a tutti i possibili tipi di seconda fase, ciascuno dei quali è costituito da  $n_0(l.h,k)$  punti ( $l = 1,2,K,17, h = 1,2,K, H_l, k = 1,2,K,21$ ).

Si ponga ora  $H_1 = 73$ ,  $H_2 = 17$ ,  $H_8 = 3$ , dal momento che i tipi 1.74, 2.20 e 8.4 sono escluse dall'indagine successiva riferendosi a punti non raggiungibili o con uso del suolo variato, mentre i tipi 2.18 e 2.19 non implicano presenza di biomasse. Nella **terza fase** di campionamento, da ciascun campione  $S(l.h,k)$  ( $l = 1,2,8, h = 1,2,K, H_l, k = 1,2,K,21$ ) e da ciascun campione  $S_0(l.h,k)$  ( $l = 1,2,8, h = 1,2,K, H_l, k = 1,2,K,21$ ), ovvero per i campioni relativi a classi e tipi di uso del suolo per cui è necessario stimare la biomassa, viene selezionato casualmente un ulteriore campione, utilizzando tassi di campionamento ancora da stabilire con precisione ma che potranno aggirarsi anche in questo caso intorno al 30%. Si indichi con  $Q(l.h,k)$  il campione estratto da  $S(l.h,k)$  e si indichi con  $m(l.h,k)$  la sua dimensione; in modo analogo si indichi con  $Q_0(l.h,k)$  il campione estratto da  $S_0(l.h,k)$  e si indichi con  $m_0(l.h,k)$  la sua dimensione. In ognuno dei punti selezionati nella terza

fase, si delimita poi un'area circolare (*plot*) di raggio  $r = 13.8$  m e superficie  $a = 600$  m<sup>2</sup> e in questa area si misurano i diametri di tutti gli alberi con circonferenza a 1.3 m dal suolo superiore a 7.85 cm. Supponendo quindi una relazione deterministica tra diametro e volume, si ricavano dalle tabelle i rispettivi volumi, da cui, per somma si ottiene la biomassa totale presente nel *plot*. Dal momento che la relazione tra diametro e volume non è in effetti deterministica ma stocastica, anche questa fase di rilevamento presenta una variabilità che sarà tuttavia ignorata in quanto per ora non quantificabile sulla base dell'informazione campionaria.

In relazione a questa terza fase è importante sottolineare che in ognuno degli  $m(l,h,k)$  *plot* centrati nei punti di  $Q(l,h,k)$  e in ognuno degli  $m_0(l,h,k)$  *plot* centrati nei punti di  $Q_0(l,h,k)$  ( $l = 1,2,8$ ,  $h = 1,2,K$ ,  $H_l$ ,  $k = 1,2,K,21$ ), viene rilevata solo la biomassa relativa al tipo di uso del suolo  $l,h$ , sebbene possano in realtà essere presenti anche biomasse relative ad altri tipi. In questo modo si assume tacitamente che le biomasse relative a tutti gli altri tipi siano nulle entro tali *plot*. Tuttavia, tale assunzione non dovrebbe causare distorsioni rilevanti in quanto, data la dimensione ridotta dei *plot*, le biomasse relative a tipi di uso del suolo diversi da  $l,h$  dovrebbero risultare trascurabili nei *plot* il cui centro cade nel tipo  $l,h$ .

Si indichi allora con  $b_i(l,h,k)$  la biomassa del tipo  $l,h$  rilevata nel *plot* centrato in  $p_i$ , con  $i \in Q(l,h,k)$  oppure  $i \in Q_0(l,h,k)$ . Si può provare allora che

$$t_i(l,h,k) = \frac{A}{N} \frac{b_i(l,h,k)}{a} = 66.6 \bar{b}_i(l,h,k)$$

costituisce, a parte effetti di bordo (che sono comunque trascurabili data la dimensione esigua dei *plot* rispetto ai quadrati), una stima corretta della biomassa del tipo di uso del suolo  $l,h$  nell'intero quadrato  $i$  (cfr. DE VRIES, 1986, pp.215-216; FATTORINI e PISANI, 1999, pp.78-85).

Quest'ultima fase conclude il campionamento, dal momento che l'informazione acquisita nelle tre fasi consente di ottenere stime delle coperture e delle eventuali biomasse di tutte le classi e tipi di uso del suolo in qualunque distretto territoriale, insieme alle stime della loro variabilità campionaria.

## 4 Formulazione delle stime

### 4.1 Stima delle coperture

L'informazione acquisita nelle prime due fasi di campionamento è completamente sufficiente per stimare le coperture dei vari tipi di uso del suolo nei vari distretti territoriali insieme alle loro varianze campionarie.

Si indichi con  $\pi(l,h,k)$  ( $l = 1,2,K,17, h = 1,2,K, H_l, k = 1,2,K, 21$ ) la copertura del tipo di uso del suolo  $l.h$  nel distretto territoriale  $k$ . Una stima corretta di tale quantità si può ottenere applicando l'espressione (14) contenuta in FATTORINI, MARCHESELLI e PISANI (2003a) che fornisce stimatori corretti delle coperture tramite campionamenti in due fasi, in situazioni molto generali in cui sono supposti errori di classificazione di prima fase corretti tramite campionamenti di seconda fase.

L'assunzione che le classificazioni di prima fase siano effettuate senza errori (cfr. paragrafo 3) semplifica notevolmente le espressioni delle stime in questione. In questo caso, infatti, per  $l = 1,2,8$ , una stima **corretta** di  $\pi(l,h,k)$  risulta

$$(1) \quad \hat{\pi}(l,h,k) = w(l,k)f(l,h,k) + w(0,k)f_0(l,h,k) \quad l = 1,2,8$$

dove

$$f(l,h,k) = \frac{n(l,h,k)}{n(l,k)}$$

e

$$f_0(l,h,k) = \frac{n_0(l,h,k)}{n(0,k)}$$

rappresentano rispettivamente la frazione di punti di  $S(l,k)$  e la frazione di punti di  $S(0,k)$  classificati in seconda fase nel tipo di uso del suolo  $l.h$ .

Si deve tuttavia sottolineare che l'espressione (1) vale solo per  $l = 1,2,8$ , ovvero vale per gli strati che, oltre allo strato 0, vengono campionati in seconda fase. Per gli strati

corrispondenti a classi non campionate in seconda fase ( $l \neq 1,2,8$ ), data l'assenza di errori di classificazione di prima fase, tutti i punti di un eventuale campione  $S(l,k)$  risulterebbero riclassificati nella stessa classe  $l$ , da cui, ovviamente  $f(l,k) = 1$ . In questi casi allora, la stima di  $\pi(l,k)$  si ottiene dalla (1) semplicemente ponendo  $f(l,k) = 1$ , ovvero

$$(2) \quad \hat{\pi}(l,k) = w(l,k) + w(0,k)f_0(l,k) \quad l \neq 1,2,8$$

dove

$$f_0(l,k) = \frac{n_0(l,k)}{n(0,k)}$$

rappresenta la frazione di punti di  $S(0,k)$  classificati in seconda fase nella classe di uso del suolo  $l$ .

