

# 应用 rioja 软件包建立有壳变形虫-环境因子转换函数

李鸿凯<sup>1,2</sup>, 李微微<sup>1</sup>, 蒲有宝<sup>1</sup>, 王从洋<sup>1</sup>, 王松梅<sup>1</sup>, 杨晓林<sup>1</sup>

(1. 东北师范大学国家环境保护湿地生态与植被修复实验室, 吉林 长春 130024; 2. 东北师范大学地理科学学院泥炭沼泽研究所, 吉林 长春 130024)

**摘要:**应用 R 语言 rioja 软件包的加权平均(Weighted Averaging, WA)和加权平均偏最小二乘(Weighted Averaging Partial Least Squares, WA-PLS)模型建立了长白山区泥炭藓泥炭地有壳变形虫与水位埋深(depth to water table, DWT)、pH 和泥炭湿度的转换函数, 为古环境定量重建奠定了基础, 也提供了 rioja 软件包应用的实例和参考。结果表明水位埋深以 WA-PLS 模型最佳(预测均方根误差 RMSEP 为 7.39 cm,  $R^2=0.74$ ); 对于 pH 和泥炭湿度, WA-PLS 第一分量和 WA.inv 都产生了最小的 RMSEP 和较高的  $R^2$  值。pH 的 RMSEP 为 0.18,  $R^2$  为 0.72。泥炭湿度的 RMSEP 为 1.95%,  $R^2$  为 0.62。如果泥炭剖面的有壳变形虫种类组成与本研究的训练样本集相同, 水位埋深、pH 和泥炭湿度可以分别以  $\pm 7.39$  cm、 $\pm 0.18$  和  $\pm 1.95\%$  的平均误差进行重建。

**关键词:** rioja 软件包; 转换函数; 有壳变形虫(*Testate amoebae*); 水位埋深; pH; 泥炭湿度

**中图分类号:** P91; P92      **文献标识码:** A      **文章编号:** 1000-0690(2013)08-1022-07

定量重建环境变化的历史是古环境研究的重要趋势。转换函数是古环境定量重建的有效方法, 它是在现代过程研究的基础上, 建立代用指标和环境因子的转换函数, 将转换函数应用于代用指标的剖面数据, 可以定量重建环境因子的变化历史。目前生物指标和环境因子转换函数模型主要有 WA、WA-PLS、ML、MAT 等。这些模型都涉及大量复杂的数据计算, 必需借助一定的软件程序来实现。C2 软件广泛应用于建立硅藻(Diatom)<sup>[1-3]</sup>、摇蚊(Chironomidae)<sup>[4,5]</sup>、介形虫(Ostracoda)<sup>[6,7]</sup>、有壳变形虫(*Testate amoebae*)<sup>[8-11]</sup>等生物指标与环境因子的转换函数建立和剖面定量重建。它功能强大, 包括了目前常用的各种转换函数模型, 集数据编辑、模型构建与检验、剖面重建、图表绘制于一体。它的试用版本虽然可以满足小样品量( $n < 75$ )的研究需求, 但样品量大时, 需购买许可, 价格不菲<sup>[12]</sup>。2012 年 C2 的开发者发布了基于 R 语言的软件包 rioja, 这为我们提供了一种新的更加经济的选择。有壳变形虫是一组陆生单细胞成壳原生动物, 作为生物指标在国外泥炭古环境研究中普遍应用<sup>[13-15]</sup>。根据各地建立的有壳变形虫与环境

因子转换函数<sup>[8-11,16-29]</sup>, 定量重建了泥炭地古水文变化<sup>[30-39]</sup>。但该指标在国内的应用相对较少<sup>[40-42]</sup>。本文尝试应用 rioja 软件包建立长白山区泥炭地有壳变形虫与水位埋深、pH、泥炭湿度的转换函数, 为古环境定量重建奠定基础, 同时也提供一个 rioja 软件包的应用实例和参考。

## 1 rioja 软件包简介

R 语言是用于统计分析、绘图的语言和操作环境。它功能强大, 内建的各种统计和数学分析工具, 优秀的制图功能, 加上数千种基于 R 语言的扩展软件包, 使得 R 几乎无所不能。更重要的是 R 所有的软件包可免费下载使用, 用户可自己根据需要编写函数命令脚本。这种自由、免费和源代码开放等特征受到用户极大的欢迎<sup>[43]</sup>。

rioja 是 C2 的开发者 Steve Juggins 博士开发的基于 R 语言的扩展软件包<sup>[44]</sup>。其功能和 C2 相当, 主要针对第四纪研究中常见的数据分析, 包括转换函数、约束聚类、剖面图绘制等。转换函数模型主要有加权平均(Weighted Averaging, WA)。包括 4 种变型: 典型回归加权平均(WA classical de-

