

面向森林生态学研究的 Web 处理服务系统^①



高仕源^{1,2}, 罗万明¹, 罗泽¹, 白晓航³, 于景华³

¹(中国科学院 计算机网络信息中心, 北京 100190)

²(中国科学院大学, 北京 100049)

³(中国科学院 沈阳应用生态研究所, 沈阳 110016)

通讯作者: 罗万明, E-mail: lwm@cnic.cn

摘要: 森林是陆地生态系统的主体, 森林生态保护与恢复是生态文明建设的重要组成部分. 面对森林生态模型单机化、重复开发部署的现状, 本文设计并实现一个面向森林生态学研究的 Web 处理服务系统. 该系统根据开放地理空间信息联盟 (Open Geospatial Consortium, OGC) 制定的空间数据共享及互操作的接口规范, 基于 WebService 技术, 实现了面向森林生态学研究的多模型整合, 并将模型计算结果进行展示, 减少了重复开发和部署工作, 提高了数据与服务的共享性.

关键词: 森林生态学; 网络处理服务; 开放地理空间信息联盟 (OGC); 地形模型; 群丛分类; 生物量; 物种多样性

引用格式: 高仕源, 罗万明, 罗泽, 白晓航, 于景华. 面向森林生态学研究的 Web 处理服务系统. 计算机系统应用, 2019, 28(5): 49-56. <http://www.c-s-a.org.cn/1003-3254/6947.html>

Web Processing Service System for Forest Ecology Research

GAO Shi-Yuan^{1,2}, LUO Wan-Ming¹, LUO Ze¹, BAI Xiao-Hang³, YU Jing-Hua³

¹(Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China)

²(University of Chinese Academy of Sciences, Beijing 100049, China)

³(Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, China)

Abstract: The forest is the mainstay of terrestrial ecosystems. So the protection and restoration of forest ecosystems are the significant parts of ecological civilization construction. In order to solve the problem that forest ecological models are deployed in local machines and developed repeatedly, we manage to design and implement a web processing system for the study on forest ecology. According to the interface specification of spatial data sharing and interoperability, which is formulated by Open Geospatial Consortium (OGC), and by combining with the technology of WebService, the system manages to realize the multi-model integration for the study on forest ecology and displays the calculation results, which facilitates the communication of researchers and the public. It could reduce the work of developing repeatedly, and increase the sharing of data and service.

Key words: forest ecology; Web processing service; Open Geospatial Consortium (OGC); topography model; association classification; biomass; biodiversity

森林是陆地生态系统的主体, 是陆地上最大的生物群落, 是人类社会赖以生存的物质基础, 是哺育人类的摇篮, 是人类不可或缺的战略资源. 全球 5 个国家拥

有超过 50% 的森林资源, 中国位列其中. 由于我国国土面积广大、人口众多, 森林分布不均匀、人均占有量少, 根据国家林业局在 2014 年 2 月发布的《中国森

① 基金项目: 中国科学院“十三五”信息化建设专项 (XXH13506-303)

Foundation item: CAS Special Fund for Informatization Construction in 13th Five-Year Plan (XXH13506-303)

收稿时间: 2018-12-08; 修改时间: 2018-12-20; 采用时间: 2019-01-15; csa 在线出版时间: 2019-05-01

林资源简况——第八次全国森林资源清查》结果显示,我国人均森林面积占有量只有世界人均占有量的四分之一,并且面临着林业生物灾害频发、森林资源采伐使用不尽合理的局面^[1]。

当前国内有一些研究森林生态保护与恢复工作的高校、科研机构,收集了一些森林资源数据并建立了一些学科模型,但相关研究工作普遍处于单机化运行,普遍存在科研人员重复构建计算模型、搭建计算环境的问题。为了减少此类重复性工作,提高数据和功能的共享与互操作,开放地理信息系统联盟 OGC(Open Geospatial Consortium) 提出了可扩展的能无缝集成各种在线空间信息处理和位置服务的框架 OWS(OGC Web Services)^[2,3]。作为 OWS 规范中面向共享的服务标准,网络处理服务 (Web Processing Service) 利用 Web Services 技术跨平台、跨语言的特性,能够提供异构平台的无缝衔接技术手段,通过封装专业性强、具有学科特色的模型,并对外提供统一标准化服务接口,其他科研机构和个人可以调用相关网络处理服务,获得相应处理结果,实现一处构建开发、处处重复使用的目的^[4-6]。

