UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Ecole de biologie

MODELLING TRENDS IN OCCUPANCY WHEN DETECTION IS IMPERFECT: A CASE STUDY ON NEWTS

Travail de Maîtrise universitaire ès Sciences en comportement, évolution et conservation
Master Thesis of Science in Behaviour, Evolution and Conservation

par

Danilo Foresti

Directeur : Prof. Nicolas Perrin
Superviseur (s) : Dr. Benedikt Ramseier Schmidt
Expert (s) : Dr. Jérôme Pellet
Département d'Écologie et Évolution

Janvier 2012

Abstract

Monitoring programs are a pillar for the conservation of species, but few of them have been active for decades. Citizen-science projects and volunteer-based observations are now able to complement the former limitation of monitoring programs, thanks to the state-of-the-art techniques for the analysis of presence-only databases relying on site occupancy models. Nevertheless, their effectiveness has mostly been demonstrated on databases and time periods relatively abundant in observations. I devised a novel application of site occupancy models to exploit the sparser information contained in most presence-only databases and early years of sampling. I illustrated the method with the Swiss amphibian database and chose the five species of newts (former genus *Triturus*) as model organism. I tested the capacity of the novel application to describe the trends in occupancy over the last three decades and additionally investigated the effect of a small set of environmental variables (connectivity, elevation, urban cover) on the dynamic parameters driving the distributional changes. The resulting estimations of occupancy trends proved to be imprecise; furthermore the expectations about the impact of the environmental variables were only partially met. Altogether, the extreme sparseness of the data represented a major limitation when describing the trends or inferring their causes, since the assumptions required by the method proposed heavily weigh on the performances of site occupancy models. Nevertheless, the results highlighted how the spatial variability of the detection probability - and not only the temporal variability - can negatively bias the inference, an issue rarely addressed in the analysis of presence-only databases. I therefore suggest accounting for this heterogeneity in future analyses of trends derived from volunteer-based observations.

Introduction

The earliest concerns about a possible decline of the amphibian populations throughout the world were expressed in 1989 at the First World Herpetology Conference in Canterbury, even though the process was already underway since many decades (Houlahan *et al.* 2000). The global amphibian assessment describing what is actually called the "global biodiversity crisis" showed that 32.5% of the known amphibian species are listed as threatened, a proportion far greater than the proportions of birds or mammals (Gardner 2001; Stuart *et al.* 2004). Monitoring programs are increasingly employed for quantifying the biodiversity and its driving factors through the study of the distribution and the abundance of species. The monitoring activity is the basis for the coordination of adequate management actions and the definition of priorities for conservation and land use, taking an active role in this process (Yoccoz *et al.* 2001; Nichols and Williams 2006). Indeed, management-oriented monitoring schemes are powerful tool thanks to which relevant hypotheses for conservation are tested and directly related to the decisions of policy-makers (Weber *et al.* 2004). This kind of study does not only uncover the factors that affected the system in the past, but also provides reliable results useful for the planning of future actions. Unfortunately, the number of monitoring programs having collected data for long time periods such as decades is low and mostly limited to bird schemes (Houlahan *et al.* 2000; Schmeller *et al.* 2009). The absence of long time data series poses a great problem, since the assessment of the biodiversity status in the past, as well as its evolution toward the present, represent a key step. For instance, the main target declared by the Governments for the Biodiversity Year was "*to achieve by 2010 a significant reduction of the current rate of biodiversity loss at the global, regional and national level* ..." (Secretariat of the Convention on Biological Diversity, 2010). This means that data describing at least three time points are required to evaluate whether the loss has effectively slowed in recent time. Although this seems an elementary relationship to test, in practice it is difficult to do. The fundamental issue therefore is: can we reconstruct the past with the knowledge and the data we have at our disposal today?

Given that long time series from monitoring activities are limited, data from complementary schemes covering the desired geographic and temporal range of inference are welcomed. The most important depositories of past biodiversity and species distribution correspond to “presence-only” databases and museum collections, containing thousands of observations going back decades in time (Lister 2011). The past datasets can be complemented by the always increasing number of records collected by naturalists and citizen-science projects in the recent years, thanks to the facilities of the web data entry (Kery *et al.* 2010b; van Strien *et al.* 2010; Wood *et al.* 2011). All these records often cover entire taxonomic groups of valuable interest (e.g. birds, herpetofauna and butterflies) and were collected following a scheme adapted to the question addressed by the observer. Thus, they can be qualified as “non-standardized”, since they are the result of multiple sampling protocols (if there are any). Drawing population trends from non-standardized data is plagued by the problem of variable observation effort, a factor that is overall increased in recent time because of the recruitment of volunteers in citizen science projects of biodiversity monitoring (Houlahan *et al.* 2000; Schmeller *et al.* 2009; Wood *et al.* 2011). Not accounting for temporal trends in the observation effort may easily lead to wrong inferences and mere sampling artefacts (van Swaay 1990; Kery and Schmidt 2008; van Strien *et al.* 2010), therefore non-standardized sources got little attention in the past to derive quantitative population trends (but see Shaffer *et al.*, 1998). Nevertheless, thanks to the recent advances in the analysis of non-standardized data, it is now possible to take them into account, since the former limitations have been overcome. The basic idea is that the changes in the observation effort can be quantified by the detection probability, allowing the recovery of the true parameters of species distribution and abundance (Kery *et al.* 2010b). However, the approach developed by Kery *et al.* (2010) performed on a database relatively dense in observations and did not try to investigate the decades with the sparse data (i.e. previous to 1990), the time when the bulk of the declines occurred for many *taxa*. For instance, Houlahan *et al.* (2000) estimated that an acceleration of the negative population trends for European amphibians took place around year 1964. Similarly, the reduction of area occupied by the tree frog *Hyla spp.* in Switzerland happened mostly before 1960 (Grossenbacher 1988). A reconstruction of the population trends in a relatively distant past therefore represents an interesting challenge, since the corresponding data are much scarcer than today and yet may carry valuable information about the state of populations.

In this study, I employ state-of-the-art methods for inferring population trends in occupancy from the non-standardized observations contained in the Swiss amphibian database. In order to estimate long-term trends in occupancy, I consider the recent years where there is a lot of data and explicitly integrate into the analysis the past years holding the sparser information. At first, I explain the devised methodological devices. Afterwards, three case studies, relying on different strategies, are going to be developed in order to evaluate the trends in occupancy of the five newt species (former genus *Triturus*) that occur in Switzerland. Two case studies describe the dynamics in occupancy without inferring any causal relationship, whereas the third one also evaluates the impact of three elementary variables known to determine the amphibian distributions: connectivity, elevation and urbanization. Finally, I discuss the implications of the analysis of highly sparse data, as well as the biological conclusions that can be drawn from such study.

Methods

Data source and site occupancy models

The karch (Koordinationsstelle für Amphibien- und Reptilienschutz in der Schweiz) is the official institution in Switzerland concerning the coordination of measures for the protection of autochthone amphibians and reptiles since 1979 (Meyer *et al.* 2009). Among other tasks, karch is in charge to collect, store and analyse the geographical information about all amphibians and reptiles in Switzerland. The data contained in the database are coded following a "presence-only" format and come from various sources, ranging from well-designed monitoring programs to volunteer observations transmitted by naturalists. Sampling effort increased in recent years because of the establishment of permanent amphibian monitoring programs (e.g. Aargau canton in 1999) and the advent of the web-based data submission for the general public.

An analysis that does not take the observation effort into account is not advised, since artificial trends in occupancy may appear or existing trends be masked as a result of trends in the effort (Kery and Schmidt 2008). However, it is possible to account for sampling effort, since it can be described by the detection probability of the target species. In other words, the ease with which an organism is found during a field trip is directly linked to the effort invested when looking for it. Presence-only databases do not provide directly the information about detectability, but Kery *et al.* (2010b) developed an approach to estimate it and subsequently correct population trends in occupancy from presence-only databases. Specifically, they exploited a novel formulation of site occupancy models, a category of models derived from the theory underlying mark-recapture studies (MacKenzie *et al.* 2002; MacKenzie *et al.* 2003). These methods integrate separate analyses about the true occupancy status and the uncertainty in its assessment caused by imperfect detection (see Appendix 1 for an introduction to occupancy modelling). The data required by site occupancy models are detection-nondetection events. Presence-only databases directly provide the detection data, while nondetections (i.e. the "zeroes") need to be recreated from a presence-only source. Assuming that observers report all the species detected during a field trip, one can declare the nondetection of a particular species if the latter is not reported on the list. Crossing the records of the multiple species contained in the database allows reconstructing time, location on which the visits took place and the originals lists of detection. Note that in this situation, nondetection does not mean unequivocally absence from a given site, exemplifying the concept of the species detectability linked to the observation effort. Site occupancy models require additional information to distinguish between the nondetection events due to imperfect detection and those arising from the true absence of the species. First, of all, an assumption of closed population over a time period must be stated, during which the occupancy status of each site is not allowed to change (either occupied or not occupied). Second, repeated visits within these periods are required: these replicated sampling occasions allow estimating the detection probability. For example, if a site was visited twice and the species was detected once, then we can infer that detection probability was 0.5. The time frame during which the occupancy status does not change is also called primary period, the repeated visits within the primary period name secondary periods (see Appendix 1 for details).

Study species

This case study is based on the taxon of the Swiss newts (*Salamandridae*, former genus *Triturus*). The five species occurring in Switzerland differ with regards to ecological traits and commonness. All but one species are listed in the Swiss amphibian Red List and for decades have been facing severe reductions of population size, area of repartition or both (Schmidt and Zumbach 2005).

The alpine newt *Ichtyosaura alpestris* (LAURENTI, 1768, Least Concern) is the most common newt and can occur at any altitude and type of pond (Meyer *et al.* 2009). The palmate newt *Lissotriton helveticus* (RAZOUMOWSKY, 1789, Vulnerable) is ecologically similar to *I.alpestris* and finds its optimum in alluvial ponds. Despite its relative commonness, it is less frequent than *I.alpestris* and occurs only in the northern part of the country (Grossenbacher 1988; Meyer *et al.* 2009). The smooth newt *Lissotriton vulgaris* (LINNAEUS, 1758, Endangered) is a low altitude, rare species occurring with its nominal form in the north and with the subspecies *L.v.meridionalis* in the southern country. Overall, *L.vulgaris* has been reported in less than 300 ponds over the whole country (Meyer *et al.* 2009). The crested newt

Triturus cristatus (LAURENTI, 1768, Endangered), once widespread in the lowland regions, is the rarest species of Switzerland and is experiencing a massive and partly unexplained decline (Grossenbacher 1988; Meyer *et al.* 2009). The Italian crested newt *Triturus carnifex* (LAURENTI, 1768, Endangered) is a southern species ecologically similar to *T.cristatus*. Even if this species is declining in its native range, it represents an invasive species in the northern country surrounding Geneva (Meyer *et al.* 2009).

Data preparation

The present study based all the inferences on square sampling units. I chose three resolutions and performed the analyses in the same way for each of them. The finer resolution corresponded to a 1 square kilometre quadrat. This choice was dictated by the overall geographical precision of the observations contained in the database, representing a good compromise between biological scale and replication of visits. Two further resolutions were investigated, corresponding to quadrats of side three kilometres (9 km^2) and five kilometres (25 km^2). The database records not reaching the geographical accuracy of 1 kilometre were discarded from the analyses at all resolutions. The attribution of the observations to their relative quadrat was performed in ArcGIS 9.3 (ESRI 2008). For clarity, the formal notation of this manuscript will always indicate a particular resolution by the length of the quadrat side.

I retained data from three time frames defining three primary periods: from 1979 to 1985, from 1992 to 1997 and from 2003 to 2008. Within these periods I assumed the status of each quadrat to be unchanged (either occupied or not occupied), yet transitions were allowed between them. The assumption of multi-year closed populations aimed at increasing the number of visits within the periods to model the detection probability, but it may not be respected due to the temporary absence of the species. This shifts the interpretation of occupancy parameters from the classic “Proportion of Area Occupied” by the species during a primary period to the less stringent “Proportion of Area Used”, the resulting unavailability for detection thus becomes one component of imperfect detection (Kery and Schaub 2011).

Hence, colonization and survival rates describe changes in the amount of quadrats used, rather than in the amount of quadrats permanently occupied by the target species.

I created detection-nondetection histories for each quadrat and period of closure following the procedure illustrated by Kery *et al.* (2010b) for presence-only databases. In this process, I considered all the pond-breeding amphibian species as background information to extrapolate the date and place of the visits. Only data from visual encounters were used when reconstructing the detection non-detection matrices. The information provided by day, month, year and identity of the observer has been exploited to deduce the original lists of detection (see Appendix 2 for the full list of species considered and a schematic example of the procedure).

The use of data from different time periods implay a supplementary challenge to deal with, since the date of old observations often missed the day and the month, causing the year to be the only information available (Figure 1). For the periods previously defined, the records with incomplete reference respectively amounted to 84%, 25% and 7%. The absence of day and month prevented the distinction of the original visits within a year, therefore underestimating the true number of visits on which the detection probability is modelled. Since dropping the incomplete records would have led to a large data loss in the period 1979-85, I decided to retain all the records and consider the missing values as unknown, constant quantities. Starting from this point, I devised a new method aiming to recover the true parameter estimates. I enlarged the definition of the secondary period aggregating the observations in each quadrat to the whole year. The secondary periods were therefore generously defined with regards to the visits carried in the field, creating what could be view as a “one-year-survey”. The replication needed for estimating the detection probabilities was provided by the multiple years included in the primary periods. Following the newly defined framework, the number of secondary periods per primary period of closure cannot exceed seven (period 1979-1985) or six (periods 1992-97 and 2003-08). Given that volunteers do not follow a sampling protocol stating that quadrats should be visited every year, the number of secondary periods for a quadrat could vary from zero (quadrat not visited during the primary period considered) up to 6 or 7 (quadrat visited every year). The particular "quadrat x year" combinations that did not carry amphibian observations were qualified as "missing data".

I derived from the database three promising covariates to model the detection probability: (i) the number of records within a quadrat in a year (i.e. the particular data employed for declaring a detection-nondetection event). I discarded the records of the target species previous to calculation in order to avoid the problem of circularity, since it is not admissible to use covariates that rely on the detection of the species under study (MacKenzie *et al.* 2002). Similarly, I used (ii) the number of distinguishable visits per quadrat and year and (iii) the number of different visitors per quadrat and year. The three covariates were used to describe the sampling effort. Indeed, detection-nondetection events declared on few records were likely to be less accurate, because arising from highly incomplete lists of observation (van Strien *et al.* 2010). The same way, a lower number of visits or visitors were supposed to decrease the detection probability. Given that the covariate (i) "number of records" allowed the value of 0 (detection event declared from a single record), all the variables were $log(x + 0.5)$ transformed (Box and Cox 1964). Prior to analysis, all the covariates were standardized to have mean 0 and standard deviation 1 (Kery 2010).

Data analysis

I used the hierarchical site occupancy model formulated by Royle and Kery (2007) to analyse the occupancy dynamics of the five newt species at three different resolutions over the three defined time periods. Dynamic site occupancy models assume that the changes in occupancy between the primary periods occur according to two transition parameters, namely survival and colonization probability. In other words, the occupancy status for quadrat i during period k depends on the occupancy status during the period $k-1$. The occupancy probability of quadrat i during primary period k therefore corresponds to the probability to get colonized if not occupied in period *k-1*, added to the probability to go extinct if occupied in period *k-1* (see Appendix 1).

I performed three case studies exploiting the same dynamic model formulation. The case studies lay on different sets of quadrats and answered to slightly different questions. Case studies 1 and 2 were complementary and described the trends in occupancy. Case study 3 aimed at identifying the factors driving the use of quadrats. For this, I integrated a small set of covariates into the ecological model. In every case study, the simpler model for detection was composed by three constant, time specific parameters (i.e. one value for each primary period, constant across all the quadrats). The simpler model describing the occupancy dynamic considered five constant, time specific vital parameters, namely: initial occupancy ψ, survival rate from period 1979-85 to period 1992-97 $\Phi 1$, survival rate from period 1992-97 to period 2003-2008 $\Phi 2$, colonization rate from period 1979-85 to period 1992-97 $\gamma 1$, colonization rate from period 1992-97 to period 2003-2008 $\gamma 2$. The previous model can be described as $p(t)\ \psi(.)\Phi(t)\gamma(t)$, its extended form corresponds to $p1(.)p2(.)p3(.)\ \psi(.)\Phi 1(.)\Phi 2(.)\gamma 1(.)\gamma 2(.)$.

I fitted the models in WinBUGS 1.4 (Spiegelhalter *et al.* 2003) using the software R 2.13 (R Development Core Team 2010) through the package R2WinBUGS (Sturtz *et al.* 2005). The Bayesian framework was preferred since maximum likelihood methods often fail to converge with large amounts of missing secondary periods (Royle and Dorazio 2008). I followed a forward stepwise model selection to determine the best detection structures (Govindan *et al.* 2011). Covariates were therefore added one by one to the null model ($p(t)\ \psi(.)\Phi(t)\gamma(t)$) and retained if the 95% credible interval of the regression slope did not contain the value 0 (2.5% and 97.5% quantiles of the posterior distribution). This process was repeated for all the three case studies. I employed non informative, uniform priors for all model parameters. For each fitted model, I ran three parallel Markov chains always discarding the first quarter as burn-in and thinned the result by 3. I assessed the convergence of the model parameters and derived quantities in every run using the *Rhat* statistics (Gelman and

Hill 2007, *Rhat* < 1.1). The number of iterations necessary to achieve the convergence of the best detection model differed for each species: usually 6000 up to 30 000 iterations were sufficient.

Case study 1 - Inferring occupancy trends from volunteer observations

The goal of this case study was to derive the trends in occupancy for the three time periods. In the present case study, I applay a variant of the potential breeding range, a basic site selection strategy proposed to partly mitigate the impact of geographic heterogeneous sampling among primary periods (Kery *et al.* 2010b). Only the quadrats known to have ever hosted the target species were selected, defining a set of quadrats that will hereafter be called "potential range". The whole karch database (from year 1970 to 2010) was analysed to recreate the potential range of each species. This way, the analysis was conditioned only on the quadrats having demonstrated at least once the suitability for the species. Note that the occupancy estimations obtained from the use of the potential range are not referable to the commonness of the species: rare species may show occupancy probabilities higher than the common species. The initial occupancy, colonization and survival rates are assumed as constant terms across the study region following the ecological model $\psi(.)\Phi(t)\gamma(t)$. The number of quadrats forming the potential range for every species and resolution is reported in Table 1.

