中国生态农业学报(中英文)
Chinese Journal of Eco-Agriculture

1952年以来我国大豆单产变异特征及其影响因素研究

秦婷婷，曹鑫悦，周泽群，褚超群，方雨桐，曲乐安，支俊俊，王　震，耿　涛

Variation characteristics of soybean yield since 1952 and its influencing factors in China

QIN Tingting, CAO Xinyue, ZHOU Zequn, CHU Chaoqun, FANG Yutong, QU Le’an, ZHI Junjun, WANG Zhen, and GENG Tao

在线阅读 View online: https://doi.org/10.12357/cjea.20210227

您可能感兴趣的其他文章

Articles you may be interested in

前茬作物秸秆还田下轮作模式和施肥对大豆产量的影响

Effects of rotational pattern and fertilization application on soybean yield under straws returning of preceding crop

中国生态农业学报(中英文). 2021, 29(9): 1493–1501

重庆市农田生态服务价值时空变化特征及其驱动因素分析

Temporal and spatial variation and driving factors of farmland ecological service value in Chongqing

中国生态农业学报(中英文). 2019, 27(11): 1637–1648

农户采纳稻虾共作模式意愿的影响因素及其异质性

Influencing factors of farmers’ willingness to adopt rice–crayfish co–culture and their heterogeneity

中国生态农业学报(中英文). 2021, 29(10): 1752–1761

河北坝上地区绿水时空变化及其驱动因素研究

Spatial–temporal changes in green water and its driving factors in the Bashang area of Hebei Province

中国生态农业学报(中英文). 2021, 29(6): 1030–1041

新疆荒漠植被的时空分布变化及其驱动因素

Spatio–temporal variation in and the driving factors of desert vegetation in Xinjiang

中国生态农业学报(中英文). 2021, 29(10): 1668–1678

福建耕地土壤磷素富集空间差异及其影响因素

Spatial variation in phosphorus accumulation and the driving factors in cultivated lands in Fujian Province

中国生态农业学报(中英文). 2018, 26(2): 274–283

关注微信公众号，获得更多资讯信息

中国生态农业学报(中英文) 2022 年 1 月 第 30 卷 第 1 期
Chinese Journal of Eco-Agriculture, Jan. 2022, 30(1): 47−56

DOI: 10.12357/cjea.20210227

秦婷婷, 曹鑫悦, 周泽群, 褚超群, 方雨桐, 曲乐安, 支俊俊, 王震, 耿涛. 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 (中英文), 2022, 30(1): 47−56

QIN T T, CAO X Y, ZHOU Z Q, CHU C Q, FANG Y T, QU L A, ZHI J J, WANG Z, GENG T. Variation characteristics of soybean yield since 1952 and its influencing factors in China[J]. Chinese Journal of Eco-Agriculture, 2022, 30(1): 47−56

1952 年以来我国大豆单产变异特征及其影响因素研究*

秦婷婷[1,2], 曹鑫悦[1], 周泽群[1], 褚超群[1], 方雨桐[1], 曲乐安[1], 支俊俊[1**], 王 震[1], 耿 涛[1]

(1. 安徽师范大学地理与旅游学院 芜湖 241002; 2. 中国矿业大学 (北京) 地球科学与测绘工程学院 北京 100083)

摘 要: 近几十年来, 我国大豆产需缺口不断扩大, 提升大豆单产水平已成为当前提高大豆总产量的首要可行举措。然而, 影响我国大豆单产的驱动因子及其地域空间差异特征并不明晰。本文通过搜集 1952 年、1965 年、1978 年、1990 年、2000 年、2010 年和 2017 年的全国各省市农业统计年鉴等数据, 从大豆种植的管理措施、自然因素、科技水平、社会因素、经济因素等方面选取与大豆生产密切相关的 13 个因子, 以大豆单产作为目标变量构建增强回归树模型, 量化各因子的相对重要性及其与大豆单产之间的关系, 分析大豆单产的变异特征, 揭示全国尺度及 4 个大豆主产区之间的大豆单产驱动力时空分异特征。研究结果表明: 1) 各年份的大豆单产变异系数为 34.1%~73.2%, 表明全国各地市大豆单产之间存在较大的差异。本研究构建的增强回归树模型可有效解释 43.3% 的大豆单产变异性, 并可量化揭示各因子与大豆单产之间的非线性关系。2) 1952 年以来影响我国大豆单产水平的最重要因素依次为大豆播种面积占农作物总种植面积的百分比 (相对重要性为 20.9%)、文盲率 (18.9%)、每公顷化肥 (折纯) 施用量 (10.7%)。3) 不同主产区的大豆单产核心驱动力存在空间差异, 北方春大豆区的最重要因素为每公顷农业机械总动力 (13.1%)、文盲率 (11.8%), 黄淮海流域夏大豆区的最重要因素为每公顷化肥 (折纯) 施用量 (25.6%)、每公顷农药 (折纯) 施用量 (18.4%), 长江流域春夏大豆区的最重要因素为研发支出占地区生产总值的百分比 (21.5%)、有效灌溉面积占农作物播种面积的百分比 (14.3%), 南方多熟大豆区的最重要因素为每公顷化肥 (折纯) 施用量 (22.7%)、第一产业占地区生产总值的百分比 (13.3%)。4) 大豆播种面积占农作物总播种面积的百分比对于全时期、改革开放前、改革开放后 3 个时期均是影响大豆单产最重要的因子, 改革开放前其他重要的因子包括文盲率和每公顷化肥 (折纯) 施用量, 改革开放后则包括每公顷农业机械总动力和年均温。总之, 我国各大豆主产区需合理施用化肥和农药, 努力提高机械化水平和农业生产者的知识水平, 本研究结果可为各省市采取有效措施提升大豆单产水平提供科学依据。

关键词: 大豆; 产量; 驱动因素; 增强回归树模型; 空间异质性

中图分类号: S210 **开放科学码(资源服务)标识码(OSID):**

* 教育部人文社会科学研究青年基金项目 (21YJCZH243)、国家自然科学基金项目 (41501229)、安徽师范大学博士科研启动金项目 (2018xjj45) 和安徽师范大学大学生创新创业训练计划项目 (202110370219) 资助

** 通信作者: 支俊俊, 主要研究方向为农业遥感与信息技术。E-mail: zhijunjun@ahnu.edu.cn
秦婷婷, 主要研究方向为资源与环境信息技术。E-mail: Shuua1999@163.com

收稿日期: 2021-04-14 接受日期: 2021-09-10

* This study was supported by the Youth Foundation Project of Humanities and Social Sciences of Ministry of Education in China (21YJCZH243), the National Natural Science Foundation of China (41501229), the Doctoral Scientific Research Foundation of Anhui Normal University (2018xjj45), and the Undergraduate Innovation and Entrepreneurship Training Program of Anhui Normal University (202110370219).

