

引用格式: 杨浩, 孟娜, 王婧, 等. 基于支持向量机的京津冀城市群热环境时空形态模拟[J]. 地球信息科学学报, 2019, 21(2): 190-200. [Yang H, Meng N, Wang J, et al. Spatial-temporal morphology simulation of Beijing-Tianjin-Hebei urban agglomeration thermal environment based on Support Vector Machine[J]. Journal of Geo-information Science, 2019, 21(2): 190-200.] DOI:10.12082/dqxxkx.2019.180271

基于支持向量机的京津冀城市群热环境时空形态模拟

杨浩^{1,2,3}, 孟娜⁴, 王婧^{1*}, 郑燕⁵, 赵莉^{6,7}

1. 北京市社会科学院, 北京 100101; 2. 北京大学经济学院, 北京 100871; 3. 北京市延庆区交通局, 北京 102100; 4. 北京师范大学政府管理学院, 北京 100875; 5. 河北民族师范学院外语学院, 承德 067000; 6. 西南科技大学文学与艺术学院, 绵阳 621000; 7. 复旦大学, 上海 200433

Spatial-temporal Morphology Simulation of Beijing-Tianjin-Hebei Urban Agglomeration Thermal Environment based on Support Vector Machine

YANG Hao^{1,2,3}, MENG Na⁴, WANG Jing^{1*}, ZHENG Yan⁵, ZHAO Li^{6,7}

1. Beijing Academy of Social Sciences, Beijing 100101, China; 2. School of Economics, Peking University, Beijing 100871, China; 3. Beijing Yanqing District Transportation Bureau, Beijing 102100, China; 4. School of Government, Beijing Normal University, Beijing 100875, China; 5. Hebei Normal University for Nationalities, Chengde 067000, China; School of Literature and Art, Southwest University of Science and Technology, Mianyang 621000, China; 7. Fudan University, Shanghai 200433, China

Abstract: As an important part of regional ecology, the thermal environment of urban agglomeration has become a research hot topic in recent years. How to choose the thermal environment quantification method for the complex geomorphological features of urban agglomeration has been a difficult technical problem to be solved. Based on this, This study proposes a solution to multi-sample, nonlinear, non-stationary and high-dimensional function fitting. The calculation method is established, and the thermal environment surface model of Beijing-Tianjin-Hebei urban agglomeration based on support vector machine (SVM) is established to reveal the temporal and spatial morphological changes of the thermal environment of urban agglomeration. The results show: ① that the SVM model has theoretical and practical feasibility in describing the spatial distribution of the thermal environment of urban agglomerations with multi-core and multi-land-use types. It can optimize the differences locally through the Gaussian Kernel Function according to the overall spatial distribution of the thermal environment, and minimize the impact of default values on the fitting results of the model. Comparing with the control method, the spatial distribution pattern of heat island in urban agglomerations with complex geomorphologic features can be simulated with higher accuracy. ② In the process of fitting the surface of SVM model, accuracy and the time of fitting are important indexes to measure the results, and original image resolution is the decisive influencing factor. ③ In 2003-2013, the most obviously change regions of urban heat island effect are Beijing and Tianjin. The heat island area of the two cities increased by 7091 km square and 4196 km square, respectively. The

收稿日期: 2018-06-04; 修回日期: 2018-12-03.

基金项目: 国家社会科学基金项目(18CGL048); 国家自然科学基金重点项目(41731286); 中国博士后科学基金资助项目(2018M631383); 北京市社会科学院青年项目(2018B5186)。[**Foundation items:** National Social Science Foundation of China, No.18CGL048; Key Program of National Natural Science Foundation of China, No.41731286; China Postdoctoral Science Foundation Funded Project, No.2018M631383; Beijing Academy of Social Sciences Youth Research Project, No.2018B5186.]

作者简介: 杨浩(1983-), 男, 吉林长春人, 博士, 助理研究员, 研究方向为土地行政、土地资源管理。E-mail: yh01@pku.edu.com

*通讯作者: 王婧(1983-), 女, 山东淄博人, 博士, 副研究员, 研究方向为区域发展与管理。E-mail: w.j.2000@163.com

spatial trend was developing continually year by year, and the trajectory of gravity center of the heat island had obvious spatial and temporal variations. Beijing's urban heat island is characterized by uneven growth in the southeast and slow growth in the west, while Tianjin's urban heat island is characterized by the expansion of city center to the surrounding areas. This study further enriches the quantitative methods of urban thermal environment assessment, and can provide quantitative and visual decision supports for urban agglomeration planning, urban construction, environmental protection and regional sustainable development practically.

Key words: urban thermal environment; temporal and spatial patterns; gaussian surface model; Support Vector Machine (SVM); Beijing-Tianjin-Hebei urban agglomeration

***Corresponding author:** WANG Jing, E-mail: w.j.2000@163.com

摘要:城市群热环境作为区域生态重要组成部分,已成为近年来的研究热点。而如何选择针对城市群这种复杂地貌特征的热环境量化工具一直是亟待解决的技术难点,基于此本研究提出了一种解决多样本、非线性、非平稳及高维函数拟合的计算方法,并建立了基于支持向量机(SVM)的京津冀城市群热环境曲面模型来揭示城市群热环境的时空形态变化。研究结果表明:① SVM模型在刻画多核心、多种土地利用类型城市群热环境的空间分布方面具有理论与实践可行性,能够根据热环境的整体空间布局通过高斯核函数进行局部优化差值,最大限度减少缺省值对模型拟合结果的影响。相比于对照方法可以模拟出更高精度的复杂地貌特征城市群热岛空间分布格局;② 在SVM模型曲面拟合的过程中,拟合精度和拟合时间是衡量拟合结果的重要指标,而原始影像的分辨率则是影响该指标的决定性因素;③ 2003–2013年区域内北京市与天津市的城市热岛效应变化最为明显,热岛面积分别增加7091 km²与4196 km²,空间上呈现出逐年接近连片发展趋势,热岛重心移动轨迹具有明显的时空分异性。北京城市热岛特征为东南部地区异速增长,西部地区缓慢增长;天津城市热岛特征为以城市中心为圆心向周围扩展。本研究进一步丰富了城市群热环境评测的定量方法,可以在实践上对城市群的城市规划、城市建设、环境保护和区域可持续发展等提供定量化、可视化的决策支持。

