Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

From census to survey – changing monitoring methodology without compromising time series integrity

Niels Martin Schmidt

nms@bios.au.dk

Aarhus University https://orcid.org/0000-0002-4166-6218

Lars Holst Hansen

Aarhus University: Aarhus Universitet

Research Article

Keywords:

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11280809/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

From census to survey – changing monitoring methodology without compromising time series integrity

Niels Martin Schmidt* and Lars Holst Hansen

Department of Ecoscience and Arctic Research Cluster, Aarhus University, Frederiksborgvej 399, 4000 Roskilde, Denmark

* Corresponding author: E-mail: nms@ecos.au.dk; Phone: +45 87 15 86 83

Authors' contributions NMS: performed field work, conceived the idea and designed methodologies, performed analyses, and lead the writing. LHH: performed field work and contributed to writing.

Abstract

Time series arising from long-term monitoring constitute a key source for ecological knowledge, and their value continues to increase as new data are being added. Yet, in many cases data collection methodologies applied are rooted in historical decisions. To adapt to new questions, opportunities or challenges, methodologies may have to be changed. Such changes, however, can not jeopardize the integrity of long-term time series. Here we present a statistical protocol that can be used to optimize and validate the switch in methodology for estimating abundances in a designated area from a full census approach to a sample survey with permanent transects. To evaluate the generality of our method, we first implemented our methodology on a suite of simulated abundance data, and second on a real-life case study on high Arctic collared lemmings (*Dicrostonyx groenlandicus*). For the simulated data series, our approach generally allowed for the transition from census-based to survey-based monitoring without compromising time series integrity, even with a relatively low number of transects. Accuracy, however, depends on overall abundance of the focal species. For the real-life data, we demonstrate that the new transect-based approach indeed was able to capture the dynamics and cyclicity of the population abundance present in the census data collected over almost 30 years. The integrity of the time series is thus maintained, whilst 94% less effort is being invested in the collection of data. Using the provided R code, the approach is easily implemented in other monitoring programs that require a change in methodology.

Keywords: Adaptive monitoring, Full census, Long-term monitoring, Sample survey, Validation

Introduction

Long-term monitoring is central for our ability to keep track of species and ecosystems in a rapidly changing world (Hughes et al. 2017). To build up long time series obviously requires time and money, and to safeguard the integrity of the time series, long-term monitoring oftentimes adheres to decisions made, and methodologies implemented, at the onset of the program. The emergence of new research questions and agendas, technology and changing funding landscapes may however require long-term monitoring programs to be changed (Buettel et al. 2025) or adapted (Lindenmayer and Likens 2009). It is, however, essential that adaptations to the new reality are conducted in such a way that the integrity of the historical time series is not compromised (Buettel et al. 2025).

In monitoring programs, the abundance of a given species can be estimated by means of either true census, where all individuals (or pairs, nests etc.) are counted within a designated area, or by sample survey, where a smaller area of the designated area is covered (Gregory et al. 2004). While the first is expected to provide a more complete abundance estimate, the latter requires less effort and allows for quantification of the uncertainties associated with the abundance estimates (Gregory et al. 2004; Moullec et al. 2017). Full census is therefore not commonly used in ecological monitoring. However, particular monitoring programs established decades ago have opted for this approach (e.g. Meltofte et al. 2020; Schmidt et al. 2021; Von Beckerath et al. 2021). The mere duration of the time series originating from such long-term programs makes them highly valuable (Hughes et al. 2017; Viblanc and Muller-Landau 2025), and the need for sustained time series comparability equally stronger. While modern technology may alleviate some of the costs associated with full census (e.g.

Rominger and Meyer 2021), more drastic methodological changes may be required, for instance to meet new funding schemes or time constraints.

Here we present a simple and effective statistical protocol that allows for the transition of abundance estimation from being based on full census to being based on sample survey. The approach relies on historical georeferenced data to balance accuracy and required effort whilst ensuring time series integrity.

Considerations

Buettel et al. (2025) presented a seven-step general framework to guide the implementation of new methods into on-going monitoring programs. For the specific case of transitioning from full census to sample survey, we will focus on evaluating the new methodology and the integrity of the resulting data, with particular attention to identifying the optimal balance between data quality (i.e. the ability of the samples survey to accurately replicate the full census data) and effort required to conduct the sample survey versus the full census. While the exact balance between the two largely hinges on the local context, some general considerations can be made.