Case study 2 - Inferring occupancy trends on permanently surveyed quadrats

The goal of this case study was to derive trends in occupancy for the three time periods without any bias due to geographic variation in the location of the visits. Indeed, a fraction of the sites forming the potential range was surveyed during only one or two periods out of the defined three. Hence, the temporal trends in occupancy of case study 1 are to some unknown extent confounded with geographic variation. In this case study, I restricted the analyses to the quadrats that were visited at least once in every primary period. I further reduced the selection selecting the quadrats belonging to the biogeographical extent of the target species. This extent was determined as the sum of all the biogeographical regions known to have ever hosted the species (Geostat, OFS). Note that the interpretation of occupancy probabilities differs from the previous case study: in this case, values equal to 100% mean that the species occurs in all the constantly visited quadrats belonging to its regions of occurrence. This way, the rare species will obligatorily show occupancy probabilities lower than the common species. Similar to case study 1, no covariates were added to the ecological model $\psi(.)\Phi(t)\gamma(t)$. The number of quadrats forming the extent for every species and resolution is reported in Table 2.

Case study 3 - Modelling the vital parameters with environmental covariates

The aim of this case study is to test a small set of environmental covariates on the vital parameters describing the occupancy dynamic of the five species. I decided to base the inference on the whole set of quadrats belonging to the biogeographical extent of the species in order to avoid problems linked to the use of a restricted range of the predictor variables (Eigenbrod *et al.* 2011). Values of occupancy equal to 100% mean that the target species occurs in the totality of quadrats having ever reported amphibian observations. The interpretation of occupancy probabilities is therefore similar to case study 2: rare species will obligatorily show lower values than the common species (for the same biogeographical region). The number of quadrats for every species and resolution is shown in Table 3. From the wide literature describing the ecology of the newts I selected the following covariates: elevation, urbanization and connectivity.

Elevation is likely to represent the best proxy explaining the distribution of species at large spatial domains in Switzerland (Kery *et al.* 2010a). Moreover, it already proved to be hardly replaceable by other covariates when building good descriptor models for the European newts (Cirovic *et al.* 2008). I therefore added this covariate into the ecological model on the initial occupancy probability. Both a linear (ALT) and quadratic term (ALTSQ) were considered, thus allowing for a peak in occurrence.

The distribution and abundance of newts is also known to respond to the degree of urbanization neighbouring the ponds (Van Buskirk 2005; Denoel and Ficetola 2008; Hartel *et al.* 2010). I therefore calculated the portion of the quadrat covered by this particular land use and tested if the urban cover had an effect on the survival or the colonization capacity of the target species. The information from the Geostat statistic of years 1992-97 (Humbel 1998; OFS 2007) has been integrated in the transition from period 1979-85 to period 1992-1997 (URBAN1); the information from the Vector 25 layer (Swisstopo 2007) has been integrated in the transition from period 1992-1997 to period 2003-08 (URBAN2). I did not test the effect of urbanization on the initial occupancy probability, because urbanization was found to be highly correlated to elevation. (see Appendix 3 for details about the preparation of land use covariates).

Spatial autocorrelation is often encountered in ecological data and causes overestimations in the importance of habitat characteristics when not properly addressed (Zanini *et al.* 2009). Measures of patch connectivity are derived from the metapopulation theory and ensure the correct treatment of non-independent observations building an autologistic model (Augustin *et al.* 1996; Zanini 2006). I therefore computed the connectivity over the whole study area using the equation

$$CONNECT_{ik} = \frac{\sum_{i \neq j} e^{(-\propto d_{ij})}\, y_{ik}}{\sum_{i \neq j} e^{(-\propto d_{ij})}}$$

where i denotes the focal quadrat, j the neighbour quadrat, k the time period and d_{ij} the Euclidean distance between the centres of the two quadrats. The scaling parameter α in the negative exponential kernel corresponds to the inverse of the mean migration distance. I assumed a value of α^{-1} = 400 metres for all the species indiscriminately (Joly *et al.* 2001; Hartel *et al.* 2010). The observed occupancy status y_{ik} (presence/absence) is usually employed in the formula, but in this study it was not acceptable: observed absences may occur in occupied quadrats because of imperfect detection and not all the quadrats were visited in the three periods. I therefore replaced the occupancy status y_{ik} for all the quadrats in each period with the product $\psi_k F_{ik}$. The first term (ψ_k) corresponds to the occupancy probability for period k estimated under the best detection model (p(covariates) ψ(.)Φ(t)γ(t)). The second term (F_{ik}) corresponds to the probability of nondetection, given occurrence. The probability of nondetection F_{ik} in the occupied site i after n visits during period k corresponds to the product (Pellet and Schmidt 2005)

$$F_{ik} = \prod_{n=1}^{n} (1 - p_{ikn})$$

I derived the detection probabilities employed for the estimation of the F values from the linear formula describing the best detection model. Trivially, F_{ik} for unsurveyed quadrats is equal to 1 and F_{ik} for quadrats with at least one positive detection event is not estimated since the occupancy status is known (the value 1 replaces the product $\psi_k F_{ik}$). The full ecological model fitted with the previously described covariates was

ψ(CONNECT1, ALT, ALTSQ)
Φ1(CONNECT2, URBAN1) γ1(CONNECT2, URBAN1)
Φ2(CONNECT3, URBAN2) γ2(CONNECT3, URBAN2)

The ecological model was fitted as a whole without a stepwise strategy, once the best detection model was found. Prior to analysis, I log-transformed the elevation covariate (original unit: meters a.s.l) and square-root transformed the covariates URBAN1 and URBAN2. Since Geostat and Vector 25 did not rely on the same format, I had to specify independent regression slopes for every transition (i.e. two slopes for colonization, two slopes for survival). On the contrary, I assumed that the impact of connectivity on survival and colonization was the same for both transitions (i.e. one slope for survival, one for colonization).

Results

Distribution of records

The visual inspection of the spatial arrangement of the records contained in the whole database revealed that the three periods partially differed in the visited quadrats (Figure 2). The period 1979-85 was characterized by an overall widespread sampling in the central and eastern part of the northern country, with the other regions discretely surveyed. The time frame 1992-97 yielded a lower number of visited quadrats, concentrating the visits to few large clusters mainly distributed in the northern country. Noteworthy, a striking decrease in sampling was observed in the extreme north-eastern part of the country. The period 2003-08 was the one carrying the most completes sampling over all the regions: the non-visited quadrats were fewer than in the two previous periods and were mostly limited to the mountainous areas of the southern and south-eastern parts of the country. The absence of observations in the north-eastern country pointed out in period 1992-97 was partially compensated, but still persisted in the period 2003-08 (Figure 2). The described variations were also certifiable in the distribution of surveys within the potential ranges of the five species (case study 1). Consider the case of *L.helveticus*, where the 3km-quadrats visited during only one time period regarded mainly the north-east for period 1979-85, whereas those from the western country belonged to the period 2003-08 (Figure 3, see Annexes Figure 1 for the other species). The distribution of the quadrats visited during the three time frames roughly reflected the spatial arrangement previously observed for the period 1992-97, with quadrats mostly located in few large clusters in the northern part of Switzerland (Figure 4).

The three case studies lay on three different strategies in the selection of quadrats; yet they did not yield substantially different scenarios with respect to the temporal distribution of the records within the primary periods. The same is valid for the five different species, I therefore report the detailed results only for *L.helveticus*, since it represents an intermediary case between rare and common species. According to case study 1, the karch database contained observations for 4690 quadrat years in the potential range of *L.helveticus* at 1km resolution. They represented 15.4% of all the possible quadrat year combinations (number of quadrats x 19 years of the study). The period 1979-85 and 1992-97 were those carrying the highest amount of quadrats visited during only one year or not at all, nevertheless some degree of replication was allowed (Figure 5). The fraction of quadrats with two or more years visited (i.e. the secondary periods to model the detection probability) increased over time and corresponded to 7%, 13%, 43% respectively to the primary periods. As expected, increasing the quadrat size led to more replicated secondary periods and a decreased number of quadrats, because of the aggregation of nearby patches. At 3km there were records for 4548 quadrat years (27.5% of all possible combinations), the fraction of quadrats visited more than in two years corresponded to 27%, 30% and 67% respectively. At 5km, the quadrat years were 3840 (39% of all possible combinations), the fraction of quadrats visited more than in two years 42%, 48% and 84% respectively (see Annexes Table 1 for all species and resolution).

Case study 1 – Inferring occupancy trends from volunteer observations

For all species and resolutions, the stepwise strategy always included at least one covariate in the detection model. The selected covariates positively affected the detection probability, except for *T.carnifex* (negative / positive: nrecords / nvisits, all resolutions), *L.helveticus* (negative / positive: nvisitors / nrecords, nvisits at 3km) and *L.vulgaris* (negative / positive: nvisits / nrecords at 1km and 5km) (Annexes Table 2). The comparison between the null ($p(t)\psi(.)\Phi(t)\gamma(t)$) and the best ($p(covariates)\psi(.)\Phi(t)\gamma(t)$) model indicated that the occupancy estimates differed for *I.alpestris* (3km and 5km) and *L.helveticus* (3km and 5km), with the best model yielding the highest values. Despite this, the inferred population trends were overall similar between null and best models for both species.

The raw proportion of occupied sites was always lower than the occupancy estimated by the models, except for *T.carnifex* (Figure 6). Moreover, poor correspondence was found between the raw and the estimated occupancy trends, indicating that not accounting for the observation effort could lead to substantially different diagnoses. For example, the raw proportion of quadrats occupied by *T.cristatus* at 3km increased over time, whereas the models assessed that the occupancy probability during the period 1979-85 was much higher than observed; the species was therefore judged as declining once corrected for the observation effort. From a general point of view, the population trends for every species were similar among resolutions. Departures from a stationary distribution on the basis of the geometric mean growth rate were found for three species out of five. Two species were reported as increasing in occupancy, that is *T.carnifex* (all resolutions) and, in a less extent, *I.alpestris* (5km). On the other side, *T.cristatus* was assessed as declining (resolutions 3km and 5km) (Table 4). Concerning the estimation of the transition parameters, it is noteworthy that their accuracy varied considerably and in general was higher for the survival rates than for the colonization rates (Table 5).

Case study 2 - Inferring occupancy trends on permanently surveyed quadrats

For all species and resolutions, the stepwise strategy always included at least one covariate in the detection model. The selected covariates positively affected the detection probability, except for *T.carnifex* (negative / positive: nrecords / nvisits, at 3km and 5km), *L.helveticus* (negative / positive: nvisitors / nrecords, nvisits at 3km) and *L.vulgaris* (negative / positive: nvisits / nrecords at 1km and 5km) (Annexes Table 3). The comparison between the null ($p(t)\psi(.)\Phi(t)\gamma(t)$) and the best ($p(\text{covariates})\psi(.)\Phi(t)\gamma(t)$) model indicated that the occupancy estimates differed for *I.alpestris* (3km and 5km) and *L.helveticus* (3km and 5km), with the best model yielding higher occupancy estimations for occupancy than the null model. Despite this, the inferred population trends were overall similar between null and best models for the five species (Figure 7).

The visual inspection of the trends in occupancy revealed that there was little correspondence between the different resolutions of three species (*I.alpestris, T.carnifex* and *L.helveticus*), thus indicating an influence of the quadrat size on the results. For example, the occupancy probability of *L.helveticus* seemed to have experienced a constant increase at 1km, a steady state at 3km and a peak in the period 1992-97 at 5km (Figure 7 and Annexes Figure

3). Overall, significant departures from a stationary population on the basis of the geometric mean growth rate were found for four species: three species increased in occupancy and one species decreased. Geometric mean growth rates (with 95% credible intervals) for the significant cases were estimated at 2.22 (1.49-2.87) and 2.60 (1.70-3.32) for *T.carnifex* (3km and 5km, respectively), 1.01 (1.00-1.03) for *I.alpestris* 5km, 0.87 (0.77-0.97) for *T.cristatus* 3km and 1.08 (1.00-1.19) for *L.helveticus* 1km (Annexes Table 4). Concerning the estimation of the transition parameters, their accuracy was variable and primarily depended on the species considered. The species showing the worst estimations for survival and colonization rates corresponded to *T.carnifex*, the other species yielded acceptable estimations in both quantities (Table 6).

Case study 3 – modelling the vital parameters with environmental covariates

For all species and resolutions, the stepwise strategy always included at least one covariate in the detection model. (Annexes Table 5). The selected covariates positively affected the detection probability, except for *T.carnifex* (negative / positive: nrecords / nvisits, at 1km and 3km), *L.helveticus* (negative / positive: nvisitors / nrecords, nvisits at 3km and 5km) and *L.vulgaris* (negative / positive: nvisits / nrecords at 1km and 5km). For all species and resolution (except *T.carnifex*), the best model ($p(\text{covariates})\psi(.)\Phi(t)\gamma(t)$) yielded higher estimations for occupancy than the null model ($p(t)\psi(.)\Phi(t)\gamma(t)$) (Annexes Table 6).

Concerning the impact of the environmental covariates, elevation (linear and quadratic terms considered together) affected the initial occupancy probability for all the species and resolutions with two exceptions, namely *T.carnifex* at 1km and *I.alpestris* at 5km, which showed no significant altitudinal profiles (Table 7). Contrarily to the expectations, a negative effect of urbanization on the colonization or survival probabilities was not found. The only significant results concerned *I.alpestris* (survival and colonization), *T.carnifex* (colonization only) and *T.cristatus* (colonization only), always indicating a positive influence of urban cover. All these results described one particular transition in one resolution only, except for

I.alpestris. For this species, the positive effect of urban cover was consistent between resolutions 3km and 5km in the transition from period 1992-97 to 2003-08 (Table 7). Connectivity increased the initial occupancy probability for all the species at every resolution analysed. The positive influence of connectivity was also found for the survival probability with one exception only, namely *T.carnifex*. The effect of connectivity on the colonization rate was positive for all the species, but *T.cristatus* and *L.vulgaris* showed a significant relationship only at the 1km resolution (Table 7).

Discussion

The present study describes a novel application of site occupancy models for the analysis of sparse data in presence-only databases. Instead of limiting the inference to the recent time, in my analysis I considered the sparse information to test whether the past state of biodiversity could be reliably reconstructed. In order to exploit the sparse data, I enlarged the time windows in which the population is assumed to be closed from one (as it was previously done; Kery *et al.*, 2010b; van Strien *et al.*, 2010) up to 7 years. For the same reason, I adapted the definition of the secondary periods on which the detection probability is modelled and additionally selected a small set of covariates to be integrated into the detection model. I performed three case studies on the Swiss amphibian database and selected the five newt species as model organisms. Two case studies aimed at describing the trends in occupancy, the third tested the influence of elevation, urbanization and connectivity.

Heterogeneity and detection probability

The three case studies showed that a model containing at least one detection covariate was better at explaining the data than a model that assumed constant detection probabilities within each primary period. As a matter of fact, the occupancy estimates between the models with or without detection covariates often differed. This may be due to heterogeneity in the detection process. Furthermore, it is known that site occupancy models underestimate the true occupancy values when heterogeneity is not accounted for (Royle and Nichols 2003; Dorazio 2007). The difference between the two model structures may therefore be explained by the covariates accounting for some heterogeneity, correcting the occupancy estimations that were negatively biased in the model without covariates.

Two reasons are generally used for explaining heterogeneity in the detection process: population size and observation effort. Population size is known to affect the detection probability via the relationship $p_i = 1 - (1 - r)^{Ni}$, where r denotes the detection probability *for a particular individual* and N_i the population size available for detection in quadrat i (Royle and Nichols 2003). As a consequence, differences in population size among quadrats may cause quadrat-specific detection probabilities and, by definition, heterogeneity. Spatial variation in the observation effort may be an additional source of heterogeneity arising from the activity of the observers, rather than biological differences between quadrats. Consider the case of two small areas, one of them being the object of an intense monitoring program by drift fence trapping and the other visited exclusively by some naturalists. Due to technical facilities, the monitored area will be characterised by a detection probability close to 1 for all the species, since every amphibian trying to cross the fence is systematically captured. On the other hand, observations from volunteers carrying torch light surveys in the second area will never reach similar values of detectability. The fact that occupancy estimates differed between the two model structures can therefore be linked to the effect of population size and/or observation effort, because the values of the detection covariates varied specifically by quadrat, secondary period (corresponding to a year) and primary period of closure. Alas, it must be highlighted that the covariates describe heterogeneity in detection probability *in senso lato*, in which the components of effort and population are blended to some extent, thus not directly distinguishable.

One could argue that accounting for heterogeneity is not necessary, as long as we can assume that the resulting trends are unaffected. For instance, why care if the results indicate that the population is declining, regardless the heterogeneity was considered or not? There are at least three answers to this: first, when we hear about a population declining by 5%, we would like to know something more about the magnitude of the phenomenon. We therefore automatically look for some absolute estimates of occupancy, to see whether the 5% lost refers to a large area or to a couple of ponds. When the absolute values of occupancy are biased, the information is not complete. Second, it has also been demonstrated that the decreases in occupancy (and not only the absolute values) are underestimated when heterogeneity is not accounted for (Dorazio 2007), which may lead to "optimistic" assessments of decline with respect to the underway situation. Third, we do not know *a priori* if accounting for heterogeneity or not will result in the same trend in occupancy, we must therefore perform both analyses.

It might be expected that increasing the quadrat size would lead to more constant detection probabilities by saturating the effects of both population size and observation effort. In other words, logic suggests that it is hard to miss all the individuals when population sizes are large and when an overall increased observation effort is exerted on a quadrat within a year. The results suggested that the expectation was wrong, since the occupancy estimates at 5km were still higher when covariates were added to the detection model. Therefore, increasing the quadrat size yielded to more replication in the primary periods to model the detection probability, but did not help in providing more constant quantities. At the moment it is not possible to decompose the heterogeneity *senso lato* that is modelled in this study, consequently it is not possible to state if the persistence of heterogeneity is due to population size, effort or both. Nevertheless, several possibilities are open for further investigations of non-standardized datasets. For example, Royle (2006) noticed that population size can be viewed as a random site effect yielding variation in the detection probability. More interestingly, past population indexes proved to be highly informative in quantifying the influence of population size on the detection probability (Tanadini and Schmidt 2011), a possibility restricted to quadrats for which previous count data are available. Despite this, other proxies for population size may be identified and measured independently: e.g. larger habitat patches may host larger population (Kindvall and Ahlen 1992; Krauss *et al.* 2004; Schmidt and Pellet 2005) and pond area represents a measure easily accessible via Geographical Information Systems. Several level of complexity may therefore be added to the model of detection probability, using covariates and site or period-specific random effects. The integration of similar strategies in the analysis of non-standardized data could greatly help in decomposing the heterogeneity *senso lato* into the components of population size and effort.

