Workshop
"Lidar application in forest inventory and related statistical issues"

Estimation of standing wood volume in forest compartments by exploiting lidar information

Chiara Torresan
Piermaria Corona, Lorenzo Fattorini, Sara Franceschi, Gianfranco Scrinzi

Viterbo - May 8, 2013

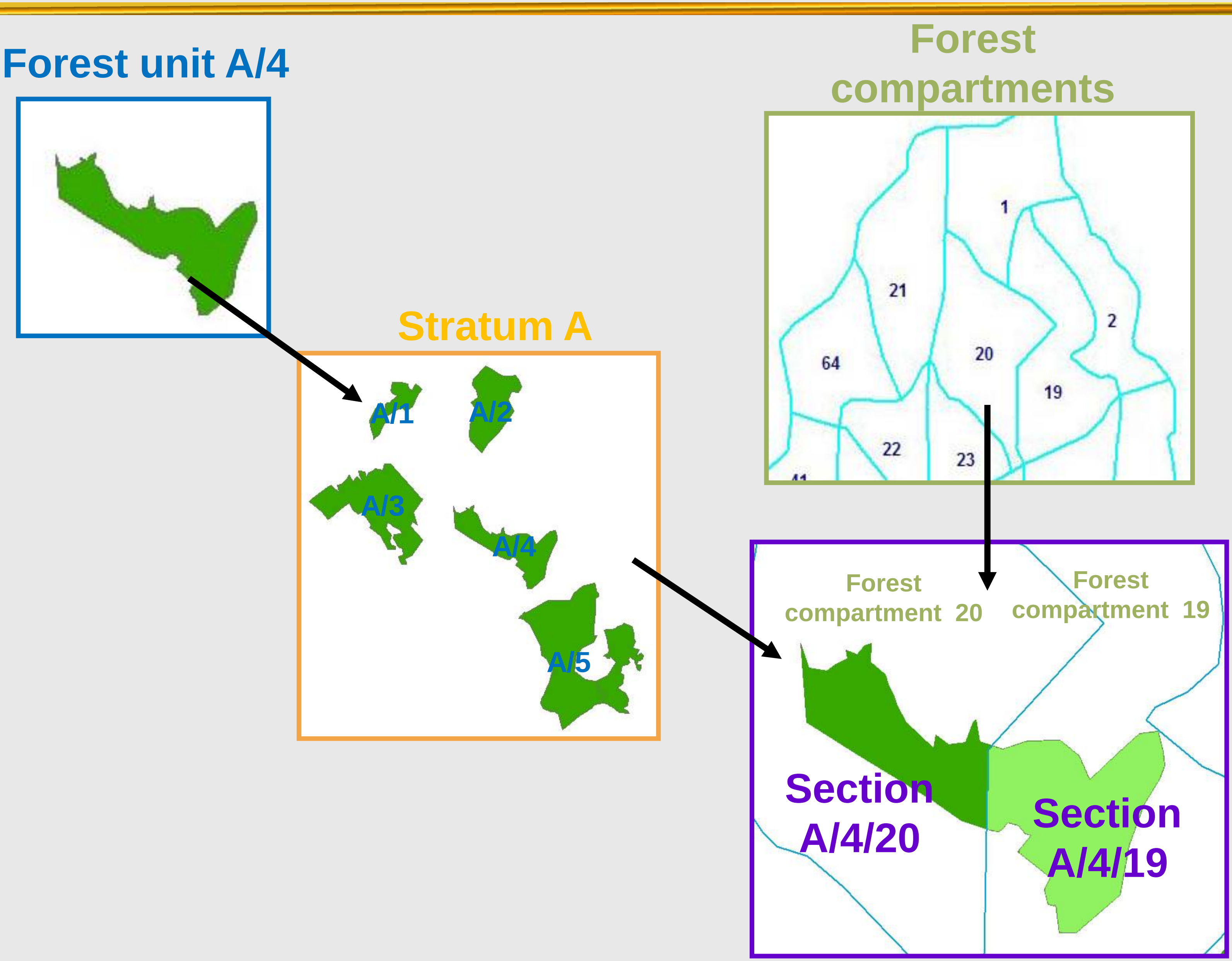
The inventory for forest planning purposes in Trentino

- since 2010, the Forest Service of Trento Province has changed its inventory method for forest management purposes from full callipering on compartment basis to stratified sampling;

- by the stratification, the forest is subjectively subdivides into homogeneous sub-populations, or strata;

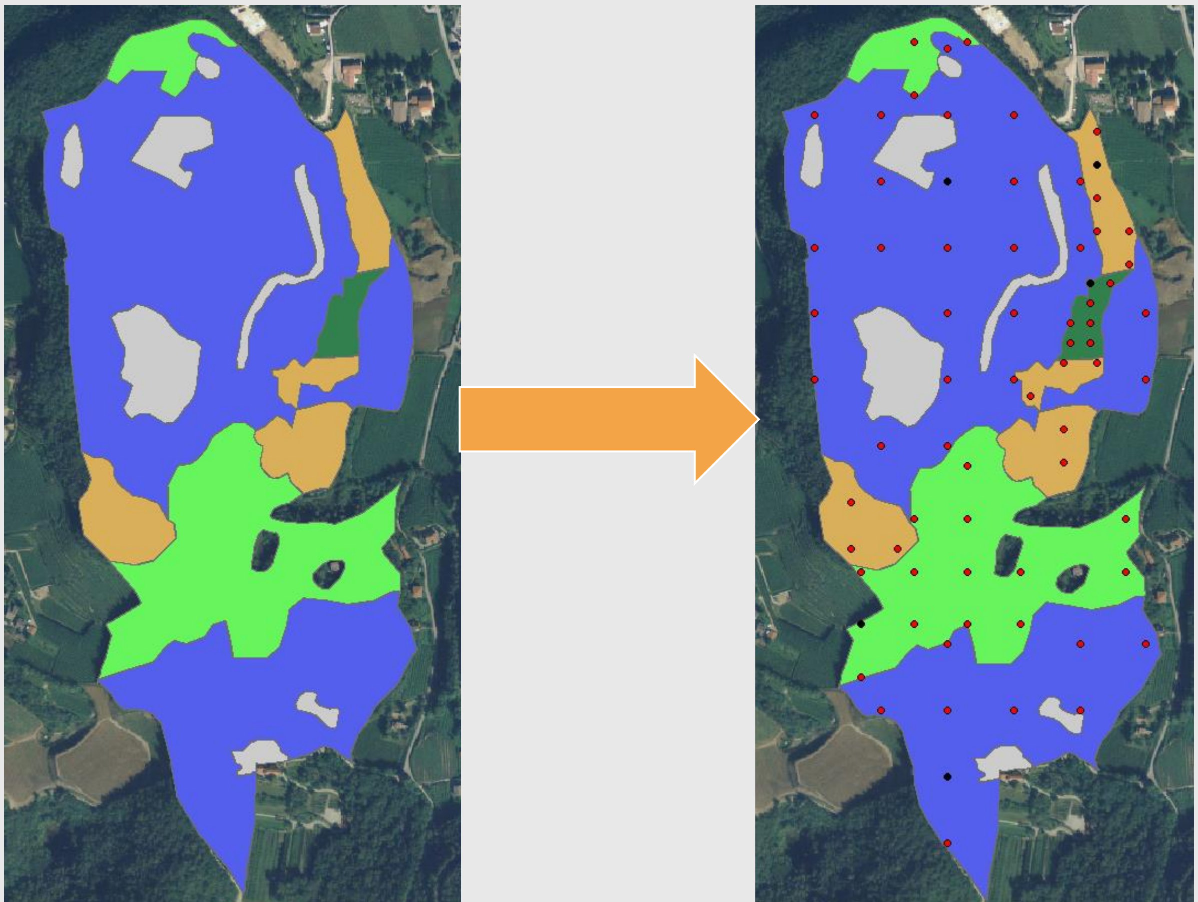
- the stratification process is based on remote sensing support analysis (orthophotos, lidar data), and ground control surveys;

