Journal of Applied Ecology 2008, **45**, 728–735 doi: 10.1111/j.1365-2664.2007.01406.x

A new method for estimating visitation rates of cryptic animals via repeated surveys of indirect signs

Bernd Gruber[1,3]*, Björn Reineking[2], Justin M. Calabrese[3], Andreas Kranz[4], Katerina Poledníková[4], Lukáš Poledník[4], Reinhard Klenke[1], Anett Valentin[1] and Klaus Henle[1]

[1]UFZ – Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstr. 15, D-04318 Leipzig, Germany; [2]Forest Ecology, Department of Environmental Sciences, ETH Zurich, Universitätsstr. 22, CH-8092 Zurich, Switzerland; [3]UFZ – Helmholtz Centre for Environmental Research, Department of Computational Landscape Ecology, Permoserstr. 15, D-04318 Leipzig, Germany; and [4]Institut für Wildbiologie und Jagdwirtschaft, Department of Integrative Biology, Universität für Bodenkultur (BOKU) Wien, Gregor Mendel-Str. 33, A –1180 Wien

---

Summary

1. Visitation rate is a measure of the frequency with which animals visit specific locations of interest. This information is crucial for various problems in the conservation and management of animal species, e.g. to determine the risk wildlife poses to human managed resources or the predation risks of nests. An important assessment tool, especially for cryptic animals, is to count indirect signs of presence, such as tracks or faeces.

2. Here we propose a maximum likelihood-based method that uses information on the age of tracks or signs, and we show that existing visitation rate/probability estimators are special cases of this more general approach.

3. Using simulated data, we compared the performance of the new visitation rate estimator to three other estimators, including the most commonly used estimator. These estimators make use of either fresh, aged or total signs, whereas our approach uses information simultaneously on both fresh and aged signs. The new estimator is, on average, in excess of three times more accurate than the next best estimator. Moreover, the new approach is very flexible and can be applied for sampling regimens with irregular time intervals between sampling.

4. We demonstrate the application of the method to field data by estimating the visitation rate of Eurasian otters (*Lutra lutra*) to a commercial fish pond. To facilitate the use of this method, we provide an easy-to-use Excel workbook and give recommendations on the most efficient sampling regimens.

5. *Synthesis and applications*. Visitation rate is an important quantity that can be estimated by repetitive sampling of indirect signs. We demonstrate the advantage of incorporating information explicitly on the age of signs over existing approaches. The new estimator can be applied to any species for which it is possible to discriminate between aged and new signs or tracks, and should be widely applicable in ecology and conservation biology.

Key-words: Eurasian otter, *Lutra lutra*, Mammalia: monitoring, maximum-likelihood estimate, track counts, visitation probability, visitation rate

Introduction

Visitation rate is a measure of the frequency with which animals visit specific locations or objects of interest. This information is important for many management and conservation issues and also for many ecological questions, such as human–wildlife conflicts, measuring risks of predation or disease transmission. For example, estimating visitation rates and their dependence on pond/landscape factors can be critically important to assessing the conflict between otters (*Lutra lutra*) and pond owners in commercial fish farms (Schwerdtner & Gruber 2007). Similarly, the attack rate of bears on sheep depends on feeding possibilities, vegetation features (Wilson *et al.* 2006) and individual type of bears. Further examples are

*Correspondence author. Bernd Gruber, Department of Computational Landscape Ecology, Permoserstr. 15, UFZ – Centre for Environmental Research, Leipzig Germany. E-mail: bernd.gruber@ufz.de

visitation rates of predators to nesting sites [e.g. lizards on turtle nests (Doody *et al.* 2006) or foxes and crows to ground breeding birds (Summers *et al.* 2004)].

Examples where the frequency of interaction and its distribution is more important than absolute numbers of individuals include visitation frequency of pollinators in different habitats (Tylianakis, Tscharntke & Klein 2006), the frequency of bites of malaria-transmitting mosquitoes (Ye *et al.* 2007), transmission of tuberculosis by badgers (*Meles meles*) (Tuyttens *et al.* 2001) and the risk of foxes (*Vulpes vulpes*) spreading rabies (Webbon, Baker & Harris 2004). In general, visitation rates convey important information whenever the exposed objects or areas have a scale at which the dispersion of individuals or their spatial use matters more than their absolute number.

Ideally, visitation rates are measured by direct observation, video surveillance (Binner, Henle & Hagenguth 1996), scent stations (Sargeant, Johnson & Berg 2003) or radio-telemetry of all individuals that might visit a location of interest. Unfortunately, these approaches are time-consuming, expensive and impossible for many systems, especially for cryptic animals. Due to these difficulties, cryptic animals are studied frequently with the use of indirect methods, such as counting tracks, faeces, hairs, dens or nests (hereafter ‘signs’). Data based on signs of animals are used to monitor the distribution and abundance (for a review see Wilson & Delahay 2001), presence/absence (Binner, Henle & Hagenguth 1996; Macdonald & Mason 1987) and the visitation probability to particular locations and their relationships to landscape factors.

Current methods to estimate visitation rates for cryptic animals are based on repeated sampling of the site(s) of interest. Usually, the number with positive records (presences) of a sign divided by the total number of sampling occasions at that location is taken as an estimate of the probability of visitation (Rowe-Rowe 1992; Madsen & Prang 2001; Klenke 2002; Rostain *et al.* 2004). Note that a similar approach is used to estimate abundance of animals, but that in this case the detection of signs is analysed across multiple sites (Caughley 1980; Wilson & Delahay 2001) instead of separately for each location, as in the case of estimating site-specific visitation rates. This visitation rate estimator has been used extensively in the literature (Tuyttens *et al.* 2001; Sadlier *et al.* 2004; Prokesova, Barancekova & Homolka 2006).