Landscape variables and the effect of connectivity

I tested whether elevation and urban cover affected the parameters describing the dynamics of occupancy of the five newts. In order to avoid spurious results due to spatially autocorrelated observations, I built an autologistic site occupancy model using an adapted measure of connectivity derived from the metapopulation theory.

The results showed that the elevation was an important term in predicting the initial occupancy for all the species (Table 7 and Annexes Figure 4). Concerning *T.carnifex*, *T.cristatus* and *L.vulgaris*, the predicted occupancy probability was highest at the lowest elevation and fell close to zero at ranges between 700-1000 metres a.s.l, confirming their overall preference for lowland mild climatic conditions (Grossenbacher 1988; Meyer *et al.* 2009). The initial occupancy of *L.helveticus* was highest at 500-600 metres a.s.l and decreased with increasing elevation to reach the zero at 1500 metres a.s.l, which corresponds to the known altitudinal distribution in Switzerland (Grossenbacher 1988). On the other hand, *I.alpestris* was characterized by a less specific elevation curve: a peak in occurrence at 1000 metres a.s.l. was predicted, but the occupancy probability was in general high over the whole altitudinal range. This result agrees with the previous knowledge describing *I.alpestris* as broadly adapted to different climatic conditions (Grossenbacher 1988; Meyer *et al.* 2009).

Altogether, the strong dependence of occupancy on elevation for four species out of five corroborated the results of Cirovic *et al.* (2008), who described the elevation as a fair proxy for building good descriptive models of newt occurrence.

The results indicated an overall strong and positive effect of connectivity on all the parameters describing the occupancy dynamics of the five newts (Table 7). The positive effect of this measure on all the parameters indicated that the occupied quadrats were spatially organized in clusters. As a matter of fact, many amphibian populations are naturally aggregated across the landscape, which may comprise large networks of interconnected patches at the local and regional level (Hamer and McDonnell 2008). Nevertheless, it is important to highlight that aggregated spatial structures are not necessarily synonymous of metapopulation systems, a term almost indiscriminately used in recent times to describe any group of amphibian populations (Smith and Green 2005). Here, connectivity should be considered only as a measure indicating the relative proximity to the network. I do not want to deny the importance of source-sink dynamics or rescue effects in maintaining the populations at the regional level, rather draw the attention on the fact that the patches of occurrence in this study are likely to describe the location of suitable ecological conditions and historical presence of the species. In light of this, it is easier to understand why connectivity influenced the dynamic parameters even at large resolution. This indicates that the survival probability is relatively low at the edge of the network and in disconnected quadrats, a shared feature with the niche centred notions of determination of ranges (Crozier 2004; Royle and Kery 2007). As a consequence, the geographical pattern of extinction is likely to evolve following a thinning process: the quadrats at the edge losing the hosted populations quicker than the quadrats at the centre of the network. In other words, two quadrats initially comprising ten populations will proceed toward the threshold of extinction at different rates. This is the opposite of a sudden extinction of all the populations in a region, as it could be engendered by a major catastrophic event. This raises interesting questions about the factors effectively driving the dynamic of the quadrats with regards to both the intrinsic characteristics of the quadrat and those of the network. Understanding the contribution of limited dispersal (within and between the quadrats) and habitat alteration deserves further investigations, which should be carried at multiple geographic scales. The results would indicate to managers the most effective conservation strategy in each particular case, e.g. increasing the density in water bodies and terrestrial habitats for maintaining a sustained flow of individuals (Rannap *et al.* 2009) or limiting the impact of anthropogenic stressors such as road traffic (Pellet *et al.* 2004).

Although urbanization is widely recognized to have a strong negative impact on many amphibians (Pellet *et al.* 2004; Price *et al.* 2005; Denoel and Ficetola 2008; Johnson *et al.* 2011), the results did not support this fact neither for the survival nor for the colonization probability (Table 7). Environmental predictors are usually assumed to have a direct influence on the ecology of the species; nonetheless they are often less correlated than expected to the real factors affecting the organism (Austin 2002). A first pragmatic explanation would therefore be that the urbanization measured in this study poorly represented the real causes of decline, but we judge this risk to be low. An alternative and more plausible hypothesis considers that the amount of urbanisation in an occupied quadrat is not relevant, as long as there is part of it covered by suitable patches to ensure the totality of the cycle of hibernation and reproduction (Zanini *et al.* 2009). Unfortunately, it was not possible to verify this last hypothesis, since the models integrating descriptors of suitable habitats (forests, hedges and marshes) did not converge. Other to the absence of an expected negative effect, some species provided transition parameters positively affected by urbanization. A possible explanation for this counterintuitive relationship suggests that the measured variable covaried with an unmeasured characteristic having a positive influence on amphibians. For example, garden ponds and newly established water bodies are highly attractive for the common species as *I.alpestris*: indeed, the alpine newt is known to quickly colonise these kinds of habitat reaching considerably high population densities in some cases (Meyer *et al.* 2009). Similarly, *T.carnifex* is known to face a major expansion phase since decades in the canton of Geneva and surrounding area, a densely urbanized region (Arntzen and Thorpe 1999).

Beyond the hypotheses stated above, several others may be found equally representing possible explanations for the strange results concerning urbanization. As previously anticipated, the models integrating more covariates would have helped in elucidating the effect of other landscape characteristics, but the models did never converge. Additionally, reaching the convergence was far more challenging than expected also for the (conceptually) simple model presented in this study, revealing a generalized difficulty in optimizing the parameters even in the powerful bayesian framework. A proof of this would be that only the autocovariate describing the spatial autocorrelation could be reliably added. This suggests that testing complex ecological models with the methodology developed to deal with sparse information is neither robust nor feasible, allowing only for a descriptive analysis of changes in occupancy. Nevertheless, I stress the fact that this study expressly took into account the data of the years 1980s to infer trends far in the past at the national level. Therefore, the possibility to test ecological hypotheses on volunteer-based records should not be excluded *a priori*. At the regional scale long temporal datasets with numerous detailed observations may be available. Similarly, large-scale distributions may be investigated using the recent years abundant in records. The use of such datasets would be useful to relax numerous assumptions and better model the detection probability, as well as provide enough “raw” material to apply post-stratifications and probabilistic selections of quadrats extending the validity of the inference beyond the analysed sample of sites (van Swaay *et al.* 2002; van Turnhout *et al.* 2008; van Strien *et al.* 2011).

Trends in occupancy and the assumption of closed population

Case studies 1 and 2 analysed the dynamic changes in occupancy of the five newts. The difference between the case studies lay into how quadrats were selected and the unknown impact of geographic bias on occupancy trends. Case study 1 accepted some geographic variation in the quadrats sampled among time periods. Describing why the activity of the observers changed across the three time periods is not the scope of this work, but it is noteworthy that case study 1 permits changes in occupancy that were caused by the total destruction of sites in a quadrat (in addition to local extinction in quadrats where the habitats were not destroyed). In case study 2, the quadrats were visited during the three time periods; hence changes in occupancy cannot be caused by the destruction of quadrats.

The visual inspection of the occupancy trends estimated from the potential range revealed that the results were fairly similar across the three resolutions for a given species (Figure 6 and Annexes Figure 2). This indicates that the quadrat size did not heavily impact on the inferred trends, allowing larger quadrats to yield more replication within the primary periods to model the detection probability. On the contrary, the set of quadrats used in case study 2 provided different estimations of occupancy trends depending on the quadrat size (Figure 7 and Annexes Figure 3). The likely explanation for resolution-dependent results lies in the spatial arrangement of the observations. Despite the geographic distance between the records is always the same, some target observations may or may not belong to a permanently visited quadrat as a function of the variable quadrat size. For this, I suggest to primarily trust the trends estimated under case study 1, because in case study 2 larger quadrats aiming at increasing the replication proved to be unreliable. A more general recommendation for future investigators would be to work with biologically-defined spatial units when the species allows it (e.g. ponds for amphibians), or to define arbitrary units of size comparatively equivalent to the area usually covered by a single population.

The results of case studies 1 and 2 showed that changes in occupancy occurred for four species out of five. Three species increased in occupancy (*I.alpestris*, *T.carnifex*, *L.helveticus*), one species decreased (*T.cristatus*) and one remained stable (*L.vulgaris*) (Table 4 and Annexes Table 4). The only species showing a significant population decline was *T.cristatus*. The decline was apparent in case studies 1 and 2 and hence independent of how quadrats were selected and of a geographic shift in the activity of observers. The results indicated that the decline occurred between periods 1979-85 and 1992-97 and that it did not continue during the transition from 1992-97 to 2003-08. The fact that the species stopped its decline in recent times could be linked to the increase of quadrat survival probabilities, rather than the colonization of new territories. Indeed, the survival probability of *T.cristatus* was considerably low and did not exceed 75% in the transition towards 1992-97, whereas in the transition towards 2003-08 it was close to 90% (Table 5). On the other hand, the colonization probability did not show noteworthy differences between the two transitions. Although the increase of the survival rates may sound as good news, it is likely to be poorly linked to a real increase in the quality of the habitat. Indeed, Grossenbacher (1988) observed that the species was facing a massive decline in the most isolated regions of the country. The loss of the precarious and small populations, together with the persistence of the larger ones, is therefore the likely explanation for the low survival rate observed in transition towards 1992-97.

The species showing the overall largest increase in occupancy was *T.carnifex*. The results indicated that this species underwent a major expansion phase in the transition from 1979-85 to 1992-97, with occupancy estimates more than doubling with regards to the initial situation. An increase in occupancy for *T.carnifex* is in agreement with the previous literature describing this species as invading the western part of Switzerland, a process started some decades ago (Grossenbacher 1988; Arntzen and Thorpe 1999; Schmidt and Zumbach 2005). Although the credible intervals of occupancy estimates are quite large, we can reasonably assess that *T.carnifex* is occupying almost completely its known potential range since the period 1992-97 (Figure 6); this means either that the expansion phase is still on the way or that it stopped. Unfortunately, case study 2 did not help in providing a definitive answer, because the estimated trends strongly depended on the quadrat size: the 5km resolution suggested that the expansion is still occurring today, whereas the 3km resolution that the expansion stopped (Figure 7 and Annexes Figure 3).

The third species showing an overall significant trend was *I.alpestris*. The results indicated that this species encountered a slight expansion phase from period 1992-97 to 2003-08, supported by a colonization capacity doubled in the transition towards 2003-08 (Table 6). Although *I.alpestris* is known to be one of the species less affected by the amphibian crisis at the national level (Schmidt and Zumbach 2005), the result was surprising and dubious at the same time. Similar to *I.alpestris*, also the results concerning the occupancy trends of *L.helveticus* and *L.vulgaris* were found to be unusually optimistic. Indeed, the results suggested an overall stationary trend for *L.vulgaris* and an increase in the period 2003-08 for *L.helveticus* (reported solely by case study 2). Litmus test, the national Red List 2005 highlighted that in the last decades *L.helveticus* disappeared from several sites in the Eastern part of the country and that *L.vulgaris* faced an even more important decline (Schmidt and Zumbach 2005).

It is difficult to provide a biological explanation for the astonishing results obtained for *I.alpestris*, *L.helveticus* and *L.vulgaris*, it is therefore appropriate to suggest that the underlying methodology may play an important role. When trends in occupancy are analysed over time, an assumption of closure over each primary period needs to be stated. This means that within each period no changes in occupancy occur, the quadrats are either occupied or not occupied, thus allowing the estimation of the Proportion of Area Occupied (hereafter PAO). However, there are cases in which the assumption of closure required to estimate the PAO may be violated. Lack of closure in the primary periods is akin to temporary emigration, so if temporary absence of the species is random, the violated assumption of closed population will be a component of imperfect detection (Kendall 1999; Kery and Schaub 2011). This kind of violation defines a new metric for “occupancy”, namely the Proportion of Area Used (hereafter PAU). The difference between PAO and PAU therefore lies in the underlying abidance of the closure assumption, distinguishing from a situation in which quadrats are constantly occupied from a situation in which quadrats are used. If the species is physically present (or absent) in a quadrat at random time points during the primary period, the PAU correctly describes the use of the quadrat (MacKenzie 2005) and temporal unavailability for detection will engender only a decrease in the detection probability. By contrast, if a switch in the occupancy status (either from occupied to unoccupied or the reverse) is followed by the persistence of the new status, the physical presence (or absence) cannot be considered as random along the primary period. In this case, a decreased detection probability is no more able to explain alone the observed absence of the species, a situation which ultimately leads to an overestimation of the occupancy values (Rota *et al.* 2009).

In the present study, the assumption of closure had to be extended up to 7 years in order to integrate the sparse data of the karch database. Nevertheless, it has not been considered that switches in the occupancy status were likely to represent permanent events, rather than temporary. As a matter of fact, the quadrats defined an area larger than the surface usually covered by a single population, making that the physical absence of the species was more likely to be explained by an extinction event, rather than the movement of all populations to another quadrat. Similarly, switches to the occupied status were more probably caused by a colonization event, rather than a temporary passage of the species. Although amphibians are able of movements that are surprising for presumably poorly dispersing animals (Smith and Green 2005), their vagility was not sufficient to ensure a temporal randomly distributed use of quadrats within the primary periods. As a result, the Proportion of Area Used estimates were biased high.

It is possible that the distribution of the visits in the three primary periods played an additional role in creating overoptimistic trends towards period 2003-08, despite the three periods have about the same length. Logic tells us that the time of the visits and the time span in which a quadrat is closed to changes are independent quantities, the first arising from the observation process and the latter from the biological process. Nevertheless, the closure assumption is violated only if the visits are spread on a temporal window longer than the period where the system is closed. This means that, omitting the variation in occupancy due to the seasonal migrations of individuals (secondary periods are defined as years in this study), visiting a quadrat for several years in a row was more likely to fall over a permanent extinction or colonization event than visiting a quadrat only in two or three subsequent years. This means that the period 2003-08 exposed a greater number of quadrats to the violation of the assumption than either period 1979-85 or 1992-97. If the number of quadrats violating the closure assumption relates to the amplitude of the bias, it might therefore be expected that the magnitude of the overestimation of occupancy values in period 2003-08 is more important than in the other primary periods. Nevertheless, this suggestion represents only a hypothesis needing a formal validation, but provides a mechanistic explanation to why the dubious results of *I.alpestris* and *L.helveticus* appeared in the transition toward period 2003-08.

In summary, the method proposed in this study to deal with the sparse data was unsafe in providing occupancy estimations and the relative trends, because the switches in occupancy occurring within the primary periods were likely to represent permanent and not temporary events. Therefore, the definition of yearly secondary periods cannot exhaustively perform without heavy consequences on the resulting trends in occupancy. Nevertheless, these findings should not discourage the researchers to go further in the analysis of non-standardized data. The framework proposed in this study did not only attempt to exploit the sparse data, but also represented the first try to integrate the observations missing day and month of the report in site occupancy models. When the complete date is available, it is no more necessary to define secondary periods of one year and simple visits can be deduced straightforwardly. As a consequence, the degree of replication within primary periods can be greatly increased. An appropriate geographic scale, in conjunction with an in-depth knowledge on the turnover rates of the species investigated (Marsh and Trenham 2001), should therefore allow the researcher to define multiple-year primary periods when required, in which the consequences of the violated assumption do not prevent the validity of the study.

Conclusion

State-of-the-art methods for the analysis of data arising from citizen-science project and naturalists made great strides in recent years. In the present study I proposed a novel application based on these methods to push the analysis as back as possible in time, where the information collected is much sparser, but more interesting for describing the past state of nature. Unfortunately, the proposed application did not succeed in providing reliable trends in occupancy or testing the causes for the observed changes, since the consequences of the violated assumptions had major impacts on the results obtained. On the other side, we should keep in mind that database were gathered by other people with their own aims, not for allowing us inferring detailed population trends or testing (our) ecological hypotheses.

Despite all of this, the study provided valuable insights in the analysis (and non-analysis) of presence-only databases. Previous inferences rarely accounted for heterogeneity in the detection probability: here we emphasize how the spatial variation in the population size and observation effort can substantially lead to negatively biased results. Future researches should therefore consider these two factors, in order to allow for insurance against distorted inference. I would also like to stress the fact that database records should not be discarded *a priori* because their collection did not follow a known sampling protocol, they have a great potential for at least three reasons: first, they are valuable in complementing the monitoring programs when describing the on-going state of nature (Schmeller *et al.* 2009). Second, they may hold extremely long time series providing retrospective insights in the evolution of a species distribution (van Strien *et al.* 2011). Third, they provide a cheap and immediate source of information for identifying the ecological factors affecting the distribution of species and communities, therefore allowing policy-makers and conservationists for targeted and valuable future management actions.

Literature cited

Arntzen, J. W. and R. S. Thorpe (1999). "Italian crested newts (Triturus carnifex) in the Basin of Geneva: Distribution and genetic interactions with autochthonous species." Herpetologica **55**(4): 423-433.

Augustin, N. H., et al. (1996). "An autologistic model for the spatial distribution of wildlife." Journal of Applay Ecology **33**(2): 339-347.

Austin, M. P. (2002). "Spatial prediction of species distribution: an interface between ecological theory and statistical modelling." Ecological Modelling **157**(2-3): 101-118.

Beebee, T. J. C. (1996). Ecology And Conservation of Amphibians. London, Chapman & Hall.

Box, G. E. P. and D. R. Cox (1964). "An Analysis of Transformations." Journal of the Royal Statistical Society Series B-Statistical Methodology **26**(2): 211-252.

Cirovic, R., et al. (2008). "Distribution patterns and environmental determinants of European newts in the Montenegrin karst area." Biologia **63**(5): 745-752.

Crozier, L. (2004). "Warmer winters drive butterfly range expansion by increasing survivorship." Ecology **85**(1): 231-241.

Cushman, S. A. (2006). "Effects of habitat loss and fragmentation on amphibians: A review and prospectus." Biological Conservation **128**(2): 231-240.

Denoel, M. and G. F. Ficetola (2008). "Conservation of newt guilds in an agricultural landscape of Belgium: the importance of aquatic and terrestrial habitats." Aquatic Conservation-Marine and Freshwater Ecosystems **18**(5): 714-728.

Dorazio, R. M. (2007). "On the choice of statistical models for estimating occurrence and extinction from animal surveys." Ecology **88**(11): 2773-2782.

Eigenbrod, F., et al. (2011). "Sub-optimal study design has major impacts on landscape-scale inference." Biological Conservation **144**(1): 298-305.

ESRI (2008). ArcGIS Desktop 9.3. Redlands, CA, Environmental System Research Institute.

