

投影寻踪分类模型在作物补偿效应评价中的应用*

郝树荣 郭相平 张展羽
(河海大学农业工程学院, 南京 210098)

【摘要】 为了解决作物旱后复水补偿效应评价和优选的不确定性,提高补偿效应评价模型的分辨率,提出了投影寻踪分类模型,并采用变异和动态信息素更新蚁群算法寻找最优的投影方向,用最佳投影方向信息研究各评价指标对补偿效应的贡献率,发现光合速率的大小直接反应了补偿生长的能力,是影响补偿效应的关键因子,与以往研究结论相符,且适度胁迫复水后第5天表现出的补偿效应最佳。

关键词: 水分胁迫 补偿效应 投影寻踪分类 综合评价
中图分类号: Q945.17 文献标识码: A 文章编号: 1000-1298(2010)01-0059-04

Comprehensive Evaluation on Compensatory Effects of Water Recovery after Drought Based on Projection Pursuit Classification Model

Hao Shurong Guo Xiangping Zhang Zhanyu
(College of Agricultural Engineering, Hohai University, Nanjing 210098, China)

Abstract

An effective projection pursuit classification model is suggested to evaluate compensatory effects of water recovery after drought in order to avoid the uncertainty of the evaluation and the optimization, and to improve the accuracy of the evaluating model. It adopts an ant colony algorithm based on mutation and dynamic pheromone updating to find the best projective direction, and then uses the best projective direction to calculate the contribution of each evaluation index to the compensatory effects. The test results showed that photosynthetic rate directly reflected the capacity of compensatory growth. Thus it is a critical factor for compensatory effects, which conformed to the previous studies. The optimal compensatory effects appeared in the fifth day after appropriate rewatering.

Key words Water stress, Compensatory effects, Projection pursuit classification, Comprehensive evaluation

引言

近年来,对水分亏缺与作物生长关系的研究中发现,作物旱后复水存在明显的补偿效应^[1~3],由于不同胁迫程度、胁迫历时的水分胁迫,对作物各项生理指标、干物质累积及产量的影响不一致^[3~4],很难确定哪个处理的补偿效应是最佳的。如何在综合考虑生理及产量等各项指标下,对多种处理方案的补偿效果作出合理评价,进而提出适宜的水分控制指标成为亟待解决的问题之一。

以往补偿效应的评价多通过单项指标的比较及各指标间的相关性,对补偿效果进行定性分析^[5],量化研究多采用综合分析法,即把各评价指标赋权后得到一个综合数值,对于权重的赋予,具有一定的主观性。基于熵的综合评价方法可根据评价指标间的变异程度确定权重,在一定程度上消除了人为赋权的缺点,但在实际应用中常出现权重的平均化现象,并缺乏各指标对总体目标贡献大小的评价^[6]。综合分析法的实质是对高维数据进行降维处理,专家给出的权重矩阵对应于每个指标的低维

收稿日期: 2009-03-09 修回日期: 2009-05-18
* 国家自然科学基金资助项目(50309003,50839002)
作者简介: 郝树荣,副教授,博士,主要从事节水灌排理论与技术研究,E-mail: srhao@ hhu. edu. cn

投影值,但无法确定权重矩阵是否属于各项指标在低维子空间的最佳投影^[6~7]。本文采用投影寻踪分类模型(projection pursuit classification,简称PPC)实现高维数据的降维过程^[7],并将基于变异和动态信息素更新的蚁群算法^[8~10]融入模型,利用该算法优化模型中的投影方向,以避免传统分析方法中人为因素对结论的影响。

1 模型建立

投影寻踪是分析高维数据的新兴统计方法。它把高维数据投影到低维子空间,通过分析低维空间数据点的变化特征来研究高维数据结构特征,具有稳健性好、抗干扰性强等优点,已在环境、农业、水利等复杂系统的优化及评价中得到了成功应用^[6~7,11]。

投影寻踪分类模型的建模包括以下步骤^[6~7,11]:

(1) 投影指标的建立。为消除各评价指标值数量级上的差异,需进行标准化处理。越大越优的指标采用式(1),越小越优的指标采用式(2)进行标准化处理,即

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

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

式中 $x_{\max}(j)$ 、 $x_{\min}(j)$ ——第 j 个指标的最大值和最小值

$x(i,j)$ ——标准化后的数据样本值
 $x^*(i,j)$ ——标准化前的数据样本值

(2) 构造投影指标函数。PPC 法是把 p 维数据 $\{x(i,j) | j=1,2,\cdots,p\}$ 综合成以 $a = \{a(1), a(2), \cdots, a(p)\}$ 为投影方向的一维投影值 $z(i)$,即

$$z(i) = \sum_{j=1}^p a(j)x(i,j) \quad (i=1,2,\cdots,n)$$
 (3)

式中 a ——单位长度向量

(3) 构造投影目标函数。为了在多维指标中找到数据的结构特征,在综合投影时,要求投影值 $z(i)$ 尽可能多地提取 $x(i,j)$ 中的变异信息,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构造为

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

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

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

其中

$$r(i,j) = |z(i) - z(j)|$$

$$u(R - r(i,j)) = \begin{cases} 1 & (R - r(i,j) \geq 0) \\ 0 & (R - r(i,j) < 0) \end{cases}$$

式中 $E(z)$ ——序列 $\{z(i) | i=1,2,\cdots,n\}$ 的均值
 R ——局部密度的窗口半径,取值范围为

$$r_{\max} + \frac{p}{2} \leq R \leq 2p$$

$r(i,j)$ ——样本间的距离

$u(R - r(i,j))$ ——单位阶跃函数

(4) 优化投影目标函数,确定最佳投影方向。当各指标值的样本集给定时,投影指标函数 $Q(a)$ 只随投影方向 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可能暴露高维数据特征的投影方向,可通过求解投影指标函数最大化问题估计最佳投影方向,即

$$\begin{cases} \max Q(a) = S_z D_z \\ \text{s. t. } \sum_{j=1}^p a^2(j) = 1 \quad (-1 \leq a(j) \leq 1) \end{cases}$$
 (7)

(5) 分类排序。把由式(7)求得的最佳投影方向 a^* 代入式(3),可得各样本点的投影值 $z^*(i)$ 。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 变异和动态信息素更新蚁群算法的运用

投影寻踪分类模型实质是一个以 $a(j)$ 为优化变量的复杂非线性优化问题,因此利用投影寻踪解决实际问题的关键是构造能够找到最优投影方向的有效算法。以往基于旋转变换思想的梯度下降法、高斯-牛顿法等传统的优化方法速度慢且易陷入局部最优。近年来人们开始大胆探索新的非经典计算途径,从生物进化的机理中受到启发,提出了遗传算法、进化规划等,从一些社会性动物(如蚁群、蜂群等)的自组织行为得到启发,提出了蚁群算法、粒子群算法等群智能算法。群智能算法作为一种新兴的演化计算技术已成为越来越多研究者的关注焦点^[12~13]。本文采用变异和动态信息素更新蚁群算法,用动态信息更新策略加速全局收敛,用独特的变异策略来加快局部寻优,更好地解决高维全局寻优问题。本算法的原理是:蚂蚁搜索路径时,若找到一段很短的子路径,就释放出相应浓度的信息素,该信息素一方面影响位于子集的两个点上的蚂蚁,另一方面会以该路径为中心向外扩散,影响路径附近其他蚂蚁的行为,使它们以更大概率在下一步选择此路径。通过这种基于信息素扩散的协作方式,其他蚂蚁在选择下一路口时选择到最优路径的干扰性会减低,从而使算法的收敛速度大大提高^[8]。

算法步骤参见文献[8]:算法中需按式(9)计算每只蚂蚁 k 将要转移的位置,假设该位置为 j ,上一

个位置为 i , 则

$$P_{ij}^k(t) = \begin{cases} \tau_{ij}^\alpha(t) \eta_{ij}^\beta(t) / \sum_{k \notin \text{tabu}_k} \tau_{ik}^\alpha(t) \eta_{ik}^\beta(t) & (j \notin \text{tabu}_k) \\ 0 & (j \in \text{tabu}_k) \end{cases} \quad (8)$$

