

doi: 10.3969/j.issn.1002-0268.2017.04.007

基于数据场理论的公路自然区划指标 分析方法研究

李超¹, 王玉兰², 黄智焱³

(1. 山东交通学院, 山东 济南 250023; 2. 山东省交通规划设计院, 山东 济南 250031;
3. 济南城市建设投资集团有限公司, 山东 济南 250101)

摘要: 为了提高公路自然区划的准确性与实用性, 基于数据场理论, 结合公路工程建设的实际需求, 进行公路自然区划指标分析方法研究。在分析数据场基本原理的基础上, 提出基于数据场理论的公路区划指标空间插值算法, 详细论述了利用场强变差函数确定各样本数据权重的方法; 将区划指标空间插值算法与地理基础数据和空间数据处理平台相结合, 同时考虑公路工程建设需求, 提出了基于数据场理论的公路区划单要素指标分析方法。利用单要素指标分析方法、反距离权重法和地统计的插值方法对施工不利日数 ($CLPD_{avg}$) 值进行预测, 对3种方法预测结果进行交叉验证。结果表明, 利用基于数据场理论的单要素指标分析方法得到的公路自然区划指标分布图具有更高的精度, 更符合实际。

关键词: 道路工程; 自然区划; 空间分析; 交叉验证; 数据场; 区划指标

中图分类号: U411

文献标识码: A

文章编号: 1002-0268 (2017) 04-0045-05

Study on Analysis Method for Highway Natural Zoning Indexes Based on Data Field Theory

LI Chao¹, WANG Yu-lan², HUANG Zhi-yan³

(1. Shandong Jiaotong University, Jinan Shandong 250023, China; 2. Shandong Communications Planning and Design Institute, Jinan Shandong 250031, China; 3. Jinan Urban Construction Investment Group Co. Ltd., Jinan Shandong 250101, China)

Abstract: In order to improve the accuracy and practicability of highway natural zoning, based on the data field theory, combining with the actual needs of highway engineering construction, the analysis method of highway natural zoning indexes is studied. On the basis of analyzing the basic principle of data field, a spatial interpolation algorithm for highway natural zoning based on the data field theory is proposed, and the method of determining the weight of each sample data with the field intensity variation function is expounded. Combining the zoning index spatial interpolation algorithm with the geographical basic data and spatial data processing platform, at the same time, considering the demand of highway engineering construction, the single factor index analysis method based on the data field theory is put forward. The number of unfavorable days for construction ($CLPD_{avg}$) is forecasted with the single factor index analysis method, inverse distance weight method and interpolation method of ground statistics to cross verify the forecast result of above 3 methods. The result shows that the highway natural zoning index distribution map generated by the method of single factor index analysis based on data field theory is more accuracy and more in line with the actual situation.

Key words: road engineering; natural zoning; spatial analysis; cross verification; data field; zoning index

收稿日期: 2016-03-29

基金项目: 山东省自然科学基金项目 (ZR2012EEL29; ZR2016EEM42); 山东省高等学校科技计划项目 (J16LA17)

作者简介: 李超 (1979-), 男, 山东济南人, 工学博士, 副教授. (lic211@163.com)

0 引言

作为公路工程自然地理条件的重要参考,公路自然区划应为公路勘察、设计、施工和定额编制提供更加具体、全面和准确的自然地理信息和工程设计参数,单要素指标对此具有重要参考价值^[1-7]。单要素指标主要反映区划研究区域的自然地理要素分布状况,以及自然地理要素对公路工程的影响程度与影响方式,如暴雨日数对公路工程的危害程度,冰冻指数、高温积温等对路基路面结构的循环破坏作用,施工不利日数对公路工程施工质量、工期及费用等方面的影响。这些区划指标在气象部门不能直接得到,必须通过对气温、降水等基础气象数据的运算分析后得出其空间分布规律,而基础气象数据是以气象站等点状要素的形式采集统计的,因此,需要一种空间估值方法,根据样点数据对未知点数据进行预测,从而得到整个研究区域的区划指标分布状况。

