Conference on Applied Statistics in Agriculture          1994 - 6th Annual Conference Proceedings

# STRATIFICATION AND CLUSTER ESTIMATOR ON AN AREA FRAME BY SQUARED SEGMENTS WITH AN ALIGNED SAMPLE

M. Fuentes

F J. Gallego

## Recommended Citation

# STRATIFICATION AND CLUSTER ESTIMATOR ON AN AREA FRAME BY SQUARED SEGMENTS WITH AN ALIGNED SAMPLE

M. Fuentes: Deparment of Statistics, University of Valladolid, 47071 Valladolid, Spain.
F.J. Gallego: JRC. 21020 Ispra (Varese), Italy

## ABSTRACT

Several European countries (Portugal, Spain, Greece, Rumania, and the Czech Republic) make crop surveys on area frame with an aligned systematic sampling of squared segments. So far crop area estimates are obtained with standard formulae for random sampling, without using the spatial structure of the sample. This is in general conservative, the estimated standard error is larger than the error actually made. Taking as clusters the set of segments with the same relative position in a block, gives often lower but very unstable variances. A more stable variance estimate is computed by repeated random permutations of the sample segments in each block before building up the clusters. There is most often a moderate variance reduction when no stratification has been performed. If the sample is stratified and the cluster estimator is applied in each stratum, the variances seem to be less unstable. In this case permutations in each block seem not to be very useful.

Keywords: *area frame sampling, cluster estimator, crop area estimation, stratification.*

## 1. INTRODUCTION

### 1.1 The Mars Project of the European Community
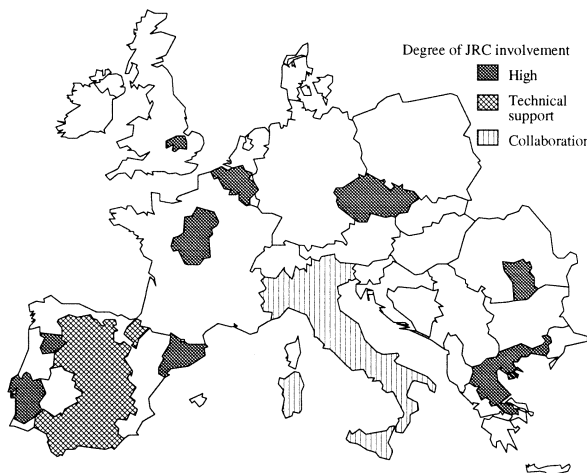


Fig. 1: European regions with segment surveys in 1992

The European Community (EC) launched in 1988 a project to assess and develop operational applications of Remote Sensing to Agricultural Statistics (Meyer-Roux, 90). The project is carried out by the Institute of Remote Sensing Applications (IRSA) of the Joint Research Centre (JRC) of the EC.

We shall focus here on a sampling of squared segments used in the frame of the so called "Regional Inventories" (fig. 1), that deal with crop area and yield estimation at the regional and national level based on sampling of aerial units, often named segments (Gallego, 93, 94). In most cases we use squared segments rather than segments that follow physical elements on the landscape (González, 91, Allen, 90, Cotter, 87).

## 1.2  Area sampling frame in Spain



□ Cluster without permutation
* Cluster after permutation
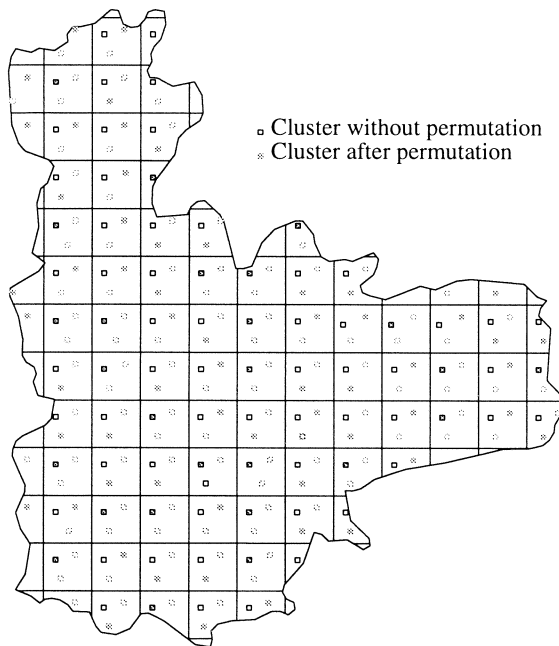
**Fig. 2: Sample of segments by repetition of a random pattern in Valladolid (Spain)**

