Turlure, C., **Pe'er, G.**, Baguette, M., Schtickzelle, N. (2018):
A simplified mark-release-recapture protocol to improve the cost effectiveness of repeated population size quantification
Methods Ecol. Evol. **9** (3), 645 – 656

http://dx.doi.org/10.1111/2041-210X.12900

Article Type: Research Article
Handling Editors : Robert Freckleton

A simplified Mark Release Recapture protocol to improve the cost-effectiveness of repeated population size quantification

Camille Turlure[1,2], Guy Pe'er[3,4], Michel Baguette[2,5] & Nicolas Schtickzelle[1]

[1] Université catholique de Louvain, Earth and Life Institute, Biodiversity Research Centre, Place Croix du Sud 4, L7.07.04, 1348 Louvain-la-Neuve, Belgium

[2] Centre National de la Recherche Scientifique and Université Paul Sabatier Toulouse III, SETE Station d'Ecologie Théorique et Expérimentale, UMR 5321, Moulis, France

[3] UFZ - Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstr. 15, 04318 Leipzig, Germany

[4] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany

[5] Muséum National d'Histoire Naturelle (Sorbonne Universités), UMR 7205 Institut de Systématique, Evolution et Biodiversité, 75005 Paris, France

Corresponding author: Camille Turlure, camille.turlure@uclouvain.be
Running title: A protocol for optimizing repeated MRR efforts

ABSTRACT

- Obtaining an accurate quantification of population size is often of prime importance in ecology and conservation biology (e.g. population viability analysis, a basic step for assessing species and population status in a given area and guiding effective conservation). When obtaining a reliable quantification of absolute (vs relative)

population size is required, Mark Release Recapture (MRR) is a method of choice for many organisms. This is a highly reliable but costly procedure in terms of time and potential impact on species and sites. Consequently, less costly alternatives are highly desirable for conservation and population ecologists.

- We present here a simplified MRR protocol to mitigate this cost of repeated MRR sampling with little compromise on the quality of the population size estimation. Using one of the largest existing butterfly MRR databases, collected on two fritillary species over a period of 20 years and >20 populations in Belgium, we assessed the possibility to reduce the effort of collecting MRR data while keeping accurate quantification of total population size. By downsampling from the full datasets and calculating a range of demographic census metrics, we specifically investigated whether marking individuals is necessary, and whether the number of sampling sessions can be reduced.
- We found that (1) counting individuals is not enough: some individual marking, even in a simplistic way to differentiate newly recorded from previously seen individuals, is essential for estimating population size. (2) A simple linear conversion function (number of "missed" individuals for each marked one) can be used to compute population size from the number of individuals marked over a small number of MRR sampling sessions (3) Parameterizing this function is system specific, because it depends on detectability of individuals, but only requires an initial effort of traditional high effort MRR in a few populations encompassing the expected range of population size, combined with previous knowledge on the species about potential factors affecting detectability.

- Our simplified MRR protocol should allow scientists to obtain absolute population size estimates almost as good as with traditional high effort MRR, but at a cost lowered in both the marking procedure and the intensity of field visits.

Keywords: Boloria eunomia, Boloria aquilonaris, bog fritillary butterfly, cranberry fritillary butterfly, Capture Mark Recapture, catch effort, Mark software, long-term monitoring, sampling efforts

INTRODUCTION

Quantifying and understanding the distribution and abundance of organisms represents the ultimate subject matter of ecology (Krebs 1972). The number of individuals (i.e. population size) is a fundamental demographic unit of a population (Williams et al. 2002; Van Dyke 2008). Obtaining an accurate estimation of population size is thus a basic step to assess species and population status and trends in a given area and to guide effective conservation (Sutherland 1996). As it is most often impossible to inventory or census all the individuals in a given population (Preston 1979), estimation methods have to be used (Williams et al. 2002). Mark Release Recapture (MRR, also known as Capture Mark Recapture, CMR) is a standard, broadly used procedure to obtain estimates of absolute population size while overcoming the problem of imperfect detection (i.e. not all individuals can be recorded). It is employed for a wide range of taxa, e.g. for small mammals (Lindenmayer et al. 1998), birds (Morrison et al. 2004), amphibians (Arnold et al. 2002) and butterflies (Schtickzelle et al. 2002). There is ample evidence that MRR gives reliable estimates (Williams et al. 2002; Haddad et al. 2007; Nowicki et al. 2008; Grimm et al. 2014) as long as some basic assumptions are respected: mainly, unique and permanent markings, no effect of marking

on behaviour, and homogeneity among individuals from the same group in terms of capture probability, survival rate and birth rate (i.e. groups with different values can exist, such as males vs females, but homogeneity must exist within groups); models relaxing some of these assumptions exist but are often tailored to specific cases (Lindberg 2012). Guidelines, common designs and statistical models are broadly available (e.g. Sandercock 2006; Lindberg 2012; Cooch and White 2017), and software is now widespread to analyse MRR data (see Bunge 2013 for a review), such as the widely used MARK program (White and Burnham 1999), marked (Laake et al. 2013), SPACECAP (Gopalaswamy et al. 2012) and others.

Alternative methods exist to estimate (relative) abundance, which include area or time-limited census methods, point counts and transect walks (see Per Douwes 1970; van Strien et al. 1997; Thomas 2005; van Swaay et al. 2008 for a description of those methods). They are less time consuming, may generate less negative impact on individuals and habitats (e.g. Gross et al. 2007; Nowicki et al. 2008), and can be more easily used for entire species' communities or a large spatiotemporal scales (van Swaay and van Strien 2005; Collier et al. 2008). However, these methods provide only indices of relative abundance (Nowicki et al. 2008): they cannot, by definition, derive estimates of absolute population sizes because this requires estimating (1) the detectability of individuals to be able to quantify the fraction of the population that remains unseen (e.g. Clobert 1995) and (2) some measure of lifetime expectancy to quantify the rate of turnover of individuals and associated probability of multiple countings (e.g. Nowicki et al. 2005). Species action plans or large scale monitoring schemes try to overcome such limitations by a high level of standardization in the count protocol. This can be very successful to produce global population trends, but requires assuming constant detectability among the conditions to be

compared (species, sites, time series…). Believing abundance indexes from count methods are automatically reliable estimates of absolute population size is misleading. This was for example illustrated in the Mitchell's satyr butterfly (Shuey and Szymanski 2010) with no correlation found between daily population size obtained by MRR and abundance estimated from transect walks (but see Thomas 1983). Distance sampling is another count method (see Buckland et al. 2005 for a description) that allows assessing detectability (as a function of the distance to the observer) to evaluate population density. Isaac et al. (2011) compared results obtained with transect counts and distance sampling applied to butterflies and found that population density estimates were highly correlated. However, we found no evidence that it can be used to estimate absolute population size, except for the sand dune lizard where counts yield consistent population size estimates with MRR (Kacoliris et al. 2009). MRR therefore remains the method of choice to estimate absolute population size when this is required.

Methods of analysing MRR work particularly well for univoltine insects with clear spatial (discrete habitat patches) and temporal (non-overlapping generations) population boundaries, characterizing many habitat-specialist and rare butterflies (Ehrlich and Hanski 2004). However, even in this case, MRR is time- consuming and laborious, with costs increasing sharply with the number of populations, generations and years to monitor (Field et al. 2005). MRR protocols can also have potential negative impacts on the viability and recapture frequency of individuals due to their manipulation/handling and possible associated physical damage (e.g. Singer and Wedlake 1981; Morton 1982; Gall 1984). Finally, intensive repeated visits to study sites can affect the habitats, e.g. through vegetation trampling, disturbing elusive animals, or even facilitating invasive species and diseases (Ruiz and Carlton 2003). Alternative individual marking techniques limiting

potential negative impacts (e.g. using individual identification based on the combination of camera traps, body marks or DNA fingerprints as the mark to recognize individuals), regularly employed with birds and mammals, can hardly be transposed to insects. Therefore, MRR implementation remains limited in nature reserves, fragile ecosystems and endangered species with a limited number of populations and individuals.

So, although butterflies in particular, and other insects in general, are considered as good indicators for which estimates of population size might be of high interest, we did not find evidence in the literature that the existing alternative methods to MRR can provide reliable estimates of absolute population size (contrary to relative abundance index). Consequently, there is a need to develop less expensive and less time-consuming methods than traditional high effort MRR that could still allow for a rigorous estimation of population size when this is needed. Here, "high effort MRR" is to be understood as MRR with unique marks for individuals, and an intensity (number and timing of sampling sessions) that is large enough to provide reliable estimates of demographic parameters; what this represents in practice depends on the study species and system, and more specifically of the recapture rate, the key to estimate detectability and use it to correct estimates of survival and population size (Cooch and White 2017).

