Estimating the size of street-dwelling populations by means of mark-resighting counts: theoretical considerations and empirical results

L. Fattorini, G. Ghellini
Dipartimento di Metodi Quantitativi – Università di Siena (Italy)

Abstract

Mark-resighting constitutes a technologically advanced approach for estimating animal abundance; a sample of animals is captured and marked and resightings are performed during subsequent occasions. The joint hypergeometric maximum likelihood estimator, the Minta-Mangel estimator and the Bowden estimator are usually adopted with mark-resighting data. Recently, the basic assumptions regarding these widely applied procedures have been investigated. In presence of any tendency of animals to aggregate into groups, theoretical considerations and simulation results demonstrate that the Bowden estimator is the sole reliable method among those usually adopted, providing that marks are quite evenly distributed among groups. In some cetacean surveys, marking disturbances are avoided by means of natural marking, a procedure in which the marked units are the animals recognizable owing to their peculiar characteristics (e.g. flipper shapes). Under natural marking, resightings can be simply performed by means of individual recognitions. This paper proposes the joint use of natural marking and Bowden criterion to estimate the abundance of street-dwelling populations. In analogy with the procedure applied for cetacean populations, the marked individuals are the persons identified and recorded in the initial part of the survey and easily recognizable in the subsequent occasions. A simulation study is performed in order to determine the performance of the Bowden estimator in terms of bias, accuracy and coverage of confidence intervals under a wide set of situations attempting to take into account some features of street-dwelling populations. As expected, simulation results demonstrate the goodness of the proposed strategy when marked individuals are evenly distributed among groups.

Keywords: natural marking, individual recognition, Bowden estimator, simulation study

1. Introduction

In studying homeless populations, Iachan and Dennis (1991, 1993) point out the necessity of surveying all the population components in such disparate settings as shelters, service systems and streets. Shelter and service systems surveys are cheaper to perform but likely to ignore relevant portions of homeless populations. On the other hand, surveys that cover both shelters and street locations would reduce the potential bias due to undercoverage and limitations of shelters and service system surveys.

Street surveys are generally difficult to perform owing to the presence of people "*who are actively hiding to avoid both victimization and being run off by authorities*" (Iachan and Dennis, 1993, p.753). At least to our knowledge, the only attempt to overcome the difficulties in sampling street-dwelling populations, not necessarily unsheltered homeless people but also street prostitutes or groups of illegal immigrants, is due to Martin et al (1997). The authors propose the estimation of the population size by means of intrusive observers referred to as the *plants*, artificially introduced in the sites where people were thought to congregate but indistinguishable from the true members of the population. Recently, the method has been applied to correct the New York city's estimates of its homeless population (Hopper et al, 2008) from street undercounts. Unfortunately, the estimation criterion proposed by Martin et al (1997) suffers from some assumptions, such as the independence of sightings between individuals in the same site, which lead to a straightforward but unrealistic likelihood of the sample data.

The purpose of this paper is to propose an estimator of street-dwelling population sizes obtained by means of mark-resighting counts when the marked individuals are persons identified and recorded in the initial part of the survey and easily recognizable in the subsequent resighting occasions. Mark-resighting methods constitute technologically advanced approaches, originally applied for the estimation of animal abundance which may be adopted as alternative to capture-recapture methods. In mark-resighting experiments a sample of units is marked and resightings are performed during subsequent occasions. The main advantage of mark-resighting methods is that resightings are generally cheaper to acquire than physically capturing and handling the units; resightings are also less intrusive. As to the disadvantage of the method, the decrease of field effort is accompanied by a reduction of the collected information. Indeed, only the resighting histories of the marked units are available, so that the huge list of capture-recapture methodologies based on the capture histories of each captured unit cannot be adopted.

In order to estimate population abundance by mark-resighting counts, the joint hypergeometric maximum likelihood estimator, the Minta-Mangel estimator and the Bowden estimator are usually adopted. These methodologies are also implemented in NOREMARK, a software which has become increasingly popular among biologists (see White, 1996). Fattorini et al. (2007 a,b) discuss the basic assumptions regarding these widely applied procedures. In presence of any tendency of units to aggregate into groups, theoretical considerations and simulation results demonstrate that the Bowden estimator is the sole reliable method among those implemented in NOREMARK, providing that marks are quite evenly distributed among groups.