Gardner, T. (2001). "Declining amphibian populations: a global phenomenon in conservation biology." Animal Biodiversity & Conservation **2**(2): 25-44.

Gelman, A. and J. Hill (2007). Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge, Cambridge University Press.

Govindan, B. N., et al. (2011). "Host selection and responses to forest fragmentation in acorn weevils: Inferences from dynamic occupancy models." Oikos.

Grossenbacher, K. (1988). Verbreitungsatlas der Amphibien der Schweiz. SBN Basel.

Guisan, A. and N. E. Zimmermann (2000). "Predictive habitat distribution models in ecology." Ecological Modelling **135**(2-3): 147-186.

Hamer, A. J. and M. J. McDonnell (2008). "Amphibian ecology and conservation in the urbanising world: A review." Biological Conservation **141**(10): 2432-2449.

Hartel, T., et al. (2010). "Using connectivity metrics and niche modelling to explore the occurrence of the northern crested newt Triturus cristatus (Amphibia, Caudata) in a traditionally managed landscape." Environmental Conservation **37**(2): 195-200.

Houlahan, J. E., et al. (2000). "Quantitative evidence for global amphibian population declines." Nature **404**(6779): 752-755.

Humbel, R. (1998). GIS use in the framework of a national statistical office: Establishment land use statistics and geocoding of national censuses. Agricultural Statistics 2000, Proceedings - an International Conference on Agricultural Statistics. T. E. Holland and M. P. R. VandenBroecke. Voorburg, Int Statistical Inst**:** 291-304.

Johnson, P. T. J., et al. (2011). "Regional Decline of an Iconic Amphibian Associated with Elevation, Land-Use Change, and Invasive Species." Conservation Biology **25**(3): 556-566.

Joly, P., et al. (2001). "Habitat matrix effects on pond occupancy in newts." Conservation Biology **15**(1): 239-248.

Kendall, W. L. (1999). "Robustness of closed capture-recapture methods to violations of the closure assumption." Ecology **80**(8): 2517-2525.

Kery, M. (2010). Introduction to WinBUGS for ecologists - A Bayesian approach to regression, ANOVA, mixed models and related analyses. 6204 Sempach, Switzerland.

Kery, M., et al. (2010a). "Predicting species distributions from checklist data using site-occupancy models." Journal of Biogeography **37**(10): 1851-1862.

Kery, M., et al. (2010b). "Site-Occupancy Distribution Modeling to Correct Population-Trend Estimates Derived from Opportunistic Observations." Conservation Biology **24**(5): 1388-1397.

Kery, M. and M. Schaub (2011). Bayesian Population Analysis using WinBUGS - A hierarchical perspective. 6204 Sempach, Switzerland.

Kery, M. and B. R. Schmidt (2008). "Imperfect detection and its consequences for monitoring for conservation." Community Ecology **9**(2): 207-216.

Kindvall, O. and I. Ahlen (1992). "Geometrical Factors and Metapopulation Dynamics of the Bush Cricket, Metrioptera bicolor Philippi (Orthoptera: Tettigoniidae)." Conservation Biology **6**(4): 520-529.

Krauss, J., et al. (2004). "Landscape occupancy and local population size depends on host plant distribution in the butterfly Cupido minimus." Biological Conservation **120**(3): 355-361.

Lister, A. M. (2011). "Natural history collections as sources of long-term datasets." Trends in Ecology & Evolution **26**(4): 153-154.

MacKenzie, D. I. (2005). "What are the issues with presence-absence data for wildlife managers?" Journal of Wildlife Management **69**(3): 849-860.

MacKenzie, D. I., et al. (2003). "Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly." Ecology **84**(8): 2200-2207.

MacKenzie, D. I., et al. (2002). "Estimating site occupancy rates when detection probabilities are less than one." Ecology **83**(8): 2248-2255.

Marsh, D. M. and P. C. Trenham (2001). "Metapopulation dynamics and amphibian conservation." Conservation Biology **15**(1): 40-49.

Meyer, A., et al. (2009). Les amphibiens et les reptiles de Suisse. Bern, Haupt.

Nichols, J. D. and B. K. Williams (2006). "Monitoring for conservation." Trends in Ecology & Evolution **21**(12): 668-673.

OFS (2007). Description de données GEOSTAT: Statistique de la superficie NOAS92. D. F. d. l'Intérieur.

Pellet, J., et al. (2004). "A concentric analysis of the impact of urbanization on the threatened European tree frog in an agricultural landscape." Conservation Biology **18**(6): 1599-1606.

Pellet, J. and B. R. Schmidt (2005). "Monitoring distributions using call surveys: estimating site occupancy, detection probabilities and inferring absence." Biological Conservation **123**(1): 27-35.

Price, S. J., et al. (2005). "The importance of spatial scale for conservation and assessment of anuran populations in coastal wetlands of the western Great Lakes, USA." Landscape Ecology **20**(4): 441-454.

R Development Core Team (2010). R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

Rannap, R., et al. (2009). "Restoring ponds for amphibians: a success story." Hydrobiologia **634**(1): 87-95.

Rota, C. T., et al. (2009). "Occupancy estimation and the closure assumption." Journal of Applay Ecology **46**(6): 1173-1181.

Royle, J. A. (2006). "Site occupancy models with heterogeneous detection probabilities." Biometrics **62**(1): 97-102.

Royle, J. A. and R. M. Dorazio (2008). Hierarchical modeling and inference in ecology: the analysis of data from populations and communities, Academic Press.

Royle, J. A. and M. Kery (2007). "A Bayesian state-space formulation of dynamic occupancy models." Ecology **88**(7): 1813-1823.

Royle, J. A. and J. D. Nichols (2003). "Estimating abundance from repeated presence-absence data or point counts." Ecology **84**(3): 777-790.

Schmeller, D. S., et al. (2009). "Advantages of Volunteer-Based Biodiversity Monitoring in Europe." Conservation Biology **23**(2): 307-316.

Schmidt, B. R. (2005). "Monitoring the distribution of pond-breeding amphibians when species are detected imperfectly." Aquatic Conservation-Marine and Freshwater Ecosystems **15**(6): 681-692.

Schmidt, B. R. and J. Pellet (2005). "Relative importance of population processes and habitat characteristics in determining site occupancy of two anurans." Journal of Wildlife Management **69**(3): 884-893.

Schmidt, B. R. and S. Zumbach (2005). Liste Rouge des amphibiens menacés en Suisse. Série OFEFP: l'environnement pratique. Berne, Switzerland, Centre de coordination pour la protection des amphibiens et des reptiles de Suisse (karch) and Office fédéral de l'environnement, des forêts et du paysage (OFEFP).

Secretariat of the Convention on Biological Diversity (2010). Global Biodiversity Outlook 3. Montreal**:** 94.

Shaffer, H. B., et al. (1998). "The role of natural history collections in documenting species declines." Trends in Ecology & Evolution **13**(1): 27-30.

Smith, M. A. and D. M. Green (2005). "Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations?" Ecography **28**(1): 110-128.

Spiegelhalter, D., et al. (2003). WinBUGS User Manual. MCR Biostatistics Unit, Institute of Public Health, Robinson Way, Cambridge CB2 2SR, UK.

Stuart, S. N., et al. (2004). "Status and trends of amphibian declines and extinctions worldwide." Science **306**(5702): 1783-1786.

Sturtz, S., et al. (2005). "R2WinBUGS: A package for running WinBUGS from R." Journal of Statistical Software **12**(3): 1-16.

Swisstopo (2007). VECTOR25: Le modèle numérique du territoire de la Suisse. OFT.

Tanadini, L. G. and B. R. Schmidt (2011). "Population Size Influences Amphibian Detection Probability: Implications for Biodiversity Monitoring Programs." PLoS ONE **6**(12): e28244.

Van Buskirk, J. (2005). "Local and landscape influence on amphibian occurrence and abundance." Ecology **86**(7): 1936-1947.

van Strien, A. J., et al. (2010). "Site-occupancy models may offer new opportunities for dragonfly monitoring based on daily species lists." Basic and Applay Ecology **11**(6): 495-503.

van Strien, A. J., et al. (2011). "Metapopulation dynamics in the butterfly Hipparchia semele changed decades before occupancy declined in The Netherlands." Ecological Applications **21**(7): 2510-2520.

van Swaay, C. A. M. (1990). "An assessment of the changes in butterfly abundance in The Netherlands during the 20th Century." Biological Conservation **52**(4): 287-302.

van Swaay, C. A. M., et al. (2002). Monitoring butterflies in the Netherlands: How to get unbiased indices. Proceedings of the Experimental & Applay Entomology, NEV Amsterdam.

van Turnhout, C. A. M., et al. (2008). "Monitoring common and scarce breeding birds in the Netherlands: applying a post-hoc stratification and weighting procedure to obtain less biased population trends." Revista Catalana d'Ornitologia **24**: 15:29.

Weber, D., et al. (2004). "Scale and trends in species richness: considerations for monitoring biological diversity for political purposes." Global Ecology and Biogeography **13**(2): 97-104.

Wood, C., et al. (2011). "eBird: Engaging Birders in Science and Conservation." PLoS Biol **9**(12): e1001220.

Yoccoz, N. G., et al. (2001). "Monitoring of biological diversity in space and time." Trends in Ecology & Evolution **16**(8): 446-453.

Zanini, F. (2006). Amphibian conservation in human shaped environments: landscape dynamics, habitat modeling and metapopulation analyses. EPFL. Lausanne. **PhD Thesis:** 1-201.

Zanini, F., et al. (2009). "The transferability of distribution models across regions: an amphibian case study." Diversity and Distributions **15**(3): 469-480.

Figures and Tables

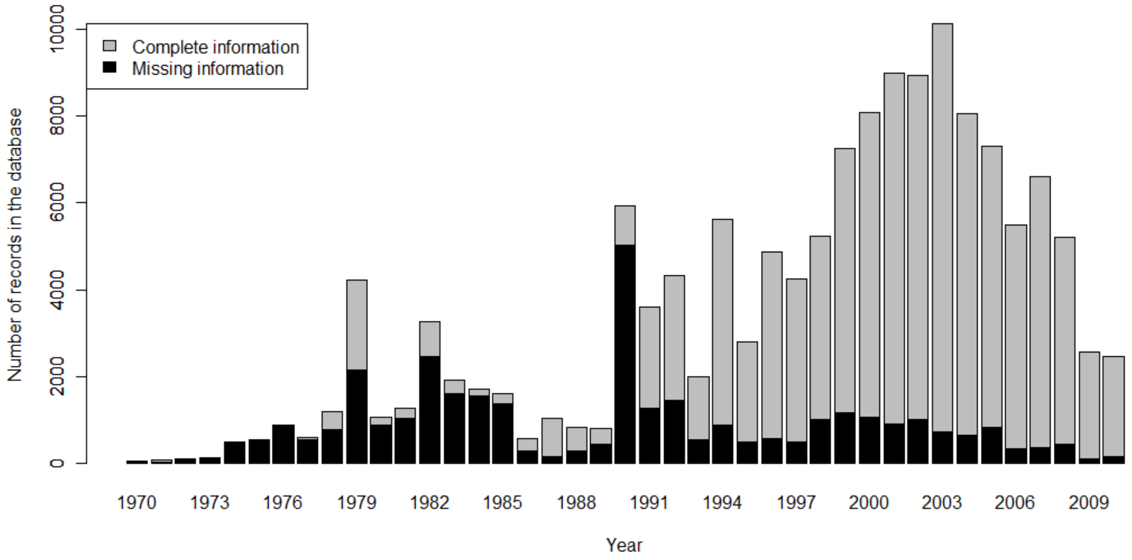


Figure 1 Distribution of records in the karch database from year 1970 to 2010. The fraction of data characterized by missing temporal information in the day or the month is represented by black bars (i.e. only the year of the observation is known).

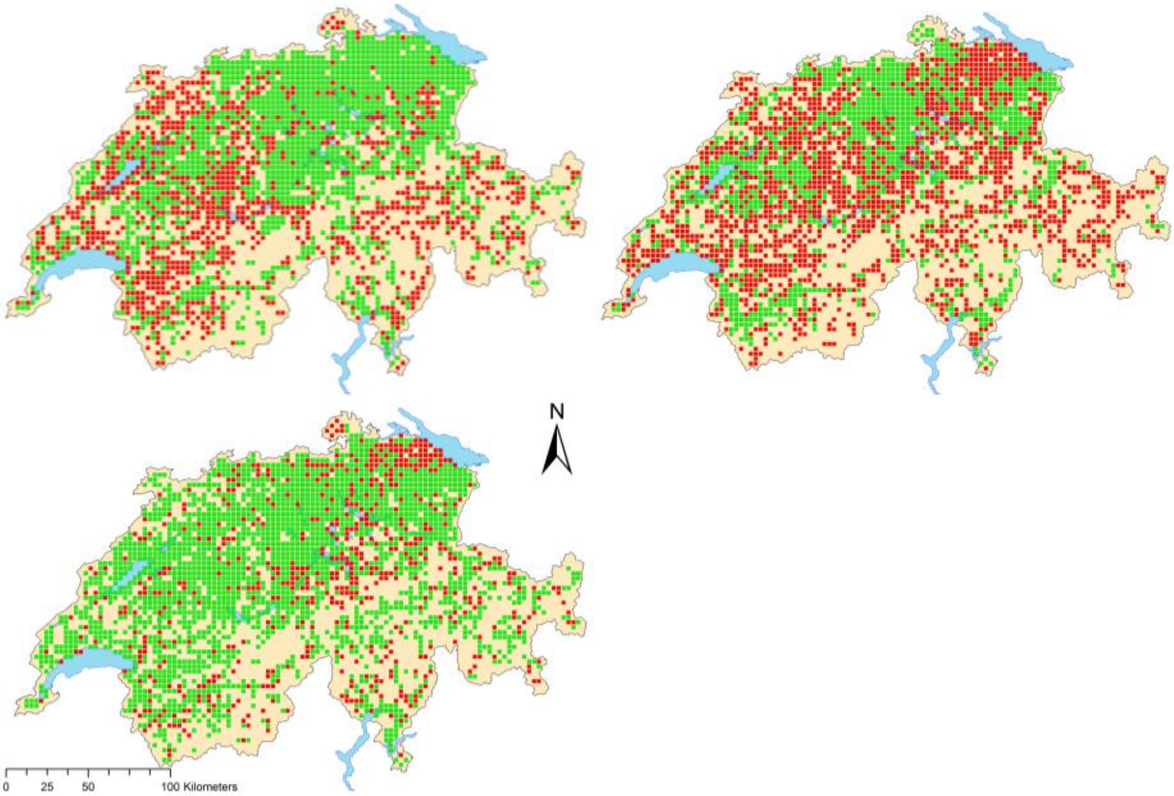


Figure 2 Geographic variation in the spatial arrangement of amphibian records over Switzerland. The showed 3-km quadrats represent the set employed in case study 3. Top left: period 1979-85; top right: period 1992-97; bottom left: period 2003-08. The green colour denotes the quadrats that were visited at least once, the red colour the quadrats that were visited during another time period.

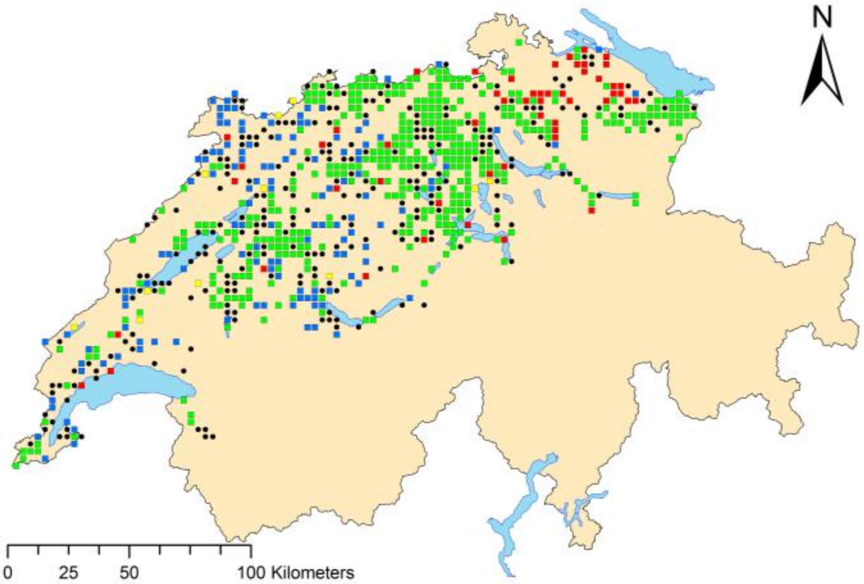


Figure 3 Geographic variation in the spatial arrangement of the data in the historical range of *L.helveticus* at 3km resolution (case study 1). Green quadrats were visited at least once during every primary period. Red, yellow and dark blue quadrats denote respectively the quadrats visited only during the period 1979-85, 1992-97 and 2003-08. The black dots indicate the quadrats visited during two periods out of three (aggregated for clarity).

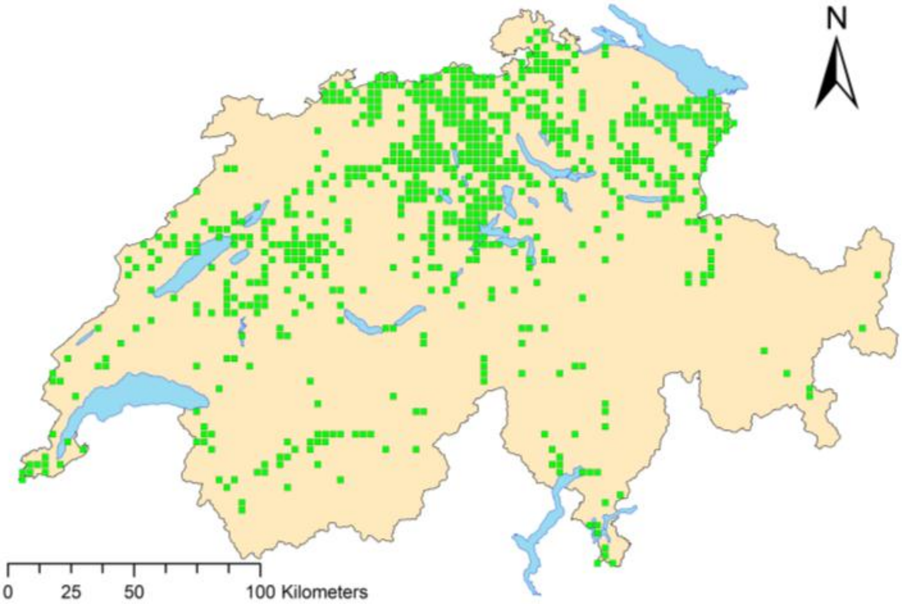


Figure 4 Spatial arrangement of the 3km-quadrats for which amphibian data were reported in every time period (basis for case study 2). The areas with a lower number of quadrats corresponded primarily to mountainous regions in the southern and eastern parts of the country, with quadrats usually located on the bottom of the valleys.