式中 $P_{ij}^k(t)$ 表示 t 时刻蚂蚁 k 由路口 i 转移到路口 j 的概率。设有路口 m 个, 蚂蚁 n 个, η_{ij} 为路口 i 转移到路口 j 的启发信息, 一般取 $\eta_{ij} = 1/d_{ij}$, d_{ij} 表示路口 i 和 j 之间的距离, β 为启发信息的重要程度; $\tau_{ij}(t)$ 表示 t 时刻在路口 i 与 j 连线上残留的信息素浓度, α 为在路径 ij 上残留信息的重要程度, 初始时刻各条路径上信息素浓度相等, 设 $\tau_{ij}(0)$ (C 为常数), 蚂蚁 k 在运动过程中, 根据各条路径上的信息素浓度决定转移方向。 tabu_k ($k = 1, 2, \cdots, n$) 为蚂蚁 k 已走过路口的集合, 开始时 tabu_k 中只有一个元素, 即蚂蚁 k 的出发路口; 随着进化的进行, tabu_k 中的元素不断增加; 随着时间的推移, 以前留在各条路径上的信息素逐渐消逝, 用参数 $1-\rho$ 表示信息素的挥发程度, 经过 m 个时刻, 蚂蚁完成一次循环, 各路径上信息素的浓度调整为

$$\tau_{ij}(t+m) = \rho \tau_{ij}(t) + \Delta \tau_{ij} \quad (\rho \in [0, 1]) \quad (9)$$

其中

$$\Delta \tau_{ij} = \sum_{k=1}^n \Delta \tau_{ij}^k$$

根据已有试验^[8], 蚂蚁在出发点附近走的路径

往往是较优的, 离出发点越远路径越复杂, 与其他蚂蚁走的路径重叠增多, 走出较优路径的概率降低。为避免信息素堆积过多及体现众多蚂蚁的共同影响, 每只蚂蚁留下的信息素应随着向远处的延伸而逐渐减弱, 由于蚂蚁间的相互联系以及对搜索的主要贡献是留下信息素, 因此动态信息素更新策略是决定收敛速度和全局寻优的关键, 它使蚁群算法的全局搜索能力增强, 可迅速收敛到一个较优解, 但从较优解到最优解所花费的收敛时间所占比重很大, 利用变异算法局部寻优能力强的特点, 在收敛到一定代数时, 对最优个体进行变异, 可大大加快局部寻优速度。

3 模型在作物旱后复水补偿效应评价中的应用

试验于 2005 年在江苏省河海大学节水园区进行, 试材为玉米—当地杂交型品种农大 108, 本试验设置了单因素 3 水平, 因素为水分胁迫时期: 玉米 5 叶期, 水分胁迫采用模拟胁迫形式, 即通过 PEG—6000 的浓度来控制胁迫程度, 胁迫历时为 5 d, 3 个胁迫水平分别为: ①非胁迫供水处理 CK、溶液水势为 0 MPa; ②溶液水势为 -0.1 MPa 的胁迫处理; ③溶液水势为 -0.2 MPa 的胁迫处理。每个处理重复 3 次。试验数据如表 1 所示。

表 1 玉米 5 叶期旱后复水补偿效应综合评价指标值及投影值

Tab.1 Indexes and projection values of comprehensive evaluation on compensatory effects of water recovery after drought on maize

