

温室作物生长环境的主成分分析

武文亮,张志斌,李王东岳

(内蒙古大学 计算机学院,内蒙古 呼和浩特 010021)

摘 要:温室大棚可以帮助瓜果蔬菜在其非时令季节提供生长环境和增加产量。由于季节对于植物自然生长的不适宜性,因此作物对于温室大棚的生长环境要求就要精确得多。通过搭建 Zigbee 无线传感器网络有效地对温室草莓生长的空气湿度、土壤湿度、土壤盐碱度、温度、CO₂ 浓度及光照六个指标进行了长期的数据采集,并尝试通过主成分分析的方法来分析影响温室草莓育苗期生长的关键因素。通过对育苗期温室草莓的上述六个环境指标进行主成分分析,分析所得的三个主分量的贡献率分别为 44.57%、29.00% 和 15.83%,累计贡献率可达到 89.40%。因此,可以用这三个主分量代替原有的六个单项指标反映原指标的绝大部分信息对温室草莓生长环境进行研究,且使得各综合指标所代表的信息不再重叠。有助于实现温室种植的再生产以及高效的、精准化的管理。

关键词:主成分分析;温室作物;生长环境;无线传感器网络

中图分类号:TP39

文献标识码:A

文章编号:1673-629X(2016)10-0165-04

doi:10.3969/j.issn.1673-629X.2016.10.036

Principal Component Analysis of Greenhouse Crop Growth Environment

WU Wen-liang, ZHANG Zhi-bin, LI Wang-dongyue

(College of Computer, Inner Mongolia University, Hohhot 010021, China)

Abstract: Greenhouses can provide fruits and vegetables with growth environment in its non-growth seasons and help increase their production. The greenhouses environment requirements are more accurate because of the inappropriate season for plants growth naturally. There mainly involves six indicators of greenhouse strawberry growth environment including air_humidity, soil_moisture, alkalinity, temperature, CO₂, and light-intensity. Relative data is collected by the wireless sensor network for a long time, and principal component analysis is applied to analyze the key factors influencing the greenhouse crop growth during the seeding period of green strawberry. The analysis shows that there are three principal components and their contribution rate are 44.57%, 29.00% and 15.83% respectively. The cumulative contribution rate can reach 89.40% and therefore these three principal components can be used to replace the original six indicators to reflect the original greenhouse strawberry growth environment information and to represent the information no longer overlap, which is very important to the reproduction of same plants and can enable them to improve the efficiency and accuracy of greenhouse cultivation management greatly.

Key words: principal component analysis; greenhouse crops; growth environment; wireless sensor network

0 引言

温室大棚可以帮助作物在不适宜生长的季节里提供生育期和增加产量,多用于低温季节喜温蔬菜、花卉、林木等植物栽培或育苗等。因此温室作物的生长环境要求相对较高。然而,目前国内大多数农户对于温室作物的生长环境控制全凭个人感觉,人感觉冷了就加温,感觉干了就浇水,感觉闷了就通风,毫无科学依据^[1]。在农业进入信息化时代后,各个科研单位对于温室作物生长环境信息的采集也越来越重视。因

此,在温室环境里,通过无线传感器网络来测量土壤湿度、空气湿度、温度、土壤盐碱度、CO₂ 浓度和光照强度等作物生长环境信息,对于调控温室温度、灌溉和施肥等作业有重要的指导价值^[2]。另外,将无线传感器网络远程监测技术应用到作物生长的不同阶段,并且对不同阶段作物的表现和环境因子进行有效分析,再将其反馈到下一轮的生产中,从而可以帮助农户实现更精准的管理,获得更优质的产品。因此,将 Zigbee 无线传感器网络应用于温室环境进行数据采集并进行有

收稿日期:2016-01-15

修回日期:2016-05-12

网络出版时间:2016-09-19

基金项目:国家自然科学基金资助项目(31360289)

作者简介:武文亮(1989-),男,硕士,研究方向为嵌入式应用和图像处理;张志斌,副教授,研究方向为机器视觉、模式识别和图像处理。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20160919.0843.064.html>

