

我国居民健康水平的地区分布差异和影响因素分析

王锴山*, 张强, 张珏

上海工程技术大学管理学院, 上海
Email: *wsjwzz18@163.com

收稿日期: 2021年8月22日; 录用日期: 2021年9月12日; 发布日期: 2021年9月24日

摘要

本文以我国34个省级行政单位为研究对象,使用2019年的数据,运用Arcgis中的空间自相关、逐步回归、加权地理回归模型等分析方法研究了我国健康水平的空间分布差异和影响因素。得出结果表明:1)我国人口健康水平整体存在地区分布差异:东部最高、中部次之、西部最低,健康水平在全国范围内的公平性较差;2)影响我国人口健康水平的因素有经济、教育、医疗、环境和社会等,其中地区卫生机构数、平均受教育年限、人均GDP、城镇化率和植被覆盖率对人口健康水平有显著正向影响,SO₂排放量对人口健康水平有显著负向影响;3)环境、教育、经济和社会因素对人口健康水平的影响具有显著的空间异质性。人均GDP对人口健康水平的正向影响呈现东强西弱的梯度分异,卫生机构数、平均受教育年限、城镇化率和植被覆盖率对人口健康水平的正向影响则呈现西强东弱的梯度分异。

关键词

健康水平, 熵值法, 空间分布差异, 加权地理回归

Analysis on the Regional Distribution Differences and Influencing Factors of the Health Level of Chinese Residents

Kaishan Wang*, Qiang Zhang, Jue Zhang

School of Management, Shanghai University of Engineering Science, Shanghai
Email: *wsjwzz18@163.com

Received: Aug. 22nd, 2021; accepted: Sep. 12th, 2021; published: Sep. 24th, 2021

*通讯作者。

文章引用: 王锴山, 张强, 张珏. 我国居民健康水平的地区分布差异和影响因素分析[J]. 应用数学进展, 2021, 10(9): 3207-3216. DOI: 10.12677/aam.2021.109335

Abstract

This article takes 34 provincial administrative units in my country as the research object, uses the data of 2019, and uses the spatial autocorrelation, stepwise regression, weighted geographic regression model and other analytical methods in Arcgis to study the spatial distribution difference and influence of my country's health level factor. The results show that: 1) the overall health level of the population in China has regional distribution differences: the eastern part is the highest, the central part is the second, and the western part is the lowest. The fairness of the health level across the country is poor; 2) factors affecting the health level of the Chinese population include economy and education, medical care, environment and society, among which the number of regional health institutions, average years of education, per capita GDP, urbanization rate and vegetation coverage have a significant positive impact on the health of the population, and SO₂ emissions have a significant negative impact on the health of the population; 3) environmental, educational, economic and social factors have significant spatial heterogeneity in the impact of population health. The positive impact of per capita GDP on the health of the population presents a gradient differentiation from the east to the weak from the east. The positive impacts of the number of health institutions, average years of education, urbanization rate and vegetation coverage on the health of the population show a gradient differentiation from the west to the weak from the east.

Keywords

Health Level, Entropy Method, Spatial Distribution Difference, Weighted Geographic Regression

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在 2019 年 6 月 24 日, 国务院印发了《国务院关于实施健康中国行动的意见》, 明确指出要加快推动以人民健康为中心的建设, 实施健康中国战略, 提高全民健康水平。人口健康研究对于促进健康学科的发展, 以及改革和重建我国的健康保障制度具有重大意义。世卫组织将健康定义为“身体、心里和社会适应方面的完好状态”。我们的研究应该从身体、心理和社会三个维度来衡量健康水平, 但总结国内外学者的研究, 考虑到数据的可获得性, 研究绝大多数集中在身体健康方面。目前, WHO 和我国的卫生健康管理部门以预期寿命、死亡率和疾病发病率来衡量身体健康水平[1]。参照前人的研究, 本文中的健康水平用居民预期寿命和死亡率相结合的方法来测量。