在东北森林屏障带生态保护与恢复的研究工作中,采集了大量森林经纬度、坡度、坡向信息,乔木、灌木、草木的胸径、高度、盖度等数据,根据调查数据研究相应学科模型,通过分析模型计算结果,对东北森林屏障带进行评估与规划。本文根据 OWS 规范中的网络处理服务相关标准,采用 Web Service 技术设计并实现一个面向森林生态学研究的 Web 处理服务系统。客户端通过向目标服务端提交符合网络处理服务规范的请求,获取该服务端所提供的网络处理服务列表、网络处理服务描述文档,根据返回信息将待处理数据提交至服务端,利用相应模型进行处理,获得模型处理结果。

1 系统总体设计

如图 1 所示,系统整体分为两部分:客户端程序和服务器端程序。客户端作为服务请求端,通过网络向服务器,即网络处理服务提供者所暴露的三个接口: GetCapabilities、DescribeProcess、Execute, 发送业务请求,服务器在业务逻辑层根据 request 请求参数对请求进行分发,调用相应逻辑模块进行处理。

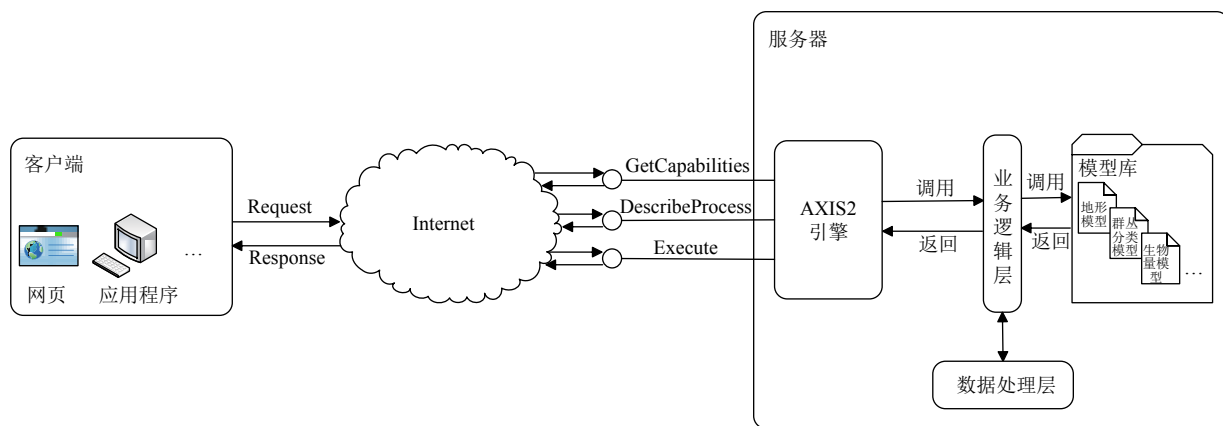


图 1 系统整体结构

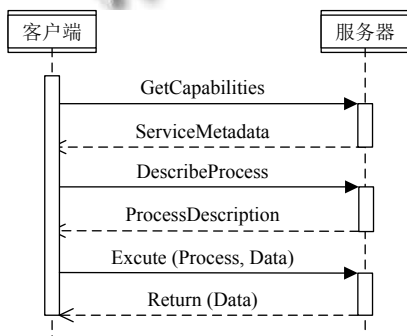


图 2 网络处理服务 UML 序列图

1.1 接口层设计

接口层负责实现客户端调用并由服务器执行的以下操作: GetCapabilities 操作允许客户端请求有关服务器提供的功能和处理能力的信息, DescribeProcess 操作允许客户端请求服务器提供的选定进程的详细元数据信息、Execute 操作允许客户端执行由进程标识符、期望的数据输入输出格式组成的执行过程, 见图 2。

