

# 利用投影寻踪分类模型评价玉米旱后生理补偿效应

郭相平, 刘展鹏, 王青梅, 郝树荣, 王为木, 褚琳琳

(河海大学现代农业工程系, 江苏 南京 210098)

**摘要:** 根据模拟干旱试验结果, 建立了作物旱后生理补偿效应的投影寻踪分类模型(PPC Mode)。采用基于实码的加速遗传算法, 对旱后复水玉米的多项生理指标的权重进行了优化。通过多维指标的降维处理, 得到了不同生理指标的最佳投影方向和投影值。模型结果表明光合速率、气孔导度对补偿效应的影响最大, 与传统综合分析法结论一致, 表明该模型具有合理性。模型避免了主观因素对评价结果的影响, 为节水灌溉条件下作物旱后补偿效应分析和评价提供了一种新的方法。

**关键词:** 投影寻踪分类模型; 玉米旱后复水; 生理指标; 补偿效应; 综合评价

**中图分类号:** S311 **文献标识码:** A **文章编号:** 1000-7601(2007)06-0022-04

旱后复水的补偿效应是节水灌溉研究的热点问题之一。作物补偿效应包括多项指标, 如何对补偿效果进行合理评价, 进而提出适宜的水分控制指标, 是水胁迫补偿效应研究的主要目的之一。传统评价方法主要采用综合分析法, 即通过单项指标的比较以及各指标之间的因果关系或相关性, 对补偿效果进行分析<sup>[1]</sup>。不同因子对补偿效应的影响程度(即权重), 主要由研究者根据经验进行评判, 必然受到研究者主观性的影响。尤其是不同处理间各指标之间差异不显著时, 对于相近的试验结果, 不同专家可能得出不同的结论<sup>[2~4]</sup>, 即评价结果具有某种程度的不确定性, 从而可能对最佳土壤水分控制指标的确定产生不利影响。

基于以上原因, 本文以玉米模拟试验结果为基础, 采用基于实码的加速遗传算法(Real coded based Accelerating Genetic Algorithm, RAGA)的投影寻踪分类模型<sup>[5]</sup>(Projection Pursuit Classification model based on RAGA, PPC-RAGA), 对高维数据(旱后复水玉米多项生理指标)进行处理, 以避免传统分析方法中人为因素对结论的影响, 试图为作物旱后补偿机理的研究提供一种定量分析方法。

## 1 投影寻踪分类模型(PPC-RAGA模型)

利用 PPC-RAGA 模型分析旱后作物补偿效应的基本思路是: 分别将各处理的多个旱后生理指标通过某种组合投影到一维空间, 对投影构形, 采用投影指标函数(目标函数)衡量投影暴露某种评价结

构的可能性大小, 寻找出使投影指标函数达到最优(即能反映高维数据结构或特征)的投影值, 投影值越大表明该处理的补偿效应就越大(与生理指标选择有关)。其中, 投影指标函数的构造及其优化问题是应用投影寻踪分类方法能否成功的关键所在。本文采用 RAGA 处理该问题, 实现投影指标函数的优化, 进而建立旱后复水补偿效应投影寻踪分类模型(PPC-RAGA 模型), 其建模过程包括如下 4 个步骤。

步骤 1: 建立合适的评价指标体系, 对胁迫后复水生理补偿指标的样本进行标准化处理。合适的评价指标体系是建模的关键所在, 其合适与否直接影响结论的正确性。根据近年来的研究成果, 选择了继补偿生长、产量和水分利用效率影响较大的根、冠生理指标, 作为旱后补偿效应的评价指标, 如表 1 所示。这些指标越大, 表示玉米的生理活性越高, 补偿效应越好。

设胁迫复水后各处理生理补偿指标数据的样本集为  $\{x^*(i, j) \mid i = 1 \sim n, j = 1 \sim p\}$ , 其中  $x^*(i, j)$  为第  $i$  个样本第  $j$  个指标值,  $n, p$  分别为玉米生理指标样本的数目(样本容量)和评价指标的数目。为了消除各评价指标的量纲影响, 统一指标值的变化范围, 采用下式对  $x(i, j)$  进行标准化处理:

对于越大越优的指标:

$$x(i, j) = \frac{x^*(i, j) - x_{\min}(j)}{x_{\max}(j) - x_{\min}(j)} \quad (1)$$

对于越小越优的指标:

$$x(i, j) = \frac{x_{\min}(j) - x^*(i, j)}{x_{\max}(j) - x_{\min}(j)} \quad (2)$$

收稿日期: 2007-04-26

基金项目: 国家自然科学基金(50309003)