** Corresponding author, E-mail: zhijunjun@ahnu.edu.cn

Received Apr. 14, 2021; accepted Sep. 10, 2021

Variation characteristics of soybean yield since 1952 and its influencing factors in China*

QIN Tingting[1,2], CAO Xinyue[1], ZHOU Zequn[1], CHU Chaoqun[1], FANG Yutong[1], QU Le'an[1], ZHI Junjun[1**], WANG Zhen[1], GENG Tao[1]

(1. School of Geography and Tourism, Anhui Normal University, Wuhu 241002, China; 2. College of Geoscience and Surveying Engineering, China University of Mining & Technology (Beijing), Beijing 100083, China)

Abstract: Over the past several decades, the consumers' demand for soybeans has grown rapidly in China, resulting in a significant increase in the gap between production and demand. Therefore, increasing the total soybean output is of critical importance to ensure food security. Given that it is difficult to increase the total area of cultivated land in China, improving soybean yield per unit area land has become the primary measure for increasing the total soybean output. However, the determinants that directly affect soybean yield, the regional spatial heterogeneity of yield remain unclear. In this study, data from agricultural statistical yearbooks at both the provincial and prefecture levels in China as well as meteorological data (e.g., temperature, precipitation, and sunshine duration) from 1952 to 2017 (comprising 1952, 1965, 1978, 1990, 2000, 2010, and 2017) were collected, whereupon 13 factors closely related to soybean production were selected from the perspective of planting management measures, natural factors, scientific and technogical levels, social factors, and economic factors. Several boosted regression tree models were built to quantify the relative importance of each factor and to determine the mechanism through which it influenced soybean yield; to analyze the variation characteristics of soybean yield; and to reveal the spatiotemporal characteristics of key driving forces across the national scale and among the four major soybean-producing areas (i.e., the northern spring soybean area, the summer soybean area in the Huang-Huai-Hai Basin, the spring and summer soybean area in the Yangtze River Basin, and the southern soybean area) over a long period since 1952. The following results were obtained. 1) The coefficient of variation of soybean yields in different years ranged from 34.1% to 73.2%, indicating that there were substantial differences in yield across the regions in China. The boosted regression tree model could effectively explain 43.3% of the soybean yield variability and quantitatively revealed the nonlinear relationship between each factor and soybean yield in the national scale. 2) The most important factor affecting soybean yield in China since 1952 was the soybean sown area as a percentage of the total crop sown area (relative importance of 20.9%), followed by the illiteracy rate (18.9%) and fertilizer consumption (pure amount) per hectare (10.7%). 3) Spatial differences existed in the dominant driving factors of soybean yield among different main production areas. The main driving factors of the northern spring soybean area were the total power of agricultural machinery per hectare (13.1%) and the illiteracy rate (11.8%); those for the summer soybean area in the Huang-Huai-Hai Basin were the fertilizer consumption (pure amount) per hectare (25.6%) and pesticide consumption (pure amount) per hectare (18.4%); those for the spring and summer soybean area in the Yangtze River Basin were the R&D expenditure as a percentage of regional GDP (21.5%) and the effective irrigation area as a percentage of the crop sown area (14.3%); and those for the southern soybean area were the fertilizer consumption (pure amount) per hectare and the primary industry as a percentage of regional GDP (13.3%). 4) The soybean sown area as a percentage of the total crop sown area was the most important factor that affected soybean yield during 1952–2017, both before and after the reformation and opening up of China. Additionally, the illiteracy rate and fertilizer consumption (pure amount) per hectare were two other important factors for the period before the reformation and opening up of the country, whereas the total power of agricultural machinery per hectare and annual average temperature were important factors afterwards. This study revealed the determinants of soybean yield and its spatiotemporal heterogeneity in China since 1952 and determined the effective measures for improving the yield of this important crop. These findings should be useful for soybean production-related departments at both the provincial and prefecture levels in China for improving the rational usage of fertilizers and pesticides, increasing the level of mechanization, and enhancing the knowledge level of agricultural producers.

Keywords: Soybean; Yield; Driving factors; Boosted regression tree model; Spatial heterogeneity

随着居民消费结构升级, 大豆 (*Glycine max*) 需求快速增加, 国内产需缺口不断扩大。2018 年我国大豆 85% 依靠国际市场进口, 主要来自于巴西、美国和阿根廷[1]。一旦国际市场出现影响大豆供给的不利因素, 国内大豆及关联产业将会受到巨大影响。2019 年农业农村部印发大豆振兴计划实施方案, 提出了提高大豆种植面积、提高大豆单产、提升品质以及绿色发展等大豆生产振兴目标。目前, 我国大豆单产水平仅是世界大豆平均单产水平的 70%, 是大豆主产国美国的 60%[2]。因此, 在当前我国耕地总量难以明显增加的现状下, 提升大豆的单产水平成为提高我国大豆总产量的首要可行举措。目前国内外学者对制约大豆单产的影响因素研究多为田间尺度对照试验或长时间序列区域尺度下某单一因素对大豆单产的影响, 如 CO_2 浓度[3-5]、套作种植[6-8]、气候变化[9-11] 等因素。研究结果表明, CO_2 浓度升高可

提高 30.9% 的大豆单株荚数和百粒重[5], 与玉米 (*Zea mays*) 间作较单作具有显著的间作优势[7], 最高温度和降水量对大豆单产有积极的显著影响[9], 1971—2016 年间霜冻对大豆产量的影响在一定程度上以负效应为主[10]。

然而, 目前有关我国大豆单产制约因素的研究缺少多个因素贡献率之间的定量化计算与比较, 大豆单产与各影响因素之间的量化关系尚不明确, 且鲜有在全国尺度上开展大豆单产影响因素及其空间异质性的研究。在此背景下, 本文以全国各省市 1952—2017 年大豆产量相关的统计年鉴数据为基础, 从大豆种植的管理措施、自然因素、科技水平、社会因素、经济因素等方面选取影响大豆单产的因子, 基于增强回归树模型, 量化各因子的相对重要性及其与大豆单产之间的关系, 揭示 1952 年以来长时间序列下全国尺度和 4 个大豆主产区之间的大豆单产驱动力时空分异特征, 以及改革开放前后大豆单产的驱动因子。研究结果对切实提高我国大豆单产水平、指导和调控农业种植结构、规划国家和地区社会经济发展, 以及确保国家粮食安全等具有重要意义。

1　材料与方法

1.1　试验区概况

为揭示我国地域的大豆单产驱动因子空间差异特征, 本文综合气候类型、耕作制度和品种类型等因素, 将我国分为 4 个大豆主产区[12-13]: 1) 北方春大豆区, 包括黑龙江、吉林、辽宁、内蒙古、新疆、甘肃、陕西北部 (榆林、延安); 2) 黄淮海流域夏大豆区, 包括山东、山西、河南、河北、安徽北部 (宿州、阜阳、亳州)、江苏北部 (宿迁、徐州、淮安、连云港); 3) 长江流域春夏大豆区, 包括湖南、湖北、江西、浙江、四川、安徽中南部、江苏中南部、陕西南部 (汉中、安康、商洛); 4) 南方多熟大豆区, 包括广东、广西、云南、贵州、福建、海南。