① GetCapabilities 操作

GetCapabilities 操作的请求用来获取目标服务器

功能的描述性文档, 请求参数包括强制和可选参数两部分, 如表 1 所示。

GetCapabilities 操作的响应返回描述该服务器功能的 XML 文档, 同样包括强制和可选参数两部分, 如表 2 所示。

表 1 GetCapabilities 操作的请求参数

参数名称	参数个数	参数描述
service	1(强制)	表示服务类型, 值为“WPS”
Request	1(强制)	表示请求类型, 值为“GetCapabilities”
AcceptVersions	0 或 1(可选)	表示网络处理服务的版本
language	0 或 1(可选)	语言标识, 需符合 RFC4646 语言编码

表 2 GetCapabilities 操作的响应文档参数

参数名称	参数描述	数据类型和值	参数个数
service	服务标识	字符串型, 非空应包括“WPS”	1(强制)
version	操作的规范版本	字符串型, 非空应包括“1.0.0”	1(强制)
update Sequence	服务元数据文档中发生更改时增加的值	字符串型, 非空值由各服务器实现所选	0 或 1(可选)
lang	语言标识	字符串型, 非空符合 RFC4646 语言编码	1(强制)
ServiceIdentification	关于指定服务器的元数据	与 OWS 相同	1(强制)
ServiceProvider	此服务器运营商的元数据	同上	1(强制)
OperationsMetadata	此服务器实现的服务和详细操作的元数据, 包括操作请求的 URL	同上	1(强制)
ProcessOfferings	此服务器提供的处理服务的简短描述无序列表	ProcessOffering 数据结构	1(强制)
Languages	服务器支持的语言	Languages 数据结构	1(强制)
WSDL	描述所有服务器所提供操作和处理功能的 WSDL 文档的位置	WSDL 数据结构	0 或 1(可选)

② DescribeProcess 操作

DescribeProcess 操作的请求用来获取服务器上一个或多个处理服务的完整描述信息, 该描述信息包括输入输出参数和格式. 请求参数包括强制参数及可选参数, 具体信息如表 3.

DescribeProcess 操作的响应是一个 Process-Descriptions data 结构的 XML 文档, 按照请求中所包含标识符 (Identifier) 将一个或多个处理服务的描述封装 XML 文档返回. 主要的强制参数及可选参数如表 4 所示.

表 3 DescribeProcess 操作的请求参数

参数名称	参数描述	数据类型和值	参数个数
service	服务类型标识	字符串型, 值为“WPS”	1(强制)
request	操作名称	字符串型, 值为“DescribeProcess”	1(强制)
version	操作的规范版本	字符串型, 非空值由各网络处理服务实现规范和模式版本确定	1(强制)
language	语言标识	字符串型符合 RFC4646 语言编码, 必须是 Capabilities Languages 中所列语言	0 或 1(可选)
Identifier	处理服务标识符	字符串型, 非空值是在服务元数据 (Capabilities) 文档中 ProcessOfferings 部分所列出的标识符	1 或多个(强制)

表 4 DescribeProcess 操作的响应文档参数

参数名称	参数描述	数据类型和值	参数个数
Process Description	完整的处理服务描述, 包括所有的输入和输出参数	ProcessDescription 数据结构	1 或多个(强制) 对应各处理标识符
service	服务标识	字符串型应包含“WPS”	1(强制)
version	操作的规范版本	字符串型, 非空值由各网络处理服务实现规范和模式版本确定	1(强制)
lang	语言标识	字符串型, 非空符合 RFC4646 语言编码	1(强制)
Identifier	处理服务标识符	ows:CodeType	1(强制)
Title	处理服务标题	字符串型	1(强制)
DataInputs	该处理服务的必须和可选输入的参数列表	DataInputs 数据结构	0 或 1(可选)
ProcessOutputs	执行处理服务得到的必须和可选输出列表	ProcessOutputs 数据结构	1(强制)

③ Executes 操作

Execute 操作允许客户端使用输入参数值去调用服务器中指定处理服务, 从而得到返回结果. 与前两个

操作不同, 该操作强制实现 HTTP POST 请求, 可选实现 HTTP GET 请求, 一般不实现 HTTP GET 请求.