单要素区划指标空间插值分析相关研究多基于空间统计学理论。刘洪等在内蒙古草地类型与生物量气候区划研究中,利用梯度距离平方反比法,推算出内蒙古自治区月平均气温、月平均空气湿度和年降水量等单要素区划指标,并在此基础上得到伊万诺夫湿润度,作为综合区划指标^[8]。吴东丽、Ninyerola等利用反距离权重法进行空间插值^[9-10]。岳文泽、Kebaili等基于地统计的插值方法对降水量和蒸发量的样本数据进行空间插值,但检验显示,此方法空间插值的精度还有待提高^[11-12];Mair等尝试反距离权重法对降雨量进行空间插值分析,并对各种方法进行比较^[13-15]。上述方法虽然在总体上能够较好地反映气候要素的空间分布格局,但是由于理论模型基于统计学,有限的样本数量难以准确估计母体分布特征^[16],同时受到地形变化过于复杂、样本空间分布不均匀等因素的影响,导致上述方法空间插值的精度不足。

本文拟借助数据场预测空间数据的优势^[17-18],充分考虑各样本点在数值与空间位置等方面对预测点的影响,并将区划指标空间插值算法与地理基础数据和空间数据处理平台相结合,提出基于数据场的公路区划单要素指标分析方法。

1 基于数据场的区划指标分析方法

基于数据场的公路区划指标分析方法主要适用于具有空间连续分布特性的单要素区划指标的计算

分析。

基于数据场的公路区划指标分析方法是对地理基础数据、公路工程技术和指标估值算法以及数据处理平台的综合应用。首先,以地理基础数据、公路工程技术和数据处理平台为基础,建立并计算公路区划指标分析模型,形成区划指标样本数据。然后,对区划指标样本数据进行数据分布、数据全局趋势、数据空间相关性和方向影响等方面的分析,从而发现数据空间规律并做适当处理,以保证区划指标空间插值运算的精度。最后,通过数据趋势移除、场强变差函数建模、空间插值计算和交叉验证等步骤,完成区划指标计算,并编制公路区划指标分布图,具体过程如图1所示。

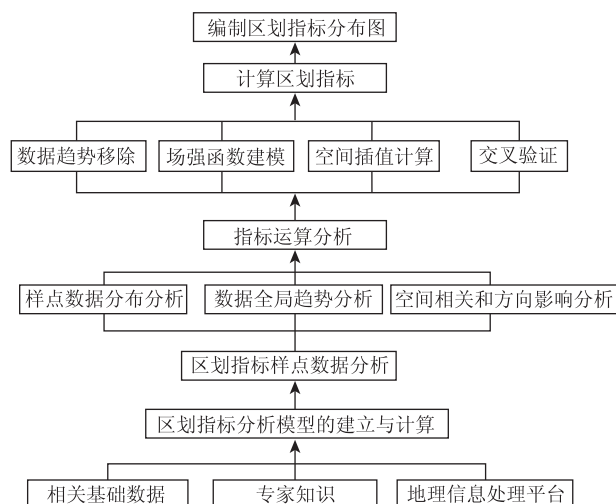


图1 基于数据场的区划指标分析方法流程图

Fig. 1 Flowchart of method of analysing zoning indexes based on data field

2 区划指标分析实例

2.1 区划指标计算

一年中大约有多长时间适宜施工,影响施工的主要气候因素有哪些,影响程度如何,这些问题与工程总工期的估计、工程费用的估算等密切相关。热量条件、水分条件和风等均对公路施工有不同程度的影响,其中各种不利气候条件影响时间的长短是工程人员最为关注的方面,因此,可以采用一年中各种不利影响因素的出现日数作为影响公路施工的气候条件评价指标。主要包括年施工热量不利日数、年施工降水不利日数和年大风日数等。其中,年施工降水不利日数是中国公路自然区划一级分区的重要参考指标,因此,本文以年施工降水不利日数为例,详细介绍基于数据场的区划指标分析方法。