Data integrity

As suggested by Buettel et al. (2025), the validity of new methods in long-term monitoring can be examined by running the old and the new method in parallel and then cross-validating the results. While this is the preferred approach in many cases, transitioning from full census to transect-based sample survey may allow for a different validation approach: If

the historical data consist of georeferenced registrations from within a well-defined census area, they can be used to evaluate the accuracy of the sample survey in comparison to the full census by means of the new methods ability to replicate past dynamics. First and foremost, the new approach must capture the temporal pattern in animal abundance within the census area, with the abundance obtained in the full census regarded as true abundance. However, ideally the new approach should also be able to capture other population dynamical features, such as population trends and cyclicity patterns. Depending on the focal species, other metrics may be (more) relevant for the evaluation of accuracy.

Effort

The number of transects as well as their length and width (i.e. total area covered) determine the effort required. Both the number and the dimensions can be adjusted, provided that it requires less effort than the full census and the required accuracy is achieved. Knowledge obtained during the full census period may help guide the layout of the transect for the sample survey. The focal species and the landscape in which the census has been conducted may require a particular layout of transects.

Methodological approach

In this section we provide a roadmap to facilitate the transition from census-based to survey -based monitoring. An R script to implement the methodology is provided in Online Resource 1. The first step is to generate the permanent transects, randomly distributed within the existing census area. Transect dimensions may vary according to focal species and

105 the characteristics of the census area and transect dimensions thus are to be based on
106 knowledge of the census area and previous experience with optimal transect width for nest
107 detection within the census area. The second step is to select all non-overlapping transects
108 from the randomly distributed transects.

109 The third step is then to select a subset of non-overlapping transects that best capture the
110 interannual fluctuations in focal species numbers (individuals, pair, nests, etc.) within the
111 census area. Here the script loops through all non-overlapping transects to find the best fit
112 (linear regression; based on R^2) between the total number found within transects in each
113 year and the true number found within the full census area. The coefficients from the linear
114 regressions are then used to calculate the expected number within the census area based on
115 the transect (survey) data. However, as accuracy is likely to increase with effort (i.e. number
116 of transects), the optimal balance between accuracy and number of transects is determined
117 by having the script looping through the above optimization process across a range of
118 transects from 1 to *n* transects. As a measure of accuracy, we use the offset between the
119 true focal species number and the numbers estimated from the transects. Specifically, we
120 use the absolute deviation between the estimated number and the true total number across
121 all years within the census area, expressed as percentage of the latter.