复水天数 /d	样本 序号	溶液水势 /MPa	根系活力 /mg·g ⁻¹ ·h ⁻¹	叶绿素含量 /mg·g ⁻¹ FW	光合速率 /μmol·m ⁻² ·s ⁻¹	气孔导度 /mol·m ⁻² ·s ⁻¹	原初光能 转化效率	干物质累 积量/g	投影值
0	1	0	11.167	1.70	28.36	0.240	0.708	3.70	1.424 0
	2	-0.1	11.522	1.63	22.99	0.210	0.715	3.32	1.140 3
	3	-0.2	12.801	1.38	20.50	0.170	0.620	2.57	0.837 2
1	4	0	7.503	1.96	22.46	0.210	0.757	4.09	1.152 7
	5	-0.1	8.648	2.46	21.08	0.195	0.738	3.45	0.987 0
	6	-0.2	8.265	1.68	12.66	0.090	0.721	2.43	0.343 3
3	7	0	5.784	2.19	25.47	0.253	0.746	4.88	1.273 5
	8	-0.1	7.118	2.37	29.97	0.410	0.780	4.42	1.562 7
	9	-0.2	6.027	1.82	20.71	0.180	0.752	2.52	0.603 2
5	10	0	5.412	2.21	30.96	0.280	0.695	7.37	1.463 4
	11	-0.1	5.795	3.01	31.47	0.500	0.721	7.55	1.814 4
	12	-0.2	5.686	2.52	21.04	0.150	0.732	3.06	0.502 0
15	13	0	2.783	2.22	11.36	0.088	0.788	21.69	0.586 9
	14	-0.1	2.958	2.44	13.18	0.085	0.789	20.78	0.612 8
	15	-0.2	2.982	3.00	12.91	0.083	0.791	8.43	0.115 6

试验选取了 $j = 15$ 个评价样本, 分别是 3 个控水处理在胁迫结束、复水 1、3、5 d 和复水 15 d 时的数据; 每个样本选取 6 个具有代表性的评价因子, 分

别是根系活力、叶绿素含量、光合速率、气孔导度、PS II 原初光能转化效率和干物质累积量, 即属于六维参数寻优问题。对表 1 中的数据建立综合效益分

类评价的投影寻踪模型,得: $\rho = 0.2, \alpha = 0.3, \beta = 5$, $Q = 1\,800, m = 35$;最大投影指标函数值为461.98;最佳投影方向为: $a^* = (0.185\,2, 0.055\,3, 0.603\,0, 0.534\,8, 0.010\,5, 0.559\,4)$,最佳投影方向各分量的大小反映了玉米旱后复水各评价指标对总体补偿效应综合评价的贡献大小,即对玉米5叶期补偿效应的影响程度大小依次为光合速率、干物质累积量、气孔导度、根系活力、叶绿素含量和PSⅡ原初光能转化效率。光合作用是补偿效应的物质基础,其大小直接反应了补偿生长的能力,它是影响补偿效应的关键因子,此与以往研究结论相符^[4]。15个样本的投影值为: $z^*(i) = (1.424, 1.140\,3, 0.837\,2, 1.152\,7, 0.987\,0, 0.343\,3, 1.273\,5, 1.562\,7, 0.603\,2, 1.463\,4, 1.814\,4, 0.502\,0, 0.586\,9, 0.612\,8, 0.115\,6)$,样本的投影值越大,表明该样本的补偿效应越好。投影值的排序如表1所示,11号样本的补偿效应最大,即适度胁迫复水后第5天表现出的补偿效应最佳。另外从图1可见,玉米旱后复水补偿效应近似呈抛物线。胁迫结束,轻旱与对照间投影值接近,而重旱与对照差异明显,表明轻旱对作物的生理活性影响不大,而重旱明显影响了作物的生理活性。复水3 d后,轻旱投影值已明显高于对照,出现超越补偿,而重旱到复水后15 d,投影值才逐渐接近对照,表现出明显的

滞后性。

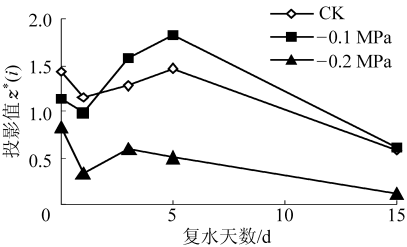


图1 旱后复水补偿效应综合评价投影值
Fig.1 Projection values of comprehensive evaluation on compensatory effects of water recovery after drought

