Image Anal Stereol 2019;38:131-139 doi:10.5566/ias.1991
Original Research Paper

VARIANCE PREDICTION FOR POPULATION SIZE ESTIMATION

Ana I. Gómez✉, Marcos Cruz and Luis M. Cruz-Orive

Department of Mathematics, Statistics and Computation, Faculty of Sciences, University of Cantabria, Avda. Los Castros 48, E-39005 Santander, Spain.
e-mail: gomezanab@gmail.com, marcos.cruz@unican.es, luis.cruz@unican.es
(Received July 28, 2018; revised March 5, 2019; accepted April 8, 2019)

ABSTRACT

Design unbiased estimation of population size by stereological methods is an efficient alternative to automatic computer vision methods, which are generally biased. Moreover, stereological methods offer the possibility of predicting the error variance from a single sample. Here we explore the statistical performance of two alternative variance estimators on a dataset of 26 labelled crowd pictures. The empirical mean square errors of the variance predictors are compared by means of Monte Carlo resampling.

Keywords: Cavalieri error variance predictor, geometric sampling, Monte Carlo resampling, particle counting, population size, Split error variance predictor, systematic quadrats.

INTRODUCTION

Population size is an important parameter for instance in ecology and social sciences. The target parameter is the size of a discrete population of objects, generally called 'particles'. A particle is a compact set separated from other particles, serving as a model for a bird, a human, *etc*. In the present context the population is projected onto an observation plane which in practice is the plane of an image. The result is a finite set $Y = \{y_1, y_2, \ldots, y_N\}$ of N particle projections contained in a bounded region of the plane, with the condition that all the particle projections are unambiguously distinguishable for counting. Thus, the effective particle population is $Y \subset \mathbb{R}^2$, and the target is N – unobsevable original particles are ignored. In practice the sampling unit is a conveniently defined subset of a particle projection, *e.g.*, a projected human head, or a distinguishable part of it, see Fig. 1a. For automatic resampling, however, each particle is replaced with an associated point, see Fig. 1b. Thus, henceforth $y_i \in Y$ may represent the ith particle in the 'real life' sampling frame, or the ith point particle in the 'computer simulation' sampling frame. For short, the former may be called the 'real mode', the latter the 'computer mode'.

A design unbiased method to estimate N, named *CountEm* (https://countem.unican.es/), based on systematic quadrats, was recently proposed (Cruz *et al.*, 2015; Cruz and González-Villa, 2018a;b). The method is familiar in quantitative microscopy, (Miles, 1978; Howard and Reed, 2005), and can be applied to any population that can be mapped onto an observation plane. In the real mode, a particle is sampled by a quadrat according to the forbidden line rule (Gundersen, 1977). In the computer mode, however, a point particle is sampled if it is contained in the quadrat, simply.

The CountEm method is based on manual counting of reasonably small samples. By means of Monte Carlo resampling, Cruz and González-Villa (2018a) have shown that manually counting about 100 individuals, in about 30 nonempty systematic quadrats, usually yield coefficients of error (*i.e.*, relative standard errors) below 10%, irrespective of population size. Apart from the bonus of design unbiasedness, with such small sample sizes *CountEm* is an attractive alternative to automatic methods, which are hitherto biased to unknown degrees (for references on the shortcomings of automatic particle detection see the latter paper and the *Introduction* of Cruz *et al.*, 2015).

An error variance predictor for the design unbiased estimator $\widehat{N}$ of N, denoted by $\mathrm{var}_{\mathrm{Cav}}(\widehat{N})$, was presented in Cruz *et al.* (2015). The estimator has a between stripes contribution handled by the Cavalieri approach, plus a contribution among quadrats within stripes, for which the splitting approach was adopted (with two subsamples per stripe). Here we propose a new, alternative estimator, $\mathrm{var}_{\mathrm{split}}(\widehat{N})$, for which the estimation of either contribution is based on a splitting design – for details and references see the *Appendix*. The variance predictors are compared and matched against the empirical variance, $\mathrm{Var}_e(\widehat{N})$, and their relative accuracy is compared via their mean square error (MSE). Automatic Monte Carlo resampling is used on 26 point particle populations manually edited from real life images, covering a variety of patterns.

COUNTEM METHOD

The idea of the method is to perform systematic sampling by superimposing on the population image a uniform random (UR) test grid $\Lambda_x \subset \mathbb{R}^2$ of quadrats, see Fig. 1b, of a fixed orientation (namely a FUR test system, a term introduced by Miles and Davy, 1977). The adopted fundamental tile is a square $J_0 = [0,T)^2$. The tiles form a partition of the plane, and any tile J_k of the test system can be brought to coincide with J_0 by a translation $-\tau_k$ that leaves the entire test system unchanged. The fundamental probe is a square $T(0) = [0,t)^2 \subseteq J_0$, that is, $0 < t \leq T < \infty$. For a description of test systems see Santaló (1976), or Gual Arnau and Cruz-Orive (1998). The probe $T(0)$ is shifted into $T(x) = T(0) + x$ by a UR translation x within J_0, dragging the entire test system into $\Lambda_x = \Lambda_0 + x$. Thus,

$$\Lambda_x = \{T(x+\tau_k),\ k \in \mathbb{Z}\}\,, \quad x \sim \mathrm{UR}(J_0). \qquad (1)$$

Under the preceding design, an unbiased estimator (UE) of the population size N is the product of the sample size Q, namely the total number of particles counted with the grid under the aforementioned sampling rules, times the area sampling period T^2/t^2, that is,

$$\widehat{N} = \frac{T^2}{t^2} \cdot Q\,. \qquad (2)$$

For a proof of the unbiasedness of $\widehat{N}$ under the computer mode (point particles, *e.g.*, Fig. 1b) see Appendix S1 from Cruz *et al.* (2015). The unbiasedness property is independent from the orientation of the test system relative to the population. If the shape of the convex hull of the population image is nearly rectangular (as in Fig. 1b), however, then it is advisable to avoid parallelism between the edges of the image and the quadrat rows, or columns, in order to avoid an unduly large error variance. This idea is suggested in Fig. (11) from Gundersen *et al.* (1999). In the present experiment, each of the population images studied was framed by a rectangle, for which we adopted a common tilting angle of 30° with respect to the base of the frame.

