

栖息地适宜度评价方法研究进展

易雨君¹, 程曦¹, 周静²

1. 北京师范大学水沙科学教育部重点实验室, 北京 100875; 2. 中国水电顾问集团成都勘测设计研究院, 四川 成都 610072

摘要: 栖息地模拟法是根据指示物种所需的物理生境条件确定河流流量, 评价栖息地适宜度, 为水生生物提供一个适宜的物理生境。与其他方法相比, 栖息地模拟法考虑了生物本身对物理生境的要求, 需要建立物种-生境评价指标。文章总结了目前用于描述物种—生境关系的栖息地适宜度评价方法, 包括栖息地适宜度指数、多元统计方法、模糊逻辑方法、人工神经网络、对多物种和群落的统计分析, 归纳了各个方法的进展和应用情况, 重点分析了统计学方法在评估栖息地适宜度时的优势和不足。栖息地适宜度指数有二元、单变量、多变量3种格式, 前2种方法只考虑单一因子, 实际应用中多变量格式应用较多。多变量格式方法直观、所需数据容易获取、实际操作性强, 是栖息地量化的经典方法, 也是目前应用最多的方法, 然而其对专家经验依赖较多, 主观性较强。近些年来, 多元统计方法在栖息地适宜度评价方面的应用不断增加, 它考虑物理变量之间的相互作用和相关性。模糊逻辑法在处理栖息地模拟中的不确定性方面具有优势, 能更好地利用专家知识, 更合理的处理建模过程中测量的不准确性和不确定性, 同时也考虑了多个变量之间的相互作用, 但是当考虑的变量数增加时, 模糊规则的数量会迅速增加, 给计算带来不便。人工神经网络能够隐性地找出响应变量和环境变量之间的复杂关系, 但是其解释能力不足, 并需要大量的实测数据对其进行训练, 实际应用受到限制。通过排序分析或梯度分析可以对多物种和群落进行统计分析。这些方法都各有优劣, 在实际应用中, 应结合实际情况选择最合适的栖息地适宜度评价方法。

关键词: 栖息地适宜度; 单变量; 多变量; 统计学方法; 模糊逻辑; 回归; 人工神经网络

中图分类号: TV213.4

文献标志码: A

文章编号: 1674-5906 (2013) 05-0887-07

引用格式: 易雨君, 程曦, 周静. 栖息地适宜度评价方法研究进展[J]. 生态环境学报, 2013, 22(5): 887-893.

YI Yujun, CHENG Xi, ZHOU Jing. Research progress in habitat suitability assessment methods [J]. Ecology and Environmental Sciences, 2013, 22(5): 887-893.

栖息地模拟法被认为是迄今为止生态需水计算方法中最复杂和最具科学依据的方法。栖息地模型主要应用于3个方面: 第一, 以非生物和生物变量为基础, 预测物种的出现频率; 第二, 加深对物种—栖息地关系的理解; 第三, 将物种对栖息地的要求量化^[1]。由于该方法可量化, 且基于生物本身对物理生境的选择, 被认为是最可信的生态需水评价方法^[2], 并被广泛的应用于受水利工程影响的河流生态系统的保护与修复。代表方法包括 IFIM 法 (Instream Flow Incremental Methodology), CASiMiR 法 (Computer Aided Simulation Model for Instream Flow Regulations) 等, 其中 IFIM 法应用最广。栖息地模拟法中关键的一步是建立栖息地适宜度评价指标。物理生境因子对生物种群与多样性有显著影响, 而生物群落对栖息地也有特定的敏感性与偏好。目前, 已经开发有多种栖息地统计学模型, 用于描述栖息地质量与生物之间的关系。本文将对世界上出现的主要方法进行全面而系统的总结。

1 栖息地适宜度指数

栖息地量化的经典方法是栖息地适宜度指数法 (HSI)。栖息地适宜度指数 (Habitat Suitability Index) 用来定量生物对栖息地偏好与栖息地生境因子之间的关系, 该方法由美国鱼类及野生动物署在栖息地评估程序 (HEP) 中率先提出, 并已得到广泛应用^[3]。栖息地适宜度曲线是物理栖息地特征与物种在该条件下生存质量的定量描述。该法被全球愈 90% 的鱼类栖息地模型所采用, 其中应用最广泛的 PHABSIM^[4-7]模型就采用了此种方法。

栖息地适宜度指数有 3 种格式^[8]: 二元格式 (binary format), 单变量格式 (univariate format), 多变量格式 (multivariate format)。

1.1 二元格式

二元格式为每一个变量建立一个适宜的范围, 它适合于感兴趣的某一生命阶段, 且以阶梯函数表示。假如一个变量的值位于标准范围之内, 其适宜度是 1; 任何标准范围之外的变量, 其适宜度值是 0; 这表明无论其它变量的值为多少, 一旦有一个

基金项目: 环境保护部环保公益项目(201209029-4); 国家自然科学基金项目(51279220); 国家科技支撑计划课题(2011BAC12B02)

作者简介: 易雨君 (1981 年生), 女, 副教授, 博士, 从事流域生态环境需水、水生栖息地数值模拟研究。E-mail: yiyujun@bnu.edu.cn