122

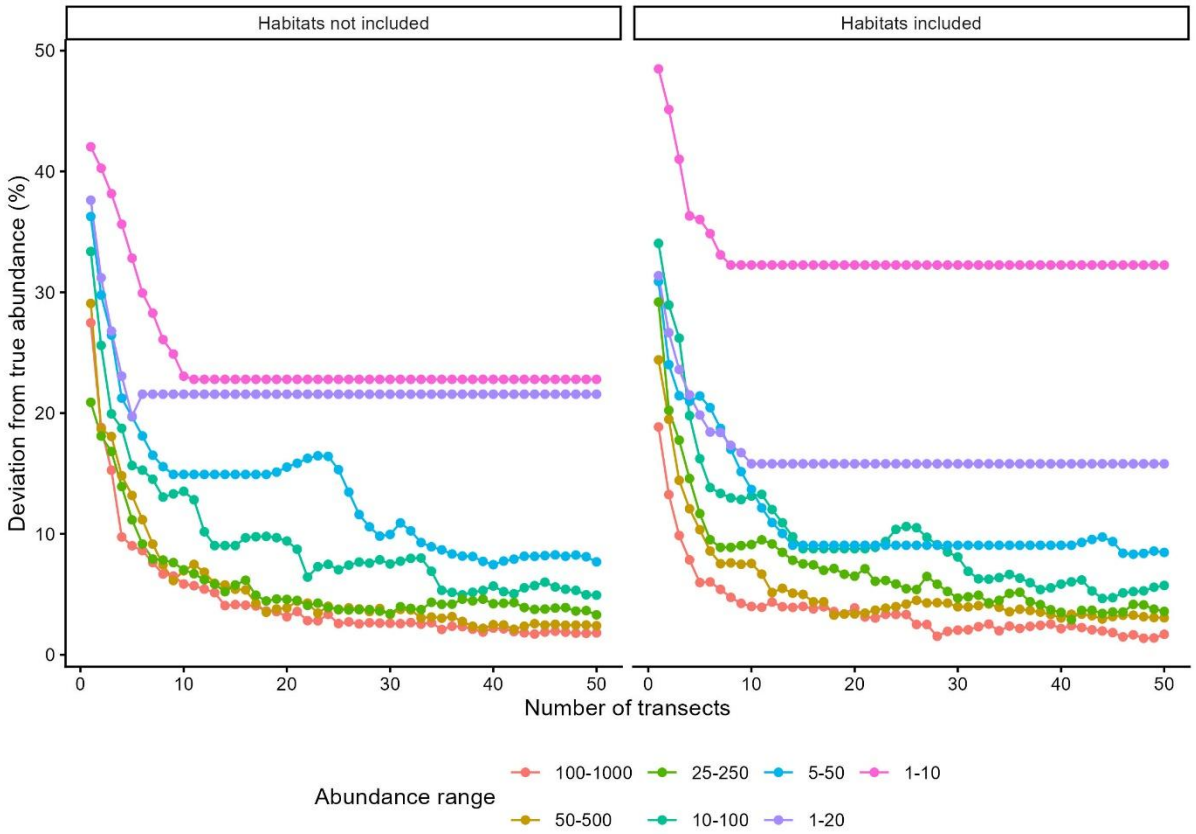
123 **Method validation**

124 ***Simulated data***

125 To examine the generality of our approach, we simulated two types of census data time
126 series (see Online Resource 3). Both data series simulated the location and overall

abundance of a focal species in a 1x1 km census area over a period of 25 years, with one census per year. In the first time series, locations of the focal species were distributed completely randomly within the census area, while in the latter we introduced five habitat classes, each with a specific preference of the focal species. For each of the two types of time series, we generated six scenarios of varying abundance ranges across the study period (i.e. 100-1000, 50-500, 25-250, 10-100, 5-50, 1-20 and 1-10). We then applied our approach to each of the simulated data series within the hypothetical census area, with transect dimensions set to 200 m long and 15 m wide (see Online Resource 3).

For both types of simulated time series, offset decreased with the number of transects included in the sample survey (Figure 1). Offset improvement levelled off after the inclusion of around 25-30 transects. This pattern was most evident in time series with the highest abundances, while offset remained high in time series with low abundance. Interestingly, in the completely random time series offset generally levelled off earlier and reached slightly lower offset levels as compared to the simulated data with habitats and habitat selection of the focal species. On the other hand, in time series with habitats the decline in offset with the inclusion of more transects was slightly faster than in the simulated random time series (Figure 1).



145

146 **Figure 1. Offset between census-based and survey-based abundances.** Left panel shows
147 offsets based on simulated time series with completely random locations (no habitats), while
148 right panel shows offsets based on simulated time series with focal species habitat selection.
149 Seven abundance range scenarios were implemented for each time series type. See Online
150 Resource 3 for further details about the simulated data series.

151

152 **Case study: Collared lemming**

153 To further validate and showcase our methodology, we used existing data from the long-
154 term monitoring program at Zackenberg in High Arctic Kalaallit Nunaat (Greenland). Here,
155 the population dynamics and habitat selection of the collared lemming (*Dicrostonyx*
156 *groenlandicus*) has been monitored since 1996 (Schmidt et al. 2021). The monitoring has

this far consisted of annual full census of a 1.06 km^2 designated census area in which all fresh lemming winter nests are located, georeferenced and examined according to the protocol (Schmidt et al. 2021). The full census is conducted by walking in 5-10 m wide transects marked by temporary poles that are successively moved across the entire census area. The lemming area covers a variety of habitat types, of which lemmings in particular select for *Salix* snow beds (Schmidt et al. 2021). The lemming population exhibit large inter-annual abundance fluctuations (from 13 to 1300 nests within the census area).