In some cetacean population surveys, the impact of the marking procedure is drastically reduced by means of an alternative protocol referred to as the *natural marking* which has the advantage to avoid captures. In these surveys, the marked animals are those readily recognizable owing to some peculiar characteristics (e.g. flipper shapes), in such a way that the resighting procedures can be simply performed by means of individual recognitions (see e.g. Hammond 1986, 1990, Wilson et al, 1999).

Natural marking seems to be particularly suitable for dealing with human populations, because it reduces the need for *formal contacts* with the units that very often refuse to be investigated. In particular, the mark phase could be made using some recognizing technique based on photos, films or some self-owned specific characteristics of the individuals (as done for cetacean populations), thereby reducing the ethical problems related to a proper mark phase for human beings. Besides, the resighting itself does not need any physical contact with the units, but only a well-structured technique that allows their effective detections.

On the basis of these previous considerations, the use of mark-resighting counts with naturally-marked individuals joined with the use of Bowden estimator seems to be a promising procedure for estimating the abundance of street-dwelling populations. In section 2, the statistical background necessary to the construction of the Bowden estimator is considered, while in section 3 a simulation study is performed in order to determine the performance of the proposed strategy in terms of bias, accuracy and coverage of confidence intervals, under a wide set of situations which attempts to take into account the main features of street-dwelling populations. Concluding remarks are performed in section 4.

2. Statistical background

Throughout the paper, N denotes the abundance of a closed population of street-dwelling individuals and constitutes the target parameter, M denotes the size of the initial portion, say $\mathcal{M}$, of individual recognized as marks and T denotes the number of resighting occasions performed in the experiment. Moreover, n denotes the overall number of resightings performed in T occasions, x_i denotes the total number of resightings for individual i in T occasions, while

$$\overline{x} = \frac{1}{M}\sum_{i\in\mathcal{M}} x_i$$

and

$$s_x^2 = \frac{1}{M-1}\sum_{i\in\mathcal{M}}(x_i - \overline{x})^2$$

respectively denote the mean and the variance of the x_i 's for $i \in \mathcal{M}$.

The Bowden estimator (BE), originally proposed by Bowden (1993) and subsequently investigated by Bowden and Kufeld (1995) is given by

$$\hat{N}_B = \left(\frac{n}{\overline{x}} + \frac{s_X^2}{\overline{x}^2}\right)\left(1 + \frac{s_X^2}{M\overline{x}^2}\right)^{-1} \tag{1}$$

In order to derive (1), Bowden (1993) supposes that

a) the total number of resightings for each unit constitutes a set of fixed values $x_1, \ldots, x_N$;

b) the M unit to be marked are selected from the population by means of simple random sampling without replacement (SRSWOR).

Under assumption *a*) the total number of resightings

$$n = \sum_{i=1}^{N} x_i$$

constitutes a finite population total which is exactly known after the T occasions. On the other hand, the finite population mean, say $\overline{X} = n / N$, and the finite population variance, say

$$S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{X})^2$$

are unknown parameters. Moreover, under assumption *b*), but also adopting other sampling schemes with inclusion probabilities equal to M / N (*e.g* systematic sampling or stratified sampling with proportional allocation), $\bar{x}$ constitutes a design-unbiased estimator of $\overline{X}$. Accordingly, in these cases $n / \bar{x}$ constitutes a very natural estimator for N.