The inventory for forest planning purposes in Trentino



The inventory for forest planning purposes in Trentino

- the sampling process defines the number and the location of the sampling points in each stratum according to its importance;



The results from the inventory

- field inventories allow data collection to update figures at forest and strata level for the estimation of key attribute (standing wood volume, in particular timber volume);

but

- the growing stock estimation is required by the Forest Service at forest compartment level;

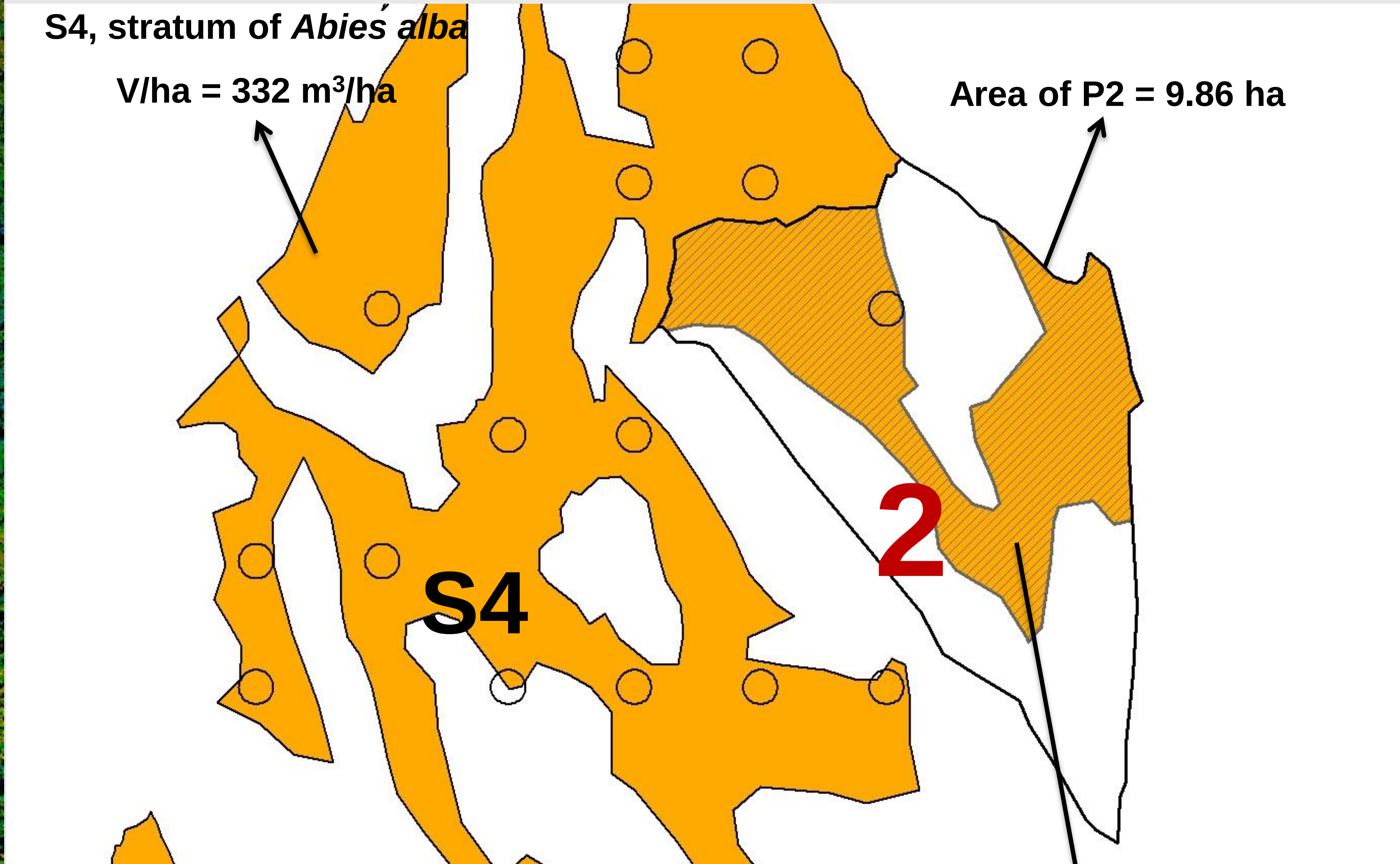
then

- the volume of a forest compartment is computed as weighted sum on the section surfaces of the average volume of the stratum;

but

- often it happens that a small number or no sampling points fall within a section, hence the estimation of the timber volume will not be accurate;

