

青海省生态脆弱度分级 1-km 栅格数据集 (2015) 的研发与内容

支泽民¹, 刘峰贵^{1,2*}, 陈 琼^{1,2}, 周 强^{1,2}

1. 青海师范大学地理科学学院, 西宁 810008;

2. 高原科学与可持续发展研究院, 西宁 810008

摘 要: 青海省作为青藏高原的重要组成部分, 对于我国生态安全屏障构建具有十分重要的作用, 生态战略意义重大。本数据利用“SPR”(生态敏感性—生态压力—生态恢复力)模型, 基于生态敏感性、生态压力以及生态恢复力 3 个准则, 选取坡度、海拔、地形起伏度、土壤侵蚀敏感性、土地退化敏感性、NDVI 等共 12 个指标构建生态脆弱度评价体系, 其中原始数据中坡度、海拔与地形起伏度分辨率为 30 m×30 m, 土壤侵蚀敏感性、土地退化敏感性、NDVI、年平均气温、年平均降水、人口密度、GDP 密度、道路密度、放牧强度数据分辨率 1 km×1 km, 基于随机森林模型进行权重确定对青海省生态脆弱度进行评估。最后, 将研究区分为微度、轻度、中度、重度、极度五个脆弱度等级, 最终数据分辨率为 1 km×1 km, 数据格式为.tif 格式, 由 5 个文件组成, 数据大小为 1.30 MB。

关键词: 青海省; 生态脆弱性; 随机森林; SPR 模型

DOI: <https://doi.org/10.3974/geodp.2022.03.11>

CSTR: <https://cstr.escience.org.cn/CSTR:20146.14.2022.03.11>

数据可用性声明:

本文关联实体数据集已在《全球变化数据仓储电子杂志(中英文)》出版, 可获取:

<https://doi.org/10.3974/geodb.2022.01.04.V1> 或 <https://cstr.escience.org.cn/CSTR:20146.11.2022.01.04.V1>。

1 前言

近年来, 随着全球气候变化的加剧与社会经济的飞速发展, 人类对于自然环境的改造作用迅速增强, 对于生态环境的影响越来越大, 人与自然之间的关系日趋紧张。在此背景下, 对于生态环境的脆弱性评估、生态环境脆弱管理研究逐渐受到生态学领域学者重视^[1]。目前, 对于生态脆弱性概念的界定较多, 但公认的生态脆弱性主要指生态系统内部对外部干扰的响应^[2], 在特定的时间和空间尺度上, 景观系统或者生态系统在外部作用条件下所

收稿日期: 2022-02-10; 修订日期: 2022-06-06; 出版日期: 2022-09-25

基金项目: 中华人民共和国科学技术部(2019YFA0606902)

*通讯作者: 刘峰贵 L-8795-2018, 青海师范大学高原科学与可持续发展研究院, lfg_918@163.com

数据引用方式: [1] 支泽民, 刘峰贵, 陈琼等. 青海省生态脆弱度分级 1-km 栅格数据集(2015)的研发与内容[J]. 全球变化数据学报, 2022, 6(3): 409–415. <https://doi.org/10.3974/geodp.2022.03.11>. <https://cstr.escience.org.cn/CSTR:20146.14.2022.03.11>.

[2] 支泽民, 刘峰贵, 陈琼等. 青海省生态脆弱度分级 1-km 栅格数据集(2015) [J/DB/OL]. 全球变化数据仓储电子杂志, 2022. <https://doi.org/10.3974/geodb.2022.01.04.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2022.01.04.V1>.

