

密级:

论文编号:

中国农业科学院 学位论文

中国花生黄曲霉毒素污染风险预警模型研究

Researches on Risk Prediction Model of Aflatoxin
Contamination in China

博士研究生: 武琳霞

指导教师: 李培武 研究员

申请学位类别: 农学博士

专 业: 农产品质量与食品安全

研 究 方 向: 质量安全标准与检测

培 养 单 位: 油料作物研究所

研究生院

2019年5月

密级:

论文编号:

中国农业科学院 学位论文

中国花生黄曲霉毒素污染风险预警模型研究

Researches on Risk Prediction Model of Aflatoxin
Contamination in China

博士研究生: 武琳霞

指导教师: 李培武 研究员

申请学位类别: 农学博士

专 业: 农产品质量与食品安全

研 究 方 向: 质量安全标准与检测

培 养 单 位: 油料作物研究所

研究生院

2019年5月

Secrecy:

No.

Chinese Academy of Agricultural Sciences Dissertation

Researches on Risk Prediction Model of Aflatoxin Contamination in China

Ph.D. Candidate: Wu Linxia

Supervisor: Professor Li Peiwu

Major: Agro-product Quality&Safety

Specialty: Standards and Testing of Quality & Safety

May 2019

独创性声明

本人声明所提交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得中国农业科学院或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

研究生签名：武林霞

时间：2019年5月29日

关于论文使用授权的声明

本人完全了解中国农业科学院有关保留、使用学位论文的规定，即：中国农业科学院有权保留送交论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。同意中国农业科学院可以用不同方式在不同媒体上发表、传播学位论文的全部或部分内容。

研究生签名：武林霞

时间：2019年5月29日

导师签名：李洪

时间：2019年5月29日

中国农业科学院

博士学位论文评阅人、答辩委员会签名表

| | | | | | | | |
|------------------|-----|---------------------|---|---|-------------|-----------|-----|
| 论文题目 | | 中国花生黄曲霉毒素污染风险预警模型研究 | | | | | |
| 论文作者 | | 武琳霞 | 专业 | 农产品质量与食品安全 | 研究方向 | 质量安全标准与检测 | |
| 指导教师 | | 李培武 | | 培养单位(研究所、中心) | | 油料作物研究所 | |
| 姓名 | | 职称 | 硕(博)导师 | 单 位 | 专 业 | 签 名 | |
| 评 阅 人 | | | 硕导 <input type="checkbox"/> 博导 <input type="checkbox"/> | | | / | |
| | | | 硕导 <input type="checkbox"/> 博导 <input type="checkbox"/> | | | / | |
| | | | 硕导 <input type="checkbox"/> 博导 <input type="checkbox"/> | | | / | |
| 答 辩 主 席 | | 喻大昭 | 研究员 | 硕导 <input type="checkbox"/> 博导 <input checked="" type="checkbox"/> | 湖北省农科院 | 农学 | 喻大昭 |
| 答 辩 委 员 | 栗茂腾 | 教授 | 硕导 <input type="checkbox"/> 博导 <input checked="" type="checkbox"/> | 华中科技大学 | 作物遗传育种 | 栗茂腾 | |
| | 张椿雨 | 教授 | 硕导 <input type="checkbox"/> 博导 <input checked="" type="checkbox"/> | 华中农业大学 | 作物遗传育种 | 张椿雨 | |
| | 李学宝 | 教授 | 硕导 <input type="checkbox"/> 博导 <input checked="" type="checkbox"/> | 华中师范大学 | 植物遗传 | 李学宝 | |
| | 马立新 | 教授 | 硕导 <input type="checkbox"/> 博导 <input checked="" type="checkbox"/> | 湖北大学 | 生物化学 | 马立新 | |
| | 冯健雄 | 研究员 | 硕导 <input type="checkbox"/> 博导 <input checked="" type="checkbox"/> | 江西省农科院 | 农产品加工与贮藏 | 冯健雄 | |
| | 陈洪 | 研究员 | 硕导 <input type="checkbox"/> 博导 <input checked="" type="checkbox"/> | 中国农科院油料所 | 食品加工与安全 | 陈洪 | |
| | 吴刚 | 研究员 | 硕导 <input type="checkbox"/> 博导 <input checked="" type="checkbox"/> | 中国农科院油料所 | 生物化学与分子生物学 | 吴刚 | |
| | 张奇 | 研究员 | 硕导 <input type="checkbox"/> 博导 <input checked="" type="checkbox"/> | 中国农科院油料所 | 农产品质量安全检测技术 | 张奇 | |
| 会议记录(秘书) | | 李慧 | | | | | |
| 论文答辩时间地点 | | 2019年5月27日 油料所三楼会议室 | | | | | |

摘 要

花生 (*Arachis hypogaea* L.) 是我国主要油料作物和优势出口农产品, 是食油兼用型经济作物, 其营养价值高兼具食疗功能, 在中国及世界油料生产中占有重要地位。然而花生易受真菌感染产生黄曲霉毒素, 严重影响人体健康, 已成为影响我国花生消费安全和世界贸易的主要制约因素。花生黄曲霉毒素污染可以发生在种植、收获、储藏、运输、销售等各个环节, 其中收获前田间污染过程既是花生黄曲霉毒素污染的主要原因, 也是储藏期污染的源头。花生黄曲霉毒素污染防控是一个世界性难题, 由于我国花生品种繁多, 土壤类型和气候类型多样, 因此了解花生黄曲霉毒素污染时空消长变化趋势, 解析品种抗性、土壤类型、真菌产毒力及气候条件对其影响, 开展花生黄曲霉毒素污染预警模型研究, 构建农产品真菌毒素污染预警系统与黄曲霉毒素污染早期甄别技术, 变事后处理为事前预防, 对花生黄曲霉毒素污染科学控制具有重要意义。

本文基于连续十年花生主产区的黄曲霉毒素污染数据和产毒真菌分布以及收集权威机构发布的品种信息、产地土壤数据与气候数据, 探明了我国花生黄曲霉毒素污染消长变化及其与影响因素的关联, 建立了适于中国花生黄曲霉毒素污染状况的规则分类—平衡取样—随机森林预警模型, 实现了对黄曲霉毒素污染的预测, 为从源头开展花生黄曲霉毒素风险治理提供关键技术支撑。

主要研究方法与研究结果如下:

1. 明确了我国花生 AFB₁ 的空间分布特征, 绘制出中国花生 AFB₁ 的污染分布地图。主要结果如下:

(1) 采用地统计分析对中国花生污染的 AFB₁ 进行时空分析, 获得了中国花生黄曲霉毒素的污染模式, 表明中国花生 AFB₁ 呈现显著的时间和空间差异;

(2) 利用 GS+10.0 软件获得全方位的半变异函数, 表明花生 AFB₁ 含量具有空间自相关, 2009-2017 年花生 AFB₁ 半变异模型为球状模型;

(3) 利用普通克里金插值获得未取样地区的待测值, 生成中国花生 AFB₁ 的污染分布地图, 表明我国花生 AFB₁ 污染高风险区域呈分散分布, 与这些地区花生品种、土壤类型、菌种产毒力、气候条件等因素相关。

2. 分析了 2009-2017 连续九年我国花生黄曲霉毒素污染消长变化, 并阐述了花生品种抗性、土壤类型、菌株产毒力及气候条件对其产生的影响。主要结果如下:

(1) 黄曲霉毒素污染存在年度间差异, 2017、2014、2015 和 2013 年黄曲霉毒素污染水平高于其它年份, 总量污染水平分别为 $26.67 \pm 116.53 \mu\text{g}/\text{kg}$ 、 $17.49 \pm 72.01 \mu\text{g}/\text{kg}$ 、 $15.33 \pm 87.19 \mu\text{g}/\text{kg}$ 和 $13.95 \pm 52.65 \mu\text{g}/\text{kg}$ 。2009 年和 2010 年黄曲霉毒素污染水平最低, 总量污染水平分别为 $2.40 \pm 32.48 \mu\text{g}/\text{kg}$ 和 $4.32 \pm 35.15 \mu\text{g}/\text{kg}$ 。

(2) 黄曲霉毒素污染存在地区间差异, 长江流域主产区的花生黄曲霉毒素污染水平最高, 不同省份总量污染范围为 $5.59-37.15 \mu\text{g}/\text{kg}$ 。其次是南方主产区(总量污染范围为 $4.23-19.23 \mu\text{g}/\text{kg}$) 和北方主产区(总量污染范围为 $1.78-9.11 \mu\text{g}/\text{kg}$)。东北主产区的污染水平最低, 黄曲霉毒素总量不超过 $1.00 \mu\text{g}/\text{kg}$ 。

(3) 黄曲霉毒素低抗品种、砂土土质、高产毒力菌株在超标样品中占比较大, 高温干旱区域与黄曲霉毒素污染严重区域重合, 验证了各影响因素对花生黄曲霉毒素污染消长变化的影响,

为建立收获前花生黄曲霉毒素污染大尺度宏观预警模型奠定基础。

3. 建立了适合中国的规则分类—平衡取样—随机森林花生黄曲霉毒素预警模型, 并对模型进行了验证。主要结果如下:

(1) 确定了中国花生黄曲霉毒素预警模型的主要气候变量, 即纬度、8-20 时降水量、平均气压和日平均气温。利用各气候参数与 AFB₁ 之间的最大互信息系数值排序, 初选得到 8 个变量: 纬度、平均地表气温、日最低地表气温、8-20 时降水量、平均气压、日最低气压、日平均气温和日最高气温; 综合考虑初选变量间的多重共线性与全面性, 本研究最终确定建模气候变量为: 纬度、8-20 时降水量、平均气压和日平均气温。

(2) 建立了模型分类规则, 即高纬度区域 ($\geq 40^{\circ}\text{N}$)、低纬度区域 ($\leq 21^{\circ}\text{N}$) 与低温区域 (日平均气温 $\leq 16.8^{\circ}\text{C}$) 适合制定规则。通过分析主要气候变量与总样本及 AFB₁ 污染样本分布的关系, 可知高纬度区域 ($\geq 40^{\circ}\text{N}$)、低纬度区域 ($\leq 21^{\circ}\text{N}$) 与低温区域 (日平均气温 $\leq 16.8^{\circ}\text{C}$) 没有超限样本分布, 即为了提高模型精度可将此区域样品直接认定为未超限样品。

(3) 建立了平衡取样—随机森林模型, 对花生 AFB₁ 是否超限预测精度达 95% 以上。以筛选得到的 4 个主要气候因子与花生品种抗性、土壤类型及真菌产毒力为模型输入参数, 利用平衡取样—随机森林方法进行分类预测。结果表明当分类阈值为 20 $\mu\text{g}/\text{kg}$ 和 5 $\mu\text{g}/\text{kg}$ 时, 未超限样本预测准确率分别为 98.94% 和 96.97%, 超限样本分类准确率均为 100%。同其他常用分类器(决策树、支持向量机、K-近邻法、BP 神经网络、径向基函数神经网络) 相比较, 随机森林能有效分离超限与未超限样品, 具有较高的预测精度。

(4) 明确了模型输入参数对花生 AFB₁ 分类预测的贡献率排序。进一步分析输入参数在分类模型中的贡献率排序, 结果表明当以 20 $\mu\text{g}/\text{kg}$ 和 5 $\mu\text{g}/\text{kg}$ 为阈值时, 7 个输入参数在随机森林分类模型中的贡献率排序依次为: 纬度 > 日平均气温 > 8-20 时降水量 > 平均气压 > 品种抗性 > 土壤类型 > 真菌产毒力。

(5) 初步验证了模型预测效果, 其预测准确率为 86.89%。利用建立的规则分类—平衡取样—随机森林预警模型对 2018 年收获的花生 AFB₁ 含量进行预测, 结果表明综合规则分类样本, 当分类阈值为 20 $\mu\text{g}/\text{kg}$ 和 5 $\mu\text{g}/\text{kg}$ 时, 样本分类准确率均为 86.89%。

关键词: 花生, 黄曲霉毒素污染, 消长变化, 空间分析, 预警模型

Abstract

Peanut (*Arachis hypogaea* L.) is not only the main oil crops and agricultural products with export superiority in China, but also a kind of food with high nutritional value and therapeutic function. It is an economic crop can be used for both direct consumption and oil extraction. It occupies an important position in peanut production in China and the world. However, peanuts are very vulnerable to fungal infection and aflatoxin contamination, which poses a threat to human health and has become a major constraint factor affecting peanut consumption safety and world trade in China. Aflatoxin contamination can occur in the whole process such as planting, harvesting, storage, transportation and marketing. The contamination happened in the field is not only the main cause of aflatoxin contamination of peanut, but also the source of contamination during storage. Aflatoxin contamination control is a worldwide problem. There are many kinds of peanut varieties, soil types and weather patterns in China. Therefore, it is of great significance to understand the variation trend of aflatoxin contamination in peanut and analyze the effects of variety resistance, soil types, aflatoxigenic potential of fungi and climatic conditions on aflatoxin contamination. Thus, it is significant for scientific control of aflatoxin contamination in peanuts to study on risk prediction model of aflatoxin contamination, develop the early warning technology, and then adopt precise and efficient methods to make post-processing as pre-prevention.

Based on the data of aflatoxin content in peanut and distribution of aflatoxigenic fungi of main peanut production areas, as well as peanut variety information, soil information and climate data released by authoritative agencies during 10 consecutive years from 2009 to 2018, the variation trend of aflatoxin contamination and its relationship with influence factors were analyzed. A rule-based pre-classification combination with balanced sampling and random forest model suitable for China was established, which can be used for prediction of aflatoxin content in peanuts. It provided crucial technical support for risk management of aflatoxin contamination in peanut.

The main research methods and results are summarized as follows:

1. The spatial distribution characteristics of AFB₁ in peanut were evaluated and AFB₁ contamination map was generated. The main results were as follows:

(1) The distribution pattern of aflatoxin contamination in peanuts from China was carried out by geostatistical analysis. It showed that there were significant temporal and spatial differences.

(2) The omnidirectional semivariogram was obtained by using GS+10.0 software. It indicated that the content of AFB₁ in peanut had spatial autocorrelation, and the semi-variogram of AFB₁ contamination from 2009 to 2017 was spherical model.

(3) The contamination map of AFB₁ was generated after the AFB₁ values in unsampled areas were obtained by ordinary kriging interpolation. It revealed that AFB₁ contamination presented scattered distribution. It may relate with the different varieties, soil types, aflatoxin producing potential of fungi and climatic conditions in these areas.

2. The trend of aflatoxin contamination of peanut in China during the nine consecutive years from 2009 to 2017 was analyzed. The effects of resistance of different peanut varieties, soil types, aflatoxin producing potential of fungi and climatic conditions on the variation of aflatoxin contamination were discussed. The main results were as follows:

(1) Aflatoxin contamination had annual difference. Aflatoxin contamination was heavier in 2017, 2014, 2015 and 2013 than that in other years, with total aflatoxins of $26.67 \pm 116.53 \mu\text{g/kg}$, $17.49 \pm 72.01 \mu\text{g/kg}$, $15.33 \pm 87.19 \mu\text{g/kg}$ and $13.95 \pm 52.65 \mu\text{g/kg}$. It was the lowest in 2009 ($2.40 \pm 32.48 \mu\text{g/kg}$) and 2010 ($4.32 \pm 35.15 \mu\text{g/kg}$).

(2) Aflatoxin contamination had regional differences. The contamination level of aflatoxin in peanut was the highest in the main producing areas of the Yangtze River Basin, with total aflatoxins range of $5.59\text{-}37.15 \mu\text{g/kg}$, followed by the southern areas ($4.23\text{-}19.23 \mu\text{g/kg}$) and the northern areas ($1.78\text{-}9.11 \mu\text{g/kg}$). It was the lowest in the northeastern areas with total aflatoxins of no more than $1.00 \mu\text{g/kg}$.

(3) Aflatoxin susceptible varieties, sandy soil, high aflatoxin producing potential strains, high temperature had higher proportion in non-compliant samples than compliant samples. High temperature regions were similar with high aflatoxin contamination level regions. The effects of influencing factors on the variation trend of aflatoxin in peanut were verified. It laid a foundation for the establishment of large-scale early warning technology of aflatoxin contamination in peanut.

3. A peanut aflatoxin prediction model suitable for China based on rule classifier, balanced sampling and random forest was established. And then it was verified by data of 2018. The main results were as follows:

(1) The main climatic variables finally were selected. According to the maximum information coefficient between climatic variables and aflatoxin, as well as considering the multi-collinearity and the comprehensiveness of the factors, the main climatic variables finally selected were latitude, precipitation from 8:00 to 20:00, average air pressure and average air temperature.

(2) The rules for preclassification were made. By analyzing the relationship between the distribution of the main climatic variables and aflatoxin content, it can be seen that the effect of rule-making was remarkable in high latitude region ($> 40^\circ \text{N}$), low latitude ($< 21^\circ \text{N}$) and low temperature region (average air temperature $\leq 16.8^\circ \text{C}$). That was to say, in order to improve the accuracy of the model, the samples in these regions can be directly identified as non-compliant samples.

(3) The results of the balanced sampling-Random forest model showed that the prediction accuracy was more than 95%. Four main climatic parameters, resistance of peanut varieties, soil types and aflatoxigenic potential of fungi were identified as the input parameters of the model. Balanced sampling-Random forest was used to classify and predict the aflatoxin contamination risk. When the classification threshold was $20 \mu\text{g/kg}$ and $5 \mu\text{g/kg}$, the cross-validation results showed that the prediction accuracy of samples with AFB_1 content below threshold was 98.94% and 96.97% respectively, and the classification accuracy of samples with AFB_1 content below threshold was 100%. Compared with other commonly used classifiers (Decision tree, Support vector machine, K-nearest

neighbor, Back propagation network and Radial basis function neural network), Random forest can effectively separated the samples whether containing AFB₁ content more than limits (20 µg/kg and 5 µg/kg).

(4) Importance of the input parameters was analyzed. The importance of the input parameters for the classification model was further analyzed. The results indicated that when the threshold values were 20 µg/kg and 5µg/kg, importance of the 8 input parameters in descending order was latitude > mean daily temperature > precipitation from 8:00 to 20:00 > mean air pressure > peanut varieties > soil types > aflatoxin producing potential of fungi.

(5) The model validation results showed that the prediction accuracy was 86.89%. The mean content of AFB₁ in peanuts collected in 2018 was predicted by the rule classifier-balanced sampling-random forest classification model. Considering pre-classified samples, when the classification threshold was 20 µg/kg and 5 µg/kg, the results showed that the prediction accuracy of all the samples was 86.89%.

Key words: Peanut, Aflatoxin contamination, Variation trend, Spatial analysis, Prediction model

目 录

| | |
|---------------------------------|----|
| 第一章 引 言 | 1 |
| 1.1 花生产业的重要性 | 1 |
| 1.1.1 花生的营养价值 | 1 |
| 1.1.2 花生的经济价值 | 2 |
| 1.2 黄曲霉毒素对花生产业的影响 | 3 |
| 1.2.1 黄曲霉毒素性质稳定 | 3 |
| 1.2.2 黄曲霉毒素危害大 | 5 |
| 1.3 花生黄曲霉毒素污染主要影响因素 | 7 |
| 1.3.1 收获前影响花生黄曲霉毒素污染的主要因素 | 7 |
| 1.3.2 收获期影响花生黄曲霉毒素污染的主要因素 | 9 |
| 1.3.3 收获后影响花生黄曲霉毒素污染的主要因素 | 9 |
| 1.4 花生黄曲霉毒素污染防控研究进展 | 10 |
| 1.4.1 物理防控 | 10 |
| 1.4.2 化学防控 | 10 |
| 1.4.3 生物防控 | 11 |
| 1.4.4 全程防控 | 12 |
| 1.4.5 预警技术 | 13 |
| 1.5 本课题研究目的与意义 | 16 |
| 1.6 本课题主要研究内容 | 16 |
| 1.7 本课题技术路线 | 17 |
| 第二章 中国花生黄曲霉毒素污染空间分布 | 18 |
| 2.1 引言 | 18 |
| 2.2 材料 | 19 |
| 2.2.1 花生样品 | 19 |
| 2.2.2 仪器与试剂 | 20 |
| 2.3 方法 | 20 |
| 2.3.1 花生黄曲霉毒素检测方法 | 20 |

| | |
|--|----|
| 2.3.2 黄曲霉毒素空间变异性分析方法 | 21 |
| 2.3.3 结果统计方法 | 23 |
| 2.4 结果与分析 | 23 |
| 2.4.1 中国花生 AFB ₁ 空间变异特征 | 23 |
| 2.4.2 中国花生 AFB ₁ 污染分布地图 | 26 |
| 2.5 讨论 | 26 |
| 2.6 本章小结 | 27 |
| 第三章 中国花生黄曲霉毒素污染消长变化趋势 | 28 |
| 3.1 引言 | 28 |
| 3.2 材料 | 29 |
| 3.2.1 花生样品 | 29 |
| 3.2.2 仪器与试剂 | 29 |
| 3.2.3 花生品种抗性数据 | 29 |
| 3.2.4 土壤类型数据 | 29 |
| 3.2.5 菌株产毒力数据 | 29 |
| 3.2.6 气候数据 | 29 |
| 3.3 方法 | 30 |
| 3.3.1 花生黄曲霉毒素检测方法 | 30 |
| 3.3.2 结果统计方法 | 30 |
| 3.3.3 数据处理方法 | 30 |
| 3.4 结果与分析 | 30 |
| 3.4.1 中国花生黄曲霉毒素污染动态变化 | 30 |
| 3.4.2 影响黄曲霉毒素污染动态变化的主要因素 | 31 |
| 3.5 讨论 | 37 |
| 3.6 本章小结 | 38 |
| 第四章 中国花生黄曲霉毒素 B ₁ 污染预警模型的构建 | 39 |
| 4.1 引言 | 39 |
| 4.2 材料 | 40 |

| | |
|---------------------------------|----|
| 4.2.1 属性变量 | 40 |
| 4.2.2 花生样品 | 40 |
| 4.3 方法 | 40 |
| 4.3.1 多重共线性诊断 | 40 |
| 4.3.2 相关性分析 | 41 |
| 4.3.3 建模方法 | 41 |
| 4.4 方法及模型 | 41 |
| 4.4.1 规则分类器 | 41 |
| 4.4.2 最大互信息系数 | 42 |
| 4.4.3 基于机器学习的分类方法 | 43 |
| 4.5 结果与分析 | 47 |
| 4.5.1 地理与气候变量筛选 | 47 |
| 4.5.2 规则分类—平衡取样—随机森林模型的建立 | 53 |
| 4.5.3 随机森林分类模型与其它分类模型精度比较 | 65 |
| 4.5.4 模型验证 | 67 |
| 4.6 讨论 | 67 |
| 4.7 本章小结 | 69 |
| 第五章 全文结论 | 71 |
| 参考文献 | 74 |
| 致 谢 | 89 |
| 作者简历 | 90 |

图目录

| | |
|--|----|
| 图 1.1 2016 年中国五大油料作物产量 | 2 |
| 图 1.2 2017 年世界花生产量（排名前 10） | 3 |
| 图 1.3 黄曲霉毒素（ AFB_1 、 AFB_2 、 AFG_1 、 AFG_2 、 AFM_1 和 AFM_2 ）的化学结构 | 4 |
| 图 2.1 2009-2017 年中国花生取样省和取样市县 | 19 |
| 图 2.2 球状模型函数图 | 22 |
| 图 2.3 2009-2017 年花生 AFB_1 含量平均值的变异函数 | 25 |
| 图 2.4 2009-2017 年中国花生 AFB_1 平均值的克里金插值图 | 26 |
| 图 3.1 2009-2017 年中国花生主产省 AFB_1 消长变化趋势 | 31 |
| 图 3.2 AFB_1 超标与未超标样品的不同抗性品种占比 | 32 |
| 图 3.3 AFB_1 超标与未超标样品的不同土壤类型占比 | 32 |
| 图 3.4 AFB_1 超标与未超标样品的不同产毒力菌株占比 | 33 |
| 图 3.5 不同时期平均气候条件与 AFB_1 含量拟合的决定系数 | 34 |
| 图 3.6 2009-2017 年中国花生主产区降水量（收获前 30 天平均值）克里金插值图 | 35 |
| 图 3.7 2009-2017 年中国花生主产区日平均气温（收获前 30 天平均值）克里金插值图 | 35 |
| 图 3.8 (a) 三大主产区花生 AFB_1 含量与降水量（收获前 1 个月平均值）的关系； (b) 三大主产区花生 AFB_1 含量与日平均气温的关系；(c) 三大主产区花生 AFB_1 含量与日最低气温（收获前 1 个月平均值）的关系；(d) 三大主产区花生 AFB_1 含量与日最高气温（收获前 1 个月平均值）的关系 | 36 |
| 图 4.1 决策树典型结构 | 44 |
| 图 4.2 BPNN 结构图 | 45 |
| 图 4.3 随机森林算法流程图 | 47 |
| 图 4.4 2009-2017 年花生 AFB_1 平均含量和气候因子矩阵散点图 | 48 |
| 图 4.5 AFB_1 分布直方图 | 51 |
| 图 4.6 花生 AFB_1 含量在不同纬度的散点分布 | 54 |
| 图 4.7 不同纬度区间样品占比 | 54 |

| | |
|--|----|
| 图 4.8 以 5 $\mu\text{g}/\text{kg}$ 和 20 $\mu\text{g}/\text{kg}$ 为阈值, 纬度区间对应超限样品的比例..... | 55 |
| 图 4.9 花生 AFB ₁ 含量在不同降水量 (8-20 时) 的散点分布..... | 55 |
| 图 4.10 不同 8-20 时降水量区间样品占比..... | 56 |
| 图 4.11 以 5 $\mu\text{g}/\text{kg}$ 和 20 $\mu\text{g}/\text{kg}$ 为阈值, 8-20 时降水量区间对应超限样品的比例..... | 56 |
| 图 4.12 花生 AFB ₁ 含量在不同平均气压的散点分布..... | 57 |
| 图 4.13 不同平均气压区间样品占比..... | 57 |
| 图 4.14 以 5 $\mu\text{g}/\text{kg}$ 和 20 $\mu\text{g}/\text{kg}$ 为阈值, 平均气压区间对应超限样品的比例..... | 58 |
| 图 4.15 花生 AFB ₁ 含量在不同日平均气温的散点分布..... | 59 |
| 图 4.16 不同日平均气温区间样品占比..... | 59 |
| 图 4.17 以 5 $\mu\text{g}/\text{kg}$ 和 20 $\mu\text{g}/\text{kg}$ 为阈值, 日平均气温区间对应超限样品的比例.. | 60 |
| 图 4.18 基于纬度和日平均气温变量的样本分布..... | 61 |
| 图 4.19 基于纬度和 8-20 时降水量变量的样本分布..... | 61 |
| 图 4.20 基于纬度和平均气压变量的样本分布..... | 62 |
| 图 4.21 基于随机森林分类模型的变量贡献率排序 (以 20 $\mu\text{g}/\text{kg}$ 为阈值对花生 AFB ₁ 含量进行分类)..... | 64 |
| 图 4.22 基于随机森林分类模型的变量贡献率排序 (以 5 $\mu\text{g}/\text{kg}$ 为阈值对花生 AFB ₁ 含量进行分类)..... | 64 |
| 图 4.23 2009-2017 年中国花生污染黄曲霉毒素污染类型占比..... | 69 |

表目录

| | |
|---|----|
| 表 1.1 黄曲霉毒素的熔点和紫外吸收 | 4 |
| 表 1.2 世界各国和地区农产品中黄曲霉毒素的最大限量 | 5 |
| 表 1.3 2007-2017 年中国出口产品遭 RASFF 预警通报批次..... | 6 |
| 表 2.1 2009-2017 年花生 AFB ₁ 平均含量的变异函数理论模型参数 | 24 |
| 表 4.1 气候因子多重共线性诊断表 | 49 |
| 表 4.2 气候因子多重共线性诊断—方差比例表 | 49 |
| 表 4.3 气候因子相关系数表 | 50 |
| 表 4.4 利用不同的线性相关分析与非线性相关分析系数比较 | 52 |
| 表 4.5 不同变量之间的线性相关分析 | 53 |
| 表 4.6 当 AFB ₁ 含量阈值 20 μg/kg 时, 训练样本的混淆矩阵 | 63 |
| 表 4.7 当 AFB ₁ 含量阈值 20 μg/kg 时, 测试样本的混淆矩阵 | 63 |
| 表 4.8 当 AFB ₁ 含量阈值 5 μg/kg 时, 训练样本的混淆矩阵 | 63 |
| 表 4.9 当 AFB ₁ 含量阈值 5 μg/kg 时, 测试样本的混淆矩阵 | 63 |
| 表 4.10 当 AFB ₁ 含量阈值为 20 μg/kg 时, 随机森林分类模型与其它分类模型精度 比较 | 66 |
| 表 4.11 当 AFB ₁ 含量阈值为 5 μg/kg 时, 随机森林分类模型与其它分类模型精度 比较 | 66 |
| 表 4.12 当 AFB ₁ 含量阈值 20 μg/kg 时, 验证样本的混淆矩阵 | 67 |
| 表 4.13 当 AFB ₁ 含量阈值 5 μg/kg 时, 验证样本的混淆矩阵 | 67 |
| 表 4.14 中国不同花生产区土壤中黄曲霉的菌落数与产毒力水平 | 69 |

英文缩略表

| 英文缩写 | 英文全称 | 中文名称 |
|-------------------|---|-----------------------|
| AFT | Aflatoxin | 黄曲霉毒素 |
| AFB ₁ | Aflatoxin B ₁ | 黄曲霉毒素 B ₁ |
| AFB ₂ | Aflatoxin B ₂ | 黄曲霉毒素 B ₂ |
| AFG ₁ | Aflatoxin G ₁ | 黄曲霉毒素 G ₁ |
| AFG ₂ | Aflatoxin G ₂ | 黄曲霉毒素 G ₂ |
| AFM ₁ | Aflatoxin M ₁ | 黄曲霉毒素 M ₁ |
| AFM ₂ | Aflatoxin M ₂ | 黄曲霉毒素 M ₂ |
| AFP ₁ | Aflatoxin P ₁ | 黄曲霉毒素 P ₁ |
| AFQ ₁ | Aflatoxin Q ₁ | 黄曲霉毒素 Q ₁ |
| AFG _{2a} | Aflatoxin G _{2a} | 黄曲霉毒素 G _{2a} |
| FAO | Food and agriculture organization | 世界粮农组织 |
| MRL | Maximum residue limit | 最大残留限量 |
| RASFF | Rapid Alert system for food and feed | 欧盟食品和饲料类快速预警系统 |
| CCCF | Codex committee on contaminants in foods | 食品污染物法典委员会 |
| EU | European Union | 欧盟 |
| US | the United States | 美国 |
| FDA | Food and drug administration | 食品药品监督管理局 |
| CAC | Codex alimentarius commission | 食品法典委员会 |
| WHO | World health organization | 世界卫生组织 |
| IPCC | Intergovernmental panel on climate change | 联合国政府间气候变化专门委员会 |
| a _w | Water activity | 水分活度 |
| GAP | Good agricultural practice | 良好农业规范 |
| GMP | Good manufacturing practice | 良好操作规范 |
| HACCP | Hazard analysis and critical control point | 危害分析与关键控制点 |
| R ² | Coefficient of determination | 决定系数 |
| GA/BPN | Genetical gorithm/back propagation neural network | 遗传算法/反相传播神经网络 |
| APSIM | Agricultural production systems simulator | 作物生长模拟方法 |
| ARID | The agricultural reference index for drought | 农业干旱参考指数 |
| NDVI | Normalized difference vegetation index | 归一化植被指数 |
| VA | Versicolorin A | 杂色曲霉素 A |

| 英文缩写 | 英文全称 (续) | 中文名称 |
|-----------------|--|---------------------|
| ST | Sterigmatocystin | 杂色曲霉素 |
| GIS | Geographic information system | 地理信息系统 |
| HPLC | High performance liquid chromatography | 高效液相色谱 |
| IAC | Immunoaffinity chromatography | 免疫亲和层析 |
| FLD | Fluorescence detection | 荧光检测 |
| LOD | Limit of detection | 检出限 |
| RSS | Residual sum of squares | 残差平方和 |
| DON | Deoxynivalenol | 脱氧雪腐镰刀烯醇 |
| OTA | Ochratoxin A | 赭曲霉毒素 A |
| FB ₁ | Fumonisin B ₁ | 伏马毒素 B ₁ |
| FB ₂ | Fumonisin B ₂ | 伏马毒素 B ₂ |
| SD | Standard deviation | 标准差 |
| LAT | Latitude | 纬度 |
| LON | Longitude | 经度 |
| ALT | Altitude | 海拔高度 |
| EVA | Evaporation | 蒸发量 |
| GSTmean | Mean soil surface temperature | 平均地表气温 |
| GSTmax | Maximum soil surface temperature | 日最高地表气温 |
| GSTmin | Minimum soil surface temperature | 日最低地表气温 |
| PRE20-8 | Precipitation from 20:00 to 8:00 | 20-8 时降水量 |
| PRE8-20 | Precipitation from 8:00 to 20:00 | 8-20 时降水量 |
| PRE20-20 | Precipitation from 20:00 to 20:00 | 20-20 时降水量 |
| WIN | Mean wind speed | 平均风速 |
| WINmax | Maximum wind speed | 最大风速 |
| WINext | Extreme wind speed | 极大风速 |
| SSD | Sunshine duration | 日照时数 |
| PRSmear | Mean air pressure | 平均气压 |
| PRSmear | Maximum air pressure | 日最高气压 |
| PRSmear | Minimum air pressure | 日最低气压 |
| TEMmean | Mean air temperature | 日平均气温 |
| TEMmax | Maximum air temperature | 日最高气温 |
| TEMmin | Minimum air temperature | 日最低气温 |

| 英文缩写 | 英文全称 (续) | 中文名称 |
|---------|---|----------------|
| RHUmean | Mean relative humidity | 平均相对湿度 |
| RHUmin | Minimum relative humidity | 最小相对湿度 |
| MIC | Maximal information coefficient | 最大互信息系数 |
| MINE | Maximal information-based nonparametric exploration | 最大的基于信息的非参数性探索 |
| DT | Decision tree | 决策树 |
| SVM | Support vector machine | 支持向量机 |
| KNN | K-nearest neighbor | k-近邻法 |
| ANN | Artificial neural network | 人工神经网络 |
| BPNN | Back propagation network | BP 神经网络 |
| RBFNN | Radial basis function neural network | 径向基函数神经网络 |
| RF | Random forest | 随机森林 |
| OOB | Out-of-bag error | 袋外错误率 |
| CFU | Colony forming units | 菌落数 |
| μg | Microgram | 微克 |
| mg | Milligram | 毫克 |
| g | Gram | 克 |
| kg | Kilogram | 千克 |
| μL | Microliter | 微升 |
| mL | Milliliter | 毫升 |
| L | Litre | 升 |
| μm | Micron | 微米 |
| nm | Nanometer | 纳米 |
| mm | Millimeter | 毫米 |
| cm | Centimeter | 厘米 |
| m | Meter | 米 |
| s | Second | 秒 |
| min | Minute | 分钟 |
| h | Hour | 小时 |
| d | Day | 天 |
| μmol | Micromole | 微摩尔 |
| mol | Mole | 摩尔 |

第一章 引言

1.1 花生产业的重要性

1.1.1 花生的营养价值

花生 (*Arachis hypogaea* L.) 营养价值高而且兼具食疗功能, 能够改善微循环, 被称为养血的第一食物。此外, 花生抗衰老能力极强, 可以保持皮肤滋润、细嫩, 因此也被誉为“长生果”, 是广受人们喜爱的坚果之一 (于琬淇 2017)。它富含蛋白质 (26-39%)、脂肪 (47-59%)、碳水化合物 (11-27%) 和纤维素 (Jabeen 2012)。花生仁中含有人体必需的 8 种氨基酸, 且比例适宜, 消化吸收率高达 90%。其中赖氨酸具有提高儿童智力的作用, 谷氨酸和天门冬氨酸能够促进细胞发育并增强大脑记忆力, 赖氨酸也是防止衰老的重要成分。在花生所含的脂肪中, 绝大部分为不饱和脂肪酸, 其中亚油酸的含量非常丰富, 而其又是人体最为重要的一种必需脂肪酸。花生四烯酸具有降低血脂和血清胆固醇的功能, 可以减少冠心病的发生。花生中还含有丰富的胆碱、卵磷脂等物质, 有助于增强记忆、延缓衰老和滋润皮肤。碳水化合物中淀粉约占 4%, 其余是游离糖, 分为可溶性和不溶性。花生纤维组织中的可溶性纤维在人体消化吸收的过程中会吸取某些毒素, 从而减少有害物质在体内的累积并减少毒性作用的发生, 进而降低肠癌的发生率 (赵贵兴 等 2011)。

除了油脂、蛋白质、碳水化合物和膳食纤维等, 花生同时富含多种微量营养素, 如矿物质: 钠 (42.0 mg/100 g)、钾 (705.11 mg/100 g)、镁 (3.98 mg/100 g)、钙 (2.28 mg/100 g)、铁 (6.97 mg/100 g)、锌 (3.2 mg/100 g)、磷 (10.55 mg/100 g) 和硼, 以及维生素 B₂、PP、A、C、D、E、K 等 (Nyirahakizimana *et al.* 2013)。花生含有的维生素 C 有降低胆固醇的作用, 可以起到防治动脉硬化、高血压和冠心病的作用。花生中的维生素 K 具有止血作用。某些花生品种还富含硒元素, 其营养丰富, 口味鲜美, 色泽诱人, 常食可调节人体生理功能、提高血红蛋白含量、护肤美容、明眼乌发、益智长寿等 (潘晓红 等 2010)。花生也富含植物固醇, 包括白藜芦醇、 β -谷固醇和植物异黄酮等, 其中白藜芦醇是一种生物活性很高的多酚类物质, 具有降血脂、抗氧化、抗肿瘤、抗癌以及防治心血管疾病等保健功能 (Sanders *et al.* 2000)。

花生的内皮含有抗纤维蛋白溶解酶, 可防治各种外伤出血、肝病出血、血友病等。而花生仁外层的红皮可以促进血小板的生成, 避免血小板聚集, 预防心血管疾病。花生红衣的止血作用比花生高出 50 倍, 对多种出血性疾病都会产生良好的止血功效。由于其高营养价值, 花生还具有多种其它用途, 比如可用作食疗性产品 (Nabuuma *et al.* 2013)、糖果配料 (Ajay *et al.* 2012) 以及动物饲料 (Grosso *et al.* 1995) 等。花生的仁、皮、壳、叶、茎、油均可入药, 可谓全身是宝。生食花生不但可防口唇干裂、烦渴、声哑失音、口腔糜烂、便秘、痔疮、肛裂、胃胀、胃酸过多、营养不良症, 而且对预防口角炎、干眼病、骨质增生、湿疥、皮炎、脚气病等维生素缺乏症效果更佳 (游克仁 2004)。花生粕是花生仁经压榨提炼油脂后的副产品, 含有丰富的植物蛋白, 其代谢能是粕类饲料中最高的, 粗蛋白含量 48% 以上, 精氨酸含量高达 5.2%, 是所有动植物饲料中最高的, 维生素及矿物质含量与其它饼粕类相近似。此外, 花生茎叶也可作为优质饲料。

1.1.2 花生的经济价值

花生是中国主要的种植作物之一，已成为我国食用植物油和蛋白质的重要来源。在我国的五大油料作物中，花生的产量高于油菜籽、葵花籽、芝麻和胡麻籽，位于产量之首（图 1.1）。部分花生可用于鲜食与加工，最主要的用途是作为食用花生油的原材料，全世界大约 60% 的花生仁都用来榨油。2014 年世界粮农组织（Food and Agriculture Organization, FAO）数据显示，花生油是世界第六大植物油，仅次于棕榈油、大豆油、菜籽油、葵花籽油和棉籽油（FAO 2014）。我国在世界花生生产中占有重要地位，近二、三十年来发展迅速，我国花生总产量自 1993 年以来就稳居世界花生生产国和消费国之首，2017 年花生总产量达到 1715 万吨（图 1.2）。2016 年中国花生出口量为 11 万吨，位居世界第五，进口量为 22 万吨，位居世界第二。

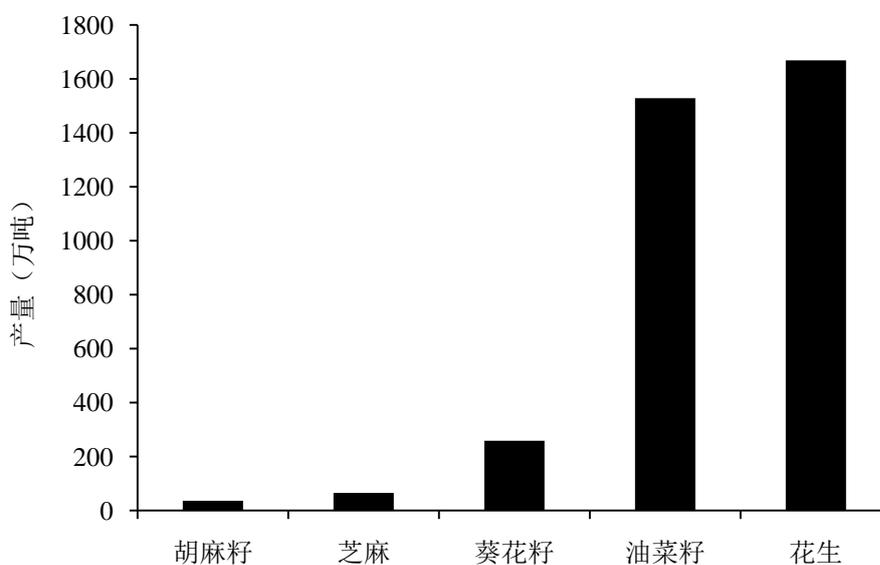


图 1.1 2016 年中国五大油料作物产量

Figure 1.1 Production of five main oil crops in China in 2016

数据来源：中华人民共和国国家统计局 <http://data.stats.gov.cn/workspace/index?m=hgnd>

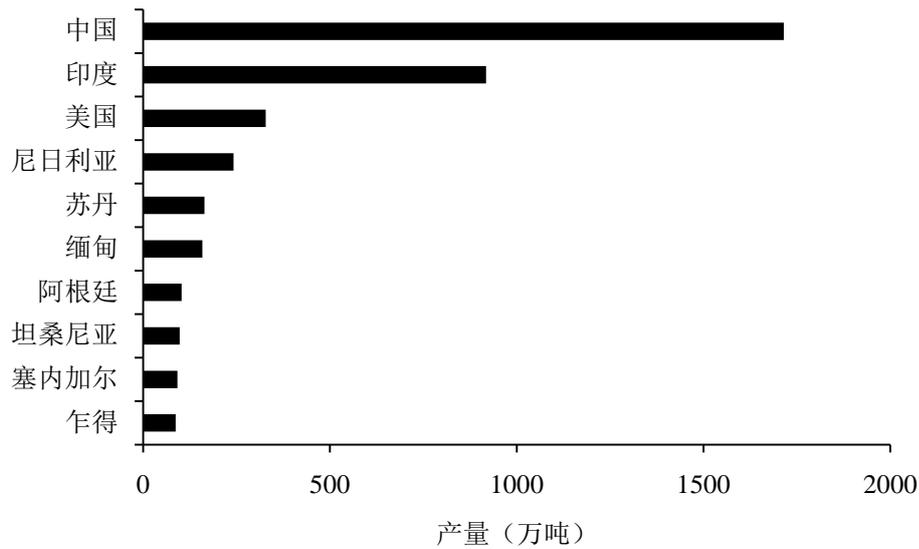


图 1.2 2017 年世界花生产量 (排名前 10)

