RC24040 (W0608-097) August 31, 2006
Other

IBM Research Report

Crossover in Scaling Regimes in Species-Area Relationships

Daniel E. Platt
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Crossover in Scaling Regimes in Species-Area Relationships

Daniel E. Platt
IBM T. J. Watson Research Center,
Yorktown Hgts, NY

Abstract –

The observation and elucidation of power-law scaling of Species-area relationships has recently emerged as an area of significant interest. This study takes guidance from some prior studies to identify some salient features of the distribution of inter-species populations over multiple scaling regimes. Particularly, we seek to explain the transition from weak to strong scaling in the species-area relationship as the scale is expanded through ecological meso-zone to inter-biome scales. We explore a Poisson-without-replacement process as a simplified model of the type of sampling that would produce such a cross-over behavior, producing weak dependence and strong-scaling regimes, as well as effective weak-scaling exponents that span several decades of scaling factors.

Introduction

The power-law scaling of the number of species with area has been described since the early 20th century, [1] which has continued in the expression of recent interest. [2][3] Indeed, power-law scaling is ubiquitous in ecological systems, [4] and habitat. [5] It is of continuing interest to catalogue and understand the reason for this ubiquity of power-law scaling behavior.

Of these, the Species-Area Relationship (SAR) is one of the more challenging to understand. One recent study to attempt to understand why this scaling form should be so common and robust [3] demonstrates this difficulty. The authors identify the importance of the clustering of species in promoting the emergence of a scaling law, yet leave to a supplementary section the actual arguments leading to this conclusion. The argument depends on a scaling of their "proximity function" $F^s(R) = F\left(\frac{R - \langle R \rangle^s}{\sigma_s}\right)$ for each species s. Defining $D(r)\Delta r = \sum_s \Theta\left(\frac{\Delta r}{2} - \left|\langle R \rangle^s - r\right|\right) = \sum_s \delta\left(\langle R \rangle^s - r\right)\Delta r$, and the correspondence $\sigma\left(\langle R \rangle^s\right) = \sigma_s$ given the premise that the individuals of a species tended to cluster, the number of species inside radius R was shown to be $S = \sum_s F^s = \int_0^{R_m} dr D(r) F\left(\frac{R - r}{\sigma(r)}\right)$. This can be shown to scale when the width satisfies $\sigma(br) = b\sigma(r)$, and $D(r) \propto r^\gamma$ out to a fixed cutoff $r = R_m = \max_s \langle R \rangle^s$ that does not scale. The authors then consider field data, and find that the assumption of scaling for $\sigma(r)$ does not describe the data as well as $\sigma(br) = \sigma(r) + k \ln b$, for which the authors could not construct a simple scaling form leaving the argument of F invariant. Further, they report their derived scaling form was not observed to hold in application to the field data, which challenges the premise that the clustering of species' individuals allows the separability of contributions F^s from each species s to S.

It is clear that the issue of scaling in species number with area demonstrates some complex features, where the challenges begin at the level of formulation. It is worth exploring the issues involved with construction of a scaling form for F, as well as the assumptions upon which the derivation was constructed. Particularly, there are the questions of segregation ("The importance of clustering,") and the question of weak scaling at smaller scales, vs. the stronger scaling observed at larger scales. This paper explores the issues involved with construction of a consistent scaling form for the proximity function, as well as a possible mechanism for the crossover from micro-ecosystem scales to the scaling observed when crossing biome boundaries.

Analysis

It is possible to explore a functional form $\frac{h(r,R)}{\sigma(r)}$ for the argument of the proximity function F to remain constant under scaling. This holds when $\frac{h(br,bR)}{\sigma(br)}=\frac{h(r,R)}{\sigma(r)}$. The functional form for h may be obtained by substituting $b=1/R$, yielding $h(r,R)=\frac{h\left(\frac{r}{R},1\right)}{\sigma\left(\frac{r}{R}\right)}\sigma(r)$. Ultimately, this implies that the most general scaling form for the argument of F is simply $\frac{h(r,R)}{\sigma(r)}=\frac{h\left(\frac{r}{R},1\right)}{\sigma\left(\frac{r}{R}\right)}=h_0\left(\frac{r}{R}\right)$.