Per quanto riguarda la valutazione della precisione di tali stime, FATTORINI, MARCHESELLI e PISANI (2003a) riportano le espressioni delle stime conservative della matrice delle varianze e covarianze, ancora nel caso molto generale in cui sono contemplati errori di classificazione di prima fase. Nel caso in questione invece, per  $l = 1,2,8$ , la stima **conservativa** della varianza di  $\hat{\pi}(l,h,k)$  si riduce a

$$(3) \quad v^2(l,h,k) = \frac{1}{N-1} \left\{ w(l,k) \frac{N(l,k)-1}{n(l,k)-1} f(l,h,k) + w(0,k) \frac{N(0,k)-1}{n(0,k)-1} f_0(l,h,k) - \right. \\ \left. w(l,k) \frac{N(l,k)-n(l,k)}{n(l,k)-1} f^2(l,h,k) - w(0,k) \frac{N(0,k)-n(0,k)}{n(0,k)-1} f_0^2(l,h,k) - \hat{\pi}^2(l,h,k) \right\}$$

mentre, per i motivi precedentemente detti, l'espressione della stima della varianza di  $\hat{\pi}(l,k)$  con  $l \neq 1,2,8$  si riduce ulteriormente a

$$(4) \quad v^2(l,k) =$$

$$\frac{1}{N-1} \left\{ w(l,k) + w(0,k) \frac{N(0,k)-1}{n(0,k)-1} f_0(l,k) - w(0,k) \frac{N(0,k)-n(0,k)}{n(0,k)-1} f_0^2(l,k) - \hat{\pi}^2(l,k) \right\}$$

Dal momento che, come già detto nel paragrafo 2, interessa stimare anche le coperture relative a tipi di uso del suolo o distretti territoriali più ampi, e dal momento che in alcuni casi tali stime si possono ottenere convenientemente come somma delle stime relative ai singoli tipi o ai singoli distretti, si rende necessario stimare le covarianze tra le stime coinvolte nella somma al fine di valutare la varianza della stima finale.

Sempre dalle formule generali derivate da FATTORINI, MARCHESELLI e PISANI (2003a), l'espressione della stima della covarianza tra  $\hat{\pi}(l.h,k)$  e  $\hat{\pi}(l'.h',k')$ , per  $k \neq k'$ , risulta

$$(5) \quad c(l.h,k; l'.h',k') = -\frac{1}{N-1} \hat{\pi}(l.h,k) \hat{\pi}(l'.h',k')$$

dove, ovviamente,  $l.h$  è sostituito da  $l$  quando  $l \neq 1,2,8$  e  $l'.h'$  è sostituito da  $l'$  quando  $l' \neq 1,2,8$ . Se invece  $k = k'$  ma  $l \neq l'$ , la covarianza risulta

$$(6) \quad c(l.h,k; l'.h',k) = -\frac{1}{N-1} \left\{ w(0,k) \frac{N(0,k)-n(0,k)}{n(0,k)-1} f_0(l.h,k) f_0(l'.h',k) + \hat{\pi}(l.h,k) \hat{\pi}(l'.h',k) \right\}$$

dove, anche in questo caso,  $l.h$  è sostituito da  $l$  quando  $l \neq 1,2,8$  e  $l'.h'$  è sostituito da  $l'$  quando  $l' \neq 1,2,8$ . Se infine  $k = k'$ ,  $l = l'$  con  $l = 1,2,8$  e  $h \neq h'$ , la covarianza risulta

$$(7) \quad c(l.h,k; l.h',k) = -\frac{1}{N-1} \left\{ w(l,k) \frac{N(l,k)-n(l,k)}{n(l,k)-1} f(l.h,k) f(l.h',k) + w(0,k) \frac{N(0,k)-n(0,k)}{n(0,k)-1} f_0(l.h,k) f_0(l.h',k) - \hat{\pi}(l.h,k) \hat{\pi}(l.h',k) \right\}$$

Ovviamente, qualora interessi stimare la superficie  $A(l.h,k)$  relativa al tipo di uso del suolo  $l.h$  nel distretto  $k$ , la stima risulta  $\hat{A}(l.h,k) = A \hat{\pi}(l.h,k)$ , mentre una stima della corrispondente varianza risulta  $A^2 v^2(l.h,k)$ .

## 4.2 Stima di somme di coperture

Dal momento che alcuni dei 112 tipi di uso del suolo di seconda fase possono essere accorpati in insiemi più ampi, come quelli elencati nella colonna 5 di tabella 2, può interessare la stima della copertura di tali insiemi a livello di singolo distretto. Per  $l = 1,2,8$ , si indichi allora con  $H$  l'insieme di indici che identificano un insieme di tipi di  $l$  e si indichi con  $\pi(l,H,k)$  la copertura di tale insieme nel distretto territoriale  $k$  ( $k=1,2,K,21$ ). Ovviamente, se  $H = \{1,2,K, H_l\}$ , l'insieme  $l.H$  coincide con tutta la classe  $l$ . Tenendo conto di quanto detto nel paragrafo 2, la stima **corretta** di tale copertura si potrebbe ottenere dalla stima delle coperture dei singoli tipi che le compongono, ovvero

$$\hat{\pi}(l,H,k) = \sum_{h \in H} \hat{\pi}(l,h,k)$$

Tuttavia, dal momento che tale stima si può ottenere classificando opportunamente i campioni  $S(l,k)$  e  $S(0,k)$ , essa può anche essere ottenuta direttamente tramite

$$(8) \quad \hat{\pi}(l,H,k) = w(l,k)f(l,H,k) + w(0,k)f_0(l,H,k)$$

dove

$$f(l,H,k) = \frac{n(l,H,k)}{n(l,k)}$$

e

$$f_0(l,H,k) = \frac{n_0(l,H,k)}{n(0,k)}$$

rappresentano rispettivamente la frazione di punti di  $S(l,k)$  e la frazione di punti di  $S(0,k)$  classificati in seconda fase nell'insieme di tipi di uso del suolo  $l.H$ . In modo analogo, la stima **conservativa** della varianza di  $\hat{\pi}(l,H,k)$  risulta

$$(9) \quad v^2(I,H,k) = \frac{1}{N-1} \left\{ w(I,k) \frac{N(I,k)-1}{n(I,k)-1} f(I,H,k) + w(0,k) \frac{N(0,k)-1}{n(0,k)-1} f_0(I,H,k) - \right. \\ \left. w(I,k) \frac{N(I,k)-n(I,k)}{n(I,k)-1} f^2(I,H,k) - w(0,k) \frac{N(0,k)-n(0,k)}{n(0,k)-1} f_0^2(I,H,k) - \hat{\pi}^2(I,H,k) \right\}$$

Per gli sviluppi successivi è necessario determinare anche l'espressione della stima della covarianza tra  $\hat{\pi}(I,H,k)$  e  $\hat{\pi}(I,H,k')$ , per  $k \neq k'$ , che risulta

$$(10) \quad c(I,H,k; I,H,k') = -\frac{1}{N-1} \hat{\pi}(I,H,k) \hat{\pi}(I,H,k')$$

Ovviamente, in tali espressioni  $I.H$  è sostituito da  $I$  se  $H = \{1,2,K, H_l\}$ . In questo caso, allora, le espressioni (8) e (9) coincidono rispettivamente con le espressioni (2) e (4).

Inoltre, dal momento che i 21 distretti territoriali in cui è ripartita l'Italia possono essere accorpati in alcune aree geografiche, può interessare la stima delle coperture dei 112 tipi di uso del suolo a livello di area geografica. Si indichi allora con  $K_j$  gli insiemi di indici che identificano i distretti che afferiscono alle  $j$  aree geografiche e con  $\pi(I.h, K_j)$  ( $I = 1,2,K, 17, h = 1,2,K, H_l$ ) la copertura del tipo di uso del suolo  $I.h$  nell'area geografica  $j$ . In questo caso, dal momento che l'informazione deriva da campionamenti indipendenti effettuati in seconda fase nei singoli distretti, la stima **corretta** di tale copertura si può ottenere **solo** come somma delle coperture nei singoli distretti, ovvero

$$(11) \quad \hat{\pi}(I.h, K_j) = \sum_{k \in K_j} \hat{\pi}(I.h, k)$$

mentre, per le note proprietà della varianza di somme di variabili casuali, una stima **conservativa** della varianza campionaria di  $\hat{\pi}(I.h, K_j)$  risulta

$$(12) \quad v^2(I.h, K_j) = \sum_{k \in K_j} v^2(I.h, k) + 2 \sum_{k \in K_j} \sum_{k' > k \in K_j} c(I.h, k; I.h, k')$$

dove, ovviamente,  $I.h$  è sostituito da  $I$  quando  $I \neq 1,2,8$ .

