Application of fundamental equations to species—area theory

Xubin Pan*

Abstract

Background: Species—area relationship (SAR), endemics-area relationship (EAR) and overlap-area relationship (OAR) are three important concepts in biodiversity study. The application of fundamental equations linking the SAR, EAR and OAR, can enrich the axiomatic framework of the species—area theory and deepen our understanding of the mechanisms of community assembly.

Results: Two fundamental equations are derived and extended to power law model and random replacement model of species—area distribution. Several important parameters, including the overlap index and extinction rate, are defined and expressed to enrich the species—area theory. For power law model, both EAR and OAR have three parameters, with one more parameter of the total area than SAR does. The EAR equation is a monotonically increasing function for parameter $c$ and $z$, and a monotonically decreasing function for parameter $A$. The extinction rate, with two parameters, is a monotonically increasing function for parameter $z$, and a monotonically decreasing function for parameter $A$. The overlap index is a monotonically increasing function for parameter $A$, and a monotonically decreasing function for parameter $z$, independent of parameter $c$.

Conclusions: The general formats of SAR, EAR, OAR, overlap index, overlap rate, sampling rate and extinction rate, are derived and extended to power law model and random replacement model as the axiomatic framework of species—area theory. In addition, if the total area is underestimated, the extinction rate will be overestimated.

Keywords: Endemics-area relationship, Overlap-area relationship, Power law, Random replacement, Real total area, Sampling rate, Overlap rate, Extinction rate, Overlap index

Background

Species—area relationship (SAR) is a core concept in biodiversity and species distribution [1], and endemics-area relationship (EAR) is a useful tool in biodiversity conservation and habitat preservation [2–5]. Besides SAR and EAR, overlap-area relationship (OAR), which refers to the number of overlap species in two areas, is also a relevant and important concept [6–8]. To link SAR and EAR and develop a complete species—area theory, two fundamental equations are established to describe species distribution and interrelation between two compensatory areas [7]. Now the species—area theory has been reconstructed by the set theory, integrating SAR, EAR, OAR, alpha diversity, beta diversity, and gamma diversity [8]. Although OAR curves for two areas of the same size are described and zeta diversity as the average number of species shared by multi-assemblages is proposed, the expanding concept that compares two or more areas of different sizes has not been fully discussed yet [8, 9]. Furthermore, to investigate the spatial characteristics of species richness, it is necessary to integrate the two fundamental equations into the species—area model with distribution information or assumption. Then more parameters can be defined and expressed with empirical data, which can enrich the axiomatic framework of the mechanisms or processes of community assembly.

In addition, debate still exists over the estimation of extinction rate based on the SAR, which is higher than observed extinction rate [10–13]. One explanation for the overestimation is that some species are "committed to
extinction” instead of going extinct due to habitat clearing [14–16]. However, another reason has been ignored in this debate [8]. According to the power law model, the SAR is a two-parameter equation, whereas the EAR is a three-parameter equation. It does not consider total area in SAR, while total area and its corresponding total species number are crucial factors to determine species disappearing and extinction rate. However, the impact of total area on the extinction rate is still unknown without the specific species–area model and sensitivity analysis.

In this paper, power law and random replacement functions, both of which are widely used species–area models, were selected for the application of two fundamental equations [17–19]. Then several important parameters were defined and expressed to enrich the species–area theory. For power law model, sensitivity analysis of parameters was conducted for EAR, extinction rate and overlap index, and the extinction rate based on different total areas were assessed for overestimate comparison.

**Methods**

The relationships among SAR, EAR and OAR have been shown in Fig. 1, where $S_A$ is the number of species in area $A$, $E_A$ is the number of species that will disappear when habitat area $A$ is cleared, $O_{A,A-a}$ is the number of overlap species in two areas $a$ and $A-a$.

$$S_A(=E_A)$$

is the number of total species in the total area $A$. These relationships can be connected by two fundamental equations [7]:

$$S_a + E_{A-a} = S_A,$$  \tag{2}

and

$$O_{a,A-a} = S_a - E_a.$$  \tag{3}

These fundamental Eqs. (2 and 3) for species–area theory were applied to power law model and random replacement model of SAR. To enrich the species–area theory, several parameters were proposed, including overlap index, overlap rate, sampling rate and extinction rate, which were defined by equations in the general format, power law model and random replacement model.

