Sub-optimal study design has major impacts on landscape-scale inference

Felix Eigenbrod [a,*], Stephen J. Hecnar [b], Lenore Fahrig [a]

[a] Geomatics and Landscape Ecology Research Laboratory (GLEL), Department of Biology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada K1S 5B6
[b] Department of Biology, Lakehead University, 955 Oliver Rd, Thunder Bay, Ontario, Canada P7B 5E1

ARTICLE INFO

ABSTRACT

Landscape-scale processes (e.g. habitat loss) are major drivers of the global biodiversity crisis, but the complexity and size of landscapes makes study design at this scale difficult. However, the impact of statistical problems associated with sub-optimal study design on inferences drawn from landscape-scale studies is poorly understood. Here, we examine how three common statistical 'pitfalls' associated with sub-optimal study design – (1) using landscapes that overlap in space; (2) using only a portion of the potential range of the landscape predictor variable(s) of interest; (3) failing to account for correlations among landscape predictor variables – affect the inferred relationships between the abundances of six species of anurans and the amount of forest in the landscape using a large ($n = 1141$) empirical dataset from Wisconsin and Michigan, USA. We show that sub-optimal study design alone can be sufficient to cause a switch in the sign of the inferred relationship between a species response and landscape structure, and that using only a portion of the potential range of a predictor variable, and correlations between predictor variables, are particularly likely to affect inferences. Our results also provide the first evidence of a non-monotonic relationship between forest amount and gray treefrog abundance, and suggest that inconsistencies in the literature about the inferred relationships between anuran presence/abundance and forest amount in the Great Lakes basin are likely largely due to sampling design issues. Increased attention to study design is therefore necessary for the development of robust generalizations in landscape ecology.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Over the past 20 years, landscape ecology has led to major strides in understanding the impacts of landscape-scale human activities on biodiversity (Turner, 2005). However, the complexity and size of landscapes means that study design at this scale is challenging (McGarigal and Cushman, 2002), making it difficult for investigators to avoid common statistical 'pitfalls' associated with sub-optimal study design. Spatial autocorrelation (Legendre, 1993), lack of replication or pseudoreplication (Hulbert, 1984), and multicollinearity of predictor variables (e.g. Graham, 2003) are statistical pitfalls that are problematic throughout ecology, while a failure to conduct a study at a scale (spatial extent) appropriate for the species and process being studied (Wiens, 1989; Holland et al., 2004) is an additional problem particularly relevant to landscape ecology. Finally, logistical restrictions in the choice of sampling units mean that variation in landscape-scale predictor variables of interest is often low relative to the potential range of the predictor (Brennan et al., 2002).

It is likely that sub-optimal study design has a major impact on inferences drawn from landscape ecological studies. McGarigal and Cushman (2002) reviewed the literature on empirical studies on landscape fragmentation and concluded that sub-optimal study design was a major contributor to the lack of a consensus within this field. However, the lack of consensus on the effects of habitat fragmentation may also be due to the large number of questions that can be asked about this subject and the many definitions of "fragmentation" in the literature (Fahrig, 2003). Thus, the question remains: to what extent are the conclusions of empirical studies in landscape ecology compromised by sub-optimal study design?

Here, we provide the first empirical examination of how different elements of sampling design can affect inferred relationships between landscape structure and species responses. As a test case, we evaluate the effects of three major statistical pitfalls on the inferred relationships between the amount of forest in the landscape and abundances of six species of anurans, using a large empirical dataset from the Great Lakes basin. The three pitfalls we consider are: (1) using landscapes that overlap in space (non-independence; pseudoreplication); (2) using only a portion of the potential range of the landscape predictor variable(s) of interest; and (3) failing to

* Corresponding author. Address: School of Biological Sciences, University of Southampton, Life Sciences Building (B85), Highfield Campus, Southampton SO17 1BJ, UK. Tel.: +44 (0)2380 597382.
E-mail addresses: f.eigenbrod@soton.ac.uk (F. Eigenbrod), stephen.hecnar@lakeheadu.ca (S.J. Hecnar), lenore_fahrig@carleton.ca (L. Fahrig).

account for correlations among landscape predictor variables. We use real ecological data rather than a simulated dataset in this study, because our goal is to determine whether sub-optimal study design is likely to lead to large errors in inferences in actual field studies. We chose these three pitfalls because these relate to three of the key aspects of experimental design in landscape ecology as identified by Brennan et al. (2002) and McGarigal and Cushman (2002).

Overlapping landscapes are a form of pseudoreplication (Hulbert, 1984) because values of predictor variables from nearly the same landscapes are used as multiple observations in the dataset; the degree of pseudoreplication depends on the degree of overlap between landscapes. Pseudoreplication will result in non-independence of residuals, thus increasing the likelihood of making Type I errors by causing a systematic underestimation of confidence intervals (Legendre, 1993). Overlapping landscapes can also lead to lower variation in the predictor variable (Brennan et al., 2002), thus reducing the statistical power to detect an effect.

Overlapping landscapes are relatively common in landscape ecological studies due to logistic constraints. For example, Gibbs et al. (2005) examined the effects of (among other things) land cover change at up to 10 km radii from sampling points on population transitions in anuran populations in upstate New York by re-sampling sites in 2001–2002 that were originally surveyed 1973–1980. As the goal of this analysis was a historical comparison, site selection was constrained by the 1973–1980 survey. Sites in this first survey were clustered into short survey routes, so many sites in this analysis are located less than 20 km apart (the minimum distance required to ensure fully non-overlapping landscapes at the scale of Gibbs et al.'s analyses).

The chance of detecting an effect of a predictor will also be greatly reduced if only a portion of its potential range is considered (Brennan et al., 2002). If there is a non-monotonic effect of a predictor, then a limited range of predictor variables could additionally lead to contradictory findings, as the slope of the relationship between the response and the predictor will vary depending on the range of the predictor value.