收稿日期: 2012-12-20

变量适宜度为零,则计算单元是不适宜的。因此,一个计算的单元能被考虑为适宜的栖息地只有所有的变量值均在适宜的范围之内。二元格式可以灵活使用,可以给一个或所有的相关变量建立二元曲线,且能在不同的物种和生命阶段的变量之间“混合和匹配”。

1.2 单变量格式

运用单变量或连续曲线可以建立每个因子的适宜范围和最佳范围。1982 年, Bovee^[5]提出使用 0 和 1 之间的数值定义鱼类栖息地的适宜性。在考虑适宜的条件范围内,鱼对某个变量的选择偏好或最佳的范围比较狭窄。这种格式将动物的行为特征表达成一系列单变量曲线而不是二元格式的台阶函数。曲线的峰值代表生物对该因子的最适宜或最喜爱的范围,曲线的端点代表了每一个因素的适宜性范围。单因素指标是目前普遍使用的格式^[9]。确定生境适宜度曲线的方法主要有 3 种:(1)专家判断法。这种方法的适宜度曲线来自专家经验或文献。该方法通常不需要野外实测数据,费用少、节省时间,但该方法不是从实测数据所推出,较大程度依赖主观判断,缺乏一定的可靠性。(2)生境利用法。该方法以测量的微生境特性频率分布为基础,直接由目标物种特定生命阶段的栖息地使用情况得出,也叫栖息地使用函数。(3)生境偏好法。基于观察频率,但同时考虑某一有限生境类型的可获取性。每个生态位通过它们在观察系统中出现的百分比加权或相乘,频率分析和生境生态位加权同时进行。

1.3 多变量格式

一般物理栖息地都依赖于多个变量,所以要将多个适宜度指数(SI)结合成为综合适宜度指数(HSI)^[10-12]。常用的计算综合适宜度指数的方法有算术平均法、几何平均法^[13-14]、乘积法^[15-16]、最小值法^[10]、加权求和法和加权乘积法^[17-18]。乘积法对单个 SI 相乘类似于将独立变量的概率相乘,只要有一个栖息地因子不适合,乘积就为零^[15]。最小值法假设限制因子决定了栖息地适宜度的上限,而具有较高 SI 值的因子不能为低 SI 值的因子补偿^[10]。算术平均的 HSI 基于假设某个变量条件较好时,可以弥补其它变量的不足。加权求和法则能体现各栖息地变量的相对重要性。加权乘积法考虑了不同因子间的补偿作用,同样在出现一个适宜度指数为零时,HSI 就为零。综合适宜度指数(HSI)确定后,将相应的综合适宜度指数与各单元面积相乘就可以计算加权可用面积(WUA)^[19]。综合适宜度指数(HSI)也可以用多元函数表达。这取决于生境因子的组合和相互关系,通常由指数为多项式的指数函

数表示。

2 多元统计方法

栖息地适宜度指数(HSI)没有考虑栖息地变量之间的相互作用和相关性。对栖息地的生境因子的独立考虑忽略了栖息地的复杂性,不能恰当的表达自然水体的情况。多元分析方法考虑了物理变量之间的相互作用,通过多个环境特征的累积效果来决定物种的响应。多元统计方法考虑环境变量间的内在相互作用和相关性,因此更适合做鱼类栖息地分析。过去 20 年中,多元统计方法在模拟物种分布和栖息地要求方面的应用不断增加,方法多样^[11, 19-24]。目前模拟物种分布和栖息地情况的多元统计方法主要包括多元线性回归法、岭回归、主成分回归、逻辑回归、广义线性模型和广义加和模型。

2.1 多元线性回归法

多元线性回归法(MLR)是用来描述一个因变量(物种丰度)和多个自变量(非生物变量)之间关系的最常用方法。利用自变量和因变量之间的联系,将一个表示物种响应的变量与单个或者多个环境变量联系起来。多元回归模型使用连续的响应数据来描述鱼类丰度和栖息地变量(如水深,底质,水温,河道宽度,以及土地利用和森林覆盖等流域尺度特征)间的关系^[11, 24-25]。

物种响应与栖息地变量之间的关系可以用下式表示:

$$Y = \beta_0 + X^T \beta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon \quad (1)$$

其中: Y 代表响应变量(如,丰度); $X = (x_1, \dots, x_m)$ 是由 m 个预测变量组成的向量; β_0 为常数; $\beta = (\beta_1, \dots, \beta_m)$ 是 m 个与预测变量相对应的回归系数组成的向量; ε 为残差。

多元线性回归可用于建立河流鱼类群落的栖息地适宜度模型,也可建立河流和湖泊中物种—栖息地关系。如果响应变量与预测变量之间的关系为非线性,可以对模型中的栖息地变量进行变换。

许多栖息地特征都彼此相关,所以,当有些预测变量之间高度相关时,对典型栖息地变量使用等式(1)就可能出现多重共线性问题。Harraway^[26]提出,两变量之间相关系数超过 0.5 就可能出现共线性问题。出现多重共线性问题时,估算出的回归系数很不稳定,标准偏差很大。在使用多元线性回归法时,需先计算自变量的两两相关性来检测多重共线性。