Obviously, $n / \bar{x}$ is a biased estimator for N, since the expectation of $1 / \bar{x}$ does not equal $1 / \overline{X}$. On the basis of the Taylor series expansion of $n / \bar{x}$ around $\bar{x} = \overline{X}$ up to the second order, it follows that

$$\mathrm{E}\left(\frac{n}{\bar{x}} \mid x_1,\ldots,x_N\right) \cong N\left(1 + \frac{\mathrm{V}(\bar{x})}{\overline{X}^2}\right)$$

which under SRSWOR reduces to

$$\mathrm{E}\left(\frac{n}{\bar{x}} \mid x_1,\ldots,x_N\right) \cong N\left(1 + \frac{S_x^2}{M\overline{X}^2}\right) - \frac{S_x^2}{\overline{X}^2} \tag{2}$$

Accordingly, the BE of type (1) is a bias-reduced estimator of N when the marked units are selected by means of SRSWOR. It must be noticed that in this framework $\mathrm{E}(\cdot \mid x_1,\ldots,x_N)$ and $\mathrm{V}(\cdot \mid x_1,\ldots,x_N)$ denote expectation and variance performed with respect to SRSWOR of the units to be marked but conditional on the values of $x_1,\ldots,x_N$, which as usual in the design-based approach are assumed to be fixed values.

As to the basic assumption *a*), it seems far from being adequate in the framework of mark-resighting surveys of animals but also in the framework of mark-resighting surveys performed on street-dwelling populations. In order to justify the fact that x_i is assumed to be a fixed characteristic of unit *i*, Bowden & Kufeld (1995, p.843) point out that "*The sighting period and process should be predetermined, fixed and defined*" and that "*The area searched during the sighting period does not need to be the entire study area*". Probably, these two sentences have the purpose of excluding the use of any sampling design, such as encounter designs, which may be adopted to resight the animals after marking. Indeed, if resightings arise from a random search, as when animals are sighted from transects or observation points randomly thrown onto the study area, the x_i's necessarily constitute realizations of random variables and assumption *a*) has no statistical sense. However, even if no sampling plan is adopted and the paths to be travelled for observing animals or street-dwelling individuals are purposively selected, it is quite unrealistic to suppose that the number of resightings is a fixed characteristic of the units, such as their body weight, sex or age. Rather, since the x_i's depend on a plethora of situations and factors, they should be more realistically considered as a realization of the sequence of random variables $X_1,\ldots,X_N$.

As to assumption *b*), the use of SRSWOR for selecting the units to be marked is unrealistic not only for animals populations but also for street dwelling populations. In both cases, populations to be sampled are without-frame populations in such a way that there is no possibility of sampling units by means of SRSWOR, just like balls from an urn. Accordingly, the bias reduction performed in (1) is likely to be poorly effective when the units to be marked are selected from the population by schemes greatly differing from SRSWOR.

By using the Taylor series expansion of $n/\bar{x}$ up to the first leading terms, it follows that

$$\mathrm{V}\left(\frac{n}{\bar{x}} \mid x_1,\ldots,x_N\right) \cong N^2 \frac{\mathrm{V}(\bar{x})}{\bar{X}^2}$$

which under SRSWOR reduces to

$$\mathrm{V}\left(\frac{n}{\bar{x}} \mid x_1,\ldots,x_N\right) \cong \frac{NU}{M}\frac{S_x^2}{\bar{X}^2} \tag{3}$$

in such a way that a trivial estimator of (3) turns out to be

$$V^2 = \frac{\hat{N}_B(\hat{N}_B - M)}{M}\frac{s_x^2}{\bar{x}^2} \tag{4}$$

Thus, in accordance with (4), Bowden (1993) suggests the use of

$$V_B^2 = \left\{\frac{\hat{N}_B\left(\hat{N}_B - M\right)}{M}\frac{s_x^2}{\bar{x}^2}\right\}\left(1+\frac{s_x^2}{M\bar{x}^2}\right)^{-2}$$

as an estimator of the variance of $\hat{N}_B$.

Finally, since under SRSWOR, $\bar{x}$ converges to normality as N grows along with M (see *e.g.* Thompson 2002, p.31), the author proposes the use of

$$\hat{N}_B \pm t_{1-\alpha/2, M-1} V_B$$

as the confidence interval for N at the nominal level $1-\alpha$, where $t_{1-\alpha/2,M-1}$ is the $1-\alpha/2$ quantile of a *t*-distribution with $(M-1)$ degrees of freedom.