Correlations among predictors are a problem throughout ecology (e.g. Freckleton, 2002; Graham, 2003), but particularly so for landscape ecology (McGarigal and Cushman, 2002; Smith et al., 2009). High correlations among predictors mean it is impossible to know which of the related predictors are in fact responsible for a given effect on a species, and can lead to erroneous inferences.

To control for correlations, investigators often use multiple regression models. This reduces the likelihood of making incorrect inferences, but can still lead to the erroneous conclusion that there is no effect of a predictor, because variation shared between predictors (where they co-vary) is not included in the estimates of the effect of each individual predictor, reducing statistical power. In addition, multicollinearity in predictors can still lead to inaccurate model parameterization even when a multiple regression is used (Graham, 2003).

We do not consider the other three important issues in experimental design in landscape ecology identified by Brennan et al. (2002) and McGarigal and Cushman (2002) here. These three issues are: (1) a failure to account for large-scale gradients in environmental variables (which can lead to problems of spatial autocorrelation even if landscapes are non-overlapping (Schooley, 2006)); (2) a failure to select the appropriate landscape extent for the study; and (3) small samples sizes. Preliminary analyses showed that large-scale environmental gradients have little effect on the relationship between anuran abundance and forest amount in our dataset, so we were unable to test the effects of this particular 'pitfall' here. Similarly, we were restricted to landscape extents of 500–5000 m radii because (1) the response variable was the number of anurans calling within hearing distance (approximately 500 m) of the survey site (Mossman et al., 1998) and (2) selecting non-overlapping landscapes at extents larger than 5000 m would have severely restricted site selection. The 500–5000 m range of radii corresponds to scale at which anurans are generally thought to respond to the amount of forest in the landscape (Cushman, 2006), but covers only one order of magnitude; Holland et al. (2004) showed shifts in the sign of the relationship between the abundance of a beetle and the amount of forest in the landscape when the latter was measured over 20–2000 m, i.e., two orders of magnitude. Preliminary analyses (Table A1 in the Appendix) confirmed that the scale at which the amount of forest cover in the landscape was measured, within 500–5000 m radius, had little effect on the relationships between anuran relative abundance and forest amount in our data set, so we were unable to test the effects of this 'pitfall'. Finally, we did not test the effect of sample size since, in isolation, reduced sample size simply leads to a loss of statistical power.

The relationships between forest cover and anurans in the Great Lakes basin are a particularly suitable choice for this study because for six anuran species there are 13 published landscape-scale studies in this region relating forest amount to anuran abundance. For two of these species (gray treefrog and American toad) the various studies give contradictory results, and for two other species (leopard frog and green frog) 'no effect' is the most common finding (Table 1). Are these differences and lack of effects real, or are they likely artefacts of study design?

2. Materials and methods

2.1. Data sources

For the response variables, we used relative anuran abundance data at 1141 survey sites (minimum distance between sites is 188 m; maximum distance is 839,200 m) from two large-scale volunteer-based anuran monitoring programs – the Wisconsin Frog and Toad Survey and the Michigan Frog and Toad Survey (Fig. 1). The two surveys use a very similar and well-established protocol (Mossman et al., 1998). Trained volunteers conduct nocturnal call surveys under suitable weather conditions (warm humid nights with little wind) at or near wetlands (10 sites per route) three times a year (early spring, late spring and early summer). Each call survey lasts a minimum of 3 min in Michigan and 5 min in Wisconsin, with up to 10 min allowed in both cases, so the observer can be confident that all calls were recorded. The estimated relative abundance of anurans is then assigned to one of four classes: 0 – not present; 1 – present, few individuals, no overlap among calls, individuals can be counted; 2 – several individuals, some overlap

Table 1
Summary of the results of 13 studies examining the effects of the amount of forest cover in surrounding landscapes on the presence or abundance of six anuran species in focal ponds/wetlands in the centre of each landscape. The superscript numbers next to each species indicate which species are covered by each reference.

Species	Summary of forest effects		
	Positive	Negative	No effect
Rana sylvatica (wood frog)[a]	8	0	3
Pseudacris crucifer (spring peeper)[b]	8	0	1
Rana pipiens (northern leopard frog)[c]	0	3	6
Bufo americanus (American toad)[d]	2	5	4
Hyla versicolor (gray treefrog)[e]	3	1	7
Rana clamitans (green frog)[f]	3	0	4