In this paper, we focus on a methodology to develop a reduced effort MRR sampling protocol providing estimates of population size that are almost as accurate but with a much lower cost. It implies to count the number of different individuals, discriminating already counted individuals via a simple marking, and to apply a conversion function to transform it into a population size estimate. The need to first calibrate the conversion function makes the protocol most useful for studies implying repeated quantification of population size. We illustrate and test it using one of the largest MRR databases existing for butterflies: it

contains 150 independent MRR datasets among which 115 were used for the present analysis with around 24,000 marked individuals and 41,500 (re)captures (appendix 1), collected yearly over two decades in a series of Belgian populations of two butterfly species, the bog fritillary *Boloria eunomia* and the cranberry fritillary *B. aquilonaris*. In particular, we investigate the following questions: (1) Is marking of individuals necessary? (2) Can the marking be simplified into a single generic mark used for all the captured individuals instead of a unique identifier, simply to distinguish previously marked and unmarked individuals? (3) Can the sampling effort be reduced while maintaining the estimates for population size as reliable as with high effort MRR? (4) Can a general conversion function be used in different contexts and/or for different species and how to estimate its parameters?

METHODS

Study species & landscapes

The bog fritillary *Boloria eunomia* (ESPER, 1799) and the cranberry fritillary *B. aquilonaris* (STICHEL, 1908) are specialist species of wet meadows and peat bogs. Their distribution in Belgium is restricted to the south of the country, and both species are considered as vulnerable in Belgium (Fichefet et al. 2008), but of least concern in Europe (van Swaay et al. 2011). We studied 15 populations of *B. eunomia* and 14 populations of *B. aquilonaris*, over the 1992-2012 period; not every population was sampled every year however.

The two species were sampled by MRR: habitat patches were regularly visited (every 4 days on average), weather permitting, during the flight period (May-June for *B. eunomia*, June-July for *B. aquilonaris*), and butterflies were netted and marked with an individual code on the underside of the left hindwing using a permanent pen. Sampling routes were kept

fixed and their lengths adjusted to the area of every habitat patch to keep catch effort homogeneous. For each (re)capture, the following information was recorded: individual code, first capture vs. recapture, sex, date and location (habitat patch). This protocol was similar for the two species and kept constant for all populations over the years. Sites supporting *B. aquilonaris* populations were classified as "open" (large bogs without tree edges) vs. "closed" (bogs surrounded by tree edges). To reduce error rate and ensure the highest possible quality to our MRR data, the marking protocol has been optimized and MRR data extensively checked (Schtickzelle 2003). First, individual codes were formed with signs and figures that were highly reading error proof even if some parts of the signs are lost (e.g. when a portion of the wing was damaged). Second, the capture histories (sequences of capture records for each individual) were checked for inconsistencies in sex, location or timestamp (e.g. an individual cannot change sex, or be recaptured before being marked).

For both species, dispersal events between populations were very rare, so we could assume that population size is not biased by dispersal events. Furthermore, each dataset (i.e. MRR data collected on one specific population and specific year) is statistically independent from all the others because they share extremely few data and very few individuals were recorded in more than one population. Accordingly, every dataset was analysed separately, and could be regarded as one independent data point in subsequent statistical analyses.

Reference population size

The total population size, #**Ntot**, corresponding to the total number of different butterflies present in a given population in a given year (i.e. over the whole yearly adult generation), was estimated using Jolly-Seber (JS) models, as implemented in the POPAN analysis in Mark

software (White and Burnham 1999). Based on capture histories of the different individuals recorded in a population, the probability of an individual to be (re)captured (a measure of detectability) is estimated, and subsequently used to correct estimates of survival, birth rates, daily and total (seasonal) population size (Cooch and White 2017). For each dataset, we computed #Ntot, its standard error and 95% confidence interval following the methodology and its implementation for butterflies' MRR datasets as initially described in Schtickzelle et al (2002).

Calculation of abundance metrics

A series of abundance metrics were computed for each MRR dataset separately:

- **#C**, the total number of captured (i.e. marked) individuals;
- **#CR**, the total number of (re)capture records;
- **#CRmax**, the maximum number of (re)capture records on any single capture session;
- #**Cadj**, the adjusted versions of #C according to sampling effort (see below);
- **#CRadj**, the adjusted versions of #CR according to sampling effort (see below).

#CR and #CRmax are proxies for simple counts that do not distinguish previously counted from newly seen individuals. #C and #CR being sums over all sampling sessions of the dataset, they are likely to increase with the sampling effort (i.e. the number of MRR sampling sessions, **#Sampling**). We therefore computed #Cadj and #CRadj as adjusted versions of the #C and #CR abundance metrics by dividing them by an inflation factor **IF**. IF is assumed to sigmoidally increase with the sampling effort from 0 to a maximum value of 1 in the high effort MRR dataset; it was therefore computed according to the following equation:

$$\text{logit}(IF) = a + b * \#Sampling.$$

The two parameters, a and b, were estimated by logistic regression of IF according to #Sampling (PROC GENMOD in SAS 9.4, www.sas.com) on the pool of MRR datasets containing at least six sampling sessions and 25 marked butterflies (i.e. 59 datasets for *B. eunomia*, 25 and 28 datasets for *B. aquilonaris* in closed and open sites, respectively). For each dataset, we computed the values of #C and #CR that would have been obtained if the population was sampled for a certain number of sessions, from three to the real number of samples. This was performed by downsampling the MRR data to keep (re)captures recorded on a subset of samplings sessions, as regularly spaced during the flight season as possible. In practice, we first determined the mean length of the flight season for each of the species, which was 28 days for *B. eunomia* and 25 for *B. aquilonaris*. We then split the flight season into time intervals of equal length, whose midpoints were the target dates for downsampling. Finally, (re)captures recorded on the sampling date closest to each midpoint was retained. #C was then computed as the total number of different butterflies recorded at least once in these samplings days, and #CR as the total number of (re)captures. Dividing #C (or #CR) by the real value observed in the full dataset gave the observed data point, i.e. the value of IF, expressing the proportion of individuals that would have been marked if sampling had been restricted to that specific number of sessions.

Statistical analysis of the power of abundance metrics as population size predictors

The five abundance metrics described above (#C, #Cadj, #CR, #CRadj and #CRmax) were individually used in a linear model to explain variations in #Ntot among the datasets. No intercept term was included because a zero population size is expected when no MRR data is recorded; this also helps avoiding problems where the intercept, hence #Ntot predicted at small #C, is estimated as a negative value given the best line fit is constrained by data points

at large #C (more information about forcing zero intercept is given in **Appendix 2**). The slope was estimated via weighed linear regression (PROC GENMOD in SAS 9.4). For *B. aquilonaris*, two variants of the model were fitted, one with a different slope for closed and open sites, and one with a single slope for all sites. The weight of each observation was 1/cv_Ntot, with cv_Ntot being the coefficient of variation associated to the estimate of #Ntot from the original MRR datasets. The rationale for using such a weighed regression is that the relative uncertainty in the estimation of #Ntot from the MRR data was different for each dataset, according to the amount of information it contained (linked to the number of marked individuals and overall frequency of recapture observed for that population that year).

The relative predictive power of the different models (5 for *B. eunomia*, 10 for *B. aquilonaris*) was compared on three criteria: (Criterion 1) R^2 and the AICc value of the model, expressing the fit/complexity ratio based on the absolute prediction error |#Ntot-#Ntot_predicted|; (Criterion 2) the average over all datasets of the relative prediction error, computed as |#Ntot-#Ntot_predicted|/#Ntot, expressing prediction error in % instead of absolute magnitude; and (Criterion 3) the proportion of datasets for which #Ntot_predicted fell within the 95% confidence interval of #Ntot. The rationale to use these criteria is to obtain a more complete picture of the prediction power of each model (beyond merely goodness of fit), with quantitative measures that consider especially population size, since obviously a given error of, let's say, 10 individuals would be far more significant if #Ntot was 30 than if it was 1000 individuals.

RESULTS

Data summary

The 63 MRR datasets for *B. eunomia* and 52 for *B. aquilonaris* total to 13,246 and 10,851 marked individuals, 26,973 and 14,489 (re)captures, respectively. Reliable estimates of the reference population size (#Ntot) could be obtained using Jolly-Seber demographic models from 61 and 36 of these datasets for the two species, respectively. #Ntot ranged from 14 to 1553 individuals (mean = 359) for *B. eunomia* and from 53 to 2482 individuals (mean = 702) for *B. aquilonaris*. In these datasets providing a #Ntot estimate, the number of sampling sessions (#Sampling) per dataset (one species, one population, one year) ranged from 6 to 35 (mean = 12.6) for *B. eunomia* and from 7 to 22 (mean = 10.2) for *B. aquilonaris*. More details on demographic metrics for all datasets are provided in **Appendix 1**.

Inflation factor: how many sampling sessions do we need?

By downscaling datasets and plotting the inflation of marked individuals with sampling intensity, we found that the slope of the saturation curve (inflation factor, IF) differed between the two species as well as between the open and closed sites for *B. aquilonaris* (**Figure 1**). Nonetheless, for both species about 80% of the population was already marked by 6-8 sampling sessions during the flight season.