关键词:城市热环境;时空形态;高斯曲面模型;支持向量机(SVM);京津冀城市群

1 引言

近年来,伴随城镇化进程的加快,城市基础设施建设强度不断增加,城市土地利用/覆被状况逐渐变化所产生的增温与高强度人为热排放在原本离散的城市内部难以消解,导致城市的热岛效应恶化趋势日渐明显^[1-4];同时,城市土地利用/覆被状况的逐渐变化与区域能量平衡过程的改变直接影响着地表-大气之间的相互作用,进而改变大气的热力学和动力学特征^[5-8]。现阶段,城市热环境问题已不再是单纯的气候环境问题,而是影响城市化进程和城市生态环境可持续发展的重大阻碍之一。京津冀作为典型的城市群,热环境问题尤为突出,研究京津冀城市群热环境时空格局变化的现状与特点、分析其对区域热环境变化的影响机制,不仅对京津冀地区的生态环境保护具有重要的参考价值,还对城市群的生态环境健康与可持续发展具有重大的战略意义。

从国内外城市热环境研究来看,相关研究主要集中在城市热岛效应及热岛强度度量、城市气候等方面,且以对单一城市热环境研究为主,对城市群

或区域内多个城市热环境效应、热环境时空变化研究相对较少。例如,Sun等^[9]通过对广州市城乡空气温度差异性的历史分析,探讨了城市发展对城市热岛的影响,还将遥感技术应用于土地利用/土地覆盖和地表温度的信息,阐述了广州市的土地利用变化及其热环境效应。陈云浩等^[10]运用RS、GIS技术,构建了热力景观空间格局的评价指标体系,用转移矩阵法对上海多年的多时相热力景观转移速率进行计算,并在分析后得出了上海市近十年热力相关类型的变化强度及趋势;张心怡等^[11]以上海为主要研究对象,运用格局、过程、机制和模型,并结合Landsat TM/ETM+影像和GIS的空间分析功能,研究了上海的土地利用/土地覆盖变化及热环境效应,为上海的城市规划发展、环境保护及生态建设等提供依据。李雪松等^[12]以城市区域面积、地理信息等详细的数据调研为依据,对武汉市东南片区的城市扩张进行了基于城市冠层模型的WRF数字模拟,导出了不同时间段气温及其差值,并进行量化比对和分析。王福海等^[13]以重庆市主城区为研究对象,基于Landsat TM遥感数据,采用人机交互解译和定量遥感技术相结合的方法,研究重庆市主城区热环

境时空变化特征与建设用地扩展之间的关系, 得出了地表温度与建设用地呈明显的正相关关系, 新增建设用地容易形成新兴热岛区等结论。张丽娟等^[13]基于黑龙江省1960–2010年的土地利用变化数据, 采用自然正交分解(EOF)、气候倾向率及 Observation Minus Reanalysis(OMR)等方法, 分析了土地利用变化对黑龙江省气温的影响。

本文选择京津冀城市群作为研究区域, 运用支持向量机(Support Vector Machines, SVM)方法研究整个区域热环境时空变化。一方面, 可揭示城市群空间发展对热环境的影响, 有助于对城市群热环境时空变化规律及问题的探索与发现; 另一方面, 对京津冀城市群热环境风险的防范起到科学支撑, 为相关部门决策提供参考价值。此外, 选择支持向量机(SVM)模型对城市群热环境进行模拟, 在研究方法上也具有较强的实用性与创新性。

2 研究区概况与数据源

2.1 研究区概况

京津冀城市群包括北京、天津、河北两市一省, 土地面积18.50万km², 占全国土地总面积的1.93%。集聚了全国7.23%的人口, 创造了11.3%的国内生产总值, 城镇化水平达到60.7%。从气候环境来看, 京津冀城市群地势为西北高、东南低, 区域内主要城市的气候变化与大气环流变化密切相关, 春季干燥少雨多风沙, 夏季炎热多雨, 秋季晴爽, 冬季寒冷少雪, 属于典型温带季风气候区; 近年来京津冀城市群以北京市、天津市为中心增长点, 引领带动本区域内各次级中心城市(河北省11个地级市)统筹协调发展。国家“十三五”规划纲要将京津冀城市群定为“以首都为核心的世界级城市群、区域整体协同发展改革引领区、生态环境改善示范区”, 这表明京津冀城市群作为中国东部核心区的地位日益重要。

2.2 数据源

本文以2003–2013年MODIS-LSTs地表温度产品为主要数据源, 通过半径法选取北京、天津、唐山、保定、石家庄5个城市为代表, 对京津冀地区城市群的热岛足迹进行研究, 建立城市热岛容量模型, 获得MODIS数据所建立的京津冀城市群温度数据库。MODIS-LSTs产品数据通过每日4次经过京津冀城市群的Terra(上午)和Aqua(下午)两颗卫

星获取。LSTs数据产品通过MOD11A1和MYD11A1产品的反演获取京津冀城市群区域内空间分辨率为1 km, 温度反演精确度小于1.0 K的地表温度数据。虽然MODIS的数据分辨率较为粗糙, 但其每日4次的温度提取使得城市热岛在不同时间尺度上的拟合变为可能。特别是MODIS-LSTs在排除大气干扰的情况下, 反演精度能在0.5 K范围之内, 符合较高精度的拟合需求。因此, 本研究中基于MODIS-LSTs产品所建立的京津冀城市群地表温度数据库能够比较客观地反映京津冀城市群热环境的现实情况。而且利用MODIS遥感数据通过分裂窗算法和白天/夜间陆表温度算法等得到的地表温度精度更高, 也具有较好的稳定性。同时, 为了减少后续研究中曲面拟合运算中的误差, 减少模拟过程中由于行政边界的不规则剪切而造成的缺省值对曲面拟合的影响, 本研究对该数据库进行网格化设置, 通过ArcGIS软件, 采用兰伯特投影(LCC), 设置中心经纬线为116.69°E, 39.35°N, 并将其转化为(–300~300 km, –372~372 km)的矩形网格(图1)。其中, 研究区域内由于云层、海洋等环境因素造成的缺省值区域被设置为空值, 不纳入模型模拟计算。

本研究采用夏季白天地表温度平均值作为原始数据输入。原因在于城区热岛形成主要是由下

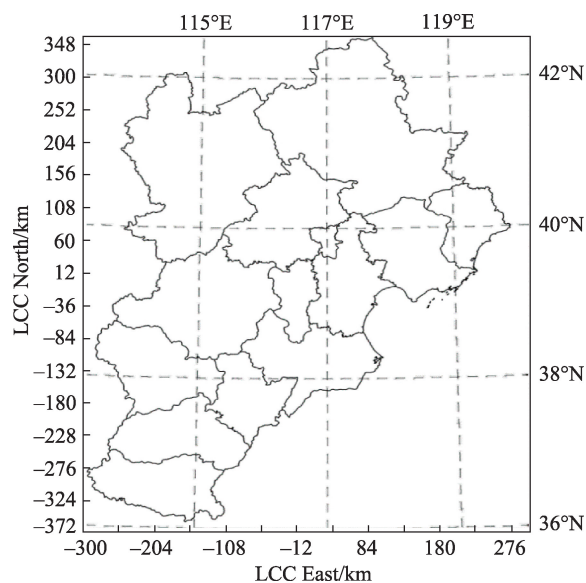


图1 网格化设置后的研究区域位置

Fig. 1 Study area after grid setting

垫面热容性质的差异所造成, 城镇区域中夹杂分布的绿地及公园会造成热岛曲面模拟的误差; 而在冬季城市热岛形成的主要来源属人为热源, 城镇区域内分布较为均匀, 模拟误差较小, 因此在该时期进