Execute 操作的请求参数如表 5 所示.

表5 Execute 操作的请求参数

参数名称	参数描述	数据类型和值	参数个数
service	服务类型标识	字符串型, 值为“WPS”	1(强制)
request	操作名称	字符串型, 值为“Execute”	1(强制)
version	操作的规范版本	字符串型, 非空值由各网络处理服务实现规范和模式版本确定	1(强制)
Identifier	明确的标识符或处理服务名	字符串型, 非空值是 Capabilities 文档中的可用的处理服务标识符	1(强制)
DataInputs	该处理服务执行所需输入列表	DataInputs 数据结构	0 或 1(可选)
ResponseForm	定义了响应类型, 原始数据或 XML 文档	ResponseForm 类型数据结构	0 或 1(可选)
language	语言标识	字符串型符合 RFC4646 语言编码, 必须是 Capabilities Languages 中所列语言	1(可选)

Execute 操作的响应文档参数如表 6 所示。

表6 Execute 操作的响应参数

参数名称	参数描述	数据类型和值	参数个数
service	服务标识	字符串型, 非空应包含“WPS”	1(强制)
version	操作的规范版本	字符串型, 非空值由各网络处理服务实现规范和模式版本确定	1(强制)
lang	语言标识	字符串型, 非空符合 RFC4646 语言编码	1(强制)
serviceInstance	GetCapabilities 调用的网络处理服务的 URL	URL 类型	1(强制)
Process	处理服务描述	ProcessBrief 数据结构	1(强制)
Status	该处理服务的执行状态	Status 数据结构	1(强制)
DataInputs	该处理服务执行所需输入列表	DataInputs 数据结构	0 或 1(可选)
OutputDefinitions	定义执行该处理服务所期望得到的输出列表	OutputDefinitions 数据结构	0 或 1(可选)
ProcessOutputs	该处理服务执行的输出值列表	ProcessOutputs 数据结构	0 或 1(可选)

1.2 业务逻辑层设计

在服务器端, 业务逻辑层负责处理客户端对不同接口的调用请求。当接收到 GetCapabilities 请求时, 向模型库请求所有模型的标识符和简单描述, 并将获得的信息封装为 XML 响应文档返回; 当接收到 DescribeProcess 请求时, 根据标识符的值向模型库请求与该标识符对应的算法完全描述, 并将相关信息封装成 ProcessDescriptions 响应文档返回; 当接收到 Execute 请求时, 将 DataInputs 下 Input 参数作为输入, 解析为系统定义的数据类型, 根据标识符的值请求对应的模型, 对输入数据进行处理, 并将处理结果封装成响应文档返回。

1.3 数据处理层设计

数据处理层负责对输入数据进行类型匹配、转化及校验, 并对输出数据进行格式转换。由于网络处理服务需要接收和返回除整型、浮点型、字符型等基本数据类型外的矢量数据、栅格数据等生态环境领域数据类型, 因此在将数据提交到模型库执行前有必要对其进行转换。同理, 当需要返回复杂数据类型的数据时, 也需要对模型处理结果进行转化。

2 模型库构建

2.1 地形模型

经纬度、海拔、坡度等地形数据, 对野外生态环境观测与保护有着重要作用。

本系统构建的地形模型, 通过输入数字高程模型 (Digital Elevation Model, DEM) 和经纬度得到给定位置海拔等信息。

数字高程模型是一组包含地表相关情况的有序数列, 是数字地形模型 (Digital Terrain Model, DTM) 的一个分支^[7]。其包含了平面坐标及高程值两部分数据, 主要有三种类型的数据: 规则格网、不规则三角网结构、等高线型, 现常用规则格网型数据。数字地形模型可以通过如下公式表示:

$$Kp = fk(up, vp), \quad k = 1, 2, 3, \dots, m, p = 1, 2, 3, \dots, n \quad (1)$$

当 m 取值为 1 时, 表示数字高程模型, 即:

$$V_i = (X_i, Y_i, Z_i), \quad i = 1, 2, \dots, n \quad (2)$$

其中, X_i, Y_i 表示表面坐标, Z_i 是该点对应的高程值。实验数据来自于“地理空间数据云”网站在线计算服务, “DEM 高程数据切割”模块中 90 m 分辨率 SRTM 数据。

2.2 群丛分类模型

群丛是植被分类的基本单位,相当于植物分类中“种”的分类等级,进行群丛分类是了解一个地区植被特征最基本、最重要方法之一^[8-10].