改革开放 40 年来, 我国人口健康水平持续提高, 但同时, 人口健康仍然具有显著的区域差异。缩小人口健康的区域差异是“健康中国”建设的重要目标, 也是全面建成小康社会的重要基础。在人口健康的影响因素方面, 已有研究认为人口健康既受自然因素(包括气候变化、环境污染等)的影响, 也受社会人文因素(包括经济发展、文化教育、医疗保健等)的影响, 其中经济、教育、医疗等因素对人口健康有促进作用, 而环境污染等因素则不利于人口健康, 这些研究采用的研究方法主要有主成分回归、加权最小二乘法、固定效应模型等[2]。但这些研究存在以下不足: 一是健康影响因素指标体系不够健全, 多是从一个或少数几个方面, 未从经济、教育、医疗、自然条件、环境污染等方面构建系统、全面的指标体系; 二是多为各因素对健康水平整体影响的分析, 缺少使用理加权回归方法分析各因素对健康水平影响的空

间差异。

本文以我国 2019 年的 34 个省际单元为研究对象，采用空间自相关方法分析我国人口健康水平的空间格局，利用逐步回归方法分析各因素对人口健康水平的整体影响，利用地理加权回归方法分析各因素对人口健康水平影响的空间差异，希望为促进全国范围的可持续发展和“健康中国”战略的实施提供科学依据。

2. 数据来源与研究方法

2.1. 数据来源

2.1.1. 人口健康数据

根据 WHO 对于健康水平的定义和测度，根据数据的可获得性，本文采用预期寿命、婴儿死亡率和孕产妇死亡率等数据来衡量健康水平。数据来自于中国统计年鉴、全国人口普查数据和国家医疗保障局统计数据，时间 2019 年。

2.1.2. 影响因素数据

代表经济水平的人均 GDP 数据来自于《中国经济统计年鉴》；代表教育水平的地区平均受教育年限数据来自于《中国人口普查资料》；反映地区医疗水平的千人医疗机构数、医疗卫生床位数、卫生人员数等数据来自于《中国统计年鉴》；代表环境水平的植被覆盖率、SO₂ 工业废气排放总量等数据来自于《中国环境统计年鉴》；反映社会水平的城镇化率数据来自于《中国城市统计年鉴》，时间 2019 年。

2.1.3. 地图数据

中国行政区图通过国家基础地理信息中心 2021 最新地图获得，全国共有 34 个省级行政单位。

2.2. 指标选取

综合以往的研究[3]，本文全国各地区的健康水平用预期寿命、婴儿死亡率和孕产妇死亡率来测量。对于健康水平的影响因素，可以从经济、教育、医疗、社会和环境五个维度进行综合考量。

通过对经济、教育、医疗、社会和环境等五个维度[4]的分析，结合数据的可获得性，选取了八项指标来分析其对我国人口健康水平的影响。具体如表 1：

Table 1. Influencing factors of my country's population health level

表 1. 我国人口健康水平的影响因素

维度	指标	单位
经济	人均 GDP	元
教育	平均受教育年限	年
医疗	千人医疗机构数	个
环境	工业废气排放总量	t
	植被覆盖率	%
社会	城镇化率	%

2.3. 研究方法

2.3.1. 全国各地区健康水平的测量方法

本文使用预期寿命、婴儿死亡率和孕产妇死亡率三项数据来测度健康水平。这些指标可以较为准确

地反映一个地区的健康水平收到地区经济、文化教育、医疗保健、环境和社会等因素的影响情况, 是国际上通用的计算健康水平的通用方法[5]。第一步, 使用熵值法计算三项指标的权重; 第二步加权求和求出地区健康水平。

