

# R 软件在森林生物多样性动态监测中的应用

陈伟<sup>1,2</sup>, 陈升晖<sup>3</sup>, 兰国玉<sup>2,4\*</sup>

(1. 海南大学教务处, 海南 海口 570228; 2. 中国热带农业科学院 橡胶研究所, 海南 儋州 571737;  
3. 海南职业技术学院, 海南 海口 570216; 4. 农业部儋州热带作物科学观测实验站, 海南 儋州 571737)

**摘要:** 森林生物多样性动态监测样地的数据巨大且繁杂, 迫切需要一种功能强大且灵活的软件来处理, R 软件具有免费、开源、强大的统计分析及其完美的作图功能。基于西双版纳 20 hm<sup>2</sup> 热带季节雨林动态监测样地的野外调查数据, 采用 R 软件处理。着重介绍了 R 软件在地形图的绘制、种面积曲线和种序列曲线的绘制、以及物种空间分布图等方面的应用, 希望能起到抛砖引玉的作用。

**关键词:** R 软件; 生物多样性; 动态监测; 应用

**中图分类号:** S718.54

**文献标志码:** A

**文章编号:** 1001-7461(2013)04-0180-04

## Application of R Software on Dynamic Monitoring of Forest Biodiversity

CHEN Wei<sup>1,2</sup>, CHEN Sheng-hui<sup>3</sup>, LAN Guo-yu<sup>2,4\*</sup>

(1. Office of Academic Affairs, Hainan University, Haikou, Hainan 571737, China;  
2. Rubber Research Institute, Chinese Academy of Tropical Agricultural Sciences, Danzhou, Hainan 571737, China;  
3. Hainan College of Vocation and Technique, Haikou, Hainan 570216, China; 4. Danzhou Investigation  
& Experiment Station of Tropical Crops, Ministry of Agriculture, Danzhou, Hainan 571737, China)

**Abstract:** Basically, data on forest biodiversity monitoring plots are huge and complicated. There is an urgent need for a powerful and flexible software to handle the data. R software can meet this demand for its free, open source, powerful statistical analysis and perfect mapping function. Based on processing the data collected from the forest biodiversity dynamic monitoring plots in tropical seasonal rainforest in Xishuangbanna with an area of 20 hm<sup>2</sup>, this paper briefly introduced some application examples of R software in drawing topographic map, species area curve, species sequence curve, and species spatial distribution map.

**Key words:** R software; biodiversity; dynamic monitoring; application

R 软件由于其免费、开源、强大的统计分析及其完美的作图功能已得到越来越多人的关注与应用<sup>[1-4]</sup>。R 的统计分析过程常常被分解成一系列步骤, 并且所有的中间结果都被保存在对象中, 以便使用 R 里面的函数做进一步的分析。虽然 SAS、SPSS 和 Minitab 也提供了丰富的屏幕输出内容, 但其中间结果很难在后续过程中分析使用。所以国外绝大多数的统计和计量经济学研究人员以及实业界人士都选择 R 软件。生态学家在 10 a 多前就在森林生物多样性监测研究方面开始应用 R 软件, 而国内学

者对于 R 软件了解和应用并不是很多。鉴于此, 本文介绍了 R 软件在森林生物多样性监测中的几个应用范例, 希望能起到抛砖引玉的作用。

## 1 森林生物多样性动态监测概况

### 1.1 国外森林生物多样性动态监测

森林的动态监测研究最早是在 1975 年建立在哥斯达黎加的被火侵入或是荒废的 13 hm<sup>2</sup> 的干旱森林<sup>[5-6]</sup>。但真正意义的热带森林动态监测样地是随后 S. Hubell 于 1980 年在巴拿马的 Barro Colo-

收稿日期: 2012-11-08 修回日期: 2012-12-15

基金项目: 中国热带农业科学院橡胶研究所基本科研业务费项目“海南岛橡胶林固定样地建立及生物多样性动态监测研究”(160022013013); 海南省自然科学基金“海南岛橡胶林生物多样性动态监测研究(313060)。