The Spanish Ministry of Agriculture is conducting a yearly survey on squared segments of 49 ha. In 1993, the size of the sample was nearly 10000 segments in an area of more than 350.000 Km$^2$. That sample was without stratification and aligned by blocks of 10Km.*10Km. Figure 2 shows an example of that sampling for a province. So far the estimates are obtained with standard formulae (González, 91, Gallego, 93) as if the sampling were purely random. This classic estimator $\bar{y}_{ran}$ does not use the particular spatial structure of the sample and is in general rather conservative, the estimated standard error is presumably larger than the error actually made.

## 2.  CLUSTER ESTIMATOR

### 2.1  Cluster Estimator by Replicates

Each squared block has $M$ segments, out of which we draw a sample of $r$ segments ($M=100$, $r=3$ in the Spanish survey) (fig. 2). The sample has units $y_i$ that can be written as *(b,j)* from replica *j (j=1,r)* in block *b (b=1,B)*. All the segments corresponding to replica *j* have the same relative position in the block. An estimator by clusters can be applied in this sampling (Ambrosio, 1993), where each cluster $C_j$ is made up of all the segments $\{(b,j),\ b=1,...,B\}$ that have the same relative position in each block. This estimator has been tested in Spain. The variances of the estimates become lower in general, but are extremely unstable because of the very small effective sample size. The relative efficiency (ratio of variances) varies between 0.4 and 50 (Tables 1, 2).

For each cluster: $\bar{y}_{\bullet i} = \dfrac{\sum\limits_{j \in C_i} y_{ji}}{B}$ . The cluster estimator $\bar{y}_{clus} = \dfrac{\sum\limits_{i=1}^{r} \bar{y}_{\bullet i}}{r}$ is unbiased :

$E(\bar{y}_{clus}) = \bar{Y}$

The expressions for the variance of the cluster estimator and the relative efficiency are:

$$Var(\bar{y}_{clus}) = \left(1 - \frac{r}{M}\right) \frac{\sum\limits_{i=1}^{M}(\bar{y}_{\bullet i} - \bar{Y})^2}{r(M-1)} \qquad Rel.eff.(\bar{y}_{clus}/\bar{y}_{ran}) = \frac{Var(\bar{y}_{ran})}{Var(\bar{y}_{clus})} \qquad (1)$$

| Sevilla | Classic estimator | | Cluster estimator | | | Cluster estimator with permutations | | |
|---|---|---|---|---|---|---|---|---|
| CROP | area % | std. error | area % | std. error | rel. eff. clus/ran | area % | std. error | rel. eff. per/ran |
| Wheat | 17.37 | 1.28 | 17.45 | 2.86 | 0.20 | 18.59 | 1.15 | 1.24 |
| Barley | 0.44 | 0.12 | 0.45 | 0.14 | 0.67 | 0.50 | 0.14 | 0.74 |
| Sugar beet | 2.38 | 0.45 | 2.37 | 0.47 | 0.94 | 2.37 | 0.49 | 0.87 |
| Sunflower | 16.17 | 1.16 | 16.13 | 1.64 | 0.50 | 17.45 | 1.00 | 1.35 |
| Fallow | 0.50 | 0.01 | 0.49 | 0.14 | 0.79 | 0.55 | 0.13 | 0.87 |

**Table 1. Classic, cluster estimator and the average of 100 cluster estimators with a random permutation in each block before building up the clusters, in Sevilla.**

| Salamanca | Classic estimator | | Cluster estimator | | | Cluster estimator with permutations | | |
|---|---|---|---|---|---|---|---|---|
| CROP | area % | std. error | area % | std. error | rel. eff. clus/ran | area % | std. error | rel. eff. per/ran |
| Wheat | 4.06 | 0.53 | 4.41 | 0.47 | 1.26 | 4.06 | 0.36 | 2.18 |
| Barley | 7.03 | 0.93 | 7.65 | 0.96 | 37.32 | 7.03 | 0.59 | 2.52 |
| Sugar beet | 0.41 | 0.15 | 0.45 | 0.16 | 0.85 | 0.41 | 0.14 | 1.20 |
| Sunflower | 1.98 | 0.44 | 2.15 | 0.48 | 0.84 | 1.98 | 0.33 | 1.78 |
| Fallow | 6.43 | 0.82 | 7.05 | 1.02 | 0.64 | 6.44 | 0.65 | 1.58 |

