

# 东北地区大豆单产空间分异特征及其影响因素分析

王 晨, 褚 琳<sup>\*</sup>, 杨 喆, 杨镇豪, 张歆雅, 王天巍, 蔡崇法

(华中农业大学资源与环境学院, 武汉 430070)

**摘 要:** 作为中国最重要的大豆主产区, 东北地区的大豆生产成为解决大豆进出口产需问题的核心。明晰东北地区大豆单产空间分异特征及其影响因素, 对于保卫“黑土粮仓”和夯实粮食安全“压舱石”具有重要意义。该研究基于谷歌地球引擎 (Google Earth Engine, GEE) 平台, 采用多特征随机森林的分类方法, 提取 2022 年大豆种植空间分布信息, 结合多时相叶面积指数 (leaf area index, LAI) 数据与实测产量建立大豆估产模型, 明晰区域大豆单产空间分异特征, 运用地理探测器定量解析大豆单产空间分异特征的影响因素。结果表明: 1) 2022 年东北地区大豆种植面积的总体提取精度达 89.48%, Kappa 系数为 0.89, 与统计数据之间的决定系数 ( $R^2$ ) 为 0.92。大豆种植面积由北向南递减, 大豆种植区主要分布于松嫩平原地区, 重心位于绥化市。2) 2022 年东北地区大豆平均单产为 2 514.08 kg/hm<sup>2</sup>, 与实测产量之间的  $R^2$  为 0.72。大豆单产空间分布集聚性显著 ( $P < 0.01$ ), 呈现出北高南低的分布特征。3) 土壤类型、土壤 pH 值和大豆补贴是解释大豆单产空间分异特征最重要的 3 种单因子,  $q$  值分别为 0.27、0.24 和 0.24。年均降雨 $\cap$ 年均积温、年均降雨 $\cap$ 大豆补贴以及土壤类型 $\cap$ 大豆补贴是解释大豆单产空间分异特征最重要的 3 对交互因子,  $q$  值分别为 0.44、0.40 和 0.40。人为因素对大豆单产差异影响显著 ( $P < 0.01$ ), 大豆补贴、大豆粮价、农业灌溉面积、农业机械总动力、肥料价格和文盲率的较佳影响范围分别为 4 801~7 500 元/hm<sup>2</sup>、5 601~5 800 元/t、13.6 $\times 10^4$ ~26.4 $\times 10^4$  hm<sup>2</sup>、252 $\times 10^4$ ~436 $\times 10^4$  kW、2 500~2 602 元/t 以及 1.4%~1.8%。东北地区大豆单产在空间上呈现由北向南递减的趋势, 具有显著的空间异质性。大豆单产空间分异受自然因素与人为因素的共同影响, 自然因素起主导作用, 人为因素起调控作用。

**关键词:** 土壤; 降雨; 作物; 空间分异特征; 地理探测器; 东北地区; 多特征随机森林; 大豆单产

doi: 10.11975/j.issn.1002-6819.202306163

中图分类号: S127

文献标志码: A

文章编号: 1002-6819(2023)-21-0108-12

王晨, 褚琳, 杨喆, 等. 东北地区大豆单产空间分异特征及其影响因素分析[J]. 农业工程学报, 2023, 39(21): 108-119.

doi: 10.11975/j.issn.1002-6819.202306163 <http://www.tcsae.org>

WANG Chen, CHU Lin, YANG Zhe, et al. Spatial heterogeneity and determinants of soybean yield in Northeast China[J].

Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(21): 108-119. (in Chinese

with English abstract) doi: 10.11975/j.issn.1002-6819.202306163 <http://www.tcsae.org>

## 0 引 言

大豆在人类饮食结构中占据重要地位, 关系饲料、油脂和食用蛋白安全, 提升大豆产量对保障粮食安全至关重要。中国是世界大豆原产国, 大豆出口贸易额曾居世界第一。随着生活水平的提高和畜牧业的快速发展, 人们对大豆的需求超过了国内生产的供应, 产需缺口逐渐扩大, 中国从大豆净出口国转变为净进口国。针对该问题, 农业农村部立足中国资源禀赋和生产实际, 于 2002 年制定了第一轮大豆振兴计划实施方案, 该方案推动形成了进口大豆与国产大豆错位竞争、相互补充的格局。为进一步提振大豆自给率, 2019 年中央一号文件实施第二轮“大豆振兴计划”<sup>[1]</sup>, 该计划以农业供给侧结构性改革为主线, 强调提高大豆补贴标准, 加强大豆优势区生产能力建设。东北地区作为最主要的大豆优势

产区, 其大豆产量占全国生产总量的 1/2<sup>[2]</sup>, 是国家粮食安全的“压舱石”。东北地区大豆稳产提质是推动大豆产业振兴的重要保障, 明晰东北地区大豆单产空间分异特征及其影响因素, 对于保障中国粮食安全具有重要作用。

大豆总产量由种植面积和单位面积产量共同决定, 其中大豆种植面积是关键性因素<sup>[3]</sup>, 准确掌握大豆种植空间分布情况是大区域估产的重要前提。农作物种植面积可以通过传统的农业调查和遥感技术获得, 遥感技术因数据获取简单、成本低和空间覆盖广等优点, 已经成为大尺度范围内大豆种植面积提取的有效途径<sup>[4]</sup>。遥感数据中的光学影像数据, 如 GF 系列<sup>[5]</sup>、MODIS<sup>[6]</sup>、Landsat<sup>[7]</sup>、Sentinel-2 以及多源融合数据<sup>[8]</sup>等, 被广泛用于农作物种植面积的提取。其中, Sentinel-2 由于其较高的光谱分辨率、时间分辨率以及 3 个红边波段, 在农作物分类中得到广泛应用<sup>[9]</sup>。此外, 雷达影像出色的穿云能力和不同作物相异的后向散射强度扩展了遥感作物识别的分类特征, 同时为多源遥感影像融合提取大范围作物提供了新思路<sup>[10]</sup>。在农作物遥感分类中, 分类算法和分类特征变量的选择是两个最重要的方面。最大似然法<sup>[11]</sup>、决策树分类<sup>[12]</sup>和随机森林分类<sup>[13]</sup>等分类算法被广泛使

收稿日期: 2023-06-24 修订日期: 2023-11-11

基金项目: 国家重点研发计划项目 (2021YFD1500703)

作者简介: 王晨, 研究方向为农业资源环境遥感。

Email: wangchen3013@163.com

<sup>\*</sup>通信作者: 褚琳, 博士, 副教授, 研究方向为资源与环境遥感、水土保持与区域生态环境。Email: chulin@mail.hzau.edu.cn

用, 随机森林算法可以通过随机抽样和特征子集的选择, 降低模型过拟合的风险, 具有较好的泛化能力, 尤其在处理高维数据和异常值方面表现出更好的适应性。此外, 分类特征变量的选择是关系分类是否成功的决定性因素之一, 光谱特征、空间特征、时相特征、后向散射特征及 DEM 等特征可以被选用于农作物遥感分类。

研究通常采用农学估产方法<sup>[14]</sup>、统计预报方法<sup>[15]</sup>、数值模拟方法<sup>[16]</sup>和遥感方法<sup>[17]</sup>估算农作物单产。特别是遥感方法。因其具有宏观、实时监测的优势, 在大区域尺度上常被用来高效、精确地评估农作物单产。基于生物学原理, 遥感方法通过光谱获取作物的生长信息, 并建立作物光谱与产量之间的相关关系<sup>[18]</sup>, 包括植被指数-气象因子综合估产模式、以热红外信息为基础的估产模式以及植被指数统计估产模式 3 种方法<sup>[19]</sup>。植被指数统计估产模式是利用作物生长部分时期或全生育期的植被指数累加值或某一时期的植被指数与产量建立统计关系。从技术可行性角度看, 玉米<sup>[20]</sup>、水稻<sup>[21]</sup>和大豆<sup>[22]</sup>等农作物都可以采用植被指数统计估产模式进行单产估算。基于植被指数的多时相估产方法, 在农作物遥感估产领域中得到广泛使用<sup>[23]</sup>。

自然环境和人为条件等多种因素相互作用、相互影响, 共同决定了农作物最终产量的形成, 造成了区域农作物产量的空间差异。自然环境包括水分含量<sup>[24]</sup>和叶绿素<sup>[25]</sup>等生物学因素, 降雨<sup>[26]</sup>和积温<sup>[27]</sup>等气候因素以及地形<sup>[28]</sup>、土壤肥力<sup>[29]</sup>和土壤 pH 值<sup>[30]</sup>等环境因素, 这些因素是决定区域作物产量的主要和基础要素。同时, 人为条件, 如病虫害防治<sup>[31]</sup>和水肥管理<sup>[32]</sup>等农业管理方式, 也对农作物产量的提高起着关键作用。影响东北地区大豆单产空间分异的因素包含多个方面, 涉及广泛的因素和复杂的驱动机制。因此, 探明大豆单产空间分异背后的影响因素显得十分重要。分析影响因素的方法包括相关性分析<sup>[33]</sup>、主成分分析<sup>[34]</sup>、结构方程<sup>[35]</sup>以及地理探测器<sup>[36]</sup>等方法。在这些方法中, 地理探测器因其能够在空间尺度上对多个因素及其相互作用进行研究, 并提供清晰的可视化分析结果, 在空间分析上具有独特地位, 而被广泛应用<sup>[37]</sup>。

研究基于谷歌地球引擎 (Google Earth Engine, GEE) 平台, 采用随机森林分类器提取大豆种植空间分布信息, 结合遥感产品和实测产量, 根据大豆物候特征建立估产模型, 获得东北地区大豆单产的空间分布, 明晰大豆单产空间分异特征, 选取自然环境和人为条件两方面因子, 使用地理探测器对大豆单产空间分异进行定量归因分析。以期促进大豆生产稳定发展, 为东北地区大豆稳产增产提供理论参考。

## 1 材料与方法

### 1.1 研究区概况

东北地区 (38°72'N~53°56'N, 115°52'E~135°09'E) 辖黑龙江省、吉林省、辽宁省和内蒙古自治区东部三市一盟 (呼伦贝尔市、通辽市、赤峰市和兴安盟), 总面积约 125 万 km<sup>2</sup> (图 1)。