对公路建设来讲,日降水量大于5 mm,一般不能进行正常施工,特别对于路基工程施工。因此称一年内日降水量 ≥ 5 mm的天数为年施工降水不利日数(CLPD),计算公式为:

$$CLPD = \sum_{i=1}^n d_i, \quad (1)$$

式中,CLPD为年施工降水不利日数; n 为年的天数; d_i 为第*i*日降水量高于 ≥ 5 mm时取1,否则取0。累年年均施工降水不利日数 $CLPD_{5\text{ mm}}$ 为:

$$CLPD_{\text{avg}} = \sum_{i=1}^n \sum_{j=1}^m d_{ij}/n, \quad (2)$$

式中, n 为累年年数, $n=30$; m 为第*i*年的天数; d_{ij} 为第*i*年第*j*日降水量高于5 mm时取1,否则取0。

首先利用由中央气象台提供的全国660个气象站(图2)采集记录的基础降水量数据和式(2)计算各气象站的CLPD值,作为年施工降水不利日数指标空间插值计算的样点数据。在此基础上,通过基于数据场的区划指标分析方法对我国疆域内的CLPD进行空间插值计算分析,得到我国CLPD指标的空间分布图。

2.2 指标空间预测模型的建立

2.2.1 场强变差函数建模

为了对场强变差函数值有一个较为清晰的描述,根据经验场强变差函数值所关联的间距,对经验场强变差函数值(样点)进行分组。通过样点对分组,每组样点对都具有一个样点对间距代表值和一个对应的场强变差函数代表值,从而使得场强变差函数云中点的数量大大减少。这样, $CLPD_{\text{avg}}$ 残差的空间自相关可以被很好地捕获,尤其是在短距离处,这对于插值精度尤为重要。由于之前移除了一个粗尺度趋势, $CLPD_{\text{avg}}$ 的场强变差函数值与间距相关联的

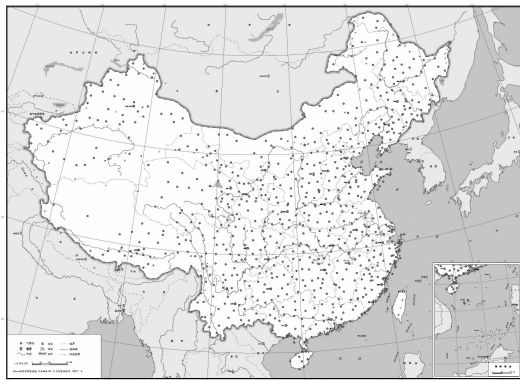


图2 气象站分布图

Fig. 2 Distribution of weather stations

散点图如图3、图4所示。

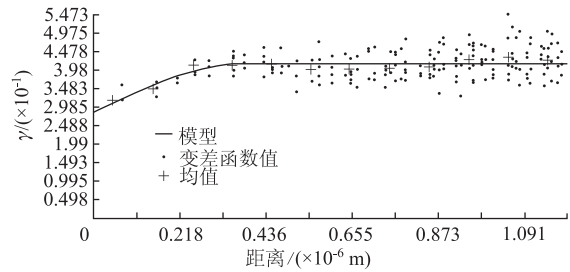


图3 $CLPD_{\text{avg}}$ 场强变差函数散点与函数模型图

Fig. 3 Scatter points and function model curve of $CLPD_{\text{avg}}$ field intensity variation function