The method in current use does not discriminate between aged and new signs in the estimation of visitation rates/probabilities. This means that in surveys in which it is possible to make such a distinction, information that could potentially improve estimates is neglected. Here, we propose a new maximum-likelihood approach, which uses additional information on the age of signs to improve the precision and efficiency of estimates. We show that the current estimation technique is a special case of our more general framework and thus our work unifies the new method with existing techniques under the same mathematical umbrella. We derive the new estimator in the context of a particular field survey design, quantify the performance of the method using simulated data sets and, finally, apply the method to an empirical example to highlight its utility. To identify study designs that allow the efficient estimation of visitation rates/probabilities, we quantified the performance of the new estimator across a wide range of different sampling schemes using simulated data.

Methods

SURVEY DESIGN

The approach is based on the repeated sampling of the presence or absence of animal signs at locations of interest over a time interval (t days). Times between sampling occasions need not be of equal duration. The motivating example for this work was a data set of weekly collections ($t = 7$) of otter spraints at commercial fish farms. At the end of each visit, signs were recorded and then removed to ensure that only new signs would be found during the next visit. Observed signs were classified as fresh (from the previous night) and aged (from the previous $t - 1$ nights). For otters, aged spraints can be determined by their dried surface and solid consistency, whereas fresh spraints from the last day have a wet and soft consistency (Macdonald & Mason 1987). This can vary slightly, depending on weather conditions and the time of day of the survey. Experienced researchers are able to determine the age of spraints correctly but a pilot study, where spraints are sampled daily, is recommended. This classification scheme was appropriate for otter spraints, but could be adjusted easily to accommodate different situations.

This sampling design results in two types of information for each location: the number of sampling occasions where (a) fresh or (b) aged spraints were detected during the survey. The use of the second piece of information requires that the sampling is sufficiently frequent so that aged signs do not become obliterated between sampling occasions, but not so long that aged signs are always present. This imposes limits on the maximum and minimum time span between sampling periods, which will depend upon the species and the type of sign used. In the case of otters, spraints are obliterated typically after about 7 weeks for *L. lutra* (Jenkins & Harper 1980) and 4–5 weeks for *L. maculicollis* (Rowe-Rowe 1992). Generally speaking, the distinction between fresh and aged signs provides information on the visitation rate on two different (a shorter and a longer) time scales, therefore estimators that use this additional information should perform better than those based on a single (shorter or longer) time scale.

DERIVATION OF ESTIMATORS

In the following, we derive maximum likelihood estimators for visitation rate and visitation probability for our new method and for new and existing estimators. Table 1 provides a summary of the notation used. Note that visitation rate and visitation probability are different manifestations of the same matter. Visitation rate is calculated as the number of visits per unit time, whereas the visitation probability is the per-unit time probability that an animal visits a site.

We assume that the visitation process to a site is a Poisson process and that the visitation rate (number of visits/day) is constant for the entire duration of a study (assumptions will be discussed below). Under these assumptions, the probability that a site will receive h visits in some time interval of length t is given by the Poisson distribution:

$$\Pr(Z = h) = \frac{e^{-\lambda t}(\lambda t)^h}{h!}, \qquad \text{eqn 1}$$

Table 1. Description of variables and parameters used in the derivation of the visitation rate/probability estimators

f	Number of sampling occasions in which fresh signs/tracks were observed at a site
a	Number of sampling occasions in which old signs/tracks were observed at a site
c	Number of sampling occasions in which signs/tracks are found without distinction between old and new signs/tracks
n	Number of sampling occasions.
t	Interval in number of days between sampling occasions
p	True (simulated) daily visitation probability
$\hat{p}$	Estimator of daily visitation probability
λ	Visitation rate
$\hat{\lambda}$	Estimator of visitation rate
$\hat{\lambda}_{fr}$, $\hat{p}_{fr}$	Estimator of visitation rate/probability using fresh signs only
$\hat{\lambda}_{ag}$, $\hat{p}_{ag}$	Estimator of visitation rate/probability using aged signs only
$\hat{\lambda}_{cm}$, $\hat{p}_{ag}$	Estimator of visitation rate/probability using aged and fresh signs without distinction (used in the literature)
$\hat{\lambda}_{f+a}$, $\hat{p}_{f+a}$	Estimator of visitation rate/probability using information on aged and fresh signs

where λ is the per-day visitation rate we wish to estimate from the data. The probability of observing no visits (absence) after t days is:

$$\Pr(Z = 0) = e^{-\lambda t}, \qquad \text{eqn 2}$$

whereas the probability of observing at least one visit (presence) during t days is simply the complement of the above:

$$\Pr(Z > 0) = 1 - e^{-\lambda t}. \qquad \text{eqn 3}$$

When only fresh signs are taken into account, $t = 1$, and the probability of observing f presences and $n - f$ absences in n observations is given by the binomial distribution:

$$\Pr(Z = f) = \binom{n}{f} p_{fr}^{f}(1 - p_{fr})^{n-f}, \qquad \text{eqn 4}$$

where $p_{fr} = 1 - e^{-\lambda}$. As we are interested in the likelihood L, the binomial coefficient, which is a normalization constant, may be dropped, yielding:

$$L(\lambda \mid data_{fr}) = p_{fr}^{f}(1 - p_{fr})^{n-f}. \qquad \text{eqn 5}$$

The negative log likelihood is then:

$$-\ln(L(\lambda \mid data_{fr}) = \Im(data_{fr} \mid \lambda) = -(f \ln(p_{fr}) + (n - f)\ln(1 - p_{fr})). \qquad \text{eqn 6}$$

