

基于多元回归分析的温室茶树腾发量预测模型

孙红严¹, 马德新^{1, 2, 3*}

(1. 青岛农业大学 动漫与传媒学院, 山东 青岛 266109; 2. 山东省计算中心 国家超级计算济南中心, 山东 济南 250014;
3. 山东大学 计算机科学与技术学院, 山东 青岛 266237)

摘要: 利用青岛农业大学温室茶园内的小型智慧监测系统近年来监测的气象资料, 将 Penman-Monteith 方程和 P-M 温室修正公式作为基础方程, 通过参考作物腾发量模型进行茶树腾发量计算。以相对湿度、土壤含水量、温度和光照强度等环境因素作为自变量, 温室茶园腾发量作为回归因变量, 将 Excel 作为计算技术平台, 建立了温室茶树腾发量的多元回归预测模型, 并且进行了回归分析的显著性检验。通过试验证明此模型具有较大可行性与有效性, 操作方法简单便捷, 对于节约水资源, 促进作物提质增产具有重要的现实意义, 可为北方温室茶树精量灌溉提供参考依据。

关键词: 温室茶树; 腾发量; 多元回归; 环境因子; 预测模型; 节水灌溉

中图分类号: S571.1 文献标志码: A 文章编号: 1001-8581(2020)07-0112-05

Prediction Model of Greenhouse Tea Tree Evapotranspiration Based on Multiple Regression Analysis

SUN Hong-yan¹, MA De-xin^{1, 2, 3*}

(1. School of Animation and Communication, Qingdao Agricultural University, Qingdao 266109, China;
2. National Supercomputer Center in Jinan, Shandong Computer Science Center, Jinan 250014, China;
3. School of Computer Science and Technology, Shandong University, Qingdao 266237, China)

Abstract: The meteorological data monitored by the small-scale intelligent monitoring system in the greenhouse tea garden of Qingdao Agricultural University in recent years was used. The Penman-Monteith equation and the P-M greenhouse correction formula were used as the basic equations, and the tea evapotranspiration calculation was carried out by referring to the crop evapotranspiration model. Taking environmental factors including humidity, temperature, soil water content, temperature and light intensity as the independent variable influence factors, the greenhouse tea garden evapotranspiration as the regression dependent variable, Excel as the computing technology platform, the greenhouse tea tree evapotranspiration multiple regression prediction model was established, and a significant test of regression analysis was performed. The experiment proves that the model has great feasibility and accuracy, and the operation method is simple and convenient. It has important practical significance for saving agricultural water resources and promoting crop quality and yield, and can provide reference for the precision irrigation of greenhouse tea trees in the north.

Key words: Greenhouse tea tree evapotranspiration; Multiple regression; Environmental factors; Prediction model; Water-saving irrigation

茶树是我国重要的经济作物之一, 是一种比较古老的双子叶植物, 在农业经济生产中占有重要的地位^[1]。茶树种植区域较广, 在北方地区冬季气温相对较低且降雨量少, 不适宜茶树的生长, 因此冬季可采用在温室中种植茶树的方法。茶树在生长的过程中需水量较多, 但是由于环境因子

的复杂性, 茶树的腾发量难以确定, 这就需要对茶树进行精量灌溉, 在有效促进茶树生长的同时也能减少对农业用水的浪费。

作物需水量的精确计算和预测是制订作物灌溉制度和促进作物有效生长的前提, 也是农业生产节约用水的关键^[2]。茶树的腾发量与温度、湿

收稿日期: 2020-02-04

基金项目: 山东省重点研发计划(2019GNC106001); 淄博市重点研发计划项目(2019gy010101); 青岛市民生科技计划项目(18-6-1-112-nsh); 山东省高等学校科技计划项目(J17KA154)。

作者简介: 孙红严(1994—), 男, 硕士研究生, 研究方向: 农业信息化。* 通信作者: 马德新。

度、土壤含水量以及光照强度等多种环境因素相关,要综合考虑各种因素对茶树腾发量的影响,才能建立起准确的温室茶树腾发量预测模型^[3]。为此,特使用智慧农业设备对温室内环境因子进行精准监控,并采用由王林林等^[4]以 Penman-Monteith 作为基础方程,借鉴 P-M 温室修正式结合温室茶园实际情况所提出的温室茶树腾发量计算模型进行茶树腾发量的计算,选用 Excel 作为技术平台进行多元回归分析和显著性检验,建立了多元因素影响条件下温室茶树腾发量的预测模型。