表现出的敏感反应和恢复状态，是系统本身所具有的属性在受到干扰作用时的表现^[3]。对于区域生态脆弱性进行评估可为区域发展提供有效参考，推动区域经济健康可持续发展。

青海省位于我国西部，是青藏高原的重要组成部分，属青藏高寒区，是长江、黄河以及澜沧江三大河流的发源地，生态战略地位十分重要。自 2017 年以来，青海省坚持“生态立省”战略，打造“生态文明高地”目的明确。因此，对于青海省全域的生态脆弱性评估就具有十分重要的意义。本数据集基于“SPR”（生态敏感性-生态压力-生态恢复力）模型，从生态敏感性、生态压力、生态恢复力三个层面，分别选取坡度、海拔、地形起伏度、土壤侵蚀敏感性、土地退化敏感性、NDVI、年平均气温、年平均降水、人口密度、GDP 密度、道路密度、放牧强度共 12 个指标构建青海省生态脆弱度评估指标体系，基于随机森林模型分别确定各分目标层及总目标权重，通过 ArcGIS 空间叠加分析得到青海省生态脆弱度空间分布数据集，该数据集可有效厘清青海省生态脆弱度的空间分布状况，为青海省区域经济发展方向提供参考依据。

2 数据集元数据简介

《青海省生态脆弱度分级 1-km 栅格数据集（2015）》^[4]的名称、作者、地理区域、数据年代、时间分辨率、空间分辨率、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表 1。

表 1 《青海省生态脆弱度分级 1-km 栅格数据集》元数据简表

| 条 目 | 描 述 |
|-----------|---|
| 数据集名称 | 青海省生态脆弱度分级 1-km 栅格数据集（2015） |
| 数据集短名 | EcoVul.Qinghai_2015 |
| 作者信息 | 支泽民 ABG-7100-2020，青海师范大学地理科学学院，zhizemin@126.com 刘峰贵 L-8795-2018，青海师范大学地理科学学院，lfg_918@163.com 陈 琼 AAB-3346-2021，青海师范大学地理科学学院，qhchenqiong@163.com 周 强 AAB-3351-2021，青海师范大学地理科学学院，zhouqiang729@163.com |
| 地理区域 | 青海省 |
| 数据年代 | 2015 年 |
| 空间分辨率 | 1 km×1 km |
| 数据格式 | .tif |
| 数据量 | 1.30 MB |
| 基金项目 | 中华人民共和国科学技术部（2019YFA0606902） |
| 数据计算环境 | R for Windows 3.3.3 环境下的 R Studio, ArcGIS 10.2 |
| 出版与共享服务平台 | 全球变化科学研究数据出版系统 http://www.geodoi.ac.cn |
| 地址 | 北京市朝阳区大屯路甲 11 号 100101，中国科学院地理科学与资源研究所 |
| 数据共享政策 | 全球变化科学研究数据出版系统的“数据”包括元数据（中英文）、通过《全球变化数据仓储电子杂志（中英文）》发表的实体数据集和通过《全球变化数据学报（中英文）》发表的数据论文。其共享政策如下：（1）“数据”以最便利的方式通过互联网系统免费向全社会开放，用户免费浏览、免费下载；（2）最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源；（3）增值服务用户或以任何形式散发和传播（包括通过计算机服务器）“数据”的用户需要与《全球变化数据学报（中英文）》编辑部签署书面协议，获得许可；（4）摘取“数据”中的部分记录创作新数据的作者需要遵循 10%引用原则，即从本数据集中摘取的数据记录少于新数据集总记录量的 10%，同时需要对摘取的数据记录标注数据来源 ^[5] |

3 数据研发方法

青海省生态脆弱度分级数据主要以“SPR”模型搭建指标体系, 主要包括生态敏感性指标层、生态压力指标层以及生态恢复力指标层, 根据前人研究, 同时结合青海省实际并考虑数据的可获得性, 选取坡度、海拔、地形起伏度、土壤侵蚀敏感性、土地退化敏感性、NDVI、年平均气温、年平均降水、人口密度、GDP 密度、道路密度、放牧强度共 12 个指标构建生态脆弱性评估体系^[6-9], 采用随机森林对评估指标进行权重赋值, 后利用 ArcGIS 空间叠加分析对指标数据进行叠加得到青海省生态脆弱性分布数据, 最后参考前人于青藏高原高寒区所做研究, 利用 ArcGIS 自然断点法将脆弱性划分为微度、轻度、中度、重度、极度五个等级^[10]。