Taking the first derivative of the negative log likelihood with respect to λ, setting this equal to zero and solving for λ yields the maximum likelihood estimator for λ given only data on fresh signs:

$$\hat{\lambda}_{fr} = -\ln\left(\frac{n - f}{n}\right). \qquad \text{eqn 7}$$

Inserting equation 2 into equation 7, the maximum likelihood estimator for the visitation probability is:

$$\hat{p}_{fr} = \frac{f}{n}. \qquad \text{eqn 8}$$

Next, we assume that a distinction can be made between fresh and aged signs. When only aged signs are used, the time interval is $t - 1$ days. Denoting a as the number of presences and $n - a$ as the number of absences in n observations, the negative log likelihood can be derived in the same manner as above, giving:

$$\Im(\lambda \mid data_{ag}) = -(a \ln(p_{ag}) + (n - a)\ln(1 - p_{ag})), \qquad \text{eqn 10}$$

where $p_{ag} = 1 - e^{-\lambda(t-1)}$. Taking the first derivative with respect to λ, setting it equal to zero and solving for λ yields the maximum likelihood estimator given aged data only:

$$\hat{\lambda}_{ag} = -\frac{\ln\left(\frac{n - a}{n}\right)}{t - 1}. \qquad \text{eqn 11}$$

The estimator for the probability of visitation over $t - 1$ days is:

$$\hat{p}_{ag}(t - 1) = \frac{a}{n}. \qquad \text{eqn 12}$$

This can be converted into a daily visitation probability by noting that $1 - a/n$ is the probability that no visit occurred in $t - 1$ days. The $(t - 1)$th root of $1 - a/n$ is the per-day probability that the location was not visited. The complement of this quantity is the daily visitation probability:

$$\hat{p}_{ag}(t = 1) = 1 - \sqrt[t-1]{1 - \frac{a}{n}}. \qquad \text{eqn 13}$$

When no distinction between new and aged tracks is possible, or when all data are combined into one presence/absence data set, simply recording the number of presences, c, and absences in n observations yields:

$$\Im(\lambda \mid data_{cm}) = -(c \ln(p_{cm}) + (n - c)\ln(1 - p_{cm})), \qquad \text{eqn 14}$$

where $p_{cm} = 1 - e^{-\lambda t}$. Taking the first derivative with respect to λ, setting it equal to zero, and solving for λ yields the maximum likelihood estimator given combined data only:

$$\hat{\lambda}_{cm} = -\frac{\ln\left(\frac{n - c}{n}\right)}{t} \qquad \text{eqn 15}$$

and for the (t)-days-visitation probability:

$$\hat{p}_{cm}(t) = \frac{c}{n}, \qquad \text{eqn 16}$$

which can be converted into a daily visitation probability:

$$\hat{p}_{cm}(t = 1) = 1 - \sqrt[t]{1 - \frac{c}{n}}. \qquad \text{eqn 17}$$

This is the only estimator for visitation rates/probabilities used previously in the literature.

We now wish to incorporate the new and aged sign data together in the same framework. Because we have assumed a Poisson process, the presence or absence of signs in an interval of time does not affect the probability of signs being deposited in the next time interval. We may therefore multiply the likelihood for the fresh data and that for aged data together to yield the likelihood for the full data set:

$$L(\lambda \mid data_{f+a}) = p_{fr}^{f}(1-p_{fr})^{n-f}p_{ag}^{a}(1-p_{ag})^{n-a}. \qquad \text{eqn 18}$$

The negative log likelihood is then:

$$\Im(\lambda \mid data_{f+a}) = -(f\ln(p_{fr}) + (n-f)\ln(1-p_{fr}) + a\ln(p_{ag}) + (n-a)\ln(1-p_{ag})). \qquad \text{eqn 19}$$

In contrast to the estimators above, this maximum likelihood estimator cannot be solved for in general. Analytical solutions are available only for $t \le 5$ days, but can be extremely cumbersome even for these cases. However, for $t = 2$, it has a particularly simple form:

$$\hat{\lambda}_{f+a}(t=2) = -\ln\left(\frac{2n-f-a}{2n}\right). \qquad \text{eqn 20}$$

Other cases can be solved numerically.

In the above derivations, the time intervals among sampling occasions have been assumed to be constant. This assumption may be relaxed and different time intervals between sampling occasions may be used. Then the likelihoods for the sampling occasions of different time intervals may be multiplied. Upon taking the negative log of this likelihood, one obtains:

$$\Im(\lambda \mid data_{f+a}) = -(f\ln(p_{fr}) + (n-f)\ln(1-p_{fr}) + \sum_{k}(a_k\ln(p_{ag,k}) + (n_k - a_k)\ln(1-p_{ag,k}))) \qquad \text{eqn 21}$$

for k groups of sampling occasions with time intervals identical within, but different among, groups. Now, a_k and n_k represent the number of presences of aged signs and the number of occasions in sampling group k, respectively. The probability $p_{ag,k}$ is calculated by substituting t in equation 1 by t_k, the length of time of sampling group k. Note that the terms dealing with fresh signs are unaffected, because by definition the time interval for fresh signs is always 1 day. Therefore, f and n remain the total number of presences and the total number of sampling occasions, respectively, across all time intervals. All the above-derived estimators are different cases of the same general framework. They rely on the same basic assumptions and they differ only in the data they use.

Approximate confidence intervals can be constructed for all estimators based on the likelihood-ratio test statistic (Hilborn & Mangel 1997: 162). The –2 log-likelihood of the maximum likelihood parameter estimate has an approximate χ^2 distribution with nr degrees of freedom, where nr is the number of fitted parameters. For example, in the case of 1 degree of freedom, the upper and lower 95% (99%) confidence limits are those parameter values for which the corresponding –2 log-likelihood is larger by 3·84 (6·63) than the minimum –2 log-likelihood.