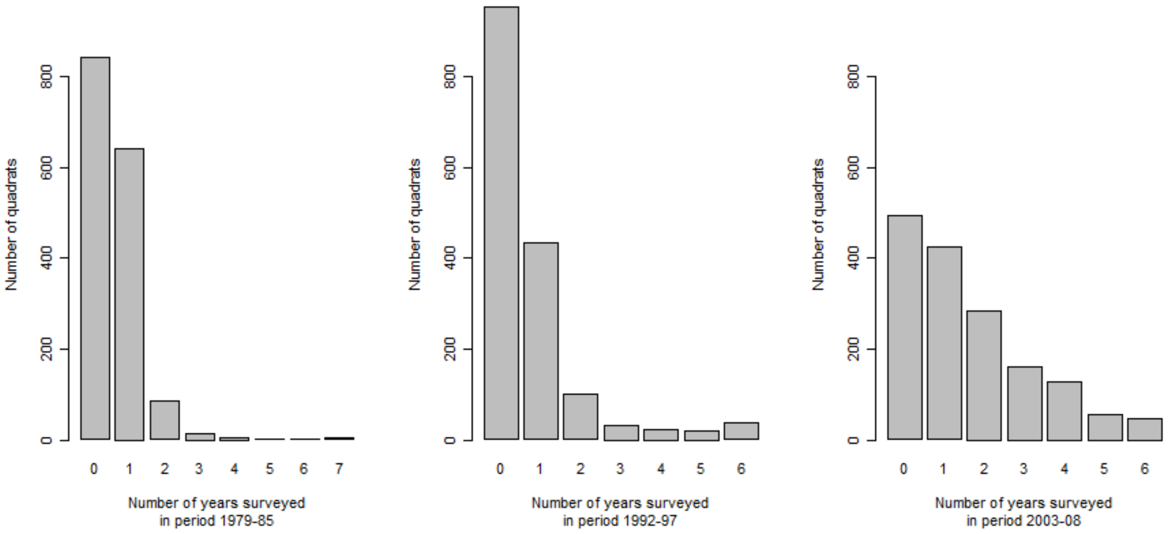


Figure 5 Distribution of the replication in the historical range of *L.helveticus* at 1km resolution (case study 1). From left to right: period 1979-85, period 1992-97 and period 2003-08. The replicates represent the quadrats grouped by the number of years surveyed in each period. The total number of quadrats corresponds to 1601 for the case shown.

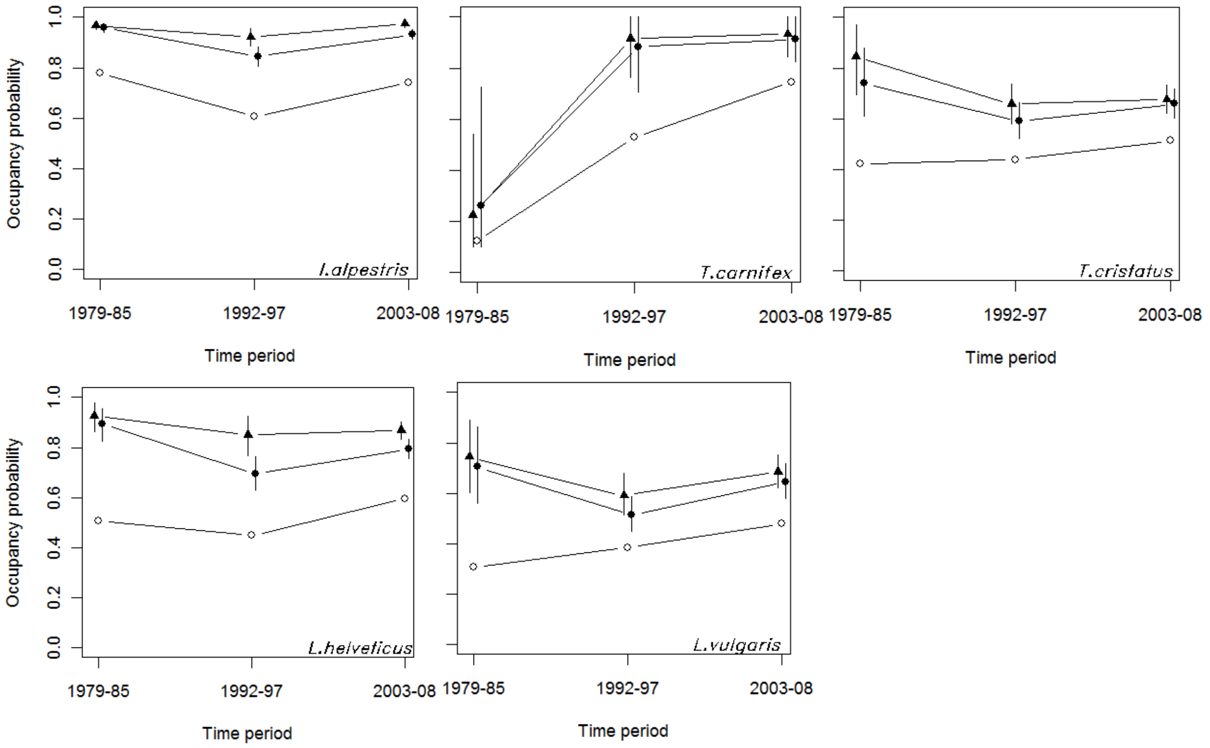


Figure 6 Observed (open circles) and estimated (filled symbols) occupancy probabilities with 95% credible intervals (vertical bars) for the three primary periods in case study 1 at 3km resolution. The occupancy probabilities estimated under the best (filled triangle) and the null (filled circle) detection models are shown.

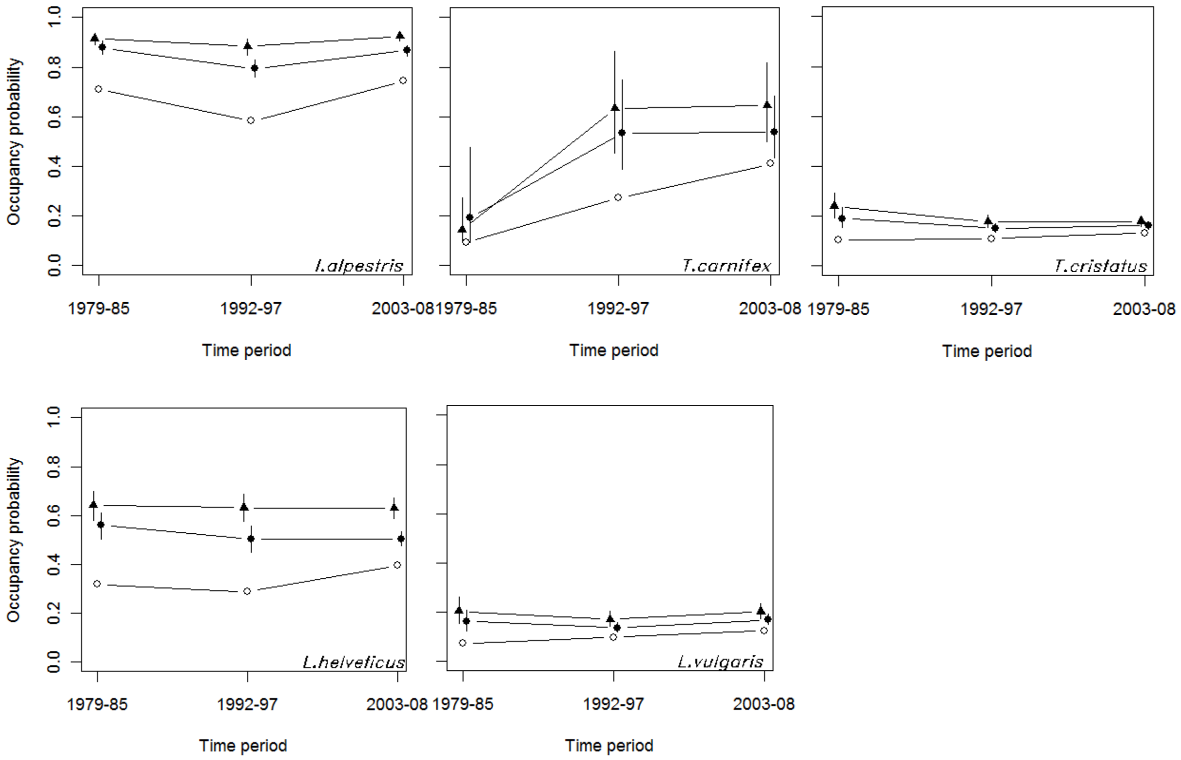


Figure 7 Observed (open circles) and estimated (filled symbols) occupancy probabilities with 95% credible intervals (vertical bars) for the three periods in case study 2 at 3km resolution. The occupancy probabilities estimated under the best (filled triangle) and the null (filled circle) detection models are shown.

Table 1 Number of quadrats forming the potential range on which the case study 1 is based. The values represent the sum of quadrats known for having ever hosted the target species according to the data contained in the karch database (years 1970 - 2010).

Species		*I.alpestris*	*T.carnifex*	*T.cristatus*	*L.helveticus*	*L.vulgaris*
Resolution	1 km	5335	90	391	1601	408
	3 km	2191	51	250	869	279
	5 km	1231	38	196	518	210

Table 2 Number of quadrats on which the case study 2 is based. The values represent the sum of quadrats visited during each of the three primary periods and located within the biogeographical extent of the target species. The sum of regions forming the biogeographical extent of each species is shown (total of Switzerland $n = 16$).

Species		*I.alpestris*	*T.carnifex*	*T.cristatus*	*L.helveticus*	*L.vulgaris*
N° regions		16	4	12	11	13
Resolution	1 km	556	19	506	511	519
	3 km	708	44	638	639	659
	5 km	595	47	530	518	554

Table 3 Number of quadrats on which the case study 3 is based. All the quadrats located in the biogeographical extent of the selected species were considered. *T.carnifex* at 5 km resolution was not investigated since the number of quadrats was insufficient. The number of biogeographic regions for each species and resolution is reported in Table 3.

Species		*I.alpestris*	*T.carnifex*	*T.cristatus*	*L.helveticus*	*L.vulgaris*
N° regions		16	4	12	11	13
Resolution	1 km	7616	471	6379	6421	6598
	3 km	2760	192	2205	2146	2307
	5 km	1264	-	1006	919	1052

Table 4 Growth rates per transition and geometric mean growth rates for all species and resolutions estimated in case study 1. The point value and the extremes of the credible interval (2.50% and 97.50% quantiles) are represented for each quantity. The numerical labels indicate the primary time periods, the arrows the transitions between them. Significant departures from a stationary population with λ=1 are indicated in **bold**.

Specie	**Resolution**	**Growth rate 1→2**	**Growth rate 2→3**	**Geometric mean λ**
I.alpestris	1km	0.97 (0.94-1.01)	1.02 (0.99-1.05)	0.99 (0.98-1.01)
I.alpestris	3km	**0.95 (0.92-0.99)**	**1.06 (1.02-1.10)**	1.00 (0.99-1.02)
I.alpestris	5km	0.99 (0.96-1.02)	**1.03 (1.01-1.06)**	**1.01 (1.00-1.02)**
T.carnifex	1km	**3.43 (1.45-7.00)**	1.03 (0.93-1.23)	**1.84 (1.23-2.68)**
T.carnifex	3km	**5.03 (1.66-10.00)**	1.03 (0.90-1.23)	**2.21 (1.31-3.16)**
T.carnifex	5km	**5.50 (2.07-11.33)**	1.18 (0.95-1.62)	**2.47 (1.58-3.46)**
T.cristatus	1km	**0.80 (0.67-0.96)**	1.09 (0.96-1.25)	0.93 (0.86-1.02)
T.cristatus	3km	**0.78 (0.65-0.97)**	1.03 (0.91-1.17)	**0.90 (0.82-0.99)**
T.cristatus	5km	**0.74 (0.63-0.90)**	1.00 (0.87-1.15)	**0.86 (0.80-0.94)**
L.helveticus	1km	0.96 (0.87-1.05)	1.03 (0.95-1.13)	0.99 (0.96-1.03)
L.helveticus	3km	0.92 (0.81-1.02)	1.03 (0.93-1.14)	0.97 (0.94-1.01)
L.helveticus	5km	1.02 (0.92-1.13)	0.94 (0.88-1.02)	0.98 (0.94-1.03)
L.vulgaris	1km	0.84 (0.68-1.05)	**1.27 (1.12-1.45)**	1.03 (0.93-1.15)
L.vulgaris	3km	0.80 (0.64-1.02)	1.16 (0.99-1.36)	0.96 (0.86-1.08)
L.vulgaris	5km	0.93 (0.71-1.22)	1.12 (0.95-1.31)	1.02 (0.89-1.15)

Table 5 Survival (Φ) and colonization (γ) probabilities for all species and resolutions estimated in case study 1. The point values (with 95% credible interval) are represented on the normal scale. The numerical labels indicate the primary time periods, the arrows the transitions between them. The highly inaccurate estimations (95% credible interval larger than 0.5 on the normal scale) are highlighted in **bold**.

Specie	**Resolution**	**Survival Φ 1→2**	**Survival Φ 2→3**	**Colonization γ 1→2**	**Colonization γ 2→3**
I.alpestris	1km	0.96 (0.93-0.98)	0.97 (0.95-0.98)	**0.41 (0.07-0.68)**	0.79 (0.64-0.90)
I.alpestris	3km	0.94 (0.90-0.98)	0.98 (0.96-0.99)	**0.33 (0.05-0.61)**	0.90 (0.76-0.99)
I.alpestris	5km	0.97 (0.94-1.00)	0.99 (0.98-1.00)	**0.43 (0.13-0.73)**	0.85 (0.62-0.99)
T.carnifex	1km	**0.82 (0.41-0.99)**	0.95 (0.86-1.00)	0.91 (0.69-1.00)	**0.55 (0.04-0.97)**
T.carnifex	3km	**0.81 (0.39-0.99)**	0.94 (0.81-1.00)	0.90 (0.68-1.00)	**0.58 (0.04-0.98)**
T.carnifex	5km	**0.77 (0.32-0.99)**	0.94 (0.82-1.00)	**0.76 (0.45-0.97)**	**0.62 (0.07-0.97)**
T.cristatus	1km	0.70 (0.59-0.81)	0.89 (0.79-0.97)	**0.51 (0.05-0.90)**	0.42 (0.23-0.62)
T.cristatus	3km	0.70 (0.59-0.82)	0.91 (0.81-0.98)	**0.38 (0.03-0.82)**	0.23 (0.06-0.42)
T.cristatus	5km	0.69 (0.57-0.81)	0.88 (0.77-0.96)	**0.51 (0.04-0.97)**	0.24 (0.04-0.45)
L.helveticus	1km	0.93 (0.84-1.00)	0.88 (0.84-0.93)	**0.24 (0.01-0.60)**	0.88 (0.73-0.98)
L.helveticus	3km	0.89 (0.79-0.97)	0.89 (0.84-0.93)	**0.36 (0.03-0.79)**	0.73 (0.49-0.89)
L.helveticus	5km	0.94 (0.86-0.99)	0.89 (0.84-0.94)	**0.73 (0.22-0.99)**	**0.57 (0.09-0.90)**
L.vulgaris	1km	0.75 (0.61-0.87)	0.94 (0.85-1.00)	**0.29 (0.01-0.65)**	0.59 (0.41-0.75)
L.vulgaris	3km	0.70 (0.56-0.85)	0.87 (0.75-0.96)	**0.25 (0.01-0.58)**	0.41 (0.25-0.57)
L.vulgaris	5km	0.74 (0.60-0.88)	0.86 (0.75-0.96)	**0.29 (0.03-0.57)**	0.35 (0.18-0.52)

Table 6 Survival (Φ) and colonization (γ) probabilities estimated for all species and resolutions in case study 2. The point values (with 95% credible interval) are represented on the normal scale. The numerical labels indicate the primary time periods, the arrows the transitions between.