[a,c,d,e]Lehtinen et al. (1999), [a,c,d,e,f]Findlay et al. (2001), [a,b,c,d,f]Guerry and Hunter (2002), [a,b,c,d,e,f]Houlahan and Findlay (2003), [a,b,c,d,e,f]Trenham et al. (2003), [a]Homan et al. (2004), [b,d,e]Knutson et al. (2004), [b,d,e,f]Price et al. (2004), [a]Porej et al. (2004), [a,b,c,d]Gibbs et al. (2005), [a,b,d,e,f]Herrmann et al. (2005), [a,b,c,d,f]Gagné and Fahrig (2007), [a,b,c,d,e,f]Eigenbrod et al. (2008).
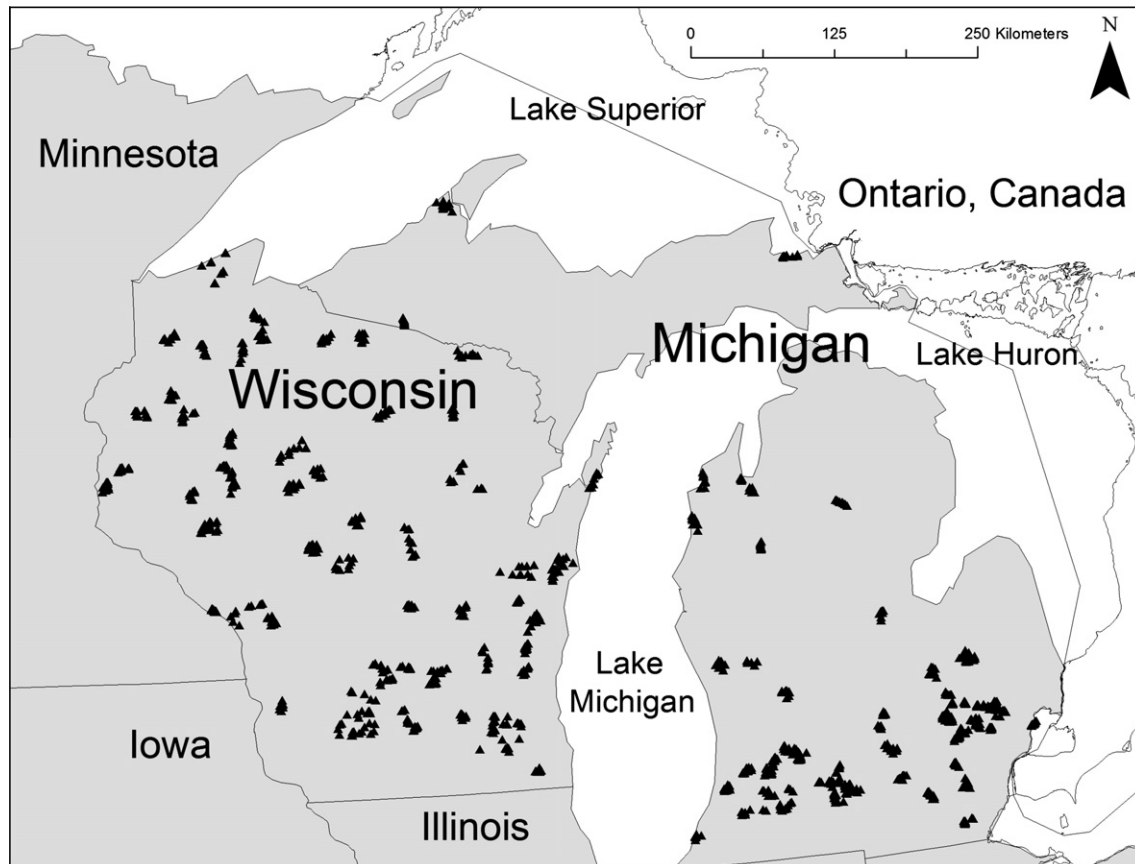
Fig. 1. Locations of 1141 surveys sites (▲) within Wisconsin (600 sites) and Michigan (541 sites), USA. Not all 1141 sites are visible, as sites are clustered by survey route, and often overlap at the scale of the figure. More details are available at http://www.michigan.gov/dnr/0,1607,7-153-10370_12143_12194—,00.html and http://www.mbr-pwrc.usgs.gov/wifrog/frog.htm.

among calls; number of individuals can be counted; 3 – many individuals, overlap among calls, number of individuals cannot be distinguished.

These two datasets are well-suited for our analyses for two main reasons. First, the survey sites are representative of the full gradient of land use intensity typically present in the Great Lakes basin, from the relatively undisturbed northern mixed-wood forests of northern Michigan and Wisconsin to largely agricultural landscapes to suburban wetlands at the edges of the major urban conglomerations of Detroit and Milwaukee. Secondly, all 1141 sites have at least 5 years of data between 1996 and 2002, allowing us to use the average of the maximum call index heard for each species at each site over time (after Knutson et al., 1999) as the response for every analysis. Using average values greatly reduces the impact of the large year-to-year fluctuations typical of anuran populations (Pechmann et al., 1991), and also increases statistical power by converting the ordinal response data from a single year's sampling to a continuous variable. In addition, using the maximum value at a site for a species in a particular year reduces the problem of seasonal fluctuations in call detectability (Royle, 2004).

We chose the six species of anurans we examine in this analysis – wood frog (*Rana sylvatica* LeConte), spring peeper (*Pseudacris crucifer* Wied–Neuwied), northern leopard frog (*Rana pipiens* Schreber), American toad (*Bufo americanus* Holbrook), eastern gray treefrog (*Hyla versicolor* LeConte), green frog (*Rana clamitans* Latreille) – both because of the inconsistencies in the literature for these species (Table 1), and because these species vary in their level of dependency on forest. The spring peeper and wood frog are both forest-obligate species and have been shown to show strong positive associations with forest cover in most landscape-scale

studies (Table 1). The remaining species are generally thought to be habitat generalists that are less sensitive to the amount of forest in the landscape though, as discussed earlier, positive and/or negative associations with forest cover have been found for all species (Table 1). Choosing species with a range of dependencies on forest allowed us to examine the effects of study design on inference for both stronger and weaker relationships between a landscape-scale predictor and a response. Natural history information on these species is available in the Appendix.

For the landscape predictor variables, we used the 2001 United States Geological Survey land cover dataset (30 m resolution) (Vogelmann et al., 2001). We reclassified this dataset to give six land cover classes: forest, wetland, crop, pasture, open water and development (roads, urbanization, quarries).

2.2. General approach

Our approach to testing the effects of the three pitfalls on the inferred relationship between anuran abundance and forest cover in this dataset was to run two analyses (or sets of analyses) per pitfall. In one analysis, we controlled for the pitfall, and in the other we did not. Unless otherwise indicated, the analyses were always simple linear regressions with anuran abundance for each of the six species as the predictor variable, and the amount of forest cover in the landscape the response variable. We then compared the values of the slope coefficients for forest cover (hereafter 'forest coefficient') from the two analyses. We used linear regressions because that was the approach taken in 12 of the 13 landscape-scale studies carried out in the Great Lakes on these species (Table 1).

To obtain robust estimates of the forest coefficients and to test for significant differences between the forest coefficients when controlling vs. when not controlling for each pitfall, we used a randomization approach. For each analysis, we randomly selected 1000 subsets of 25 or 50 sites and ran the same statistical models for all 1000 subsets. To obtain a distribution of the forest coefficient, we used the program FOCUS 2.3 (Holland et al., 2004) to select the random subsets with the restriction that all sites were located at least 10 km from their nearest neighbour (non-overlapping landscapes at the 5000 m scale) for all analyses except where we were testing for the impact of overlapping landscapes. The 5000 m scale was chosen to test all pitfalls as preliminary analyses (Table A1 in the Appendix) showed that this was the scale at which the six species we examined responded most strongly, and that running the analyses at scales other than 5000 m did not qualitatively affect our findings.