Qualora interessi invece la stima delle coperture dei 112 tipi di uso del suolo a livello nazionale, si indichi con  $\pi(l.h)$  ( $l = 1,2,K,17, h = 1,2,K,H_l$ ) la copertura del tipo di uso del suolo  $l.h$  nell'intera nazione. La stima **corretta** di  $\pi(l.h)$  si ottiene tramite

$$(13) \quad \hat{\pi}(l.h) = \sum_{k=1}^{21} \hat{\pi}(l.h, k)$$

mentre una stima **conservativa** della varianza campionaria di  $\hat{\pi}(l.h)$  risulta

$$(14) \quad v^2(l.h) = \sum_{k=1}^{21} v^2(l.h, k) + 2 \sum_{k=1}^{21} \sum_{k' > k=1}^{21} c(l.h, k; l.h, k')$$

dove, anche in questo caso,  $l.h$  è sostituito da  $l$  quando  $l \neq 1,2,8$ . Si noti che nelle espressioni (12) e (14), che sono analoghe, i valori di  $v^2(l.h, k)$  e  $c(l.h, k; l.h, k')$  si ottengono rispettivamente dalle espressioni (3) e (5).

Inoltre, per  $l = 1,2,8$ , può interessare la stima della copertura a livello nazionale dell'insieme di tipi di uso del suolo  $l.H$ . Si indichi allora con  $\pi(l.H, K_j)$  la copertura dell'insieme  $l.H$  nell'area geografica  $j$ . La stima **corretta** di  $\pi(l.H, K_j)$  si ottiene tramite

$$(15) \quad \hat{\pi}(l.H, K_j) = \sum_{k \in K_j} \hat{\pi}(l.H, k)$$

mentre una stima **conservativa** della varianza campionaria di  $\hat{\pi}(l.H, K_j)$  risulta

$$(16) \quad v^2(l.H, K_j) = \sum_{k \in K_j} v^2(l.H, k) + 2 \sum_{k \in K_j} \sum_{k' > k \in K_j} c(l.H, k; l.H, k')$$

con  $l.H$  è sostituito da  $l$  quando  $H = \{1,2,K, H_l\}$ .

Infine, sempre per  $l = 1,2,8$ , qualora interessi la stima della copertura a livello nazionale dell'insieme di tipi di uso del suolo  $l.H$ , si indichi con  $\pi(l.H)$  la copertura dell'insieme  $l.H$  nell'intera nazione. La stima **corretta** di  $\pi(l.H)$  si ottiene tramite



$$(17) \quad \hat{\pi}(I.H) = \sum_{k=1}^{21} \hat{\pi}(I.H, k)$$

mentre una stima **conservativa** della varianza campionaria di  $\hat{\pi}(I.H)$  risulta

$$(18) \quad v^2(I.H) = \sum_{k=1}^{21} v^2(I.H, k) + 2 \sum_{k=1}^{21} \sum_{k' > k=1}^{21} c(I.H, k; I.H, k')$$

con  $I.H$  sostituito da  $I$  quando  $H = \{1, 2, K, H_i\}$ . Si noti che nelle espressioni (15) e (17), che sono analoghe, i valori di  $\hat{\pi}(I.H, k)$  si possono ottenere direttamente dall'espressione (8), mentre nelle espressioni (16) e (18), anch'esse analoghe, i valori di  $v^2(I.H, k)$  e  $c(I.H, k; I.H, k')$  si ottengono rispettivamente dalle espressioni (9) e (10).

Ovviamente in questo paragrafo sono state considerate per brevità solo le aggregazioni spaziali o tipologiche che più comunemente interessano. Tuttavia, qualora interessino altri tipi di aggregazioni, il modo di procedere risulta, *mutatis mutandi*, del tutto simile a quello delineato in questa sezione. Ovvero, quando non è possibile operare direttamente, la stima della copertura aggregata si ottiene come somma delle stime delle singole coperture mentre la stima della varianza si ottiene come somma delle varianze e delle covarianze delle singole stime coinvolte nella somma.

#### 4.3 Stima di rapporti di coperture

Per i tipi di uso del suolo classificati in seconda fase, ovvero per  $I = 1, 2, 8$ , sia  $\pi(I.H, k)$  la copertura dell'insieme di tipi di uso del suolo  $I.H$  nel distretto territoriale  $k$  e, per un dato  $h \in H$ , sia  $\pi(I.h, k)$  la copertura della tipologia  $I.h$  nel medesimo distretto territoriale. Si può allora essere interessati a stimare il rapporto

$$\theta(I.h, k | H) = \frac{\pi(I.h, k)}{\pi(I.H, k)}$$

che rappresenta ovviamente la copertura del tipo  $l.h$  all'interno dell'insieme  $l.H$  nel distretto  $k$ .

Risulta quindi naturale stimare  $\theta(l.h, k | H)$  tramite il rapporto delle stime

$$(19) \quad \hat{\theta}(l.h, k | H) = \frac{\hat{\pi}(l.h, k)}{\hat{\pi}(l.H, k)}$$

dove  $\hat{\pi}(l.H, k)$  può essere ottenuta direttamente tramite la (8). Ovviamente, essendo  $\hat{\theta}(l.h, k | H)$  un rapporto di stime, non sono note in modo esatto le sue proprietà. Tuttavia, approssimando  $\hat{\theta}(l.h, k | H)$  in serie di Taylor sino al termine di primo grado, risulta immediato provare che  $\hat{\theta}(l.h, k | H)$  è una stima **approssimativamente corretta** di  $\theta(l.h, k | H)$ , mentre una stima **approssimativamente conservativa** della sua varianza risulta

$$(20) \quad v^2(l.h, k | H) = \frac{v^2(l.h, k)}{\hat{\pi}^2(l.H, k)} + \frac{\hat{\pi}^2(l.h, k)}{\hat{\pi}^4(l.H, k)} v^2(l.H, k) - 2 \frac{\hat{\pi}(l.h, k)}{\hat{\pi}^3(l.H, k)} c(l.h, k; l.H, k)$$

dove

$$(21) \quad c(l.h, k; l.H, k) = v^2(l.h, k) + \sum_{h' \neq h \in H} c(l.h, k; l.h', k)$$

costituisce una stima della covarianza tra  $\hat{\pi}(l.h, k)$  e  $\hat{\pi}(l.H, k)$ . Si noti che nelle espressioni (20) e (21), i valori di  $v^2(l.h, k)$ ,  $c(l.h, k; l.h', k)$  e  $v^2(l.H, k)$  si ottengono rispettivamente dalle espressioni (3), (7) e (9).