For power law model, sensitivity analysis of parameters was conducted for EAR, extinction rate and overlap index, and the extinction rate based on different total areas was assessed for overestimate comparison. The data can be downloaded from the Supplementary of Data.

**Results**

**Application of two fundamental equations to power law model**

Power law model has been widely used for species–area relationship:

$$S_a = ca^z,$$  \tag{4}

where $S_a$ is the number of species in area $a$, and $c$ ($a = 1$, $c = S_j$) and $z$ ($0 \leq z \leq 1$) are fitted constants [11, 17]. However, the corresponding function of endemics-area relationship (EAR) has not received much attention, such as in He and Hubbell’s paper [12],

$$S_{loss} = S_A - S_{A-a} = cA^z - c(A-a)^z,$$  \tag{5}

where $S_{loss}$ is the number of species that disappear when habitat area $a$ is cleared, $A$ is the total area, and $S_A$ is the total number of the species in area $A$. In fact,

$$E_a = S_{loss} = cA^z - c(A-a)^z,$$  \tag{6}

where $E_a$ is the number of species that exist only in area $a$, but not in area $A - a$. Because

$$a + (A - a) = A, \quad S_a + E_{A-a} = cA^z + cA^z - c(A - (A - a))^z = cA^z = S_A = E_A$$  \tag{7}

($E_A$ is the total number of specific species in area $A$), another method to get EAR is to derive it based on the relationship between SAR and EAR. If area $A - a$ is cleared while area $a$ remains, the number of species that will disappear is $cA^z - cA^z$. Thus, we can get the endemics-area curves,

$$E_{A-a} = cA^z - cA^z, \quad E_a = cA^z - c(A - a)^z,$$  \tag{8}

the same as those that are derived with the former method. SAR and EAR are rotationally symmetrical, with the center at $(A/2, S/2)$ [7]. The OAR ($O_{a,A-a}$) between the area $a$ and $A - a$ can be calculated as

$$O_{a,A-a} = S_a + S_{A-a} - S_A = ca^z + c(A-a)^z - cA^z$$

$$= S_a - E_a.$$  \tag{9}

When it comes to $a = A - a$, $O_{a,A-a}$ attains its peak maximum value of $2c(\frac{A}{2})^z - cA^z$. When $A = 1280$, $c = 25$, and $z = 0.25$, the SAR, EAR and OAR curves are shown in Fig. 2.

For power law format of SAR, both EAR and OAR equations have three parameters: $c$, $z$ and $A$. The EAR equation derived from power law function is different.
from the previous power law format of “endemics-area relationship”:

\[ E_a = c' a^{z'}, \quad (10) \]

where

\[ c' = c A^{z-z'} \quad (11) \]

\[ 1/2^z = 1 - 1/2^z \quad (12) \]

which is based on the community-level fractal model [4, 12, 20]. A comparison of these two EAR curves is shown in Fig. 3, where three intersecting points are: (0, 0), \((A/2, c A^z-c A^z/2^z)\), and \((A, c A^z)\). In the interval \((0, A/2)\), power law format of “endemics-area relationship” based on fractal model underestimates the number of endemic species; in the interval \((A/2, A)\), however, the power law format overestimates the number of endemic species. Both formats are three-parameter equations, but the latter equation is derived based on community-level fractal and power-law assumptions, which has decreased the accuracy of the model. Additionally, the former equation is simple and easy for parameter fitting. In the EAR curve, the species number decreases slowly at the beginning of habitat loss. Due to accumulation effect of habitat loss, the extinction rate of endemic species will speed up until all species disappear. Although the species extinction rate seems small at the beginning of land clearing, it will be too late to conserve biodiversity when most of the habitat disappears.