作者简介: 陈伟, 女, 硕士, 讲师, 研究方向: 计算机应用。

\* 通信作者: 兰国玉, 男, 副研究员, 研究方向: 橡胶林生物多样性。E-mail: languoy@gmail.com

rado Island (BCI)地区建立了 50 hm<sup>2</sup> 的动态监测样地<sup>[7]</sup>。P. S. Ashton<sup>[8]</sup>认为在东南亚的热带森林物种物种更为丰富,含有更多的同属植物,其多样性的维持机制与巴拿马的是不同的,于是 1983 年,马来西亚森林研究所在马来西亚半岛的 Pasoh 用同样的技术规范建立了第 2 个 50 hm<sup>2</sup> 动态监测样地。第 3 个 50 hm<sup>2</sup> 动态监测样地是印度科学研究院在印度南部干旱落叶林中建立的。随后各国科学家根据研究需要,应用同样的技术规范在各自的研究区域建立了 16~52 hm<sup>2</sup> 不同尺度的动态监测样地。截止目前,全球共有建立了 47 个动态监测样地,共监测了全球已知热带树种的 14% 的物种,约 8 500 个物种的 4 500 000 植株(www.ctfs.si.edu)。基于这些动态监测样地的野外调查数据,许多生态学家致力于发展和验证森林物种多样性维持的诸多机制,在中性理论<sup>[7]</sup>、生态位分化理论<sup>[8-9]</sup>、Janzen-Connell 假说<sup>[10-12]</sup>、更新限制假说<sup>[13-14]</sup> 等方面发表了大量的科学论文。

## 1.2 国内森林生物多样性动态监测

2004 年,加拿大阿尔波特大学何芳良教授和台湾东海大学孙义方教授建议在我国境内的典型纬度带上建设大尺度的森林动态监测样地。为揭示中国森林生物多样性的形成和维持机制、以及研究森林生物多样性对全球变化的响应,中国科学院生物多样性委员会于 2004 年组织并启动了“中国森林生物多样性监测网络”建设<sup>[15]</sup>,目前,在温带、亚热带和热带分别建立了 3 个、5 个、2 个 5~25 hm<sup>2</sup> 的森林群落动态监测样地。仅仅几年,其研究成果便在 Ecology Letters, Ecology, Journal of Ecology, American naturalist, PLoS ONE, Okios 等国际著名期刊发表。中国森林生物多样性监测网络是继美国史密森热带森林中心热带雨林生物多样性监测网络以后又一大型生物多样性监测的地区性网络体系(<http://www.cfbiodiv.org/wlgs.asp>)。中国森林生物多样性网络的建立,对于验证和发展森林树种多样性维持机制的诸多理论和假说等具有非常重要的科学意义。

## 2 几个简单的 R 函数

R 软件中含有很多函数可以实现数据的读取、浏览、筛选、排序、绘图等,功能较为强大。为了便于读者能够快速了解 R 软件在森林生物多样性应用范例,在介绍范例前,先介绍 3 个重要的函数。

### 2.1 读取数据函数

将数据读入 R 软件通常用 read.table() 函数,括号内为读入数据的位置及文件名,需要注意的是

要用引号括起来。例如读取 H 盘根目录下名称为 bb0 的文件则用 read.table(“h:/bb0”)。

### 2.2 绘图函数

在 R 软件绘图常用的 plot(x,y,...)。括号内为 x 为绘制图形 x 轴的坐标值,y 为 y 轴坐标值。plot 函数有许多参数如图形类型参数 type,内有 p (点图),l(线图);图形的颜色参数 col。例如绘制 x 和 y 的散点图,要求点为红色。则输入 plot(x,y,type=p,col=“red”)

### 2.3 R 帮助

R 软件最大特点和优点便是帮助功能强大。如不了解某个函数的具体参数如何设置,则在 R 里输入“? + 函数名称”则会弹出该函数的帮助信息。例如想了解 plot 函数的用法,则输入“? plot”则可。文中应用范例中出现的一些函数也可用 R 帮助来了解其具体功能与参数设置。