With 1000 subsets for each analysis, we expected up to 50 significant (25 positive and 25 negative) associations by chance alone at the 95% confidence level. Thus, we only considered a species to have a statistically significant positive or negative association with forest cover in an analysis if 25 or more of the subsets showed significant (α = 0.05) positive or negative associations with forest cover. We considered a change from 25 or more significant positive associations with forest cover to 25 or more significant negative associations with forest cover (or vice versa) to signify a shift in the sign of the forest coefficient. All statistical analyses were run using R 2.6 (R Development Core Team, 2007).

2.3. Pitfall 1: using landscapes that overlap in space

To determine the effect of overlapping landscapes on the forest coefficient, we compared the results of 1000 random subsets that were selected from the full dataset such that there was no overlap at the 5000 m scale (i.e., the minimum distance from each point to its nearest neighbour was 10,000 m), with 1000 random subsets selected to maximize the overlap of landscapes.

To obtain sets of landscapes with maximum overlap, we selected 1000 subsets of 50 sites, with each subset made up of several clusters of sites that were within 10,000 m of each other. For each subset, we first selected one site at random out of the full dataset and then selected all other sites that were within 10,000 m of this first site. The distribution of sites (Fig. 1) meant that there were never 50 sites within 10,000 m of any given site, so we repeated this random selection procedure until we had picked the necessary 50 sites for a subset, subject to the restriction that sites could not be selected more than once. This approach gave a mean nearest neighbour distance of 1638 m for the 1000 subsets of 50 points, as compared to a mean nearest neighbour distance of 37,753 m in the 1000 subsets where there was no overlap at the 5000 m scale. We then regressed the abundance of each of the six species on the amount of forest cover in the landscape at the 5000 m scale for each subset in both analyses (overlapping and non-overlapping landscapes).

2.4. Pitfall 2: using only a portion of the potential range of the landscape predictor variable of interest

To evaluate the effect of the range across landscapes of the amount of forest in the landscape on the relationship between anurans and the amount of forest cover, we compared the results of six datasets (five subsets of the full dataset (n = 1141) and the full dataset itself) with varying ranges of forest cover in the landscapes at the 5000 m scale. These were: (1) 0.5–15% forest (264 sites); (2) 40–60% forest (200 sites); (3) 30–70% forest (433 sites); (4) 50–97% forest (219 sites); (5) 0.5–50% (922 sites); (6) 0.5–97% forest (no restrictions; therefore all 1141 sites in the full dataset). We se-

lected these datasets to examine the effects of the range in the amount of forest cover at low, intermediate, and high levels of forest cover. At the high-forest end of the gradient we were unable to examine a narrower range than 50–100% forest cover due to the relatively small number of sites with very high forest cover in our dataset. For each dataset we then randomly selected 1000 subsets of 25 sites, and regressed anuran abundance against forest amount for each subset. We used 25 sites rather than 50 because of the relatively small total number of sites in some of the six categories.

2.5. Pitfall 3: failing to account for correlations among landscape predictor variables

All six species of anurans examined in this study except for the green frog have also been shown to be negatively affected by both urbanization (Gagné and Fahrig, 2007) and road traffic (Eigenbrod et al., 2008, 2009), so correlations between roads and/or urban development and forest are particularly likely to affect the magnitude and sign of the forest coefficient. To test for the effects of correlations between forest cover and development (and/or roads; road density is highly correlated with development in our dataset; $r > 0.8$ over all 1141 sites), we first created a subset of the full 1141 sites in which we attempted to maximize the strength of the (negative) Pearson correlation between the amount of forest and the amount of development at the 5000 m scale, while still retaining a large sample size. This 'high correlation' dataset consists of 500 sites and has an r of -0.77. We then did the opposite, minimizing the correlation between forest amount and development while still retaining a large samples size to create the 'low correlation' dataset (727 of the original 1141 sites and $r = -0.05$). We then randomly selected 1000 subsets of 50 sites from each of the 'high correlation' and 'low correlation' datasets, and ran both a simple linear regression of anuran abundance on forest amount and a multiple regression model of anuran abundance on forest amount and development for each subset of 50. The inclusion of the development term meant there we had slightly less statistical power to detect a significant relationship in the forest coefficient in the multiple regression than in the simple linear regression.

3. Results

3.1. Pitfall 1 – overlapping landscapes

While overlapping landscapes increased the variability around the forest coefficient for all species, there was no shift in the sign of the forest coefficient for any species (Table 2). Overlapping landscapes decreased the strength of the positive association with the forest coefficient for the wood frog and spring peeper, and increased the negative association between forest cover and leopard frog abundance. The mean forest coefficient also increased for the gray treefrog and decreased for the American toad in the subsets with overlapping landscapes.

3.2. Pitfall 2 – range of predictor

Reductions in the range of forest cover led to large changes in the forest coefficient for every species, and switches in the sign of the coefficient for three species – the gray treefrog, leopard frog and green frog (Table 3). In addition to these switches in the signs of the forest coefficient, there were shifts from a significant association to no association for all other species. There was also a notably strong positive association with forest cover in the 0–15% forest cover dataset for the gray treefrog and spring peeper, which became much weaker (spring peeper) and switched to negative (gray

Table 2
Comparison of results of simple linear regression analyses from two sets of analyses – 'overlapping landscapes', and 'no overlap at the 5000 m scale' (see methods for details). 1000 random subsets (each with n = 50 sites) were selected from each dataset, and anuran abundance was regressed on forest cover at the 5000 m scale for all analyses. + indicates that there were at least 25 significant (α = 0.05) positive associations with forest cover for an analysis; ++ indicates at least 250 significant positive associations, and +++ at least 500 significant positive associations with forest cover. –, – –, or – – – indicates at least 25, 250 or 500 significant negative associations, respectively. NF (no effect) indicates that there were less than 25 significant positive or negative associations per analysis.