1.2　研究数据

1.2.1　数据来源

本文所使用的相关农业统计数据主要来源于全国各省市 (不含港、澳、台地区) 统计部门正式出版的统计年鉴、农村统计年鉴、农业年鉴、农业统计资料、农村经济年鉴和城市年鉴等, 并利用“国家农业科学数据共享中心”平台获取的数据对缺失数据进行了补充完善。具体搜集了全国各省市 (不含港、澳、台地区)1952—2017 年 (1952 年、1965 年、1978 年、1990 年、2000 年、2010 年、2017 年共 7 个年份) 农作物总播种面积、大豆种植面积、大豆产量、农用化肥施用量、农药施用量、有效灌溉面积、农用机械总动力、第一产业生产总值、地区生产总值、文盲率、城市化率、农作物受灾面积、研发经费支出等数据。气象数据来源于国家农业科学数据共享中心网站 (http://data.cma.cn/) 发布的各气象观测站点 (共 2421 个站点) 数据, 包括年均温、年均降水量和年均日照时间, 对这些数据使用 ArcGIS 软件空间插值统计分析, 获得了 1952—2017 年全国各省市 500 m × 500 m 的气象要素 (年均值) 空间插值栅格数据。对数据进行单位统一换算、指标计算、涉及价格数据基准转换 (如将所有 GDP 数据统计换算成可比价)、异常值删除等处理, 得到共 2008 个样本数据。

1.2.2　指标选择

本文从大豆种植的管理措施[14]、自然因素[15-16]、科技水平[12-13, 17]、社会因素[18-19]、经济因素[20] 等方面选取了 13 个驱动因子 (表 1) 作为自变量, 以大豆单产作为目标变量, 构建增强回归树模型。

1.3　研究方法

1.3.1　增强回归树模型

增强回归树模型 (boosted regression trees, BRT) 也称为促进回归树模型。BRT 采用无放回抽样技术从训练样本中抽取一定比例 (一般≥50%) 的样本用于建模, 分析自变量对因变量的影响程度, 未被抽取的样本用于精度验证, 最后对生成的全部分类回归树取均值并输出, 从而提高模型的稳定性和预测精度[21-24]。BRT 的最大优点是不需要考虑自变量之间的交互作用, 数据允许存在缺省值并支持多种数据类型, 可量化各自变量因子对目标变量的影响程度 (相对贡献率) 以及与目标变量之间的非线性关系, 清晰直观、易于解释。

采用开源编程软件 R (version 3.6.3), 基于统计整理得到的 1952—2017 年 (1952 年、1965 年、1978 年、1990 年、2000 年、2010 年、2017 年共 7 个年份) 全国各省市共 2008 个样本点数据, 将大豆单产作为因变量, 13 种驱动因子作为自变量, 分别建立全时期 (1952—2017)、改革开放前 (1952 年、1965 年、1978 年) 和改革开放后 (1990 年、2000 年、2010 年、2017 年) 3 个模型, 使用 gbm 软件包 (version 2.1.8) 建立增强回归树模型[22]。

1.3.2　模型参数设置

模型中需要调试的参数有 4 个, 分别是: bagging

表 1 用于模型构建的大豆单产影响因子

Table 1 Selected influencing factors using for soybean yield modeling

划分依据 Division basis	驱动因子 Driving factor	简写 Abbreviation	单位 Unit	说明 Instruction
管理措施 Management measures	有效灌溉面积占农作物播种面积的百分比 Effective irrigation area as a percentage of crop sown area	EIAP	%	指示灌溉水平 Indicates irrigation level
	大豆播种面积占农作物总种植面积的百分比 Soybean sown area as a percentage of total crop sown area	SAP	%	指示大豆的生产规模 Indicates soybean production scale
自然因素 Natural factors	受灾害面积占农作物总播种面积的百分比 Disaster area as a percentage of total crop sown area	DAP	%	指示自然灾害 Indicates natural disasters
	年平均气温 Annual average temperature	AVT	℃	指示温度气候因子 Indicates temperature factor
	年平均日照时间 Annual sunshine duration	AST	h	指示光照时长气候因子 Indicates illumination duration factor
	年平均降水量 Annual average precipitation	AVP	mm	指示降水量气候因子 Indicates precipitation factor
科技水平 Scientific and technological level	每公顷农业机械总动力 Total power of agricultural machinery per hectare	AMP	$\times10^7$ W·hm^{-2}	指示机械化水平 Indicates the level of mechanization
	每公顷化肥(折纯)施用量 Fertilizer consumption (pure amount) per hectare	FCP	t·hm^{-2}	指示重要农业生产资料 Indicates important materials of agricultural production
	每公顷农药(折纯)施用量 Pesticides consumption (pure amount) per hectare	PCP	t·hm^{-2}	指示重要农业生产资料 Indicates important materials of agricultural production
	研发支出占地区生产总值的百分比 R&D expenditure as percentage of regional GDP	RDIG	%	指示研发投入 Indicates research and development investment
社会因素 Social factors	人口城镇化率 Population urbanization rate	PUR	%	指示地区城镇化水平 Indicates regional urbanization level
	文盲率 Illiteracy rate	ILR	%	指示地区受教育水平 Indicates regional educational level
经济因素 Economic factors	第一产业占地区生产总值的百分比 Primary industry as a percentage of regional GDP	PIG	%	指示农业在本地的经济地位 Indicates the economic status of agriculture in the local area

fraction (抽样率), 指的是随机抽取训练样本的比例, 用于构建单个决策树; learning rate (学习率), 决定了每棵决策树对最终模型的贡献大小; tree complexity (树复杂性), 决定了模型中预测因子交互作用的水平; ntree (树数量), 指的是模型中决策树的数量。本研究将 4 个参数分别设置为 0.60、0.005、3 和 20 000。最佳参数配置由基于 10-折交叉验证 (10-fold cross-validation) 的最高预测精度来确定, 模型的预测精度通过重复运行 50 次计算取平均值。

1.3.3 模型性能指标

模型性能通过 4 个指标进行评估: 决定系数 (coefficient of determination, R^2)、相关系数 (Pearson's correlation coefficient, r)、平均绝对误差 (mean absolute prediction error, MAE) 和均方根误差 (root mean square error, RMSE)。其中, R^2 和 r 的值越大、MAE 和 RMSE 的值越小代表模型的性能越优越。具体计算公式如下:

$$R^2 = \frac{\left\{\sum_{i=1}^{n}[(O_i-\overline{O})\times(P_i-\overline{P})]\right\}^2}{\sum_{i=1}^{n}(O_i-\overline{O})^2\times\sum_{i=1}^{n}(P_i-\overline{P})^2} \tag{1}$$

$$r = \frac{\sum_{i=1}^{n}[(O_i-\overline{O})\times(P_i-\overline{P})]}{\sqrt{\sum_{i=1}^{n}(O_i-\overline{O})^2}\times\sqrt{\sum_{i=1}^{n}(P_i-\overline{P})^2}} \tag{2}$$

$$\mathrm{MAE} = \sum_{i=1}^{n}\frac{|P_i-O_i|}{n} \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{n}\frac{(P_i-O_i)^2}{n}} \tag{4}$$

式中: O_i表观测值, P_i代表回归值, $\overline{O}$代表观测值的均值, $\overline{P}$代表回归值的均值。

2 结果与分析

2.1 1952—2017 年大豆单产变异特征

图 1 展示了 1952—2017 年期间 7 个年份大豆单产的变异特征。该时期内大豆单产总体呈增长趋势, 1952 年大豆单产均值为 0.816 t·hm^{-2}, 2017 年增长至 1.853 t·hm^{-2}, 是 1952 年的 2.27 倍, 年均增长率为 1.27%。值得注意的是, 1965 年的大豆单产均值为 0.715 t·hm^{-2}, 相比 1952 年下降 12.4%。从 1965 年开始, 大豆单产进入快速增长期, 至 2000 年增长至