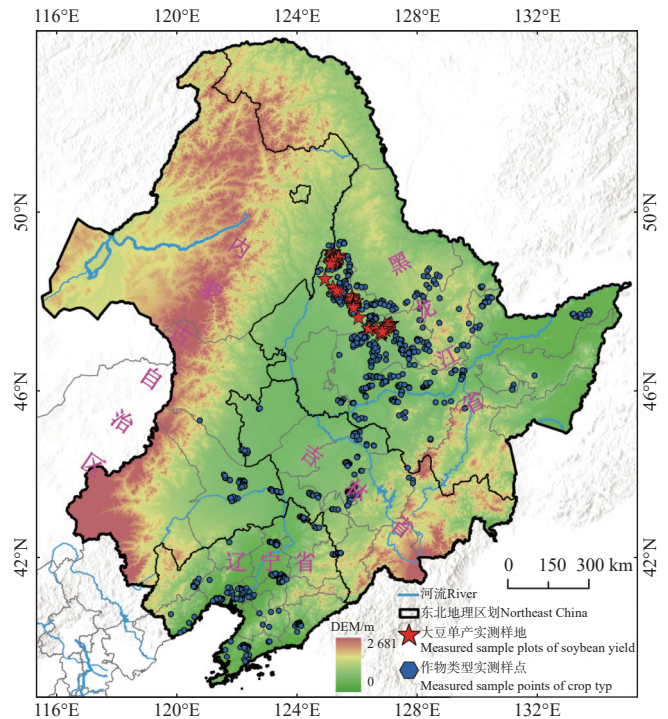


图 1 东北地区研究范围

Fig.1 Location of Northeast China (NEC)

研究区属温带大陆性季风气候, 10℃ 以上的年积温为 1600~3600℃, 年降水量约为 510 mm, 其中 80% 的降水集中在 6—9 月。地形由三面环山的盆地和中间的平原组成, 平原包括松嫩平原、三江平原和辽河平原。境内以黑土、棕壤、暗棕壤、草甸土、白浆土、栗钙土和沼泽土等土类为主。大豆种植为一年一季, 生长期介于 150~180 d 之间, 包括出苗期、幼苗期、花芽分化期、开花结荚期、鼓粒期和成熟期 6 个时期。每年 4 月底至 5 月初播种, 9 月底收获。大豆播种后 4~6 d 即可出苗, 随后幼茎不断伸长, 根系初步形成, 进入幼苗期, 开始需要较多的水分和养料。当复叶出现 4~5 片时, 主茎下部开始分枝, 同时分化花芽, 根、茎、叶开始旺盛生长, 大豆进入花芽分化期。从始花到幼荚, 大豆的开花与结荚往往是并进的, 此时可称为开花结荚期, 是生育最旺盛的时期。8 月上旬大豆进入鼓粒期, 干物质开始累积, 豆粒逐渐膨大。成熟期叶片变黄, 豆粒脱水, 摇动植株时英内有轻微响动。

### 1.2 数据收集

#### 1.2.1 气象数据

积温数据来源于中国科学院资源环境科学与数据中心 (<http://www.resdc.cn/>) 发布的中国气象要素逐日站点观测数据集, 针对 2022 年研究区内各气象站  $\geq 10^\circ\text{C}$  的逐日平均温度进行清洗与统计, 采用经验贝叶斯克里金插值法将 2022 年的点值积温扩展至面。

降雨数据来源于地球资源数据云 (<http://www.gis5.g.com>) 提供的 2022 年中国 1 km 分辨率年均降水量数据集。

太阳辐射数据来源于中国科学院资源环境科学与与数

据中心 (http://www.resdc.cn) 发布的中国气象要素逐日站点观测数据集, 根据研究区内的 247 个气象站的逐日日照时数, 计算 2022 年的年均太阳辐射量。

### 1.2.2 遥感影像

以 Sentinel-1 和 Sentinel-2 影像作为大豆种植面积提取的数据源, 根据粮食作物的物候特征, 选取 2022 年 5—9 月的多时相影像数据集。对 Sentinel-1 数据进行均值与 Refine Lee 滤波处理, 对 Sentinel-2 数据进行去云和中值滤波处理, 通过采用阈值分割法, 剔除水域、道路和城镇区, 以降低非农用地等地物类型对分类结果的影响。

叶面积指数 (leaf area index, LAI) 数据来自 MODIS 的 MOD15A2H 产品, 该产品空间分辨率为 500 m, 时间分辨率为 8 d (https://ladsweb.modaps.eosdis.nasa.gov/search/)。该数据集在全球主要的 6 套 LAI 产品的交叉验证中显示出最低的不确定性<sup>[38]</sup>, 较其他植被指数产品更具优势。

夜间灯光数据来源于 NPP-VIIRS 产品, 该数据是由美国国家海洋和大气管理局以及美国国家海洋和大气研究所的地球观测组 (https://payneinstitute.mines.edu/eog-2/viirs/) 利用可见光红外成像辐射仪套件 (VIIRS) 昼/夜夜段的夜间数据制作的一套平均辐射复合图像, 空间分辨率为 750 m。

### 1.2.3 实测样点

结合 Google Earth 高清地图与 GPS-RTK 手持设备获取植被类型样点数据和大豆实测产量数据。作物类型样点数据记录了作物类型和准确的经纬度坐标, 包括水稻、大豆、玉米、林地和草地采样点共 872 个。大豆实测产量数据记录了样地的经纬度坐标, 共计 53 个样地, 择取 2/3 样地用于大豆单产估算, 剩余样地用于验证 (图 1)。

在大豆测产过程中, 每个测产地块采用对角线取样法采集 3 个样方, 单个样方的面积为 1 m<sup>2</sup>。采样过程中, 确保取样点距离地头 5 m 以上, 避开道路、水渠、地埂边缘、树荫、电线杆、水井等可能对测量结果产生影响的设施或物体。记录标准样方大豆株数, 将植株连根拔起 (注意地上部的完整性, 尽量避免炸荚和田间损失), 装袋并贴上标签。

室内进行百粒质量的测定。将大豆风干后脱粒 (籽粒含水量应小于 13%), 去除虫食粒和空瘪粒。将籽粒混匀, 用四分法分出 3 份 100 粒称量 (质量差不超过 0.2 g, 若超过 0.2 g, 则用四分法重分), 其平均值为样点的百粒质量, 取样地内 3 个样点的百粒质量平均值为该样地的百粒质量。

### 1.2.4 地理与统计数据

土壤属性数据来源于中国科学院南京土壤研究所 (http://soildata.issas.ac.cn/) 的 1:100 万中国土壤数据集, 该数据集包含了土壤类型、土壤氮磷钾含量、有机质含量、机械组成和 pH 值等一系列土壤属性数据, 分辨率为 30 m。

DEM 数据来源于地理空间数据云 (http://www.

gscloud.cn/) 的 ASTER GDEM 数字高程数据集, 空间分辨率为 30 m。基于 DEM 数据, 分别采用分段坡长法和汇流面积法计算坡长和坡度。

人口密度数据来源于 WorldPop 人口栅格数据集 (https://hub.worldpop.org/), 空间分辨率为 1 km, 该数据集基于夜间灯光数据、土地利用数据等通过随机森林算法来估算人口空间分布, 具有较高的精度; 文盲率数据来源于 2020 年第七次人口普查数据, 根据研究区各市县总人口数和受教育情况计算获得; 城镇化率来源于 2017—2021 年各市级国民经济和社会发展统计公报。

市级大豆补贴价格来自各省财政厅下发的生产者补贴工作实施方案与下辖各市人民政府公布的有关大豆生产者补贴发放公告和实地问卷调查; 市级大豆粮价来源于各省农业农村厅、农牧厅和农牧局公布的生资准备情况信息和全国粮价价格监测系统中的市级收购价; 市级肥料价格来自各省农业农村厅的各主要农资产品市场价格咨询。

市级化肥施用量、农业总产值、农田灌溉面积和农业机械总动力均来源于 2018—2022 年的《黑龙江省统计年鉴》《吉林省统计年鉴》《辽宁省统计年鉴》以及《内蒙古统计年鉴》。市级大豆种植面积来源于 2022 年各市级国民经济和社会发展统计公报、人民政府网和中国食品安全网。

表 1 随机森林分类特征

特征类型 Feature types	特征变量 Feature variables	来源 Source
地形特征 Terrain features	高程 坡度	SRTM 数据
光谱特征 Spectral signatures	B5 波段 B6 波段 B7 波段 B8 波段 绿度指数 (GI) 归一化水体指数 (NDWI) 修正型比值植被指数 (MSRI) 绿度归一化植被指数 (GNDVI) 叶绿素吸收比值指数 (MCARI) 三角植被指数 (TVI) 绿度植被指数 (GVI) 叶绿素植被指数 (GCVI) 比值植被指数 (SRI) 增强植被指数 (EVI) 植被水分含量指数 (LSWI) 转化叶绿素吸收指数 (TCARI) 重归一化植被指数 (RDVI) 归一化植被指数 (NDVI)	Sentinel-2-L2A
纹理特征 Texture features	对比度 (Con) 相关性 (Cor) 方差 (Var) 角二阶矩 (ASM) 反差分矩阵 (IDM) 熵 (Ent)	Sentinel-2 (灰度共生矩阵)
后向散射 Back scattering	后向散射系数 (VV) 后向散射系数 (VH)	Sentinel-1-GRD

## 1.3 研究方法

### 1.3.1 多特征随机森林

随机森林是一种集成机器学习方法, 具有不易过拟合、学习特征全面、可快速准确处理海量数据等优势,

满足在大区域尺度上区分作物类型的条件。基于 Sentinel-2 数据的红光和近红外反射计算全区域内的归一化植被指数 (normalized difference vegetation index, NDVI) 值, 采用 NDVI 阈值分割法对遥感影像进行预处理, 剔除非植被覆盖地类对后续分类的影响<sup>[39]</sup>。构建决策树样本集, 共计 1771 个样点, 包括水稻样点 442 个、大豆样点 430 个、玉米样点 550 个、森林样点 272 个和草地样点 77 个。综合考虑大豆生育期的光谱特征、纹理特征、后向散射特征和地形特征 4 种特征类型的 28 个分类特征变量, 建立特征集 (表 1)。随机构建样本集和特征集生成多棵决策树, 构建随机森林。

