引用格式:姚可桢,岳书平.网络大数据下的中国现代食甜习惯空间分布特征及其影响因素研究[J].地球信息科学学报,2020,22(6):1202-1215.
[Yao K Z, Yue S P. Study on spatial distribution of modern sweet diet and its impact factors in China based on big data from internet[J]. Journal of Geo-information Science, 2020,22(6):1202-1215.] **DOI:**10.12082/dqxxkx.2020.190432

网络大数据下的中国现代食甜习惯空间分布特征及其影响因素研究

姚可桢,岳书平*
南京信息工程大学地理科学学院,南京 210044

Study on Spatial Distribution of Modern Sweet Diet and its Impact Factors in China based on Big Data from Internet

YAO Kezhen, YUE Shuping*
School of Geographical Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China

Abstract: As the most important element with local characteristics in regional culture, dietary geographical culture presents a new diversified situation under the background of large scale population movement. However, up to now, the domestic research on the distribution characteristics of sweet diet based on traditional cognition is still lack of objective data. Based on the web crawler technology, this paper obtained about 20 million pieces of gourmet consumption data in 31 provincial capitals in mainland China. The degree of sweetness in our diets in urban areas was calculated for traditional dishes, main food dishes, drinks, and dessert dishes. Based on ArcGIS and MySQL softwares, spatial analysis and mathematical statistics were used to understand the spatial distribution characteristics of the modern Chinese sweet diet and identify its affecting factors. The results show that there were dramatically regional differences in the spatial distribution of sweet diet in China, especially in the southeastern coastal areas and the central inland areas, with the evaluation parameter (R^2) of spatial grouping analysis reaching 0.88. The distribution of modern sweet diet generally presented a surrounding pattern of "High East, Middle North, Micro-low West and Low Inside". From either the overall or local point of view, the Moran indexes were positive at 1% significance level, and there was a significant positive spatial autocorrelation for sweet diet habits at different areas in China rather than an obvious trend of dispersion. There were three distinct geographical agglomeration areas: the high-sweetness agglomeration areas along the southeast coast of Jiangsu, Zhejiang, Shanghai, and Fujian, the low-sweetness agglomeration areas in southwest areas of Chongqing, Guizhou, and Sichuan, and the northwest inland low-sweetness agglomeration areas in Shanxi and Ningxia. The accuracy of the stepwise regression model of sweet diet habit distribution was 0.82, and results suggest that meteorological elements such as precipitation, humidity, temperature, and geographical location were important factors that influence the spatial distribution of sweet diet habit in modern China. Moreover, we found that geographical

收稿日期:2019-08-08;**修回日期:**2020-01-16.
基金项目: 国家自然科学基金项目(41901355);江苏省自然科学基金青年项目(BK20160953);江苏省一流本科专业项目。
[**Foundation items:** National Natural Science Foundation, No.41901355; Natural Science Foundation of Jiangsu Province, No.BK20160953; First class undergraduate program of Jiangsu Province.]
作者简介:姚可桢(1997—),男,江苏南通人,硕士生,主要从事网络大数据和GIS应用研究。E-mail: kezhen_yao@qq.com
***通讯作者:**岳书平(1979—),女,山东德州人,博士,讲师,主要从事大数据分析与GIS应用及GIS在土壤侵蚀中的应用研究。E-mail: yueshuping@nuist.edu.cn

location had a regulating effect on the influence of sunshine duration on sweet diet habit. Specifically, the sweetness of inland cities generally increased with the increase of sunshine duration, while the sweetness of coastal cities usually decreased with the decrease of sunshine duration. This study aims to reveal the regional disparity of sweet culture in modern China, which provides reference for the planning of the category structure in urban catering industry and better understanding of the new trend of the development of modern sweet food consumption.

Key words: sweetness of food in urban; big data; web crawler; spatial autocorrelation analysis; hot spot analysis; spatial grouping analysis; stepwise regression analysis; GIS

***Corresponding author:** YUE Shuping, E-mail: yueshuping@nuist.edu.cn

摘要：饮食地理文化作为地域文化中最具地方特色的重要元素，在现代人口大规模流动背景下呈现出全新的多样化局面，而基于传统认知的“南甜北咸”的地域分异已然不能代表中国现代食甜分布的空间特征。因此，本文采用网络爬虫技术，获取我国大陆31个省会城市共计约2000万条美食消费数据，从传统类菜品、主食类菜品、饮料类和甜品类菜品4个方面计算城市食甜度，在ArcGIS、MySQL软件支持下，借助GIS空间分析和数理统计方法探究我国现代食甜习惯的空间分布特征，分析影响食甜分布的因素。研究发现：① 中国食甜在空间分布上存在显著的地域分异特征，聚类分析评价参数R^2高达0.88，现代食甜习惯总体呈现“东高北中，西微内低”的包围式格局；② 从整体抑或局部角度，在1%显著性水平上莫兰指数均为正，中国食甜分布呈现显著的空间正相关关系，形成特色鲜明的3个地理集聚区，即以苏浙沪闽为主的东南沿海高甜集聚区，以渝黔川为主的西南内陆低甜集聚区和以陕宁为主的西北内陆低甜集聚区；③ 构建了中国现代食甜习惯分布影响因素模型，其拟合精度为0.82，分析结果显示降水、湿度、气温等气象要素及地理位置是影响现代我国食甜空间分布的重要因素。

关键词：城市食甜度；大数据；网络爬虫；空间自相关分析；热点分析；空间聚类分析；逐步回归分析；GIS

1 引言

中国地域幅员辽阔，气候和地貌条件差异很大，导致饮食文化存在显著的地域分异特征，在烹调与口味方面表现尤为突出。作为5种基本味觉之一的甜，不仅能带来味蕾上的愉悦，更是一种诠释美好的文化象征。因此，甜食在中国具有广大的消费群体和旺盛的消费需求。关于食甜的传统地理认知上，以南方喜甜为主流认识，八大菜系中苏菜、浙菜口味偏甜。但是，随着经济的快速发展和人口的大量流动，原有的饮食习俗与观念也随之发生了巨大变化。

作为人地关系的产物，饮食文化的形成及其时空变化与地理环境、民族特色密不可分。王汐牟等[1]通过分析宋元明时期的饮食谱录，探讨不同时代的社会饮食风尚及其时代特征。丁继月[2]以历史变迁为视角，探索了大连建市至今五个历史时期餐饮业的变化，并对不同时期的菜系特色、著名餐馆、影响因素及其空间分布等进行了总结。陈传康[3]通过对比研究中国四大菜系的形成与发展，阐述了中国饮食文化的区域特征与地域分化的时空发展规律。近年来，随着大数据时代的到来，利用真实可信的大数据样本进行饮食文化的相关研究也已取得一定成效。秦萧等[4]利用爬取的南京市大众点评网的用户点评数据，建立口碑评价指标体系，通过核密度分析方法划分南京餐饮业的空间分布格局；刘琪等[5]对采用大众点评网的餐饮商家信息数据，基于ArcGIS的核密度分析和综合评价对深圳市餐饮业整体和餐饮分级下的空间分布格局进行了研究。除此之外，夏令军等[6]将视角放眼全国335个地级及以上城市，详细分析了中国餐饮业在市级尺度的空间分布格局、空间集聚特征及其影响因素。顾秋实等[7]以全国范围内的餐馆为主要数据源，以文化地理学为视角，采用GIS核密分析和热点探测模型为主要方法，探索传统八大菜系的地理分布特征及其空间扩散效应。与此同时，鉴于餐饮大数据存在数据质量参差不齐的问题，黄恒君等[8]将多个团购网站和GIS等异源异构数据整合，给出餐饮业名录库更新的数据质量评估实例；秦之湄[9]等采用数据清洗和数据应用的方法，将城市餐饮业排放源按照POI的分布进行了高时空分辨分配，探索出餐饮源高时空分辨清单新方法。

从饮食口味角度出发，蓝勇[10]将中国传统食辣划分为长江上中游辛辣重区、北方微辣区和东南沿海淡味区共3个地区；张仁军等[11]探讨了中国现代食辣的空间分布及影响分布的因素，找出食辣的3个集聚区，同时指出日照和地理位置对食辣的空间分布特征起重要作用。目前，关于食甜的研究主