To make this method as accessible as possible, we provide an Exel workbook, 'DoubleTrack', which contains a worksheet-based user interface and embedded macros that implement the numerical procedures described above. The workbook and a detailed description of how to use it to estimate visitation rates, visitation probabilities and confidence intervals for both regular and irregular time intervals between sampling occasions may be found for download in the online appendix (see Supplementary material, Appendices S1, S2) and on the authors' website [http://www.ufz.de/index.php?en = 1902].

SIMULATION EXPERIMENT

We used simulations to compare the performance of the different estimators. A Poisson arrival process was simulated to generate data sets using weekly sampling occasions (constant time intervals $t = 7$). We varied the number of sampling occasions (n) between 2 and 42 (21 levels), and the daily visitation probability (P) between 0·05 and 0·95 (19 levels). Although the Poisson process is defined in terms of a rate, we used visitation probabilities in the following examples to maintain consistency with the literature (Marques *et al.* 1987; Tuyttens *et al.* 2001; Webbon, Baker & Harris 2004; Prokesova, Barancekova & Homolka 2006). The visitation rate can be calculated directly from the visitation probability using equation 2. All estimators showed an almost exact fit when more than 42 sampling occasions were used; hence we restrict our analyses to a maximum of 42 weeks. The trivial cases of $P = 0$ and $P = 1$ are exact for all estimators and were also omitted.

For each of the 21 × 19 = 399 scenarios (combinations of parameter values), we simulated $r = 100$ repetitions and calculated relative bias, relative precision and relative accuracy for each estimator (Hellmann & Fowler 1999). We used relative performance metrics to be able to combine all runs in a meaningful way because, for example, absolute differences of 0·05 are much more severe when the true visitation probability is $P = 0{\cdot}1$ compared to when the true probability is $P = 0{\cdot}9$.

We calculated relative bias for each scenario as:

$$bias(\hat{p}) = \frac{\sum_{i=1}^{r}\frac{\hat{p}_i - p}{p}}{r}, \qquad \text{eqn 22}$$

where P is the true visitation probability and $\hat{p}$ is the estimated visitation probability in the ith simulation. The relative bias can be interpreted as percentage away from the true value, i.e. a relative bias of +0·1 means that, on average, the estimator is 10% higher than the true value. A perfect estimator has a relative bias of zero and the distribution of relative biases over all 399 runs should be symmetrical around zero.

Relative precision is measured by the relative variance:

$$\mathrm{var}(\hat{p}) = \left[\frac{\sum_{i=1}^{r}\left(\frac{\hat{p}_i - p}{p}\right)^2 - \left(\sum_{i=1}^{r}\left(\frac{\hat{p}_i - p}{p}\right)\right)^2}{r}\right]\Bigg/(r-1). \qquad \text{eqn 23}$$

Relative precision is a measure of how much variation in estimates exists around the true parameter value. Note that precision is expressed in a somewhat counterintuitive way, such that a value of zero indicates the highest precision. We combined relative bias and relative precision into a single quantity, relative accuracy (Hellmann & Fowler 1999), which is expressed as a mean square error (MSE):

$$accuracy = MSE(\hat{p}) = \mathrm{var}(\hat{p}) + (bias(\hat{p}))^2. \qquad \text{eqn 24}$$

A perfect estimator would result in a relative accuracy value of zero. The closer the relative accuracy of an estimator is to zero across the whole parameter space, the better its overall performance.

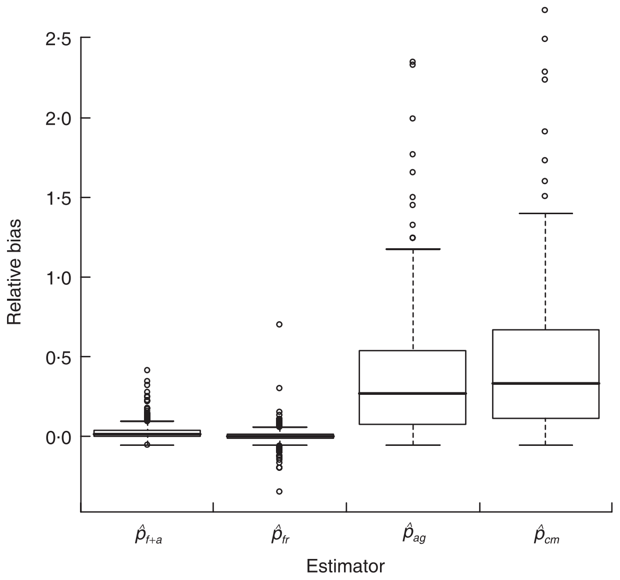

Fig. 1. Boxplot of relative bias for all estimators based on all 21 × 19 = 399 scenarios. Boxes indicate quantiles, whiskers indicate the 1·5 times the interquartile distance and the dots represent outliers larger than 1·5 times the interquartile distance.

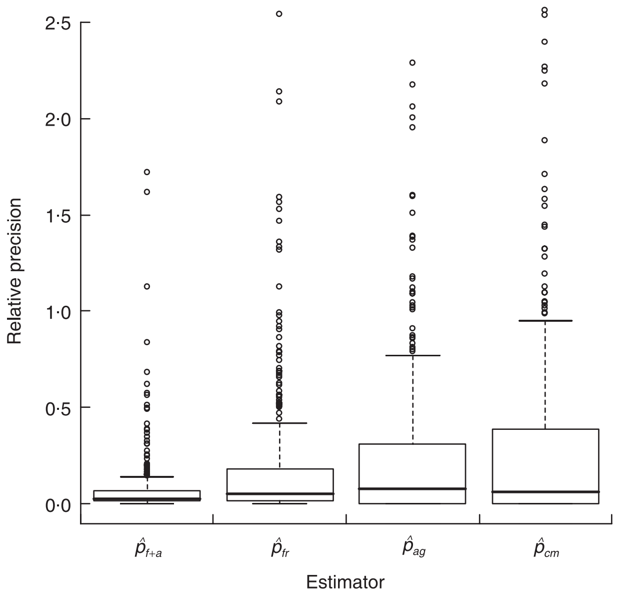

Fig. 2. Boxplot of relative precision for all four estimators based on all 399 scenarios. Boxes indicate quantiles, whiskers indicate 1·5 times the interquartile distance and the dots represent outliers larger than 1·5 times the interquartile distance.