Interestingly, as already argued by Fattorini et al (2007a,b), assumption *a*) may be relaxed with no detrimental effects on estimation. Indeed, assumption *a*) may be interpreted more realistically as if the estimation were performed conditional on the resulting values of the random variables $X_1,\ldots,X_N$. Under SRSWOR (ensured by assumption *b*) and conditional to $x_1,\ldots,x_N$, the sample mean $\bar{x}$ constitutes an approximately unbiased estimator of $\bar{X}$. Then from the well known properties on conditional expectation and variance and from (2) and (3), it follows that

$$\mathrm{E}\left(\hat{N}_B\right) = \mathrm{E}_{x_1,\ldots,x_N}\left\{\mathrm{E}\left(\hat{N}_B \mid x_1,\ldots,x_N\right)\right\} \cong \mathrm{E}_{x_1,\ldots,x_N}\{N\} = N \tag{5}$$

and

$$\mathrm{V}\left(\hat{N}_B\right) = \mathrm{E}_{x_1,\ldots,x_N}\left\{\mathrm{V}\left(\hat{N}_B \mid x_1,\ldots,x_N\right)\right\} + \mathrm{V}_{x_1,\ldots,x_N}\left\{\mathrm{E}\left(\hat{N}_B \mid x_1,\ldots,x_N\right)\right\} \cong$$

$$\cong \mathrm{E}_{x_1,\ldots,x_N}\left\{\left(\frac{NU}{M}\frac{S_x^2}{\bar{X}^2}\right)\left(1+\frac{S_x^2}{M\bar{X}^2}\right)^{-2}\right\} \tag{6}$$

where now $\mathrm{E}(\cdot)$ and $\mathrm{V}(\cdot)$ denote expectation and variance performed with respect to both the SRSWOR of the animals to be marked and the joint distribution of $X_1,\ldots,X_N$. It is worth noting that the relations (5) and (6) hold whenever the joint distribution of $X_1,\ldots,X_N$ is.

Practically speaking, relations (5) and (6) ensure that $\hat{N}_B$ and V_B^2, being conditionally unbiased, turn out to be unbiased estimators even if $x_1,\ldots,x_N$ constitute realizations of random variables rather than fixed values. Accordingly, if the marks are evenly apportioned among the units, as should happen under SRSWOR, a good level of robustness should be expected for BE and the related confidence intervals.

3. Simulation study

The performance of BE was checked by means of a simulation study. The study was planned to take into consideration: *i*) the presence of sites where street- dwelling units congregated in groups of various sizes; *ii*) the number of units identified as marks in order to consider various levels of mark effort; *iii*) the distributions of marks among groups in order to generate more or less representative marking of units within population; *iv*) the movement of units among sites in order to generate more or less stable populations; *v*) the number of resighting occasions as well as the number of sites visited at each occasions in order to consider various levels of survey effort; *vi*) the missing of units due to their absence during a single resighting occasion or during the whole survey; *vii*) the missing of units due to the fact that enumerators were unable to find and counts all the units in a site.

To this purpose an artificial population of 1,000 street dwelling individuals was supposed to be spread out onto a city. A frame of 37 sites (city parks, area under bridges, bus and train stations and other locations) was supposed to be identified by previous investigations. The sites were labelled from 1 to 37 and randomly located onto a torus (see Figure 1) in such a way that the distance between sites j and k was quantified by means of the discrete metric