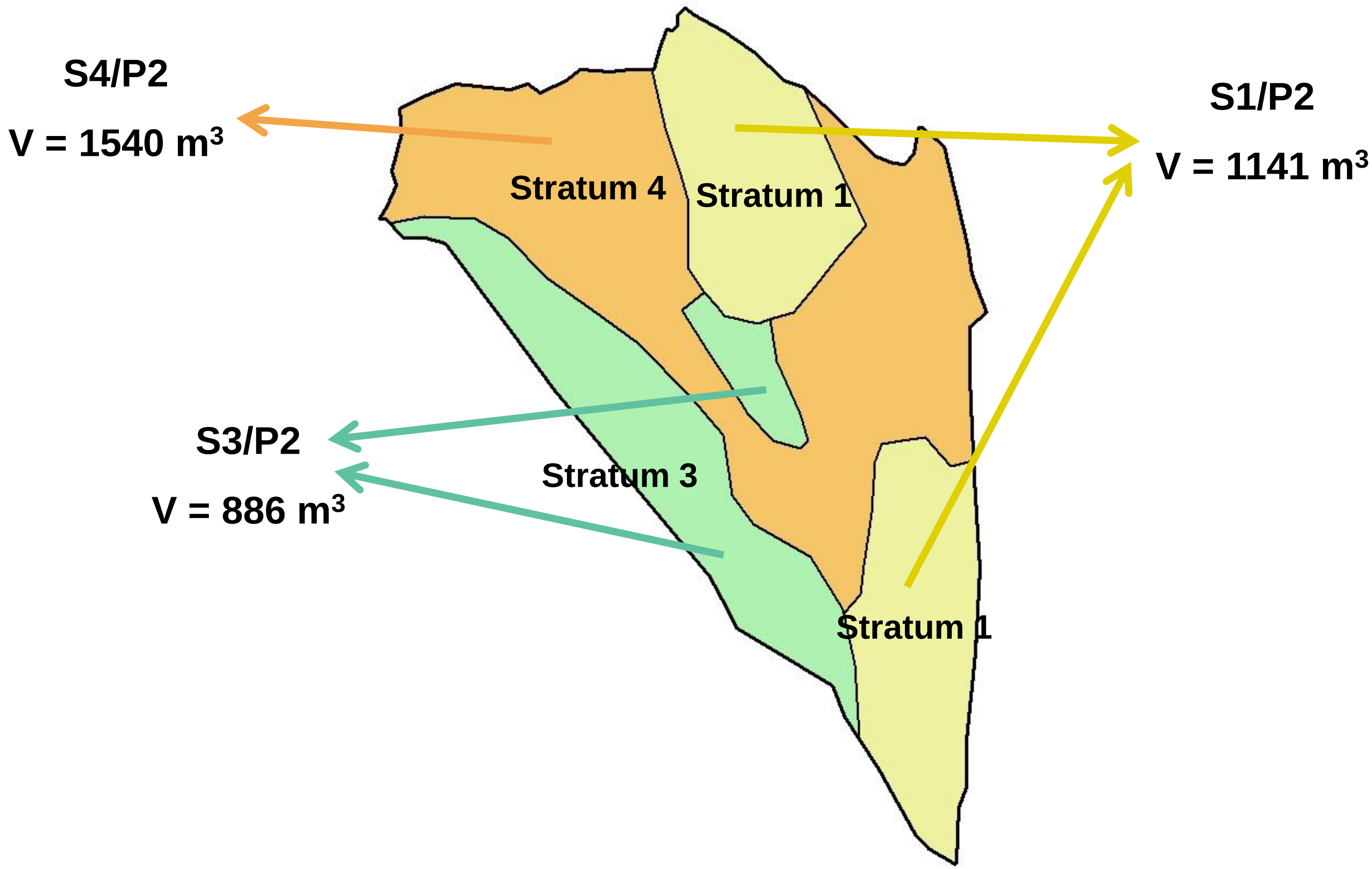
The timber volume estimation at forest compartment level



Area of stratum S4 in P2 = 4.64 ha

V = 332 m^3/ha * 4.64 ha = 1540 m^3

The timber volume estimation at forest compartment level



V (forest compartment) = 3567 m^3

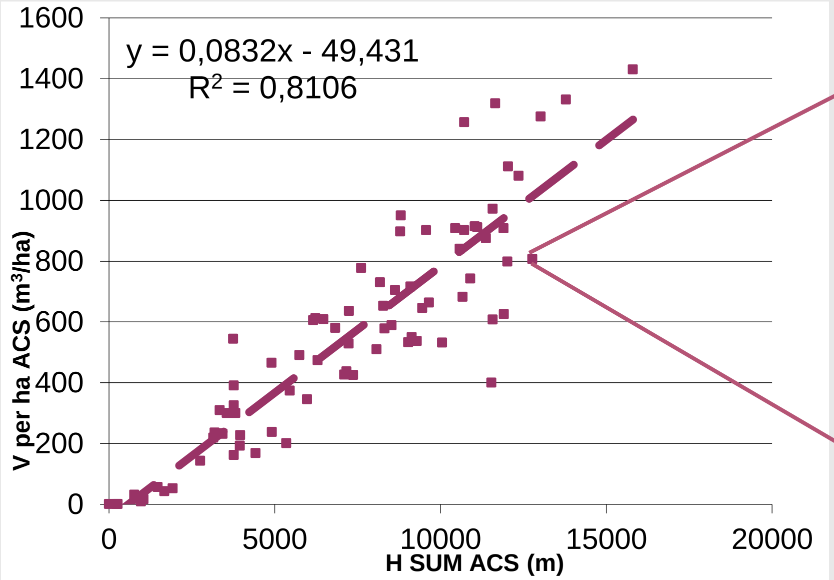
Our questions, and hyphotesis

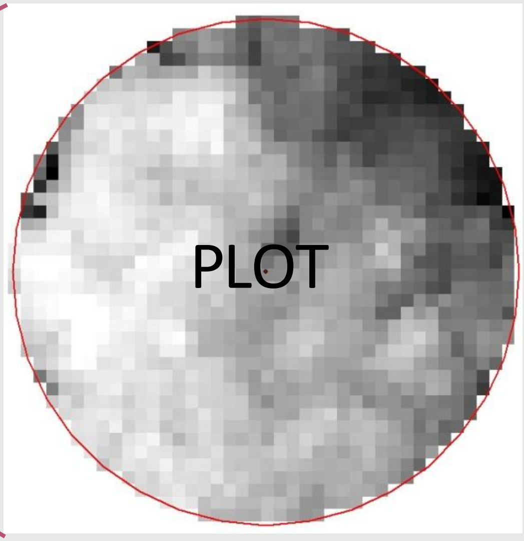
- we wondered:

\- can lidar data be used as auxiliary information and assist in improving the estimate of the wood volume at forest compartment level?