1.656 t·hm^{-2}, 是 1965 年的 2.32 倍; 该期间年均增长率为 2.43%, 约为 1952—2017 年期间年均增长率的两倍。2000 年之后, 大豆单产增长趋于平缓, 至 2017 年期间的年均增长率降至 0.67%, 仅为 1952—2017 年期间年均增长率的一半。图 1 的箱形图展示了各年份大豆单产的地区变异, 总体而言, 各年份的大豆单产变异系数为 34.1%~73.2%, 属于中等程度变异, 离散程度相对较高, 表明全国各地市大豆单产之间存在较大的差异。

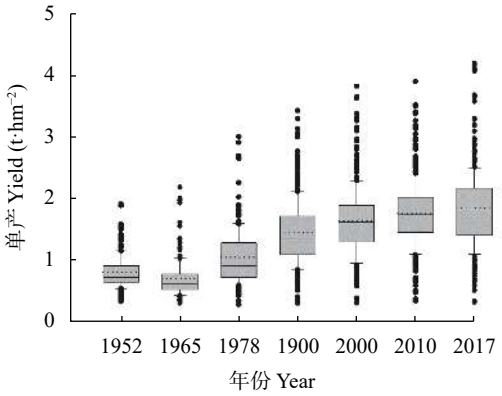

图 1　1952—2017 年中国大豆单产的变异特征

Fig. 1　Variations of soybean yield in different years from 1952 to 2017 in China

箱形图中的虚线表示均值, 实线表示中值, 上边框和下边框分别代表上四分位数和下四分位数, 上边线和下边线代表 10%~90% 的样本区间, 离散点代表异常值。Dotted lines in boxes represent mean values. Solid lines in boxes represent median values. Boxes show the 25%–75% quartiles. Whisker caps show the 10%–90% percentiles. Discrete points represent outlier values.

2.2　大豆单产模型的性能分析

全国尺度及 4 个大豆主产区的各模型性能结果如表 2 所示。全国尺度大豆单产驱动力 BRT 模型的性能结果 R^2 为 0.433, r 为 0.639, RMSE 为 0.097, MAE 为 0.467, 表明模型可有效解释全国尺度下 43.3% 的大豆单产变异性。最高解释度出现在黄淮海流域夏大豆区 (67.8%), 南方多熟大豆区的解释度也在 60.0% 以上, 长江流域春夏大豆区的解释度则接近 60.0%, 而北方春大豆区的解释度最低, 为 43.1%。北方春大豆区跨度大, 气候、土壤和种植结构等方面差异明显, 区域内大豆单产影响机制差异较大, 导致该模型性能低于其他地区。4 个大豆主产区模型性能指标中的相关系数 r 均高于 0.63, MAE 和 RMSE 均保持较低值, 表明所构建的 BRT 模型性能良好。将全国尺度数据划分改革开放前后两个时期分别建模, 改革开放前 R^2 为 0.486, 改革开放后 R^2 为 0.494, 模型性能较为稳定。

2.3　大豆单产驱动因子的相对重要性

由模型得到各驱动因子的相对影响, 即各因子对大豆单产的贡献率 (图 2)。1952 年至 2017 年, 对大豆单产影响最大的因子为大豆播种面积占农作物总播种面积的百分比 (SAP), 相对重要性为 20.9%。我国文盲率 (ILR) 从 78.6% 下降至 3.0%, 其对大豆单产的影响程度为 18.9%, 是促使大豆单产提升的重要因素。每公顷化肥 (折纯) 施用量 (FCP) 和每公顷农药 (折纯) 施用量 (PCP) 影响程度分别为 10.6% 和 8.5%。年均降水量 (AVP) 的值变化幅度不大, 有效灌溉面积占农作物播种面积中百分比 (EIAP) 增加明显, 两因子影响程度分别为 7.6% 与 7.4%。2017 年全国平均每公顷农业机械总动力 (AMP) 为 5.9 kW·hm^{-2}, 是 1952 年 0.0013 kW·hm^{-2} 的 4538 倍, 但其影响程度仅为 6.1%。年均温 (AVT) 因子紧随其后, 影响程度为 6.0%。其他因子影响程度较低, 依次为研发支出所占地区生产总值的百分比 (RDIG, 4.2%)、年均日照时间 (AST, 3.8%)、人口城镇化率 (PUR, 2.8%)、第一产业占地区生产总值的百分比 (PIG, 2.0%)、受灾害面积占农作物总播种面积的百分比 (DAP, 1.3%)。

大豆单产主要贡献因子, 改革开放前为大豆播种面积占农作物总播种面积的百分比 (SAP, 23.1%)、

表 2　用于大豆单产驱动力分析的增强回归树模型 (boosted regression trees, BRT) 的性能

Table 2　Performance of boosted regression trees (BRT) models for soybean yield analysis

指标 Index	全国 Nationwide	北方春大豆区 Northern spring soybean area	黄淮海流域夏大豆区 Summer soybean area in the Huang-Huai-Hai Basin	长江流域春夏大豆区 Spring and summer soybean area in the Yangtze River Basin	南方多熟大豆区 Southern soybean area	改革开放前 Before the reform and opening up	改革开放后 After the reform and opening up
R^2	0.433	0.431	0.678	0.586	0.608	0.486	0.494
r	0.639	0.639	0.821	0.753	0.769	0.581	0.614
MAE	0.467	0.467	0.301	0.408	0.350	0.506	0.334
RMSE	0.097	0.101	0.090	0.100	0.075	0.133	0.069

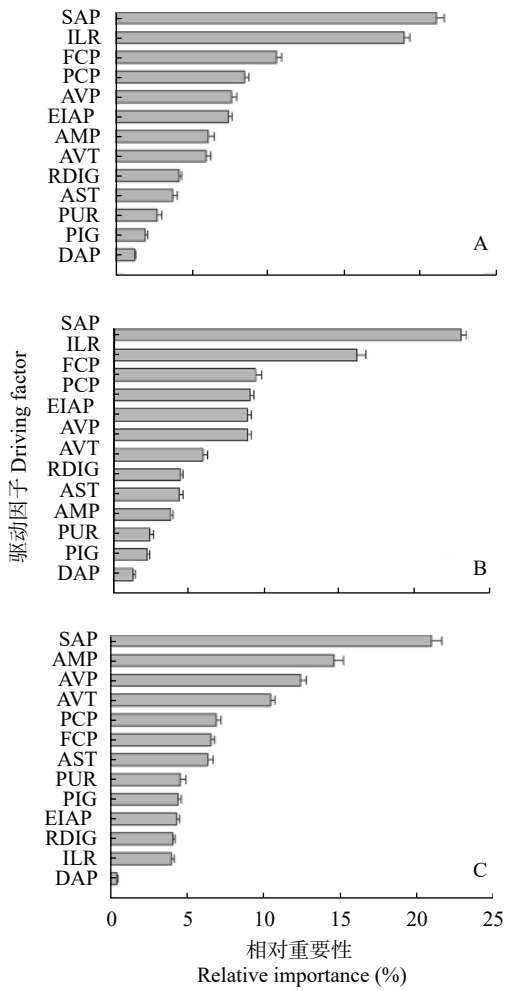

图 2 1952—2017 年 (A)、改革开放前 (1952—1978 年, B) 和改革开放后 (1985—2017 年, C) 大豆单产各驱动因子的相对重要性 (因子简写见表 1, 图中误差线表示模型运行 50 次计算求得的各因子重要性的标准差)