行城市热岛曲面模拟的精度较差,易于筛选适用于该区域热岛曲面拟合的模型。

3 研究方法

为实现对城市群热环境时空形态的测定、计算,需要对城市热环境空间形态的几何特征和布局特点有细致详尽的认识。目前,城市热环境曲面拟合模拟因其精确高、直观性强等特点成为目前进行城市空间形态研究的重要手段之一^[14-15]。

3.1 高斯曲面模型

从城市热岛形成的机理来看,影响地表温度的因素主要有初始地温、地物热惯量和辐射总量等。如果将研究区域视为地表热力系统,那么初始地温相当于整个地表热力系统的背景温度,且地物热惯量的差异则是导致地表热力系统内地表温度差异的主要影响因素。因此,热力系统的地表温度可以表示为^[16]:

$$T_s(x, y) = T_{s1}(x, y) + T_{s2}(x, y) \quad (1)$$

式中: (x, y) 分别为像元坐标; $T_s(x, y)$ 为地表温度; $T_{s1}(x, y)$ 为背景温度; $T_{s2}(x, y)$ 为热力系统局部热异常导致的温度增量。

在研究城市热岛效应时,通常考虑乡村地区地物热惯性相对稳定,将其作为背景场,即地表温度为 $R(x, y)$ 。如果在 $R(x, y)$ 温度背景场上叠加城市地区热异常则会表现出城市热岛效应,那么把由城市热岛效应导致的温度增量 $T_{s2}(x, y)$ 定义为热岛信号函数 $UHI(x, y)$ 。因此,热力系统的地表温度可转换为:

$$T_s(x, y) = R(x, y) + UHI(x, y) \quad (2)$$

因此,根据式(2)分别抽象出乡村地表温度场 $R(x, y)$ 和热岛信号函数 $UHI(x, y)$,即可拟合出传统城市热环境曲面。然而,拟合城市群热环境曲面时,多核心、城乡分界不明显、空间分布规律更复杂等热环境空间特征因素会制约传统曲面模型的精度。

二维高斯曲面模型相较于传统曲面模型能够有效地拟合区域空间内具有孤岛效应热值空间分布,并且可以将不同城市之间以及单城市中不同时期的地表城市热岛(Surface Urban Heat Island, SUHI)的变化趋势进行比对,能够更好地测度城市在各个

时期热岛的形状和方向。近年来,二维高斯曲面模型被广泛地运用到 SUHI 拟合或城市 LST 空间形态分布研究中,特别是当城市区域相对于周围的农村地区凸显出热岛效应时测量更加准确。在高斯曲面模型中,可通过式(3)对城市中 SUHI 进行拟合。

$$SUHI(x, Y) = a_0 \times \exp \left[\frac{((x-x_0)\cos\varphi + (y-y_0)\sin\varphi)^2}{0.5a_x^2} - \frac{((y-y_0)\cos\varphi - (x-x_0)\sin\varphi)^2}{0.5a_y^2} \right] \quad (3)$$

式中: x 和 y 为像元值的位置; a_0 为热岛强度的最大值; a_x 和 a_y 为热岛空间扩展范围(即热岛椭圆的 x 轴和 y 轴); φ 为热岛椭圆的偏向角度; x_0 和 y_0 为模拟热岛椭圆的中心位置。

在本研究中,通过 MATLAB10 进行高斯曲面拟合的具体建模过程如下:

(1)输入:通过 MATLAB 来读取 MODIS 网格数据矩阵(包括原始坐标 $(x$ 和 $y)$ 和温度数据)。

(2)高斯拟合:通过对数转换将原始方程转换为线性二阶方程,并通过最小二乘方程实现非线性数据拟合,通过 lsqcurvefit.m (MATLAB 优化工具箱)实现迭代拟合过程,最终得到 $a_1 \sim a_6$ 共 6 个参数:

$$\ln(SUHI(x, y)) = a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6 \quad (4)$$

(3)结果输出:从步骤(1)、(2)得到的拟合系数 $a_1 \sim a_6$,通过式(5)~(10)进行转换,得到高斯曲面方程中的关键性参数,并在此基础上进行 MATLAB 绘图,可得到该城市群中 SUHI 二维或三维分布图。

$$a_1 = \ln a_0 - \frac{x_0^2 \cos^2\varphi + 2x_0y_0 \sin\varphi \cos\varphi + y_0^2 \sin^2\varphi}{2a_x^2} - \frac{y_0^2 \cos^2\varphi - 2x_0y_0 \sin\varphi \cos\varphi + x_0^2 \sin^2\varphi}{2a_y^2} \quad (5)$$

$$a_2 = \frac{x_0 \cos^2\varphi + y_0 \sin\varphi \cos\varphi}{a_x^2} - \frac{x_0 \sin^2\varphi - y_0 \sin\varphi \cos\varphi}{a_y^2} \quad (6)$$

$$a_3 = \frac{y_0 \sin^2\varphi + x_0 \sin\varphi \cos\varphi}{a_x^2} - \frac{y_0 \cos^2\varphi - x_0 \sin\varphi \cos\varphi}{a_y^2} \quad (7)$$

$$a_4 = \left(-\frac{1}{a_x^2} + \frac{1}{a_y^2} \right) \sin\varphi \cos\varphi \quad (8)$$

$$a_5 = -\left(\frac{\cos^2\varphi}{2a_x^2} + \frac{\sin^2\varphi}{2a_y^2} \right) \quad (9)$$

$$a_6 = -\left(\frac{\sin^2 \varphi}{2a_x^2} + \frac{\cos^2 \varphi}{2a_y^2}\right) \quad (10)$$

3.2 支持向量机拟合模型

支持向量机方法是20世纪90年代初Vapnik等^[17]根据统计学习理论提出的一种用于解决模式识别以及函数拟合等方面的智能算法。SVM能够有效地解决小样本、非线性、非平稳及高维函数拟合等方面的问题^[18]。因此,SVM智能方法在拟合具有局地突变特征(例如多核心、非平稳)的城市热岛拟合中存在较大优势。

SVM方法的核心在于升维和线性化。SVM通过引入核函数,将样本空间映射到高维特征空间中,从而使在起始的样本空间非线性可分转化为在高维特征空间中线性可分。SVM所引入核函数的优势并不需要非线性映射的显性表达式,而是直接在高维特征空间中建立线性学习机,因此不会增加计算量和复杂性。经过升维和线性化,低维空间的输入向量变换为高维空间的向量内积值,即:

$$K(x, y) = \phi(x) \phi(y) \quad (11)$$

常用的核函数有以下4种:

(1)线性核函数: $K(x, y) = x \cdot y$ 。

(2)多项式核函数: $K(x, y) = [(x \cdot y) + 1]^d$ 。

(3)径向基函数(RBF): $K(x, y) = \exp(-|x - y|^2 / d^2)$ 。

(4)二层神经网络核函数(Sigmoid)
 $K(x, y) = \tanh(a(x \cdot y) + b)$ 。

由于RBF核函数计算量较小,适合多维非线性空间映射,并且核函数RBF为高斯函数形式,而高斯函数由于具有平滑及精准等特点而被广泛应用于具有城市热岛曲面的模拟。因此,在本研究中SVM方法采用RBF核函数对研究区域城市群热岛曲面进行拟合。

运用SVM进行非线性函数曲面拟合的算法如下:

给定城市热环境样本空间 $(x_i, y_i), x_i \in \mathbb{R}^n, y_i \in \mathbb{R}^n, i=1, \dots, n$ 。其中: j_i 代表城市地理坐标向量; y 表示地表温度向量; n 为城市地理坐标向量的维数; 设线性回归函数为: $f(x) = w^T \cdot x + b$, 则SVM中最佳曲面拟合函数可以转化为求解下列函数的极值:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (12)$$

$$\begin{cases} f(x_i) - y_i \leq \xi_i^* + \varepsilon, & i=1, \dots, n \\ f(x_i) - y_i \leq \xi_i + \varepsilon, & i=1, \dots, n \\ \xi_i, \xi_i^* \geq 0, & i=1, \dots, n \end{cases} \quad (13)$$

式中: w 为广义线性函数的斜率; C 为惩罚因子; ξ_i 和 ξ_i^* 为松弛变量; ε 为精度控制因子。

使用拉格朗日乘数法将原问题变换为对偶问题,经过求解对偶问题,拟合函数可简化成以下形式:

$$f(x) = \sum_{i=1}^n (a_i - a_i^*) k(x, x_i) + b \quad (14)$$

式中: a_i 和 a_i^* 为对应的样本支持向量; $K(x_i, x)$ 为核函数。

当选定相应的SVM模型后,模型中惩罚性因子 C , 以及核函数变量 g , 成为影响模型精度的关键性变量。目前主要通过“KCV-交叉验证法(K-fold cross-validation, K-CV)”结合“网络搜寻法”来进行关键性变量的取值。

本文基于MATLAB 10环境,选用LIBSVM-3.20软件进行城市群热环境曲面拟合,具体建模过程如下:①编译,在MATLAB 10环境下编译LIBSVM-3.20软件包;②核函数选取,本文选取径向基函数(RBF)作为SVM中的核函数,原因在于其适用多维复杂环境、计算量小、函数形式为高斯形式且适合城市热岛曲面的模拟;③参数优化,主要选用“KCV-交叉验证法”结合“网络搜寻法”对模型中关键性参数(C, g)进行优化模拟;④模型训练及预测,选用LIBSVM中模型训练及预测模块,结合研究区域热环境数据进行城市群热环境的曲面拟合;⑤输出,在前面建模的基础上进行MATLAB绘图便可得到该城市群中SUHI二维或三维分布图。

对于单个城市热环境空间曲面模拟,SVM方法具有“曲面整体把握性”以及“曲面细节刻画性”据对现有模型进行参数率定和模型改进,通过敏感性实验论证改进模型的精确性,构建能够充分表现城市群热环境动态变化特征的空间形态曲面模型,提高现有模型的精度是解决城市群热环境空间形态模拟的关键。

3.3 模型验证

在研究区域范围内按照 $24 \text{ km} \times 24 \text{ km}$ 绘制网格,在每个网格内随机选取1个数据(共计208个数据)作为验证值,不参与模型拟合运算。模型验证方法主要通过分差系数 f 表示,并结合回归系数 R^2

进行计算,具体如下:

$$f = \frac{V_{\text{mdl}} - V_{\text{obs}}}{V_{\text{mdl}} + V_{\text{obs}}} \quad (15)$$

式中: V_{mdl} 表示拟合数据; V_{obs} 表示观测数据。当 $|f|$ 小于 0.33, 表示模拟误差小于 2, 在可接受范围之内。

模型精度主要通过均方差根误差 (RMSE), 即模拟值与真值偏差的平方和观测次数 n 比值的平方根) 来表示。

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{\text{obs}, i} - X_{\text{mdl}, i})^2}{n}} \quad (16)$$

式中: X_{obs} 为原始观测输入的数据; X_{mdl} 为模型模拟数据; n 为数据的个数。

3.4 城市群热环境分级

为定量测算城市群热环境时空形态, 本研究利用京津冀城市群热环境模拟空间曲面作为输入数据, 因不同季节地表温度有明显差异, 为揭示其空间分布特征规律, 首先需要将地表温度正规化处理并通过密度分割技术将城市热环境进行分级处理^[19]。

$$T_{ni} = \frac{T_{si} - T_{s \min}}{T_{s \max} - T_{s \min}} \quad (17)$$

式中: T_{ni} 表示第 i 个像元正规化后的值; T_{si} 为第 i 个像元的地表温度; $T_{s \max}$ 表示京津冀城市群地表温度最大值; $T_{s \min}$ 表示京津冀城市群地表温度最小值。

通过密度分割技术将地表温度划分为低温、次低温、中温、次高温和高温 5 个热力等级 (表 1)。

表 1 地表温度等级区间划分标准

Tab. 1 Criteria for dividing the level of surface temperature

地表温度等级	温度范围
低温	$T_{ni} < T_{\text{mean}} - 1.5s$
次低温	$T_{\text{mean}} - 1.5s \leq T_{ni} < T_{\text{mean}} - 0.5s$
中温	$T_{\text{mean}} - 0.5s \leq T_{ni} < T_{\text{mean}} + 0.5s$
次高温	$T_{\text{mean}} + 0.5s \leq T_{ni} < T_{\text{mean}} + 1.5s$
高温	$T_{ni} \geq T_{\text{mean}} + 1.5s$

注: T_{ni} 为正规化后的像元值; T_{mean} 为正规化后所有像元平均值。

3.5 城市群热环境重心测算

重心是用于描述地理对象空间分布位移的一个重要指标^[20-21]。重心表示地理要素空间分布的倾向性, 也反映了整体上的非均衡性和“高密度”部位。重心动态转移反映了地理要素分布的总体变化趋势。热岛区域重心是城市热岛效应发展的重要空间特征指标, 在城市热环境调控决策过程中有