① 数据标准化: 对三项数据的量纲、数量级和正负方向进行标准化处理。

正向指标标准化的公式:

$$X_{ij} = \frac{X_{ij} - X_{j\min}}{X_{j\max} - X_{j\min}} \quad (1)$$

负向指标标准化的公式:

$$X_{ij} = \frac{X_{j\max} - X_{ij}}{X_{j\max} - X_{j\min}} \quad (2)$$

② 第 i 地区的第 j 指标比重:

$$Y_{ij} = \frac{X_{ij}}{\sum_{i=1}^m X_{ij}} \quad (3)$$

③ 计算熵值:

$$E_j = \frac{\sum_{i=1}^m (Y_{ij} \ln Y_{ij})}{\ln m} \quad (4)$$

④ 计算指标权重:

$$W_j = \frac{1 - E_j}{\sum_{j=1}^n (1 - E_j)} \quad (5)$$

⑤ 加权计算地区健康水平:

$$H_{ij} = \sum_{j=1}^3 X_{ij} W_j \quad (6)$$

在上述公式中, m 为地区数量(34 个), E_j 为 j 指标的熵值, X_{ij} 为 i 地区的 j 指标标准化后的取值, W_j 为 j 指标指标的权重, H_{ij} 为 i 地区 j 指标的健康水平。

其中用熵值法计算出来的预期寿命、孕妇死亡率和婴儿死亡率权重分别为: 0.55、0.20、0.25。

2.3.2. 人口健康水平空间格局的测量方法

① 全局空间自相关。全局空间自相关是空间计量经济学中检验空间关系的重要方法[6], 可以利用 Moran's I 的结果来判断空间统计上具有离散或者集聚现象, 计算公式为:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n (X_i - \bar{X})(X_j - \bar{X})}{S^2 \sum_{i=1}^n \sum_{j=1}^n W_{ij}} \quad (7)$$

式中: n 为研究的空间单元数; X_i 、 X_j 为第 i 、 j 区域的属性值, \bar{X} 为区域内各个空间单元的均值, S^2 为方差, W_{ij} 为空间权重矩阵, 采用车相邻, 相邻取值为 1, 不相邻取值为 0。

② 局部空间自相关。局部空间自相关，是反映不同要素指标在局部呈现的空间关系，本文采用 Getis-Ord G_i^* 来衡量局部空间自相关，也可以识别出不同地区的高值区与低值区的空间分布，其效果类似于 Arcgis 中的冷热点分析。计算公式如下：

$$G_i^*(d) = \frac{\sum_{j=1}^n W_{ij}(d) X_j}{\sum_{j=1}^n X_j} \quad (8)$$

2.3.3. 人口健康水平影响因素的测量方法

综合考虑各个因素对于健康水平的影响，本文构建了健康水平决定模型。如下：

$$H_i = f(E_i, T_i, M_i, N_i, S_i)$$

其中 H_i 代表 i 地区的健康水平， E_i , T_i , M_i , N_i , S_i 分别代表经济、教育、医疗、环境和社会五个维度的影响因素。

① 地理加权回归模型

普通最小二乘法回归它可以表明多个自变量对一个因变量的影响强度，但是体现的是平均值影响或者全局估计，不能体现不同区域空间中的数据非稳定性和差异性[7]，所以本文建立空间异质性模型 GWR 模型，这可以更好的解释我国健康水平的空间差异以及产生差异的影响因素。加权地理回归对每一个观测的空间单元都进行局回归，考虑地理位置的不同，对空间上每个参数进行估计，研究变量之间的关系，更好地反映变量之间的空间依赖性和空间差异的变化情况(方远平等, 2012)。GWR 模型引入地理区位坐标，具体结构表示为：

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1, n} \beta_k(u_i, v_i) x_{ik} + \varepsilon_i$$

空间权重矩阵是 GWR 模型的核心，空间权函数的选取对于模型结果的影响很大。本文采用高斯函数法，保证空间权函数的连续性：

$$W_{ij} = \exp \left[- \left[\frac{d_{ij}}{b} \right]^2 \right]$$

国际上普遍采用 Bowan 提出的交叉确认方法(CV)来确定带宽 b ，Fotheringasm 提出：“当 CV 达到最小值，对应的就是所需的最佳带宽。”

$$CV = \sum_{i=1}^n [y_i - \hat{y}_i(b)]^2$$

3. 实证结果分析

3.1. 人口健康水平的空间格局

为了直观的反映全国各个地区人口健康水平的空间分布情况，使用 Arcgis 中的自然断裂法将所有省级行政单位分为五个级别[8]，绘制出健康水平的空间分布图(如图 1)。可以看出，全国的人口健康水平呈现空间分布差异，总体上来看，人口健康水平高的地区为北京、上海、山东、浙江、江苏、广东、海南、香港和澳门等沿海省份或行政单位。整体的趋势从东到西，健康水平逐渐降低。说明全国各个地区人口健康水平存在显著的区域差异，健康公平性也有待改善。