This imposes a restriction on the form of σ such that $h(r,R)=r-R$. This form is necessary if σ is designed to represent an rms variation of r. But this implies $r-R=\frac{\frac{r}{R}-1}{\sigma\left(\frac{r}{R}\right)}\sigma(r)=\frac{r-R}{R\sigma\left(\frac{r}{R}\right)}\sigma(r)$, so that $\sigma(r)=R\sigma\left(\frac{r}{R}\right)$. Substituting $R=r$, it follows that $\sigma(r)=r\cdot\sigma(1)$. Therefore, it is not possible to construct a scaling form for h and σ that would satisfy the observed $\sigma(r)=\sigma(1)+k\ln r$ form (which may be obtained from the scaling form $\sigma(br)=\sigma(r)+k\ln b$ by setting $b=1/r$). The proposed *ansatzen* for the scaling of the proximity function are extremely restrictive. It is possible to construct scaling forms that are far more general. [6][7]

If species clustered, this form might explain such a scaling behavior. Indeed, even if such clustering, or segregation between species' territories is not observed, this may be important in understanding human culturally enforced geographic variations, as well as scaling in numbers of species in geographic isolation. However, there is a significant literature that seeks to explain [8-13] species enrichment where numbers of species are observed with overlapping ranges within any particular biome that has been observed among a wide variety of biomes [14-19]. The whole field of species interactions and modeling, with its long history, seeks to describe and explain the interactions between species that occupy the same ecological zones.[20-22]

On the other hand, the gross features of the scaling behavior may simply be due to considerations of sampling. Those features in question are as follows. The range of exponents z in the species-area relationship $S \propto A^z$ has been reported to vary from nearly zero for microbes in controlled experiments, through ranges of 0.1-0.4 over smaller scales, among birds and plants, and on islands, with significantly larger values up to nearly 1 over intercontinental, or inter-biome, scales. [23-24] Within a biome, the species occupation is fairly uniform, with only minor regional variations. This leads to the observation that there appears to be two regimes of scaling: weak scaling (or simply weak dependence), observed over shorter scales, and strong scaling observed at larger inter-biome scales.

Consider the standard Poisson process is modified to exclude "replacement" in the sampling of species. In this case, the sampling of an area a, where the average area in which an animal is expected to be found is A, and while the number of animals that could be found is expected to increase with area without bound following a normal Poisson process, each animal is modeled as being one of a number N of species, each species containing an equal fraction $1/N$ of animals. As each example of a species is encountered, the number n of species identified is increased, and the chances of encountering a new species in the next animal is reduced. Following the traditional derivation of the probability distribution governing Poisson processes, the probability $P_n(a+\Delta a)$ of having found n species in area $a+\Delta a$ is equal to the probability of having found n species in area a times the probability of finding no new species in area Δa plus the probability of having found $n-1$ species in area a times the probability of finding one more species in area Δa. The probability of finding a new species is the probability of finding an animal, $\Delta a / A$ times the probability that the animal is a new species, $1 - n/N$. This leads to the condition

$$P_n(a + \Delta a) = \left(1 - \frac{\Delta a}{A}\frac{N-n}{N}\right)P_n(a) + \frac{\Delta a}{A}\frac{N-n+1}{N}P_{n-1}(a) + O\left(\Delta a^2\right),$$

subject to the traditional initial conditions

$$P_n(0) = \begin{cases} 1 & \text{where } n = 0 \\ 0 & \text{where } n > 0 \end{cases}$$