Cruz and González-Villa (2018a) propose a simple way to choose a grid compatible with preestablished values of the sample size and the number of nonempty quadrats.

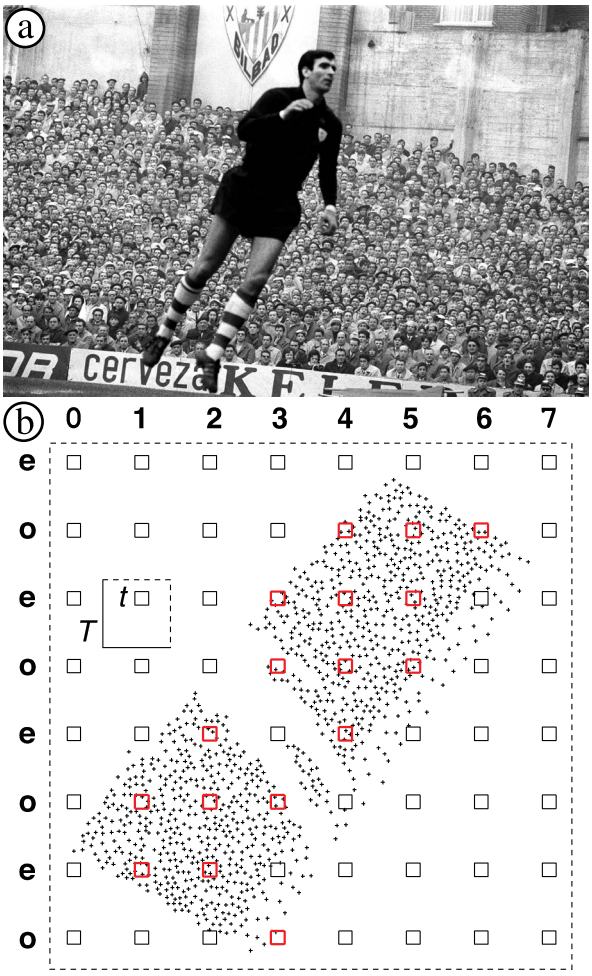

Fig. 1. *(a): Spectators in a football match (Bilbao, 1966), taken from Cancio (2010) with permission from the author. (b): Corresponding associated point particles, as used for Monte Carlo automatic resampling, with a tilted, FUR test grid of quadrats, superimposed on them. The fundamental tile is a square of side T, the fundamental probe is a square quadrat of side* $t \in (0,T]$.

VARIANCE PREDICTORS

CAVALIERI VARIANCE PREDICTOR

The aforementioned variance predictor $\mathrm{var}_{\mathrm{Cav}}(\widehat{N})$ takes into account quadrat dependence using G. Matheron's transitive theory (Matheron, 1971). Hints on the derivations, and earlier references, are given in the *Appendix*. Here the idea is to regard the quadrat sample as a two stage sample. The first stage involves planar Cavalieri stripes of thickness $t > 0$ a constant period $T > t$ apart (*e.g.*, columns numbered from 0 to 7 in Fig. 1b). In the second stage, each stripe is subsampled in turn by a perpendicular series of Cavalieri stripes with the same parameters $\{t,T\}$ (*e.g.*,

even and odd rows, labeled e and o respectively in Fig. 1b). The result is equivalent to a grid of systematic quadrats with the latter parameters.

Next we define the necessary notation:

- $\tau = t/T \in (0,1]$, stripe sampling fraction.
- n: number of stripes encompassing the particle population, $(n > 2)$.
- n_i: number of quadrats subsampled within the ith stripe section, $i = 1, 2, \ldots, n$.
- q_{ij}: number of particles captured by the jth quadrat within the ith stripe, $j = 1, 2, \ldots, n_i$.
- Q_{oi}, Q_{ei}: total numbers of particles captured by the odd numbered, and by the even numbered quadrats, respectively, within the ith stripe.
- $Q_i = \sum_{j=1}^{n_i} q_{ij}$, total number of particles sampled in the ith stripe. Note that $Q_i = Q_{oi} + Q_{ei}$.
- $Q_o = \sum_{i \text{ odd}} Q_i$, $Q_e = \sum_{i \text{ even}} Q_i$, total number of particles on the odd numbered stripes and on the even numbered stripes, respectively.
- $Q = \sum_{i=1}^{n} Q_i$, total number of sampled particles.