**Table 2. Classic, cluster estimator and the average of 100 cluster estimators with a random permutation in each block before building up the clusters, in Salamanca.**

Estimation by cluster seems to give a significant improvement on the classic estimate (Tables 1, 2) because of the large variability among elements of a cluster, i.e. among segments in different blocks, compared to the difference among segments inside a block (Cochran 77, chapter 8.3).

## 2.2 Cluster Estimator with Permutations by Blocks.

In order to get a more stable variance estimate, we propose a modification to this estimator based on a random permutation of the sample segments in each block before building up the clusters: If we make a random permutation $\pi_b$ of the elements in each block $b$, independent from block to block, the clusters will have another configuration (fig. 2), the new clusters are $C'_j=\{(b,\pi_b(j),\ b=1,..,B\}$. The variance $V_{per}$ that would result could be calculated as the average of all the cluster variances. We have $M!^{B-1}$ different cluster configurations. $V_{per} = \dfrac{\sum_{i=1}^{(M!)^{B-1}} Var(\bar{y}_{clus})_i}{(M!)^{B-1}}$ where $Var(\bar{y}_{clus})_i$ is the cluster variance that results after the $i^{th}$ permutation, with the new configurations of the clusters.

Tables 1 and 2 show that cluster estimator improves the variance in general, but not always. In particular, worse variances are obtained for non dominant crops in Sevilla. The main reason seems to be that the variability inside the block is high because of the low number of fields per segment due to the relatively large size of fields in this province.

## 3. CLUSTER VARIANCES ESTIMATED WITH AND WITHOUT PERMUTATIONS.

### 3.1 The study of *correlation* among segments.

We can also write the variance of a cluster estimator as a function of the *intracluster correlation*, that can be looked at as the correlation among the segments that belong to the same cluster:

$$Var(\bar{y}_{clus}) = \frac{S^2(BM-1)}{B^2 r(M-1)}\left(1-\frac{r}{M}\right)\left[1+(B-1)\rho\right] \qquad (2)$$

This expression is a generalisation of the intra-cluster variance (Cochran 77, page 209),

where the factor $\rho$ represents the *correlation*: $\quad \rho = \dfrac{E\left(y_{ji}-\bar{Y}\right)\left(y_{ui}-\bar{Y}\right)}{E\left(y_{ji}-\bar{Y}\right)^2}$

$y_{ji}$ is the value in the segment of the block $j$ and the cluster $i$. We can write:

$$\rho = \frac{2\displaystyle\sum_{i=1}^{M}\sum_{j<u}^{B}\left(y_{ji}-\bar{Y}\right)\left(y_{ui}-\bar{Y}\right)}{(B-1)\displaystyle\sum_{j,i=1}^{B,M}\left(y_{ji}-\bar{Y}\right)^2} \qquad (3)$$

The *correlation* in this context is not a real correlation but a relation of dependence among segments. If we make a random permutation of all the elements in each block, the clusters will have another configuration, the variance that would result could be calculated as the average of all the cluster variances. We have demonstrated that:

$$V_{per} = \frac{S^2(BM-1)\left(1-\dfrac{r}{M}\right)(1+(B-1)\rho)}{B^2 r(M-1)} \qquad (4)$$

$\rho$ is the average of the $(M!)^{B-1}$ correlations $\rho_w$ among the segments that belong to the

same cluster after the $w^{\underline{th}}$ permutation, $\rho_w = \dfrac{E\left(y_{ji}^w-\bar{Y}\right)\left(y_{ui}^w-\bar{Y}\right)}{E\left(y_{ji}^w-\bar{Y}\right)^2}$ where $y_{ji}^w$ is the

value that the function takes in the segment that belongs to the block $j$ and to the cluster $i$ after the $w^{\underline{th}}$ permutation, we can also write the correlation:

$$\rho_w = \frac{2\displaystyle\sum_{i=1}^{M}\sum_{j<u}^{B}\left(y_{ji}^w-\bar{Y}\right)\left(y_{ui}^w-\bar{Y}\right)}{(B-1)\displaystyle\sum_{j,i=1}^{B,M}\left(y_{ji}^w-\bar{Y}\right)^2} \qquad (5)$$

then the average of these correlations will be   $\rho = \dfrac{1}{(M!)^{B-1}} \displaystyle\sum_{w=1}^{(M!)^{B-1}} \rho_w$ .

