

冬小麦赤霉病流行程度 长期预报模型

高 革 居为民

(江苏省农业气象与卫星遥感中心,南京 210008)

提 要

小麦赤霉病的发生流行程度主要由天气条件决定,而500hPa大气环流因子是较好地表征天气形势和控制天气条件的物理量。利用最优化相关分析方法,分区挑选了与小麦赤霉病显著、稳定相关的500hPa大气环流因子,在此基础上建立了冬小麦赤霉病流行程度长期预报环流模型。

关键词: 小麦赤霉病 大气环流 最优化因子 相关分析 预报模型

引 言

小麦赤霉病是江苏省小麦的主要病害之一,它不仅造成小麦严重减产,而且降低小麦品质甚至不能食用。经研究表明^[1],赤霉病菌源是感染的基础,而花期气象条件“暖雨湿天”是左右赤霉病发病流行程度的主导因素。因此常称冬小麦赤霉病为“气象型”病害。本文根据500hPa大气环流特征量控制天气条件这一特性,利用最优化因子相关普查分析^[2]方法,挑选一批与小麦赤霉病相关极其显著的500hPa大气环流因子,并对所选因子进行稳定性和复共线性检验,提高了预报因子的可靠性,在此基础上,建立了江苏省淮河以南各地区冬小麦赤霉病预报的环流模型。

1 预报因子的筛选

1.1 资料的选择

1.1.1 因变量的选择

江苏省不同地区之间,赤霉病的发生流行程度存在着差异,根据历史上小麦赤霉病的发生情况,可将全省分为4个区^[3],即太湖和沿江东部地区(I区);宁镇扬低山丘陵地

区(II区);里下河和东部沿海地区(III区);淮北地区(IV区)。其中淮北地区很少发生赤霉病,故只对其它3个地区进行赤霉病分析和预报。

资料来源于江苏省植保站的实测值。在I、II、III区内选择代表性好,资料年代较长的代表站,以其1965~1994年的逐年平均病穗率来代表各区历年赤霉病的发生、流行状况,作为因变量。

1.1.2 自变量的选择

由于500hPa大气环流系统是由超长波、长波和短波系统组成,它表征了大气环流系统的活动情况,从长期天气预报的观点看,它们对一般天气条件的影响也存在一定的滞后效应^[4]。因此,本文选择影响我国天气过程的74项主要的500hPa大气环流特征量作为自变量,分析其与小麦赤霉病发生程度之间的关系。江苏省小麦赤霉病的感病关键期为4月下旬至5月上旬,考虑资料来源和预报时效,选取上一年1月到当年2月份的环流特征量进行相关分析。

1.2 最优化因子相关分析^[2]

如何选择到相关显著、稳定并且物理意义明确的预报因子是回归预报的关键和基础。为了增加因子的信息量,首先对因子进行了膨化处理^[3],然后再进行最优化相关普查。

自变量与因变量之间存在着不同形式(线性、非线性)的相互关系,找到两者之间最合适的相关形式非常重要,但也十分困难。而最优化相关普查方法可以解决这一问题。

因子 x 的线性和非线性(含单调的和非单调的单峰(谷)型)化处理可归纳为一种通用变换形式:

$$Q = (|X - b|/B + 0.5)^a$$

式中 a, b 为待定参数,且 $X_{\min} \leq b \leq X_{\max}, B = \max(X_{\max} - b, b - X_{\min})$ 。经上述变换后, Q 与 Y (Y 为因变量)必为单调关系,且 $(|X - b|/B + 0.5)$ 的值在区间 $[0.5, 1.5]$ 内变化。对于单峰(谷)型关系的因子,为了避免 x 在最低或最高值附近出现的个别样本的偶然误差影响, b 的取值以 $(X_{\min} + (X_{\max} - X_{\min})/4) \leq b \leq (X_{\max} - (X_{\max} - X_{\min})/4)$ 为宜。至于 a 值,根据我们的实际工作经验,一般在 $(-10, -1/10)$ 和 $(1/10, 10)$ 两个区间内取值,效果较好。待定参量 a, b 可用最优化技术求出。令目标函数为:

$$f(a, b) = 1 - R^2 = \min$$

R 为 a, b 取一定值时, Q 与 Y 的相关系数。应用二维寻优的变量转换思路将其分解为一元问题逐步处理^[5]。

经过上述方法处理普查后获得的因子,是一批与因变量相关最显著的因子。

1.3 稳定性检验

通过最优化相关普查方法可以找到相关显著因子,但它不能保证选择到的因子与因变量之间相关的平稳性。所以,还应对所选因子用滑动相关检验法进行稳定性检验^[6],以淘汰掉一些相关程度前好后差或波动变化较大的因子,保证所选因子与赤霉病之间具有