$$d(j,k) = \min\left(|j-k|, 37-|j-k|\right),\ j \neq k = 1,\ldots,37$$

Sites were assumed to have different sizes. At the start of the survey (i.e. before the first sighting occasion) the following sizes were assumed for the 37 sites: 1 site of size 200, 2 sites of size 100, 4 sites of size 50, 10 sites of size 20 and 20 sites of size 10. Since street dwelling populations are usually mobile, three levels of mobility were assumed: none (people did not move from their initial sites); medium (a 90% of people did not move from their initial sites, a remaining 10% were allowed to move to the two nearest sites or to rest on the same site with equal probability 1/3); high (a 60% of the people did not move from their initial sites, a 30% were allowed to move to the two nearest sites or rest on the same site with equal probability 1/3 and a remaining 10% were allowed to move to the six nearest sites or rest on the same site with equal probability 1/7). Previous investigation were also assumed to recognize easily identifiable units (marks) in each site. As to the mark distribution among sites, three situations were considered: in the first situation, marks were unevenly distributed among sites just identifying one mark per site; in the second situation a more even distribution of mark was assumed: since the exact sizes of sites were actually unknown to the enumerators, they were just assumed to be able to rank sites in accordance to their size; it was assumed that rank 5 was assigned to the site of size 200, rank 4 to the sites of size 100, rank 3 the sites of size 50, rank 2 to the sites of size 20 and rank 1 to the sites of size 10; thus one mark per rank score was identified in each site for a total of 65 marks; in the third situation an even distribution of mark was assumed by identifying one mark per ten individuals in each site for a total of 100 marks. As to the distribution of marks within each site, two situations were assumed: in the first situation marks were identified among all the units of a site, while the second situation tooks into consideration the fact that least mobile individuals had more chance to be identified (marked) than people frequently moving among sites; thus in the second situation marks were identified among those individuals which did not move among sites. As to the counting effort, various effort levels were considered by varying the number of sighting occasions and the number of sites visited at each occasion: 3 or 5 sighting occasions were assumed and at each occasion five sites randomly selected without replacement from the site frame were visited or, alternatively, all the sites were visited. Moreover, units were supposed to have a probability of 0.05 to be absent for the whole survey period while the remaining units were supposed to have a probability of 0.05 to be absent in any single occasion. Finally, owing to the difficulties to detect all the units present at a site, a detection probability exponentially decreasing with the site size was assumed. Thus, the probability of spotting a unit in a site containing s individuals was

$$p(s) = \exp\{-0.0025(s-1)\} \ , \ s \geq 1 .$$

decreasing from about 0.98 in sites of size 10 to about 0.60 in sites of size 200.

For each scenario resulting from combining mark distribution among sites, mobility of population, mark distribution within sites, number of sites visited at each occasion and number of occasions, 100,000 replications were performed. For each replication, the Bowden estimate of the population size $\hat{N}_B$ was computed together with the corresponding estimate of the sampling variance $\hat{V}_B^2$. In turn, the variance estimate gave rise to an estimate of the relative standard error say $R\hat{S}E = \hat{V}_B / \hat{N}_B$, as well as to the confidence interval at nominal level of 95%, say $\hat{N}_B \pm t_{0.975, M-1} \hat{V}_B$.

Subsequently the relative bias (RB) and the relative root mean squared error (RRMSE) of BE, together with the expectation of the relative standard error estimator (ERSEE) and the actual coverage of the 95% confidence intervals (COVER) were empirically determined from the Monte Carlo distributions arose from each scenario. The simulation results are reported in Table 1 to 3.

4. Discussion

The analysis of the tables motivates the following conclusions:

i) in the presence of an even distribution (Table 3) of marked individuals among sites, BE provides satisfactory performance with negligible RBs and with RRMSE's ranging from 3% to 13% in accordance with the decrease of the sampling effort; moreover $R\hat{S}E$ proves to be a conservative estimator of the actual RRMSE and the confidence interval show coverage very near, sometimes superior, to the nominal level of 95%; the performance of BE deteriorates as the mark become unevenly distributed among sites (Tables 1 and 2); these results about BE are akin to those achieved in a simulation study performed by Fattorini et al (2007 a,b) to comparing the performance of NOREMARK methods in estimating the size of animal populations;

ii) while the mark distribution among sites is proven to have an heavy impact on the properties of BE, the mobility of individuals among sites as well as the distribution of marks within sites seems to have a slight impact; moreover the number of occasions (3 *vs* 5) seems to have a small effects on the percentage of detected people and hence on the estimator performance; on the other hand, the sampling of sites drastically reduces the accuracy of BE; in accordance with these consideration a suitable survey design should perform few sighting occasions with all the sites visited at each occasion;

iii) interestingly, in presence of an even distribution of marks among sites, the absence of some units in some occasions or during the whole survey, does not cause underestimation.