The next expression for $V_{per}$ does not depend on the permutations, then we can interpret the correlation that results:

$$V_{per} = \frac{S^2(BM-1)\left(1-\dfrac{r}{M}\right)\left(1+(B-1)\rho^*\right)}{B^2 r(M-1)} \qquad (6)$$

the only difference with respect to (4) is the expression of the factor of correlation, now we have:

$$\rho^* = \frac{2}{(B-1)(BM-1)S^2 M! M}\left[\underbrace{\sum_{i=1}^{M}\sum_{j<u}^{B}\left(y_{ji}-\overline{Y}\right)\left(y_{ui}-\overline{Y}\right)}_{A}+\underbrace{\sum_{i\neq l}^{M}\sum_{j<u}^{B}\left(y_{ji}-\overline{Y}\right)\left(y_{ul}-\overline{Y}\right)}_{B}\right] \quad (7)$$

**A . -** this factor is due to the intra-cluster correlation.
**B . -** this factor can be seen as due to correlation among segments that are in different blocks and have *different* relative position.

If the correlation is 0 among segments that belong to different blocks, variances with and without permutations become the same, $\rho^*$ and $\rho$ are null. With our data we can estimate the factors of correlation and test the hypothesis that the correlation is 0.

$$\left.\begin{array}{l} H_0:\rho=0 \\ H_1:\rho\neq 0 \end{array}\right\} \;\Rightarrow\; \left\langle V(\overline{y}_{clus})=V_{per}\right\rangle$$

We could estimate the variance of the cluster estimator with permutations, with the cluster estimator associated to the new configuration of clusters, although we should condition by the $r$ sampling segment drawn in each block, because we don't have any information about the others segments.

## 4. COMPARISON OF CLUSTER VARIANCE AND VARIANCE AFTER STRATIFICATION.

### 4.1 Considering each block as a stratum.

Although there is not any stratification, we can consider each squared sampling block as a stratum. Then the variance of a stratified estimator with the same sampling units, but when a stratum is formed by a block, is:

$$V(\overline{y}_{st}) = (1-r/M)\frac{1}{r(M-1)B^2}\sum_{j=1}^{B}\sum_{i=1}^{M}\left(y_{ji}-\overline{y}_{j\bullet}\right)^2 \qquad (8)$$

so it is not very difficult to see that:

$$Re\,l.eff.(\overline{y}_{st}/\overline{y}_{clus}) = \frac{Var(\overline{y}_{clus})}{Var(\overline{y}_{st})} = 1 + \tilde{\rho}(B-1) \qquad (9)$$

the interest now is the study of the factor of correlation $\tilde{\rho}$.

$$\tilde{\rho} = \frac{2\sum_{i=1}^{M}\sum_{j<u}^{B}(y_{ji}-\overline{Y}_{j\bullet})(y_{ui}-\overline{Y}_{u\bullet})}{(B-1)\sum_{j,i=1}^{B,M}(y_{ji}-\overline{Y}_{j\bullet})^2} \qquad (10)$$

this factor is the correlation between the deviation from the stratum means of the couples that are in the same cluster. So if we study this factor we can know when the cluster estimator is better than the stratified estimator.

## 4.2 Area frame with stratification in Czech Republic and Portugal.

In the Czech Republic and Portugal, a stratification has been made by photo-interpretation. In Czechoslovakia the blocks are very big, each block has 400 segments of 400 Ha., and there are 5 strata that have different number of clusters: 12, 12, 8, 4 and 2 respectively, with the standard stratified estimator giving better results (Tables 4, 6). This is because blocks are not very homogeneous *intra*, and not very heterogeneous *inter*. With smaller blocks like in Portugal, where blocks are 10Km*10Km and there are 3 strata with different number of clusters: 4, 2 and 1 respectively, cluster estimator performs again worse than the classic stratified estimator (Table 5). Cluster estimator is a good option when there is no stratification like in Spain (Table 3), where blocks are of the same size as in Portugal but each block has 3 sampling segments.