Tuttavia, la copertura del tipo di uso del suolo  $l.h$  rispetto all'insieme di tipi  $l.H$  può interessare non solo a livello dei singoli distretti ma anche a livello di area geografica o nazionale. In altri termini si può essere interessati a stimare

$$\theta(l.h, K_j | H) = \frac{\pi(l.h, K_j)}{\pi(l.H, K_j)}$$

oppure si può essere interessati a stimare

$$\theta(l.h | H) = \frac{\pi(l.h)}{\pi(l.H)}$$

In questi casi le formule delle stime e delle stime delle varianze relative sono analoghe, *mutatis mutandi*, a quelle ottenute a livello di singolo distretto. In particolare una stima **approssimativamente corretta** di  $\theta(l.h, K_j | H)$  risulta

$$(22) \quad \hat{\theta}(l.h, K_j | H) = \frac{\hat{\pi}(l.h, K_j)}{\hat{\pi}(l.H, K_j)}$$

dove  $\hat{\pi}(l.h, K_j)$  e  $\hat{\pi}(l.H, K_j)$  si ottengono rispettivamente dalle espressioni (11) e (15). Inoltre, una stima **approssimativamente conservativa** della sua varianza risulta

$$(23) \quad v^2(l.h, K_j | H) = \frac{v^2(l.h, K_j)}{\hat{\pi}^2(l.H, K_j)} + \frac{\hat{\pi}^2(l.h, K_j)}{\hat{\pi}^4(l.H, K_j)} v^2(l.H, K_j) - 2 \frac{\hat{\pi}(l.h, K_j)}{\hat{\pi}^3(l.H, K_j)} c(l.h, K_j; l.H, K_j)$$

dove

$$(24) \quad c(l.h, K_j; l.H, K_j) = v^2(l.h, K_j) + \sum_{h' \neq h \in H} c(l.h, K_j; l.h', K_j)$$

costituisce una stima della covarianza tra  $\hat{\pi}(l.h, K_j)$  e  $\hat{\pi}(l.H, K_j)$ , mentre, a sua volta,

$$(25) \quad c(l.h, K_j; l.h', K_j) = \sum_{k \in K_j} \sum_{k' \in K_j} c(l.h, k; l.h', k')$$

costituisce una stima della covarianza tra  $\hat{\pi}(l.h, K_j)$  e  $\hat{\pi}(l.h', K_j)$ . Si noti che nelle espressioni (23), (24) e (25) i valori di  $v^2(l.h, K_j)$ ,  $v^2(l.H, K_j)$  e  $c(l.h, k; l.h', k')$  sono dati rispettivamente dalle espressioni (12), (16) e (5).

In modo del tutto analogo, una stima **approssimativamente corretta** di  $\theta(l.h | H)$  risulta

$$(26) \quad \hat{\theta}(l.h | H) = \frac{\hat{\pi}(l.h)}{\hat{\pi}(l.H)}$$

dove  $\hat{\pi}(l.h)$  e  $\hat{\pi}(l.H)$  si ottengono rispettivamente dalle espressioni (13) e (17). Inoltre una stima **approssimativamente conservativa** della sua varianza risulta

$$(27) \quad v^2(l.h | H) = \frac{v^2(l.h)}{\hat{\pi}^2(l.H)} + \frac{\hat{\pi}^2(l.h)}{\hat{\pi}^4(l.H)} v^2(l.H) - 2 \frac{\hat{\pi}(l.h)}{\hat{\pi}^3(l.H)} c(l.h; l.H)$$

dove

$$(28) \quad c(l.h; l.H) = v^2(l.h) + \sum_{h' \neq h \in H} c(l.h; l.h')$$

costituisce una stima della covarianza tra  $\hat{\pi}(l.h)$  e  $\hat{\pi}(l.H)$ , mentre, a sua volta,

$$(29) \quad c(l.h; l.h') = \sum_{k=1}^{21} \sum_{k'=1}^{21} c(l.h, k; l.h', k')$$

costituisce una stima della covarianza tra  $\hat{\pi}(l.h)$  e  $\hat{\pi}(l.h')$ . Si noti che nelle espressioni (27), (28) e (29) i valori di  $v^2(l.h)$ ,  $v^2(l.H)$  e  $c(l.h, k; l.h', k')$  sono dati rispettivamente dalle espressioni (14), (18) e (5).

Ovviamente, in tutte le espressioni di questo paragrafo,  $l.H$  è sostituito da  $l$  quando  $H = \{1, 2, K, H_l\}$

#### 4.4 Stima delle biomasse

Si tenga presente che in terza fase  $H_1 = 73$ ,  $H_2 = 19$  e  $H_8 = 3$ . Dato allora il campione  $Q(l,h,k)$  ( $l = 1,2,8$ ,  $h = 1,2,K$ ,  $H_l$ ,  $k = 1,2,K,21$ ) selezionato in terza fase e date le stime delle biomasse  $t_i(l,h,k)$  con  $i \in Q(l,h,k)$ , si indichi con

$$\bar{t}(l,h,k) = \frac{1}{m(l,h,k)} \sum_{i \in Q(l,h,k)} t_i(l,h,k)$$

la loro media campionaria e con

$$s^2(l,h,k) = \frac{1}{m(l,h,k) - 1} \sum_{i \in Q(l,h,k)} \{t_i(l,h,k) - \bar{t}(l,h,k)\}^2$$

la loro varianza campionaria.

In modo del tutto analogo, dato il campione  $Q_0(l,h,k)$  ( $l = 1,2,8$ ,  $h = 1,2,K$ ,  $H_l$ ,  $k = 1,2,K,21$ ) e date le stime delle biomasse  $t_i(l,h,k)$  con  $i \in Q_0(l,h,k)$ , si indichi con

$$\bar{t}_0(l,h,k) = \frac{1}{m_0(l,h,k)} \sum_{i \in Q_0(l,h,k)} t_i(l,h,k)$$

la loro media campionaria e con

$$s_0^2(l,h,k) = \frac{1}{m_0(l,h,k) - 1} \sum_{i \in Q_0(l,h,k)} \{t_i(l,h,k) - \bar{t}_0(l,h,k)\}^2$$

la loro varianza campionaria.

L'informazione acquisita nelle tre fasi di campionamento è completamente sufficiente per stimare le biomasse relative ai vari tipi di uso del suolo nei vari distretti territoriali insieme alle loro varianze campionarie. Si indichi con  $T(l,h,k)$  ( $l = 1,2,8$ ,  $h = 1,2,K$ ,  $H_l$ ,  $k = 1,2,K,21$ ) la biomassa del tipo di uso del suolo  $l,h$  nel distretto territoriale  $k$ . Una stima corretta di tale

quantità si può ottenere applicando i risultati teorici contenuti in FATTORINI, MARCHESELLI e PISANI (2003b), che forniscono stimatori corretti delle biomasse tramite campionamenti in tre fasi, in situazioni molto generali in cui sono supposti errori di classificazione di prima fase corretti tramite campionamenti di seconda fase ed in cui biomasse relative ad un dato tipo possono essere rilevate anche in *plot* il cui centro cade in altri tipi di uso del suolo.