Species	Overlapping landscapes		No overlap at 5000 m scale	
	Mean forest coefficient ± 1 SD	Sign of forest coefficient	Mean forest coefficient ± 1 SD	Sign of forest coefficient
Wood frog	0.54 ± 1.351	+	0.99 ± 0.422	+++
Spring peeper	2.52 ± 1.666	++	1.91 ± 0.548	+++
Leopard frog	−1.43 ± 1.217	– –	−0.23 ± 0.247	–
American toad	−1.05 ± 1.401	–	−0.46 ± 0.397	–
Gray treefrog	1.80 ± 1.838	+	0.47 ± 0.558	+
Green frog	0.39 ± 1.198	NF	0.18 ± 0.455	+

Table 3
Comparison of results of simple linear regression analyses of anuran abundances on forest cover using five different ranges in the percentage of forest cover in the landscape at the 5000 m scale. 1000 random subsets (each with n = 25 sites) were selected from each species dataset. + indicates that there were at least 25 significant (α = 0.05) positive associations with forest cover for an analysis; ++ indicates at least 250 significant positive associations, and +++ at least 500 significant positive associations with forest cover. –, – –, or – – – indicates at least 25, 250 or 500 significant negative associations, respectively. NF (no effect) indicates that there were fewer than 25 significant positive or negative associations per analysis. Note that there were over 25 significant positive and negative associations for the green frog when the range of forest cover was 0–50%.

Range in forest cover (%)	Wood frog Forest coefficient		Spring peeper Forest coefficient		Leopard frog Forest coefficient		American toad Forest coefficient		Gray treefrog Forest coefficient		Green frog Forest coefficient	
	Mean ± 1 SD	Sign	Mean ± 1 SD	Sign	Mean ± 1 SD	Sign	Mean ± 1 SD	Sign	Mean ± 1 SD	Sign	Mean ± 1 SD	Sign
0–15	−0.69 ± 2.270	NF	13.33 ± 4.536	+++	−0.80 ± 2.105	NF	−2.60 ± 3.082	–	10.40 ± 3.814	+++	−1.83 ± 3.380	–
0–50	1.19 ± 1.096	+	3.23 ± 1.443	+++	−0.26 ± 0.666	–	−0.71 ± 0.990	–	1.71 ± 1.503	+	0.05 ± 1.154	+–
40–60	1.15 ± 1.938	+	1.29 ± 1.722	+	1.05 ± 1.198	+	−0.45 ± 1.854	NF	−0.80 ± 2.679	–	−0.29 ± 2.201	NF
30–70	0.15 ± 1.148	NF	0.64 ± 1.233	+	−0.25 ± 0.618	–	0.17 ± 0.945	NF	−0.66 ± 1.404	–	0.57 ± 1.213	+
50–100	1.18 ± 1.190	+	0.32 ± 0.816	NF	−0.69 ± 0.639	–	−0.43 ± 1.052	NF	−0.24 ± 1.457	NF	−0.08 ± 1.214	NF
0–100	0.97 ± 0.647	++	1.94 ± 0.859	+++	−0.23 ± 0.385	–	−0.47 ± 0.598	–	0.50 ± 0.968	+	0.20 ± 0.710	+

Table 4
Comparison of results of analyses from two datasets with high and low correlations, respectively, between development and forest cover at the 5000 m scale. "Forest only" were simple linear regressions of anuran abundances on forest cover, and 'Devel. + forest' were multiple regressions of anuran abundances on development and forest cover. + indicates that, of the 1000 random subsets (each with n = 50 sites) there were at least 25 significant (α = 0.05) positive associations with forest cover; ++ indicates at least 250 significant positive associations, and +++ at least 500 significant positive associations with forest cover. –, – –, or – – – indicates at least 25, 250 or 500 significant negative associations, respectively. NF (no effect) indicates that there were fewer than 25 significant positive or negative associations per analysis.

Species	High correlation dataset				Low correlation dataset			
	Mean forest coefficient ± 1 SD		Sign of forest coefficient		Mean forest coefficient ± 1 SD		Sign of forest coefficient	
	Forest only	Devel. + forest	Forest only	Devel. + forest	Forest only	Devel. + forest	Forest only	Devel. + forest
Wood frog	1.14 ± 0.317	0.68 ± 0.478	+++	+	0.56 ± 0.519	0.57 ± 0.523	+	+
Spring peeper	1.73 ± 0.411	0.79 ± 0.621	+++	+	2.27 ± 0.625	2.27 ± 0.627	+++	+++
Leopard frog	0.12 ± 0.157	−0.22 ± 0.230	+	–	−0.62 ± 0.325	−0.63 ± 0.324	– –	– –
American toad	−0.08 ± 0.293	−0.42 ± 0.445	NF	–	−1.03 ± 0.463	−1.05 ± 0.457	– –	– –
Gray treefrog	0.65 ± 0.457	−0.29 ± 0.678	+	–	0.29 ± 0.666	0.29 ± 0.669	+	+
Green frog	0.55 ± 0.353	0.48 ± 0.487	+	+	−0.16 ± 0.547	−0.17 ± 0.538	–	–

treefrog) when the lower end of the possible range of forest cover was excluded from the analysis (Table 3).

3.3. Pitfall 3 – correlations between predictors

High negative correlations between the amount of forest cover and the amount of development in the landscape inflated the positive association with forest cover for every species except the spring peeper, and led to a switch in the sign of the forest coefficient for the leopard frog, green frog, and, after controlling for development through multiple regression, the gray treefrog. Controlling for development by including it in the multiple regression analysis in the 'high correlation' dataset brought the results in the 'high correlation' dataset more in line with the 'low correlation' dataset for the wood frog, leopard frog, and American toad, but less so for the spring peeper, gray treefrog and green frog (Table 4).