收稿日期: 2013-01-02; 修订日期: 2013-04-17

基金项目: 国家自然科学基金青年基金项目(41001121)资助。

作者简介: 李鸿凯(1977-), 男, 河南许昌人, 博士, 讲师, 主要从事湿地生态与全球环境变化研究。E-mail: lihk431@nenu.edu.cn

shrinking, WA.cla)、反向回归加权平均(WA inverse deshrinking, WA.inv)、典型回归耐受值降权加权平均(Tolerance downweighted WA classical deshrinking, WA.cla.tol)和反向回归耐受值降权加权平均(Tolerance downweighted WA inverse deshrinking, WA.inv.tol)),加权平均偏最小二乘(Weighted Averaging Partial Least Squares, WA-PLS),偏最小二乘(Partial Least Squares, PLS),最大似然(Maximum Likelihood, ML),现代模拟技术(包括 Modern Analogue Technique, MAT 和 Weighted Average Modern Analogue Technique, WMAT),Imbrie & Kipp 因子分析(Imbrie and Kipp Factor Analysis, IKFA)。

下载 R 语言的安装源程序可访问 <http://cran.r-project.org>。R 安装完成后,可有以下 2 种方式安装 rioja 软件包。① 在线安装:从 R 界面菜单栏依次点击“程序包>安装程序包…”,选择任一镜像点,再从程序包列表中选择 rioja 后自动下载安装。也可在 R 界面控制台中直接输入 `install.packages("rioja")`,选择镜像点自动下载安装;② 本地安装:从 <http://cran.r-project.org/web/packages/rioja/index.html> 下载 zip 包存储于本地磁盘,下载的 zip 包不需解压,然后依次点击菜单栏的“程序包>从本地 zip 文件安装程序包…”,选择本地磁盘上存储的 zip 包文件安装 rioja。R 各种功能都是通过函数命令实现的。由于 rioja 不属于 R 的内建标准包,每次打开 R 后需要重新加载才可以调用其中的函数。加载可通过 `library(rioja)` 或 `require(rioja)` 命令实现。

## 2 样品采集与室内分析

### 2.1 样品采集

2008 年 7~8 月期间,在长白山区的哈泥(42°12'50"N, 126°31'05"E)、金川(42°20'47"N, 126°21'35"E)、赤池(42°03'16"N, 128°03'22"E)和圆池(42°01'55"N, 128°25'58"E)4 个泥炭地取样。取样时根据水文梯度沿剖面线选择不同泥炭藓(*Sphagnum*)丘,每个藓丘选择丘顶、丘中和丘下等不同生境作为取样点,用 GPS 记录每个取样点的位置。取样时首先用剪刀将取样点表面修剪整齐,在表面做一个直径为 5 cm 的圆形标记,用刀沿标记边缘垂直向下切开至约 10 cm,从下部切断小心地取出表层的 10 cm 样品,修剪整齐,装入封口袋中标记为样品 A,待运回实验室烘干称质量,计算

泥炭湿度;在紧邻取 A 样的地方用同样的方法,再取样品 B,样品 B 取出后修剪整齐,去除顶部约 2~3 cm 泥炭藓绿色部分,留下 3~10 cm 部分装入封口袋,运回实验室冷冻保存,用于有壳变形虫的鉴定分析。这样根据泥炭地面积和生境多样性不同,哈泥取样 46 个,金川取样 15 个,赤池取样 7 个,圆池取样 13 个,累计采集样品 81 个。取样同时现场测量水位埋深,在取 A 和 B 样的位置向下挖一小洞,0.5 h 后测量洞内水位距表面的深度即水位埋深(DWT)。同时,取相邻表层泥炭藓,用力挤压释出其中水分,测量释出水的 pH 和电导率。测量采用 PD-501 型便携式多功能测量仪, pH 电极为 201T-F 型塑壳 pH 三复合电极,自动温度补偿;电导率电极为 DJS-1 铂黑电导电极。

### 2.2 有壳变形虫样品处理与鉴定

有壳变形虫样品前处理参考文献[45]和[46],但选择的是 20  $\mu\text{m}$  孔径的细筛网,具体方法如下:① 取采集的样品 B 置于烧杯中,蒸馏水浸泡 4 h 以上,再煮沸约 5 min,并轻轻搅拌,使有壳变形虫壳体与附着基质充分脱离;② 用蒸馏水冲洗样品过 300  $\mu\text{m}$  孔径标准筛,去除粗粒杂质,筛下样品过 20  $\mu\text{m}$ (625 目)标准筛,去除细粒杂质,保留 20~300  $\mu\text{m}$  部分;③ 蒸馏水冲洗细筛上样品入试管中,在离心机上以 3 000 rpm 的速度离心 5 min;④ 去除上部清液,封闭保存以备制作玻片观察、鉴定。