研究采用随机森林的袋外误差 (out-of-bag error, OOB 误差) 计算特征变量的重要性, 并进行归一化处理。

$$V(X^j) = \frac{1}{L} \sum_{i=1}^m (e_i^j - e_i) \quad (1)$$

$$V(X^j)_j = \frac{V(X^j) - V(X^j)_{\min}}{V(X^j)_{\max} - V(X^j)_{\min}} \quad (2)$$

式中  $X^j$  为第  $j$  个特征变量;  $e_i$  为每个决策树的袋外误差;  $e_i^j$  为随机为  $X^j$  赋值后计算的新袋外误差;  $V(X^j)$  为特征变量的重要性;  $V(X^j)_{\min}$  和  $V(X^j)_{\max}$  分别为特征变量重要性的最小值和最大值;  $V(X^j)_j$  为归一化后的特征变量重要性。

### 1.3.2 遥感估产

结合大豆生育期 LAI 和地面实测产量数据, 基于指数模型、幂函数模型、线性模型和对数模型构建产量估测模型, 最大决定系数 ( $R^2$ ) 被用来筛选与保留最佳估产模型。如表 2 所示, 夏季最大 LAI 数据构建的线性模型效果最好,  $R^2$  为 0.75, 因此选用该模型对研究区的大豆产量进行估算。

表 2 基于叶面积指数时序曲线特征参数的大豆估产模型

Table 2 Soybean yield estimation model based on leaf area index (LAI) temporal curve feature parameters

特征参数 Characteristic parameters	估产模型 Yield estimation models	决定系数 $R^2$
7 月平均 LAI	$Y = 662.86X + 1185.15$	0.54
7 月最大 LAI	$Y = -91.13X^2 + 1565.40X - 2180.52$	0.73
8 月平均 LAI	$Y = -367.20X^2 + 3370.87X - 3542.41$	0.58
8 月最大 LAI	$Y = -271.37X^2 + 2995.38X - 4330.84$	0.65
夏季最大 LAI	$Y = 531.31X + 585.33$	0.75
生育期平均 LAI	$Y = -207.67X^2 + 2318.12X - 1350.62$	0.64
生育期 LAI 标准差	$Y = -1557.99X^2 + 6749.27X - 3235.46$	0.58
夏季 LAI 增加值	$Y = 1179.3 \ln(X) + 5621.40$	0.41
生育期累积 LAI	$Y = -10.36X^2 + 730.68X - 8586.30$	0.36

### 1.3.3 空间自相关分析

采用全局莫兰指数 (Global Moran's I) 和局部莫兰指数 (Local Moran's I) 检验大豆单产的空间自相关性及空间聚集情况。全局空间自相关是对属性值在整个区域空间特征的描述, 也是最常用的全局自相关统计指标, 可从整体上反映区域空间相关性。局部莫兰指数可进一步量化空间要素与周围要素之间的差异程度及显著性, 能够识别出较大空间范围内的空间聚集、热点和空间异常点。计算式如下:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n \sum_{j=1}^n W_{ij}} \quad (3)$$

$$I_i = \frac{(x_i - \bar{x})}{S^2} \sum_{j=1}^n W_{ij} (x_j - \bar{x}) \quad (4)$$

式中  $I$  为全局莫兰指数;  $I_i$  为局部莫兰指数;  $n$  为空间单元数目;  $\bar{x}$  为变量的平均值;  $x_i$  和  $x_j$  分别为变量  $x$  在相邻配对空间点的取值;  $S$  为标准差;  $W_{ij}$  为邻接或距离空间权重矩阵。

### 1.3.4 标准差椭圆

标准差椭圆是一种常见的度量地理数据空间分布方向的空间统计方法, 其参数有重心、长轴和短轴等。重心反映地理要素的中心位置, 长轴表示地理要素的主要方向, 短轴表示地理要素分布的范围。计算式如下:

$$(X, Y) = \left( \frac{\sum_{i=1}^n w_i a_i}{\sum_{i=1}^n w_i}, \frac{\sum_{i=1}^n w_i b_i}{\sum_{i=1}^n w_i} \right) \quad (5)$$

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (w_i \hat{a}_i \cos \theta - w_i \hat{b}_i \sin \theta)^2}{\sum_{i=1}^n w_i^2}} \quad (6)$$

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^n (w_i \hat{a}_i \sin \theta - w_i \hat{b}_i \cos \theta)^2}{\sum_{i=1}^n w_i^2}} \quad (7)$$

式中  $(a_i, b_i)$  为点要素坐标;  $w_i$  为权重;  $(X, Y)$  为加权平均中心, 即重心;  $\hat{a}_i$ 、 $\hat{b}_i$  为不同点要素离平均中心的坐标偏差;  $\sigma_x$ 、 $\sigma_y$  分别为沿横轴和纵轴的标准差;  $\theta$  为椭圆方位角, 即正北方向顺时针至椭圆长轴的夹角。

### 1.3.5 地理探测器

地理探测器主要用于探测地理要素的差异性及其对研究对象空间分布的影响因素, 其功能包括因子探测器、交互作用探测器、风险探测器和生态探测器。因子探测器探测因变量的空间分异特征, 以及自变量对因变量的解释力, 满足分析大豆单产空间分异影响因素的需求。因此研究采用因子探测器, 将 25 种因子对大豆单产空间分异的影响进行量化 (表 3), 因子解释力用  $q$  值来衡量。计算式如下:

$$q = 1 - \frac{\sum_{h=1}^L N_h \sigma_h^2}{N \sigma^2} = 1 - \frac{S_{SW}}{S_{ST}} \quad (8)$$

$$S_{SW} = \sum_{h=1}^L N_h \sigma_h^2 \quad (9)$$

$$S_{ST} = N \sigma^2 \quad (10)$$

式中  $h=1, 2, \dots, L$  为因子的分层;  $N_h$  和  $N$  分别为  $h$  层和全区的单元数;  $\sigma_h^2$  和  $\sigma^2$  分别代表  $h$  层和全区的方差;  $S_{SW}$  和  $S_{ST}$  分别代表层内方差之和以及全区总方差;  $q$  的值域范围为  $[0, 1]$ ,  $q$  值越大表明因子对属性的解释力越强,  $q$  值等于 0 表明因子与属性之间无任何关联,  $q$  值等于 1 表明因子完全控制属性的空间分布。

表 3 大豆单产空间分异特征的影响因子

Table 3 Influencing factors of spatial heterogeneity of soybean yield per unit

因素类型 Factor types	各个因子 Each factors
自然因素 Natural factors	高程 ( $X_1$ )、坡度 ( $X_2$ )、 坡长 ( $X_3$ )、土壤类型 ( $X_4$ )、 土壤氮含量 ( $X_5$ )、土壤磷含量 ( $X_6$ )、 土壤 pH 值 ( $X_7$ )、黏土含量 ( $X_8$ )、 壤土含量 ( $X_9$ )、砂土含量 ( $X_{10}$ )、 有机质含量 ( $X_{11}$ )、年均降雨 ( $X_{12}$ )、 太阳辐射量 ( $X_{13}$ )、年均积温 ( $X_{14}$ )
	夜间灯光 ( $X_{15}$ )、人口密度 ( $X_{16}$ )、 化肥施用量 ( $X_{17}$ )、农业总产值 ( $X_{18}$ )、 农业灌溉面积 ( $X_{19}$ )、农业机械总动力 ( $X_{20}$ )、 文盲率 ( $X_{21}$ )、大豆补贴 ( $X_{22}$ )、 大豆粮价 ( $X_{23}$ )、肥料价格 ( $X_{24}$ )、 城镇化率 ( $X_{25}$ )

交互作用探测器可以识别影响因素是否对大豆单产具有独立影响，或者它们的相互作用是否增强（削弱）了大豆单产的解释力，是地理探测器相较于其他统计方法的最大优势。分别计算单因子和双因子的  $q$  值并比较其大小，基于判断依据确定两因子间交互作用的大小和方向（表 4）。

表 4 大豆单产空间分异特征影响因子 ( $X_1, X_2$ ) 交互作用类型Table 4 Interaction types of influencing factors ( $X_1, X_2$ ) of spatial heterogeneity of soybean yield of per unit

判断区间 Criterion of interval	交互作用 Interaction
$q(X_1 \cap X_2) < \min(q(X_1), q(X_2))$	非线性减弱
$\min(q(X_1), q(X_2)) < q(X_1 \cap X_2) < \max(q(X_1), q(X_2))$	单因子非线性减弱
$q(X_1 \cap X_2) > \max(q(X_1), q(X_2))$	双因子增强
$q(X_1 \cap X_2) = q(X_1) + q(X_2)$	独立
$q(X_1 \cap X_2) > q(X_1) + q(X_2)$	非线性增强

注： $q(X_1 \cap X_2)$  为因子  $X_1$  和  $X_2$  的交互作用； $\min(q(X_1), q(X_2))$  为因子  $X_1$  和  $X_2$  的作用两者取最小值； $\max(q(X_1), q(X_2))$  为因子  $X_1$  和  $X_2$  的作用两者取最大值； $q(X_1) + q(X_2)$  为因子  $X_1$  和  $X_2$  的作用两者求和。

Note:  $q(X_1 \cap X_2)$  means the interaction of the driving factor  $X_1$  and  $X_2$ ;  $\min(q(X_1), q(X_2))$  means the minimum of the driving force of  $X_1$  and driving force of  $X_2$ ;  $\max(q(X_1), q(X_2))$  means the maximum of the driving force of  $X_1$  and driving force of  $X_2$ ;  $q(X_1) + q(X_2)$  means the sum of the driving force of  $X_1$  and driving force of  $X_2$ .