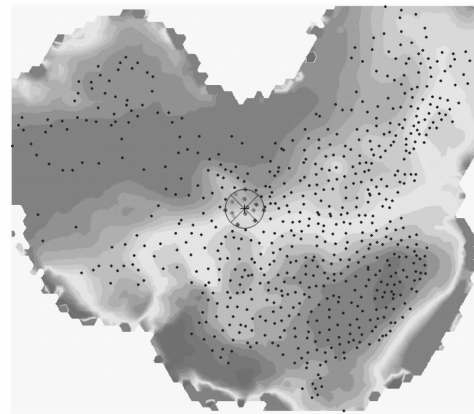


图4 邻域搜索计算图

Fig. 4 Neighborhood search calculation map

计算最佳的参数值得到 $CLPD_{\text{avg}}$ 全向稳定场强变差函数模型:

$$\gamma(h) = \begin{cases} 0 & h = 0 \\ C_0 + C \left(\frac{3}{2} \frac{h}{a} - \frac{1}{2} \frac{h^3}{a^3} \right) & 0 < h \leq a, \\ C_0 + C & h > a \end{cases} \quad (3)$$

式中, C_0 为块金值,表示 h 很小时两空间数据点间的变化幅度; C 为基台值,它是先验方差与块金效应之差。在较短距离处,场强变差函数模型曲线骤然上升,当间距达到变程值后趋于平稳。变程是场强变差函数模型开始趋于平稳的距离。场强变差函数趋于平稳表明变程以外的属性(即 $CLPD_{\text{avg}}$ 值)几乎不存在自相关。

3 空间插值计算

根据第2节介绍的指标估值算法以及样本点 $CLPD_{\text{avg}}$ 场强变差函数及其各方向分量分别进行计算,然后将结果叠加得到 $CLPD_{\text{avg}}$ 分布图。

一般情况下,带估值点周围都分布有若干已知样点,参与估值的样点多对于提高估值精度是有帮助的,但根据指标估值算法,在估值计算期间每个估值点都要进行大量的 $CLPD_{avg}$ 场强变差函数值的计算,形成关于参与计算的样点数据权重的多项式,然后求解多项式,得到各样点数据权重,如果参与估值的样点数量太大,形成的多项式规模会过于庞大,势必会影响到计算效率,甚至无法求解。因此,在具体样点的估值计算时,首先定义一个搜索圆,将用于估算未知位置的点用适当数量的点包围起来,圆外的点将不参与计算。此外,为避免特定方向出现偏差,可以将圆分成若干个扇形,在每个扇形中都选择相同数量的已知点,同时指定点的数量、搜索圆半径以及搜索圆的扇形数量,如图 4 中小圆所示。

搜索圆内有多个参与估值的样点数据,其灰度代表了在估值过程中的权重。其权重体现了数据在估值过程中所起作用的大小。利用此方法生成 $CLPD_{avg}$ 指标分布图,如图 5 所示。

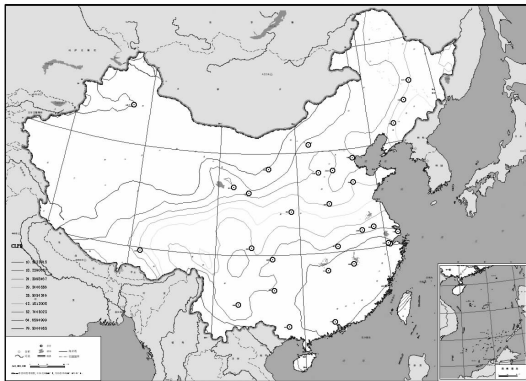


图 5 $CLPD_{avg}$ 指标分布图
Fig. 5 $CLPD_{avg}$ index distribution map

4 交叉验证

交叉验证的目的在于了解模型预测未知值的准确程度。首先将样点数据集中的某些点作为未知点,利用样点数据集的其余部分估算这些点的值,然后将测量值与估计值进行比较(测量值与估计值的差值称为预测误差)。根据估计误差计算得到的统计数据可用于判断数据场函数模型是否适于区划指标 $CLPD_{avg}$ 的空间分布拟合。本文同时利用反距离权重法和地统计的插值方法对 $CLPD_{avg}$ 值进行预测,并进行交叉验证。通过对比表 1 中 3 种方法交叉验证指标值可知,利用基于数据场区划指标分析方法具有更高的预测精度。

