

引用格式:

代秋芳, 熊诗路, 李震, 宋淑然, 陈梓蔚, 王元. 基于 LSTM 的柑橘幼苗蒸发量预测[J]. 湖南农业大学学报(自然科学版), 2023, 49(6): 743-747.

DAI Q F, XIONG S L, LI Z, SONG S R, CHEN Z W, WANG Y. The prediction of evaporation for citrus seedlings based on LSTM[J]. Journal of Hunan Agricultural University(Natural Sciences), 2023, 49(6): 743-747.

投稿网址: <http://xb.hunau.edu.cn>



基于 LSTM 的柑橘幼苗蒸发量预测

代秋芳^{1,2,3}, 熊诗路^{1,2}, 李震^{1,2,3*}, 宋淑然^{1,2,3}, 陈梓蔚^{1,2}, 王元^{1,2}

(1.华南农业大学工程学院(人工智能学院), 广东 广州 510642; 2.国家柑橘产业技术体系机械化研究室, 广东 广州 510642; 3.广东省农情信息监测工程技术研究中心, 广东 广州 510642)

摘要: 将柑橘幼苗作为试验对象, 利用传感器采集空气相对湿度和温度, 以基质相对湿度、温度和 EC 值作为环境因子, 采用称重法实时采集作物的质量变化量作为作物的蒸发量; 以环境因子为模型输入, 作物蒸发量为模型输出, 构建长短期记忆神经网络(LSTM)预测模型, 优化后的最优模型结构以及训练参数包括 LSTM 模型的隐藏层 1 层, 隐藏层节点数为 120, 迭代样本数为 128, 训练迭代次数为 175, 网络的激活函数选择 tanh 函数, 学习率为 0.001, 时间步长为 72。LSTM 预测模型的决定系数(R^2)、均方根误差(RMSE)、平均绝对误差(MAE)分别为 0.993 9、0.015 5 g、0.011 3 g; 与循环神经网络(RNN)、门控循环单元(GRU)的预测效果进行对比, LSTM 预测模型的预测蒸发值更接近真实蒸发值, 预测结果相对误差范围波动最小, RMSE、MAE 最小, R^2 最大, 说明这 3 种模型中 LSTM 预测模型的预测效果最佳。

关键词: 柑橘幼苗; 蒸发量; 环境因子; 长短期记忆神经网络(LSTM)

中图分类号: S666; S126

文献标志码: A

文章编号: 1007-1032(2023)06-0743-05

The prediction of evaporation for citrus seedlings based on LSTM

DAI Qiufang^{1,2,3}, XIONG Shilu^{1,2}, LI Zhen^{1,2,3*}, SONG Shuran^{1,2,3}, CHEN Ziwei^{1,2}, WANG Yuan^{1,2}

(1.College of Electronic Engineering(College of Artificial Intelligence), South China Agricultural University, Guangzhou, Guangdong 510642, China; 2.Mechanization Research Office of National Citrus Industry Technology System, Guangzhou, Guangdong 510642, China; 3.Guangdong Agricultural Information Monitoring Engineering Technology Research Center, Guangzhou, Guangdong 510642, China)

Abstract: In this study, citrus seedlings were selected to estimate the predictions of evaporation. The air relative humidity and temperature were collected by sensors and mass method was used to collect the mass change of crops in real time as crop evaporation. The substrate relative humidity, temperature and EC value were used as environmental factors. With environmental factors as model input and crop evaporation as model output, a long short-term memory neural network(LSTM) prediction model was constructed. The optimized model structure and training parameters included 1 hidden layer of the LSTM model, 120 hidden layer nodes, 128 iteration samples, and 175 training iterations. The activation function of the network is tanh function, the learning rate was 0.001, and the time step was 72. The coefficient of determination(R^2), root mean square error(RMSE) and mean absolute error(MAE) of LSTM prediction model were 0.993 9, 0.015 5 g and 0.011 3 g, respectively. Compared with the prediction effect of recurrent neural network(RNN) and gated cycle unit(GRU), the predicted evaporation value from LSTM prediction model was closer to the real evaporation value, and the relative error range of prediction results had the smallest fluctuation, RMSE and MAE were