作者简介: 郭相平(1968-), 男, 山东成武人, 副教授, 博士后, 主要从事灌排理论研究。

其中,  $x_{\max}(j)$ 、 $x_{\min}(j)$  分别为第  $j$  个指标值的最大值和最小值,  $x(i, j)$  为标准化后的数据样本值。

步骤2:构造投影指标函数  $z(i)$ 。投影寻踪分类方法就是把  $p$  维数据  $\{x^*(i, j) \mid j = 1 \sim p\}$  综合成以  $a = \{a(1), a(2), \dots, a(p)\}$  为投影方向的一维投影值  $Z(i)$ , 设  $a$  为单位长度向量, 即

$$Z(i) = \sum_{j=1}^p a(j)x(i, j), (i = 1 \sim n) \quad (3)$$

然后根据  $\{Z(i) \mid i = 1 \sim n\}$  的一维散布图进行分类。综合投影指标值时, 要求投影值  $Z(i)$  的散布特征应为: 局部投影点尽可能密集, 而在整体上投影点团之间尽可能散开。即  $Z(i)$  的标准差  $S_z$  尽可能大, 同时投影值  $Z(i)$  的局部密度  $D_z$  达到最大, 基于此, 投影目标函数可构造为

$$Q(a) = S_z D_z \quad (4)$$

$$S_z = \left[ \sum_{i=1}^n (Z(i) - E(z))^2 / (n-1) \right]^{0.5} \quad (5)$$

$$D_z = \sum_{i=1}^n \sum_{j=1}^n (R - r(i, j)) \cdot u(R - r(i, j)) \quad (6)$$

式中,  $E(z)$  为序列  $\{Z(i) \mid i = 1 \sim n\}$  的均值;  $R$  为局部密度的窗口半径, 其选取既要使包含在窗口内的投影点的平均个数不太少, 避免滑动平均偏差太大, 又不能使它随着  $n$  的增大而增加太快。  $R$  的取值范围为  $r_{\max} + \frac{p}{2} \leq R \leq 2p$ ;  $r(i, j)$  表示样本之间的距离,  $r(i, j) = |Z(i) - Z(j)|$ ;  $u(R - r(i, j))$  为单位阶跃函数, 当  $R - r(i, j) \geq 0$  时, 其函数值为 1, 当  $R - r(i, j) < 0$  时其函数值为 0<sup>[6,7]</sup>。

步骤3:优化投影指标函数, 确定最佳投影方向。当玉米各处理生理补偿评价指标样本数据给定时, 投影指标函数  $Q(a)$  只随着投影方向  $a$  的变化而变化。不同的投影方向反映不同的数据结构特征, 最佳投影方向就是最大的可能暴露高维数据某类特征结构的投影方向。因此, 可以通过求解投影指标函数最大化问题来估计最佳投影方向, 即:

$$\text{最大化目标函数: Max: } Q(a) = S_z \cdot D_z \quad (7)$$

$$\text{约束条件: s.t: } \sum_{j=1}^p a^2(j) = 1, -1 \leq a(j) \leq 1 \quad (8)$$

上述问题是一个以  $\{a(j) \mid j = 1 \sim p\}$  为优化变量的非线性优化问题, 常规优化方法处理较复杂。因此, 本文应用模拟生物优胜劣汰规则与群体内部染色体信息交换机制的基于实数编码的加速遗传算法(RAGA)来解决其高维全局寻优问题。

步骤4:分类(优序排列)。把由步骤3求得的最

佳投影方向  $a^*$  代入式(3)后可得各样本点的投影值  $Z^*(i)$ , 然后对各个样本进行分类评价。将  $Z^*(i)$  与  $Z^*(j)$  进行比较, 二者越接近, 表示样本  $i$  与  $j$  越倾向于分为同一类。若按  $Z^*(i)$  值从大到小排序, 则可以将样本从优到劣进行排序。

## 2 实例分析

试验于2005年5~8月在河海大学节水园区温室玻璃大棚内进行模拟干旱胁迫试验, 供试验品种为高产、杂交型玉米品种——农大108。试验以高聚分子聚乙二醇(PEG-6000)作为渗透剂, 在5叶期(苗期)进行胁迫处理。处理共设3个水平, 即充分供水对照(CK)、轻度胁迫(溶液水势-0.1 MPa)、重度胁迫(溶液水势-0.2 MPa), 对应的全营养液中PEG浓度分别为0、40、82 g/L。胁迫持续时间为5 d。分别在复水后0、1、3、5、15 d取样, 测试各处理根、冠的生理指标。各处理均重复3次。各生理指标的测试结果如表1所示。