Overlap index \(h\) is proposed here,

\[ h = \frac{O_{aA-a}}{S_a} = \frac{S_a - E_a}{S_a} = 1 - \frac{E_a}{S_a}. \quad (13) \]

When \(a = A/2\), \(h\) reaches \(2 - 2^z\), the same value as obtained with bisection scheme [6]. We can get the overlap index for power law model as

\[ h = \frac{O_{aA-a}}{S_a} = \frac{S_a - E_a}{S_a} = 1 - \frac{E_a}{S_a}. \quad (14) \]

where \(x = \frac{a}{A}\) is the ratio of area \(a\) over total area \(A\). The overlap index \(h\) is the function of \(x\) and \(z\), independent of \(c\). Then parameter \(z\) reflects the overlapping or self-similarity properties of species number in power law function of SAR [6]. The ratio of overlapping species number over the total species number, overlap rate, is

\[ h' = \frac{O_{aA-a}}{S_A} = \frac{S_a - E_a}{S_A}. \quad (15) \]

Since \(S_A\) is constant, \(h'\) has a similar shape with \(O_{aA-a}\) and will reach its peak value at \(2^{1-z} - 1\) when \(a = A/2\). The ratio of \(S_a\) over \(S_A\), sampling rate, is

\[ \eta = \frac{S_a}{S_A} = c \left( \frac{a}{A} \right)^z = c x^z. \quad (16) \]

The extinction rate is

\[ \lambda = \frac{E_a}{S_A} = \frac{c A^z - c (A-a)^z}{c A^z} = 1 - (1 - x)^z, \quad (17) \]

the format of which coincides with the previous extinction rate in the species–area curves [3].
Application of two fundamental equations to random replacement model

Random replacement models for SAR and EAR are

\[ S_a = S_a^1 = S_A - \sum_{i=1}^{S_A} \left( 1 - \frac{a}{A} \right)^{N_i}, \]  

(18)

\[ E_a = S_a^N = \sum_{i=1}^{S_A} \left( \frac{a}{A} \right)^{N_i}, \]  

(19)

respectively, where \( N_i \) is the number of individuals of species \( i \), and \( S_a^1 \) and \( S_a^N \) are species-area curve and endemics-area curve across all species in \( A \), respectively [12, 17].

\[ S_a + E_{A-a} = S_A - \sum_{i=1}^{S_A} \left( 1 - \frac{a}{A} \right)^{N_i} + \sum_{i=1}^{S_A} \left( \frac{A-a}{A} \right)^{N_i} \]

(20)

corresponding with the fundamental equation.

\[ O_{a,A-a} = S_a - E_a \]

\[ = S_A - \sum_{i=1}^{S_A} \left( 1 - \frac{a}{A} \right)^{N_i} - \sum_{i=1}^{S_A} \left( \frac{a}{A} \right)^{N_i} \]

\[ = S_A - \sum_{i=1}^{S_A} \left[ \left( 1 - \frac{a}{A} \right)^{N_i} - \left( \frac{a}{A} \right)^{N_i} \right] \]

(21)

\[ O_{a,A-a} = S_a + S_{A-a} - S_A = S_A - \sum_{i=1}^{S_A} \left( 1 - \frac{a}{A} \right)^{N_i} + S_A - \sum_{i=1}^{S_A} \left( \frac{a}{A} \right)^{N_i} \]

\[ = S_A - \sum_{i=1}^{S_A} \left[ \left( 1 - \frac{a}{A} \right)^{N_i} - \left( \frac{a}{A} \right)^{N_i} \right] \]

(22)

\[ h = \frac{O_{a,A-a}}{S_a} = \frac{S_a - E_a}{S_a} \]

\[ = 1 - \frac{E_a}{S_a} \]

\[ = 1 - \frac{\sum_{i=1}^{S_A} \left( \frac{a}{A} \right)^{N_i}}{S_A - \sum_{i=1}^{S_A} \left( 1 - \frac{a}{A} \right)^{N_i}} \]

(23)

\[ h' = \frac{O_{a,A-a}}{S_A} = \frac{S_a + S_{A-a} - S_A}{S_A} \]

\[ = \frac{2S_A - \sum_{i=1}^{S_A} \left[ \left( 1 - \frac{a}{A} \right)^{N_i} + \left( \frac{a}{A} \right)^{N_i} \right]}{S_A} - 1 \]

(24)