In each stratum segments are relatively similar, so it is not expected that cluster estimator by strata gives very good results, because of the homogeneity among segments in each cluster. The variance is the average of the variances of each different stratum, then if the variance is very low in a stratum this will compensate higher variances in other strata. In general, this doesn't occur with estimator with permutations due to the stability of its variance, in fact global efficiency shows clearly which the real situation is. If we have a stratification we should either avoid using cluster estimator or study carefully the variance in each stratum.

| *PROVINCE* | *WHEAT* | | *BARLEY* | | *SUNFLOWER* | | *FALLOW* | | *SUGAR BEET* | | *VINEYARD* | | *OLIVE* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *(in Spain)* | **area** | **eff.** | **area** | **eff.** | **area** | **eff.** | **area** | **eff.** | **area** | **eff.** | **area** | **eff.** | **area** | **eff.** |
| **Albacete** | 17.5 | 0.96 | 184.2 | 1.16 | 24.1 | 0.94 | 172.1 | 1.48 | 1.4 | 1.03 | 93.7 | 1.60 | 16.8 | 1.55 |
| **Almeria** | 7.8 | 1.02 | 12.8 | 2.56 | - | - | 60.9 | 0.96 | - | - | 2.8 | 1.01 | 20.3 | 1.01 |
| **Avila** | 4.9 | 1.00 | 40.8 | 1.47 | 10.4 | 1.19 | 18.8 | 1.55 | 3.3 | 1.47 | - | - | - | - |
| **Burgos** | 144.7 | 1.52 | 197.1 | 1.60 | 43.2 | 1.13 | 53.3 | 1.10 | 4.1 | 1.10 | 12.8 | 1.31 | - | - |
| **Cadiz** | 105.4 | 1.24 | 4.7 | 0.98 | 78.8 | 1.30 | 3.5 | 1.21 | 22.4 | 0.89 | 18.5 | 1.34 | 16.2 | 1.14 |
| **Cuidad Real** | 28.1 | 1.16 | 158.6 | 1.31 | 6.9 | 0.77 | 158.6 | 1.24 | 2.1 | 1.68 | 190.2 | 1.29 | 75.9 | 1.14 |
| **Cordoba** | 130.5 | 1.74 | 23.8 | 1.00 | 108.6 | 1.92 | 32.0 | 0.98 | 3.5 | 1.49 | 17.2 | 1.15 | 250.9 | 1.56 |
| **Cuenca** | 51.7 | 1.39 | 213.7 | 1.15 | 165.3 | 1.40 | 45.3 | 1.07 | - | - | 95.5 | 1.39 | 24.6 | 1.03 |
| **Granada** | 22.6 | 1.31 | 92.6 | 1.65 | 23.5 | 1.62 | 81.6 | 1.45 | - | - | 4.0 | 1.05 | 146.9 | 1.24 |
| **Guadalajara** | 49.5 | 0.98 | 99.8 | 1.45 | 12.1 | 1.24 | 41.1 | 1.29 | - | - | 3.3 | 1.06 | 16.5 | 1.28 |
| **Huelva** | 21.9 | 1.27 | 4.1 | 1.08 | 28.7 | 1.58 | 9.8 | 1.08 | 2.9 | 1.04 | 4.8 | 1.00 | 21.2 | 1.18 |
| **Jaen** | 29.5 | 1.30 | 28.4 | 0.91 | 15.4 | 1.12 | 36.4 | 1.35 | 2.6 | 1.20 | - | - | 507.5 | 1.43 |
| **Leon** | 20.0 | 1.51 | 19.7 | 1.45 | 1.8 | 1.02 | 70.2 | 1.17 | 3.3 | 1.03 | 4.1 | 0.96 | - | - |
| **Madrid** | 9.7 | 0.86 | 42.7 | 1.18 | 8.3 | 1.25 | 36.5 | 1.05 | - | - | 14.2 | 1.34 | 14.6 | 1.5 |
| **Malaga** | 32.9 | 1.16 | 19.4 | 1.08 | 33.8 | 1.21 | 16.2 | 0.96 | - | - | 2.7 | 1.35 | 88.7 | 1.49 |
| **Palencia** | 73.0 | 1.14 | 153.5 | 1.50 | 32.1 | 1.10 | 63.0 | 1.38 | 8.8 | 1.02 | 1.5 | 1.03 | - | - |
| **Segovia** | 54.5 | 1.44 | 100.1 | 1.56 | 34.9 | 0.94 | 53.0 | 1.51 | 4.4 | 0.68 | 1.0 | 1.16 | - | - |
| **Soria** | 69.0 | 1.41 | 89.8 | 1.29 | 37.1 | 1.91 | 46.3 | 1.24 | - | - | - | - | - | - |
| **Toledo** | 57.6 | 1.02 | 226.2 | 1.34 | 9.0 | 0.97 | 209.2 | 1.28 | - | - | 186.8 | 1.74 | 93.7 | 1.20 |
| **Zamora** | 50.1 | 1.22 | 85.3 | 1.38 | 22.8 | 1.28 | 147.5 | 1.08 | 8.5 | 1.13 | 15.8 | 1.02 | - | - |