将表1中数据作为玉米5叶期胁迫5d后复水的15个评价样本, 每个样本选有6个生长生理补偿评价指标, 即属于6维数据。对表1中数据建立综合效益分类评价的PPC-RAGA模型, 过程中选定父代初始种群规模为  $n=400$ , 交叉概率  $p_c=0.80$ , 变异概率  $p_m=0.80$ , 优秀个体数目选定为20个,  $\alpha=0.05$ , 加速次数为7。把该样本按式(1)进行标准化处理后, 依次代入式(3)、(5)、(6)和(4), 即得比例的投影指标函数, 然后用RAGA优化式(7)和式(8)所定的问题, 得密度窗宽  $R=4.8306$ , 最大投影指标函数值为1166.3, 最佳投影方向为:  $a^* = (0.0032, 0.4560, 0.5918, 0.5667, 0.2672, 0.2218)$ 。把  $a^*$  代入式(4)后即得各测量点的样本建立的综合效益分类评价样本的投影值:  $Z^*(i) = (0.8539, 0.8088, 0.0416, 0.9800, 0.9958, 0.3230, 1.1550, 1.6829, 0.8342, 1.2155, 1.8722, 0.9520, 0.8147, 0.8433, 0.8819)$ , 结果见表1, 投影值的相对值(与对照处理的比值)见图1。

样本的投影值越大, 表明该样本的生理活性越高, 亦即旱后生理指标的补偿效果越好。从图1还可以发现, 五叶期胁迫处理的玉米, 在复水后其生理指标的综合投影值近似呈抛物线变化。5 d胁迫结束时, 轻度胁迫处理的玉米幼苗与对照之间投影值接近, 复水3 d后超过对照, 5 d后达到最大 [ $Z^*(11)=1.8700$ ], 其后与对照逐渐接近。这表明轻度胁迫对作物的生理活性影响不大, 复水后生理活性逐渐升高并出现超越补偿效应, 但超越补偿效应的

持续时间并不太长,即生理指标的超越补偿效应具有时间性。重度胁迫处理在 5d 胁迫结束时,其相对综合投影值达到最小值( $Z^*(3)=0.0416$ ),在其后的约 2 周内仍然低于对照,超越补偿效应不明显。这表明重度胁迫严重抑制了玉米的生理活性,复水后需要较长时间才能恢复。

最佳投影方向各分量的大小实质上反映了玉米苗期旱后复水各生理补偿评价指标对总体补偿效应的综合效益评价影响程度,见图 2。本例中,最佳投影方向  $\alpha^*=(0.0032, 0.4560, 0.5918, 0.5667, 0.2672, 0.2218)$ ,表明各生理指标对旱后补偿效应的影响程度大小依次为净光合速率、气孔导度、叶绿素含量、PS II 原初光能转化效率、PS II 的潜在活性、根系活力。这是因为光合作用是干物质累积的最终

来源,其大小直接反映了旱后同化物生产能力的高低,从而决定了补偿生长的程度。从投影方向的分量大小也可以看出,复水后短时间内,补偿效应主要与冠部生理指标的补偿效应有关,而根系活力对补偿效应的影响较小,这与复水后生长中心向地上部分转移有关<sup>[8]</sup>。

通过对投影值的分析,可得到旱后生理指标的总变化趋势以及各指标对补偿效应的影响权重,据此对旱后补偿效应进行分析,所得结论与采用综合分析法所得之结论基本一致<sup>[9,10]</sup>,表明 PPC-RAGA 模型可从整体上反映植物旱后补偿效应的实际情况,避免了主观性和人为因素的干扰,具有合理性。

表 1 复水后玉米生理指标及综合效益分类评价样本的投影值

Table 1 Physiological indexes of maize after rewatering and comprehensive values of projection function