Models for predicting population size

Among the tested metrics and models for estimating the total population size (#Ntot), #Cadj performed best for *B. eunomia*, while a model considering site type (open vs. closed) and #C performed best for *B. aquilonaris* (**Table 1**). For the two species, any of the metrics for estimating the total population size (#Ntot) could seemingly explain a high proportion of the

variation in #Ntot among datasets (R^2 = 71% to 98%, **Table 1**). However, there were striking differences in predictive performance when assessing the different models in terms of prediction error: abundance metrics that are based on marked individuals (#C and #Cadj) had low error with respect to the estimated #Ntot (18-19% for both species), whereas metrics that are based on counts only (#CR, #CRadj and #CRmax) yielded double or nearly-double error values for both species. Better performance was reflected also by substantially lower AICc values for models containing #Cajd and #C compared to all other models. For *B. eunomia* it was evident also when we considered the proportion of (downsampled) datasets included in the 95% CI of #Ntot: 62% and 54% the models considering #C and #Cadj were included in the 95% CI, compared to 26-38% for the other models (**Table 1**). For *B. aquilonaris* the difference was less profound, but still, any model with marked individuals performed better compared to the same model/metric with unmarked individuals (counts).

Calculating the conversion function: how many individuals are missed per marked or observed one?

Plotting the predicted population size according to the number of marked individuals illustrates that the "conversion function" ("marked to real") differed between the two species and between the two site types for *B. aquilonaris* (**Figure 2**). The slope of the function was 1.49 for *B. eunomia*, meaning than for any two marked butterflies, circa 1 individual was "missed" in the population. For *B. aquilonaris*, the slope of the relation was significantly higher for open (3.23) compared to closed sites (1.89), meaning that for every marked individual, either circa 2 or 1 individuals were missed in open versus closed sites.

DISCUSSION

Because high effort MRR sampling, as usually performed when estimating population size is the aim, is costly in terms of time and money, and potentially impacts sampled sites and species due to high catch effort, several alternative methods have been developed. Many of these methods involve replacing marking and recapturing individuals by simply counting individuals that are seen. As detailed in the introduction, these methods may offer good estimates of relative population abundance indices, and they have indeed been used successfully for large surveys, such as butterfly monitoring schemes. However, they are not suitable or designed for quantifying absolute population size, which remains a key for population viability analyses (Morris and Doak 2002; Schtickzelle and Baguette 2009; Pe'er et al. 2013) or other estimations of risks to species' populations under (anthropogenic) pressures. Here, we did not aim at comparing MRR with count methods: their objectives are different, and dropping individual identity information does not make MRR data equal to transect counts because of missing standardization steps, such as the moving box around the experimenter where individual are counted. We addressed the question of how to reduce the efforts to be invested in MRR without compromising the quality of the population size estimation.

The new finding that clearly arises from our study, performed on one of the largest collections of MRR datasets existing for butterflies, is that it is possible to get a reliable quantification of population size from a simplified MRR protocol via a simple linear conversion function encapsulating all aspects of detectability and rate of turnover of individuals into a slope quantifying the "number of existing (i.e. marked + missed) individuals for each marked one" (**Figure 2**). Our results indicate three clear specific

conclusions: (1) individual marking, even in a simplistic way, is needed for estimating population size; (2) the conversion function can be reliably applied on the number of individuals marked in a limited number of sampling sessions (around 6-8 in our case), largely reducing the overall cost of the sampling; (3) this function is species-specific (and potentially also habitat- or sex-specific) and an initial effort of high effort MRR in sites covering the range of expected population sizes is needed to parameterize it. Next, we will discuss these conclusions in details, and then provide a methodology for reducing MRR efforts in future studies.

To quantify absolute population size, individuals must be marked

On the three cases studied here, the predictive power of models involving marked individuals (captures only) was very good, and largely better than models based on counting the number of observed animals (i.e. captures and recaptures pooled). This confirms that a reliable and precise quantification of population size implies to estimate two parameters: (1) the detectability of individuals, which is known already from a long time as a required quantity to convert number of individuals observed into number of individuals present in the population (e.g. Clobert 1995; Gross et al. 2007; Ry and Schmidt 2008; Isaac et al. 2011) and (2) the rate of turnover of individuals, which influences the probability of multiple counts of the same individual. Since detectability may largely vary in space and time, between species and even sexes (e.g. due to the movement behaviour of species; Turlure et al. 2011), MRR studies used to sample every population of interest with an effort (number of sampling sessions and capture intensity) large enough to estimate it adequately, via the knowledge of the capture histories of the individuals (Schtickzelle et al. 2002). Furthermore, contrary to marking, simply counting the individuals does not prevent from multiple counts

of the same individual, whose probability depends on its lifetime. Marking is then necessary to quantify absolute population size, but our results show that it does not imply to record the complete capture history for every individual separately, which requires intensive and repeated MRR with unique individual identifiers. This means that a simplified marking protocol can be used, which can be as simple as a single mark applied to all individuals, greatly simplifying and lightening marking and data recording processes. Furthermore this marking protocol is also suitable for species too small to allow marking with an individual identifier, such as many of the Lycaenids or Hesperids.

The MRR sampling effort can be reduced to a few sampling sessions per population only

Marking is a necessary, but not sufficient, condition to obtain a reliable quantification of population size. A minimal catch effort, in quantity and quality, is needed too in order to obtain a reliable estimate of the number of marked individuals to be translated into population size using the conversion function. In the case of *B. eunomia* and *B. aquilonaris*, the inflation factor curves (**Figure 1**) indicate that after 6-8 sampling sessions, 60%-80% of the individuals that could be marked with many sampling sessions were already marked. It is important however to spread these sampling sessions over the flight season so that every individual present in the population has a chance to be marked. Otherwise, some individuals may be born and die during a "no sampling" period, meaning they cannot be marked or counted, leading to the underestimation of the population size. Notably, sampling frequency is indeed an issue also in systematic monitoring (as shown by e.g. Schmucki et al. 2016) indicating a need to consider sampling frequency with respect to the anticipated life-span and asynchronous emergence of adults and sexes, also when individuals are not marked. In a similar MRR simplification attempt, Nowicki et al. (2005) provided a reduced

effort protocol (at least five sampling sessions) based on the conversion of peak daily population size into total population size. It uses a formula containing both the lifetime expectancy (based on recording full capture histories of individually marked butterflies) and the duration of the flight period and has been calibrated on several species. Longcore et al. (2003) also proposed a method that takes lifespan of individuals into account; based on Zonneveld (1991), it uses the death rate of individuals to correct daily count data.

The conversion function must be parameterized with some initial high effort MRR data

Before it can be used to translate a number of marked individuals into population size, the conversion function must be calibrated with the adequate slope for the study system. This is also true for the inflation factor according to the number of MRR sampling sessions. Indeed, the sampling effort needed to accurately record individuals, mark them or even notice species presence obviously varies greatly from one species to another, and even within a species; this is because detectability varies among species but also sexes and contexts. In their simulation study, Archaux et al. (2012) showed that a small detectability difference (4-8%) can lead to the miscalculations of population sizes in 50-90% of the cases. Detectability can greatly vary between species but also sexes and contexts. For instance, Pellet (2008) found detection probabilities ranging from 50% to 77% during transect counts while comparing four butterfly species. This is because individuals hiding in the vegetation or using a perching strategy are probably less easily detected than constantly patrolling ones, or because species can have cryptic coloration. Also, detectability was for example assessed at 48% in woodland edges vs. 88% in open fens for the butterfly *M. nausithous* (Pellet et al. 2012).

In our results, the slope of the conversion function, expressing the number of individuals present in the population for every one that was marked, ranged from 1.49 for *B. eunomia*, to 3.23 for *B. aquilonaris* in open sites, with an in-between 1.89 for *B. aquilonaris* in closed sites. These can easily be related to behavioural differences: studies of the flight behaviour within habitat indicated a rather tortuous and slow flight in *B. eunomia* versus rather straight and rapid flight in *B. aquilonaris* (Schtickzelle and Baguette 2003; Turlure et al. 2011). The behaviour of *B. aquilonaris* in flight, differing between open and closed habitats, also translates into differences in detectability according to site configuration: open areas are often windswept, reinforcing the flight speed of individuals, while closed areas are wind protected by trees.

Fortunately, the conversion function turned out to give estimates of population size very close to those obtained with high effort MRR even with an extremely simple equation, namely a linear relationship without intercept (**Figure 2**). This means that it can be parameterized quite easily for the study system with only a few data points, i.e. populations for which the real population size has been estimated as precisely as possible using high effort MRR. These populations chosen for fitting the conversion function should be as much as possible spread over the range of expected population size in the study area to improve the estimation of the slope by linear regression and the ability to check if the linearity assumption holds over that range. **Figure 3** illustrates how the estimation of the slopes rapidly stabilizes as the number of data points (#Ntot estimates from high effort MRR) used to fit the conversion function increases in our three case studies. Such high effort MRR datasets are also suitable to estimate how the inflation factor increases with the number of sampling sessions.