多元回归分析预测法是进行数据统计分析与数据挖掘处理的重要方法之一,在各行各业中均有广泛的应用^[5-6]。通过多元回归分析预测模型的建立可找出各影响因子对温室茶树腾发量的作用规律,选用适合的数学模型建立相应的方程加以计算^[7]。根据此模型种植户可人为地操控某些环境因子,从而实现作物提质增产,并且可以为北方温室茶树灌溉提供重要决策。

1 试验数据及计算方法

1.1 温室茶园及数据监测站情况简介

数据监测系统是由山东省计算中心国家超级计算济南中心智能控制技术创新团队开发研制的小型智慧农业监测站,可以进行实时监测。监测站位于山东青岛城阳区东旺疃村北温室茶园,共有 2 台,该茶园处于暖温带,常年平均气温约为

$$ET_0(T) = \frac{0.408\Delta(R_n - G) + \frac{\gamma 39611(e_a - e_d)}{(T + 273) \left[\ln\left(\frac{Z - 0.64h_t}{0.13h_t}\right) \right]}}{\Delta + \gamma + \frac{15\gamma}{\left[\ln\left(\frac{Z - 0.64h_t}{0.13h_t}\right) \right]^2}} \quad (1)$$

公式(1)中: $ET_0(T)$ 为温室茶树腾发量 (mm/d); Δ 为饱和水汽压曲线斜率 (kPa/°C); R_n 为地表净辐射通量 [MJ/(m²·d)]; G 为土壤热通量 [MJ/(m²·d)]; γ 为干湿表常数 0.0646 kPa/°C; e_a 和 e_d 分别为饱和水汽压和实际水汽压 (kPa); T 为 2 m 高度处的平均温度 (°C); h_t 为茶树冠层高度 (m); Z 为该地的海拔高度 (m)。

2 多元回归分析

2.1 温室茶树腾发量的多元回归分析

多元回归预测模型是数学模型中常用的一种

12 °C,但冬天气温低于 0 °C,茶园内主要种植的茶树品种为黄山小种,每株平均高度在 0.23~0.45 m,计算时取平均值 0.34 m,其他有少量崂山茶品种茶树。监测项目主要为温室茶园灌溉用水量、土壤含水量、温度、相对湿度、光照强度。该智慧监测站是基于物联网技术的监测站,具有数据传输稳定、数据获取准确的优点,并且能够及时上传云端,用户可通过 App 以及 PC 端登陆网站进行数据获取,简单方便。数据样本选用 2017 年 11 月~2019 年 5 月小型智慧农业监测站上传至云端的数据,在此期间选取茶树春梢生长期内的数据(每年 3~5 月)用于温室茶树腾发量的计算以及多元回归预测模型数据的使用。

1.2 温室茶树腾发量的计算方法及其计算结果

对于参考作物腾发量的计算,联合国粮农组织的专家组提出了 Penman-Monteith 方程,其是国际上唯一推荐的计算方法。但由于实际大田情况与温室环境有较大的差异性,在实际应用时要进行修正^[8]。陈新明等^[9]以 Penman-Monteith 作为基础方程,运用空气动力学原理对风速以及地表净辐射规律进行了修正,提出了适用于温室参考作物腾发量的 P-M 修正式。其中青岛农业大学的王林林等^[4]通过在温室茶园中进行实验与探索提出了温室茶树腾发量计算模型 $ET_0(T)$ 。本研究将利用该公式进行计算:

预测方法,通过线性代数方程表达出未知量与多个自变量之间的规律,建立多个变量之间的数学模型方程式的统计方法^[10-11]。通过多元回归分析的原理,可建立如下的多变量的线性回归方程:

$$Y = a + \partial_1 X_1 + \partial_2 X_2 + \partial_3 X_3 + \partial_4 X_4 \quad (2)$$

公式(2)中, Y 表示预测要素; a 表示常数项; X_1 、 X_2 、 X_3 、 X_4 分别代表自变量相对湿度(%)、温度(°C)、土壤含水量(%)、光照强度(lx); ∂_1 、 ∂_2 、 ∂_3 、 ∂_4 分别代表 4 个自变量的斜率。通过 Excel 在回归分析中会计算出复相关系数、标准差、残差等一系列参数,这些参数可用于回归分析的 t 检验和 F