| 取样日期<br>Sampling date   | 测量序号<br>NO. | 处理号<br>Treatment | TTC    | CHL   | $P_n$  | $G_s$ | $F_v/F_m$ | $F_v/F_0$ | $Z^*(i)$ |
|-------------------------|-------------|------------------|--------|-------|--------|-------|-----------|-----------|----------|
| 复水 0 天<br>Rewater 0 d   | 1           | CK               | 11.167 | 1.586 | 23.930 | 0.260 | 0.708     | 2.451     | 0.854    |
|                         | 2           | -0.1 MPa         | 11.522 | 2.216 | 19.770 | 0.190 | 0.715     | 2.509     | 0.809    |
|                         | 3           | -0.2 MPa         | 12.801 | 1.428 | 13.880 | 0.090 | 0.620     | 1.658     | 0.042    |
| 复水 1 天<br>Rewater 1 d   | 4           | CK               | 7.503  | 1.956 | 22.350 | 0.210 | 0.757     | 3.148     | 0.980    |
|                         | 5           | -0.1 MPa         | 8.648  | 2.460 | 20.640 | 0.210 | 0.738     | 2.826     | 0.996    |
|                         | 6           | -0.2 MPa         | 8.265  | 1.682 | 12.660 | 0.090 | 0.721     | 2.587     | 0.323    |
| 复水 3 天<br>Rewater 3 d   | 7           | CK               | 5.784  | 2.187 | 25.470 | 0.250 | 0.746     | 2.942     | 1.155    |
|                         | 8           | -0.1 MPa         | 7.118  | 2.368 | 29.970 | 0.410 | 0.780     | 3.559     | 1.683    |
|                         | 9           | -0.2 MPa         | 6.027  | 1.817 | 20.710 | 0.180 | 0.752     | 3.070     | 0.834    |
| 复水 5 天<br>Rewater 5 d   | 10          | CK               | 5.412  | 2.211 | 30.020 | 0.290 | 0.695     | 2.323     | 1.216    |
|                         | 11          | -0.1 MPa         | 5.795  | 3.165 | 31.470 | 0.500 | 0.721     | 2.608     | 1.872    |
|                         | 12          | -0.2 MPa         | 5.686  | 2.525 | 21.040 | 0.170 | 0.732     | 2.752     | 0.952    |
| 复水 15 天<br>Rewater 15 d | 13          | CK               | 2.783  | 2.619 | 12.970 | 0.100 | 0.788     | 3.732     | 0.815    |
|                         | 14          | -0.1 MPa         | 2.958  | 2.629 | 13.220 | 0.110 | 0.789     | 3.757     | 0.843    |
|                         | 15          | -0.2 MPa         | 2.982  | 2.905 | 12.820 | 0.090 | 0.791     | 3.789     | 0.882    |

注: TTC 为根系活力,  $\text{mg}/(\text{g FW} \cdot \text{h})$ ; CHL 为叶绿素含量,  $\text{mg}/\text{g FW}$ ;  $P_n$  为净光合速率,  $\text{CO}_2 \mu\text{mol}/(\text{m}^2 \cdot \text{s})$ ;  $G_s$  为气孔导度,  $\mu\text{mol}/(\text{m}^2 \cdot \text{s})$ ;  $F_v/F_m$  为 PS II 原初光能转化效率;  $F_v/F_0$  为 PS II 的潜在活性;  $Z^*(i)$  为样本投影值。

Note: TTC—Root activity,  $\text{mg}/(\text{g FW} \cdot \text{h})$ ; CHL—chlorophyll content,  $\text{mg}/\text{g FW}$ ;  $P_n$ —Net Photosynthesis rate,  $\text{CO}_2 \mu\text{mol}/(\text{m}^2 \cdot \text{s})$ ;  $G_s$ —Stoma conductivity,  $\mu\text{mol}/(\text{m}^2 \cdot \text{s})$ ;  $F_v/F_m$ —Efficiency of excitation energy capture by open PS II reaction centers;  $F_v/F_0$ —Potential activity of PS II;  $Z^*(i)$ —Projection values of samples.

### 3 结 语

玉米旱后复水所产生的生理补偿效应包括多个方面,不同指标的变化趋势各异,传统的综合分析较为困难,尤其是当各处理生理指标差异水平不显著,或者不同指标间关系不易确定时,受研究者认知结构等主观因素影响,所得结论具有一定的不确定性,据此确定的最佳水分控制指标对生产实践的指导性

较差,而 PPC 模型克服了上述缺点,使结论具有唯一性。只要合理选择评价指标,其评价结果符合实际,令人满意。且 PPC 模型克服了传统投影寻踪方法计算复杂、编程实现困难的缺点,解决了高维数据全局寻优的难题,大大减少了寻优工作量。为节水灌溉条件下作物生理、生态指标研究提供了一条值得探索的新径。

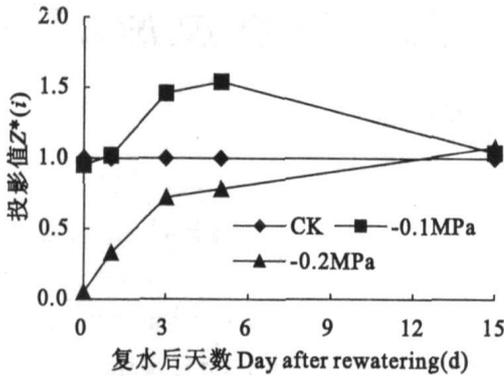


图 1 复水后生理综合指标投影值(相对值)