效分析对于实现温室种植的高效和精准化管理具有非常重要的研究和指导意义。

对于不同阶段所采集的生长环境指标应用主成分分析的方法,可以帮助农户了解作物在不同生长阶段的关键因素,从而获得植物在不同阶段生长的最佳条件以有效指导下一轮的生产种植。目前,利用多元统计在杨梅、草莓、小麦、水稻等作物的品种上进行的研究较多^[3-6],利用植物的光合特性和光、温度等单个环境因子关系的研究也已较为深入^[7-9],对影响植物生长的土壤元素进行主成分分析的研究也非常多^[10-12],而对于作物各个生长环境指标综合起来应用主成分分析的研究未见报道。

1 数据采集

目前,已知的能够对作物的生长影响的因素有五十多种^[13]。其中,温度、湿度、盐碱度、CO₂浓度和光照强度这五个因素对作物的生长影响较大,因此下面重点分析这五种因素对温室草莓育苗期生长的影响。

温度主要通过影响植物的光合过程、呼吸过程、蒸腾作用等代谢活动来左右作物的生长,除此之外温度对有机物的合成和运输,吸收和运输水分等代谢过程也有一定作用。同样,作物只有在合适的温度范围内才能正常生长,温度过高或过低都会抑制作物的生长速度,进而影响作物的产量。影响作物生长的湿度包括土壤湿度和空气湿度,当然两者之间相互影响,密切相关。相对湿度主要通过蒸腾作用来影响作物生长:相对湿度在合适的范围内时,作物蒸腾较旺盛,有利于

作物生长,过高或过低都会打破作物体内的水分平衡,阻碍作物生长而造成作物减产。同时,较大的湿度也会影响作物发生病虫害的几率^[14]。土壤盐碱度是土壤最重要的化学性质,它是土壤各种化学性质的综合反映,它与土壤微生物的活动、有机质的合成和分解、各种营养元素的转化与释放及有效性、土壤保持养分的能力都有关系。土壤盐碱度对土壤养分有效性有重要影响。CO₂是植物进行光合作用,形成有机物不可或缺的元素。在光照充裕的情况下,高浓度的 CO₂会促进植物的生长;而在光照不足的情况下,植物不进行光合作用,高浓度的 CO₂反而会抑制植物的生长。光照是作物进行光合作用的中心因素,是作物制造有机物的基础,是作物生存必不可少的环境因素。作物只有在合适的光照强度下才能正常进行光合作用,合理控制温室环境内的光照强度对植物生长有重要意义。

文中通过搭建 Zigbee 无线传感器网络在内蒙古自治区鄂尔多斯市准格尔旗十二连城乡五家尧农业示范园对温室草莓生长选择了土壤湿度(soil_moisture)、空气湿度(humidity)、温度(air_temperature)、盐碱度(alkalinity)、CO₂浓度(carbon_dioxide)和光照强度(light_intensity)六个指标进行了为期6个月的数据采集。图1为草莓育苗期某一个周期内的各项数据经过MATLAB编程处理后得到的动态变化曲线。可以看出,除了土壤湿度以外,其他五个指标都呈周期性变化,这主要是受到昼夜交替的影响,而酸碱度和土壤湿度整体呈下降趋势主要受到灌溉时间的影响。