Fig. 2 Relative influence of each driving factor on soybean yield during the periods from 1952 to 2017 (A), before the reform and opening up (B, 1952—1978), and after the reform and opening up (C, 1985—2017) (See Table 1 for factors abbreviations, error bars represent standard deviations of the variable importance averaged over 50 model runs)

文盲率 (ILR, 16.2%) 和每公顷化肥 (折纯) 施用量 (FCP, 9.4%), 改革开放后为大豆播种面积占农作物总播种面积的百分比 (SAP, 20.9%)、每公顷农业机械总动力 (AMP, 14.5%) 和年均温 (AVT, 12.4%)。可以看出, 无论改革前、改革后, 还是全时期, SAP 都是各个时期对大豆单产影响最大的因子, 其重要性均超过 20%。

2.4 各驱动因子与大豆单产之间的非线性交互关系

各驱动因子与大豆单产之间的关系均呈非线性, 大豆单产与各驱动因子之间的定量关系如局部依赖图 (图 3) 所示。1952—2017 年, 全国大豆播种面积占农作物总种植面积的百分比 (SAP) 由 8.3% 下降至 6.0%, 大豆单产随大豆 SAP 的升高而下降直至趋于平稳。文盲率 (ILR) 高于 20% 时对大豆单产影响微弱, 但在 0~20% 范围内对随着 ILR 升高, 大豆单产显著降低。随着每公顷化肥 (折纯) 施用量 (FCP) 提升, 大豆单产明显提升, 当施用量超 0.5 $t \cdot hm^{-2}$ 后, 大豆单产基本保持不变; 每公顷农药 (折纯) 施用量 (PCP) 因子、每公顷农业机械总动力 (AMP) 因子影响机制与其类似。

大豆单产随着年均降水量 (AVP) 的增加波动变化, 年均降水量在 500~1000 mm 区间时, 年均降水量与单产呈正相关; 大豆单产与有效灌溉面积占农作物总播种面积的百分比 (EIAP) 因子整体呈正相关; 年均温 (AVT) 在 0~5 ℃ 时, 大豆单产最高, 随着温度上升大豆单产总体呈下降趋势; 日照时间 (AST) 的增加亦不能直接提高大豆单产。研发支出占地区生产总值 (RDIG) 的增加对大豆单产提升具有积极影响, 对大豆单产的贡献率为 4.2%。城镇化 (PUR) 的提升对大豆单产总体呈负面影响。第一产业占地区总产值的百分比 (PIG) 对大豆单产不仅贡献率较低, 对大豆单产提升也较为微弱。

2.5 大豆各主产区的核心驱动因子

大豆各主产区主要驱动因子不同, 但文盲率在各主产区均为主要驱动因子之一, 主要驱动因子空间分异如图 4 所示。影响北方春大豆区大豆单产的最重要因素为每公顷农业机械总动力 (AMP, 13.1%)、文盲率 (ILR, 11.8%)、大豆播种面积占农作物总种植面积的百分比 (SAP, 11.7%), 影响黄淮海流域夏大豆区大豆单产的最重要因素为每公顷化肥施用量 (FCP, 25.6%)、每公顷农药 (折纯) 施用量 (PCP, 18.4%)、文盲率 (ILR, 15.1%), 影响长江流域春夏大豆区大豆单产的最重要因素依次为研发支出占地区生产总值的百分比 (RDIG, 21.5%)、有效灌溉面积占农作物播种面积的百分比 (EIAP, 14.3%)、文盲率 (ILR, 11.8%), 影响南方多熟大豆区大豆单产的最重要因素依次为文盲率 (ILR, 35.4 %)、每公顷化肥 (折纯) 施用量 (FCP, 22.9%)、每公顷农药 (折纯) 施用量 (PCP, 7.8%)。

3 讨论

3.1 全国尺度下大豆单产的影响因子

研究结果表明, 全时期、改革开放前、改革开放后 3 个时期大豆单产的最主要影响因子均为大豆播种面积占农作物总播种面积的百分比。改革开放前大豆单产的其他重要影响因子还包括文盲率与每

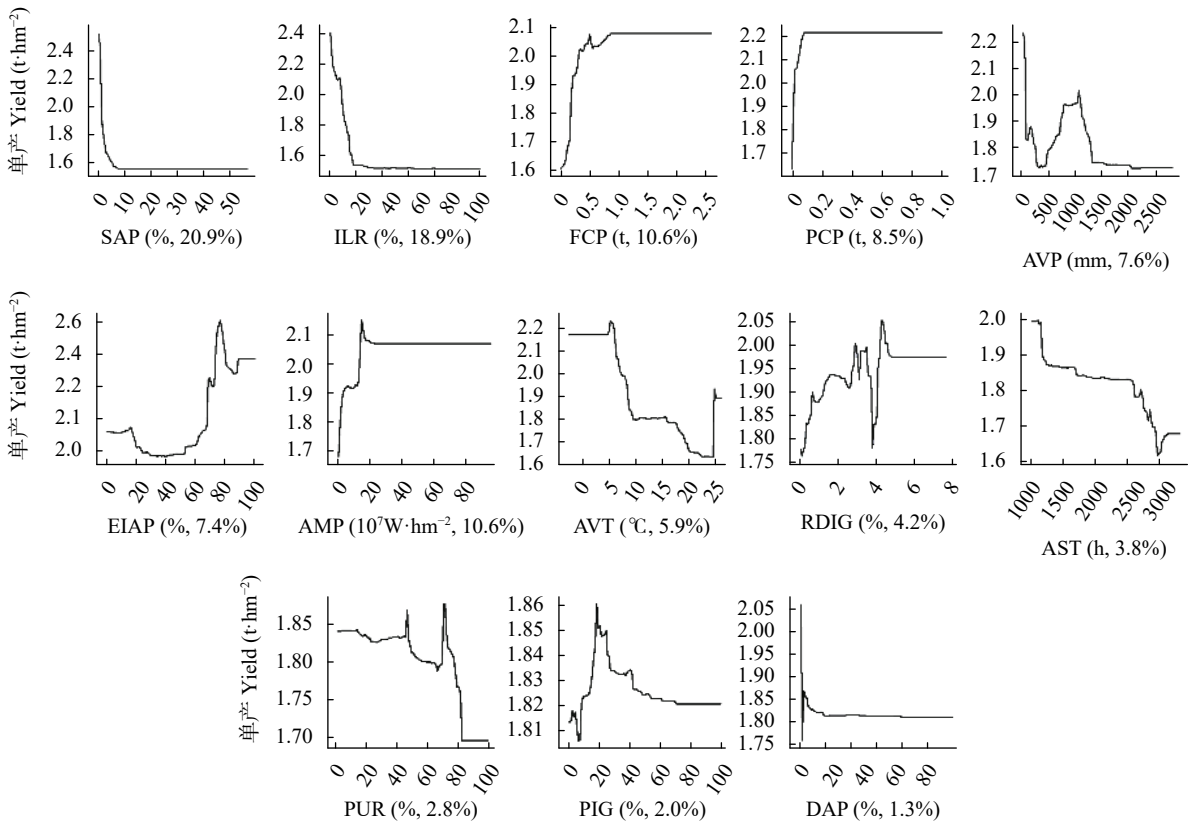

图 3　各驱动因子与大豆单产之间的非线性交互关系 (指标全称见表 1, 因子名后括号内分别为因子的单位及相对重要性)

Fig. 3　Interactive nonlinear relationships between soybean yield and its impact driving factors (See Table 1 for factors abbreviations, the unit and relative importance of each factor were showed in the bracket after factor's names).