要集中在健康医学领域和甜食作物的种植技术等方面，如分析甜食膳食模式与人体健康的关系[12]、青少年及老年人的饮食健康问题[13-14]、利用GIS手段研究适合甜橙生长的气象条件等[15]。当前对于食甜地域分异的认识，仍以“南甜北咸，东辣西酸”的传统认知占据主流。然而，随着社会经济的快速发展和人口流动呈现常态化趋势，不同饮食口味在空间上广泛传播、交融，势必会影响中国食甜的空间分异。因此，基于传统认识的现代食甜习惯已无法科学客观地反映现代食甜的分布状况。另外，目前的研究主要基于大众点评网站的点评数据，很少或并未涉及到用户实际的消费数据，同时也未利用现今美食团购的2大主流平台——美团或饿了么的消费数据。因此，一定程度上说，所获取的大数据缺乏一定时效性，不能充分反映用户真实的饮食习惯。

现代城市外来人口流动大，势必会对当地饮食文化造成影响，而网络美食团购数据本身建立在多样化人口背景下，因此采用美食团购数据更能反映现代食甜习惯的特征。鉴于此，本文基于网络美食团购数据，以实际消费行为作为参考，研究现代城市人口在饮食口味上对甜这一味觉的喜爱程度。该“食甜”习惯不限于传统饮食的“菜甜”，同时包含对甜品、奶茶等现代甜食的延伸，是针对现代食甜习惯的一次全新探究。因此，文中以主流外卖平台“饿了么”上抓取到的2000多万条菜单消费数据为基础，以GIS空间技术分析方法和数理统计方法为依托，选取我国大陆地区31个省会城市的饮食消费数据，借助大数据海量、多样、真实的特性，探究我国现代食甜习惯的空间分布特征，探寻影响食甜习惯的可能原因。通过该研究，旨在客观有效地揭示现代社会变更下我国甜食文化的地域特征，进一步掌握我国食甜消费发展的新趋势，以期为城市餐饮行业门类结构的分布与规划提供参考依据。

2 数据源与数据处理

2.1 数据源

本研究所采用的数据源包括网络美食消费数据和基础地理数据。由于“饿了么”平台的消费份额占据网络餐饮市场第二位，并可根据用户经纬度定位附近商家，这有利于按城市进行定向爬取，故本研究选择“饿了么”平台作为城市现代食甜数据的来源。本文数据获取对象为全国大陆地区22个省份（台湾省及港澳台地区没有采集数据）、5个自治区的省会城市及4个直辖市共计31个地级以上城市，数据抓取时段为2019年1月1日—2019年1月31日，爬取内容包括：商家信息、菜谱信息以及销量数据，最终累计爬取约40万条商家数据，2000多万条菜单数据。

饮食作为人地关系的产物，其分布与地理环境及历史文化密不可分，因此地理环境的空间异化也是影响食甜分布的重要成因。为了识别影响中国现代食甜习惯空间分布的主要因素，本研究选取了气候要素、地理位置等环境要素。气候数据来源于中国气象数据网（http://data.cma.cn/）“中国地面累年值年值数据集（1981—2010年），包括气温、降水、日照时数等相关要素。该数据以各气象站点为基准获取，为了能更好地表达所研究城市的气候状况，文中选取与城市地理位置上相近的若干站点的数据均值作为该城市的气象要素数据。另外，本研究的区域范围为中国大陆地区（不包括台湾省及港澳台地区），采用国家测绘地理信息局标准地图服务网站（http://bzdt.ch.mnr.gov.cn/index.html）提供的中国行政界线数据，糖料播种面积数据、GDP数据来自中国统计年鉴（2019）。将上述数据进行几何纠正及投影转换，统一采用WGS1984大地坐标系，Albers等面积割圆锥投影，中央经线为110°E，双标准纬线为25° N和47° N。

2.2 爬取流程与数据结构

由于饿了么是基于用户位置定位附近商家，所以在爬取一个城市之前首先需要标记点位。本文通过HTML代码在地图上对每个城市自动打点，即以某经纬度坐标为圆心，2 km为半径做缓冲区，依次等间距排开，直至充满城市行政边界为止，具体爬取过程如图1所示。

由图1可以看出，前端网页爬取的同时需要实现后端数据批量入库。首先，根据点位表中点位顺

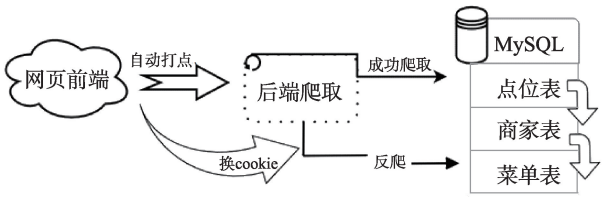


图1 “饿了么”2019年1月商家菜单数据抓取及存储流程

Fig. 1 Crawling and storing flow chart of merchant's menu data from Ele.me in January 2019

序爬取商家信息并存入商家表；其次，根据商家表中商家顺序爬取菜单信息并存入菜单表。其中，编程语言为java，在IntelliJ IDEA集成环境下开发，套用轻量级的Spring Boot快速开发平台—renren-fast v3.0。该开源项目前后端分离，后端部署的环境要求为JDK 1.8、数据库MySQL 5.5+。接口测试工具为Postman，登陆时生成token凭证，服务端解析token，并确定用户身份及用户权限，数据通过json进行交互。

为了提高数据的存储或检索效率，对于点位表，以id字段为主键，包含地址、经纬度等非空字段；对于商家表，以id字段为主键，包含商家id、点位id、商家名、商家评分、评分人数、图片路径、商家销量、商家地址、电话、经纬度、描述信息等非空字段；对于菜单表，以id字段为主键，包含菜单id、菜单名称、月销量、商家id、菜单评分、评分人数、描述信息等非空字段。表间可根据唯一的共同字段实现高效索引和增删改等操作。

2.3 菜单数据处理及食甜度计算方法

由于根据点位爬取到的商家表和菜单表中存在重复数据，故首先需要对两表进行去重操作，依据是具有唯一性的商家id和菜单id字段。去重后，按月销量字段对菜单表进行降序排列，发现来自不同城市的菜单表存在以下共性现象：① 表中出现许多不能反映菜肴特征但销量靠前的单样品菜单，如“金针菇”、“土豆片”、“羊肉串”等；② 表中出现许多主食类菜品且销量靠前，如“米饭”、“手工馍”、“苏州汤包”、“甜烧饼”等；③ 表中出现甜品、饮料类菜单，如“蛋糕”、“豆浆”、“奶茶”等。

鉴于共性现象①会对甄别菜单甜度产生干扰，所以剔除具有明显辛辣口味特征的商家菜单，这类商家的菜单基本都为单样品菜单，对于食甜研究不具备参考价值，如火锅店、麻辣香锅、冒菜、烧烤等。由共性现象②、③可看出，基于“饿了么”数据源的菜谱信息不全是诸如糖醋排骨、番茄炒蛋等类似的传统菜肴，实际上主食类、甜品类（主要为蛋糕、地方特色糕点等）、饮料类菜品（主要为豆浆、奶茶等零售店饮品，不包含具有批发性质的常见饮料，如可乐、果汁等）的销量不容小觑，这也反映了现代饮食对食甜需求和表现形式的多样化趋势，故在衡量食甜时应计入考虑。鉴于此，本文所关注的“食甜”不局限于通常理解的传统菜肴中的食甜习俗，同时包含甜品、奶茶等现代甜食的延伸，本质上是反映现代城市人口对纯粹甜味的喜爱程度。由于不同类别的菜品其甜度特征不一样，需采用不同的计算方式，故需通过数据库模糊查询筛选语句分类提取出主食类菜品、甜品类菜品、饮料类菜品，剩下的菜品定义为传统类菜品，将这4类菜品作为衡量食甜度的依据。