重要的应用价值。对于单个地理对象的重心坐标, 可用以下公式计算:

$$X_t = \frac{\sum_{i=1}^n (C_{ii} \times X_i)}{\sum_{i=1}^n C_{ii}} \quad (18)$$

$$Y_t = \frac{\sum_{i=1}^n (C_{ii} \times Y_i)}{\sum_{i=1}^n C_{ii}} \quad (19)$$

式中: X_t 、 Y_t 分别为 t 时刻城市热岛区域重心; X_i 、 Y_i 分别为第 i 块城市热岛区域的几何中心坐标; C_{ii} 为第 i 个方位区面积。

在此基础上, 进一步计算重心转移距离 (L) 和重心转移角度 (α) 2 个变量。转移距离指示了不同时期城市热岛的均衡程度, 偏离方向指示了城市热岛的“高密度”部位。

$$L_{t+1} = \sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2} \quad (20)$$

式中: L_{t+1} 表示从 t 到 $t+1$ 时刻城市热岛重心转移距离。

$$\alpha_{t+1} = \begin{cases} \arctan\left(\frac{y_{t+1} - y_t}{x_{t+1} - x_t}\right) & x_{t+1} \geq x_t, y_{t+1} \geq y_t \\ \pi - \arctan\left(\frac{y_{t+1} - y_t}{x_{t+1} - x_t}\right) & x_{t+1} < x_t, y_{t+1} \geq y_t \\ \pi + \arctan\left(\frac{y_{t+1} - y_t}{x_{t+1} - x_t}\right) & x_{t+1} < x_t, y_{t+1} < y_t \\ 2\pi + \arctan\left(\frac{y_{t+1} - y_t}{x_{t+1} - x_t}\right) & x_{t+1} \geq x_t, y_{t+1} < y_t \end{cases} \quad (21)$$

式中: α_{t+1} 表示从 t 到 $t+1$ 时刻城市热岛转移方向与正东方向的夹角。

4 结果与分析

4.1 京津冀城市群热环境曲面拟合

本研究采用二维高斯曲面拟合方法和支持向量机方法对京津冀城市群进行热环境曲面模拟。由图 2 可得知, 2 个模型的预测结果均在可接受范围之内 ($|f| < 0.33$)。这表明 2 个模型从整体上均较好地拟合出城市群地面温度。同时, 支持向量机模型拟合精度远大于高斯二维模型的拟合精度即支持向量机模型拟合误差系数 $|f| = 0.016$ 小于二维高斯模型拟合误差系数 $|f| = 0.043$ 。从拟合数据的相关性来看, 支持向量机模型中相关系数平方 ($R^2 = 0.906$) 远大于二维高斯模型拟合的验证数据 ($R^2 = 0.41$)。这表明相对于二维高斯模型来讲, 支持向量机模型对于城市群热环境的模拟在局部细节上也同样具有较高的精度。以上结果表明, 对于京

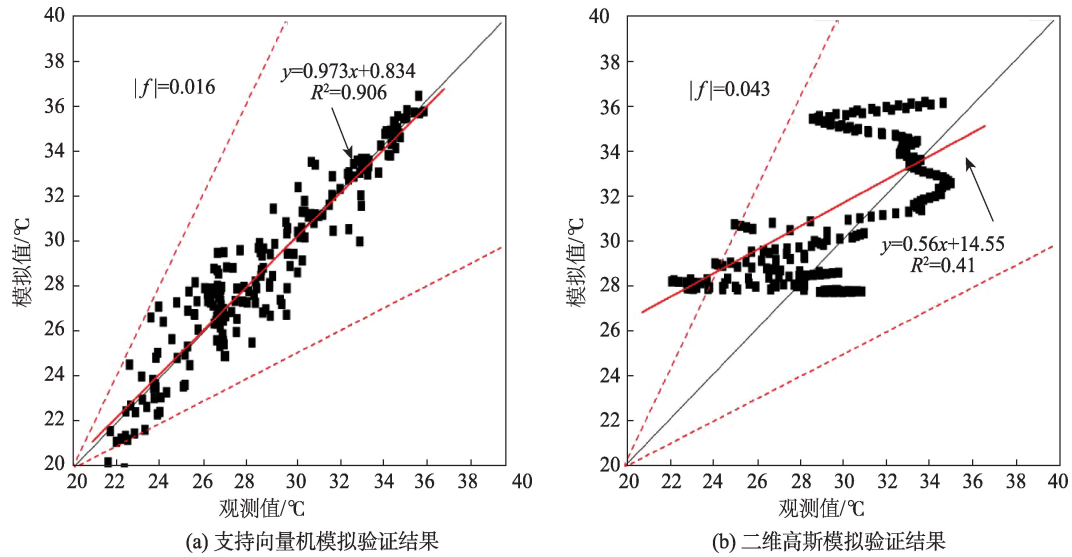


图2 模型模拟验证结果

Fig. 2 Verification of simulation results

京津冀城市群热环境的模拟,无论是在整体还是在局部的拟合方面,支持向量机模型比高斯二维模型的模拟精度要高。

为进一步验证2个模型在京津冀地区热环境空间模拟的适用性,本研究分别将整个区域数据输入2个模型,得到京津冀地区热环境空间二维图(图3)及三维分布图(图4)。图3和图4同样验证了模型验证时期的结论(图2)。从整体上来看,支持向量机模型的拟合值与观测值的均方差根误差(Root Mean Square Error, RSME)为1.264,而二维高斯模

型拟合结果($RSME=3.217$)远低于该结果,同样证明支持向量机在京津冀热环境空间拟合的准确性。从热环境空间分布拟合来看,支持向量机模型拟合结果与现实情况符合较好($R^2=0.837$,图3),而二维高斯模型拟合结果较差,不能反映该区域热环境的空间分布($R^2<0.01$,与观测值不相关)。从空间热环境曲面构建结果(图5)来看,支持向量机模型曲面重构结果能够反映现实环境中的热环境空间分布,而二维高斯模型拟合结果显示关键性参数 a_i 值在模型中的优化结果为0,难以构建完整的曲面。