Historical georeferenced data on lemming nest location were available from 1996 to 2024 from Greenland Ecosystem Monitoring. Data from 2024 were, however, excluded from the analyses as polar bear presence prevented the complete coverage of the census area. Data from the other years were complete.

Laying out transects

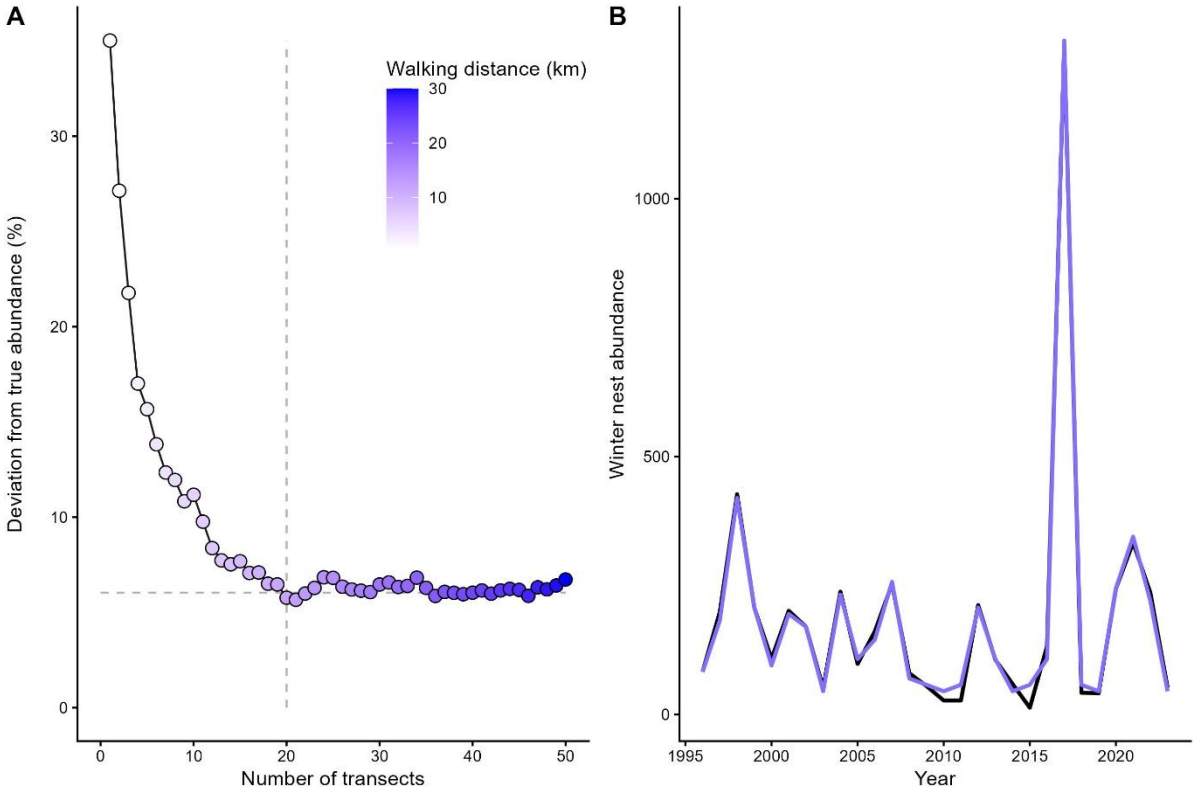
We first generated randomly distributed transects within the census area (see Online Resource 4). Each transect was set to be 300m long and 15m wide. From these transects, we then extracted the non-overlapping transects within the census area (n = 71). As outlined above, we then quantified the offset (% absolute deviance from total true abundance (across all years; n=5155 winter nests)) for 1 to 50 transects to select the optimal number of transects to balance accuracy and effort. In the lemming case study, offset declined markedly with increasing number of transects but levelled out to about 6.0% from around 20 transects onwards, yet with some variation (Figure 2A). Even though similar or slightly lower offset was found when both 38 (6.0%) and 50 transects (5.4%) were included (Figure 2A), improvement of accuracy was limited. Combined with the linear increase in required effort with more transects, we chose 20 non-overlapping, permanent transects as a cost-

efficient set-up. Using these 20 transects, we found a high correlation between the census-based nest number and the survey-based nests ($F_{1,26}$ = 7244, P < 0.001, R^2 = 0.996):

$$Nest\ number_{Census} = 33.458 + 12.018 \cdot Nest\ number_{Survey} \quad \text{(Eq. 1)}$$