对于传统类菜品，将相同（同一品牌商家不同分店所致）或相近菜名合并，再按销量进行排名，最后筛选出排名前30名的菜品作为甜度衡量依据。对每道菜采用虚拟变量进行标识，0代表非甜菜，1代表甜菜。判断是否为甜菜的方法为：① 根据爬取到的介绍菜的口味及做法的字段信息（存储在数据库中）进行判断，② 通过“香哈菜谱”APP查询菜的具体做法来进行判断，如糖三勺。传统类菜品甜度用每个城市点餐数据中甜菜占总菜肴销售量的比率来表示，具体计算公式如下：

$$R_1=\frac{\sum_{i=1}^{30}S_i\times V_i}{\sum_{i=1}^{30}V_i} \tag{1}$$

式中：S_i表示第i个传统菜品的甜味特征且$S_i\in\{0,1\}$；V_i表示第i个传统菜品的销量；R_1表示传统类菜品的甜度指数。

对于主食类菜品，其甜度不能简单进行二值化处理，原因主要有2个方面：① 主食是碳水化合物特别是淀粉摄入的主要来源，从食物组分上看存有甜度但从味觉角度达不到“甜菜”的标准，比如汤包、馒头等淀粉类主食，存在一定的甜度特征，但甜度为1显然不合理，而糊辣汤等主食的甜度特征又基本为0；② 南北方对于同一主食的甜度把控也存在区别，如南方馒头加工的配料比中糖含量为15.0%[16]，而北方馒头加工配料中辅料较少[17]。故主食类菜品的甜度取值介于0~1之间，本研究采用平均甜度指数特征作为城市主食类菜品的甜度指数。平均甜度指数特征定义为：

$$R_2=\frac{\sum_{j=1}^{n}S_j\times V_j}{\sum_{j=1}^{n}V_j} \tag{2}$$

式中：S_j表示第j个主食菜品的甜味特征且$S_j\in[0,1]$；V_j表示第j个主食菜品的销量；n表示主食菜品的个数；R_2表示主食类菜品的平均甜度指数。

对于甜品类和饮料类菜品，如糕点、豆浆等，具有明显的甜度特征，故甜度指数为1，即R_3=R_4=1。

最终，城市总食甜度为4类菜品因子的甜度得分之和。其中，每个因子的甜度得分为该类的甜度指数与其权重之积，权重用该类菜品的销量占4类菜品销量之和的比率来表示，公式如下：

$$R=\sum_{t=1}^{4}R_t\times W_t \tag{3}$$

$$W_t=\frac{Q_t}{\sum_{t=1}^{4}Q_t} \tag{4}$$

式中：R表示城市总甜度指数；R_t表示第t个菜品因子的甜度指数；W_t表示第t个菜品因子的权重；Q_t表示第t个菜品因子的销量。由此可见，本文所研究的食甜习惯不限于传统食甜范围，而是保留城市餐饮特征的现代食甜习惯。整个数据处理流程如图2所示。

2.4 空间分析方法

2.4.1 空间自相关分析

空间自相关分析（Spatial Autocorrelation Analysis）以经典统计学为基础，用来检验在空间上具有一定规律性的空间变量在不同空间位置上潜在的依赖关系[18]。作为空间统计学的重要研究领域，以全局空间自相关分析和局部空间自相关分析相结合的研究方法在各类空间问题中得到深入运用[19]。因此，本文选取全局空间自相关、局部空间自相关及热点分析方法，深入剖析不同城市食甜度的空间集聚性，探究中国现代食甜习惯的空间分布特征。

（1）全局空间自相关

全局空间自相关分析用于验证区域整体的空间模式，主要对区域整体范围的内部关联特征进行分析[20]，判断该现象在空间是否有聚集特性存在。一般来讲，采用Moran's I指数进行衡量，其计算公式为：

$$I=\frac{n\sum_{i=1}^{n}\sum_{j=1}^{n}W_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n}W_{ij}\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{5}$$

式中：I为Moran指数；n为城市个数；x_i、x_j分别为第i个和第j个城市的甜度；$\bar{x}$表示所有城市甜度的均值；W_{ij}为空间权重矩阵。空间权重矩阵是观测对象之间空间依赖的正式表达，主要类型包括基于邻接关系和基于距离关系的权重矩阵，在不同的空间矩阵条件下自相关性可能会出现显著空间差异[21]。

Moran's I的取值范围是，正值代表正相关，负值代表负相关，零代表不相关。由于Moran's I没有显著性检验的功能，所以将其转化成正态检验统计量Z参照正态分布表进行显著性检验：

$$Z=\frac{I-E(I)}{\sqrt{VAR(I)}} \tag{6}$$

式中：$VAR(I)$为全局Moran's I的方差；$E(I)$为全局Moran's I的期望值[22]。

（2）局部空间自相关

局部空间自相关反映研究要素在一个区域单元上与其邻近单元的相关程度，主要对空间各子区域的关联特征进行分析，文中即表示城市之间食甜习惯相互集聚的倾向。本文采用LISA（Local Indicators of Spatial Association）对局部自相关进行分析，对某一特定城市单元：

$$I_i=\frac{(x_i-\bar{x})}{\sum_{i}(x_i-\bar{x})^2}\sum_{j}W_{ij}(x_i-\bar{x}) \tag{7}$$

若局部空间自相关显著，表明该城市食甜状况与其周边城市之间存在一定的空间关联程度或空间差异程度，具体空间关系包含4种[23]：“高—高相关”（H-H），即高属性区域同样被高属性区域围绕；“低—低相关”（L-L），即低属性区域仍被低属性区

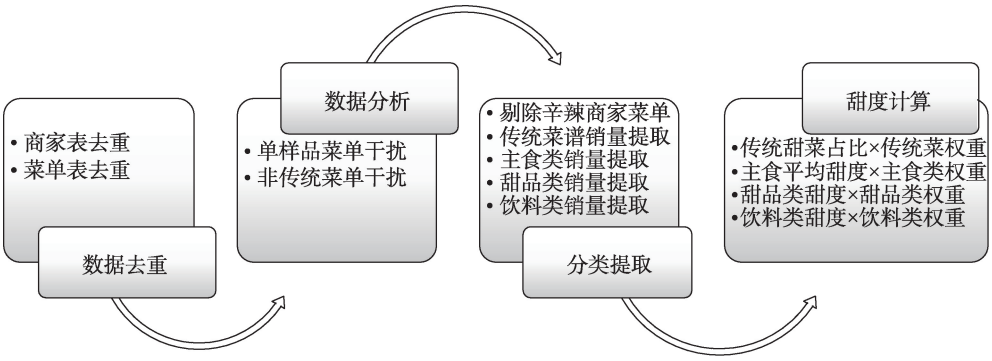


图2 “饿了么”2019年1月商家菜单数据处理流程

Fig. 2 Data processing flow chart of merchant's menu from Ele.me in January 2019

域围绕；“高—低相关”（H-L），即高属性区域被低属性区域围绕；“低—高相关”（L-H），即低属性区域被高属性区域围绕。同样可用Z值来检验局部Moran's I的显著性程度，方法如式（5）所示。在给定的显著性水平下，I_i的取值范围为[-1, 1]，正值代表正相关，负值代表负相关，零则代表不相关。

（3）热点分析

整体或者局部的Moran指数只能检测出具有相近值的要素是否呈现聚类，但并不能说明该聚类是否由高值或低值组成。换言之，在局部Moran中，高低值更容易引起关注，其周围可能存在的其他聚类情况可能会被忽视，不能判断是否具有统计学意义上的冷热点。而与Moran指数类似，G-统计量的局部版本Getis-Ord Gi*（常被称为热点分析）的Z值能有效识别冷热点区域，如果Z得分高且为正数，表明存在高值聚类或热点；如果Z得分低且为负数，表明存在低值聚类或冷点[24]。G_i^*的计算公式为：

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j} \right)^2 \right]}{n-1}}} \tag{8}$$

式中：x_j是要素j的属性值；$w_{i,j}$是要素i和j的空间权重；n为要素总数。且有：

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n} \tag{9}$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\bar{X}\right)^2} \tag{10}$$

按式（5）计算完成后显著性检验同样取决于p值和Z值，与Moran类似，p值表示数据的可靠性，Z值反映聚类情况，在此不再赘述。

2.4.2 空间聚类分析

空间聚类分析是空间模式识别和空间挖掘的重要手段之一[25]，本研究采用聚类分析按照喜好食甜的高低程度对我国现代食甜习惯进行区域划分，利用ArcGIS的分组分析方法实现食甜习惯的自然聚类。为确保能较好实现组内相似性和组间差异性最大化，需采用不同的空间约束对要素进行分组。因此，本文就组数和空间约束2个方面进行分组探索。分组分析效果的评价参数为R^2，计算公式为：

$$R^2 = \frac{TSS - ESS}{TSS} \tag{11}$$

式中：TSS是总平方和；ESS是回归平方和。R^2值反映在分组完成后原始数据中变化的保留程度，故R^2值越大代表要素区分效果越好。

2.4.3 逐步回归分析

逐步回归方法通过对可能影响食甜度的多个地理气候要素进行引入、筛选、剔除，最终拟合出最优解。本质上是建立最优的多元线性回归方程，基本思想是将各因子逐个引入回归方程，同时对模型和所有已有因子在给定的显著性水平上检验，若不显著则剔除，直到既没有其余显著因子选入回归方程，也没有不显著的因子从回归方程中剔除为止，此时即为最优解。本文选择的逐步回归过程采用向前法，即变量由少到多，每次增加一个，由一元回归模型到二元回归模型再到多元回归模型，直至没有可引入的显著性变量为止[26]，分析变量主要包括年均气温、日照时长、降水量、相对湿度及糖料作物播种面积、人均GDP等指标。