**Table 3.** Area estimates (in 1000 Ha.) and relative efficiency of cluster estimates with permutation (100 repetitions) in several provinces in Spain.

| CROP | STRATIFIED ESTIMATOR | | | CLUSTER ESTIMATOR | | | PERMUTATION ESTIMATOR | | |
|------|-----------|-----------|-----------------|-----------|-----------|-----------------|-----------|-----------|-----------------|
| | area (ha.) | std. error | rel.ef. str/ran | area (ha.) | std. error | rel.ef. clu/str | area (ha.) | std. error | rel.ef. per/str |
| Wheat | 824731 | 44488 | 1.19 | 822530 | 33338 | 1.78 | 828806 | 47988 | 0.86 |
| Barley | 654669 | 33677 | 1.07 | 659901 | 32672 | 1.06 | 659088 | 36637 | 0.84 |
| Rapeseed | 140200 | 16262 | 1.31 | 142453 | 16044 | 1.03 | 139270 | 18159 | 0.80 |
| Sug. Beet | 119522 | 13776 | 1.08 | 116496 | 13771 | 1.00 | 118162 | 14639 | 0.89 |
| Potatoes | 108838 | 14751 | 1.16 | 112545 | 13118 | 1.26 | 108260 | 17582 | 0.70 |
| Maize | 360810 | 26994 | 1.17 | 360016 | 28087 | 0.92 | 364803 | 27933 | 0.93 |
| Forest | 1895724 | 84641 | 0.57 | 1889710 | 67664 | 1.56 | 1881747 | 89136 | 0.90 |
| Oth.Plant | 2927335 | 133198 | 0.81 | 2940234 | 72726 | 3.35 | 2933805 | 146916 | 0.82 |

Table 4.- Comparison of classic stratified and cluster estimators in the **Czech Republic**. Global results.

| CROP | STRATIFIED ESTIMATOR | | | CLUSTER ESTIMATOR | | | PERMUTATION ESTIMATOR | | |
|------|-----------|-----------|-------------------|-----------|-----------|-------------------|-----------|-----------|-------------------|
| | area (ha.) | std. error | rel.ef. (str/ran) | area (ha.) | std. error | rel.ef. (clu/str) | area (ha.) | std. error | rel.ef. (per/str) |
| Wheat | 20556 | 3518 | 2.10 | 20244 | 3757 | 0.88 | 20655 | 4264 | 0.68 |
| Rye | 28984 | 3472 | 0.77 | 28653 | 4077 | 0.73 | 29219 | 3605 | 0.93 |
| Maize | 12124 | 1834 | 1.22 | 11986 | 2128 | 0.74 | 12079 | 1900 | 0.93 |
| Potatoes | 12299 | 1686 | 1.26 | 12316 | 1937 | 0.76 | 12317 | 1809 | 0.87 |
| Annual | 9682 | 1935 | 2.05 | 9619 | 2222 | 0.76 | 9755 | 2395 | 0.65 |
| Fallow | 84677 | 8481 | 2.19 | 85019 | 8458 | 1.01 | 85028 | 8694 | 0.95 |
| Vineyard | 91878 | 12204 | 1.33 | 91929 | 3688 | 10.95 | 90423 | 15962 | 0.58 |
| Olive | 62296 | 8196 | 1.48 | 62473 | 8355 | 0.96 | 61818 | 8755 | 0.88 |
| Perman. | 49618 | 6409 | 1.61 | 49594 | 6334 | 1.02 | 49813 | 6877 | 0.87 |
| Forest | 194463 | 23966 | 0.49 | 192640 | 38598 | 0.39 | 194473 | 24544 | 0.95 |

Table 5.- Stratified and cluster estimators in **Tras-os-Montes (Portugal)**.