\- which could be the statistical methods more suitable to attain rigorous inference of forest resources (volume) at forest compartment level?

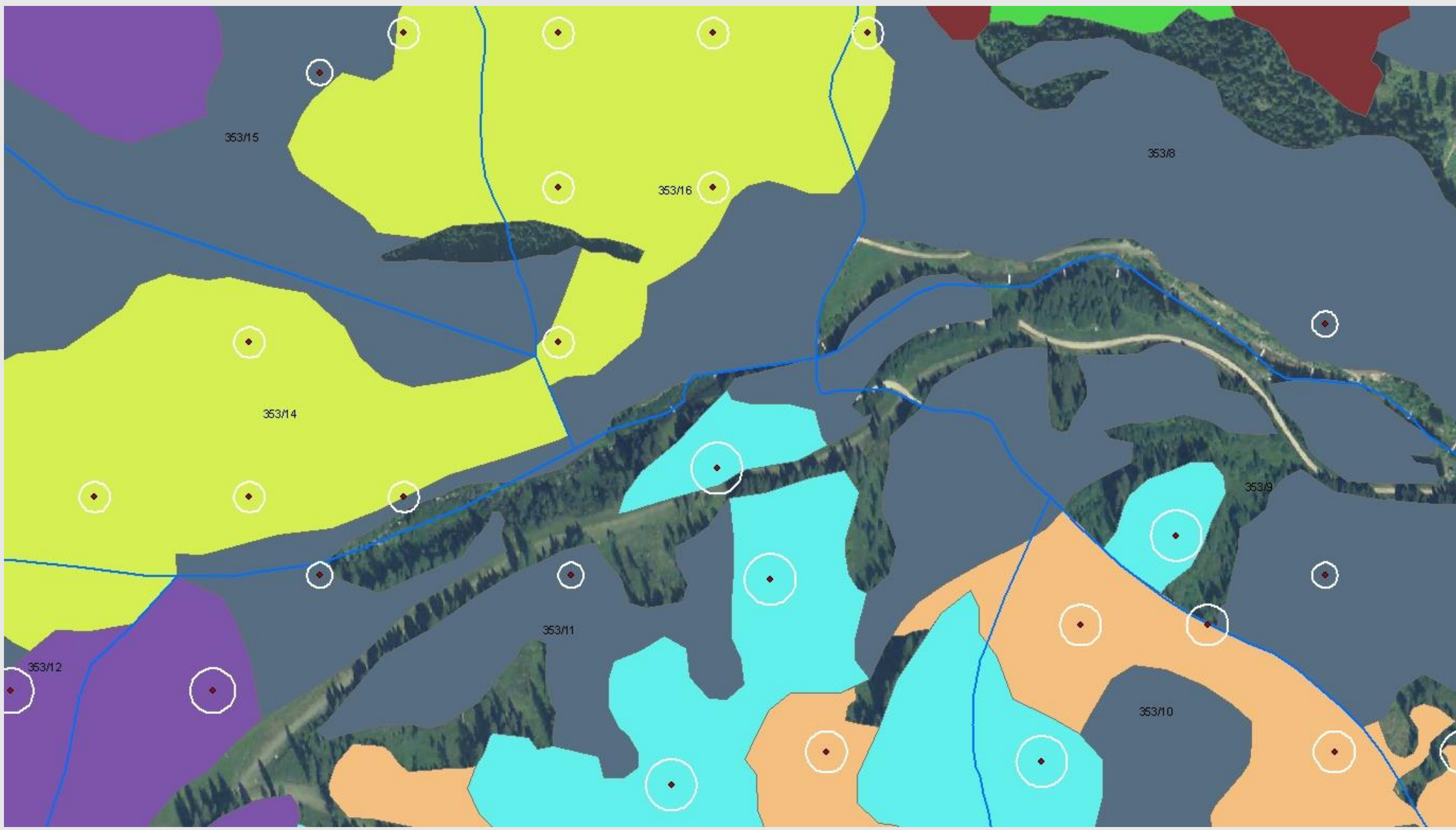
- we hypothesized:





The relevance of the methodology

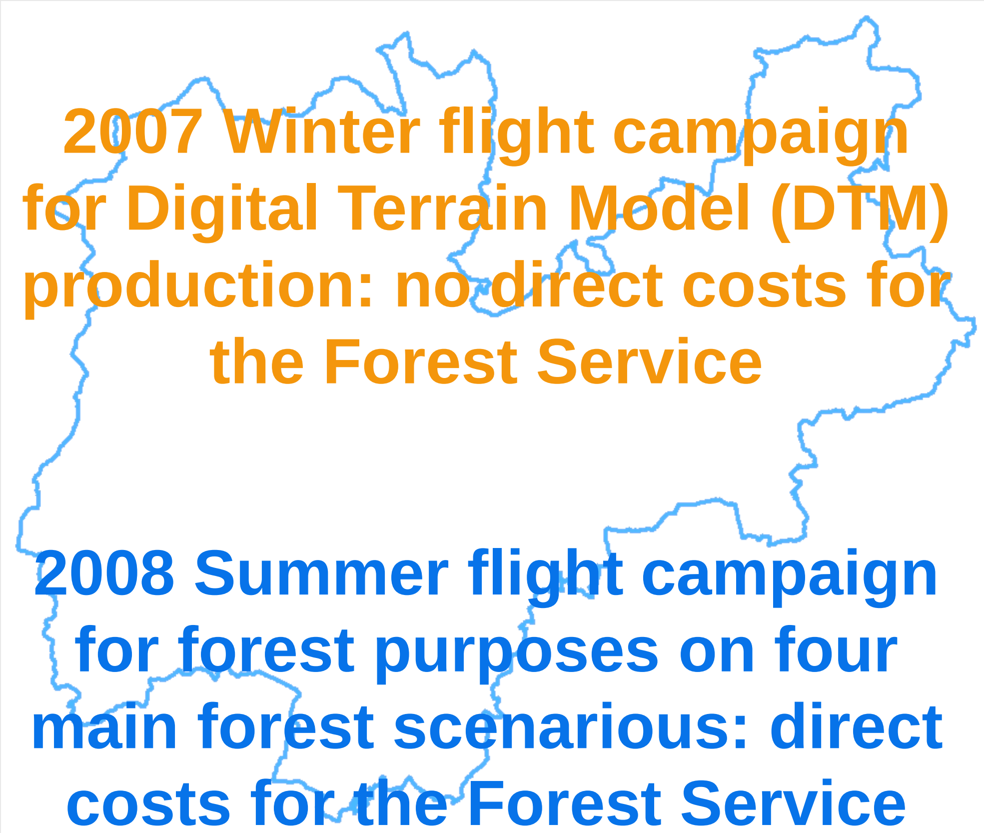
• it has the purpose to improve the volume estimation accuracy at forest compartment level conserving the inventory system already tested and in use in the Forest Service of Trento province exploiting the available lidar data as auxiliary information;