Time series integrity

As mentioned above, there may be several metrics that can be used to validate the ability of the sample survey to reflect the full census abundances and dynamics. Here we examined three metrics relevant to lemmings. First, we evaluated how well the interannual variations in lemming winter nest numbers were captured by the sample survey in comparison to the full census. Due to the above high correlation between the two approaches (Eq. 1), the survey-based data adequately represent the census-based data and capture the interannual variation and amplitude in lemming abundance (Figure 2B). Yet, the estimated minimum number of nests in the census area may be somewhat biased as indicated by the intercept value in Eq. 1. Specifically, at low abundances the likelihood of transects holding no nests is high, which would then translate to an estimated minimum number of 33 nests in the census area (Eq. 1). This is about 2.5 times higher than the actual lowest number observed in the census area (n=13), suggesting that nest numbers may be overestimated when abundances are low (Figure 2B). Still, the fluctuations and amplitude in sample survey and full census abundances are remarkably similar, suggesting that this may be a minor problem.



199

200 **Figure 2. Selection of number of transects and evaluation of sample survey performance.**
201 A) shows the offset (%) between the true nest abundances from the full census and
202 abundances estimated based on the sample survey (recalculated using Eq. 1). The vertical
203 dotted line indicates the selected number of transects, with the horizontal dotted line
204 indicating the associated percentage offset. The total walking distance in km required to
205 cover the transects is indicated by the colour code. In B) the recalculated sample survey
206 abundances (using Eq. 1) are shown in blue and the full census abundances in black during
207 the period 1996 to 2023.

208

209 As a second validation metric, we compared the periodicity of the census-based and survey-
210 based data series. Cyclic abundance fluctuations are a key feature of lemming population
211 dynamics (Gauthier et al. 2024) and thus an important metric to be able to capture. To do

so, we performed wavelet analysis on both the census-based and the survey-based time series (Figure 3). As expected from the results in Figure 2B, the periodicity observed using the two census methods was similar, with no qualitative nor quantitative differences between the two (wavelet coherence ranged between 0.901 and 0.999, with a mean of 0.994; Figure 3).

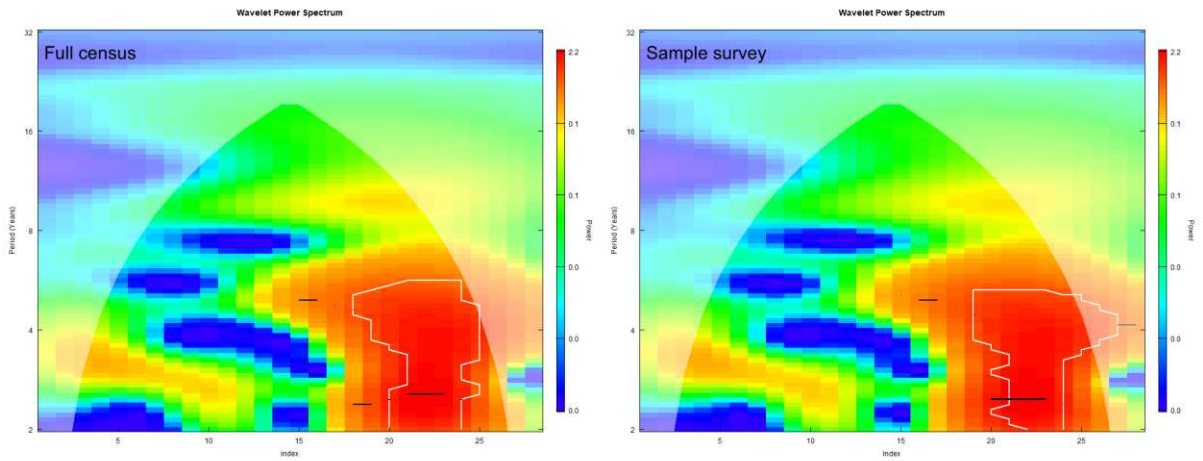


Figure 3. Outputs of wavelet analyses on the Zackenberg collared lemming time series. Wavelet analysis performed on the full census data (left panel) and sample survey data (right panel), the latter recalculated using regression coefficients from Eq. 1.