制作好的玻片放在生物显微镜下放大 100 倍观察鉴定并计数,个别需放大至 400~1 000 倍。有壳变形虫种类鉴定标准和命名主要依据文献[54],每个样品累计观察不少于 150 个壳体(最大为 248,最小为 150),部分样品由于壳体数过少(<100),在最终数据分析中被删除,这样获得有效样品 75 个,其中哈泥 42 个,金川 15 个,赤池 5 个,圆池 13 个。

## 3 转换函数的建立与优化

### 3.1 数据处理

本文的物种数据为物种的相对丰度。物种数据和环境因子数据以每样一行的方式分别输入,并读取到 R 工作空间。排序分析表明水位埋深(depth to water table, DWT), pH 和泥炭湿度是影响长白山区有壳变虫的组成的主要环境因子,可作为目标变量进行转换函数的构建<sup>[41]</sup>。转换函数模型采用 WA 和 WA-PLS, rioja 中分别对应函数 `WA()` 和 `WAPLS()`。结果的交叉验证采用 leave-one-out,

也称 jack-knifing, 通过 crossval(..., cv.method="loo")实现。评价模型预测性能的优劣主要根据预测均方根误差(root mean square error of prediction, RMSEP)和 $R^2$ (预测值和实测值相关系数平方)。

### 3.2 模型建立与优化

首先选择全部 75 个样品作为训练样本进行转换函数构建。水位埋深 WA.inv 和 WA-PLS 模型的 RMSEP 最小, 为 10.93,  $R^2$  为 0.57 (表 1)。pH 的 WA.inv.tol 模型 RMSEP 最小, 为 0.28, 但  $R^2$  小于 WA.cla.tol, 类似的情况也出现在泥炭湿度, WA.inv 和 WA-PLS 的 RMSEP 最小, 但  $R^2$  小于 WA.cla 和 WA.cla.tol。

许多研究表明, 残差值大的异常样品对转换函数模型的性能有显著影响, 为降低预测误差, 常把残差值大于环境梯度宽度 20% 的样品从训练样本集中剔除<sup>[8-11, 16-29]</sup>。经过这样的数据过滤, 有 9 个样品从水位埋深的训练样本集中剔除, 剩余 66 个。pH 剔除 12 个样品, 剩余 63 个, 泥炭湿度剔除 17 个样品, 剩余 58 个。异常样品剔除后, 3 个环境因子的转换函数预测性能都明显提高。水位埋深的 WA-PLS 第二分量 RMSEP 最小, 为 7.39,  $R^2$  也最大, 为 0.74。对于 pH 和泥炭湿度, WA.inv 和 WA-PLS 第一分量 RMSEP 最小, 但  $R^2$  也略小于 WA.cla。pH 的 RMSEP 为 0.18,  $R^2$  为 0.72。泥炭湿度的 RMSEP 为 1.95%,  $R^2$  为 0.62。

为了更直观地表示模型的预测性能, 通常需要绘制环境因子的模型预测值-实测值散点图和残差值-实测值散点图。为此 rioja 软件包针对不同的转换函数模型分别定义了 plot() 函数。但这些函数有时不能满足用户的个性化需求, 这时我们可以使用 R 内建的各种高级和低级绘图函数, 更加自由灵活地绘制符合自己要求的图表。图 1 为优化前后, 水位埋深、pH 和泥炭湿度 3 个环境因子

转换函数预测性能比较。图中可以看出, 模型优化以后, 预测值-实测值图中样点更集中于 1:1 附近, 残差值相对更小, 更集中于 0 轴上下。

## 4 讨 论

水位埋深、泥炭湿度和 pH 值是目前有壳变形虫-环境因子转换函数研究中常用的目标变量。特别是水位埋深是唯一在所有有壳变形虫转换函数研究中都被选择作为目标变量的环境因子。不同研究中水位埋深的 RMSEP 在数值上存在较大差距, 造成这种差距的原因主要有两方面。首先在这些研究中, 泥炭地的类型不同, 有的是雨养型泥炭地<sup>[7, 11, 20, 21, 29]</sup>, 有的虽以泥炭藓占优势, 但却具显具有矿养型的特点<sup>[9, 10, 16-19, 22-28]</sup>。这些不同类型泥炭地的水文状况存在明显差异, 因此水位埋深、pH 和泥炭湿度等我们感兴趣的环境因子对有壳变形虫种类组合变化的解释量及显著水平也各不相同, 这必然会导致转换函数预测精度上的差异。其次, 不同研究采用的训练样本集的环境梯度宽度不同, 梯度宽度越大, 模型的 RMSEP 也相应越大。我们总结了过去研究中水位埋深的 RMSEP 与梯度宽度(图 2), 分析表明二者呈极显著正相关( $r=0.72, p<0.01$ )。以图中的拟合直线作为模型的平均预测精度水平的话, 本研究结果在该线下方, 说明转换函数的预测性能比较理想。