A simplified Mark-Release-Recapture protocol

Based on these conclusions, we propose a simple approach to obtain estimates of population size that are almost as good as those yielded by high effort MRR, but with a much lower sampling intensity. Our proposed simplified MRR is split into four steps (**Figure 4**). First (Step 1: Site selection), one must identify a few (say, 4-5) sites hosting populations with a range of different expected sizes; if detectability differences are expected, e.g. among landscape contexts, this selection should be repeated for each context. Secondly (Step 2: MRR data collection), high effort MRR is conducted to obtain precise estimates of population size using classical demographic analyses based on full capture histories of the individuals. MRR data previously collected and/or published could be reused here, as we did in this study with *B. eunomia* and *B. aquilonaris*. The third step (Step 3: Inflation factor) involves estimating the inflation factor by downsampling these high effort MRR datasets. At that stage, it is possible to analyse the impact of the number of sampling sessions and their temporal occurrence so as to determine the optimal MRR sampling design. Finally (Step 4: Conversion function), one needs to parameterise the conversion function by calibrating its slope using the number of marked individuals and the estimated population size in this set of sites. The linearity assumption can be easily checked (**Figure 3**) and extra high effort MRR datasets collected if in doubt about the estimate of the slope. In all four stages, it is important to employ good knowledge of species' biology and behaviour, to consider context-specific effects that could affect these conversion ratios.

Such a protocol will decrease significantly the cost of MRR studies aiming at estimating precisely the population size by allowing several major simplifications: (1) a simple group marking, even a single mark, can be used; (2) only a few MRR sampling sessions (here 6-8, to be compared to 10-13 on average and up to 35 in our high effort MRR)

are required to get the metric estimate to be converted into population size; (3) initially, a one-shot high effort MRR campaign must be done to parameterize the conversion function and the inflation factor, but sampling a few populations is enough. Such a decrease in the sampling effort can significantly reduce the costs associated to each demographic survey and/or allow surveying more populations for the same cost. As a practical example, let's imagine one would like to sample all *B. aquilonaris* populations in Belgium (i.e. 54 populations), with on average 1 h of sampling per site by one or two persons simultaneously between 09:00 and 18:00, and excluding the journey between populations. It would take 36 days (324 h) in reduced effort MRR (6 sampling sessions) vs 60 days (540 h) in a high effort MRR (10 sampling sessions). Combined via the simple marking of the individuals, which accelerates MRR on the field and data coding, this makes it possible to survey all these 54 populations during the flight period of the species (roughly 5 weeks in Belgium).

Nevertheless, it must be kept in mind that this conclusion holds true for the estimate of population size, but high effort MRR studies are useful to study other aspects of (meta)population dynamics. In these cases, the simplified protocol we present here might not be the best solution. For example, to record dispersal events, individual specific (or at least site specific) marking is necessary, and more intense and more frequent MRR sampling sessions mean more movement data with a finer spatiotemporal resolution (Baguette et al. 2011). Another example where our simplified protocol is not adequate is to address questions involving the estimation of vital rates of adults, such as survival or lifetime expectancy (e.g. Vandewoestijne et al. 2008).

Conclusion

With this study, we add to the existing evidence that counting individuals does not allow to estimate absolute population size because detectability and rate of turnover of individuals remain unknown; individuals need to be marked. Count methods, and the relative abundance indexes they provide, are very useful in some contexts, e.g. to give the big picture of abundance trend trough time (see the many successful examples of butterfly monitoring schemes), but are not aimed at, and cannot be used for, quantifying absolute population size. Obtaining a reliable quantification of absolute population size is still of prime importance in other contexts, e.g. quantitative modelling of population viability analysis or definition of IUCN threat status, and MRR is the method of choice for this purpose. We offer here a simple and efficient simplified MRR protocol as a way to reduce its cost and potential impact on species and sites with a limited effect on the reliability of the population size estimate. We believe this protocol, in its approach but not especially its specific details (such as the linearity assumption of the conversion function), can be extended to cases with similar characteristics, i.e. mainly aiming at estimating true population size for species (1) with non-overlapping generations and (2) whose populations can be reasonable well delimited in space. Only with these two conditions fulfilled, the estimate of a total size is meaningful for a population because it is finite in time and in space; otherwise, only instantaneous population size is to be estimated, as done for many birds or mammals. However, the generalisation power of our simplified protocol has still to be formally tested on other taxonomic groups.

Acknowledgments

CT (post-doctoral fellow) and NS (research associate) acknowledge financial support from the F.R.S.-FNRS. GP acknowledges support by the EU FP7 project EU BON (project #308454). MB is part of the 'Laboratoire d'Excellence' (LABEX) entitled TULIP (ANR-10-LABX-41), and acknowledges support by the TenLamas project (EU FP6 BiodivERsA Eranet). CT, GP and MB acknowledge support by the EU FP7 project SCALES (project #226852). We thank Julie Choutt, Sofie Vandewoestijne, Guillaume Senterre, Quentin Dubois, Guillaume Latour, Gabriel Nève and Luc Mousson for their valuable help with field work. We thank Piotr Nowicki and an anonymous reviewer for their in-depth constructive comments, which helped us to further improve this manuscript. Site access and a permission to study the species in the field were granted by the Ministère de la Région Wallonne (now Service Public de Wallonie). This is publication no. BRC405 of the Biodiversity Research Centre of the UCL. We have no conflict of interest to declare.

Data Accessibility

Data are presented in Appendix 1 and archived in Pangaea (https://pangaea.de; ref: PDI-15846).

Author contribution

CT, MB and NS designed the study. CT, MB and NS performed data collection. CT and NS performed data analyses; all authors commented, interpreted and participated to the improvement of these analyses. CT and NS wrote the first draft of the manuscript, and all

authors contributed substantially to revisions. All authors read and approved the final version of the paper.

References

1. Archaux, F., Henry, P.Y. & Gimenez, O. (2012) When can we ignore the problem of imperfect detection in comparative studies? *Methods in Ecology and Evolution,* 3, 188-194.
2. Arnold, T.W., Anderson, M.G., Sorenson, M.D. & Emery, R.B. (2002) Survival and philopatry of female redheads breeding in Southwestern Manitoba. *The Journal of Wildlife Management,* 66, 162-169.
3. Baguette, M., Clobert, J. & Schtickzelle, N. (2011) Metapopulation dynamics of the bog fritillary butterfly: experimental changes in habitat quality induced negative density-dependent dispersal. *Ecography,* 34, 170-176.
4. Buckland, S.T., Anderson, D.R., Burnham, K.P. & Laake, J.L. (2005) Distance sampling. *Encyclopedia of Biostatistics*), John Wiley & Sons, Ltd,
5. Bunge, J.A. (2013) A survey of software for fitting capture–recapture models. *Wiley Interdisciplinary Reviews: Computational Statistics,* 5, 114-120.
6. Clobert, J. (1995) Capture-recapture and evolutionary ecology: A difficult wedding? *Journal of Applied Statistics,* 22, 989-1008.
7. Collier, N., MacCay, D.A. & Bekendorff, K. (2008) Is relative abundance a good estimator of population size? Evidence from fragmented populations of a specialist butterfly (Lepidoptera: Lycaenidae). *Population Ecology,* 50, 17-23.
8. Cooch, E.G., White, G.C. (2017) Program Mark: A gentle introduction. Online Source.
9. Ehrlich, P.R. & Hanski, I. (2004) On the wings of checkerspots. A model system for population biology. Oxford University Press, New York.
10. Fichefet, V., Barbier, Y., Baugnée, J.-Y., Dufrêne, M., Goffart, P., Maes, D. & Van Dyck, H. (2008) Papillons de jour de Wallonie (1985-2007). Publication du Groupe de Travail Papillons de jour Lycaena et du Centre de Recherche de la Nature, des Forêts et du Bois (MRW-DGRNE), Gembloux.
11. Field, S.A., Tyre, A.J., Possingham, H.P. & Lubow (2005) Optimizing allocation of monitoring effort under economic and observational constraints. *Journal of Wildlife Management,* 69, 473-482.
12. Gopalaswamy, A.M., Royle, J.A., Hines, J.E., Singh, P., Jathanna, D., Kumar, N.S. & Karanth, K.U. (2012) Program SPACECAP: software for estimating animal density using spatially explicit capture–recapture models. *Methods in Ecology and Evolution,* 3, 1067-1072.

13. Grimm, A., Gruber, B. & Henle, K. (2014) Reliability of different Mark-Recapture methods for population size estimation tested against reference population sizes constructed from field data. *Plos One,* 9, e98840-

14. Gross, K., Kalendra, E.J., Hudgens, B.R. & Haddad, N.M. (2007) Robustness and uncertainty in estimates of butterfly abundance from transects counts. *Population Ecology,* 49, 191-200.

15. Haddad, N.M., Hudgens, B.R., Damiani, C., Gross, K. & Kuefler, D. (2007) Determining optimal population monitoring for rare butterflies. *Conservation Biology,* 22, 929-940.

16. Isaac, N.J.B., Cruickshanks, K.L., Weddle, A.M., Marcus Rowcliffe, J., Brereton, T.M., Dennis, R.L.H., Shuker, D.M. & Thomas, C.D. (2011) Distance sampling and the challenge of monitoring butterfly populations. *Methods in Ecology and Evolution,* 2, 585-594.

17. Kacoliris, F.P., Berkunsky, I. & Williams, J.D. (2009) Methods for assessing population size in sand dune Lizards (*Liolaemus multimaculatus*). *Herpetologica,* 65, 219-226.