When \( a = A/2 \), \( O_{a,A-a} \) and \( h' \) reaches its maximum value

\[ O_{a,A-a} = S_A - 2 \sum_{i=1}^{S_A} \left( \frac{1}{2} \right)^{N_i} \]

(25)

\[ h = \frac{2S_A - S_A}{S_{A/2}} \]

\[ = 2 - \frac{S_A}{S_{A/2}} \]

\[ = 2 - \frac{S_A}{S_A - \sum_{i=1}^{S_A} \left( 1 - \frac{a}{2} \right)^{N_i}} \]

\[ = 2 - \frac{1}{1 - \sum_{i=1}^{S_A} \left( \frac{1}{2} \right)^{N_i}} \]

(26)

\[ h' = \frac{2 \left[ S_A - \sum_{i=1}^{S_A} \left( \frac{1}{2} \right)^{N_i} \right]}{S_A} - 1 \]

(27)

\[ \eta = \frac{S_a}{S_A} = 1 - \frac{E_{A-a}}{S_A} = 1 - \frac{\sum_{i=1}^{S_A} \left( 1 - \frac{a}{A} \right)^{N_i}}{S_A} \]

(28)

\[ \lambda = \frac{E_a}{S_A} = \frac{\sum_{i=1}^{S_A} \left( \frac{a}{A} \right)^{N_i}}{S_A} \]

(29)

The general format, power law model and random replacement model for SAR, EAR and OAR are shown in Table 1.

Sensitivity analysis for power law model

The EAR equation is a monotonically increasing function for parameter \( c \). In Fig. 4a, when \( c \) increases from 10 to 50, the number of extinct species increases from 32 to 158. The EAR equation is a monotonically increasing function for parameter \( z \). In Fig. 4b, when \( z \) increases from 0.1 to 0.5, the number of extinct species increases from 40 to 250. The EAR equation is a monotonically decreasing function for parameter \( A \). In Fig. 4c, when \( A \) increases from 9 to 100, the number of extinct species in the same area decreases, providing an important theoretical support for large habitat preservation.

Extinction rate is a monotonically increasing function for parameter \( z \) in the interval (0, 1). In Fig. 5a, when the fraction of habitat loss \( x \) increases from 0 to 1, the extinction rate also increases from 0 to 1. The extinction rate derived from EAR equation displays a similar pattern for parameter \( A \). In Fig. 5b, when \( A \) increases from 9 to 100, the percentage of extinct species in the same area decreases.
Table 1 General format, power law and random placement models for SAR, EAR and OAR

| Parameters | General format | Power law model | Random placement model |
|------------|----------------|----------------|------------------------|
| SAR        | \( S_a + E_a = S_E = E_a \) | \( S_a = ca^2 \) | \( S_a = S_{Oa} = S_a = \frac{S_a}{\sum_{i=1}^{N_a} (1 - \frac{x}{a})^N} \) |
| EAR        | \( O_{A \rightarrow a} = O_{A} - O_{a} \) | \( O_{A} = ca^2 + c(A - a)^2 - cA^2 \) | \( O_{A} = \frac{S_{OA}}{\sum_{i=1}^{N_a} (1 - \frac{x}{a})^N + (\frac{x}{a})^N} \) |
| OAR        | \( h, a = A/2 \) | \( h = 2 - 2^z \) | \( h' = \frac{2S_{OA} - S_a}{S_{OA}} \) |
|            | \( h', a = A/2 \) | \( h' = 2^{1-z} - 1 \) | \( h' = 1 - \frac{\sum_{i=1}^{N_a} (1 - \frac{x}{a})^N}{S_a + \sum_{i=1}^{N_a} (1 - \frac{x}{a})^N} \) |
|            | \( \eta = \frac{S_a}{S_{OA}} \) | \( \eta = c(\frac{x}{a})^2 \) | \( \eta = 1 - \frac{\sum_{i=1}^{N_a} (1 - \frac{x}{a})^N}{S_a + \sum_{i=1}^{N_a} (1 - \frac{x}{a})^N} \) |
|            | \( \lambda = \frac{S_a}{S_{OA}} \) | \( \lambda = 1 - c(1 - \frac{x}{a})^z \) | \( \lambda = \frac{\sum_{i=1}^{N_a} (1 - \frac{x}{a})^N}{S_a + \sum_{i=1}^{N_a} (1 - \frac{x}{a})^N} \) |