Figure 1.2 Countries ranking the top ten in peanut production in 2017

数据来源: FAOSTAT <http://www.fao.org/faostat/en/#data/QC>

1.2 黄曲霉毒素对花生产业的影响

1.2.1 黄曲霉毒素性质稳定

作为一种地下结果的作物,花生易受真菌侵染并产生黄曲霉毒素(Aflatoxin, AFT)。黄曲霉毒素是一种由黄曲霉(*Aspergillus flavus*)、寄生曲霉(*A.parasticus*)和集峰曲霉(*A.nomius*)等真菌的产毒菌株所产生的一类次生代谢产物(Iqbal *et al.* 2012)。自从 20 世纪 60 年代英国十万只火鸡死亡事件发生以来,黄曲霉毒素已经引起了全世界的关注(Stoev 2013)。目前已经发现了近 20 种黄曲霉毒素,包括黄曲霉毒素 B₁(Aflatoxin B₁, AFB₁)、黄曲霉毒素 B₂(Aflatoxin B₂, AFB₂)、黄曲霉毒素 G₁(Aflatoxin G₁, AFG₁)、黄曲霉毒素 G₂(Aflatoxin G₂, AFG₂),以及它们的主要代谢物黄曲霉毒素 M₁(Aflatoxin M₁, AFM₁)、黄曲霉毒素 M₂(Aflatoxin M₂, AFM₂)、黄曲霉毒素 P₁(Aflatoxin P₁, AFP₁)、黄曲霉毒素 Q₁(Aflatoxin Q₁, AFQ₁) 和黄曲霉毒素 G_{2a}(Aflatoxin G_{2a}, AFG_{2a})等(Bara *et al.* 2010),它们是二氢呋喃氧杂萘邻酮的衍生物,含有一个双呋喃环和一个氧杂萘邻酮(香豆素)(图 1.3)。其中,AFB₁、AFB₂、AFG₁和 AFG₂是最常见的四种自然发生的黄曲霉毒素(Ifeanacho *et al.* 2017)。与其他杂环化合物类似,黄曲霉毒素具有荧光性,并且可以通过荧光性质进行区分,例如在紫外光照射下 AFB₁和 AFB₂呈蓝色荧光、AFG₁呈绿色荧光、AFG₂呈黄绿色荧光、AFM₁呈蓝紫色荧光。黄曲霉毒素的熔点与吸收光谱见表 1.1,其极微溶于水(10-30 μg/mL),不溶于非极性溶剂,可溶于中等极性有机溶剂(如氯仿和甲醇),极溶于二甲亚砜(Cole *et al.* 1981)。在有氧、酸碱性 pH (<3 或 >10)和氧化剂存在的条件下,紫外光照射会使黄曲霉毒素变得不稳定。其内酯环易受碱性水解,也可与氨或次氯酸钠反应降解反应(IARC, 2002)。

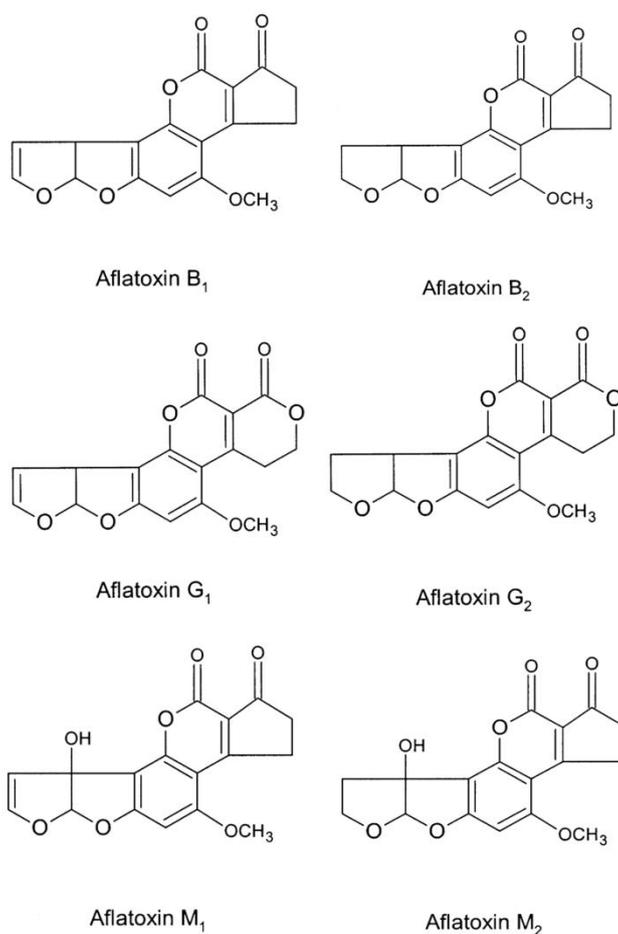
图 1.3 黄曲霉毒素 (AFB₁、AFB₂、AFG₁、AFG₂、AFM₁ 和 AFM₂) 的化学结构Figure 1.3 Chemical structures of aflatoxins (AFB₁, AFB₂, AFG₁, AFG₂, AFM₁ and AFM₂)

表 1.1 黄曲霉毒素的熔点和紫外吸收

Table 1.1 Melting-points and ultraviolet absorption of aflatoxins

| 黄曲霉毒素 | 分子式 | 分子量 | 熔点 (°C) | 紫外线吸收 (乙醇) | |
|----------------|--|-------|-------------------------------|-----------------------|--|
| | | | | λ_{\max} (nm) | ϵ (L mol ⁻¹ cm ⁻¹) |
| B ₁ | C ₁₇ H ₁₂ O ₆ | 312.3 | 268-269 (分解温度) (不溶于氯仿) | 223 | 25 600 |
| | | | | 265 | 13 400 |
| | | | | 362 | 21 800 |
| B ₂ | C ₁₇ H ₁₄ O ₆ | 314.3 | 286-289 (分解温度) (不溶于氯仿-正戊烷) | 265 | 11 700 |
| | | | | 363 | 23 400 |
| | | | | | |
| G ₁ | C ₁₇ H ₁₂ O ₇ | 328.3 | 244-246 (分解温度) (不溶于氯仿甲烷) | 243 | 11 500 |
| | | | | 257 | 9 900 |
| | | | | 264 | 10 000 |
| G ₂ | C ₁₇ H ₁₄ O ₇ | 330.3 | 237-240 (分解温度) (不溶于乙酸乙酯) | 362 | 16 100 |
| | | | | 265 | 9 700 |
| | | | | 363 | 21 000 |
| M ₁ | C ₁₇ H ₁₂ O ₇ | 328.3 | 299 (分解温度) (不溶于甲醇) | 226 | 23 100 |
| | | | | 265 | 11 600 |
| | | | | 357 | 19 000 |

注: 摘自 O'Neil *et al.* 2001

1.2.2 黄曲霉毒素危害大

根据暴露时间和暴露剂量的不同,黄曲霉毒素具有急慢性毒性、致突变性、致癌性和致畸性等 (Nesci *et al.* 2016), 并且具有免疫抑制性和基因毒性 (Theumer *et al.* 2018)。其中 AFB₁ 毒性最强,是氯化钾的 10 倍,砒霜的 68 倍,被世界卫生组织的癌症研究机构列为一级致癌物 (Huang *et al.* 2017)。此外,黄曲霉毒素可减少可溶性营养物质的运输 (Du *et al.* 2017),破坏蛋白质、碳水化合物和脂类的新陈代谢 (Rotimi *et al.* 2017),危害儿童成长 (Mupunga *et al.* 2017)。黄曲霉毒素污染多种农产品,包括谷物(玉米、稻米、小麦、大麦、高粱)、油料(花生、向日葵、棉籽)、香料(辣椒、黑胡椒、芫荽、姜黄、姜)和坚果(杏仁、椰子、巴西坚果、核桃、开心果)等。动物在饲用受黄曲霉毒素污染的饲料后,黄曲霉毒素也会在其奶、蛋和肉产品中检出。

黄曲霉毒素污染严重威胁人体健康,已造成巨大的经济损失。近年来,国内外发生过多起因花生产品中黄曲霉毒素超标而导致的人畜中毒死亡恶性事件。2010 年国家质检局质量监督抽查查出广东神符山花生油等 14 批次食用植物油黄曲霉毒素超标;2011 年因黄曲霉毒素等安全因素全年共吊销注销了 77 家企业的食用油生产许可证;2015 年 11 月 16 日,广西壮族自治区食品药品监督管理局对 80 批次食品进行抽检,其中 4 批次不合格产品均为 AFB₁ 超标引起,且不合格产品全部为花生油产品。据报道,美国食品因黄曲霉毒素和其他真菌毒素污染造成的损失达 5 亿美元至 15 亿美元;在中低收入国家,其直接和间接危害已花费巨额医疗费 (Robens *et al.* 2003)。由于黄曲霉毒素对人体健康的影响,多国已设立黄曲霉毒素的最大残留限量 (Maximum residue limit, MRL),范围多在 2-20 $\mu\text{g}/\text{kg}$ (表 1.2)。其中,欧盟的限量标准最为严格,规定花生黄曲霉毒素 B₁ 不超过 2 $\mu\text{g}/\text{kg}$,黄曲霉毒素 B₁、B₂、G₁ 和 G₂ 的总和不超过 4 $\mu\text{g}/\text{kg}$ 。在我国花生国际贸易中,因黄曲霉毒素超标遭到欧盟食品和饲料类快速预警系统 (Rapid Alert System for Food and Feed, RASFF) 预警通报的花生产品批次居高不下,2017 年占全部预警通报产品的 27.4%。因此,开展花生黄曲霉毒素污染防控研究,降低黄曲霉毒素污染风险,是保障农产品和食品消费安全,维护国家经济利益的迫切需求。

表 1.2 世界各国家和地区农产品中黄曲霉毒素的最大残留限量

Table 1.2 Maximum residue levels of aflatoxins in different agricultural products of different countries and regions

| 地区 | 产品种类 | MRL/ ($\mu\text{g}/\text{kg}$) | |
|----------|--|----------------------------------|---------|
| | | AFB ₁ | 黄曲霉毒素总量 |
| CCCF | 未加工的花生、杏仁、巴西坚果等 | - | 15 |
| | 可直接食用的巴西坚果、开心果、杏仁、榛子等 | | 20 |
| EU | 供直接食用或作为配料的花生、坚果、干果及其加工产品，除玉米和大米以外的谷物及谷物产品 | 2 | 4 |
| | 玉米和大米 | 5 | 10 |
| | 婴幼儿谷物食品 | 0.1 | - |
| | 调味品 | 5 | 10 |
| US (FDA) | 食品 (不包括牛奶) | - | 20 |
| 加拿大 | 坚果 | - | 15 |
| 土耳其 | 调味品 | 5 | 10 |
| 巴西 | 玉米和玉米制品 | - | 20 |
| 越南 | 供直接食用的食品 | 5 | 10 |
| 津巴布韦 | 供直接食用的食品 | 5 | - |
| 坦桑尼亚 | 未加工玉米 | 5 | 10 |
| 马来西亚 | 用于加工的花生 | - | 15 |
| 日本 | 所有食品 | - | 10 |
| CAC | 未加工花生 | - | 15 |
| 澳大利亚及新西兰 | 花生及其它坚果 | - | 15 |
| 中国 | 玉米和玉米制品，花生和花生制品，花生油，辐照花生 | 20 | - |
| | 大米、辐照大米、食用植物油 | 10 | - |
| | 小麦、大麦制品、酱油、豆类等 | 5 | - |
| WHO | | 5 | - |
| FAO/WHO | 未加工的花生和其它坚果 | - | 15 |
| | 可直接食用的坚果 (JECFA) | - | 10 |

CCCF: Codex Committee on Contaminants in Foods 食品污染物法典委员会; US: the United States 美国; FDA: Food and Drug Administration 食品药品监督管理局; EU: European Union 欧盟; CAC: Codex Alimentarius Commission 食品法典委员会; WHO: World Health Organization 世界卫生组织

表 1.3 2007-2017 年中国出口产品遭 RASFF 预警通报批次

Table 1.3 The bulletin batches of export products notified by RASFF during 2007-2017

| 年份 | 被通报产品总批次 | 被通报花生批次 | 因 AFT 被通报花生批次 | 因 AFT 被通报花生产品批次占被通报产品总批次的比例% | 因 AFT 被通报花生产品批次占被通报花生产品总批次的比例% |
|------|----------|---------|---------------|------------------------------|--------------------------------|
| 2007 | 314 | 57 | 37 | 11.78 | 64.91 |
| 2008 | 516 | 166 | 160 | 31.01 | 96.39 |
| 2009 | 317 | 60 | 40 | 12.62 | 66.67 |
| 2010 | 384 | 76 | 68 | 17.71 | 89.47 |
| 2011 | 594 | 76 | 63 | 10.61 | 82.89 |
| 2012 | 516 | 62 | 54 | 10.47 | 87.10 |
| 2013 | 367 | 63 | 46 | 12.53 | 73.02 |
| 2014 | 368 | 43 | 36 | 9.78 | 83.72 |
| 2015 | 401 | 124 | 104 | 25.94 | 83.87 |
| 2016 | 196 | 41 | 34 | 17.35 | 82.93 |
| 2017 | 263 | 79 | 72 | 27.38 | 91.14 |

数据来源: 食品伙伴网 <http://db.foodmate.net/rasff/index.php?cname=&lname=&chanpin=&ctime>

1.3 花生黄曲霉毒素污染主要影响因素

由于花生黄曲霉毒素污染受到收获前、收获时和收获后多环节诸多因素影响，黄曲霉毒素污染控制是个系统性工程。收获前影响黄曲霉毒素污染的因素主要包括花生品种、土壤类型、污染菌株类型、气候条件和农艺措施。合适的收获时间、花生的水分活度和成熟度是收获期的主要影响因素。收获后的储藏和运输条件，例如温湿度、包装、卫生条件等是影响黄曲霉毒素产生的主要条件。

1.3.1 收获前影响花生黄曲霉毒素污染的主要因素

1.3.1.1 花生品种

自 20 世纪 80 年代开始，多国研究者试图找到抗黄曲霉侵染和抗黄曲霉毒素产生的花生品种 (Azaizeh *et al.* 1989; Blankenship *et al.* 1985)。近年来，转基因技术已被广泛应用于品种改良。在对转化数据的初步分析中，含有 Bt (苏云金芽孢杆菌) 基因的转基因花生的黄曲霉毒素污染水平显著低于非 Bt 花生 (Ozias-Akins *et al.* 2002)。Guo 等 (2008) 鉴定了花生中抗寄生曲霉侵染和抗黄曲霉毒素产生的相关抗性基因 (iso ara h3 和 LEA)，并研发了花生微阵列用于鉴定抗黄曲霉侵染的候选基因 (Guo *et al.* 2011)。此外，抗病、抗虫或抗旱花生品种也能够显著降低真菌污染的发生率 (Mutegi *et al.* 2012)。中花 6 号是我国中部地区应用最广泛的优良品种，具有优良的抗黄曲霉毒素污染、抗青枯病、抗根腐病能力，其对黄曲霉毒素产生抗性的原因可能是曲霉菌高感染胁迫使白藜芦醇的含量升高 (Liao *et al.* 2010)。

1.3.1.2 土壤类型

花生可以在不同类型的土壤中生长，如轻质沙土、壤土和粘土等。轻砂质土壤有利于黄曲霉孢子的快速传播与增殖，特别是在生长后期遇到高温干旱的条件，花生受黄曲霉毒素污染的风险较高。相反，由于粘度较高的土壤具有较好的持水力，因此受到干旱胁迫的可能性较低，从而可以减少花生黄曲霉毒素的污染水平。严重缺钙的土壤会使花生果壳中果胶钙的积累受到影响，造成花生果壳组织变松，易产生烂果，从而造成黄曲霉菌侵染 (李新国 等 2011)。此外，黄曲霉毒素污染还会引起土壤毒化，损害土壤功能，致使土壤理化性质变坏，微生物的生命活动受限，肥力下降，土壤利用率降低，农作物生长发育不良，造成作物减产 (王海鸥 等 2015)。

1.3.1.3 土壤菌株类型

土壤中存在丰富的微生物菌群，而花生作为地下结果的作物，与土壤中黄曲霉毒素产毒真菌直接接触，从而容易受到黄曲霉毒素的污染。在花生根际土壤分离的真菌中，最常见的有曲霉属 (*Aspergillus*)、青霉属 (*Penicillium*)、酒曲菌属 (*Rhizopus*) 及镰刀菌属 (*Fusarium*) (Gachomo *et al.* 2004; Youssef *et al.* 2008)。黄曲霉 (*Aspergillus flavus*) 和寄生曲霉 (*A. parasiticus*) 是侵染花生并产生黄曲霉毒素 B₁、B₂、G₁、G₂ 的最常见曲霉属真菌 (Dorner 2002a; Vaamonde *et al.* 2003; Mutegi *et al.* 2012; Zhang *et al.* 2017)。其他真菌如青霉菌和镰刀菌等会通过竞争性抑制而减少黄曲霉毒素的产生 (Horn *et al.* 1998)。此外，同一菌种的不同形态类型也会影响花生的黄曲霉毒素污染水平，例如黄曲霉 (*A. flavus*)，它有两种类型，S-和 L-菌株。研究表明黄曲霉 S 型菌株的发生率与花生黄曲霉毒素污染密切相关，而 L 型菌株与花生黄曲霉毒素总量无显著相关性 (Mutegi *et al.* 2012)。而不产毒的黄曲霉和寄生曲霉也可以通过竞争性抑制对黄曲霉毒素进行生物防控。

1.3.1.4 气候条件

工业革命以来,全球平均温度呈显著的上升趋势,1970年以后温度化进程加快。20世纪80年代以来,气候变化对环境、生态、经济等可能造成的问题成为人类关注的焦点,气候问题已成为全球公认的环境问题。气候变化已成为当今世界最大的国际政治议题之一,在2009年丹麦举行的哥本哈根气候变化大会上,温家宝总理同基础四国(中国、印度、南非、巴西)首脑和美国总统奥巴马磋商“哥本哈根协定”。在2013年9月IPCC(Intergovernmental Panel on Climate Change,联合国政府间气候变化专门委员会)发布的第五次评估报告中指出,在过去的112年里(1901-2012),全球平均气温升高0.89℃,并指出未来气候变暖仍将继续,到21世纪末气温将比1850年至1900年间上升1.5-2.0℃以上。这均表明气候变化已经成为事实,天气状况发生变化,温度不断升高,降雨发生转移。

黄曲霉分布区域较广,在高温高湿的地区(纬度26-35°)比在热带或低温地区分布广泛,花生黄曲霉毒素污染通常发生在纬度35°以下的地区(Logrieco *et al.* 2004),在纬度45°以上的地区则非常罕见(Klich 2002)。Sanders等(1985)发现黄曲霉毒素污染并不是与黄曲霉侵染率直接相关。Cole等(1985)指出,食品或农产品在被黄曲霉侵染后,只有在受到环境胁迫(干旱和高温)的条件下,自然抗性机制打破后才会发生真菌的生长和黄曲霉毒素的产生。即使是在受到真菌侵染后,收获前保持花生仁较高的水分活度也会保护花生的天然防御机制(植保素产生)(Dorner *et al.* 1989; Dorner 2008)。在花生生长过程中影响黄曲霉毒素产生的因素很多,干旱和高温胁迫是花生污染黄曲霉毒素最重要的气候因素(Cole *et al.* 1984; Sanders *et al.* 1993; Craufud *et al.* 2006)。花生结荚期的干旱胁迫和土壤高温(>22℃)会促进黄曲霉毒素污染(Rachaputi *et al.* 2002; Bankole *et al.* 2006; Cole *et al.* 1982; Cole *et al.* 1989; Horn 2005; Rachaputi *et al.* 2002),这是因为干旱胁迫引起植物脯氨酸含量的增加,从而增加了黄曲霉毒素的产生(Barnett *et al.* 1966; Payne *et al.* 1983)。有研究表明与灌溉条件下种植的花生相比,在干旱条件下种植的花生受黄曲霉毒素污染更为严重(Reddy *et al.* 2003)。因此,充足的降雨量可以控制或减少花生黄曲霉毒素的产生。此外,收获前一个月灌溉可以帮助作物抵御高温和干旱胁迫。

1.3.1.5 农艺措施

不同的农艺措施,包括轮作、套作、种植时间以及灌溉和施肥情况,也会影响花生受黄曲霉的侵染和黄曲霉毒素的污染(Torres *et al.* 2014)。在同一土地上连续种植花生会加重真菌的侵染和黄曲霉毒素的污染(Ortiz *et al.* 2011)。作物轮作,特别是与不易被曲霉属侵染的作物轮作可降低真菌的存活率(Mutegi *et al.* 2012)。然而,作物轮作对黄曲霉毒素污染的影响取决于种植环境,例如,在半干旱环境中,由于曲霉的种群密度本身可能很高,作物轮作对真菌活性的影响则可能很小(CAC 2004)。在不使用杀虫剂的地区,作物的种植密度较大或施肥不足有可能增加被黄曲霉毒素污染的风险(Anderson *et al.* 1995)。害虫对黄曲霉毒素的污染起到重要的作用,因为害虫破坏花生组织,从而加速了真菌的侵染,导致花生黄曲霉毒素污染加重(Waliyar *et al.* 2008)。已有研究报道杀虫剂和杀菌剂的使用会抑制产毒真菌的生长和黄曲霉毒素的产生(D'Mello *et al.* 1998; Lee *et al.* 2001; Dorner *et al.* 2003)。此外,适时播种、合理种植、科学施肥、加强田间管理对减少花生黄曲霉毒素污染是非常必要的。

1.3.1.6 植物抗毒素的产生

植物抗毒素是由植物在病原体感染部位迅速合成的抗菌物质。虽然还不确定植物抗毒素的化

学性质,但研究表明当花生受到真菌侵染时会产生植物抗毒素 (Aguamah *et al.* 1981; Wotton *et al.* 1985)。早在 1972 年就证实了收获前花生荚果对真菌的抗性是由于真菌侵染使其产生大量植物抗毒素 (Vidhyasekaran *et al.* 1972)。只要花生具有产生植物抗毒素的能力,其就不会被黄曲霉毒素污染 (Dorner *et al.* 1989)。然而当干旱胁迫造成植物抗毒素产生停止时,则会发生黄曲霉毒素的污染 (Dorner *et al.* 1989)。水分活度是影响花生产生植物抗毒素的重要因素,花生在高水分活度条件下 ($a_w > 0.97$) 能够产生足够的植物抗毒素,从而抑制黄曲霉的生长和黄曲霉毒素的产生。随着花生籽粒 a_w 的降低,长期干旱 ($a_w < 0.95$) 导致花生籽粒产生植物抗毒素的能力下降,直至最终丧失。

1.3.2 收获期影响花生黄曲霉毒素污染的主要因素

如果不及时收获或机械收获花生,黄曲霉毒素的污染则可能会加重 (Heathcote *et al.* 1978; Cole *et al.* 1995; Torres *et al.* 2014)。花生的收获时间取决于其成熟度和收获期的气候条件。

在适宜成熟度收获花生是非常重要的,因为过熟或不熟收获都会导致最终花生黄曲霉毒素含量的升高 (Diao *et al.* 2014)。此外,由于真菌侵染和黄曲霉毒素的产生,延迟收获的种子质量会变差 (CAC 2004)。已有研究表明,黄曲霉侵染和黄曲霉毒素污染水平在较小、未成熟籽粒中高于成熟的籽粒 (Sanders *et al.* 1981; Cole *et al.* 1982; Hill *et al.* 1983; Cole *et al.* 1985; Sanders *et al.* 1985)。这是由于未成熟花生产生植物抗毒素较慢,从而会较早地受到黄曲霉毒素的污染。此外,即使是在无法产生植物抗毒素的条件下,成熟花生对黄曲霉毒素污染仍然有较强的抵抗力 (Dorner *et al.* 1989)。

在干燥天气收获的花生通常不会发生黄曲霉毒素污染,相反经常发生在潮湿天气收获的花生 (Kabak *et al.* 2006)。在机械收获和脱粒过程中造成的损伤,会使花生籽粒更易受到真菌的侵袭和黄曲霉毒素的污染 (Heathcote *et al.* 1978)。受到虫害的花生种子污染黄曲霉毒素的风险显著高于完整的花生种子。即使是在不利于真菌感染和黄曲霉毒素产生的条件下,病虫害也会增加花生种子受黄曲霉毒素污染的风险 (Sobolev *et al.* 2007)。

1.3.3 收获后影响花生黄曲霉毒素污染的主要因素

一般来说,当收获后花生籽粒含水量大于 10% 时,容易产生黄曲霉毒素,因此收获后及时干燥并保持在安全含水量水平,可有效控制花生黄曲霉毒素的污染 (Torres *et al.* 2014)。Diener 等 (1970) 报道,收获后将花生迅速干燥并保持 a_w 低于 0.83 可以有效抑制黄曲霉毒素的产生。黄曲霉在花生中生长的最低含水量为 8-10%,环境相对湿度为 82%。通常籽粒含水量高于 10% 时容易产生黄曲霉毒素 (Diener *et al.* 1970)。堆放花生可以引起热量积聚和水分凝聚,继而导致真菌生长和黄曲霉毒素污染。有研究报道未脱壳花生的最大储藏含水量为 9%,而脱壳花生的最大储藏含水量为 7%。在这样的水分含量下,如果环境相对湿度保持在 70%,温度保持在 25-27 °C,则可以保证花生安全储存一年左右 (Odogola 1994; Waliyar *et al.* 2007; Waliyar *et al.* 2008)。

为了防止在储存和运输过程中污染黄曲霉毒素,最重要的是控制温、湿度和卫生条件 (Dickens 1977; CAC 2004)。屋顶漏水、仓库通风不当导致的冷凝以及未及时干燥的花生都会造成储藏期间较高的籽粒水分 (Davidson *et al.* 1982)。因此,储藏和运输条件是控制花生黄曲霉毒素污染的

重要因素。Dickens (1977) 建议在干净、干燥、低湿度 (约为 8%) 的条件下贮藏花生, 并且要保持花生免受害虫和真菌的侵害, 这至今仍是合理存储花生的基础条件。

储藏前分离受污染的花生是减少黄曲霉毒素产生的有效途径 (Cole *et al.* 1995; Dorner 2008)。储藏花生必须定期检查是否污染有霉菌, 如果发霉明显则应立即确定和消除发霉的原因, 同时将霉变花生立即去除。由于产生真菌毒素的真菌是专性需氧菌, 因此通过气调保存, 例如加入二氧化碳、氮气、一氧化碳和二氧化硫气体, 可以防止或减少真菌毒素的产生 (Kabak *et al.* 2006)。Heathcote 等 (1978) 研究指出增加储存仓库中 CO₂ 的浓度会显著减少黄曲霉毒素的产生。

1.4 花生黄曲霉毒素污染防控研究进展

由于花生黄曲霉毒素污染不能完全避免, 因此有必要研究其防控措施。花生黄曲霉毒素的防控措施主要有两类, 一类是脱毒, 可以去除或降解花生黄曲霉毒素, 包括物理、化学和生物学方法; 另一类是预警, 基于影响因素对黄曲霉毒素污染风险进行预测。预警方法是目前和今后研究的热点, 有助于预防和减少收获前污染和优化收获后管理。

1.4.1 物理防控

1.4.1.1 分离挑选

花生收获后首先需要进行分离挑选, 即将带有可见霉菌、虫蚀粒、坏果、烂果、未成熟粒挑出, 不得用于食用。之后利用电子色选去除脱色及疑似污染样品, 这是由于黄曲霉毒素污染极不均匀, 通常只发生在少数种子, 而大部分花生种子并没有被污染。而且通常污染样品与未污染样品颜色差别很大, 例如污染样品会颜色较深、较浅或发生霉变, 与该品种花生正常种仁色差很大的样品则需要进行分离。

常用的分离挑选方法包括手动分拣、大小和密度分离或电子色选 (Dorner 2008), 其中电子色选是脱壳花生加工企业使用的最有效的分离技术, 可以减少 70% 的黄曲霉毒素 (Cole *et al.* 1995)。这种方法存在的不足之处是分离出的样品并不一定全部受到黄曲霉毒素污染 (Waliyar *et al.* 2008)。近年来, 虽然电子分选技术得到了不断的提高, 然而并非所有受黄曲霉毒素污染的种仁都发生了色变, 因此该技术并不能完全有效去除黄曲霉毒素污染样品。因此, 为了有效降低花生仁中黄曲霉毒素的整体污染水平, 最好的方法是在光电分选和人工分离之后进行漂洗 (Dorner 2008)。但主要缺点是成本比较高, 包括直接费用、漂洗过程中的重量损失以及分拣而损失的样品等 (Dorner *et al.* 2006)。

1.4.2 化学防控

花生黄曲霉毒素污染是不可避免的, 且黄曲霉毒素理化性质十分稳定, 耐高温, 一般的烹调 and 加工方法难以破坏其毒性。因此, 需要采取新的方法对污染产品进行脱毒以减少其对人体健康的影响。化学脱毒措施主要包括:

1.4.2.1 杀虫剂和杀菌剂的使用

在田间植物生长期间使用杀虫剂的主要目的是控制虫害, 从而减少真菌侵染的风险 (EC 1999)。然而杀菌剂在鲜花生或通风储藏花生上的应用效果却不明确。一些研究表明在大田实验

中杀菌剂可以控制真菌毒素的产生, 然而由于杀菌剂的抗药性导致产毒真菌的抗性突变, 间接降低了杀菌剂对真菌毒素的控制效果。就产黄曲霉毒素真菌而言, Markoglou 等 (2011) 发现了对杀菌剂苯胺基嘧啶具有抗性的寄生曲霉菌株, 其产毒力远高于野生型菌株。

1.4.2.2 臭氧防控

Proctor 等 (2004) 研究了臭氧和热处理对花生仁和花生粉中黄曲霉毒素的影响, 结果表明臭氧脱毒效率随着温度的升高和处理时间的延长而增加。其中黄曲霉毒素 B₁ 和 G₁ 的降解率最高, 花生仁的降解率高于面粉。随着处理时间的延长, 升温对黄曲霉毒素的降解效率降低, 表明在室温下臭氧处理 10-15min 的降解率与较高温度下短时间内的降解率相同, 然而低温降解却更经济。de Alencar 等 (2012) 为证实臭氧对花生黄曲霉毒素的杀菌和脱毒作用, 研究了花生仁在 13 和 21 mg/L 的臭氧浓度下处理 0、24、48、72 和 96h 之后的脱毒效率。结果表明臭氧能有效控制花生中的真菌菌落总数, 在臭氧浓度为 21 mg/L 处理花生样品 96h 后, 菌落数较少约为 10³/g。

1.4.2.3 抗氧化剂防控

从人体健康角度来说, FDA 允许使用抗氧化剂如丁基羟基茴香醇、丁基羟基甲苯和对苯二甲酸丙酯作为抗菌剂, 并且被认为是安全的化学物质。通过含有这些抗氧化剂的物质可以延迟氧化酸败从而达到防止花生氧化的目的, 并且已经在天然污染黄曲霉毒素和接种黄曲霉的花生中得到验证。研究表明使用食品级抗氧化剂三元混合物处理储藏花生, 可以抑制花生中的微生物菌群尤其是黄曲霉菌群的发生, 并且未检测到黄曲霉毒素 (Passone *et al.* 2009)。此外, 食品企业和消费者越来越多地倾向于使用天然植物化学物质来抑制产毒真菌的生长和黄曲霉毒素的合成 (Etcheverry *et al.* 2011)。

1.4.3 生物防控

1.4.3.1 不产毒真菌防控

在收获前, 基于不产毒黄曲霉和寄生曲霉菌株的生物防治方法已经成功应用于大田作物, 并且减少了在储藏过程中黄曲霉毒素的污染 (Dorner *et al.* 2002b)。这种方法的原理是, 当在土壤中加入不产毒菌株孢子时, 它们将与自然存在的产毒菌株竞争侵染部位, 以获得所必需的养分。此外, 研究证明土壤接种不产毒菌株具有延续效应, 可减少花生在贮藏期间受黄曲霉毒素的污染 (Dorner 2004; Dorner *et al.* 2002b)。基于这一原理, 美国已成功将不产毒菌株 AF36 和 *A.flavus* K49 应用于田间条件下棉花和玉米黄曲霉毒素的生物防治 (Abbas *et al.* 2011b; Cotty 2006; Cotty *et al.* 2003; Horn *et al.* 2000)。同样, 尼日利亚利用 4 株不产毒的黄曲霉菌株制成生物防制剂, 并命名为 Aflasafe™ (Atehnkeng *et al.* 2008)。在花生中, 不产毒黄曲霉菌株 NRRL 21882 已成功商业化制成生物制剂, 品牌名称为 Afla-Guard™ (Dorner *et al.* 2002b; Dorner *et al.* 1998; Horn *et al.* 2009)。Afla-Guard™ 在美国的首次使用便对黄曲霉毒素得到了成功防控, 农民的花生库存中黄曲霉毒素平均含量减少 85%, 去壳花生中黄曲霉毒素含量减少率高达 98% (Dorner 2009)。澳大利亚和阿根廷等国家也将这一方法成功应用于花生的黄曲霉毒素污染防治 (Alaniz Zanon *et al.* 2013; Pitt *et al.* 2006)。目前, 为了确定对黄曲霉毒素的防控程度以及如何将生物防治很好地纳入标准农业生产, 多机构、多国家一直在进行大规模的田间试验 (Abbas *et al.* 2011a)。Horn 等 (2009) 报道了黄曲霉的有性繁殖阶段, 并且正在探索在自然界有性重组的应用 (Moore *et al.* 2009; Olarte

et al. 2010)。虽然黄曲霉的不产毒菌株已成功应用于黄曲霉毒素的生物防控，但是多年重复使用其有可能发生基因转移，从而导致原始菌株的遗传组成发生变化，这些新的黄曲霉基因型的产毒力及其对作物的侵染力还需要进一步研究（*Olarte et al.* 2012）。

1.4.3.2 细菌防控

Kong 等（2010）在中国黄海分离得到一株巨型海洋芽孢杆菌，并通过体外和体内试验研究了其对黄曲霉引起花生腐坏的抑制作用。结果表明，该拮抗剂的生物防治效果显著，能够降低黄曲霉毒素的生物合成及 *aflR* 基因和 *aflS* 基因的表达（*Kong et al.* 2010）。

1.4.4 全程防控

花生黄曲霉毒素污染是一个世界性的难题，对人类健康构成威胁，已成为世界贸易的主要制约因素。多种影响因素影响花生黄曲霉毒素污染，这些因素存在地区差异和年份差异。针对影响花生黄曲霉毒素污染的因素，已有研究报道了抑制毒素生物合成的相应对策。但是花生黄曲霉毒素的污染可以发生在种植、收获、储藏、运输、销售等各环节，针对单一环节已不足以解决花生黄曲霉毒素的污染问题，除了选用合适的花生品种、地块，合理密植、灌溉、施肥以及适时收获，还需要对不同的环节进行全程有效的监控，将农产品安全管理的重点从最终产品的检测转到农产品生产的全程监控，实现“从农田到餐桌”的全程防控，对于减少花生黄曲霉毒素是必要的。

1.4.4.1 基于良好农业规范的预防措施

美国规定花生的生产和加工应遵循适用于所有人类可食用食品的《国际推荐操作规范—食品卫生一般原则》的要求，其从储藏、运输、人员及包装等方面对整个生产链卫生做了要求（CAC 2003）。良好农业规范（Good agricultural practice, GAP）从广义上是指用以阐述生产健康安全农业产品和非农业产品的农业生产和产后加工的环境、经济和社会的可持续性发展（吕婕 等 2009）。姜宗亮等（2012）对花生收获前品种、土壤、种植方式、田间灌溉及病虫害防治，收获设备、收获时间、病株处理及干燥清洁，运输容器、条件控制、储存环境等提出推荐措施，可以基于 GAP 对花生黄曲霉毒素进行防控。

在实施 GAP 源头治理方面，我国科学研究数据尚不充分，国内花生业乃至世界花生产业在应用“良好农业规范（GAP）”这一农产品安全控制技术研究亟待开展。

1.4.4.2 基于良好操作规范的预防措施

良好操作规范（Good manufacturing practice, GMP）要求食品生产企业具备良好的生产设备、合理的生产工艺、完善的质量管理和严格的检测系统，以保障最终产品的质量符合标准（陈小霖 2008）。姜宗亮等（2012）对花生从接收脱壳、分级、脱皮、成品的分装与储存等方面提出推荐措施，可以基于 GMP 对花生黄曲霉毒素进行防控。王耀波等（2002）通过对出口花生全过程可能造成黄曲霉毒素污染的因素进行分析，有针对性地提出了监控措施和建议，提出建立 GMP，以期最大限度从源头上预防花生黄曲霉毒素的污染，但这一设想尚未得到实施。

1.4.4.3 危害分析与关键控制点

危害分析与关键控制点（Hazard analysis and critical control point, HACCP），是指通过对食品“从田间到餐桌”全过程关键危害的分析，设立关键控制点并采取控制措施预防危害发生，从而达到保障花生消费安全的目的（陈小霖 2008）。其作为一种应用广泛的综合性食品安全管理系统，

自从 20 世纪 70 年代在美国形成以来，逐渐为世界众多国家所采纳。HACCP 作为食品安全控制体系，能够对食品中的危害进行系统的识别与评估，当结果和 HACCP 计划有别时，则会采取纠偏措施进行纠正，是一个预防性而不是应急性的工具，其旨在通过危害识别、实施监控和及时防控将风险降到最低 (Park *et al.* 1999)。

张春新等 (2005) 运用 HACCP 原理，从分析产品危害产生的原因、关键控制点的确立和控制措施的设置三个主要方面对出口花生黄曲霉毒素的全过程控制进行了探讨。曾义等 (2007) 运用 HACCP 原理，对炒制花生“从农田到加工出厂”的过程进行了黄曲霉毒素的危害分析，确定了关键控制点，并提出了预防和控制措施。Gorayeb 等 (2009) 在巴西典型花生产后企业中，将 HACCP 计划运用于从带壳原料接收到去壳花生运输的全过程，并确定了带壳原料接收，干燥、贮藏，脱壳花生的贮藏和在贮藏过程中对温湿度的控制这四个黄曲霉毒素的关键控制点，并进行科学管理以保证消费安全。

在收获前、收获后干燥和储存期强调良好农业规范 (GAP)，在不同产品的加工和销售阶段突出良好操作规范 (GMP)，HACCP 系统应当建立在完善的良好农业规范 (GAP) 与良好操作规范 (GMP) 基础上 (姜宗亮 等 2012)。

1.4.5 预警技术

自 1960 年“火鸡 X 病”事件后，花生黄曲霉毒素污染问题受到了全世界的广泛关注，但如何对其进行有效控制仍然是一个难题，开展花生不同环节黄曲霉毒素污染预警技术研究，在早期发现并进行预防，可以有效减少花生及其制品中黄曲霉毒素污染，是花生黄曲霉毒素污染控制的高效手段，也是研究的一个新方向和热点。已有的研究表明，开展花生黄曲霉毒素污染产前、产后预警是可行的，但是不同国家、不同地区、不同环节花生黄曲霉毒素污染影响因素、关键控制点等都不同，即使同一地块分离得到的黄曲霉或寄生曲霉，其产毒力也存在较大差异，因此要建立科学、高效的预警技术和模型，难度很大。系统开展不同产地、不同环节花生黄曲霉毒素污染来源、发生机制、迁移转化与累积分布规律研究，探明花生黄曲霉毒素污染关键环节与关键控制点，并基于多年多点花生黄曲霉毒素污染和相关品种、土壤、真菌产毒力和气候环境等大数据对预警模型进行校准和修正，建立适合不同地区、不同环节花生黄曲霉毒素的预警技术与模型，将是今后花生黄曲霉毒素预警技术研究的重点和发展方向。

剔除霉变粒、脱毒控制措施成本高、成效低。例如，一般的化学脱毒措施反应可逆，有的降解产物毒性未知，部分脱毒后的花生产品不能再用作食品，会造成大量的浪费 (Chauhan *et al.* 2010)。因此根据科学预警结果进行针对性防控是花生黄曲霉毒素控制最经济有效的方法，成为花生黄曲霉毒素污染控制研究的热点。气象环境和土壤条件是影响花生污染黄曲霉毒素的主要因素。通过对经验数据进行规律总结，可以将这些因素造成的黄曲霉毒素污染风险计算出来，进行风险预警，进而根据预警结果在收获前采取调整收获期、灌溉等农艺措施，或收获后进行温湿度调控等措施来减少花生污染黄曲霉毒素的风险。因此花生黄曲霉毒素预警是在以往花生种植、收获、储藏、运输、加工过程中，对黄曲霉毒素随着温湿度等外界条件的变化产生和积累的规律进行监测与分析，当出现黄曲霉毒素超标的前兆时，则采取一定的调控措施以降低黄曲霉毒素污染，减少或消除损失的方法。国外食品和农产品中危害因子预警系统研究起步较早，我国预警技术起

步于 20 世纪 80 年代, 且多集中于粮食产量和病虫害的预测预报。随着公众对食品和农产品质量的高度重视, 食品及农产品质量安全预警研究引起广泛关注。根据花生种植、生长及储藏等不同阶段, 将花生黄曲霉毒素的预警方法分为收获前预警和收获后预警。

1.4.5.1 收获前的预警方法

气候环境和花生根际的土壤条件是田间阶段影响花生黄曲霉毒素污染的主要因素。国内外学者研究发现, 收获期高温干旱是花生受黄曲霉侵染和黄曲霉毒素累积的主要气候风险因子 (Blankenship *et al.* 1984; Craufurd *et al.* 2006; Cotty *et al.* 2007)。在收获前 4-6 周, 即花生结荚期, 干旱和高温 (22-35 °C) 是触发花生产生黄曲霉毒素污染的主要因素 (Hill *et al.* 1983; Cole *et al.* 1985; Dorner *et al.* 1989; 张建成 等 2006; 蔡骥业 1996; 张毅 等 2010)。基于这些研究结果, 美国、澳大利亚等国家已研究出针对气象参数及农艺参数构建的田间阶段花生黄曲霉毒素预警模型。通过预警模型, 可以将气象因素造成的黄曲霉毒素污染风险计算出来, 从而对其风险进行预警, 根据预警结果进而采取相应的农艺措施来减少花生黄曲霉毒素污染风险。