18. Krebs, C.J. (1972) Ecology. Harper & Row, New York.

19. Laake, J.L., Johnson, D.S. & Conn, P.B. (2013) marked: an R package for maximum likelihood and Markov Chain Monte Carlo analysis of capture–recapture data. *Methods in Ecology and Evolution,* 4, 885-890.

20. Lindberg, M.S. (2012) A review of designs for capture–mark–recapture studies in discrete time. *J Ornithol,* 152, 355-370.

21. Lindenmayer, D.B., Lacy, R.C. & Viggers, K.L. (1998) Modelling survival and capture probabilities of the mountain brushtail possum (Trichosurus caninus) in the forests of south-eastern Australia using trap-recapture data. *Journal of Zoology,* 245, 1-13.

22. Longcore, T., Mattoni, R., Zonneveld, C. & Bruggeman, J. (2003) Insect Count Analyzer: a tool to assess responses of butterflies to habitat restoration. *Ecological Restoration,* 21, 60-61.

23. Morris, W.F. & Doak, D.F. (2002) Quantitative conservation biology. Theory and practice of population viability analysis. Sinauer Associates, Sunderland, Massachusetts.

24. Morrison, S.A., Bolger, D.T. & Sillett, T.S. (2004) Annual Survivorship of the Sedentary Rufous-Crowned Sparrow (Aimophila ruficeps): No Detectable Effects of Edge or Rainfall in Southern California. *The Auk,* 121, 904-916.

25. Nowicki, P., Richter, A., Glinka, U., Holzschuh, A., Toelke, U., Henle, K., Woyciechowski, M. & Settele, J. (2005) Less input same output: simplified approach for population size assessment in Lepidoptera. *Population Ecology,* 47, 203-212.

26. Nowicki, P., Settele, J., Henry, P.-Y. & Woyciechowski, M. (2008) Butterfly monitoring methods: the ideal and the real world. *Israel Journal of Ecology and Evolution,* 54, 69-88.

27. Pe'er, G., Matsinos, Y.G., Johst, K., Franz, K.W., Turlure, C., Radchuk, V., Malinowska, A.H., Curtis, J., Naujokaitis-Lewis, I., Wintle, B.A. & Henle, K. (2013) A protocol for better design, application, and communication of Population Viability Analyses. *Conservation Biology,* 27, 644-656.

28. Pellet, J. (2008) Seasonal variation in detectability of butterflies surveyed with Pollard walks. *Journal of Insect Conservation,* 12, 155-162.

29. Per Douwes (1970) Size of, gain to and loss from a population of adult *Heodes virgaureae* L. (Lep., Lycaenidae). *Insect Systematics & Evolution,* 1, 263-281.

30. Preston, F.W. (1979) The Invisible Birds. *Ecology,* 60, 452-454.

31. Ruiz, G.M. & Carlton, J.T. (2003) Invasive Species: Vectors and management strategies. Island Press, Chicago.

32. Ry, M. & Schmidt, B.R. (2008) Imperfect detection and its consequences for monitoring for conservation. *Community Ecology,* 9, 207-216.

33. Sandercock, B.K. (2006) Estimation of demographic parameters from live-encounter data: a summary review. *Journal of Wildlife Management,* 70, 1504-1520.

34. Schmucki, R., Pe'er, G., Roy, D.B., Stefanescu, C., van Swaay, C.A.M., Oliver, T.H., Kuussaari, M., van Strien, A.J., Ries, L., Settele, J., Musche, M., Carnicer, J., Schweiger, O., Brereton, T.M., Harpke, A., Heliölä, J., Kühn, E. & Julliard, R. (2016) A regionally informed abundance index for supporting integrative analyses across butterfly monitoring schemes. *Journal of Applied Ecology,* 53, 501-510.

35. Schtickzelle, N. (2003) Metapopulation dynamics and viability of the bog fritillary butterfly *Proclossiana eunomia*. Thesis/Dissertation.

36. Schtickzelle, N. & Baguette, M. (2009) (Meta)population viability analysis: a crystal ball for the conservation of endangered butterflies? *Ecology of butterflies in Europe* (eds J. Settele, T.G. Shreeve, M. Konvicka & H. Van Dyck), Cambridge University Press,

37. Schtickzelle, N. & Baguette, M. (2003) Behavioural responses to habitat patch boundaries restrict dispersal and generate emigration-patch area relationships in fragmented landscapes. *Journal of Animal Ecology,* 72, 533-545.

38. Schtickzelle, N., Le Boulengé, E. & Baguette, M. (2002) Metapopulation dynamics of the bog fritillary butterfly: demographic processes in a patchy population. *Oikos,* 97, 349-360.

39. Shuey, J. & Szymanski, J. (2010) Modified Pollard transects do not predict estimated daily population size for the secretive butterfly, *Neonympha mitchellii mitchellii* French. *Journal of the Lepidopterist's Society,* 66, 221-224.

40. Sutherland, W.J. (1996) Ecological census techniques, a handbook. Cambridge University Press, Cambridge.

41. Thomas, J.A. (1983) A quick method of assessing butterfly numbers during survey. *Biological Conservation,* 27, 195-211.

42. Thomas, J.A. (2005) Monitoring change in the abundance and distribution of insects using butterflies and other indicator groups. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 360, 339-357.

43. Turlure, C., Baguette, M., Stevens, V.M. & Maes, D. (2011) Species- and sex-specific adjustments of movement behavior to landscape heterogeneity in butterflies. *Behavioral Ecology,* 22, 967-975.

44. Van Dyke, F. (2008) The conservation of populations: concept, theory and analysis. *Conservation biology. Fundations, concepts, applications* (ed F. Van Dyke), pp. 213-242. Springer,

45. van Strien, A.J., Van De Pavert, R., Moss, D., Yates, T.J., van Swaay, C.A.M. & Vos, P. (1997) The statistical power of two butterfly monitoring schemes to detect trends. *Journal of Applied Ecology,* 34, 817-828.

46. van Swaay, C.A.M., Maes, D., Collins, S., Munguira, M.L., Sasic, M., Settele, J., Verovnik, R., Warren, M., Wiemers, M., Wynhoff, I. & Cuttelod, A. (2011) Applying IUCN criteria to invertebrates: How red is the Red List of European butterflies? *Biological Conservation,* 144, 470-478.

47. van Swaay, C.A.M., Nowicki, P., Settele, J. & van Strien, A.J. (2008) Butterfly monitoring in Europe: methods, applications and perspectives. *Biodiversity and Conservation,* 17, 3455-3469.

48. van Swaay, C.A.M. & van Strien, A.J. (2005) Using butterfly monitoring data to develop a European grassland butterfly indicator. *Studies on the Ecology and Conservation of Butterflies in Europe.* (eds E. Kühn, J.A. Thomas, R. Feldman & J. Settele), pp. 106-108. Pensoft publishers, Sofia.

49. Vandewoestijne, S., Schtickzelle, N. & Baguette, M. (2008) Positive correlation between genetic diversity and fitness in a large, well-connected metapopulation. *BMC Biology,* 6, 46-

50. White, G.C. & Burnham, K.P. (1999) Program MARK: survival estimation from populations of marked animals. *Bird Study,* 46 Supplement, S120-S139.

51. Williams, B.K., Nichols, J.D. & Conroy, M.J. (2002) Analysis and management of animal populations: modeling, estimation and decision making. Academic press, San Diego, California.

52. Zonneveld, C. (1991) Estimating death rates from transect counts. *Ecological Entomology,* 16, 115-121.

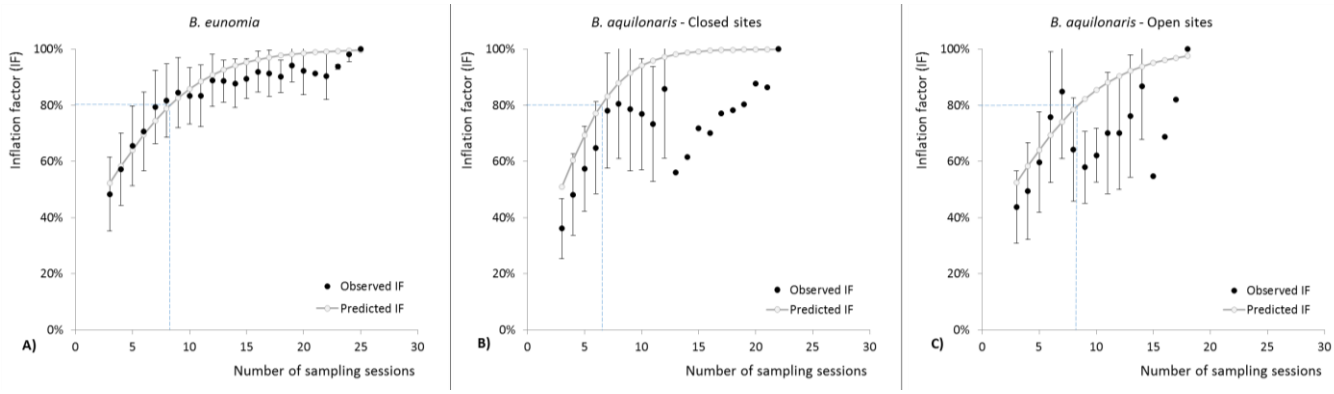
Tables

Table 1. Fit and predictive power of the different models tested (5 for *B. eunomia*, 10 for *B. aquilonaris*) to predict population size (#Ntot) from the following demographic metrics: #C = total number of marked individuals; #CR = total number of (re)captures; #CRmax = maximum number of (re)captures on a single day; #Cadj = adjusted version of #C according to sampling effort (number of MRR sampling sessions); #CRadj = adjusted versions of #CR according to sampling effort; type = site configuration (open vs. closed), for *B. aquilonaris* only. See text for details on how these metrics were estimated. Prediction error is reported both as mean relative prediction error, computed as |#Ntot-#Ntot_predicted|/#Ntot, and as the proportion of datasets for which the prediction was classified as correct, i.e. when #Ntot_predicted felt within the 95% confidence interval of the observed #Ntot estimate.

