

基于关键气象因子对嘉禾烟叶产量影响的研究

曾祥忠¹ 李辉¹ 何娜² 李井军³ 周璇¹ 任周琴⁴

1 湖南省嘉禾县气象局

2 湖南省资兴市气象局

3 湖南省嘉禾县烟草局

4 湖南省临武县气象局

DOI:10.32629/as.v9i8.4253

[摘要] 利用嘉禾国家基本气象站(区站号57974)2005—2024年逐日气象资料与同期烟叶产量资料,采用直线滑动平均法分解产量,构建124个候选气象因子,经相关普查、偏相关分析筛选主控因子,构建4类预报模型并以留一法交叉验证(LOOCV)检验。结果表明:移栽—伸根期 $\geq 10^{\circ}\text{C}$ 积温与相对气象产量呈显著负相关($r=-0.457, P<0.05$),为最强因子;4月上旬日照($r=-0.447$)、3月上旬气温日较差($r=-0.442$,即该旬最高与最低气温差值的平均,是寡照的指示指标)次之,6月中旬均温为正相关关键时段;PLS回归为最优模型(LOOCV $R^2=0.267$, RMSE=6.24%),MLR固定方程丰歉年方向预报准确率达83%。早春热量条件、成熟期水分胁迫与旺长期暴雨日数是嘉禾烟叶产量的核心主控因子。

[关键词] 烤烟; 气象因子; 产量分解; 偏最小二乘回归; 产量预报

中图分类号: S572 文献标识码: A

Effects of Key Meteorological Factors on Tobacco Yield in Jiahe

Xiangzhong Zeng¹ Hui Li¹ Na He² Jingjun Li³ Xuan Zhou¹ Zhouqin Ren⁴

1 Jiahe County Meteorological Bureau

2 Zixing Meteorological Bureau

3 Jiahe County Tobacco Bureau

4 Linwu County Meteorological Bureau

[Abstract] Using daily meteorological data (2005—2024) from the Jiahe National Basic Meteorological Station and tobacco yield records, the yield per mu was decomposed into trend and meteorological yields, 124 candidate factors were screened by correlation survey and partial correlation analysis, and four forecast models were tested by leave-one-out cross-validation(LOOCV). The $\geq 10^{\circ}\text{C}$ accumulated temperature in the transplanting-to-rooting stage was the strongest factor ($r=-0.457, P<0.05$); sunshine duration in early April ($r=-0.447$) and the mean diurnal temperature range in early March ($r=-0.442$) ranked next. The PLS model performed best (LOOCV $R^2=0.267$, RMSE=6.24%), and the fixed MLR equation achieved 83% accuracy in predicting the direction of yield-anomaly years.

[Key words] flue-cured tobacco; meteorological factors; yield decomposition; PLS regression; yield forecast

引言

嘉禾县地处湘南,属亚热带季风湿润气候,是湖南省烤烟主产区之一。烤烟大田生长期(3-7月)冷暖空气交汇频繁,低温冷害、暴雨洪涝、冰雹大风等对产量影响显著^[1-2]。量化关键气象因子影响、明确主控因子与关键时段并建立业务预报模型,是烟草生产气象服务的重要基础^[3-4]。本文以2005—2024年烟叶产量与同期逐日气象资料,按“数据清洗—产量分解—因子库构建—因子筛选—模型验证”路线开展研究。

1 资料与方法

气象资料为嘉禾国家基本气象站2005年3月1日—2024年7月20日逐日观测,要素含日照、气温(平均/最高/最低)、相对湿度、降水量(08—08时)、极大风速、冰雹及大风;产量与灾情资料来源于县烟草部门统计及县气象灾情档案。资料经口径统一、并行补缺与线性插值(缺测不足1%)、冰雹“观测+灾情”互证等清洗,研究窗口(3月1日—7月20日)内核心要素无缺测,共20a×142 d。