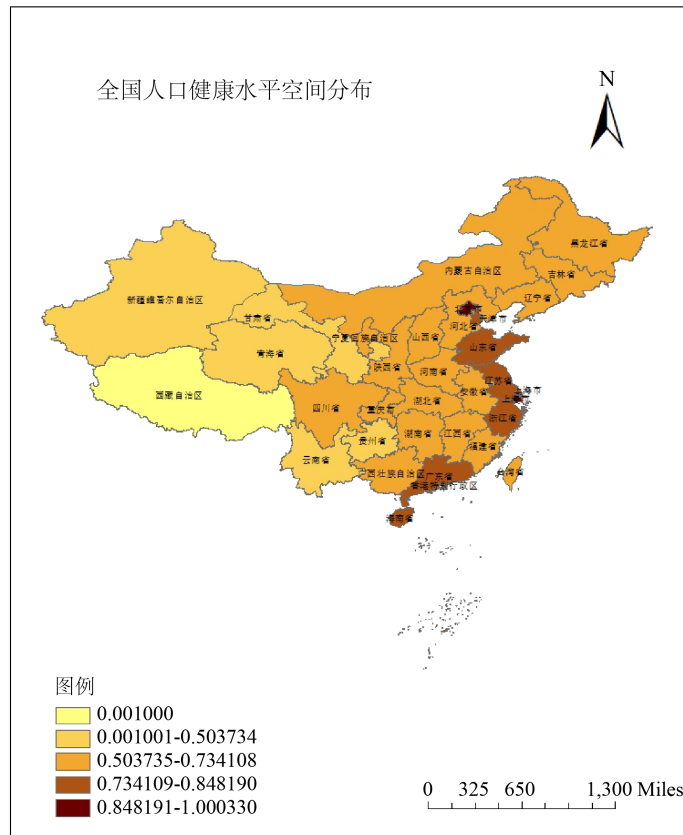


Figure 1. Spatial distribution of national population health level
图 1. 全国人口健康水平的空间分布

为更准确地识别我国各个地区人口健康水平的空间集聚特征，采用空间自相关方法进行分析。如图 2，全局空间自相关分析表明，全都通过了显著性检验。这足以说明全国人口健康水平存在空间集聚，即人口健康水平高的地级单元聚集，人口健康水平低的地级单元相邻。局部空间自相关分析表明，人口健康水平的空间格局表现为明显的高值区集聚，低值区集聚的两极分化特征。其中，落在第一象限的高—高集聚的地区数 22 个，占总数的 64.7%；落在第三象限的低—低集聚的地级单元数 7 个，占总数的 20.59%。

利用 Arcgis 对 2019 年全国各地区的人口健康水平进行热点分析，也按照自然段点法分为六个等级，分为高热点区、热点区、次热点区、次冷点区、冷点区和高冷点区(如图 3)。总体来看，全国的健康水平呈现明显的地区差异，即“东—中—西”分布差异，从东到西依次是高热点区向高冷点区变化明显，呈现西热东冷的分布格局，本文接下来深入对于探究此现象成因。

3.2. 人口健康水平的影响因素分析

基于 GWR 模型的分析

使用 ols 最小二乘法回归会出现多重共线性的问题，会影响结果的估计。为了解决这一问题，本文先采用逐步回归模型解决这一问题，根据表 2 可以看出，人均 GDP、平均受教育年限、地区医疗机构数、SO₂ 排放量、植被覆盖率和城镇化率的六个变量都通过了显著性检验和共线性检验。模型的拟合度为 R² = 0.853，由结果可以看出，人均 GDP、平均受教育年限、地区医疗机构数、植被覆盖率和城镇化率都是正相关关系，标准化系数为正，说明在其他控制变量保持不变的前提下，这几个因素都对健康水平有正向影响。