		Model fit			Model selection		Prediction error	
Species	**Demographic metric**	**Nb of parameters**	**Residual sum of squares**	**R^2**	**AICc**	**Δ AICc**	**Mean |prediction error|**	**Proportion of datasets with correct prediction**
B. eunomia	**#Cadj**	**2**	**3949213**	**98%**	**679.98**	**0.00**	18.4%	62%
	#C	2	4568310	97%	688.86	8.88	18.8%	54%
	#CRadj	2	23266000	86%	788.16	108.18	35.2%	38%
	#CR	2	27870228	83%	799.17	119.19	37.7%	26%
	#CRmax	2	30870246	81%	805.41	125.43	35.7%	31%
B. aquilonaris	**type*#C**	**3**	**4705355**	**98%**	**430.85**	**0.00**	17.5%	78%
	type*#Cadj	3	5873174	97%	438.84	7.99	19.4%	75%
	type*#CRadj	3	7019273	96%	445.25	14.40	28.5%	61%
	type*#CR	3	10639577	95%	460.23	29.38	28.6%	47%
	#Cadj	2	13650636	93%	466.81	35.96	36.0%	53%
	type*#CRmax	3	15444905	92%	473.64	42.79	33.5%	56%
	#C	2	19534423	90%	479.71	48.86	37.1%	53%
	#CRadj	2	35292646	82%	501.00	70.15	59.9%	39%
	#CRmax	2	47519807	76%	511.71	80.86	71.2%	33%
	#CR	2	56142692	71%	517.72	86.87	55.5%	36%

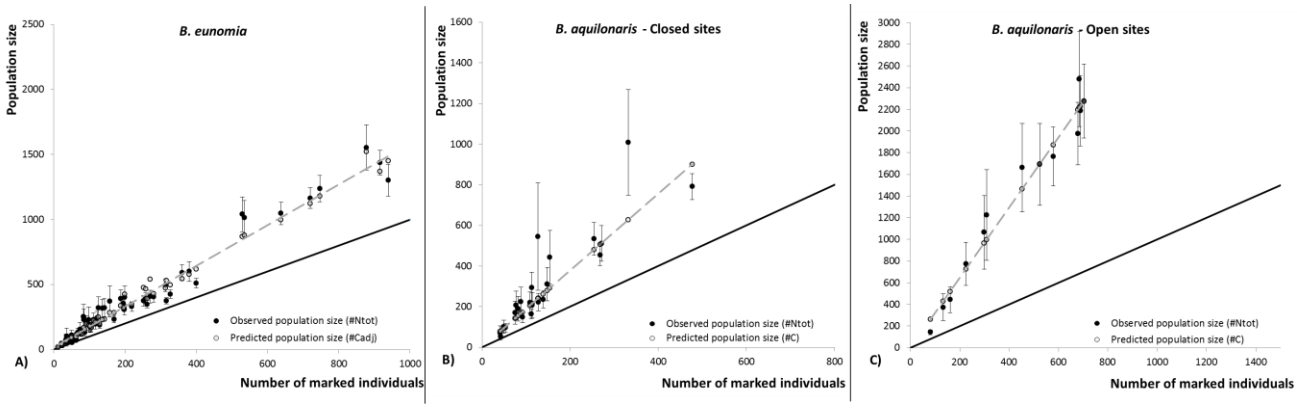
Figures

Figure 1.



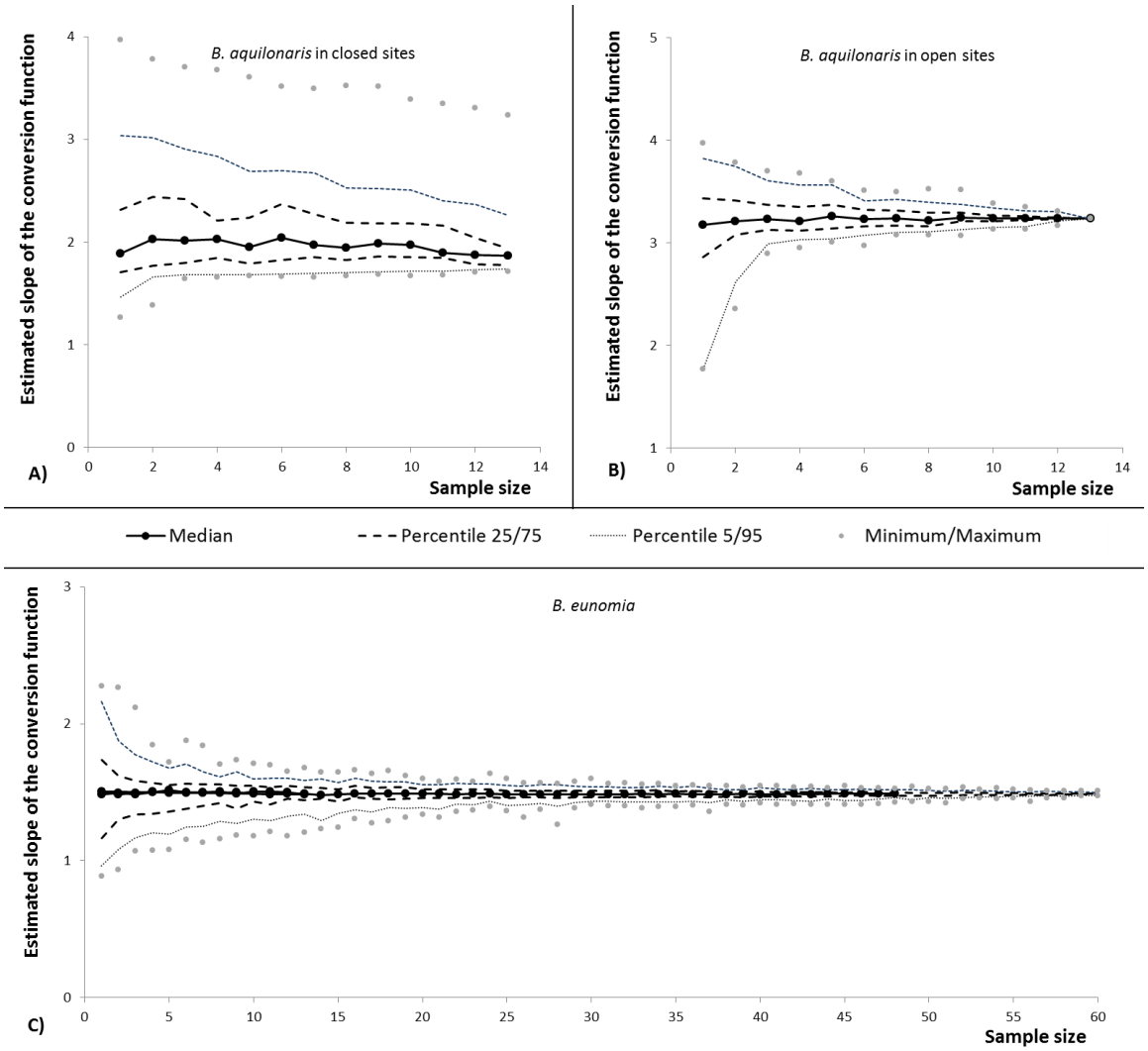
Inflation factor IF estimated as a sigmoidally increasing function of the number of sampling sessions; IF represents the proportion of the marked individuals in the full MRR dataset that would have been marked if sampling had been restricted to that specific number of sessions. This figure illustrates this for A) *B. eunomia*, B) *B. aquilonaris* in closed sites and C) *B. aquilonaris* in open sites. Black dots show the observed values (mean + SD) for the datasets (59, 24 and 28, respectively; see **Appendix 1**) containing at least three sampling sessions. Grey curves with dots represent the sigmoid regression curve; note that, for each panel, the number of datasets decreases as the number of sampling days increases, which explains why the best fit curves might not always closely match observed data for high values of sampling sessions, containing a comparatively lower amount of data points. The dotted lines indicate the number of sampling sessions necessary to IF = 80% (arbitrary level chosen for illustration).

Figure 2.