SAR species—area relationship; EAR endemic area relationship; OAR overlap area relationship; \( S_a \) is the number of species in area \( a \); \( E_a \) is the number of species only in the area \( a \), but not in the area \( A \); \( O_{A \rightarrow a} \) is the number of species both in the area \( a \) and \( A \); \( c (a = 1, c = S_{OA}) \) and \( z (0 < z < 1) \) are constants; \( A \) is the total area, \( S_a \) is the total number of the species in the area \( A \), and \( E_a \) is the total number of specific species in area \( A \); \( h \) is overlap index; \( h' \) is the ratio of overlapping species number over the total species number, overlap ratio; \( \eta \) is the ratio of \( S_a \) over \( S_{OA} \) sampling rate; \( \lambda \) is extinction rate; \( N_i \) is the number of individuals of the specific species \( i \); \( S_{OA} \) and \( S_a \) are species—area curve and endemics-area curve across all species in \( A \) [3, 6, 7, 10, 12, 16, 17].

A sensitivity analysis of two parameters (\( z \) and \( A \)) is conducted for the overlap index \( h \). The \( h \) equation is a monotonically increasing function for parameter \( A \) in Fig. 6a, when \( A \) increases from 320 to 1280. The \( h' \) equation is a monotonically decreasing function for parameter \( z \) in Fig. 6b, when \( z \) increases from 0.15 to 0.5.

Discussion

Table 2 shows the extinction rate estimate and overestimate comparison. If the real total area is 100, overestimate would occur when the total area is set to, say, 9 or 49. If the total area \( A \) is underestimated for 51 %, the overestimate of extinction rates will be 1.04 and 1.36 for 0.52 and 25 % of real total habitat loss, respectively [13, 21, 22]. However, the power law format of EAR will underestimate the extinction rate for the 0.52 and 25 % of real total habitat loss. If the total area \( A \) is underestimated for 91 %, the overestimate of extinction rates will be 10.34 and 40.91 for 0.52 and 9 % of real total habitat loss, respectively. Thus for field estimate of extinction rate, species—area relationship is an important tool, plus boundary identification of real total area, which can be assisted by remote sensing and geographic information system. If a smaller total area is adopted compared with the actual area, the species extinction rate estimate will be overstated.

Based on EAR, however, small total area for habitat preservation does lead to potential high species extinction rate. Thus large total area should be adopted for the Natural Protected Areas (NPAs). UNESCO-MAB World Network of Biosphere Reserves, suggests to apply a zonation system to NPAs, which consists of a core zone, a buffer zone and a transition zone. Normally, both the buffer zone and the transition zone do not have any different or concerned species that are not in the core zone, thus the total number of species will decrease when the protected area is expanded from core zone to include the buffer zone and transition zone. But both the buffer zone and transition zone can relax the influence of anthropic activities on the core zone, and this result can be derived from the species—area theory.

Since EAR and OAR involve species in two complementary areas, one more parameter, the total area, has been added in their expressions compared with SAR. If the concepts of EAR and OAR are expanded to arbitrary two areas (they can be treated as complementary in the point of mathematics), then the \( h' \) will be
transferred to the Jaccard index, and Sørensen index can also be expressed by \( 2O_{aA} - O_{aA} + O_{aA} \) [23–25]. If the concepts of EAR and OAR are expanded to more areas, then zeta diversity and new beta diversity can handle this circumstance [8, 9].

Conclusions

Fundamental equations for species–area theory are applied to power law model and random replacement model of SAR. To enrich the species–area theory, several parameters are proposed, including overlap index, overlap rate, sampling rate and extinction rate, which are defined by equations in the general format, power law model and random replacement model. For power law model, both EAR and OAR have three parameters, with one more parameter of the total area than SAR does. If the total area is underestimated, the extinction rate will be overestimated. The EAR equation is a monotonically increasing function for parameter \( c \) and \( z \), and a monotonically decreasing function for parameter \( A \). Extinction rate, which has two parameters, is a monotonically increasing function for parameter \( z \), and a monotonically decreasing function for parameter \( A \). The overlap index is a monotonically increasing function for parameter \( A \),
and a monotonically decreasing function for parameter $z$, independent of parameter $c$.