转换函数研究要求获得最低的 RMSEP 值和最高的  $R^2$  值。为此研究人员通常有以下 2 方面的努力。第一、尝试不同的转换函数模型。目前文献报道的模型有 WA、WA-PLS、ML、MAT、IKFA 等。这些模型为我们提供了更多的选择, 但仍以 WA 和 WA-PLS 最受青睐, 这是本研究选择这两种模型的原因。其实这些模型并无本质上的优劣之分, 而且在许多情况下, 它们的预测性能虽有差

表 1 水位埋深、pH 和泥炭湿度转换函数预测性能

Table 1 Predication performance of transfer functions for DWT, pH and peat moisture

模型	水位埋深		pH		泥炭湿度	
	RMSEP	$R^2$	RMSEP	$R^2$	RMSEP	$R^2$
WA.inv	10.93 (7.75)	0.57 (0.72)	0.29 (0.18)	0.48 (0.72)	3.40% (1.95%)	0.34 (0.62)
WA.cla	12.99 (8.31)	0.59 (0.72)	0.36 (0.19)	0.50 (0.73)	4.85% (2.26%)	0.37 (0.63)
WA.inv.tol	11.28 (8.33)	0.55 (0.68)	0.28 (0.18)	0.51 (0.70)	3.42% (1.97%)	0.34 (0.62)
WA.cla.tol	13.35 (9.10)	0.56 (0.68)	0.36 (0.23)	0.52 (0.68)	4.76% (2.19%)	0.36 (0.62)
WA-PLS	10.93 (7.39)	0.57 (0.74)	0.29 (0.18)	0.48 (0.72)	3.40% (1.95%)	0.34 (0.62)