4. Discussion

Our results indicate that study design alone can lead to both shifts in the sign and large decreases in the strength of the inferred relationship between a species response and a landscape predictor. The range of the predictor variable – forest cover in this study – had the largest effect on the sign of the relationship and on the strength of the inferred relationships in our dataset. Correlations between forest and development also led to shifts in the sign of the forest coefficient for some species, and affected the strength of the inferred relationship for all species. While the degree to which the three pitfalls affect ecological relationships in other datasets will depend on the spatial structure of the landscapes and strengths and forms of the ecological relationships being examined, our results are sufficiently striking that they have major implications for future landscape-scale studies. Interactions among

the pitfalls are also possible, and could increase their effects on landscape-scale inferences.

The range of the predictor variable (Pitfall 2) had a particularly large impact on the inferred relationship between forest cover and anuran abundance in our dataset. The main reasons for this are the curvilinear relationships between the amount of forest cover in the landscape and abundances of all species (Fig. 2); many of these only became apparent when the full range of values of the predictor variable was considered. In particular, the non-monotonic relationships shown by the gray treefrog, leopard frog, and green frog (Fig. 2) were sufficient to cause shifts in the sign of the forest coefficients depending on the range of forest cover considered in the analysis. This pitfall is likely very common, as many land cover types or anthropogenic stressors will be regionally uncommon, making it impossible to find sites that encompass the full *potential* range (0–100%) of the predictor variable of interest. For example, it is (fortunately!) still difficult to find landscapes where urban development is 100% in most regions (e.g. in a study examining the effects of urbanization on stream amphibians, Riley et al. (2005) found that the maximum level urban development in watersheds in southern California was 37%). On the other hand, in human-dominated regions sites with very high amounts of natural land cover types may not exist (e.g. Bellamy et al. (1996)'s study on woodland birds in England was conducted in a region with only 2% forest cover). The opposite can also be true – e.g. forest cover was never *less* than 23% in Herrmann et al. (2005)'s study examining the effects of forest cover on anurans in the heavily forested US state of New Hampshire; Herrmann et al. suggest this lack of variation may be the reason for the low explanatory power of forest cover in predicting species richness. Clearly, such studies can be very valuable and should continue to be conducted; however, investigators should (1) ensure the full *possible* range of the predictor variable in the region is included in the analysis, (2) clearly state that extrapolation beyond the range the predictor variable considered in the study should be avoided, and (3) discuss the impact that limitations in the range of the predictor variable may have on their findings.

Correlations among predictor variables (Pitfall 3) also led to shifts in the sign of the inferred relationship for two species – the leopard frog and the green frog. The shift from an apparent positive relationship with forest cover when this was correlated with development to a negative relationship with forest once this correlation was removed suggests that the negative associations of both species with development are stronger than their negative associations with forest cover. This finding is supported by the literature, particularly for the leopard frog (Gagné and Fahrig, 2007; Eigenbrod et al., 2008). Also of note was the switch in the sign of the association with forest cover of the gray treefrog after controlling for development from positive to negative when the correlation between forest and development was high. The latter result was probably due to the non-monotonic relationship between this species and forest cover discussed earlier, and illustrates how high correlations between predictor variables can decrease interpretability of results even after statistically controlling for such collinearities through multiple regression models. Again, correlations between predictor variables are likely to be extremely common in landscape ecology. McGarigal and Cushman (2002)'s review of the effects of habitat fragmentation showed that 94% of 134 studies were designed in such a way that habitat area and fragmentation were highly correlated (and that this was a major part of the reason for the confusion about whether habitat fragmentation or area had a greater effect on organisms), while Smith et al. (2009)'s review and analysis of the same topic suggests that there have been no recent improvements in this.

Our results also demonstrate how correlations among predictor variables can lead to large reductions in statistical power in multiple regressions, confirming the results of numerical simulations by Graham (2003). For example, the spring peeper showed a much
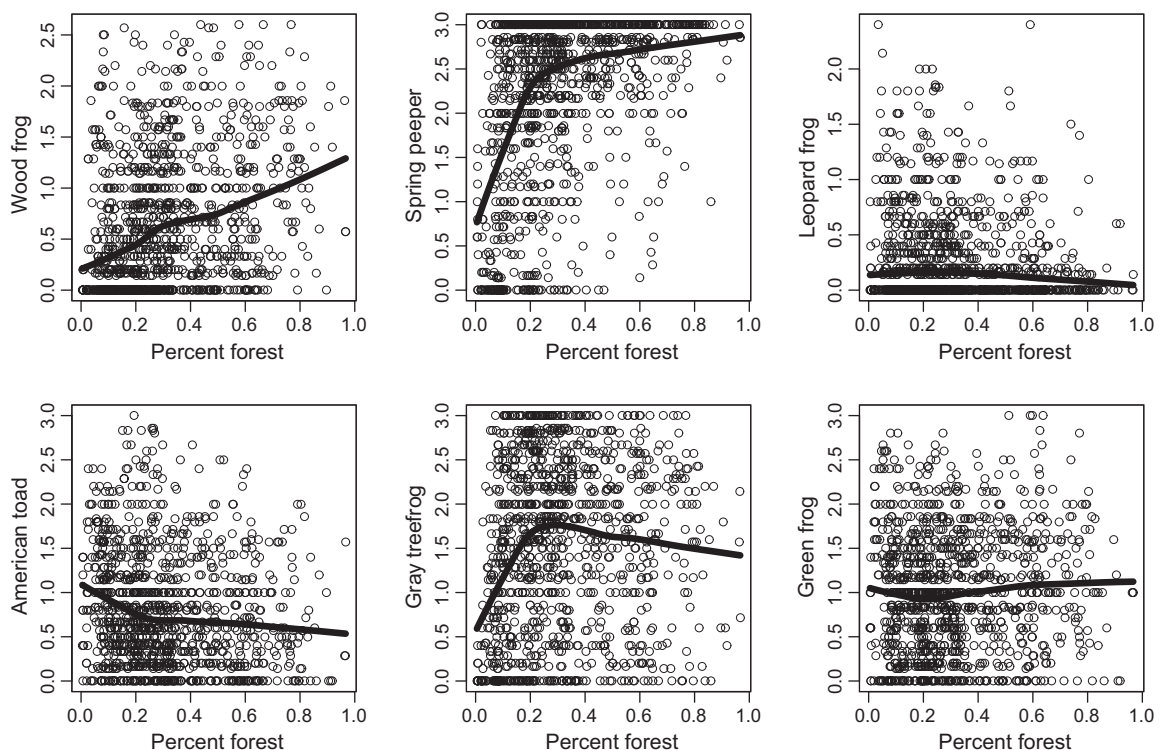
Fig. 2. Relationships between anuran relative abundance at 1141 sites (ponds/wetlands) and the percentage forest cover in the surrounding landscapes measured at the 5000 m scale (radius). The lowess smoothing line is shown for each species.

weaker positive association with forest cover (after controlling for the effect of development using a multiple regression) when the correlation between development and forest cover was high than when development and forest cover were uncorrelated.