检验,临界值检验可检验整体参数的显著性, F 统计值检验是检验自变量之间的线性或非线性关系^[11]。

2.2 温室茶树腾发量多元回归分析预测模型的建立

表 1 2018 年 3 月 7 日~3 月 15 日温室茶树相关样本数据

日期	温室茶树腾发量/mm	温度/℃	湿度/%	光照强度/lx	土壤含水量/%
3 月 7 日	3.10	25.42	57.70	35747.61	11.98
3 月 8 日	3.43	29.89	61.23	38147.98	12.23
3 月 9 日	3.00	30.20	56.41	31133.00	11.90
3 月 10 日	3.80	31.11	56.40	31660.83	24.92
3 月 11 日	3.20	27.48	66.54	22359.61	17.63
3 月 12 日	2.80	24.03	78.42	13027.28	16.49
3 月 13 日	3.40	26.00	72.66	24148.09	17.00
3 月 14 日	3.70	28.17	52.22	32814.43	16.29
3 月 15 日	3.60	31.94	61.38	3238.56	15.16

2.2.2 通过 Excel 建立预测模型 通过 Microsoft Excel 为技术计算平台,建立温室茶树腾发量预测模型,在进行计算处理的过程中主要包含了回归统计表、方差分析表以及回归参数表 3 个部分^[12]。

2.2.2.1 回归统计表 在回归统计表中分别给出了 Multiple R(相关系数)、R Square(测定系数)和 Adjusted R Square(校正测定系数)。其中,相关系数 R 的值为 0.806777192,这就说明各自变量环境因子与因变量温室茶树腾发量的大小变化显著相关;测定系数 R^2 的值为 0.650889438,因为测定系数也被称为拟合优度,它是相关系数平方的值,这也说明这些变量之间具有很大的相关性;校正测定系数的值为 0.301778876,在回归分析中可不作参考依据;标准误差是回归方程的估计标准误差,它的值为 0.281013895;观测值代表的是样本数量,在本样本数据中共有 9 个,为 2018 年 3 月 7 日~2018 年 3 月 15 日温室内各环境因子数据和通过计算得出的茶树腾发量数据。

2.2.2.2 方差分析表 方差分析表主要由 4 个参数部分组成,分别是自由度、误差平方和、均方差、 F 值、 P 值,具体值如表 3 所示。在表 3 中第一行 df 代表的是自由度的大小, SS 代表的是误差平方和(变差)。第 1 行中 $df=4$,表示回归自由度 dfr ,即 4 个环境因子变量; SS 是回归平方,它表示温室茶树腾发量的预测值对其平均值的总偏差。第 2 行中 df 表示 dfe (残差自由度); SS 表示 SSe (标准误差),它的值可说明各环境因子对温室茶树腾发量预测值的总偏差,如果这个数值越大,则说明其预测误差就越大,在此 $SSe=0.315875$,说明其误差

2.2.1 建立样本数据 将公式(1)在 Matlab 中编写成计算程序,并将监测的数据代入进行计算,选择 2018 年 3 月 7 日~2018 年 3 月 15 日之间的数据,共 9 d 进行逐日计算。计算结果及其相关计算数据如表 1 所示。

较小,预测度较高。第 3 行中 df 表示 dft (总自由度),等于变量数目减 1; SS 表示 SSt (总变量差)。第 4 列 F 所对应统计量的数值可用于评价回归方程的显著性。第 5 列 Significance F 代表着弃真率,它的意思就是预测模型为假的概率。

2.2.2.3 预测方程 根据回归分析原理结合回归参数表可知其回归预测方程:

$$Y = 2.4489 + 0.04X_1 - 0.01X_2 + 0.01X_3 + 0.04X_4 \quad (3)$$

从公式(3)可以得出温度和土壤含水量与温室茶树腾发量呈相关的关系,湿度和光照强度与温室茶树腾发量呈负相关的关系。在实际的作物生长过程中可根据此方程来影响或者预测温室茶树腾发量的大小,从而对相关影响因子进行调控,同时也可以为精准灌溉决策提供参考依据,进而有效地促进茶树的生长。

表 2 回归统计结果

指标	参数
Multiple R	0.806777192
R Square	0.650889438
Adjusted R Square	0.301778876
标准误差	0.281013895
观测值	9