Le assunzioni che le classificazioni di prima fase sono effettuate senza errori e che biomasse relative ad un dato tipo possono essere rilevate solo nei *plot* il cui centro cade nello stesso tipo (cfr. paragrafo 3), semplifica notevolmente le espressioni delle stime in questione. In questo caso, infatti, una stima **corretta** di  $T(l,h,k)$  risulta

$$(30) \quad \hat{T}(l,h,k) = N(l,k)f(l,h,k)\bar{t}(l,h,k) + N(0,k)f_0(l,h,k)\bar{t}_0(l,h,k)$$

Per quanto riguarda la valutazione della precisione di tali stime, FATTORINI, MARCHESELLI e PISANI (2003b) riportano le espressioni delle stime conservative della matrice delle varianze e covarianze, ancora nel caso generale. Nel caso in questione, invece, la stima **conservativa** della varianza di  $\hat{T}(l,h,k)$  risulta

$$(31) \quad V^2(l,h,k) = \frac{N}{N-1} \left\{ \frac{N(l,k)[N(l,k)-1]}{n(l,k)-1} \frac{n(l,h,k)-1}{m(l,h,k)} f(l,h,k) s^2(l,h,k) + \right. \\ + \frac{N(0,k)[N(0,k)-1]}{n(0,k)-1} \frac{n_0(l,h,k)-1}{m_0(l,h,k)} f_0(l,h,k) s_0^2(l,h,k) + \\ + \frac{N(l,k)[N(l,k)-1]}{n(l,k)-1} f(l,h,k)[1-f(l,h,k)] \bar{t}^2(l,h,k) + \\ + \frac{N(0,k)[N(0,k)-1]}{n(0,k)-1} f_0(l,h,k)[1-f_0(l,h,k)] \bar{t}_0^2(l,h,k) + \\ \left. + N(l,k)f^2(l,h,k) \bar{t}^2(l,h,k) + N(0,k)f_0^2(l,h,k) \bar{t}_0^2(l,h,k) - \frac{\hat{T}^2(l,h,k)}{N} \right\}$$

Dal momento che, come già detto nel paragrafo 2, interessa stimare anche le biomasse relative a tipi di uso del suolo o distretti territoriali più ampi, e dal momento che l'informazione proviene da campioni di terza fase tutti indipendenti, tali stime si possono ottenere solo come somma delle stime relative ai singoli tipi o ai singoli distretti. Si rende quindi necessario stimare le covarianze tra le stime coinvolte nella somma al fine di valutare la varianza della stima finale.

Sempre dalle formule generali derivate da FATTORINI, MARCHESELLI e PISANI (2003b), l'espressione della stima della covarianza tra  $\hat{T}(l.h,k)$  e  $\hat{T}(l'.h',k')$ , per  $k \neq k'$ , risulta

$$(32) \quad C(l.h,k; l'.h',k') = -\frac{\hat{T}(l.h,k)\hat{T}(l'.h',k')}{N-1}$$

Se invece  $k = k'$  ma  $l \neq l'$ , la covarianza risulta

$$(33) \quad C(l.h,k; l'.h',k) = -\frac{N}{N-1} \left\{ \frac{N(0,k)[N(0,k)-n(0,k)]}{n(0,k)-1} f_0(l.h,k)f_0(l'.h',k) \bar{t}_0(l.h,k)\bar{t}_0(l'.h',k) + \frac{\hat{T}(l.h,k)\hat{T}(l'.h',k)}{N} \right\}$$

Se infine  $k = k'$ ,  $l = l'$  e  $h \neq h'$ , la covarianza risulta

$$(34) \quad C(l.h,k; l.h',k) = -\frac{N}{N-1} \left\{ \frac{N(l,k)[N(l,k)-n(l,k)]}{n(l,k)-1} f(l.h,k)f(l.h',k) \bar{t}(l.h,k)\bar{t}(l.h',k) + \frac{N(0,k)[N(0,k)-n(0,k)]}{n(0,k)-1} f_0(l.h,k)f_0(l.h',k) \bar{t}_0(l.h,k)\bar{t}_0(l.h',k) + \frac{\hat{T}(l.h,k)\hat{T}(l.h',k)}{N} \right\}$$

#### 4.5 Stima di somme di biomasse

In analogia a quanto fatto nel paragrafo 4.2, per  $l = 1,2,8$  si indichi con  $H$  l'insieme di indici che identificano un insieme di tipi di  $l$  e si indichi con  $T(l,H,k)$  la copertura di tale insieme nel distretto territoriale  $k$  ( $k = 1,2,K,21$ ). Si tenga ancora presente che se  $H = \{1,2,K, H_l\}$ , l'insieme  $l,H$  coincide con tutta la classe  $l$ . Dal momento che le stime delle biomasse  $T(l,h,k)$  sono tutte ottenute da campionamenti indipendenti di terza fase, la stima **corretta** di tale quantità si può ottenere solo come somma dalle stime delle biomasse dei singoli tipi che compongono l'insieme  $l,H$ , ovvero

$$(35) \quad \hat{T}(l,H,k) = \sum_{h \in H} \hat{T}(l,h,k)$$

mentre la stima **conservativa** della varianza di  $\hat{T}(l,H,k)$  risulta

$$(36) \quad V^2(l,H,k) = \sum_{h \in H} V^2(l,h,k) + 2 \sum_{h \in H} \sum_{h' \neq h \in H} C(l,h,k; l,h',k)$$

Per gli sviluppi successivi è necessario determinare anche l'espressione della stima della covarianza  $\hat{T}(l,H,k)$  e  $\hat{T}(l,H,k')$ , per  $k \neq k'$ , che risulta

$$(37) \quad C(l,H,k; l,H,k') = \sum_{h \in H} \sum_{h' \in H} C(l,h,k; l,h',k')$$

Si noti che nelle espressioni (36) e (37) i valori di  $V^2(l,h,k)$ ,  $C(l,h,k; l,h',k')$  e  $C(l,h,k; l,h',k)$  sono forniti rispettivamente dalle espressioni (31), (32) e (34).

Se inoltre interessa la stima delle biomasse dei vari tipi di uso del suolo a livello di area geografica, si indichi con  $T(l,h,K_j)$  ( $l = 1,2,8$ ,  $h = 1,2,K, H_l$ ) la biomassa del tipo di uso del suolo  $l,h$  nell'area geografica  $j$ . Anche in questo caso la stima **corretta** di tale quantità si può ottenere solo come somma delle biomasse nei singoli distretti, ovvero

$$(38) \quad \hat{T}(l,h,K_j) = \sum_{k \in K_j} \hat{T}(l,h,k)$$



mentre una stima **conservativa** della varianza campionaria di  $\hat{T}(I,h,K_j)$  risulta

$$(39) \quad V^2(I,h,K_j) = \sum_{k \in K_j} V^2(I,h,k) + 2 \sum_{k \in K_j} \sum_{k' > k \in K_j} C(I,h,k; I,h,k')$$

Qualora interessi invece la stima delle biomasse dei vari tipi di uso del suolo a livello nazionale, si indichi con  $T(I,h)$  ( $I = 1,2,8, h = 1,2,K, H_j$ ) la biomassa del tipo di uso del suolo  $I,h$  nell'intera nazione. La stima **corretta** di  $T(I,h)$  si ottiene tramite

$$(40) \quad \hat{T}(I,h) = \sum_{k=1}^{21} \hat{T}(I,h,k)$$

mentre una stima **conservativa** della varianza campionaria di  $\hat{T}(I,h)$  risulta

$$(41) \quad V^2(I,h) = \sum_{k=1}^{21} V^2(I,h,k) + 2 \sum_{k=1}^{21} \sum_{k' > k=1}^{21} C(I,h,k; I,h,k')$$

Si noti che nelle espressioni (39) e (41), che sono analoghe, i valori di  $V^2(I,h,k)$  e  $C(I,h,k; I,h,k')$  sono forniti rispettivamente dalle espressioni (31) e (32).