注: 括号里为数据过滤后的结果。



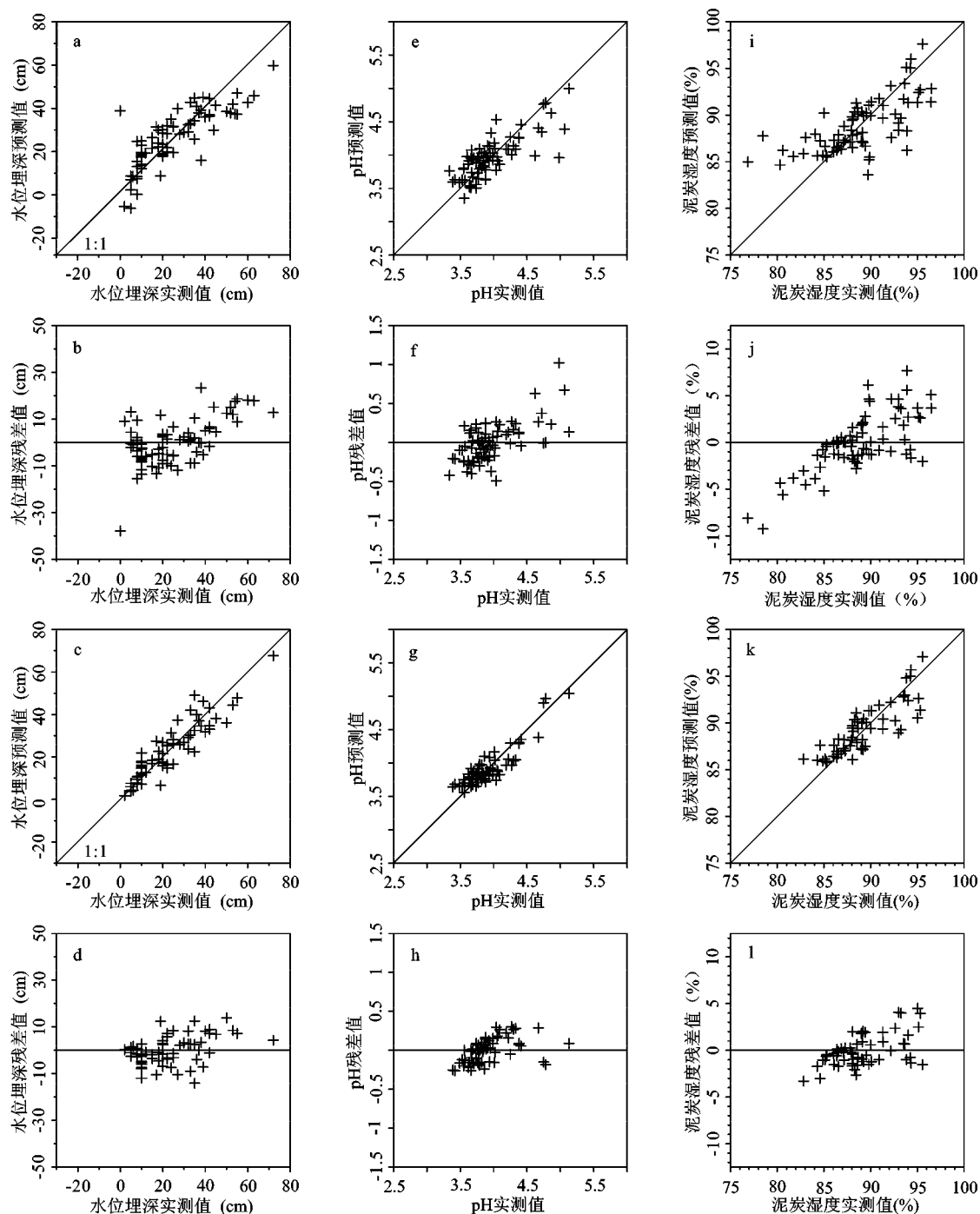


图1 优化前(a,b,e,f,i,j)后(c,d,g,h,k,l)水位埋深、pH和泥炭湿度转换函数模型的预测性能

Fig.1 Performance of transfer function models for DWT, pH and peat moisture, pre-( a,b,e,f,i,j) and post-( c,d,g,h,k,l) data filtering

异,但差异不大。通过比较,我们选择RMSEP值相对最低、 $R^2$ 相对最高的模型作为最优模型。第二、通过适当的数据过滤,剔除残差值大的异常样品。数据过滤虽然可以明显提高转换函数的预测性能,却以缩短训练样本集的环境梯度宽度为代价,因为被剔除的样品大部分位于环境梯度的两端,这一方面由于预测模型普遍存在轴端残差增

大的现象,另一方面样本集中环境梯度两端的样品量相对较少,代表性不足,造成残差较大。本研究中数据过滤后,水位埋深的梯度宽度由72 cm降低为70 cm。表面上看似降低幅度不大,但对比图1中a,b与c,d可以看出水位埋深低值端和高值端的样品量明显减少,特别是高值端,水位埋深大于60 cm的样品仅剩余YC1\_4(DWT=72 cm)。同

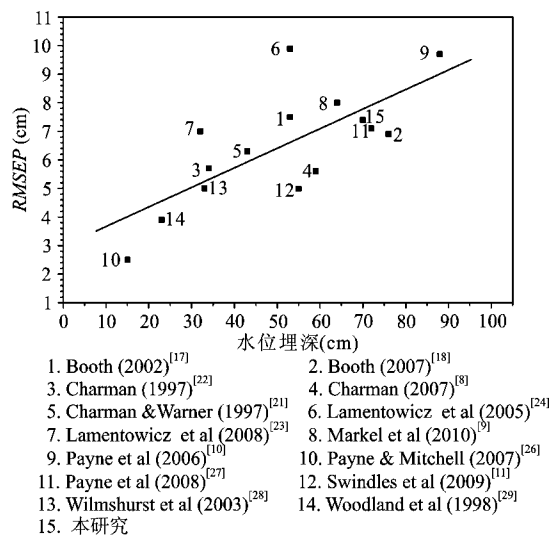


图2 不同研究中的RMSEP与水位梯度宽度关系

Fig.2 The RMSEP versus DWT gradient range in different studies

样的pH值和泥炭湿度经数据过滤后的梯度宽度也分别有不同幅度的降低,pH由1.79降为1.74,泥炭湿度由19.64%降低为12.72%。针对这种情况研究者需要在训练样本集的代表性与预测误差之间寻找平衡。本研究采用的20%切割点是普遍接受和采纳的标准,也便于与其它研究进行对比。建立转换函数为的是古环境的定量重建,本区如果泥炭剖面的有壳变形虫种类组成与现代训练样本集相同,水位埋深、pH和泥炭湿度可以分别以 $\pm 7.39$  cm、 $\pm 0.18$ 和 $\pm 1.95\%$ 的平均误差进行重建。