4 结论

- (1) 光合速率的大小直接反应了补偿生长的能力,是影响补偿效应的关键因子,此与以往研究结论相符,并且玉米5叶期短历时适度胁迫复水后补偿效应最优。
- (2) 投影寻踪分类模型应用到玉米旱后复水补偿效应评价中,只要合理选择评价指标,其评价结果符合实际,并能客观地反映各评价指标的贡献率。
- (3) 将基于变异和动态信息素更新的蚁群算法与投影寻踪分类模型相结合,解决了高维数据全局寻优的难题,减少了寻优工作量,使模型能够得以更为广泛的应用。

参 考 文 献

1 山仑,苏珮,郭礼坤,等. 不同类型植物对干湿交替环境的反应[J]. 西北植物学报, 2000,20(2):164~170.
Shan Lun, Su Pei, Guo Likun, et al. The response of different crops to drying wetting cycle in field[J]. Acta Botanica Boreali-occidentalia Sinica, 2000,20(2):164~170. (in Chinese)

2 Messina F J, Durham S L. Trade-off between plant growth and defense? A comparison of sagebrush populations[J]. Oecologia, 2002,131(1):43~51.

3 赵丽英,邓西平,山仑. 水分亏缺下作物补偿效应类型及机制研究概述[J]. 应用生态学报,2004,15(3):523~526.
Zhao Liying, Deng Xiping, Shan Lun. A review on types and mechanisms of compensation effect of crops under water deficit [J]. Chinese Journal of Applied Ecology, 2004,15(3):523~526. (in Chinese)

4 Hao Shurong, Guo Xiangping, Wang Weimu, et al. Research on compensatory effects of water-recovery on rice at tillering stage[C] // Effective Utilization of Agricultural Soil and Water Resources and Protection of Environment. Nanjing: Hohai University Press, 2006.

5 王密侠,康绍忠,蔡焕杰,等. 调亏对玉米生态特性及产量的影响[J]. 西北农林科技大学学报,2000,28(1):31~36.
Wang Mixia, Kang Shaozhong, Cai Huanjie, et al. The effect of regulated deficit irrigation on ecological characteristics and yield of corn[J]. Journal of Northwest Sci-Tech University of Agriculture and Forestry, 2000,28(1):31~36. (in Chinese)

6 封志明,郑海霞,刘宝勤. 基于遗传投影寻踪模型的农业水资源利用效率综合评价[J]. 农业工程学报,2005,21(3):66~67.
Feng Zhiming, Zheng Haixia, Liu Baoqin. Comprehensive evaluation of agricultural water use efficiency based on genetic projection pursuit model[J]. Transactions of the CSAE, 2005,21(3):66~67. (in Chinese)

7 付强,杨广林,金菊良. 基于PPC模型的农机选型与优序关系研究[J]. 农业机械学报,2003,34(1):101~107.
Fu Qiang, Yang Guanglin, Jin Juliang. Selection of agricultural machinery types and their optimum order based on PPC model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2003,34(1):101~107. (in Chinese)