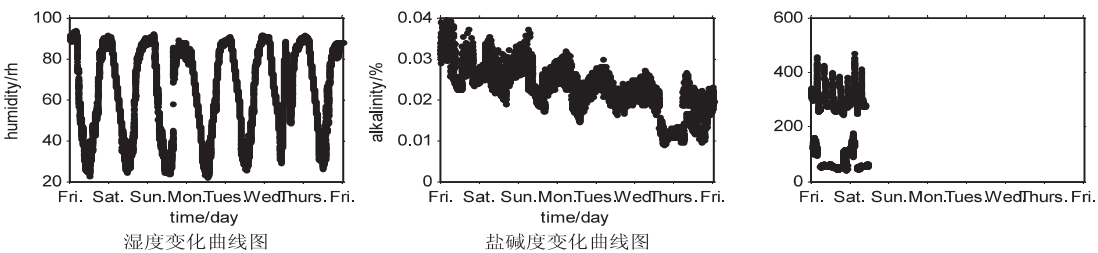


图 1 温室环境信息动态变化曲线图

2 数据处理与分析

2.1 主成分分析方法

主成分分析是把多个指标简化为少数几个综合指标的一种统计方法,它采取的是一种降维处理技

术,找出几个综合因子来代表原来众多的变量,这些综合因子尽可能地反映原始变量的信息量,而彼此之间又互不相关,从而达到简化的目的^[15]。主成分分析法的建模步骤大体上可以分为原始数据标准化,计算相关系数,计算特征值、特征向量,根据累计贡献率选取

主成分,计算主成分载荷,构建主成分表达式等过程^[16]。具体过程如下:

(1)原始指标数据的标准化。假设用 p 个变量 X_1, X_2, \cdots, X_p 来描述研究对象,那么 n 个样品就可以表示为 $X_i = (x_{i1}, x_{i2}, \cdots, x_{ip})^T (i = 1, 2, \cdots, n)$ 。数据标准化的公式为:

$$x_{ik} = \frac{x_{ik} - \bar{x}_k}{S_k} (i = 1, 2, \cdots, n; k = 1, 2, \cdots, p) \quad (1)$$

其中, $\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ik}, S_k^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2$ 。

(2)根据标准化后的数据求特征之间的相关系数矩阵 R 。相关系数的计算公式为:

$$r_{ij} = \frac{\text{COV}(X_i, X_j)}{\sqrt{D(X_i)} + \sqrt{D(X_j)}} (i = 1, 2, \cdots, p; j = 1, 2, \cdots, p) \quad (2)$$

其中, $\text{COV}(X_i, X_j) = E[(X_i - \bar{X}_i)(X_j - \bar{X}_j)]$, $D(X_i) = \text{COV}(X_i, X_i)$ 。

(3)求 R 的特征根 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ 及相应的标准正交化特征向量 u_1, u_2, \cdots, u_p 。主成分的计算公式为:

$$y_i = u_i^T x (i = 1, 2, \cdots, p) \quad (3)$$

(4)计算贡献率和累计贡献率。贡献率的计算公式为:

$$\text{CR} = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \quad (4)$$

累计贡献率的计算公式为:

$$\text{CRA} = \sum_{k=1}^p \left(\frac{\lambda_k}{\sum_{i=1}^p \lambda_i} \right) \quad (5)$$

2.2 温室草莓生长环境主成分分析

在所采集的温室育苗期草莓生长环境的六个指标中,空气湿度和土壤湿度的单位是% rh,表示空气中水蒸气压与其空气相同情况下饱和水蒸气压的百分比,采集到的空气湿度范围为 22.171 4~93.585 2,土壤湿度的范围为 20.827 7~49.954 2;土壤盐碱度指的是土壤含盐量,其单位是%,土壤含盐量超过 0.3% 时会造成作物低产或者不能生长,采集到的范围为 0.009 1~0.039 6;CO₂浓度的单位是 ppm,表示百万分比浓度,采集到的范围为 43.032 4~541.567 5;空气温度的单位是℃,表示相对标准大气压下纯净的冰水混合物的温度到水的沸点温度的百分之一,采集到的范围为 10.000 0~40.080 0;光照强度的单位是 lux,表示 1 流明的光通量均匀分布在 1 平方米面积上产生的照度,采集到的范围为 0~65 535。由于各参数的单位不同,为了使各个参数具有可比性,首行要对原始数据进行标准化处理,以消除量纲的影响^[17]。在经过对原始数据标准化处理后,对各温室育苗期草莓生长环境的各个指标进行相关性分析,相关系数矩阵如表 1 所示。可以看出,只有空气温度和空气湿度的相关性极显著($\alpha = 0.01$),其他指标间都不显著,这对于开展下一步的分析是极为有利的。因为,对于主成分分析而言,指标间越不相关,分析的结果就会越理想^[3]。

