

基于主成分分析的温室内水面蒸发量估算模型构建及验证

王科^{1,2}, 李银坤¹, 郑文刚¹, 刘美英^{2*}, 武佳乐^{1,2}, 纪玉如³, 陈菲¹, 侯升林⁴

(1.北京市农林科学院 智能装备技术研究中心, 北京 100097; 2.内蒙古农业大学
草原与资源环境学院 内蒙古自治区土壤质量与养分资源重点实验室, 呼和浩特 010018;
3.中国农业大学 园艺学院, 北京 100094; 4.河北省农林科学院, 石家庄 050051)

摘要:【目的】估算温室无风环境下的水面蒸发量(E_p)。【方法】基于温室内 2020 年与 2022 年 3—7 月的实测水面蒸发量(E_p)与气象数据,采用主成分分析法对 E_p 的影响因素进行分析,利用提取出的主成分与 E_p 构建多元回归模型,并对估算结果进行验证。【结果】试验期间 E_p 随试验时间的延长呈上升趋势,2020 年 3 月与 2022 年 3 月的 E_p 日平均值分别为 1.84 mm 与 1.94 mm,6 月分别增加至 3.77 mm 与 5.15 mm;辐射、湿度对温室无风环境下水面蒸发量的影响占主要地位,其中光合有效辐射与水面蒸发量的相关性最高,相关系数为 0.852 ($P<0.01$),其次为太阳辐射与湿度,相关系数分别为 0.811 与 -0.770 ($P<0.01$)。第一主成分的太阳辐射、光合有效辐射以及湿度对水面蒸发量影响较大,特征值为 4.44,其中太阳辐射对水面蒸发量影响最明显,得分系数最高,为 0.328;湿度与光合有效辐射次之,得分系数分别为 0.311 与 -0.321。基于主成分分析结果建立了水面蒸发量估算模型,水面蒸发量估算值与实测值显著正相关($P<0.01$),方程相关系数 R^2 为 0.908, MBE 为 0.10, RMSE 为 0.48 mm/d,一致性指数较高($d=0.94$)。【结论】在温室无风环境下太阳辐射、光合有效辐射与湿度对水面蒸发量影响较高。

关键词: 温室; 水面蒸发量; 气象要素; 多重共线性; 主成分分析

中图分类号: P332.2

文献标志码: A

doi: 10.13522/j.cnki.ggps.2022543

OSID: 

王科, 李银坤, 郑文刚, 等. 基于主成分分析的温室内水面蒸发量估算模型构建及验证[J]. 灌溉排水学报, 2023, 42(5): 60-66.

WANG Ke, LI Yinkun, ZHENG Wengang, et al. Construction and Validation of a Model for Estimating Surface Water Evaporation in Greenhouse Based on Principal Component Analysis[J]. Journal of Irrigation and Drainage, 2023, 42(5): 60-66.

0 引言

【研究意义】我国设施蔬菜种植面积已达 410 hm^2 , 总产值占蔬菜产业的 50%^[1]。灌溉水是温室蔬菜水分需求的主要来源,灌溉决策不合理是我国目前盲目灌溉现象依旧普遍存在的主要原因^[2],灌溉量不足或过多不仅影响了蔬菜正常生产、浪费水资源,还导致了温室蔬菜经济效益下降等问题^[3]。水面蒸发量是指单位时间内从单位(水)表面面积蒸发的水量,常表示为单位时间内从全部(水)面积上所蒸发的液态水的深度^[4],精确获取温室内水面蒸发量信息,对构建科学合理灌溉制度,促进温室蔬菜高产优质生产至关重要^[5]。20 cm 标准蒸发皿是目前直接测量不同生产条件下水面蒸发量(E_p)的主要方式,但人工测

量的误差大、效率低,而采用自动称质量法的条件要求高、测量成本高^[6]。模型估算法也是获取 E_p 的主要途径,该方法需要根据区域气候特点进行 E_p 模型构建与验证,具有应用范围广、成本低等优点^[7]。建立适宜的 E_p 估算模型,对制定科学合理的灌溉制度、发展节水农业等具有十分重要的意义^[2]。【研究进展】目前, E_p 估算模型主要包括经验公式模型与机理公式模型。机理公式模型结构通常较为复杂,不利于推广应用,有研究^[9]提出了气-水二维湍流流动理论模型,模型考虑了风在经过水体而产生的环流,以及空气流速、水温、气温等物理因素对水面蒸发的影响。经验公式模型包括 2 类,一类是考虑太阳、水与空气之间质量和能量的交换过程来估算 E_p ,例如 Penman 法^[9]与波文比法^[8];另一类是以风速、水汽压差等其他气象因子建立的经验公式,例如 Blaney-Criddle 公式^[7]。闵骞^[8]对彭曼模型中的水面辐射平衡值与水面附近的空气干燥力 2 个参数进行修正,针对露天环境建立了以风速、湿度和饱和水汽压为主要参数的适用于全国的 E_p 模型。也有研究^[10]以 Blaney-Criddle 公式为基础,建立了精度在月、旬、日的 E_p 经验模型,且 E_p