表 1 模型评价指标表
Tab. 1 Model evaluation index table

指标名称	数据场 计算值	反距离权重 计算值	地统计的 插值方法	指标标准值
平均误差	0.126	0.265	0.173	0
均方根	6.608	9.728	7.401	越小越好
标准化均值	0.019 554 8	0.042 7	0.036 1	0
标准化均方根	1.095 548	1.250	1.150	1
平均标准差	5.99	7.697	6.987	越小越好

5 结论

本文在分析数据场基本原理的基础上,提出基于数据场理论的区划指标空间插值算法,详细论述了利用场强变差函数计算各参与未知数据点空间估值的各数据权重。将区划指标空间插值算法与地理基础数据和空间数据处理平台相结合,同时考虑公路工程建设需求,提出了基于数据场的公路区划单要素指标分析方法。方法首先以地理基础数据、公路工程技术和数据处理平台为基础,建立并计算公路区划指标分析模型,形成区划指标样本数据;然后对区划指标样本数据进行数据分布、数据全局趋势、数据空间相关性和方向影响等方面的分析,从而发现数据空间规律并做适当处理,以保证区划指标空间插值运算的精度;最后通过数据趋势移除、场强变差函数建模、空间插值计算和交叉验证等步骤,完成区划指标计算,并编制公路区划指标分布图。利用基于数据场的公路区划单要素指标分析方法编制单要素区划指标年施工不利日数分布图,以对上述方法进行分析验证。

参考文献:

References:

[1] 齐洪亮,田伟平,舒延俊.基于坡面径流强度的中国公路水文区划研究[J].公路交通科技,2010,27(3):12-16.
QI Hong-liang, TIAN Wei-ping, SHU Yan-jun. Research on Hydrological Regionalization for Highways in China Based on Slope Runoff Intensity [J]. Journal of Highway and Transportation Research and Development, 2010, 27(3): 12-16.

[2] 齐洪亮,田伟平,舒延俊.基于河川径流强度的中国公路水文区划研究[J].公路交通科技,2011,28(3):55-60.
QI Hong-liang, TIAN Wei-ping, SHU Yan-jun. Research on Hydrological Regionalization for Highway in China Based on River Runoff Intensity [J]. Journal of Highway