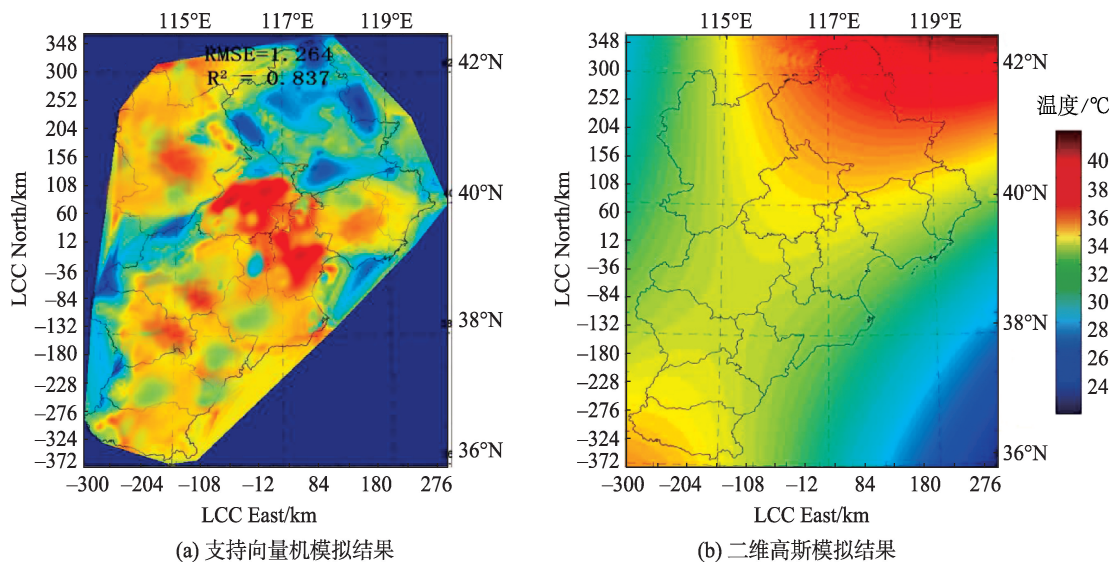


图3 模型模拟的二维结果

Fig. 3 Two dimensional simulation results

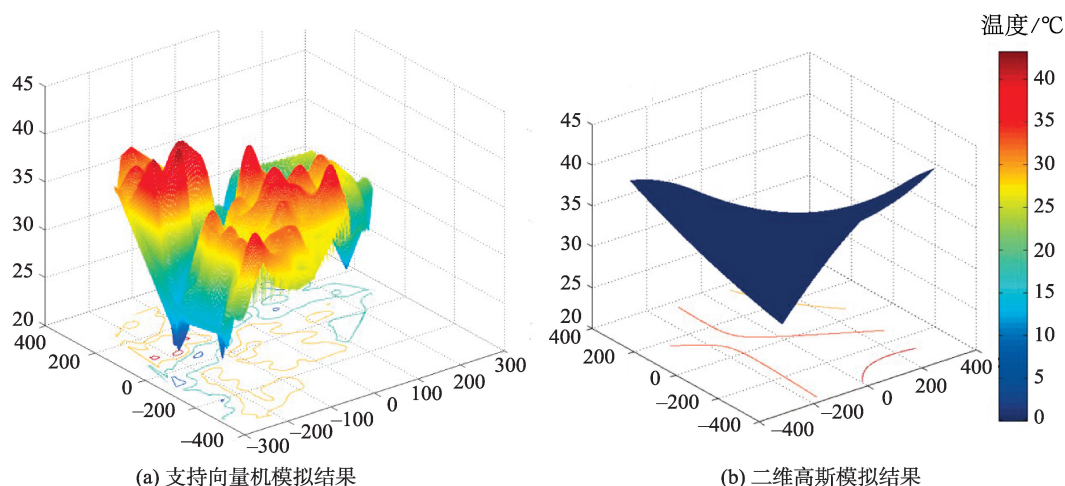


图4 模型模拟的三维结果

Fig. 4 Three dimensional simulation results

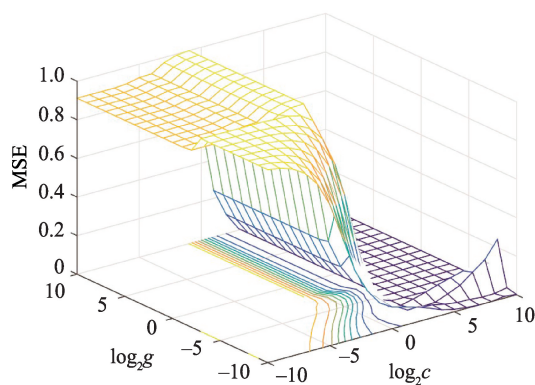


图5 支持向量机参数优化

Fig. 5 Support vector machine parameter optimization

究其原因:① 高斯曲面模型要求数据整体上符合正态分布(即存在单个热岛核心)。因此,该模型在城市与乡村集聚区域差异明显的小型单一热岛核心的城市拟合效果较好。而对于多个城市集聚区域(常常具有多个热岛核心)数据波动性较大,没有明显的正态分布趋势,空间分布拟合结果较差。但支持向量机模型通过高斯核函数致使单一曲面分解为线性加和的多个曲面,从而对具有多个热岛核心的京津冀城市群整体曲面拟合较好;② 京津冀城市群存在多种土地利用交错相间的情况,而且夏季不同土地利用的热容性质是造成区域热环境差异的重要因素,交错相间的土地利用类型致使该区域热环境输入数据波动性较大,影响二维高斯曲面的模拟效果。但支持向量机通过高斯核函数能够进行局部优化,然后再进行曲面线性加和,从而能够在局部进行数据滤波的效果,使结果更精确;

③ 由于研究区域内受云层、海洋等环境因素影响而造成的缺省值区域被设置为空值、不纳入模型模拟计算,致使部分拟合数据在空间上不连续,影响了高斯模型的拟合效果。但支持向量机模型能够根据热环境整体空间布局,通过高斯核函数进行局部优化差值,最大限度减少缺省值对模型拟合结果的影响。

综上所述,通过对支持向量机模型和二维高斯模型在模型验证结果以及空间拟合精度的比较发现,支持向量机模型不仅能在整体上较好地模拟具有多核心、多种土地利用类型的京津冀城市群热环境的空间分布,而且能够根据热环境的整体空间布局通过高斯核函数进行局部优化差值,最大限度减少缺省值对模型拟合结果的影响。

4.2 支持向量机模型参数优化

由支持向量机模型原理可知,控制支持向量机的拟合精度最重要的2个参数为:惩罚性因子 C 和径向基函数(RBF)即高斯核函数中参数 g 。为了提高支持向量机的模拟精度,本研究采用“KCV-交叉验证法”结合“网络搜寻法”对模型中关键性参数(C, g)进行优化模拟。在MATLAB中 C 及 g 的寻值范围设置为 $2n$,其中 n 取值范围为 $-10 \sim 10$,步长间隔为1,同时将输入数据集分为6个子集进行交叉验证,即 $V=6$ 。参数优化结果见图5。由图5可发现,参数 C 值在 $(2 \sim 10) \sim 23$ 之间取值时模型均方误差(Mean Squared Error, MSE)远大于 C 值在 $24 \sim 210$ 之间。对于高斯核参数取值最优范围为 $(2 \sim 5) \sim (2 \sim 8)$ 之间。整个模型模拟值与验证值的均方根