Fig.1 Relative comprehensive projection values after rewatering

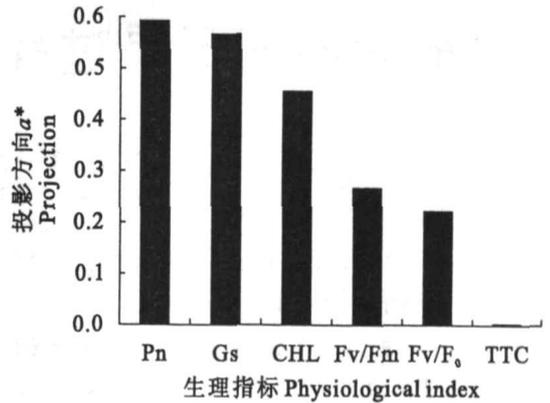


图 2 复水后生理补偿效应综合评价各评价指标的投影方向

Fig.2 Projection of physiological indexes after rewatering

生理活性的补偿是旱后作物生长补偿的基础。PPC 模型的评价结果表明,轻度胁迫处理的玉米,其生理活性可在复水后很快赶上并超过对照,出现超补偿效应。这种效应可保持 2 周左右,然后与对照逐渐接近。对于重度胁迫处理,复水后虽有恢复,但仍然表现为抑制效应,且需要较长的恢复时间才能恢复到对照水平。因此,在玉米 5 叶期,水分控制应以轻度胁迫为宜,避免重度胁迫,以利于后期的补偿生长。

参考文献:

[1] 郭相平,张烈君,王 琴,等.作物水分胁迫补偿效应研究进展[J].河海大学学报,2005,33(6):634-637.  
 [2] 丁端锋,蔡焕杰,王 健,等.玉米苗期调亏灌溉的复水补偿效应[J].干旱地区农业研究,2006,24(3):64-67.  
 [3] 刘庚山,郭安红,任三学,等.夏玉米苗期有限水分胁迫拔节期

复水的补偿效应[J].生态学杂志,2004,23(3):24-29.  
 [4] 郭相平,王 琴,刘展鹏,等.旱后复水对玉米后继新生叶片生理特性的影响[J].农业科学研究,2006,27(2):20-23.  
 [5] 付 强,赵小勇.投影寻踪模型原理及其应用[M].北京:科学出版社,2006.  
 [6] 赵小勇,付 强,邢贞相,等.投影寻踪模型的改进及其在生态农业建设综合评价中的应用[J].农业工程学报,2006,22(5):222-225.  
 [7] 吴承祯,洪 伟,洪 滔.基于改进的投影寻踪的森林生态系统生态位分级模型[J].应用生态学报,2006,17(3):357-361.  
 [8] 吕金印,李成龙,张俊林,等.干旱对不同基因型春小麦碳同化物分配的影响[J].干旱地区农业研究,2005,26(3):84-87.  
 [9] 郭相平,康绍忠.玉米调亏灌溉的后效性[J].农业工程学报,2000,16(4):58-60.  
 [10] 蔡焕杰,康绍忠,张振华,等.作物调亏灌溉的适宜时间与调亏程度的研究[J].农业工程学报,2000,16(3):24-27.

## Comprehensive Evaluation on Compensatory Effects of Rewatering after PEG Stress Based on PPC-RAGA Model

GUO Xing-ping, LIU Zhan-peng, WANG Qing-mei, HAO Shu-rong, WANG Wei-mu, CHU Lin-lin  
 (Department of Irrigation and Drainage, Hohai University, Nanjing, Jiangsu 210098, China)

**Abstract:** A projection pursuit classification model(PPC Model) was suggested in this paper to evaluate the comprehensive compensatory effects of water recovery after drought based on simulation water stress experiments on maize. The optimum projection direction and the value of the physiological properties of maize leaves were obtained based on real coded accelerating genetic algorithm(RAGA) and the comprehensive index which synthesized the characteristics of compensatory effects of the physiological properties was also got by projecting high dimensional data to low dimensional space. The results of the model indicated net photosynthetic rate and stomata conductivity had the greatest effect on compensatory effects, which was consistent with the conclusions from traditional comprehensive analysis, indicating the model was reasonable if physiological properties were chosen correctly. The model avoided subjective ingredient influences of investigator on conclusion and could be used to investigate the effects of water stress and water recovery on compensatory effects under water saving irrigation.

**Keywords:** PPC model; water recovery after drought; compensatory effect; physiological properties comprehensive assessment