3 中国现代食甜习惯的空间分布特征分析

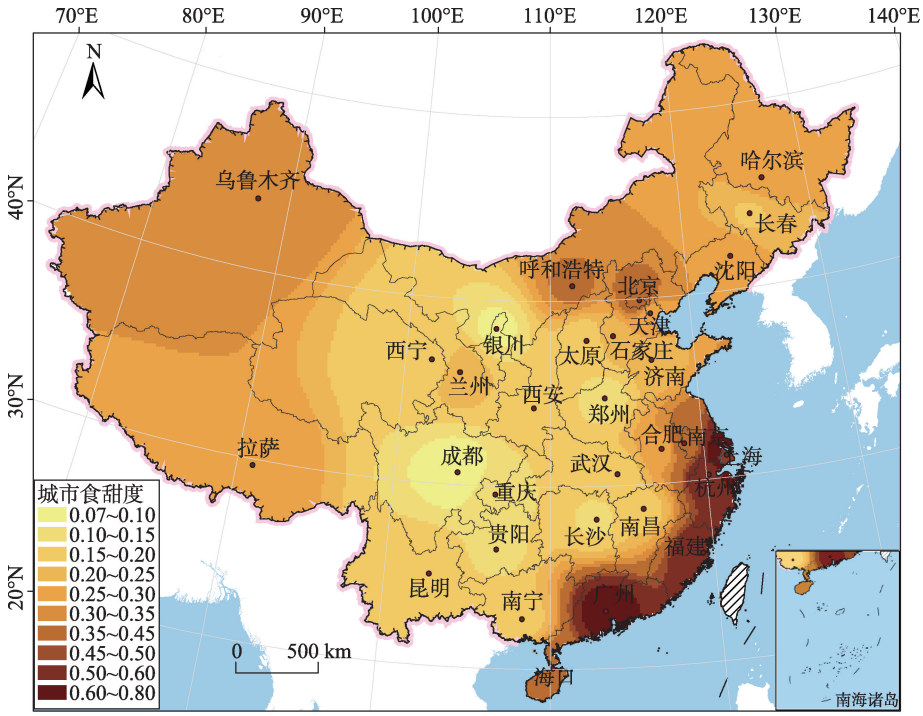
3.1 中国现代食甜习惯的区域差异分析

为了更好地反映中国不同地区食甜习惯的差异，本研究对由式（3）计算得出的不同城市食甜度数据进行空间插值，从而得到呈全域分布的食甜习惯空间分布数据。从表1可看出，根据对多种插值方法精度的交叉验证结果，反距离加权法均方根误差较小，表明其插值精度较高。由于普通克里格本身更关注空间整体，故其平均偏差较小，但标准均方根误差较大，而且其插值结果在空间上表现出呈东北-西南向的条带状属性分层，省域内部的甜度特征分异程度也较大，甚至出现新疆、拉萨等西部地区达到高度食甜的情况，这显然与实际情况不符。因此，本文选取反距离加权法对各城市的食甜度进行空间插值，从而得到我国大陆地区现代食甜习惯空间分布图，如图3所示。

表1　不同插值方法精度对比一览表

Tab. 1　Comparison of accuracy of different interpolation methods

插值方法	均方根误差（RMS）	平均偏差	标准均方根误差（标准化RMS）
反距离加权	0.130	-0.016	—
普通克里格	0.138	0.008	1.366



注：该图基于自然资源部标准地图服务网站下载的审图号为GS(2016)1600号的标准地图制作，底图无修改。由于数据获取困难，本研究不包括香港、台湾和澳门。

图3 中国现代食甜空间分布

Fig. 3 Distribution map of Sweet Habits in modern China

如图3所示，中国现代食甜习惯整体上表现出明显的集聚性分布特征，东南沿海一带普遍爱好食甜，属于典型的高值集聚区，其中：上海、广州、福州、杭州等城市的食甜度分别为0.77、0.65、0.60、0.57；中国东北和西北地区（特别是新疆）也属于食甜较多的地区，食甜度介于0.30~0.45之间；中国广大的中部地区，特别是以成都为中心的渝黔川地区和以银川为代表的河套平原区食甜度均较低，银川、成都、重庆、贵阳等城市的食甜度分别为0.10、0.07、0.13和0.14，存在显著的空间自相关性。

3.2 中国现代食甜习惯的空间自相关分析

3.2.1 基于莫兰指数的点位自相关分析

针对基于反距离加权插值得到的空间分布数据能否客观反映城市食甜习惯的真实集聚情况，需要进一步开展全局空间自相关性分析来进行验证。本研究借助ArcGIS 10.2的空间统计功能，对城市点图层的甜度字段分析其全局自相关性，依据各城市的经纬度坐标采用反距离方法生成空间权重矩阵。研究结果表明：在显著水平p=0.000 45时标准化Z值为4.0803，超过1%显著性的临界值2.58，全局Moran's I指数为0.1747，远大于$E(I)$值，这表明中国现代食甜习惯呈现显著的空间正相关关系，即食甜程度较高（较低）的城市在空间上相邻，食甜习惯的单元集聚特征比较明显，在整体上具有空间依赖关系。但是，全局自相关仅能反映区域整体的空间关系，并不能揭示局部区域的食甜分布特征。故本文采用局部空间自相关作进一步分析。对点图层的甜度字段进行局部自相关分析，得出城市食甜度的LISA分析结果（表2）。

从表2可以看出，中国31个省会城市中只有5个城市存在显著的局部空间自相关，且均为空间正自相关，其余城市均不显著。上海、杭州、福州关联类型为H-H型，即高食甜区同样被高食甜区包围，并且通过了1%水平的显著性检验；成都、银川

表2 5个代表城市食甜度的LISA结果

Tab. 2 LISA of Urban Eating Sweetness

城市	食甜度	Moran's I	标准化Z值	p值	关联类型
上海	0.7697	0.000 035	5.1219***	0.0000	H-H
杭州	0.5724	0.000 032	4.4924***	0.0000	H-H
福州	0.6000	0.000 011	2.8835***	0.0039	H-H
成都	0.0794	0.000 013	3.0953***	0.0019	L-L
银川	0.1000	0.000 010	2.3447**	0.0190	L-L

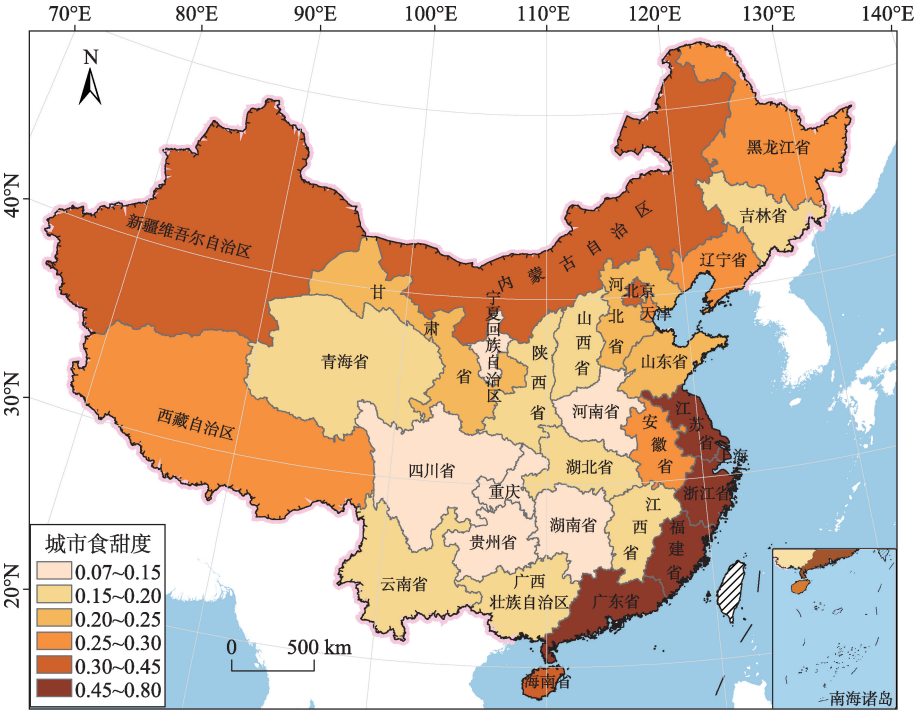
注：**代表在5%水平上显著，***代表在1%水平上显著。

关联类型为L-L型，即低食甜区同样被低食甜区包围，在5%和1%水平上均通过显著性检验；但是，其余26个城市的局部I指数均未通过显著性检验，不存在空间自相关性。结合图3的分析结果，可以得到以下结论：① 中国食甜在空间上形成了3个集聚区：以上海、杭州、福州为中心的东南沿海高甜集聚区，以成都为中心的渝黔川低甜集聚区和以银川为中心的西北内陆低甜集聚区；② 中国食甜在空间上不存在明显的分散趋势，即（不）爱好食甜城市周围会存在具有相近食甜口味的城市；③ 中国食甜习惯在局部上存在显著差异，但与传统认识上的“南甜北咸”存在一定程度的吻合，这表明在现代人口加速流动和饮食“重口味”的趋势推动下传统食甜习惯并未完全改变。