Accordingly, providing that the marks are evenly distributed among sites, the strategy involving the use of Bowden criterion with mark-resighting data performed with natural marks constitutes a promising procedure for estimating street-dwelling population sizes, accomplishing robustness, computational simplicity and ethical propriety regarding both mark and resighting phases. It is worth noting that the strategy can be applied also when marks are plants artificially introduced in the street-dwelling population under study, as proposed by Martin et al (1997). In this case the only problem is to ensure that plants behave like the population units in all their patterns, absences included.

References

Bowden DC (1993) A simple technique for estimating population size. Department of Statistics, Colorado State University, Forth Collins.

Bowden DC, Kufeld RC (1995) Generalized mark-sigh population size estimation applied to Colorado moose. Journal of Wildlife Management, 59, 840-851.

Fattorini L., Marcheselli M., Monaco A., Pisani C. (2007a) Estimating population size by means of mark-resighting counts: theoretical considerations and empirical results. Working Paper n.70, Dipartimento di Metodi Quantitativi, Università di Siena, Siena (Italy).

Fattorini L, Marcheselli M, Monaco A, Pisani C (2007b) A critical look at some widely used estimators in mark-resighting experiments. Journal of Animal Ecology, 76, 957-965.

Hammond PS (1986) Estimating the size of naturally marked whale populations using capture-recapture techniques. Report of the International Whaling Commission. Special Issue, 8, 253-282.

Hammond PS (1990) Capturing whales on film – estimating cetacean population parameters from individuals recognition data. Mammal Review, 20, 17-22.

Hopper K, Shinn M, Laska E, Meisner M, Wanderling J (2008) Estimating number of unsheltered homeless people through plant-capture and postcount survey methods. American Journal of Public Health, 98, 1438-1442.

Iachan R, Dennis ML (1991) The design of homeless surveys. Proceedings of the Survey Research Method Section, American Statistical Association, 181-185.

Iachan R, Dennis ML (1993) A multiple frame approach to sampling the homeless and transient population. Journal of Official Statistics, 9, 747-764

Martin E, Laska E, Hopper K, Meisner M, Wanderling J (1997) Issues in the use of a plant-recapture method for estimating the size of the street dwelling population. Journal of Official Statistics, 13, 59-73.

Thompson SK (2002) Sampling, 2nd edition. New York, Wiley

White GC (1996) NOREMARK: population estimation from mark-resighting surveys. Wildlife Society Bulletin, 24, 50-52.
Wilson B, Hammond PS, Thompson PM (1999) Estimating size and assessing trend in a coastal bottlenose dolphin population. Ecological Applications, 9, 288-300.

TABLE 1. Performance of the Bowden estimator for an artificial street dwelling population partitioned into 37 sites in which marks are unevenly distributed among sites (one mark per site)

mobility	mark distribution	visited sites	occasions	RB	RRMSE	ERSEE	COVER
none	random	Five at occasion	3 (0.31)	-0.14	0.26	0.21	0.75
			5 (0.51)	-0.12	0.18	0.14	0.75
		All	3 (0.93)	-0.11	0.11	0.05	0.38
			5 (0.95)	-0.11	0.11	0.05	0.33
medium	random	five at occasion	3 (0.31)	-0.12	0.25	0.22	0.78
			5 (0.51)	-0.10	0.17	0.14	0.79
		All	3 (0.93)	-0.09	0.10	0.05	0.48
			5 (0.95)	-0.09	0.10	0.05	0.44
	random among stable individuals	five at occasion	3 (0.31)	-0.12	0.24	0.21	0.79
			5 (0.51)	-0.10	0.16	0.14	0.80
		All	3 (0.93)	-0.09	0.10	0.05	0.48
			5 (0.95)	-0.09	0.10	0.05	0.44
high	random	five at occasion	3 (0.32)	-0.06	0.25	0.23	0.85
			5 (0.51)	-0.05	0.17	0.16	0.87
		All	3 (0.94)	-0.05	0.07	0.05	0.73
			5 (0.95)	-0.05	0.06	0.04	0.71
	random among stable individuals	five at occasion	3 (0.31)	-0.08	0.18	0.22	0.91
			5 (0.51)	-0.06	0.13	0.14	0.91
		All	3 (0.94)	-0.05	0.07	0.05	0.73
			5 (0.95)	-0.05	0.06	0.05	0.71