## 5 结 论

基于R语言的rioja软件包专用于第四纪研究的数据分析,其中的各种数学模型可以用来建立环境因子和代用指标的转换函数。有壳变形虫是一种重要的有潜力的生物指标。应用rioja软件包的WA和WA-PLS模型建立了长白山区泥炭地有壳变形虫与水位埋深、pH和泥炭湿度的转换函数。经20%数据过滤,水位埋深的WA-PLS模型第二分量最佳,RMSEP为7.39 cm, $R^2$ 为0.74;对于pH和泥炭湿度,WA-PLS第一分量和WA.inv都产生了最小的RMSEP和较高的 $R^2$ 值。pH的RMSEP为0.18, $R^2$ 为0.72。泥炭湿度的RMSEP为1.95%, $R^2$ 为0.62。这意味着在本区如果泥炭剖面的有壳变形虫种类组成与现代训练样本集相同,水位埋深、pH和泥炭湿度可以分别以 $\pm 7.39$  cm、 $\pm 0.18$ 和 $\pm 1.95\%$ 的平均误差进行重建。

## 参考文献:

- [1] Li D, Knudsen M F, Jiang H, et al. A diatom-based reconstruction of summer sea-surface salinity in the Southern Okinawa Trough, East China Sea, over the last millennium[J]. Journal of Quaternary Science,2012,27(8):771-779.
- [2] Mills K, Ryves D B. Diatom-based models for inferring past water chemistry in western Ugandan crater lakes[J]. Journal of Paleolimnology,2012,48(2):383-399.
- [3] Pestryakova L A, Herzschuh U, Wetterich, S. et al. Present-day variability and Holocene dynamics of permafrost-affected lakes in central Yakutia (Eastern Siberia) inferred from diatom records[J]. Quaternary Science Reviews,2012,51:56-70.
- [4] Engels S, Cwynar L C, Rees A B H, et al. Chironomid-based water depth reconstructions: An independent evaluation of site-specific and local inference models[J]. Journal of Paleolimnology,2012,48(4):693-709.
- [5] Cwynar L C, Rees A B H, Pedersen C R, et al. Depth distribution of chironomids and an evaluation of site-specific and regional lake-depth inference models: A good model gone bad[J]. Journal of Paleolimnology,2012,48(3):517-533.
- [6] Gouramanis C, Dodson J, Wilkins D, et al. Holocene palaeoclimate and sea level fluctuation recorded from the coastal Barker Swamp, Rottnest Island, south-western Western Australia[J]. Quaternary Science Reviews,2012,54:40-57.
- [7] Mischke S, Sun Z, Herzschuh U, et al. An ostracod-inferred large Middle Pleistocene freshwater lake in the presently hyper-arid Qaidam Basin (NW China)[J]. Quaternary International,2010,218(1-2):74-85.
- [8] Charman D J, Blundell A, ACCROTELM Members. A new European testate amoebae transfer function for palaeohydrological reconstruction on ombrotrophic peatlands[J]. Journal of Quaternary Science,2007,22(3):209-221.
- [9] Markel E R, Booth R K, Qin Y. Testate amoebae and  $\delta^{13}C$  of sphagnum as surface-moisture proxies in Alaskan peatlands[J]. The Holocene,2010,20(3):463-475.
- [10] Payne R J, Kishab K, Blackford J J, et al. Ecology of Testate amoebae (Protista) in south-central Alaska peatlands: building transfer function models for palaeoenvironmental studies[J]. The Holocene,2006,16(3):403-414.
- [11] Swindles G T, Charman D J, Roe, H M, et al. Environmental controls on peatland Testate amoebae (Protozoa: Rhizopoda) in the North of Ireland: Implications for Holocene palaeoclimate studies[J]. Journal of Paleolimnology,2009,42(1):123-140.
- [12] Juggins S. C2 Version 1.5 User guide. Software for ecological and palaeoecological data analysis and visualization[M]. Newcastle, UK: Newcastle University,2007.
- [13] Charman D J. Biostratigraphic and palaeoenvironmental applications of Testate amoebae[J]. Quaternary Science Reviews, 2001,20: 1753-1764.
- [14] Mitchell E A D, Charman D J, Warner B G. Testate amoebae