3.2.2　基于莫兰指数的区域自相关分析

上述基于呈点状分布的城市食甜度的空间自相关分析，多以统计检验的数字形式来验证空间关系，缺乏直观有效的可视化表达与分析。但是，由于基于点位插值得到的城市食甜空间分布栅格数据，因城市单元数量的有限性不能较好地转化为面状矢量数据进行分析。因此，本研究利用含有甜度属性字段的面状矢量数据作进一步探究。省会城市作为一省的政治、经济、文化、交通中心以及省内流动人口的高度聚集地，虽然其饮食文化特征无法表征该省的全部食甜特征，但在一定程度上能够反映出该省的总体食甜习惯。因此，本文依据省会城市的食甜度对各省份进行定性分类，得到图4所示的分级渲染后的各省食甜分布数据。在此基础上，利用ArcGIS和GeoDa软件探索城市食甜面状数据的空间自相关性，进一步分析中国现代食甜的区域集聚情况。

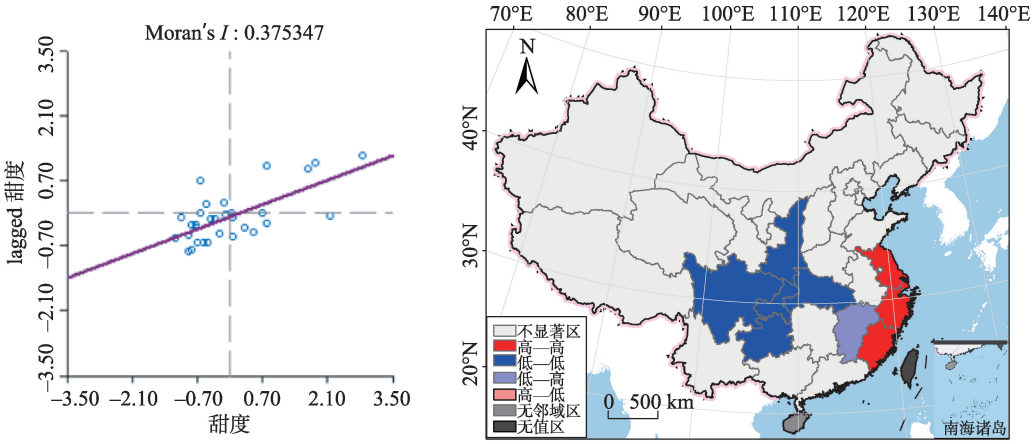
首先对不同省份面状食甜数据进行全局Moran's *I*统计，基于邻接关系生成空间权重矩阵，得到Moran's *I*为0.3753，标注化*Z*值为2.799，超过1%显著性的临界值2.58，再次证明我国现代食甜习惯在空间上呈集聚效应。然后，采用Moran散点图和LISA集聚图研究城市食甜度在局部区域的集聚与分散情况（图5）。在本研究中，Moran散点图的第一象限（HH象限）和第三象限（LL象限）代表区域内食甜度存在较强的空间正相关，即具有空间均质性；第二象限（HL象限）和第四象限（LH象限）代表区域内食甜度存在较强的空间负相关，即具有空间异质性。空间权重矩阵选择邻接矩阵，类型为



注：该图基于自然资源部标准地图服务网站下载的审图号为GS(2016)1600号的标准地图制作，底图无修改。由于数据获取困难，本研究不包括香港、台湾和澳门。

图4　中国不同省份现代食甜空间分布

Fig. 4　Distribution map of Sweet Habits in different province of China



(a) Moran 散点图 (b) LISA集聚图

注：该图基于自然资源部标准地图服务网站下载的审图号为GS(2016)1600号的标准地图制作，底图无修改。由于数据获取困难，本研究不包括香港、台湾和澳门。

图5 中国不同省份面状食甜数据的Moran散点图和LISA集聚图

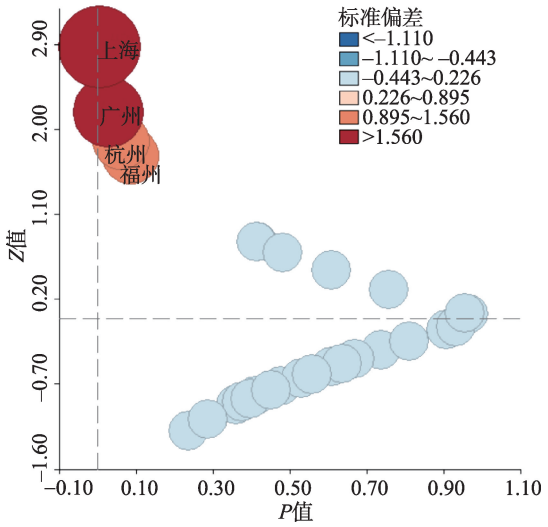
Fig. 5 Moran scatter graph and LISA agglomeration graph of polygon data in China

Queen型，共边共点皆视为邻接要素。

在Moran散点图中，研究单元的趋势线偏向HH象限和LL象限，说明中国食甜呈现高值与高值集聚、低值与低值集聚的模式。同样，LISA集聚图也印证了该点。江苏、浙江、福建3个省及上海市呈现高—高集聚模式，该区属于食甜“热点”区域，反映东南沿海区域在喜好食甜这一习惯上具有极大的相似性，而四川、贵州、陕西、湖北4个省及重庆市呈现低—低集聚模式，该区为食甜“冷点”区域，反映包含渝黔川在内的西南部分地区、包含陕西在内的西北部分地区及包含湖北在内的华中部分地区食甜程度均偏低，这与上述区域差异分析和点位自相关分析的结论相吻合，进一步验证了我国现代食甜习惯在局部地域上呈现集聚状态。此外，江西省表现为低-高集聚，由于该区毗邻东南沿海高甜区域和内陆低甜区，且自身食甜度较低。但是，由于只存在一个面元，不足以构成大范围的连片区域，故认为整体上食甜不存在低—高集聚区。

3.2.3 基于Gi-统计量的热点分析

通过整体和局部的Moran指数虽然能检测出具有相近食甜值的区域在空间上的集聚情况，但不能说明该聚类是否由高值或低值组成，因此可能存在一些特殊情况。例如，由Moran发现的高甜聚集区可能是由高于邻近区域的低甜值面积单元构成，并不是实际意义上的高甜聚集区，这种集聚应视为食甜的冷点区域。作为G-统计量的局部版本，Gi-统计量(Getis-Ord Gi*)能避免过度关注高值要素，可用于识别具有统计显著性的高值(热点)和低值(冷点)的空间聚类。因此，本研究采用Gi-统计量，借助ArcGIS软件对城市点图层的甜度字段进行热点分析，采用反距离空间要素定义方式，根据热点分析得到的P值、Z值以及G_i(d)值生成城市食甜度冷热点气泡图(图6)，对聚类结果和显著效果在规范坐标系中进行可视化表达。在图6中，



注：横轴代表P值，水平方向上P值越小代表置信度越高，聚类效果越显著；纵轴代表Z值，正值代表高值聚类，垂直向上方向气泡越大代表聚类程度越高，负值代表低值聚类，垂直向下方向气泡越小代表聚类程度越高；气泡颜色为$G_i(d)$值的标准偏差。

图6 中国不同城市食甜度冷热点气泡

Fig. 6 Cold point and hot Point bubble diagrams of sweetness in different cities of China

越靠近垂直虚线表示聚类效果越显著。在水平虚线之上，越向上且气泡越大代表具有越高聚类程度的热点区域；在水平虚线之下，越向下且气泡越小代表具有越高聚类程度的冷点区域。