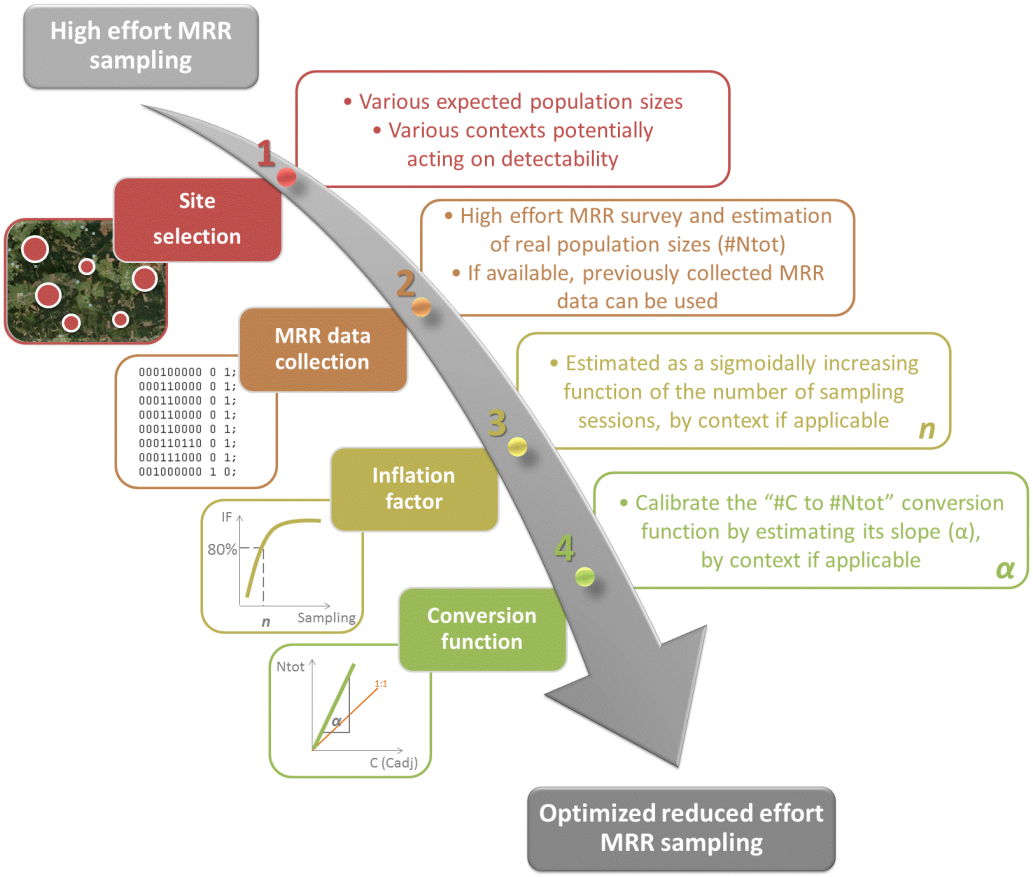
Total population size as estimated on original MRR data (#Ntot with its 95% confidence interval; black dots) and as predicted with the best model (grey dots and grey dashed lines) for A) *B. eunomia*, B) *B. aquilonaris* in closed sites and C) *B. aquilonaris* in open sites, as a function of the number of marked individuals (#C). The solid black line indicates the 1:1 line, i.e. the ideal case where every existing individual would have been marked (detectability = 100%), to illustrate the differences between the three cases in the proportion of missed individuals, reflected in how the slope of the linear regression differs from this ideal case.

Figure 3.



Sensitivity of the estimation of the slope of the conversion function to the sample size (i.e. number of data points from high effort MRR used to estimate the conversion function such as on **Figure 2**) for A) *B. aquilonaris* in closed sites, B) *B. aquilonaris* in open sites and C) *B. eunomia*. Displayed are the median (black line), 25%-75% (dashed black lines), 5%-95% (dotted grey line), minimum and maximum (grey dots) of 100 slope estimates, each obtained on a random downsample of the original dataset (as seen on **Figure 2**) as a function of the sample size. The slope estimation is more variable when the relation is less linear as is the case for *B. aquilonaris* in closed sites (vs *B. eunomia* and *B. aquilonaris* in open sites). Note that this represents a worst case scenario as we did not control for how the data points were spread along the X (sample size) axis.

Figure 4.



Schematic representation of the method to design a simplified, reduced effort MRR sampling scheme to estimate population size. In step 1, several sites should be carefully selected as hosting populations with a range of different expected size (as represented by circles of various sizes in the map) and contexts. In step 2, intensive high effort MRR data are collected on those selected sites and analysed using classical demographic methods to estimate the absolute population size (#Ntot). Those data can of course be complemented by already available (non-)published data. In step 3, MRR data are used to assess the inflation factor and the minimum number (n)of sampling sessions needed to catch a predefined threshold (e.g. at least 80% as illustrated here and on Figure 1) of the possibly marked individuals. In step 4, the slope α of conversion function relating #Ntot to the number of marked individuals (#C or #Cadj) can be obtained. Finally, a reduced effort MRR sampling design can be selected, largely reducing the cost of MRR without sacrificing the quality of population size estimates.

APPENDIX

Appendix 1. List of MRR datasets used for the analyses. See main text for the meaning of the variables.