收稿日期: 2022-08-30

修回日期: 2023-10-20

基金项目: 国家自然科学基金项目(31971797); 广东省现代农业产业技术体系创新团队建设项目(2022KJ108); 财政部和农业农村部国家现代农业产业技术体系(CARS-26)

作者简介: 代秋芳(1979—), 女, 湖南常德人, 博士, 讲师, 主要从事喷雾技术及电子信息技术研究, daiqiufang@scau.edu.cn; *通信作者, 李震, 博士, 教授, 主要从事机电一体化技术应用研究, Lizhen@scau.edu.cn

the smallest, and R^2 was the largest, indicating that the prediction effect of LSTM prediction model was the best among these three models.

Keywords: citrus seedlings; evaporation; environmental factor; long short-term memory neural network(LSTM)

作物的蒸发量对于实现作物的精准灌溉具有关键意义,作物的蒸发量能够反映作物的生长状态以及生长环境的变化。精准获取作物蒸发量是作物水分管理中不可或缺的一步。关于作物蒸发量的研究大多是根据联合国粮农组织推荐的 FAO-56 Penman-Monteith(P-M)公式计算蒸散量(ET_0)以及利用热平衡法计算蒸腾速率^[1]。陈新明等^[2]利用温室大棚内计算蒸发量公式对 P-M 公式进行修正改进。但 P-M 公式需要参数较多,参数获取难度较大,公式复杂,实际应用难度大^[3-4]。

李灶鹏等^[5]利用无线传感器网络采集数据以及 GA-BP 神经网络预测大棚棉花蒸发量。冯培存等^[6]利用优化后的 BP 神经网络预测温室甜瓜的 ET_0 , 预测的相关系数在 0.8 以上,具有较好的预测效果。张庆玉等^[7]利用热平衡法测定柑橘树日间蒸腾速率并用人工神经网络对蒸腾速率进行预测。随着深度学习的发展应用,对非线性问题的拟合效果不断提升。MAJHI 等^[8]提出了不同输入组合的 Deep-LSTM (长短期记忆神经网络)模型来模拟印度中东部恰蒂斯加尔邦 3 个农业气候区的每日蒸发损失量并与人工神经网络进行比较,Deep-LSTM 模型预测精度更高。LAKMINI PRARTHANA JAYASINGE 等^[9]利用混合 LSTM 模型预测了澳大利亚昆士兰州 3 个干旱易发地区的蒸发量,取得较好的预测效果。高鹏等^[10]发现土壤含水率以及电导率和环境因子(蒸散量、空气相对湿度等)具有高度相关性,利用 LSTM 模型预测土壤含水率和电导率并与广义回归神经网络(GRNN)预测效果进行对比,发现 LSTM 模型预测精度高,性能稳定。李珊珊等^[11]基于卷积神经网络-双向长短期记忆神经网络(CNN-BiLSTM)模型预测温室甘蓝潜在蒸散量,利用卷积神经网络(CNN)提取数据特征,再结合 LSTM 网络得到最终模型预测值,结果表明 CNN-BiLSTM 模型预测效果优于随机森林(RF)、支持向量机(SVR)等模型。李莉等^[12]基于 LSTM 预测了温室内的番茄植株蒸发量,结果表明 LSTM 预测模型的预测精度和预测速度优于其他模型的。上述研究结果表明 LSTM 模

型预测效果较好,能够提升作物精准灌溉的效果。但针对幼苗这类蒸发量变化量差异较小的作物的蒸发量预测的研究较少。

针对以上研究现状,笔者制作柑橘幼苗环境与生长信息采集系统,建立 LSTM 预测模型,并通过模型对比、指标评价等验证预测模型的性能,以期实现对柑橘幼苗蒸发量的精准预测。