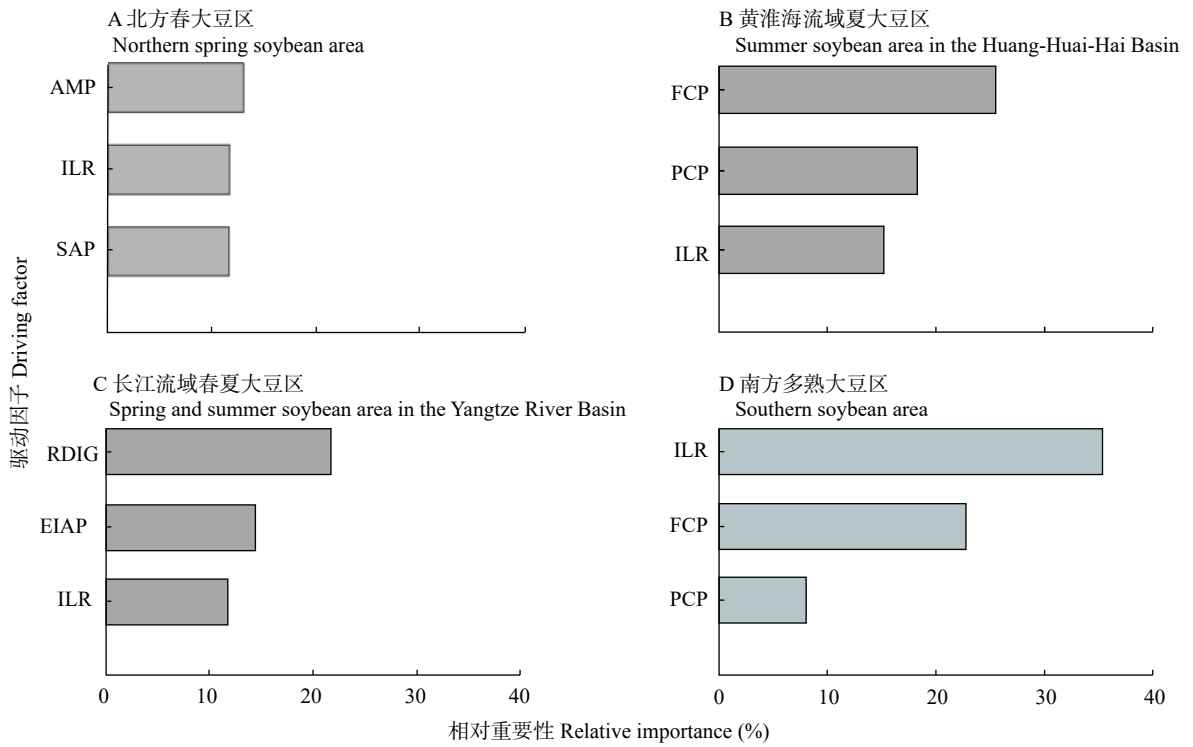

图 4　1952—2017 年中国大豆主产区的大豆单产主要驱动因子 (指标全称见表 1)

Fig. 4　Top three determinants of soybean yield in the four major soybean producing areas of China during the period from 1952 to 2017 (See Table 1 for factor abbreviations)

公顷化肥 (折纯) 施用量, 改革开放后还包括每公顷农业机械总动力与年均温。文盲率降低, 农民知识水平提高, 有助于提高单产。1949 年我国文盲率高达 80%, 改革开放以后, 国家出台一系列政策来支持农村义务教育事业发展, 2000 年我国已经实现了“基本普及九年义务教育、基本扫除青壮年文盲”的历史性任务, 2017 年我国文盲率仅 3%。义务教育的普及提高了农村劳动力的知识水平, 职业教育培养了一批农村实用型人才[25], 显著降低了文盲率, 使得农民在优质品种选择、作物生长管理以及农业科技成果应用等方面的知识水平大幅提升, 并进一步提升农业生产效率和农业技术水平, 从而提高大豆单产。

3.2 影响因子的相对重要性及非线性关系

扩大大豆种植面积、提高大豆在农作物播种总面积中的占比能够有效提高大豆单产[26-28], 其机理在于提高大豆在当地农作物生产中的地位能够促使农民提高耕作管理水平 (如减少重迎茬比例、增加轮作), 增加农业生产资料的投入 (如优良品种、化肥和农药施用量、机械化作业等), 以及吸引政府加大耕作补贴、农业保险、贷款等方面的扶持力度, 从而降低生产成本和提高自然灾害风险的应对水平, 进一步提高大豆单产。因此, 在提高大豆种植面积的同时, 需增加其他生产要素的投入力度, 以进一步形成规模化效益[29]。本研究结果中大豆单产与大豆播种面积占农作物总种植面积的百分比之间的交互作用关系反映的主要是大豆播种面积占比在时间尺度上的变化, 与 1952—2017 年期间大豆播种面积由 1.1679 万 hm^2 下降至 0.8245 万 hm^2, 但大豆单产从 0.816 $t \cdot hm^{-2}$ 公顷增长至 1.853 $t \cdot hm^{-2}$ 的趋势一致。为进一步揭示同一年份不同地区之间大豆单产存在差异性的关键影响因子, 在今后研究中可考虑搜集县级统计资料以显著扩大样本数量, 从而构建单个年份的大豆单产与各指标之间的相互作用关系来进行分析。

从大豆生长情况看, 多年使用有机肥料土壤中钾、磷、氮等物质的含量已经接近饱和, 仅增加氮肥、钾肥等有机型的肥料施用量而不提高其他微量元素如锌、锰等的投入, 则很难对大豆生长起促进作用。农药对防病治虫、促进粮食和农业稳产高产至关重要, 但区域中也存在滥用现象。因此, 化肥与农药须合理施用, 否则容易造成土壤养分失调、地力下降, 导致大豆抗逆性差、产量降低。大豆单产与生育期的水分、光照、热量等条件关系密切[30], 大豆单产可能受积温和种植制度等的影响较大, 因此, 有必要进一步分区进行统计分析[31]。

3.3 大豆主产区驱动因子空间分异特征

由于大豆四大主产区所处地理环境、种植规模、经济发展水平等条件不同, 大豆单产的核心驱动力存在空间差异。北方春大豆区, 大豆种植规模大, 地形辽阔平整, 区域内各省份大豆全程机械化水平略有不同。其中, 东北春大豆区中集中分布的松花江和辽河流域平原地带基本实现了大豆全程机械化生产, 在耕整地、播种、收获等各个环节均有适合该地区生产的成熟机具, 大豆生产综合机械化水平高。黄土高原春大豆种植区域气候寒凉, 土质瘠薄, 品种中耐瘠薄、耐旱的黑豆较多。西北春大豆种植区域年均降水量少, 土壤蒸发量大, 种植大豆须人工灌溉。