Results

The relative bias was lowest for $\hat{p}_{f+a}$ and $\hat{p}_{fr}$ (Fig. 1). $\hat{p}_{ag}$ and $\hat{p}_{cm}$ showed considerable positive bias, which would result in overestimation of the visitation probability. $\hat{p}_{fr}$ was the only unbiased estimator, whereas $\hat{p}_{f+a}$ is biased only to a slight degree. The outliers of relative bias, where estimates are bigger than 0·1 (hence a relative bias of more than 10%) for $\hat{p}_{f+a}$ and $\hat{p}_{fr}$, are due to cases with very low sampling effort ($n \leq 4$) and/or very low visitation probability ($P \leq 0{\cdot}1$).

Precision is a measurement of the magnitude of the variation in an estimate. In principle, an estimator can be completely unbiased but, due to its high variation, still be very imprecise. The relative precision of $\hat{p}_{f+a}$ is lowest of all the estimators (Fig. 2). All other estimators show a considerable variance across the parameter space. Again, $\hat{p}_{ag}$, the estimator that uses only aged spraints, and $\hat{p}_{cm}$, the estimator that does not distinguish between aged and new spraints, score last. The least biased estimator $\hat{p}_{fr}$ has a relatively high variance and is less precise than $\hat{p}_{f+a}$. The outlying values of relative precision for $\hat{p}_{f+a}$, where estimates are higher than 0·1, are due to low sampling effort ($n < 6$) and low visitation probabilities ($P < 0{\cdot}15$), whereas the variance of $\hat{p}_{fr}$ remains high ($> 0{\cdot}2$) even for higher sampling efforts when visitation probabilities are fairly low ($P < 0{\cdot}2$).

The mean values for relative bias, precision and accuracy are shown in Table 2. Overall relative bias of $\hat{p}_{f+a}$ and $\hat{p}_{fr}$ are low (3% and 0·15%, respectively), in contrast to $\hat{p}_{ag}$ and $\hat{p}_{cm}$ (36% and 44%, respectively). Overall relative precision is lowest for $\hat{p}_{f+a}$, followed by $\hat{p}_{fr}$, $\hat{p}_{ag}$ and $\hat{p}_{cm}$, and the same ranking results when relative accuracy is considered. In summary, $\hat{p}_{f+a}$ performs best (three times more accurate than the next best estimator $\hat{p}_{fr}$) and is only slightly more biased than $\hat{p}_{fr}$.

Table 2. Summary statistics for all estimators. Mean relative bias, mean relative precision, and mean relative accuracy over all 399 scenarios

Estimator	$\hat{p}_{f+a}$	$\hat{p}_{fr}$	$\hat{p}_{ag}$	$\hat{p}_{cm}$
Bias	0·0300	0·0015	0·3606	0·4356
Precision	0·0731	0·2465	0·3551	0·4609
Accuracy	0·0768	0·2501	0·6062	0·8153

In addition to average performance, it is important to know how an estimator performs across the entire parameter space (Fig. 3). A perfect estimator would exhibit a perfectly flat plane with all values of relative accuracy close to zero. $\hat{p}_{f+a}$ shows the best performance and is inaccurate only for either low sampling effort or low visitation probabilities (Fig. 3a). It also shows the preferable characteristic that its relative accuracy approaches zero at moderate sampling effort. $\hat{p}_{f+a}$ shows the flattest plane and is the most accurate across all scenarios. In contrast, $\hat{p}_{ag}$ and $\hat{p}_{cm}$ do not have values of zero at even high sampling efforts, which reflects the consistent lack of performance of these estimators.

To assist the study design we further simulated the effect of different time intervals (t) on the accuracy of the estimators (Table 3). This table can be used to design studies, if the visitation probability can be estimated roughly from previous studies or by a pilot study. In accordance with intuition, the optimal sampling interval is longer for low values of P. Here, more time is required to obtain a large enough number of presences to yield precise estimates. If P is higher than about 0·8, one must sample every other day to extract the maximum amount of information from fresh and aged samples (otherwise all aged samples are likely to be presences). Whereas the