- and Transportation Research and Development, 2011, 28 (3): 55-60.
- [3] 王福恒,李家春,马保成,等.公路岩土区划中岩石分类体系与类型分布[J].公路交通科技,2010,27(6):102-106.
WANG Fu-heng, LI Jia-chun, MA Bao-cheng, et al. Classification System and Type Distribution of Rock in Highway Rock and Soil Zoning [J]. Journal of Highway and Transportation Research and Development, 2010, 27 (6): 102-106.
- [4] 王福恒,李家春,马保成,等.公路岩土区划方法与区划指标的研究[J].公路交通科技,2010,27(12):48-52.
WANG Fu-heng, LI Jia-chun, MA Bao-cheng, et al. Method and Index for Zoning of Highway Rock and Soil [J]. Journal of Highway and Transportation Research and Development, 2010, 27 (12): 48-52.
- [5] 张志超,李萍,许金良.基于ArcGIS的山东省公路岩土区划研究[J].公路工程,2014,39(1):263-268.
ZHANG Zhi-chao, LI Ping, XU Jin-liang. Research on the Highway Geotechnical Division of Shandong Province Based on ArcGIS [J]. Highway Engineering, 2014, 39 (1): 263-268.
- [6] 汪双杰,王佐,袁堃,等.青藏公路多年冻土地区公路工程地质研究回顾与展望[J].中国公路学报,2015,28(12):1-8.
WANG Shuang-jie, WANG Zuo, YUAN Kun, et al. Qinghai-Tibet Highway Engineering Geology in Permafrost Regions: Review and Prospect [J]. China Journal of Highway and Transport, 2015, 28 (12): 1-8.
- [7] 毛雪松,王富春,王秉纲.中国公路沙漠自然区划研究[J].公路交通科技,2009,26(1):57-60.
MAO Xue-song, WANG Fu-chun, WANG Bing-gang. Desert Natural Division of Highway in China [J]. Journal of Highway and Transportation Research and Development, 2009, 26 (1): 57-60.
- [8] 刘洪,郭文利,权维俊.内蒙古草地类型与生物量气候区划[J].应用气象学报,2011,22(3):329-335.
LIU Hong, GUO Wen-li, QUAN Wei-jun. Climatic Division of the Types and Yields of Grassland in Inner Mongolia [J]. Journal of Applied Meteorological Science, 2011, 22 (3): 329-335.
- [9] 吴东丽,王春乙,薛红喜,等.华北地区冬小麦干旱风险区划[J].生态学报,2011,31(3):760-769.
WU Dong-li, WANG Chun-yi, XUE Hong-xi, et al. The Drought Risk Zoning of Winter Wheat in North China [J]. Acta Ecologica Sinica, 2011, 31 (3): 760-769.
- [10] NINYEROLA M, PONS X, ROURE J M. Objective Air Temperature Mapping for the Iberian Peninsula using Spatial Interpolation and GIS [J]. International Journal of Climatology, 2007 (27): 1231-1242.
- [11] 岳文泽,徐建华,徐丽华.基于地统计方法的气候要素空间插值研究[J].高原气象,2005,24(6):974-980.
YUE Wen-ze, XU Jian-hua, XU Li-hua. A Study on Spatial Interpolation Methods for Climate Variables Based on Geostatistics [J]. Plateau Meteorology, 2005, 24 (6): 974-980.
- [12] BARGAOUI Z K, CHEBBI A. Comparison of Two Kriging Interpolation Methods Applied to Spatiotemporal Rainfall [J]. Journal of Hydrology, 2009, 365 (1): 56-73.
- [13] MAIR A, FARES A. Comparison of Rainfall Interpolation Methods in a Mountainous Region of a Tropical Island [J]. Journal of Engineering Physics and Thermophysics, 2011, 84 (1): 13-22.
- [14] 戢晓峰,郝晶晶.基于多源数据的货物空间运输联系特征及形成机制[J].公路交通科技,2016,33(7):153-158.
JI Xiao-feng, HAO Jing-jing. Study on Freight Spatial Transport Connection Characteristics and Its Formation Mechanism Based on Multi-source Data [J]. Journal of Highway and Transportation Research and Development, 2016, 33 (7): 153-158.
- [15] 卢守峰,王杰,刘改红,等.基于流量和出租车GPS数据的城市道路网络宏观基本图[J].公路交通科技,2014,31(9):138-144.
LU Shou-feng, WANG Jie, LIU Gai-hong, et al. Macroscopic Fundamental Diagram of Urban Road Network Based on Traffic Volume and Taxi GPS Data [J]. Journal of Highway and Transportation Research and Development, 2014, 31 (9): 138-144.
- [16] 李德仁,王树良,李德毅.空间数据挖掘理论与应用[M].北京:科学出版社,2006.
LI De-ren, WANG Shu-liang, LI De-yi. Theory and Application of Spatial Data Mining [M]. Beijing: Science Press, 2006.
- [17] 韩绍泽,刘嘉庆,王凡,等.数据场聚类在信息安全中的应用研究[J].计算机与数字工程,2009,37(1):109-110,195.
HAN Shao-ze, LIU Jia-qing, WANG Fan, et al. Research on Application of Data Site in Information Security [J]. Computer and Digital Engineering, 2009, 37 (1): 109-110, 195.
- [18] 龚小鹏.基于数据场的空间数据建模研究[D].北京:中国地质大学,2009.
GONG Xiao-peng. Research on Spatial Data Modeling Based on Data Field [D]. Beijing: China University of Geosciences, 2009.