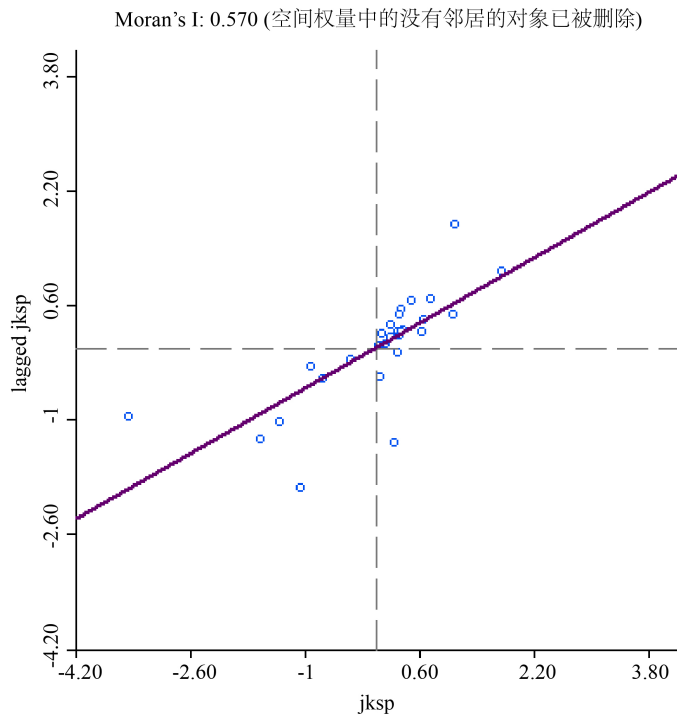


Figure 2. Lisa scatter plot of the health level of the national population
图 2. 全国人口健康水平的 Lisa 散点图

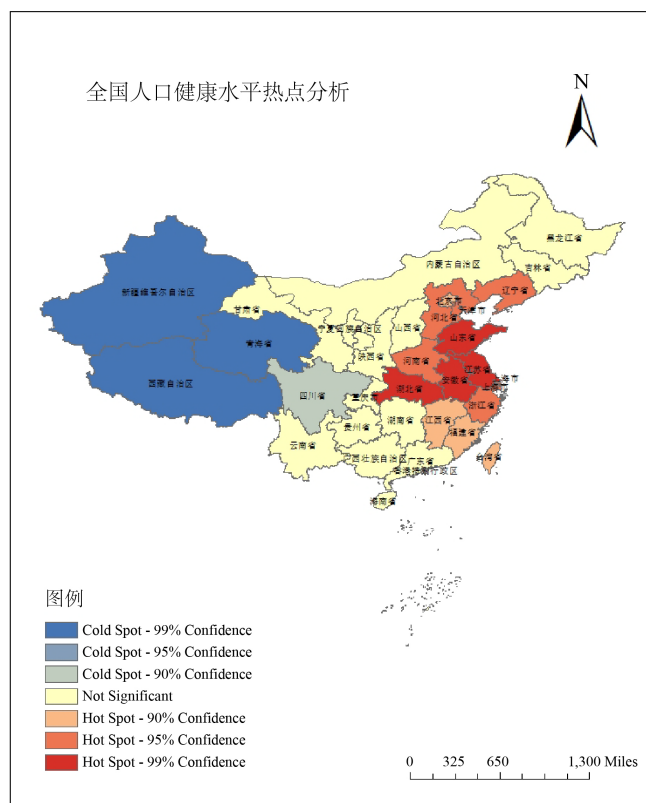


Figure 3. Hot spot analysis of the national population health level
图 3. 全国人口健康水平的热点分析

Table 2. Results of stepwise regression analysis of health levels in various regions of the country
表 2. 全国各地区健康水平逐步回归分析结果

常量、自变量	标准误差	标准化系数	t	VIF	F	R ²
常量	0.136	-0.559	-4.124			
人均 GDP	9.26149e-007	0.195	-0.102	2.315		
平均受教育年限	0.026	0.065	2.486	1.914		
地区医疗机构数	6.21424e-007	0.135	3.791	1.657	26.1022	0.853
SO ₂ 排放量	0.047	-0.012	-0.304	2.198		
植被覆盖率	0.086	0.185	2.157	2.041		
城镇化率	0.338	0.182	2.451	1.816		