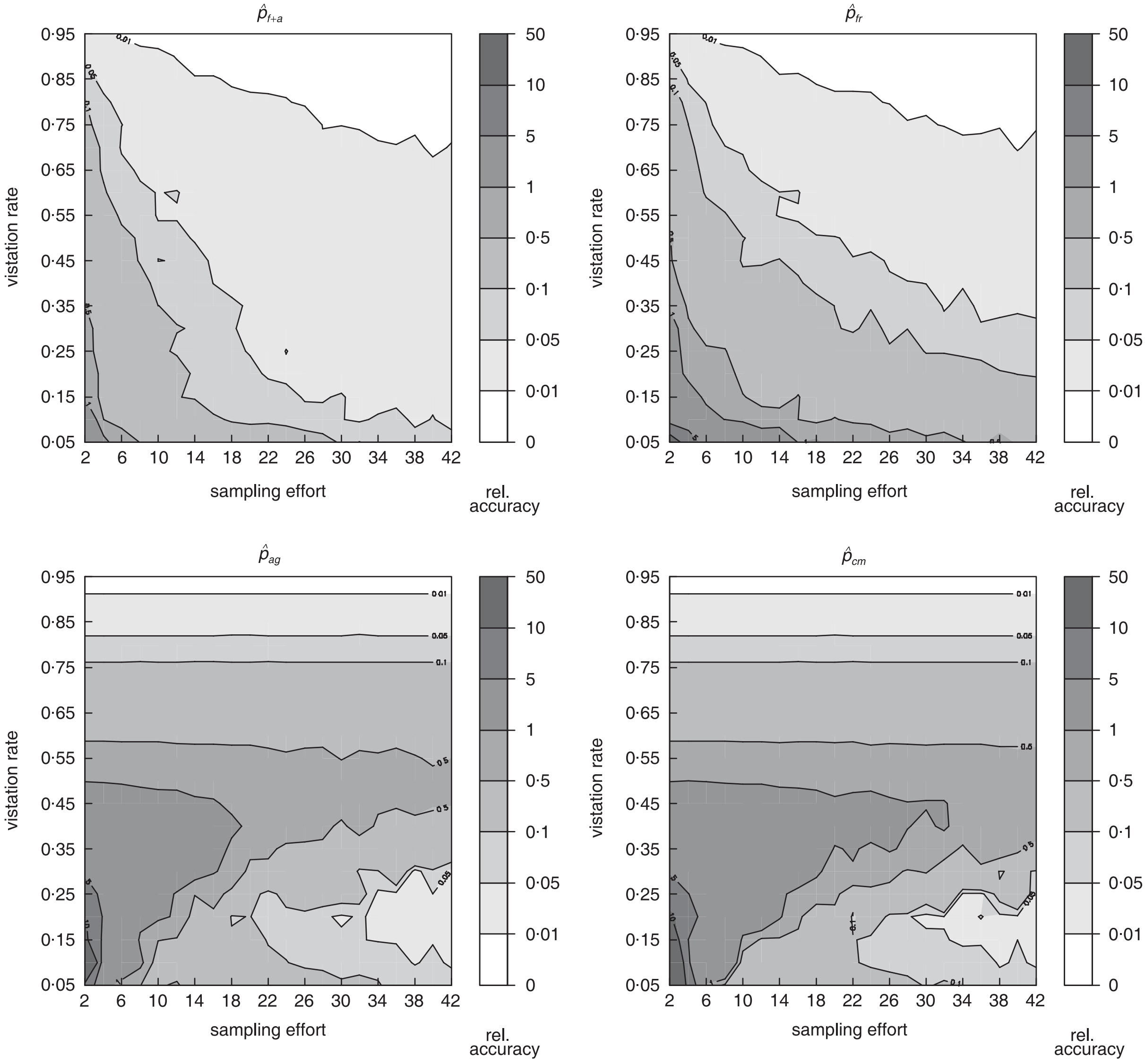

Fig. 3. Relative accuracy of all four estimators over the complete simulated parameter space. An optimal estimator has a flat plane of zero across the whole parameter space.

number of sampling occasions increases the precision of estimates, it has only a minor influence on the optimal time interval. At very low values of P (< 0·1), the optimal interval is shortened, when sampling effort is low.

EMPIRICAL EXAMPLE

We applied the new method to an empirical example. The data set consists of observations of otter spraints at a pond in Upper Lusatia, Germany. The survey was carried out for 12 weeks ($n = 12$) and the time interval between sampling occasions t was 7 days ($t = 7$). The complete data set is provided in Table 4. We used 'DoubleTrack' (see Supplementary material, Appendix S2) to calculate the daily visitation rate, the daily visitation probability and the 95% confidence intervals on these estimates (Table 5).

The results show more precise estimates (smaller confidence intervals) when aged and new spraints are used.

Table 3. Optimal time interval of surveys for different visitation probabilities

P	Optimal interval (time units)
< 0·05	=30
0·05–0·15	10–20
0·15–0·25	5–10
0·25–0·55	4–5
0·55–0·8	3
> 0·80	2

Another way to demonstrate the gain in precision is to determine how many more sampling occasions are needed to achieve the same level of precision. When using the estimator for fresh and aged signs, we achieve a similar precision after 10 sampling occasions compared to the precision achieved by the other estimators after 12 sampling occasions.

Table 4. Example data set on visits of the European otter (*Lutra lutra*) to a fish pond. The study lasted 12 weeks ($n = 12$), with weekly sampling occasions ($t = 7$) for fresh and aged spraints. 0: absence, 1: presence

Survey no.	Fresh	Aged	Time interval
1	0	1	7
2	0	0	7
3	0	1	7
4	1	1	7
5	0	1	7
6	1	1	7
7	0	1	7
8	0	1	7
9	1	1	7
10	0	0	7
11	0	1	7
12	0	1	7

Table 5. Results are shown for the example on the European otter (*Lutra lutra*). For each method, the point estimates for visitation probability and visitation rate and the associated confidence intervals (CI) are presented

Method	Visitation probability (95% CI)	Visitation rate (95% CI)
fr	0·250 (0·069–0·528)	0·288 (0·071–0·750)
ag	0·258 (0·131–0·443)	0·299 (0·140–0·585)
cm	0·258 (0·131–0·443)	0·299 (0·142–0·585)
f + a	0·256 (0·144–0·405)	0·295 (0·156–0·520)

Discussion

Overall, the new maximum likelihood estimator that combines information from new and aged spraints performed best across the entire examined parameter space. In the case of weekly sampling occasions, it provides precise, almost unbiased estimates. It fails only in cases of low sampling effort ($n < 4$) and low visitation probabilities ($P < 0{\cdot}1$). In such cases, none of the other estimators perform better. The estimator based on fresh tracks provides accurate estimates only when the survey lasts at least 8 weeks and the visitation probabilities are larger than 0·3. When this is the case it is possible to use this estimator, although one should be cautious in doing so as relying solely on fresh signs can lead to imprecise estimates when the visitation probability is low. When the visitation probability is low, it is possible that no fresh spraints will be observed leading to a $\hat{p}_{fr}$ estimate of zero, thus underestimating the true visitation probability.