从图6可以看出，上海（最上方红色气泡）属于偏爱食甜的热点区域，P值为0.004（<0.01），属于高度聚类状态，其周围较大区域极大可能同样偏爱食甜。广州（较小的红色气泡）亦属于爱好食甜的热点区域，但其P值为0.03（0.01<P<0.05），表明置信度为95%，属于中度聚类状态，其周围部分区域很有可能偏爱食甜。另外，杭州、福州（橘色气泡）也属于偏爱食甜的热点区域，P值均小于0.1，属于低度聚类状态，其周围一定区域较为可能偏爱食甜。由图6可知，中国现代食甜度不存在冷点区域。综上所述，食甜热点区域依然集中在东南沿海一带，这与上文通过Moran's I指数分析得到的结论相一致。

3.3 中国现代食甜习惯的空间聚类分析

由上述分析可以看出，中国现代食甜习惯的确存在显著的区域集聚效应。因此，基于此可对中国食甜习惯进行聚类分析，将各省份按照食甜程度高低分成若干组，生成我国现代食甜分区图。考虑到不同地区的食甜度存在空间自相关性，食甜习惯亦存在空间约束关系，因此本研究对CONTIGUITY面邻接类型、DELAUNAY三角测量及K_NEAREST邻接类型3种方法进行了对比，具体分组结果如表3所示。从表3可以看出，DELAUNAY三角测量在不同组数情况下的R^2值均大于其他两种方法，理论上聚类效果最佳，但其孤元个数过大，在分组效果上出现较多割裂单元，实际聚类效果不佳，故舍弃；K_NEAREST同样存在孤元个数过多的问题，而CONTIGUITY在组数为4时，R^2达到0.88，说明分组效果较为合理，同时不存在孤元，故最终采用分组为4，空间约束关系为CONTIGUITY时的聚类结果作为中国不同省份食甜分区（图7）。

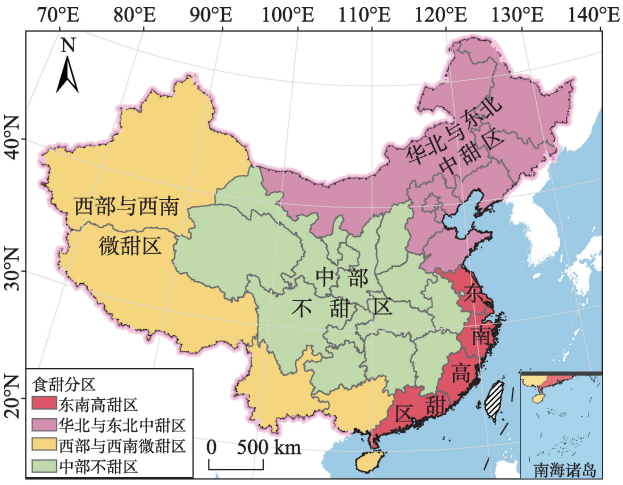
如图7所示，中国现代食甜习惯可分为4个区域：东南高甜区、华北与东北中甜区、西部与西南微甜区和中部不甜区，各区域的食甜度均值分别为0.60、0.32、0.28、0.16。对比直接分级渲染食甜分布（图4），可以发现聚类结果能反映出中国总体食甜状况。虽然中甜区和微辣区与图4有所区别，但对高度食甜区和不食甜区具有较好的分类效果。

表3　中国不同省份食甜度的分组结果表

Tab. 3　Grouping analysis results of sweetness for different provinces in China

空间约束类型	评价参数		
	组数	R^2	孤元个数
CONTIGUITY面邻接类型	3	0.78	0
	4	0.88	0
	5	0.87	1
DELAUNAY三角测量	3	0.87	2
	4	0.90	3
	5	0.92	4
K_NEAREST邻接类型	3	0.80	1
	4	0.88	3
	5	0.93	4

注：在分组分析时，会出现某省单独被归为一类或其周围所有相邻接的省份均与之不为一类的情况，这里将该省份定义为孤元。孤元的出现代表实际分类效果不佳，原因在于组数过多或违背空间自相关性。



注：该图基于自然资源部标准地图服务网站下载的审图号为GS（2016）1600号的标准地图制作，底图无修改。由于数据获取困难，本研究不包括香港、台湾和澳门。

图7　中国现代食甜空间聚类分析结果

Fig. 7　Distribution Map of Modern Sweet Habits in China based on grouping analysis

4　中国现代食甜习惯的影响因素分析

4.1 食甜影响因素模型的构建

4.1.1 筛选因素

上述研究表明，中国现代食甜习惯呈现出“东高北中，西微内低”的包围式格局，沿海省份的食甜度普遍高于内陆省份。通过查阅相关资料，这种分布格局实质上是多种因素共同作用的结果，具体包

括：① 自然因素方面，多雨且光热条件好的地区糖料作物产量高，而且人体在高温高湿的环境中能量的消耗更大，所以为了维持身体的能量平衡，当地居民自然养成吃甜的习惯[3,28]。② 社会经济因素方面，糖料作物生产条件不足的区域，自身糖料资源相对缺乏，且糖的运输成本高，价格高，因此在经济发展状况不足的情况下，当地居民摄糖量少，主要以咸调剂口味[28]。③ 历史因素方面，古代制糖成本与工序比制盐困难许多，因此不同地区制糖技术差异显著，这势必会对历史时期饮食文化的形成具有重要影响，进而影响我国现代食甜的分布状况[3,11]。总体来看，物产的丰富程度主要受地理位置的影响。鉴于上述原因，本文选择年均气温、日照时长、年降水量、相对湿度4个气象要素和糖料作物播种面积、人均GDP 2个社会经济要素作为影响中国现代食甜分布的因素。另外，文中在自变量中加入地理位置这个二值变量，即依据其所在省份是否有海岸线进行取值，有海岸线省份用k=1表示，反之k=0。由于沿海和内陆地区的气候条件差异显著，城市食甜度受其影响可能存在差异，为了验证差异性的大小，本文在模型中加入二值变量与气温、日照、降水要素的交叉项，同时也作为自变量加入模型。

4.1.2 建立回归模型

由于不同要素的数据内部差异过大，由此导致异方差性较大。因此，为了提高模型的拟合精度，本文采用双对数回归模型，回归方程如下：

$$\begin{aligned}\ln y = {} & c+\beta_1\times\ln t+\beta_2\times\ln p+\beta_3\times\ln s+\beta_4\times \\ & \ln h+\beta_5\times\ln\overline{GDP}+\beta_6\times area+\beta_7\times \\ & k\times\ln t+\beta_8\times k\times\ln p+\beta_9\times k\times \\ & \ln s+\beta_{10}\times k\times\ln h+\varepsilon\end{aligned} \tag{12}$$

式中：y为城市甜度；β(i=1,2,⋯,10)为影响系数；t代表年平均气温；p代表年降水量；s代表日照时长；h代表相对湿度；$\overline{GDP}$代表人均GDP；$area$代表各地区糖料播种面积；k为虚拟变量；c、ε为常数项。

4.2 食甜习惯影响因素分析

文中利用SPSS软件中的逐步回归方法对模型进行求解，计算出的各模型参数结果如表4所示。从表中可以看出，各模型均通过置信度95%显著性水平检验，其中模型6的R^2值最大，达到0.82，且各变量均在0.01显著水平及以上相关，这说明模型拟合效果较好。所以，模型6所对应的回归方程即为最优回归方程：

表4 逐步回归分析参数

Tab. 4 Stepwise regression analysis parameters

变量	模型1	模型2	模型3	模型4	模型5	模型6
	β_i	β_i	β_i	β_i	β_i	β_i
C	-1.67^{***}	-6.17^{***}	-7.10^{***}	-7.49^{***}	-6.78^{***}	-6.68^{***}
$k\times\ln p$	0.12^{***}	0.13^{***}	0.11^{***}	0.72^{**}	3.54^{**}	3.77^{***}
$\ln s$		0.87^{***}	0.88^{***}	0.99^{***}	0.99^{***}	0.99^{***}
$\overline{GDP}$			0.43^{**}	0.35^{**}	0.05	
$k\times\ln h$				-1.01^{*}	-3.58^{**}	-3.81^{***}
$k\times\ln t$					-3.19^{*}	-3.45^{***}
F值	21.89^{***}	22.18^{**}	20.70^{***}	17.95^{***}	16.86^{***}	22.18^{***}
R^2	0.50	0.68	0.76	0.79	0.82	0.82