2.2 岭回归

在多重共线性存在的情况下,必须使模型中存在多余变量的可能性最小,同时注意不删除任何重要的变量。解决多重共线性问题的一个方法是使用

岭回归。岭回归由 Hoerl 等^[27]提出, 是一种偏估计方法, 能够对回归系数给出更稳健的估计。这些估计通常具有偏差, 但当使用在不同样本上时, 得出的估计值的均方差往往很小, 回归系数也更稳定。岭回归的原则是以增加偏差为代价减小波动性。通过这两个数之间的协调, 可以得出一个最优的预测模型^[28]。因此, 具有偏差但方差较小的估计比方差大的无偏估计更具有优势。当数据数量相对变量数目不大时, 岭回归具有优势。由于观测的数据数量往往受到限制, 岭回归的这种特性使得其在栖息地模拟中十分有用。目前, 岭回归在栖息地模拟中的应用还很少, 但当研究中的某些环境变量高度相关时可以作为一种替代方法。

2.3 主成分回归

主成分分析 (PCA) 是能用于多重共线性数据上的另一种多变量统计方法。PCA 主要是通过一系列原始变量的线性组合来解释数据集的方差-协方差结构。PCA 可以得出互不相关的“新隐含变量”, 也能有效地减少变量数目。栖息地模拟中一种利用 PCA 的方法是主成分回归。通过非生物变量的数据找出主成分, 然后用选出的主成分对响应变量进行回归。

主成分回归被用于许多水生栖息地研究: 分析支流汇合点处鱼类丰度^[29], 分析河流中鱼类丰度^[21,30], 池塘中鱼类的季节性和空间变化^[31], 和对水生栖息地进行分区^[32-33]。PCA 还能被用来检验哪些变量能够最大程度地解释数据中的变异性。在这一方法中, 需要计算主成分, 使用与前几个主成分高度相关的环境变量来模拟响应变量。如果前几个主成分能解释大部分的变异, 则与前几个主成分无相关性的环境变量就可以被剔除。

2.4 逻辑回归

多元线性回归适用于响应变量 (e.g. 物种丰度) 连续时, 但不适用于响应变量为二元变量的情况, 比如说物种“有—无”的数据。对二元响应变量进行多元线性回归无法给出值仅为 1 或 0 的变量, 反而会产生 0 和 1 之间甚至是大于 1 的值。在这种情况下, 适合的模型是逻辑回归。逻辑回归可以被用于分析一个伯努利型或二元响应变量 (适合—不适合) 和许多解释变量 (即表示栖息地质量的环境因素, 如水深、流速和底质) 之间的关系。逻辑回归允许同时分析底质、覆盖物等分类变量和水深、流速等连续性变量^[20]。在给出解释变量集合后, 该模型能预测正响应出现的概率^[34-35]。基于有—无数据, 可以绘制物种的响应曲线, 表示出物种存在的概率 p 和环境变量之间的函数关系^[36]。对逻辑回归模型的参数估计通常使用最大似然法。逻辑回归被广泛用

于研究各种水生生物的栖息地利用。

在近 30 年来, 一个非常重要的统计学进展就是将线性回归改进为广义线性模型 (GLM) 和广义加性模型 (GAM)。这两个模型由于 (由于) 能够处理多种常见的分布形态, 而在生态学研究广泛应用。

2.5 广义线性模型

广义线性模型中, 假设数据来自多种概率分布形态, 其中多数的误差都不符合正态分布。广义线性模型由 3 部分组成: 响应变量分布, 对应响应变量 Y ; 线性预测变量, 对应模型中充当预测变量的环境变量; 链接函数 g , 描述了线性预测变量和响应变量的期望值 $\mu=E(Y)$ 之间的函数关系。广义线性模型通过线性预测方程, 将一个平均值函数与环境变量联系起来^[34,37]。广义线性模型的表达式与多重线性回归模型相同:

$$g(\mu(x)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m \quad (2)$$

广义线性模型适用于仅有单一响应变量且变量符合指数族的拟合模型。指数族包括正态分布, 二项式分布, 泊松分布, 几何分布, 负二项式分布, 指数分布, 和反正态分布^[37]。上面讨论的常规线性模型和逻辑回归都是广义线性模型的特例。一般的线性回归基于正态分布假设, 并直接来模拟平均值。广义线性模型在两个方面将一般的线性回归模型推广: 首先, 它允许响应变量为非正态分布; 其次, 它可以模拟函数的平均值^[34]。

在阐述生态学数据的分布规律上, 广义线性模型作为一种有力的模型范式, 正在受到越来越广泛的关注^[28,38]。许多生态学中的广义线性模型应用都是在植物研究和植物物种分析上, 将广义线性模型用于分析水生栖息地的做法还很有限。Labonne 等^[39]用广义线性模型分析了 Beaulieu 河中的鱼类与栖息地的关系。得出广义线性模型是一种分析水生栖息地—种群关系的有力工具, 尤其是在更简单的方法所需假设无法满足时。

2.6 广义加和模型

广义加和模型是广义线性模型的非参数扩展^[40-41], 实际上也是一种广义化的多元线性回归 (MLR)。广义线性模型将经典回归的应用扩展到二项式分布、泊松分布、 γ 分布和逆二项式分布等其他统计分布形式上, 广义加和模型则使用非参数光滑函数, 而不是参数方程 ($ax + bx^2$) 来估计响应曲线。