收稿日期: 2022-09-30

基金项目: 河北省重点研发计划项目(21327005D); 国家重点研发计划项目政府间国际科技创新合作专项(2019YFE0125100); 河北省现代农业产业技术体系项目(HBCT2021200201, HBCT2021200202)

作者简介: 王科(1999-),男,硕士研究生,主要从事作物水肥高效利用研究。E-mail: wke1020@163.com

通信作者: 刘美英(1974-),女,内蒙古清水河人。副教授,硕士生导师,主要从事土壤肥力与植物营养方面的教学科研工作。

E-mail: liumeiyingimau@163.com

模型精度较原始模型有所提高；濮培民^[11]通过分析水面蒸发过程中的水汽界面上的质量和能量传递过程，建立了基于温度、水汽压差、风速、相对湿度等因子的 E_p 估算模型。Jafari 等^[12]利用主成分回归法，建立了基于风速、湿度、水汽压差等气象因子的 E_p 估算模型，且精度明显高于回归模型；陆云燕等^[13]利用主成分分析法，从诸多气象因子之间提取出 3 个主成分作为变量，与其建立的 PCA-BP 神经网络算法模型精度较高。【切入点】精确获取水面蒸发量是温室内水资源高效利用与农业灌溉制度确定的关键前提^[14]。目前， E_p 估算模型的研究主要针对露天环境，并将风速作为主要模型参数，但对温室无风环境下的 E_p 模型估算研究较少。【拟解决的关键问题】为此，本研究利用标准蒸发皿获取实测 E_p ，分析其在温室内无风环境下的主要变化趋势，利用主成分分析法探究影响 E_p 的主要因素，分析提取出的主成分与 E_p 的关系，进而建立基于主成分的温室内水面蒸发量估算模型并验证，为温室节水灌溉和绿色生产提供科学理论支持。

1 试验与方法

1.1 试验区概况

试验于 2020 年与 2022 年的 3—7 月在北京市农林科学院三连跨多功能连栋温室开展。试验温室南北走向，长 39 m、宽 11 m，试验布置于三连跨多功能连栋温室的中跨。试验温室位于东经 116.29°，北纬 39.94°，海拔 56 m，多年平均气温为 11.1℃，年均降水量为 500~600 mm。该区属典型的温带大陆性季风气候。温室内透光性良好，且温室内气流较弱，风速近似为 0。

试验期间种植作物为茄子，2020 年 3 月 1 日定植，2020 年 7 月 15 日收获，2022 年 3 月 3 日定植，2022 年 7 月 15 日收获。依据直径 20 cm 蒸发皿测定的水面蒸发量的 80% 进行灌水^[15]，灌水周期为 7~10 d。供试茄子品种为“京茄黑宝”，采用畦栽模式，栽培畦宽 0.75 m，高 0.1 m，每畦种植 2 行，株距为 0.45 m，行距为 0.5 m，试验区基施有机肥 30 000 kg/hm²，试验期间追施水溶性肥料($\omega(\text{N}) : \omega(\text{P}_2\text{O}_5) : \omega(\text{K}_2\text{O}) = 3 : 1 : 6$) 3 次，每次纯氮施量为 75 kg/hm²。

1.2 数据处理与分析

采用 Excel 2016 软件整理分析数据，利用 SPSS 20.0 对气象数据进行主成分分析，再将主成分与水面蒸发量进行多元回归分析。同时采用均方根误差 (RMSE)、平均偏差 (MBE) 与一致性指数 (d) 等指标对估算值和实测值的符合度进行统计学分析。误差 RMSE 越小，MBE 越接近 0，一致性指标 d 越大，则模拟效果越好。

1.3 数据的选取与监测

利用 2020 年 3 月 1 日—7 月 15 日实测逐日水面蒸发量与气象数据探讨温室内水面蒸发量变化规律及影响因素，并建立模型；使用 2022 年 3 月 2 日—7 月 15 日的对模型进行准确性检验。

1.3.1 水面蒸发量

利用直径为 20 cm 的称质量式标准蒸发皿测定水面蒸发量 (E_p)，蒸发皿位于温室中央，精度为 0.1 mm，每 10 min 测量 1 次数据。

1.3.2 气象数据

气象数据来源于温室内小型气象站 (型号 AG1000，美国生产)，气象站位于温室中央，监测要素包括平均气温 (T)、日最高/低气温 ($T_{\max}/T_{\min}$)、太阳辐射 (R_a)、光合有效辐射 (PAR) 和相对湿度 (RH)，数据采集间隔为 10 min。通过 FAO-56 推荐方法计算饱和水汽压差 (VPD)^[16]。