- analysis in ecological and paleoecological studies of wetlands: past, present and future[J]. *Biodiversity and Conservation*, 2008, **17**(9):2115-2137.
- [15] Chambers F M, Booth R K, et al. Development and refinement of proxy-climate indicators from peats[J]. *Quaternary International*, 2012, **268**:21-23.
- [16] Booth R K. Ecology of testate amoebae (Protozoa) in two Lake Superior coastal wetlands: implications for paleoecology and environmental monitoring[J]. *Wetlands*, 2001, **21**(4):564-576.
- [17] Booth R K. Testate amoebae as paleoindicators of surface-moisture changes on Michigan peatlands: Modern ecology and hydrological calibration[J]. *Journal of Paleolimnology*, 2002, **28**(3):329-348.
- [18] Booth R K. Testate amoebae as proxies for mean annual water-table depth in Sphagnum-dominated peatlands of North America[J]. *Journal of Quaternary Science*, 2007, **23**:43-57.
- [19] Booth R K, Sullivan M E, Sousa V A. Ecology of testate amoebae in a North Carolina pocosin and their potential use as environmental and paleoenvironmental indicators[J]. *Ecoscience*, 2008, **15**(2):277-289.
- [20] Charman D J, Warner B G. Relationship between testate amoebae (Protozoa: Rhizopoda) and microenvironmental parameters on a forested peatland in northeastern Ontario[J]. *Canadian Journal of Zoology*, 1992, **70**:2474-2482.
- [21] Charman D J, Warner B G. The ecology of testate amoebae (Protozoa: Rhizopoda) in oceanic peatlands in Newfoundland, Canada: modelling hydrological relationships for palaeoenvironmental reconstruction[J]. *Ecoscience*, 1997, **4**(4):555-562.
- [22] Charman D J. Modelling hydrological relationships of testate amoebae (Protozoa: Rhizopoda) on New Zealand peatlands[J]. *Journal of The Royal Society of New Zealand*, 1997, **27**(4):465-483.
- [23] Lamentowicz Ł, Lamentowicz M, Gabka M. Testate amoebae ecology and a local transfer function from a peatland in western Poland[J]. *Wetlands*, 2008, **28**(1):164-175.
- [24] Lamentowicz M, Mitchell, E A D. The ecology of testate amoebae (protists) in Sphagnum in north-western Poland in relation to peatland ecology[J]. *Microbial Ecology*, 2005, **50**(1):48-63.
- [25] Mitchell E A D, Buttler A J, Warner B G et al. Ecology of testate amoebae (Protozoa: Rhizopoda) in Sphagnum peatlands in the Jura mountains, Switzerland and France[J]. *Ecoscience*, 1999, **6**(4):565-576.
- [26] Payne R J, Mitchell E A D. Ecology of testate amoebae from mires in the central Rhodope Mountains, Greece and development of a transfer function for palaeohydrological reconstruction[J]. *Protist*, 2007, **158**:159-171.
- [27] Payne R J, Charman D J. and Matthews S. Testate amoebae as paleohydrological proxies in Sürmene Ağaçbaşı Yaylasi peatland (Northeast Turkey)[J]. *Wetlands*, 2008, **28**(2):311-323.
- [28] Wilmshurst J M, Wiser S K, Charman D J. Reconstructing Holocene water tables in New Zealand using testate amoebae: differential preservation of tests and implications for the use of transfer functions[J]. *The Holocene*, 2003, **13**(1):61-72.
- [29] Woodland W A, Charman D J, Sims P C. Quantitative estimates of water tables and soil moisture in Holocene peatlands from Testate amoebae[J]. *The Holocene*, 1998, **8**(3):261-273.
- [30] Booth R K, Jackson S T. A high-resolution record of late Holocene moisture variability from a Michigan raised bog, USA[J]. *The Holocene*, 2003, **13**(6):863-876.
- [31] Charman D J. Summer water deficit variability controls on peatland water-table changes: Implications for Holocene palaeoclimate reconstructions[J]. *The Holocene*, 2007, **17**(2):217-227.
- [32] Charman D J, Hohl V, Blundell A, et al. A 1000-year reconstruction of summer precipitation from Ireland: Calibration of a peat-based palaeoclimate record[J]. *Quaternary International*, 2012, **268**:87-97.
- [33] Elliott S M, Roe H M, Patterson R T. Testate amoebae as indicators of hydroseral change: An 8500 year record from Mer Bleue Bog, eastern Ontario, Canada[J]. *Quaternary International*, 2012, **268**:128-144.
- [34] Hendon D, Charman D J, Kent M. Palaeohydrological records derived from testate amoebae analysis from peatlands in northern England: within-site variability, between-site comparability and palaeoclimatic implications[J]. *The Holocene*, 2001, **11**(2):127-148.
- [35] Hendon D, Charman D J. High-resolution peatland water-table changes for the past 200 years: the influence of climate and implications for management[J]. *The Holocene*, 2004, **14**(1):125-134.
- [36] Lamarre A, Garneau M, Asnong H, et al. Holocene paleohydrological reconstruction and carbon accumulation of a permafrost peatland using testate amoeba and macrofossil analyses, Kuujuaupik, subarctic Québec, Canada[J]. *Review of Palaeobotany and Palynology*, 2012, **186**:131-141.
- [37] Lamentowicz M, Cedro A, Galka M, et al. Last millennium palaeoenvironmental changes from a Baltic bog (Poland) inferred from stable isotopes, pollen, plant macrofossils and testate amoebae[J]. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 2008, **265**(1-2):93-106.
- [38] Mitchell E A D, van der Knaap W O, van Leeuwen J F N, et al. The palaeoecological history of the Praz-Rodet bog (Swiss Jura) based on pollen, plant macrofossils and testate amoebae (Protozoa)[J]. *The Holocene*, 2001, **11**(1):65-80.
- [39] Schoning K, Charman D J, Wastegård S. Reconstructed water tables from two ombrotrophic mires in eastern central Sweden compared with instrumental meteorological data[J]. *The Holocene*, 2005, **15**(1):111-118.
- [40] 李鸿凯,王升忠,卜兆君,等.长白山区典型泥炭藓沼泽的有壳变形虫[J].*湿地科学*, 2010, **8**(3):249-255.
- [41] 李鸿凯,卜兆君,王升忠,等.长白山区泥炭地现代有壳变形虫环境意义探讨[J].*第四纪研究*, 2009, **29**(4):817-824.
- [42] 秦养民,顾延生,王银善,等.天鹅洲长江故道近代环境变化的