误差最小值时, C 值为 64 即 $\log_2 C = 6$, g 值为 0.0078, $\log_2 g = -7$, 此时的 C, g 值为模型最优参数值。

4.3 分辨率对曲面拟合效果的影响

在进行曲面拟合的过程中, 拟合精度和拟合时间是衡量拟合结果的重要指标, 而原始影像的分辨率则是影响该指标的决定性因素。为了衡量原始影像分辨率对曲面拟合效果的影响, 本研究将原始 MODIS 数据在 ArcGIS 中进行重采样, 分辨率为 3、6、12、和 24 km, 热值数据取其平均值, 并将其在支持向量机模型中进行建模, 从曲面拟合结果可知: 随着遥感影像分辨率的降低, SVM 模拟精度大幅下降。其中, 拟合值与观测值之间的均方根误差 (RMSE) 由 1 km 分辨率时的 1.264 上升到 1.449 (3 km)、2.23 (6 km)、3.764 (12 km)、4.664 (24 km), 分别升高了 14.63%、76.42%、197.78% 和 253.16%。拟合值与观测值之间的相关性由 1 km 分辨率时的 0.837 降到 0.803 (3 km)、0.751 (6 km)、0.648 (12 km)、0.589 (24 km), 分别降低了 4.06%、10.28%、22.58% 和 29.63%。同时, 支持向量机拟合所需时间大大降低, 由 1 km 分辨率时的 129 min 降低到 (3 km)、22 min (6 km)、10 min (12 km)、3 min (24 km), 分别降低了 74.42%、82.17%、92.25% 和 97.67%。

4.4 京津冀城市群热环境时空形态变化

根据上文对京津冀城市群的热环境模拟分析可知河北省的唐山、保定、石家庄、沧州、廊坊 5 个典型城市的热岛效应虽然呈现上升趋势, 但整体来看变化幅度并不大, 而北京市和天津市作为京津冀城市群的核心城市所呈现的热岛效应最为明显。对

北京市和天津市热环境时空形态变化进行测算后的结果显示: 2003 年北京市城市热岛区域面积为 1096.98 km², 天津市城市热岛区域面积为 370.32 km²。2003–2013 年北京市与天津市的城市热岛区域面积分别增加了 7.46 倍与 12.33 倍。10 年间北京市和天津市的城市热岛面积增长率分别为 709.11 km²/a 和 349.77 km²/a。而从空间分布形态来看, 北京市和天津市城市热岛区域从空间上呈现逐年接近连片发展的趋势。

2003–2013 年北京市和天津市城市热岛重心移动轨迹具有明显的时空分异性 (表 2)。2003–2006 年, 北京市城市热岛重心向东 (东北或东南) 缓慢移动, 2008 年以后主要向东南方向移动, 移动速度明显加快, 其中 2012–2013 年城市热岛重心移动 3.13 km。天津市城市热岛重心移动速度较北京缓慢, 重心移动方向以南方 (东南或西南) 为主。对比 2 个城市热岛时空格局特征, 北京市城市热岛特征为东南部地区异速增长, 西部地区缓慢增长; 天津市城市热岛特征表现为以城市中心为圆心向周围扩展。

5 结论与讨论

本文通过对比高斯曲面模型和支持向量机拟合模型的算法和优缺点, 遴选城市群热环境空间形态模拟的最佳拟合模型, 并对 2003–2013 年 MODIS-LSTs 地表温度数据、支持向量机拟合模型等进行参数优化和曲面拟合分辨率设置, 揭示了 2003–2013 年的京津冀城市群主要城市的热环境空间分布特征规律。主要结论如下:

表 2 2003–2013 年北京市和天津市城市热岛重心移动距离和移动方位角

Tab. 2 The moving distance and azimuth of the center of gravity of urban heat island in Beijing and Tianjin from 2003 to 2013

时段	北京市		天津市	
	移动距离/km	移动方位角/°	移动距离/km	移动方位角/°
2003–2004	0.36	302.72	0.46	287.07
2004–2003	0.74	17.84	0.26	270.35
2003–2006	0.78	200.71	0.16	226.91
2006–2007	0.57	29.83	1.44	350.37
2007–2008	0.87	328.22	0.38	351.31
2008–2009	1.36	0.86	0.73	228.19
2009–2010	2.26	333.89	0.95	227.89
2010–2011	0.85	333.66	0.31	339.49
2011–2012	2.39	332.72	0.45	74.08
2012–2013	3.13	345.89	0.59	255.50

(1)支持向量机(SVM)模型可以较好地模拟城市群热岛的空间格局分布。对于京津冀城市群热环境的模拟,无论是在整体还是在局部的拟合方面,SVM模型都比高斯二维模型的模拟精度高。高斯曲面模型对于多个城市聚集区域没有明显的正态分布趋势,空间分布拟合结果较差。SVM模型不仅能在整体上较好地模拟具有多核心、多种土地利用类型的京津冀城市群热环境的空间分布,而且能够根据热环境的整体空间布局通过高斯核函数进行局部优化差值,最大限度减少缺省值对模型拟合结果的影响。同时,研究结果还证明了SVM模型对于刻画京津冀城市群这种囊括多种地貌特征的区域热岛效应理论与实践方面的可行性。

(2)遥感影像分辨率是影响SVM拟合精度和拟合时间的重要因素。研究发现:拟合精度和拟合时间作为衡量拟合结果的重要指标,分辨率正是影响该指标的决定性因素。随着遥感影像分辨率的降低,SVM的模拟精度明显下降。3、6、12、和24 km时的分辨率与1 km时分辨率的各数值相比,拟合值与观测值之间的均方根误差分别升高了14.63%、76.42%、197.78%和253.16%。拟合值与观测值之间的相关性分别降低了4.06%、10.28%、22.58%和29.63%。同时SVM拟合所需时间大大降低,分别降低了74.42%、82.17%、92.25%和97.67%。

(3)通过模拟京津冀城市群2003–2013年的热环境时空形态变化发现:在京津冀城市群区域内,河北省的唐山、保定、张家口、沧州、廊坊5个主要城市的热岛效应变化幅度不明显,而北京市和天津市的热岛效应变化较为显著。北京市和天津市城市热岛区域大幅增加,重心移动轨迹具有明显的时空分异性。对比2个城市热岛时空格局特征发现:北京市城市热岛特征为东南部地区异速增长,西部地区缓慢增长;天津市城市热岛特征为以城市中心为圆心向周围扩展,这与天津市近年来“摊大饼”式的城市建设发展模式密切相关。

本研究进一步丰富了城市生态环境评测的定量方法,并且可以在实践上对城市群的城市规划、城市建设、环境保护和区域可持续发展等提供量化、可视化的决策支持。但本研究仍存在一定的不足,特别是在数据精度方面,MODIS数据在刻画较为细致的空间特征方面还存在一定的局限性,而Landsat-8、Sentinel-2等新型卫星生产的LST产品可以较好地克服这些问题,基于此类数据的应用将是后续研究中的一个重要发展方向。