$$P_n(a) = 0 \quad \text{where } n < 0$$

This can be reduced to a system of differential equations

$$\frac{dP_n(a)}{da} = -\frac{N-n}{NA}P_n(a) + \frac{N-n+1}{NA}P_{n-1}(a),$$

which may be shown to satisfy the conservation of total probability $\frac{d}{da}\left(\sum_n P_n(a)\right) = 0$. The total will continue to be the same as those of the initial conditions regardless of the amount of area sampled. The solution to these equations is

$$P_n(a) = \binom{N}{n}\left(e^{\frac{a}{NA}} - 1\right)^n e^{-\frac{a}{A}},$$

with the expected number of species having been observed in an area a being

$$\langle n\rangle(a) = N \cdot \left(1 - e^{-\frac{a}{NA}}\right).$$

If N is very large, this reduces to the standard Poisson distribution results. The expected scaling "exponent" is then

$$z = \frac{d\ln\langle n\rangle}{d\ln a} = \frac{\frac{a}{NA}}{e^{\frac{a}{NA}} - 1}.$$

This shows two regimes. When $a << NA$, the standard Poisson distribution results apply, and $z \approx 1$. When $a >> NA$, the total number of species are expected to have been found, and $z \approx 0$. The transition is displayed in figure 1.

An extension to uneven proportions σ_i for each species i such that $\sum_i \sigma_i = 1$ is a little more complicated. The probability of detecting any number of species i is

$$P_{x_i} = (1 - x_i)e^{-\frac{\sigma_i a}{A}} + x_i\left(1 - e^{-\frac{\sigma_i a}{A}}\right)$$

where $x_i = 0$ if the species is not detected, and $x_i = 1$ if it was detected. Consider the total set S of vectors $\vec{x}$ composed of all 2^N combinations of the x_i's 0 and 1 values. This set may be partitioned into subsets S_n, each with $\begin{pmatrix} N \\ n \end{pmatrix}$ components comprised of those vectors such that $\sum_i (\vec{x})_i = 1$. Then

$$P_n(a) = \sum_{\vec{x} \in S_n} \prod_i \left[(1 - (\vec{x})_i)e^{-\frac{\sigma_i a}{A}} + (\vec{x})_i\left(1 - e^{-\frac{\sigma_i a}{A}}\right) \right].$$

When all the proportions are equal, so that $\sigma_i = \frac{1}{N}$ for all species, there are n terms contributing $\left(1 - e^{-\frac{\sigma_i a}{A}}\right)$ to the product, and $N - n$ terms contributing $e^{-\frac{\sigma_i a}{A}}$, and $\begin{pmatrix} N \\ n \end{pmatrix}$ of these for all the members of S_n, so the expression reducees to the form for the simple Poisson-without-replacement. The expectation value $\langle n \rangle$ is simply

$$\langle n \rangle = N - \sum_i e^{-\frac{\sigma_i a}{A}},$$

which also reduces to the simple Poisson-without-replacement form in the equal σ_i's case. The expression for the effective scaling exponent will then be

$$z = \frac{d \ln\langle n \rangle}{d \ln a} = \frac{\sum_i \frac{\sigma_i a}{A} e^{-\frac{\sigma_i a}{A}}}{\sum_i \left(1 - e^{-\frac{\sigma_i a}{A}}\right)}.$$

In the case of nearly equal σ_i's, saturation occurs when $a \approx NA$, roughly the scale needed to detect N individual animals, which can happen in a fairly small area compared to the size of a biome. The transition is fairly rapid from a large z down to a very weak dependence. Even so, as is seen in figure 1, this transition can require more than a decade to occur. The scale over which z stays in a range of 0.2 or 0.3 is much smaller, however.

The effect of including a range of σ_i's that spans several decades does increase the region of cross-over, and can produce a region of nearly constant scaling exponents z spanning several decades (Figure 2). Once the sampling has become saturated, it will continue to be saturated until the length scale starts to sample biomes, at which time, the numbers of species being accumulated scales as the number of biomes, and the sampling is back to the $a << NA$ regime.