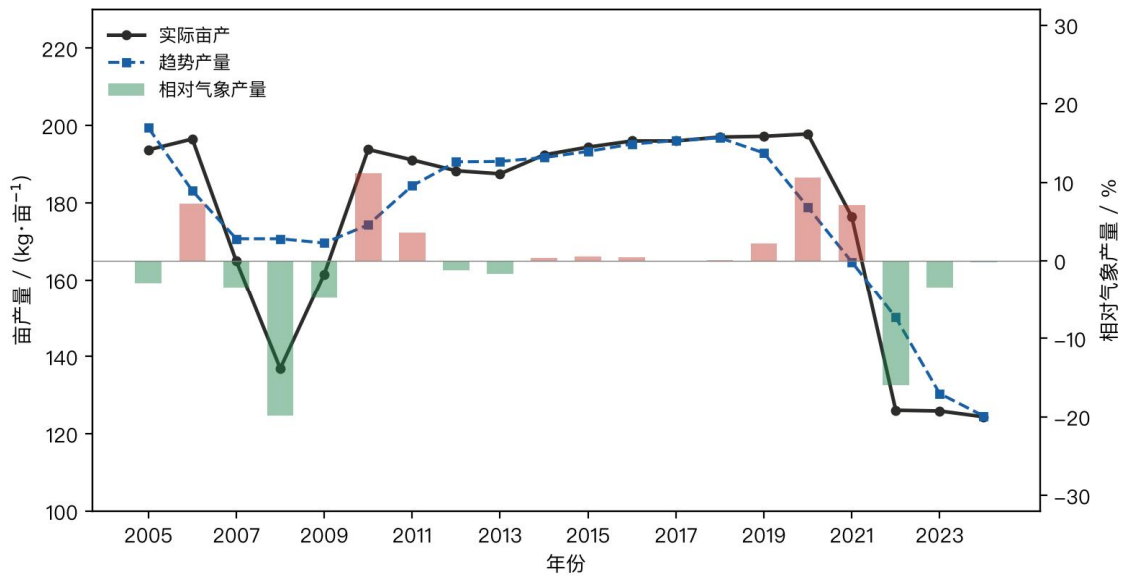


图1 2005—2024年嘉禾县烤烟亩产、趋势产量与相对气象产量

采用5年直线滑动平均法(窗口内线性拟合,端点缩窗)将亩产 Y 分解为趋势产量 Y_t 与气象产量 Y_w ,以Hodrick-Prescott滤波($\lambda=6.25$)验证稳健性,取相对气象产量 $Y_w'=(Y-Y_t)/Y_t\times 100\%$ 为分析对象。按生育进程划分移栽—伸根期(3月1日—4月20日)、旺长期(4月21日—5月20日)、成熟期(5月21日—7月20日),逐期统计气温、最低气温、日较差、相对湿度、 $\geq 10^\circ\text{C}$ 积温、降水、日照及低温冷害、霜冻(日最低 $\leq 2^\circ\text{C}$)、暴雨($\geq 50\text{ mm}$)、连阴雨、无降水日、大风、冰雹等18个因子,共54个;另按旬(共14旬)统计旬均温、旬最低温、旬日较差、旬降水、旬日照共70个因子,合计124个候选因子。

计算124个因子与相对气象产量的Pearson相关系数($n=20$, $P<0.10$ 、 $P<0.05$ 临界 r 为0.378、0.444),Spearman相关佐证;对共线因子组作偏相关分析。分别构建MLR、PLS、岭回归与随机森林模型,以留一法交叉验证(L00CV,变量选择置于循环内)评价,指标为 R^2 、RMSE、MAE及丰歉年方向准确率($|Y_w'|>5\%$)。

2 结果与分析

2005—2024年嘉禾烤烟亩产124.4~197.8 kg,趋势产量“先升后降”,2019年后由197 kg降至125 kg左右,与种植面积快速扩张(2022年扩大38%)有关。剔除趋势后相对气象产量标准差7.29%:2008年(-19.81%)、2022年(-16.00%)极端减产,2010年(+11.17%)、2020年(+10.56%)极端增产(图1)。

124个候选因子中14个达 $P<0.10$ 显著水平(表1)。生育期尺度以移栽—伸根期 $\geq 10^\circ\text{C}$ 积温负相关最强($r=-0.457$, $P<0.05$),平均日较差($r=-0.427$)、平均气温($r=-0.433$)同为负相关;旬尺度上4月上旬日照($r=-0.447$, $P<0.05$)、3月上旬气温日较差($r=-0.442$)、3月上旬日照($r=-0.380$)呈负相关,5月上旬旬最低温($r=0.430$)、6月中旬旬均温($r=0.430$)呈正相关,成熟期无降水日数($r=0.425$)指示偏干利于采烤,旺长期暴雨日数($r=0.398$)