参考文献(References):

- [1] 刘伟东,尤焕苓,孙丹. 1971-2010年京津冀大城市热岛效应多时间尺度分析[J]. 气象,2016,42(5):598-606. [Liu W D, You H L, Sun D. Multi-time scale analysis of megacities heat island effect in Beijing-Tianjin-Hebei region from 1971 to 2010[J]. Meteorological Monthly, 2016,42(5):598-606.]
- [2] 陈彬辉,冯瑶,袁建国,等. 基于MODIS地表温度的京津冀地区城市热岛时空差异研究[J]. 北京大学学报(自然科学版),2016,52(6):1134-1140. [Chen B H, Fen Y, Yuan J G, et al. Spatiotemporal difference of urban heat island in Jing-Jin-Ji area based on MODIS land surface temperature[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2016,52(6):1134-1140.]
- [3] 金鑫,金彦香,杨登兴. SWAT模型在土地利用/覆被变化剧烈地区的改进与应用[J]. 地球信息科学学报,2018,20(8):1064-1073. [Jin X, Jin Y X, Yang D X. Improved SWAT and its application to a region with severe land use/land cover changes[J]. Journal of Geo-information Science, 2018,20(8):1064-1073.]
- [4] Camila M, Nakata-Osaki, Léa Cristina L S, et al. This tool for heat island simulation: A GIS extension model to calculate urban heat island intensity based on urban geometry[J]. Computers, Environment and Urban Systems, 2018,67:157-168.
- [5] 彭保发,石忆邵,王贺封,等. 城市热岛效应的影响机理及其作用规律——以上海市为例[J]. 地理学报,2013,68(11):1461-1471. [Peng B F, Shi Y S, Wang H F, et al. The impacting mechanism and laws of function of urban heat islands effect: A case study of Shanghai[J]. Acta Geographica Sinica, 2013,68(11):1461-1471.]
- [6] Chen A L, Zhao X F, Yao L, et al. Application of a new integrated landscape index to predict potential urban heat islands[J]. Ecological Indicators, 2016,69:828-835.
- [7] Jose L F R, Filho A J Pereira, Karam H A. Estimation of long term low resolution surface urban heat island intensities for tropical cities using MODIS remote sensing data[J]. Urban Climate, 2016,17:32-66.
- [8] 侯浩然,丁凤,黎勤生. 近20年来福州市城市热环境变化遥感分析[J]. 地球信息科学学报,2018,20(3):385-395. [Hou H R, Ding F, Li Q S. Remote sensing analysis of changes of urban thermal environment of Fuzhou city in China in the past 20 years[J]. Journal of Geo-information Science, 2018,20(3):385-395.]
- [9] Sun Q, Wu Z, Tan J. The relationship between land surface temperature and land use/land cover in Guangzhou, China[J]. Environmental Earth Sciences, 2012,65(6):1687-1694.

- [10] 陈云浩,史培军,李晓兵,等.城市空间热环境的遥感研究——热场结构及其演变的分形测量[J].测绘学报,2002,31(4):322-326. [Chen Y H, Shi P J, Li X B, et al. Research on urban spatial thermal environment using remote sensing image[J]. Acta Geodaetica et Cartographica Sinica, 2002,31(4):322-326.]
- [11] 张心怡,刘敏,孟飞.基于RS和GIS的上海城建用地扩展研究[J].长江流域资源与环境,2006,15(1):29-33. [Zhang X Y, Liu M, Meng F. Expansion of urban construction land in Shanghai City based on RS and GIS. Resources and Environment in the Yangtze Basin, 2006,15(1):29-33.]
- [12] 王福海,陈丹.建设用地扩展对城市热环境影响研究[J].中国人口·资源与环境,2016,26(S1):194-197. [Wang F H, Chen D. Temporal-spatial variation of the thermal environment impact research by distribution characteristics of construction land[J]. Chinese Journal of Population Resources and Environment, 2016,26(S1):194-197.]
- [13] 张丽娟,于洋,栗练灵,等. 1960-2010年黑龙江省土地利用变化对气温的影响[J].地理科学,2017,37(6):952-959. [Zhang L J, Yu Y, Li L L, et al. Effects of land use change on air temperature of Heilongjiang Province in 1960-2010 [J]. Chinese Geographical Science, 2017,37(6):952-959.]
- [14] Liu X, Wang C L, Jing Y S, et al. Annual variations of temperature on four urban underlying surfaces and fitting analysis[J]. Journal of Tropical Meteorology, 2013,19(2): 138-144.
- [15] Pramod N Chivate, Andrei G Jablokow. Review of surface representations and fitting for reverse engineering[J]. Computer Integrated Manufacturing Systems, 1995,8(3): 193-204.
- [16] 占文凤,陈云浩,周纪等.支持向量机的北京城市热岛模拟:热岛强度空间格局曲面模拟及其应用[J].测绘学报,2011,40(1):96-103. [Zhan W F, Chen Y H, Zhou J, et al. Spatial simulation of urban heat island intensity based on the support vector machine technique: A case study in Beijing[J]. Acta Geodaetica et Cartographica Sinica, 2011, 40(1):96-103.]
- [17] Vapnik V. The nature of statistical learning theory[M] New York: Springer-Verlag,1995.
- [18] 丁世飞,齐丙娟,谭红艳.支持向量机理论与算法研究综述[J].电子科技大学学报,2011,40(1):2-10. [Ding S F, Qi B J, Tan H Y. An overview on theory and algorithm of support vector machines[J]. Journal of University of Electronic Science and Technology of China, 2011,40(1):2-10.]
- [19] 乔治,田光进.北京市热环境时空分异与区划[J].遥感学报,2014,18(3):715-734. [Qiao Z, Tian G J. Spatiotemporal diversity and regionalization of the urban thermal environment in Beijing[J]. Journal of Remote Sensing, 2014, 18(3):715-734.]
- [20] 郭兵,陶和平,刘斌涛,等.基于GIS、Delaunay三角形的复杂地形区经济重心获取新方法[J].地理与地理信息科学,2012,28(2):55-59. [Guo B, Tao H P, Liu B, et al. Research on acquisition method of economic gravity center based on GIS and Delaunay in complicated terrain area [J]. Geography and Geo-information Science, 2012,28(2): 55-59.]
- [21] 王劲峰,葛咏,李连发,等.地理学时空数据分析方法[J].地理学报,2014,69(9):1326-1345. [Wang J F, Ge Y, Li L F, et al. Spatiotemporal data analysis in geography[J]. Acta Geologica Sinica, 2014,69(9):1326-1345.]