在美国, 1990 年 Thai 等研究发现在干旱条件下, 黄曲霉毒素污染过程呈近似一级反应动力学曲线, 且生长速率取决于花生根际 (土壤表层以下 2.54-5.08 cm) 土壤温度, 但以此建立的模型仅基于干旱条件下的土壤温度参数, 且没有得到实际应用。1997 年, Parmar 等利用神经网络和线性回归方法建立黄曲霉毒素污染预测模型, 模型参数包括土壤温度、干旱时间、作物生长时间和积温。结果表明当计算积温的土壤温度阈值为 25 °C 并且神经网络有 8 个隐藏层时, 模型的训练集和测试集的决定系数 (Coefficient of determination, R^2) 分别达到 0.9250 和 0.9522。2000 年, Henderson 等根据遗传算法/反向传播神经网络方法 (Genetical gorithm/back propagation neural network, GA/BPN), 利用相同的参数建立了花生黄曲霉毒素预测模型, 可以为 GA/BPN 自动确立最优参数, 然而由于美国花生营销系统规定花生中出现黄曲霉时才会进行经济处罚, 而对只含有黄曲霉毒素的花生没有规定, 因此花生黄曲霉毒素预警模型在美国并未得到推广应用。

美国在 1998 年发布了 CROPGRO 模型, 其中 CROPGRO-peanut 是一个可以根据输入的气候条件、土壤特征、植物生长特征管理措施来模拟花生水分平衡、荚果周围土壤温度、叶片表面温度和植物缺水量的模型 (Boote *et al.* 1998), 在花生结荚期适宜黄曲霉生长和产毒的条件下, 根据其模拟的土壤湿度与花生黄曲霉毒素的关系可以预测黄曲霉毒素含量, 并成功应用于尼日尔的黄曲霉毒素风险预测 (Craufurd *et al.* 2006)。基于对 CROPGRO-peanut 模型的应用, 根据黄曲霉毒素含量与繁殖期气候参数 (降水量、最低温、最高温和日照时长) 的关系, 马里实现了对花生黄曲霉毒素含量的监控与预测 (Boken *et al.* 2008)。

澳大利亚 2010 年开展了基于气候环境条件的黄曲霉毒素风险预警研究, 并将其作为降低花生黄曲霉毒素污染的主要策略。昆士兰大学基于作物生长模拟方法 (Agricultural production systems simulator, APSIM), 利用环境温度、辐射量、降水量及土壤温度等气候环境因子建立了田间生长阶段花生黄曲霉毒素污染预测模型, 农户可以将日降水量、环境温度和土壤温度输入一个互联网交互页面—“AFLAMAN”, APSIM 中的花生黄曲霉毒素污染预测模型便可做出计算与预测, 并将田地里土壤温度、湿度及黄曲霉毒素污染的趋势图再通过“AFLAMAN”反馈给农户, 以便农户采取相应的预防措施 (Chauhan *et al.* 2010), 此模型在澳大利亚小范围区域得到实际应用, 但澳大利亚花生种植面积小, 土壤气候类型单一, 花生品种少, 此模型并不适合中国这样的花生主产大国。

此外,也有针对其它作物的收获前预警模型相关研究, Battilani 等 (2013; 2016) 应用理论模型 AFLA-Maize, 根据收获前每小时气温、相对湿度和降水量预测了欧洲玉米和小麦中的黄曲霉毒素污染风险。Chauhan 等 (2015) 应用 APSIM 理论模型预测肯尼亚地区收获前玉米中黄曲霉毒素污染风险。Salvacion 等 (2015) 利用模糊逻辑算法预测当前与将来气候变化对玉米中黄曲霉毒素污染的影响。Damianidisa 等 (2018) 利用农业干旱参考指数 (the agricultural reference index for drought, ARID), 根据多元逻辑回归模型对玉米中黄曲霉毒素污染进行预测, 计算 ARID 所利用的参数为日最高气温、日最低气温、露点温度、降水量、风速、潜在蒸发散射量和太阳辐射量。Smith 等 (2016) 利用有关降雨的遥感数据、植被覆盖 (由归一化植被指数表示, Normalized difference vegetation index, NDVI)、土壤特性 (有机质含量、pH、盐基交换总量、盐度、质地和土壤类型) 来预测肯尼亚玉米黄曲霉毒素污染。在收获前对黄曲霉毒素进行预警进而采取预防性农艺措施, 不仅省时省力而且大大减少了产后脱毒的费用, 因此是目前研究热点。

1.4.5.2 收获后的预警方法

收获后的预警方法包括:

(1) 模型预警

花生黄曲霉毒素污染不仅可发生在收获前, 只要条件适合, 在收获后的干燥、储藏和加工等环节产毒真菌都可以侵染合适的寄主进而产生黄曲霉毒素, 并进入食物链 (Brown *et al.* 1999)。已有国内外学者利用预测微生物学手段对不同温度和湿度条件下花生中的产毒真菌建立了预测模型, 一旦出现适合产毒的条件, 便可以采取预防措施, 将损失降到最低。

Troeger 等 (1989) 根据批量花生烘烤模型对烘烤过程中热风的温湿度条件进行管控, 根据热风的温湿度来模拟花生的水分含量变化, 当条件适合霉菌生长和毒素产生时, 模型就会给出预警指示。Whitaker 等 (2002) 研究了利用电子制表模型预测在储藏、脱壳、漂洗过程中花生黄曲霉毒素的含量分布, 可以更好地对污染花生进行管理。李瑞芳 (2007) 用 Boltzmann 和 Logistic 两种模型模拟了储藏期间黄曲霉的生长规律, 并初步探索了其黄曲霉毒素含量之间的关系, 即当储藏温度低于 15 ℃、相对湿度低于 85% 时, 可以保证花生避免黄曲霉污染, 预防黄曲霉毒素产生。Zhang 等 (2011) 用示差扫描量热法和热重分析法得到 100-190 ℃ 条件下花生黄曲霉毒素动力学模型, 预测相关系数平均值分别为 0.857 和 0.750, 此模型可以成功应用于花生热加工过程中黄曲霉毒素的预警。

产毒真菌的产毒力存在种间及种内差异, 且目前对其生物合成途径差异及影响因素仍知之甚少, 因此收获后产毒预警模型的研究较少。

(2) 前体物预警

由于 AFB₁ 有剧毒性, 会给人体健康带来危害, 如果找到在生物合成过程中与它的产生有密切相关的前体化合物, 并且该化合物更容易被检测, 或者具有更小的毒性, 则利用该化合物对花生黄曲霉毒素污染进行预警将具有重要的意义。

杂色曲霉素 A (Versicolorin A, VA) 是 AFB₁ 的生物合成前体物, 是黄曲霉毒素生物合成过程中首个出现毒性作用基团双呋喃环的化合物。国内外对以黄曲霉毒素生物合成前体代谢物作为 AFB₁ 的预警检测进行了初步研究。研究表明, VA 能够指示 AFB₁ 的产生且毒性较 AFB₁ 小, 检测方法也较安全、简便, 因此可以达到预警 AFB₁ 的目的 (谭辉勇 2010)。方明英 (2008) 通过研究表明在液体察氏培养基、酵母浸膏蔗糖培养基、大豆培养基和大米中培养黄曲霉和寄生曲霉,

利用薄层色谱法可以在 AFB₁ 产生前检出 VA。谭辉勇（2010）探讨了黄曲霉/寄生曲霉产毒的预警报告生物标志分子，结果表明样品中的 VA 和 AFB₁ 含量之间具有中度相关性（ $R^2=0.405$ ），并且当样品中 VA 含量较高，AFB₁ 含量较低时，在湿度为 80%，温度为 28℃ 避光放置 10d 后 AFB₁ 的含量与 10d 前样品中 VA 的含量呈曲线相关关系，相关系数 $R=0.615$ 。因此 VA 作为黄曲霉毒素的预警报告分子是可行的。Alkadri 等（2014）指出作为黄曲霉毒素的前体物，杂色曲霉素（Sterigmatocystin, ST）可以在储藏期间转化为黄曲霉毒素。

（3）分子预警

黄曲霉毒素分子预警研究还处于起步阶段，目前对 AFB₁ 的分子预警研究主要是通过水分和温度对黄曲霉和寄生曲霉关键基因簇的表达、生长速率及黄曲霉毒素产生的影响和结构基因的表达量与 AFB₁ 的相关性。Schmidt 等（2010）研究表明关键基因 aflR/aflS 的表达量之比与温度×水分活度之间的交互作用相关，该比率越大 AFB₁ 含量越高（Schmidt *et al.* 2009; Schmidt *et al.* 2010），此方法可应用于这两类真菌产毒的预警。Medina 等（2015）通过改进的 Luedeking-Piret 模型研究了基因表达量、温度、水分活度和真菌生长速率与黄曲霉毒素产生之间的关系。这种方法解释了环境胁迫因子与黄曲霉毒素生物合成途径中相关基因之间的关系，并且可通过基因活化来调节产毒力并用于现有模型的修正。虽然 AFB₁ 的含量归根结底受气候变化的调控，但这也为利用基因表达预测毒素含量提供了一个新思路。

1.5 本课题研究目的与意义

花生 (*Arachis hypogaea* L.) 是我国食油兼用型经济作物和优势出口农产品，在中国及世界花生生产中占有重要的地位。然而花生易受真菌侵染并产生黄曲霉毒素，已成为影响我国花生消费安全和世界贸易的主要制约因素。花生黄曲霉毒素的污染可以发生在花生生产的各个环节，其中收获前田间污染过程既是花生黄曲霉毒素污染的主要原因，也是储藏期污染的源头。而我国花生品种繁多，土壤类型和气候类型多样，综合利用花生品种抗性、土壤类型、真菌产毒力及收获前气候条件对黄曲霉毒素进行预测，开展收获前花生黄曲霉毒素污染预警模型研究，是花生黄曲霉毒素污染控制的高效手段，也是目前和今后黄曲霉毒素污染控制技术研究的一个新方向和热点。本研究将会为确保花生质量的进一步提高和花生产业健康发展，为监管我国农产品质量安全提供有力的技术支撑，对花生等主要农产品质量安全监管、保障安全消费以及国际贸易等都具有重要意义。

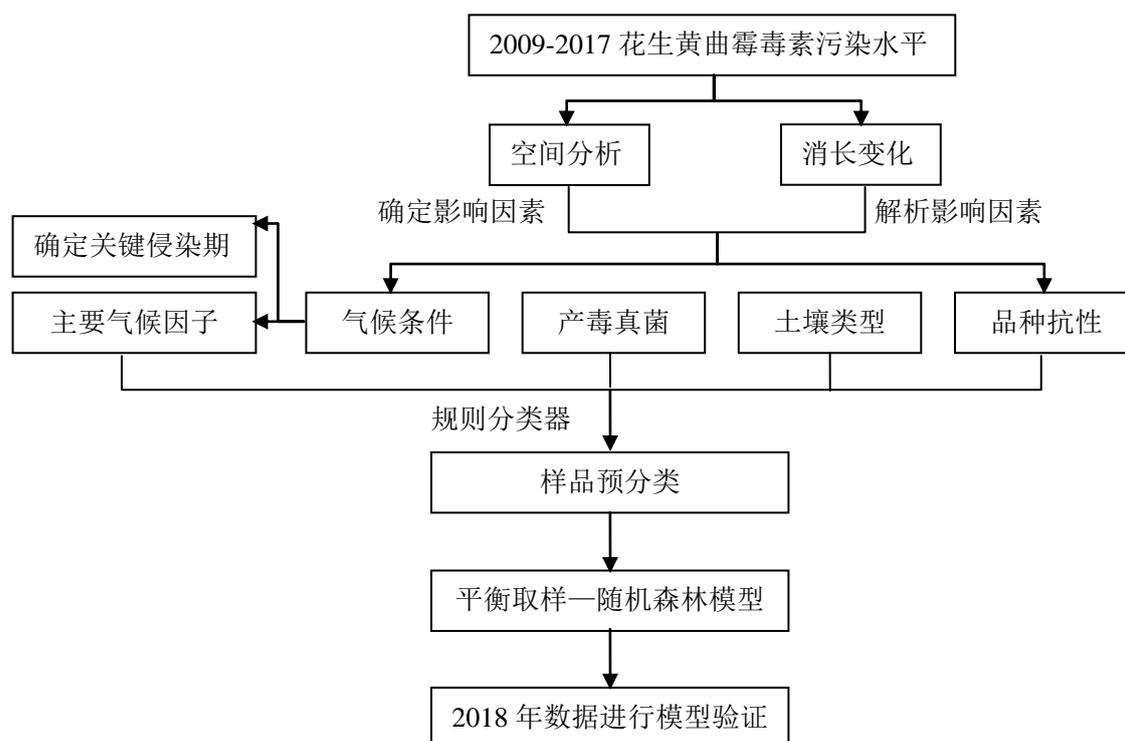
1.6 本课题主要研究内容

本课题的研究内容主要包括以下三个部分：

1. 利用地统计分析和普通克里格插值，绘制中国花生黄曲霉毒素污染分布地图，分析花生黄曲霉毒素污染的空间分布特征，明确污染高风险产区及其影响因素；
2. 纵观 2009-2017 连续九年花生黄曲霉毒素污染消长变化趋势，解析不同花生品种、土壤类型、菌株产毒力及气候条件对黄曲霉毒素污染消长变化的影响；
3. 基于最大互信息系数，建立了适合中国的规则分类—平衡取样—随机森林花生黄曲霉毒素预警模型，实现了对黄曲霉毒素的定性预测。

1.7 本课题技术路线

本课题技术路线如下所示：



第二章 中国花生黄曲霉毒素污染空间分布

2.1 引言

研究花生黄曲霉毒素的空间变异及空间分布特征可揭示花生黄曲霉毒素污染的空间分布规律,对花生黄曲霉毒素污染预测制图、精准农业和长期定位监测等具有重要意义,是研究高效采样策略、黄曲霉毒素精准预测和黄曲霉毒素污染防治的重要基础。花生黄曲霉毒素的空间变异是指在一定的地区内,在同一时间、不同地点的花生黄曲霉毒素的分布特点存在明显的差异性。花生黄曲霉毒素污染是在各种因素综合作用下形成的,涉及到花生品种、土壤和气象因素等,使得其在区域分布上存在很大的时空变异性。空间变异性的主要研究方法有:传统统计学方法、地统计学方法、地理信息系统(Geographic information system, GIS)和地统计相结合的方法以及遥感监测的方法(郭安廷等 2018)。GIS和地统计学相结合的方法是当前最常用的方法,自Matheron提出以后,地统计学方法得到了广泛的应用与发展,在空间数据分布格局建模以及数学化表达特征上具有优势。而GIS在空间数据管理以及精准可视化表达上具有优势,两者相互结合可以有效地揭示变量的空间变异特征(郭安廷等 2018)

GIS是一种计算机控制的方法,可以根据地理坐标(经度和纬度)来集合、储存、处理和显示数据(Battilani *et al.* 2006)。GIS和地统计分析可以用来描述、分析和展示多种变量的空间分布,从而揭示其与空间变量的关系,并解决实际问题。这种方法已经应用于多个农业领域,包括土壤学(White *et al.* 1997)、昆虫学(Byrne *et al.* 1996; Liebhold *et al.* 1993)、医学昆虫学(Nicholson *et al.* 1996)和植物病理学(Jaime-Garcia *et al.* 2001; Nelson *et al.* 1999; Wu *et al.* 2001)。地统计学可以用来描述空间的连续性,以揭示许多自然现象的本质特征,并结合经典的统计回归技术估计未取样位置的值(Cressie 1993),而且空间连续性的模型可以用来评定估计值的不确定度。半变异函数(Variograms)反映了空间自相关,半变异函数的值越低,空间自相关或空间连续性越强(Isaaks *et al.* 1991)。

克里金插值法(Kriging)是重要的地统计工具,是利用回归方法对空间邻近的点进行插值的方法,其在统计相似性的基础上根据取样点的数据对未取样点的数据进行权重赋值,依据估计方差最小化给出了未取样点数据的无偏估计值(Borgelt *et al.* 1997)。当数据分布是正态时,克里金法是最优的估计方法,普通克里金插值法是其中应用较为广泛的一种(葛跃等 2016),可根据采样位置之间的半方差函数确定权重。地统计分析方法已经应用于植物流行病害(Chellemi *et al.* 1988; Gottwald *et al.* 1995; Larkin *et al.* 1995; Lecoustre *et al.* 1989; Stein *et al.* 1994; Webster *et al.* 1992; Wu *et al.* 2001)、番茄病毒管理区域风险评价(Nelson *et al.* 1994)、病原体的遗传结构(Jaime-Garcia *et al.* 2001)、菌群密度模式和黄曲霉毒素污染的菌株组成(Orum *et al.* 1999)。

本章采用地统计学方法(Nelson *et al.* 1999)与GIS相结合的方法分析了花生黄曲霉毒素的空间变异性,以获得中国花生黄曲霉毒素的污染模式。首先利用GS+10.0软件进行半变异函数分析,获得全方位的半变异函数,然后利用克里金插值法(Kriging)获得未取样地区的待测值,得到一个无偏估计值和最小化的误差方差,继而生成花生黄曲霉毒素空间分布栅格图,确定了黄曲霉毒素污染高风险地区,为我国花生黄曲霉毒素污染风险监管和有效防控奠定了重要的基础。

2.2 材料

2.2.1 花生样品

2.2.1.1 花生取样范围与取样量

鉴于气候类型、土壤类型、花生品种和种植方式等的差异,将中国的花生主产区分为四部分,分别为东北主产区(包括黑龙江省、吉林省和辽宁省)、北方主产区(包括河北省、山东省和河南省)、长江流域主产区(包括安徽省、江西省、江苏省、湖南省、湖北省和四川省)和南方主产区(包括广东省、广西壮族自治区和福建省)(图 2.1)。花生抽样范围覆盖四大花生主产区,抽取的花生样品覆盖中国花生生产总面积的 94%以及总产量的 96%。采用分层随机取样方法,在每个主产区根据种植面积和产量抽取 10-30 份花生样品,每个样品 30 公斤。为抽取的花生样本能充分代表该地区花生黄曲霉毒素污染状况,采用四分法分取样,至少每点取样样品 3kg。在抽样过程中要及时准确地记录样品信息(包括花生样品的抽样地点、品种名称、土壤类型、生产情况及管理情况等)。将抽取的带壳花生样品在一周内晾晒至水分含量低于 11%之后运至实验室,并储存于通风干燥的环境,待测黄曲霉毒素含量。

2.2.1.2 花生样品的制备

经过去杂、去尘及脱壳后,用切片机切片(厚度为 0.5cm),再经样品粉碎机粉碎,研磨至可通过 0.9mm 孔径的样品筛,然后将粉碎后的样品装于样品瓶内 4℃ 储存,采用高效液相色谱法(High performance liquid chromatography, HPLC)检测黄曲霉毒素含量。

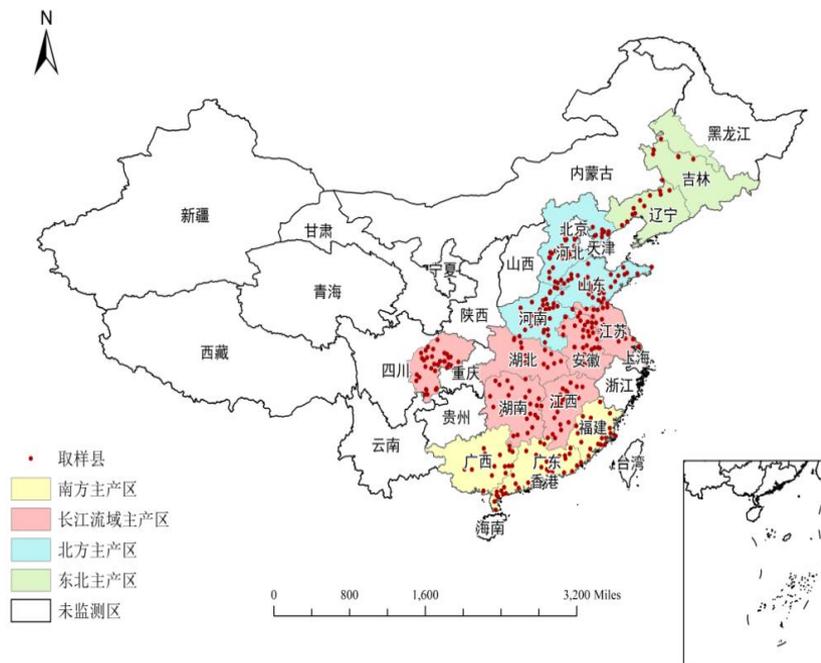


图 2.1 2009-2017 年中国花生取样省和取样市县

Figure 2.1 Sampling provinces and counties of peanuts from 2009 to 2017 in China

2.2.2 仪器与试剂

| | |
|---|-----------------|
| Agilent1200 液相色谱仪, 配荧光检测器 | 美国 Agilent 公司 |
| 切片机 | 温岭市林大机械有限公司 |
| 样品粉碎机 | 日本 NIKAI 有限公司 |
| 电子分析天平 | 德国 Sartorius 公司 |
| 旋涡混合器 | 意大利 VELP 公司 |
| 超声波清洗器 | 昆山市超声仪器有限公司 |
| CT6E 台式离心机 | 日本 HITACHI 公司 |
| 干热氮吹仪 | 北京八方世纪科技有限公司 |
| 中速定性滤纸 | 杭州特种纸业有限公司 |
| 电热鼓风干燥箱 | 上海特讯实业有限公司医疗设备厂 |
| 0.45 μm 有机相滤膜 | 上海市新亚净化器件厂 |
| Nova-pak C18 色谱柱 (3.9mm \times 150mm) | 美国 Agilent 公司 |
| 黄曲霉毒素 B ₁ 、B ₂ 、G ₁ 、G ₂ 标准样品 | 美国 Sigma 公司 |
| 甲醇, 色谱纯 | 国药集团化学试剂有限公司 |
| 乙腈, 色谱纯 | 国药集团化学试剂有限公司 |
| 三氟乙酸, 分析纯 | 国药集团化学试剂有限公司 |
| 石油醚 (60-90 $^{\circ}\text{C}$), 分析纯 | 国药集团化学试剂有限公司 |
| 三氯甲烷, 色谱纯 | 国药集团化学试剂有限公司 |
| 黄曲霉毒素免疫亲和柱 | 实验室自制 |

2.3 方法

2.3.1 花生黄曲霉毒素检测方法

样品通过免疫亲和层析 (Immunoaffinity chromatography, IAC) —HPLC—荧光检测法 (Fluorescence detection, FLD) 进行花生 AFB₁、AFB₂、AFG₁、AFG₂ 含量的检测。具体步骤为: 取研磨后样品 (5g), 用 15mL 4% 的 NaCl 甲醇水萃取。超声提取 (50 $^{\circ}\text{C}$, 10min) 后用双层慢速定量滤纸过滤, 将滤液 (3mL) 收集在 50mL 离心管中。然后与 8mL 双蒸馏水 (dd-H₂O) 混合, 静置过夜。收集 3mL 上清液, 用 8mL 纯水稀释。提取物以 1 滴/s 的流速通过免疫亲和柱进行纯化, 然后用 1mL 甲醇洗脱, 洗脱液过有机膜 (0.22 μm) 过滤后, 进行 HPLC 定量分析。

HPLC 分析在高效液相色谱系统 (Agilent 1200, NYSE: A, Palo Alto, CA, USA) 上进行, 荧光检测器激发光波长和发射光波长分别为 360nm 和 440nm。在 Capcell Pak C-18 色谱柱 (4.6 mm \times 150 mm \times 5 μm) 上进行色谱分离, 以水—甲醇 (55:45) 为流动相, 流速为 0.7mL/min, 进样量为 10 μL 。

该方法灵敏度高, 准确性和重现性好, 加标回收率为 95%-105%, 相对标准偏差 \leq 10%。黄曲霉毒素的检出限为 3 倍信噪比, 本方法的检出限 (Limit of detection, LOD) 分别为: AFB₁: 0.01 $\mu\text{g}/\text{kg}$, AFB₂: 0.01 $\mu\text{g}/\text{kg}$, AFG₁: 0.09 $\mu\text{g}/\text{kg}$, AFG₂: 0.05 $\mu\text{g}/\text{kg}$, 范围内线性关系良好,

$R^2 > 0.999$ 。

2.3.2 黄曲霉毒素空间变异性分析方法

采用地统计学方法分析花生黄曲霉毒素的空间变异性,采用 GIS 空间插值法绘制黄曲霉毒素空间分布图。首先每个采样点分析地理参考利用高斯克吕格投影坐标系。投影坐标系以米为单位,是一种十进制的用经纬度表示的直角坐标系,以便在一个平面(二维)上正确计算采样位置之间的距离(Jaime-Garcia *et al.* 2003)。然后用 GS+10.0 (Plainwell, Michigan, USA) 对黄曲霉毒素数据进行正态分布检验和对数转化并进行半变异函数分析,得到空间插值参数。最后根据空间插值参数,利用 ArcGIS (version 10.2 for Windows; ESRI, Redlands, CA) 空间分析模块中的普通克里金法进行空间插值。

在地统计学中,变异函数值基于由变程、块金值和基台值拟合得到的半变异函数。区域变量 $Z(x)$ 在满足二阶平稳或本征假设的条件下,变异函数存在。在一维条件下,它的定义为当空间点 x 在一维 x 轴上变化时,区域变量 $Z(x)$ 在点 x 和 $x+h$ 处的值 $Z(x)$ 与 $Z(x+h)$ 差的方差一半定义为区域化变量 $Z(x)$ 在 x 轴方向上的变异函数。变异函数 $r(h)$ 的计算公式为(赵越等 2018):

$$r(h) = 1/2E[Z(x) - Z(x+h)]^2 \quad (\text{公式 2.1})$$

变异函数的理论模型主要分为球状模型、指数模型、高斯模型、纯块金效应模型和线性有/无基台值模型。其中:

(1) 球状模型 (spherical model) 的一般公式为:

$$r(h) = \begin{cases} 0 & h=0 \\ C_0 + C[3h/2a - 1/2(h/a)^3] & 0 < h \leq a \\ C_0 + C & h > a \end{cases} \quad (\text{公式 2.2})$$

式中, C_0 为块金常数, $C_0 + C$ 为基台值, C 为拱高, a 为变程。球状模型适用于三维空间,是地统计学应用最广的理论模型(图 2.2)。当 $C_0=0$, $C=1$ 时,称为标准球状模型。

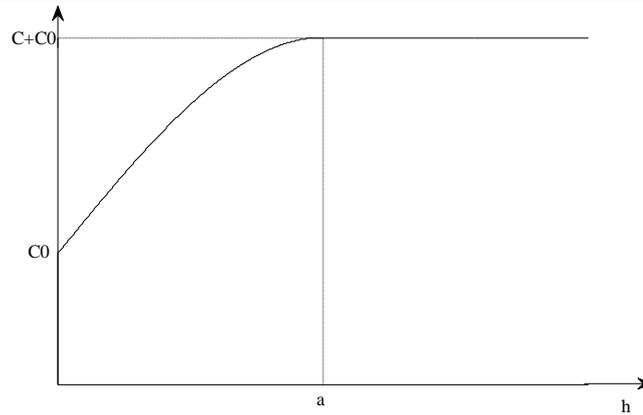


图 2.2 球状模型函数图

Figure 2.2 Histogram of spherical model

(2) 指数模型 (Exponential model) 的一般公式为:

$$r(h) = \begin{cases} 0 & h=0 \\ C_0 + C(1 - e^{-h/d}) & h < 0 \end{cases} \quad (\text{公式 2.3})$$

式中 C_0 和 C 意义同上, d 是控制方程空间范围的距离参数, 其仅在无穷远时相关性完全消失。当 $h=3d$ 时, $r(h)$ 近似等于 C_0+C , 因此指数模型的变程一般取 $3d$ 。当 $C_0=0$, $C=1$ 时, 称为标准指数模型。

(3) 高斯模型 (Gaussian model) 的一般公式为:

$$r(h) = \begin{cases} 0 & h=0 \\ C_0 + C(1 - e^{-(h/d)^2}) & h > 0 \end{cases} \quad (\text{公式 2.4})$$

式中 C_0 和 C 意义同上, 当 $h=\sqrt{3}d$ 时, $r(h)$ 近似等于 C_0+C , 因此高斯模型的变程一般取 $\sqrt{3}d$ 。

(4) 纯块金效应模型的一般公式为:

$$r(h) = \begin{cases} 0 & h=0 \\ C_0 & h > 0 \end{cases} \quad (\text{公式 2.5})$$

式中 C_0 和 C 意义同上。

(5) 线性有基台值模型的一般公式为:

$$r(h) = \begin{cases} C_0 & h=0 \\ Ah & 0 < h \leq a \\ C_0 + C & h > a \end{cases} \quad (\text{公式 2.6})$$

式中 C_0 和 C 意义同上。

(6) 线性无基台值模型的一般公式为:

$$r(h)=\begin{cases} C_0 & h=0 \\ Ah & h>0 \end{cases} \quad (\text{公式 2.7})$$

式中 C_0 和 C 意义同上。

本研究中用到球状模型 C_0 表示块金方差,是由随机因素引起的空间异质性; C 表示结构方差,是由系统因素引起的空间异质性;基台值 $C+C_0$ 表示系统内总的变异; $C/(C+C_0)$ 称为基底效应,表示由自相关部分引起的空间变异性程度(赵明松等 2013)。根据基底效应的大小可对空间变异性进行分级,当 $C/(C+C_0)$ 的值小于 25% 时,表示变量空间自相关性较弱,区域化变量主要受自然因素的影响; $C/(C+C_0)$ 的值在 25-75% 之间时,表示变量具有中等程度的空间自相关性,区域化变量受人为因素和自然因素影响相当;当 $C/(C+C_0)$ 的值大于 75% 时,表示变量具有强烈的空间自相关性,区域化变量主要受人为因素的影响(葛跃等 2016)。球状模型的变程是变异函数达到基台值的样本间隔距离,间隔距离 $>a$ 后,区域化变量空间自相关性消失。

2.3.3 结果统计方法

本章中的平均值均为算术平均值,即使用某一变量所有值的总和除以该变量数目。

2.4 结果与分析

2.4.1 中国花生 AFB_1 空间变异特征

中国的花生黄曲霉毒素污染是一个长期存在的问题,利用 Kolmogorov-Smirnov 检验分别对每年的花生 AFB_1 平均值进行正态分布检验,其呈左偏分布,而经对数转换后,符合正态分布,所以符合地统计分析的条件。把每个县取样点的经纬度和样品 AFB_1 数据分别输入 GS+10.0 软件,进行地统计分析。根据残差(Residual sum of squares, RSS) 越小越好和决定系数(R^2) 越大越好的原则,2009-2017 年 AFB_1 平均值的最佳拟合模型为球状模型,空间变异性参数见表 2.1,半变异函数见图 2.3。

从表 2.1 可以看出,2009-2017 年中国花生 AFB_1 平均值的 $C/(C+C_0)$ 的值在 50-70% 之间,表示具有中等程度的空间自相关性,说明中国花生 AFB_1 的空间异质性是由结构性因素(如土壤类型、土壤母质、气候、地形地貌等)和随机性因素(如花生品种、农业土地利用方式、耕作管理措施、施肥及取样分析误差等)(郭天文等 2007; 宋莎等 2011)共同作用的,由于 $C/(C+C_0)$ 的值越小受气候等自然因素的影响越大,因此结果表明气候条件对花生黄曲霉毒素污染的影响作用相对较大。2017 年的基台值最大为 6.77,表明 AFB_1 含量空间变异性大,2016 年基台值最小为 1.05,表明 AFB_1 含量空间变异性小。模型中块金值均小于基台值,表明 AFB_1 含量受县市场距离的强烈影响。为了了解区域化变量的空间变异特征,变程可作为研究区域内监测站间隔的布设依据。2017 年的变程最大为 5110000m,而 2016 年的 AFB_1 含量的变程最短为 1070000m,由于采样点的布点距离不能超过半变异函数的变程,且最佳布点距离应该为最短变程值的 1/4 至 1/2(葛跃等 2016),因此布点间隔应考虑设置为 26750-53500 m。

表 2.1 2009-2017 年花生 AFB₁ 平均含量的变异函数理论模型参数Table 2.1 Variogram model parameters of mean AFB₁ content in peanut during 2009-2017

| 年份 | 理论模型 | 块金值 | 基台值 | 空间相关性 $C/(C+C_0)$ | 变程 | R ² | RSS |
|------|------|------|-------|----------------------|---------|----------------|------|
| 2009 | 球状模型 | 4.12 | 8.24 | 0.500 | 5110000 | 0.325 | 5.22 |
| 2010 | 球状模型 | 5.63 | 11.26 | 0.500 | 4513000 | 0.432 | 10.6 |
| 2011 | 球状模型 | 5.70 | 12.11 | 0.529 | 4868000 | 0.385 | 24.9 |
| 2012 | 球状模型 | 3.64 | 11.57 | 0.685 | 5110000 | 0.518 | 15.0 |
| 2013 | 球状模型 | 4.07 | 9.90 | 0.589 | 5110000 | 0.525 | 8.40 |
| 2014 | 球状模型 | 6.17 | 15.48 | 0.601 | 4600000 | 0.314 | 53.8 |
| 2015 | 球状模型 | 4.74 | 9.48 | 0.500 | 4110000 | 0.082 | 18.6 |
| 2016 | 球状模型 | 1.05 | 3.478 | 0.699 | 107000 | 0.122 | 38.0 |
| 2017 | 球状模型 | 6.77 | 13.55 | 0.500 | 5110000 | 0.233 | 17.9 |
| 总体 | 球状模型 | 4.61 | 17.22 | 0.732 | 6110000 | 0.861 | 5.23 |

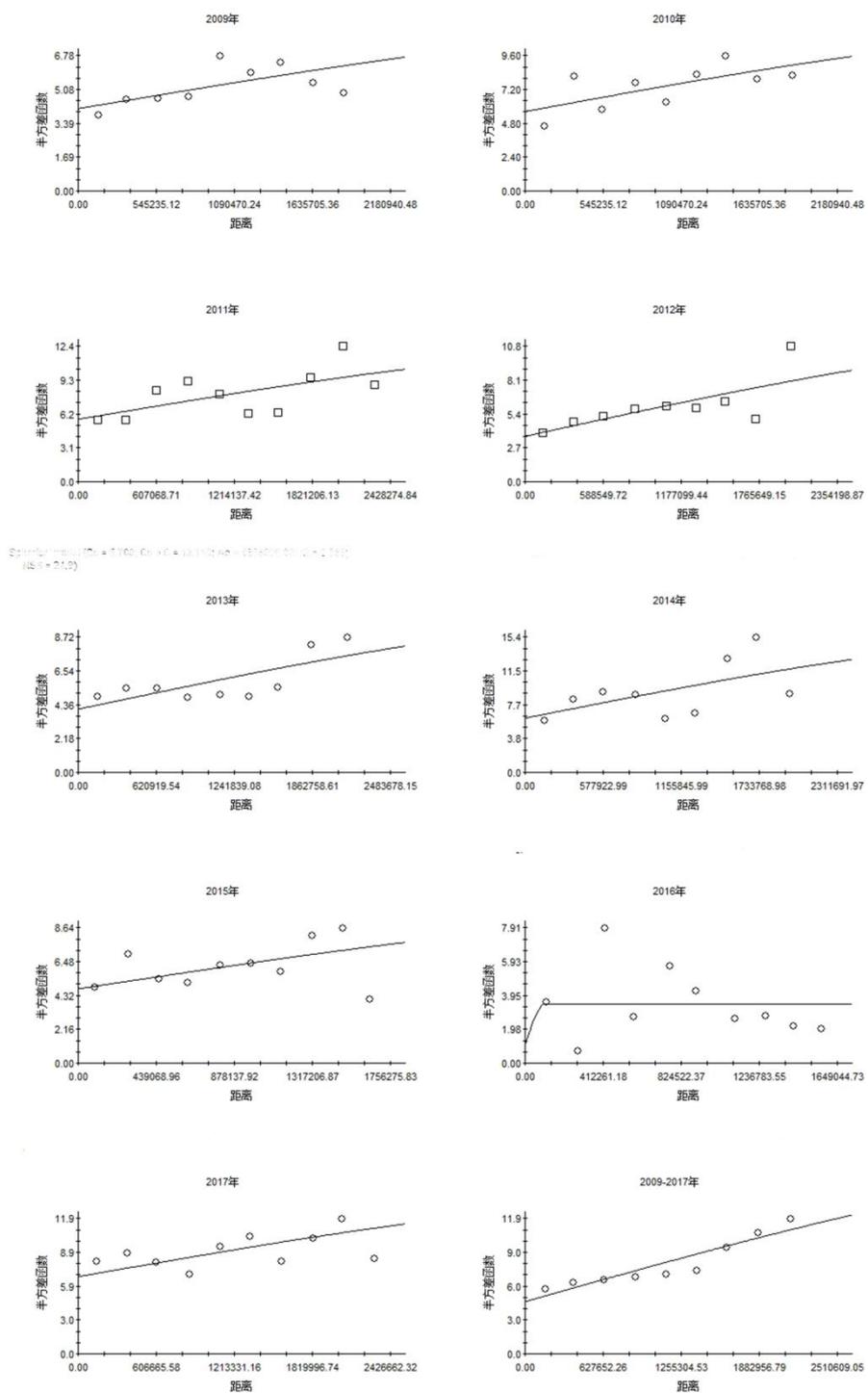


图 2.3 2009-2017 年花生 AFB₁ 含量平均值的变异函数

Figure 2.3 Variograms of mean AFB₁ content in peanut during 2009-2017

2.4.2 中国花生 AFB₁ 污染分布地图

中国花生 AFB₁ 的克里金插值图表明 AFB₁ 平均污染风险呈现离散分布的特点，在长江流域以南出现高值区分布，且存在年度间差异（见图 2.4）。从克里金插值结果看出中国花生 AFB₁ 污染呈现相近的模式，具体来说在 2009-2017 年，长江流域主产区的湖南省、江西省和安徽省以及南方主产区的广西和广东省呈现较为严重的污染程度；而东北主产区的 AFB₁ 污染程度一直都很轻。克里金插值图也会显示出一些 AFB₁ 检测结果不容易发现的细节，在长江流域主产区的湖南省，2009-2010 和 2012-2013 四年都显示出较高的 AFB₁ 污染水平，2012 年南方主产区的福建省显示出较高的 AFB₁ 污染水平。中国花生 AFB₁ 的时空分布差异表明黄曲霉毒素污染受多种因素的显著影响，例如 2009-2014 年长江流域主产区的花生 AFB₁ 污染最严重，但 2015 年最严重地区为南方主产区，这种状况与降水量和高温天气的转移相关。其它因素可以解释花生 AFB₁ 污染的其余变异度，例如不同的土壤类型、不同的花生品种、不同类型的产毒菌种、产毒菌种不同的产毒力（Battilani *et al.* 2006）以及不同的播种和收获时间等。

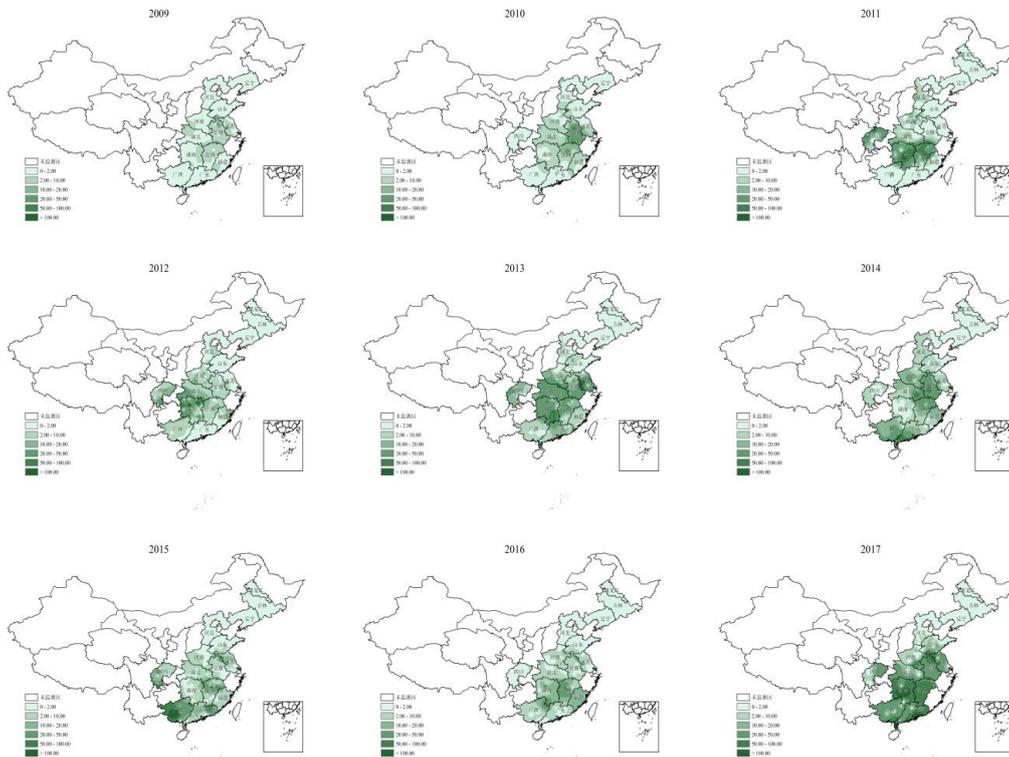


图 2.4 2009-2017 年中国花生 AFB₁ 平均值的克里金插值图

Figure 2.4 Kriging interpolation map of mean AFB₁ content in peanuts from China

2.5 讨论

目前，很多学者利用地统计和 GIS 相结合的方法来研究不同对象的空间变异情况，通过对多年的具有空间自相关的变量进行克里金插值，可以确定在区域范围内的多年农业问题（Nelson *et al.* 1999; Jaime-Garcia *et al.* 2001; Nelson *et al.* 1994）。Jaime-Garcia 等（2003）研究了 1997-2001