表 1 相关系数矩阵

环境指标	空气湿度	土壤盐碱度	CO ₂ 浓度	空气温度	光照强度	土壤湿度
空气湿度	1.000 0					
土壤盐碱度	0.061 1	1.000 0				
CO ₂ 浓度	0.204 5	-0.027 9	1.000 0			
空气温度	-0.913 0**	-0.156 1	-0.178 5	1.000 0		
光照强度	-0.705 6	-0.055 3	-0.121 0	0.786 1	1.000 0	
土壤湿度	-0.126 7	0.727 2	-0.041 2	-0.006 6	0.094 6	1.000 0

注:**表示 $\alpha = 0.01$ 显著水平。

计算温室育苗期草莓生长环境参数矩阵的特征值、特征向量、贡献率和累计贡献率等特征参数,结果如表 2 和表 3 所示。

由表 2 可知,主分量 1、2、3 的贡献率分别为 44.57%、29.00%、15.83%,累计贡献率达到 89.40%;根据 85% 累计贡献率的原则,故可以略去后 3 个主分量;因此可以用前 3 个主分量反映原指标的绝大部分信息,可以代替原来的 6 个单项指标对温室草莓生长环境进行研究,且使得各综合指标所代表的信息不再

重叠。

表 2 环境信息各主成分的特征值和百分率

主分量	特征值	百分率	累计百分率
1	2.674 46	0.445 7	0.445 7
2	1.739 78	0.290 0	0.735 7
3	0.949 961	0.158 3	0.894 0
4	0.314 017	0.052 3	0.946 3
5	0.256 101	0.042 7	0.989 0
6	0.065 683 9	0.010 9	1

表 3 各主成分的特征向量

主分量	空气湿度	土壤盐碱度	CO ₂ 浓度	空气温度	光照强度	土壤湿度
1	0.573 9	0.041 4	0.054 0	-0.620 1	0.525 3	0.074 1
2	0.071 3	-0.701 8	-0.021 2	0.023 5	0.106 0	-0.700 1
3	0.171 5	0.068 6	-0.979 3	0.010 5	-0.075 1	-0.032 6
4	-0.589 2	0.055 5	-0.0851	-0.752 7	-0.216 3	-0.171 1
5	-0.535 7	-0.029 5	-0.159 0	0.180 4	0.801 8	0.107 2
6	-0.044 6	-0.705 1	-0.071 0	-0.125 5	-0.132 5	0.680 1

由表 3 可以对主分量 1、2、3 列出如下方程：

$$Y_1=0.573\ 9x_1+0.041\ 4x_2+0.054\ 0x_3-0.620\ 1x_4+0.525\ 3x_5+0.074\ 1x_6$$
$$Y_2=0.071\ 3x_1-0.701\ 8x_2-0.021\ 2x_3+0.023\ 5x_4+0.106\ 0x_5-0.700\ 1x_6$$
$$Y_3=0.171\ 5x_1+0.068\ 6x_2-0.979\ 3x_3+0.010\ 5x_4-0.075\ 1x_5-0.032\ 6x_6$$

对上面三个方程可以做如下的解释：主成分 1 的特征向量中空气温湿度起着决定性的作用，因此第 1 主成分主要代表育苗期草莓温室环境的空气温湿度特点；主成分 2 的特征向量中土壤盐碱度和土壤湿度起着决定性的作用，因此第 2 主成分主要代表育苗期草莓温室环境的土壤特点，土壤是温室作物生长的关键因素之一；主成分 3 的特征向量中 CO₂ 浓度起着决定性的作用，因此第 3 主成分主要代表温室环境植物光合作用和呼吸作用，光合作用中 CO₂ 浓度直接影响作物有机物形成的多少。