Inoltre, può interessare la stima delle biomasse a livello di area geografica dell'insieme di tipi di uso del suolo  $I,H$ . Si indichi allora con  $T(I,H,K_j)$  la biomassa dell'insieme  $I,H$  nell'area geografica  $j$ . La stima **corretta** di  $\hat{T}(I,H,K_j)$  si ottiene tramite

$$(42) \quad \hat{T}(I,H,K_j) = \sum_{k \in K_j} \hat{T}(I,H,k)$$

mentre una stima **conservativa** della varianza campionaria di  $\hat{T}(I,H,K_j)$  risulta

$$(43) \quad V^2(I,H,K_j) = \sum_{k \in K_j} V^2(I,H,k) + 2 \sum_{k \in K_j} \sum_{k' > k \in K_j} C(I,H,k; I,H,k')$$

Infine, qualora interessi la stima della biomassa a livello nazionale dell'insieme di tipi di uso del suolo  $I.H$ , si indichi con  $T(I.H)$  la biomassa dell'insieme  $I.H$  nell'intera nazione. La stima **corretta** di  $T(I.H)$  si ottiene tramite

$$(44) \quad \hat{T}(I.H) = \sum_{k=1}^{21} \hat{T}(I.H, k)$$

mentre una stima **conservativa** della varianza campionaria di  $\hat{T}(I.H)$  risulta

$$(45) \quad V^2(I.H) = \sum_{k=1}^{21} V^2(I.H, k) + 2 \sum_{k=1}^{21} \sum_{k' > k=1}^{21} C(I.H, k; I.H, k')$$

Si noti che nelle espressioni (42) e (44), che sono analoghe, i valori di  $\hat{T}(I.H, k)$  sono dati dalla espressione (35), mentre nelle espressioni (43) e (45), ancora analoghe, i valori di  $V^2(I.H, k)$  e  $C(I.H, k; I.H, k')$  sono dati rispettivamente dalle espressioni (36) e (37).

Ovviamente, in tutte le formule di questa sezione,  $I.H$  è sostituito da  $I$  quando  $H = \{1, 2, K, H_i\}$ .

Anche in questo paragrafo sono state considerate per brevità solo le aggregazioni spaziali o tipologiche che più comunemente interessano. Tuttavia, qualora interessino altri tipi di aggregazioni, il modo di procedere risulta, *mutatis mutandi*, del tutto simile a quello delineato in questa sezione.

#### 4.6 Stima di densità di biomasse

In genere la biomassa relativa a un certo tipo di uso del suolo è riportata non in termini assoluti ma è rapportata alla superficie dello stesso tipo. Sia allora  $A(I.h, k) = A\pi(I.h, k)$  la superficie del tipo  $I.h$  nel distretto territoriale  $k$  e sia  $T(I.h, k)$  la biomassa relativa. Si può allora essere interessati a stimare il rapporto

$$D(I.h, k) = \frac{T(I.h, k)}{A(I.h, k)} = \frac{1}{A} \frac{T(I.h, k)}{\pi(I.h, k)}$$

che rappresenta ovviamente la quantità per ettaro della biomassa relativa al tipo  $l.h$  nel distretto  $k$ .

Risulta quindi naturale stimare  $D(l.h,k)$  tramite il rapporto delle stime

$$(46) \quad \hat{D}(l.h,k) = \frac{1}{A} \frac{\hat{T}(l.h,k)}{\hat{\pi}(l.h,k)}$$

Ovviamente, in analogia a quanto detto nel paragrafo 4.3, essendo  $\hat{D}(l.h,k)$  un rapporto di stime, non sono note in modo esatto le sue proprietà. Tuttavia, approssimando anche in questo caso  $\hat{D}(l.h,k)$  in serie di Taylor sino al termine di primo grado, risulta immediato provare che  $\hat{D}(l.h,k)$  è una stima **approssimativamente corretta** di  $D(l.h,k)$ , mentre una stima **approssimativamente conservativa** della sua varianza risulta

$$(47) \quad V_D^2(l.h,k) = \frac{1}{A^2} \left\{ \frac{V^2(l.h,k)}{\hat{\pi}^2(l.h,k)} + \frac{\hat{T}^2(l.h,k)}{\hat{\pi}^4(l.h,k)} v^2(l.h,k) - 2 \frac{\hat{T}(l.h,k)}{\hat{\pi}^3(l.h,k)} C_{T\pi}(l.h,k) \right\}$$

dove

$$(48) \quad C_{T\pi}(l.h,k) = \frac{1}{N-1} \left\{ \frac{N(l,k)[N(l,k)-1]}{n(l,k)-1} f(l.h,k)[1-f(l.h,k)] \bar{t}(l.h,k) + \right. \\ \left. + \frac{N(0,k)[N(0,k)-1]}{n(0,k)-1} f_0(l.h,k)[1-f_0(l.h,k)] \bar{t}_0(l.h,k) + \right. \\ \left. + N(l,k)f^2(l.h,k) \bar{t}(l.h,k) + N(0,k)f_0^2(l.h,k) \bar{t}_0(l.h,k) - \hat{T}(l.h,k)\hat{\pi}(l.h,k) \right\}$$

costituisce una stima della covarianza tra  $\hat{T}(l.h,k)$  e  $\hat{\pi}(l.h,k)$ , mentre i valori di  $v^2(l.h,k)$  e  $V^2(l.h,k)$  sono dati rispettivamente dalle espressioni (3) e (31). Per gli sviluppi successivi è necessario determinare anche l'espressione della stima della covarianza tra  $\hat{T}(l.h,k)$  e  $\hat{\pi}(l.h,k')$ , che risulta

$$(49) \quad C_{T\pi}(l.h, k; l.h, k') = -\frac{\hat{T}(l.h, k)\hat{\pi}(l.h, k')}{N-1}$$

Tuttavia, la densità della biomassa relativa al tipo  $l.h$  può interessare non solo a livello dei singoli distretti ma anche a livello di area geografica o nazionale. In altri termini si può essere interessati a stimare

$$D(l.h, K_j) = \frac{1}{A} \frac{T(l.h, K_j)}{\pi(l.h, K_j)}$$

oppure si può essere interessati a stimare

$$D(l.h) = \frac{1}{A} \frac{T(l.h)}{\pi(l.h)}$$

In questi casi le formule delle stime e delle stime delle varianze relative sono analoghe, *mutatis mutandi*, a quelle ottenute a livello di singolo distretto. In particolare una stima **approssimativamente corretta** di  $D(l.h, K_j)$  risulta

$$(50) \quad \hat{D}(l.h, K_j) = \frac{1}{A} \frac{\hat{T}(l.h, K_j)}{\hat{\pi}(l.h, K_j)}$$

dove i valori di  $\hat{\pi}(l.h, K_j)$  e  $\hat{T}(l.h, K_j)$  si ottengono rispettivamente dalle espressioni (11) e (38). Inoltre una stima **approssimativamente conservativa** della sua varianza risulta

$$(51) \quad V_D^2(l.h, K_j) = \frac{1}{A^2} \left\{ \frac{V^2(l.h, K_j)}{\hat{\pi}^2(l.h, K_j)} + \frac{\hat{T}^2(l.h, K_j)}{\hat{\pi}^4(l.h, K_j)} v^2(l.h, K_j) - 2 \frac{\hat{T}(l.h, K_j)}{\hat{\pi}^3(l.h, K_j)} C_{T\pi}(l.h, K_j) \right\}$$

dove

$$(52) \quad C_{T\pi}(l.h, K_j) = \sum_{k \in K_j} C_{T\pi}(l.h, k) + \sum_{k \in K_j} \sum_{k' \neq k \in K_j} C_{T\pi}(l.h, k; l.h, k')$$

costituisce una stima della covarianza tra  $\hat{T}(l.h, K_j)$  e  $\hat{\pi}(l.h, K_j)$  mentre nelle espressioni (51) e (52) i valori di  $v^2(l.h, K_j)$ ,  $V^2(l.h, K_j)$ ,  $C_{T\pi}(l.h, k)$  e  $C_{T\pi}(l.h, k; l.h, k')$  si ottengono rispettivamente dalle espressioni (12), (39), (48) e (49).