Conclusions

Given the ubiquity of power-law scaling behavior in biological systems demands consideration of the possibility of such scaling behavior in the species-area relationship. This type of question is more than significant given the current interest in quantifying biodiversity and its importance for ecosystem stability. Yet it has been challenging to identify which features are most important in characterizing scaling in the SAR. [3][26-28] It is possible that several of those features identified in the literature play a role.

One such model [3] proposes a particular character of species clustering to explain how such scaling arises simply from geometric considerations. We have shown that the scaling issues detailed in the supplementary material are overly restrictive in form to accommodate observed variations, and that a simple repair is not possible given the proposed scaling form. More general scaling forms might produce consistent results. More to the point, the question of clustering and segregation in scale sizes smaller than a biome (this could be taken as a definition of biome) runs counter to a significant volume of literature. Yet, it is the nature of biomes that species segregation by biome is expected, and scaling as described by species clustering could apply there. There are also domains where such clustering could lead to scaling behavior, such as the geographical distribution of human cultures segregating by language and custom.

The sampling models described above are oversimplified. Yet, we have shown that a range of species densities that spans several decades can produce some regimes where approximate scaling behavior is expected. Scaling may also emerge in the relationship of the scaling [25-26] of micro- and meso-zone to the species they contain, as well as the various ways the species interact with each other [27-29] within their zones. The problem of empirically establishing scaling behavior is made more complicated by the presence of weak relationships at smaller length scales. Exponents of 0.2 or 0.3 require at least 4 decades in scale of the independent variable to see scaling of one decade in scale of the dependent variable. Scaling should be established over several decades of scale. Confounding contributions, such as logarithmic additions, may contribute to a false or apparent exponent. However, simple sampling of random mixed-species populations with densities spanning several orders of magnitude can also account for such scaling. All of these factors have been observed in the wild, and any of them may contribute to scaling behavior.

The issue of more homogenous mixing of species' ranges does suggest a sampling model that could explain some of the changes in the strength of scaling that was observed. The models presented here are based on several simplifying assumptions, yet it demonstrate crossover from a strong scaling regime to a weak dependence regime as the number of species detected within a habitat type saturates over almost all scales within that biome, yet will return to stronger scaling dependence as the length scale sampled starts to incorporate multiple biomes. Moreover, when a range of species abundances was included that spans multiple decades of scale, an effective scaling of the number of species with area was produced. While the details of such crossover behavior and the geographical distribution of each species may be required to make detailed predictions of the behavior in the transition scaling region, the model is sufficient to demonstrate the character of such a transition, produces scaling behavior in the observed numbers of species with area scaling, and explains the transition from weak scaling within regional habitats to strong scaling as habitat type boundaries are crossed, one of the more significant features previously observed in the field. [23,24]

References

1. Arrhenius, O. (1921) *J. Ecol.* **9**, 95–99.
2. Southwood, T. R. E, May, R. M, and Sugihara, G. (2006) “Observations on related ecological exponents,” *PNAS*, **103**, 6931-33.
3. Martin, H. G. and Goldenfeld, N. (2006) “On the origin and robustness of power-law species–area relationships in ecology,” *PNAS*, **103**, 10310-15.
4. Marquet, P. A, Quiñones, R. A, Abades, S, Labra, F, Tognelli, M, Arim, M, and Rivadeneira, M. (2005), “Scaling and power-laws in ecological systems,” *J. Exp. Biol.* **208**, 1749-1769.
5. Lathrop, Jr. R. G and Peterson, D. L. (1992) “Identifying structural self -similarity in mountainous landscapes,” *Landscape Ecology*, **6**, 233-38.
6. Platt, D. E. and Family, F. (1993) “Consistent scaling of multifractal measures: Multifractal spatial correlations” *Phys. Rev. E* **47**, 2281–2288.
7. Platt, D. E. and Family, F. (1993). “Scaling of the Point-Point Correlation Function of DLA,” *Fractals* **1**, 229-237.
8. Colwell, R. K., and D. C. Lees. (2000). “The mid-domain effect: geometric constraints on the geography of species richness.” *Trends in Ecology and Evolution* **15**:70-76.
9. Pimm, S. L., and J. H. Brown. (2004). “Domains of diversity.” *Science* **304**:831-833.
10. Colwell, R. K., C. Rahbek, and N. Gotelli. (2004). “The mid-domain effect and species richness patterns: what have we learned so far?” *American Naturalist* **163**:E1-E23.
11. Zapata, F. A., K. J. Gaston, and S. L. Chown. (2005). “The mid-domain effect revisited.” *American Naturalist* **166**:E144-E148.
12. Godfray, H. C., and J. H. Lawton. (2001). “Scale and species numbers.” *Trends in Ecology & Evolution* **16**:400-404.