广义加和模型基于的唯一假设是函数具有可加和性, 且各项均光滑。GAMs 与 GLMs 相似, 都是使用一个链接函数来建立响应变量平均值和预

测变量光滑函数之间的关系。广义加和模型在模型形式上更为灵活。事实上,可以将标准广义线性模型和非参数回归同时应用于预测变量上^[37,42]。GAMs的优势在于能够处理响应变量和预测变量之间的高度非线性、非单调的函数关系。与GLMs类似,这一工具可以用于处理非线性的数据结构,从而使模型更好地反映数据特性,加深对生态系统的理解。Costa等^[40]假设响应变量遵循泊松分布,用对数形的GAMs模型在中生境尺度上建立了西班牙Cabriel河濒危鱼类Júcar nase的栖息地适宜度模型,很好地模拟了鱼类的栖息地情况。该应用表明GAMs在评价鱼类栖息地方面具有应用价值。广义加和模型在生态学研究中被用来预测物种分布与环境的关系,例如模拟森林生物区系、预测随着环境梯度变化的植被分布,也被用来模拟水生栖息地。但是广义加和模型以及其他如人工神经网络等非参数方法都具有一个潜在缺陷,即无法给出模型的参数函数。当想要利用GIS建立一个空间预测模型时,这就会是一个问题。Myers等^[37]认为,仅当线性回归等简单的模型无法适当地拟合数据时,才使用广义加和模型。

3 模糊逻辑方法

生态学模拟中往往存在不确定性,此前介绍的用确定数值定义生境因子适宜度指数的方法有以下不足:(1)物种偏好不明确的情况下仍然用精确的数学表达式来描述;(2)多变量分析中输入变量相互独立,与现实不符;(3)难以在已完成的模型中加入新的附加参数。

生态学数据中内在的不确定性包括随机变量的存在、不完整或不准确的测量数据、使用估计而不是直接测量的结果。处理这种不确定性的一个方法是模糊数学。与传统方法相比,模糊逻辑能更好地利用不精确、不确定的测量结果和模糊的专家知识。首先,使用模糊集表示和处理不精确数据;其次,使用语言学规则表示和处理模糊的专家知识,其中的不精确项被定义为模糊规则。

模糊集理论是经典集合理论的扩展,其中心概念是模糊集或隶属度函数。模糊逻辑法模仿人脑的不确定性概念判断、推理思维方式,对于模型未知或不能确定的描述系统,以及强非线性、大滞后的控制对象,应用模糊集合和模糊规则进行推理,表达过渡性界限或定性知识经验,模拟人脑方式,实行模糊综合判断,推理解决常规方法难于对付的规则型模糊信息问题。模糊方法在表示栖息地模拟的不确定性时十分有用,可以利用非精确或者模糊的信息,也可以将容易获得的专家知识方便地表示为偏好数据集。这使其在栖息地模拟上具有明显优势,

因此,模糊模型是近年来栖息地模拟评价方法中广受关注以及应用较多的方法^[43-46]。要在栖息地模型中应用模糊逻辑,必须要考虑水深、底质、流速和覆盖物等不同物理变量的组合。

模糊逻辑法在栖息地模拟中的优势有:①模糊逻辑接受不精确信息的数值处理,如仅对鱼类栖息地适宜度的定性认识;②模糊逻辑法考虑了多变量影响,但不需要输入变量相互独立;③少量的野外观测数据即能够满足模糊规则的制订;④便于引入新参数,各种物理参数的组合都可以被引入栖息地模拟中,便于应用;⑤数值解决方案是看得见的(没有黑箱),并且得到最后结果的路径是可以理解的。

模糊逻辑法已经在模拟鱼类栖息地和底栖动物上得到较多应用^[47-48]。李若男等^[49]用模糊逻辑方法模拟了青狮潭水库运行影响下光倒刺鲃的栖息地变化;Mouton等^[50]用模糊逻辑法建立了比利时北部Zwalm河大头鱼的栖息地模型。模糊逻辑法也具有一定的局限性:模糊规则的数量随物理参数的增加而迅速增加;接近于人类语言的规则能给出错误的表达,即很容易被将自己作为专家的人定义。完全使用专家知识法建立模糊规则也较为困难费时,目前许多研究使用数据驱动的优化方法生成模糊规则^[45,51-52],以弥补专家知识的缺乏。

4 人工神经网络

生物栖息地质量与环境变量之间常常为复杂的非线性关系。为处理这种复杂性,使用对数、幂及指数函数各种变形能较好地改善结果,但是往往与数据不匹配^[25,53]。人工神经网络(ANNs)被引入解决这些问题^[54]。

在生态学和环境科学领域,ANNs的应用自从90年代初就有报道^[55]。在河流生态学中,前馈ANNs曾被用于模拟鱼类物种丰度,密度和鱼类种群生物量^[53-54,56]。一般结论是ANNs较适合用来预测鱼类丰度和分析水生栖息地数据的非线性关系。在所有应用中,仅需一个隐藏层就足以模拟响应变量。根据应用类型不同,隐藏层中的节点数在2~8个之间。由于单隐藏层的网络被认为可以拟合所有的连续函数而广受欢迎,同时与多层ANNs相比,计算时间也大为缩短。