黄淮海流域夏大豆区是我国大豆的第二大主产区, 是高蛋白大豆产区[32]。每公顷化肥施用量是该地区的最重要驱动因子, 土壤养分供应、氮磷钾养分吸收规律、田间施肥技术、肥料损失等的区域异质性高, 不同区域应根据当地的气候状况、土壤特征、种植制度和大豆肥效等因素, 制定专用复合 (混) 肥料配方[13]。黄淮海流域夏大豆产区在耕种环节仍然存在上茬作物秸秆处理难的问题, 大豆收获主要依靠稻麦联合收获机, 机具适应性差、收获损失大、破碎程度高。

四大产区中长江流域最注重科技投入, 1952—2017 年期间研发支出占地区生产总值的百分比平均为 0.8%, 大豆单产的提高需要加大科技投入, 实施大豆重大科研联合攻关, 培育出适合地区种植的优质大豆种苗, 并加快大豆产业领域新技术、新成果的应用与转化。

与北方及黄淮海大豆产区相比, 南方基本以种植水稻 (Oryza *sativa*) 为主, 旱地则有玉米 (*Zea mays*)、甘薯 (*Dioscorea esculenta*)、油菜 (*Brassica nupus*) 等作物, 耕作制度比较复杂[33]。大豆种植方式多样, 如与旱地作物实行间套混作或利用田埂隙地栽培等。南方大豆产区机械化生产受品种、地形、种植模式等多方面因素制约, 大豆机械化收获推进缓慢, 机械化生产水平相对较低, 尤其是收获环节, 缺乏专用的大豆收获机具。可加大大豆耕种收获等环节机具购置补贴力度, 推广适用的大豆全程机械化技术及装备, 提升大豆生产科技水平。第一产业占地区生产总值百分比和年均降水量是该地区除化肥外的主要影响因子, 南方多熟大豆产区雨量充沛, 为大豆生长提供了丰富的水分, 1952—2017 年间年均降水量为 1599.4 mm, 是北方春大豆产区的 3.5 倍。

3.4 增强回归树模型的特点及结果不确定性

与前人研究通常使用的线性回归模型和决策树模型等相比, 本文所构建的增强回归树模型可量化各驱动因子对大豆单产的影响 (相对重要性) 以及大豆单产与单个驱动因子之间的非线性关系 (局部依赖图), 其拟合多个树的过程能最大程度地弥补单一树模型预测能力弱的缺点, 比大多传统模型方法有更强的预测能力, 采用的无放回抽样技术可提升模型稳定性与模拟精度[22], 并能有效评估模型结果的不确定性 (由经过多次模型重复运行计算结果的标准差表达)。

本文结果的不确定性主要来源于数据的搜集质量, 由于统计数据时间跨度大、覆盖范围广, 早期的统计数据与实际数据可能存在一定偏差; 其次, 尽管本文对全部数据进行了统一基准等处理, 但各时期统计口径及标准等方面存在的差异可能增大模型结果的不确定性; 再次, 由于本文使用的统计数据是以地级市为单位进行的汇总数据, 模型拟合结果只能细化表达至地级市之间的空间差异特征。

4 结论

本研究构建的增强回归树模型不仅量化了各影响因子对于大豆单产的相对重要性, 并且有效揭示了大豆单产与各影响因子之间的非线性关系。全国尺度下, 各时期大豆单产的最重要影响因子均为大豆播种面积占农作物总播种面积的百分比, 对大豆单产有重要影响的因子在改革开放前还包括文盲率和每公顷化肥 (折纯) 施用量, 在改革开放后还包括每公顷农业机械总动力和年均温。不同大豆主产区的大豆单产驱动力存在空间异质性, 但文盲率在各主产区均为主要驱动因子之一。研究结果可为国家及各省市今后提升大豆单产水平所应采取的有效措施以及调整农业种植结构等提供决策依据, 各省市需合理施肥和播撒农药, 努力提高机械化水平和农业生产者的知识水平。

参考文献 References

[1] 中华人民共和国国家统计局. 中国统计年鉴[M]. 北京: 中国统计出版社, 2019
National Bureau of Statistics of the People's Republic of China. China Statistical Yearbook[M]. Beijing: China Statistics Press, 2019

[2] 张晓峰, 王宏志, 刘洛, 等. 近50年来气候变化背景下中国大豆生产潜力时空演变特征[J]. 地理科学进展, 2014, 33(10): 1414−1423
ZHANG X F, WANG H Z, LIU L, et al. Spatial-temporal characteristics of soybean production potential change under the background of climate change over the past 50 years in China[J]. Progress in Geography, 2014, 33(10): 1414−1423

[3] MILLER J E, HEAGLE A S, PURSLEY W A. Influence of ozone stress on soybean response to carbon dioxide enrichment: Ⅱ. biomass and development[J]. Crop Science, 1998, 38(1): 122−128

[4] MORGAN P B, BOLLERO G A, NELSON R L, et al. Smaller than predicted increase in aboveground net primary production and yield of field-grown soybean under fully open-air [CO_2] elevation[J]. Global Change Biology, 2005, 11(10): 1856−1865

[5] 郝兴宇, 林而达, 杨锦忠, 等. 自由大气CO_2浓度升高对夏大豆生长与产量的影响[J]. 生态学报, 2009, 29(9): 4595−4603
HAO X Y, LIN E D, YANG J Z, et al. Effects of free air CO_2 enrichment (FACE) on growth and yield of summer soybean[J]. Acta Ecologica Sinica, 2009, 29(9): 4595−4603

[6] 杨峰, 娄莹, 廖敦平, 等. 玉米–大豆带状套作行距配置对作物生物量、根系形态及产量的影响[J]. 作物学报, 2015, 41(4): 642−650
YANG F, LOU Y, LIAO D P, et al. Effects of row spacing on crop biomass, root morphology and yield in maize–soybean relay strip intercropping system[J]. Acta Agronomica Sinica, 2015, 41(4): 642−650

[7] 任媛媛, 王志梁, 王小林, 等. 黄土塬区玉米大豆不同间作方式对产量和经济收益的影响及其机制[J]. 生态学报, 2015, 35(12): 4168−4177
REN Y Y, WANG Z L, WANG X L, et al. The effect and mechanism of intercropping pattern on yield and economic benefit on the Loess Plateau[J]. Acta Ecologica Sinica, 2015, 35(12): 4168−4177

[8] 杨文钰, 雍太文, 任万军, 等. 发展套作大豆, 振兴大豆产业[J]. 大豆科学, 2008, 27(1): 1−7
YANG W Y, YONG T W, REN W J, et al. Develop relay-planting soybean, revitalize soybean industry[J]. Soybean Science, 2008, 27(1): 1−7

[9] SATARI YUZBASHKANDI S, KHALILIAN S. On projecting climate change impacts on soybean yield in Iran: an econometric approach[J]. Environmental Processes, 2020, 7(1): 73−87

[10] 解文娟, 杨晓光, 杨婕, 等. 气候变化背景下东北三省大豆干旱时空特征[J]. 生态学报, 2014, 34(21): 6232−6243
XIE W J, YANG X G, YANG J, et al. Spatio-temporal characteristics of drought for soybean under climate change in the three provinces of Northeast China[J]. Acta Ecologica Sinica, 2014, 34(21): 6232−6243