在该模型中,需要了解样地划分以及样地内所有胸径大于1 cm的木本植物.通常在生态学中,用重要值(Important Value, IV)表示物种在样地群落中的相对重要性^[11],计算公式如下:

$$IV = (P + S_{DBH}) / 2 \quad (3)$$

式中, P 表示物种相对多度, S_{DBH} 表示物种相对胸高断面面积.根据样方划分与样方内各物种重要值形成二维矩阵.

目前学界常用的是双向指示种分析法(Two-Way Indicator SPecies ANalysis, TWINSpan),是基于指示种分析法的改进算法^[12],算法流程如下:

算法1. TWINSpan 算法

- 1) 对数据进行相互平均排序(Reciprocal Averaging, RA).
- 2) 在形成的第一排序轴以轴心为中心将所有样方划分为正负两组.
- 3) 选择排序轴两端的种作为指示种(indicator species),其指示意义的大小用指示值(indicator value)来衡量:

$$D(i) = |n_1(i)/N_1 - n_2(i)/N_2| \quad (i=1,2,\dots,p)$$

其中, $D(i)$ 为种 i 的指示值, N_1 和 N_2 分别表示步骤2中正负组中的样方数, $n_1(i)$ 和 $n_2(i)$ 分别表示种 i 在正负组中出现的样方数.

- 4) 根据选择的指示种计算样方的指示分(indicator score),并按指示得分将样方分组,选择适当的阈值将样方分为正负两组.比较相互平均排序的结果,如果不完全一致,将排序轴中心向两侧扩展出的中性带(indifferent zone)中错分类的调整过来.
- 5) 重复上述过程,直到每个组内样方数降到设定值,分类结束.

2.3 生物量模型

森林生物量是指在某一时刻在单位面积内存在的所有林木有机质总量.由于森林生物量是森林生态系统重要的状态特征,是研究全球碳循环、森林系统碳储量等生态问题研究的基础,因此了解森林生物量有十分重要的意义^[13].目前可以测定树木生物量的方法有皆伐法、平均木标准木法、径级标准木法和异速生长模型法等,其中直接测量获得的生物量最为准确.但是由于该方法需要耗费大量人力物力,且对生态环境造成一定破坏,不能对整个调查区域内的林木进行全部砍伐,因此普遍采用生物量模型法进行森林生物量的估算.

在建立生物量模型的过程中,主要考虑与生物量计算有关且易获取的林木特征统计量.根据既有的研

究成果以及实际采集林木特征数据,采用以下生物量计算模型^[14-16]:

$$W = \alpha D^\beta H^\gamma \quad (4)$$

其中, α 、 β 、 γ 是模型参数, D 是林木的胸径, H 是树高.考虑到不同用户对不同林木使用的 α 、 β 、 γ 系数不尽相同,因此在系统设计时允许用户自行输入相应数值.

2.4 物种多样性模型

本文所研究的物种多样性主要是指森林生态系统中,植物物种多样性问题,主要表征了该生态系统的复杂性、衡量其植物资源的丰富度,可以为生态系统稳定性评估、未来发展的预测提供数据支持.

本文实现了以下指数进行物种多样性评估^[17-20]:

① 物种丰富度指数

Margalef 指数:

$$R = \frac{S - 1}{\ln N} \quad (5)$$

② 物种多样性指数

Simpson 指数:

$$D = 1 - \sum_{i=1}^S \left(\frac{N_i}{N} \right)^2 \quad (6)$$

ShannonWiener 指数:

$$H = - \sum_{i=1}^S \left(\frac{N_i}{N} \right) \ln \frac{N_i}{N} \quad (7)$$

③ 物种均匀度指数

Pielou 指数:

$$J = \frac{H}{\ln S} \quad (8)$$

其中, S 是样地内物种总数, N 为所有物种的个体总数, N_i 为物种 i 在样地内的物种个体数.