The Cavalieri variance predictor (Eq. 3 of Cruz *et al.*, 2015) is:

$$\begin{aligned} \mathrm{var}_{\mathrm{Cav}}(\widehat{N}) &= \frac{c(\tau)}{\tau^4}\,[3(C_0 - \widehat{v}_n) - 4C_1 + C_2] + \frac{\widehat{v}_n}{\tau^4}\,, \\ c(\tau) &= \frac{(1-\tau)^2}{6(2-\tau)}\,, \\ C_k &= \sum_{j=1}^{n-k} Q_j Q_{j+k}\,, \quad k = 0, 1, 2, \\ \widehat{v}_n &= \sum_{i=1}^{n} \mathrm{var}(Q_i)\,. \end{aligned} \tag{3}$$

The first term in the right hand side of the first Eq.3 estimates the between stripes variance contribution, whereas $\widehat{v}_n/\tau^4$ estimates the within stripes contribution, which can be estimated by splitting the quadrat counts within the ith stripe into two subsamples with total counts Q_{oi}, Q_{ei}, respectively. Thus,

$$\begin{aligned} \widehat{v}_n &= s(\tau) \sum_{i=1}^{n} (Q_{oi} - Q_{ei})^2\,, \\ s(\tau) &= \frac{(1-\tau)^2}{3-2\tau}\,. \end{aligned} \tag{4}$$

SPLIT VARIANCE PREDICTOR

Here the between and the within stripes contributions are both handled by means of the splitting design adopted for v above. Thus, the corresponding predictor, alternative to Eq. 3, reads,

$$\mathrm{var}_{\mathrm{split}}(\widehat{N}) = \frac{s(\tau)}{\tau^4}[(Q_o - Q_e)^2 - \widehat{v}_n] + \frac{\widehat{v}_n}{\tau^4}\,. \tag{5}$$

NEW MODIFICATION

In Cruz *et al.* (2015), the first Eq. 3, and Eq. 5, were computed for a single direction of the stripes (*e.g.*, along the colums 0–7 in Fig. 1b). Here we have introduced the following, new modification: For each population, the mentioned predictors were computed first for a given direction of the stripes, and then for the perpendicular direction. In each case, the final variance predictor was the average of the corresponding two predictors.

In the figures, the predictors $\mathrm{var}_{\mathrm{Cav}}(\widehat{N})$ and $\mathrm{var}_{\mathrm{split}}(\widehat{N})$ are denoted by $C1$ and $S1$, respectively, when computed along a single direction, and by $C2$ and $S2$ when computed along two directions.

REMARKS ON NOTATION

In the sequel, true variances will be denoted by $\mathrm{Var}(\cdot)$, whereas their predictors, or estimators, will be denoted by $\mathrm{var}(\cdot)$. The true mean or expected value of a random variable (*e.g.*, an estimator) will be denoted by $\mathbb{E}(\cdot)$. For a positive random variable, the square coefficient of variation is $\mathrm{CV}^2(\cdot) = \mathrm{Var}(\cdot)/\{\mathbb{E}(\cdot)\}^2$. If the random variable is an estimator, then the equivalent notation $\mathrm{CE}^2(\cdot)$, called the square coefficient of error, is used.

DATASET OF POINT PARTICLES

The variance predictors $\mathrm{var}_{\mathrm{Cav}}(\widehat{N})$ and $\mathrm{var}_{\mathrm{split}}(\widehat{N})$ were compared on the aforementioned dataset of 26 point particle populations, see Fig. 2 for a subset, using a Monte Carlo resampling procedure described in the next section. One of the images was Fig. (1a) from Cruz *et al.* (2015). The 25 remaining images were borrowed from the *UCF dataset* (Idrees *et al.*, 2013). The original *UCF dataset* consists of 50 images, but we have selected those with $N > 1000$ since these are the population sizes for which *CountEm* is more useful. The images were sorted by increasing N and numbered from 1 to 26. A subsample of six images is displayed in Fig. 2.

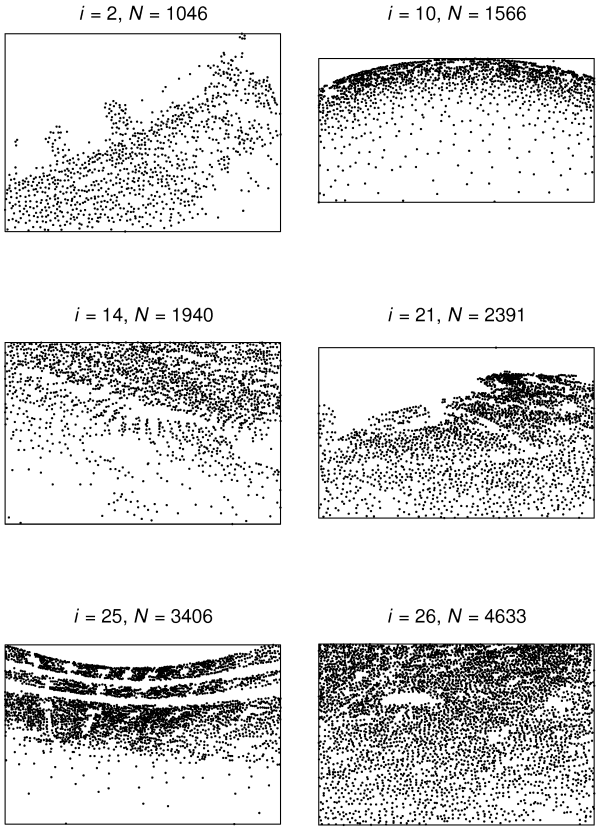

Fig. 2. *Six of the 26 point particle populations constituting the dataset, ready for automatic Monte Carlo resampling. Each point was associated with a real particle from the original image, as in Fig. 1.*

EMPIRICAL ASSESSMENT OF THE VARIANCE PREDICTORS BY MONTE CARLO RESAMPLING

The empirical variance of $\widehat{N}$ and the performances of $\mathrm{var}_{\mathrm{Cav}}(\widehat{N})$ and $\mathrm{var}_{\mathrm{split}}(\widehat{N})$ were evaluated by Monte Carlo resampling on each of the 26 point particle populations available with the aid of the R-spatstat package (Baddeley *et al.*, 2015).