1.4 研究方法

1.4.1 多重共线性分析

“多重共线性”最早于 1934 年提出^[17]，是指线性回归模型中的 2 个或多个变量之间，存在高度相关关系。共线性的存在会使回归模型剔除有用的自变量，导致模型失真或难以准确估计；或会使回归模型的系数符号与实际情况相反，使得模型中的变量无法得到合理的解释。

诊断多重共线性时，最常用的方法为方差膨胀因子法 (VIF) 法^[18]，在多元回归中，设有 k 个自变量 $X_1, X_2, \dots, X_k$ ，则：

$$VIF = \frac{1}{1 - R_i^2} \quad (1)$$

R_i^2 为 X_i 自变量与其他 $k-1$ 个自变量间的复相关系数。 $VIF=1$ ，则表明自变量 X_i 与其他 $k-1$ 个自变量之间没有共线性关系，一般情况下认为^[19]，当 $VIF \geq 10$ 时自变量之间存在较为严重的共线性，且 VIF 值越大，共线性越强。

还有研究^[17]对 VIF 的倒数，即容忍值 (TOL) 法对回归模型的共线性进行进一步探讨：

$$TOL = 1 - R_i^2 \quad (2)$$

TOL 的取值在 0~1 之间，一般认为该值越接近 0，说明自变量之间的共线性越强。

1.4.2 主成分回归

主成分回归是 Massy^[20]于 1965 年提出的一种解决多重共线性问题的新方法。基本思路是将原来多个自变量通过线性组合，转化为少数几个主成分，然后再将主成分与因变量建立回归方程，依据主成分与自变量的关系，求得基于原自变量的回归模型。

1.4.3 主成分分析

主成分分析^[19]是通过降维将多个变量转化为几个互不相关的变量,其数学模型为:

对于 k 个观测变量 $X_1, \dots, X_k$, n 个观测的样本数矩阵:

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1k} \\ X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix} = (X_1, X_2, \dots, X_k). \quad (3)$$

将 k 个预测变量综合成为 p 个新变量(综合变量),即:

$$\begin{cases} F_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1k}X_k \\ F_2 = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2k}X_k \\ \vdots \\ F_p = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pk}X_k \end{cases} \quad (4)$$

模型满足: ① F_i, F_j 互不相关 ($i \neq j, i, j = 1, 2, \dots, p$); ② F_1 的方差 $> F_2$ 的方差 $> F_3$ 的方差, 以此类推; ③ $a_{j1}^2 + a_{j2}^2 + \cdots + a_{jk}^2 = 1$ ($j = 1, 2, \dots, k$)。模型中称 F_1 为第一主成分, F_2 为第二主成分, 以此类推。 a_{ij} 称为主成分系数。

选取主成分的个数取决主成分的累计方差贡献率, 通常选取累计贡献率大于 85% 以上等主成分, 这

些主成分更好地保留了原来因子的信息, 减少了因子数量, 从而起到筛选因子的作用。

1.4.4 多元回归分析

多元线性回归^[19]模型是研究一个变量 Y 与 2 个或 2 个以上的变量 $X_1, X_2, \dots, X_n$ 之间是否存在确定的函数关系, 其表达式一般为:

$$Y = a_0 + a_1X_1 + \cdots + a_nX_n + \mu, \quad (5)$$

式中: a_0 为回归常数; $a_1, \dots, a_n$ 为回归系数; Y 为因变量; X 为自变量; μ 为误差项。

2 结果与分析

2.1 试验期间水面蒸发量变化规律

图 1 为 2020 年与 2022 年试验期间水面蒸发量 (E_p) 的逐日变化规律。温室内 E_p 年内变化规律一致, 2020 年和 2022 年温室内 E_p 的变化范围分别为 0.19~6.85 mm 和 0.37~7.01 mm, E_p 均值分别为 3.32 mm 和 3.27 mm。随着试验时间的延长 E_p 呈波动上升趋势, 2020 年和 2022 年 3 月的 E_p 均值分别为 1.84 mm 和 1.94 mm, 6 月 E_p 分别增加至 3.77 mm 和 5.15 mm。

