Supplementary Material

Appendix 1 - Model description.

Log-normal model

The log-normal model (LNM), first proposed by Preston (1948) is based on the central limit theorem. Because species diversity is affected by simultaneous influences of many biotic and abiotic factors, the random variation of these factors will cause species abundances to be a normally distributed (May 975). The logarithmic form (LNM) is:

$$S_{(R)} = S_0 e^{-(\alpha^2 R^2)} \tag{1}$$

where $S_{(R)}$ is the number of species in the Rth octave to the left and to the right of the symmetrical curve, S_0 is the number of species within modal abundance octave and $1/\alpha$ is the width of the distribution (Kevan & Belaoussoff 1997).

Broken stick model

The broken stick model (BSM) was proposed by MacArthur (1957). The environment is compared with a stick of unit length on which S-1 points are thrown at random. The stick is broken at these points and the lengths of the S resulting segments are proportional to the abundances of S species. Under the assumption that all species in the community sharing close taxonomic status and with similar competitive capacity, the expected abundance of the ith rarest species among S species and N individual trees is:

$$N_i = \frac{N}{S} \sum_{k=i}^{S} \frac{1}{k} \tag{2}$$

where k equals to i. The model assumes that the resource allocation among competitive species follows a one-dimensional gradient (MacArthur 1957).

Zipf model

The Zipf model (ZM) was first introduced by Frontier (1985) who assumed that species occupancy is dependent on the actual environmental and physical conditions and the species present. The cost of occupancy of a pioneer species is rather lower than that of late successional species which have higher

requirements of forest ecosystem structure. This assumption reflects a successional process that late settlers are more difficult to survive than first arrivals. The abundance of the ith species is:

$$N_i = Nqi^y \tag{3}$$

where N is the number of individual trees, q the predicted relative abundance of the species with the highest frequency in the community and γ a constant representing the average probability of species occupancy, suggesting as well a priority effect (Frontier 1985).

Niche preemption model

The niche preemption model (NPM) was proposed by Motomura (1932). When species enter an unsaturated habitat at the same time intervals and occupy a part of the niche, the SADs are more likely to form a NPM, which is particularly common for communities affected by only one or few factors. This model assumes that α percent of the total niches is occupied by the most frequent species, the second one occupies α percent of the remaining niches, i.e., $\alpha(1-\alpha)$; by this analogy, the niche occupied by ith species is $\alpha(1-\alpha)^{i-1}$. The species abundance is ranked by the amount of niches they occupy. The expected abundance for the ith species is:

$$N_i = N\alpha(1-\alpha)^{i-1} \tag{4}$$

where N is the total number of individual trees in the community.

Neutral model

The neutral theory (NM) was introduced by Hubbell (2001). The neutral theory makes the following two assumptions: (1) the total number of individual trees species of a community is constant. In a saturated community, the increase of one species will result, to some extent, in a decrease of another species, and (2) there is no difference in functional and physiological traits between species. Species under the neutral theory have identical natality, mortality, immigration and speciation rates. On the basis of these two assumptions, species abundance shows a zero-sum multinomial distribution. The probability distribution of the species follows the following equations (Etienne 2005):

$$P[D|(\theta,J,m)] = \frac{J!}{\prod_{i=J}^{S} n_i \prod_{j=1}^{J} \Phi_j!} \frac{\theta^S}{(\theta)_S} \cdot \sum_{A=S}^{J} \left[K(D,A) \frac{(\theta)_j}{(\theta)_A} \frac{I^A}{(I)_J} \right] \tag{5}$$

$$I = \frac{m(J-1)}{1-m} \tag{6}$$

$$A = \sum_{i=1}^{S} a_i \leq J \tag{7}$$

$$K(D,A) = \sum_{\left[a_i,\ldots,a_s | \sum_{i=1}^{S} a_i = A\right]} \prod_{i=1}^{S} \frac{\overline{S}(n_i, a_i)\overline{S}(a_{i,1})}{\overline{S}} \tag{8}$$

where S is the total number of species, n_S is the abundance of sth species and $D = (n_1, n_2, k, n_s)$ is the SADs. θ is an index represents a "fundamental diversity" number the greater the θ, the larger the number of species in a community. j is the total number of individual trees. m the immigration rate per unit area. Φ the number of species with j individual trees. i the number of individual trees which migrated to the local community. A is the number of individual trees of the parent generation and a_i the number of individual trees of the parent generation of the ith species. $K(D,A)$ is a multinomial coefficient. The expected species abundance was estimated by taking the means of 100 simulations of neutral communities using the estimated θ and m predicted by a maximum likelihood estimation method and the total number of observed individual trees j (Walker & Cyr 2007).