3.1 研究区概况

青海省 (31°36'N–39°19'N, 89°35'E–103°04'E) 位于我国西部, 青藏高原东北部。地势呈西部地区较高东部较低, 南北部较高中部较低的地形特征, 整体呈梯形下降的趋势。东部地区由黄土高原向青藏高原过渡, 东北部为青海省重要的农业区, 往南地形复杂多样, 80%以上区域为高原, 多山; 西部主要地貌为高原与盆地, 我国四大盆地之一的柴达木盆地即位于青海省西部, 全省整体属于大陆性气候, 区域内河流众多, 包括黄河、澜沧江、长江、湟水河以及黑河 5 大水系, 同时是黄河、长江以及澜沧江的发源地。青海省分布有多个国家级保护区, 包括“三江源”自然保护区、可可西里自然保护区、柴达木梭梭林自然保护区等。因此, 受其地理位置、气候等自然因素的影响, 青海省生态环境及其脆弱。又由于其生态战略地位的重要性, 青海省的生态安全直接影响到全国的生态安全, 对青海省的生态脆弱性评估就变得十分重要。

3.2 算法原理

本文主要参考前人所开发的生态脆弱性评估模型^[2,3], 对模型中涉及的评估指标进行本地化处理, 基于随机森林模型对各准则层指标分别进行权重计算, 利用 ArcGIS 空间叠加分析工具将指标层、准则层依次叠加, 得到青海省生态脆弱度空间分布数据集。

3.2.1 SPR 模型

SPR 模型由乔青等人于 2008 年提出^[11], 该模型主要用于对某一地区生态脆弱性进行综合评估。在一定区域内的某一特定时间段, 系统内部的框架结构具有不稳定性, 这种不稳定性在受到外界干扰时表现为敏感性, 并且由于没有足够的应对能力而向不利于系统稳定的方向发展, 系统内部受压力的作用会呈现出恢复力。SPR 模型包含生态敏感性、生态压力以及生态恢复力 3 个准则层, 其指标的选取可根据区域的实际特征确定, 其基本运行方法为: 基于研究区实际环境特征, 对 3 个准则层下的模型指标进行本地化, 通过对不同指标进行权重赋值, 对不同指标基于权重值进行空间叠加, 得到区域生态脆弱性指数, 并依据分类方法进行脆弱性分级。

生态脆弱度由生态脆弱性指数 (ecological fragility index, EFI) 分级得到, EFI 由生态敏感性指数 (ecological sensitivity index, ESI)、生态压力指数 (ecological stress index, EPI) 以生态恢复力指数 (ecological resilience index, ERI) 计算得到, 其中, ESI、EPI 为正向指标, ERI 为负向指标, 其计算公式为:

$$EFI=ESI\cdot Ws+EPI\cdot Wp-ERI\cdot Wr$$

(1)

式中， Ws 代表生态敏感性的权重， Wp 代表生态压力的权重， Wr 代表生态恢复力的权重。

3.2.2 随机森林模型

随机森林 (random forests, RF) 模型作为一种较为常用的机器学习方法，属于组合式的深度学习方法^[12]，对于解决非线性问题具有较大的优势^[13]，作为一种分类回归树的数据挖掘方法^[12]，其原理是通过集成多个决策树，分类投票获取类别。同时，该方法对于运算数据的要求较低，对于多要素类的问题研究，可定量提供要素对于结果的重要性程度，相对更具有客观性，准确性也较高^[14]。与其他机器学习算法相比，RF 具有一定的抗过拟合能力，即其可通过对决策树进行平均，降低过拟合风险，对于模型运算中产生的噪声与异常值具有较高的容忍度^[14,15]，能够度量变量的重要性程度，该特点对于权重类问题的解决具有极大的优势。因此，本文使用 RF 对生态脆弱性进行评价，主要使用 R 3.3.3 环境下的 R studio 编译器调用 RF 软件包进行迭代运算。以随机森林提供的各因子的重要性程度为指标的权重，利用 ArcGIS 空间叠加分析对处理后的数据进行叠加得到青海省生态脆弱度分级数据。