风险探测器用于确定影响因子两个子域的大豆单产平均值是否具有显著差异，并通过  $t$  统计量检验。

$$t_{\bar{y}_{h=1}-\bar{y}_{h=2}} = \frac{\bar{y}_{h=1} - \bar{y}_{h=2}}{\left( \frac{\text{Var}(\bar{y}_{h=1})}{N_{h=1}} + \frac{\text{Var}(\bar{y}_{h=2})}{N_{h=2}} \right)^{1/2}} \quad (11)$$

式中  $\bar{y}_{h=1}$  为子区域  $h$  内的大豆单产均值；Var 代表方差。

在进行影响因素分析前需对全部因子进行格式转化以满足地理探测器的运行格式。首先，将人为因素进行栅格化处理，采用自然间断点法将所有因子重分类为 5 个类别。其次，考虑研究区范围和各因子的空间分辨率，布设并提取出与大豆单产空间分布相交的 10 km 渔网。利用分区统计功能，筛选出每个网格单元内各因子和大豆单产等级的众数，并将其作为该网格各因子及大豆单产的主导等级。最后，将全部因子与大豆单产进行空间连接，以完成格式转化。

## 2 结果与分析

### 2.1 东北地区大豆种植空间分布特征

#### 2.1.1 大豆种植提取及精度验证

根据随机森林算法的 OOB 误差，对参与分类的 28

个特征变量进行重要性计算，以评估各个特征变量对分类的贡献程度（图 2）。特征重要性排在前 8 位的依次为高程、VV、VH、LSWI、B5、NDVI、B6 和坡度，这与研究区地形起伏明显、植被类型丰富的情况相符合。其次是 GVI、TCARI、Con、MSR、GI、Var、B8、GNDVI、EVI、NDWI、B7、Cor、TVI、SRI、MCARI、Ent、RDVI、ASM 和 GCVI，IDM 对分类的贡献最小。从特征类型来看，特征变量的重要性贡献率由大到小依次为：地形、后向散射、光谱特征、纹理特征。

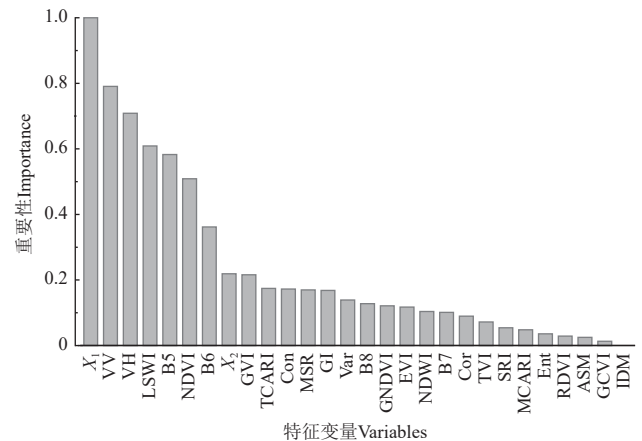


图 2 分类特征变量的重要性

Fig.2 Importance of classification feature variables

表 5 展示了植被类型验证样点对随机森林分类的精度评价结果。模型总体分类精度为 89.48%，Kappa 系数为 0.89。从不同地物类型来看，模型提取的大豆生产者精度和用户精度分别达到 90.32% 和 89.47%，误分主要出现在大豆与玉米之间。水稻和林地的生产者精度与用户精度均达到 89% 以上。玉米和草地的生产者精度与用户精度处于 89% 以下，其精度均低于其他植被类别，其中玉米的主要误分出现在玉米与大豆之间，草地的误分出现在草地与林地之间。

表 5 不同植被类型验证数据分类精度评价结果  
Table 5 Evaluation of classification precision for different vegetation type of the validate the data

项目 Item	混淆矩阵 Confusion matrix					生产者精度 Producer accuracy/%	用户精度 User accuracy/%
	水稻 Rice	大豆 Soybean	玉米 Maize	林地 Forest land	草地 Grass land		
水稻 Rice	79	8	1	0	0	89.77	90.34
大豆 Soybean	3	84	6	0	0	90.32	89.47
玉米 Maize	5	9	86	0	0	86	88.21
林地 Forest land	0	0	0	53	3	94.64	91.85
草地 Grassland	0	0	0	2	13	86.67	88.45

总体分类精度 Overall accuracy/% 89.48 Kappa 系数 Kappa coefficient 0.89

研究以地级市为统计单位，对遥感提取的 2022 年大豆种植面积与当年的统计数据进行了验证。结果显示，遥感数据的面积总量精度为 84.88%，与统计数据呈现显著的线性关系， $P < 0.01$ ， $R^2$  可达 0.92（图 3）。

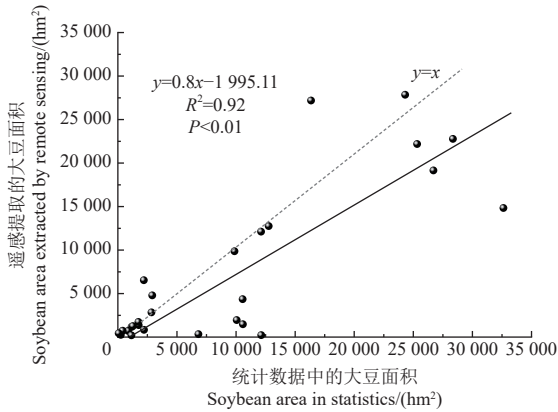


图 3 遥感提取的大豆面积与统计数据面积比较

Fig.3 Comparison with soybean area extracted by remote sensing and statistical data

2.1.2 大豆种植空间分布特征

图 4 展示的是 2022 年东北地区的大豆空间分布，种植面积整体上由北向南递减。黑龙江省的大豆种植面积最多，占大豆种植总面积的 62.37%，主要分布于黑河市、齐齐哈尔市和绥化市。其次是吉林省以及内蒙古自治区的呼伦贝尔市、通辽市、赤峰市和兴安盟地区。相比之下，辽宁省的大豆种植面积最少，仅占大豆种植总面积的 2.27%。值得注意的是，在县级尺度范围内，嫩江市、克山县和海伦市的大豆种植面积相对较多。



图 4 东北地区大豆空间分布图

Fig.4 Spatial distribution map of soybean in NEC

使用标准差椭圆对大豆种植区的重心、分布范围和分布方向进行分析。结果显示，大豆集中分布于松嫩平原地区，重心位于绥化市。标准椭圆差的长轴与短轴之间存在较大差异，表明其空间分布方向性显著，大致呈现以“黑河—齐齐哈尔—绥化—大庆—松原”的走向分布。沿短轴走向，朝西北方向延伸至呼伦贝尔市、兴安盟和白城市，向东南方向可延伸至伊春市、哈尔滨市和

长春市。

2.2 东北地区大豆单产空间分布特征

基于最佳估产模型获得东北地区的大豆单产，图 5 展示的是实测数据对模拟产量的验证结果。结果显示，实测数据与模拟产量之间呈现相关性， $R^2$  为 0.72，经显著性检验  $P<0.01$ 。

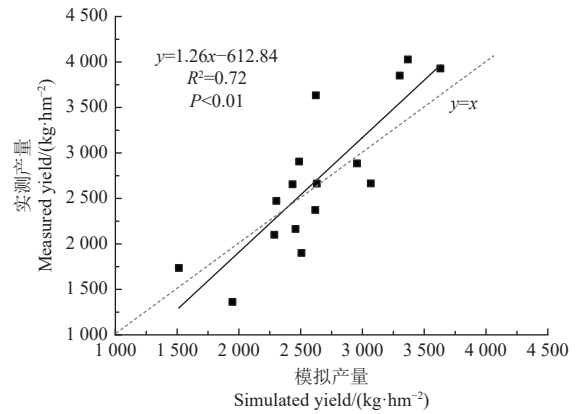


图 5 模拟的产量与实测产量比较

Fig.5 Comparison of simulated yield data and measured data points

当产量小于 2 381.18 kg/hm<sup>2</sup> 时，模拟产量值大于实测产量值。结果表明结合多时相 LAI 与实测产量建立的大豆估产模型，能够预测东北地区的大豆单产空间分布。

图 6 展示的是东北地区的大豆单产空间分布。结果显示，2022 年东北地区的大豆平均单产为 2 514.08 kg/hm<sup>2</sup>，大豆单产空间分布集聚性显著，呈现出北高南低的分布格局。单产高值区主要位于松嫩平原和三江平原地区，包括黑龙江省西部的齐齐哈尔市和绥化市、中北部的黑河市以及东部的佳木斯市等。辽河平原和松嫩平原南部的单产水平相对较低，以锦州市和通辽市为代表。值得注意的是，与其他地区相比，嫩江市、克山县和海伦市 3 个县级行政区的单产水平相对较高，其平均单产分别为 2 907.83、2 880.60 和 2 901.15 kg/hm<sup>2</sup>。

全局空间自相关分析结果显示，Moran's I 为 0.96，Z 得分为 8.69 远大于 2.58，且  $P<0.01$  通过显著性检验。结果表明，东北地区的大豆单产呈现显著的正相关性，在空间上存在明显的正向溢出影响，即区域大豆单产的提升或下降会导致临近地区大豆单产产生同方向的变动趋势。

研究区大豆单产的局部空间自相关以聚合性为主要聚类特征见图 7，HH 型和 LL 型占比分别为 39.16% 和 12.48%，不显著聚集特征占比可达 47.35%，离散型（HL 型和 LH 型）占比最低，仅为 1.01%。HH 型主要集中于研究区北部的黑河市、齐齐哈尔市和绥化市，少量分布于中部的松原市，该地区土壤肥沃，自然条件优越，农业机械化水平程度较高，易于带动周边地区发展，形成高耦合协调集聚区。LL 型主要分布于研究区西南部的通辽市、赤峰市、白城市、阜新市以及锦州市，结合大豆空间分布可知，该区域的大豆单产水平较低，极易形成 LL 集聚区。不显著聚集特征呈现全局分布，无明显的集中区。