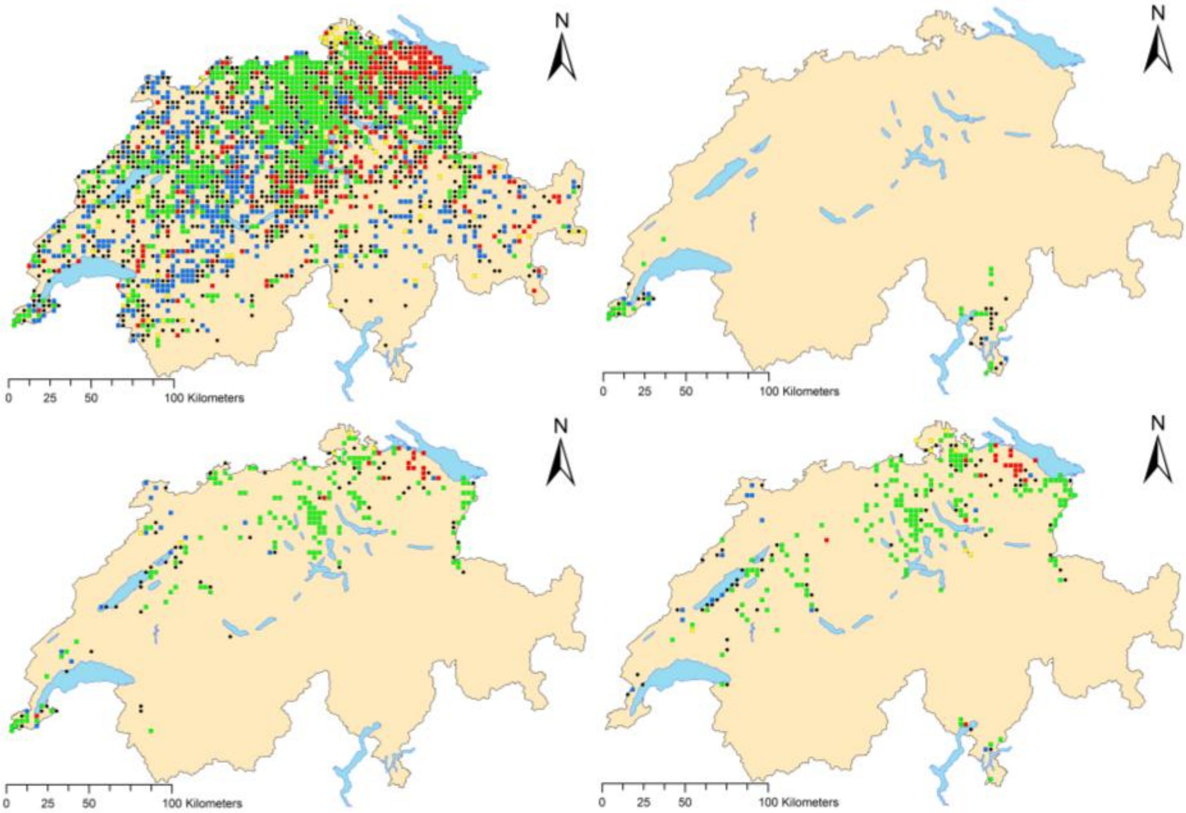
Species	**Resolution**	**Survival Φ 1→2**	**Survival Φ 2→3**	**Colonization γ 1→2**	**Colonization γ 2→3**
I.alpestris	1km	0.97 (0.93-1.00)	0.92 (0.88-0.96)	0.13 (0.00-0.36)	0.27 (0.12-0.42)
I.alpestris	3km	0.95 (0.91-0.99)	0.98 (0.96-1.00)	0.14 (0.01-0.34)	0.47 (0.29-0.63)
I.alpestris	5km	0.98 (0.96-1.00)	0.99 (0.98-1.00)	0.18 (0.05-0.34)	0.33 (0.16-0.51)
T.carnifex	1km	0.67 (0.18-0.99)	0.71 (0.28-0.99)	0.44 (0.06-0.92)	0.28 (0.01-0.79)
T.carnifex	3km	0.80 (0.38-0.99)	0.89 (0.64-1.00)	0.58 (0.32-0.87)	0.22 (0.01-0.67)
T.carnifex	5km	0.77 (0.34-0.99)	0.92 (0.72-1.00)	0.38 (0.17-0.64)	0.29 (0.03-0.63)
T.cristatus	1km	0.76 (0.61-0.90)	0.87 (0.72-0.98)	0.04 (0.00-0.08)	0.02 (0.00-0.04)
T.cristatus	3km	0.66 (0.52-0.79)	0.90 (0.78-0.98)	0.03 (0.00-0.07)	0.03 (0.00-0.05)
T.cristatus	5km	0.68 (0.53-0.82)	0.85 (0.73-0.95)	0.04 (0.00-0.11)	0.05 (0.01-0.10)
L.helveticus	1km	0.96 (0.88-1.00)	0.85 (0.76-0.94)	0.11 (0.00-0.26)	0.31 (0.18-0.43)
L.helveticus	3km	0.93 (0.84-0.99)	0.87 (0.80-0.93)	0.10 (0.01-0.24)	0.22 (0.11-0.34)
L.helveticus	5km	0.97 (0.93-1.00)	0.89 (0.83-0.94)	0.22 (0.05-0.39)	0.07 (0.00-0.19)
L.vulgaris	1km	0.69 (0.50-0.87)	0.93 (0.79-1.00)	0.03 (0.00-0.08)	0.07 (0.04-0.12)
L.vulgaris	3km	0.69 (0.51-0.85)	0.84 (0.70-0.96)	0.04 (0.00-0.09)	0.07 (0.04-0.11)
L.vulgaris	5km	0.67 (0.52-0.82)	0.88 (0.75-0.98)	0.04 (0.00-0.10)	0.08 (0.04-0.13)

Table 7 Posterior summaries of the results obtained testing the environmental covariates in case study 3. The values are on the logit scale and denote the mean value of the regression slope (with standard deviation of the posterior distribution). $\psi 1$ denotes initial occupancy, Φ survival and γ colonization. The numerical labels indicate the primary time periods, the arrows the transitions between them. ALT denotes elevation, ALTSQ square elevation, CON connectivity and URB urban cover (see Methods for further details on the model structure and the covariates). The regression slopes whose 95% credible interval did not contain 0 are highlighted in **bold**.

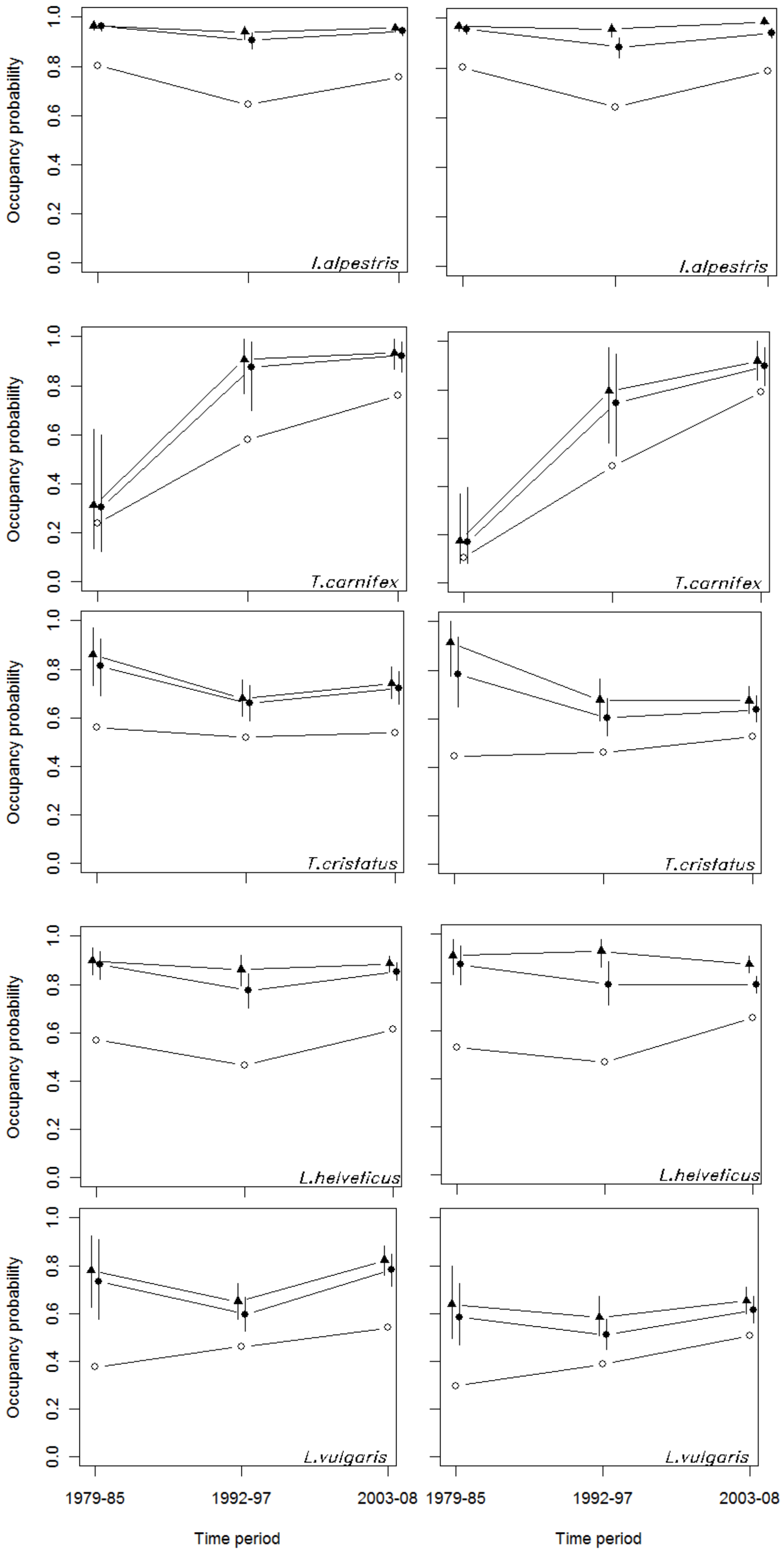
Species	**Resol.**	**CON $\psi 1$**	**CON Φ**	**CON γ**	**ALT $\psi 1$**	**ALTSQ $\psi 1$**	**URB $\Phi 1\rightarrow 2$**	**URB $\Phi 2\rightarrow 3$**	**URB $\gamma 1\rightarrow 2$**	**URB $\gamma 2\rightarrow 3$**
I.alpestris	1km	**0.70 (0.07)**	**1.03 (0.09)**	**1.87 (0.54)**	**0.43 (0.08)**	**-0.47 (0.09)**	0.18 (0.18)	0.13 (0.14)	0.44 (0.29)	0.55 (0.34)
I.alpestris	3km	**1.03 (0.10)**	**0.94 (0.13)**	**2.10 (0.80)**	0.13 (0.12)	**-0.32 (0.14)**	-0.19 (0.26)	-0.08 (0.26)	**2.64 (1.66)**	**2.33 (1.31)**
I.alpestris	5km	**1.23 (0.18)**	**1.67 (0.26)**	**1.48 (1.03)**	-0.27 (0.20)	-0.39 (0.25)	-0.58 (0.34)	**0.93 (0.53)**	0.59 (2.56)	**1.67 (1.32)**
T.carnifex	1km	**6.54 (2.05)**	3.64 (3.29)	**2.22 (0.86)**	0.08 (4.84)	-5.15 (3.12)	2.37 (5.56)	2.10 (3.24)	0.40 (0.32)	-2.02 (1.88)
T.carnifex	3km	**2.90 (2.23)**	-1.12 (2.00)	**6.12 (2.37)**	**-5.81 (2.46)**	-2.00 (2.30)	-2.64 (4.65)	-3.73 (4.04)	0.10 (2.01)	**5.71 (2.10)**
T.cristatus	1km	**0.99 (0.37)**	**0.31 (0.13)**	**1.46 (0.26)**	**-2.27 (0.60)**	-0.29 (0.40)	-0.04 (0.17)	0.34 (2.06)	-0.87 (0.84)	**1.11 (0.56)**
T.cristatus	3km	**0.22 (0.08)**	**0.53 (0.32)**	0.02 (0.73)	**-2.16 (0.64)**	0.11 (0.39)	-0.01 (0.32)	4.12 (4.30)	0.17 (1.03)	0.13 (1.25)
T.cristatus	5km	**0.35 (0.14)**	**1.15 (1.00)**	0.15 (1.18)	**-2.44 (1.19)**	-0.34 (0.68)	0.64 (0.60)	0.19 (2.24)	-0.11 (1.47)	-1.17 (1.98)
L.helveticus	1km	**0.78 (0.09)**	**0.82 (0.18)**	**1.73 (0.59)**	**-0.85 (0.11)**	**-0.57 (0.11)**	0.80 (2.24)	-0.32 (0.16)	0.39 (0.60)	0.33 (0.93)
L.helveticus	3km	**1.22 (0.18)**	**0.51 (0.15)**	1.13 (1.33)	**-1.53 (0.20)**	**-1.02 (0.22)**	3.33 (3.16)	0.04 (0.22)	1.80 (2.11)	2.36 (2.42)
L.helveticus	5km	**1.19 (0.21)**	**1.21 (1.04)**	**3.85 (1.71)**	**-1.77 (0.27)**	**-1.07 (0.32)**	3.18 (2.92)	1.14 (1.46)	0.27 (1.87)	0.81 (2.32)
L.vulgaris	1km	**0.67 (0.15)**	**2.68 (1.30)**	**1.15 (0.25)**	**-2.33 (0.67)**	**-0.98 (0.43)**	-0.52 (0.55)	-0.47 (1.88)	-0.13 (0.75)	-0.13 (0.59)
L.vulgaris	3km	**0.33 (0.10)**	**1.89 (0.82)**	0.19 (0.51)	**-3.91 (1.44)**	-1.25 (0.83)	-0.73 (0.56)	-1.03 (1.71)	0.23 (1.06)	0.47 (0.43)
L.vulgaris	5km	**0.26 (0.13)**	**0.77 (0.42)**	0.93 (0.88)	**-3.29 (1.01)**	-0.79 (0.61)	-1.16 (0.97)	1.34 (2.21)	0.24 (1.27)	-0.03 (1.46)

Appendix and Annexes

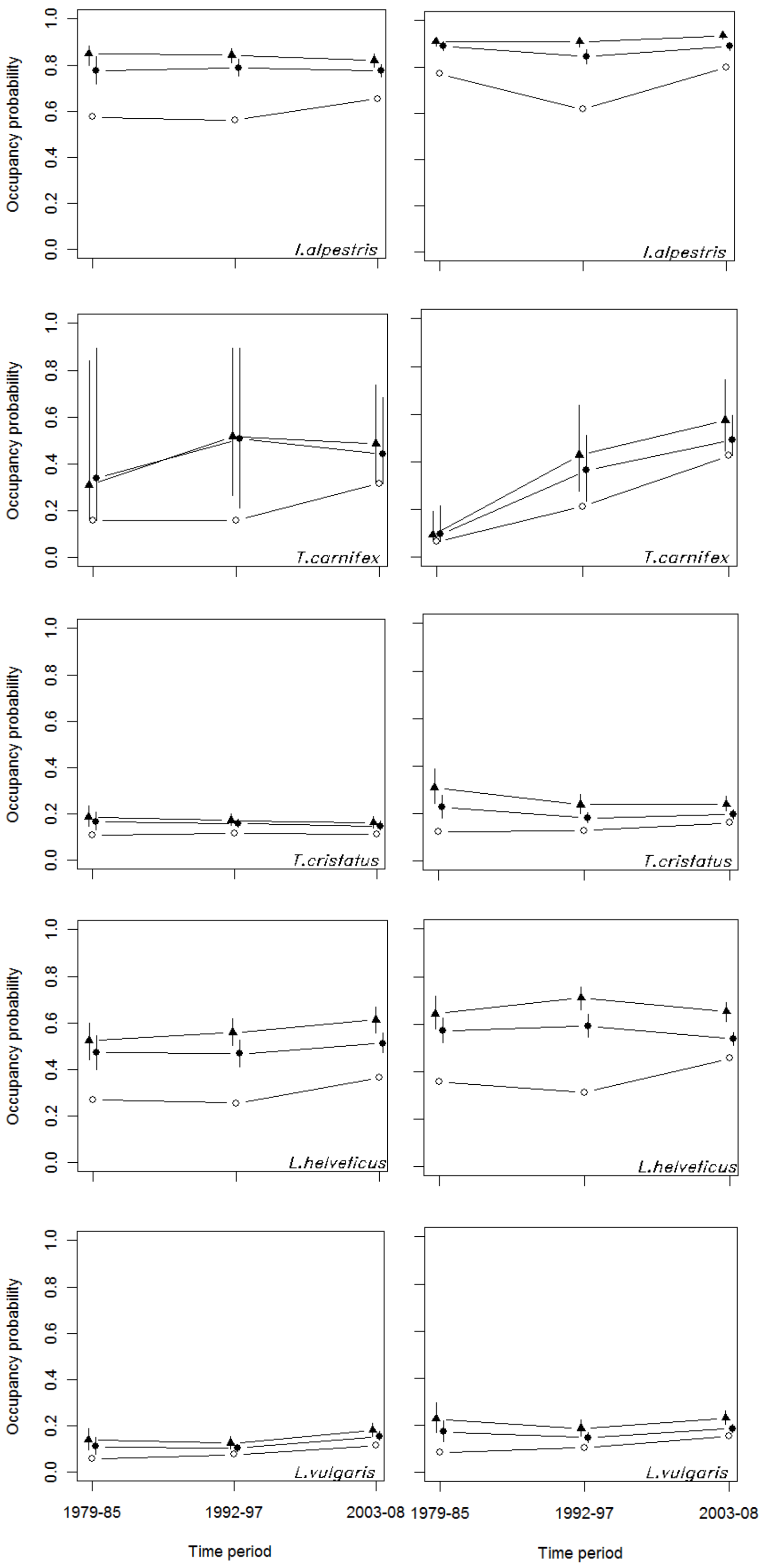
Annexes figures and tables



Annexes Figure 1 Geographic variation in the spatial arrangement of amphibian surveys at 3km resolution observed in case study 1. *I.alpestris* (top left), *T.carnifex* (top right), *T.cristatus* (bottom left), *L.vulgaris* (bottom right). Green quadrats have been visited at least once during every primary period. Red, yellow and dark blue quadrats denote the quadrats visited only during periods 1979-85, 1992-97 and 2003-08, respectively. The black dots indicate the quadrats visited during two periods out of three (aggregated for clarity).

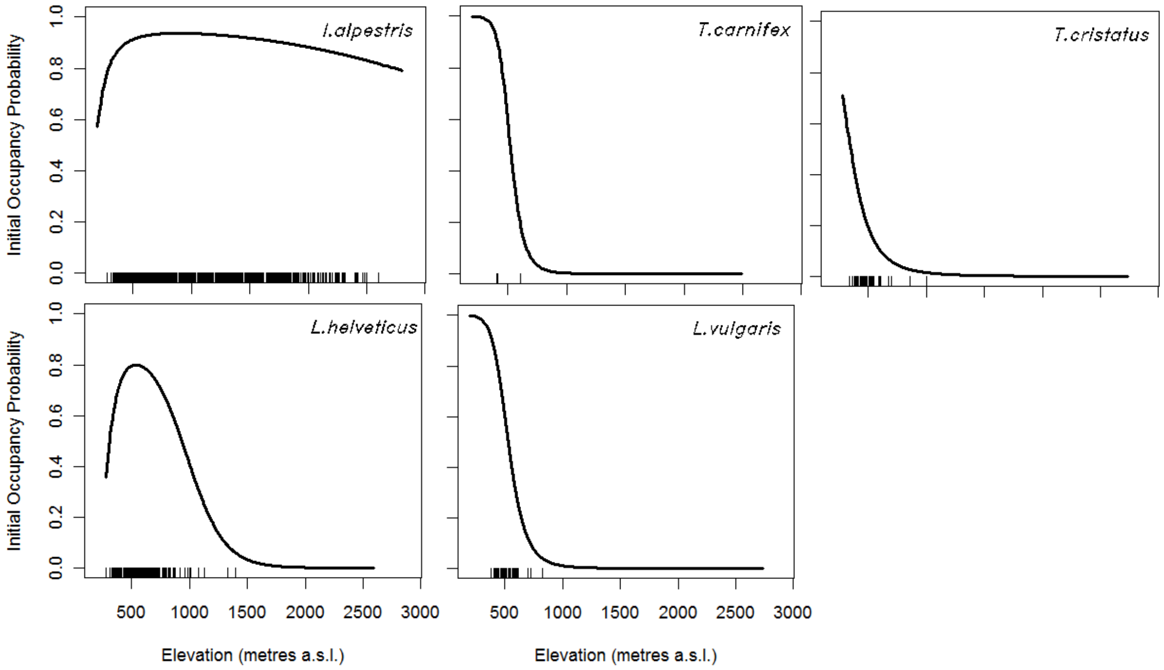


Annexes Figure 2 Observed (open circles) and estimated (filled symbols) occupancy probabilities with 95% credible intervals (vertical bars) for the three periods in case study 1. From top to bottom: *I.alpestris, T.carnifex, T.cristatus, L.helveticus, and L.vulgaris.* The left panels represent the 1km resolution, the right panels the 5km resolution. The occupancy probabilities estimated under the best (filled triangle) and the null (filled circle) detection models are shown.