2.3 多元回归预测模型的显著性检验

回归系数显著性检验是对 t 临界值和 F 统计值进行检验。 t 临界值在回归分析中主要是用来检验单个参数的显著性,通过 t 检验明确被检验的参数是否显著有效^[13-14]。在此多元回归模型的预测中通过 Excel 进行数据分析,得到了回归参数表(表 4)。在表中第 4 列 t Stat 的值代表临界 t 值,

用于对温室茶树腾发量多元回归预测模型的显著性检验, t 值是回归系数与其标准误差的比值。结合表 4 且经查表可知, 在本模型中土壤含水量、温度、相对湿度和光照强度 4 个影响因子的观察值分别为: 0.74、-0.174、-0.022、1.52, 当前自由度为 4, 可以明显得出 4 个影响因子对温室茶树腾发量预测结果影响显著, 所以可用来预测温室茶树腾发量。

F 统计值是检验自变量与因变量总体之间的线性关系, 通过 F 统计值的检验能说明整体参数

中至少有一个是显著的, 但不一定所有的都是显著性相关^[7]。在表 3 中 F 列对应的为 F 统计值的数值, 这个值用于评价预测方程的显著性, 在表中 $F = 1.864$, 因此预测方程通过了检验, 显著性较好。

表 3 方差分析结果

指标	df	SS	MS	F	Significance F
回归分析	4	0.588925	0.147231	1.864422	0.280536631
残差	4	0.315875	0.078969		
总计	8	0.9048			

表 4 回归参数表

指标	Coefficients	标准误差	t Stat	P -value	Lower 95%	Upper 95%	下限 95.0%	上限 95.0%
Intercept	2.45	2.80	0.87	0.43	-5.34	10.23	-5.34	10.23
X Variable 1	0.04	0.05	0.75	0.50	-0.11	0.19	-0.11	0.19
X Variable 2	-0.01	0.02	-0.70	0.52	-0.07	0.04	-0.07	0.04
X Variable 3	0.01	0.02	-0.02	0.98	-3.30	3.20	-3.32	0.00
X Variable 4	0.04	0.03	1.57	0.19	-0.03	0.11	-0.03	0.11

3 讨论

3.1 预测模型的应用与拓展

温室茶园腾发量多元回归预测模型主要采用土壤含水量、温度、相对湿度和光照强度 4 种影响因子作为输入样本的信息源, 通过 Excel 进行回归建模分析, 建立多元回归预测方程。对于北方温室茶园来说, 该模型对于安装有相关气象因子监测设备的茶园代表性以及利用价值较高, 对于没有相关气象监测设备的温室茶园, 也可以通过了解相邻安装有检测设备的气象条件或者通过茶园各气象因子对茶树腾发量影响的大小来进行相关灌溉决策, 所以此模型应用对于北方大部分温室茶园有较大的参考价值, 具有区域预测性^[15]。

在此预测模型中选定了土壤含水量、温度、相对湿度和光照强度 4 种气象因子对温室茶树腾发量进行回归分析, 结果表明各影响因子对茶树腾发量影响程度不同, 可通过调整相关因子大小来改变温室茶树腾发量, 并且可准确预测温室茶树腾发量^[16]。该模型预测主要基于 Excel 进行多元回归分析, 在操作上更加简单方便, 根据回归分析得出的回归预测方程进行了显著性检验。在温室茶树实际的生产中可准确预测茶树腾发量, 同时也可通过调节相关因子来进行节水灌溉, 降低农业用水资源的浪费, 因此该模型实用性较高, 可广泛推广。

3.2 预测模型的优缺点

3.2.1 预测模型的优点

首先, 该预测模型在操

作上简单方便, 对通过现代信息技术设备采集的相关气象因子进行建模, 在建模时选用 Excel 作为计算平台, 不需要进行专门模型软件的研究开发, 极大地降低了成本。其次, 该预测模型采用多种气象因子进行多元回归分析, 应用性较为广泛和全面, 创建的预测模型也方便改造。最重要的是该预测模型具有相应的回归系数计算, 并进行了显著性检验, 相关性较高, 具有较高的推广应用价值。

3.2.2 预测模型的缺点 作物腾发量的大小不仅仅受环境因子的影响, 同时还受作物本身的生长生理过程、农业栽培技术以及作物品种等多种因素的共同影响。此模型仅仅通过气象因子的影响对其预测, 具有局限性^[17-19]。且当影响因子为极端数据时, 预测结果的误差也会增大。