[11] 郝兴宇, 韩雪, 居煇, 等. 气候变化对大豆影响的研究进展[J]. 应用生态学报, 2010, 21(10): 2697−2706
HAO X Y, HAN X, JU H, et al. Impact of climatic change on soybean production: a review[J]. Chinese Journal of Applied Ecology, 2010, 21(10): 2697−2706

[12] 乔金友, 姜岩, 王博, 等. 我国大豆主产区农业机械化现状及发展策略研究[J]. 农机化研究, 2017, 39(4): 1−6
QIAO J Y, JIANG Y, WANG B, et al. Research on current status and developing strategy of agricultural mechanization of soybean main production area in China[J]. Journal of

Agricultural Mechanization Research, 2017, 39(4): 1−6

[13] 车升国. 区域作物专用复合(混)肥料配方制定方法与应用[D]. 北京: 中国农业大学, 2015
CHE S G. Design method and application of formula of regional crop-based compound fertilizer[D]. Beijing: China Agricultural University, 2015

[14] 李秀芬, 郭昭滨, 朱海霞, 等. 黑龙江省大豆生长季旱涝时序特征及其对产量的影响[J]. 应用生态学报, 2020, 31(4): 1223−1232
LI X F, GUO Z B, ZHU H X, et al. Time-series characteristics of drought and flood in spring soybean growing season and its effect on soybean yield in Heilongjiang Province, China[J]. Chinese Journal of Applied Ecology, 2020, 31(4): 1223−1232

[15] 宫丽娟, 李秀芬, 田宝星, 等. 黑龙江省大豆不同生育阶段干旱时空特征[J]. 应用气象学报, 2020, 31(1): 95−104
GONG L J, LI X F, TIAN B X, et al. Spatio-temporal characteristics of drought in different growth stages of soybean in Heilongjiang[J]. Journal of Applied Meteorological Science, 2020, 31(1): 95−104

[16] 苏营, 张逸飞, 牟文雅, 等. 大豆主要株型和产量指标对大气CO_2和温度升高的响应[J]. 生态学报, 2016, 36(9): 2597−2606
SU Y, ZHANG Y F, MOU W Y, et al. Morphological traits and yield of soybean under elevated atmospheric CO_2 concentration and temperature[J]. Acta Ecologica Sinica, 2016, 36(9): 2597−2606

[17] 张伟明, 管学超, 黄玉威, 等. 生物炭与化学肥料互作的大豆生物学效应[J]. 作物学报, 2015, 41(1): 109−122
ZHANG W M, GUAN X C, HUANG Y W, et al. Biological effects of biochar and fertilizer interaction in soybean plant[J]. Acta Agronomica Sinica, 2015, 41(1): 109−122

[18] 王玎. 山东省城镇化对粮食生产的影响机制研究[D]. 济南: 山东大学, 2014
WANG D. Mechanism of impact of urbanization on food production in Shandong Province[D]. Jinan: Shandong University, 2014

[19] 杨志海, 王雅鹏. 城镇化影响了粮食安全吗?——来自1462个县(市)面板数据的实证检验[J]. 城市发展研究, 2012, 19(10): 1−5
YANG Z H, WANG Y P. Does urbanization influence food safety? Empirical test based on the panel data of 1462 counties (cities)[J]. Urban Development Studies, 2012, 19(10): 1−5

[20] 杨树果. 产业链视角下的中国大豆产业经济研究[D]. 北京: 中国农业大学, 2014
YANG S G. Economics of soybean industry in China from industry chain perspective[D]. Beijing: China Agricultural University, 2014

[21] 李春林, 刘淼, 胡远满, 等. 基于增强回归树和Logistic回归的城市扩展驱动力分析[J]. 生态学报, 2014, 34(3): 727−737
LI C L, LIU M, HU Y M, et al. Driving forces analysis of urban expansion based on boosted regression trees and Logistic regression[J]. Acta Ecologica Sinica, 2014, 34(3): 727−737

[22] ELITH J, LEATHWICK J R, HASTIE T. A working guide to boosted regression trees[J]. Journal of Animal Ecology, 2008, 77(4): 802−813

[23] PRASAD A M, IVERSON L R, LIAW A. Newer classification and regression tree techniques: bagging and random forests for ecological prediction[J]. Ecosystems, 2006, 9(2): 181−199

[24] 蔡文华, 杨健, 刘志华, 等. 黑龙江省大兴安岭林区火烧迹地森林更新及其影响因子[J]. 生态学报, 2012, 32(11): 3303−3312
CAI W H, YANG J, LIU Z H, et al. Controls of post-fire tree recruitment in Great Xing'an Mountains in Heilongjiang Province[J]. Acta Ecologica Sinica, 2012, 32(11): 3303−3312

[25] 周济. 大力办好农村教育事业[J]. 中小学校长, 2008(12): 4–7
ZHOU J. We will do a good job in rural education[J]. China's Rural Education, 2008(12): 4–7

[26] 李立军. 中国耕作制度近50年演变规律及未来20年发展趋势研究[D]. 北京: 中国农业大学, 2004
LI L J. Involving regularities in recent 50 years and future trends in 2020 on farming system in China[D]. Beijing: China Agricultural University, 2004

[27] LIOUBIMTSEVA E, HENEBRY G M. Grain production trends in Russia, Ukraine and Kazakhstan: New opportunities in an increasingly unstable world?[J]. Frontiers of Earth Science, 2012, 6(2): 157−166

[28] PRISHCHEPOV A V, PONKINA E, SUN Z L, et al. Revealing the determinants of wheat yields in the Siberian breadbasket of Russia with Bayesian networks[J]. Land Use Policy, 2019, 80: 21−31

[29] 王济民. 我国的大豆经济: 供给与需求的重点分析[D]. 北京: 中国农业科学院, 2000
WANG J M. Soybean economy in China: with focus on supply and demand[D]. Beijing: Chinese Academy of Agricultural Sciences, 2000

[30] 王畅, 赵海东, 冯乃杰, 等. 两个生态区大豆光热资源利用率和产量的差异及对化控剂的响应[J]. 应用生态学报, 2018, 29(11): 3615−3624
WANG C, ZHAO H D, FENG N J, et al. Differences in light and heat utilization efficiency and yield of soybean in two ecological zones and their response to chemical control regulators[J]. Chinese Journal of Applied Ecology, 2018, 29(11): 3615−3624

[31] 程叶青, 张平宇. 东北商品粮基地粮食生产的区域分异[J]. 自然资源学报, 2005, 20(6): 925−931
CHENG Y Q, ZHANG P Y. Regional differentiation of grain production of commodity grain bases in the northeast China[J]. Journal of Natural Resources, 2005, 20(6): 925−931

[32] 成雪峰, 张凤云. 黄淮海夏大豆生产现状及发展对策[J]. 大豆科学, 2010, 29(1): 157−160
CHENG X F, ZHANG F Y. Present conditions and countermeasures of soybean production in Huang-Huai-Hai regions[J]. Soybean Science, 2010, 29(1): 157−160

[33] 蒋慕东. 二十世纪中国大豆改良、生产与利用研究[D]. 南京: 南京农业大学, 2006
JIANG M D. Study on improvement, production and utilization of soybean in China during the 20th Century[D]. Nanjing: Nanjing Agricultural University, 2006