## 3 应用举例

R 软件在森林生物多样性监测中的应用比较广泛,本文着重介绍以下几个方面的应用,地形图的绘制,种面积曲线的拟合,种序列曲线的绘制,物种空间分布图等方面的应用。数据主要来自西双版纳热带季节雨林动态监测样地<sup>[16]</sup>。

### 3.1 地形图的绘制

地形图的绘制在许多软件中都可以实现,如 ARCVIEW, SIGMPLOT 等软件都可以。同样 R 软件中也可以用简单的 image 命令和 persp 命令来实现。在 R 软件初始安装后,已经具有一些基本的软件包(package)。用 R 软件绘制样地地形图是森林生物多样性监测研究最简单的应用。首先是样地数据库的建立,样地数据可以直接输入在 excel 表格(保存为 csv 格式)中或 txt 文本中,其次将样地的地形数据(mapdata.csv)读入 R 软件内。然后用高级绘图命令中的 contour 来绘制样地的等值线图(等高线图)。样地地形图绘制的 R 语言如下(图 1):

```
par(fin=c(4.7,5.7))
par(col.axis=“black”,cex.axis=0.8,cex.main=0.8,cex.lab=0.8,font.lab=1,font.axis=1,font.main=1)
mapdata=read.csv(“f:/mapdata.csv”,header=F)
map=as.matrix(mapdata)
altitude=matrix(map,nrow=41,ncol=51,dimnames=NULL)
x=10*(1:nrow(altitude))-10
y=10*(1:ncol(altitude))-10
```

```
image(x,y,altitude,col=topo.colors(100),axes=
F,xlab="x 轴 x-coordinates(m)",ylab="y 轴
y-coordinates(m)")
contour(x,y,altitude,nlevels=40,add=T,col="
peru")
axis(1,at=seq(0,400,by=100))
axis(2,at=seq(0,500,by=100))
box()
```

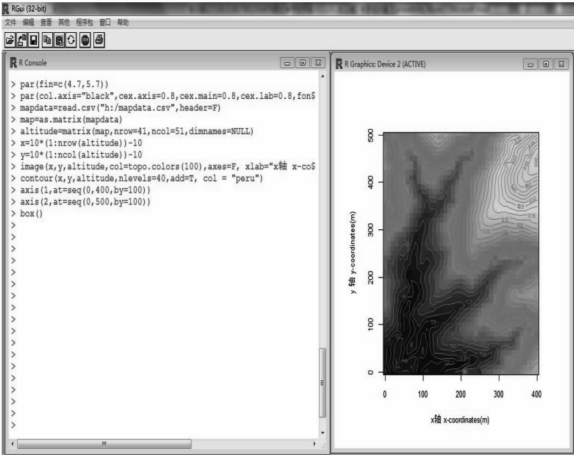


图 1 样地地形图及其 R 语言界面

Fig. 1 The surface of contour map of the dynamic plot and its R code

3.2 种-面积曲线的绘制

种-面积曲线关系是描述物种数量随空间尺度而增加的函数,也是保护生物学的一个核心理论<sup>[17-19]</sup>。传统种-面积的曲线的绘制比较麻烦,首先按照要求进行“巢式”取样,分别记录个方格面积内的物种数量,在以面积为  $x$  轴,累计物种数为  $y$  轴绘制,便得到种-面积曲线。在森林动态监测样地的研究中,应用 R 软件进行种-面积曲线的绘制则较为容易,只要写简单的程序就可以。两者的区别是,R 软件绘制的种-面积曲线更为细致,因为在 R 里面可以设定取样面积的大小,而不像传统的取样,按照正倍数增加。种-面积曲线绘制的 R 语言如下(图 2):

```
par(col.axis="black",cex.axis=1.5,cex.main=1.5,cex.lab=1.5,font=
2,mar=c(1,1,1,1))
x=seq(0,400,8)
y=seq(0,500,10)
nsize=length(x)
sp.no=numeric()
for(i in 1:nsize){
data.subset=subset(bb0,bb0$gx<=x[i]& bb0$gy<=y[i])
sp.no[i]=length(unique(data.subset$sp))
}
```

```
ha=x*y/10 000
plot(x*y/10 000,sp.no,col="dark green",type="o",xlab="面积 Area (ha)",ylab="物种数
No. of species",ylim=c(0,500),cex=1.5)
```