年在德克萨斯州南部抽取的棉籽中黄曲霉毒素的平均值与含量大于 20 $\mu\text{g}/\text{kg}$ 的样品百分比的空间分布, 发现其空间分布规律具有重复性, 且确定了重点监测区域。克里金插值图表明黄曲霉毒素污染呈片状分布, 且具有年度间的差异性, 所有的半变异函数为球状模型。Casado 等 (2009) 研究了小麦中脱氧雪腐镰刀烯醇 (Deoxynivalenol, DON) 和赭曲霉毒素 A (Ochratoxin A, OTA) 的空间分布特征, 结果表明 DON 呈空间连续分布, OTA 呈随机分布; 且抽样时除了需要考虑取样点的数量外, 还需要考虑取样点的位置, 以便得到可靠的污染水平。Casado 等 (2010) 利用地统计学方法研究了储藏玉米中伏马毒素 B₁ (Fumonisin B₁, FB₁) 和伏马毒素 B₂ (Fumonisin B₂, FB₂) 的空间分布, 结果表明两种伏马毒素之间没有空间相关性。Kerry 等 (2017) 利用 GIS 与地统计结合的方法研究了美国佐治亚州南部玉米黄曲霉毒素污染的影响因素。Battilani 等 (2006) 应用地统计学方法对欧洲南部和以色列的葡萄黑曲霉 (*Aspergillus Section Nigri*) 和炭疽病菌 (*A. carbonarius*) 的早熟期发病情况进行了空间分析。研究发现黑曲霉的空间变异性与经度和纬度显著相关, 且具有年度间的差异, 收获期的高温干旱条件显著影响黑曲霉的分布。本章中 AFB₁ 平均值的年度间差异主要是受花生品种、土壤、气候等条件的影响。这些研究表明, 根据地统计学方法与风险因子结合生成的插值图, 可以有效确定高风险区域, 为风险预警提供基础支撑。然而目前我国在此方面的研究多针对于土壤性质 (养分、重金属等)、天气变化 (降水量) 及大气污染物 (PM_{2.5}) 等, 对黄曲霉毒素空间变异性的研究还未见报道。

2.6 本章小结

中国花生 AFB₁ 呈现显著的时间和空间差异, 尽管长江流域主产区和南方主产区黄曲霉毒素污染较严重, 但是大部分地区的花生黄曲霉毒素污染水平都很低。由地统计分析结果可知花生 AFB₁ 平均值的 $C/(C+C_0)$ 值在 50-70% 之间, 表明具有中等程度的空间自相关性, 是由结构性因素 (如土壤类型、土壤母质、气候、地形地貌等) 和随机性因素 (如花生品种、农业土地利用方式、耕作管理措施、施肥及取样分析误差等) 共同作用, 且受气候条件影响作用相对较大。2009-2017 年花生 AFB₁ 半变异模型为球状模型, 2017 年的 AFB₁ 含量空间变异性最大 (基台值为 6.77), 2016 年 AFB₁ 含量空间变异性最小 (基台值为 1.05)。且模型中块金值均小于基台值, 表明 AFB₁ 含量受县市间距离的强烈影响。2016 年的变程最小为 107000m, 而 2017 年的变程最大为 511000m。最短变程可作为研究区域内监测站间隔的布设依据, 由于采样点的最佳布点距离应该为最短变程值的 1/4 至 1/2, 因此布点间隔应考虑设置为 26750-53500 m。

2009-2017 年中国花生 AFB₁ 平均值的克里金插值图表明 AFB₁ 污染风险呈现离散分布的特点, 东北主产区的黄曲霉毒素污染风险低, 而长江流域主产区和南方主产区的黄曲霉毒素污染风险较高。这与不同地区的花生品种、土壤类型、菌种产毒力、气候条件等因素相关。确定常年具有黄曲霉毒素污染高风险地区为重点防控地区来实施防控措施, 例如与不易污染黄曲霉毒素的作物进行轮作, 种植黄曲霉毒素高抗品种, 及时收获与灌溉, 利用不产毒菌株进行生物防控等。

第三章 中国花生黄曲霉毒素污染消长变化趋势

3.1 引言

花生是最易受产毒真菌侵染并产生黄曲霉毒素的农产品之一(Horn 2005; Monyo *et al.* 2012; Murphy *et al.* 2006; Ma *et al.* 2014)。黄曲霉毒素污染受多种因素的影响,由于透气性及理化性质的差异,不同质地的土壤会影响黄曲霉孢子的增殖。花生品种抗性会影响黄曲霉的侵染及黄曲霉毒素的产生,同时抗虫性与抗旱性也会间接影响黄曲霉毒素的污染水平。产毒真菌定殖于花生并产生黄曲霉毒素,污染状况受到真菌产毒力的影响。黄曲霉毒素通常由曲霉属黄绿组(*Aspergillus* section *Flavi*)的菌株产生(Matthias *et al.* 2009)。其中最常见产毒菌种为黄曲霉(*A. flavus*)、寄生曲霉(*A. parasiticus*)和集峰曲霉(*A. nomius*),而集峰曲霉(*A. nomius*)相对较为罕见(Houshyarfard *et al.* 2014; Ehrlich *et al.* 2007)。黄绿组的菌种在产毒种类、产毒力、菌核产生、菌核形态和孢子形成等方面存在很大的差异。一般来说,黄曲霉(*A. flavus*)只产生B族黄曲霉毒素,寄生曲霉(*A. parasiticus*)和集峰曲霉(*A. nomius*)既可以产生B族黄曲霉毒素又可以产生G族黄曲霉毒素。而且几乎全部的寄生曲霉(*A. parasiticus*)均有较高的产毒力(Astoreca *et al.* 2011; Donner *et al.* 2015)。黄曲霉(*A. flavus*)不产生G族黄曲霉毒素(AFG₁和AFG₂),是因为黄曲霉毒素合成基因中有一个1.5-2.2 kb的缺失,导致G族黄曲霉毒素合成必需的基因 *cypA* 的缺失(Ehrlich *et al.* 2004)。此外, S_{BG} 具有类似黄曲霉(*A. flavus*)的菌核形态,但是却可以同时产生B族和G族黄曲霉毒素。有研究指出, S_{BG} 在系统树上属于黄曲霉(*A. flavus*)和寄生曲霉(*A. parasiticus*)的分支(Torres *et al.* 2006)。但是 S_{BG} 的确切分类归属尚不明确(Matthias *et al.* 2009)。研究表明,在较低的温度条件下(一定范围内),G族黄曲霉毒素产生或累积较多,而在较高的温度条件下,G族黄曲霉毒素代谢较快(Schoroeder *et al.* 1967)。而我国花生根际土壤只分离得到黄曲霉(*A. flavus*)和寄生曲霉(*A. parasiticus*),且多数为黄曲霉(*A. flavus*)(朱婷婷 2018; 张初署 2013)。

气温和降水量(或湿度)对黄曲霉毒素污染有显著影响(Wu *et al.* 2016)。据报道气候变化将会加重作物受黄曲霉(*A. flavus*)的侵染和黄曲霉毒素的污染(Medina *et al.* 2014)。许多研究表明,收获花生样品中自然产生的黄曲霉毒素受气候条件的影响,在花生生长后期的高温干旱会加速花生黄曲霉毒素的产生(Hill *et al.* 1983)。在干旱条件下,当根际土壤温度为29.6-31.3℃时,花生污染黄曲霉毒素的风险较高(Cole *et al.* 1985)。Williams等(1983)发现花生生长后期的平均温度为25.7到27℃时,可以产生黄曲霉毒素。Klich(2007)研究表明产毒真菌存在于广泛区域,但在纬度16°到35°的温暖区域较多,高于45°则较少。中国作为花生生产大国,且气候类型、花生品种、土壤类型多样,摸清中国花生黄曲霉毒素污染水平及其发生规律,对掌握其污染分布、建立黄曲霉毒素预警模型和实施精准防控具有非常重要的意义。收获前的高温干旱对增加产毒真菌的分布及黄曲霉毒素污染在多国家、多作物中已有报道,例如玉米(Pleadin *et al.* 2014; Abbas *et al.* 2002; Pleadin *et al.* 2015; Kos *et al.* 2013),花生(Waliyar *et al.* 2015; Rachaputi *et al.* 2002; Bircan 2009)中污染风险会随着降雨量的减少而增加。为了阐明由于干旱胁迫和高温引起的真菌及黄曲霉毒素污染高风险的机制,许多学者进行了一系列研究。有一种说法是水分胁迫诱导游离氨基酸如脯氨酸的形成,而这种氨基酸会在持续干旱条件下刺激黄曲霉和寄生曲霉产生黄曲霉毒素。

这些氨基酸可能作为产毒真菌的碳源，用来合成黄曲霉毒素。还有一些研究者认为，在干旱胁迫的条件下，植物对黄曲霉毒素产生的内在机制，如植物抗毒素，会丧失活性。收获前花生中植物抗毒素的产生抑制了真菌的生长和黄曲霉毒素的积累，都其高与水分活度密切相关。长期干旱使花生的 a_w 降低到不能产生植物抗毒素的水平，从而增加黄曲霉毒素的产生。此外，由于黄曲霉具有耐干性，可以在相当高的温度（25-42℃）和低的水分（-2 到 32MPa）下生长，比其它真菌甚至其它曲霉属菌种都具有竞争优势（Bircan *et al.* 2008）。因此，气候条件尤其是温度和降雨量会影响花生黄曲霉毒素的污染水平。

本章在实验室前期对花生黄曲霉毒素监测的基础上，对黄曲霉毒素污染的消长变化趋势进行分析，并解析了花生品种、土壤类型、真菌产毒力及气候条件对其消长变化的影响。

3.2 材料

3.2.1 花生样品

同 2.2.1。

3.2.2 仪器与试剂

同 2.2.2。

3.2.3 花生品种抗性数据

花生品种的抗性数据来自国家花生数据中心（<http://www.peanutdata.cn>）。根据花生品种对逆境和病虫害的抗性，本章将采集的花生品种对黄曲霉毒素的抗性分为高、中、低三类。

3.2.4 土壤类型数据

花生主产县的土壤类型信息来自中国土壤数据库（<http://vdb3.soil.csdb.cn>），根据不同地点 0-20cm 发生层质地，将其分为砂土、壤土和粘土三类。

3.2.5 菌株产毒力数据

在本实验室前期的工作基础上（朱婷婷 2018），将花生根际土壤中菌株（黄曲霉与寄生曲霉）的产毒力（ $\mu\text{g/kg AFB}_1/\text{g 土壤}$ ）分为高、中、低三类。其中高产毒力菌株可产生 $>100000 \mu\text{g/kg AFB}_1/\text{g 土壤}$ ，中产毒力菌株可产生 $10000-100000 \mu\text{g/kg AFB}_1/\text{g 土壤}$ ，低产毒力菌株可产生 $<10000 \mu\text{g/kg AFB}_1/\text{g 土壤}$ 。

3.2.6 气候数据

2009-2017 年全国 140 多个气象站收获前降水量、日平均气温、日最高气温与日最低气温数据来源于中国气象数据网（<http://data.cma.cn/site/index.html>）—中国地面气候资料日值数据集（V3.0）。

3.3 方法

3.3.1 花生黄曲霉毒素检测

同 2.3.1。

3.3.2 方差分析

采用 SPSS 统计分析软件 (IBM SPSS Statistics Version 20, IBM Corporation, Armonk, NY, USA) 的单因素方差分析 (One-Way ANOVA) 来判断差异显著性, 当 $P < 0.05$ 时为差异显著。本章中的平均值均指算术平均值。

3.3.3 关键时期的确定

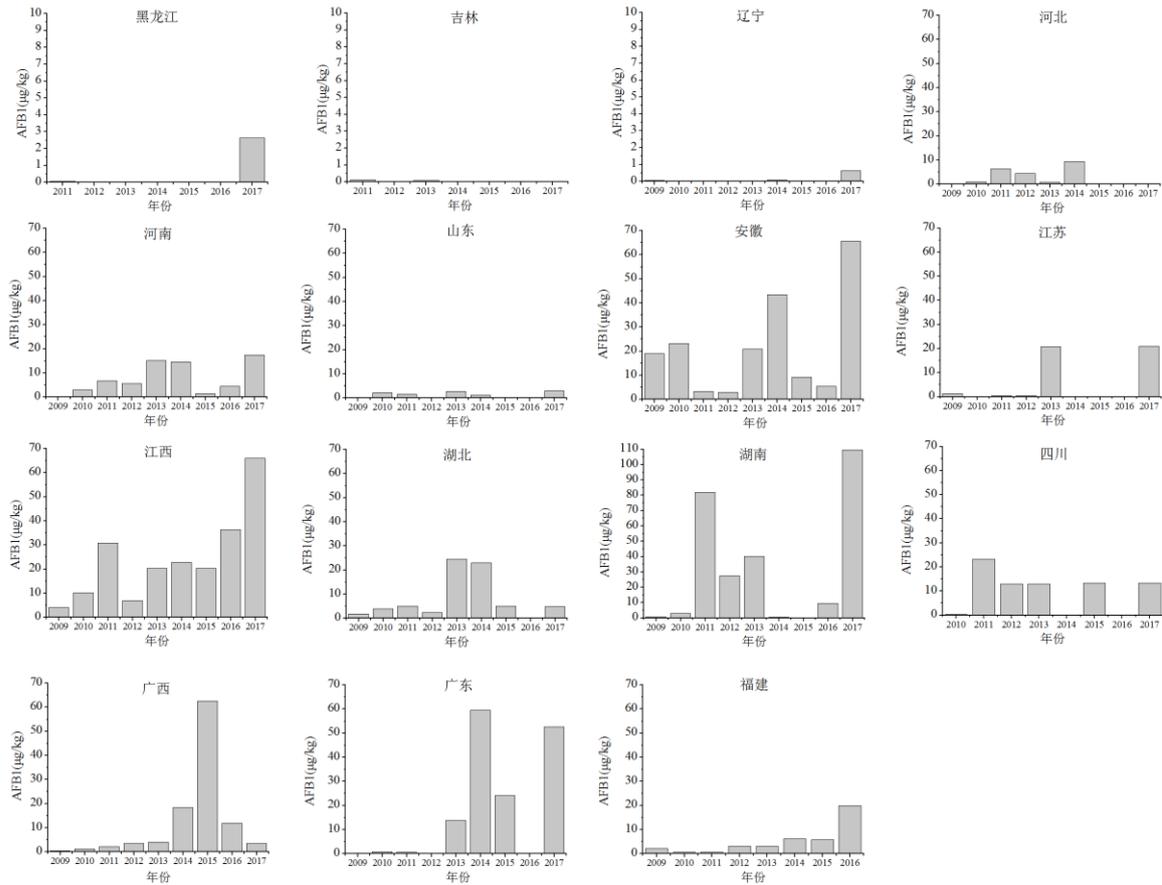
利用 SPSS 软件对 AFB_1 含量平均值与收获前 1 个月、2 个月和 3 个月降水量和日平均气温的平均值进行多元非线性回归。

3.4 结果与分析

3.4.1 中国花生黄曲霉毒素污染动态变化

从 2009 到 2017 年, 自 15 个花生主产省、337 个花生主产市县共采集花生样品 12980 份, 根据种植面积的不同, 利用五点取样法在每个市县取 10-30 份样品。采样点覆盖纬度为 20-46°N, 包括不同的气候类型, 例如温带气候和亚热带气候。2009-2017 年不同省份花生黄曲霉毒素污染水平及其消长变化趋势如图 3.1 所示。结果表明, 在 12980 份花生样品中大部分花生黄曲霉毒素含量低于检出限, 污染样品占总样品量的 20.82%, 在总样品中, 花生黄曲霉毒素总量 ($AFB_1+AFB_2+AFG_1+AFG_2$) 的平均值±标准差 (Standard deviation, SD) 为 $11.54 \pm 67.18 \mu\text{g/kg}$, 最大检出值为 $978.07 \mu\text{g/kg}$ 。根据《GB 2761-2017 食品安全国家标准食品中真菌毒素限量》规定, 花生 AFB_1 含量的最大残留限量 (MRL) 为 $20 \mu\text{g/kg}$ (GB 2761-2011), 则超标样品仅占总样品量的 5.41%。

利用方差分析可知中国花生黄曲霉毒素年份与地区间差异显著 ($P < 0.05$)。除 2016 年只检出 B 族黄曲霉毒素以外, 四种黄曲霉毒素在其他年份均有检出。2017、2014、2015 和 2013 年黄曲霉毒素污染水平最高, 总量分别为 $26.67 \pm 116.53 \mu\text{g/kg}$ 、 $17.49 \pm 72.01 \mu\text{g/kg}$ 、 $15.33 \pm 87.19 \mu\text{g/kg}$ 和 $13.95 \pm 52.65 \mu\text{g/kg}$ 。2009 年和 2010 年黄曲霉毒素污染水平最低, 总量分别为 $2.40 \pm 32.48 \mu\text{g/kg}$ 和 $4.32 \pm 35.15 \mu\text{g/kg}$ 。而 2009 年和 2011 年黄曲霉毒素的检出率最高, 均大于 30%。2014 年和 2017 年的超标率最高, 分别为 9.36% 和 8.61%。就不同主产省来说, 黑龙江省只检出 AFB_1 , 吉林省只检出 AFB_1 、 AFB_2 和 AFG_1 , 其他省份四种黄曲霉毒素均有检出。黄曲霉毒素在长江流域主产区的湖南省、江西省、安徽省、四川省和南方主产区的广东省、广西最为严重。2009-2017 年间, 这些地区的黄曲霉毒素总量均接近或高于 $20 \mu\text{g/kg}$ 。北方主产区 (包括河南省、河北省和山东省) 的黄曲霉毒素污染较轻, 总量范围为 $2.22-9.11 \mu\text{g/kg}$ 。包括黑龙江省、吉林省和辽宁省在内的东北主产区污染最轻, 黄曲霉毒素总量均不超过 $1.00 \mu\text{g/kg}$ 。

图 3.1 2009-2017 年中国花生主产省 AFB₁ 消长变化趋势Figure 3.1 Variation trend of AFB₁ content in different peanut production provinces of China

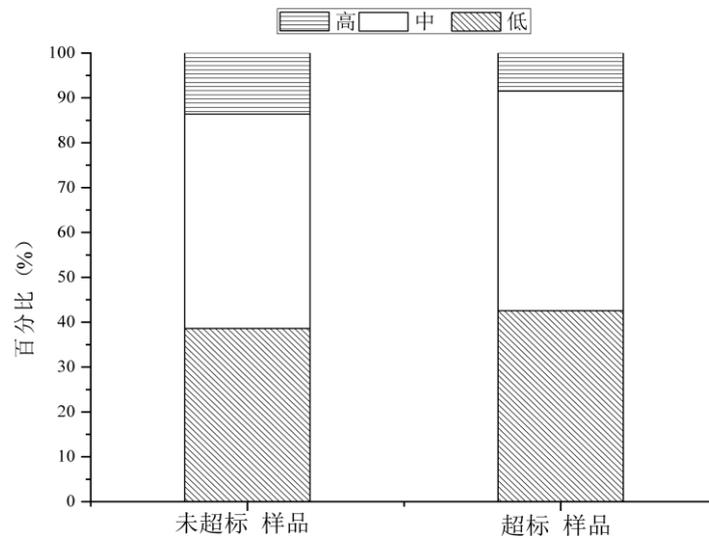
3.4.2 影响黄曲霉毒素污染动态变化的主要因素

造成中国花生黄曲霉毒素污染程度呈现年度与地区间差异的因素主要包括花生品种抗性、土壤类型、菌株产毒力及气候条件。具体表现为：

3.4.2.1 花生品种抗性

中国花生品种多样，本文涉及 300 多个花生品种，包括白沙、花育、鲁花、丰花、海花、冀花、远杂、豫花、天府、中花、湘花、粤油、湛油等将近 30 个系列。其对黄曲霉毒素的抗性目前无官方资料可查，而抗旱、抗病虫害的特性会间接影响花生品种对黄曲霉毒素的抗性。根据花生对逆境和病虫害的抗性文献查阅，将采集的花生品种对黄曲霉毒素的抗性分为高、中、低三类。

根据超标与未超标样品中对黄曲霉毒素具有不同抗性的花生品种占比（图 3.2）可知，未超标样品中黄曲霉毒素高抗品种占比较高（13.61%），约为超标样品的 1.5 倍（8.51%）；中抗品种占比相当，在未超标样品与超标样品中占比分别为 47.77% 和 48.94%；低抗品种在超标样品（42.55%）中占比稍高于未超标样品（38.61%）。此外，湖南、福建、四川等地多使用自留种，而品种的退化有可能会造成抗性的降低，从而导致这些地区黄曲霉毒素污染的持续加重。因此结果表明花生品种对黄曲霉毒素的抗性影响了花生黄曲霉毒素的污染水平。

图 3.2 AFB₁ 超标与未超标样品中不同抗性品种占比Figure 3.2 Proportion of varieties with different resistance to AFB₁ in compliant and non-compliant samples

3.4.2.2 土壤类型

根据不同花生主产区 0-20cm 发生层质地，将土壤类型其分为砂土、壤土和粘土三类。根据超标与未超标样品中不同土壤类型占比（图 3.3）可知，超标样品的土壤类型中砂土占比较高（4.33%），而未超标样品仅为（0.81%）；超标样品与未超标样品中壤土占比相当，分别为 61.94% 和 62.90%；粘土类型在未超标样品中占比（36.29%）高于超标样品（33.72%）。由于砂土透气性好，孢子传播速度快（王后苗等 2012），所以会加重花生黄曲霉毒素的污染，而粘土正好相反。因此结果表明土壤类型会影响花生黄曲霉毒素的污染水平。

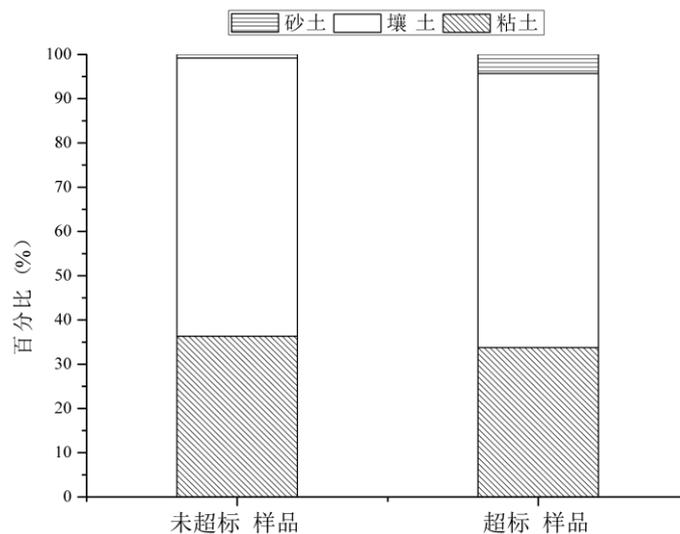
图 3.3 AFB₁ 超标与未超标样品的不同土壤类型占比

Figure 3.3 Proportion of different soil types in compliant and non-compliant samples

3.4.2.3 菌株产毒力

根据超标与未超标样品中不同产毒力菌株（黄曲霉与寄生曲霉）的占比（图 3.4）可知，未超标样品中低产毒力菌株占比（33.33%）高于超标样品（23.08%）；中、高产毒力菌株均在超标样品中占比较多，分别为 53.85% 和 23.08%，在未超标样品中占比分别为 50.00% 和 16.67%。表明根际土壤菌株产毒力会影响了花生黄曲霉毒素的污染水平。研究表明长江流域主产区的土壤中具有最多的黄曲霉数量和最高的 AFB₁ 产毒力，其次是南方主产区，而东北主产区最少（朱婷婷 2018）。此外，江西樟树、江西丰城、安徽固镇均是高产毒菌株分布地区，从而也造成这些地区黄曲霉毒素污染较为严重。

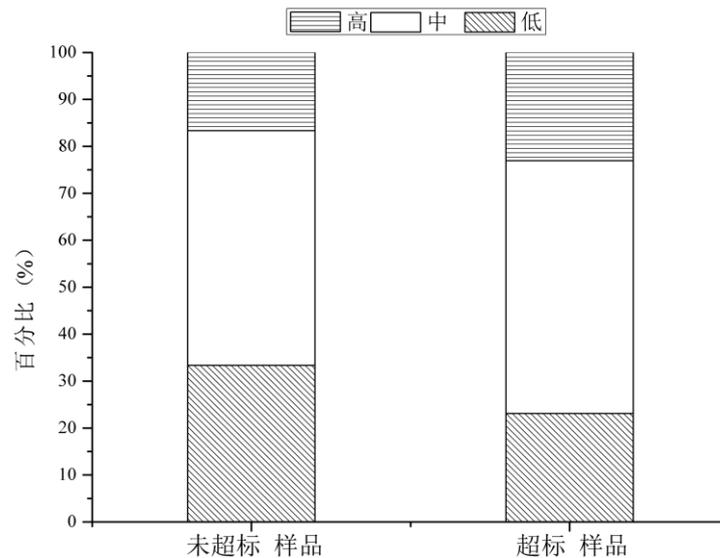


图 3.4 AFB₁ 超标与未超标样品的不同产毒力菌株占比

Figure 3.4 Proportion of fungal strains with different aflatoxin producing potential in compliant and non-compliant samples

3.4.2.4 气候条件

气温和降水量（或湿度）对黄曲霉毒素污染有显著影响。通过确定其作用于黄曲霉毒素的关键时期，即可分析关键时期的气候条件对花生黄曲霉毒素污染的影响。

(1) 黄曲霉毒素污染关键时期的确定

由于收获前 4-6 周的高温（25.7-31.3 °C）和干旱通常会引起花生黄曲霉毒素污染的发生（Wu *et al.* 2016），对 AFB₁ 含量平均值与干旱条件下收获前 1 个月、2 个月和 3 个月的降雨量和日平均气温进行多元非线性回归拟合，以湖南省为例，拟合后的决定系数（R²）结果如图 3.5 所示。结果表明收获前 1 个月（7 月）的平均气候与 AFB₁ 含量平均值拟合最好，即收获前 1 个月的气候条件对 AFB₁ 污染水平的影响最大。

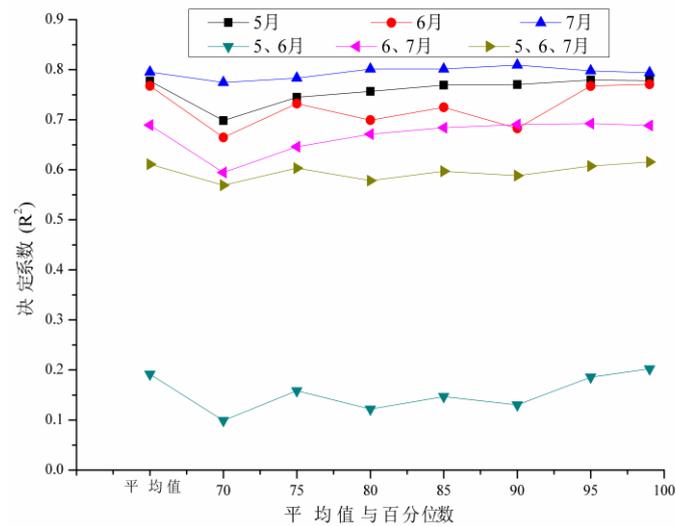


图 3.5 不同时期平均气候条件与 AFB₁ 含量拟合的决定系数

Figure 3.5 Multiple regression determination coefficient (R^2) fitted by the AFB₁ content with average climatic conditions in different periods of time (Hunan province)

(2) 气候条件对黄曲霉毒素污染的影响

根据 2009-2017 年花生主产区降水量 (图 3.6) 与日平均气温 (图 3.7) 可以看出, AFB₁ 污染严重区域基本都处于高温干旱区域。2013、2015 与 2017 年的高温干旱天气造成花生受 AFB₁ 污染较为严重; 2009 年与 2012 年日平均气温较低 (<30℃)、2009 年与 2010 年雨水充沛, 因此 2009、2010 及 2012 年 AFB₁ 污染程度较低; 2016 年虽日平均气温较高, 但降水较多, 使得 AFB₁ 污染并不严重。我国长江流域主产区及南方主产区花生收获前正值高温天气, 如遇干旱则会加重黄曲霉毒素的污染, 而东北主产区的常年低温天气不利于产毒真菌的生长以及黄曲霉毒素的产生, 从而减轻了黄曲霉毒素的污染。

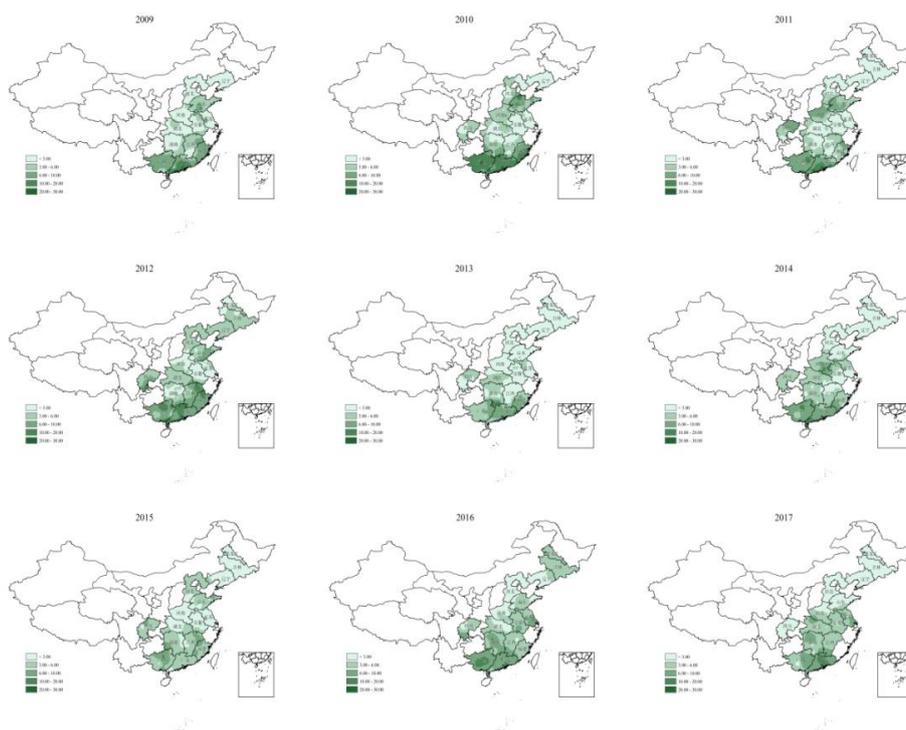


图 3.6 2009-2017 年中国花生主产区降水量（收获前 30 天平均值）克里金插值图

Figure 3.6 Kriging interpolation map of precipitation (mean value of 30d before harvest) in peanut production areas of China

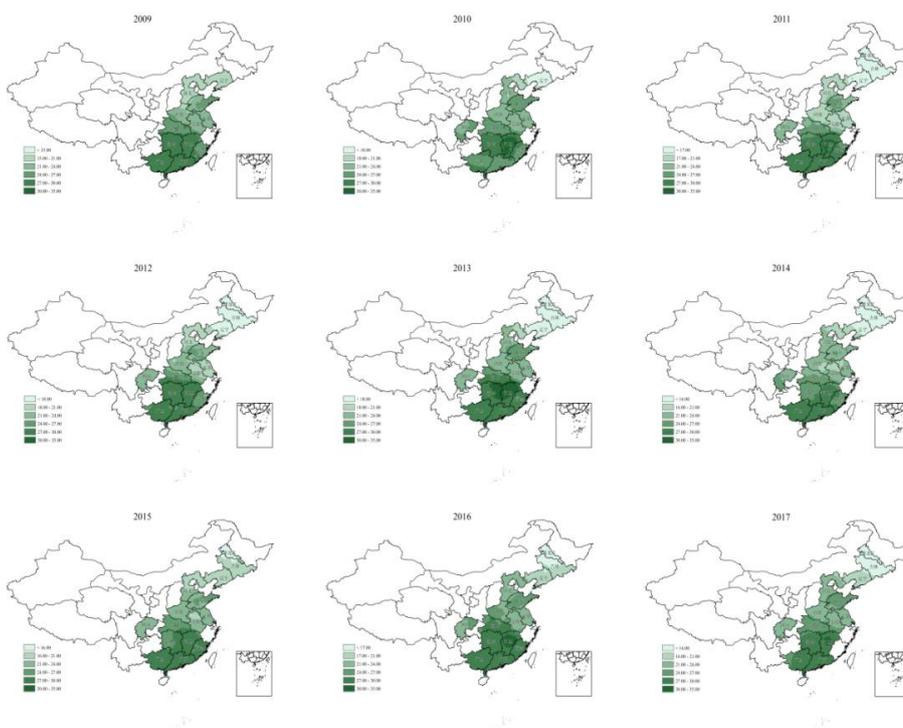


图 3.7 2009-2017 年中国花生主产区日平均气温（收获前 30 天平均值）克里金插值图

Figure 3.7 Kriging interpolation map of average air temperature (mean value of 30d before harvest) in peanut production areas of China

(3) 不同主产区气候条件对黄曲霉毒素污染的影响

将 2009-2017 连续九年北方主产区、长江流域主产区和南方主产区收获前 30 天降水量、日平均气温、日最高气温和日最低气温的平均值与花生 AFB₁ 的平均含量进行关联, 结果如图 3.8 所示。

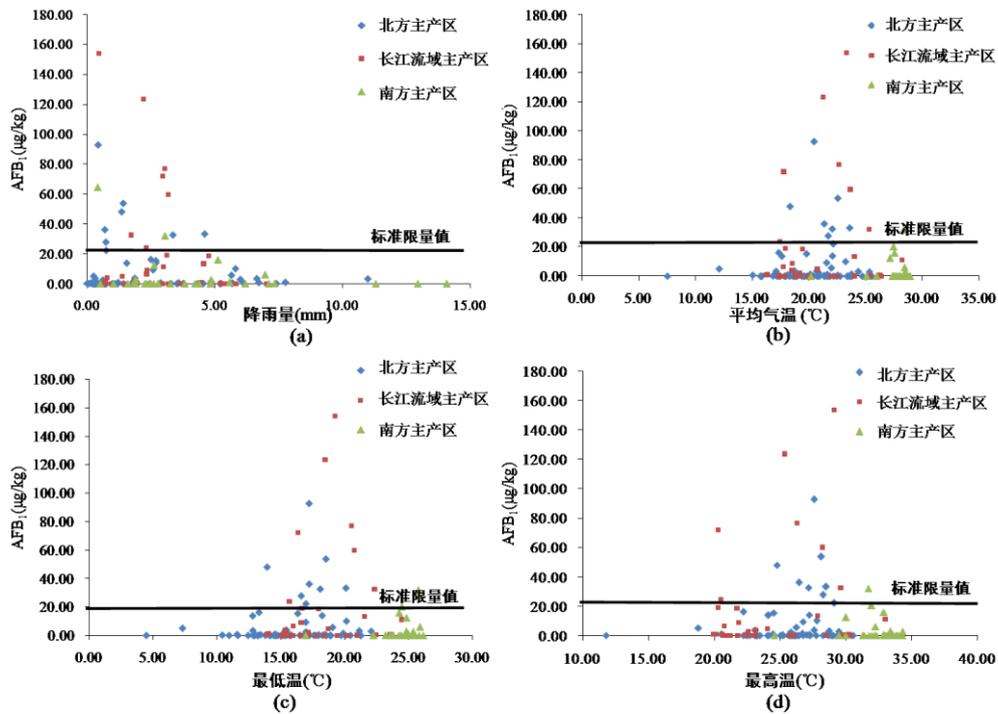


图 3.8 (a) 三大主产区花生 AFB₁ 含量与降水量 (收获前 1 个月平均值) 的关系; (b) 三大主产区花生 AFB₁ 含量与日平均气温的关系; (c) 三大主产区花生 AFB₁ 含量与日最低气温 (收获前 1 个月平均值) 的关系; (d) 三大主产区花生 AFB₁ 含量与日最高气温 (收获前 1 个月平均值) 的关系

Figure 3.8 (a) AFB₁ content in peanut varied with the precipitation (the average value of one month before harvest) in 3 main production areas; (b) AFB₁ content in peanut varied with the mean air temperature (the average of one month before harvest) in 3 main production areas; (c) AFB₁ content in peanut varied with the minimum air temperature (the average of one month before harvest) in 3 main production areas; (d) AFB₁ content in peanut varied with the maximum air temperature (the average of one month before harvest) in 3 main production areas

图3.8 (a) 显示了2009-2017年中国不同的花生主产县(区、市)花生AFB₁含量与降水量(收获前1个月平均值)之间的关系, 结果表明AFB₁含量随降水量的减少而增多, 当降水量低于4mm时, 超过最大残留限量20 µg/kg的可能性较大。

图 3.8 (b) 显示了 2009-2017 年中国不同的花生主产县(区、市)花生 AFB₁ 含量与日平均气温(收获前 1 个月平均值)之间的关系, 结果表明在北方主产区, 当日平均气温接近 20 °C 时, AFB₁ 含量增加; 在长江流域主产区, 当日平均气温接近 21 °C 时, AFB₁ 含量增加; 在南方主产区, 当日平均气温接近 24 °C 时, AFB₁ 含量增加。在干旱条件下, 当日平均气温分别为 15-25 °C, 17-25 °C 和 21-28 °C 时, 北方主产区、长江流域主产区和南方主产区超过最大残留限量 20 µg/kg 的可能性较大。

图 3.8 (c) 显示了 2009-2017 年中国不同的花生主产区 (区、市) 花生 AFB₁ 含量与日最低气温 (收获前 1 个月平均值) 之间的关系, 结果表明在北方主产区, 当日最低气温接近 17 °C 时, AFB₁ 含量增加; 在长江流域主产区, 当日最低气温接近 19 °C 时, AFB₁ 含量增加; 在南方主产区, 当日最低气温接近 21 °C 时, AFB₁ 含量增加。在干旱条件下, 当日最低气温分别为 13-21 °C, 15-24 °C 和 17-26 °C 时, 北方主产区、长江流域主产区和南方主产区超过最大残留限量 20 µg/kg 的可能性较大。

图 3.8 (d) 显示了 2009-2017 年中国不同的花生主产区 (区、市) 花生 AFB₁ 含量与日最高气温 (收获前 1 个月平均值) 之间的关系, 结果表明在北方主产区, 当日最高气温接近 27 °C 时, AFB₁ 含量增加; 在长江流域主产区, 当日最高气温接近 31 °C 时, AFB₁ 含量增加; 在南方主产区, 当日最高气温接近 32 °C 时, AFB₁ 含量增加。在干旱条件下, 当日最高气温分别为 22-29 °C, 20-30 °C 和 25-33 °C 时, 北方主产区、长江流域主产区和南方主产区超过最大残留限量 20 µg/kg 的可能性较大。

总体来说, 随着降水量的减少, 日平均气温接近 23 °C, 日最低气温接近 20 °C, 日最高气温接近 29 °C 时, AFB₁ 的含量升高。在干旱条件下, 当日平均气温范围为 15-28 °C, 日最低气温范围为 13-26 °C, 日最高气温范围为 20-33 °C 时, 花生 AFB₁ 超标的风险较大。结果表明花生黄曲霉毒素污染水平与气候条件密切相关。

3.5 讨论

花生黄曲霉毒素污染问题在全世界多地均有发生。在马来西亚, Leong 等 (2010) 发现 14 份带壳花生样品中, 42.9% 污染有黄曲霉毒素, AFB₁ 的含量值范围为 15.5-618 µg/kg。在埃塞俄比亚东部, 93/100 的花生样品检出黄曲霉毒素, 阳性样品中总量的范围为 15-11900 µg/kg (Chala *et al.* 2013)。喀麦隆黄曲霉毒素污染较为严重, 97% 的花生样品污染有黄曲霉毒素, 检出样品的平均值为 47 µg/kg, 最大值高达 210 µg/kg (Abia *et al.* 2013)。在巴西, 29/359 的花生样品污染有黄曲霉毒素, 其含量范围为 7.9-1496 µg/kg (Andrade *et al.* 2013)。在津巴布韦, 18 份花生样品中有 3 份检出黄曲霉毒素, 含量范围为 6.6-622 µg/kg (Mupunga *et al.* 2014)。在伊拉克, 5 份花生样品全部检出黄曲霉毒素, 含量高达 782 µg/kg (Thalij 2015)。在尼日利亚, 花生黄曲霉毒素的检出率约为 39.29%, 阳性样品中黄曲霉毒素总量含量范围和平均值分别为 0.4-2076 µg/kg 和 216.1±496.7 µg/kg (Oyedele *et al.* 2017)。相比以上研究, 尽管本研究中有 5.41% 的样品 AFB₁ 含量高于 20 µg/kg, 但是总体污染水平低于其他国家, 尤其是低于非洲国家。这与各个国家地区环境、气候、土壤、花生品种和农艺措施等多种因素相关。为了降低黄曲霉毒素污染水平, 农民应实施良好的田间操作, 比如合理轮作、正确使用化肥、合理灌溉、防治害虫、及时收获和干燥、种植抗性品种以及使用杀菌剂或生物拮抗剂等以减少产毒真菌的侵染及黄曲霉毒素的产生。

鉴于中国花生种植面积广, 花生品种、土壤类型、气候类型等多种多样, 花生黄曲霉毒素问题已经成为一个值得关注的问题。本章探讨了不同花生品种抗性、土壤类型、菌株产毒力及气候条件与花生黄曲霉毒素污染的关系, 并探明了中国不同花生主产区黄曲霉毒素含量与收获前 1 个月日平均气温、日最低气温、日最高气温和降水量之间的关系, 得到了不同花生主产区最易受黄曲霉毒素污染的温度范围, 但是其似乎低于其他文献报道的数据, 原因是本章中黄曲霉毒素是自

然发生的，也就是说影响因素不是实验室的可控条件，这反映了中国花生在自然条件下产生黄曲霉毒素的含量与收获前 1 个月气候条件的真正关系，了解这其中的规律对于采取产前防控措施有着非常重要的意义。

3.6 本章小结

本章首先分析了 2009-2017 连续九年花生黄曲霉毒素污染消长变化趋势，结果表明黄曲霉毒素污染有年度间和地区间差异，2017、2014、2015 和 2013 年黄曲霉毒素污染水高于其它年份，2009 年和 2010 年黄曲霉毒素污染水平最低。长江流域主产区的花生黄曲霉毒素污染水平最高，南方主产区和北方主产区污染水平较高，东北主产区的污染水平最低。其次阐述了不同花生品种抗性、土壤类型、菌株产毒力及气候条件对黄曲霉毒素污染消长变化的影响，结果表明黄曲霉毒素低抗品种、砂土土质、高产毒力菌株在超标样品中占比较大，高温干旱区域与花生污染严重区域重合，验证了其会加重花生黄曲霉毒素污染。最后详述了气候条件与花生黄曲霉毒素污染之间的关系，确定了收获前 1 个月的气候对 AFB₁ 污染水平的影响最大。当降水量低于 4mm 时，超过最大残留限量 20 μg/kg 的可能性较大。在干旱条件下，当日平均气温分别接近 20℃、21℃ 和 24℃ 时，北方主产区、长江流域主产区和南方主产区的花生 AFB₁ 含量增加，当日平均气温分别为 15-25℃，17-25℃ 和 21-28℃ 时，北方主产区、长江流域主产区和南方主产区超过最大残留限量 20 μg/kg 的可能性较大；当日最低气温分别接近 17℃、19℃ 和 21℃ 时，北方主产区、长江流域主产区和南方主产区的花生 AFB₁ 含量增加，在干旱条件下，当日最低气温分别为 13-21℃，15-24℃ 和 17-26℃ 时，北方主产区、长江流域主产区和南方主产区超过最大残留限量 20 μg/kg 的可能性较大；当日最高气温分别接近 27℃、31℃ 和 32℃ 时，北方主产区、长江流域主产区和南方主产区的花生 AFB₁ 含量增加，在干旱条件下，当日最高气温分别为 22-29℃，20-30℃ 和 25-33℃ 时，北方主产区、长江流域主产区和南方主产区超过最大残留限量 20 μg/kg 的可能性较大。