1 材料与方法

1.1 试验装置与供试对象

于 2022 年 5—6 月,在华南农业大学工程学院(113°20'44"E, 23°9'30"N)的日光温室进行试验。试验地点属于亚热带季风气候区。试验装置包括柑橘幼苗环境与生长信息采集系统及柑橘潮汐灌溉育苗系统。采集系统包括 E95-DTU(433Lxx)型 LoRa 无线数字传输电台、PR-300BYH-M 型气象多要素百叶箱(空气相对湿度精度 $\pm 3\%$ 、空气温度精度 $\pm 0.5\text{ }^\circ\text{C}$)、PR-3000-ECTH-N01 型三合一传感器(电导率精度 $\pm 3\%$ 、水分精度 $\pm 0.1\%$ 、温度精度 $\pm 0.5\text{ }^\circ\text{C}$)、CHS-D 型电子秤(量程 30 kg、精度 $\pm 0.1\text{ g}$)等。柑橘潮汐灌溉育苗系统包括灌溉控制系统、苗床、营养液池等。试验柑橘幼苗品种为‘明日见’,植株高 50 cm,叶片数为 16,处于幼苗期到开花坐果期。采用无纺布育苗袋培育,育苗袋尺寸为直径 20 cm、高 30 cm。培育柑橘幼苗的栽培基质选用黄土、椰糠和泥炭土,按照 1:1:1 的体积配比进行均匀混合,基质容重约为 0.64 g/cm^3 ,总孔隙度约为 70%。该混合基质的容重和总孔隙度适宜柑橘幼苗的培育^[13-14]。

1.2 试验方法

采用称重法^[15]对柑橘幼苗的蒸发量进行测定。在处理前,采用柑橘潮汐育苗灌溉系统对植株进行充分灌溉,使得基质水分达到充分饱和。运用采集系统每隔 10 min 采集幼苗和承托容器的质量。在盛放柑橘幼苗的容器顶部,需要围绕幼苗的茎秆覆盖薄膜,以防止基质水分蒸发,避免对测量的精度造成影响。

使用采集系统采集温室内空气相对湿度，基质的电导率、体积含水率、温度等数据。数据采集时间为 2022 年 5 月 19 日至 2022 年 6 月 19 日每天的 07:00—19:00，每间隔 10 min 上传 1 次，总共 30 d 的数据，共 2160 条数据。每个数据值测量 3 次并取平均值，以减小误差的影响。

1.3 数据采集分析

运用 Python 的 Pandas 和 Numpy 以及 Matplotlib 对数据进行处理以及绘图。空气相对湿度为 76%~88%，空气温度为 25~31 °C，基质温度约为 24.5~28.0 °C，低于空气温度的范围。另外，根据数据处理的结果分析，蒸发量在每天 12:00—14:00 最大(图 1)，也符合作物的日间蒸发规律^[16]。

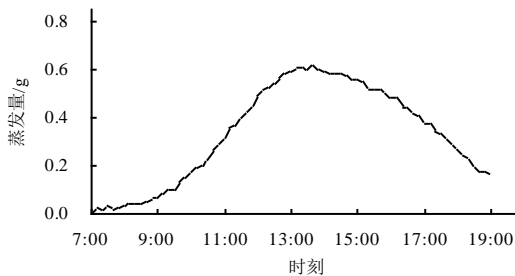


图 1 不同时刻柑橘幼苗的蒸发量
Fig.1 Evaporation data at different times

2 数据处理与模型构建

2.1 数据预处理

采集过程中数据的传输会产生丢包，导致采集数据失败。针对缺少的数据，采用前 1 天和后 1 天同一时刻的数据来填充数据。对采集过程出现的异常数据需要进行剔除，再采用同类均值插补法^[17]进行处理。为了消除样本数据的量级差异的影响以及加快模型训练收敛速度，数据进行归一化处理。

2.2 模型构建

蒸发量预测流程如图 2 所示。主要包括数据采集及处理、模型构建训练、预测结果评估 3 部分。具体为：①利用传感器定时采集柑橘幼苗的生长环境信息以及柑橘幼苗的生长信息，然后对数据进行异常值和缺失值处理，并对其进行归一化操作，最后再把归一化后的数据作为样本，划分训练集和测试集；②构建 LSTM 模型，对模型进行初始参数设置，将训练集样本的数据转换成 LSTM 模型的输入维度，再输入到模型中进行训练，根据训练得到的