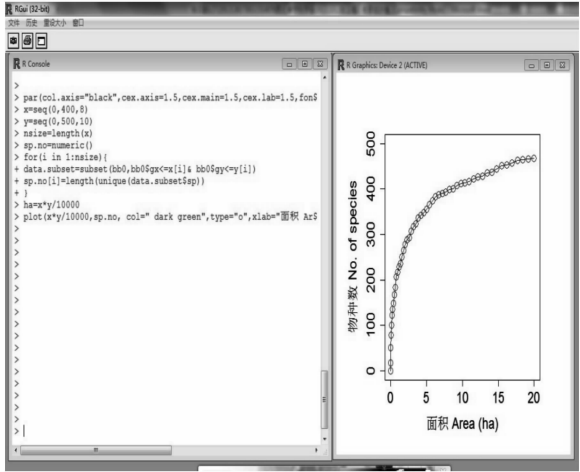


图 2 种-面积曲线图及其 R 语言界面

Fig. 2 The surface of species-area curve of the dynamic plot and it R code

3.3 种序图的生成

种序列曲线可以在 Excel 文档中生成,但过程较为麻烦,而在 R 软件中则比较容易实现。首先计算每个物种的个体数量,用 table 命令,然后将这些物种按照个体数量的多少进行排序,用 sort 命令,最后用 plot 命令绘制出样地的种序图。

样地种序列绘制的 R 语言如下(图 3):

```
par(cex.axis=1.5,cex.lab=1.5,font.lab=1,font.axis=1,cex.main=1.5)
abund=table(bb0$sp)
abundance=sort(abund,decreasing=T)
plot(log(abundance),col="red",xlab="种序
Species sequence",ylab="个体数(取对数)ln(abundance)",cex=1.2)
```

3.4 树种空间分布图的绘制

绘制某一树种不同径级在空间的分布图,主要包括以下几个步骤,第一步,绘制样地的等高线图,第二步,用低级命令函数将样地的树种的信息添加在等高线图上。请绘制样地内望天树在 3 个不同径阶( $1 < dbh < 5$ ,  $5 < dbh < 30$ ,  $dbh > 30$ )的树种分布图,3 个不同径阶的个体用不同的符号和颜色表示。绘制树种空间分布 R 语言为(图 4):

```
par(fin=c(4.7,5.7))
par(col.axis="black",cex.axis=1,cex.lab=1,font=2,mar=c(1,1,1,1))
mapdata=read.csv("h:/mapdata.csv",header=F)
map=as.matrix(mapdata)
```