本章基于我国产后花生黄曲霉毒素污染的消长变化趋势以及不同影响因素（花生品种抗性、土壤类型、菌株产毒力及气候条件）对其产生的影响，表明它们之间存在密切关系，为建立收获前花生黄曲霉毒素污染大尺度宏观预警模型奠定了基础。

第四章 中国花生黄曲霉毒素 B₁ 污染预警模型的构建

4.1 引言

食品安全问题不容忽视，而真菌毒素的污染影响粮食及饲料的质量，在许多领域受到了的关注（Lake *et al.* 2012; Miraglia *et al.* 2009; Tao *et al.* 2011; Wheeler *et al.* 2013）。由于人为因素以及气候变化的影响，了解农作物中黄曲霉毒素污染的发生规律，从而根据预期的变化采取防控措施变得尤为重要。作物生长及其与有益和致病和/或产毒微生物的相互作用存在年度差异，这主要取决于天气状况、作物和微生物之间的互作（Moore *et al.* 2015）。研究真菌毒素的发生规律具有重要的经济和社会意义。

影响花生黄曲霉毒素产生的因素很多，包括真菌类型、花生品种、花生成熟度、土壤类型、耕种密度、虫害、气象条件、农艺措施、植物抗毒素、水分活度、收获和干燥时间等（Torres *et al.* 2014; Diao *et al.* 2014）。其中，气候变化会改变产毒真菌的菌群结构及黄曲霉毒素的污染情况，例如在之前无需关注黄曲霉毒素污染风险的欧洲，近年来也开始发现黄曲霉毒素的污染（Piva *et al.* 2006; Dobolyi *et al.* 2013; Levic *et al.* 2013）。气候变化还会改变作物生长发育周期和昆虫的活动，从而影响黄曲霉的增殖和宿主受污染程度。在未来 10-20 年，大气 CO₂ 浓度预计以 1.5 μmol/年的速度增长，加之其它温室气体浓度的增加，全球气温将以 0.03 °C/年的速度增长，这将会极大影响作物的生长（Magan *et al.* 2011），而温度升高往往伴随着干旱胁迫的发生。在发展中国家，干旱对食品安全的影响更大，例如，在干旱胁迫下容易感染黄曲霉毒素的玉米和花生，黄曲霉毒素的污染风险将会增加、产品的营养质量将会降低并影响其消费和出口贸易。嗜干菌，如 *Walleimia sebi*、双孢霉（*Xeromyces bisporus*）和黄孢霉属（*Chrysosporium spp.*）等真菌，能在非常干燥的条件下生长（*a_w* 为 0.65-0.75），其也可能成为侵染食品的主要微生物（Magan *et al.* 2007）。其中一些真菌如 *W. sebi* 可以产生对动物和人体有毒的代谢物 walleminol 和 walleminone（Pieckova' *et al.* 2002）。有研究表明，在极端干燥条件下次生代谢产物在嗜干菌之间的竞争作用中发挥作用（Leong *et al.* 2011）。这些均要求作物品种具有更高的抗旱性，以适应环境的变化。

目前国内外花生黄曲霉毒素预警技术的热点是模型预警，美国建立了遗传算法/反向传播神经网络模型（GA/BPN），但该技术和模型还未能花生生产中得到广泛应用；美国在 1998 年发布了 CROPGRO 模型，其中在花生结荚期间，并且在适宜黄曲霉生长和产毒的条件下，CROPGRO-peanut 可以根据黄曲霉毒素含量与繁殖期气候参数（降水量、最低温、最高温和日照时长）的关系预测花生黄曲霉毒素含量，但此模型只在当地适用。澳大利亚昆士兰大学采用作物生长模拟方法（APSIM），建立了气候环境因子等对田间生长阶段花生黄曲霉毒素污染预测模型，但因澳大利亚花生种植面积小，气候土壤类型简单，模型仅针对当地 10 个主要花生品种，很难适用于中国。而中国花生种植面积大，产区经纬度跨度 20° 以上，气候类型、土壤类型和花生品种繁多，建立适合中国的花生黄曲霉毒素预警模型对于提高花生质量安全及监管具有重要意义。因此本章首先分析了气候因子与花生 AFB₁ 含量平均值之间的关系，利用相关性分析及最大互信息系数对气候变量进行筛选，基于筛选后的主要气候因子结合花生品种抗性、土壤类型及真菌产毒力因素，建立了适合中国的花生黄曲霉毒素预警模型，以为花生黄曲霉毒素污染控制提供有

效的途径。为了验证模型的有效性，将建立的模型用于 2018 年采自 61 个花生主产区 1181 份花生样品 AFB₁ 含量的预测，通过分析预测误差评估本章的随机森林预测模型在长时间序列研究中的各影响因子对花生 AFB₁ 含量影响的科学性。

4.2 材料

4.2.1 属性变量

4.2.1.1 花生品种抗性

同 3.2.3。

4.2.1.2 土壤类型

同 3.2.4。

4.2.1.3 菌株产毒力

同 3.2.5。

4.2.1.4 地理与气候数据

2009-2017 年全国 140 多个气象站的数据资料源自于中国气象数据网—中国地面气候资料日值数据集 (V3.0)。参数包括 3 个地理位置参数：纬度 (Latitude, LAT)、经度 (Longitude, LON)、海拔高度 (Altitude, ALT) 及 19 个气象参数：蒸发量 (Evaporation, EVA)、平均地表气温 (Mean soil surface temperature, GSTmean)、日最高地表气温 (Maximum soil surface temperature, GSTmax)、日最低地表气温 (Minimum soil surface temperature, GSTmin)、20-8 时降水量 (Precipitation from 20:00 to 8:00, PRE20-8)、8-20 时降水量 (Precipitation from 8:00 to 20:00, PRE8-20)、20-20 时降水量 (Precipitation from 20:00 to 20:00, PRE20-20)、平均风速 (Mean wind speed, WIN)、最大风速 (Maximum wind speed, WINmax)、极大风速 (Extreme wind speed, WINext)、日照时数 (Sunshine duration, SSD)、平均气压 (Mean air pressure, PRSmean)、日最高气压 (Maximum air pressure, PRSmax)、日最低气压 (Minimum air pressure, PRSmin)、日平均气温 (Mean air temperature, TEMmean)、日最高气温 (Maximum air temperature, TEMmax)、日最低气温 (Minimum air temperature, TEMmin)、平均相对湿度 (Mean relative humidity, RHUmean)、最小相对湿度 (Minimum relative humidity, RHUmin)。

4.2.2 花生样品

同 2.1.1。

4.3 方法

4.3.1 多重共线性诊断

本章采用统计学软件 SPSS20.0 对气候因子进行了多重共线性诊断，具体操作过程为：分析—回归—线性—统计量—共线性诊断。

4.3.2 相关性分析

相关系数是用来评价两个变量之间线性相关程度的一个统计指标，本章利用 matlab 2014a 计算各变量与目标变量的相关系数。为了度量气候因子与花生 AFB₁ 之间的相关关系，首先利用不同的方法（Pearson、Spearman 和 Kendall）来计算线性相关系数；其次，利用最大互信息系数（Maximum information coefficient, MIC）来衡量所有变量与目标变量之间的非线性相关性，初步筛选气候变量；最后根据初选气候变量之间的线性相关性再次进行变量筛选，确定主要气候因子。

4.3.3 建模方法

本章采用规则分类—平衡取样—随机森林方法，以 2009-2017 年的数据为基础，利用花生品种抗性、土壤类型、菌株产毒力和主要气候因子预测花生 AFB₁ 含量是否超限（以 5 μg/kg 和 20 μg/kg 为阈值），并进一步分析了参数在分类模型中的贡献率大小。为考察模型的精确度，与常用的分类预测方法（决策树、支持向量机、K-近邻法、BP 神经网络、径向基函数神经网络）进行结果比较，并利用 2018 年数据进行模型验证。

4.4 方法及模型

4.4.1 规则分类器

4.4.1.1 算法

模型规则的表达采用析取范式 $R = (r_1 \vee r_2 \vee \dots \vee r_k)$ ，其中 R 是规则集， r_i 是分类规则或析取项。

每一个分类规则的表达形式如下： $r_i: (\text{条件 } i) \rightarrow y_i$

规则左边称为规则前件或前提。它是属性测试的合取：

$$\text{条件 } i = (A_1 \text{ op } v_1) \wedge (A_2 \text{ op } v_2) \wedge \dots \wedge (A_n \text{ op } v_n) \quad (\text{公式 4.1})$$

其中 (A_j, v_j) 是属性值对，op 是比较运算符，取自集合 $\{=, \neq, <, >, \leq, \geq\}$ 。每一个属性测试 $(A_j \text{ op } v_j)$ 称为一个合取项。规则右边称为规则后件，包含预测类 y_i 。

如果规则 r 的前件和记录 x 的属性匹配，则称 r 覆盖 x 。当 r 覆盖给定的记录时，称 r 被激发或触发。度量分类规则质量的指标包括覆盖率（Coverage）和准确率（Accuracy）。给定数据集 D 和分类规则 $r: A \rightarrow y$ ，规则的覆盖率是指 D 中触发规则 r 的记录所占的比例。准确率或置信因子是指触发 r 的记录中类标号等于 y 的记录所占的比例。

$$\begin{aligned} \text{Coverage}(r) &= |A|/|D| \\ \text{Accuracy}(r) &= |A \cap y|/|A| \end{aligned} \quad (\text{公式 4.2})$$

其中 $|A|$ 表示满足规则前件的记录数， $|A \cap y|$ 表示同时满足规则前件和后件的记录数， D 表示记录总数。

4.4.1.2 基于规则分类器的特征

基于规则的分类器有如下特点：

(1) 规则集的表达能力几乎等价于决策树，因为决策树可以用互斥和穷举的规则集表示。基

于规则分类器和决策树分类器都对属性空间进行直线划分，并将类指派到每个划分。然而，如果基于规则的分类器允许一条记录触发多条规则的话，就可以构造一个更加复杂的决策边界。

(2) 基于规则的分类器通常被用来产生更易于解释的描述性模型，而模型的性能却可与决策树分类器相媲美。

(3) 被很多基于规则的分类器所采用的基于类的规则定序方法非常适于处理类分布不平衡的数据集。

4.4.2 最大互信息系数

最大互信息系数 (MIC) 源自 2011 年 Reshef 等在 Science 发表地一篇文章，它是用来度量变量相互关系 (线性、非线性与非函数) 的统计量。MIC 的算法将信息论和概率的概念应用于连续型数据。该方法不需要对数据分布做任何假设，便可估计变量之间的相关性 (包括线性和非线性)，因此可以有效应用于现实环境中，例如在雷达回波检测、系统识别中，非线性是广泛存在的，包括周期性、指数关系等。

通常相关关系的度量方法有皮尔逊 (积矩) 相关系数 (Pearson product-moment correlation coefficient)、斯皮尔曼 (秩次, 等级) 相关系数 (Spearman's rank correlation coefficient) 以及决定系数 (Coefficient of determination) 等。皮尔逊相关系数也就是通常所说的 (线性) 相关系数，为一个介于 1 和 -1 之间的值，其中 1 表示变量完全正相关，0 表示线性无关，-1 表示完全负相关。绝对值越大表示线性关系越强，但它不能度量线性关系的斜率以及非线性关系，当然也不能度量非函数关系。斯皮尔曼 (秩次, 等级) 相关系数，即把 X 中的 n 个数从小到大排列，若 X_i 排在了第 A_i 个位置，则 X_i 的秩次为 A_i，从而可得 X 所对应的秩次向量 A，同理可得 Y 的所对应的秩次向量 B，A 和 B 之间的皮尔逊相关数就是 X 与 Y 之间的斯皮尔曼相关系数。斯皮尔曼秩相关系数可以在一定程度上反应出二者的非线性关系，若 Y 是 X 的严格单调递减 (增) 函数，则斯皮尔曼秩相关系数是 -1 (+1)。可决系数是指设 Y_i 是观测值，f_i 是线性回归对应的预测值 (f_i(X₁, X₂, ..., X_n))，则 R² 是拟合优度 (Goodness of fit) 的一个统计量，是指回归直线对观测值的拟合程度。R² 越接近 1，说明回归直线对观测值的拟合程度越好 (在总变差中由模型作出解释的部分所占比重越大)；反之，R² 越接近 0，说明回归直线对观测值的拟合程度越差 (在总变差中由模型作出解释的部分所占比重越小)。

MIC 属于最大的基于信息的非参数性 (Maximal information-based nonparametric exploration, MINE) 探索，用于衡量两个变量 X 和 Y 的线性或非线性的强度。MIC 度量具有普适性，其不仅可以发现变量间的线性函数关系，还能发现非线性函数关系 (指数的，周期的)；不仅能发现变量间的函数关系，还能发现非函数关系 (比如函数关系的叠加，或者有趣的图形模式)。MIC 度量具有均衡性。对于相同噪声水平的函数关系或者非函数关系，MIC 度量具有近似的值。所以 MIC 度量不仅可以用来纵向比较同一相关关系的强度，还可以用来横向比较不同关系的强度。MIC 度量计算的方法为：假设 D 是有序对的有限集合，即 D = { (x, y) | x 与 y 相关或不相关 }。n 表示数据集 D 的数据量，即 n = |D|。D 的 x 值被划分成 s 份，y 值被划分成 t 份，形成 s × t 的网格划分数据空间，使落在第 G 格子中的数据点的频率作为 DG。假设有有限数据集 D ⊂ R²，s、t 为正整数：

$$I^*(D, s, t) = \max I(DG) \quad (\text{公式 4.3})$$

其中互信息最大的 $s \times t$ 网格划分, $I(DG)$ 是 DG 的交互信息 (Reshef *et al.* 2011)。用 $\log_2 \min\{s, t\}$ 进行归一化, 特征矩阵的计算如下:

$$M(D)_{s,t} = \frac{I^*(D,s,t)}{\log_2 \min\{s,t\}} \quad (\text{公式 4.4})$$

最后, 找到能使归一化互信息最大的网格分辨率, 作为 MIC 的度量值

$$MIC(D) = \max_{s,t \leq B(n)} \{M(D)_{s,t}\} \quad (\text{公式 4.5})$$

其中网格的分辨率限制为 $s, t \leq B(n)$, 且 $\omega(1) < B(n) < O(n^{1-\epsilon})$, $0 < \epsilon < 1$ 。数据集 D 的 MIC 值用于测量数据集 D 中两个因素 x 和 y 之间的相关性。MIC 值越大, x 和 y 之间的相关性越高。从 MIC 的定义可知, MIC 的值在 $(0, 1)$ 区间之内 (Reshef *et al.* 2011)。

利用 MIC 探寻两变量之间的关系已应用于许多领域, 如汽车交通 (Shao *et al.* 2016)、药物基因组学 (Zhou *et al.* 2012)、生物学 (Liu *et al.* 2016) 以及环境 (Eiler *et al.* 2014) 等方面。

4.4.3 基于机器学习的分类方法

机器学习主要是研究如何通过数据的学习来改善系统自身的泛化性能。在计算机系统中, “经验” 通常以 “数据” 的形式存在, 因此, 机器学习所研究的主要内容是关于在计算机上从数据中产生 “模型” (Model) 的算法, 即 “学习算法” (Learning algorithm)。通过对机器学习方法的研究, 既有助于深化对数据读出过程的理解, 也使得整个数据读出流程更加智能, 这体现在两个方面: (1) 学习能力 (Learning): 数据读出系统可通过学习过程 (训练) 来提升自身的性能, 并通过所学特征参数形式更新分类器模型, 使其对输入数据进行最佳分类输出; (2) 泛化能力 (Generalization): 经过学习后的数据读出系统, 能够对未标注的新输入数据进行识别和分类, 即利用学习所获得的知识, 解决未知问题的能力。机器学习的算法种类繁多, 如决策树、支持向量机、神经网络、随机森林等, 每种算法又可以通过调节参数来优化其性能。针对不同的需求, 需要选择合适的机器学习算法 (蒋荻 2017)。

4.4.3.1 决策树方法

决策树 (Decision tree, DT) 是机器学习中用于分类的一种方法, 基于树结构来对被测对象进行分类这恰好是人类在面临对问题进行决策时的一种自然的处理机制, 即对于问题的每一个已知条件进行判断最终得到该问题的结论 (张耀鹏 2018)。根据决策树学习到的知识可表示为多个 if-then 过程以提高可读性。决策树的结构如图 4.1 所示。

一棵决策树仅包含一个根节点、若干个内部节点和若干个叶子节点。根节点包含样本全集, 每个分支为该属性的可能取值, 每个非叶子节点根据属性取值的不同向下生成一个分支。决策树中对于被预测的对象, 将其各属性按照取值从根节点开始一路往下进行判断即可预测出被测对象属于的类别。决策树在每一次的分裂过程中做出当前的最优划分特征, 但并不关心能否达到全局最优。所以在决策树的构造过程中, 数据中的噪声非常容易导致过拟合的出现。

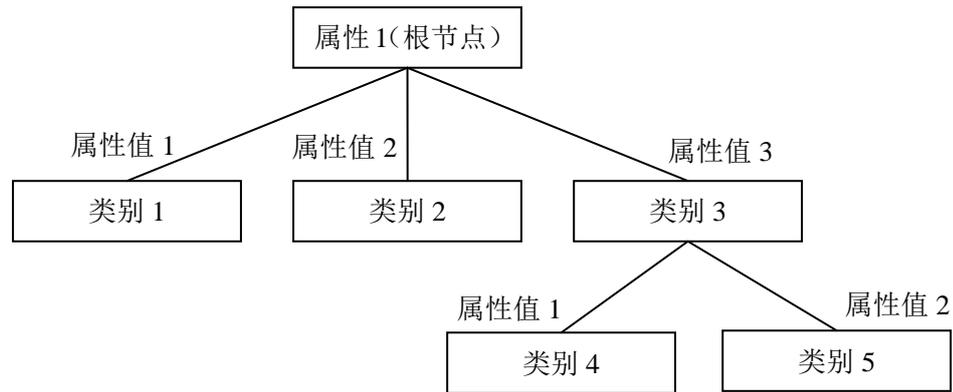


图 4.1 决策树典型结构

Figure 4.1 Typical structure of decision tree

4.4.3.2 支持向量机

支持向量机 (Support vector machine, SVM) 是一种通用学习机器, 它较好地实现了结构风险最小化思想。SVM 是将输入向量映射到一个高维的特征空间, 并在该特征空间中构造最优分类面, 它能够避免在多层前向网络中无法克服的一些缺点, 并且理论证明了: 当选用合适的映射函数时, 大多数输入空间线性不可分的问题在特征空间可以转化为线性可分问题来解决 (张国云 2006)。

支持向量机有以下显著特征: (1) 结构简单; (2) 凸优化问题有关的优化问题无局部极小点; (3) 稀疏性, 最优超平面之法向量是输入数据的线性组合, 每个数据的系数在某种意义上反映了该数据的重要性; (4) 模块化可以清晰地分成两个模块: 一个通用的学习机和与具体问题有关的核函数; (5) 本质上是线性学习机, 是在核函数诱导的 (隐含的) 特征空间上的线性函数。

4.4.3.3 k—近邻法

k—近邻法 (K-nearest neighbor, KNN) 是机器学习中一种简单常用且性能较好的算法, 它是由最近邻法扩展而来。最近邻法是将待分类样本点的特征矢量与训练集中所有样本点的特征矢量计算某一距离度量, 训练集中与待定位样本点距离最小的样本点所属类别即为待定位样本点的分类结果。但是由于在实际应用中, 训练集的采集往往是存在误差的, 所以可能存在以下情况: 最近邻法找到的与待分类样本距离最近的训练样本采集时存在较大误差, 实际与待分类样本距离最近的训练样本并不是选中样本而导致错误的分类结果。为了降低上述风险, 就可以选取与待分类样本距离最近的多个训练样本, 然后用投票方式进行决策而不是直选出距离最近的一个, 这就是 k—近邻法 (李运 2015)。KNN 分类方法是一种非参数的分类技术, 对于未知和非正态分布的数据可以取得较高的分类准确率, 具有概念清晰、易于实现等诸多优点。但同时也存在分类过程中计算量过大、对样本库过于依赖和度量相似性的距离函数不适用等问题。

4.4.3.4 人工神经网络

人工神经网络 (Artificial neural network, ANN), 也常简称为神经网络或类神经网络, 是 20

世纪 80 年代以来人工智能领域兴起的研究热点。它从信息处理角度对人脑神经网络进行抽象，建立某种简单模型，按不同的连接方式组成不同的网络。ANN 是一种非线性数学模型，能够模拟任意复杂程度的非线性过程，最大的优点是能够反复训练、测试网络结构，最终逼近任意复杂的非线性函数。最近十多年来，ANN 的研究工作不断深入，已经取得了很大的进展，其在模式识别、智能机器人、自动控制、预测估计、生物、医学、经济等领域已成功地解决了许多现代计算机难以解决的实际问题，表现出了良好的智能特性。

BP 神经网络 (Back propagation network, BPNN) 是应用最广泛的一种 ANN (卢志宏 2016)，最早由 Rumelhart 和 McClelland 等人提出，能够学习和存储大量输入—输出模式分映射关系，而不需要准确描述这种映射关系的数学方程。BPNN 使用最速下降法进行学习，通过反向传播来不断调整网络的权值和阈值，使网络的误差和平方和最小，一般包括输入层、隐含层和输出层，层与层之间相互连接，层内的神经元之间不可连接 (图 4.2)。BPNN 具有的优点是：(1) 高度并行性，具有联想功能；(2) 高度非线性和容错性；(3) 较强的自学能力；(4) 系统的稳定性。但是 BP 算法也存在以下局限性：(1) 网络训练的收敛速度慢；(2) 网络训练容易陷入局部极小值点；(3) 网络的学习和记忆不稳定；(4) 网络的隐含层的层数和每个隐含层节点数的选取没有完整统一的理论指导，是根据经验确定，试凑法工作量大，网络确定过程中产生大量的冗余，从而增加了学习和仿真的时间。

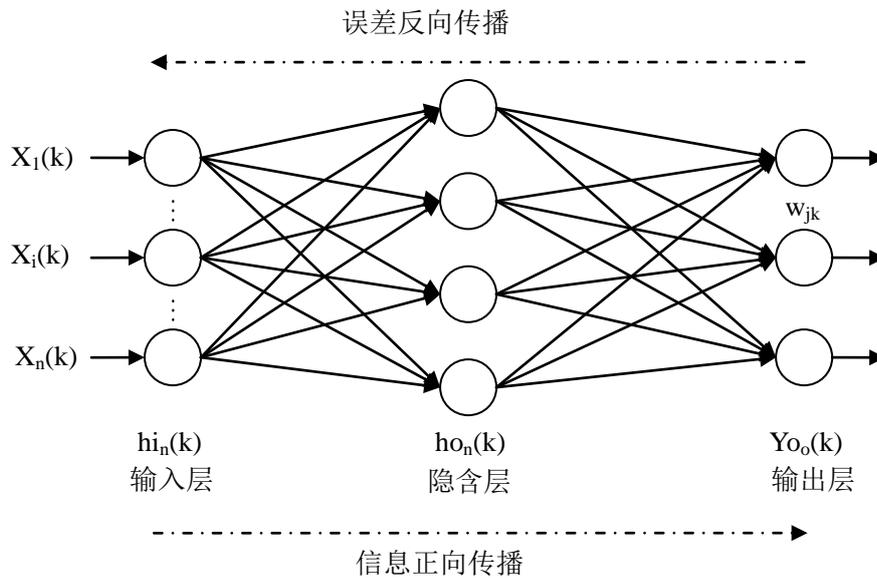


图 4.2 BPNN 结构图

Figure 4.2 Structure diagram of BPNN

径向基函数神经网络 (Radial basis function neural network, RBFNN)，在 20 世纪 80 年代末由 J.Moody 和 C.Darken 提出的一种含有局部特性的神经网络，是一种具有单隐层的 H 层前馈式网络。和同为前馈式网络的 BP 全局逼近网络相比，RBFNN 具有的优点为：(1) 结构简单，输入输出层都是线性映射关系，易于结构设计、参数调节并且参数学习过程的收敛速度很快；(2) 具有很强的拟合输入输出映射关系的功能，在前向网络中 RBFNN 是能够完成映射功能最优的网络

结构；(3) RBFNN 的学习速度快，且避免了局部极小问题；(4) 能够实现“动态的学习”；(5) 可以通过聚类的方式选择基函数的中心位置，大大降低了对人工经验调参数的依赖。缺点是：(1) 无法解释自己的推理过程和推理依据；(2) 不能向用户提出必要的询问，而且当数据不充分时，神经网络就无法进行工作；(3) 把一切问题的特征都变为数字，把一切推理都变为数值计算，其结果势必是丢失信息；(4) 理论和算法还有待于进一步完善和提高；(5) RBFNN 用于非线性系统建模需要解决的关键问题是样本数据的选择。另外优选过程会出现数据病态现象。

4.4.3.5 随机森林

随机森林 (Random forest, RF) 指的是利用多棵树对样本进行训练并预测的一种分类器。就是通过集成学习的思想将多棵树集成的一种算法，它的基本单元是决策树，而它的本质属于机器学习的一大分支—集成学习 (Ensemble learning) 方法。决策树是一种树形结构，其中每个内部节点表示一个属性上的测试，每个分支代表一个测试输出，每个叶节点代表一种类别。随机森林算法流程图如图 4.3 所示。常见的决策树算法有 C4.5、ID3 和 CART。集成学习通过建立几个模型组合来解决单一的预测问题。它的工作原理是生成多个分类器/模型，各自独立地学习和作出预测。这些预测最后结合成单预测，因此优于任何一个单分类的做出预测。随机森林是集成学习的一个子类，它依靠于决策树的投票选择来决定最后的分类结果。随机森林是一种重要的基于 Bagging 的集成学习方法，可以用来做分类、回归等问题。

随机森林有许多优点：(1) 具有极高的准确率；(2) 随机性的引入，使得随机森林不容易过拟合且具有很好的抗噪声能力；(3) 能处理很高维度的数据，并且不用做特征选择；(4) 既能处理离散型数据，也能处理连续型数据，数据集无需规范化；(5) 速度快，可以得到变量贡献率排序；(6) 容易实现并行化。随机森林的缺点是：(1) 随机森林中的决策树个数很多时，训练时需要的空间和时间会较大；(2) 随机森林模型还有许多不好解释的地方，有点像黑盒模型。

随机森林的关键问题就是如何选择最优的输入特征数目，要解决这个问题主要依据计算袋外错误率 (Out of bag error, OOB)。随机森林有一个重要的优点就是，没有必要对它进行交叉验证或者用一个独立的测试集来获得误差的一个无偏估计。它可以在内部进行评估，也就是说在生成的过程中就可以对误差建立一个无偏估计。在构建每棵树时，由于训练集使用了不同的 Bootstrap sample (随机且有放回地抽取)。所以对于每棵树而言 (假设对于第 k 棵树)，大约有 $1/3$ 的训练实例没有参与第 k 棵树的生成，它们称为第 k 棵树的 OOB 样本。它的计算方式如下：(1) 对每个样本，计算它作为 OOB 样本的树对它的分类情况 (约 $1/3$ 的树)；(2) 然后以简单多数投票作为该样本的分类结果；(3) 最后用误分个数占样本总数的比率作为随机森林的 OOB 误分率。

随机森林拥有广泛的应用前景，目前已应用于医学 (De *et al.* 2016)、遥感 (Belgiu *et al.* 2016)、生物测定学 (Finely *et al.* 2013)、食品科学 (Zhang *et al.* 2016) 和金融保险等 (Chasidim *et al.* 2016)。

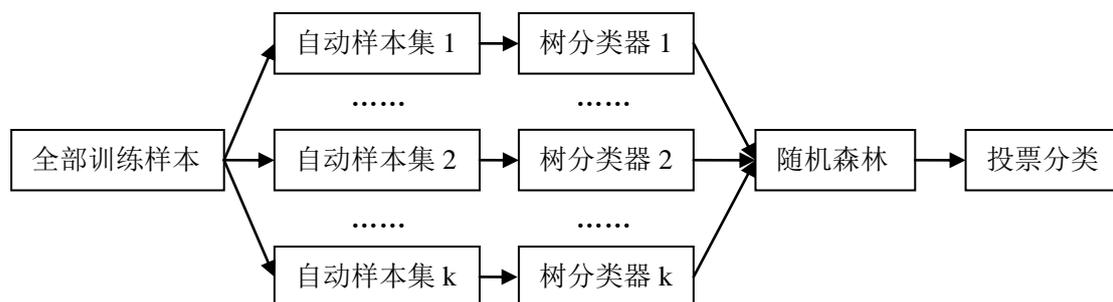


图 4.3 随机森林算法流程图

Figure 4.3 Algorithm schematic diagram of RF

4.5 结果与分析

4.5.1 气候变量筛选

4.5.1.1 数据特点

(1) 气候因子与花生 AFB₁ 平均含量之间的关系

由于图形显示限制, 部分气候因子(蒸发量、20-20 时降水量、日平均气温及平均相对湿度湿度)与四个产区花生 AFB₁ 平均含量散点图绘入图 4.4。图 4.4 表明气候因子与花生 AFB₁ 平均含量之间存在着非常复杂的非线性关系。

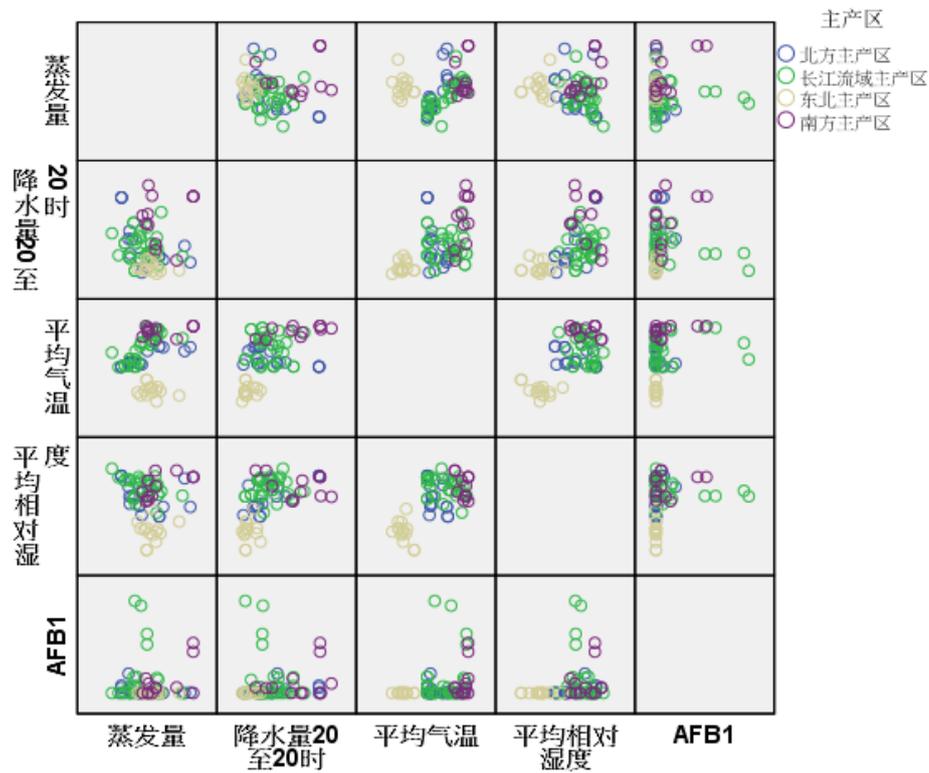
图 4.4 2009-2017 年花生 AFB₁ 平均含量和气候因子矩阵散点图

Figure 4.4 Matrix scatter plot between mean AFB₁ content in peanuts and climatic factors in China from 2009 to 2017

(2) 气候因子的多重共线性检验

在统计学中多重共线性是指自变量之间存在着相关性。当自变量之间存在高度相关性时，就不能精确确定模型中任意一个自变量对因变量的单独影响。因此，为确保模型的精确度，本章首先对气候因子进行多重共线性检验，诊断结果见表 4.1、表 4.2。统计学中共线性诊断的主要依据是“特征值”，主要考察的是自变量的方差。具体计算过程为先得自变量之间的相关系数矩阵，然后计算相关矩阵的特征值，最后得到相应成分。一般情况下，数据经过标准化后方差值为 1，因此每一个特征值都可以解释某一个自变量的一定比例成分。当条件索引值超过 10 时，则认为自变量之间存在多重共线性。

观察气候因子多重共线性诊断表(表 4.1)可知：表中第 4 维度条件索引值为 8.747，接近 10，第 5 维度条件索引值为 11.115，大于 10，说明待研究的气候因子之间存在着较强共线性。观察表 4.2 可知：表中第 7 个特征值分别解释了 0.01 的纬度，0.02 的蒸发量；第 8 个特征值分别解释了 0.01 的纬度，0.76 的蒸发量；第 9 个特征值分别解释了 0.04 的经度，0.03 的最大风速，0.02 的最小相对湿度等。也即存在着 2 个或以上特征值，可同时解释多个变量，说明气候因子之间存在着一定的多重共线性。

为了研究气候因子之间的多重共线性是否会影响模型的精确性，本章进一步对气候因子之间的相关性做了研究。同样采用 SPSS 20.0 软件计算气候因子之间的相关系数，操作过程为：分析—相关—双变量，相关系数结果见表 4.3。通过显著性 0.01 以及 0.05 检验，由表 4.3 可知气候参

数之间具有较大相关性。因此认为气候因子之间的多重共线性会影响普通模型的精度。

表 4.1 气候因子多重共线性诊断表

Table 4.1 Multi-collinearity diagnostic table for climatic factors

| 模型 | 维度 | 特征值 | 条件索引 |
|----|----|----------|-----------|
| 1 | 1 | 19.023 | 1 |
| | 2 | 0.768 | 4.976 |
| | 3 | 0.531 | 5.986 |
| | 4 | 0.249 | 8.747 |
| | 5 | 0.154 | 11.115 |
| | 6 | 0.13 | 12.116 |
| | 7 | 0.081 | 15.358 |
| | 8 | 0.028 | 26.005 |
| | 9 | 0.014 | 37.228 |
| | 10 | 0.01 | 43.628 |
| | 11 | 0.006 | 58.471 |
| | 12 | 0.003 | 81.628 |
| | 13 | 0.003 | 85.45 |
| | 14 | 0.001 | 153.56 |
| | 15 | 0 | 222.651 |
| | 16 | 0 | 241.411 |
| | 17 | 0 | 278.216 |
| | 18 | 7.38E-05 | 507.856 |
| | 19 | 2.29E-05 | 912.44 |
| | 20 | 6.76E-06 | 1677.074 |
| | 21 | 6.87E-08 | 16645.487 |

表 4.2 气候因子多重共线性诊断—方差比例表

Table 4.2 Multi-collinearity diagnosis-variance scale of climate factors

| 常量 | LAT | LON | EVA | GST mean | GST max | GST min | WIN max | PRS max | TEM max | TEM min | RHU min |
|------|------|------|------|-------------|------------|------------|------------|------------|------------|------------|------------|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0.08 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0.01 | 0 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0.01 | 0 | 0.76 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0.04 | 0 | 0 | 0 | 0 | 0.03 | 0 | 0 | 0 | 0.02 |
| 0 | 0.02 | 0.08 | 0.1 | 0 | 0 | 0.01 | 0.02 | 0 | 0 | 0 | 0.02 |
| 0 | 0.15 | 0.01 | 0 | 0 | 0.03 | 0.01 | 0 | 0 | 0 | 0 | 0.01 |
| 0 | 0.53 | 0.04 | 0.02 | 0 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0.02 |
| 0 | 0.03 | 0.69 | 0 | 0 | 0 | 0.01 | 0.57 | 0 | 0 | 0 | 0.01 |
| 0 | 0 | 0.09 | 0 | 0 | 0 | 0.36 | 0.33 | 0 | 0.02 | 0.01 | 0.23 |
| 0 | 0 | 0 | 0 | 0.02 | 0.01 | 0 | 0 | 0 | 0.05 | 0.02 | 0.07 |
| 0 | 0.09 | 0 | 0.01 | 0 | 0 | 0.06 | 0.01 | 0 | 0 | 0.02 | 0.02 |
| 0 | 0.06 | 0.02 | 0 | 0.01 | 0.03 | 0.35 | 0 | 0 | 0.14 | 0.17 | 0.47 |
| 0 | 0.06 | 0 | 0 | 0.72 | 0.69 | 0.07 | 0.01 | 0 | 0.08 | 0 | 0.07 |
| 0 | 0 | 0 | 0 | 0.24 | 0.21 | 0.07 | 0 | 0 | 0.7 | 0.77 | 0.03 |
| 0.97 | 0.02 | 0 | 0 | 0 | 0 | 0 | 0.02 | 0 | 0 | 0.01 | 0.01 |
| 0.02 | 0.02 | 0.01 | 0 | 0 | 0 | 0.06 | 0 | 1 | 0.02 | 0.01 | 0.02 |

表 4.3 气候因子相关系数表

Table 4.3 Correlation coefficient table of climate factors

| | LAT | LON | EVA | GSTmean | GSTmax | GSTmin | WINmax | PRSmean | PRSmax | TEMmax | TEMmin | RHUmin |
|---------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| LAT | 1 | -0.094** | -0.222** | -0.759** | -0.532** | -0.359** | -0.124** | 0.334** | 0.344** | -0.736** | -0.856** | -0.549** |
| LON | -0.094** | 1** | 0.412** | 0.202** | 0.286** | -0.117** | 0.928** | 0.018** | 0.019** | 0.179** | 0.065** | -0.088** |
| EVA | -0.222** | 0.412** | 1 | 0.582** | 0.677** | -0.340** | 0.367** | -0.103** | -0.106** | 0.563** | 0.386** | -0.226** |
| GSTmean | -0.759** | 0.202** | 0.582** | 1 | 0.890** | 0.070** | 0.178** | -0.298** | -0.309** | 0.965** | 0.924** | 0.264** |
| GSTmax | -0.532** | 0.286** | 0.677** | 0.890** | 1 | -0.233** | 0.243** | -0.211** | -0.218** | 0.832** | 0.667** | -0.093** |
| GSTmin | -0.359** | -0.117** | -0.340** | 0.070** | -0.233** | 1** | -0.020** | -0.065** | -0.074** | 0.024** | 0.295** | 0.898** |
| WINmax | -0.124** | 0.928** | 0.367** | 0.178** | 0.243** | -0.020** | 1 | 0.022 | 0.023 | .139** | .067* | 0.003 |
| PRSmean | 0.334** | 0.018** | -0.103** | -0.298** | -0.211** | -0.065** | 0.022 | 1 | 1.000** | -0.307** | -0.331** | -0.163** |
| PRSmax | 0.344** | 0.019** | -0.106** | -0.309** | -0.218** | -0.074** | 0.023 | 1.000** | 1 | -0.318** | -0.344** | -0.172** |
| TEMmax | -0.736** | 0.179** | 0.563** | 0.965** | 0.832** | 0.024** | 0.139** | -0.307** | -0.318** | 1 | 0.912** | 0.189** |
| TEMmin | -0.856** | 0.065** | 0.386** | 0.924** | 0.667** | 0.295** | 0.067* | -0.331** | -0.344** | 0.912** | 1 | 0.522** |
| RHUmin | -0.549** | -0.088** | -0.226** | 0.264** | -0.093** | 0.898** | 0.003 | -0.163** | -0.172** | 0.189** | 0.522** | 1 |

注: **在 0.01 水平上显著相关; *在 0.05 水平上显著相关。

(3) 黄曲霉毒素含量呈偏态分布

通过绘制不同 AFB₁ 含量的花生样品频数和占比分布直方图（图 4.5），可知大约 85% 的样品 AFB₁ 含量范围为 0-5 $\mu\text{g}/\text{kg}$ ，超过国家限量标准（20 $\mu\text{g}/\text{kg}$ ）的样品量占比很低，大约为 10%。

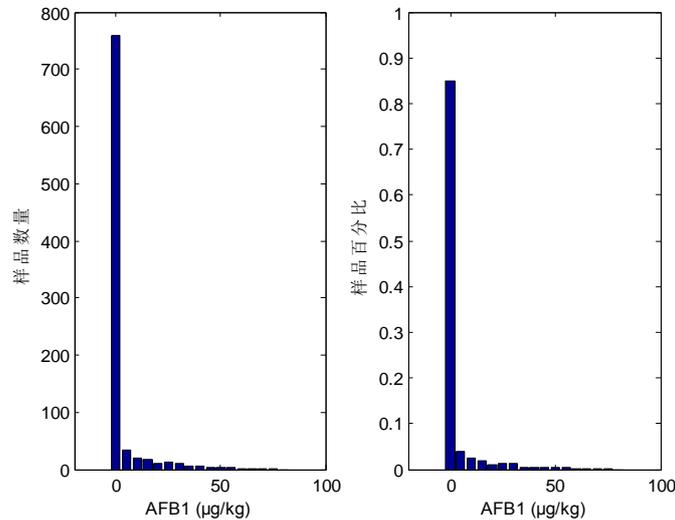


图 4.5 AFB₁ 分布直方图

Figure 4.5 AFB₁ distribution histogram

综合以上分析可知，在具体的建模过程中存在三点问题：（1）气候因子与花生 AFB₁ 含量之间呈现非常复杂的非线性关系；（2）气候因子之间存在着很强的多重共线性问题；（3）黄曲霉毒素污染分布呈左偏分布，即使相同条件下的样品也同时存在污染值的高低，无法做出完全地定量预测；且偏态分布导致超限样本量远低于未超限样本量，在使用传统分类器的时候，效果无法保证。因此首先需要对气候变量进行筛选；其次，为了获取更好的风险预测效果，在预测时以“牺牲未超限样本预测准确度来提升超限样本预测准确度”为原则。这样可以有效的保证超限样本的预测准确度。同时，分析样本关于各个气候变量的分布规则，建立基于未超限样本分布规律的规则分类器，在整体上保证未超限样本准确度最终还能保持在一个合理的水平；再次，采用基于抽样的方法，重复复制超限样本，使得超限样本总量与未超限样本量大体相当，甚至超过未超限样本数量。这样做的目的是在分类预测的时候，增大超限样本的权重，从而在两种样本高度交织的区域内，把更多的样本归结为超限样本。这样可以保证超限样本预测的准确度。