与需水高峰一致。3月上旬—4月上旬寡照低温与6月中旬以后温度条件为两个关键时段。

表1 与相对气象产量显著相关($P<0.10$)的气象因子

因子	时段/生育期	r	P 值	显著性
$\geq 10^\circ\text{C}$ 积温	移栽—伸根期	-0.457	0.043	**
旬日照时数	4月上旬	-0.447	0.048	**
旬平均日较差	3月上旬	-0.442	0.051	*
平均气温	移栽—伸根期	-0.433	0.056	*
旬最低温	5月上旬	0.430	0.059	*
旬均温	6月中旬	0.430	0.058	*
平均日较差	移栽—伸根期	-0.427	0.060	*
无降水日数	成熟期	0.425	0.062	*
低温冷害日数	移栽—伸根期	0.402	0.079	*
暴雨日数	旺长期	0.398	0.082	*
旬日照时数	3月上旬	-0.380	0.099	*

注: $n=20$; *, **分别表示通过 $P<0.10$ 、 $P<0.05$ 检验。

偏相关分析表明,早春热量—光照类因子(积温、均温、日较差、日照等)两两相关达0.50~0.97,构成同一信号簇,控制簇内因子后仅4月上旬日照偏相关接近显著(偏 $r=-0.424$, $P=0.063$),取3月上旬旬日较差与积温为簇代表;旺长期暴雨日数偏相关0.404($P=0.077$),有独立贡献。最终确定11个候选预报因子。

模型检验(表2)显示:逐步回归在L00CV循环内选变量时20次迭代出现13种组合,小样本下不稳定;随机森林亦低于气候基准确。PLS回归(2成分)L00CV精度最优($R^2=0.267$, RMSE=6.24%, MAE=5.10%),定为主模型;引入最低气温后 R^2 由0.180明显提高。MLR固定方程为 $Y_w'=-52.42-1.004\cdot\text{DTR3上}+2.045\cdot\text{T7上}$ (DTR3上为3月上旬旬平均日较差, T7上为7月上旬旬均温),全样本 $R^2=$

0.390 ($P=0.015$), 丰歉年方向准确率83% (5/6): 日较差每偏小1℃产量约降1.0个百分点, 7月上旬均温每偏高1℃产量约升2.05个百分点。亩产预报按“趋势产量 $\times$ (1+预报相对气象产量/100)”换算。两个极端减产年(2008、2022)各模型均低估, 提示含冰灾灾后恢复、面积骤扩等非气象成因。

表2 产量预报模型留一法交叉验证 (LOOCV) 精度对比

模型	LOOCV R^2	RMSE/%	MAE/%	丰歉年方向准确率
PLS 回归 (2 成分)	0.267	6.24	5.10	67%
岭回归	0.105	6.89	5.43	67%
MLR 固定方程 (日较差+7 月上均温)	0.057	7.07	6.05	83%
随机森林	-0.170	7.88	6.06	67%
逐步回归 (循环内选变量)	-1.536	11.60	10.13	33%
气候平均 (基准)	0.000	7.29	4.92	33%

3 结论与讨论

(1) 嘉禾烟叶产量气象变幅达 $\pm 20\%$, 核心主控因子为早春热量—光照条件 (移栽—伸根期 $\geq 10^\circ\text{C}$ 积温, $r=-0.457$; 3月上旬日较差, $r=-0.442$)、成熟期水分胁迫 (无降水日数, $r=0.425$) 与旺长期暴雨日数 ($r=0.398$); 关键影响时段为3月上旬—4月上旬 (寡照负效应) 与6月中旬以后 (温度正效应)。

(2) 小样本条件下PLS回归为最优产量预报模型 (LOOCV $R^2=0.267$, RMSE=6.24%), MLR固定方程丰歉年方向预报准确率83%,

二者结合可满足县域产量预报业务需求。

(3) 早春热量—光照因子与产量呈负相关: 日较差是寡照的直接指示 (3月上旬日较差与旬日照相关达0.93), “暖而寡照”不利于烟苗光合积累与缓苗, 日较差因子使“热量—光照耦合”机制获得更直接证据, 其普适性有待验证。此外, 样本仅20年、县级产量统计存在平滑化, 因子结论与模型精度待序列延长后复核; 冰雹致灾宜结合灾情案例 (2013年3月20日全县烤烟受灾5545亩) 单独分析。

[参考文献]

[1]王馥棠.粮食产量气象预测预报[M].北京:气象出版社, 1991.

[2]张养才,何维勋,李世奎.中国农业气象灾害概论[M].北京:气象出版社,1991.

[3]魏凤英.现代气候统计诊断与预测技术(第2版)[M].北京:气象出版社,2007.

[4]黄嘉佑.气象统计分析与预报方法(第4版)[M].北京:气象出版社,2016.

[作者简介:]

曾祥忠(1999—),男,汉族,湖南省郴州市人,本科,助理工程师,研究方向:农业气象。