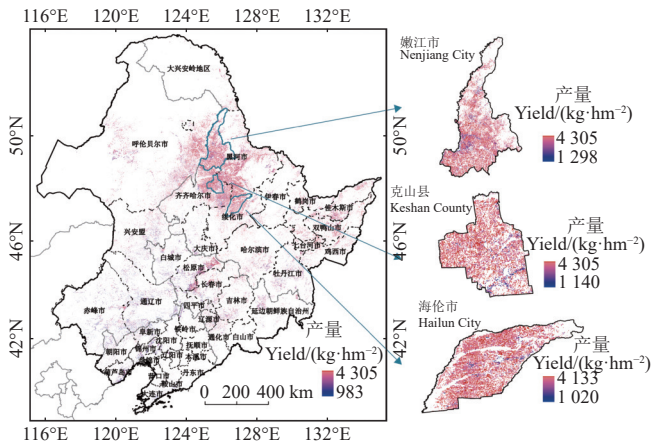
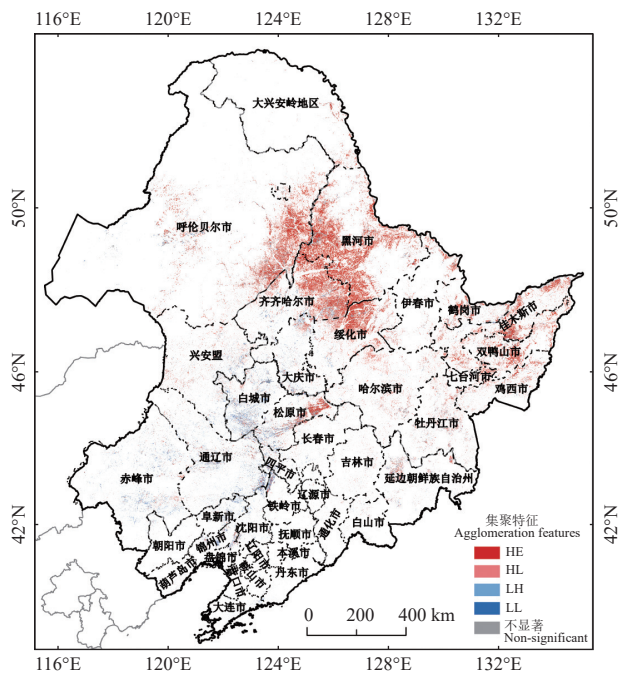


图6 东北地区大豆单产空间分布

Fig.6 Spatial distribution of soybean yield per unit in NEC



注: HH 为高高集聚; HL 为高低集聚; LH 为低高集聚; LL 为低低集聚。  
Note: HH represents high and clustering; HL represents high and low clustering; LH represents low and high clustering; LL refers to low and low clustering.

图7 东北地区大豆单产集聚特征空间分布

Fig.7 Spatial distribution of aggregated characteristics of soybean yield per unit in NEC

### 2.3 东北地区大豆单产空间分异的影响因素

#### 2.3.1 因子探测分析

图8展示的是影响因子对大豆单产空间分异的解释力,  $q$  值越大表明单因子的解释力越强。因子探测结果显示,  $X_4$  (土壤类型)、 $X_7$  (土壤 pH 值) 和  $X_{22}$  (大豆补贴) 对大豆单产空间分异的解释力最高,  $q$  值分别为 0.27、0.24 和 0.24; 其次为  $X_{23}$  (大豆粮价) 和  $X_5$  (土壤氮含量),  $q$  值均大于 0.22;  $X_3$  (坡长) 和  $X_{15}$  (夜间灯光) 的解释力最弱,  $q$  值分别为 0.01 和 0.004, 对大豆单产空间分异的影响最小。对大豆单产空间分异解释力最强的 6 项因子中, 包括 4 项自然因子:  $X_4$  (土壤类型)、 $X_7$  (土壤 pH 值)、 $X_5$  (土壤氮含量) 和  $X_{11}$  (有机质含量) 以及 2 项人为因子:  $X_{22}$  (大豆补贴) 和  $X_{23}$  (大豆粮价)。表明研究区的大豆单产空间分异特征受自然因素与

人为因素的共同影响, 其中自然因素起主要作用, 人为因素起调控作用。

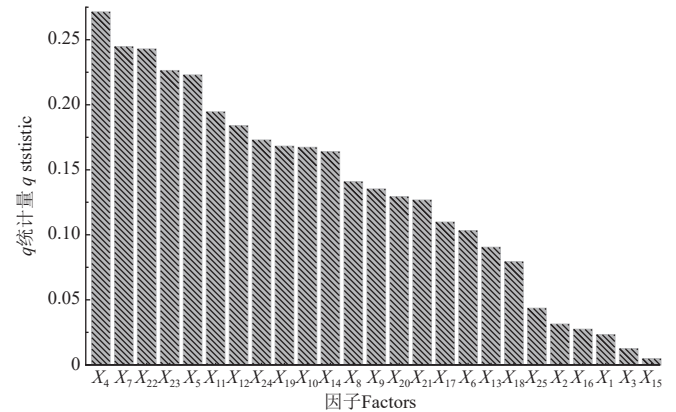


图8 单因子的解释力

Fig.8 Explanatory power of independent

#### 2.3.2 交互因子探测分析

图9展示的是交互因子对大豆单产空间分异的  $q$  值统计量,  $q$  值越大表明交互因子的解释力越强。结果显示,  $X_4$  (土壤类型)、 $X_5$  (土壤氮含量)、 $X_7$  (土壤 pH 值)、 $X_{11}$  (有机质含量)、 $X_{12}$  (年均降雨) 和  $X_{23}$  (大豆粮价) 6 种因子与其余因子交互的  $q$  值均高于 0.18, 解释力相对较高, 其中  $X_{12}$  (年均降雨) 和  $X_{14}$  (年均积温) 的交互作用对大豆单产的解释力最强,  $q$  值为 0.44。而  $X_2$  (坡度)、 $X_3$  (坡长)、 $X_{15}$  (夜间灯光) 和  $X_{16}$  (人口密度) 与其余因子的交互解释力普遍偏低,  $q$  值均小于 0.32, 其中  $X_3$  (坡长) 与  $X_{15}$  (夜间灯光) 的交互作用解释力最弱,  $q$  值仅为 0.02, 这与单一影响因子中  $X_3$  (坡长) 与  $X_{15}$  (夜间灯光) 对大豆单产空间分异解释能力最弱的结果相一致。

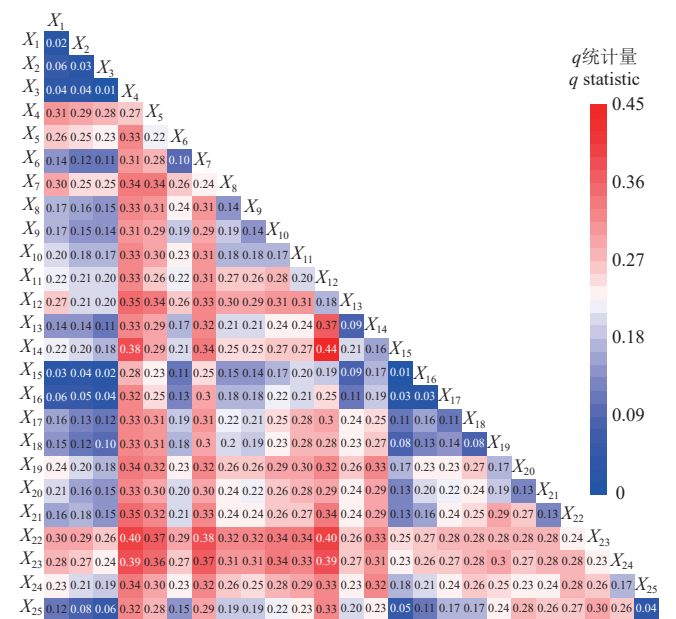


图9 交互影响因子的解释力

Fig.9 Explanatory power of an interactive driving factor

因子交互探测结果中, 任意两因子交互后的  $q$  值均大于单个因子的  $q$  值, 因子间的交互作用表现为双因子增强和非线性增强效果, 不存在独立起作用或交互作用

减弱的因子, 说明东北地区大豆单产空间分异是多因子共同作用的结果。

对大豆单产空间分异解释力最高的 8 对因子, 依次是:  $X_{12}$  (年均降雨) 和  $X_{14}$  (年均积温) 的交互作用,  $q$  值达 0.44;  $X_{12}$  (年均降雨) 和  $X_{22}$  (大豆补贴)、 $X_4$  (土壤类型) 和  $X_{22}$  (大豆补贴) 的交互作用,  $q$  值均为 0.40;  $X_4$  (土壤类型) 和  $X_{23}$  (大豆粮价) 以及  $X_{12}$  (年均降雨) 和  $X_{23}$  (大豆粮价) 的交互作用,  $q$  值皆为 0.39;  $X_7$  (土壤 pH 值) 和  $X_{22}$  (大豆补贴)、 $X_4$  (土壤类型) 和  $X_{14}$  (年均积温) 的交互作用,  $q$  值均为 0.38;  $X_{14}$  (年均积温) 和  $X_{13}$  (太阳辐射) 的交互作用,  $q$  值为 0.37。可以发现, 解释力最高的 7 对因子都涉及到  $X_{12}$  (年均降雨)、 $X_4$  (土壤类型) 和  $X_7$  (土壤 pH 值) 与其他因子的交互作用, 这 3 项自然类因子与人为因素中的  $X_{22}$  (大豆补贴) 和  $X_{23}$  (大豆粮价) 交互后的解释力均高于其单个自然类因子的解释力, 且单个自然类因子的解释力均高于单个人为类因子的解释力。

进一步表明自然因素对于大豆单产空间分异的形成起着主导作用, 人为因素起调控作用。

### 2.3.3 风险探测分析

为进一步探究人为因素对于大豆单产空间分异的最佳影响范围, 选取对大豆单产空间分异解释力较高的前 6 种人为因素, 包括大豆补贴、大豆粮价、农业灌溉面积、肥料价格、农业机械总动力和文盲率, 在一定的因子数值范围内进行风险探测分析。

由探测结果 (图 10) 可知, 大豆补贴和大豆粮价对大豆单产的影响趋势总体一致, 即因子的增加伴随着大豆单产均值的增加, 分别在 4 801~7 500 元/hm<sup>2</sup> 和 5 601~5 800 元/t 范围内达到最大值。与此相反的是肥料价格, 随着肥料价格的上涨, 大豆单产的均值下降, 当肥料价格处于 2 500~2 602 元/t 时, 大豆单产均值达到最大 (2 931.6 kg/hm<sup>2</sup>)。此外, 大豆单产均值大致随农业灌溉面积、农业机械总动力和文盲率的升高先增加后减小, 分别在 13.6×10<sup>4</sup>~26.4×10<sup>4</sup> hm<sup>2</sup>、252×10<sup>4</sup>~436×10<sup>4</sup> kW 和 1.4%~1.8% 范围内时, 大豆单产均值最高。由此说明, 合理控制农业灌溉面积和农业机械总动力有利于推动大豆增产稳产。