Annexes Figure 3 Observed (open circles) and estimated (filled symbols) occupancy probabilities with 95% credible intervals (vertical bars) for the three periods in case study 3.

From top to bottom: *I.alpestris, T.carnifex, T.cristatus, L.helveticus, and L.vulgaris*. The left panels represent the 1km resolution, the right panels the 5km resolution. The occupancy probabilities estimated under the best (filled triangle) and the null (filled circle) detection models are shown.



Annexes Figure 4 Elevation profiles obtained from case study 1 at 3km resolution. Only the significant regression slopes from Table 7 were employed (95% credible interval not covering 0). The small ticks inside the box indicate the observed detection events in period 1979-85.

Annexes Table 1 Distribution of the within primary periods replication for every species and resolution inferred in case study 1. The replicates represent the fraction of quadrats grouped by the amount of years surveyed in each primary period: not visited (column 0), visited one year (column 1), and visited more than in two years (column ≥2).

		Period 1979-85			**Period 1992-97**			**Period 2003-08**		
Specie	**Resolution**	0	1	≥2	0	1	≥2	0	1	≥2
I.alpestris	1km	0.56	0.39	0.05	0.71	0.22	0.07	0.46	0.29	0.24
I.alpestris	3km	0.38	0.42	0.20	0.57	0.25	0.18	0.25	0.30	0.45
I.alpestris	5km	0.29	0.37	0.34	0.42	0.27	0.31	0.11	0.23	0.66
T.carnifex	1km	0.77	0.17	0.07	0.58	0.28	0.14	0.08	0.38	0.54
T.carnifex	3km	0.35	0.41	0.24	0.29	0.39	0.31	0.00	0.20	0.80
T.carnifex	5km	0.24	0.37	0.39	0.24	0.37	0.39	0.00	0.08	0.92
T.cristatus	1km	0.48	0.40	0.12	0.46	0.29	0.25	0.23	0.26	0.51
T.cristatus	3km	0.18	0.43	0.38	0.28	0.31	0.41	0.10	0.16	0.73
T.cristatus	5km	0.13	0.31	0.57	0.17	0.19	0.64	0.04	0.10	0.86
L.helveticus	1km	0.53	0.40	0.07	0.59	0.27	0.13	0.31	0.27	0.43
L.helveticus	3km	0.30	0.43	0.27	0.41	0.29	0.30	0.13	0.20	0.67
L.helveticus	5km	0.20	0.38	0.42	0.26	0.26	0.48	0.04	0.11	0.84
L.vulgaris	1km	0.50	0.37	0.13	0.44	0.32	0.24	0.25	0.29	0.46
L.vulgaris	3km	0.20	0.39	0.41	0.24	0.32	0.44	0.14	0.15	0.71
L.vulgaris	5km	0.10	0.36	0.54	0.16	0.23	0.61	0.07	0.09	0.84

Annexes Table 2 Best detection models obtained in case study 1 following a stepwise model selection. The values are on the logit scale (with standard deviation of the posterior distribution). Period 1-3: intercept for period 1-3; nrecords: number of records; nvisits: number of distinguishable visits; nvisitors: number of visitors. See the main text for details one the detection covariates. Non-significant covariates are indicated with "ns".

Species	**Resolution**	**Period 1**	**Period 2**	**Period 3**	**nrecords**	**nvisits**	**nvisitors**
I.alpestris	1km	1.45 (0.06)	0.40 (0.06)	0.53 (0.03)	ns	0.55 (0.03)	ns
I.alpestris	3km	1.08 (0.06)	0.16 (0.06)	0.33 (0.04)	0.16 (0.04)	0.78 (0.04)	ns
I.alpestris	5km	1.07 (0.06)	0.14 (0.06)	0.41 (0.04)	0.23 (0.05)	0.83 (0.05)	ns
T.carnifex	1km	2.80 (1.37)	-0.09 (0.33)	-0.13 (0.18)	-1.16 (0.25)	1.51 (0.25)	ns
T.carnifex	3km	1.10 (0.82)	-0.08 (0.29)	-0.57 (0.19)	-0.83 (0.24)	1.50 (0.27)	ns
T.carnifex	5km	1.39 (0.85)	-0.03 (0.35)	-0.50 (0.21)	-0.81 (0.30)	1.37 (0.33)	ns
T.cristatus	1km	0.33 (0.22)	0.36 (0.17)	-0.20 (0.11)	0.38 (0.07)	ns	ns
T.cristatus	3km	-0.52 (0.19)	-0.13 (0.17)	-0.51 (0.11)	0.40 (0.11)	0.35 (0.12)	ns
T.cristatus	5km	-1.09 (0.16)	-0.56 (0.16)	-0.81 (0.12)	0.61 (0.13)	0.30 (0.12)	ns
L.helveticus	1km	0.35 (0.11)	-0.32 (0.10)	-0.17 (0.06)	0.64 (0.06)	0.13 (0.06)	ns
L.helveticus	3km	-0.32 (0.10)	-0.63 (0.11)	-0.51 (0.06)	0.90 (0.07)	0.38 (0.08)	-0.15(0.05)
L.helveticus	5km	-0.59 (0.11)	-0.94 (0.10)	-0.55 (0.07)	1.21 (0.05)	ns	ns
L.vulgaris	1km	-0.31 (0.22)	0.32 (0.18)	-0.26 (0.11)	1.05 (0.13)	-0.33 (0.12)	ns
L.vulgaris	3km	-0.97 (0.18)	-0.17 (0.17)	-0.62 (0.12)	0.93 (0.08)	ns	ns
L.vulgaris	5km	-1.24 (0.21)	-0.55 (0.17)	-0.58 (0.12)	1.22 (0.14)	-0.34 (0.14)	ns

Annexes Table 3 Best detection models obtained in case study 2 following a stepwise model selection. The values are on the logit scale (with standard deviation of the posterior distribution). Period 1-3: intercept for period 1-3; nrecords: number of records; nvisits: number of distinguishable visits; nvisitors: number of visitors. See the main text for details one the detection covariates. Non-significant covariates are indicated with "ns".

Species	**Resolution**	**Period 1**	**Period 2**	**Period 3**	**nrecords**	**nvisits**	**nvisitors**
I.alpestris	1km	0.60 (0.12)	0.23 (0.09)	0.50 (0.08)	0.56 (0.06)	0.43 (0.07)	ns
I.alpestris	3km	0.92 (0.08)	0.24 (0.08)	0.42 (0.05)	0.46 (0.06)	0.76 (0.06)	ns
I.alpestris	5km	1.10 (0.08)	0.17 (0.07)	0.51 (0.05)	0.41 (0.06)	0.78 (0.06)	ns
T.carnifex	1km	1.10 (1.60)	-1.90 (0.93)	-1.08 (0.54)	ns	0.91 (0.39)	ns
T.carnifex	3km	1.48 (0.87)	-0.92 (0.40)	-1.36 (0.35)	-1.18 (0.36)	2.07 (0.43)	ns
T.carnifex	5km	1.34 (0.86)	-0.35 (0.44)	-0.82 (0.33)	-0.90 (0.35)	1.37 (0.40)	ns
T.cristatus	1km	0.22 (0.32)	-0.33 (0.22)	-0.57 (0.19)	ns	0.49 (0.10)	ns
T.cristatus	3km	-0.84 (0.20)	-0.52 (0.18)	-0.78 (0.14)	0.38 (0.12)	0.41 (0.12)	ns
T.cristatus	5km	-1.38 (0.20)	-1.02 (0.17)	-1.27 (0.14)	0.52 (0.13)	0.45 (0.13)	ns
L.helveticus	1km	-0.27 (0.19)	-0.71 (0.13)	-0.60 (0.10)	0.97 (0.07)	ns	ns
L.helveticus	3km	-0.62 (0.12)	-0.86 (0.11)	-0.79 (0.08)	1.02 (0.09)	0.40 (0.10)	-0.15 (0.07)
L.helveticus	5km	-0.69 (0.13)	-1.19 (0.09)	-0.80 (0.08)	1.11 (0.09)	0.18 (0.09)	ns
L.vulgaris	1km	-0.90 (0.31)	-0.32 (0.27)	-0.67 (0.19)	1.33 (0.19)	-0.47 (0.16)	ns
L.vulgaris	3km	-1.31 (0.21)	-0.61 (0.20)	-0.93 (0.14)	0.93 (0.09)	ns	ns
L.vulgaris	5km	-1.57 (0.22)	-0.90 (0.18)	-1.05 (0.14)	0.98 (0.08)	ns	ns

Annexes Table 4 Growth rates per transition and geometric mean growth rates estimated in case study 2. The point value and the extremes of the credible interval (2.50% and 97.50% quantiles) are represented for each quantity. The numerical labels indicate the primary time periods, the arrows the transitions between them. Significant departures from a stationary population with λ=1 are highlighted in **bold**.

Species	**Resolution**	**Growth rate λ 1→2**	**Growth rate λ 2→3**	**Geometric mean λ**
I.alpestris	1km	0.99 (0.94-1.06)	0.97 (0.93-1.02)	0.98 (0.96-1.02)
I.alpestris	3km	0.97 (0.93-1.00)	**1.05** (1.01-1.09)	1.01 (0.99-1.02)
I.alpestris	5km	1.00 (0.97-1.02)	**1.03** (1.01-1.05)	**1.01** (1.00-1.03)
T.carnifex	1km	2.12 (0.55-5.00)	1.04 (0.46-2.00)	1.38 (0.73-2.00)
T.carnifex	3km	**4.99** (2.07-8.50)	1.04 (0.76-1.45)	**2.22** (1.49-2.87)
T.carnifex	5km	**5.16** (2.00-9.33)	1.40 (0.94-2.19)	**2.60** (1.70-3.32)
T.cristatus	1km	0.93 (0.71-1.22)	0.95 (0.79-1.11)	0.93 (0.82-1.08)
T.cristatus	3km	**0.74** (0.59-0.93)	1.02 (0.88-1.18)	**0.87** (0.77-0.97)
T.cristatus	5km	**0.77** (0.60-0.99)	1.01 (0.85-1.20)	0.88 (0.78-1.00)
L.helveticus	1km	1.07 (0.94-1.27)	1.10 (0.96-1.26)	**1.08** (1.00-1.19)
L.helveticus	3km	0.99 (0.89-1.10)	1.00 (0.90-1.10)	0.99 (0.94-1.05)
L.helveticus	5km	1.11 (0.99-1.23)	**0.92** (0.86-0.99)	1.01 (0.95-1.07)
L.vulgaris	1km	0.93 (0.64-1.35)	**1.46** (1.17-1.82)	1.16 (0.96-1.40)
L.vulgaris	3km	0.84 (0.62-1.13)	1.20 (0.98-1.45)	1.00 (0.86-1.16)
L.vulgaris	5km	0.83 (0.63-1.12)	**1.25** (1.02-1.50)	1.01 (0.87-1.17)

Annexes Table 5 Best detection models obtained in case study 3 following a stepwise model selection. The values are on the logit scale (with standard deviation of the posterior distribution). Period 1-3: intercept for period 1-3; nrecords: number of records; nvisits: number of distinguishable visits; nvisitors: number of visitors. See the main text for details one the detection covariates. Non-significant covariates are indicated with “ns”. The point values were employed to compute the detection probabilities in the calculation of the variable CONNECT (see main text).

Species	**Resolution**	**Period 1**	**Period 2**	**Period 3**	**nrecords**	**nvisits**	**nvisitor**
I.alpestris	1km	0.68 (0.06)	-0.14 (0.05)	0.04 (0.04)	ns	0.73 (0.03)	ns
I.alpestris	3km	0.86 (0.06)	-0.07 (0.06)	0.09 (0.04)	0.17 (0.04)	0.86 (0.04)	ns
I.alpestris	5km	1.08 (0.07)	0.08 (0.07)	0.30 (0.05)	0.18 (0.05)	0.94 (0.06)	ns
T.carnifex	1km	2.40 (1.40)	-0.68 (0.37)	-0.65 (0.26)	-0.71 (0.20)	1.18 (0.23)	ns
T.carnifex	3km	-0.54 (1.13)	0.33 (0.43)	-0.82 (0.32)	-0.63 (0.28)	1.59 (0.36)	ns
T.cristatus	1km	-1.24 (0.22)	-0.59 (0.19)	-1.19 (0.14)	0.50 (0.09)	0.27 (0.08)	ns
T.cristatus	3km	-1.35 (0.24)	-0.53 (0.20)	-1.22 (0.15)	0.40 (0.11)	0.50 (0.11)	ns
T.cristatus	5km	-1.73 (0.23)	-1.05 (0.20)	-1.73 (0.17)	0.41 (0.14)	0.61 (0.14)	ns
L.helveticus	1km	-1.03 (0.12)	-1.22 (0.09)	-1.06 (0.06)	0.85 (0.05)	0.18 (0.05)	ns
L.helveticus	3km	-1.03 (0.11)	-1.29 (0.10)	-1.13 (0.07)	0.92 (0.07)	0.43 (0.08)	-0.15 (0.05)
L.helveticus	5km	-0.91 (0.12)	-1.30 (0.10)	-1.09 (0.08)	1.03 (0.08)	0.37 (0.09)	-0.12 (0.06)
L.vulgaris	1km	-1.68 (0.26)	-0.96 (0.19)	-1.22 (0.13)	1.16 (0.10)	-0.28 (0.09)	ns
L.vulgaris	3km	-1.93 (0.24)	-0.86 (0.21)	-1.53 (0.15)	1.02 (0.08)	ns	ns
L.vulgaris	5km	-1.88 (0.25)	-1.29 (0.21)	-1.57 (0.16)	1.01 (0.08)	ns	ns

Annexes Table 6 Occupancy probabilities estimated from the best ($p(\text{covariates})\psi(.)\Phi(t)\gamma(t)$) and the null ($p(t)\psi(.)\Phi(t)\gamma(t)$) detection models in case study 3. Only the extremes of the 95% credible intervals without the point value are shown. Bold indicates the cases of significant differences between occupancy estimates from null and best models on the basis of the 95% credible interval. The point values of the occupancy probabilities obtained by the best models were employed in the calculation of the CONNECT covariate (see main text).

		Period 1979-85				**Period 1992-97**				**Period 2003-08**			
		Best model		Null model		Best model		Null model		Best model		Null model	
Species	**Resolution**	2.50%	97.50%	2.50%	97.50%	2.50%	97.50%	2.50%	97.50%	2.50%	97.50%	2.50%	97.50%
I.alpestris	1km	0.78	0.85	0.71	0.78	**0.78**	**0.84**	**0.66**	**0.73**	**0.75**	**0.80**	**0.65**	**0.70**
I.alpestris	3km	0.86	0.91	0.81	0.87	**0.83**	**0.90**	**0.70**	**0.79**	**0.86**	**0.91**	**0.74**	**0.79**
I.alpestris	5km	0.87	0.92	0.83	0.88	**0.86**	**0.92**	**0.76**	**0.84**	**0.89**	**0.94**	**0.80**	**0.85**
T.carnifex	1km	0.02	0.11	0.02	0.11	0.18	0.40	0.15	0.32	0.21	0.37	0.18	0.31
T.carnifex	3km	0.03	0.47	0.04	0.44	0.16	0.36	0.15	0.30	0.18	0.36	0.15	0.28
T.cristatus	1km	0.09	0.17	0.06	0.11	**0.08**	**0.12**	**0.05**	**0.07**	**0.08**	**0.11**	**0.05**	**0.07**
T.cristatus	3km	0.12	0.23	0.08	0.15	0.08	0.13	0.06	0.09	**0.09**	**0.13**	**0.06**	**0.09**
T.cristatus	5km	0.20	0.39	0.12	0.22	0.15	0.24	0.10	0.15	**0.14**	**0.22**	**0.10**	**0.14**
L.helveticus	1km	**0.38**	**0.51**	**0.25**	**0.33**	**0.41**	**0.52**	**0.26**	**0.33**	**0.41**	**0.49**	**0.27**	**0.32**
L.helveticus	3km	**0.50**	**0.64**	**0.38**	**0.48**	**0.54**	**0.65**	**0.36**	**0.46**	**0.50**	**0.59**	**0.34**	**0.39**
L.helveticus	5km	0.55	0.69	0.45	0.56	**0.62**	**0.73**	**0.46**	**0.57**	**0.57**	**0.66**	**0.41**	**0.47**
L.vulgaris	1km	0.08	0.16	0.04	0.09	**0.09**	**0.13**	**0.05**	**0.07**	**0.10**	**0.14**	**0.06**	**0.09**
L.vulgaris	3km	0.12	0.26	0.08	0.17	**0.10**	**0.17**	**0.06**	**0.10**	**0.12**	**0.18**	**0.09**	**0.12**
L.vulgaris	5km	0.14	0.29	0.10	0.20	0.14	0.25	0.09	0.15	**0.16**	**0.25**	**0.12**	**0.16**

1. Short introduction to site occupancy models

The so-called "presence-absence" data are usually employed in combination with Generalized Linear Models when modelling species distribution as a function of environmental characteristics with a variety of purposes (Guisan and Zimmermann 2000). These data arise from two, distinct stochastic processes: the first determining the true state of occurrence and non-occurrence at a given site, and the second one yielding the observed detection-nondetection data (MacKenzie *et al.* 2002). Thus, presence-absence data should be called detection-nondetection data.

Assume a set of *i* sites in which a species is distributed with probability of occurrence ψ, we can formulate the true occupation status z ($z = 1$ for occupied, $z = 0$ for unoccupied) at a given spatial unit as a Bernoulli trial equal to a single coin flip (Royle and Kery 2007)

$$z_i \sim Bernoulli(\psi_i)$$

However, the observation of the previous process is not free of errors due to imperfect detection. Thus, some of the sampled zeroes may represent undetected, occupied sites with probability $1 - p_{ij}$, where p_{ij} corresponds to the detection probability for site i during visit j, given occurrence. The observed status of occupancy y for site i during visit j can be therefore described as another coin flip (Royle and Kery 2007)

$$y_{ij}|z_i \sim Bernoulli(z_i\ x\ p_{ij})$$

When simple logistic regressions are employed among the set of sites considered, the product of both the occurrence and the detectability of the species is modelled. Repeated visits, attempting to measure the maximal response, fail to reach a reasonable threshold when detection probabilities are too low (i.e. declaring that a site is occupied if at least one visit reveals a detection event). Moreover, a priori information about the required number of visits is not always available or comparable. Indeed, even the studies relying on a standardized field protocol may fail to produce constant detection probabilities over time (Schmidt 2005).