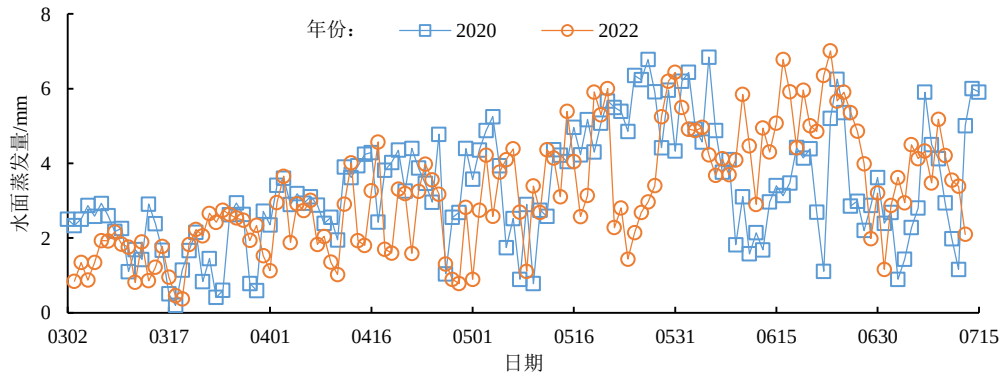


图 1 温室内水面蒸发量动态

Fig.1 Dynamic of pan evaporation in greenhouse

2.2 水面蒸发量影响因子筛选及影响因子间共线性诊断

表 1 为水面蒸发量 (E_p) 与气象因子相关分析结果。由表 1 可知, 各气象因子均与水面蒸发量极

显著相关 ($P < 0.01$), 其中, RH 与其负相关, 其他环境因子均与 E_p 正相关。 E_p 与 PAR 的相关性最高, 相关系数为 0.852, 与 T_{min} 相关性最低, 相关系数为 0.428。

表 1 温室蒸发皿水面蒸发量与气象因子的相关分析

Table 1 Correlation between pan evaporation and various environmental factors in greenhouse

指标	E_p	T	RH	R_a	PAR	T_{min}	T_{max}	VPD
E_p	1							
T	0.753**	1						
RH	-0.770**	-0.317**	1					
R_a	0.811**	0.437**	-0.773**	1				
PAR	0.852**	0.497**	-0.760**	0.968**	1			
T_{min}	0.428**	0.868**	0.004	0.042	0.102	1		
T_{max}	0.705**	0.812**	-0.474**	0.581**	0.615**	0.562**	1	
VPD	0.751**	0.713**	-0.614**	0.581**	0.616**	0.436**	0.928**	1

注 **表示表示在 0.01 水平上显著相关。

由表 1 还可以看出, 各气象因子之间存在明显相关性。其中, PAR 与 R_a 相关性最高, 相关系数为 0.968;

而 $T_{\min}$ 与 RH 、 R_a 、 PAR 相关性均不显著。表 2 为采用方差膨胀因子法与容忍值法对各气象要素的共线性定量分析结果，可以看出，除 RH 外，各气象因子的容忍值均小于 0.1，膨胀系数大于 10，说明除 RH 以外的气象因子之间均存在较为严重的共线性问题，需要对数据进行主成分分析以解决共线性问题。

表 2 共线性诊断

Table 2 The diagnosis of collinearity

指标	T	RH	R_a	PAR	$T_{\min}$	$T_{\max}$	VPD
TOL	0.05	0.25	0.05	0.05	0.09	0.07	0.08
VIF	18.75	3.99	18.63	18.56	11.21	14.09	12.41

2.3 水面蒸发量主成分多元回归模型建立

对水面蒸发量影响的气象因子进行主成分分析表明，提取出 2 个主成分 (F_1 和 F_2)， F_1 综合了原始变量的 63.4% 的变异信息，明显高于第二主成分 F_2 的 24.5%，二者合计贡献率大于 85%。这表明 2 个主成分在包含全部变量的同时，对水面蒸发量 (E_p) 的影响最为明显。依据表 3 各主成分的得分系数可知， F_1 中 R_a 、 PAR 和 RH 的载荷较大，得分系数分别为 0.328、0.311 和 -0.321； F_2 中 T 与 $T_{\min}$ 的载荷较大，得分系数分别为 0.353 和 0.431。故 E_p 主要受 R_a 、 PAR 、 T 、 $T_{\min}$ 的正向影响，受 RH 的负向影响。

依据表 3 提取出的 2 个互相独立的主成分 F_1 与 F_2 ，与实测 E_p 进行多元回归分析，详见表 4，建立的主成分回归模型参数均通过 $P=0.05$ 水平的显著性检验；模型决定系数 R^2 为 0.854，表明 E_p 变化的 85.4% 是由 F_1 和 F_2 所代表的影响因子所引起的。

表 3 主成分系数和贡献率

Table 3 Principal component score coefficient and contribution rates

自变量	主成分	
	F_1	F_2
T	-0.060	0.353
RH	-0.321	0.124
R_a	0.328	-0.102
PAR	0.311	-0.071
$T_{\min}$	-0.220	0.431
$T_{\max}$	0.062	0.241
VPD	0.112	0.179
特征值	4.44	1.71
贡献率/%	63.4	24.5

表 4 水面蒸发量主成分多元回归模型

Table 4 The pan evaporation and main components regression model

因变量	主成分回归模型	决定系数
E_p	$E_p=1.195F_1+0.804F_2+3.320$	0.854**

注 **表示表示在 0.01 水平上显著相关。

2.4 水面蒸发量估算值与实测值比较及相关性分析

将验证数据进行标准正态变换后，代入表 4 主成分回归模型中，水面蒸发量 (E_p) 估算值与实测值变化趋势如图 2 所示，估算值与实测值全生育期变化规律一致，峰值较为接近，分别为 6.88 mm 和 7.01 mm。整个生育期的 E_p 估算值为 441.16 mm，与实测 E_p 相比，高估了 14.82 mm。 E_p 的实测值与估算值相关性如图 3 所示，模型估算值与实测值极显著正相关，方程决定系数 R^2 为 0.908 ($P<0.01$)，回归趋势线与 1:1 线较为接近，可见模型估算值与实测值有较高的线性相关。