Applied Statistics in Agriculture

| CROP | | STRATIFIED ESTIMATOR | | CLUSTER ESTIMATOR | | | PERMUTATION ESTIMATOR | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *c.v. (%)* | *std. error* | *c.v. (%)* | *std. error* | *rel.ef. (clu/str)* | *c.v. (%)* | *std. error* | *rel.ef. (per/str)* |
| **Wheat** | Stratum 4 | 24.91 | 11226 | 29.23 | 12678 | - | 26.88 | 12269 | - |
| | Stratum 5 | 61.64 | 26538 | 24.16 | 10739 | - | 62.17 | 25626 | - |
| | **Global** | 5.39 | 44488 | 4.05 | 33338 | 1.78 | 5.79 | 47988 | 0.86 |
| **Maize** | Stratum 1 | 8 | 14622 | 6.86 | 12432 | - | 8.49 | 15702 | - |
| | Stratum 5 | 64.60 | 17530 | 69.85 | 20860 | - | 62.85 | 16959 | - |
| | **Global** | 7.48 | 26994 | 7.80 | 28087 | 0.92 | 7.66 | 27933 | 0.93 |
| **Other** | Stratum 4 | 9.25 | 35318 | 10.75 | 42474 | - | 9.48 | 35690 | - |
| | Stratum 5 | 9.39 | 113400 | 2.11 | 25518 | - | 10.39 | 126517 | - |
| | **Global** | 4.55 | 133198 | 2.47 | 72726 | 3.35 | 5.01 | 146916 | 0.82 |

**Table 6-** Classic stratified and cluster estimators in the **Czech Republic.** Results for some strata.

## SUMMARY

In an area frame of squared segments on squared blocks, where clusters are built with segments that have the same relative position in each block, if there is no stratification and blocks are relatively small, cluster estimator gives lower variances than standard formulae, although there is the problem of its unstable variance. Making independent permutations of the sampling segments in each block, if blocks are complete, gives the same estimation as using cluster estimator, but improves the stability of the variance. When there is a stratification, cluster estimator is not a good choice due to the correlation among segments in each stratum: The cluster estimator works well when segments in each cluster are quite heterogeneous. Area frame in Spain is without stratification; then cluster estimator with permutations is a good option to improve the variance.

## REFERENCES:

Allen, J.D., 1990, A Look at the Remote Sensing Applications Program of the National Agricultural Statistics Service. Jou. of Official Stat. Vol 6, n. 4, pp. 393-409.

Ambrosio L., Alonso R., Villa A., 1993, Estimación de superficies cultivadas por muestreo de áreas y teledetección. Precisión relativa. Estadística Española, Vol 35, pp. 91-103.

Cochran W., 1977, Sampling Techniques. New York: John Wiley and Sons

Cotter, J. Nealon J. (1987), Area Frame design for Agricultural Surveys. U.S. Dept. of Agriculture. Nat. Agr. Stat. Serv.

Gallego F.J., Delincé J., Rueda C., 1993, Crop area estimates through remote sensing: stability of the regression correction. Int. J. Remote Sensing. Vol 14, n.18, pp 3433-3445.

Gallego, F.J., Delincé, J.,Carfagne E., 1994 Two-Staged Area Frame Sampling On Squared Segments For Farm Surveys. Survey Methodology (in print).

González F., López S., Cuevas J.M., 1991, Comparing Two Methodologies for Crop Area Estimation in Spain Using Landsat TM Images and Ground Gathered Data. Remote Sens. Environ. no 32, pp. 29-36.

Meyer-Roux J., 1990. Présentation du Projet Pilote de Télédétection Appliquée aux Statistiques Agricoles. Conference on the Appl. of Remote Sensing to Agricultural Statistics. Office for Publications of the E.C. Luxembourg.