The lidar data availability

- forestry technicians can often exploit the CHM produced by the data from airborne laser scanning surveys;

province of Trento



national territory

The model for timber volume estimation

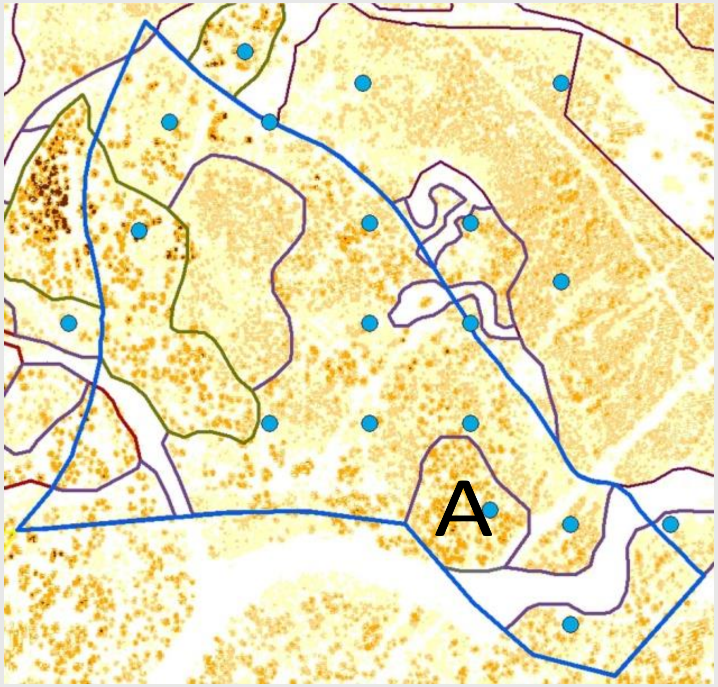
- for the spatial unit A (a section), it counts the relationship:

$$Y_A = \beta X_A + \varepsilon_A$$

$$E_m(\varepsilon_A) = 0 \qquad V_m(\varepsilon_A) = \sigma^2 X_A$$

Y_A amount of timber volume within A

X_A known population of CHM pixel in A



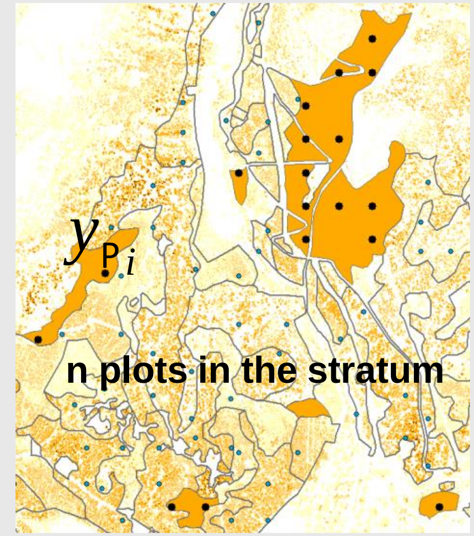
Y_A and X_A are normal random variables with:

$$E_m(Y_A | X_A) = \beta X_A$$

$$V_m(Y_A | Y_A) = \sigma^2 X_A$$

The estimation strategy

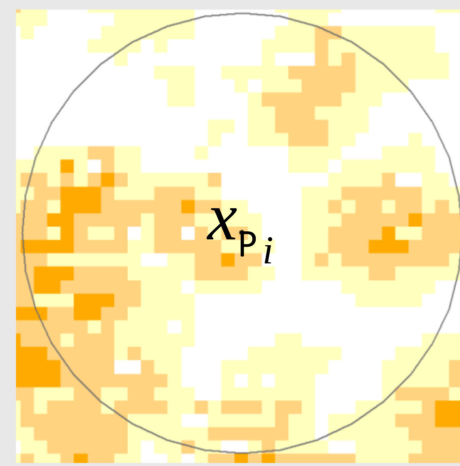
- a model-unbiased estimator for Y_A is achieved by a model-unbiased estimator for β :



$$\hat{b}_n = \frac{\bar{y}_n}{\bar{x}_n} = \frac{\bar{y}_P}{\bar{x}_P}$$ maximum likelihood estimator of β

$\bar{y}_n$ arithmetic means of wood volume within n plots in the stratum ($\bar{y}_P$)

$\bar{x}_n$ arithmetic means of heights from the CHM within n plots ($\bar{x}_P$)



- a model-dependent predictor of the random variable Y_A is given by:

$$\hat{Y}_A = \hat{b}_n X_A = \frac{\bar{y}_P}{\bar{x}_P} X_A$$

Three inference approaches

- model-based inference;
- design-based inference (delta and jackknife methods);
- hybrid inference (delta and jackknife methods for the design-based component);

Model-based inference

- the model:

$$Y_A = \beta X_A + \varepsilon_A$$

is used to justify the estimator $\hat{Y}_A$;

- all the variability of the estimator arises from the model;