如图 3 所示, 各因素对人口健康水平的影响具有显著的空间异质性。具体空间差异如下:

① 人均 GDP 对人口健康水平的正向影响呈现东部强、中部次之、西部最弱的梯度分异。由图 4 可知, 人均 GDP 的回归系数全部为正值, 且回归系数东部地区 > 中部地区 > 西部地区, 说明人均 GDP 的增长对全国人口健康水平的提高有促进作用, 但作用强度东部地区强于中部地区, 中部地区强于西部地区。经济水平是保障人口健康的基础, 经济水平决定了地区教育、医疗、社会保障等保障健康的必要条件, 我国各地区地区的人均 GDP 与地区人口健康水平的高低梯度分异一致, 表明现阶段我国经济水平对人口健康水平的基础性影响。

② 平均受教育年限对人口健康水平的正向影响呈现东部强、中部次之、西部最弱的梯度分异。由图 4 可知, 平均受教育年限的回归系数东部地区 > 中部地区 > 西部地区, 说明平均受教育年限的提高对全国人口健康水平的提升有促进作用, 但作用强度东部地区强于中部地区, 中部地区强于西部地区。平均受教育年限反映的是不同人口受教育水平, 教育通过提高人口健康素养和意识, 优化健康行为和改善经济社会条件而对人口健康产生影响, 受教育年限高的人群, 往往更易获得健康。受教育年限提高带来的健康效应在中、西部地区更显著, 而东地区教育水平已较高, 提高受教育年限带来的健康效应就不够明显, 应通过完善社会保障体系等其它方式带来继续提升人口健康水平。

③ 地区卫生机构数对人口健康水平的正向影响亦呈现东部地区强、中部次之、西部最弱的梯度分异。由图 4 可知, 地区卫生机构数的回归系数东部地区 > 中部地区 > 西部地区, 说明卫生机构数的提高对全国人口健康水平的提升有促进作用, 但作用强度也是东部地区强于中部地区, 中部地区强于西部地区。地区卫生机构数代表地区卫生资源水平, 卫生资源水平对人口健康有正向影响, 我国地区卫生机构数与人口健康水平高低分异趋势一致, 与平均受教育年限的影响相似, 由于西部地区卫生机构数的配置落后于东中部地区, 所以, 西部地区增加卫生机构数的配置带来的健康效应较中西部地区更显著。

④ SO₂排放量对人口健康水平的负向影响呈现西部强、中部次之、西部最弱的梯度分异。由图 4 可知, SO₂排放量的回归系数为负值, 且回归系数绝对值东部地区 > 中部地区 > 西部地区, 说明 SO₂排放量对全国人口健康水平的提升有不利影响, 影响强度为东部地区强于中部地区, 中部地区强于西部地区。研究表明, SO₂排放量损害人体健康, 使婴儿及老年人死亡率提高。我国各地区的 SO₂排放量一定程度上代表着地区环境因素, 这是影响人口健康水平的重要指标, 结果显示, 东部地区 SO₂排放量本已较高, 虽然人口健康也受 SO₂排放量制约, 但由于当地政府采取一定的政策和居民对这一环境因素的适应性, 其受 SO₂排放量升高的不利影响要小于中西部地区。