In modo del tutto analogo, una stima **approssimativamente corretta** di  $D(l.h)$  risulta

$$(53) \quad \hat{D}(l.h) = \frac{1}{A} \frac{\hat{T}(l.h)}{\hat{\pi}(l.h)}$$

dove i valori di  $\hat{\pi}(l.h)$  e  $\hat{T}(l.h)$  si ottengono rispettivamente dalle espressioni (13) e (40). Inoltre, una stima **approssimativamente conservativa** della sua varianza risulta

$$(54) \quad V_D^2(l.h) = \frac{1}{A^2} \left\{ \frac{V^2(l.h)}{\hat{\pi}^2(l.h)} + \frac{\hat{T}^2(l.h)}{\hat{\pi}^4(l.h)} v^2(l.h) - 2 \frac{\hat{T}(l.h)}{\hat{\pi}^3(l.h)} C_{T\pi}(l.h) \right\}$$

dove

$$(55) \quad C_{T\pi}(l.h) = \sum_{k=1}^{21} C_{T\pi}(l.h, k) + \sum_{k=1}^{21} \sum_{k' \neq k=1}^{21} C_{T\pi}(l.h, k; l.h, k')$$

costituisce una stima della covarianza tra  $\hat{T}(l.h)$  e  $\hat{\pi}(l.h)$ , mentre nelle espressioni (54) e (55) i valori di  $v^2(l.h)$ ,  $V^2(l.h)$ ,  $C_{T\pi}(l.h, k)$  e  $C_{T\pi}(l.h, k; l.h, k')$  si ottengono rispettivamente dalle espressioni (14), (41), (48) e (49).

Si deve far notare che in questo paragrafo sono state considerate per brevità solo le quantità per ettaro relative ad aggregazioni spaziali. Qualora interessino le quantità per ettaro relative ad alcuni insiemi di tipi di uso del suolo, il modo di procedere risulta, *mutatis mutandi*, del tutto simile a quello delineato in questa sezione.

## 5 Problemi inerenti l'accorpamento di alcune classi e tipi di uso del suolo

In questo paragrafo sono trattati alcuni problemi specifici che, per motivi di concisione, non sono stati trattati in dettaglio nella parte generale.

I tipi di uso del suolo che afferiscono all'insieme delle *formazioni forestali rade*, ovvero *boschi di larice e cembro, boschi di abete rosso, .....*, *sugherete, altri boschi di latifoglie sempreverdi* - tutti ovviamente caratterizzati da bassissima copertura -, possono essere identificati in prima fase nell'insieme *Formazioni forestali*, ed in questo caso sono associati ai codici da 1.52 a 1.68, oppure nell'insieme *Formazioni forestali rade*, ed in questo caso sono associati ai codici da 2.1 a 2.17. Tuttavia dal momento che si tratta delle stesse categorie forestali (o tipi di uso del suolo) occorre considerare la stima della copertura e della biomassa complessiva.

Sia allora  $\hat{\pi}(1.h,k)$  la stima della copertura di uno dei 17 tipi identificati in prima fase nella classe 1 e sia  $\hat{\pi}(2.h',k)$  la stima della copertura del medesimo tipo identificato in prima fase nella classe 2, nel distretto  $k$ . In tal caso, per quanto detto nel paragrafo 4.2, la stima della copertura complessiva di tale tipo di uso del suolo nel distretto  $k$  risulta

$$\hat{\pi}(1.h + 2.h', k) = \hat{\pi}(1.h, k) + \hat{\pi}(2.h', k)$$

dove i valori di  $\hat{\pi}(1.h,k)$  e  $\hat{\pi}(2.h',k)$  si ottengono dall'espressione (1). Inoltre, una stima conservativa della varianza di  $\hat{\pi}(1.h + 2.h', k)$  risulta

$$v^2(1.h + 2.h', k) = v^2(1.h, k) + v^2(2.h', k) + 2c(1.h, k; 2.h', k)$$

dove i valori di  $v^2(1.h,k)$  e  $v^2(2.h',k)$  si ottengono dall'espressione (3) mentre il valore di  $c(1.h,k;2.h',k)$  si ottiene dall'espressione (6).

In modo del tutto analogo, se  $\hat{T}(1.h,k)$  è la stima della biomassa di uno dei 17 tipi identificati in prima fase nella classe 1 e  $\hat{T}(2.h',k)$  è la stima della biomassa del medesimo tipo identificato in prima fase nella classe 2, sempre nel distretto  $k$ , per quanto detto nel paragrafo 4.5, la stima della biomassa complessiva di tale insieme nel distretto  $k$  risulta

$$\hat{T}(1.h + 2.h', k) = \hat{T}(1.h, k) + \hat{T}(2.h', k)$$

dove i valori di  $\hat{T}(1.h, k)$  e  $\hat{T}(2.h', k)$  si ottengono dall'espressione (30). Inoltre, una stima conservativa della varianza di  $\hat{T}(1.h + 2.h', k)$  risulta

$$V^2(1.h + 2.h', k) = V^2(1.h, k) + V^2(2.h', k) + 2C(1.h, k; 2.h', k)$$

dove i valori di  $V^2(1.h, k)$  e  $V^2(2.h', k)$  si ottengono dall'espressione (31) mentre il valore di  $C(1.h, k; 2.h', k)$  si ottiene dall'espressione (33).

Un problema analogo si pone, sebbene solo a livello di stima della copertura, per le classi *praterie e pascoli incolti e zone aperte con vegetazione rada o assente*, le quali in prima fase possono essere identificate nella classe 1 e classificate rispettivamente come 1.72 e 1.73, nella classe 2 e classificate rispettivamente come 2.18 e 2.19, oppure possono essere identificate direttamente nelle classi 6 o 7. Anche in questo caso, dal momento che si tratta delle stesse categorie occorre considerare la stima della copertura complessiva.

Sia allora  $\hat{\pi}(1.h, k)$  con  $h=72$  o  $h=73$  la stima della copertura di una delle 2 categorie identificate in prima fase nella classe 1, sia inoltre  $\hat{\pi}(2.h', k)$  con  $h'=18$  o  $h'=19$  la stima della copertura di una delle 2 categorie identificate in prima fase nella classe 2 e sia  $\hat{\pi}(l, k)$  con  $l=6$  o  $l=7$  la stima della copertura delle medesime categorie identificate in prima fase nella classe 6 o 7, nel distretto  $k$ . In tal caso, sempre per quanto detto nel paragrafo 4.2, la stima della copertura complessiva di tali categorie nel distretto  $k$  risulta

$$\hat{\pi}(1.h + 2.h' + l, k) = \hat{\pi}(1.h, k) + \hat{\pi}(2.h', k) + \hat{\pi}(l, k)$$

dove i valori di  $\hat{\pi}(1.h, k)$  e  $\hat{\pi}(2.h', k)$  si ottengono dall'espressione (1) e il valore di  $\hat{\pi}(l, k)$  si ottiene dall'espressione (2). Inoltre, una stima conservativa della varianza di  $\hat{\pi}(1.h + 2.h' + l, k)$  risulta

$$v^2(1.h + 2.h' + l, k) = v^2(1.h, k) + v^2(2.h', k) + v^2(l, k) + 2c(1.h, k; 2.h', k) + 2c(1.h, k; l, k) + 2c(2.h', k; l, k)$$

dove i valori di  $v^2(1.h,k)$  e  $v^2(2.h',k)$  si ottengono dall'espressione (3), il valore di  $v^2(l,k)$  si ottiene dall'espressione (4) mentre i valori di  $c(1.h,k;2.h'k)$ ,  $c(1.h,k;l,k)$  e  $c(2.h',k;l,k)$  si ottengono dall'espressione (6), avendo cura di sostituire negli ultimi due casi  $l.h'$  con 6 o 7.

