Sampling principles for biodiversity study

Xubin Pan

Institute of Plant Quarantine, Chinese Academy of Inspection and Quarantine, Beijing 100029, China

E-mail: xubin.hu.pan@gmail.com

Abstract

Sampling is a fundamental tool in ecology and critical for biodiversity measurement. However, basic principles of biodiversity sampling have been overlooked for many years. In this paper, I proposed and explored five principles of sampling for a specific area and biodiversity study. The first principle of sampling, *species increasing with area*, is that the number of species increases with the area. The second principle of sampling, *individuals increasing with area*, is that the number of individuals increases with the area. The third principle of sampling, *sum of species number*, is that the sum of species number in one area and species number in another area is no less than the total species number in the two areas. The fourth principle of sampling, *individual complement*, is that the sum of the mathematical expectation of individual number of one or several species in the area a and that of the same one or several species in the area A-a is the total individual number N of the same one or several species in the total area A. The fifth principle of sampling, *species-area theory*, is that the sum of the mathematical expectation of number of species in the area a and that of number of species lost if area A-a is cleared is the total species number M in the total area A.

Keywords: Individuals-area relationship, Species-area relationship, Endemics-area relationship

Introduction

Sampling is a fundamental tool in ecology. Because the complete investigation of one ecological question maybe involves a large area and long time, it will cost too much time and money, and leads to infeasible in practice. Thus, sampling is unavoidable in the ecology study. However, due to the complexity of ecosystem, such as interactions among organisms and environment and their large spatial-temporal heterogeneity, how to sampling should be very carefully, which will decide how to analyze the data from the sampling and the credibility of the results. Thus, solid sampling theory is very important for ecology, although it didn't get too much concern recently. As the core of the ecology, the spatial relationship between the species and the area can be a good start for rethink the sampling theory.

Individual–area relationship (IAR), species-area relationship (SAR) and endemics-area relationship (EAR) are important concepts in biodiversity conservation and habitat preservation (WCMC, 1992; Kinzig and Harte, 2000; Connor et al., 2000; Green and Ostling, 2003; Chris et al., 2004; World Resources Institute, 2005). However, there is still much debating over the estimation of extinction rate based on SAR, which tends to make overestimation when compared with observed extinction (Pimm and Askins, 1995; Rosenzweig, 1995; Harte and Kinzig, 1997; He and Hubbell, 2011; Pan, 2013). One explanation is that such estimation includes certain species that are "committed to extinction" instead of going extinct right away after habitat clearing (Heywood, 1994; Tilman, 1994; Mace et al., 2003). Recently, He and Hubbell (2011) suggested that the reason was the difference of between sampling areas based SAR and EAR, because a sample area of a species for extinction is often larger than a sample

bioRxiv preprint doi: https://doi.org/10.1101/000364; this version posted November 14, 2013.

area of the same species for existence. Among these controversies, the core problem is the sampling of biodiversity measurement, the basic principles of which have been overlooked for many years.

However, much work has been focused on statistical or mathematical calculation based on SAR, rather than the biological and ecological implication, especially the basic principles for biodiversity sampling (Turner and Tjørve, 2005). Are there common laws for sampling in biodiversity measurement? Here, we proposed five sampling principles for biodiversity study in a specific area and the last two were proved, focusing on the change of species number and individuals for one or several species with a changing area. This analysis will be helpful for the establishment of theory platform for biodiversity and other ecological discussion.

Theoretical Frame

Biodiversity not just study the number of species, but also the amount of individuals. For the sampling problems of biodiversity, is the relationship between the area and the number of species and amount of individuals. Usually, it is thought the relationship between the number of species and the area is saturation curve, and the relationship between the amount of individuals and the area is increasing curve. The shape of this curve is influenced by the spatial distribution and sampling collection/statistic.

Sampling principles for biodiversity measurement

The first principle of sampling, species increasing with area, is that the number of species

bioRxiv preprint doi: https://doi.org/10.1101/000364; this version posted November 14, 2013.

increases with the area. Assume the number of species is m in the area a, when the sampling area increases from a to a', the number of species in area a' is m', and $m' \geq m$. If no new species emerges in the area $a' - a$, $m' = m$.

The second principle of sampling, individuals increasing with area, is that the number of individuals increases with the area. Assume the number of individuals is n in the area a, when the sampling area increases from a to a', the number of individuals in area a' is m', and $n' \geq n$. If there no new individuals emerges in the area $a' - a$, $n' = n$.