基于以上考虑，本章采用规则分类、平衡取样结合随机森林的方法，利用花生品种抗性、土壤类型、真菌产毒力和主要气候因子对花生 AFB₁ 含量是否超限（以 5 $\mu\text{g}/\text{kg}$ 和 20 $\mu\text{g}/\text{kg}$ 为分类阈值）进行预测。为考察模型的精确度，在利用随机森林法建立模型时进行了交叉验证，并与常用的分类预测方法（决策树、支持向量机、K-近邻法、BP 神经网络、径向基函数神经网络）进行结果比较，且进一步分析了各变量在分类模型中的贡献率大小。

4.5.1.2 地理与气候变量筛选

由于数据涉及的属性变量众多，包括 22 个参数，因此首先需要降维，减少变量个数。具体步骤可以分为：

(1) 相关性分析

为了度量各气候因子与花生 AFB₁ 之间的相关关系, 首先利用不同的方法 (Pearson、Spearman 和 Kendall) 来计算线性相关系数。通过计算可以得到, 各变量与目标变量的相关系数都非常低, 多数都不足 0.1, 基本可以认定各属性变量与 AFB₁ 含量之间不存在线性相关 (表 4.4)。Spearman 相关系数可以在一定程度上反应出非线性关系, 而本研究中 Spearman 相关系数较 Pearson 系数和 Kendall 系数高, 也可在一定程度上表明各属性变量与 AFB₁ 含量之间存在非线性关系。

(2) 基于 MIC 进行变量初步筛选

利用 MIC 对所有变量与目标变量之间进行非线性相关分析, 可得知: 所有变量与 AFB₁ 含量之间的相关度整体仍然比较低, 显示因素和指标间的相关度仍然不高, 常规分析建模难度大。但是根据上述 MIC 系数排序, 可以看出该系数还是可以在一定程度上反映各个因素与 AFB₁ 之间相关性 (见表 4.4)。

根据所有指标的 MIC 系数值排序, 考虑因素的全面性, 初选得到的 8 个变量为: 纬度 (LAT)、平均地表气温 (GSTmean)、日最低地表气温 (GSTmin)、8-20 时降水量 (PRE8-20)、平均气压 (PRSmean)、日最低气压 (PRSmin)、日平均气温 (TEMmean) 和日最高气温 (TEMmax)。

表 4.4 利用不同的线性相关分析与非线性相关分析系数比较

Table 4.4 Comparison of coefficients by different linear and non-linear correlation analysis

| 参数 | MIC | Pearson | Spearman | Kendall |
|----------|--------|---------|----------|---------|
| LAT | 0.1654 | -0.0742 | -0.1757 | -0.1297 |
| LON | 0.1373 | -0.0754 | -0.1624 | -0.1189 |
| ALT | 0.1367 | 0.0690 | 0.0294 | 0.0205 |
| EVA | 0.1492 | 0.0861 | 0.0822 | 0.0574 |
| GSTmean | 0.1632 | 0.1034 | 0.1688 | 0.1194 |
| GSTmax | 0.1492 | 0.1092 | 0.1334 | 0.0932 |
| GSTmin | 0.1542 | 0.0826 | 0.1513 | 0.1090 |
| PRE20-8 | 0.1459 | -0.0173 | 0.0554 | 0.0392 |
| PRE8-20 | 0.1573 | -0.0448 | 0.0738 | 0.0526 |
| PRE20-20 | 0.1293 | -0.0355 | 0.0706 | 0.0498 |
| WIN | 0.1305 | -0.0100 | -0.0185 | -0.0133 |
| WINmax | 0.1297 | -0.0182 | -0.0710 | -0.0509 |
| WINext | 0.1342 | -0.0066 | -0.0765 | -0.0558 |
| SSD | 0.1406 | 0.0382 | -0.0206 | -0.0147 |
| PRSmean | 0.1709 | -0.0678 | -0.0456 | -0.0334 |
| PRSmax | 0.1391 | -0.0680 | -0.0481 | -0.0352 |
| PRSmin | 0.1556 | -0.0684 | -0.0441 | -0.0325 |
| TEMmean | 0.1726 | 0.0966 | 0.1606 | 0.1141 |
| TEMmax | 0.1554 | 0.1055 | 0.1407 | 0.0993 |
| TEMmin | 0.1421 | 0.0871 | 0.1591 | 0.1141 |
| RHUmean | 0.1353 | -0.0633 | -0.0417 | -0.0304 |
| RHUmin | 0.1309 | -0.0295 | 0.0397 | 0.0281 |

(3) 基于线性相关性对变量进行再次筛选

初步筛选得到的 8 个变量仍然较多, 进一步分析这些初选变量间的相关系数。通过分析发现有些变量, 比如平均地表气温 (GSTmean)、日最低地表气温 (GSTmin)、日平均气温 (TEMmean)、日最高气温 (TEMmax) 之间高度线性相关, 平均气压 (PRSmean)、日最低气压 (PRSmin) 之间相关度高, 相关系数都超过 0.8, 最高的为 0.97, 接近 1, 具有多重共线性, 这些变量仍可进一步缩减 (见表 4.5)。

由于所有变量与目标变量之间的相关度都比较低, 过少的变量未必对建模有利, 故在考虑

MIC 系数值排序和地理位置、降水量、气温等因素的全面性基础上, 最终选择建模气候变量为: 纬度 (LAT)、8-20 时降水量 (PRE8-20)、平均气压 (PRSmean) 和日平均气温 (TEMmean)。

表 4.5 不同变量之间的线性相关分析

Table 4.5 Linear correlation analysis coefficients between different variables

| | LAT | GST mean | GST min | PRE 8-20 | PRS mean | PRS min | TEM mean | TEM max |
|---------|---------|-------------|------------|-------------|-------------|------------|-------------|------------|
| LAT | 1.0000 | -0.7539 | -0.8620 | -0.5352 | 0.3401 | 0.3339 | -0.8120 | -0.7316 |
| GSTmean | -0.7539 | 1.0000 | 0.9173 | 0.3381 | -0.2986 | -0.2927 | 0.9655 | 0.9645 |
| GSTmin | -0.8620 | 0.9173 | 1.0000 | 0.4913 | -0.3602 | -0.3510 | 0.9743 | 0.9045 |
| PRE8-20 | -0.5352 | 0.3381 | 0.4913 | 1.0000 | -0.2353 | -0.2306 | 0.4409 | 0.3760 |
| PRSmean | 0.3401 | -0.2986 | -0.3602 | -0.2353 | 1.0000 | 0.9997 | -0.3195 | -0.3088 |
| PRSmin | 0.3339 | -0.2927 | -0.3510 | -0.2306 | 0.9997 | 1.0000 | -0.3120 | -0.3042 |
| TEMmean | -0.8120 | 0.9655 | 0.9743 | 0.4409 | -0.3195 | -0.3120 | 1.0000 | 0.9685 |
| TEMmax | -0.7316 | 0.9645 | 0.9045 | 0.3760 | -0.3088 | -0.3042 | 0.9685 | 1.0000 |

4.5.2 规则分类—平衡取样—随机森林模型的建立

4.5.2.1 基于规则的预分类

四个筛选得到的气候变量与全部样品及 AFB₁ 超限样品之间的分布情况如下:

(1) 纬度与样品分布关系

图中红线表示 AFB₁ 含量分别为 5 $\mu\text{g}/\text{kg}$ 和 20 $\mu\text{g}/\text{kg}$ 两种阈值线, 由于 AFB₁ 含量过大的样品数量比较小, 为了视图效果, 将纵坐标上限确定为 60 $\mu\text{g}/\text{kg}$ 。

从图 4.6 中可以看出, 大部分样品集中在中纬度地区, 具体各纬度区间的样品占比可以从图 4.7 中进一步体现。同时, 超限样品呈现一定的集中趋势, 以中纬度地区为主。四分图 (图 4.8) 分别展示了在每个分区间内超限样品占区间内所有样品的比例, 可以认为该地区的超限率。区间样品数量和超限样品占比呈现高度一致性, 也就是说在中纬度地区样品量最多, 而这些地区的样品超限率也最高, 最高能超过 40%。而低纬度和高纬度地区, 尤其是高纬度地区, 样品超限率为 0。

因此对于纬度变量的规则制定为: 将高纬度区域 ($\geq 40^\circ\text{N}$) 和低纬度区域 ($\leq 21^\circ\text{N}$) 样品直接确定为未超限样品, 减少未超限样品规模, 也降低误分类概率。

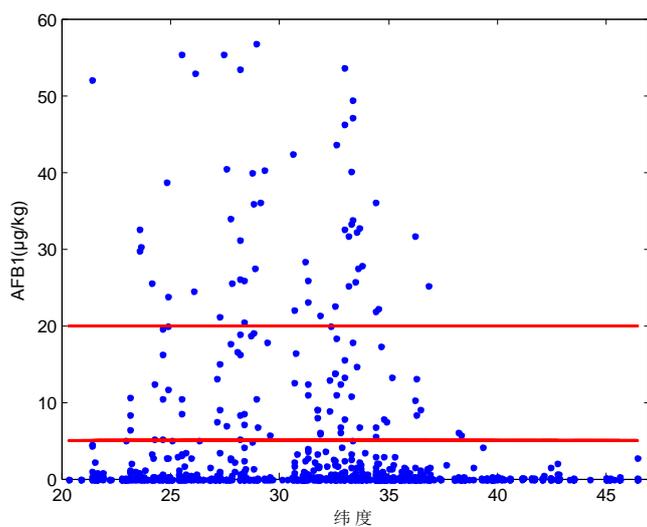


图 4.6 花生 AFB₁ 含量在不同纬度的散点分布

Figure 4.6 Scatter diagram of AFB₁ in peanuts with different latitudes

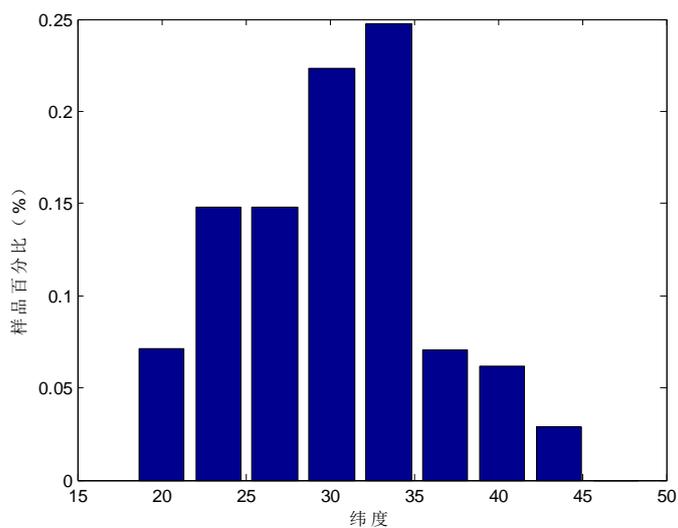


图 4.7 不同纬度区间样品占比

Figure 4.7 Histogram diagram of AFB₁ with different latitudes

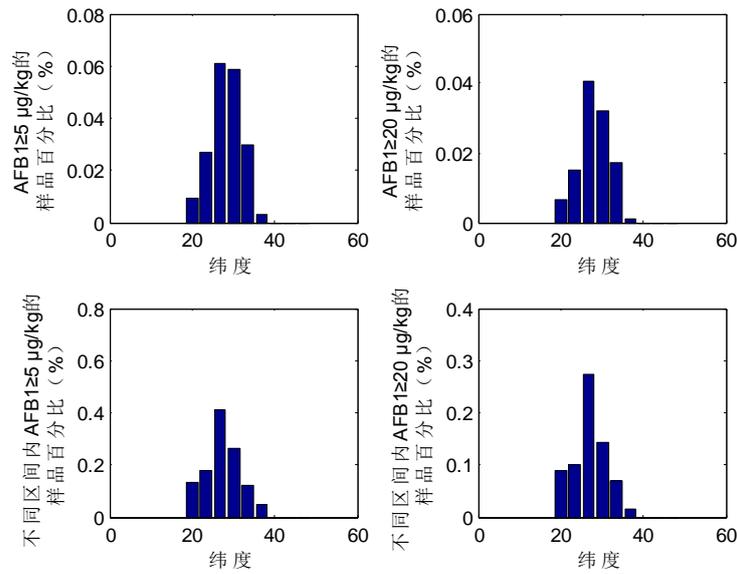


图 4.8 以 5 μg/kg 和 20 μg/kg 为阈值，纬度区间对应超限样品的比例

Figure 4.8 Histogram diagram of proportion of samples with AFB₁ ≥ 5 μg/kg and ≥ 20 μg/kg in all samples and according interval with different latitudes

(2) 8-20 时降水量与样品分布关系

各降水量区间样品的占比可以从图 4.9 中体现。同时，超限样品在低降水量区域较多，超限样品也以低降水量区间为主（图 4.10），四分图（图 4.11）分别展示了在每个分区间内超限样品占区间内所有样品的比例，各降水量区间内均有超限样品分布。因此对降水量规则制定不明显。

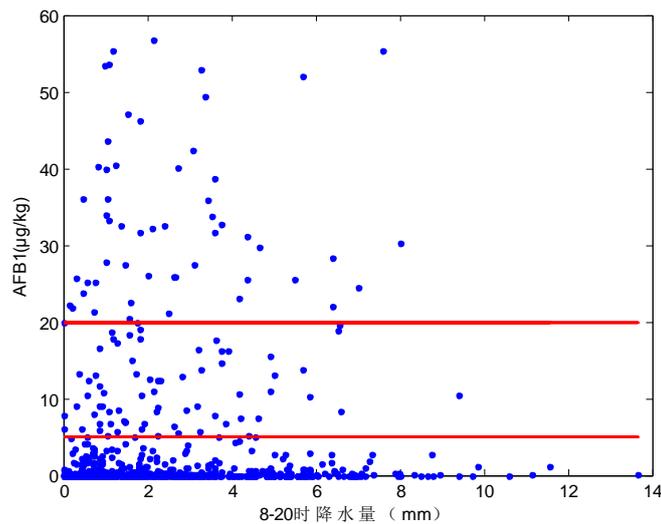


图 4.9 花生 AFB₁ 含量在不同降水量（8-20 时）的散点分布

Figure 4.9 Scatter diagram of AFB₁ in peanuts with different precipitation (from 8:00 to 20:00)

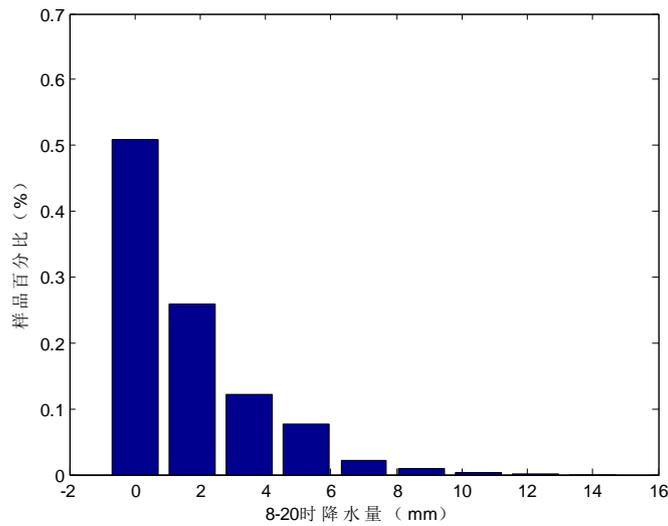


图 4.10 不同 8-20 时降水量区间样品占比

Figure 4.10 Histogram diagram of AFB₁ with different precipitation from 8:00 to 20:00

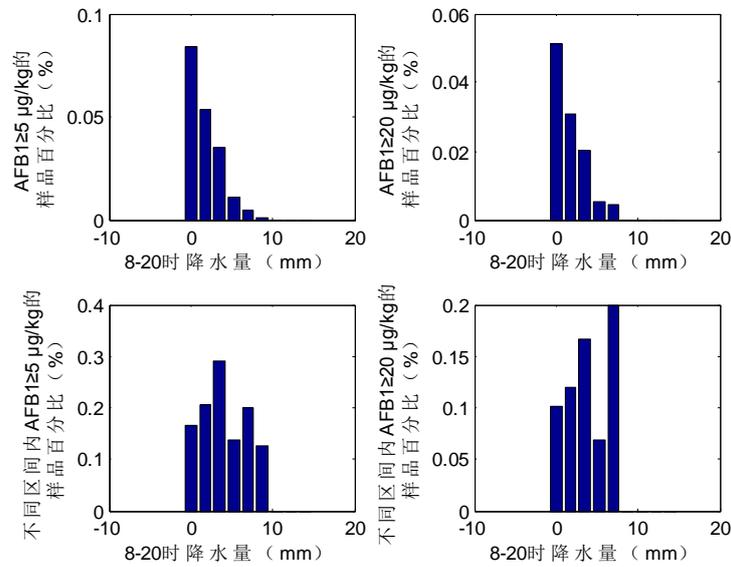


图 4.11 以 5 μg/kg 和 20 μg/kg 为阈值，8-20 时降水量区间对应超限样品的比例

Figure 4.11 Histogram diagram of proportion of samples with AFB₁ ≥ 5 μg/kg and ≥ 20 μg/kg in all samples and according interval with different precipitation from 8:00 to 20:00

(3) 平均气压与样品分布关系

各平均气压区间样品的占比可以从图 4.12 中体现。同时，超限样品在高平均气压区域较多，超限样品也以高平均气压区间为主（图 4.13），四分图（图 4.14）分别展示了在每个分区间内超

限样品占区间内所有样品的比例，各平均气压区间均有超限样品分布。因此对平均气压规则制定不明显。

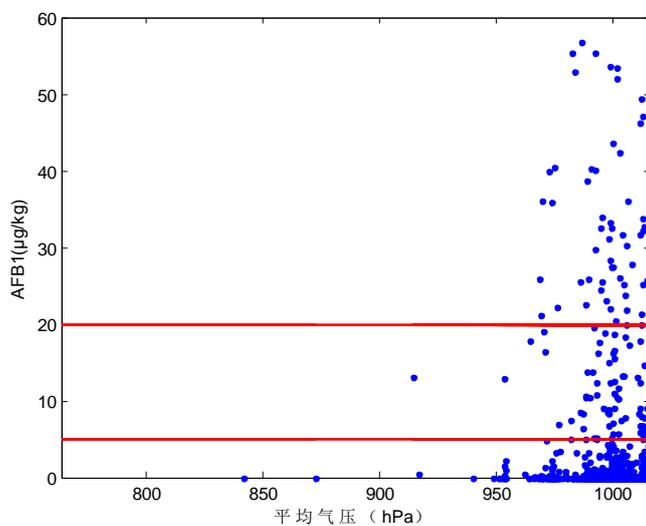


图 4.12 花生 AFB₁ 含量在不同平均气压的散点分布

Figure 4.12 Scatter diagram of AFB₁ in peanuts with different mean air pressure

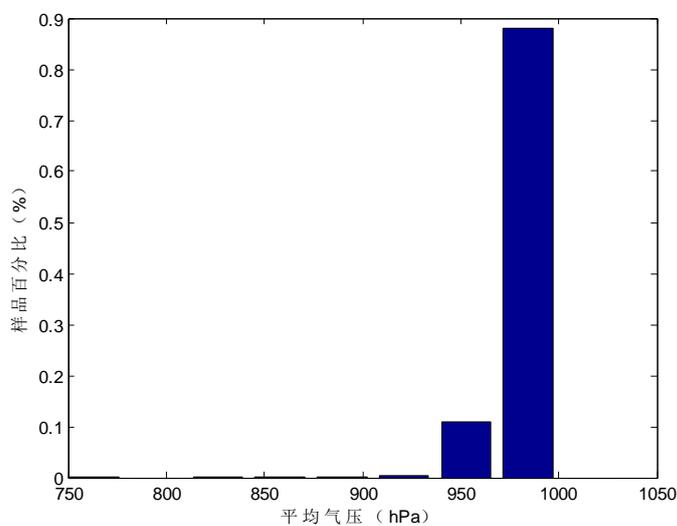


图 4.13 不同平均气压区间样品占比

Figure 4.13 Histogram diagram of AFB₁ with different mean air pressure

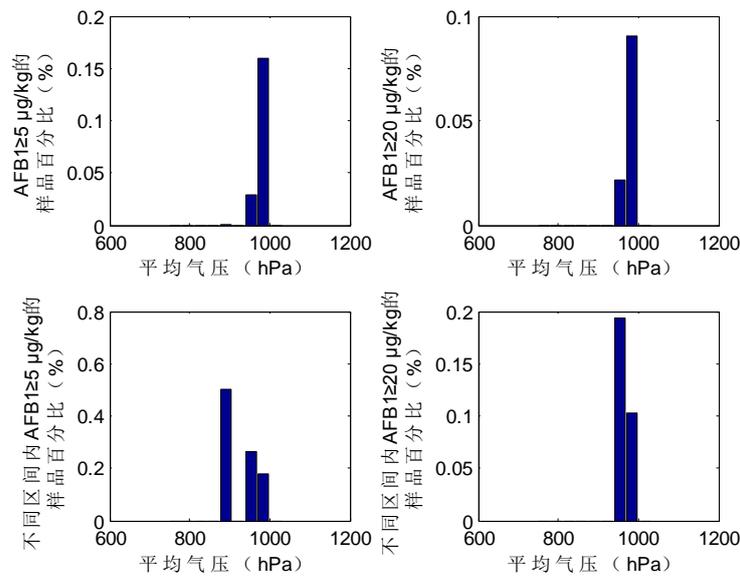


图 4.14 以 5 μg/kg 和 20 μg/kg 为阈值，平均气压区间对应超限样品的比例

Figure 4.14 Histogram diagram of proportion of samples with AFB₁ ≥ 5 μg/kg and ≥ 20 μg/kg in all samples and according interval with different mean air pressure

(4) 日平均气温与样品分布关系

从图 4.15 中可以看出，大部分样品集中在中高温地区，具体各温度区间样品的占比可以从图 4.16 中进一步体现。同时，超限样品呈现一定的集中趋势，超限样品也以中高温地区为主，四分图（图 4.17）分别展示的在每个分区间内超限样品占区间内所有样品的比例，区间样品数量和超限样品占比呈现高度一致性，也就是说在中高温地区样品量最多，这些地区的样品超限率也最高，而低温地区样品超限率为 0。

因此对于日平均气温变量的规则制定为：找到中高温集中区间，将低温地区（日平均气温 ≤ 16.8 ℃）的样品直接确定为未超限样品，减少超限样品规模，也降低误分类概率。

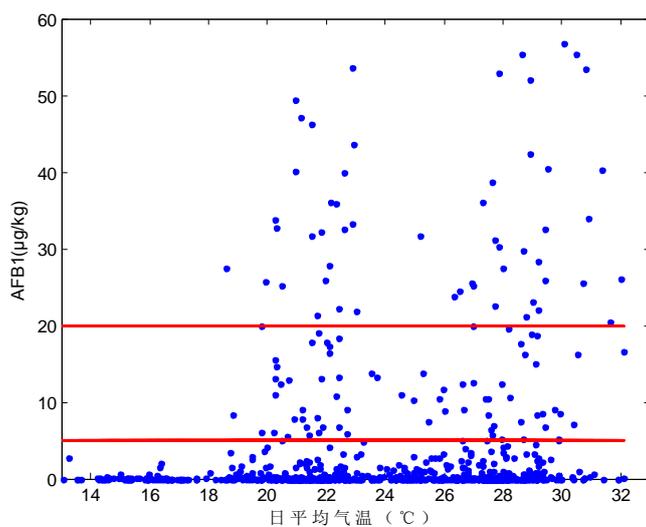
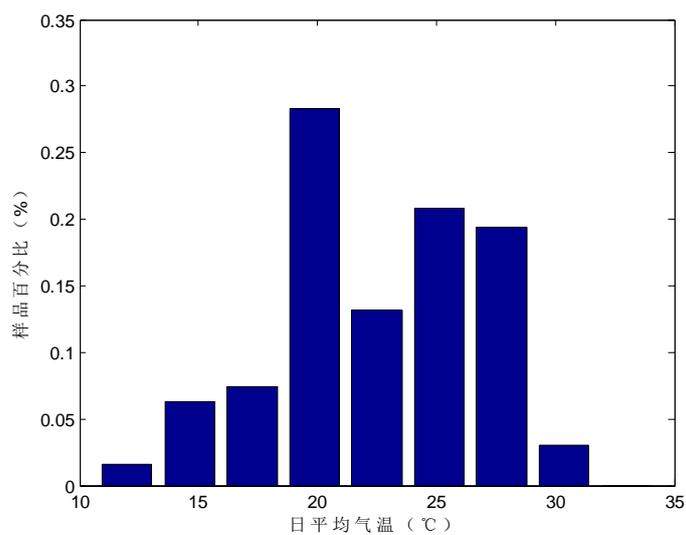
图 4.15 花生 AFB₁ 含量在不同日平均气温的散点分布Figure 4.15 Scatter diagram of AFB₁ in peanuts with different mean air temperature

图 4.16 不同日平均气温区间样品占比

Figure 4.16 Histogram diagram of AFB₁ with different mean air temperature

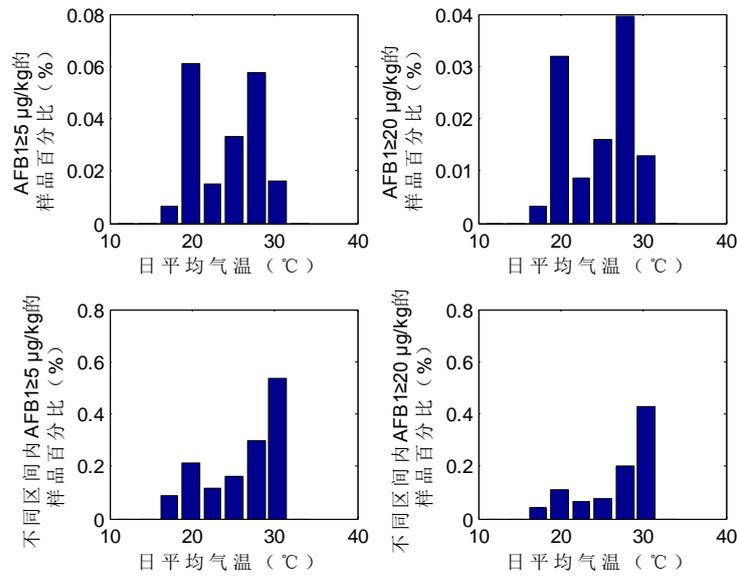


图 4.17 以 5 µg/kg 和 20 µg/kg 为阈值，日平均气温区间对应超限样品的比例

Figure 4.17 Histogram diagram of proportion of samples with AFB₁ ≥ 5 µg/kg and ≥ 20 µg/kg in all samples and according interval with different mean air temperature

总体来说，分别以 5 µg/kg 和 20 µg/kg 为阈值，高纬度区域 (≥40°N)、低纬度区域 (≤21°N) 和低温区域 (日平均气温 ≤ 16.8 °C) 的样品超限率为 0，制定规则效果显著。对于 8-20 时降水量和平均气压，各区间均有超限样品分布，且不存在某一区间无超限样品的情况，因此不适合制定规则。

综合上述分析，只有纬度和日平均气温适合规则的制定，简单易行，便于操作。具体规则为：

(纬度 > 临界高纬度) ∨ (纬度 < 临界低纬度) → 样品未超限

(日平均气温 < 临界低温) → 样品未超限

4.5.2.2 样本偏态分布解决策略

基于规则的分类器虽然可以使未超限样本量适当降低，但是没有从根本上解决超限样本量远低于未超限样本量的问题。下图是以 20 µg/kg 为阈值，分别利用纬度和日平均气温 (图 4.18)、纬度和 8-20 时降水量 (图 4.19) 和纬度与平均气压 (图 4.20) 两两变量绘制的未超标样本和超标样本的分布情况。

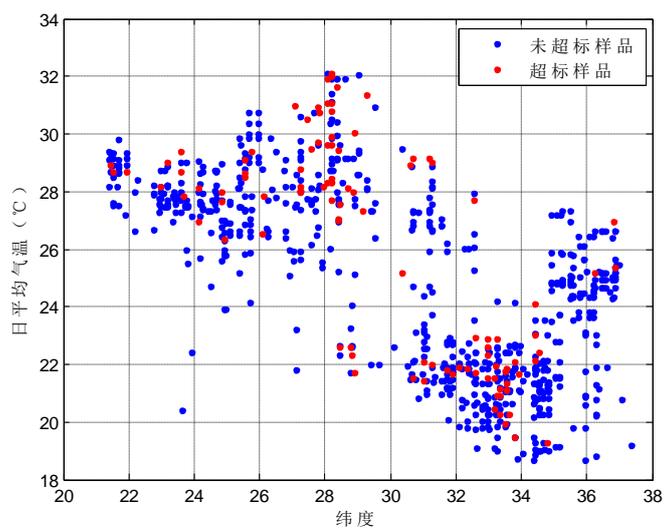


图 4.18 基于纬度和日平均气温变量的样本分布

Figure 4.18 Sample distribution based on latitude and mean air temperature

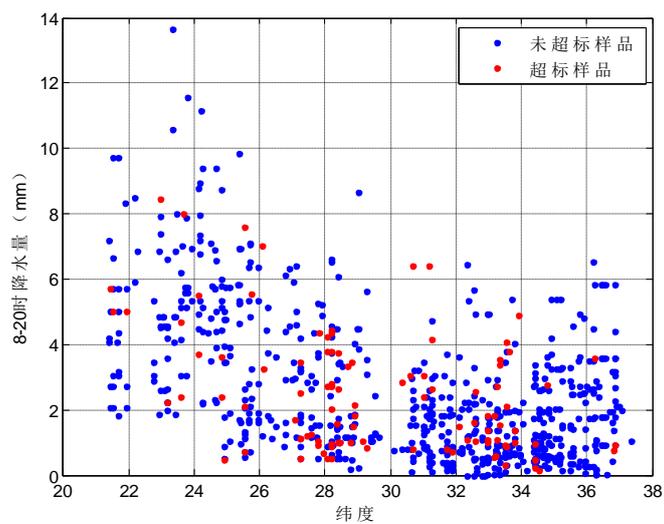


图 4.19 基于纬度和 8-20 时降水量变量的样本分布

Figure 4.19 Sample distribution based on latitude and precipitation from 8:00 to 20:00

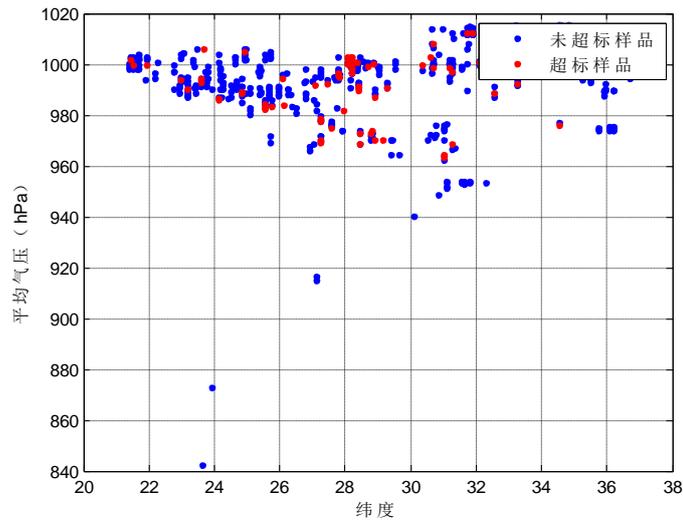


图 4.20 基于纬度和平均气压变量的样本分布

Figure 4.20 Sample distribution based on latitude and mean air pressure

从图中可以看出，超标样本相对集中在两三个区域，但是同时未超标样本也在这部分聚集，由于未超标样本量远超过超标样本，在使用传统分类器的时候，效果无法保证。为了解决这个问题，可以采用基于抽样的方法，复制超标样本，使得超标样本总量与未超标样本量大体相当，甚至超过未超标样本数量。在分类预测的时候增大超标样本的权重，使得在两种样本高度交织的区域内，把更多的样本归结为超标样本，这样可以保证超标样本预测的准确度。

4.5.2.3 随机森林

根据超标样品与未超标样品分布相互交织的特点，基于线性和曲线的分类效果将大打折扣，而随机森林是一类专门为决策树分类器设计的组合算法。它组合多棵决策树做预测，组合分类器一般都优于单个决策树分类器。基于筛选出来的主要气候因子（纬度、8-20 时降水量、平均气压、日平均气温），结合花生品种抗性分类变量、土壤类型分类变量和菌株产毒力分类变量，利用随机森林进行建模。为了说明算法有效性，采用交叉验证的方法。将样本随机分为 10 份，取 8 份做训练样本，生成随机森林，然后 2 份做测试，说明模型的准确性。

计算流程如下：

原始样本 937 个，当 AFB₁ 含量阈值为 20 $\mu\text{g}/\text{kg}$ 时：

- (1) 利用规则分类实现 152 个未超标样本的 0 误差分类；
- (2) 剩余样本 785 个，其中未超标样本 680 个，超标样本 105 个；
- (3) 重复复制超标样品 5 次，超标样本达到 630 个，整体量与未超标样本相当；
- (4) 利用随机森林（森林规模为 50），进行分类预测。

为了预测花生样品的 AFB₁ 含量是否超过 20 $\mu\text{g}/\text{kg}$ ，随机森林分类模型分别选择 252 和 63 个样品作为训练集和测试集。训练样本的混淆矩阵结果见表 4.6，训练数据集中，未超标样本分类预测准确率为 98.64%，超标样本分类准确率为 100.00%，总的样本分类预测准确率为 99.21%。测试样本的混淆矩阵结果见表 4.7，测试数据集中，未超标样本分类预测准确率为 94.44%，超标

样本分类准确率为 100.00%，总的样本分类预测准确率为 96.83%。综合规则分类样本，未超标样本预测准确率为 98.94%，超标样本分类准确率为 100.00%，总的样本分类预测准确率为 99.07%。

表 4.6 当 AFB₁ 含量阈值 20 μg/kg 时，训练样本的混淆矩阵

Table 4.6 The OOB error of training set by RF classification of samples with AFB₁ threshold of 20 μg/kg

| | AFB ₁ 含量 < 20 μg/kg | AFB ₁ 含量 ≥ 20 μg/kg | OOB 误判率 |
|--------------------------------|--------------------------------|--------------------------------|---------|
| AFB ₁ 含量 < 20 μg/kg | 145 | 2 | 0.014 |
| AFB ₁ 含量 ≥ 20 μg/kg | 0 | 105 | 0.000 |
| 总计 | | | 0.008 |

表 4.7 当 AFB₁ 含量阈值 20 μg/kg 时，测试样本的混淆矩阵

Table 4.7 The OOB error of testing set by RF classification of samples with AFB₁ threshold of 20 μg/kg

| | AFB ₁ 含量 < 20 μg/kg | AFB ₁ 含量 ≥ 20 μg/kg | OOB 误判率 |
|--------------------------------|--------------------------------|--------------------------------|---------|
| AFB ₁ 含量 < 20 μg/kg | 34 | 2 | 0.056 |
| AFB ₁ 含量 ≥ 20 μg/kg | 0 | 27 | 0.000 |
| 总计 | | | 0.032 |

当 AFB₁ 含量阈值 5 μg/kg 时：

- (1) 利用规则分类实现 131 个未超限样本的 0 误差分类；
- (2) 剩余样本 806 个，其中未超限样本 629 个，超限样本 177 个；
- (3) 重复复制超限样品 3 次，超限样本达到 708 个，整体量与未超限样本相当；
- (4) 利用随机森林（森林规模为 50），进行分类预测。

为了预测花生样品的 AFB₁ 含量是否超过 5 μg/kg，随机森林分类模型分别选择 275 和 68 个样品作为训练集和测试集。准确率和以 20 μg/kg 为阈值时大体相当。训练样本的混淆矩阵结果见表 4.8，训练数据集中，未超限样本分类预测准确率为 97.08%，超限样本分类准确率为 100.00%，总的样本分类预测准确率为 98.55%。测试样本的混淆矩阵结果见表 4.9，测试数据集中，未超限样本分类预测准确率为 85.29%，超限样本分类准确率为 100.00%，总的样本分类预测准确率为 92.65%。综合规则分类样本，未超限样本预测准确率为 96.97%，超限样本分类准确率为 100.00%，总的样本分类预测准确率为 97.49%。

表 4.8 当 AFB₁ 含量阈值 5 μg/kg 时，训练样本的混淆矩阵

Table 4.8 The OOB error of training set by RF classification of samples with AFB₁ threshold of 5 μg/kg

| | AFB ₁ 含量 < 5 μg/kg | AFB ₁ 含量 ≥ 5 μg/kg | OOB 误判率 |
|-------------------------------|-------------------------------|-------------------------------|---------|
| AFB ₁ 含量 < 5 μg/kg | 133 | 4 | 0.029 |
| AFB ₁ 含量 ≥ 5 μg/kg | 0 | 138 | 0.000 |
| 总计 | | | 0.015 |

表 4.9 当 AFB₁ 含量阈值 5 μg/kg 时，测试样本的混淆矩阵

Table 4.9 The OOB error of testing set by RF classification of samples with AFB₁ threshold of 5 μg/kg

| | AFB ₁ 含量 < 5 μg/kg | AFB ₁ 含量 ≥ 5 μg/kg | OOB 误判率 |
|-------------------------------|-------------------------------|-------------------------------|---------|
| AFB ₁ 含量 < 5 μg/kg | 29 | 5 | 0.147 |
| AFB ₁ 含量 ≥ 5 μg/kg | 0 | 34 | 0.000 |
| 总计 | | | 0.074 |

4.5.2.4 输入参数在分类模型中的贡献率排序

根据以上分析,已经基于规则分类—平衡取样—随机森林方法得出了主要气候因子(纬度、8-20时降水量、平均气压和日平均气温)、品种抗性、土壤类型及真菌产毒力对花生 AFB₁ 含量平均值的分类预测模型。进而通过在分类器中的选择频率对输入参数进行排序,来确定它们对花生 AFB₁ 分类预测的相对贡献率顺序。由图 4.21 和图 4.22 可以看出,当以 20 $\mu\text{g}/\text{kg}$ 和 5 $\mu\text{g}/\text{kg}$ 为阈值时,7 个模型变量在随机森林分类模型中的贡献率排序依次为:纬度>日平均气温>8-20 时降水量>平均气压>品种抗性>土壤类型>真菌产毒力。这是由于在收获前 1 个月,花生黄曲霉毒素污染主要受气候因素的影响,在此阶段其它因素的影响程度相对较小。

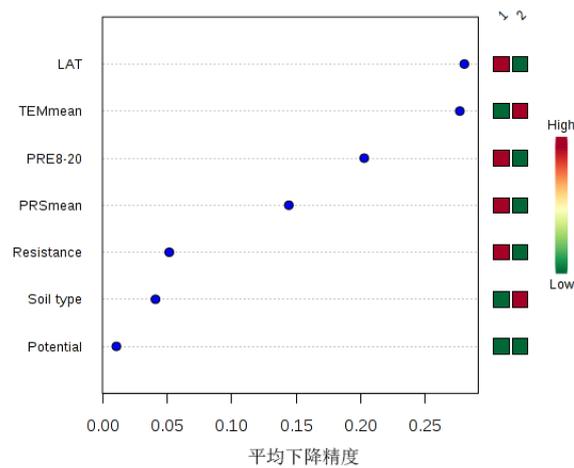


图 4.21 基于随机森林分类模型的变量贡献率排序 (以 20 $\mu\text{g}/\text{kg}$ 为阈值对花生 AFB₁ 含量进行分类)

Figure 4.21 Features ranked by their contributions to classification accuracy (Classification of AFB₁ content in peanut with threshold value of 20 $\mu\text{g}/\text{kg}$)

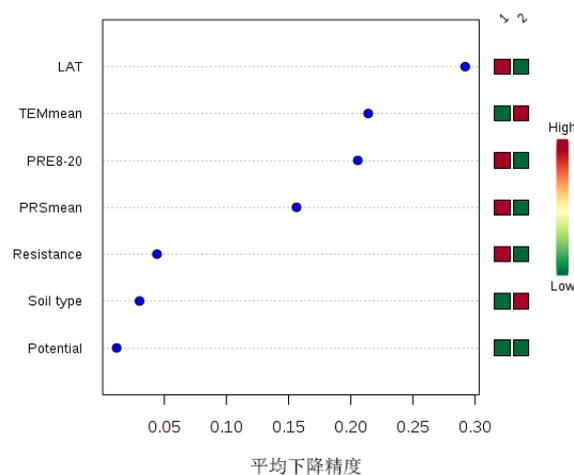


图 4.22 基于随机森林分类模型的变量贡献率排序 (以 5 $\mu\text{g}/\text{kg}$ 为阈值对花生 AFB₁ 含量进行分类)

Figure 4.22 Features ranked by their contributions to classification accuracy (Classification of AFB₁ content in peanut with threshold value of 5 $\mu\text{g}/\text{kg}$)

4.5.3 随机森林分类模型与其它分类模型精度比较

为考察所构建的随机森林分类模型精确度,以 2009-2017 年的数据为基础,80%的数据作为训练集,20%的样品作为测试集,分别采用决策树方法(DT)、支持向量机(SVM)、k-邻近算法(KNN)、BP 神经网络(BPNN)、RBF 神经网络(RBFNN)对数据进行模型拟合,观测同种数据系列下不同方法的预测精度。当以 20 $\mu\text{g}/\text{kg}$ 和 5 $\mu\text{g}/\text{kg}$ 为分类阈值时,测试集预测结果见表 4.10 和表 4.11。