Finally, as the third validation metric, we examined whether inference about collared lemming habitat selection depends on whether it is based on data from the full census or from the sample survey. We first extracted habitat information for the entire lemming census area using a 5x5m habitat map of the census area (Bay 1998), and the survey transects and quantified the availability of the various habitat types (m^2) in the census area and in the area covered by the transects. The availability of habitat types in the survey transects versus availability in the full census area was highly correlated ($F_{1,6}$ = 775.6, P <

0.001, $R^2 = 0.992$), suggesting that the transects captures the habitat distribution within the full census area well. This is expected as the layout of transects is based on the location of observed nests and hence implicitly includes lemming habitat preferences.

All lemming nests were then annotated with habitat type, and lemming habitat selection was estimated using the isodar approach, following Schmidt et al. (2021). For comparison, isodar analysis was performed on census-based data and survey-based data, the latter recalculated using Eq. 1.

As evident from the isodar results in Table 1, the shift to sample survey does change the estimates of lemming habitat selection as compared to the full census. This is particularly true for the intercept, which is indicative of the quantitative difference between habitats, but also for the slope, which is indicative of qualitative difference between habitats (Morris 1988). Hence, estimating lemming habitat selection based on sample survey is not comparable to estimates based on the full census baseline data. While this is unfortunate, it was not unexpected as the small spatial coverage of the sample survey compared to the full census lowers sample size and increases stochasticity.

Table 1. Summary statistics of isodar analyses of collared lemming (*Dicrostonyx groenlandicus*) habitat selection at Zackenberg 1996 and 2023. Values are intercepts and slopes ± 95% confidence limits, with bold indicating significant deviations from either zero intercept or unity slope ($P<0.05$). R^2 and P values are for the linear regressions. “Full census” indicates analysis of full census area abundances, while “Sample survey” indicates analysis of estimated nest abundances within the census area based on Eq. 1.

	Intercept			Slope			R^2	P
Salix snow bed *vs Grassland*								
Full census	-0.04	±	0.26	**1.33**	±	**0.14**	0.93	<0.001
Sample survey	**-0.42**	**±**	**0.36**	**2.11**	**±**	**0.53**	0.75	<0.001
Salix snow bed *vs Cassiope* heath								
Full census	**0.33**	**±**	**0.19**	**1.24**	±	**0.11**	0.95	<0.001
Sample survey	0.15	±	0.25	**1.45**	**±**	**0.19**	0.89	<0.001
Grassland *vs Cassiope* heath								
Full census	**0.27**	**±**	**0.21**	0.93	±	0.12	0.88	<0.001
Sample survey	**0.27**	**±**	**0.14**	**0.68**	**±**	**0.13**	0.77	<0.001

Effort

It requires approximately 200 km of walking to cover the full lemming census area at Zackenberg. Using 20 transects, the total distance to be covered was reduced to 12km (20 transects á 15x300m, each walked in two 7.5m wide sub-transects to resemble past census technique and to ensure nest detection). This corresponds to a 94% reduction in walking distance and thus constitutes a marked reduction in effort.

Discussion

Changing methodology in long-term monitoring poses a threat to the consistency of time series (Beard et al. 1999), and considerable effort must be made to ensure that the integrity of the time series remain (e.g. Buettel et al. 2025; Ellingsen et al. 2017; Jones et al. 2022). In this article, we have shown that even a substantial methodological change to an existing

long-term monitoring program can be made without compromising the consistency of the time series of abundance. Hence, time series integrity was maintained whilst effort was markedly reduced. However, the decrease in sampling sizes when using sample survey warrants caution when estimating other population-related metrics, such as habitat selection analysis and demographic parameters.