尽管ANNs在生态学研究是一个有力的建模工具,研究者也常常批判其解释能力不足,是模拟生态学现象的一种“黑箱”模型^[53-54]。之所以出现这一观点,是因为输入变量对于输出变量的预测能力往往必须依赖于网络本身,与线性回归方法相比,在解释各个变量之间的相对重要性上比较困难^[54]。

该方法的优点是,它可以整合任何数量的描述物理环境的输入参数;缺点是需要大量的数据来建

立模型和进行计算。

5 多物种和群落的统计分析

以上的统计预测模型一般适合单物种与环境的关系,要同时分析多个物种,就必须对每个物种分别建立模型,需要多个模型,很不方便^[35]。因此需要其他类型的多变量统计方法来分析多个地点多个物种的数据。在生态学领域,这些方法一般被称作排序分析或梯度分析,用来量化环境变量和鱼类种群结构之间的关系。

排序分析是一种多变量统计方法,通过对物种、地点等环境实体进行排序,将相似的实体就近排列,不同的实体排列较远,就可用于分析多个地点物种出现情况的数据集和环境变量。通过分析这些实体之间的相对邻近度,可以总结出环境变量对物种的影响情况。排序分析分为直接法和间接法。间接排序分析不直接考虑环境变量,而是通过提取物种丰度指数变异性的正交主轴,来分析多地点的多物种数据。再通过定性或者回归方法,使用环境数据间接解释排序轴。主成分分析和对应分析(CA)是两种使用较广泛的间接方法。CA对群落的生态数据进行分析时,性能比主成分分析更好。与间接梯度分析方法相反,直接梯度法在计算排序轴时直接考虑环境变量。直接梯度分析方法用来检测可以被观察到的环境变量最佳解释的物种丰度分布模式^[35],并使用所提供的预测变量矩阵(即环境变量)的变化,来定量计算响应变量矩阵的变异性^[57-59]。这种特性使直接排序方法在多个物种水生栖息地建模时成为十分有用的工具。典范对应分析(CCA)和冗余分析(RDA)这两个直接梯度分析方法都被用于统计学栖息地模拟^[35]。Grift等^[59]使用典范对应分析分析了鱼苗沿环境变量梯度的空间分布,找出了一系列与物种出现高度相关的环境变量(流速,深度,温度,潮间带陆地植被和底质类型等)。

6 结论与研究展望

本文总结了目前用于分析物种—生境关系的栖息地适宜度评价方法,重点分析了统计学方法在评估栖息地适宜度时的优势和劣势。目前栖息地适宜度评价方法已有多种,栖息地定量的经典方法是栖息地适宜度指数(HSI)法。一般而言,物理栖息地依赖于水深、流速、底质、覆盖物等多个变量,必须要结合多个适宜度指数来定义一个综合指数。因此,在水生栖息地分析上,多变量方法考虑了环境变量中的相互关系和相关性,更为合适。

统计学方法在利用生境因子预测物种分布和丰度方面也得到广泛应用。在模拟水生生物物种—环境关系上已出现多种统计方法。大多数统计学模型都是基于多元线性回归和逻辑回归,这两种回归都已

经被用于分析水生生物密度(主要是鱼类)和分布。近年来模糊逻辑法发展较为迅速,成为HSI法之后的第二普遍方法,该方法在处理栖息地模拟中的不确定性方面具有优势,能更好地利用专家知识,更合理的处理建模过程中测量的不准确性和不确定性。要选择适合的物种—环境关系模型,必须考虑到研究目标和数据来源,尤其是考虑测量环境和变量类型。当环境变量很多且高度相关时(多重共线性),可以使用主成分回归,主成分回归允许使用由环境变量得到的新的正交不相关主成分来模拟物种丰度。岭回归是一种处理多重共线性数据的新方法,允许使用彼此相关的环境变量,得出的回归系数比常规的最小二乘法回归更为稳定。回归模型如广义线性模型,在处理非正态环境变量时具有优势。广义线性模型允许响应变量具有正态分布以外的分布形式。广义加性模型和人工神经网络更适合分析物种分布和环境变量之间的非线性关系。在统计学栖息地模拟预测中,人工神经网络是较有前景的方向,具有非线性,适合多变量系统,具有自适应、自组织、自学习能力,容错和数据压缩这些良好的算法特性。广义加性模型被用来将非参数光滑函数应用于回归模型中。该方法将不同的光滑函数分别应用于各预测变量,最后计算各部分的响应并相加。广义加性模型和人工神经网络的缺陷在于两种都是非参数方法,不能给出传统的数学函数,导致模型难以输出,在GIS上建立空间预测时,该缺陷十分明显。排序方法可以用来分析多物种数据,其中应用最广的是典范对应分析,属直接梯度分析,允许存在单峰响应,并直接使用环境变量来计算排序轴。

综上所述,目前栖息地适宜度评价方法较多,各有优缺点。有研究将两种方法同时应用到一个案例上,将两种方法的有效性进行对比,然而,这样的比较并不多见,主要集中在栖息地适宜度指数法与模糊逻辑法、或多元线性回归方面的比较。其他方法具体应用方面的比较还不多见。栖息地适宜度评价方法在以后的应用中,应根据具体条件,对比不同方法的优势和不足,然后选择合适的统计学方法。

参考文献:

- [1] AHMADI-NEDUSHA B, ST-HILAIRE A, BERUBE M, et al. A review of statistical methods for the evaluation of aquatic habitat suitability for instream flow assessment[J]. *River Res. Applic.*, 2006, 22: 503-523.
- [2] Yi Y J, WANG Z Y, YANG Z F. Impact of the Gezhouba and Three Gorges Dams on habitat suitability of carps in the Yangtze River[J]. *Journal of Hydrology*, 2010a, 387: 283-291.
- [3] Yi Y J, WANG Z Y, YANG Z F. Two-dimensional habitat modeling of Chinese sturgeon spawning sites[J]. *Ecological Modelling*, 2010b,

- 221(5):864-875.
- [4] BOVEE K D. Development and Evaluation of Habitat Suitability Criteria for Use in the Instream Flow Incremental Methodology[C] // U.S. Fish and Wildlife Service Biological Report. Washington: U.S. Dept. of the Interior, 1986, 86(7): 1-235.
- [5] BOVEE K D. A guide to Stream Habitat Analysis using the Instream Flow Analysis Incremental Methodology[C] // Fort Collins. U.S. Fish and Wildlife Service. Colorado: 1982.
- [6] BOCKELMANN B N, FENRICH E K, LIN B, et al. Development of an ecohydraulics model for stream and river restoration[J]. *Ecological Engineer*, 2004, 22, 227-235.
- [7] 郑超磊, 刘苏峡, 舒畅, 等. 泥曲河道内最小生态需水研究[J]. *长江流域资源与环境*, 2010, 19(3): 329-334.
- [8] 英晓明, 李凌. 河道内流量增加方法 IFIM 研究及其应用[J]. *生态学报*, 2006, 26(5): 1567-1573.
- [9] 班璇. 中华鲟产卵栖息地的生态需水量[J]. *水利学报*, 2011, 42(1): 47-55.
- [10] 易雨君, 王兆印, 陆永军. 长江中华鲟栖息地适合度模型研究[J]. *水科学进展*, 2007, 18(4): 538-543.
- [11] VADAS R L, ORTH D J. Formulation of habitat suitability models for stream fish guilds: do the standard methods work?[J]. *Transactions of the American Fisheries Society*, 2001, 130: 217-235.
- [12] 陈新军, 冯波, 许柳雄. 印度洋大眼金枪鱼栖息地指数研究及其比较[J]. *中国水产科学*, 2008, 15(2): 269-278.
- [13] 易雨君, 乐世华. 长江四大家鱼产卵场的栖息地适宜度模型方程[J]. *应用基础与工程科学学报*, 2011, 19: 117-122.
- [14] KORMAN J, PERRIN C J, LEKSTRUM T. A Guide for the Selection of Standard Methods for Quantifying Sportfish Habitat Capability and Suitability in Streams and Lakes of British Columbia[C] // Vancouver. Report to B.C. Environment Fisheries Branch. British Columbia: 1994: 5-32.
- [15] 张文鸽, 黄强, 蒋晓辉. 基于物理栖息地模拟的河道内生态流量研究[J]. *水科学进展*, 2008, 19(2): 192-197.
- [16] BEECHER H A, CALDWELL B A, DEMOND S B. Evaluation of depth and velocity preferences of juvenile coho salmon in Washington streams[J]. *North American Journal of Fisheries Management*, 2002, 22: 785-795.
- [17] LI F Q, CAI Q H, FU X P, et al. Construction of habitat suitability models (HSMs) for benthic macroinvertebrate and their applications to instream environmental flows: A case study in Xiangxi River of Three Gorges Reservoir region, China[J]. *Progress in Natural Science*, 2009, 19: 359-367.
- [18] 康鑫, 张远, 张楠, 等. 太子河洛氏鲮幼鱼栖息地适宜度评估[J]. *生态毒理学报*, 2011, 6(3): 310-320.
- [19] VISMARA R, AZZELLINO A, BOSI R, et al. Habitat suitability curves for brown trout (*Salmo trutta fario* L.) in the River Adda, northern Italy: comparing univariate and multivariate approaches[J]. *Regulated Rivers: Research and Management*, 2001, 17: 37-50.
- [20] FILIPE A F, COWX I G, COLLARES-PEREIRA M J. Spatial modelling of freshwater fish in semi-arid river systems: a tool for conservation[J]. *River Research and Applications*, 2002, 18: 123-136.
- [21] JUTILA E, AHVONEN A, JULKUNEN M. Instream and catchment characteristics affecting the occurrence and population density of brown trout, *salmo trutta* L., in forest brooks of a boreal river basin[J]. *Fisheries Management and Ecology*, 2001, 8: 501-511.
- [22] MALLET J P, LAMOUREUX P, SAGNES P, et al. Habitat preferences of European grayling in a medium size stream, the Ain river, France[J]. *Journal of Fish Biology*, 2000, 56: 1312-1326.
- [23] MERIGOUX S, DOLEDEC S, STATZNER B. Species traits in relation to habitat variability and state: neotropical juvenile fish in floodplain creeks[J]. *Freshwater Biology*, 2001, 46: 1251-1267.
- [24] YU S, LEE T. Habitat preference of the stream fish, *Sinogastromyzon Puliensis* (Homalopteridae)[J]. *Zoological Studies*, 2002, 41(2): 183-187.