Site occupancy models are a category of models derived from the theory underlying the mark recapture studies and allow separate inferences about the two processes previously described (MacKenzie *et al.* 2002). They exploit the data resulting from the multiple visits (i.e. secondary periods) carried to some sites within a period of closure, namely a primary period in which the occupancy status of the site does not change. The disaggregated strings of detection-nondetection are directly analysed and the resulting estimations of detection probability are used for correcting the estimations of occupancy. Consider a site that is visited four times and the species of interest goes undetected during the second and third surveys: the resulting history corresponds to “1001” with probability $\psi p^2(1-p)^2$. Another site visited four times in which none of the surveys reported a positive result holds the history “0000”: the probability of the latter equals to $(1 - \psi) + \psi(1-p)^4$, which is the probability that the site is not occupied added to the probability of nondetection over four visits for an occupied site. Relevant covariates for detection according to the ecology of the species can be integrated via the logit link function to obtain estimations of the detection probability that vary for each site and visit. Similarly, ecological covariates describing the occupancy probability in each site can be added in the same manner to the parameter ψ.

The previous framework has firstly been described by MacKenzie *et al.* (2002) and applies to a single period of closure. Extensions of the ecological model allowing changes in the status of occupancy are possible. In this case, multiple primary periods are defined. The changes in occupancy from an occupied status to unoccupied and vice versa are not allowed within, but can occur between them. Dynamic formulations of these open population models among primary periods can be described by (MacKenzie *et al.* 2003)

$$\Psi_{i,k} = \Psi_{i,k-1} * \Phi_{\mathrm{i,k-1}} + (1 - \Psi_{i,k-1}) * \gamma_{\mathrm{i,k-1}}$$

where $\Psi_{i,k}$ denotes the occupancy probability during period *k*, $\Phi_{\mathrm{i,k-1}}$ the survival and $\gamma_{i,k-1}$ the colonization probability of site *i* from period *k-1* to *k* for the same site. Note that the occupancy estimations for every period are calculated recursively from the previous one as a function of survival and colonization probabilities; only the occupancy in the first period is an independent initial state. Similar to the single season model, replicated visits within every period of closure are required to estimate the detection probability. More interestingly, estimations of the biological processes underlying the distributional changes are directly estimable using the dynamic formulation. Furthermore, the factors thought to affect the survival or the colonization rates are testable integrating the relevant covariates using the logit link function. Govindan *et al.* (2011) provide a case study of the application of these models on the dynamic of acorn weevils, as well as a full description of the likelihood specification in a bayesian modelling framework. For the most interested reader, Royle and Kery (2007) described in detail the state-space formulation of dynamic population models, explaining the basis for the computation of secondary population parameters and highlighting the subtle difference between classic and finite sample estimators of those quantities.

2. Species considered in the creation of detection-nondetection histories

Genus and species of the permanent pond amphibians employed for checklist reconstruction.

Genus	species	Genus	species	Genus	species
Alytes	*obstetricans*	*Hyla*	*intermedia*	*Rana*	*temporaria*
Bombina	*variegata*	*Pelophylax*	*esculentus*	*Ichtyosaura*	*alpestris*
Bufo	*bufo*	*Pelophylax*	*lessonae*	*Triturus*	*carnifex*
Bufo	*calamita*	*Pelophylax*	*ridibundus*	*Triturus*	*cristatus*
Bufo	*viridis*	*Rana*	*dalmatina*	*Lissotriton*	*helveticus*
Hyla	*arborea*	*Rana*	*latastei*	*Lissotriton*	*vulgaris*

Schematic representation of the method creating the detection-nondetection matrices from presence-only sources (modified from Kery *et al.*, 2010b). Observations from three hypothetical quadrats in a single primary period are represented. TRAL: *I.alpestris*, TRHE: *L.helveticus*, BUBU: *B.bufo*, RATE: *R.temporaria*. Missing visits are denoted by the mark "-" and do not contribute to the analysis. The detection-nondetection history for *I.alpestris* in quadrat 1 corresponds to "10001"; the species was detected during the first and the last visit, whereas other species were observed in visits 2, 3 and 4. Quadrat 2 was sampled only twice during the whole period, therefore the detection-nondetection history is shorter than the previous, corresponding to "- 0 - 1 -". The third quadrat has not been visited during the time period considered, but it has been sampled at least once during another period.

Quadrat	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5
1	RATE,BUBU,TRAL	TRHE	BUBU	TRHE,BUBU	TRAL
2	-	TRVU	-	TRVU,TRAL	-
3	-	-	-	-	-

3. Case study 3 – Landscape variables for occupancy modelling

The aim of this case study is to test a small set of environmental covariates on the vital parameters describing the occupancy dynamic for the three periods. The tested covariates corresponded to elevation, urban cover and connectivity. The latter was prepared into the software R 2.13 (R-Development-Core-Team 2010), whereas the first two in a Geographic Information System. I performed all the spatial analyses using ArcGIS 9.3 (ESRI 2008). The data source is the Unilgis server – the server for geographical data for the University of Lausanne. I calculated the landscape covariates using a focal averaging moving window of size equivalent to the quadrat: e.g. at 3 km resolution, all the variables have been derived by a square window of side 3x3 km.

In this particular case study, I considered only the quadrats entirely covering the country, since the older source for land use is restricted to the Swiss territory (i.e. Geostat layer as97cent). Geostat (OFS 2007) and Vector25 (Swisstopo 2007) layers followed two different reclassification procedures, because they do not rely on the same codification. The direct comparison of land use values between the two data sources is therefore not allowed. For the same reason, their use as predictors in distribution models is subject to the constraint of a source specific slope estimate. This problem is likely to find a solution in year 2012, when the complete information about land use in a compatible Geostat format will be available for the whole country and the defined time periods.

Landscape variables for occupancy modelling

I created reclassified maps describing the urban cover to be analysed by focal averaging window, thus deriving the % of the quadrat covered by the particular land use.

GEOSTAT layer

Update time: 1992-1997 (as97cent).

Resolution: Original data is a raster layer at 100 metres resolution.

Reference(s):

- OFS (2007). Description de données GEOSTAT: Statistique de la superficie NOAS92. D. F. d. l'Intérieur.

Geostat classes chosen for reclassification: 21,24-29,37,41-49

Vector25 layer

Update time: 2003-2008.

Resolution: Original data employed in this study include shapefiles, polygons and polylines. Precision is estimated at 3-8 metres.

Reference(s):

- Swisstopo (2007). VECTOR25: Le modèle numérique du territoire de la Suisse. OFT.
- Swisstopo (2008). Mise à jour VECTOR25: Rapport de release 2008. OFT.

Note: In order to allow the analysis of a vector layer, the reclassified maps were rasterized to a resolution of 10 metres prior to the focal statistics.

Vector 25 classes chosen for reclassification: from the layer describing the primary surface (pri25_a), I retained all the polygon objects belonging to the Z_Siedl class.

Elevation layer

Update time: NA.

Resolution: The layer is coded as a raster at 25 metres resolution. It has been derived from the 1 : 25 000 national map and reflects its accuracy (maximal average error estimated at 3 metres in the central Alps).

Reference(s):

- Office Fédéral de la Topographie (Swisstopo), I analysed the original Digital Elevation Model (layer mnt25_ch) by focal averaging moving window to derive the mean value for each quadrat.

Several other land use and climatic variables were prepared for their integration into the vital parameters of each species; unfortunately the resulting models were too complex to achieve the convergence. For a full list of the original covariates and the relative reclassification maps, see the CD furnished to the supervisors Dr. B.R.Schmidt (Benedikt.Schmidt@unine.ch) and Prof. N.Perrin (Nicolas.Perrin@unil.ch).

Literature cited

Arntzen, J. W. and R. S. Thorpe (1999). "Italian crested newts (Triturus carnifex) in the Basin of Geneva: Distribution and genetic interactions with autochthonous species." Herpetologica **55**(4): 423-433.

Augustin, N. H., et al. (1996). "An autologistic model for the spatial distribution of wildlife." Journal of Applied Ecology **33**(2): 339-347.

Austin, M. P. (2002). "Spatial prediction of species distribution: an interface between ecological theory and statistical modelling." Ecological Modelling **157**(2-3): 101-118.

Box, G. E. P. and D. R. Cox (1964). "An Analysis of Transformations." Journal of the Royal Statistical Society Series B-Statistical Methodology **26**(2): 211-252.

Cirovic, R., et al. (2008). "Distribution patterns and environmental determinants of European newts in the Montenegrin karst area." Biologia **63**(5): 745-752.

Crozier, L. (2004). "Warmer winters drive butterfly range expansion by increasing survivorship." Ecology **85**(1): 231-241.

Denoel, M. and G. F. Ficetola (2008). "Conservation of newt guilds in an agricultural landscape of Belgium: the importance of aquatic and terrestrial habitats." Aquatic Conservation-Marine and Freshwater Ecosystems **18**(5): 714-728.

Dorazio, R. M. (2007). "On the choice of statistical models for estimating occurrence and extinction from animal surveys." Ecology **88**(11): 2773-2782.

Eigenbrod, F., et al. (2011). "Sub-optimal study design has major impacts on landscape-scale inference." Biological Conservation **144**(1): 298-305.

ESRI (2008). ArcGIS Desktop 9.3. Redlands, CA, Environmental System Research Institute.

Gardner, T. (2001). "Declining amphibian populations: a global phenomenon in conservation biology." Animal Biodiversity & Conservation **2**(2): 25-44.

Govindan, B. N., et al. (2011). "Host selection and responses to forest fragmentation in acorn weevils: Inferences from dynamic occupancy models." Oikos.

Grossenbacher, K. (1988). Verbreitungsatlas der Amphibien der Schweiz. SBN Basel.

Guisan, A. and N. E. Zimmermann (2000). "Predictive habitat distribution models in ecology." Ecological Modelling **135**(2-3): 147-186.

Hamer, A. J. and M. J. McDonnell (2008). "Amphibian ecology and conservation in the urbanising world: A review." Biological Conservation **141**(10): 2432-2449.

Hartel, T., et al. (2010). "Using connectivity metrics and niche modelling to explore the occurrence of the northern crested newt Triturus cristatus (Amphibia, Caudata) in a traditionally managed landscape." Environmental Conservation **37**(2): 195-200.

Houlahan, J. E., et al. (2000). "Quantitative evidence for global amphibian population declines." Nature **404**(6779): 752-755.

Humbel, R. (1998). GIS use in the framework of a national statistical office: Establishment land use statistics and geocoding of national censuses. Agricultural Statistics 2000, Proceedings - an International Conference on Agricultural Statistics. T. E. Holland and M. P. R. VandenBroecke. Voorburg, Int Statistical Inst**:** 291-304.

Johnson, P. T. J., et al. (2011). "Regional Decline of an Iconic Amphibian Associated with Elevation, Land-Use Change, and Invasive Species." Conservation Biology **25**(3): 556-566.

Joly, P., et al. (2001). "Habitat matrix effects on pond occupancy in newts." Conservation Biology **15**(1): 239-248.

Kendall, W. L. (1999). "Robustness of closed capture-recapture methods to violations of the closure assumption." Ecology **80**(8): 2517-2525.

Kery, M. (2010). Introduction to WinBUGS for ecologists - A Bayesian approach to regression, ANOVA, mixed models and related analyses. 6204 Sempach, Switzerland.

Kery, M., et al. (2010a). "Predicting species distributions from checklist data using site-occupancy models." Journal of Biogeography **37**(10): 1851-1862.

Kery, M., et al. (2010b). "Site-Occupancy Distribution Modeling to Correct Population-Trend Estimates Derived from Opportunistic Observations." Conservation Biology **24**(5): 1388-1397.

Kery, M. and M. Schaub (2011). Bayesian Population Analysis using WinBUGS - A hierarchical perspective. 6204 Sempach, Switzerland.

Kery, M. and B. R. Schmidt (2008). "Imperfect detection and its consequences for monitoring for conservation." Community Ecology **9**(2): 207-216.

Kindvall, O. and I. Ahlen (1992). "Geometrical Factors and Metapopulation Dynamics of the Bush Cricket, Metrioptera bicolor Philippi (Orthoptera: Tettigoniidae)." Conservation Biology **6**(4): 520-529.

Krauss, J., et al. (2004). "Landscape occupancy and local population size depends on host plant distribution in the butterfly Cupido minimus." Biological Conservation **120**(3): 355-361.

Lister, A. M. (2011). "Natural history collections as sources of long-term datasets." Trends in Ecology & Evolution **26**(4): 153-154.

MacKenzie, D. I. (2005). "What are the issues with presence-absence data for wildlife managers?" Journal of Wildlife Management **69**(3): 849-860.

MacKenzie, D. I., et al. (2003). "Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly." Ecology **84**(8): 2200-2207.

MacKenzie, D. I., et al. (2002). "Estimating site occupancy rates when detection probabilities are less than one." Ecology **83**(8): 2248-2255.

Marsh, D. M. and P. C. Trenham (2001). "Metapopulation dynamics and amphibian conservation." Conservation Biology **15**(1): 40-49.

Meyer, A., et al. (2009). Les amphibiens et les reptiles de Suisse. Bern, Haupt.

Nichols, J. D. and B. K. Williams (2006). "Monitoring for conservation." Trends in Ecology & Evolution **21**(12): 668-673.

OFS (2007). Description de données GEOSTAT: Statistique de la superficie NOAS92. D. F. d. l'Intérieur.

Pellet, J., et al. (2004). "A concentric analysis of the impact of urbanization on the threatened European tree frog in an agricultural landscape." Conservation Biology **18**(6): 1599-1606.

Pellet, J. and B. R. Schmidt (2005). "Monitoring distributions using call surveys: estimating site occupancy, detection probabilities and inferring absence." Biological Conservation **123**(1): 27-35.

Price, S. J., et al. (2005). "The importance of spatial scale for conservation and assessment of anuran populations in coastal wetlands of the western Great Lakes, USA." Landscape Ecology **20**(4): 441-454.

R-Development-Core-Team (2010). R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

Rannap, R., et al. (2009). "Restoring ponds for amphibians: a success story." Hydrobiologia **634**(1): 87-95.

Rota, C. T., et al. (2009). "Occupancy estimation and the closure assumption." Journal of Applied Ecology **46**(6): 1173-1181.

Royle, J. A. and R. M. Dorazio (2008). Hierarchical modeling and inference in ecology: the analysis of data from populations and communities, Academic Press.

Royle, J. A. and M. Kery (2007). "A Bayesian state-space formulation of dynamic occupancy models." Ecology **88**(7): 1813-1823.

Royle, J. A. and J. D. Nichols (2003). "Estimating abundance from repeated presence-absence data or point counts." Ecology **84**(3): 777-790.

Schmeller, D. S., et al. (2009). "Advantages of Volunteer-Based Biodiversity Monitoring in Europe." Conservation Biology **23**(2): 307-316.

Schmidt, B. R. (2005). "Monitoring the distribution of pond-breeding amphibians when species are detected imperfectly." Aquatic Conservation-Marine and Freshwater Ecosystems **15**(6): 681-692.

Schmidt, B. R. and J. Pellet (2005). "Relative importance of population processes and habitat characteristics in determining site occupancy of two anurans." Journal of Wildlife Management **69**(3): 884-893.

Schmidt, B. R. and S. Zumbach (2005). Liste Rouge des amphibiens menacés en Suisse. Série OFEFP: l'environnement pratique. Berne, Switzerland, Centre de coordination pour la protection des amphibiens et des reptiles de Suisse (karch) and Office fédéral de l'environnement, des forêts et du paysage (OFEFP).

Smith, M. A. and D. M. Green (2005). "Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations?" Ecography **28**(1): 110-128.

Spiegelhalter, D., et al. (2003). WinBUGS User Manual. MCR Biostatistics Unit, Institute of Public Health, Robinson Way, Cambridge CB2 2SR, UK.

Stuart, S. N., et al. (2004). "Status and trends of amphibian declines and extinctions worldwide." Science **306**(5702): 1783-1786.

Sturtz, S., et al. (2005). "R2WinBUGS: A package for running WinBUGS from R." Journal of Statistical Software **12**(3): 1-16.

Swisstopo (2007). VECTOR25: Le modèle numérique du territoire de la Suisse. OFT.

Tanadini, L. G. and B. R. Schmidt (2011). "Population Size Influences Amphibian Detection Probability: Implications for Biodiversity Monitoring Programs." PLoS ONE **6**(12): e28244.

Van Buskirk, J. (2005). "Local and landscape influence on amphibian occurrence and abundance." Ecology **86**(7): 1936-1947.

van Strien, A. J., et al. (2010). "Site-occupancy models may offer new opportunities for dragonfly monitoring based on daily species lists." Basic and Applied Ecology **11**(6): 495-503.

van Strien, A. J., et al. (2011). "Metapopulation dynamics in the butterfly Hipparchia semele changed decades before occupancy declined in The Netherlands." Ecological Applications **21**(7): 2510-2520.

van Swaay, C. A. M. (1990). "An assessment of the changes in butterfly abundance in The Netherlands during the 20th Century." Biological Conservation **52**(4): 287-302.

van Swaay, C. A. M., et al. (2002). Monitoring butterflies in the Netherlands: How to get unbiased indices. Proceedings of the Experimental & Applied Entomology, NEV Amsterdam.

van Turnhout, C. A. M., et al. (2008). "Monitoring common and scarce breeding birds in the Netherlands: applying a post-hoc stratification and weighting procedure to obtain less biased population trends." Revista Catalana d'Ornitologia **24**: 15:29.

Weber, D., et al. (2004). "Scale and trends in species richness: considerations for monitoring biological diversity for political purposes." Global Ecology and Biogeography **13**(2): 97-104.

Wood, C., et al. (2011). "eBird: Engaging Birders in Science and Conservation." PLoS Biol **9**(12): e1001220.

Yoccoz, N. G., et al. (2001). "Monitoring of biological diversity in space and time." Trends in Ecology & Evolution **16**(8): 446-453.

Zanini, F. (2006). Amphibian conservation in human shaped environments: landscape dynamics, habitat modeling and metapopulation analyses. EPFL. Lausanne. **PhD Thesis:** 1-201.

Zanini, F., et al. (2009). "The transferability of distribution models across regions: an amphibian case study." Diversity and Distributions **15**(3): 469-480.

1424