The use of aged signs alone does not provide accurate estimates unless the visitation probabilities are either very low (< 0·25) or very high (> 0·8). However, crude estimates may still be possible for other visitation probabilities, provided the survey lasts at least approximately 30 weeks. The estimator $\hat{p}_{cm}$ showed the worst performance as it is both heavily biased and very imprecise, which is alarming as it is the estimator used in the literature. It performs well only with a very high sampling effort (> 30 sampling occasions) or if visitation probability is very high (> 0·8).

It is intuitively obvious that an estimator that uses both kinds of information should produce superior estimates. Our analyses support this intuition, showing that the new estimator performs robustly across the widest range of parameter space. As fresh signs give unbiased information on the last day only, this source of information is superior in cases of high daily visitation probabilities, whereas aged signs accumulate information over a longer time period and are therefore beneficial if visitation probabilities are low. The length of the sampling interval is also important if visitation rates are extreme (either very low or very high). In the case of low visitation rates, fresh signs alone contain little information and lead to highly imprecise results if the number of sampling occasions is low. In such cases, the information gained from aged signs is better suited to estimate visitation probability. In contrast, if the true visitation probability is extremely high (around 0·8) and the sampling interval fairly long ($t > 5$), aged spraints will be found in almost all sampling occasions, hence the visitation rate based on $\hat{p}_{ag}$ is nearly always 1. Here, information based on fresh spraints is crucial for a more precise estimate. Thus, if possible, any effort should be made to distinguish between aged and new tracks/signs and to use this information with the combined maximum likelihood estimator. Although care should be taken to classify the age of signs as accurately as possible, our simulations demonstrate that the new estimator performs well even if the classification is uncertain, as long as it is not systematically biased.

The independent increments assumption of the Poisson arrival process we have modelled (i.e. the presence or absence of visits in one time interval does not depend upon the presence or absence of visits in another, disjoint time interval) is a relatively strong assumption. Current methods – both for visitation rates and for abundance – also rely on this assumption, although this is often not stated explicitly. Note, however, that this assumption does not require that each individual animal moves randomly among the site of interest – a much stronger assumption. Rather, the aggregate pattern of arrivals of all animals of the focal species to a site of interest over time must resemble a Poisson process. This can happen for a variety of different, non-random individual behaviours if individuals act independently of one another and the total number of individuals is not too small. It is beyond the scope of the present paper to explore the range of individual behaviours and population sizes that would lead to arrivals at a site approximating a Poisson process. If a strong violation of the independent increments assumption is suspected from the data (e.g. all of the presences in a series of surveys are clustered together in time), another arrival process model could be used in place of the Poisson. In the case of the Eurasian otter, visitation rate estimates obtained using our approach and an approach based on radio-tracking data (Polednik 2005) were very similar, suggesting that the Poisson processes and its independent increments assumption is reasonable. Nevertheless, in some cases it may be preferable to use the estimators only as an ordering index when comparing the visitation rate among sites, if strong violations of the assumption of independence are expected. In conclusion, although these methods may be inaccurate (Stander 1998) and/or rely on assumptions that are sometimes difficult to meet (Sargeant, Johnson & Berg 1998; Stephens *et al*. 2006), they are often the best that can be achieved and are still suitable as relative indices (Caughley 1980). The method can, in principle, be applied to any species that leaves traces of its presence which can be dated to a certain time span. For example, it may be possible to date the remains of eaten fruits, leaves or animal carcasses or burrows as new or old.

We can think of several ways to extend the method. First, as mentioned previously, it is possible to derive estimators for sampling regimens with irregular time intervals. In principle, it is also possible to use more than two different age classes if objective criteria are available for such a classification of tracks. This should further improve estimates of visitation rates and probabilities. Also, an accumulation period of > 1 day for fresh signs may be more appropriate in some specific systems. Finally, it is also possible to assign the signs at each occasion to a number of individuals (as opposed to presence/absence) when such an assignment can be made reliably and then multiply the visitation rate by the average number of individuals (similar to line-transect methods) to obtain the average number of animals visiting a site (please note this extension implies a further assumptions on the independence of visits of groups of animals). In summary, we believe our approach has the potential to improve considerably the precision of visitation rate/probability estimates across a wide range of situations in which cryptic animals must be monitored.

References

Binner, U., Henle, K. & Hagenguth, A. (1996) Raumnutzung und Dismigration des Fischotters (Habitat use and dispersal of the European Otter). *Schriftenreihe Des Landesamts für Umwelt und Natur Mecklenburg-Vorpommern*, **1**, 43–47.

Caughley, G. (1980) *Analysis of Vertebrate Populations*. John Wiley, Chichester/New York/Brisbane.

Doody, J.S., Green, B., Sims, R., Rhind, D., West, P. & Steer, D. (2006) Indirect impacts of invasive cane toads (*Bufo marinus*) on nest predation in pig-nosed turtles (*Carettochelys insculpta*). *Wildlife Research*, **33**, 349–354.

Hellmann, J.J. & Fowler, G.W. (1999) Bias, precision, and accuracy of four measures of species richness. *Ecological Applications*, **9**, 824–834.

Hilborn, R. & Mangel, M. (1997) *The Ecological Detective: Confronting Models with Data*. Princeton University Press, Princeton, NJ.