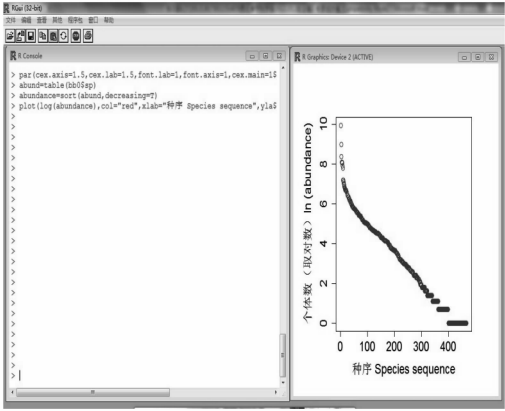


图 3 种序列曲线图及其语言界面

Fig. 3 The surface of species-sequence curve of the dynamic plot and it R code

```
altitude = matrix (map, nrow = 41, ncol = 51, dimnames=NULL)
x = 10 * (1:nrow (altitude)) - 10
y = 10 * (1:ncol (altitude)) - 10
contour (x,y,altitude,axes = F,col = "grey",nlevels = 40,lwd = 0.05,drawlabels = F)
xdim = 400
ydim = 500
axis (side = 1,pos = 0,cex = 0.8)
axis (side = 2,pos = 0,cex = 0.8)
axis (side = 3,pos = ydim,at = c (seq (0,xdim,by = 100),xdim),labels = F)
axis (side = 4,pos = xdim,at = c (seq (0,ydim,by = 100),ydim),labels = F)
SHORWA0 = subset (bb0,sp == "SHORWA")
SHORWA1 = subset (SHORWA0,dbh < 5 & dbh >= 1)
SHORWA2 = subset (SHORWA0,dbh < 18 & dbh >= 5)
SHORWA3 = subset (SHORWA0,dbh >= 43)
points (SHORWA1 $ gx, SHORWA1 $ gy, col = "dark green",pch = 3,cex = 0.4,xlim = c (0,400),ylim = c (0,500))
points (SHORWA2 $ gx, SHORWA2 $ gy, col = "blue",pch = 16,cex = 0.7)
points (SHORWA3 $ gx, SHORWA3 $ gy, col = "red",pch = 22,bg = "red",cex = 0.7)
```

4 结论

R 软件虽然是一个比较好的统计分析软件,但 R 软件对于新学者来说,上手较慢。主要是因为要在 R 软件中得到自己想要得到的结果,通常需要我们自己编写 R 程序或 R 函数,这就使得一些初学者望而

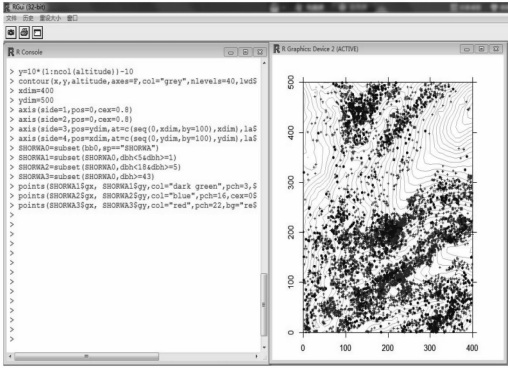


图 4 树种空间分布图及其语言界面

Fig. 4 The surface of spatial distribution map of tree of the dynamic plot and it R code

生畏。然而幸运的是从第 1 个热带森林动态监测样地建立到现在已经有近 30 a 时间了,在此期间,各国科学家为计算而开发了诸多 R 程序和软件包,如 B. L. Zimmerman<sup>[20]</sup> 等研发的 CTFS 软件包可以进行森林动态监测样地的各类计算,J. Oksanen<sup>[21]</sup> 等人开发的 Vegan 软件包也包含针对群落生态学的研究的许多程序和命令。R 软件在国外森林动态监测应用较为广泛,国内仅是少数科研所在近几年建立的动态监测样地的相关研究中有应用,且与国外相比,应用处于较低水平,尚未进行相关软件包的开发工作。另外需要说明的一点是 R 软件不仅在森林生物多样性动态监测中应用比较广泛,在整个生态学或生物统计学中,R 软件都拥有广泛的应用前景。

致谢:文中的数据主要来源于西双版纳热带季节雨林动态监测样地调查结果,在此特感谢全体在样地调查中付出辛勤劳动的工作人员及导师曹敏研究员。

参考文献:

[1] 王斌会. 多元统计分析及 R 语言建模[M]. 广州:暨南大学出版社,2010.

[2] 汤银才. R 语言与统计分析[M]. 北京:高等教育出版社,2005.

[3] 王怀亮. R 软件在系统聚类分析中的应用[J]. 合作经济与科技,2011(14):126-127.

[4] 马克平. 生物多样性研究进展Ⅸ[M]. 北京:气象出版社,2011:351-362.

[5] CONDIT R. Research in large, long-term tropical forest plot [J]. Tree, 1995, 10: 18-22.

[6] CONDIT R. Tropical forest census plots: Methods and results from Barro Colorado Island, Panama and a comparison with other plot[M]. Berlin: Springer, 1998.

[7] HUBBELL S. P. Neutral theory in community ecology and the hypothesis of functional equivalence[J]. Functional Ecology, 2005(19):166-172.