- 2 Hewko L O, Weber T R. Hydraulic energy storage based hybrid propulsion system for a terrestrial vehicle[C] // Proceedings of the Intersociety Energy Conversion Engineering Conference, 1990: 99 ~ 105.
- 3 何玉彬,刘艳秋,徐立勤,等. 电液伺服系统神经网络在线自学习自适应控制[J]. 中国电机工程学报,1998,18(6): 434 ~ 437.
He Yubin, Liu Yanqiu, Xu Liqin, et al. Neural network based on-line self-learning adaptive control for electro-hydraulic servo system[J]. Proceedings of the CSEE, 1998, 18(6): 434 ~ 437. (in Chinese)
- 4 刘金琨. 滑模变结构控制 MATLAB 仿真[M]. 北京:清华大学出版社,2005.
- 5 吴立刚,王常虹,高会军. 不确定线性 2-D 离散系统的鲁棒滑模控制[J]. 控制与决策,2007,22(2): 148 ~ 154.
Wu Ligang, Wang Changhong, Gao Huijun. Robust sliding mode control for a class of uncertain linear 2-D discrete-time systems[J]. Control and Decision, 2007, 22(2): 148 ~ 154. (in Chinese)
- 6 米阳,潘伟,井元伟. 一类不确定时滞系统的模糊滑模控制[J]. 控制与决策,2006,21(11): 1 280 ~ 1 283.
Mi Yang, Pan Wei, Jing Yuanwei. Fuzzy sliding mode control for a class of uncertain time-delay system[J]. Control and Decision, 2006, 21(11): 1 280 ~ 1 283. (in Chinese)
- 7 张颖伟,王小刚,王福利. 基于模糊自适应滑模的非线性系统的故障调节[J]. 控制与决策,2005,20(4): 408 ~ 411.
Zhang Yingwei, Wang Xiaogang, Wang Fuli. Fault accommodation for nonlinear system using fuzzy adaptive sliding control [J]. Control and Decision, 2005, 20(4): 408 ~ 411. (in Chinese)
- 8 蔡素芬,李文林. 存在外界干扰的连续系统离散自适应滑模控制[J]. 控制与决策,2005,20(11): 1 317 ~ 1 320.
Cai Sufen, Li Wenlin. Discrete sliding mode control for continuous system with disturbance[J]. Control and Decision, 2005, 20(11): 1 317 ~ 1 320. (in Chinese)
- 9 白寒,王庆九,徐振,等. 阀控非对称缸系统多级滑模鲁棒自适应控制[J]. 农业机械学报,2009,40(10):193 ~ 198.
Bai Han, Wang Qingjiu, Xu Zhen, et al. Multiple sliding mode robust adaptive control for valve controlled asymmetric cylinder system[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(10): 193 ~ 198. (in Chinese)
- 10 陈宝瑞,马彪,李和言,等. 高速履带车辆静液传动改进模糊控制[J]. 农业机械学报,2009,40(11):18 ~ 21.
Chen Baorui, Ma Biao, Li Heyan, et al. Improved fuzzy control of high speed tracked vehicle hydrostatic transmission[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(11): 18 ~ 21. (in Chinese)

(上接第 62 页)

- 8 朱庆保,杨志军. 基于变异和动态信息素更新的蚁群优化算法[J]. 软件学报,2004,15(2):185 ~ 192.
Zhu Qingbao, Yang Zhijun. An ant colony optimization algorithm based on mutation and dynamic pheromone updating[J]. Journal of Software, 2004, 15(2): 185 ~ 192. (in Chinese)
- 9 Merkle D, Middendorf M, Schmeck H. Ant colony optimization for resource-constrained project scheduling[J]. IEEE Trans. on Evolutionary Computation, 2002, 6(4): 333 ~ 339.
- 10 Parpinelli R S, Lopes H S, Freitas A A. Data mining with an ant colony optimization algorithm[J]. IEEE Trans. on Evolutionary Computation, 2002, 6(4): 321 ~ 328.
- 11 邵光成,张展羽,刘娜,等. 投影寻踪分类模型在膜下滴灌模式评价中的应用[J]. 水利学报,2007,38(8):944 ~ 947.
Shao Guangcheng, Zhang Zhanyu, Liu Na, et al. Application of projection pursuit model to evaluate under mulch drip irrigation scheme[J]. Journal of Hydraulic Engineering, 2007, 38(8): 944 ~ 947. (in Chinese)
- 12 陈广洲,汪家权,解华明. 粒子群算法在投影寻踪模型优化求解中的应用[J]. 计算机仿真,2008,25(8):159 ~ 165.
Chen Guangzhou, Wang Jiaquan, Xie Huaming. Application of particle swarm optimization for solving optimization problem of projection pursuit modeling[J]. Computer Simulation, 2008, 25(8): 159 ~ 165. (in Chinese)
- 13 高尚,杨静宇. 群智能算法及其应用[M]. 北京:中国水利水电出版社,2006.