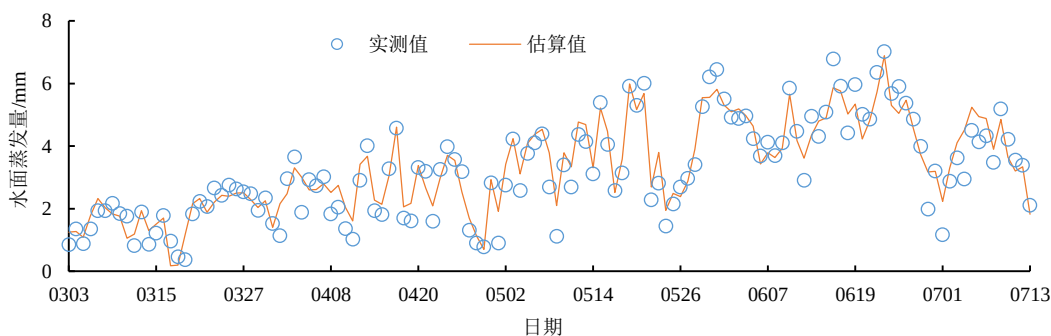


图 2 水面蒸发量估算值与实测值比较

Fig.2 Changes of measured values of pan evaporation and estimated values of the model

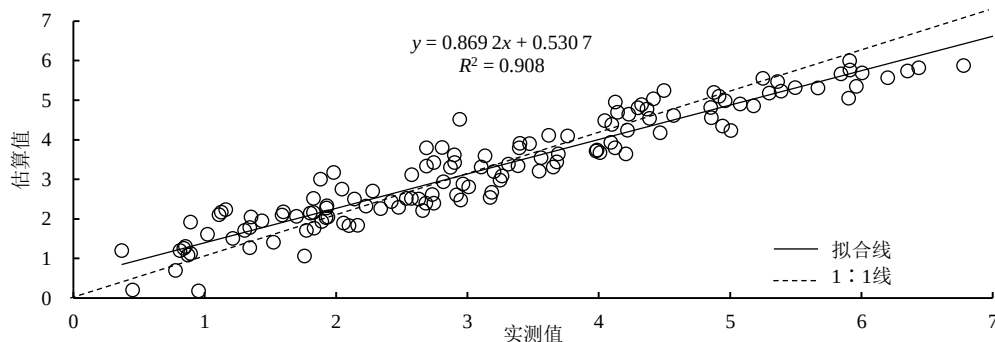


图 3 水面蒸发量估算值与实测值相关性比较

Fig.3 Comparison between measured values of pan evaporation and estimated values of the model

2.5 水面蒸发量估算值与实测值的统计分析

为进一步评价水面蒸发量 (E_p) 估算模型的估算精度, 对模型的估算值与实测值进行统计分析,

详见表 5。基于模型估算的日均 E_p 为 3.32 mm/d, 与实测 E_p 相对误差为 3.51%, $RMSE$ 为 0.48 mm/d, MBE 为 0.10, 一致性指数为 0.94, 综合来看模型误差较低。

表 5 模型估算值与蒸发皿实测水面蒸发量比较的统计特征值

Table 5 Statistical comparison between measured values of pan evaporation and estimated values of the model

日均估算水面蒸发量/(mm d ⁻¹)	日均实测水面蒸发量/(mm d ⁻¹)	比值	相对误差/%	$RMSE$ /(mm d ⁻¹)	MBE	d
3.32	3.21	1.03	3.51	0.48	0.10	0.94

注 比值为日均估算水面蒸发量与日均实测水面蒸发量之比, d 为一致性指数。

3 讨论

本试验中, 水面蒸发量 (E_p) 整体呈波动上升趋势, 这与环境温度不断上升, 光照逐渐增强等有关^[8]。陆云燕等^[13]研究表明, 露天环境下的 E_p 从 4—7 月呈逐渐升高趋势。水面蒸发主要因为水面水汽压大于上空水汽压, 而使逸出水面的水分子多于大气返回的水分子量, 其蒸发强度与气象因子相关性较高, 且受各气象因子影响程度不同^[21]。本研究表明, E_p 与 R_a 的相关系数为 0.811, 与 PAR 的相关系数为 0.852 ($P < 0.01$), 明显高于 E_p 与 VPD 的相关系数 0.751, E_p 与 RH 的相关系数 -0.770 ($P < 0.01$)。这与李银坤等^[2]在温室环境中的研究结果类似。但露天环境条件下, RH 与 VPD 是影响 E_p 的主要环境因素^[8]。这可能是露天环境中风速较强而造成的研究结果差异。孙凤贤等^[22]、刘昌宇^[9]对静止水面下 E_p 的主要影响因素的研究得出, 在风速较低时, 太阳辐射强度越高, E_p 越大, 且在无风环境下, 光照强度对 E_p 的影响可以提高 21 倍。本研究中主成分分析的结果也进一步说明了这个问题, 对 E_p 影响最大第一主成分的主要影响因素为 R_a 、 PAR 和 RH , 主成分贡献率可达 63.4%; 而第二主成分的主要影响因素为 T 和 T_{min} , 主成分贡献率仅为 24.5%。可见, 温室环境条件下 R_a 与 PAR 对 E_p 的影响, 明显强于 VPD 与 RH 对 E_p 的影响。