- 有壳变形虫记录[J]. 地质科技情报, 2007, 26(3): 37~42.
- [43] R Core Team. R: A Language and Environment for Statistical Computing [CP/OL]. Vienna, Austria: R Foundation for Statistical Computing, 2012 [2012-12-15]. <http://www.r-project.org>.
- [44] Juggins S. rioja: Analysis of Quaternary Science Data [CP/OL]. 2012 [2012-12-15]. <http://cran.r-project.org/package=rioja>.
- [45] Hendon D, Charman D J. The preparation of testate amoebae (Protozoa: Rhizopoda) samples from peat[J]. The Holocene, 1997, 7(2): 199-205.
- [46] Charman D J, Hendon D, Woodland W A. The identification of testate amoebae (Protozoa: Rhizopoda) in peats[M]. London: Quaternary Research Association, 2000.

## Building Transfer Functions Between Testate amoeba and Environmental Variables with ‘rioja’ Package

LI Hong-kai<sup>1,2</sup>, LI Wei-wei<sup>1</sup>, PU You-bao<sup>1</sup>, WANG Cong-yang<sup>1</sup>, WANG Song-mei<sup>1</sup>, YANG Xiao-lin<sup>1</sup>

(1. State Environmental Protection Key Laboratory of Wetland Ecology and Vegetation Restoration, Northeast Normal University, Changchun, Jilin 130024, China; 2. Institute for Peat and Mire Research, School of Geographical Science, Northeast Normal University, Changchun, Jilin 130024, China)

**Abstract:** R language, as an open source programming language and software environment, is widely used in statistics for its free availability. The ‘rioja’ package of R specially deals with the analysis of Quaternary science data, containing functions for constrained clustering, transfer functions and plotting stratigraphic data. Testate amoebae are a group of unicellular protists living in terrestrial habitats. Their decayed resistant and morphologically diagnostic shells (tests) allow them to be extensively used as proxy in peat based paleoenvironmental reconstruction. This study aimed to: ① Present an example of application of ‘rioja’ package; ② build Testate amoeba-based transfer functions for quantitatively reconstructing paleoenvironment changes in Changbai Mountains with peat archive. The training set was constituted by 75 samples collected from four peatlands, Hani (42°12′50″N, 126°31′05″E), Jinchuan (42°20′47″N, 126°21′35″E), Chichi (42°03′16″N, 128°03′22″E) and Yuanchi (42°01′55″N, 128°25′58″E), in Changbai Mountains, northeast China. Three factors, depth to water table (DWT), pH and peat moisture, were selected as the target environmental variables. The models of Weighted Averaging (WA) and Weighted Averaging Partial Least Squares (WA-PLS) were used to build transfer functions. Leave-one-out was chosen as cross validation method. The results showed that the second component of WA-PLS is the best models for DWT producing a RMSEP of 7.39 and  $R^2$  of 0.74. For pH and peat moisture, both first component of WA-PLS and WA with inverse deshrinking could be regarded as the best models for they have the lowest RMSEP and relatively higher  $R^2$ . The RMSEP of pH is 0.18 and  $R^2$  is 0.72, while for peat moisture RMSEP is 1.95% and  $R^2$  is 0.62. The performances of the transfer function were comparable with other studies in the world. DWT, pH and peat moisture could be quantitatively reconstructed with the mean errors of  $\pm 7.39$  cm,  $\pm 0.18$  and  $\pm 1.95\%$ , respectively, if Testate amoebae assemblage of profiles was the same as the training set in this study.

**Key words:** ‘rioja’ package; transfer function; Testate amoebae; depth to water table; pH; peat moisture