**Additional file**

Additional file 1. For power law model, sensitivity analysis of parameters was conducted for EAR, extinction rate and overlap index, and the extinction rate based on different total areas was assessed for overestimate comparison.

**Abbreviations**

SAR: species-area relationship; EAR: endemics-area relationship; OAR: overlap-area relationship.

**Acknowledgements**

Thanks go to Mr. Fengqiao Liu for language polishing, and the editor and two anonymous reviewers for their insightful comments on the manuscript.

**Competing interests**

The author declares that he has no competing interests.

**Availability of data and materials**

The dataset supporting the conclusions of this article is included within its Additional file 1.

**Funding**

This work is supported by Beijing Nova Programme (Z151100003150107). The author also wants to thank the financial support for his study from Chinese Government Award for Outstanding Self-Financed Students Abroad, Texas A&M University-Kingsville, Tsinghua University and Chinese Academy of Inspection and Quarantine.

**References**

1. Tjørve E, Turner WR. The importance of samples and isolates for species-area relationships. Ecography. 2009;32:391–400.
2. WCMC. Global biodiversity: status of the earth’s living resources. London: Chapman & Hall; 1992.
3. Kinzig AP, Harte J. Implications of endemics-area relationships for estimates of species extinctions. Ecology. 2000;81:3305–11.
4. Green JL, Ostling A. Endemics-area relationships: the influence of species dominance and spatial aggregation. Ecology. 2003;84:3090–7.
5. Millennium Ecosystem Assessment (2005) Ecosystems and human well-being: biodiversity synthesis. World Resources Institute.
6. Tjørve E, Tjørve KMC. The species-area relationship, self-similarity, and the true meaning of the $z$-value. Ecology. 2008;89:3528–33.
7. Pan XB. Fundamental equations for species-area theory. Sci Rep. 2013;3:1334.
8. Pan XB. Reconstruct species-area theory using set theory. Natl Acad Sci Lett. 2015;38:173–7.
9. Hui C, McGeoch MA. Zeta diversity as a concept and metric that unifies incidence-based biodiversity patterns. Am Nat. 2014;184:884–94.
10. Pimm SL, Askins RA. Forest losses predict bird extinctions in eastern North America. Proc Natl Acad Sci USA. 1995;92:9343–7.
11. Rosenzweig ML. Species diversity in space and time. Cambridge: Cambridge University Press; 1995.
12. Harte J, Kinzig AP. On the implications of the species-area relationships for endemism, spatial turnover, and food web patterns. Okios. 1997;80:417–27.
13. He F, Hubbell SP. Species-area relationships always overestimate extinction rates from habitat loss. Nature. 2011;473:368–71.
14. Heywood VH, Mace GM, May RM, Stuart SN. Uncertainties in extinction rates. Nature. 1994;368:105.
15. Tilman D, May RM, Lehman CL, Novak MA. Habitat destruction and the extinction debt. Nature. 1994;371:365–6.
16. Mace GM, Gittleman JL, Purvis A. Preserving the tree of life. Science. 2003;300:1707–9.
17. Arthenuis O. Species and area. J Ecol. 1921;99:95–9.
18. Coleman B. On random placement and species-area relationships. Math Biosci. 1981;54:191–215.
19. Tjørve E, Kunin WE, Police C, Tjørve KMC. The species–area relationship: separating the effects of species-abundance and spatial distribution. J Ecol. 2008;96:1141–51.

20. Harte J, Kinzig AP, Green J. Self-similarity in the distribution and abundance of species. Science. 1999;284:334–6.

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

22. World Resources Institute. Millennium ecosystem assessment. Ecosystems and human well-being: biodiversity synthesis; 2005.

23. Jaccard P. The distribution of the flora in the alpine zone. New Phytol. 1912;11:37–50.

24. Tuomisto H. A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. Ecography. 2010;33:2–22.

25. Sørensen TA. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content, and its application to analyses of the vegetation on Danish commons. Kongelige Danske Videnskabernes Selskabs Biol Skr. 1948;5:1–34.