Overlapping landscapes (Pitfall 1) did not lead to shifts in the sign of the forest coefficient here, but did lead to increased variation around the mean of the forest coefficient among subsets as compared to the subsets where there was no overlap. Such variability is due to the reduction in the number of independent landscapes caused by overlap of landscapes – there were effectively only 5–10 landscapes (at the 5000 m scale) in the 'overlapping' analyses (each with a cluster of 5–10 overlapping landscapes within them) as compared with the 50 landscapes in the non-overlapping analyses. This reduction in independent landscapes means that a study is less representative of the region being sampled than its sample size suggests (pseudoreplication), increasing the probability of Type I errors. It is, for example, less likely that 50 sites distributed over five landscapes are as representative of the distribution of forest cover in Michigan and Wisconsin than are 50 sites distributed over 50 independent landscapes. The reduction in 'true' sample size caused by overlapping landscapes also increases the likelihood of failing to select sites across the full range of the predictor variables of interest (Pitfall 3), which, as we have shown, can easily lead to incorrect inferences.

This research has far-reaching conservation implications, given that the pitfalls investigated here are common, and, as our results show, can lead to misidentification of the true direction of the relationship between a landscape-scale predictor and a response. Such misidentification can lead to incorrect management actions. For example, a failure to differentiate between the effects of different types of human modifications of the landscapes can lead managers to: (a) miss opportunities for conservation in areas with low amounts of habitat surrounded by relatively benign matrix, such as low-intensity agriculture (Gibbs et al., 2009) or (b) prioritize conservation in areas with high amounts of habitat but with small amounts of very detrimental matrix such as high-traffic roads (Eigenbrod et al., 2008). Even studies that show no significant effect of a biologically important predictor variable due to issues in sampling design can have implications for conservation, as they can be taken by policy makers as evidence that a given predictor is irrelevant. Poorly designed studies that find no effect of a predictor can thus serve as justification for not initiating important management actions. Proper study design is arguably most important for under-studied species, as there will be little a priori knowledge to aid the interpretation of results.

In addition to its general implications, this study provides convincing evidence for the 'true' relationships between forest cover and anuran abundance for the two species of anurans for which there are conflicting findings in the literature (Table 1). The mixture of positive and negative associations observed for the gray treefrog is almost certainly a result of the non-monotonic relationship with forest cover we observed in this study for this species; all of the studies on this species in Table 1 assumed straight line relationships with forest cover. We show for the first time that gray treefrogs have a very strong requirement for some forest in the landscape (20–30% cover), but there is little additional benefit beyond this amount, and that abundance may decrease as forest cover nears 100%. Post-hoc analysis showed that this relationship holds even if only sites where there is no correlation with development are considered (Fig. A1 in the Appendix). This non-monotonic relationship is not surprising given this species' natural history – the gray treefrog prefers breeding ponds with little forest cover, but lives and forages in small trees/shrubs and prefers forests to clearings for dispersal (Hocking and Semlitsch, 2007). The positive relationships observed in the literature between American toads and forest amount, and the large number of 'no effects' of forest

cover observed for the leopard frog and green frog (Table 1) are most likely due to strong correlations between forest amount and development, which, as we show, can mask the negative associations between these species and forest cover.

The most important step for investigators to avoid the above pitfalls is to quantify the spatial structure of the potential landscapes (experimental units) in their study at the design stage. GIS software should be used to screen all potential landscapes for their suitability (McGarigal and Cushman, 2002), and then select a subset of non-overlapping sites at scales relevant to the organism and process of interest for which the full range of the predictor variables of interest occur, and for which the degree of collinearity of the predictors of interest with other potentially important environmental variables is low (e.g. Guerry and Hunter, 2002). Such an approach – outlined in greater detail in Brennan et al. (2002) – will maximize the potential of a study to detect the effects of one or two predictors of particular interest. However, in exploratory analyses in which the goal is to examine relationships between many environmental predictor variables and some measure of biodiversity, collinearity between predictors and limited ranges in the distributions of some predictor variables will be nearly unavoidable. Investigators should therefore provide descriptive statistics of the distributions of predictor variables and the relationships between them in such studies and explicitly discuss the likely impacts of these pitfalls on their findings (e.g. Houlahan and Findlay, 2003). Sophisticated statistical modelling techniques may also help to disentangle the effects of multiple predictor variables in spatial datasets, but can also introduce biases (Smith et al., 2009), and should therefore not be viewed as a panacea.

Acknowledgements

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.biocon.2010.09.007.

References

Bellamy, P.E., Hinsley, S.A., Newton, I., 1996. Factors influencing bird species numbers in small woods in south-east England. Journal of Applied Ecology 33, 249–262.

Brennan, J.M., Bender, D.J., Contreras, T.A., Fahrig, L., 2002. Focal patch landscape studies for wildlife management. In: Wu, J., Taylor, W.W. (Eds.), Optimizing Sampling Effort Across Scales. Integrating Landscape Ecology into Natural Resource Management. Cambridge University Press, Cambridge., pp. 68–91.

Cushman, S.A., 2006. Effects of habitat loss and fragmentation on amphibians: a review and prospectus. Biological Conservation 128, 231–240.

Eigenbrod, F., Hecnar, S.J., Fahrig, L., 2008. The relative effects of road traffic and forest cover on anuran populations. Biological Conservation 141, 35–46.