For comparative purposes, the sampling fraction was prescribed as $f = Q/N$, where $Q = 100$, and N was the known population size, namely the number of manually annotated points in each image. The initial numbers of quadrats considered were $n_0 = 50$ and $n_0 = 100$. Practical prescriptions for the real case in which N is unknown are provided by Cruz and González-Villa (2018b).

As indicated in the preceding paper, with the prescribed pair $\{f, n_0\}$, the grid parameters $\{t, T\}$ were computed as follows,

$$T = \sqrt{\frac{B_x B_y}{n_0}}\,, \qquad (6)$$
$$t = T\sqrt{f}\,,$$

where B_x, B_y represent the width and the height of the image frame, respectively. As mentioned in the second section, the resulting grid was tilted by 30° with respect to the horizontal axis. Next we recall the necessary notation to describe the resampling procedure:

- $Y = \{y_1, y_2, \ldots, y_N\} \subset \mathbb{R}^2$: finite set – called the 'population' – of N point particles in a bounded area. Here $y_i \in Y$ denotes the ith point particle.
- Λ_x: FUR grid of quadrats, see Eq. 1 .
- Q: random number of point particles captured by the quadrats.

For each pair $\{t, T\}$ a total of $M \equiv K^2 = 32^2 = 1024$ replicated superimpositions of the grid Λ_x onto Y were generated, corresponding to M systematic replications of the point x within J_0, arranged into a $K \times K$ UR subgrid within J_0 – this design should be expected to be more efficient than one of independent random replications (Cruz *et al.*, 2015). We relabel the M systematic sites of the point x as $\{x_k,\ k = 1, 2, \ldots, M\}$.

For each k, the corresponding sample total,

$$Q_k = Q(Y \cap \Lambda_{x_k})\,, \qquad (7)$$

was computed automatically. From Eq. 2, the kth UE of N is:

$$\widehat{N}_k = (T/t)^2 \cdot Q_k\ . \qquad (8)$$

For each population, the empirical mean, variance and square coefficient of error of $\widehat{N}$ were computed respectively as follows,

$$\mathbb{E}_e(\widehat{N}) = \frac{1}{M}\sum_{k=1}^{M} \widehat{N}_k\,, \qquad (9)$$

$$\mathrm{Var}_e(\widehat{N}) = \frac{1}{M}\sum_{k=1}^{M} \left[\widehat{N}_k - \mathbb{E}_e(\widehat{N})\right]^2, \qquad (10)$$

$$\mathrm{CE}_{\mathrm{e}}^2(\widehat{N}) = \mathrm{Var}_e(\widehat{N})/N^2\ . \qquad (11)$$

On the other hand, to check the performance of the variance predictors we also computed the corresponding sets of M replicates $\{\mathrm{var}_{\mathrm{Cav}}(\widehat{N}_k)\}$ and $\{\mathrm{var}_{\mathrm{split}}(\widehat{N}_k)\}$ of the variance predictors, as well as the corresponding square coefficients of error, namely,

$$\mathrm{ce}_{(\cdot)}^2(\widehat{N}_k) = \mathrm{var}_{(\cdot)}(\widehat{N}_k)/N^2\,, \qquad (12)$$

where $(\cdot)$ stands for either 'Cav' or 'split'. In the denominator we use N instead of $\widehat{N}_k$ because we are

interested in the behaviour of the numerator (*i.e.*, in the variance predictor itself). The corresponding summary measure was computed as,

$$\mathrm{ce}^2_{(\cdot)}(\widehat{N}) = \frac{1}{M}\sum_{k=1}^{M}\mathrm{ce}^2_{(\cdot)}(\widehat{N}_k)\,. \tag{13}$$

The preceding three equations are useful to illustrate the performance of the variance predictors graphically, see the next section. To describe the statistical performance of the variance predictors, however, we computed their mean square errors, namely,

$$\mathrm{MSE}_e(\mathrm{var}_{(\cdot)}(\widehat{N})) = \frac{1}{M}\sum_{k=1}^{M}[\mathrm{var}_{(\cdot)}(\widehat{N}_k) - \mathrm{Var}_e(\widehat{N})]^2. \tag{14}$$

Recall that, for an estimator '$\cdot$', $\mathrm{MSE}(\cdot) = \mathrm{Var}(\cdot) + \mathrm{Bias}^2(\cdot)$. The first term in the right hand side of the preceding identity measures the variation of the estimator about its own mean, whereas the bias is the mean deviation of the estimator about the true value of the target parameter.

We will also use the normalized MSE, which we call the coefficient of MSE because it is analogous to the square coefficient of error, namely,

$$\mathrm{CMSE}_e(\mathrm{var}_{(\cdot)}(\widehat{N})) = \frac{\mathrm{MSE}_e(\mathrm{var}_{(\cdot)}(\widehat{N}))}{\mathrm{Var}^2_e(\widehat{N})}\,. \tag{15}$$

As advanced in the *Introduction*, for a given set of sampling parameters we want to compare the statistical performance of the Cavalieri and the Split variance predictors on each of $n = 26$ point particle populations of sizes $N_1, N_2, ..., N_n$. For the ith population we define the relative empirical accuracy of the latter to the former predictor as follows (*e.g.*, Lindgren, 2017),

$$\mathrm{RA}_i(Cav, Split) = \frac{\mathrm{MSE}_e(\mathrm{var}_{\mathrm{split}}(\widehat{N}_i))}{\mathrm{MSE}_e(\mathrm{var}_{\mathrm{Cav}}(\widehat{N}_i))}\,. \tag{16}$$