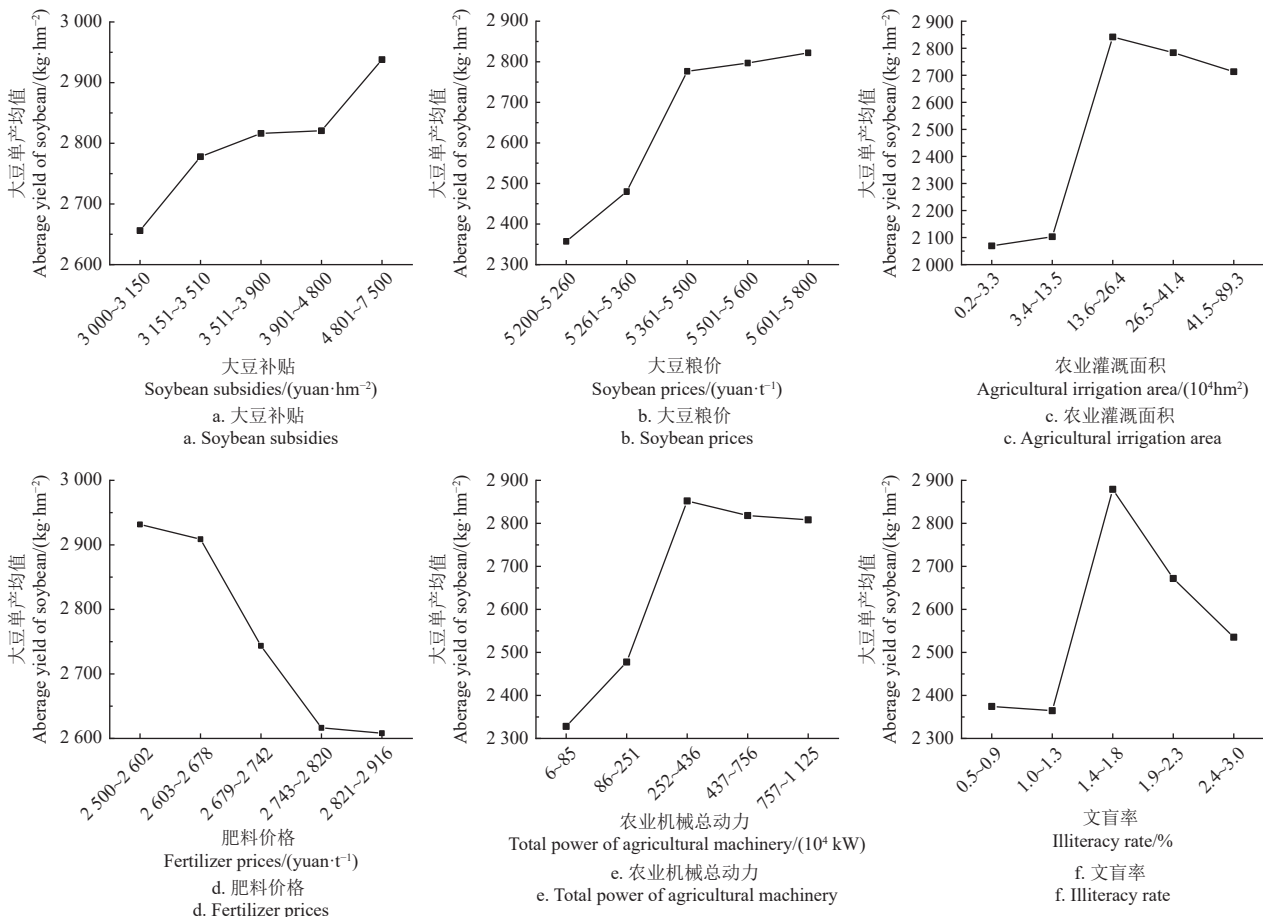


图 10 6 种人为因子的风险探测结果

Fig.10 Risk detector results for six anthropogenic factors

## 3 讨论

研究基于谷歌地球引擎 (Google Earth Engine, GEE) 平台, 采用多特征随机森林分类提取大豆种植空间分布的方法仍有不足之处。作物生长状态不一致导致的“同

物异征”会对分类结果产生影响。例如, 由于播种时间和地理位置的差异, 同一类作物可能具有多种生长期特征, 导致特征值相似的现象<sup>[40]</sup>。此外, 基于独立像元分类所产生的“椒盐现象”普遍存在于地物信息提取中, 会对最终分类精度造成一定影响<sup>[41]</sup>。值得注意的是, 大

豆玉米带状复合种植是东北地区常见的种植模式, 该种植模式引起的混合像元现象将影响大豆分类精度, 基于 Sentinel-1、Sentinel-2 和 Sentinel-3 对 LAI 进行反演的的方法或许可以用来解决这个问题<sup>[42]</sup>。同时, 不同数据源的遥感影像数据在时间和空间分辨率上的不匹配现象, 也会对结果造成一定的偏差<sup>[43-44]</sup>。

东北地区大豆单产空间分异特征受自然因素与人为因素的交互作用, 自然因素起主导作用, 人为因素起调控作用。受各地自然条件的影响, 不同地区种植的大豆品种存在差异, 大豆耐肥性、抗旱性和抗逆性等生长发育特性会对产量产生影响。

人为因素中的重要影响因子包括大豆补贴、肥料价格、农业灌溉面积、农业机械总动力和文盲率等。随着国家“大豆振兴计划”的实施和政府相关配套政策的执行, 如调高大豆补贴标准、支持农民参与大豆完全成本保险、加强高效节水灌溉设施建设和推广先进适用的大豆全程机械化技术及装备等, 农民种植大豆的积极性显著提高。当农业灌溉面积和农业机械总动力增加到一定程度时, 大豆单产反而出现略微的下降趋势。随着农业灌溉面积的增加, 大豆单产出现略微的下降趋势可能涉及到多个复杂因素, 这种现象通常被称为“灌溉诱导产量下降”或“灌溉效应递减”<sup>[45]</sup>。农业机械总动力增加, 大豆单产略有下降的现象通常被称为“机械化递减效应”或“机械化陷阱”<sup>[46]</sup>。文盲率降低, 农民在作物管理以及优质品种选择等方面的知识水平不断增强, 大豆单产水平得到提升<sup>[47]</sup>。当文化水平提高, 文盲率降低到一定水平时, 部分农民更倾向于追求非农就业, 农业劳动力流失现象加剧<sup>[48]</sup>, 导致部分地区的大豆单产水平降低。值得注意的是, 受影响的地区仅占东北大豆种植区面积的 4.34%, 对大豆种植区总体单产水平的影响并不显著。自 2019 年大豆振兴计划实施以来, 2022 年中国大豆产量创下历史新高, 但大豆价格的持续走低使得农民面临了“种大豆不如种玉米挣钱”的尴尬。针对这一问题, 要从全产业链角度出发, 既管“种”, 更管“收”。一方面, 政府应出台相应政策, 调节大豆和玉米的种植效益比; 另一方面, 政府通过实施储备吞吐机制, 引导价格保持在合理的范围内。此外, 还需进一步从市场机制和产业链发展等方面寻求创新策略。

## 4 结 论

研究基于谷歌地球引擎 (Google Earth Engine, GEE) 平台, 采用多特征随机森林的分类方法, 对 2022 年东北地区的大豆种植空间分布信息进行提取。结合多时相叶面积指数 (leaf area index, LAI) 数据与实测产量建立大豆估产模型, 明晰区域大豆单产空间分异特征。运用地理探测器进行定量分析, 明确大豆单产空间分异特征的影响因素。结论如下:

1) 大豆种植面积总体提取精度为 89.48%, Kappa 系数为 0.89, 与统计数据之间的  $R^2$  达 0.92。大豆种植面积由北向南递减, 主要分布于松嫩平原地区, 空间分布的重心位于绥化市。

2) 东北地区的大豆平均单产为 2 514.08 kg/hm<sup>2</sup>, 与实测产量的  $R^2$  为 0.72。大豆单产空间分布集聚性显著,

呈现出北高南低的格局特征。

3) 大豆单产空间分异特征受自然因素与人为因素的共同影响, 自然因素起主导作用, 人为因素起调控作用。土壤类型、土壤 pH 值和大豆补贴是解释大豆单产空间分异特征最重要的 3 种单因子,  $q$  值分别为 0.27、0.24 和 0.24; 年均降雨 $\cap$ 年均积温、年均降雨 $\cap$ 大豆补贴以及土壤类型 $\cap$ 大豆补贴是解释大豆单产空间分异特征最重要的 3 对交互因子,  $q$  值分别为 0.44、0.40 和 0.40。人为因素对大豆单产差异影响显著, 大豆补贴、大豆粮价、农业灌溉面积、农业机械总动力、肥料价格和文盲率的最佳影响范围分别为 4 801~7 500 元/hm<sup>2</sup>、5 601~5 800 元/t、13.6 $\times$ 10<sup>4</sup>~26.4 $\times$ 10<sup>4</sup> hm<sup>2</sup>、252 $\times$ 10<sup>4</sup>~436 $\times$ 10<sup>4</sup> kW、2 500~2 602 元/t 以及 1.4%~1.8%。