由表 4.10 可知,不同方法对超标样品与未超标样品分类结果参差不齐。当 AFB₁ 含量阈值为 20 $\mu\text{g}/\text{kg}$ 时,对未超标样品来说,KNN 和 RBFNN 的分类准确率最高为 100.00%,其次是 RF (94.44%)、SVM (94.44%) 和 DT (88.89%),BPNN 分类的准确率最低为 80.56%;对超标样品来说,RF 的分类准确率最高为 100.00%,其次是 DT (88.89%)、BPNN (81.48%) 和 SVM (11.110%),RBFNN 和 KNN 的分类准确率最低为 0.00%;总体来说,RF 的分类准确率最高为 96.83%,其次是 DT (88.89%)、BPNN (80.95%)、SVM (58.73%),KNN 和 RBFNN 分类的准确率最低均为 57.14%。当 AFB₁ 含量阈值为 5 $\mu\text{g}/\text{kg}$ 时,结论与分类阈值为 20 $\mu\text{g}/\text{kg}$ 结论近似(表 4.11)。对未超限样品来说,RBFNN 的分类准确率最高为 100.00%,其次是 SVM (97.06%)、RF (85.29%)、DT (76.47%)、BPNN (73.53%),KNN 的分类的准确率最低为 0.00%;对超限样品来说,RF 和 KNN 的分类准确率最高为 100.00%,其次是 DT (94.12%)、BPNN (82.35%)、SVM (8.82%),RBFNN 的分类准确率最低为 0.00%;总体来说,RF 的分类准确率最高为 92.65%,其次是 DT (85.29%)、BPNN (77.94%)、SVM (52.94%),KNN 和 RBFNN 的分类的准确率最低为 50.00%。

由于超限样品与未超限样品数据分布相互交织,BPNN 和 RBFNN 对未超限样品分类效果较差,SVM 这类基于线性和曲线的分类划分效果也大打折扣,因此基于决策树分类器(DT、KNN 和 RF)的方法分类结果较好。经过对比尝试,结果表明随机森林是分类效果最为理想的方法。因为随机森林是一类专门为决策树分类器设计的组合算法。它组合多棵决策树做预测,组合分类器一般都优于单个决策树分类器。

表 4.10 当 AFB₁ 含量阈值为 20 µg/kg 时，随机森林分类模型与其它分类模型精度比较

Table 4.10 The error of testing set by different classification methods of samples with AFB₁ threshold of 20 µg/kg

| | DT | | | SVM | | | KNN | | | BPNN | | | RBFNN | | | RF | | |
|----------------|----------------|----------------|-------|-----|---|-------|-----|---|-------|------|----|-------|-------|---|-------|----|----|-------|
| | 1 ^a | 2 ^b | 误判率 | 1 | 2 | 误判率 | 1 | 2 | 误判率 | 1 | 2 | 误判率 | 1 | 2 | 误判率 | 1 | 2 | 误判率 |
| 1 ^a | 32 | 4 | 0.111 | 34 | 2 | 0.056 | 36 | 0 | 0.000 | 29 | 7 | 0.194 | 36 | 0 | 0.000 | 34 | 2 | 0.056 |
| 2 ^b | 3 | 24 | 0.111 | 24 | 3 | 0.889 | 27 | 0 | 1.000 | 5 | 22 | 0.185 | 27 | 0 | 1.000 | 0 | 27 | 0.000 |
| 总 | | | 0.111 | | | 0.413 | | | 0.429 | | | 0.190 | | | 0.429 | | | 0.032 |

a: 0 表示 AFB₁ 含量 < 20 µg/kg; b: 1 表示 AFB₁ 含量 ≥ 20 µg/kg

表 4.11 当 AFB₁ 含量阈值为 5 µg/kg 时，随机森林分类模型与其它分类模型精度比较

Table 4.11 The error of testing set by different classification methods of samples with AFB₁ threshold of 5 µg/kg

| | DT | | | SVM | | | KNN | | | BPNN | | | RBFNN | | | RF | | |
|----------------|----------------|----------------|-------|-----|---|-------|-----|----|-------|------|----|-------|-------|---|-------|----|----|-------|
| | 1 ^a | 2 ^b | 误判率 | 1 | 2 | 误判率 | 1 | 2 | 误判率 | 1 | 2 | 误判率 | 1 | 2 | 误判率 | 1 | 2 | 误判率 |
| 1 ^a | 26 | 8 | 0.235 | 33 | 1 | 0.029 | 0 | 34 | 1.000 | 25 | 9 | 0.265 | 34 | 0 | 0.000 | 29 | 5 | 0.147 |
| 2 ^b | 2 | 32 | 0.059 | 31 | 3 | 0.912 | 0 | 34 | 0.000 | 6 | 28 | 0.176 | 34 | 0 | 1.000 | 0 | 34 | 0.000 |
| 总 | | | 0.147 | | | 0.471 | | | 0.500 | | | 0.221 | | | 0.500 | | | 0.074 |

a: 0 表示 AFB₁ 含量 < 5 µg/kg; b: 1 表示 AFB₁ 含量 ≥ 5 µg/kg

4.5.4 模型验证

利用建立的模型对 2018 年采自 61 个花生主产县的 1181 份花生样品进行验证。

计算流程如下：

原始样本共 61 个，当 AFB₁ 含量阈值 20 μg/kg 时：

- (1) 利用规则分类实现 10 个未超标样本的 0 误差分类；
- (2) 剩余样本 51 个；
- (3) 利用建立的随机森林预警模型，对验证样本是否超标进行分类预测。

验证样本的混淆矩阵结果见表 4.12，验证数据集中，未超标样本分类预测准确率为 84.44%，超标样本分类准确率为 83.33%，总的样本分类预测准确率为 84.31%。综合规则分类样本，未超标样本预测准确率为 87.27%，超标样本分类准确率为 83.33%，总的样本分类预测准确率为 86.89%。

表 4.12 当 AFB₁ 含量阈值 20 μg/kg 时，验证样本的混淆矩阵

Table 4.12 The OOB error of validation set by RF classification of samples with AFB₁ threshold of 20 μg/kg

| | AFB ₁ 含量 < 20 μg/kg | AFB ₁ 含量 ≥ 20 μg/kg | OOB 误判率 |
|--------------------------------|--------------------------------|--------------------------------|---------|
| AFB ₁ 含量 < 20 μg/kg | 38 | 7 | 0.156 |
| AFB ₁ 含量 ≥ 20 μg/kg | 1 | 5 | 0.167 |
| 总计 | | | 0.157 |

当 AFB₁ 含量阈值 5 μg/kg 时：

- (1) 利用规则分类实现 9 个未超限样本的 0 误差分类；
- (2) 剩余样本 52 个；
- (3) 利用建立的随机森林预警模型，对验证样本是否超限进行分类预测。

验证样本的混淆矩阵结果见表 4.13，验证数据集中，未超限样本分类预测准确率为 83.33%，超限样本分类准确率为 90.00%，总的样本分类预测准确率为 84.62%。综合规则分类样本，未超限样本预测准确率为 86.27%，超限样本分类准确率为 90.00%，总的样本分类预测准确率为 86.89%。

表 4.13 当 AFB₁ 含量阈值 5 μg/kg 时，验证样本的混淆矩阵

Table 4.13 The OOB error of validation set by RF classification of samples with AFB₁ threshold of 5 μg/kg

| | AFB ₁ 含量 < 5 μg/kg | AFB ₁ 含量 ≥ 5 μg/kg | OOB 误判率 |
|-------------------------------|-------------------------------|-------------------------------|---------|
| AFB ₁ 含量 < 5 μg/kg | 35 | 7 | 0.167 |
| AFB ₁ 含量 ≥ 5 μg/kg | 1 | 9 | 0.100 |
| 总计 | | | 0.154 |

根据验证集（2018 年数据）预测结果可知，模型准确率较训练集与测试集降低，其可能与样品分布不均匀，取样与分析检测误差以及收获后的处理方式相关。

4.6 讨论

自 1960 年“火鸡 X 病”事件后，花生黄曲霉毒素污染问题受到了全世界的广泛关注，但如

何对其进行有效控制仍然是一个难题,开展基于品种抗性、土壤类型、真菌产毒力、地理信息、气候气象等多因素的花生黄曲霉毒素大尺度风险预警模型及基于产毒相关标识物的分子预警技术研究,是花生黄曲霉毒素污染控制的高效手段,也是今后黄曲霉毒素污染控制技术研究的一个新方向和热点。通过提前预警、科学防控,可以有效减少花生及其制品黄曲霉毒素污染,保障花生消费安全。

在模型目标变量的选择上,本章选择对花生 AFB₁ 含量进行预测,首先是因为中国花生限量标准是根据 AFB₁ 的含量进行制定。其次,丁小霞(2011)研究发现 AFB₁ 和黄曲霉毒素总量具有很好的相关性,相关系数达到 0.99。再次,中国花生主要受黄曲霉(*A.flavus*)的侵染,原因如下:(1)黄曲霉(*A.flavus*)主要产生 B 族黄曲霉毒素,寄生曲霉(*A.parasiticus*)和集峰曲霉(*A.nomius*)不仅产生 B 族黄曲霉毒素,还产生 G 族黄曲霉毒素(Ehrlich 2003; Ehrlich *et al.* 2004; Nagarajan *et al.* 1973),因此根据不同的黄曲霉毒素污染类型可以间接确定中国花生受侵染的产毒真菌类型;(2)在黄曲霉毒素阳性样品中, AFB₁ 的检出率最高为 89.12%,其次是 AFB₂、AFG₂ 和 AFG₁,阳性样品检出率分别为 59.60%、13.73%和 7.70%;(3)在黄曲霉毒素阳性样品中,同时污染有 AFB₁ 和 AFB₂,是最常见的污染类型,约占阳性样品的 37.99%。其次为只检出 AFB₁ (36.40%) (见图 4.23),即绝大多数样品的黄曲霉毒素污染类型与寄生曲霉(*A. parasiticus*)和集峰曲霉(*A.nomius*)的性质不符。这些均表明中国花生易受黄曲霉(*A.flavus*)的侵染,是产生黄曲霉毒素的主要菌种。各国的研究也表明,花生和花生制品中分离的产黄曲霉毒素菌种主要为黄曲霉(*A.flavus*) (Joffe 1969; Lisker *et al.* 1993)。Mutitu 等(2004)研究表明寄生曲霉(*A. parasiticus*)和黄曲霉(*A.flavus*)是从肯尼亚市售脱壳花生中分离到的主要黄绿组真菌。在来自西非的 436 份花生样品中,黄曲霉(*A.flavus*)是主要的黄绿组真菌 (Mutege *et al.* 2012)。在巴西,花生中分离得到的主要黄绿组真菌为黄曲霉(*A.flavus*)和寄生曲霉(*A. parasiticus*) (Martins *et al.* 2017)。在印度 (Mohana *et al.* 2016) 和巴基斯坦 (Jabeen *et al.* 2012),花生产品中分离得到的黄绿组真菌只有黄曲霉(*A.flavus*)。这些研究表明黄曲霉(*A.flavus*)是侵染花生的主要产毒真菌,其次是寄生曲霉(*A. parasiticus*),而在亚洲寄生曲霉(*A. parasiticus*)基本很少分布。已有研究表明在中国,黄曲霉(*A.flavus*)是种植花生的土壤中的主要产黄曲霉毒素真菌(表 4.14)。张初署(2013)研究表明,在分离得到的花生土壤产毒真菌中,黄曲霉(*A.flavus*)数量占 94.2%,而寄生曲霉只占 5.8%。尽管在长江流域地区,高产毒黄曲霉(*A.flavus*)较少,低产毒和不产毒菌株较多,但是黄曲霉毒素污染却是很严重的,并且存在显著的年度间差异 (Ding *et al.* 2012; Ding *et al.* 2014)。朱婷婷(2018)对采自我国四大花生主产区的 580 份花生根际土壤进行了菌株分离、纯化以及形态学和分子生物学方法鉴定,分离得到黄曲霉(*A.flavus*) 470 株(占所鉴定菌株的 98%),寄生曲霉菌 7 株。因此中国花生受 G 族黄曲霉毒素污染程度较低,最终选择对 AFB₁ 含量进行预测而非黄曲霉毒素总量。

表 4.14 中国不同花生产区土壤中黄曲霉的菌落数与产毒力水平

Table 4.14 CFU (Colony forming units) and aflatoxin producing potential of *Aspergillus flavus* isolated from soils of different peanut production regions in China

| 地区 | CFU/g | 黄曲霉数量(%) | 不同产毒力水平黄曲霉数量(%) | | | |
|-----------------|-------|-----------|-----------------|----------|----------|----------|
| | | | 不产毒 | 低产毒 | 中产毒 | 高产毒 |
| 辽宁 ^a | 11 | 9(81.8) | 4(36.4) | 3(27.3) | 1(9.1) | 3(27.3) |
| 山东 ^a | 85 | 73(85.9) | 24(28.2) | 30(35.3) | 4(4.7) | 27(31.8) |
| 湖北 ^a | 127 | 126(99.1) | 51(40.2) | 46(36.2) | 17(13.4) | 13(10.2) |
| 广东 ^a | 121 | 117(96.6) | 19(15.7) | 26(21.5) | 9(7.4) | 67(55.4) |

^a摘自张 (2013)

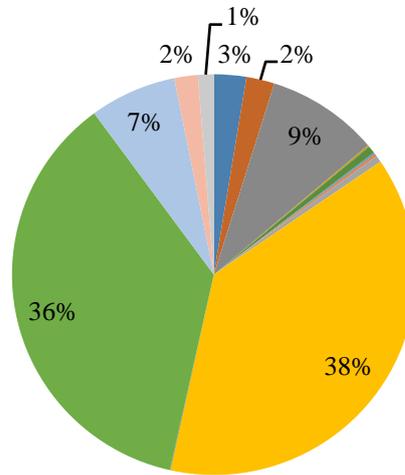


图 4.23 2009-2017 年中国花生污染黄曲霉毒素污染类型占比

Figure 4.23 Proportion of aflatoxin chemotype co-occurred in peanuts collected from China during 2009-2017

在收获前对黄曲霉毒素进行预警进而采取预防性农艺措施, 不仅省时省力而且大大减少了产后脱毒的费用。但是由于真菌与农作物之间的关系具有特异性, 且黄曲霉毒素与气象因子和农艺措施之间关系复杂, 收获前预警模型研究较少且只集中在几个国家或地区, 在花生上的研究更是有限, 因此, 为了应对气候变化对食品安全的影响及减少收获前黄曲霉毒素污染, 迫切需要建立针对不同地区和不同作物的黄曲霉毒素预警模型。

4.7 本章小结

本章首先以 2009-2017 年的数据为基础, 分 3 步对气象参数进行变量筛选, 通过降维减少变量个数。(1) 通过相关性 (Pearson、Spearman 和 Kendall) 分析可知, 各气候因子与花生 AFB₁ 之间的线性相关系数都非常低 (多数都不足 0.1), 基本可以认定各气候属性变量与 AFB₁ 含量之间不存在线性相关; (2) 通过 MIC 对所有气候变量与目标变量之间进行非线性相关分析, 可知

其在一定程度上反映各因素与 AFB₁ 含量之间的相关性。根据 MIC 系数值排序, 考虑因素的全面性, 初选得到的 8 个变量为: 纬度、平均地表气温、日最低地表气温、8-20 时降水量、平均气压、日最低气压、日平均气温和日最高气温; (3) 基于线性相关性对初步筛选的 8 个变量进行再次筛选, 在考虑 MIC 系数值排序和地理位置、降水量、气温等因素的全面性基础上, 最终选择建模气候变量为: 纬度、8-20 时降水量、平均气压、日平均气温。

其次, 基于规则制定对样本进行预分类。根据 4 个主要气候变量与总样本及 AFB₁ 污染样本之间的分布情况, 选取自变量区间样本数量和超标样本占比呈现高度一致性, 且存在某一区间无超限样品的情况, 进行规则制定。最终高纬度区域 ($\geq 40^{\circ}\text{N}$)、低纬度区域 ($\leq 21^{\circ}\text{N}$) 与低温区域 (日平均气温 $\leq 16.8^{\circ}\text{C}$) 制定规则效果显著, 8-20 时降水量和平均气压不适合制定规则。具体规则为:

(纬度 $>$ 临界高纬度) \vee (纬度 $<$ 临界低纬度) \rightarrow 样品未超限

(日平均气温 $<$ 临界低温) \rightarrow 样品未超限

最后, 以筛选得到的 4 个主要气候因子 (纬度、8-20 时降水量、平均气压、日平均气温)、花生品种抗性、土壤类型及真菌产毒力为输入参数建模。针对超限样本量远低于未超限样本量的问题, 对超限样品重复取样后利用随机森林进行分类预测, 模型森林规模为 50。综合规则分类样本, 交叉验证结果表明, 当分类阈值为 20 $\mu\text{g}/\text{kg}$ 时, 未超标样本预测准确率为 98.94%, 超标样本分类准确率为 100.00%; 当分类阈值为 5 $\mu\text{g}/\text{kg}$ 时, 未超限样本预测准确率为 96.97%, 超限样本分类准确率为 100.00%。基于得到的规则分类—平衡取样—随机森林模型, 进一步分析了输入参数在分类模型中的贡献率排序。结果表明, 当以 20 $\mu\text{g}/\text{kg}$ 和 5 $\mu\text{g}/\text{kg}$ 为阈值时, 7 个参数在随机森林分类模型中贡献率排序均为: 纬度 $>$ 日平均气温 $>$ 8-20 时降水量 $>$ 平均气压 $>$ 品种抗性 $>$ 土壤类型 $>$ 真菌产毒力。同其他常用分类器 (DT、SVM、KNN、BPNN、RBFNN) 相比较, RF 能有效分离超限与未超限样品, 且具有较高的预测精度, 因此最终确定 RF 为适合中国花生黄曲霉毒素预警的建模方法。为了验证模型的有效性, 对 2018 年收获的花生 AFB₁ 含量进行预测, 结果表明综合规则分类样本, 当分类阈值为 20 $\mu\text{g}/\text{kg}$ 和 5 $\mu\text{g}/\text{kg}$ 时, 样本分类准确率均为 86.89%。

第五章 全文结论

1. 基于地统计分析与克里金插值绘制了黄曲霉毒素污染分布地图, 探明中国花生 AFB₁ 呈现显著的时间和空间差异。主要结论如下:

(1) 地统计分析结果表明花生 AFB₁ 平均值的 $C/(C+C_0)$ 的值在 50-70% 之间, 具有中等程度的空间自相关性, 表明花生黄曲霉毒素污染受结构性因素 (如土壤类型、土壤母质、气候、地形地貌等) 和随机性因素 (如花生品种、农业土地利用方式、耕作管理措施、施肥及取样分析误差等) 共同作用, 且受气候条件影响作用相对较大。

(2) 2009-2017 年花生 AFB₁ 半变异模型为球状模型, 2017 年的 AFB₁ 含量空间变异性最大, 基台值为 6.77, 2016 年 AFB₁ 含量空间变异性最小, 基台值为 1.05。且模型中块金值均小于基台值, 表明 AFB₁ 含量受县市间距离的强烈影响。2016 年的变程最小为 107000 米, 而 2017 年的变程最大为 511000 米。最短变程可作为研究区域内监测站间隔的布设依据, 由于采样点的最佳布点距离应该为变程值的 1/4 至 1/2, 因此布点间隔应考虑设置为 26750-53500 m。

(3) 2009-2017 年中国花生 AFB₁ 平均值的克里金插值图表明 AFB₁ 污染风险呈现离散分布的特点, 东北主产区的黄曲霉毒素污染风险低, 而长江流域主产区和南方主产区的黄曲霉毒素污染风险较高。这与不同地区的花生品种、土壤类型、菌种产毒力、气候条件等因素相关。确定常年具有黄曲霉毒素污染高风险地区为重点防控地区来实施防控措施, 例如与不易污染黄曲霉毒素的作物进行轮作, 种植黄曲霉毒素高抗品种, 及时收获与灌溉, 利用不产毒黄曲霉菌株进行生物防控等。

2. 分析了 2009-2017 连续九年花生黄曲霉毒素污染消长变化趋势, 并解析了不同花生品种抗性、土壤类型、菌株产毒力及气候条件对其产生的影响。主要结论如下:

(1) 黄曲霉毒素污染存在年度间和地区间差异, 2017、2014、2015 和 2013 年黄曲霉毒素污染水高于其它年份, 2009 年和 2010 年黄曲霉毒素污染水平最低。长江流域主产区的花生黄曲霉毒素污染水平最高, 其次是南方主产区和北方主产区, 东北主产区的污染水平最低。

(2) 黄曲霉毒素低抗品种、砂土土质、高产毒力菌株在超标样品中占比较大, 高温干旱区域与花生污染严重区域重合, 验证了各影响因素对花生黄曲霉毒素污染消长变化的影响。

(3) 收获前 1 个月为花生受黄曲霉毒素污染的关键时期, 得到不同主产区最易受黄曲霉毒素污染的降雨量和气温及可能超标的范围。

通过多元非线性回归确定了收获前 1 个月的平均气候对 AFB₁ 污染水平的影响最大, 并详述了气候条件与花生黄曲霉毒素污染之间的关系。即当降水量低于 4mm 时, 超过最大残留限量 20 $\mu\text{g}/\text{kg}$ 的可能性较大。在干旱条件下, 当日平均气温分别接近 20 $^{\circ}\text{C}$ 、21 $^{\circ}\text{C}$ 和 24 $^{\circ}\text{C}$ 时, 北方主产区、长江流域主产区和南方主产区的花生 AFB₁ 含量增加; 当日平均气温分别为 15-25 $^{\circ}\text{C}$, 17-25 $^{\circ}\text{C}$ 和 21-28 $^{\circ}\text{C}$ 时, 北方主产区、长江流域主产区和南方主产区超过最大残留限量 20 $\mu\text{g}/\text{kg}$ 的可能性较大。当日最低气温分别接近 17 $^{\circ}\text{C}$ 、19 $^{\circ}\text{C}$ 和 21 $^{\circ}\text{C}$ 时, 北方主产区、长江流域主产区和南方主产区的花生 AFB₁ 含量增加; 在干旱条件下, 当日最低气温分别为 13-21 $^{\circ}\text{C}$, 15-24 $^{\circ}\text{C}$ 和 17-26 $^{\circ}\text{C}$ 时, 北方主产区、长江流域主产区和南方主产区超过最大残留限量 20 $\mu\text{g}/\text{kg}$ 的可能性较大。当日最高气温分别接近 27 $^{\circ}\text{C}$ 、31 $^{\circ}\text{C}$ 和 32 $^{\circ}\text{C}$ 时, 北方主产区、长江流域主产区和南方主产区的花生 AFB₁ 含量增加; 在干旱条件下, 当日最高气温分别为 22-29 $^{\circ}\text{C}$, 20-30 $^{\circ}\text{C}$ 和 25-33 $^{\circ}\text{C}$ 时, 北方主产区、

长江流域主产区和南方主产区超过最大残留限量 20 $\mu\text{g}/\text{kg}$ 的可能性较大。为建立收获前花生黄曲霉毒素污染大尺度宏观预警模型奠定了基础。

3. 建立了适合中国的规则分类—平衡取样—随机森林花生黄曲霉毒素预警模型, 并对模型进行了验证, 预测准确率为 86.89%。主要结论如下:

(1) 确定了中国花生黄曲霉毒素预警模型的主要气候变量。以 2009-2017 年数据为基础, 分 3 步对气候参数进行变量筛选, 确定纬度、8-20 时降水量、平均气压、日平均气温为终选气候变量。首先, 通过相关性 (Pearson、Spearman 和 Kendall) 分析可知, 各气候因子与花生 AFB_1 之间不存在线性相关; 其次, 通过 MIC 对所有气候变量与目标变量之间进行非线性相关分析, 得到 8 个初选变量: 纬度、平均地表气温、日最低地表气温、8-20 时降水量、平均气压、日最低气压、日平均气温和日最高气温; 最后, 基于线性相关性对初步筛选的 8 个变量进行再次筛选, 最终选择建模气候变量为: 纬度、8-20 时降水量、平均气压、日平均气温。

(2) 建立了模型预分类规则。通过分析 4 个主要气候变量与总样本及 AFB_1 污染样本分布的关系, 可知高纬度区域 ($\geq 40^\circ\text{N}$)、低纬度区域 ($\leq 21^\circ\text{N}$) 与低温区域 (日平均气温 $\leq 16.8^\circ\text{C}$) 没有超限样本分布, 即为了提高模型精度可将此区域样品直接认定为未超限样品。

(3) 建立平衡取样—随机森林预警模型, 对花生 AFB_1 是否超限预测精度达 95% 以上。以筛选得到的 4 个主要气候因子 (纬度、8-20 时降水量、平均气压、日平均气温)、花生品种抗性、土壤类型及真菌产毒力为输入参数建模。针对超限样本量远低于未超限样本量的问题, 对超限样品重复取样后利用 RF 进行分类预测, 模型森林规模为 50。综合规则分类样本, 交叉验证结果表明, 当分类阈值为 20 $\mu\text{g}/\text{kg}$ 时, 未超标样本预测准确率为 98.94%, 超标样本分类准确率为 100.00%; 当分类阈值为 5 $\mu\text{g}/\text{kg}$ 时, 未超限样本预测准确率为 96.97%, 超限样本分类准确率为 100.00%。同其他常用分类器 (DT、SVM、KNN、BPNN、RBFNN) 相比较, RF 能有效分离超限与未超限样品, 且具有较高的预测精度, 因此最终确定 RF 为适合中国花生黄曲霉毒素预警的建模方法。

(4) 明确了模型输入参数对花生 AFB_1 分类预测的贡献率排序结果。基于得到的规则分类—平衡取样—随机森林模型, 进一步分析了输入参数在分类模型中的贡献率排序。结果表明, 当以 20 $\mu\text{g}/\text{kg}$ 和 5 $\mu\text{g}/\text{kg}$ 为阈值时, 7 个参数在随机森林分类模型中贡献率排序均为: 纬度 > 日平均气温 > 8-20 时降水量 > 平均气压 > 品种抗性 > 土壤类型 > 真菌产毒力。

(5) 初步验证了模型预测效果, 其预测准确率为 86.89%。为了验证模型的有效性, 利用建立的规则分类—平衡取样—随机森林预警模型对 2018 年收获的花生 AFB_1 含量进行预测, 结果表明综合规则分类样本, 当分类阈值为 20 $\mu\text{g}/\text{kg}$ 和 5 $\mu\text{g}/\text{kg}$ 时, 样本分类准确率均为 86.89%。

自 1960 年“火鸡 X 病”事件后, 花生黄曲霉毒素污染问题受到了全世界的广泛关注, 但如何对其进行有效控制仍然是一个难题。本文建立了花生黄曲霉毒素产前预警模型, 实现了利用花生品种抗性、土壤类型、真菌产毒力和产前气候条件对黄曲霉毒素污染进行预测的目的, 但其中仍有可补充完善之处。本文的未来研究思路和存在的不足之处有:

(1) 综合更多因素建立花生黄曲霉毒素产前预警模型。影响花生黄曲霉毒素污染的因素很多, 除了本文建模用到的花生品种抗性、土壤类型、真菌产毒力和产前气候条件, 还包括病虫害以及施肥、灌溉、轮作、套作、覆膜等农艺措施。然而由于中国花生种植分散, 多数农民没有进行程序化生产, 造成许多信息不明或对其采取顺其自然的管理方式。此外已有研究表明 CO_2 浓度也会影响黄曲霉毒素的污染 (Schmidt *et al.* 2010)。因此为了提高产前预警模型的精度, 综合利用

更多的因素对模型进行校正将是行之有效的。

(2) 建立花生黄曲霉毒素全程预警模型。本文建立了花生黄曲霉毒素产前预警模型，然而花生黄曲霉毒素污染在各个环节均有可能发生，实现对花生黄曲霉毒素的全程防控，需针对花生不同环节开展黄曲霉毒素污染预警模型研究。利用产前预警模型在早期发现并预防，结合产后预警模型实施控制，可以有效减少花生及其制品的黄曲霉毒素污染。

(3) 建立适合不同地区的花生黄曲霉毒素预警模型与技术。开展花生黄曲霉毒素污染产前及产后预警研究是可行的，但是不同国家、不同地区、不同环节花生黄曲霉毒素污染影响因素、关键控制点等都不同，即使是同一块农田分离的黄曲霉或寄生曲霉，其产毒力也存在较大差异。因此要建立科学高效的预警模型和技术，需要系统开展不同产地、不同环节花生黄曲霉毒素污染来源、发生机制、迁移转化与累积分布规律研究，探明花生黄曲霉毒素污染关键环节与关键控制点，并基于多年多点花生黄曲霉毒素污染和相关气候环境等大数据对预警模型进行校准和修正，建立适合不同地区、不同环节的花生黄曲霉毒素预警模型与技术。

(4) 建立基于产毒相关标识物的分子预警技术。近年来，随着分子生物学的快速发展，真菌毒素的分子预警技术得到了广泛关注。因此筛选和识别更有效的分子标记物，通过关键基因或次级代谢产物的表达对黄曲霉毒素进行预测，可以有效减少花生及其制品黄曲霉毒素污染，保障花生消费安全。

(5) 建立基于前体物的预警技术。除了 ST 和 VA，黄曲霉毒素合成通路的其它前体物质还包括杂色曲霉毒素 C、奥佛尼红素、奥佛鲁凡素及降散盘衣酸等。这些代谢物的发生通常意味着产毒真菌的污染 (Sulyok *et al.* 2015)，因此基于前体物的预警技术也是降低黄曲霉毒素污染的有效途径和今后的研究方向及重点。

(6) 建立黄曲霉毒素风险交流与管理智能平台。建立便捷友好的平台，如手机应用程序，向农民和风险管理者提供作物信息、污染风险或水平、实际解决方案、问题咨询及反馈、预测模型的输出及建议等关键信息，使农民不仅可以了解作物的田间生长状况，还可以及时、经济、有效地采取措施，预防或补救收获、储存、加工和运输过程中的风险。

综上所述，建立集合高效的风险预警技术对于降低成本、提高效率、最大限度地发挥风险评估和风险预警的作用，以及切实提高农产品质量和安全是非常有必要的。宏观、分子、生态生理和次级代谢产物数据集的有效整合，不仅可以更准确地预测不同地区的黄曲霉毒素污染风险，也为潜在新型毒素污染风险的预测提供了借鉴。

参考文献

1. 蔡骥业. 花生黄曲霉毒素污染及去毒方法综述. 广西农学报, 1996(1): 7-12.
2. 陈小霖. 供应链环境下的农产品质量安全保障研究 [博士学位论文]. 南京: 南京理工大学, 2008.
3. 丁小霞. 中国产后花生黄曲霉毒素污染与风险评估方法研究 [博士学位论文]. 北京: 中国农业科学院, 2011.
4. 方明英. VA 分离、检测与毒性分析及其作为 AFT 预警报告分子可行性研究 [硕士学位论文]. 广州: 暨南大学, 2008.
5. 葛跃, 王明新, 孙向武, 等. 长三角地区秋冬季大气 PM_{2.5} 含量空间变异特征. 环境化学, 2016, 35(08): 1698-1706.
6. 郭安廷, 崔锦霞, 许鑫, 等. 基于 GIS 与地统计的土壤养分空间变异研究. 中国农学通报, 2018, 34(23): 72-79.
7. 郭天文, 赖丽芳, 胡志桥, 等. 甘肃省河西走廊中部新垦荒地土壤养分空间变异特征. 干旱地区农业研究, 2007(04): 151-156.
8. 蒋荻. 基于机器学习的新型数据读出方法研究 [博士学位论文]. 合肥: 中国科学技术大学, 2017.
9. 姜宗亮, 门爱军, 张艺兵. 花生中黄曲霉毒素污染和控制. 北京: 中国标准出版社, 2012.
10. 李瑞芳, 韩北忠, 陈晶瑜. 黄曲霉生长预测模型的建立及毒素含量变化. 食品研究与开发, 2007(12): 129-132.
11. 李新国, 万书波. 钙对花生生长发育调控的研究进展. 山东农业科学, 2011(08): 65-67+74.
12. 李运. 机器学习算法在数据挖掘中的应用 [博士学位论文]. 北京: 北京邮电大学, 2015.
13. 卢志宏. 基于 BP 神经网络的荒漠啮齿动物群落对气候变化滞后响应 [博士学位论文]. 呼和浩特: 内蒙古农业大学, 2016.
14. 吕婕, 吕青, 李成德. 良好农业规范 (GAP) 的现状及应用研究. 安徽农业科学, 2009(12): 5812-5813, 5816.
15. 潘晓红, 陈宇瑶, 史文卫. 富硒黑花生的营养价值与栽培技术. 现代农业科技, 2010(13): 68.
16. 宋莎, 李廷轩, 王永东, 等. 县域农田土壤有机质空间变异及其影响因素分析. 土壤, 2011, 43(01): 44-49.
17. 谭辉勇. VA 作为 AFB₁ 预警指示物的研究及其酶传感器检测方法的探讨 [硕士学位论文]. 广州: 暨南大学, 2010.
18. 王海鸥, 陈守江, 胡志超, 等. 花生黄曲霉毒素污染与控制. 江苏农业科学, 2015, 43(01): 270-273.
19. 王后苗, 廖伯寿. 农作物收获前黄曲霉毒素污染与控制措施. 作物学报, 2012, 38(01): 1-9.
20. 王耀波, 门爱军, 张鹏, 等. 建立防范出口花生黄曲霉毒素污染的 GMP 设想. 花生学报, 2002(2): 18-22.
21. 游克仁. 花生的营养价值及增产措施. 特种经济动植物, 2004(07): 28.

22. 于琬淇. 花生芽的培育与营养价值. 农家参谋, 2017(24): 93.
23. 曾义, 孙国防, 茅力. 运用 HACCP 原理分析花生黄曲霉毒素的预防和控制. 南京医科大学学报: 自然科学版, 2007(4): 402-406.
24. 张初署. 中国四个生态区花生土壤中黄曲霉菌分布、产毒特征及遗传多样性研究 [博士学位论文]. 北京: 中国农业科学院, 2013.
25. 张春新, 李建军. HACCP 在控制花生黄曲霉毒素污染中的应用. 中国国门时报, 2005-02-01.
26. 张国云. 支持向量机算法及其应用研究. 湖南大学, 2006.
27. 张建成, 江晨, 李双铃, 等. 花生黄曲霉毒素污染状况及防控技术. 作物杂志, 2006(1): 67-68.
28. 张耀鹏. 基于决策树算法的扼流适配变压器故障诊断系统研究与设计. 北京交通大学, 2018.
29. 张毅, 华福平, 童燕, 等. 花生黄曲霉毒素的危害及预防. 现代农业科技, 2010(1): 340-343.
30. 赵贵兴, 陈霞, 刘昊飞, 等. 花生的功能成分、营养价值及其开发利用研究. 安徽农学通报(下半月刊), 2011, 17(12): 39-42.
31. 赵明松, 张甘霖, 王德彩, 等. 徐淮黄泛平原土壤有机质空间变异特征及主控因素分析. 土壤学报, 2013, 50(01): 1-11.
32. 赵越, 罗志军, 赵杰, 等. 南方丘陵区耕地土壤养分空间特征及其影响因素. 浙江农业学报, 2018, 30(6): 1035-1043.
33. 中华人民共和国卫生部. 食品中真菌毒素限量. GB/T2761-2011.
34. 朱婷婷. 花生土壤中产黄曲霉毒素菌的分布、产毒力与毒素污染研究 [硕士学位论文]. 北京: 中国农业科学院, 2018.
35. Abbas H.K., Weaver M.A., Horn B.W., et al. Special double issue-*Aspergillus*, aflatoxins, cyclopiazonic acid and biological control of aflatoxin. Toxin Reviews, 2011a, 30(2-3): 59-70.
36. Abbas H.K., Williams W.P., Windham G.L., et al. Aflatoxin and fumonisin contamination of commercial corn (*Zea mays*) hybrids in Mississippi. Journal of Agricultural and Food Chemistry, 2002, 50(18): 5246-5254.
37. Abbas H.K., Zablotowicz R.M., Horn B.W., et al. Comparison of major biocontrol strains of non-aflatoxigenic *Aspergillus flavus* for the reduction of aflatoxins and cyclopiazonic acid in maize. Food Additives and Contaminants, 2011b, 28(2), 198-208.
38. Abia W.A., Warth B., Sulyok M., et al. Determination of multi-mycotoxin occurrence in cereals, nuts and their products in Cameroon by liquid chromatography tandem mass spectrometry (LC-MS/MS). Food Control, 2013, 131(2): 438-453.
39. Aguamah G.E., Langcake P., Leworthy D.P., et al. Two novel stilbene phytoalexins from *Arachis hypogaea*. Phytochemistry, 1981, 20: 1381-1383.
40. Ajay B.C., Gowda M.V.C., Rathnakumar A.L., et al. Improving genetic attributes of confectionary traits in peanut (*Arachis hypogaea* L.) using multivariate analytical tools. Journal of Agricultural Science, 2012, 4(3): 247-258.
41. Alaniz Zanon M.S., Chiotta M.L., Gaj-Merlera G., et al. Evaluation of potential biocontrol agent for aflatoxin in Argentinean peanuts. International Journal of Food Microbiology, 2013, 162(3): 220-225.

42. Alkadri D., Rubert J., Prodi A., et al. Natural co-occurrence of mycotoxins in wheat grains from Italy and Syria. *Food Chemistry*, 2014, 157: 111-118.
43. Anderson W.F., Holbrook C.C., Wilson D.M., et al. Evaluation of preharvest aflatoxin contamination in several potentially resistant peanut genotypes. *Peanut Science*, 1995, 22(1): 29-32.
44. Andrade P.D., Mello M.H., Frana J.A., et al. Aflatoxins in food products consumed in Brazil: A preliminary dietary risk assessment. *Food Additives & Contaminants Part A Chemistry Analysis Control Exposure & Risk Assessment*, 2013, 30(1): 127-36.
45. Astoreca A.L., Dalcero A.M., Pinto V.F., et al. A survey on distribution and toxigenicity of *Aspergillus* section *Flavi* in poultry feeds. *International Journal of Food Microbiology*, 2011, 146(1): 38-43.
46. Atehnkeng J., Ojiambo P.S., Ikotun T., et al. Evaluation of atoxigenic isolates of *Aspergillus flavus* as potential biocontrol agents for aflatoxin in maize. *Food Additives and Contaminants Part A*, 2008, 25(10): 1264-1271.
47. Azaizeh H.A., Pettit R.E., Smith O.D., et al. Reaction of peanut genotypes under drought stress to *Aspergillus flavus* and *A. parasiticus*. *Peanut Science*, 1989, 16(2): 109-113.
48. Bankole S., Schollenberger M., Drochner W. Mycotoxins in food systems in Sub Saharan Africa: a review. *Mycotoxin Research*, 2006, 22(3): 163-169.
49. Bara C., Bara L., Bara V., et al. Aflatoxin B₁ and its influence in the liver cancer in the small mammals, 2010: 85-92.
50. Barnett N.M., Naylor A.W. Amino acid and protein metabolism in Bermuda grass during water stress. *Plant Physiology*, 1966, 41(7): 1222-1230.
51. Battilani P., Barbano C., Marin S., et al. Mapping of *Aspergillus* Section *Nigri* in Southern Europe and Israel based on geostatistical analysis. *International Journal of Food Microbiology*, 2006, 111(4): S72-S82.
52. Battilani P., Leggieri M.C., Rossi V., et al. AFLA-maize, a mechanistic model for *Aspergillus flavus* infection and aflatoxin B₁ contamination in maize. *Computers & Electronics in Agriculture*, 2013, 94(11): 38-46.
53. Battilani P., Toscano P., Van der Fels-Klerx H.J., et al. Aflatoxin B₁ contamination in maize in Europe increases due to climate change. *Scientific Reports*, 2016, 6: 24328.
54. Belgiu M., Drăguț L. Random forest in remote sensing: A review of applications and future directions. *Isprs Journal of Photogrammetry & Remote Sensing*, 2016, 114: 24-31.
55. Bircan C., Barringer S.A., Ulken U., et al. Increased aflatoxin contamination of dried figs in a drought year. *Food Additives & Contaminants*, 2008, 25(11): 1400-1408.
56. Bircan C. Comparison of homogenization techniques and incidence of aflatoxin contamination in dried figs for export. *Food Additives & Contaminants Part B*, 2009, 2(2): 171-177.
57. Blankenship P.D., Cole R.J., Sanders T.H. Comparative susceptibility of four experimental peanut lines and the cultivar Florunner to preharvest aflatoxin contamination. *Peanut Science*, 1985, 12(2):

- 70-72.
58. Blankenship P.D., Cole R.J., Sanders T.H., et al. Effect of geocarposphere temperature on pre-harvest colonization of drought-stressed peanuts by *Aspergillus flavus* and subsequent aflatoxin contamination. *Mycopathologia*, 1984, 85(1-2) : 69-74.
 59. Boken V.K., Hoogenboom G., Williams J.H., et al. Monitoring peanut contamination in Mali (Africa) using AVHRR satellite data and a crop simulation model. *International Journal of Remote Sensing*, 2008, 29 (1): 117-129.
 60. Boote K.J., Jones J.W., Hoogenboom G. Simulation of crop growth: CROPGRO model. *Agricultural Systems and Simulation*, 1998(suppl.) : 651-692.
 61. Borgelt, Wieda, Sudduth. Geostatistical analysis of soil chemical properties from nested grids. *Journal of Physical Chemistry*, 1997,13: 477-483.
 62. Brown R.L., Chen Z.Y., Clevel T.E., et al. Advances in the development of host resistance in corn to aflatoxin contamination by *Aspergillus flavus*. *Phytopathology*, 1999, 89(2) : 113-117.
 63. Byrne D.N., Rathman R.J., Orum T.V., et al. Localized migration and dispersal by the sweet potato whitefly, *Bemisia tabaci*. *Oecologia*, 1996, 105(3): 320-328.
 64. Casado M.R., Parsons D.J., Magan N., et al. A short geostatistical study of the three-dimensional spatial structure of fumonisins in stored maize. *World Mycotoxin Journal*, 2010, 3(1): 95-103.
 65. Casado M.R., Parsons D.J., Weightman R.M., et al. Geostatistical analysis of the spatial distribution of mycotoxin concentration in bulk cereals. *Food Additives & Contaminants*, 2009, 26(6): 867-873.
 66. Chala A., Mohammed A., Ayalew A., et al. Natural occurrence of aflatoxins in groundnut (*Arachis hypogaea* L.) from eastern Ethiopia. *Food Control*, 2013, 30(2): 602-605.
 67. Chauhan Y., Tatnell J., Krosch S., et al. An improved simulation model to predict pre-harvest aflatoxin risk in maize. *Field Crops Research*, 2015, 178 :91-99.
 68. Chauhan Y.S., Wright G.C., Rachaputi R.C.N., et al. Application of a model to assess aflatoxin risk in peanuts. *The Journal of Agricultural Science*, 2010, 148(3) : 341.
 69. Chellemi D.O., Rohrbach K.G., Yost R.S., et al. Analysis of the spatial pattern of plant pathogens and diseased plants using geostatistics. *Phytopathology*, 1988,78(2): 221-226.
 70. Codex Alimentarius Commission (CAC). Code of Practice for the prevention and reduction of aflatoxin contamination in peanuts. 2004, CAC/RCP 55.
 71. Codex Alimentarius Commission (CAC). General principles of food hygiene. 2003, CAC/RCP1-1969.
 72. Cole R.J., Blankenship P.D., Hill R.A., et al. Effect of geocarposphere temperature on preharvest colonisation of drought stressed peanuts by *Aspergillus flavus* and subsequent aflatoxin contamination. In Kurata H. & Ueno Y. (Eds), *Toxigenic Fungi-Their Toxins and Health Hazard* 1984: 44-52. Tokyo: Kodansha Ltd.
 73. Cole R.J., Cox R.H. *Handbook of toxic fungal metabolites*, New York, Academic Press, 1981: 1-66.