多元回归模型中, 多重共线性问题的来源主要是自变量自身的性质, 即多个自变量之间高度相关^[17]。本研究中, 除 RH 外, 各气象因子之间容忍值皆小于 0.1, 膨胀系数大于 10, 这说明除 RH 以外的气象因子之间存在较为严重的共线性^[19], 而共线性强的各气象因子之间通常存在较为紧密的联系。如表 1 所示, R_a 与 RH 显著负相关 ($P < 0.01$), 相关系数可达 -0.773, 而其他气象因子显著正相关 ($P < 0.01$), 相关系数在 0.4 以上。这主要是因为环境中热量的主要来源是太阳辐射, 随着光照增强, 温度升高, 饱和水汽压差也随之升高, 空气随着温度的升高而膨胀, 因此单位体积内的水分子减少, 湿度随之降低^[21]。许莹莹等^[19]研究发现, 平均气温、地温、绝对湿度和水汽压差间彼此密切相关, 相关系数在 0.75~0.97 之间, 共线性

分析表明, 地温、绝对湿度和平均气温的容忍度均小于 0.1, 方差膨胀因子均大于 10, 说明各影响因子之间存在较为严重的共线性。可见各自变量之间联系较为紧密时, 会在建立回归方程时产生共线性问题。本研究利用主成分回归法建立的 E_p 估算模型的相对误差为 3.51%, $RMSE$ 值为 0.48 mm/d, 一致性指数为 0.94。主成分分析可以很好地解决共线性问题, 具有降低模型误差的效果^[17]。Jafari^[12]研究发现, 使用主成分回归模型来预测 E_p 的模型 R^2 为 0.82, $RMSE$ 为 0.52 mm/d, 而多元回归模型的 R^2 为 0.79, $RMSE$ 为 2.28 mm/d。还有学者^[17]利用了主成分分析与相关分析相结合的方式, 重新提取出 7 个各自独立的变量, 消除了传统最小二乘法建立的多元回归模型中存在的共线性, 并利用多元回归分析建立了模型。陆云燕等^[13]使用机器学习算法的建模方式, 将提取出的主成分, 通过 BP 神经网络建立 E_p 的估算模型, 得到了较好的结果。本试验利用不同年份数据, 建立了基于主成分分析的 E_p 估算模型, 为估算温室环境下的 E_p 提供了可靠的依据。

4 结论

1) 温室 2020 年与 2022 年春茬茄子种植期间水面蒸发量 (E_p) 的变化范围分别为 0.19~6.85 mm 和 0.37~7.01 mm, 且随试验时间的延长, E_p 呈上升趋势。

2) 温室无风环境内, 辐射因素与湿度对 E_p 的影响占主导地位, 其中 PAR 与 E_p 的相关性最高, 相关系数为 0.852 ($P < 0.01$), PAR 与 R_a 、 RH 的相关系数分别为 0.811 和 -0.770 ($P < 0.01$); 在主成分分析中得出, F_1 对 E_p 影响最明显, 特征值最大, 为 4.44, 其主要代表 R_a 、 PAR 和 RH 对 E_p 的影响, 各因子得分系数分别为 0.328、0.311 和 -0.321。

3) 建立了基于主成分分析的多元回归 E_p 估算模型, 该模型消除了各变量之间的共线性, 其估算值与实测值显著正相关 ($P < 0.01$), R^2 为 0.908, MBE 为 0.10, $RMSE$ 为 0.48 mm/d, 一致性指数较高 ($d=0.94$), 模型估算误差较低。

参考文献:

- [1] 周杰, 夏晓剑, 胡璋健, 等. “十三五”我国设施蔬菜生产和科技进展及其展望[J]. 中国蔬菜, 2021(10): 20-34.
ZHOU Jie, XIA Xiaojian, HU Zhangjian, et al. Technological development and production of protected vegetable in China during ‘The thirteenth five-year plan’ and future prospect[J]. China Vegetables, 2021(10): 20-34.
- [2] 李银坤, 詹保成, 郭文忠, 等. 基于蒸发皿水面蒸发量的温室生菜适宜灌溉量研究[J]. 灌溉排水学报, 2022, 41(4): 13-19.
LI Yinkun, ZHAN Baocheng, GUO Wenzhong, et al. Optimizing irrigation amount for greenhouse lettuce production based on pan-measured evaporation[J]. Journal of Irrigation and Drainage, 2022, 41(4): 13-19.
- [3] LIU H, DUAN A W, LI F S, et al. Drip irrigation scheduling for tomato grown in solar greenhouse based on pan evaporation in North China plain[J]. Journal of Integrative Agriculture, 2013, 12(3): 520-531.
- [4] 段莉珠. 障碍物遮挡率对水面蒸发量的影响浅析[J]. 黑龙江水利科技, 2019, 47(9): 75-77.
DUAN Lizhu. Preliminary analysis of influence of barrier shielding ratio on evaporation quantity of water surface[J]. Heilongjiang Hydraulic Science and Technology, 2019, 47(9): 75-77.
- [5] 郭文忠, 陈青云, 高丽红, 等. 设施蔬菜生产节水灌溉制度研究现状及发展趋势[J]. 农业工程学报, 2005, 21(S2): 24-27.
GUO Wenzhong, CHEN Qingyun, GAO Lihong, et al. Present situation and developmental tendency on system of water-saving irrigation of vegetable production in protective cultivation[J]. Transactions of the Chinese Society of Agricultural Engineering, 2005, 21(S2): 24-27.
- [6] 周书影. TEZ-601 全自动称重式蒸发系统在杜陵水文站的应用比测分析[J]. 广东水利水电, 2020(12): 50-53.
ZHOU Shuying. Application research of TEZ-601 automatic weighing evaporation system in duling hydrology station[J]. Guangdong Water Resources and Hydropower, 2020(12): 50-53.
- [7] 武金慧, 李占斌. 水面蒸发研究进展与展望[J]. 水利与建筑工程学报, 2007, 5(3): 46-50.
WU Jinhui, LI Zhanbin. Advances and prospect of research on water surface evaporation[J]. Journal of Water Resources and Architectural Engineering, 2007, 5(3): 46-50.
- [8] 闵蓁. 利用彭曼公式预测水面蒸发量[J]. 水利水电科技进展, 2001, 21(1): 37-39.
MIN Qian. Prediction of water surface evaporation by penman formula[J]. Advances in Science and Technology of Water Resources, 2001, 21(1): 37-39.
- [9] 刘昌宇. 太阳光谱吸收对静止水体蒸发影响的研究[D]. 哈尔滨: 哈尔滨工程大学, 2014.
LIU Changyu. Effect of spectral absorption of solar energy on evaporation of still water[D]. Harbin: Harbin Engineering University, 2014.
- [10] 闵蓁. 湖泊(水库)水面蒸发量预测方法的探讨[J]. 水文, 1997, 17(2): 39-43
MIN Qian. Discussion on prediction method of lake (reservoir) water surface evaporation[J]. Hydrology, 1997, 17(2): 39-43.
- [11] 濮培民. 水面蒸发与散热系数公式研究(一)[J]. 湖泊科学, 1994, 6(1): 1-12.
PU Peimin. Studies on the formulae for calculating evaporation and heat loss coefficient from water surface in China(I)[J]. Journal of Lake Science, 1994, 6(1): 1-12.
- [12] JAFARI Majid, DINPASHOH Y. Derivation of regression models for pan evaporation estimation[J]. Environmental Resources Research, 2019, 7(1): 30-42.
- [13] 陆云燕, 王振龙, 吕海深, 等. 基于主成分分析和 BP 神经网络的五道沟水面蒸发计算研究[J]. 水文, 2022, 42(1): 35-39.
LU Yunyan, WANG Zhenlong, LYU Haishen, et al. Calculation of water surface evaporation in Wudaogou based on principal component analysis and BP neural network[J]. Journal of China Hydrology, 2022, 42(1): 35-39.
- [14] 刘海军, 黄冠华, 王明强, 等. 基于蒸发皿水面蒸发量制定冬小麦喷灌计划[J]. 农业工程学报, 2010, 26(1): 11-17.
LIU Haijun, HUANG Guanhua, WANG Mingqiang, et al. Sprinkler irrigation scheme of winter wheat based on water surface evaporation of a 20 cm standard pan[J]. Transactions of the Chinese Society of Agricultural Engineering, 2010, 26(1): 11-17.
- [15] 王湛, 李银坤, 郭文忠, 等. 不同灌水量对温室茄子蒸腾规律及水分利用的影响[J]. 中国农村水利水电, 2019(7): 6-10.
WANG Zhan, LI Yinkun, GUO Wenzhong, et al. The effect of different irrigation amount on the transpiration rate and water use efficiency of autumn eggplant in greenhouses[J]. China Rural Water and Hydropower, 2019(7): 6-10.
- [16] 胡晨, 葛继稳, 许向南, 等. 基于 FAO56 Penman-Monteith 公式估算神农架大九湖泥炭湿地蒸散及作物系数[J]. 应用生态学报, 2020, 31(5): 1 699-1 706.
HU Chen, GE Jiwen, XU Xiangnan, et al. Estimation of evapotranspiration and crop coefficient in Dajiuhu peatland of Shennongjia based on FAO56 Penman-Monteith[J]. Chinese Journal of Applied Ecology, 2020, 31(5): 1 699-1 706.
- [17] 杨梅, 肖静, 蔡辉. 多元分析中的多重共线性及其处理方法[J]. 中国卫生统计, 2012, 29(4): 620-624.
YANG Mei, XIAO Jing, CAI Hui. Multiple collinearity in multivariate analysis and its processing methods[J]. Chinese Journal of Health Statistics, 2012, 29(4): 620-624.
- [18] 梁素梅, 吴迪. 基于多重共线性的气象观测站优化模型[J]. 价值工程, 2020, 39(32): 157-160.
LIANG Sumei, WU Di. Optimization model of meteorological observation station based on multicollinearity[J]. Value Engineering, 2020, 39(32): 157-160.
- [19] 许莹莹, 李薇, 王振龙, 等. 基于主成分分析的裸地潜水蒸发与气象要素关系模拟[J]. 水文, 2020, 40(4): 7-13, 39.
XU Yingying, LI Wei, WANG Zhenlong, et al. Simulation of relationship between evaporation and meteorological elements of bare ground diving based on principal component analysis[J]. Journal of China Hydrology, 2020, 40(4): 7-13, 39.
- [20] MASSY W F. Principal components regression in exploratory statistical research[J]. Journal of the American Statistical Association, 1965, 60(309): 234-256.
- [21] 李万义. 适用于全国范围的水面蒸发量计算模型的研究[J]. 水文, 2000, 20(4): 13-17, 63.
LI Wanyi. A study on the generalized model of water surface evaporation[J]. Hydrology, 2000, 20(4): 13-17, 63.
- [22] 孙凤贤, 刘昌宇, 夏新林, 等. 太阳辐射对静止水面稳态蒸发的影响[J]. 计算物理, 2014, 31(6): 699-705.
SUN Fengxian, LIU Changyu, XIA Xinlin, et al. Effect of solar radiation on evaporation of still water surface[J]. Chinese Journal of Computational Physics, 2014, 31(6): 699-705.