3 系统展示

本系统界面如图3所示,根据前文所述主要有三部分功能: GetCapabilities 操作对应的获取网络处理服务信息功能、 DescribeProcess 操作对应的获取处理服务信息、 Execute 接口对应的模型计算应用.数据来源于中国科学院沈阳应用生态研究所,辽宁省部分森林样地调查数据.



图3 系统界面

选择获取服务基本信息,系统会弹出当前系统基本信息及可提供所有处理服务的描述信息,如图4所示.

根据业务需要选择所需模型,查看计算结果.在地形模型中,在下拉列表选择某一地区数据,通过输入一个数字高程数据集 DEM 和给定位置的经纬度,获取到该点的海拔以及整个 DEM 数据集包含区域内的最高海拔、最低海拔、平均海拔等数据,结果如图6所示.



图4 服务描述信息

根据业务需要选择查看某个或多个处理服务的具体信息,包括服务标识、数据的输入输出格式等信息,如图5.



图5 处理服务详细信息



图6 地形模型计算结果

在群从分类模型中,同样选择某一样地数据集,通过向后台发送样方和物种的二元数据,后台根据前文所述的 TWINSpan 算法的相关原理,经过计算后返回该样地划分的各个样方所属群从种类结果,前端自动弹出相应计算结果.

在生物量模型中,选择某一物种数据集,输入 α 、 β 、 γ 的取值,客户端会将返回结果组装成 echarts 进行展示,可以自行选取以折线图、柱状图等形式进行显示.如图8所示.

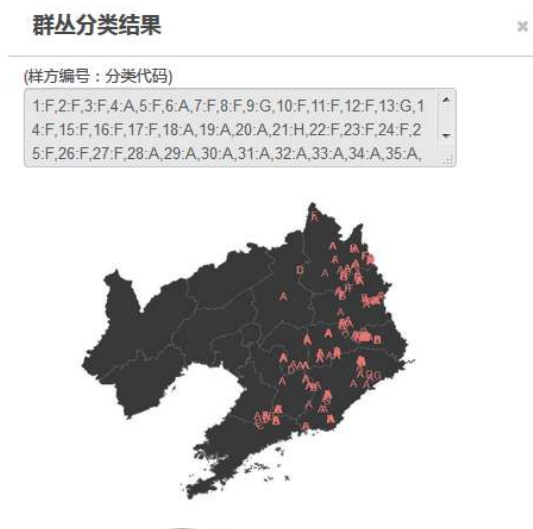


图7 群落分类模型计算结果

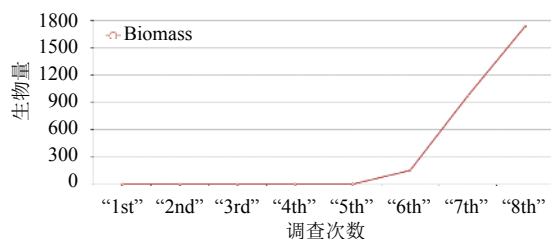


图8 生物量模型计算结果

在物种多样性模型中,选择需要查看的样地编号,服务器将调取该样地的所有林木的统计数据,计算出各评估指数返回给客户端进行展示.如图9所示.

4 总结

面向森林生态学研究的 Web 处理服务系统根据 OGC 相关规范,以 Web Service 技术为基础,整合相关森林生态学模型,改变传统单机独立重复开发与部署模型计算环境的方式,将计算模型集中化、网络化,通过提供网络处理服务的方式,减少了研究人员重复开发生态学模型和搭建相应处理环境的时间及精力.

系统各层之间松耦合度,以及模型库中各模型相互独立,方便今后系统的改进和拓展.未来将进一步增加生态学研究模型,并研究拓展系统可处理的空间数

据类型.