3.3 技术路线

本数据主要通过选取坡度、海拔、地形起伏度、土壤侵蚀敏感性、土地退化敏感性、NDVI、年平均气温、年平均降水、人口密度、GDP 密度、道路密度、放牧强度共 12 个指标分别构建生态敏感性、生态压力以及生态恢复力三个分目标，基于随机森林模型对分目标进行权重确定，利用 ArcGIS 空间分析工具分别计算生态敏感性、生态压力以及恢复力三个分目标，后再次使用随机森林模型确定三个分目标对于主目标的权重，利用 ArcGIS 空间叠加分析得到青海省生态脆弱性指数空间分布，基于 ArcGIS 自然裂点分类方法将其划分为极度、重度、中度、轻度、微度五个等级（图 1）。

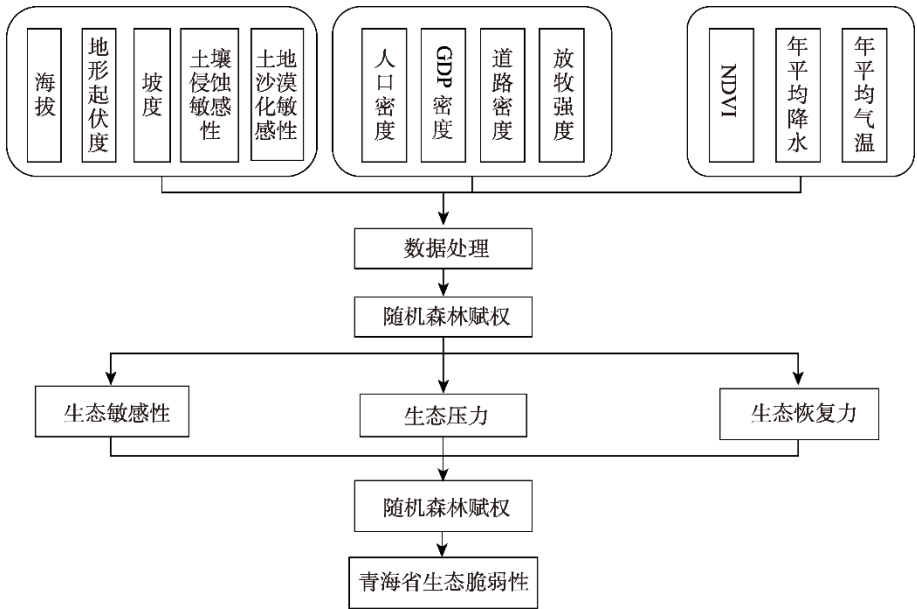


图 1 数据研发流程图

4 数据结果与验证

4.1 数据集组成

本数据的原始数据包括：青海省行政区划数据、道路数据（矢量）来源于全国地理资源目录服务系统，道路密度数据（栅格）基于 ArcGIS “密度分析”工具制备，数据空间分辨率为 1 km×1 km；年末存栏牲畜数量数据来源于青海省各市（州）统计年鉴以及国民经济与社会发展统计公报（2016 年），将放牧强度定义为单位草地面积承载的牲畜数量；地形特征数据主要包括海拔、坡度以及地形起伏度，海拔数据来源于地理空间数据云平台，坡度与地形起伏度通过海拔数据提取获得，数据分辨率为 30 m×30 m，通过“重采样”采样为 1 km×1 km；土壤侵蚀敏感性与土地沙漠化敏感性数据来源于中科院生态环境研究中心，数据分辨率为 1 km×1 km，数据时间为 2015 年；归一化植被指数（NDVI）数据、年均降水、年均气温数据、人口密度数据以及 GDP 密度数据来源于中科院资源环境科学数据中心，数据分辨率均为 1 km×1 km，数据时间均为 2015 年。