- [25] BROSE S, LEK S. Relationships between environmental characteristics and the density of age-0 Eurasian perch *perca fluviatilis* in the littoral zone of a lake: a nonlinear approach[J]. *Transactions of the American Fisheries Society*, 2002, 131: 1033-1043.
- [26] HARRAWAY J. *Regression Methods Applied*[M]. Dunedin, New Zealand: University of Otago Press, 1995:180-272.
- [27] HOERL A E, KENNARD R W. Ridge regression: applications to nonorthogonal problems[J]. *Technometrics*, 1970, 12(1): 55-67.
- [28] GUISAN A, EDWARDS T C, HASTIE T. Generalized linear and generalized additive models in studies of species distributions: setting the scene[J]. *Ecological Modelling*, 2002, 157: 89-100.
- [29] FREEMAN S M, ROGERS S I. A new analytical approach to the characterization of macro-epibenthic habitats: linking species to the environment[J]. *Estuarine, Coastal and Shelf Science*, 2003, 56: 749-764.
- [30] CORBACHO C, SANCHEZ J M. Patterns of species richness and introduced species in native freshwater fish faunas of a Mediterranean-type basin: the Guadiana River (southwest Iberian Peninsula)[J]. *Regulated Rivers: Research and Management*, 2001, 17: 699-707.
- [31] ELSO J I, GILLER P S. Physical characteristics influencing the utilization of pools by brown trout in an afforested catchment in Southern Ireland[J]. *Fish Biology*, 2001, 58: 201-221.
- [32] BAIN M B. Habitat at the local scale: Multivariate patterns for stream fishes[J]. *Bulletin Francais De La Peche Et De La Pisciculture*, 1995, 337, 338, 339: 165-177.
- [33] COHEN P, ANDRIAMAHEFA H, WASSON J. Towards a regionalization of aquatic habitat: distribution of mesohabitat at the scale of a large basin[J]. *Regulated Rivers Research and Management*, 1998, 14: 391-404.
- [34] AGRESTI A. *An Introduction to Categorical Data Analysis*[M]. New York: John Wiley & Sons, 1996: 99-121.
- [35] JONGMAN R H G, TER-BRAAK C J F, TONGEREN O F R. *Data Analysis in Community and Landscape Ecology*[M]. Cambridge: Cambridge University Press, 1995: 29-72.
- [36] FLADUNG M, SCOLTEN M, THIEL R. Modelling the habitat preferences of preadult and adult fishes on the shoreline of the large, lowland Elbe River[J]. *Journal of Applied Ichthyology*, 2003, 19: 303-314.
- [37] MYERS R H, MONTGOMERY D C, VINING G G. *Generalized Linear Models with Applications in Engineering and the Sciences*[M]. New York: John Wiley & Sons, 2002: 202-264.
- [38] LE P O, BAULIER L, CLOAREC A, et al. Habitat suitability for juvenile common sole (*Solea solea*, L.) in the Bay of Biscay (France): A quantitative description using indicators based on epibenthic fauna[J]. *Journal of Sea Research*, 2007, 57: 126-136.
- [39] LABONNE J, ALLOUCHE S, GAUDIN P. Use of a generalised linear model to test habitat preferences: the example of Zingel asper, an endemic endangered percid of the River Rhône[J]. *Freshwater Biology*, 2003, 48: 687-697.
- [40] COSTA R M S, MARTÍNEZ-CAPEL F, MUÑOZ-MAS R, et al. Habitat Suitability Modeling at Mesohabitat Scale and Effects of Dam Operation on the Endangered JÚCAR NASE, *Parachondrostoma Arrigonis* (River Cabriel, Spain)[J]. *River Research and Applications*, 2012, 28: 740-752.
- [41] LEHMAN A. GIS modelling of submerged macrophyte distribution using Generalized Additive Models[J]. *Plant Ecology*, 1998, 139: 113-124.
- [42] PLEYDELL D R J, CHRETIEN S. Mixtures of GAMs for habitat suitability analysis with overdispersed presence/absence data[J]. *Computational Statistics and Data Analysis*, 2010, 54: 1405-1418.
- [43] FUKUDA S. Assessing the applicability of fuzzy neural networks for habitat preference evaluation of Japanese medaka (*Oryzias latipes*)[J]. *Ecological Informatics*, 2011, 6: 286-295.
- [44] MOUTON A M, ALCARAZ-HERNANDEZ J D, DE-BAETS B, et al. Data-driven fuzzy habitat suitability models for brown trout in Spanish Mediterranean rivers[J]. *Environmental Modelling & Software*, 2011, 26(5): 615-622.
- [45] MOUTON A M, DE-BAETS B, GOETHALS P L M. Knowledge-based versus data-driven fuzzy habitat suitability models for river management[J]. *Environmental Modeling & Software*, 2009, 24: 982-993.
- [46] MUNOZ-MAS R, MARTINEZ-CAPEL F, SCHNEIDER M, et al.