## [参 考 文 献]

- [1] 崔宁波, 董晋. 新时代粮食安全观: 挑战、内涵与政策导向[J]. 求是学刊, 2020, 47(6): 56-65.  
CUI Ningbo, DONG Jin. Food security outlook in the new era: challenges, connotation and policy[J]. Seeking Truth, 2020, 47(6): 56-65. (in Chinese with English abstract)
- [2] HAO C L, XIAO W H, ZHOU Y Y, et al. Phosphorus balance in typical rainfield of black soil region in Northeast China[J]. *Geosciences Journal*, 2019, 23(4): 637-648.
- [3] 王伟. 大豆种植全程解决方案[M]. 哈尔滨: 黑龙江人民出版社, 2021: 6-8.
- [4] 王丽美, 靳国旺, 熊新, 等. 耕地细碎化农业区冬小麦遥感制图方法[J]. *农业工程学报*, 2022, 38(22): 190-198.  
WANG Limei, JIN Guowang, XIONG Xin, et al. Winter wheat mapping in land fragmentation areas using remote sensing data[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(22): 190-198. (in Chinese with English abstract)
- [5] 孙中平, 刘素红, 姜俊, 等. 中高分辨率遥感协同反演冬小麦覆盖度[J]. *农业工程学报*, 2017, 33(16): 161-167.  
SUN Zhongping, LIU Suhong, JIANG Jun, et al. Coordination inversion methods for vegetation cover of winter wheat by multi-source satellite images[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2017, 33(16): 161-167. (in Chinese with English abstract)
- [6] 郭昱杉, 刘庆生, 刘高焕, 等. 基于 MODIS 时序 NDVI 主要农作物种植信息提取研究[J]. *自然资源学报*, 2017, 32(10): 1808-1818.  
GUO Yushan, LIU Qingsheng, LIU Gaohuan, et al. Extraction of main crops in Yellow River Delta based on MODIS NDVI time series[J]. *Journal of Natural Resources*, 2017, 32(10): 1808-1818. (in Chinese with English abstract)
- [7] DORAISWAMY P C, HATFIELD J L, JACKSON T J. Crop condition and yield simulations using Landsat and MODIS[J]. *Remote Sensing of Environment*, 2004, 4(92): 548-559.
- [8] 解毅, 张永清, 苟兰, 等. 基于多源遥感数据融合和 LSTM 算法的作物分类研究[J]. *农业工程学报*, 2019, 35(15): 129-137.  
XIE Yi, ZHANG Yongqing, XUN Lan, et al. Crop classification based on multi-source remote sensing data fusion and LSTM algorithm[J]. *Transactions of the Chinese Society*

- of Agricultural Engineering (Transactions of the CSAE), 2019, 35(15): 129-137. (in Chinese with English abstract)
- [9] MMAMOKOMA G M, ADRIAAN V N, ZAMA E M. Pre-harvest classification of crop types using a Sentinel-2 time-series and machine learning[J]. Computers and Electronics in Agriculture, 2020(169): 105164.
- [10] LUO K, LU L L, XIE Y H, et al. Crop type mapping in the central part of the North China Plain using Sentinel-2 time series and machine learning[J]. Computers and Electronics in Agriculture, 2023(205): 107577.
- [11] 赵静, 潘方红, 兰玉彬, 等. 无人机可见光遥感和特征融合的小麦倒伏面积提取[J]. 农业工程学报, 2021, 37(3): 73-80.  
ZHAO Jing, PAN Fanghong, LAN Yubin, et al. Wheat lodging area extraction using UAV visible light remote sensing and feature fusion[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(3): 73-80. (in Chinese with English abstract)
- [12] FERNANDO S R, CARLOS A, JOSE F O. Identification of tillage for soybean crop by spectro-temporal variables, GEOBIA, and decision tree[J]. Remote Sensing Applications: Society and Environment, 2020(19): 100356.
- [13] LI H, ZHANG C, ZHANG S. Crop classification from full-year fully-polarimetric L-band UAVSAR time-series using the random forest algorithm[J]. International Journal of Applied Earth Observation and Geoinformation, 2020(87): 102032.
- [14] 秦元伟, 赵庚星, 姜曙千, 等. 基于中高分辨率卫星遥感数据的县域冬小麦估产[J]. 农业工程学报, 2009, 25(7): 118-123.  
QIN Yuanwei, ZHAO Gengxing, JIANG Shuqian, et al. Winter wheat yield estimation based on high and moderate resolution remote sensing data at county level[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2009, 25(7): 118-123. (in Chinese with English abstract)
- [15] 赵明燕, 熊黑钢, 陈西玫. 新疆奇台县化肥施用量变化及其与粮食单产的关系[J]. 中国生态农业学报(中英文), 2009, 17(1): 75-78.  
ZHAO Mingyan, XIONG Heigang, CHEN Ximei. Effect of chemical fertilizer application on grain yield in Qitai County[J]. Chinese Journal of Eco-Agriculture, 2009, 17(1): 75-78. (in Chinese with English abstract)
- [16] 余强毅, 吴文斌, 陈羊阳, 等. 农作物空间格局变化模拟模型的 MATLAB 实现及应用[J]. 农业工程学报, 2014, 30(12): 105-114.  
YU Qiangyi, WU Wenbin, CHEN Yangyang, et al. Model application of an agent-based model for simulating crop pattern dynamics at regional scale based on MATLAB[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2014, 30(12): 105-114. (in Chinese with English abstract)
- [17] ALI M, HAMIDEH N, MAJID V. Improving crop yield estimation by assimilating LAI and inputting satellite-based surface incoming solar radiation into SWAP model[J]. Agricultural and Forest Meteorology, 2018, 250(3): 159-170.
- [18] 杨妮, 邓树林, 樊艳红, 等. 叶绿素荧光遥感反演及其农业监测应用研究进展[J]. 江苏农业科学, 2023, 51(14): 1-13.  
YANG Ni, DENG Shulin, FAN Yanhong, et al. Research progress of chlorophyll fluorescence retrieval by remote sensing and its application in agricultural monitoring[J]. Jiangsu Agricultural Sciences, 2023, 51(14): 1-13. (in Chinese with English abstract)
- [19] 孟庆岩, 李强子, 吴炳芳. 农作物单产预测的运行化方法[J]. 遥感学报, 2004, 8(6): 602-610.  
MENG Qingyan, LI Qiangzi, WU Bingfang. Operational method for crop yield prediction[J]. National Remote Sensing Bulletin, 2004, 8(6): 602-610. (in Chinese with English abstract)
- [20] RAI A S, TELMO J C, LUCIANA N, et al. Forecasting maize yield at field scale based on high resolution satellite imagery[J]. Biosystems Engineering, 2018(171): 179-192.
- [21] 张悦琦, 李荣平, 穆西晗, 等. 基于多时相 GF-6 遥感影像的水稻种植面积提取[J]. 农业工程学报, 2021, 37(17): 189-196.  
ZHANG Yueqi, LI Rongping, MU Xihan, et al. Extraction of paddy rice planting areas based on multi-temporal GF-6 remote sensing images[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(17): 189-196. (in Chinese with English abstract)
- [22] HASSANIJALILIAN O, IGATHINATHANE C, DAY S. Measuring soybean iron deficiency chlorosis progression and yield prediction with unmanned aerial vehicle[J]. Smart Agricultural Technology, 2023(4): 100204.
- [23] 祁向前, 孙德浩, 贾连星. 哨兵 2 号多时相植被指数作物分类及监测[J]. 测绘工程, 2022, 31(6): 47-53.  
QI Xiangqian, SUN Dehao, JIA Lianxing. Crop classification and monitoring of sentinel-2 multi-temporal vegetation index[J]. Engineering of Surveying and Mapping, 2022, 31(6): 47-53. (in Chinese with English abstract)
- [24] DAVID D T, BRADLEY A K, Dave L B. Maize grain yield and crop water productivity functions in the arid Northwest U S[J]. Agricultural Water Management, 2022(264): 107513.
- [25] 畅灼卓, 王雅情, 冯累, 等. 谷子精量播种机播种参数对晋谷 21 号生长特性与产量的影响[J]. 农业工程学报, 2020, 36(12): 22-29.  
CHANG Zhuozhuo, WANG Yaqing, FENG Lei, et al. Effects of sowing parameters of millet precision seeder on growth characteristics and yield of Jingu 21[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(12): 22-29. (in Chinese with English abstract)
- [26] 闫春娟, 王文斌, 曹永强, 等. 不同耐旱型大豆生理特性对不同降雨条件的响应[J]. 大豆科学, 2018, 37(3): 359-365.  
YAN Chunjuan, WANG Wenbin, CAO Yongqiang, et al. Response of physiological characteristics of different drought-tolerant soybean varieties to different rainfall climatic conditions[J]. Soybean Science, 2018, 37(3): 359-365. (in Chinese with English abstract)
- [27] 曹庆军, 杨粉团, 陈喜凤, 等. 播期对吉林省中部玉米生长发育、产量及品质的影响[J]. 玉米科学, 2013, 21(5): 71-75.  
CAO Qingjun, YANG Fentuan, CHEN Xifeng, et al. Effects of sowing date on growth, yield and quality of spring maize in the