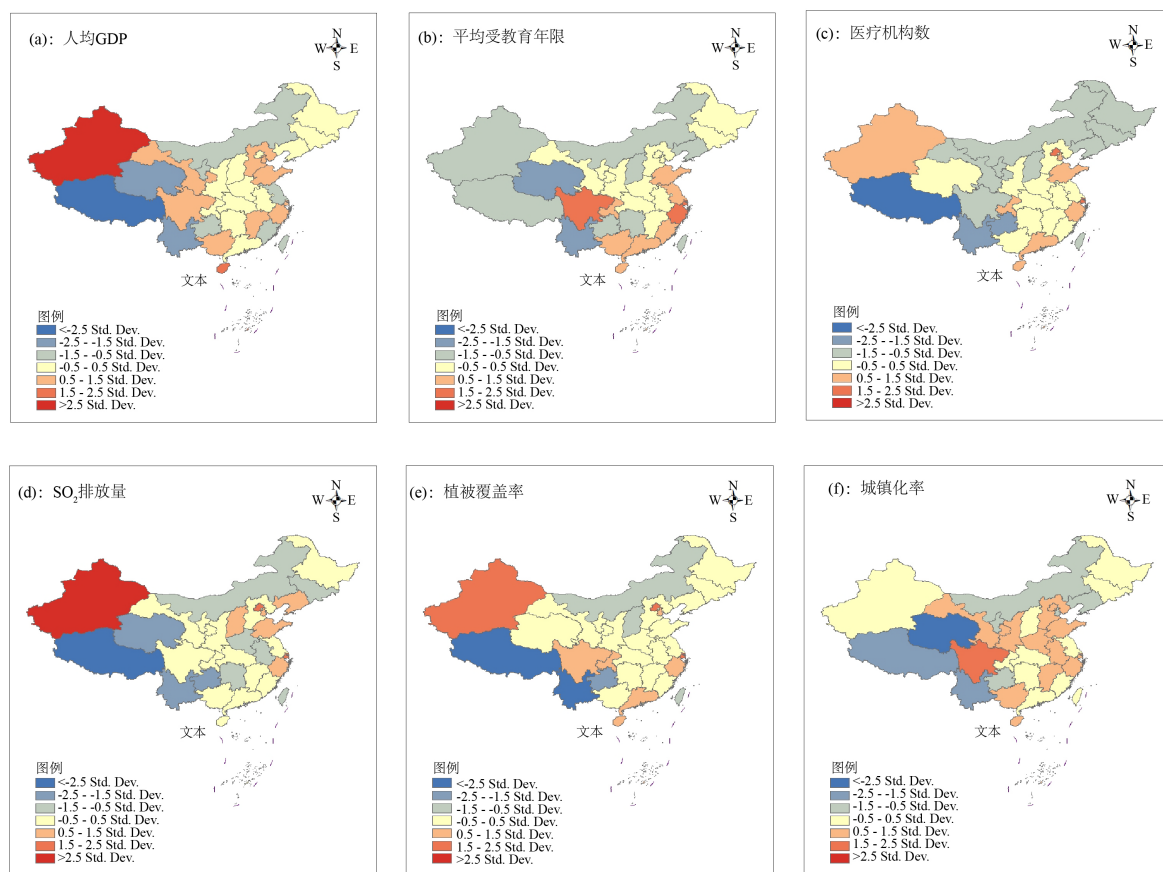


Figure 4. The spatial distribution of the regression coefficients of the factors affecting the health level of the national population

图 4. 全国人口健康水平的影响因素回归系数的空间分布

⑤ 植被覆盖率对人口健康水平的正向影响亦呈现东部强、中部次之、西部最弱的梯度分异。由图 4 可知，植被覆盖率的回归系数东部地区 > 中部地区 > 西部地区，说明植被覆盖率的升高对整个我国地区人口健康水平的提升有正向影响，影响强度为东地区强于中部地区，中部地区强于西部地区。我国植被覆盖率的地区差异与人口健康水平的分异相一致。植被覆盖率高的地区环境质量相对较高，利于人体健康。东中部地区植被覆盖率本来已较高，而西部地区多高寒山区，植被覆盖率相对较低，因此，随着植被覆盖率的升高，西部地区人口健康受到的有利影响将会强于东中地区。

⑥ 城镇化率对人口健康水平的正向影响亦呈现东部强、中部次之、西部最弱的梯度分异。由图 4 可知，城镇化率的回归系数东部地区 > 中部地区 > 西部地区，说明城镇化率的升高对整个我国地区人口健康水平的提升有正向影响，影响强度为东地区强于中部地区，中部地区强于西部地区。我国城镇化率的地区差异与人口健康水平的分异相一致。城镇化率高的地区基础设施相对完善，这有利于人体健康。东中部地区城镇化率本来已较高，而西部地区城镇化率相对较低，因此，随着城镇化率的升高，西部地区人口健康受到的有利影响将会强于东中地区。