For instance, if $\mathrm{RA}_i(Cav, Split) > 1$, then the Cavalieri variance predictor is 'better', *i.e.*, more accurate, than the Split one for the ith population, in the sense that it has the smaller MSE there. If the variance predictors involved were unbiased, then the RA index would be the usual relative efficiency ratio. As a global numerical summary from the n populations we propose the relative empirical accuracy index as

$$\mathrm{RA}(Cav, Split) = \frac{\sum_{i=1}^{n}\mathrm{MSE}_e(\mathrm{var}_{\mathrm{split}}(\widehat{N}_i))}{\sum_{i=1}^{n}\mathrm{MSE}_e(\mathrm{var}_{\mathrm{Cav}}(\widehat{N}_i))}\,. \tag{17}$$

RESULTS

For each of the 26 populations, for each variance predictor, and for each of two different sampling intensities, the descriptors $\mathrm{CE}^2_e(\widehat{N}_i)$ and $\mathrm{ce}^2_{(\cdot)}(\widehat{N}_i)$ defined above are represented in Fig. 3 by a black and a coloured line, respectively. When computed using a single direction, the latter values are linked by a coloured broken line. Here $\widehat{N}_i$ denotes the UE of N_i , $(i = 1, 2, ..., n = 26)$. In addition, for each predictor $C2$ and $S2$ the $M = 1024$ individual replications given by the square root of Eq. 12 are enclosed within the corresponding 95% confidence bands.

For each of the mentioned cases, Fig. 4 displays the corresponding $\mathrm{CMSE}^{1/2}$'s, namely the square roots of Eq. 15. Further, Fig. 5 displays the individual relative accuracy indices computed via Eq. 16. Finally, the summary RA indices computed with Eq. 17 are displayed in Table 1.

Table 1. *Summary indices of relative accuracy (RA) computed via Eq. 17. The abbreviations $C1$, $S1$, $C2$, and $S2$ are defined in the subsection 'New modification'.*

	$n_0 = 50$	$n_0 = 100$
$RA(C2, S2)$	1.17	1.13
$RA(C1, S1)$	1.21	1.16
$RA(C2, C1)$	1.01	1.08
$RA(S2, S1)$	1.05	1.11

DISCUSSION AND CONCLUSIONS

A new error variance predictor, called the Split predictor, $\mathrm{var}_{\mathrm{split}}(\widehat{N})$, was proposed for the *CountEm* population size estimator. Its performance was tested against the current Cavalieri predictor, $\mathrm{var}_{\mathrm{Cav}}(\widehat{N})$, on a dataset of 26 point particle populations.

The Cavalieri variance predictor was statistically better than the Split predictor for the dataset considered, in the sense that the individual relative accuracy ratios, see Eq. 16, were always greater than 1, see Fig. 5. In this graph the variance predictors were $C2$, $S2$, as defined in the subsection *New modification*. The corresponding summary index $\mathrm{RA}(C2, S2)$ also reveals a superiority of $C2$ over $S2$, see Table 1. However, the mean values of $\mathrm{ce}_{(\cdot)}(\widehat{N}_i)$ for the individual populations, represented by red and blue continuous lines in Fig. 3, were rather similar among $C2$ and $S2$ for each value of n_0. This is compatible with the fact both predictors are obtained under the same model assumptions. A similar result was obtained by Cruz-Orive and Geiser (2004) for FUR

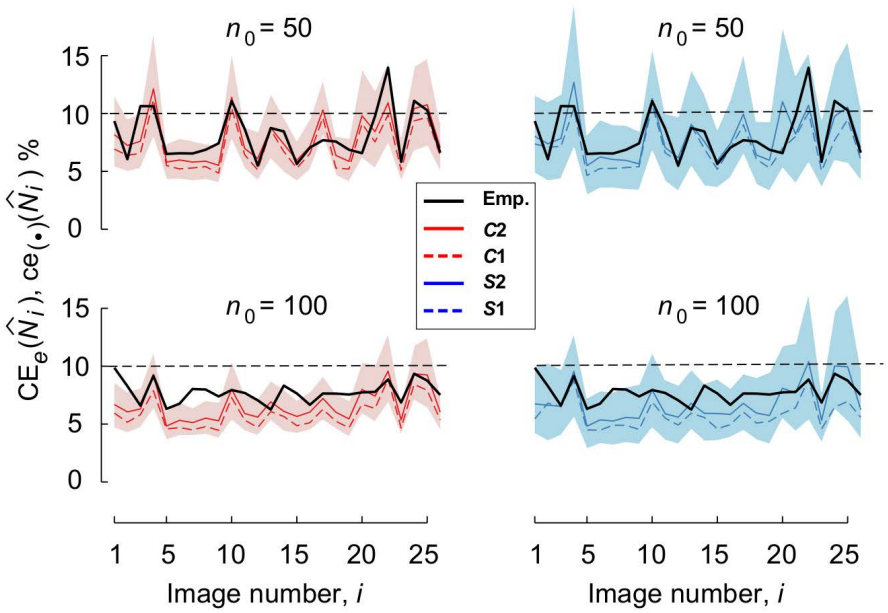

Fig. 3. *Behaviour of the empirical, and the mean predicted, coefficients of error of the number estimators, computed for each particle population as the square roots of Eq. 11 and Eq. 13 respectively. The coloured bands enclose the* 95% *of the Monte Carlo replications computed for each population as the square root of Eq. 12. The inset symbols are defined in the subsection 'New modification'.*

Cavalieri disectors under the Cavalieri and the Split designs, see Fig. 8 of the latter paper.