注：*代表通过90%的显著性检验；**代表通过95%的显著性检验；***代表通过99%的显著性检验；σ为标准误差。

$$\begin{aligned}\ln y = {} & -6.68+3.77\times(k\times\ln p)+0.99\times\ln s- \\ & 3.81\times(k\times\ln h)-3.45\times(k\times\ln t)\end{aligned} \tag{13}$$

由方程可知，人均GDP、糖料作物播种面积均被自动剔除，降水量、相对湿度和平均温度3个气象要素对城市食甜度的影响较为显著，影响系数分别高达3.77、−3.81、−3.45，这种影响在沿海地区尤为突出。在高温高湿的沿海地区，人们对于甜食的需求量较大，再加上这些地区属于糖料主产区，丰富的糖类产物势必会影响人们的饮食习惯。但是，随着交通和物流行业的发展，糖类运输能力增大且成本降低，因此本地糖料作物的产量和经济发展水平对当地居民现代食甜习惯的限制性影响大大减弱。

此外，方程中日照一项并未剔除，这说明其对食甜度也具有一定的正向影响。但是，日照对食甜的影响并不是简单的线性关系。为了验证日照与食甜的相关关系，本文根据虚拟变量将城市分为沿海和内陆城市，通过绘制散点图来研究这2类城市的日照与食甜关系（图8）。从图8可看出，随着日

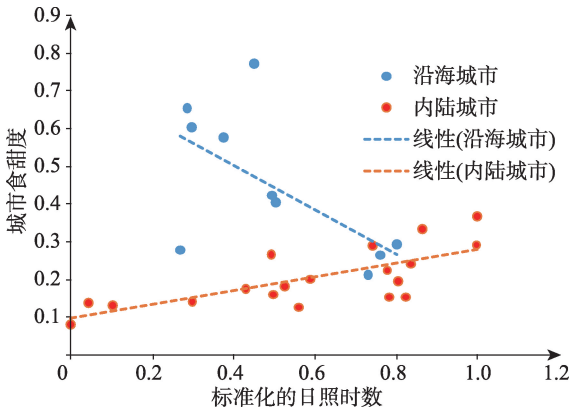


图8 沿海与内陆城市食甜度与日照时数的关系

Fig. 8 Relationship between Sweetness and Sunshine Hours in coastal and inland cities

照时数的增加，沿海城市的食甜度呈现下降趋势，而内陆城市的食甜度却呈现增加趋势。这种截然相反的变化趋势，表明在不同的地理位置上，日照对食甜的作用效果是不同的。由此说明，回归方程尽管有较高的拟合精度，且揭示了气象要素对于现代食甜分布的重要影响，但限于样本点数量，其具体的作用机制还有待进一步的探究。

5 结论与讨论

5.1 结论

本文基于美食外卖主流网站“饿了么”平台，爬取全国大陆地区共计31个地级以上城市2019年1月1日—1月31日的海量饮食消费数据，通过MySQL实现数据存储及处理，制定出一套较为便捷合理的甜度评价方案，最终得到各城市的食甜度。基于城市食甜数据，借助空间插值、空间自相关分析、空间聚类分析及逐步回归分析法对我国现代食甜习惯的空间分布特征进行分析，识别影响我国现代食甜习惯的因素，主要得到以下结论：

（1）中国现代食甜习惯的空间分布特征。我国食甜在空间分布上存在显著的地域分异特征，如东南沿海一带普遍偏爱食甜（南京、杭州、上海、广州等城市食甜度均在0.5以上），渝黔川地区食甜较少（成都、重庆、贵阳等城市食甜度均在0.2以下），与传统“南甜北咸”的食甜认知不完全吻合。我国现代食甜习惯可分为4个区域：东南高甜区（平均食甜度0.60）、华北和东北中甜区（平均食甜度0.32）、西部和西南微甜区（平均食甜度0.28）及中部不甜区（平均食甜度0.16）。该分区结果能反映总体食甜状况，R^2达到0.88，尤其是对高度食甜区和不食甜区具有较好的分类效果，对于中甜区和微辣区存在可接受范围内的差异。

（2）中国现代食甜习惯的空间集聚性。从整体空间上看，我国现代食甜习惯呈现显著的空间正相关关系，（不）偏爱食甜区域在空间上集聚，具有空间依赖关系，不存在明显的分散趋势；从局部空间上看，我国现代食甜习惯呈现显著的局部空间正相关，形成特色鲜明的3个地理集聚区，即以苏浙沪闽为主的东南沿海一带高甜集聚区，以渝黔川为主的西南地区和以陕宁为主的西北内陆低甜集聚区；从集聚程度上看，我国现代食甜习惯的热点区域集中在苏浙沪闽粤东南沿海一带，集聚状态属于中高度水平。

（3）中国现代食甜习惯的影响因素。食甜影响因素模型拟合精度达到0.82，分析结果表明：气温、湿度、降水等气象要素及地理位置是影响我国现代食甜空间分布的重要因素，且单个要素与食甜度之间不是简单的线性关系。日照对食甜的影响过程与地理位置有关，具体表现为：内陆地区城市的食甜度一般随着日照时长的增长而增大，沿海地区城市的食甜度往往随着日照时长的增长而减小。

5.2 研究不足与展望

尽管本研究是依托大数据优势进行的，但在数据、研究方法和研究范围方面仍存在一些局限和问题，需要在今后的研究中进一步完善：

（1）数据质量方面。数据来源于主流外卖平台“饿了么”，其菜品构成与普通百姓日常饮食存在一定出入，且外卖客户群体主要是中青年人群，所获数据不能完全代表现代饮食特征；为确保数据能反映城市食甜的真实情况，每个城市的数据爬取量非常可观，最终只爬取全国共计31个城市数据，导致数据样本较少；数据抓取时段也仅为2019年1月，时间尺度过短，忽视了可能存在的季节效应，即食甜程度可能随季节发生周期性变化。

（2）数据处理方面。由于数据量过大，无法依靠每道菜的具体成分含量来衡量甜度，如衡量传统菜品甜度时按甜菜占传统菜的比例来计算，导致处理过程存在一定的主观性，最终结果可能会与城市实际甜度产生微弱的偏差。

（3）研究方法方面。限于数据样本数量有限，本文仅从空间自相关性和食甜影响因素2个方面进行分析与验证，缺少深入的探讨与剖析。例如：论证气象要素对食甜分布的影响时，并未对其详细影响机制以及产生作用的原因进行深入分析。现代食甜空间分布的形成本身是一个多维问题，用简单的线性方程进行分析可能无法全面展现其本质原因。

（4）研究范围方面。中国现代食甜习惯可能受到多方面因素影响，包括：现代人口大规模流动带来的饮食习惯交融，生活水平提高带来对饮食健康而非口味的更高要求，饮料蛋糕类的消费受限于年龄及性别特征等，这些潜在的影响都值得后续进行深入探讨。

（5）研究结果方面。本文的研究对象为省会城市，由于省会城市人口构成多样化，外来人口比例大，这会在很大程度上影响当地饮食文化的整体特征，因此省会城市的食甜水平不能完全代表该省的

食甜水平,研究结果可能与实际情况存在一定的出入。在今后的研究中将会增加其他城市的数据,进一步提高研究结果的准确性。

参考文献(References):

[1] 王汐牟,王强.时空与食事:宋元明时期的饮食谱录及其所蕴含的社会饮食风尚[J].北方从论,2018,268(2):97-101. [Wang X M, Wang Q. The space-time and the diet: The systematic dietary records and the diet custom during the Song Yuan and Ming Dynasties[J]. The Northern Forum, 2018,268(2):97-101.]

[2] 丁继月.大连市区餐饮业时空变迁及其机制研究[D].大连:辽宁师范大学,2019. [Ding J Y. Spatial- temporal change and mechanism of catering industry in Dalian city [J]. Dalian: Liaoning Normal University, 2019.]

[3] 陈传康.中国饮食文化的区域分化和发展趋势[J].地理学报,1994,61(3):226-235. [Chen C K. The culture of Chinese diet: Regional differentiation and developing trends [J]. Acta Geographica Sinica, 1994,61(3):226-335.]

[4] 秦萧,甄峰,朱寿佳,等.基于网络口碑度的南京城区餐饮业空间分布格局研究——以大众点评网为例[J].地理科学,2014,34(7):810-817. [Qin X, Zhen F, Zhu S J, et al. Spatial pattern of catering industry in Nanjing urban area based on the degree of public praise from internet: A case study of Dianping.com[J]. Scientia Geographica Sinica, 2014,34(7):810-817.]