Reference

Etienne RS (2005). A new sampling formula for neutral biodiversity. Ecology Letters 8: 253-60.

Frontier S (1985). Diversity and structure in aquatic ecosystems. Oceanography and Marine Biology 23: 53-312.

Hubbell SP (2001). The unified neutral theory of biodiversity and biogeography. Princeton University Press.

Kevan PG, Belaoussoff S (1997). Log-normality of biodiversity and abundance in diagnosis and measuring of ecosystemic health: pesticide stress on pollinators on blueberry heaths. Journal of Applied Ecology 34:80-80.

MacArthur RH (1957). On the relative abundance of bird species. Proceedings of the National Academy of Sciences of the United States of America 43: 93-295.

May RM (1975). Patterns of species abundance and diversity. Ecology & Evolution of Communities: 81-120.

Motomura I (1932). On the statistical treatment of communities. Zoology Management 44: 379-383.

Preston FW (1948). The commonness, and rarity, of species. Ecology 29: 254-283

Walker SC, Cyr H (2007). Testing the standard neutral model of biodiversity in lake communities. Oikos 116:143-155.

Appendix 2 - R script.

```
####################### Dividing the plot into small sampling scales####################

# The raw data is organized according to the data format of CTFS

scale.data.matrix= function(data,xsize,ysize,gridsize){#xsize and ysize are the side lengths of the
plot, girdsize is the side length of sampling unit

rowcol.to.index=function (rowno, colno, gridsize, plotdim = c(xsize, ysize))
    {
  badrc = (rowno <= 0 | colno <= 0 | rowno > plotdim[2]/gridsize |
      colno > plotdim[1]/gridsize)
  rowno = rowno - 1
  colno = colno - 1
  maxrow = floor(plotdim[2]/gridsize)
  index = colno * maxrow + rowno + 1
  if (length(badrc[badrc > 0]))
      index[badrc] = NA
  return(index)
    }

gxgy.to.index=function (gx, gy, gridsize, plotdim = c(xsize, ysize))
    {
  badgxgy = (gx < 0 | gy < 0 | gx >= plotdim[1] | gy >= plotdim[2] |
      is.na(gx) | is.na(gy))
  colno = 1 + floor(gx/gridsize)
  rowno = 1 + floor(gy/gridsize)
  if (length(badgxgy[badgxgy > 0]))
      colno[badgxgy] = rowno[badgxgy] = NA
  return(rowcol.to.index(rowno, colno, gridsize, plotdim))
    }

  no=gxgy.to.index(data$x,data$y,plotdim=c(xsize,ysize),gridsize=gridsize)
  data1=na.omit(cbind(data,no))
  qrat=sort(unique(no))
  n=length(qrat)
  sp=unique(data$sp)
  m=length(sp)
  abu.qrat=matrix(0,n,m)
  colnames(abu.qrat)=sp
  rownames(abu.qrat)=qrat
  nobrance.data=data1
  for(i in 1:n)
      {
      data2=nobrance.data[nobrance.data$no==qrat[i],]
      for(j in 1:m)
          {
          abu.qrat[i,j]=length(data2[data2$sp==sp[j],]$sp)}
          }
      total.abu=apply(abu.qrat,2,sum)
      abu.qrat.t=t(abu.qrat)
      abu.qrat.t2=data.frame(abu.qrat.t,total.abu)
      abu.qrat.r=abu.qrat.t2[order(-abu.qrat.t2$total.abu),]
      abu.qrat.f=t(abu.qrat.r)
      result=abu.qrat.f[-(n+1),]
      rownames(result)=qrat
    return(result)
        }

############################# Calculating the observed SAD, taking analysis based on data of
Jioahe at 10m × 10m scale as an example ########################
data.matrix10=scale.data.matrix(data, 800, 500, 10)

obsSAD.fn=function(data){
  abundorder=matrix(NA,nrow=dim(data)[1],ncol = dim(data)[2])
  for(i in 1:dim(data)[1]){
  abundorder[i,]=sort(data[i,],decreasing =TRUE)
  }
```

```
  obsSAD=apply(abundorder,2,mean)
  return(obsSAD)
}
obsSAD10=obsSAD.fn(data.matrix10)

####################Calculating the expected SAD by neutral model ################
library(untb)
theta=optimal.theta(obsSAD10)
m=optimal.prob(obsSAD10)

SADdata<-as.count(obsSAD10)

count=600  # 600 times simulations
data.result=matrix(0,count,43)
for(i in 1:count){
  line=untb(SADdata,m)
  len=length(line)
  for(j in 1:len){
    data.result[i,j]=line[[j]]
  }
}
neutalexp10=apply(data.result,2,mean)

#Calculating the expected SAD by other modes
library(vegan)
mod<-radfit(obsSAD10)
a<-fitted(mod)
BSMexp10=a[,1]
NPMexp10=a[,2]
LNMexp10=a[,3]
ZMexp10=a[,4]

###########End of the code###############
```

Tab. S1 - The abundance of each species in Changbaishan.