With the new modification the variance predictors were globally at least as accurate as the ones computed in a single direction, as shown in Table 1. For individual populations, however, there were some exceptions, see Fig. 4. Table 1 also suggests that the gain in precision due to this modification was more important for $n_0 = 100$ than for $n_0 = 50$.

Both variance predictors often underestimated the empirical empirical error variance, see Fig. 3. As shown in Fig. 4, however, the relative $\text{MSE}_e(\text{var}_{(\cdot)}(\widehat{N}_i))$, see Eq. 15, tended to be larger for $n_0 = 100$ than for $n_0 = 50$. We conjecture that this may be due to the fact that with $n_0 = 100$ the mean number of particles per quadrat decreases, this tending to induce independence among quadrat counts. Recall that either variance predictor relies on modelling quadrat count dependence using G. Matheron's transitive theory – hence the predictors tend to be less accurate under quadrat count independence.

In short, based on the RA_i and RA relative accuracy indices, for the dataset studied the preferred predictor was $C2$, namely $\text{var}_{\text{Cav}}(\widehat{N})$ computed as the average among the two contributions along the two mutually perpendicular stripe directions of the quadrat grid.

ACKNOWLEDGEMENTS

The authors acknowledge financial support from the AYA-2015-66357-R (MINECO/FEDER) project.

REFERENCES

Baddeley A, Rubak E, Turner R (2015). Spatial Point Patterns: Methodology and Applications with R. New York: Chapman and Hall/CRC.

Cancio R (2010). Simplemente ... Periodismo. Madrid: Ediciones APM.

Cruz M, Gómez D, Cruz-Orive LM (2015). Efficient and unbiased estimation of population size. PLoS ONE 10:e0141868.

Cruz M, González-Villa J (2018a). Simplified procedure for efficient and unbiased population size estimation. PloS ONE 13:e0206091.

Cruz M, González-Villa J (2018b). Unbiased population size estimation on still gigapixel images. Sociol Method Res 0049124118799373.

Cruz-Orive LM (2004). Precision of the fractionator from Cavalieri designs. J Microsc 213:205–11.

Cruz-Orive LM (2006). A general variance predictor for Cavalieri slices. J Microsc 222:158–65.

Cruz-Orive LM, Geiser M (2004). Estimation of particle number by stereology: An update. J Aerosol Med 17:197–212.

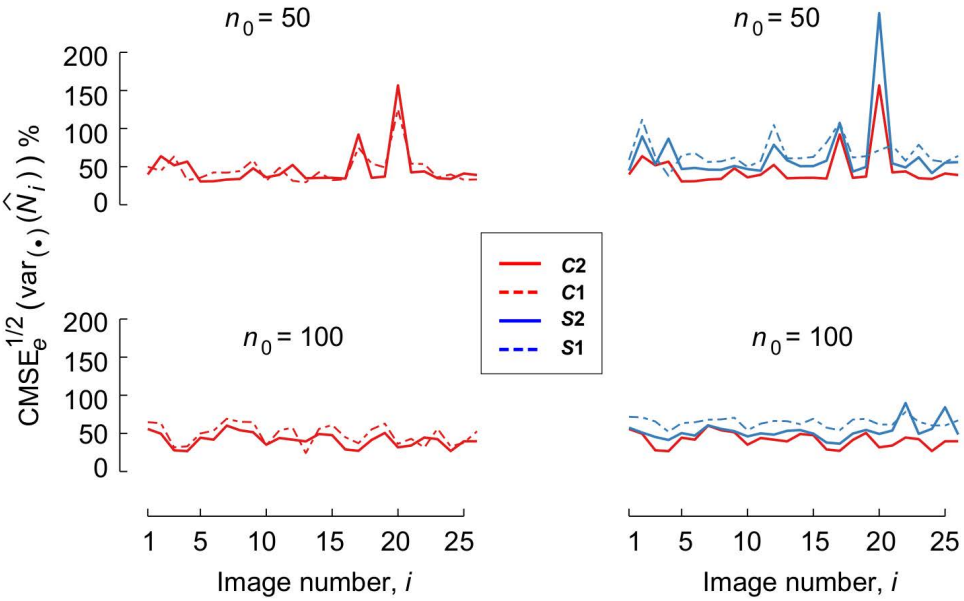

Fig. 4. *Behaviour of the coefficient of square root mean square error, computed for each particle population as the square root of Eq. 15, in percent. The inset symbols are as in Fig. 3.*

Gual Arnau X, Cruz-Orive LM (1998). Variance prediction under systematic sampling with geometric probes. Adv Appl Probab 30:889–903.

Gundersen HJG (1977). Notes on the estimation of the numerical density of arbitrary profiles: the edge effect. J Microsc 111:219–23.

Gundersen HJG, Jensen EBV, Kiêu K, Nielsen J (1999). The efficiency of systematic sampling in stereology — reconsidered. J Microsc 193:199–211.

Howard CV, Reed MG (2005). Unbiased Stereology. Three-dimensional Measurement in Microscopy, 2nd Ed. Oxford: Bios/Taylor & Francis.

Idrees H, Saleemi I, Seibert C, Shah M (2013). Multi-source multi-scale counting in extremely dense crowd images. In: Proc IEEE Conf Comput Vision Pattern Recogn (CVPR) 2547–54.

Lindgren B (2017). Statistical Theory. Routledge.