74. Cole R.J., Dorner J.W., Holbrook C.C. Advances in mycotoxin elimination and resistance. In Pattee H.E. & Stalker H.T. (Eds.), *Advances in peanut science*, 1995: 456-474. Stillwater, OK, USA: American Peanut Research Educational Society.
75. Cole R.J., Hill R.A., Blankenship P.D., et al. Influence of irrigation and drought stress on invasion by *Aspergillus flavus* of corn kernels and peanut pods. *Developments in Industrial Microbiology*, 1982, 23: 229-236.
76. Cole R.J., Sanders T.H., Dorner J.W., et al. Aflatoxin contamination of groundnuts. Environmental conditions required to induce preharvest aflatoxin contamination of groundnuts: Summary of six years' research, 1989: 279-287. Patancheru, India: ICRISAT.
77. Cole R.J., Sanders T.H., Hill R.A., et al. Mean geocarposphere temperatures that induce preharvest aflatoxin contamination of peanuts under drought stress. *Mycopathologia*, 1985, 91(1): 41-46.
78. Cotty P.J., Antilla L. *Managing aflatoxins in Arizona*. USA: U.S. Department of Agriculture, Agricultural Research Service, 2003: 9-13.
79. Cotty P.J. Biocompetitive exclusion of toxigenic fungi. In Barug D., Bhatnagar D., van Egmond H.P., et al (Eds.). *The mycotoxin factbook. Food and feed topics*. 2006: 179-192. The Netherlands: Wageningen Academic Publishers.
80. Cotty P.J., Jaime G.R. Influences of climate on aflatoxin producing fungi and aflatoxin contamination. *International Journal of Food Microbiology*, 2007, 119 (1-2) : 109-115.
81. Craufurd P.Q., Prasad P.V.V., Waliyar. F., et al. Drought, pod yield, preharvest *Aspergillus* infection and aflatoxin contamination on peanut in Niger. *Field Crops Research*, 2006, 98: 20-29.
82. Cressie N. *Geostatistics: a tool for environmental modelers*. 1993.
83. Damianidisa D., Ortizb B.V., Windhamc G.L., et al. Evaluating a generic drought index as a predictive tool for aflatoxin contamination of corn: From plot to regional level. *Crop Protection* 2018, 113: 64-74.
84. Davidson J.I., Hill R.A., Cole R.A., et al. Field performance of two peanut cultivars relative to resistance to invasion by *A. flavus* and subsequent aflatoxin contamination. *Proceedings of the American Peanut Research and Education Society*, 1982, 14: 74-78.
85. de Alencar E.R., Faroni L.R., Soares, N. et al. Efficacy of ozone as a fungicidal and detoxifying agent of aflatoxins in peanuts. *Journal of Science of Food Agricultural*, 2012, 92(4): 899-905.
86. De S.E.M.F., Wang X., Jansen M., et al. Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. *Nature Medicine*, 2016, 19(5): 614-618.
87. Diao E., Dong H., Hou H., et al. Factors Influencing aflatoxin contamination in before and after harvest peanuts: A Review. *Journal of Food Research*, 2014, 4(1): 148.
88. Dickens J.W. Aflatoxin control program for peanuts. *Journal of American Oil Chemical Society*, 1977, 54(3): 225-228.
89. Diener U.L., Davis N.D. Limiting temperature and relative humidity for aflatoxin production by *Aspergillus flavus* in stored peanuts. *Journal of American Oil Chemical Society*, 1970, 47(9):

- 347-351.
90. Ding X.X., Li P.W., Bai Y.Z., et al. Aflatoxin B-1 in post-harvest peanuts and dietary risk in China. *Food Control*, 2012, 23(1): 143-148.
 91. Ding X.X., Wu L.X., Li P.W., et al. Risk assessment on dietary exposure to aflatoxin B₁ in post-harvest peanuts in the Yangtze River ecological region. *Toxins*, 2014, 7: 4157-4174.
 92. D'Mello J.P.F., Macdonald A.M.C., Postel D., et al. Pesticide use and mycotoxin production in *Fusarium* and *Aspergillus* phytopathogens. *European Journal of Plant Pathology*, 1998, 104(8): 741-751.
 93. Dobolyi C., Sebok F., Varga J., et al. Occurrence of aflatoxin producing *Aspergillus flavus* isolates in maize kernel in Hungary. *Acta Alimentaria*, 2013, 42: 451-459.
 94. Dorner J.W. Biological control of aflatoxin contamination of crops. *Journal of Toxicology, Toxin Reviews*, 2004, 23(2&3): 425-450.
 95. Dorner J.W., Cole R.J., Blankenship P. D. Effect of inoculum rate of biological control agents on preharvest aflatoxin contamination of peanuts. *Biological Control*, 1998, 12(3): 171-176.
 96. Dorner J.W., Cole R.J., Connick W. J., et al. Evaluation of biological control formulations to reduce aflatoxin contamination in peanuts. *Biological Control*, 2003, 26(3): 318-324.
 97. Dorner J.W., Cole R.J. Effect of application of nontoxigenic strains of *Aspergillus flavus* and *A. parasiticus* on subsequent aflatoxin contamination of peanuts in storage. *Journal of Stored Products Research*, 2002b, 38(4): 329-339.
 98. Dorner J.W., Cole R.J., Sanders T.H., et al. Interrelationship of kernel water activity, soil temperature, maturity, and phytoalexin production in preharvest aflatoxin contamination of drought-stressed peanuts. *Mycopathologia*, 1989, 105(2): 117-128.
 99. Dorner J.W. Development of biocontrol technology to manage aflatoxin contamination in peanuts. *Peanut Science*, 2009, 36(1): 60-67.
 100. Dorner J.W., Lamb M.C. Development and commercial use of afla-guard, an aflatoxin biocontrol agent. *Mycotoxin Research*, 2006, 22(1): 33-38.
 101. Dorner J.W. Management and prevention of mycotoxins in peanuts. *Food Additives and Contaminants*, 2008, 25(2): 203-208.
 102. Dorner J.W. Simultaneous quantitation of *Aspergillus flavus*/*A. parasiticus* and aflatoxin in peanuts. *Journal of AOAC International*, 2002a, 85(4): 911-916.
 103. Donner M., Lichtemberg P.S.F., Doster M., et al. Community structure of *Aspergillus flavus* and *A. parasiticus* in major almond-producing areas of California, United States. *Plant Disease*, 2015, 99(8): 1161-1169.
 104. Du K., Liu P., Li Y., et al. Effects of dietary mycotoxins on gut microbiome. *Protein & Peptide Letters*, 2017, 24: 1-9.
 105. Ehrlich K.C., Chang P.K., Yu J., et al. Aflatoxin biosynthesis cluster gene *cypA* is required for G aflatoxin formation. *Applied and Environmental Microbiology*, 2004, 70(11): 6518-6524.
 106. Ehrlich K.C., Kobbeman K., Montalbano B.G., et al. Aflatoxin-producing *Aspergillus* species from

- Thailand. *International Journal of Food Microbiology*, 2007, 114(2): 153-159.
107. Ehrlich K. Sequence comparison of aflR from different *Aspergillus* species provides evidence for variability in regulation of aflatoxin production. *Fungal Genetics and Biology*, 2003, 38(1): 63-74.
108. Eiler A., Zaremba-Niedzwiedzka K., Garcia M.M., et al. Productivity and salinity structuring of the microplankton revealed by comparative freshwater metagenomics. *Environmental Microbiology*, 2014,16(9): 2682-2698.
109. Etcheverry M., Nesci A., Passone A. Chemical strategies applied to maize and peanut storage agroecosystem in Argentina to prevent aflatoxin contamination. In Cohen G E, Levin C M (Eds.), *Food storage*, 2011: 65-98. NY, USA: Nova Science Publishers, Inc.
110. European Commission (EC) report. Opinion on the relationship between the use of plant protection products on food plants and the occurrence of mycotoxins in foods. Belgium: European Commission SCP/RESI/063, 1999.
111. Fanelli G., Dantone M., Gall J., et al. Random forests for real time 3D face analysis. *International Journal of Computer Vision*, 2013, 101(3): 437-458.
112. FAOSTAT. Available at <http://www.fao.org/faostat/en/#data/QC.2014>, Accessed 10.01.17.
113. Gachomo E.W., Mutitu E.W., Kotchoni O.S. Diversity of fungal species associated with peanuts in storage and the levels of aflatoxins in infected samples. *International Journal of Agriculture and Biology*, 2004, 6: 955-959.
114. Gorayeb T.C.C., Casciatori F.P., Bianchi V.L., et al. HACCP plan proposal for a typical Brazilian peanut processing company. *Food Control*, 2009, 20 (7) : 671-676.
115. Gottwald T.R., Avinent L., Llacer G., et al. Analysis of the spatial spread of sharka (plum-poxvirus) in apricot and peach orchards in eastern Spain. *Plant Disease*, 1995, 79(3): 266.
116. Grosso N.R., Guzman C.A. Chemical Composition of Aboriginal Peanut (*Arachis hypogaea* L.) Seeds from Peru. *Journal of Agricultural & Food Chemistry*, 1995, 43(1): 102-105.
117. Guo B., Chen X., Dang P., et al. Peanut gene expression profiling in developing seeds at different reproduction stages during *Aspergillus parasiticus* infection. *BMC Developmental Biology*, 2008, 8(1):12.
118. Guo B.Z., Fedorova N.D., Chen X.P., et al. Gene expression profiling and identification of resistance genes to *Aspergillus flavus* infection in peanut through EST and microarray strategies. *Toxins*, 2011, 3(7): 737-753.
119. Heathcote J.G., Hibbert J.R. *Aflatoxin chemical and biological aspects*. Elsevier Scientific Publishing Company, Amsterdam. 1978.
120. Henderson C.E., Potter W.D., McClendon R.W., et al. Predicting aflatoxin contamination in peanuts: a genetic algorithm/neural network approach. *Applied Intelligence*, 2000, 12(3) : 183-192.
121. Hill R.A., Blankenship P.D., Cole R.J., et al. Effect of soil moisture and temperature on preharvest invasion of peanuts by the *Aspergillus flavus* group and subsequent aflatoxin development. *Applied and Environmental Microbiology*, 1983, 45: 628-633.
122. Horn B.W. Colonization of wounded peanut seeds by soil fungi: selectivity for species from

- Aspergillus* section *Flavi*. *Mycologia*, 2005, 97(1): 202-217.
123. Horn B.W., Dorner J.W. Effect of nontoxicogenic *Aspergillus flavus* and *A.parasiticus* on aflatoxin contamination of wounded peanut seeds inoculated with agricultural soil containing natural fungal populations. *Biocontrol Science and Technology*, 2009, 19(3): 249-262.
124. Horn B.W., Dorner J.W. Soil populations of *Aspergillus* species from section *Flavi* along a transect through peanut growing regions of the United States. *Mycologia*, 1998, 90(5): 767-776.
125. Horn B.W., Greene R.L., Dorner J.W. Inhibition of aflatoxin B₁ production by *Aspergillus parasiticus* using nonaflatoxigenic strains: Role of vegetative compatibility. *Biological Control*, 2000, 17(2): 147-154.
126. Houshyarfard M., Rouhani H., Falahatirastegar M., et al. Characterization of *Aspergillus* section *Flavi* from pistachio soils in Iran. *Journal of Plant Protection Research*, 2014, 54(4): 354-362.
127. Huang M.N., Yu W., Teoh W.W., et al. Genome-scale mutational signatures of aflatoxin in cells, mice, and human tumors. *Genome Research*, 2017, 27(9): 1475-1486.
128. Ifeanacho M.O., Agomuo E.N., Amadi P.U. Aflatoxin contamination of various carbohydrate rich foods, legumes, and vegetables, and implications of targeted processing techniques. *Universal Journal of Chemistry*, 2017, 5(4): 59-67.
129. International Agency for Research on Cancer (IARC). Aflatoxins. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans 2002* (82): 171-245. Lyon, France: IARC.
130. Iqbal S.Z., Asi M.R., Akram N., et al. Aflatoxin contamination in different fractions of rice from Pakistan and estimation of dietary intakes. *Mycotoxin Research*, 2012, 28(3): 175-180.
131. Isaaks E.H., Srivastava M.R., Isaaks E.H., et al. An introduction to applied geostatistics. *Technometrics*, 1991, 33: 483-485.
132. Jabeen Z. Incidence of *Aspergillus flavus* and extent of aflatoxin contamination in peanut samples of Pothwar region of Pakistan. *African Journal of Microbiology Research*, 2012, 6(9): 1942-1946.
133. Jaime-Garcia R., Cotty P.J. Aflatoxin contamination of commercial cottonseed in South Texas. *Phytopathology*, 2003, 93(9):1190-1200.
134. Jaime-Garcia R.O., Felix-Gastelum T.V., Trinidad-Correa R., et al. Spatial analysis of *Phytophthora infestans* genotypes and late blight severity on tomato and potato in the Del Fuerte Valley using geostatistics and geographic information systems. *Phytopathology*, 2001, 91(12): 1156-1165.
135. Joffe A.Z. Aflatoxin produced by 1,626 Isolates of *Aspergillus flavus* from groundnut kernels and soils in Israel. *Nature*, 1969, 221(5179): 492-492.
136. Kabak B., Dobson A.D.W., Var I. Strategies to prevent mycotoxin contamination of food and animal feed: A review. *Critical Reviews in Food Science and Nutrition*, 2006, 46(8): 593-619.
137. Kerry R., Ortiz B.V., Ingram B.R., et al. A Spatio-Temporal investigation of risk factors for aflatoxin contamination of corn in southern Georgia, USA using geostatistical methods. *Crop Protection*, 2017, 94(Complete):144-158.
138. Khaidem L., Saha S., Dey S.R. Predicting the direction of stock market prices using random forest. *Applied Mathematical Finance*, 2016: 1-20.

139. Klich M.A. *Aspergillus flavus*: the major producer of aflatoxin. *Molecular Plant Pathology*, 2007, 8(6): 713-722.
140. Klich M.A. Biogeography of *Aspergillus* species in soil and litter. *Mycologia*, 2002, 94(1): 21-27.
141. Kong Q., Shan S., Liu Q., et al. Biocontrol of *Aspergillus flavus* on peanut kernels by use of a strain of marine *Bacillus megaterium*. *International Journal of Food Microbiology*, 2010, 139(1-2): 31-35.
142. Kos J., Mastilović J., Hajnal E.J., et al. Natural occurrence of aflatoxins in maize harvested in Serbia during 2009-2012. *Food Control*, 2013, 34(1): 31-34.
143. Lake I., Hooper L., Abdelhamid A., et al. Climate change and food security: health impacts in developed countries. *Environ. Health Perspect*, 2012, 120(11): 1520-1526.
144. Larkin R.P., Gumpertz M.L., Ristaino J.B. Geostatistical analysis of *Phytophthora* epidemic development in commercial bell pepper fields. *Phytopathology*, 1995, 85(2): 191-203.
145. Lecoustre R., Fargette D., Fauquet C., et al. Analysis and mapping of the spatial spread of African cassava mosaic virus using geostatistics and the kriging technique. *Phytopathology*, 1989, 79(9): 913-920.
146. Lee S.E., Campbell B.C., Molyneux R.J., et al. Inhibitory effects of naturally occurring compounds on aflatoxin B₁ biotransformation. *Journal of Agricultural and Food Chemistry*, 2001, 49(11): 5171-5177.
147. Leong S., Pettersson O.V., Rice T., et al. The extreme xerophilic mould *Xeromyces bisporus*-growth and competition at various water activities. *International Journal of Food Microbiology*, 2011, 145(1): 57-63.
148. Leong Y.H., Ismail N., Latif A.A., et al. Aflatoxin occurrence in nuts and commercial nutty products in Malaysia. *Food Control*, 2010, 21(3):334-338.
149. Levic J., Gosic-Dondo S., Ivanovic D., et al. An outbreak of *Aspergillus* species in response to environmental conditions in Serbia. *Pestic. Fitomed.* 2013, 28(3): 167-179.
150. Liao B., Lei Y., Jiang H., et al. Management of aflatoxin contamination in groundnut by enhancing host-plant resistance. 2010.
151. Liebhold A.M., Rossi R.E., Kemp W.P. Geostatistics and geographic information systems in applied insect ecology. *Annual Review of Entomology*, 1993, 38(1): 303-327.
152. Lisker D.N., Michaeli R., Frank Z.R. Mycotoxigenic potential of *Aspergillus flavus* strains isolated from groundnuts growing in Israel. *Mycopathologia*, 1993, 122(3):177-183.
153. Liu H.M., Rao N., Yang D., et al. Modified bagging of maximal information coefficient for genome-wide identification. *International Journal of Data Mining & Bioinformatics*, 2016, 14(3): 229-257.
154. Logrieco A., Visconti, A. An Overview on Toxigenic Fungi and Mycotoxins in Europe. Dordrecht Netherlands: Kluwer Academic Publishers. 2004.
155. Magan N., Aldred D. Environmental fluxes and fungal interactions: maintaining a competitive edge. In: van West P., Avery S., Stratford M., eds. *Stress in Yeasts and Filamentous Fungi*.

- Amsterdam, the Netherlands: Elsevier Ltd, 2007: 19-35.
156. Magan N., Medina A., Aldred D. Possible climate-change effects on mycotoxin contamination of food crops pre-and postharvest. *Plant Pathology*, 2011, 60(1): 150-163.
157. Ma H.H., Zhang N.Y., Sun L.H., et al. Effects of different substrates and oils on aflatoxin B₁ production by *aspergillus parasiticus*. *European Food Research and Technology*, 2014, 240(3): 627-634.
158. Markoglou A.N., Doukas E.G., Malandrakis A.A. Effect of anilinopyrimidine resistance on aflatoxin production and fitness parameters in *Aspergillus parasiticus* Speare. *International Journal of Food Microbiology*, 2011, 146(2): 130-136.
159. Martins L.M., Sant'Ana A.S., Fungaro M. H.P., et al. The biodiversity of *Aspergillus* section *Flavi* and aflatoxins in the Brazilian peanut production chain. *Food Research International*, 2017, 94: 101-107.
160. Matthias D., Joseph A., Richarda S., et al. Distribution of *Aspergillus* section *Flavi* in soils of maize fields in three agroecological zones of Nigeria. *Soil Biology & Biochemistry*, 2009, 41(1): 37-44.
161. Medina A., Rodriguez A., Magan N. Effect of climate change on *Aspergillus flavus* and aflatoxin B₁ production. *Frontiers in Microbiology*, 2014, 5: 348-348.
162. Medina A., Schmidtheydt M., Parra R., et al. Impacts of environmental stress on growth, secondary metabolite biosynthetic gene clusters and metabolite production of xerotolerant/xerophilic fungi. *Current Genetics*, 2015, 61(3): 325-334.
163. Miraglia M., Marvin H.J.P., Kleter G.A., et al. Climate change and food safety: an emerging issue with special focus on Europe. *Food Chemical Toxicology*, 2009, 47(5): 1009-1021.
164. Mohana D.C., Thippeswamy S., Abhishek R.U., et al. Studies on seed-borne mycoflora and aflatoxin B₁ contaminations in food based seed samples: Molecular detection of mycotoxigenic *Aspergillus flavus* and their management, 2016, 23(6): 2689-2694.
165. Monyo E.S., Njoroge S.M.C., Coe R., et al. Occurrence and distribution of aflatoxin contamination in groundnuts (*Arachis hypogaea* L) and population density of Aflatoxigenic *Aspergilli* in Malawi. *Crop Protect*, 2012, 42: 149-155.
166. Moore F.C., Lobell D.B. The fingerprint of climate trends on European crop yields. *Proceedings of the National Academy of Sciences, USA*, 2015, 112(9): 2670-2675.
167. Moore G.G., Horn B.W., Elliott J.L., et al. Sexual reproduction influences aflatoxin chemotype diversity in worldwide populations of *Aspergillus flavus* and *A. parasiticus*. *Phytopathology*, 2009, 99: S88.
168. Mupunga I., Lebelo S.L., Mngqawa P., et al. Natural Occurrence of aflatoxins in peanuts and peanut butter from Bulawayo, Zimbabwe. *Journal of Food Protection*, 2014, 77(10): 1814-1818.
169. Mupunga I., Mngqawa P., Katerere D.R. Peanuts, aflatoxins and undernutrition in children in Sub-Saharan Africa. *Nutrients*, 2017, 9(12): 1287.
170. Murphy P.A., Hendrich S.L., Landgren C., et al. Food mycotoxins: an update. *Journal of Food*

- Science, 2006, 71: 51-65.
171. Mutegi C.K., Ngugi H.K., Hendriks S.L., et al. Factors associated with the incidence of *Aspergillus* section *Flavi* and aflatoxin contamination of peanuts in the Busia and Homa bay districts of western Kenya. *Plant Pathology*, 2012, 61(6): 1143-1153.
172. Mutitu E.W., Kotchoni O.S. Diversity of fungal species associated with peanuts in storage and the levels of aflatoxins in infected samples. *International Journal of Agriculture & Biology*, 2004, 8530(6): 6.
173. Nabuuma D., Nakimbugwe D., Byaruhanga Y.B., et al. Formulation of a drinkable peanut-based therapeutic food for malnourished children using plant sources. *International Journal of Food Sciences & Nutrition*, 2013, 64(4): 467-475.
174. Nagarajan V., Bhat R.V. Aflatoxin production in peanut varieties by *Aspergillus flavus* Link and *Aspergillus parasiticus* Speare. *Applied Microbiology*, 1973, 25: 319-321.
175. Nelson M.R., Felixgastelum R., Orum T.V., et al. Geographic information systems and geostatistics in the design and validation of regional plant virus management programs. *Phytopathology*, 1994, 84(84): 898-905.
176. Nelson M.R., Orum T.V., Jaime-Garcia R., et al. Applications of geographic information systems and geostatistics in plant disease epidemiology and management. *Plant Disease*, 1999, 83(4): 308-319.
177. Nesci A., Passone M.A., Barra P., et al. Prevention of aflatoxin contamination in stored grains using chemical strategies. *Current Opinion in Food Science*, 2016,11: 56-60.
178. Nicholson M.C., Mather T.N. Methods for evaluating lyme disease risks using geographic information systems and geospatial analysis. *Journal of Medical Entomology*, 1996, 33(5): 711-720.
179. Nyirahakizimana H., Mwamburi L., Wakhisi J., et al. Occurrence of species contamination in raw and roasted peanuts from formal and informal markets in Eldoret and and aflatoxin Kericho Towns, Kenya. *Advances in Microbiology*, 2013, 3(4): 333-342.
180. Odogola W.R. Postharvest management and storage of food legumes. Technical systems for agriculture. UNDP/OPS Regional Programme, RAF/92/R51. Harare. 1994.
181. Olarte R.A., Horn B.W., Dorner J.W., et al. Effect of sexual recombination on population diversity in aflatoxin production by *Aspergillus flavus* and evidence for cryptic heterokaryosis. *Molecular Ecology*, 2012, 21:1453-1476.
182. Olarte R.A., Horn B.W., Monacell J.T., et al. Insights into sexual reproduction in *Aspergillus flavus* from variation in experimental crosses and natural populations. *Phytopathology*, 2010, 100: S92.
183. O'Neil M.J., Smith A. Heckelman P. E. The Merck Index, 13th Ed., Whitehouse Station, NJ, Merck & Co., 2001: 34-35.
184. Ortiz M.P., Barros G.G., Reynoso M.M., et al. Soil populations of *Aspergillus* section *Flavi* from the main and new peanut growing areas in Argentina. ISM Conference 2011 "Strategies to reduce the impact of mycotoxins in Latin America in a global context". Abstract Book. 2011.

185. Orum T.V., Bigelow D.M., Cotty P.J., et al. Using predictions based on geostatistics to monitor trends in *Aspergillus flavus* strain composition. *Phytopathology*, 1999, 89(9): 761-769.
186. Oyedele O.A., Ezekiel C.N., Sulyok M., et al. Mycotoxin risk assessment for consumers of groundnut in domestic markets in Nigeria. *International Journal of Food Microbiology*, 2017, 251: 24-32.
187. Ozias-Akins P., Yang H., Perry E., et al. Transgenic peanut for preharvest aflatoxin reduction. *Mycopathologia*, 2002, 155: 98.
188. Park D.L., Njapau H., Boutrif E. Minimizing risks posed by mycotoxins utilizing the HACCP concept. *Food Nutrition and Agriculture*, 1999(23): 49-54.
189. Parmar R.S., McClendon R.W., Hoogenboom G., et al. Estimation of aflatoxin contamination in preharvest peanuts using neural networks. *Transactions of the ASAE*, 1997, 40(3): 809-813.
190. Passone M.A., Ruffino M., Ponzio V., et al. Postharvest control of peanut *Aspergillus* section *Flavi* populations by a formulation of food-grade antioxidants. *International Journal of Food Microbiology*, 2009, 131 (2-3): 211 -217.
191. Payne G.A., Hagler W.M. Effect of specific amino acids on growth and aflatoxin production by *Aspergillus parasiticus* and *Aspergillus flavus* in defined media. *Applied and Environmental Microbiology*, 1983, 46(4): 805-812.
192. Pieckova E., Kunova Z. Indoor fungi and their ciliostatic metabolites. *Annals of Agricultural and Environmental Medicine*, 2002, 9: 59-63.
193. Pitt J.I., Hocking A.D. Mycotoxins in Australia: Biocontrol of aflatoxin in peanuts. *Mycopathologia*, 2006, 162(3): 233-243.
194. Piva G., Battilani P., Pietri A. In *The mycotoxin factbook. Food & feed topics* (eds Barug D., et al.) 139-153 (Wageningen Academic Publishers, 2006).
195. Pleadin J., Vulić A., Perši N., et al. Aflatoxin B1 occurrence in maize sampled from Croatian farms and feed factories during 2013. *Food Control*, 2014, 40, 286-291.
196. Pleadin J., Vulic A., Persi N., et al. Annual and regional variations of aflatoxin B-1 levels seen in grains and feed coming from Croatian dairy farms over a 5-year period. *Food Control*, 2015, 47: 221-225.
197. Proctor A.D., Ahmedna M., Kumar J.V., et al. Degradation of aflatoxins in peanut kernels/flour by gaseous ozonation and mild heat treatment. *Food Additives and Contaminants*, 2004, 21: 786-793.
198. Rachaputi N.R., Wright G.C., Krosch S. Management practices to minimize pre-harvest aflatoxin contamination in Australian groundnuts. *Australian Journal of Experimental Agriculture*, 2002, 42(5): 595-605.
199. Reddy T.Y., Sulochanamma B.N., Subramanyam A., et al. Influence of weather, dry spells and management practices on aflatoxin contamination in groundnut. *Indian Phytopathology*, 2003, 56(3): 262-265.
200. Reshef D.N., Reshef Y.A., Finucane H.K., et al. Detecting novel associations in large data sets. *science*, 2011, 334(6062): 1518-1524.

201. Robens J., Cardwell K. The costs of mycotoxin management to the USA: Management of aflatoxins in the United States. *Toxin Reviews*, 2003, 22(2-3): 139-152.
202. Rotimi O.A., Rotimi S.O., Duru C.U., et al. Acute aflatoxin B₁ ̈“ Induced hepatotoxicity alters gene expression and disrupts lipid and lipoprotein metabolism in rats. *Toxicology Reports*, 2017, 4: 408.
203. Salvacion A.R., Pangga I.B., Cumagun C.J. Assessment of mycotoxin risk on corn in the Philippines under current and future climate change conditions. *Reviews on Environmental Health*, 2015, 30(3): 135-142.
204. Sanders T.H., Cole R.J., Blankenship P.D., et al. Aflatoxin contamination of peanuts from plants drought stressed in pod or root zones. *Peanut Science*, 1993, 20(1): 5-8.
205. Sanders T.H., Cole R.J., Blankenship P.D., et al. Relation of environmental stress duration to *Aspergillus flavus* invasion and aflatoxin production in preharvest peanuts. *Peanut Science*, 1985, 12(2): 90-93.
206. Sanders T.H., Hill R.A., Cole R.J., et al. Effect of drought on occurrence of *Aspergillus flavus* in maturing peanuts. *Journal of American Oil Chemical Society*, 1981, 58(12): 966-970.
207. Sanders T.H., McMichael R.W., Jr, Hendrix K.W. Occurrence of resveratrol in edible peanuts. *Journal of Agricultural and Food Chemistry*, 2000, 48(4): 1243-1246.
208. Schmidt H.M., Abdel H.A., Magan N., et al. Complex regulation of the aflatoxin biosynthesis gene cluster of *Aspergillus flavus* in relation to various combinations of water activity and temperature. *Int J Food Microbiol*, 2009, 135(3) : 231-237.
209. Schmidt H.M., Rufer C.E., Abdel H.A., et al. The production of aflatoxin B₁ or G₁ by *Aspergillus parasiticus* at various combinations of temperature and water activity is related to the ratio of aflS to aflR expression. *Mycotoxin Research*, 2010, 26(4) : 241-246.
210. Schroeder H.W., Hein H. Aflatoxins: production of the toxins in vitro in relation to temperature. *Applied Microbiology*, 1967, 15: 441-445.
211. Shao F., Li K. A graph model for preventing railway accidents based on the maximal information coefficient. *International Journal of Modern Physics B*, 2016, 31(3): 1750010.
212. Smith L. Examining environmental drivers of spatial variability in aflatoxin accumulation in Kenyan maize: Potential utility in risk prediction models. *African Journal of Food Agriculture Nutrition & Development*, 2016, 16(3): 11086-11105.
213. Sobolev V.S., Gou B.Z., Holbrook C.C., et al. Interrelationship of phytoalexin production and disease resistance in selected peanut genotypes. *Journal of Agricultural and Food Chemistry*, 2007, 55(6): 2195-2200.
214. Stein A., Kocks C.G., Zadoks J.C., et al. A geostatistical analysis of the spatio-temporal development of downy mildew epidemics in cabbage. *Phytopathology*, 1994, 84(10): 1227-1239.
215. Stoev S.D. Food safety and increasing hazard of mycotoxin occurrence in foods and feeds. *Critical reviews in food science and nutrition*, 2013, 53(9): 887-901.
216. Sulyok M., Beed F., Boni S., et al. Quantitation of multiple mycotoxins and cyanogenic glucosides

- in cassava samples from Tanzania and Rwanda by an LC-MS/MS-based multi-toxin method. *Food Additives & Contaminants*, 2015, 32(4): 488-502.
217. Tao F.L., Zhang Z. Impacts of climate change as a function of global mean temperature: maize productivity and water use in China. *Climate Change*, 2011, 105: 409-432.
218. Thai C.N., Blankenship P.D., Cole R.J., et al. Relationship between aflatoxin production and soil temperature for peanuts under drought stress. *Transactions of the ASAE*, 1990, 33(1): 324-329.
219. Thalij K.M., Hajeej J. M., Mohammed M.J., et al. Study the occurrence of aflatoxins in some crops and dry fruits in Iraqi markets. *Journal of Natural Sciences Research*, 2015, 5(11): 22-26.
220. Theumer M.G., Henneb Y, Khoury L., et al. Genotoxicity of aflatoxins and their precursors in human cells. *Toxicology Letters*, 2018, 287: 100-107.
221. Torres A.M., Barros G.G., Palacios S.A., et al. Review on pre- and post-harvest management of peanuts to minimize aflatoxin contamination. *Food Research International*, 2014, 62(8): 11-19.
222. Torres A.M., Rodriguez M.I., Chulze S.N. Genetic diversity within *Aspergillus flavus* strains isolated from peanut-cropped soils in Argentina. *Soil Biology & Biochemistry*, 2006, 38(1): 145-152.
223. Troeger J.M. Modeling quality in bulk peanut curing. *Peanut Science*, 1989, 16(2) : 105-108.
224. Vaamonde G., Patriarca A., Pinto V.F., et al. Variability of aflatoxin and cyclopiazonic acid production by *Aspergillus* section *Flavi* from different substrates in Argentina. *International Journal of Food Microbiology*, 2003, 88(1): 79-84.
225. Vidhyasekaran P., Lalithakumari D., Govindaswamy C.V. Production of a phytoalexin in groundnut due to storage fungi. *Indian Phytopathology*, 1972, 25: 240-245.
226. Waliyar F., Kumar P.L., Traore A., et al. Pre and post harvest management of aflatoxin contamination in peanuts. In Leslie J. F., Bandyopadhyay R. & Visconti A. (Eds.) *Mycotoxins: Detection Methods, Management, Public Health and Agricultural Trade* (2008: 209-218). Trowbridge: Cromwell Press.
227. Waliyar F., Ntare B.R., Diallo A.T., et al. On farm management of aflatoxin contamination of groundnut in west Africa. A synthesis report. International Crops Research Institute for the Semi-rid Tropics. 2007: 24.
228. Waliyar F., Umeh V.C., Traore A., et al. Prevalence and distribution of aflatoxin contamination in groundnut (*Arachis hypogaea* L.) in Mali, West Africa. *Crop Protection*, 2015, 70: 1-7.
229. Webster R., Boag B. Geostatistical analysis of cyst nematodes in soil. *European Journal of Soil Science*, 1992, 43: 583-595.
230. Wheeler T., von Braun J. Climate change impacts on global food security. *Science*, 2013, 341(6145): 508-513.
231. Whitaker T.B., Giesbrecht F.G., Slate A.B. Market system model to predict the effects of regulatory and processing practices on the removal of aflatoxin from peanuts. *Peanut Science*, 2002, 29(2) : 128-136.
232. White J.G., Welch R.M., Norvell W. A. Soil zinc map of the USA using geostatistics and

- geographic information systems. *Soil Science Society of America Journal*, 1997, 61(1):185-194.
233. Williams R.J., McDonald D. Grain molds in the tropics: problems and importance. *Annual Review of Phytopathology*, 1983, 21: 153-178.
234. Wotton H.R., Strange R.N. Circumstantial evidence for phytoalexin involvement in the resistance of peanuts to *Aspergillus flavus*. *Journal of General Microbiology*, 1985, 131(3): 487-494.
235. Wu B.M., Bruggen A.H., Van Subbarao K.V., et al. Spatial analysis of lettuce downy mildew using geostatistics and geographic information systems. *Phytopathology*, 2001, 91(2): 134-142.
236. Wu L.X., Ding X.X., Li P.W., et al. Aflatoxin contamination of peanuts at harvest in China from 2010 to 2013 and its relationship with climatic conditions. *Food Control*, 2016, 60:117-123.
237. Youssef M.S., El-Maghraby O.M.O., Ibrahim Y.M.. Mycobiodata and mycotoxins of Egyptian peanut (*Arachis hypogaea* L.) seeds. *International Journal of Botany*, 2008, 4: 349-360.
238. Zhang C., Ma Y., Zhao X., et al. Kinetic modelling of aflatoxins B₁ conversion and validation in corn, rice, and peanut during thermal treatments. *Food Chemistry*, 2011, 129(3): 1114-1119.
239. Zhang C., Selvaraj J.N., Yang Q., et al. A survey of aflatoxin-producing *Aspergillus* sp. from peanut field soils in four agroecological zones of China. *Toxins*, 2017, 9(1): 40.
240. Zhang L., Qian S., Li P., et al. Ion mobility spectrometry fingerprints: A rapid detection technology for adulteration of sesame oil. *Food Chemistry*, 2016, 192: 60-66.
241. Zhou D.N., Chen L., Wu D., et al. Maximal information coefficient for feature selection for clinical document classification. *Acta Physico-Chimica Sinica*, 2012,28(8): 963-970.

致 谢

时间如白驹过隙，硕博研究生的时光已接近尾声。六年的求学生涯，从最初的懵懂至如今对科学研究心怀敬畏之情，在老师、同学和朋友的全力支持下，虽过程辛苦却也收获颇丰。值此论文付梓之际，回眸过去，思绪万千。一路走来，需要感谢的人太多，文字着实不能够表达与承载。

感谢我的导师李培武研究员，导师渊博的专业知识、严谨的治学态度和精益求精的工作作风对我影响深远；循循善诱的教导、开阔的视野和不拘一格的思路给予我深深的启迪。感谢丁小霞副研究员，不仅使我树立了良好的学术目标、掌握了基本的科研方法、坚定了人生方向，还使我明白了许多待人接物与为人处世的道理。

不积跬步何以至千里，本课题及论文的顺利完成，也归功于实验室各位老师的认真负责与谆谆教导，感谢课题组岳晓凤博士、张良晓博士、张兆威博士为我充实多方面的知识储备，让我开阔了眼界，加深了对科研的理解与认识，使我能够更好地掌握和运用专业知识，并在数据处理、论文写作、图表表达等方面得以体现。感谢张奇主任、张文主任、甘冬生书记、陈小媚老师等对我生活和实验上的照顾，正是有了他们的悉心帮忙和支持，才使我的课题得以顺利完成。能够接触到这些不仅仅传授我知识和学问，而且从更高层次指导我的人生与价值追求的良师，我为自己感到庆幸。

感谢十三五国家重点研发计划项目（2017YFC1601201），国家自然科学基金（31701726、31601573、31540048），国家农产品质量安全风险评估重大项目（GJFP2015007、GJFP2016007、GJFP2016001），国家粮油作物产品质量安全风险专项（GJFP2018001）及国家科技支撑计划（2012BAB19B09）等对本论文的资助。

感谢一路上关心与支持我的同学和朋友们，感谢他们的鼓励和帮忙。感谢董雪丽、许琳、杨瑞楠、王晓、戴显红、朱婷婷等实验室的师姐妹，以及王忠政、尤思聪等师兄们，你们不仅令我在科研路上不感觉枯燥与孤单，更让我的生活过得无比充实与美好。在科研道路上我们共同进步，感谢你们给予我的所有关心和帮忙。同窗之谊，我将终生难忘！

感谢代碌碌六年来对我的照顾与理解，在我遇到困难和喜悦的时候，同我分担，与我分享。

另外，还要感谢我的家人，是家人的鼓励和支持使我在漫长的人生旅途中拥有巨大的前行动力和温暖的港湾，使我能够全身心投入到学习和工作之中，顺利完成学业。在未来的日子里，我会更加努力地学习和工作，不辜负父母和姐弟对我的殷殷期望！

最后衷心感谢在百忙之中评阅论文和参加答辩的各位专家、教授！

武琳霞
2019年5月于武汉

作者简介

武琳霞，女，汉族，中共预备党员，1989年8月出生，山西大同人。2012年毕业于山东农业大学食品科学与工程专业，获工学学士学位；2013年至今，在中国农业科学院油料作物研究所硕博连读，攻读农产品质量与食品安全专业博士学位，师从李培武研究员和丁小霞副研究员。主要从事中国花生黄曲霉毒素污染风险预警模型研究，在中国农业科学院油料作物研究所、国家农业检测基准实验室（生物毒素）、农业农村部油料产品质量安全风险评估中心（武汉）、农业农村部生物毒素检测重点实验室完成所有实验。主笔和参与发表论文9篇（包括SCI论文3篇）、编著1部。

攻读博士期间发表的论文：

1. **Linxia Wu**, Li Yu, Xiaoxia Ding, Peiwu Li, Xianhong Dai, Xiaomei Chen, Haiyan Zhou, Yizhen Bai, Jian Ding. Magnetic solid-phase extraction based on graphene oxide for the determination of lignans in sesame oil. *Food Chemistry*, 2017, 217: 320-325.
2. **Linxia Wu**, Xiaoxia Ding, Peiwu Li, Xiaohui Du, Haiyan Zhou, Yizhen Bai, Liangxiao Zhang. Aflatoxin contamination of peanuts at harvest in China from 2010 to 2013 and its relationship with climatic conditions. *Food Control*, 2016, 60: 117-123.
3. **Linxia Wu**, Haiyan Zhou, Tingting Zhu, Hongling Yang, Xueli Dong, Lin Chen, Yizhen Bai, Nanri Yin, Xiaoxia Ding, Peiwu Li. Fungal diversity and mycotoxins identification of mouldy rapeseed in China. *Oil Crop Science*, 2017, 2 (1): 38-48.
4. Xiaoxia Ding, **Linxia Wu** (共同第一), Peiwu Li, Zhaowei Zhang, Haiyan Zhou, Yizhen Bai, Xiaomei Chen, Jun Jiang. Risk Assessment on dietary exposure to aflatoxin B₁ in post-harvest peanuts in the Yangtze River ecological region. *Toxins*, 2015, 7 (10): 4157-4174.
5. Xiaohui Du, Xiaowei Deng, Peiwu Li, Xiaoxia Ding, **Linxia Wu**, Haiyan Zhou, Yizhen Bai. Peanut aflatoxin contamination and soil *Aspergillus flavus* isolates in different climate regions. *Oil Crop Science*, 2016 (1): 60-66.
6. **武琳霞**, 李培武, 丁小霞, 岳晓凤, 白艺珍. 农产品真菌毒素混合污染与累积风险评估研究进展. *食品安全质量检测学报*, 2018, 9(14): 3553-3560.
7. **武琳霞**, 丁小霞, 李培武, 姜俊, 周海燕, 印南日, 白艺珍, 陈小媚. 我国油菜镉污染及菜籽油质量安全性评估. *农产品质量与安全*, 2016 (1) : 41-46.
8. **武琳霞**, 都晓慧, 丁小霞, 周海燕, 陈琳, 李培武. 花生黄曲霉毒素污染预警技术研究进展. *中国油料作物学报*, 2016, 38 (1): 120-125.
9. 都晓慧, 丁小霞, 周海燕, 李培武, **武琳霞**, 印南日, 白艺珍. 产后花生黄曲霉毒素污染监测抽样方法研究. *中国油料作物学报*, 2015, 37(06): 876-880.

攻读博士期间出版的编著:

1. Peiwu Li, Xiaoxia Ding, Yizhen Bai, **Linxia Wu**, Xiaofeng Yue, Liangxiao Zhang. Risk Assessment and prediction of aflatoxin in agro-products. In Valentina Svalova (Eds.), Risk Assessment. 2018, 319-334. Russia: Intech.

攻读博士期间获得的成果奖励:

1. 2019 年获得中国农业科学院油料作物研究所—岛津合作实验室二等奖学金
2. 2017 年获得中期考核优秀奖
3. 2016 年获得研究生国家奖学金
4. 2016 年获得“金龙鱼”优博奖学金
5. 2016 年获得一等研究生学业奖学金
6. 2015 年获得一等研究生学业奖学金
7. 2014 年获得一等研究生学业奖学金
8. 2013 年获得中国农业科学院全国硕士研究生统一入学考试中成绩优异奖