[5] 刘琪,基于大数据的深圳市餐饮业空间分布格局研究——以大众点评网为例[C].中国城市规划学会、东莞市人民政府.持续发展理性规划——2017中国城市规划年会论文集(05城市规划新技术应用).中国城市规划学会、东莞市人民政府:中国城市规划学会,2017:611-620. [Liu Q, Study on spatial distribution pattern of catering industry in Shenzhen city based on big Data[C]. UPSC、Dongguan Municipal People's Government. Rational Planning for Sustainable Development——Papers Collection of the 2017 Annual Conference on Urban Planning in China(05 Application of New Urban Planning Technology). UPSC、Dongguan Municipal People's Government: UPSC, 2017:611-620.]

[6] 夏令军,刘艳芳,刘国炜.中国地级城市餐饮业分布格局及影响因素——基于“大众点评网”数据的实证研究[J].经济地理,2018,38(5):133-141. [Xia L J, Liu Y F, Liu G W. Spatial distribution pattern and influencing factors of catering industry in prefecture level and above cities in China: An empirical study based on the Dianping.com data[J]. Economic Geography, 2018,38(5):133-141.]

[7] 顾秋实,张海平,周星星,等.中国传统八大菜系的地理分布及其扩散效应——基于大数据视角的实证分析[J].浙江学刊,2019(5):47-53. [Gu Q S, Zhang H P, Zhou X X, et al. Geographical distribution and diffusion effect of eight traditional chinese cuisine: An empirical analysis based on data perspection[J]. Zhejiang Academic Journal, 2019 (5):47-53.]

[8] 黄恒君,陶然,傅德印.单位名录库更新:互联网大数据源及其数据质量评估[J].统计研究,2017,34(1):12-22. [Huang H J, Tao R, Fu D Y. Business register database revision: Internet data sources and data quality assessment[J]. Statistical Research, 2017,34(1):12-22.]

[9] 秦之湄.基于互联网大数据的城市餐饮源清单高时空分辨研究[C].中国环境科学学会环境信息化分会、《中国环境管理》杂志.2016全国环境信息技术与应用交流大会暨中国环境科学学会环境信息化分会年会论文集,2016:139-146. [Qin Z M. Research on high temporal and spatial resolution of urban catering source list based on Internet big data[C]. Environmental information branch of Chinese Society for Environmental Sciences、Chinese Journal of Environmental Management、2016 national environmental information technology and application exchange conference and the annual meeting of environmental information branch of Chinese Environmental Science Society, 2016:139-146.]

[10] 蓝勇,中国饮食辛辣口味的地理分布及其成因研究[J].人文地理,2001,16(5):84-88. [Lan Y. On the reasons and distribution of pungent flavor in Chinese food[J]. Human Geography, 2001,16(5):84-88.]

[11] 张仁军,刘聪.中国食辣的空间分布及影响因素——基于美食团购的大数据分析[J].重庆师范大学学报(自然科学版),2016,33(1):143-148. [Zhang R J, Liu C. The spatial distribution and the reason of the spice diet in China: based on the data mining of group purchase website[J]. Journal of Chongqing Normal University: Natural Science Edition, 2016,33(1):143-148.]

[12] 洪忻,李解权,王志勇,等.南京城区居民膳食模式及其影响因素研究[J].中华疾病控制杂志,2015,19(1):54-57. [Hong X, Li J Q, Wang Z Y, et al. Study on dietary patterns and their influencing factors among urban residents in Nanjing City[J]. Chinese Journal of Disease Control & Prevention, 2015,19(1):54-57.]

[13] Su H, Shan X W. Research on the influencing factors of dietary behaviors among school students[J]. Zhonghua liu xing bing xue za zhi, 2011,32(8):751.

[14] Katsarou A, Tyrovolas S, Psaltopoulou T, et al. Socio-economic status, place of residence and dietary habits among the elderly: The Mediterranean islands study[J]. Public Health Nutrition, 2010,13(10):1614-1621.

[15] 赵彤.基于GIS的重庆市柑橘农业气候区划[D].重庆:重

庆师范大学,2018. [Zhao T. The study on agricultural climateregionalization of citrus in Chongqing Provincebased on GIS[D]. Chongqing: Chongqing Normal University, 2018.]

[16] 杨丹,杨学明,张鹏,等.南方馒头加工中的辅料配比优化研究[J].麦类作物学报,2014,34(8):1061-1066. [Yang D, Yang X M, Zhang P, et al. Study on optimzing the of excipients in the processing of southern- style Chinese steamed bread[J]. Journal of Triticeae Crops, 2014,34(8): 1061-1066.]

[17] 周素梅,张国增,张梅红,等.北方馒头实验室制作与评价[J].科技导报,2011,29(29):75-79. [Zhou S M, Zhang G Z, Zhang M H, et al. Processing procedure in the laboratory and quality evaluation method of northern style chinese steamed bread[J]. Science & Technology Review, 2011,29(29):75-79.]

[18] 狄乾斌,王亮,邱煜焜.中国沿海省份城市发展水平空间自相关分析[J].资源开发与市场,2018,34(5): 649-653. [Di Q B, Wang L, Qiu Y K. Spatial correlation analysis of urban development level of prefecture-cities in China's coastal provinces[J]. Resource Development & Market, 2018, 34(5): 649-653.]

[19] Rong H F, Fang B, University N N. Measurement of the matching degree between urbanization and ecology in Anhui based on Barycenter model[J]. China Land Sciences, 2017,31(6):34-41

[20] 臧良震,张彩虹,郝佼辰.中国林下经济发展的空间分布特征研究[J].林业经济问题,2014,34(5):442-446. [Zang L Z, Zhang C H, Hao J C. Research on spatial distribution of Under- forest economic development in China[J]. Issues of Forestry Economics, 2014,34(5):442-446.].

[21] 梁洪运,周其龙,孙亚林.空间权重矩阵对空间自相关影响分析[J].科技资讯,2013(9):141-143. [Liang H Y, Zhou Q L, Sun Y L. Spatial weight matrix influences on spatial autocorrelation[J]. Science & Technology Information, 2013(9):141-143.]

[22] 杨青林,赵荣钦,邢月,等.中国城市碳排放的空间分布特征研究[J].环境经济研究,2017,2(1):70-81. [Yang Q L, Zhao R X, Xing Y, et al. Spatial distribution characteristics of carbon emissions in China's cities[J]. Journal of Environmental Economics, 2017,2(1):70-81.]

[23] 张生玲,王雨涵,李跃,等.中国雾霾空间分布特征及影响因素分析[J].中国人口·资源与环境, 2017,27(9):15-22. [Zhang S L, Wang Y H, Li Y, et al. Spatial distribution of haze pollution and its influencing factors[J]. China Population, Resources and Environment, 2017,27(9):15-22.]

[24] 张康聪,陈健飞.地理信息系统导论(第3版)[M].北京:清华大学出版社,2009. [Zhang K C, Chen J F, Introduction to geographic information systems(3thEdition) [M]. Beijing: Tsinghua University Press, 2009.]

[25] 景楠,颜波.东北区域物流与经济的空间自相关和空间聚类分析[J].东北师大学报(哲学社会科学版),2015(1): 134-140. [Jing N, Yan B. Spatial autocorrelation and spatial clustering analysis of regional logistics and economy in northeast China[J]. Journal of Northest Normal University (Philosophy and Social Sciences), 2015(1):134-140.]

[26] 汪紫云,李援亚.基于逐步回归分析法的湖北省农林牧渔产值增长因素分析[J].粮食科技与经济,2018,43(5):33-36. [Wang Z Y, Li Y Y. Analysis of factors affecting the growth of agricultural, forestry, animal husbandry and fishery output value in Hubei province based on stepwise regression analysis[J]. Grain Science and Technology and Economy, 2018,43(5):33-36.]

[27] 林正秋,糖的历史与糖业祖师杂谈[J].上海调味品,2004(5):30-31. [Lin Z Q. The history of sugar and sugar industry[J]. Shanghai Seasoning, 2004(5):30-31.]

[28] 蓝勇,陈姝.历史时期中国甜食的空间格局及其成因研究[J].云南大学学报(社会科学版),2019,18(6):57-67. [Lan Y, Chen S. A historical study of the spatial pattern of Chinese sweet foods and its causes[J]. Journal of Yunnan University (Social Sciences Edition), 2019,18(6):57-67.]