- the variance of the estimator is:

$$\hat{V}_m = \hat{\sigma}_c^{\,2}\left(\frac{X_A^{2}}{n\bar{x}_n} - 2\frac{X_A}{n\bar{x}_n}\sum_{i=1}^{n}\sum_{j \in P_i \cap A} x_j + X_A\right)$$

with the bias-correct maximum-likelihood estimator of σ^2:

$$\hat{\sigma}_c^{\,2} = \frac{1}{n-1}\sum_{i=1}^{n}\frac{\left(y_{P_i} - \hat{b}_n x_{P_i}\right)^2}{x_{P_i}}$$

Design-based inference

- the model:

$$Y_A = \beta X_A + \varepsilon_A$$

is used to justify the estimator;

- the variability of the estimator depends on the design;
- the variance of the estimator is:

A. $\hat{V}_d = \frac{X_A^{\ 2}}{n}\left(\frac{s_Y^{\ 2}}{\bar{x}_P^{\ 2}} - 2\bar{y}_P^{\ 2}\frac{s_{XY}}{\bar{x}_P^{\ 3}} + \bar{y}_P^{\ 2}\frac{s_X^{\ 2}}{\bar{x}_P^{\ 4}}\right)$ (delta method)

$s_X^{\ 2}$ and $s_Y^{\ 2}$ variances of the x_{P_i} and y_{P_i}

s_{XY} X,Y covariances

B. $\hat{V}_{d(jack)} = \frac{1}{n(n-1)}\sum_{i=1}^{n}(\hat{b}_{n-1}^{\ (i)} - \hat{b}_{n(jack)})^2$ (jackknife method)

$\hat{b}_{n-1}^{\ (i)}$ i-th pseudo value

$\hat{b}_{n(jack)}^{\ (i)}$ arithmetic mean of the pseudo-values

The forest inventory
The proposal
The strategy
The application of the methodology
Conclusion

Hybrid inference

- the model:

$$Y_A = \beta X_A + \varepsilon_A$$

is used to justify the estimator;

- the variability of the estimator depends both on the model and the design;

$$\hat{V}_{md} = \hat{V}_d + \hat{\sigma}_c^{\,2}(1 - \frac{X_A}{X})X_A$$

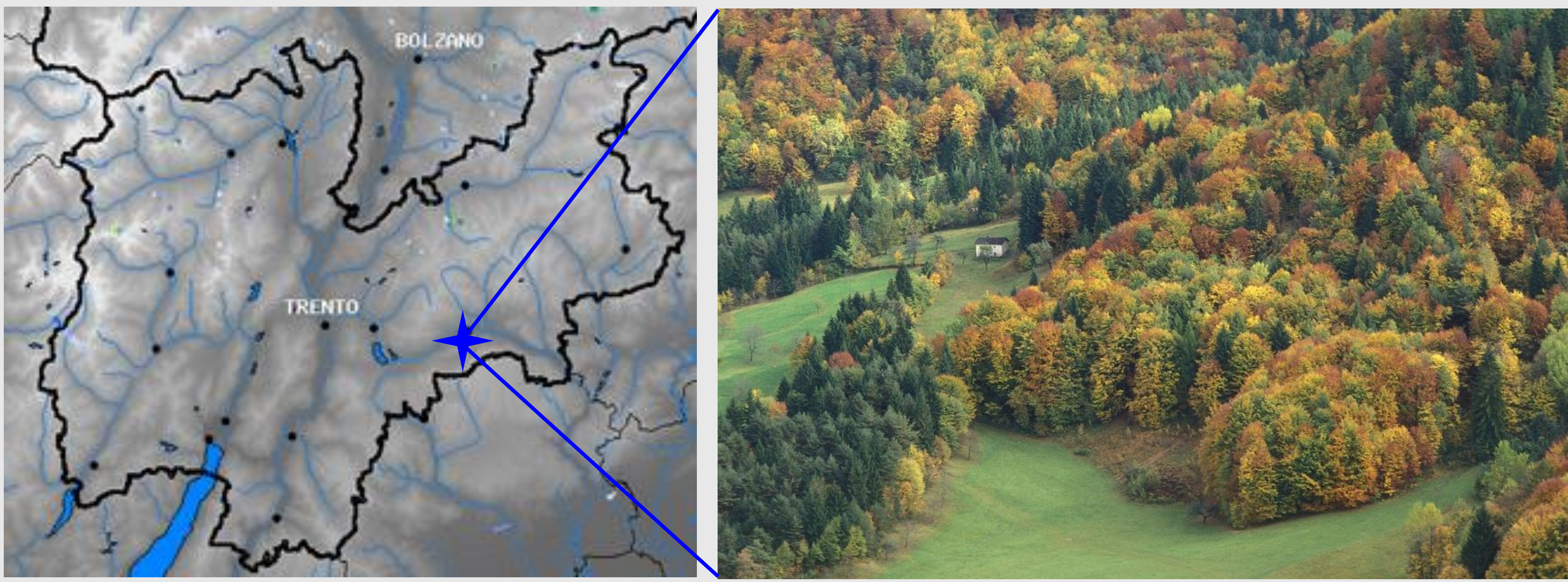
with:

X_A sum of the heights from the CHM in A

X sum of the heights from the CHM in the stratum

The simulation: the forest

- forest located in Val di Sella (Eastern of Trento Province);
- forest species composition: silver fir (*Abies alba*) and beech (*Fagus sylvatica*), together with Norway spruce (*Picea abies*), larch (*Larix decidua*) and other minor species of conifers and hardwoods;



The simulation: the lidar data

- 2008 Summer campaign, Optech ALTM 3100C sensor, 70,000 Hz laser pulse frequency, 4 returns per pulse, 5 points/m^2, 25-30 cm footprint diameter for 1000 m height of flight;

- raw data filtered and classified using Terrascan software (Axelsson algorithm);