- Assessment of brown trout habitat suitability in the Jucar River Basin (SPAIN): Comparison of data-driven approaches with fuzzy-logic models and univariate suitability curves[J]. *Science of the Total Environment*, 2012, 440: 123-131.
- [47] ADRIAENSSENS V, GOETHALS P L M, PAUW N D. Fuzzy knowledge-based models for prediction of *Aesellus* and *Gammarus* in watercourses in Flanders (Belgium)[J]. *Ecological Modelling*, 2006, 195: 3-10.
- [48] KERLE F, ZÖLLNER F, SCHNEIDER M, et al. Modelling of long-term fish habitat changes in restored secondary floodplain channels of the River Rhine [C] // Fourth International Ecohydraulics Symposium. Cape Town, South Africa, 2002.
- [49] 李若男, 陈求稳, 吴世勇, 等. 模糊数学方法模拟水库运行影响下鱼类栖息地的变化[J]. *生态学报*, 2010, 30(1): 128-137.
- [50] MOUTON A M, SCHNEIDER M, DEPESTELE J, et al. Fish habitat modelling as a tool for river management[J]. *Ecological Engineering*, 2007, 29: 305-315.
- [51] BROEKHOVEN E V, ADRIAENSSENS V, BAETS B D, et al. Fuzzy rule-based macroinvertebrate habitat suitability models for running waters [J]. *Ecological Modelling*, 2006, 198: 71-84.
- [52] MOUTON A M, SCHNEIDER M, PETER A, et al. Optimisation of a fuzzy physical habitat model for spawning European grayling (*Thymallus thymallus* L.) in the Aare river (Thun, Switzerland)[J]. *Ecological Modelling*, 2008, 215: 122-32.
- [53] LEK S, GUEGAN J F. Artificial neural networks as a tool in ecological modelling, an introduction[J]. *Ecological Modelling*, 1999, 120: 65-73.
- [54] OLDEN J D, JACKSON D A. Illuminating the 'black box': understanding variable contributions in artificial neural networks[J]. *Ecological Modelling*, 2002, 154: 135-150.
- [55] LEE J H W, HUANG Y, DICKMAN M, et al. Neural network modelling of coastal algal blooms[J]. *Ecological Modelling*, 2003, 159: 179-201.
- [56] YUAN H C, SHEN X Q, CHEN X J. Prediction of Fishing Ground based on RBF Neural Network[J]. *Procedia Engineering*, 2011, 15: 3240-3244.
- [57] WANG L, LYONS J, RASMUSSEN P, et al. Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest Ecoregion, U.S.A.[J]. *Canadian Journal of Fish and Aquatic Sciences*, 2003, 60(5): 491-505.
- [58] TER-BRAAK C J F, YERDON SCHOT D F M. Canonical correspondence analysis and related multivariate methods in aquatic ecology[J]. *Aquatic Sciences*, 1995, 57: 255-289.
- [59] GRIFT R E, BUIJSE A D, VAN-DENSEN W L T, et al. Suitable habitats for 0-group fish in rehabilitated floodplains along the River Rhine[J]. *River Research and Applications*, 2003, 19: 353-374.

Research progress in habitat suitability assessment methods

YI Yujun¹, CHENG Xi¹, ZHOU Jing²

1. Ministry of Education Key Laboratory for Water and Sediment Sciences, Beijing Normal University, Beijing 100875, China;

2. Chengdu Hydroelectric Investigation & Design Institute of SPC, Chengdu 610072, China

Abstract: Based on the requirement of indicator species on physical habitat conditions, habitat suitability simulation is considered to be a most reliable method to determine the environmental flow. The purpose of evaluating habitat suitability is to provide a suitable habitat for aquatic organisms. Compared with other evaluation methods of environmental flow, habitat suitability simulation considers the species-habitat relationships, and quantifies habitat requirements by evaluation indices. This paper aims to provide an overview of the current habitat suitability assessment methods for describing the relationship between species and habitat. Methods including habitat suitability index, multivariate statistical methods, fuzzy logic method, artificial neural networks, statistical analysis of multi-species and community are introduced. The development and application of these methods are described in this paper. At the same time, the advantages and disadvantages of statistical methods are compared. Habitat suitability index is a classical method for quantifying habitat Binary format, univariate format and multivariate format are the three formats of habitat suitability index. The first two methods only consider single factor. Because of the intuitive, required less field survey data, and can be easily used, multivariate format is the most popular method at present. However, it strongly relies on the experience of expert and tends to be subjective easily. Multivariate statistical methods include multiple linear regressions, ridge regression method, principal component regression, logistic regression, generalized linear model approach, generalized additive model method, and so on. More and more applications of multivariate statistical methods in habitat suitability assessment have appeared in recent years. Multivariate statistical methods considered the interaction between the physical factors. The fuzzy logic method is better dealing with uncertainties in the process of habitat simulation. It uses expert knowledge and dispose inaccuracies and uncertainties of measurement in the modeling process. Fuzzy logic method takes into account the interactions between multiple variables, but the number of fuzzy rules will increase rapidly when the number of variables increases. Artificial neural network presents advantages in identifying complicated relationship between response variables and environment variables implicitly, but it suffers from inadequate explanatory power. Statistical analysis of multi-species and community can be conducted by sequencing analysis or gradient analysis. These methods both have their advantages and disadvantages. The most suitable habitat suitability evaluation method should be chosen by combining actual conditions.

Key words: habitat suitability; univariate method; multivariate method; statistical method; fuzzy logic; regression; artificial neural networks