稳定、显著的相关关系。

本文对各区中与赤霉病病穗率相关系数大于等于0.55(达到0.01信度的显著相关水平)的环流因子进行稳定性检验。

经过以上最优化相关分析和稳定性检验后,在 I、II、III 区分别获得34、54、12个相关信度高、稳定性较好的环流因子。在这些因子中,单相关系数最大达0.80以上,具体单相关系数分布情况如表1。

表1 各区所选与赤霉病病穗率稳定显著相关环流因子数分布

| 相关系数 | $R \geq 0.55$ | $R \geq 0.60$ | $R \geq 0.65$ | $R \geq 0.70$ |
|------|---------------|---------------|---------------|---------------|
| 区号 | I II III | I II III | I II III | I II III |
| 因子数 | 34 54 12 | 13 18 8 | 6 7 6 | 2 1 5 |

1.4 独立性检验^[5]

统计回归模型的拟合,首先要估计其系数,但在任何情况下系数都可估计,对于线性回归方程的最小二乘法拟合来说,如果自变量数据矩阵 $X_{n \times p}$ 有以下关系

$$|X'X| = 0$$

即自变量相关阵的行列式等于0,则系数就无法估计,这表明 X 中有多元共线性存在。为了解决这一问题,本文选用主成分识别法进行因子的独立性检验,剔除共线性因子。

主成分分析的主要作用是有关自变量线性组合成一些新的相互间正交的综合变量 Z_j 。 Z_j 亦称为主成分, P 个自变量可组成 P 个主成分,因为各个 $Z_j (j=1, 2, 3, \dots, P)$ 是两两正交的, Z 方差的协方差阵是个对角阵。对角线元素(即特征根)按大小排列,可记为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, Z_j 的方差就等于 λ_j 。如果先使各变量 X_j 标准化(标准化并不改变各变量的相关关系),若 X_j 相互正交,那么 λ_j 值都将等于1。但当原变量中隐藏着共线性组合时,那么必有特征根 λ_j 接近于0。

由于 λ_j 是主成分 Z_j 方差估计,如果 $\lambda_j \approx 0$,就等于说, Z_j 近似为一个常数。又由于 X_j

都已标准化,其均值为0,因而主成分 Z_j 的均值也等于0,利用这些性质,就可以方便地找到存在于共线性组合中的具体变量。

本文利用这一方法,在 I、II、III 区的相关因子中分别剔除掉19、39、3个因子,保留下

15、15、9个因子,可以认为这些因子是分别与各区赤霉病病穗率相关显著、稳定并且相对独立的因子。表2给出 I 区因子情况,II、III 区因子情况省略,以它们建模都是有效的。

表2 I 区相关因子内容、时段和相关系数

| 项目 | 内 容 | 时 段 | 单相关系数 |
|----|----------------------------|-----------|---------|
| 1 | 大西洋副高面积指数(55°W~25°W) | 上一年7月 | -0.6146 |
| 2 | 大西洋副高面积指数(55°W~25°W) | 上一年7~9月 | 0.5824 |
| 3 | 太平洋副高强度指数(110°E~115°W) | 上一年1月 | -0.5803 |
| 4 | 北半球副高脊线(5°E~360°) | 上一年3~4月 | -0.5997 |
| 5 | 北非副高脊线(5°E~60°E) | 上一年4月 | -0.5973 |
| 6 | 北非大西洋北美副高脊线(110°W~60°E) | 上一年3~5月 | -0.5579 |
| 7 | 西太平洋副高脊线(110°E~150°E) | 上一年3月 | -0.5524 |
| 8 | 北半球副高北界(5°E~360°) | 上一年3~4月 | -0.6244 |
| 9 | 北美大西洋副高北界(110°W~20°E) | 上一年3月 | 0.5665 |
| 10 | 亚洲区极涡面积指数(1区60°E~150°E) | 上一年9~10月 | 0.6718 |
| 11 | 北半球极涡面积指数(5区0~360°) | 上一年9月 | 0.5949 |
| 12 | 北半球极涡强度指数(5区0~360°) | 上一年10~12月 | 0.5668 |
| 13 | 北半球极涡中心位置(JW) | 上一年7月 | 0.6721 |
| 14 | 北半球极涡中心位置(JW) | 上一年7~8月 | 0.5609 |
| 15 | 西藏高原(30°N~40°N 75°E~105°E) | 上一年6~7月 | 0.6404 |

2 预报模型的创建

2.1 建立预报模式

由于已考虑了因子相关的最优化、显著性、稳定性和独立性。因此,由自变量组合的联立方程可以达到非奇异的。利用逐步回归方法,从各区保留下15、15、9个因子中,筛选出贡献大的因子,建立稳定可靠的预报模型。