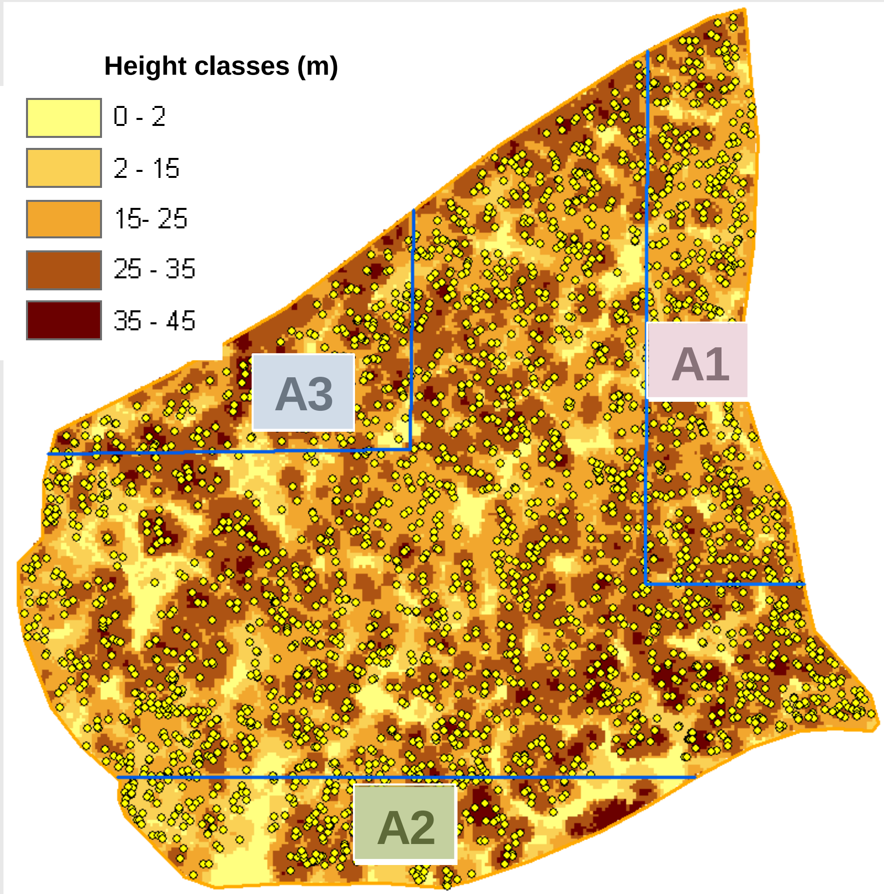
- Digital Surface Model (DSM) and the Digital Terrain Model (DTM) obtained, respectively, from first and last returns;

- CHM computed as difference between DSM and DTM at a spatial resolution of 1 m^2;

- canopy height values greater than 2 m were assumed to be vegetation hits;

The simulation: the forest field data

Val di Sella, stratum of 4.8 ha



A1, A2 e A3 portions of forest compartment 1, 2, and 3 generated by the stratum A

- stratum of 4.8 ha,
- on 2008 full callipering, recording of the geographic position of each tree;
- artificial delineation of three sections in the edges of the stratum (10% of the area of the whole stratum, 4800 m^2);

The simulation: the data processing

- for each section: simulation by 10,000 runs;

- for each run: 25, 50, 100 plots with radius of 10 m randomly and independently selected within the stratum;

- for each plot: computation of total heights ($x_{\mathrm{P}i}$) and volume ($y_{\mathrm{P}i}$);

- for each section estimation of:
- total volume ($\hat{Y}_{\mathrm{A}}$), relative bias (RB), and relative root mean squared error (RRMSE) from the Monte Carlo distributions of $\hat{Y}_{\mathrm{A}}$;
- variance of $\hat{Y}_{\mathrm{A}}$ by model-based, design-based and hybrid approach;
- percentage of cases that the 0.95 confidence intervals (C95) included the true volume values;

The results of the simulation

				Model-based		Design-based (delta-method)		Design-based (jackknife)		Hybrid (delta-method)		Hybrid (jackknife)	
Compartment portion	Number of plots in the stratum	RB	RRMSE	ERSE	C95	ERSE	C95	ERSE	C95	ERSE	C95	ERSE	C95
A1 (b1 = 0.00249)	25	-0.2	7.4	10.9	99.2	7.5	94.1	7.3	93.6	11.4	99.4	11.3	99.4
	50	0.1	5.2	9.7	99.9	5.2	94.9	5.2	94.6	10.1	100	10.1	100
	100	0.0	3.7	9.0	100	3.7	94.9	3.7	94.8	9.4	100	9.4	100
A2 (b2 = 0.00237)	25	4.7	8.6	11.5	98.7	7.5	89.9	7.3	89.3	12.0	99.2	11.9	99.1
	50	5.0	7.0	10.3	99.8	5.2	85.3	5.2	84.9	10.7	99.9	10.7	99.9
	100	4.9	5.9	9.7	100	3.7	75.6	3.7	75.3	10.1	100	10.0	100
A3 (b3 = 0.00213)	25	16.9	16.2	9.5	69.5	7.5	49.6	7.3	48.3	11.0	81.5	11.0	81.0
	50	17.2	15.6	8.0	56.4	5.2	19.0	5.2	18.6	9.7	77.8	9.7	77.5
	100	17.1	15.1	7.2	42.4	3.7	2.0	3.7	2.0	8.9	77.6	8.9	77.5

(b = 0.00246)

- bias negligible when b_A and b are very similar (A1);
- low values of bias determine RRMSE lower than 7.5% (also with few plots per stratum);
- the estimator $\hat{Y}_A$ combined with the use of the hybrid estimator of variance is a procedure for conservative evaluation of accuracy;

The case study: the forest and the field data

- forest owned by the San Martino di Castrozza municipality;
- 604.5 ha, typical sub-alpine *Picea abies* forest, located between 1500 and 1900 m of altitude;
- forest area partitioned into 8 homogeneous strata (from 20.5 to 160.4 ha; from 281 to 558 m^3/ha);
- sampling points (314 in total) selected according to the criteria used in the inventory system for forest planning in Trentino;
- field surveys carried out during the 2008 Summer;
- wood volume computed for each plot;

The study case: the lidar data

The forest inventory

The proposal

The strategy

The application of the methodology

Conclusion

- 2007 Winter flight campaign, Optech ALTM 3100C sensor, 33,000 Hz laser pulse frequency, 2 returns per pulse, 0.5 points/m^2, 25-30 footprint diameter cm for 1500 m height of flight;

- raw data filtered and classified using Terrascan software (Axelsson algorithm);