Our methodology is contingent on existing georeferenced data as baseline, which is used to ensure that the transition from census-based to survey-based monitoring is valid. The approach aims to replicate, not predict, past population dynamics, and therefore allows for all available data to be used to optimize the selection of the best performing, permanent transects. It is important to stress that laying out transects in a sample survey to capture past population dynamics differs fundamentally from laying out transects for a new sample study without georeferenced baseline data. In the latter case, one would oftentimes stratify according to relevant parameters such as habitat types, elevation etc. (Gregory et al. 2004). However, in the case presented here, georeferenced baseline census data holds detailed information about where and where not e.g. animals or nests were observed and can thus be seen as implicitly including the relevant stratification. Interestingly, the simulated data suggested that our approach worked marginally better on completely random data (lower offset) as compared to the simulated data with habitats and habitat selection of the focal species. Still, the difference was only marginal.

The overall abundance of the focal species within the census area influences the suitability of the methodology, with offset generally decreasing with increasing abundances, ultimately resulting in a decrease in the number of transects needed to arrive at an acceptable offset. One can, however, also argue that the sample survey is only valid for the range of

abundances observed in the historical data (in the lemming case study between 13 to 1300 nests). While this is true, 1300 nests can be regarded as an extreme outlier as abundances in all other years fluctuate at a much lower level (Figure 2B). This reduces the risk of extrapolating outside the observed abundance range, at least under the range of conditions experienced over the last three decades. Nonetheless, it is important to stress that our approach may only be valid within or close to the abundance ranges encountered in the historical data.

The protocol outlined here and associated R script were initially developed for a specific lemming monitoring program in Kalaallit Nunaat (Greenland). However, as exemplified in the simulated data the approach can be implemented in other monitoring programs based on full census of designated areas, thereby helping to ensure the viability of long-term monitoring programs and their valuable time series. Moreover, a similar approach of finding which sub-areas represent temporal abundance patterns within a larger area the best can be used in situations where for instance the number of monitoring plots or sites must be reduced.

Acknowledgements We thank Naalakkersuisut (Government of Greenland) for allowing us to conduct research and monitoring at Zackenberg and to Greenland Ecosystem Monitoring for access to the ecosystem data. We are indebted to the many field assistants and to the Danish Environmental Protection Agency for their support over the years.

Funding The long-term monitoring at Zackenberg in Kalaallit Nunaat Greenland has been supported financially by the Danish Environmental Protection Agency since 1996.

Data availability Data for the lemming case study are freely available at https://data.g-e-m.dk/datasets/10.17897/WJQB-X013).

Supplementary Information

Online Resource 1: R script providing the functions needed to create random, non-overlapping transects. Functions to be called from Online Resource 3 and 4.

Online Resource 2: Shape file used in Online Resource 2 and collared lemming abundance data and census area used in Online Resource 3.

Online Resource 3: R script to generate simulated time series of abundance and to run the analyses related to the simulated data shown in Figure 1.

Online Resource 4: R script to run all analyses related to the collared lemming case study as shown in Figure 2, Figure 3 and Table 1.

References

Bay, C. 1998. Vegetation mapping of the Zackenberg valley, Northeast Greenland. Danish Polar Center & Botanical Museum, University of Copenhagen, Copenhagen, Denmark. pp. 1-29.

Beard, G.R., Scott, W.A., and Adamson, J.K. 1999. The value of consistent methodology in long-term environmental monitoring. Environ. Monit. Assess. **54**(3): 239-258. doi:10.1023/a:1005917929050.

Buettel, J.C., Lindenmayer, D.B., Scheele, B.C., and Evans, M.J. 2025. Maintaining robust terrestrial ecological monitoring amid technological advancements. Trends Ecol. Evol. **40**(7): 651-662. doi:10.1016/j.tree.2025.04.003.

Ellingsen, K.E., Yoccoz, N.G., Tveraa, T., Hewitt, J.E., and Thrush, S.F. 2017. Long-term environmental monitoring for assessment of change: measurement inconsistencies over time and potential solutions. Environ. Monit. Assess. **189**(11): 595. doi:10.1007/s10661-017-6317-4.