Jenkins, D. & Harper, R.J. (1980) Ecology of otters in Northern Scotland II. Analyses of otter (*Lutra lutra*) and Mink (*Mustela vison*) faeces from Deeside, N.E. Scotland in 1977–78. *Journal of Animal Ecology*, **49**, 737–754.

Klenke, R. (2002) Habitat suitability and apparent density of the Eurasian otter (*Lutra lutra*) in Saxony. *Otter Conservation – an Example for a Sustainable Use of Wetlands Proceedings Viith International Otter Colloquium* (eds R. Dulfer, J.H. Conroy, J. Nel & A.C. Gutleb), vol. IUCN OSG Bulletin 19A, pp. 167–171. IUCN, Trebon.

Macdonald, S.M. & Mason, C.F. (1987) Seasonal marking in an otter population. *Acta Theriologica*, **32**, 449–462.

Madsen, A.B. & Prang, A. (2001) Habitat factors and the presence or absence of otters *Lutra lutra* in Denmark. *Acta Theriologica*, **46**, 171–179.

Marques, F.F.C., Buckland, S.T., Goffin, D., Dixon, C.E., Borchers, D.L., Mayle, B.A., &. Mason, C.F. & Macdonald, S.M. (1987) The use of spraints for surveying otter *Lutra lutra* populations – an evaluation. *Biological Conservation*, **41**, 167–177.

Mason, C.F. & Macdonald, S.M. (1987) The use of spraints for surveying otter *Lutra lutra* populations – An evaluation. *Biological Conservation*, **41**, 167–77.

Polednik, L. (2005) Otters (*Lutra lutra* L.) and fishponds in the Czech Republic: interactions and consequences. PhD thesis, Palacky University, Olomouc.

Prokesova, J., Barancekova, M. & Homolka, M. (2006) Density of red and roe deer and their distribution in relation to different habitat characteristics in a floodplain forest. *Folia Zoologica*, **55**, 1–14.

Rostain, R.R., Ben-David, M., Groves, P. & Randall, J.A. (2004) Why do river otters scent-mark? An experimental test of several hypotheses. *Animal Behaviour*, **68**, 703–711.

Rowe-Rowe, D.T. (1992) Survey of South African otters in a freshwater habitat, using sign. *South African Journal of Wildlife Research*, **22**, 49–55.

Sadlier, L.M.J., Webbon, C.C., Baker, P.J. & Harris, S. (2004) Methods of monitoring red foxes *Vulpes vulpes* and badgers *Meles meles*: are field signs the answer? *Mammal Review*, **34**, 75–98.

Sargeant, G.A., Johnson, D.H. & Berg, W.E. (1998) Interpreting carnivore scent-station surveys. *Journal of Wildlife Management*, **62**, 1235–1245.

Sargeant, G.A., Johnson, D.H. & Berg, W.E. (2003) Sampling designs for carnivore scent-station surveys. *Journal of Wildlife Management*, **67**, 289–298.

Schwerdtner, K. & Gruber, B. (2007) A conceptual framework for damage compensation schemes. *Biological Conservation*, **134**, 354.

Stander, P.E. (1998) Spoor counts as indices of large carnivore populations: the relationship between spoor frequency, sampling effort and true density. *Journal of Applied Ecology*, **35**, 378–385.

Stephens, P.A., Zaumyslova, O.Y., Miquelle, D.G., Myslenkov, A.I. & Hayward, G.D. (2006) Estimating population density from indirect sign: track counts and the Formozov–Malyshev–Pereleshin formula. *Animal Conservation*, **9**, 339–348.

Summers, R.W., Green, R.E., Proctor, R., Dugan, D., Lambie, D., Moncrieff, R., Moss, R. & Baines, D. (2004) An experimental study of the effects of predation on the breeding productivity of capercaillie and black grouse. *Journal of Applied Ecology*, **41**, 513–525.

Tuyttens, F.A.M., Long, B., Fawcett, T., Skinner, A., Brown, J.A., Cheeseman, C.L., Roddam, A.W. & MacDonald, D.W. (2001) Estimating group size and population density of Eurasian badgers *Meles meles* by quantifying latrine use. *Journal of Applied Ecology*, **38**, 1114–1121.

Tylianakis, J.M., Tscharntke, T. & Klein, A.M. (2006) Diversity, ecosystem function, and stability of parasitoid host interactions across a tropical habitat gradient. *Ecology*, **87**, 3047–3057.

Webbon, C.C., Baker, P.J. & Harris, S. (2004) Faecal density counts for monitoring changes in red fox numbers in rural Britain. *Journal of Applied Ecology*, **41**, 768–779.

Wilson, G.J. & Delahay, R.J. (2001) A review of methods to estimate the abundance of terrestrial carnivores using field signs and observation. *Wildlife Research*, **28**, 151–164.

Wilson, S.M., Madel, M.J., Mattson, D.J., Graham, J.M. & Merrill, T. (2006) Landscape conditions predisposing grizzly bears to conflicts on private agricultural lands in the western USA. *Biological Conservation*, **130**, 47–59.

Ye, Y., Kyobutungi, C., Louis, V.R. & Sauerborn, R. (2007) Micro-epidemiology of *Plasmodium falciparum* malaria: is there any difference in transmission risk between neighbouring villages? *Malaria Journal*, **6**. DOI: 10.1186/1475-2875-6-46.

Received 8 March 2007; accepted 14 August 2007
Handling Editor: Philip Stephens

Supplementary material

The following supplementary material is available for this article.

Appendix S1. DoubleTrack manual

Appendix S2. DoubleTrack Exel sheet

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/full/10.1111/j.1365-2664.2007.01406.x
(This link will take you to the article abstract.)