3 结束语

把 Zigbee 无线传感器网络应用到温室作物生长环境的检测中对于推进农业的信息化发展有着十分重要的作用，是调控温室温度、灌溉和施肥等作业的重要趋势。利用主成分分析方法能够有效地分析影响温室作物阶段性生长的关键因素，这对于实现温室相同作物的再种植及其高效和精准化管理具有非常重要的研究和指导意义。文中通过搭建 Zigbee 无线传感器网络对温室草莓育苗期生长环境的六个指标进行了准确的数据采集，并通过主成分分析的方法对温室草莓育苗期的生长环境指标进行了有效分析，并对影响温室育苗期草莓生长因素的综合能效进行了排序，使温室育苗期草莓的生长环境能效评价更加全面。将此方法应用到温室草莓的其他生长阶段以及其他温室作物的不同生长阶段，对于建立完整的温室作物生长模型是极具指导意义的。尤其是应用于作物的再种植时，对于调控不同阶段作物生长的最优生长条件以获得最优质的果蔬来讲，具有很好的应用价值。

万方数据

参考文献：

[1] 陈 静. 物联网技术在日光温室大棚中的应用研究[J]. 科技信息, 2012(36): 78-79.

[2] 王 曦. 物联网在现代化农业中的应用[J]. 云南农业, 2012(1): 10-11.

[3] 陈守智, 吴兴恩, 钟瑞芳, 等. 大树杨梅果实数量性状的主成分分析[J]. 云南农业大学学报, 2003, 18(2): 163-166.

[4] 马鸿翔, 盛炳成, 戴子林, 等. 草莓品种的遗传距离研究[J]. 江苏农业学报, 1995, 11(3): 41-45.

[5] 郭爱民, 曹照春. 锦橙数量性状的主成分分析及分类研究[J]. 四川农业大学学报, 1997, 15(4): 499-502.

[6] 刘来福. 作物数量性状的遗传距离及其测定[J]. 遗传学报, 1979, 6(1): 30.

[7] 刘甲祥, 彭佳红, 曹晓兰. 植物生长灯光源能效的综合评价模型[J]. 农机化研究, 2012(10): 46-49.

[8] 刘 静, 王连喜, 戴小笠, 等. 枸杞叶片净光合速率与其它生理参数及环境微气象因子的关系[J]. 干旱地区农业研究, 2003, 21(2): 95-98.

[9] 冷平生, 杨晓红, 胡 悦, 等. 5 种园林树木的光合和蒸腾特性的研究[J]. 北京农学院学报, 2000, 15(4): 13-18.

[10] 郭笃发, 王秋兵. 主成分分析法对土壤养分与小麦产量关系的研究[J]. 土壤学报, 2005, 42(3): 523-527.

[11] 何尧启. 主成分分析在喀斯特土壤环境退化研究中的初步应用—以贵州麻山地区紫云县宗地乡为例[J]. 贵州师范大学学报: 自然科学版, 1999, 17(1): 12-19.

[12] 王 瑞, 陈永忠, 王玉娟, 等. 油茶林地不同间种处理土壤养分及生长量的主成分分析[J]. 中国农学通报, 2011, 27(4): 30-35.

[13] 张 慧, 乌兰巴干. 生态环境对植物生长的影响及其环境的监测[J]. 农村牧区机械化, 2011(4): 47-48.

[14] 刘民静. 基于 ZigBee 无线传感器网络的作物生长环境监控系统[D]. 济南: 济南大学, 2014.

[15] 唐启义, 冯明光. 实用统计分析及其 DPS 数据处理系统[M]. 北京: 科学出版社, 2002: 280-285.

[16] 谢恒星, 蔡焕杰, 张振华, 等. 基于主成分分析法的温室甜瓜生长方程研究[J]. 干旱地区农业研究, 2009, 27(6): 134-138.

[17] 高志海, 胡 影, 张吨明. 木蓼属 6 种植物水分生理特征的主成分分析[J]. 甘肃林业科技, 1991(3): 15-18.