Species	Abundance	Species	Abundance
Corylus sieboldiana	27221	*Viburnum opulus*	520
Acer pseudosieboldianum	18366	*Maackia amurensis*	399
Acer barbinerve	16817	*Ulmus davidiana*	304
Acer tschonoskii	12477	*Malus baccata*	191
Acer tegmentosum	5470	*Rhamnus davurica*	111
Acer pictum	4175	*Euonymus verrucosus*	105
Pinus koraiensis	3047	*Lonicera japonica*	52
Eleutherococcus senticosus	2594	*Sorbus aucuparia*	51
Tilia amurensis	2502	*Euonymus phellomanus*	32
Syringa reticulata	2102	*Sorbaria sorbifolia*	31
Philadelphus schrenkii	1856	*Populus ussuriensis*	29
Acer ukurunduense	1845	*Populus koreana*	17
Prunus maximowiczii	1774	*Phellodendron amurense*	9
Quercus mongolica	1402	*Betula platyphylla*	8
Acer mandshuricum	1348	*Sambucus williamsii*	7
Sorbus alnifolia	942	*Aralia elata*	5
Prunus padus	924	*Ribes mandshuricum*	5
Ulmus laciniata	753	*Populus davidiana*	2
Fraxinus mandshurica	742	*Rosa davurica*	2
Betula costata	581	*Tilia mandshurica*	2
Euonymus alatus	525	*Deutzia parviflora*	1

Tab. S2 - The abundance of each species in Jiaohe.

Species	Abundance	Species	Abundance
Carpinus cordata	6576	*Ulmus davidiana*	62
Acer barbinerve	6001	*Betula dahurica*	60
Acer mandshuricum	5494	*Sorbus aucuparia*	58
Corylus sieboldiana	4525	*Rhamnus ussuriensis*	51
Acer pictum	3733	*Euonymus alatus*	37
Ulmus laciniata	2637	*Populus koreana*	36
Syringa reticulata	2449	*Eleutherococcus senticosus*	26
Tilia amurensis	1825	*Betula platyphylla*	24
Acer ukurunduense	1525	*Lonicera japonica*	23
Acer tegmentosum	1458	*Aralia elata*	20
Pinus koraiensis	1303	*Philadelphus schrenkii*	16
Sorbus alnifolia	998	*Rhamnus davurica*	12
Juglans mandshurica	784	*Fraxinus chinensis*	12
Betula costata	755	*Populus davidiana*	10
Phellodendron amurense	743	*Lonicera chrysantha*	9
Abies holophylla	656	*Malus baccata*	7
Euonymus phellomanus	629	*Maackia amurensis*	5
Prunus padus	588	*Rhamnus schneideri*	4
Fraxinus mandshurica	385	*Deutzia parviflora*	3
Euonymus verrucosus	320	*Abies nephrolepis*	3
Prunus maximowiczii	223	*Sambucus williamsii*	2
Ulmus macrocarpa	186	*Salix koreensis*	2
Quercus mongolica	151	*Lonicera ruprechtiana*	2
Tilia mandshurica	99	*Acer triflorum*	1

Tab. S3 - The abundance of each species in Liangshui.

Species	Abundance	Species	Abundance
Abies nephrolepis	5393	*Prunus padus*	81
Tilia amurensis	4233	*Picea jezoensis*	63
Acer pictum	3812	*Rhamnus davurica*	50
Pinus koraiensis	2606	*Quercus mongolica*	37
Acer tegmentosum	2039	*Corylus sieboldiana*	33
Acer ukurunduense	2012	*Phellodendron amurense*	14
Betula costata	1694	*Populus cathayana*	9
Alnus hirsuta	1477	*Populus ussuriensis*	8
Fraxinus mandshurica	1112	*Eleutherococcus senticosus*	7
Betula platyphylla	709	*Salix matsudana*	5
Picea asperata	707	*Abies holophylla*	4
Picea koraiensis	577	*Lonicera japonica*	3
Syringa reticulata	494	*Populus nigra*	3
Ulmus davidiana	318	*Rhamnus diamantiaca*	3
Populus davidiana	275	*Maackia amurensis*	2
Amygdalus davidiana	146	*Juglans mandshurica*	1