Gauthier, G., Ehrich, D., Belke-Brea, M., Domine, F., Alisauskas, R., Clark, K., Ecke, F., Eide, N.E., Framstad, E., Frandsen, J., Gilg, O., Henttonen, H., Hörnfeldt, B., Kataev, G.D., Menyushina, I.E., Oksanen, L., Oksanen, T., Olofsson, J., Samelius, G., Sittler, B., Smith, P.A., Sokolov, A.A., Sokolova, N.A., and Schmidt, N.M. 2024. Taking the beat of

the Arctic: are lemming population cycles changing due to winter climate? Proc. Royal Soc. B **291**(2016): 20232361. doi:doi:10.1098/rspb.2023.2361.

Gregory, R.D., Gibbons, D.W., and Donald, P.F. 2004. Bird census and survey techniques. *In* Bird Ecology and Conservation: A Handbook of Techniques. Oxford University Press. pp. 17-56.

Hughes, B.B., Beas-Luna, R., Barner, A.K., Brewitt, K., Brumbaugh, D.R., Cerny-Chipman, E.B., Close, S.L., Coblentz, K.E., De Nesnera, K.L., Drobnitch, S.T., Figurski, J.D., Focht, B., Friedman, M., Freiwald, J., Heady, K.K., Heady, W.N., Hettinger, A., Johnson, A., Karr, K.A., Mahoney, B., Moritsch, M.M., Osterback, A.-M.K., Reimer, J., Robinson, J., Rohrer, T., Rose, J.M., Sabal, M., Segui, L.M., Shen, C., Sullivan, J., Zuercher, R., Raimondi, P.T., Menge, B.A., Grorud-Colvert, K., Novak, M., and Carr, M.H. 2017. Long-Term Studies Contribute Disproportionately to Ecology and Policy. Bioscience **67**(3): 271-281. doi:10.1093/biosci/biw185.

Jones, C.S., Duncan, D.H., Morris, W.K., Robinson, D., and Vesk, P.A. 2022. Using data calibration to reconcile outputs from different survey methods in long-term or large-scale studies. Environ. Monit. Assess. **194**(3): 185. doi:10.1007/s10661-021-09727-2.

Lindenmayer, D.B., and Likens, G.E. 2009. Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends Ecol. Evol. **24**: 482-486. doi:10.1016/j.tree.2009.03.005.

Meltofte, H., Hansen, J., and Rigét, F. 2020. Increased variability in bird populations at Zackenberg, high Arctic Greenland 1996-2018, but few unidirectional trends or correlations with local conditions. Dansk Ornithologisk Forenings Tidsskrift **114**: 151-167.

Morris, D.W. 1988. Habitat-dependent population regulation and community structure. Evol. Ecol. **2**(3): 253-269.

Moullec, M.L., Pedersen, Å.Ø., Yoccoz, N.G., Aanes, R., Tufto, J., and Hansen, B.B. 2017. Ungulate population monitoring in an open tundra landscape: distance sampling versus total counts. Wildl. Biol. **1**: 1-11. doi:10.2981/wlb.00299.

Rominger, K.R., and Meyer, S.E. 2021. Drones, Deep Learning, and Endangered Plants: A Method for Population-Level Census Using Image Analysis. Drones **5**(4): 126. doi:10.3390/drones5040126.

Schmidt, N.M., van Beest, F.M., Dupuch, A., Hansen, L.H., Desforges, J.P., and Morris, D.W. 2021. Long-term patterns in winter habitat selection, breeding and predation in a density-fluctuating, high Arctic lemming population. Oecologia **195**: 927–935. doi:10.1007/s00442-021-04882-2.

Viblanc, V.A., and Muller-Landau, H.C. 2025. Insights Into Ecology, Evolution and Global Change Responses From Very Long-Term Studies. Ecol. Lett. **28**(4): e70072. doi:10.1111/ele.70072.

Von Beckerath, X., Benadi, G., Gilg, O., Sittler, B., Yannic, G., Klein, A.-M., and Eitzinger, B. 2021. Long-term monitoring reveals topographical features and vegetation explain winter habitat use of an arctic rodent. Arctic Science **8**(2): 349-361. doi:10.1139/as-2021-0010.