3.3 结语

通过探索实践结合相关信息技术对温室茶园腾发量建立了相关预测模型, 并进行相关性检验, 进一步确定此预测模型具有一定的可靠性, 预测结果是可信的。但是由于温室内微气候环境相对比较复杂, 在计算时对一些因素进行假定, 计算结果相对比较理想, 因此该预测模型在计算结果上存在一定的误差^[20, 21]。对于影响茶树腾发量的因素测定还不够全面, 所以在今后会继续加强对作物腾发量预测的研究, 以求能为节约农业资源用水和提高作物产量产生更大的经济利益做出更大的贡献。

参考文献:

- [1] Suthwell I, Robert L. Tea tree [M]. Taylor and Francis, 2004.
- [2] 常晓敏,高占义,王少丽,等.基于气温预报和 HS 公式的不同生育期参考作物腾发量预报[J].节水灌溉, 2016(8): 169-174.
- [3] 刘梦,罗玉峰,汪文超,等.基于天气预报的漳河灌区参考作物腾发量预报方法比较[J].农业工程学报, 2017, 33(19): 156-162.
- [4] 王林林,马文杰,马德新.基于 Penman-Monteith 方程的温室茶树蒸腾蒸发模型研究[J].节水灌溉, 2017(8): 30-33, 43.
- [5] Seghouane A K. New AIC corrected variants for multivariate linear regression model selection [J]. IEEE Transactions on Aerospace and Electronic Systems, 2011, 47(2): 1154-1165.
- [6] 代亮,许宏科,陈婷,等.基于 MapReduce 的多元线性回归预测模型[J].计算机应用, 2014, 34(7): 1862-1866.
- [7] 邹文安,姜波,张薇.基于 Excel 平台土壤含水量多元回归预测模型[J].水文, 2015, 35(2): 65-69, 89.
- [8] 陈石磊,霍再林.黑河流域近 53 年气候变化对参考作物腾发量影响研究[J].中国农村水利水电, 2016(5): 81-87, 91.
- [9] 陈新明,蔡焕杰,李红星,等.温室内作物腾发量计算与验证[J].水科学进展, 2007(6): 812-815.
- [10] 章培军,杨友社.基于多元回归和灰色关联的粮食种植面积影响因素分析[J].科技促进发展, 2019, 15(3): 326-330.
- [11] Rupa M. Regression analysis Microsoft® Excel® [M]. Software Quality Professional, 2016.
- [12] 游士兵,严研.逐步回归分析法及其应用[J].统计与决策, 2017(14): 31-35.
- [13] 杨利雄,张春丽,李庆男.平稳的平滑转移自回归过程之间的虚假回归问题研究[J].统计与决策, 2017(6): 15-19.
- [14] 范廷宾,王明.火烧云铅锌矿矿石小块体重多元线性回归模型建立及检验[J].地质学刊, 2017, 41(3): 409-414.
- [15] 邹文安,马俊英,孙晓梅.基于 Excel+VBA 技术条件的土壤含水量预测模型[J].中国防汛抗旱, 2013, 23(2): 49-52.
- [16] 罗童元,罗玉峰,吕辛未,等.基于 Web 的江苏省逐日参考作物腾发量预报系统[J].中国农村水利水电, 2018(3): 77-80.
- [17] 徐颖,张皓杰,崔宁博,等.基于不同 ELM 的西北旱区参考作物蒸散量模拟模型[J].中国农村水利水电, 2019(1): 6-12.
- [18] 闫浩芳,毋海梅,张川,等.基于修正双作物系数模型估算温室黄瓜不同季节腾发量[J].农业工程学报, 2018, 34(15): 117-125.
- [19] 符娜,宋孝玉,夏露,等.云南省不同生态水文分区参考作物蒸散量算法适用性评价[J].农业机械学报, 2017, 48(5): 208-217.
- [20] 陈新明,蔡焕杰,李红星,等.温室大棚内作物蒸发蒸腾量计算[J].应用生态学报, 2007(2): 317-321.
- [21] 潘永安,范兴科.气象资料缺测时 Penman-Monteith 温室修正式的应用[J].西北农林科技大学学报: 自然科学版, 2015, 43(1): 117-124.

(责任编辑:许晶晶)