4.2 数据结果

青海省生态脆弱度分级评估数据定义为 2015 年，空间分布为青海省全省。由脆弱度分级评估结果可得，青海省生态脆弱度分级以轻度脆弱占比最高，主要分布于青海省南部青南高原以及西部唐古拉山地区，其面积占比达到青海省面积的 36.33%；极度脆弱区占比最小，主要分布于东部河湟谷地农业区以及柴达木盆地，河湟谷地作为青海省人类活动最为强烈的区域，人为干扰对于生态环境的影响较大，生态压力较为突出，其次，柴达木盆地自然环境恶劣，全年降水短缺，沙漠化严重，也成为了极度脆弱的生态区，总体上极度脆弱区面积占比为 8.51%；其余微度脆弱区与中度脆弱区面积占比相差较小，其分别占比 19.69%、21.43%；重度脆弱区面积占比相对较小，为 14.05%^[16]（图 2）。总体来看，青海省生态脆弱度整体呈现北高南低，东高西低的分布特征。东部地区人口密度相对较大，人类活动强度较高，人地关系矛盾突出，系统所承受生态压力较大，脆弱性等级较高；西部柴达木盆地为青海省甚至是为全国重要的能源供应基地，GDP 密度相对较高，且环境较为恶劣，脆弱度等级较高；南部地区主要为牧业区，人口分布较少，自然环境相对较好，脆弱度相对较低。

4.3 数据结果验证

对于数据的验证主要以随机森林模型的整体精度为主，在模型迭代运算中的生态压力、生态恢复力以及生态敏感性三个分目标，和最终的生态脆弱性目标计算中，随机森林模型 R^2 均在 0.95 以上，误差总体小于 0.05，因此模型整体精度较高，数据结果具有一定的可靠性。