- Digital Surface Model (DSM) and the Digital Terrain Model (DTM) obtained, respectively, from first and last returns;

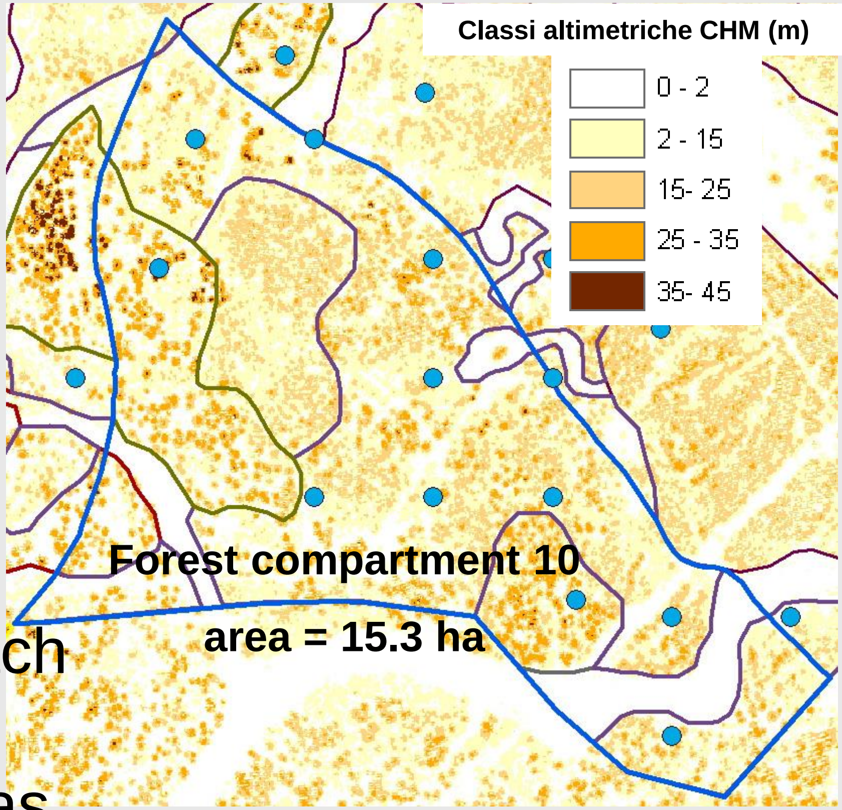
- Canopy Height Model (CHM) computed as DSM-DTM at a spatial resolution of 1 m^2;

- canopy height values greater than 2 m were assumed to be vegetation hits;

The case study: the data processing

- 4 forest compartments selected (15.0, 15.3, 22.7, and 24.5 ha);
- in each section, the volume was estimated by $\hat{Y}_{A} = \frac{\bar{y}_{P}}{\bar{x}_{P}} X_{A}$;
- the volume of the entire compartment was computed as the sum of the values estimated for the single section;
- the hybrid inference approach was used, the design-based component of the variance was estimated by jackknife;

San Martino di Castrozza
forest compartments 5, 10, 24, 27



The case study: the results (1)

The forest inventory
The proposal
The strategy
The application of the methodology
Conclusion

Forest compartment 5 (22.7 ha)				
Section	Number of plots	Volume (m^3)	RSE	0.95 confidence interval (m^3)
1/5	0	873	11.48%	673-1072
2/5	1	1870	6.37%	1632-2108
4/5	7	8947	6.38%	7805-10089
7/5	1	366	20.80%	213-519
Total		12056	4.95%	10863-13249

Forest compartment 10 (15.3 ha)				
Section	Number of plots	Volume (m^3)	RSE	0.95 confidence interval (m^3)
1/10	0	1004	11.30%	777-1231
2/10	1	570	7.84%	481-659
3/10	0	165	20.51%	97-233
5/10	9	4247	10.53%	3352-5142
6/10	1	721	15.14%	503-939
Total		6706	7.12%	5751-7661

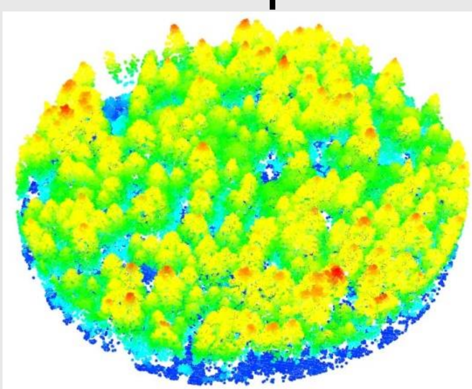
The case study: the results (2)

Forest compartment 24 (15.0 ha)				
Section	Number of plots	Volume (m^3)	RSE	0.95 confidence interval (m^3)
1/24	3	4651	10.31%	3692-5610
2/24	4	3401	6.04%	2990-3812
5/24	1	373	19.02%	231-515
6/24	1	292	17.14%	192-392
8/24	1	186	18.66%	117-255
Total		8904	5.95%	7844-9964

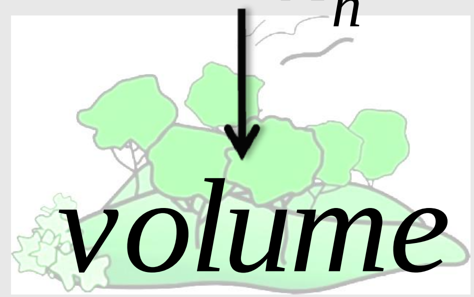
Forest compartment 27 (24.5 ha)				
Section	Number of plots	Volume (m^3)	RSE	0.95 confidence interval (m^3)
1/27	1	1191	11.11%	926-1456
3/27	4	2800	9.09%	2291-3309
4/27	3	4775	6.80%	4126-5424
6/27	1	545	15.60%	375-715
8/27	10	199	18.23%	126-272
Total		9511	4.66%	8625-10397

Conclusion

- the uncertainty of the timber volume estimate is comparable to that achievable with full callipering: standard error of 4% in the previous forest inventories of Trento province;
- when the CHM is available, the proposed strategy may represent a strong and highly effective support to forest managers improving the accuracy of volume estimation at forest compartment level;



$$\hat{\beta} = \frac{\bar{y}_n}{\bar{x}_n}$$



Thanks
for your attention

chiara.torresan@entecra.it