预测结果来调整模型结构以及训练的参数，包括隐藏层节点数、训练迭代次数、迭代样本数、时间步长、学习率等参数；③通过对模型参数的调整得到最优的 LSTM 模型，然后通过最优模型训练得出预测的结果。

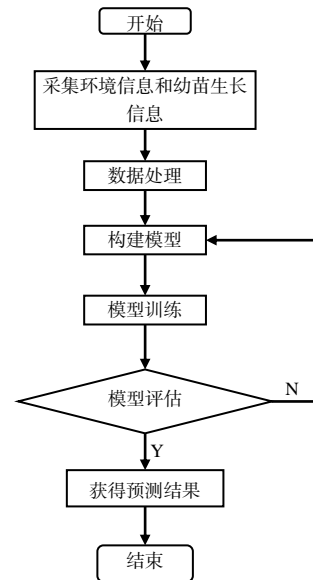


图 2 柑橘幼苗蒸发量预测流程
Fig.2 Flowchart of citrus seedling evaporation prediction

2.3 模型评价

为了验证模型预测柑橘幼苗蒸发量的性能，参照文献[18]，采用决定系数(R^2)、均方根误差(RMSE)、平均绝对误差(MAE)等指标对模型的效果和性能进行评价。针对 LSTM 模型拟合结果的置信度，采用标准差作为指标进行评价。为了探究不同预测模型对柑橘幼苗蒸发量的预测效果，同时利用 LSTM、循环神经网络(RNN)、门控循环单元(GRU)对柑橘幼苗蒸发量进行预测分析并对比预测效果。

3 结果与分析

3.1 模型训练结果

利用 2022 年 5—6 月在华南农业大学工程学院日光温室采集的柑橘幼苗生长数据进行模型仿真，经过数据的有效处理后，总计有效数据为 2088 条。把数据按照时间顺序进行排序，选取 1728 条数据作为预测模型的训练集数据，剩余 360 条数据作为预测模型的测试数据。

将时间步长设置为 1、6、72，分别代表利用样本的 0~10 min、0~1 h、0~1 d 的数据预测当前的结果。隐藏层的节点数设置为 32、60、90、120，对

其进行组合,选取 RMSE 最小的值对应的时间步长和节点数等参数作为 LSTM 网络的结构设置。从表 1 可以看出,当网络的隐含层节点数为 120,时间步长为 72 时,LSTM 预测模型测试样本训练的

RMSE 最小,为 0.016 g,MAE 为 0.012 5 g,故 LSTM 柑橘幼苗蒸发量预测模型选择隐含层节点数为 120,时间步长为 72。

表 1 LSTM 模型不同网络结构和参数下的 RMSE 和 MAE

Table 1 RMSE and MAE of LSTM models under different network structures and parameters								
步长	RMSE				MAE			
	30	60	90	120	30	60	90	120
1	0.025	0.024	0.024	0.024	0.018 6	0.017 8	0.017 5	0.017 4
6	0.021	0.020	0.018	0.019	0.014 9	0.015 0	0.012 3	0.014 0
72	0.019	0.016	0.016	0.016	0.016 7	0.015 9	0.013 1	0.012 5

通过多次的模型参数调整并验证预测结果,得到调整后的 LSTM 预测模型的最优网络参数。最优模型结构以及训练参数包括 LSTM 模型的隐藏层 1 层,隐藏层节点数为 120,迭代样本数为 128,训练迭代次数 175,网络的激活函数选择 tanh 函数,学习率为 0.001,时间步长为 72。

3.2 模型评价结果

LSTM、RNN、GRU 预测模型在相同的运算环

境以及训练集样本数据一致的情况下,预测柑橘幼苗蒸发量的变化曲线如图 3 所示。由图 3 可知:在蒸发量变化峰值的区间内,LSTM 的预测曲线更接近真实的曲线,预测蒸发值更接近真实蒸发值,不会出现过高或者过低的峰值抖动情况;RNN 在预测蒸发量的数据变化较大时会出现较大的波动;GRU 预测结果普遍偏低,有些甚至出现负值情况,造成严重偏差。