生物多样性评估结果



图9 群落分类模型计算结果

参考文献

- 1 国家林业局. 第八次全国森林资源清查结果. 林业资源管理, 2014, (1): 1-2. [doi: 10.13466/j.cnki.lyzygl.2014.01.001]
- 2 Open Geospatial Consortium Inc(OGC). OpenGIS Web Processing Service vl. 0. 0. <http://www.opengeospatial.org/>, 2007.
- 3 Open Geospatial Consortium Inc(OGC). OGC WPS 2.0.2 Interface Standard: Corrigendum 2. <http://www.opengeospatial.org/>, 2018.
- 4 李光师. 基于 OGC Web 服务规范的空间信息共享. 鞍山师范学院学报, 2013, 15(4): 50-53. [doi: 10.3969/j.issn.1008-2441.2013.04.012]
- 5 李德卿. 基于 WebService 技术的代理商管理系统的设计与实现[硕士学位论文]. 青岛: 青岛科技大学, 2018.
- 6 宋东泽, 罗万明, 阎保平. 一个生态传感网的 Web 处理服务平台. 科研信息化技术与应用, 2015, 6(2): 37-46. [doi: 10.11871/j.issn.1674-9480.2015.02.005]
- 7 程维明, 周成虎, 柴慧霞, 等. 中国陆地地貌基本形态类型定量提取与分析. 地球信息科学学报, 2009, 11(6): 725-736. [doi: 10.3969/j.issn.1560-8999.2009.06.007]
- 8 张文静, 张钦弟, 王晶, 等. 多元回归树与双向指示种分析在群落分类中的应用比较. 植物生态学报, 2015, 39(6): 586-592. [doi: 10.17521/cjpe.2015.0056]
- 9 Sun XW, Yang QS, Liu HM, et al. Classification of plant associations based on a 20 hm² dynamics plot of evergreen broad-leaved forest in Mt. Tiantong, Zhejiang, China.

- Chinese Journal of Plant Ecology, 2018, 42(5): 550–561. [doi: 10.17521/cjpe.2017.0263]
- 10 邢韶华, 于梦凡, 杨立娟, 等. 关于植物群丛划分的探讨. 生态学报, 2013, 33(1): 310–315. [doi: 10.5846/stxb201108241235]
- 11 杨庆松, 马遵平, 谢玉彬, 等. 浙江天童 20ha 常绿阔叶林动态监测样地的群落特征. 生物多样性, 2011, 19(2): 215–223. [doi: 10.3724/SP.J.1003.2011.09013]
- 12 王健, 张韬. TWINSpan 算法的实现和应用. 内蒙古大学学报(自然科学版), 2007, 38(1): 79–83.
- 13 曾伟生, 陈新云, 蒲莹, 等. 基于国家森林资源清查数据的不同生物量和碳储量估计方法的对比分析. 林业科学研究, 2018, 31(1): 66–71. [doi: 10.13275/j.cnki.lykxyj.2018.01.008]
- 14 方精云, 刘国华, 徐嵩龄. 我国森林植被的生物量和净生产量. 生态学报, 1996, 16(5): 497–508.
- 15 Zeng WS. Developing tree biomass models for eight major tree species in China. Biomass Volume Estimation and Valorization for Energy. 2017. 1–21. [doi: 10.5772/65664]
- 16 符利勇, 唐守正, 张会儒, 等. 东北地区两个主要树种地上生物量通用方程构建. 生态学报, 2015, 35(1): 150–157. [doi: 10.5846/stxb201403310601]
- 17 马菁, 张学俭. 红松针阔混交林乔灌木生物多样性研究. 林业调查规划, 2018, 43(5): 56–59, 63. [doi: 10.3969/j.issn.1671-3168.2018.05.011]
- 18 张燕夫. 人工营造的阔叶树红松混交林生物多样性. 林业科技通讯, 2018, (9): 15–17. [doi: 10.13456/j.cnki.lykt.2018.09.004]
- 19 魏忠平, 刘阳, 王婷, 等. 辽宁老秃顶子自然保护区典型森林群落生物多样性研究. 中南林业科技大学学报, 2016, 36(4): 68–72. [doi: 10.14067/j.cnki.1673-923x.2016.04.013]
- 20 韩风林, 布仁仓, 常禹, 等. 大兴安岭白桦-兴安落叶松林火烧迹地林下植被群落恢复过程的动态分析. 生态学杂志, 2015, 34(2): 312–318.