Matheron G (1971). The Theory of Regionalized Variables and its Applications, vol. 5. Les Cahiers du Centre de Morphologie Mathèmatique de Fontainebleau, No. 5. Fontainebleau: École national supérieure des mines de Paris.

Miles RE (1978). The sampling, by quadrats, of planar aggregates. J Microsc 113:257–67.

Miles RE, Davy P (1977). On the choice of quadrats in stereology. J Microsc 110:27–44.

Santaló LA (1976). Integral Geometry and Geometric Probability. Addison-Wesley, Reading, Massachusetts.

APPENDIX: HINTS ON THE DERIVATION OF THE ERROR VARIANCE PREDICTORS

Let $Y \subset \mathbb{R}^2$ represent a finite population of N point particles contained in a bounded region of the plane. Fix a sampling axis Ox, and let $L(x)$ denote a straight line normal to Ox at a point of abscissa x. A stripe $L_t(x)$ of thickness $t > 0$ is the portion of the plane comprised between the two parallel lines $L(x)$ and $L(x+t)$. Let $Q_t(x)$ denote the total number of particles captured by $L_t(x)$, namely the cardinality of $Y \cap L_t(x)$. For Matheron's transitive theory to work, it is necessary to assume that there exists a function $f : \mathbb{R} \to \mathbb{R}^+$ such that,

$$Q_t(x) = \int_x^{x+t} f(y)\,\mathrm{d}y\,, \quad x \in \mathbb{R}, \tag{18}$$

in which case,

$$N = \frac{1}{t}\int_{\mathbb{R}} Q_t(x)\,\mathrm{d}x\,, \tag{19}$$

see Cruz-Orive (2006), who gives a chronological account on the evolution of the problem.

Consider a FUR test system Λ_1 of Cavalieri stripes of thickness t and period T, $0 < t \le T < \infty$, normal to the sampling axis. Let $\{N_1, N_2, \ldots, N_n\}$, $n \ge 3$, denote the successive numbers, in their natural order, of point particles captured by the n stripes encompassing Y, where N_1, N_n are the first, and the last, nonzero

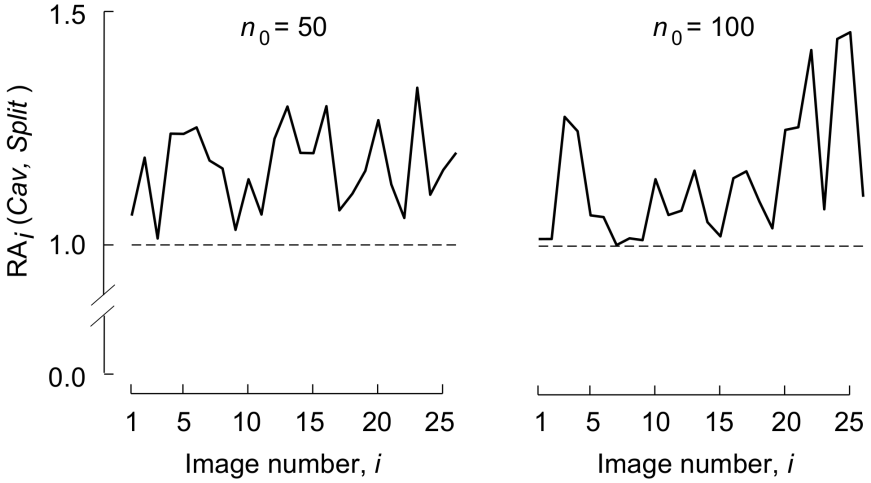

Fig. 5. *Relative accuracy ratio for each particle population, computed via Eq. 17. The variance predictors involved were C2 and S2.*

numbers, respectively. For each $i = 1, 2, ..., n$, let $\widehat{N}_i$ denote an UE of N_i. An UE of N is,

$$\widehat{N} = \frac{T}{t} \sum_{i=1}^{n} \widehat{N}_i \,. \tag{20}$$

A predictor of $\mathrm{Var}(\widehat{N})$ is,

$$\begin{aligned}
&\mathrm{var}(\widehat{N}) = \frac{c(\tau)}{\tau^2} \left[3(\widehat{C}_0 - \widehat{v}_n) - 4\widehat{C}_1 + \widehat{C}_2 \right] + \frac{\widehat{v}_n}{\tau^2} \,, \\
&\tau = t/T \,, \\
&c(\tau) = \frac{(1-\tau)^2}{6(2-\tau)} \,, \\
&\widehat{C}_k = \sum_{j=1}^{n-k} \widehat{N}_j \widehat{N}_{j+k} \,, \quad k = 0, 1, 2, \\
&\widehat{v}_n = \sum_{i=1}^{n} \mathrm{var}(\widehat{N}_i) \,,
\end{aligned} \tag{21}$$

which is Eq. 3.3 from Cruz-Orive (2006) with the smoothness constant $q = 0$, because the function f is expected to have finite jumps – thus, $c(\tau) \equiv \alpha(0, \tau)$, see the first Eq. 3.7 of the same paper. The predictor is based on a fractional model of the covariogram of f near the origin, see Eq. II.1 of that paper. The last Eq. 21 assumes that the within stripe deviations $\{\widehat{N}_i - N_i\}$ are mutually uncorrelated. Note that the symbol $\widehat{v}_n$ in 21 is different from $\widehat{v}_n$ in Eqs. 3 and 5.