下面给出 I 区预报模式,II、III 区模式略。

$$Y = -0.1784X'_1 - 3.0701X'_2 - 5.0623X'_3 + 3.6749X'_4 + 4.6176X'_5 + 33.2443$$

$$F=30.56934, S=7.693681, R=0.9296716$$

其中

$$X'_1 = (|X_1 - 21.6|/7.6 + 0.5)^{-7.33}, X_1 \text{ 为上一年7月大西洋副高月平均面积指数}(55 \sim 25^\circ \text{W});$$

$$X'_2 = (|X_2 - 17.3|/44.70 +$$

$0.5)^{-2.78}, X_2 \text{ 为上一年1月太平洋副高月平均强度指数}(110^\circ \text{E} \sim 115^\circ \text{W});$

$X'_3 = (|X_3 - 5.58|/9.92 + 0.5)^{-3.06}, X_3 \text{ 为上一年3~4月月平均北半球副高脊线}(5^\circ \text{E} \sim 360^\circ) \text{ 的平均值};$

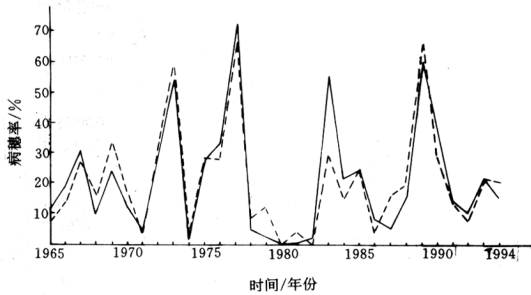
$X'_4 = (|X_4 - 178.58|/14.42 + 0.5)^{5.36}, X_4 \text{ 为上一年9~10月月平均亚洲区极涡面积指数}(1 \text{区} 60 \sim 150^\circ \text{E}) \text{ 的平均值};$

$X'_5 = (|X_5 - 84.92|/220.09 + 0.5)^{-2.70}, X_5 \text{ 为上一年7月月平均北半球极涡中心位置}(JW)。$

2.2 预报模型拟合结果与两年预报情况

利用上述预报模型进行回代检验,作出赤霉病病穗率的拟合结果,附图是 I 区的拟合情况,II、III 区拟合情况相似(图略)。可以看出,模式的历史拟合效果较好,预报值和实际值趋势一致,误差不大。(下转封三)

(上接57页)



附图 I 区的小麦赤霉病穗率的拟合结果

在1995年和1996年,对所建模式进行了业务试用。1995、1996年病穗率预报值 I 区为 23.8%、24.3%, II 区为 6.7%、7.9%, III 区为 14.4%、14.3%,实际值 I 区为 25%、20%, II 区为 8%、5%, III 区为 18%、10%,预报值与实际值基本相符。

3 小结

3.1 大气环流特征量作为长期天气过程的重要因子,对冬小麦赤霉病这一“气象型”病害也存在着间接的影响,以大气环流特征量对天气条件影响的滞后性作为赤霉病预报的因子藉以延长预报时效是行之有效的办法。

3.2 利用最优化因子相关分析方法,对与赤霉病发生发展有指示作用的大气环流因子进行寻优相关普查,可获得显著相关的因子。

3.3 文中的计算结果不仅说明大气环流特征量与病穗率的显著相关关系,而且还表明了大气环流模式的历史拟合率较高,预报的效果也较佳。

3.4 由于本文所采用的还是统计预报方法,所以它具有统计预报的通病,对一些反常年份难以报准,实际业务应用时必须综合各种模式和方法的预报结果。

参考文献

- 1 张汉琳. 气象因素与麦类赤霉病群体流行波动研究. 气象学报, 1987, 45(3): 338~345.
- 2 汤志成等. 最优化因子处理及加权多重回归模型. 气象学报, 1992, 50(4): 514~517.
- 3 汤志成等. 冬小麦赤霉病与北太平洋海温场的相关分析及其预报. 气象, 1988, 14(8): 31~35.
- 4 章基嘉等. 中长期天气预报基础. 北京: 气象出版社, 1983年: 218~253.
- 5 汤志成等. 作物产量预报系统. 中国农业气象, 1996, 17(2): 49~52.
- 6 朱盛明. 相关系数稳定性分析方法及其应用. 气象学报, 1982, 40(4): 497~501.

The Long-range Prediction Models of Winter Wheat Gibberellin Occurrence Extent

Gao Ping Ju Weimin

(Agrometeorological & Satellite Remote Sensing Center of Jiangsu Province, Nanjing 210008)

Abstract

The occurrence of winter wheat gibberellin mainly depends on the weather condition. The general circulation factors at 500hPa are the physical quantities that represent weather course of large area and the influential weather conditions as well. Some factors of general circulation at 500hPa correlated with winter wheat gibberellin in significance and stability, were found out by optimum analysis for different regions in Jiangsu province. Finally the long range prediction models for influence extent of winter wheat gibberellin were developed based on these factors.

Key Words: winter wheat gibberellin general circulation optimum factor analysis prediction model