13. Grytnes, J. A. (2003). “Ecological interpretations of the mid-domain effect.” *Ecology Letters* **6**:883-888.
14. Bellwood, D. R., T. P. Hughes, S. R. Connolly, and J. Tanner. (2005). “Environmental and geometric constraints on Indo-Pacific coral reef biodiversity.” *Ecology Letters* **8**:643-651.
15. Brayard, A., G. Escarguel, and H. Bucher. (2005). “Latitudinal gradient of taxonomic richness: combined outcome of temperature and geographic mid-domains effects?” *Journal of Zoological Systematics and Evolutionary Research* **43**:178–188.
16. Cardelus, C., R. K. Colwell, and J. E. Watkins, Jr. (2006). “Vascular epiphyte distribution patterns: explaining the mid-elevation richness peak.” *J. Ecology* **94**:144-156.
17. Carpenter, C. (2005). “The environmental control of plant species density on a Himalayan elevation gradient.” *Journal of Biogeography* **32**:999-1018.
18. Connolly, S. R., and D. R. Bellwood. (2003). “Indo-Pacific biodiversity of coral reefs: deviations from a mid-domain model.” *Ecology* **84**:2178–2190.
19. Dunn, R. R., R. K. Colwell, and C. Nilsson. (2006). “The river domain: why are there more species halfway up the river?” *Ecography* **29**:251-259.
20. Lotka, AJ, (1925). *Elements of Physical Biology*. Williams and Wilkins, Baltimore.
21. Volterra, V, (1931). *Animal Ecology*. Chapman, New York and London.
22. Tilman, D, (1982) Resource Competition and Community Structure, Princeton University Press, Princeton, NJ.
23. Rosenzweig, M. (1995) *Species Diversity in Space and Time* (Cambridge Univ. Press, Cambridge, U.K.).
24. Williamson, M. (1988) in *Analytical Biogeography*, eds. Myers, A. A. & Giller, P. S. (Chapman & Hall, New York), pp. 91–115.
25. Lathrop Jr, R. G. and Peterson D. L, (1992) “Identifying structural self -similarity in mountainous landscapes,” *Landscape Ecology* vol. **6** no. **4**, 233-238.
26. Takayasu, H. and Inaoka, H. (1992), “New Type of Self-Organized Criticality in a model of Erosion,” *Phys. Rev. Lett.* **68**, 966-969.

27. Maurer, B. A. and Taper, M. L. (2002) “Connecting geographical distributions with population processes,” *Ecology Letters* **5**, 223-231.
28. Pasqual, M, Roy, M, Guichard, F, and Flierl, G. (2002). “Cluster size distributions: signatures of self-organization in spatial ecologies,” *Phil. Trans. R. Soc. Lond. B* **357**, 657–666.
29. Brown, J, Gillooley J. F, Allen A. P, Savage, V. M, and West, G. B. (2004). “Toward a Metabolic Theory of Ecology,” *Ecology*, **85(7)**, 1771–1789.