Consider now an independent FUR test system Λ_2 of Cavalieri stripes with the same parameters $\{t, T\}$, perpendicular to Λ_1. Clearly $\Lambda_1 \cap \Lambda_2$ is the FUR test system of quadrats Λ_x defined by Eq. 1. It is now easy to see that

$$\widehat{N}_i = \tau^{-1} Q_i \,, \tag{22}$$

is an UE of N_i, where Q_i is as defined for Eq. 3. Therefore,

$$\widehat{C}_k = \tau^{-2} C_k \,, \tag{23}$$

where C_k is given by the third Eq. 3. Moreover,

$$\mathrm{var}(\widehat{N}_i) = \tau^{-2} \mathrm{var}(Q_i) \,. \tag{24}$$

Substitution of the preceding two results into the right hand side of the first Eq. 21, the predictor given by the first Eq. 3 is obtained.

An alternative approach to treat FUR Cavalieri stripes is the Split design. The first stage sample $\{N_1, N_2, ..., N_n\}$ generated by Λ_1 is split into two systematic subsamples consisting of the odd and even numbered particle counts, respectively. Let N_o, N_e denote the corresponding total point particle counts, and $\widehat{N}_o$, $\widehat{N}_e$ the respective unbiased estimators. The development follows Matheron's theory, as above, with smoothness constant $q = 0$, and the corresponding variance predictor is given by Eq. 2.27 from Cruz-Orive (2004), which was derived for a number $n \geq 2$ of subsamples (here 'n' is not to be confused with the same symbol used above for the number of stripes). Setting $n = 2$ in that equation we obtain,

$$\begin{aligned} &\mathrm{var}(\widehat{N}) = \frac{s(\tau)}{\tau^2}\left[(\widehat{N}_o - \widehat{N}_e)^2 - \widehat{\nu}_2)\right] + \frac{\widehat{\nu}_2}{\tau^2}\,, \\ &\tau = t/T\,, \\ &s(\tau) = \frac{(1-\tau)^2}{3-2\tau}\,, \\ &\widehat{\nu}_2 = \mathrm{var}(\widehat{N}_o) + \mathrm{var}(\widehat{N}_e)\,. \end{aligned} \tag{25}$$

When the perpendicular FUR stripes from Λ_2 intersect Λ_1 we have,

$$\widehat{N}_o = \tau^{-1} Q_o,\ \widehat{N}_e = \tau^{-1} Q_e\,, \tag{26}$$

whereby,

$$\widehat{\nu}_2 = \tau^{-2} \sum_{i=1}^{n} \mathrm{var}(\widehat{Q}_i) = \tau^{-2}\, \widehat{v}_n\,. \tag{27}$$

Substitution of the preceding two results into the right hand side of the first Eq. 25, yields Eq. 5.

The Split design is now used to predict each within stripe variance by means of the first Eq. 25 with $\widehat{\nu}_2 = 0$, because the quadrat counts Q_{oi} and Q_{ei} are directly observable. Thus,

$$\mathrm{var}(\widehat{Q}_i) = s(\tau)(Q_{oi} - Q_{ei})^2\,, \tag{28}$$

which, combined with Eq. 27, yields Eq. 4.

The same results apply when point particles are replaced with arbitrary particles, the stripes with stripe disectors, and the quadrats with unbiased frames, (this is the 'real mode' mentioned in the *Introduction*). For illustrations of both the Cavalieri and the Split designs see also Cruz-Orive and Geiser (2004). The combinations given by Eq. 3 and Eq. 5 for quadrats, however, were first proposed by Cruz *et al.* (2015), (without proof). The idea was to treat the between stripe contribution with the Cavalieri model, and the within stripe contribution with the Split model. The individual quadrat counts $\{q_{ij}\}$ will tend to be small (usually between 0 and 5) – hence it is sensible to add them up into two groups (consisting of the odd and even numbered quadrat counts), which is what the Split method does for the within stripe variance contribution.

Generalization. The generalization of Eq. 21 to higher dimensions is relatively straightforward. Consider for instance a series of FUR Cavalieri slabs in $\mathbb{R}^3$, and hit this series with a perpendicular FUR series. The result is a three dimensional grid of FUR bars whose cross section is a square of side length t. In Fig. 1b, the bars would be perpendicular to the plane of the paper, and they would be viewed as quadrats. The particle contents N_i of the ith slab would be unbiasedly estimated from the contents of the Cavalieri bars subsampled within it. Thus, up to this second stage the error variance would be predicted by the first Eq. 21, and $\widehat{\nu}_n$ would be predicted by the last Eq. 21. Now, however, $\widehat{N}_i$ would be an UE of N_i computed as T/t times the total contents of all the Cavalieri bars subsampling the ith slab. Consequently, $\mathrm{var}(\widehat{N}_i)$ could be predicted again by the first Eq. 21 with a proper readaptation of the symbols. Thus, the two stage predictor of $\mathrm{var}(\widehat{N})$ would be a nested version of two Cavalieri components, one accounting for the variation among slabs, and the other for the variation among bars within slabs. The third stage probe would be a FUR Cavalieri series of slabs perpendicular to the other two: in Fig. 1b, this last series would be parallel to the plane of the paper. The intersection between the three mutually orthogonal slab series would be a three dimensional cubic grid of cubic blocks of side length t, and the sampling volume period of the grid would be $(T/t)^3$. The final variance predictor of $\widehat{N}$ would be the aforementioned nested predictor with the two Cavalieri components, plus a third nested component accounting for the variation of blocks within bars, which could now be computed using the Split predictor formula. Thus, within each bar the third component would be analogous to Eq. 28 with Q_{oi}, Q_{ei} representing the total numbers of particles captured by the odd and the even numbered cubic blocks within the bar, respectively. The complete predictor could be computed along three mutually perpendicular directions, and the final predictor would be the corresponding average ($C3$, say).