4. 结论与建议

4.1. 结论

1) 我国人口健康水平呈现东部地区高、中部地区次之、西部地区低的东西梯度分异格局，东部地区

高值集聚明显，西部地区低值集聚明显，健康水平的区域差异显著，居民健康公平性有待提高。

2) 我国居民人口健康水平受经济、教育、医疗、环境和社会因素综合影响。地区卫生机构数、平均受教育年限、人均 GDP、城镇化率和植被覆盖率对人口健康水平有显著正向影响，影响强度由大到小依次为卫生机构数、平均受教育年限、人均 GDP、城镇化率和植被覆盖率，SO₂ 排放量对人口健康水平有显著负向影响。

3) 环境、教育、经济和社会因素对人口健康水平的影响具有显著的空间异质性。人均 GDP 对人口健康水平的正向影响呈现东强西弱的梯度分异，卫生机构数、平均受教育年限、城镇化率和植被覆盖率对人口健康水平的正向影响则呈现西强东弱的梯度分异，SO₂ 排放量对人口健康水平的负向影响呈现东强西弱的梯度分异。

4.2. 建议

依据以上对我国各个地区人口健康水平影响因素的分析，为促进我国人口健康水平的提高，缩小区域差距，实现健康公平，提出以下建议：

第一，统筹考虑经济、教育、卫生资源、环境条件等影响健康的各因素，提高我国人口健康综合保障水平。我国各地区人口健康受自然和经济社会条件综合影响，单一方面的努力不能形成合力，只有从多层面、多角度出发，深化体制机制改革，优化资源配置，才能全方位促进和保障人口健康，提高健康水平，促进人口健康与经济、社会、环境的良性协调发展。

第二，兼顾各因素对人口健康水平影响的区域差异，有针对性地制定我国各地区差异化的人口健康促进政策。具体而言，对于东中部地区，在继续提高经济水平的同时，通过完善社会保障体系、深化医药卫生体制改革、倡导健康生活方式等途径，促进人口健康水平的持续提高；对于西部地区，应努力克服环境条件的限制，在发展经济的同时，着重增加医疗卫生资源的供给，提高卫生资源可及性，大力提高教育水平，提升居民健康素养，以缩小与中东部地区健康水平的差距，实现我国各地区人口健康公平。

参考文献

- [1] 杨振, 刘会敏, 王晓霞. 中国人口健康分布的时空变化与影响因素[J]. 世界地理研究, 2017, 26(2): 161-168.
- [2] 崔晓东, 周海花. 老年健康水平的区域差异分解及收敛性分析[J]. 统计与决策, 2021, 37(15): 52-56.
- [3] 敖荣军, 李浩慈, 杨振, 张涛. 老年人口健康的空间分异及影响因素研究——以湖北省为例[J]. 地理科学进展, 2017, 36(10): 1218-1228.
- [4] 杨振, 丁启燕, 周晴雨, 刘会敏. 长江中下游地区人口健康水平空间分异特征与地理影响因子[J]. 地理与地理信息科学, 2018, 34(6): 77-84+2.
- [5] 王延赏, 顾钿钿, 初海超, 杜灼. 环境状况对我国城乡居民健康水平影响[J]. 中国公共卫生, 2020, 36(9): 1264-1267.
- [6] 武冠甲, 顾宗倪, 罗小龙, 王绍博. 中国城市化与老年人健康的空间耦合特征研究[J]. 现代城市研究, 2021(4): 64-71.
- [7] 董惠玲, 任桂芳, 宁佩, 吴炳义. 基于地理加权回归的老年人口健康预期寿命影响因素分析[J]. 中国卫生政策研究, 2020, 13(2): 73-80.
- [8] 陈云, 吴雨. 长江经济带人口健康水平的空间格局及影响因素[J]. 经济地理, 2020, 40(9): 152-159.