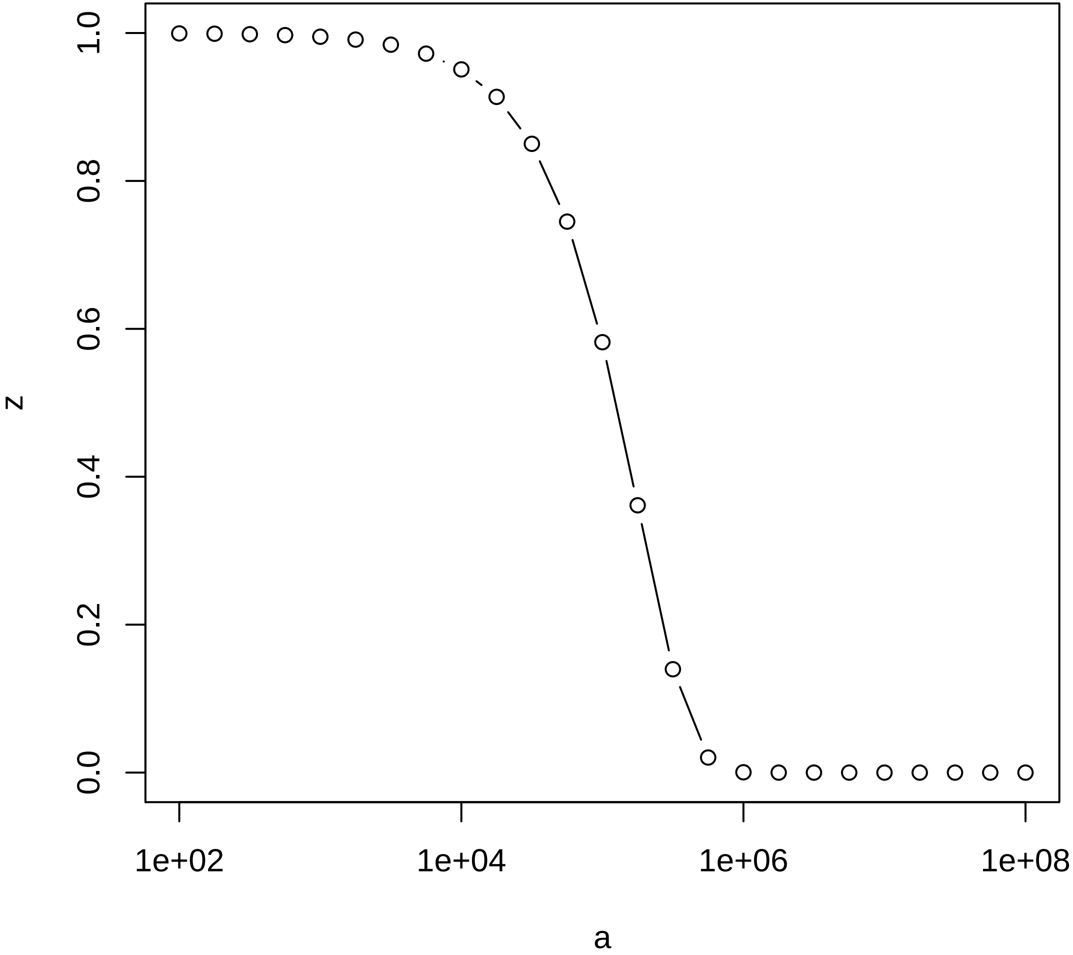
Figures

Figure 1. The crossover behavior of the scaling exponent $z = \frac{d\ln\langle n\rangle}{d\ln a} = \frac{\frac{a}{NA}}{e^{\frac{a}{NA}} - 1}$ for sampling finite numbers of species within a habitat zone. $N = 100$ and $A = 1000$.