The third principle of sampling, sum of species number, is that the sum of species number in one area and species number in another area is no less than the total species number in the two areas. Assume the number of species is m_1 in the area a_1, the number of species is m_2 in the area a_2, the number of species in area $a_1 + a_2$ is m', and $m' \leq m_1 + m_2$. If there is no overlapping species between areas a_1 and a_2, $m' = m_1 + m_2$.

The fourth principle of sampling, individual complement, is that the sum of the mathematical expectation of individual number of one or several species in area *a* and that of individual number of the same one or several species in area *A*-*a* is the total individual number *N* of the same one or several species in the total area *A*.

The fifth principle of sampling, species-area theory, is that the sum of the mathematical expectation of number of species in area *a* and that of number of species lost if area *A*-*a* is cleared is the total species number M in the total area A.

Mathematical proof for the fourth and fifth principles of biodiversity sampling

While the first, second and third principles of biodiversity sampling are straightforward and easy to understand, a complementary method in a specific area is used to verify the fourth and fifth principles.

The individual-area relationship (IAR) I_a in area a and I_{A-a} in area A-a, for one sampling, if the sampling area is a and the number of individuals in this area is n, there is always one-one corresponding area of N-n individuals of the species in area A-a (A is the total area, and N is the total individuals for one or several species in area A).

The mathematical expectation of individual-area relationship in area a is

$$ME(I_a) = \sum_{i=1}^{k} I_{a,i} P_i$$

where $ME(I_a)$ are the expected individuals for the individuals-area relationship in area a in the ith sampling, k is the total samples, and the P_i is the possibility of the corresponding ith sampling.

The mathematical expectation of individual-area relationship in area A-a is

$$ME(I_{A-a}) = \sum_{i=1}^{k} I_{A-a,i} Q_i$$

where $ME(I_{A-a})$ are the expected individuals for the individual-area relationship in area A-a in the ith sampling, and the Q_i is the possibility of the corresponding ith sampling.

For any specific sampling i, there is always a one-one corresponding $I_{a,i} + I_{A-a,i} = N$, and the

$P_i = Q_i$ in this situation.

$$ME(I_a) + ME(I_{A-a}) = \sum_{i=1}^{k} I_{a,i}P_i + \sum_{i=1}^{k} I_{A-a,i}Q_i = \sum_{i=1}^{k}\left(I_{a,i} + I_{A-a,i}\right) * P_i = \sum_{i=1}^{k} NP_i = \mathrm{N}\sum_{i=1}^{k} P_i$$
$$= \mathrm{N}$$

The number of species in area *a* is *m*, then $S_a = m$, where species-area relationship (SAR) S_a is the function of the species number with area *a*. The number of species lost in area *A-a* if the area *A-a* is cleared is *M-m*, endemics-area relationship (EAR) E_{A-a} is the number of species disappearing if area *A-a* is cleared.

The mathematical expectation of Species-area relationship in area *a* is

$$ME(S_a) = \sum_{i=1}^{k} S_{a,i}P_i$$

where $S_{a,i}$ is the species number in area *a* in the *i*th sampling, *k* is the total samples, and the P_i is the possibility of the corresponding *i*th sampling.

The mathematical expectation of Endemics-area relationship in area *A-a* is

$$ME(E_{A-a}) = \sum_{i=1}^{k} E_{A-a,i}Q_i$$

where $E_{A-a,\, i}$ is the number of species disappear if area *A-a* is lost in the *i*th sampling, and the Q_i is the possibility of the corresponding *i*th sampling.

For any specific sampling *i*, there is always a one-one corresponding $S_{a,i} + E_{A-a,i} = M$, and the $P_i = Q_i$ in this situation.

$$ME(S_a) + ME(E_{A-a}) = \sum_{i=1}^{k} S_{a,i}P_i + \sum_{i=1}^{k} E_{A-a,i}Q_i = \sum_{i=1}^{k}(S_{a,i} + E_{A-a,i})P_i$$
$$= \sum_{i=1}^{k} MP_i = M\sum_{i=1}^{k} P_i = M$$