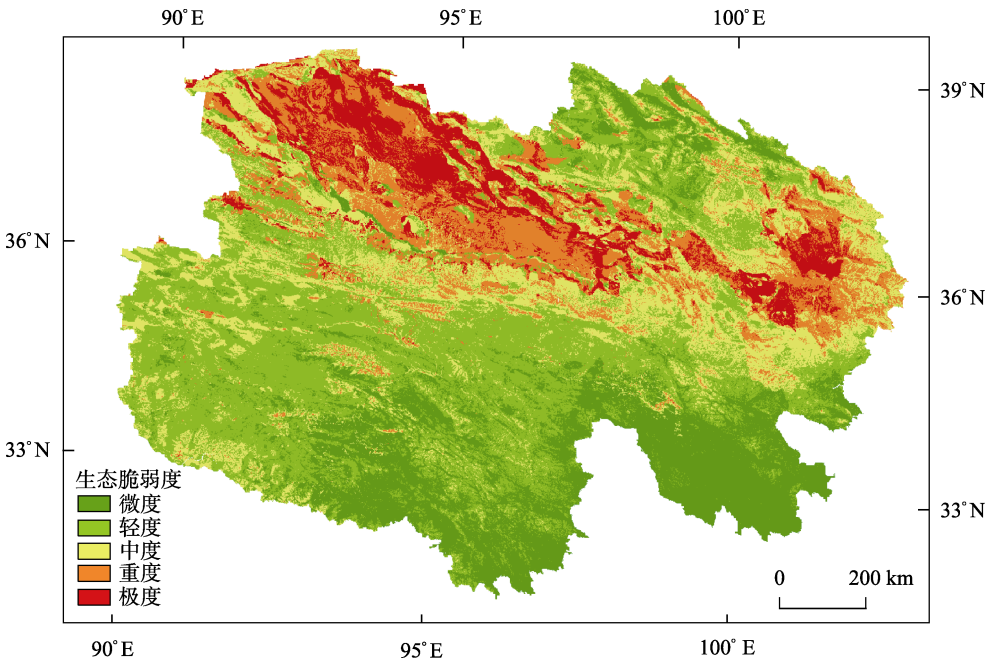


图 2 青海省生态脆弱度分级图

5 讨论与总结

为构建青海省生态脆弱度分级数据，基于生态敏感性、生态压力、生态恢复力三个指标层，选取海拔、坡度、地形起伏度、土壤侵蚀敏感性、土地沙漠化敏感性、NDVI、年平均气温、年平均降水、人口密度、GDP 密度、道路密度、放牧强度构建生态脆弱性评价指标体系，利用随机森林为各指标体系赋予权重，通过 ArcGIS 空间分析对指标进行叠加得到青海省生态脆弱性分布数据。总体来看，随着生态脆弱性评价研究的深入，脆弱性评价指标的选取逐渐成熟，但受制于某些指标的可获取性，本数据仅计算了 2015 年青海省生态脆弱度的空间分布特征。在未来的研究中，对于某些获取性较差的指标可利用其他相近指标进行替代，或通过一定的方法进行制备，从而使得研究能够具有一定的时效性。

作者分工：刘峰贵首先提出了论文的思路与想法；陈琼、周强对论文总体框架进行调整；支泽民对数据集的开发做了总体设计并采集和处理了原始底层数据，撰写了数据论文。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献

[1] 田亚平, 常昊. 中国生态脆弱性研究进展的文献计量分析[J]. 地理学报, 2012, 67(11): 1515–1525.

[2] 何彦龙, 袁一鸣, 王腾等. 基于 GIS 的长江口海域生态系统脆弱性综合评价[J]. 生态学报, 2019, 39(11): 3918–3925.

[3] 王让会, 樊自立. 干旱区内陆河流域生态脆弱性评价——以新疆塔里木河流域为例[J]. 生态学杂志,

- 2001, 20(3): 63–68.
- [4] 支泽民, 陈琼, 周强等. 青海省生态脆弱度分级 1-km 栅格数据集(2015)[J/DB/OL]. 全球变化数据仓储电子杂志, 2022. <https://doi.org/10.3974/geodb.2022.01.04.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2022.01.04.V1>.
- [5] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策[OL]. <https://doi.org/10.3974/dp.policy.2014.05> (2017 年更新).
- [6] 薛联青, 王晶, 魏光辉. 基于 PSR 模型的塔里木河流域生态脆弱性评价[J]. 河海大学学报(自然科学版), 2019, 47(1): 13–19.
- [7] 刘正佳, 于兴修, 李蕾等. 基于 SRP 概念模型的沂蒙山区生态环境脆弱性评价[J]. 应用生态学报, 2011, 22(8): 2084–2090.
- [8] 陆海燕, 孙桂丽, 李路等. 基于 VSD 模型的新疆生态脆弱性评价[J]. 新疆农业科学, 2020, 57(2): 292–302.
- [9] 张启, 李明玉. 基于 VSD 模型的区域生态脆弱性评价——以延边朝鲜族自治州为例[J]. 延边大学农学学报, 2018, 40(4): 7–15.
- [10] 于伯华, 吕昌河. 青藏高原高寒区生态脆弱性评价[J]. 地理研究, 2011, 30(12): 2289–2295.
- [11] 乔青, 高吉喜, 王维等. 生态脆弱性综合评价方法与应用[J]. 环境科学研究, 2008, 130(5): 117–123.
- [12] Breiman, L. Random Forests [J]. *Machine Learning*, 2001, 45: 5–14.
- [13] Reichstein, M., Camps-Valls, G., Stevens, B., *et al.* Deep learning and process understanding for data-driven Earth system science [J]. *Nature*, 2019, 566(7743): 195–204.
- [14] 方匡南, 吴见彬, 朱建平. 随机森林方法研究综述[J]. 统计与信息论坛, 2011, 26(3): 32–38.
- [15] 廖小平, 徐风光, 蔡旭东等. 香丽高速公路边坡地质灾害发育特征与易发性区划[J]. 中国地质灾害与防治学报, 2021, 32(5): 121–129.
- [16] 支泽民, 陈琼, 刘飞. 基于 SPR 模型的生态脆弱性评估——以青海省为例[J]. 青海环境, 2020, 30(4): 186–190.