Construction and Validation of a Model for Estimating Surface Water Evaporation in Greenhouse Based on Principal Component Analysis

WANG Ke^{1,2}, LI Yinkun¹, ZHENG Wengang¹, LIU Meiyang^{2*},
WU Jiale^{1,2}, JI Yuru³, CHEN Fei¹, HOU Shenglin⁴

(1. Intelligent Equipment Research Center, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China;
2. Inner Mongolia Key Laboratory of Soil Quality and Nutrient Resource, College of Grassland, Resources and Environment,
Inner Mongolia Agricultural University, Hohhot 010018, China; 3. College of Horticulture, China Agricultural University,
Beijing 100094, China; 4. Hebei Academy of Agriculture and Forestry Sciences, Shijiazhuang 050051, China)

Abstract: 【Objective】 Surface water evaporation (E_p) in greenhouse is often used as a reference in irrigation management. The purpose of this paper is to present a model to estimate it. 【Method】 The model was derived based on meteorological data and pan-evaporation measured from March to July in 2020 and 2022. The relationship between E_p and meteorological data was analyzed using the principal component method, from which a multiple linear regression model was developed to estimate E_p . 【Result】 E_p increased as time elapsed, with its average increasing from the range of 1.84~1.94 mm in March to the range of 3.77~5.15 mm in June in both 2020 and 2022. E_p was influenced by radiation and relative humidity the most. Photosynthetically active radiation had the highest correlation with E_p , with their correlation coefficient being 0.852 ($P<0.01$), followed by solar radiation and relative humidity with their associated correlation coefficient being 0.811 and -0.770, respectively ($P<0.01$). The first principal component of solar radiation, photosynthetically active radiation, and relative humidity has a great effect on E_p , with the eigenvalue being 4.44. The solar radiation affected E_p significantly, with the highest score coefficient (0.328), followed by relative humidity and photosynthetically active radiation, whose score coefficients were 0.311 and -0.321, respectively. Principal component analysis and verification showed that the estimated E_p using the proposed model agreed well with measured data, with $P<0.01$, $R^2=0.908$, $MBE=0.10$, $RMSE=0.48$ mm/d, and the consistency index=0.94. 【Conclusion】 Solar radiation, photosynthetically active radiation, and relative humidity are the main factors influencing E_p in the greenhouse. The multiple linear regression model derived from the principal component analysis is accurate and can provide real-time estimates of surface water evaporation in the greenhouse.

Key words: greenhouse; pan evaporation; meteorological factors; multiple collinearity; principal component analysis

责任编辑: 白芳芳