Conclusions

Sampling is a fundamental tool in ecology and critical for biodiversity measurement. However, basic principles of biodiversity sampling have been overlooked for many years. In this paper, I proposed and explored five principles of sampling for a specific area and biodiversity study. The first principle of sampling, *species increasing with area*, is that the number of species increases with the area. The second principle of sampling, *individuals increasing with area*, is that the number of individuals increases with the area. The third principle of sampling, *sum of species number*, is that the sum of species number in one area and species number in another area is no less than the total species number in the two areas. The fourth principle of sampling, *individual complement*, is that the sum of the mathematical expectation of individual number of one or several species in the area a and that of the same one or several species in the area A-a is the total individual number N of the same one or several species in the total area A. The fifth principle of sampling, *species-area theory*, is that the sum of the mathematical expectation of number of species in the area a and that of number of species lost if area A-a is cleared is the total species number M in the total area A.

Acknowledgments

I warmly thank Fengqiao Liu at Arizona State University and Dr. James Rosindell in University of Leeds for the helpful suggestions for the manuscript revision.

References

Achard F et al. 2002. Determination of deforestation rates of the world's humid tropical forests. Science 297: 999–1002.

Connor EF, Courtney AC and Yoder JM. 2000. Individuals–area relationships: the relationship between animal population density and area. Ecology 81: 734–748.

Green JL and Ostling A. 2003. Endemics-area relationships: the influence of species dominance and spatial aggregation. Ecology 84: 3090-3097.

Green JL, Harte J and Ostling A. 2003. Species richness, endemism, and abundance patterns: tests of two fractal models in a serpentine grassland. Ecology Letters 6: 919-928.

Harte J and Kinzig AP. 1997. On the implications of the species-area relationships for endemism, spatial turnover, and food web patterns. Okios 80: 417-427.

Harte J, Kinzig AP and Green J. 1999. Self-similarity in the distribution and abundance of species. Science 284: 334-336.

He F and Hubbell SP. 2011. Species-area relationships always overestimate extinction rates from habitat loss. Nature 473: 368-371.

Heywood VH, Mace GM, May RM and Stuart SN. 1994. Uncertainties in extinction rates. Nature 368: 105.

Kinzig AP and Harte J. 2000. Implications of endemics-area relationships for estimates of species extinctions. Ecology 81: 3305-3311.

Mace GM, Gittleman JL and Purvis A. 2003. Preserving the tree of life. Science 300: 1707-1709.

Pan, XB. 2013. Fundamental equations for species-area theory. Scientific Reports. 2013, 3: 1334.

bioRxiv preprint doi: https://doi.org/10.1101/000364; this version posted November 14, 2013.

Pimm SL and Askins RA. 1995. Forest losses predict bird extinctions in eastern North America. Proceedings of National Academic of Sciences 92: 9343–9347.

Rosenzweig ML. 1995. Species Diversity in Space and Time. Cambridge University Press, Cambridge.

Thomas CD. et al. 2004. Extinction risk from climate change. Nature 427: 145-148.

Tilman D, May RM, Lehman CL and Nowak MA. 1994. Habitat destruction and the extinction debt. Nature 371: 365–366.

Turner WR and Tjørve E. 2005. Scale-dependence in species-area relationships. Ecography 28: 721-730.

WCMC. 1992. Global Biodiversity: Status of the Earth's Living Resources. Chapman & Hall, London.

World Resources Institute. 2005. Millennium Ecosystem Assessment. Ecosystems and Human Well-Being: Biodiversity Synthesis.

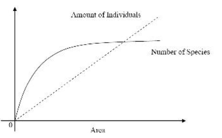
bioRxiv preprint doi: https://doi.org/10.1101/000364; this version posted November 14, 2013.

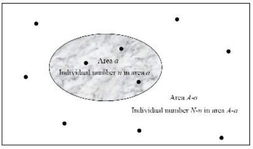
Figures Captions

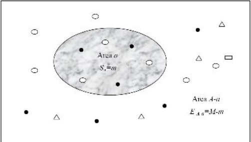
Figure 1. Relationship between the number of species and amount of individuals and the area

Figure 2. Individual-area relationship I_a in area a and $I_{A\text{-}a}$ in area A-a.

Figure 3. Species-area relationship S_a in area a and endemics-area relationship $E_{A\text{-}a}$ in area A-a.


Amount of Individuals
Number of Species
0
Area


Area a
Individual number n in area a
Area A-a
Individual number N-n in area A-a



Area a
$S_a=m$
Area A-a
$E_{A\text{-}a}=M$-m