Figure 2. Crossover with a distribution of species' member densities σ_i yielding an effective scaling exponent

$$z = \frac{d\ln\langle n\rangle}{d\ln a} = \frac{\sum_i \frac{\sigma_i a}{A} e^{-\frac{\sigma_i a}{A}}}{\sum_i \left(1 - e^{-\frac{\sigma_i a}{A}}\right)}.$$

figure1



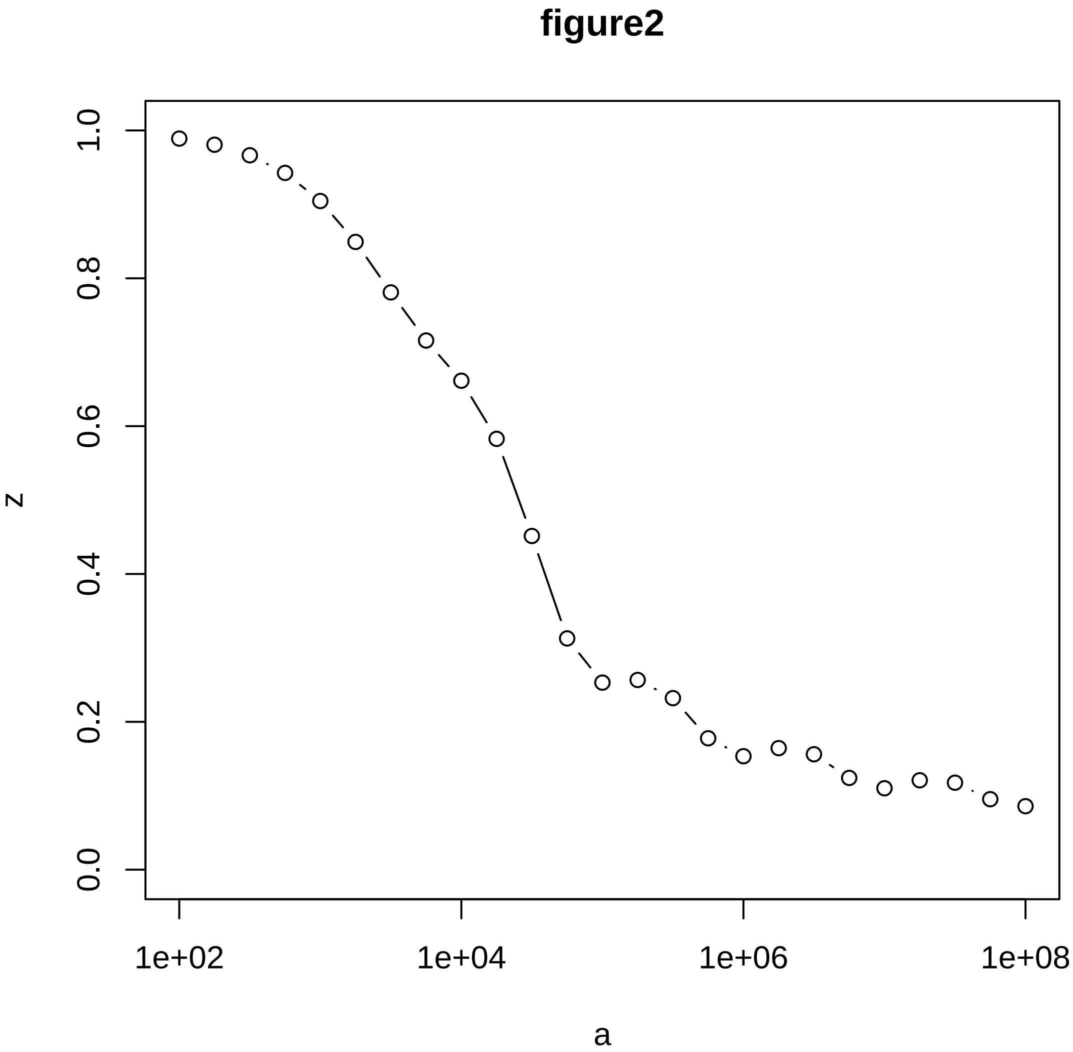


figure2
z
1.0
0.8
0.6
0.4
0.2
0.0
1e+02
1e+04
1e+06
1e+08
a