Species	Population	Year	Type	#Sampling	#C	#Cadj	#CR	#CRadj	#CRmax	#Ntot	95% CI	Used to estimate IF (Fig. 1)	Used to estimate conversion function (Fig. 2)
B. eunomia	Prés de la Lienne	1992	.	18	638	670	1207	1484	150	1047	88	x	x
B. eunomia	Hébronval	1993	.	8	63	80	144	248	34	75	7	x	x
B. eunomia	Hierlot	1993	.	11	36	44	64	103	12	99	66	x	x
B. eunomia	Prés de la Lienne	1993	.	20	360	365	659	726	89	592	60	x	x
B. eunomia	Hébronval	1994	.	14	52	56	101	133	16	84	23	x	x
B. eunomia	Prés de la Lienne	1994	.	26	262	263	702	734	76	348	29	x	x
B. eunomia	Prés de la Lienne	1995	.	17	129	133	268	312	44	188	27	x	x
B. eunomia	Prés de la Lienne	1996	.	11	316	357	926	1334	144	376	24	x	x
B. eunomia	Hébronval	1997	.	8	41	55	78	144	22	64	18	x	x
B. eunomia	Prés de la Lienne	1997	.	16	400	415	1071	1280	163	510	38	x	x
B. eunomia	Prés de la Lienne	1998	.	7	199	287	243	485	124	400	89	x	x
B. eunomia	Bellemeuse	1999	.	12	24	26	69	95	10	29	6		x
B. eunomia	Bérisménil1	1999	.	15	51	54	161	205	28	55	5	x	x
B. eunomia	Chapons	1999	.	16	60	62	189	226	21	.	.	x	
B. eunomia	Mormont	1999	.	14	50	55	141	194	25	61	10	x	x
B. eunomia	Prés de la Lienne	1999	.	18	280	289	722	842	103	404	39	x	x
B. eunomia	Hierlot	2000	.	7	83	112	104	192	33	253	94	x	x
B. eunomia	Prés de la Lienne	2000	.	16	219	233	439	558	63	330	38	x	x
B. eunomia	Bellemeuse	2001	.	14	23	24	57	72	9	33	11		x
B. eunomia	Bérisménil1	2001	.	17	62	64	242	289	22	69	5	x	x
B. eunomia	Chapons	2001	.	12	12	13	30	41	4	14	4		x
B. eunomia	Hierlot	2001	.	10	34	41	49	79	13	70	43	x	x
B. eunomia	Prés de la Lienne	2001	.	18	78	80	179	209	21	119	17	x	x
B. eunomia	Pisserotte	2002	.	16	939	975	2515	3006	258	1301	122	x	x
B. eunomia	Prés de la Lienne	2002	.	12	169	191	415	598	75	232	26	x	x
B. eunomia	Prés de la Lienne	2003	.	20	327	333	1057	1184	112	425	34	x	x
B. eunomia	Pisserotte	2004	.	25	916	921	1969	2073	196	1437	94	x	x
B. eunomia	Prés de la Lienne	2004	.	19	313	319	718	804	101	494	50	x	x
B. eunomia	Pisserotte	2005	.	16	380	389	866	988	109	601	76	x	x
B. eunomia	Prés de la Lienne	2005	.	14	105	113	222	293	47	170	36	x	x
B. eunomia	Pisserotte	2006	.	35	142	157	205	282	34	322	69	x	x
B. eunomia	Prés de la Lienne	2006	.	9	73	88	132	213	23	155	53	x	x
B. eunomia	Pisserotte	2007	.	6	91	131	123	246	32	204	56	x	x
B. eunomia	Prés de la Lienne	2007	.	10	87	101	185	281	34	132	27	x	x
B. eunomia	Pisserotte	2008	.	8	252	320	419	720	111	375	41	x	x
B. eunomia	Prés de la Lienne	2008	.	9	67	81	114	184	37	133	32	x	x
B. eunomia	Troufferies de Libin	2008	.	7	270	363	438	809	91	407	50	x	x
B. eunomia	Bérisménil1	2009	.	8	98	125	153	263	27	230	98	x	x
B. eunomia	Bièvres	2009	.	7	111	149	162	299	55	230	50	x	x
B. eunomia	Grande Fange	2009	.	7	35	50	45	90	16	82	40	x	x
B. eunomia	Mormont	2009	.	6	22	32	42	84	11	28	7		x
B. eunomia	Pisserotte	2009	.	12	529	583	913	1255	181	1040	134	x	x
B. eunomia	Prés de la Lienne	2009	.	12	198	218	347	477	56	308	40	x	x
B. eunomia	Bérisménil1	2010	.	11	103	125	185	298	31	167	29	x	x
B. eunomia	Bérisménil2	2010	.	10	48	58	93	150	20	60	12	x	x
B. eunomia	Bièvres	2010	.	10	157	190	247	398	46	371	120	x	x
B. eunomia	BihainA	2010	.	8	84	113	110	203	20	230	79	x	x
B. eunomia	BihainB	2010	.	7	125	168	166	306	39	321	83	x	x
B. eunomia	Grande Fange	2010	.	12	102	119	171	260	26	176	33	x	x
B. eunomia	Pisserotte	2010	.	12	878	1024	1574	2391	268	1553	174	x	x
B. eunomia	Prés de la Lienne	2010	.	13	535	590	892	1226	217	1015	130	x	x
B. eunomia	Bérisménil1	2011	.	12	136	154	205	295	35	316	68	x	x
B. eunomia	Bérisménil2	2011	.	8	22	30	39	72	11	29	4	x	x
B. eunomia	Bièvres	2011	.	12	195	215	340	467	48	354	55	x	x
B. eunomia	BihainA	2011	.	8	50	64	75	129	15	111	37	x	x
B. eunomia	BihainB	2011	.	12	121	141	186	283	34	243	49	x	x
B. eunomia	Grande Fange	2011	.	9	188	228	268	432	59	393	68	x	x
B. eunomia	Langlire B	2011	.	8	30	38	35	60	8	.	.	x	
B. eunomia	Mormont	2011	.	10	35	44	49	84	12	75	32	x	x
B. eunomia	Pisserotte	2011	.	15	720	756	1353	1664	350	1164	80	x	x
B. eunomia	Prés de la Lienne	2011	.	14	747	794	1416	1799	190	1238	102	x	x
B. eunomia	Pisserotte	2012	.	9	258	313	466	751	78	394	41	x	x
B. eunomia	Prés de la Lienne	2012	.	9	116	141	218	351	40	195	38	x	x
B. aquilonaris	Mirenne	1995	closed	22	477	477	920	923	119	791	64	x	x
B. aquilonaris	Commanster	1996	closed	11	114	125	227	276	37	208	63	x	x
B. aquilonaris	Mirenne	1996	closed	11	148	162	215	261	53	311	83	x	x
B. aquilonaris	Quatre Vents	1996	closed	9	41	47	79	102	14	70	34	x	x
B. aquilonaris	Arbrefontaine	1997	closed	9	53	64	84	118	18	96	21	x	x
B. aquilonaris	Commanster	1997	closed	9	92	100	159	193	36	148	28	x	x
B. aquilonaris	Logbiermé	1997	closed	10	112	119	248	287	37	164	24	x	x
B. aquilonaris	Mirenne	1997	closed	12	128	132	260	282	40	222	44	x	x
B. aquilonaris	Pisserotte	1997	closed	13	254	261	380	412	57	534	80	x	x
B. aquilonaris	Quatre Vents	1997	closed	9	42	48	112	145	23	53	10	x	x
B. aquilonaris	Pisserotte	2004	closed	16	49	51	80	89	15	101	32	x	x
B. aquilonaris	Pisserotte	2005	closed	9	75	97	95	148	30	170	57	x	x
B. aquilonaris	Pisserotte	2006	closed	9	127	139	151	184	35	545	265	x	x
B. aquilonaris	Mirenne	2009	closed	8	113	128	157	204	37	294	76	x	x
B. aquilonaris	Commanster	2010	closed	7	271	326	393	554	102	512	87	x	x
B. aquilonaris	Mirenne	2010	closed	8	154	175	189	245	34	443	132	x	x
B. aquilonaris	Pisserotte	2010	closed	7	82	99	107	151	26	189	60	x	x
B. aquilonaris	Quatre Vents	2010	closed	6	29	38	38	59	11	.	.	x	
B. aquilonaris	Arbrefontaine	2011	closed	7	109	131	149	210	46	221	51	x	x
B. aquilonaris	Commanster	2011	closed	8	268	305	412	535	101	455	54	x	x
B. aquilonaris	Logbiermé	2011	closed	7	139	167	199	280	65	236	42	x	x
B. aquilonaris	Mirenne	2011	closed	10	332	353	432	500	88	1009	259	x	x
B. aquilonaris	Pisserotte	2011	closed	8	77	93	98	138	20	206	71	x	x
B. aquilonaris	Quatre Vents	2011	closed	8	88	106	118	166	27	225	72	x	x
B. aquilonaris	Grande Fange	1995	open	20	524	537	728	777	141	1692	377	x	x
B. aquilonaris	Grande Fange	1996	open	8	299	432	381	599	87	1065	339	x	x
B. aquilonaris	Grande Fange	1997	open	15	309	342	395	472	66	1227	419	x	x
B. aquilonaris	Massotais	1997	open	16	679	736	906	1054	154	1976	289	x	x
B. aquilonaris	Mochettes	1997	open	7	134	181	154	228	51	.	.	x	
B. aquilonaris	Crépale	2009	open	7	453	611	540	798	144	1662	407	x	x
B. aquilonaris	Grande Fange	2009	open	7	225	304	266	393	72	773	199	x	x
B. aquilonaris	Massotais	2009	open	7	33	45	34	50	10	.	.	x	
B. aquilonaris	Mochettes	2009	open	6	111	160	121	190	36	.	.	x	
B. aquilonaris	Nazieufa	2009	open	6	50	72	57	90	20	.	.	x	
B. aquilonaris	Robiéfa	2009	open	7	78	105	88	130	24	.	.	x	
B. aquilonaris	Sacrawé	2009	open	7	79	107	86	127	18	.	.	x	
B. aquilonaris	Bovigny	2010	open	6	120	173	140	220	42	.	.	x	
B. aquilonaris	Crépale	2010	open	7	689	930	830	1227	185	2188	325	x	x
B. aquilonaris	Grande Fange	2010	open	11	579	657	747	922	122	1765	273	x	x
B. aquilonaris	Massotais	2010	open	6	146	211	149	234	65	.	.	x	
B. aquilonaris	Mochettes	2010	open	6	316	456	339	533	92	.	.	x	
B. aquilonaris	Nazieufa	2010	open	7	123	166	137	203	39	.	.	x	
B. aquilonaris	Robiéfa	2010	open	7	133	192	162	255	39	374	123	x	x
B. aquilonaris	Sacrawé	2010	open	6	146	211	244	384	68	.	.	x	
B. aquilonaris	Bovigny	2011	open	8	161	205	203	284	50	444	119	x	x
B. aquilonaris	Crépale	2011	open	9	685	925	798	1180	169	2482	439	x	x
B. aquilonaris	Grande Fange	2011	open	14	703	750	923	1049	118	2276	341	x	x
B. aquilonaris	Massotais	2011	open	6	90	130	91	143	48	.	.	x	
B. aquilonaris	Mochettes	2011	open	7	247	333	254	376	110	.	.	x	
B. aquilonaris	Nazieufa	2011	open	6	91	131	102	160	35	.	.	x	
B. aquilonaris	Robiéfa	2011	open	7	81	109	98	145	26	144	23	x	x
B. aquilonaris	Sacrawé	2011	open	7	193	260	214	316	56	.	.	x	

Appendix 2. Rationale for fixing the intercept to zero when estimating the conversion function.

In case of limited catch effort, a very small number of (re)captures will be recorded / of individuals. This is whatever the population size, because even a very small population can be sampled efficiently (even in totality in theory) by MRR provided an adequate catch effort.

If the conversion function is estimated only on such data with limited catch effort, we might expect that the intercept of the conversion function is higher than zero because the population size will indeed be higher than 0 even with very low numbers of marked individuals. In that context, why should the intercept be forced to 0 when estimating the conversion function, as we did in this study? The rationale is twofold:

1. For the conversion function to be reliably used for prediction, some datasets with a reasonable to high catch effort are needed in order to have reliable estimates of real Ntot. In such case, fixing the intercept to 0 does not alter the estimate of the slope because the effect mentioned above vanishes.
2. Failure to force the intercept to 0 might even create an additional bias. The best regression line is mainly determined by the fit to points with high #C (contributing proportionally more to the residual sum of squares). This may lead the line to poorly fit points with low #C values and the intercept to be estimated as lower than 0. Using such a conversion function will lead to predictions of population size that are negative for datasets with few individuals marked (low #C values), which is obviously impossible.

Considering that in theory the intercept should be 0, and that there is a danger to estimate it as a negative value, we advocate fixing the intercept to 0. More complex procedures to fit the conversion function could be used, such as constraining the intercept to be >= 0, using other regression methods to avoid data points with higher #C values have a higher weight in the estimation of parameters. Estimating a non-zero intercept, provided it makes sense for the study system, is a simple adaptation of our protocol that may surely be part of the application process that end users should do according to the specifics of their own study system.