- central area of Jilin province[J]. *Journal of Maize Sciences*, 2013, 21(5): 71-75. (in Chinese with English abstract)
- [28] 侯淑涛, 崔杨, 孟令华. 地形对漫川漫岗黑土区大豆产量的影响[J]. *农业工程学报*, 2020, 36(8): 88-95.  
HOU Shutao, CUI Yang, MENG Linghua. Effects of terrain on soybean yields in rolling hilly black soil areas[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2020, 36(8): 88-95. (in Chinese with English abstract)
- [29] JI L F, NI K, WU Z D, et al. Effort of organic substitution rates on soil quality and fungal community composition in a tea plantation with long-term fertilization[J]. *Biology and Fertility of Soils*, 2020, 56(5): 633-646.
- [30] 魏丹, 蔡姗姗, 王伟. 黑土肥力与大豆产量及品质的途径分析[J]. *大豆科学*, 2021, 40(1): 89-97.  
WEI Dan, CAI Shanshan, WANG Wei. Path analysis on black soil fertility via soybean yield and quality[J]. *Soybean Science*, 2021, 40(1): 89-97. (in Chinese with English abstract)
- [31] 潘开文, 何燕, 唐亚伟. 青藏高原农田生态系统与粮食安全科学考察研究专题进展[J]. *应用与环境生物学报*, 2022, 28(4): 813-821.  
PAN Kaiwen, HE Yan, TANG Yawei. Progress of the scientific expedition and research on farmland ecosystem and food security in Qinghai-Tibet Plateau[J]. *Chinese Journal of Applied and Environment*, 2022, 28(4): 813-821. (in Chinese with English abstract)
- [32] 刘长远, 王磊, 齐玉鑫, 等. 播期、密度及施肥量对牡豆 15 主要农艺性状和光合特性的影响[J]. *大豆科学*, 2023, 42(2): 212-219.  
LIU Changyuan, WANG Lei, QI Yuxin, et al. Effects of sowing date, density and fertilizer application amount on main agronomic traits and photosynthetic characters of Mudou 15[J]. *Soybean Science*, 2023, 42(2): 212-219. (in Chinese with English abstract)
- [33] 魏鹏飞, 徐新刚, 李中元, 等. 基于无人机多光谱影像的夏玉米叶片氮含量遥感估测[J]. *农业工程学报*, 2019, 35(8): 126-133.  
WEI Pengfei, XU Xingang, LI Zhongyuan, et al. Remote sensing estimation of nitrogen content in summer maize leaves based on multispectral images of UAV[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2019, 35(8): 126-133. (in Chinese with English abstract)
- [34] 邓陈威, 雷亚柯, 展世杰. 基于多种分析方法的花生农艺性状综合评价[J]. *山东农业大学学报 (自然科学版)*, 2023, 54(2): 194-200.  
DENG Chenwei, LEI Yake, ZHAN Shijie. Comprehensive evaluation of peanut agronomic traits based on multiple methods[J]. *Journal of Shandong Agricultural University(Natural Science Edition)*, 2023, 54(2): 194-200. (in Chinese with English abstract)
- [35] 白智媛, 陈向阳, 郑阿香. 1991—2019 年美国大豆区试品种 (系) 农艺和品质性状时空变化特征[J]. *作物学报*, 2023, 49(1): 177-187.  
BAI Zhiyuan, CHEN Xiangyang, ZHENG Axiang. Spatial-temporal variations for agronomic and quality characters of soybeans varieties(strains) tested in America from 1991 to 2019[J]. *Acta Agronomica Sinica*, 2023, 49(1): 177-187. (in Chinese with English abstract)
- [36] 王劲峰, 徐成东. 地理探测器: 原理与展望[J]. *地理学报*, 2017, 72(1): 116-134.  
WANG Jinfeng, XU Chengdong. Geo-detector: Principle and prospective[J]. *Acta Geographica Sinica*, 2017, 72(1): 116-134. (in Chinese with English abstract)
- [37] 叶妍君, 齐清文, 姜莉莉. 基于地理探测器的黑龙江垦区农场粮食产量影响因素分析[J]. *地理研究*, 2018, 37(1): 171-182.  
YE Yanjun, QI Qingwen, JIANG Lili. Impact factors of grain output from farms in Heilongjiang reclamation area based on geographical detector[J]. *Geographical Research*, 2018, 37(1): 171-182. (in Chinese with English abstract)
- [38] HUANG D H, SU L, FAN H M, et al. Identification of topographic factors for gully erosion susceptibility and their spatial modelling using machine learning in the black soil region of Northeast China[J]. *Ecological Indicators*, 2022(143): 1-13.
- [39] 杨喆, 褚琳, 肖焱彬, 等. 2017—2021 年东北地区主要粮食作物覆盖类型时空演变及其影响因素[J]. *资源科学*, 2023, 45(5): 966-979.  
YANG Zhe, CHU Lin, XIAO Yibin, et al. Spatiotemporal changes and driving forces of main grain crop planting types in Northeast China from 2017 to 2021[J]. *Resources Science*, 2023, 45(5): 966-979.
- [40] 戴建国, 张国顺, 郭鹏, 等. 基于无人机遥感可见光影像的北疆主要农作物分类方法[J]. *农业工程学报*, 2018, 34(18): 122-129.  
DAI Jianguo, ZHANG Guoshun, GUO Peng, et al. Classification method of main crops in northern Xinjiang based on UAV visible waveband images[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2018, 34(18): 122-129. (in Chinese with English abstract)
- [41] 杜培军, 夏俊士, 薛朝辉, 等. 高光谱遥感影像分类研究进展[J]. *遥感学报*, 2016, 20(2): 236-256.  
DU Peijun, XIA Junshi, XUE Zhaohui, et al. Review of hyperspectral remote sensing image classification[J]. *Journal of Remote Sensing*, 2016, 20(2): 236-256. (in Chinese with English abstract)
- [42] 周西嘉, 张悦, 王鹏新, 等. 基于 Sentinel 的时间序列田块尺度 LAI 重建与冬小麦估产[J]. *农业机械学报*, 2022, 53(8): 173-185.  
ZHOU Xijia, ZHANG Yue, WANG Pengxin, et al. Reconstruction of times series LAI and winter wheat yield estimation at field scales based on Sentinel satellites[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(8): 173-185. (in Chinese with English abstract)
- [43] 贾永红, 李德仁, 孙家柄. 多源遥感影像数据融合[J]. *遥感技术与应用*, 2000(1): 41-44.  
JIA Yonghong, LI Deren, SUN Jiabing. Data fusion techniques for multisources remotely sensed imagery[J]. *Remote Sensing Technology and Application*, 2000(1): 41-44. (in Chinese with English abstract)

- [44] 张良培, 武辰. 多时相遥感影像变化检测的现状与展望[J]. 测绘学报, 2017, 46(10): 1447-1459.  
ZHANG Liangpei, WU Chen. Advance and future development of change detection for multi-temporal remote sensing imagery[J]. Acta Geodaetica et Cartographica Sinica, 2017, 46(10): 1447-1459. (in Chinese with English abstract)
- [45] 王博, 王恒, 朱玉春. 制度能力对农田灌溉系统治理绩效的影响——基于黄河灌区六省调研数据的研究[J]. 中国人口·资源与环境, 2019, 29(8): 122-129.  
WANG Bo, WANG Heng, ZHU Yuchun, et al. Influence of institutional capacity on the governance performance of irrigation Systems: A study based on the survey data from six provinces of the Yellow River Irrigation Area[J]. China Population, Resources and Environment, 2019, 29(8): 122-129. (in Chinese with English abstract)
- [46] 罗敏, 曾以禹. 两型农业背景下的粮食生产——以贵州为例[J]. 农业技术经济, 2012(10): 52-58.
- [47] 秦婷婷, 曹鑫悦, 周泽群. 1952 年以来我国大豆单产变异特征及其影响因素研究[J]. 中国生态农业学报, 2022, 30(1): 47-56.  
QIN Tingting, CAO Xinyue, ZHOU Zequn. Variation characteristics of soybean yield since 1952 and its influencing factors in China[J]. Chinese Journal of Eco-Agriculture, 2022, 30(1): 47-56. (in Chinese with English abstract)
- [48] 姚成胜, 肖雅雯, 杨一单. 农业劳动力转移与农业机械化对中国粮食生产的关联影响分析[J]. 农业现代化研究, 2022, 43(2): 306-317.  
YAO Chengsheng, XIAO Yawen, YANG Yidan. The integrated impacts of rural labor transfer and agricultural mechanization on China's grain production[J]. Research of Agricultural Modernization, 2022, 43(2): 306-317. (in Chinese with English abstract)

## Spatial heterogeneity and determinants of soybean yield in Northeast China

WANG Chen, CHU Lin<sup>\*</sup>, YANG Zhe, YANG Zhenhao, ZHANG Xinya, WANG Tianwei, CAI Chongfa

(College of Resources and Environment, Huazhong Agricultural University 430070, China)

**Abstract:** Northeast China (NEC) has been the major soybean-producing region in China. Hence, it is very necessary to explore the spatial heterogeneity of soybean yield per unit in the NEC, in order to fully meet the current import and export production and demand. In this study, a multi-feature random forest (RF)-based classification was used to extract the spatial pattern of soybeans in 2022 using the Google Earth Engine (GEE) platform. The time series leaf area index (LAI) data was also combined with the field-measured yield. A soybean yield estimation model was established to characterize the spatial heterogeneity of regional soybean yield per unit. A geographic detector model was used to quantitatively explore the influencing factors. The results show that: 1) The overall accuracy of the soybean planting area reached 89.48% after extraction, with the Kappa coefficient of 0.89, and the coefficient of determination  $R^2$  was 0.92 between the soybean planting areas extracted from remote sensing and the statistical data of prefecture-level city. There was a marked spatial decrease in the planting area of soybeans from the northern to the southern NEC. The soybean planting areas were concentrated mainly in the Songnen Plain. Suihua City was found in the center of gravity for the soybean planting areas in the NEC. 2) The average soybean yield per unit was 2 514.08 kg/hm<sup>2</sup> in the NEC. The coefficient of determination  $R^2$  was 0.72, compared with the actual measured yield per unit. There was a significantly clustered spatial distribution of soybean yield per unit in the NEC. The areas with the high values were located mainly in the northern part of the NEC, whereas, the areas with the low values were in the southern. 3) Three dominant independent factors with the most pronounced spatial heterogeneity of soybean yield per unit were ranked in the descending order of the soil type, soil pH, and soybean subsidies, with q values of 0.27, 0.24, and 0.24, respectively. The three most significant interaction factors were to explain the spatial heterogeneity in the soybean yield per unit, including the interaction between mean annual rainfall and mean annual cumulative temperature, the interaction between mean annual rainfall and soybean subsidies, and the interaction between soil type and soybean subsidies, with q values of 0.44, 0.40 and 0.40, respectively. Six anthropogenic factors presented the significant impacts on the spatial heterogeneity of soybean yield per unit, namely soybean subsidies, soybean prices, agricultural irrigation area, total power of agricultural machinery, fertilizer prices, and illiteracy rate. Their optimal impact ranges varied significantly, where the optimal impact ranges were from 4 801 to 7 500 yuan/hm<sup>2</sup>, from 5 601 to 5 800 yuan/t, from  $13.6 \times 10^4$  to  $26.4 \times 10^4$  hm<sup>2</sup>, from  $252 \times 10^4$  to  $436 \times 10^4$  kW, from 2 500 to 2 602 yuan/t and from 1.4% to 1.8%, respectively. There was a significant spatial heterogeneity of soybean yield per unit in the NEC, with an overall decreasing trend from the north to the south. This variation trend can be primarily driven by natural factors also subjected to human intervention.

**Keywords:** soils; rainfall; crops; spatial heterogeneity; geo-detector; Northeast China; multi-feature random forest; soybean yield per unit