Eigenbrod, F., Hecnar, S.J, Fahrig, L., 2009. Quantifying the road-effect zone: threshold effects of a motorway on anuran populations in Ontario, Canada. Ecology and Society, 14, 24. <http://www.ecologyandsociety.org/vol14/iss1/art24/>.

Fahrig, L., 2003. Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution and Systematics 34, 487–515.

Findlay, C.S., Lenton, J., Zheng, L., 2001. Land-use correlates of anuran community richness and composition in southeastern Ontario wetlands. Écoscience 8, 336–343.

Freckleton, R., 2002. On the misuse of residuals in ecology: regression of residuals vs. multiple regression. Journal of Applied Ecology 71, 542–545.

Gagné, S.A., Fahrig, L., 2007. Effect of landscape context on anuran communities in breeding ponds in the National Capital Region, Canada. Landscape Ecology 22, 205–215.

Gibbs, J.P., Whiteleather, K.K., Schueler, F.W., 2005. Changes in frog and toad populations over 30 years in New York state. Ecological Applications 15, 1148–1157.

Gibbs, K.E., Mackey, R.L., Currie, D.J., 2009. Human land use, agriculture, pesticides and losses of imperiled species. Diversity and Distributions 15, 242–253.

Graham, M.H., 2003. Confronting multicollinearity in ecological multiple regression. Ecology 84, 2809–2815.

Guerry, A.D., Hunter, M.L., 2002. Amphibian distributions in a landscape of forests and agriculture: and examination of landscape composition and configuration. Conservation Biology 16, 745–754.

Herrmann, H.L., Babbitt, K.J., Baber, M.J., Congalton, R.G., 2005. Effects of landscape characteristics on amphibian distribution in a forest-dominated landscape. Biological Conservation 123, 139–149.

Hocking, D.J., Semlitsch, R.D., 2007. Effects of timber harvest on breeding-site selection by gray treefrogs (Hyla versicolor). Biological Conservation 138, 506–513.

Holland, J.D., Bert, D.G., Fahrig, L., 2004. Determining the spatial scale of species' response to habitat. BioScience 54, 227–233.

Homan, R.N., Windmiller, B.S., Reed, M., 2004. Critical thresholds associated with habitat loss for two vernal pool-breeding amphibians. Ecological Applications 14, 1547–1553.

Houlahan, J.E., Findlay, C.S., 2003. The effects of adjacent land use on wetland amphibian species richness and community composition. Canadian Journal of Fisheries and Aquatic Science 60, 1078–1094.

Hulbert, S.H., 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54, 187–211.

Knutson, M.G., Sauer, J.M., Olsen, D.A., Mossman, M.J., Hemesath, L.M., Lannoo, M.J., 1999. Effects of landscape composition and wetland fragmentation on frog and toad abundance and species richness in Iowa and Wisconsin, USA. Conservation Biology 13, 1437–1446.

Knutson, M.G., Richardson, W.B., Reineke, D.M., Gray, B.R., Parmelee, J.R., Weick, S.E., 2004. Agricultural ponds support amphibian populations. Ecological Applications 14, 669–684.

Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74, 1659–1673.

Lehtinen, R.M., Galatowitsch, S.M., Tester, J.R., 1999. Consequences of habitat loss and fragmentation for wetland amphibian assemblages. Wetlands 19, 1–12.

McGarigal, K., Cushman, S.A., 2002. Comparative evaluation of experimental approaches to the study of habitat fragmentation effects. Ecological Applications 12, 335–345.

Mossman, M.J., Hartman, L.M., Hay, R., Sauer, J.R., Dhuey, B.J., 1998. Monitoring long term-trends in Wisconsin frog and toad populations. In: Lanoo, M.J. (Ed.), Status and Conservation of Midwestern Amphibians. University of Iowa Press, Iowa City, USA, pp. 169–198.

Pechmann, J.H.K., Scott, D.E., Semlitsch, R.D., Caldwell, J.P., Vitt, L.J., Gibbons, J.W., 1991. Declining amphibian populations: the problem of separating human impacts from natural fluctuations. Science 253, 892–895.

Porej, D., Micacchion, M., Hetherington, T.E., 2004. Core terrestrial habitat for conservation of local populations of salamanders and wood frogs in agricultural landscapes. Biological Conservation 120, 399–409.

Price, J.S., Marks, D.R., Howe, R.W., Hanowski, J.M., Niemi, G.J., 2004. The importance of spatial scale for conservation and assessment of anuran populations in coastal wetlands in the western Great Lakes, USA. Landscape Ecology 20, 441–454.

R Development Core Team, 2007. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.

Riley, S.P.D., Busteed, G.T., Kats, L.B., Vandergon, T.L., Lee, L.F.S., Dagit, R.G., Kerby, J.L., Fisher, R.N., Sauvajot, R.M., 2005. Conservation Biology 19, 1894–1907.

Royle, J.A., 2004. Modeling abundance index data from anuran calling surveys. Conservation Biology 18, 1378–1385.

Schooley, R.L., 2006. Spatial heterogeneity and characteristic scales of species–habitat relationships. BioScience 56, 533–537.

Smith, A.C., Koper, N., Francis, C.M., Fahrig, L., 2009. Confronting collinearity: comparing methods for disentangling the effects of habitat loss and fragmentation. Landscape Ecology 24, 1271–1285.

Trenham, P.C., Koenig, W.D., Mossman, M.J., Stark, S.L., Jagger, L.A., 2003. Regional dynamics of wetland-breeding frogs and toads: turnover and synchrony. Ecological Applications 13, 1522–1532.

Turner, M., 2005. Landscape ecology: what is the state of the science? Annual Review of Ecology. Evolution and Systematics 36, 319–344.

Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., Van Driel, J.N., 2001. Completion of the 1990s National Land Cover Data Set for the conterminous United States. Photogrammetric Engineering and Remote Sensing 67, 650–662.

Wiens, J.A., 1989. Spatial scaling in ecology. Functional Ecology 3, 385–397.