(*) Values in brackets represent the percentage of people detected at least once

TABLE 2. Performance of the Bowden estimator for an artificial street dwelling population partitioned into 37 sites in which marks are distributed among sites proportionally to their ranks (one mark per rank score)

mobility	mark distribution	visited sites	occasions	RB	RRMSE	ERSEE	COVER
none	random	five at occasion	3 (0.31)	-0.09	0.18	0.17	0.83
			5 (0.51)	-0.08	0.13	0.11	0.83
		All	3 (0.93)	-0.07	0.08	0.04	0.57
			5 (0.95)	-0.07	0.07	0.04	0.51
medium	random	five at occasion	3 (0.31)	-0.08	0.18	0.17	0.86
			5 (0.51)	-0.06	0.12	0.11	0.86
		All	3 (0.93)	-0.06	0.07	0.04	0.65
			5 (0.95)	-0.06	0.06	0.04	0.60
	random among stable individuals	five at occasion	3 (0.31)	-0.08	0.16	0.17	0.89
			5 (0.51)	-0.06	0.11	0.11	0.88
		All	3 (0.93)	-0.05	0.06	0.04	0.68
			5 (0.95)	-0.05	0.06	0.04	0.64
high	random	five at occasion	3 (0.31)	-0.04	0.18	0.17	0.90
			5 (0.51)	-0.03	0.12	0.12	0.91
		All	3 (0.94)	-0.03	0.05	0.04	0.81
			5 (0.95)	-0.03	0.04	0.03	0.80
	random among stable individuals	five at occasion	3 (0.32)	-0.04	0.12	0.17	0.99
			5 (0.51)	-0.02	0.08	0.11	0.98
		All	3 (0.94)	-0.02	0.04	0.04	0.89
			5 (0.95)	-0.02	0.04	0.04	0.88

(*) Values in brackets represent the percentage of people detected at least once

TABLE 3. Performance of the Bowden estimator for an artificial street dwelling population partitioned into 37 sites in which marks are evenly distributed among sites (one mark per ten individuals)

mobility	mark distribution	visited sites	occasions	RB	RRMSE	ERSEE	COVER
none	random	five at occasion	3 (0.31)	-0.02	0.08	0.14	1.00
			5 (0.51)	-0.01	0.06	0.09	0.99
		All	3 (0.93)	0.00	0.03	0.04	0.96
			5 (0.95)	0.00	0.03	0.03	0.96
medium	random	five at occasion	3 (0.31)	-0.01	0.10	0.14	0.99
			5 (0.51)	0.00	0.07	0.09	0.98
		All	3 (0.93)	0.00	0.03	0.04	0.96
			5 (0.95)	0.00	0.03	0.03	0.96
	random among stable individuals	five at occasion	3 (0.31)	0.00	0.08	0.14	1.00
			5 (0.51)	0.01	0.06	0.09	0.99
		All	3 (0.93)	0.01	0.04	0.04	0.97
			5 (0.95)	0.01	0.03	0.03	0.97
high	random	five at occasion	3 (0.31)	0.00	0.13	0.14	0.96
			5 (0.51)	0.00	0.09	0.10	0.96
		All	3 (0.94)	0.00	0.03	0.03	0.95
			5 (0.95)	0.00	0.03	0.03	0.95
	random among stable individuals	five at occasion	3 (0.31)	0.04	0.14	0.14	0.97
			5 (0.51)	0.04	0.10	0.09	0.98
		All	3 (0.94)	0.03	0.05	0.04	0.93
			5 (0.95)	0.03	0.04	0.03	0.92

(*) Values in brackets represent the percentage of people detected at least once

Figure 1. Graphical representation of the 37 sites hosting the artificial population of 1,000 individuals and their initial sizes.