## **Bibliografia**

- DE VRIES, P.G., 1986 - *Sampling Theory for Forest Inventory*. Berlin, Springer-Verlag.
- FATTORINI L., PISANI C., 1999 - *Metodi di Campionamento per le Indagini Ambientali*, Dispense del Corso di Statistica per l'Ambiente, Facoltà di Economia "R. Goodwin", Università di Siena.
- FATTORINI L., MARCHESELLI M., PISANI C., 2003a - *Two-phase estimation of coverages with second phase corrections*. *Environmetrics* 14: 1-12.
- FATTORINI L., MARCHESELLI M. E PISANI C., 2003b - *A three-phase sampling strategy for multi-resource forest inventories*. Submitted to *Journal of Biological, Agricultural and Environmental Statistics*.



Tabella 1 - *Suddivisione in strati territoriali (distretti) del territorio italiano.*

<b>Regione o Provincia Autonoma</b>	<b>Sigla ( k )</b>
Piemonte	1
Valle d'Aosta	2
Lombardia	3
Trentino	4
Alto Adige	5
Veneto	6
Friuli Venezia Giulia	7
Liguria	8
Emilia Romagna	9
Toscana	10
Umbria	11
Marche	12
Lazio	13
Abruzzo	14
Molise	15
Campania	16
Puglia	17
Basilicata	18
Calabria	19
Sicilia	20
Sardegna	21

Tabella 2 - Stratificazione del territorio italiano per classi e per tipi di uso del suolo.

CLASSIFICAZIONE DI 1° FASE	SIGLA (l)	CLASSIFICAZIONE DI 2° FASE	SIGLA (l.h)	ALTRE CLASSIFICAZIONI
Formazioni forestali	1	boschi di larice e cembro	1.1	Boschi, boschi giovani
		boschi di abete rosso	1.2	
		boschi di abete bianco	1.3	
		pinete di pino silvestre e pino montano	1.4	
		pinete di pino nero, p. laricio e p. loricato	1.5	
		pinete di pini mediterranei	1.6	
		altri boschi di conifere, pure o miste	1.7	
		Faggete	1.8	
		boschi di rovere, roverella e farnia	1.9	
		boschi di cerro, farnetto, fragno, vallonea	1.10	
		Castagneti	1.11	
		ostrieti, carpiteti	1.12	
		boschi idrofilii	1.13	
		altri boschi caducifogli	1.14	
		Leccete	1.15	
		Sugherete	1.16	
		altri boschi di latifoglie sempreverdi	1.17	
		boschi di larice e cembro	1.18	Boschi bassi
		boschi di abete rosso	1.19	
		boschi di abete bianco	1.20	
		pinete di pino silvestre e pino montano	1.21	
		pinete di pino nero, p. laricio e p. loricato	1.22	
		pinete di pini mediterranei	1.23	
		altri boschi di conifere, pure o miste	1.24	
		faggete	1.25	
		boschi di rovere, roverella e farnia	1.26	
		boschi di cerro, farnetto, fragno, vallonea	1.27	
		castagneti	1.28	
		ostrieti, carpineti	1.29	
		boschi igrofilii	1.30	
		altri boschi caducifogli	1.31	
		leccete	1.32	
		sugherete	1.33	
		altri boschi di latifoglie sempreverdi	1.34	
		boschi di larice e cembro	1.35	Boscaglie
		boschi di abete rosso	1.36	
		boschi di abete bianco	1.37	
		pinete di pino silvestre e pino montano	1.38	
		pinete di pino nero, p. laricio e p. loricato	1.39	
		pinete di pini mediterranei	1.40	
		altri boschi di conifere, pure o miste	1.41	
		faggete	1.42	
		boschi di rovere, roverella e farnia	1.43	
		boschi di cerro, farnetto, fragno, vallonea	1.44	
		castagneti	1.45	
		ostrieti, carpineti	1.46	
		boschi igrofilii	1.47	
		altri boschi caducifogli	1.48	

		leccete	1.49	
		sugherete	1.50	
		altri boschi di latifoglie sempreverdi	1.51	
		boschi di larice e cembro	1.52	Boschi radi
		boschi di abete rosso	1.53	
		boschi di abete bianco	1.54	
		pinete di pino silvestre e pino montano	1.55	
		pinete di pino nero, p. laricio e p. loricato	1.56	
		pinete di pini mediterranei	1.57	
		altri boschi di conifere, pure o miste	1.58	
		faggete	1.59	
		boschi di rovere, roverella e farnia	1.60	
		boschi di cerro, farnetto, fragno, vallonea	1.61	
		castagneti	1.62	
		ostrieti, carpineti	1.63	
		boschi igrofili	1.64	
		altri boschi caducifogli	1.65	
		leccete	1.66	
		sugherete	1.67	
		altri boschi di latifoglie sempreverdi	1.68	
		arbusteti subalpini	1.69	Arbusteti
		arbusteti di clima temperato	1.70	
		macchia, arbusteti mediterranei	1.71	
		praterie, pascoli, incolti	1.72	Aree naturali senza soprassuolo for.
		zone aperte con vegetazione rada o assente	1.73	
		punti non raggiungibili o con variazioni di uso del suolo	1.74	Aree escluse dall'inventario
Formazioni forestali rade	2	boschi di larice e cembro	2.1	Boschi radi
		boschi di abete rosso	2.2	
		boschi di abete bianco	2.3	
		pinete di pino silvestre e pino montano	2.4	
		pinete di pino nero, p. laricio e p. loricato	2.5	
		pinete di pini mediterranei	2.6	
		altri boschi di conifere, pure o miste	2.7	
		faggete	2.8	
		boschi di rovere, roverella e farnia	2.9	
		boschi di cerro, farnetto, fragno, vallonea	2.10	
		castagneti	2.11	
		ostrieti, carpineti	2.12	
		boschi igrofili	2.13	
		altri boschi caducifogli	2.14	
		leccete	2.15	
		sugherete	2.16	
		altri boschi di latifoglie sempreverdi	2.17	
		praterie, pascoli, incolti	2.18	Aree naturali senza soprassuolo for.
		zone aperte con vegetazione rada o assente	2.19	
		punti non raggiungibili o con variazioni di uso del suolo	2.20	Aree escluse dall'inventario
Aree temporaneamente prive di soprassuolo	3	<i>idem</i>	3	
Inclusi in superfici forestali	4	<i>idem</i>	4	

Boschetti e filari in praterie, pascoli e incolti	5	<i>idem</i>	5	Aree naturali senza soprassuolo forestale
Altre praterie, pascoli e incolti	6	<i>idem</i>	6	
Zone aperte con vegetazione rada o assente	7	<i>idem</i>	7	
Arboricoltura da legno	8	pioppeti artificiali	8.1	Superfici agricole
		piantagioni a prevalenza latifoglie	8.2	
		piantagioni a prevalenza conifere	8.3	
		punti non raggiungibili o con variazioni di uso del suolo	8.4	Aree escluse dall'inventario
Altre superfici agricole	9	<i>idem</i>	9	Superfici agricole
Boschetti e filari in sup.agric.	10	<i>idem</i>	10	
Boschi in aree urbane	11	<i>idem</i>	11	Superfici artificiali
Altre superfici artificiali	12	<i>idem</i>	12	
Boschetti e filari in sup.artif.	13	<i>idem</i>	13	
Boschetti e filari in zone umide	14	<i>idem</i>	14	Zone umide
Altre zone umide	15	<i>idem</i>	15	
Boschetti e filari in acque	16	<i>idem</i>	16	Acque
Altre acque	17	<i>idem</i>	17	