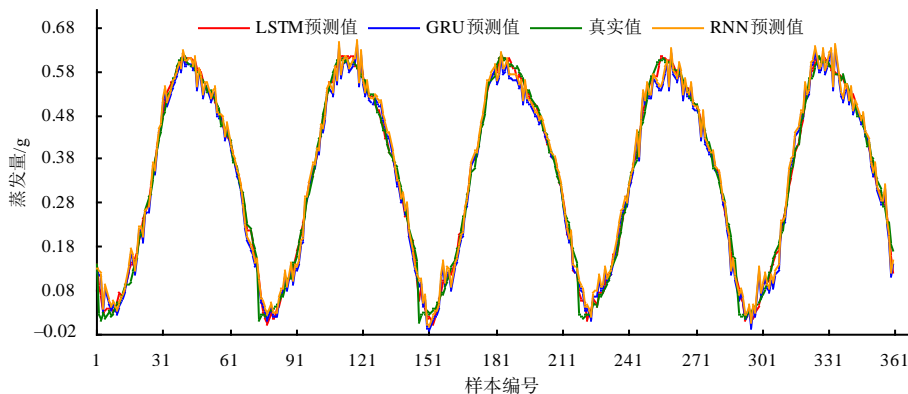


图 3 3 种模型测试集的蒸发量预测曲线

Fig.3 Evaporation prediction plot for the test set of the three models

从图 4 可知,对比模型中,RNN 模型预测结果

相对误差范围波动最大,LSTM 模型的最小,表明

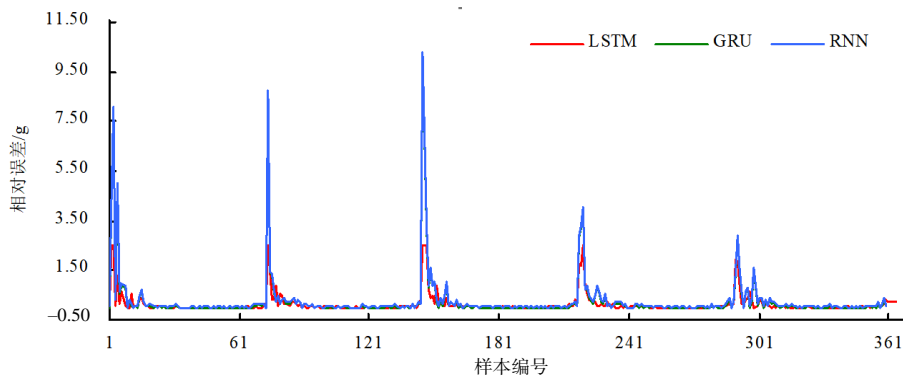


图 4 3 种模型测试集的蒸发量预测的相对误差

Fig.4 Relative error comparison of evaporation prediction for the test set of the three models

基于 LSTM 预测模型的相关性更好。表 2 说明了 LSTM、RNN、GRU 预测模型的评价指标对比效果。RNN、GRU 预测模型的 R^2 分别为 0.975 0、0.979 0，LSTM 预测模型的 R^2 为 0.993 9，表明在同一条件下基于 LSTM 的柑橘幼苗蒸发量预测模型的预测精度要高于其他 2 个预测模型的，该模型可更好地提取数据的深度特征，获得更优的非线性拟合效果。

表 2 3 种预测柑橘幼苗蒸发量模型的预测效果

Table 2 The prediction effects of the three models on citrus seedling evaporation

模型	RMSE/g	MAE/g	R^2
LSTM	0.015 5	0.011 3	0.993 9
RNN	0.024 6	0.018 8	0.975 0
GRU	0.019 1	0.014 5	0.979 0

把环境因子分为3组：空气的温度和相对湿度为1组；土壤的温度、相对湿度和EC值为1组；空气和土壤全部因素相加为1组。3组数据分别输入到模型中，得到LSTM预测模型的标准差分别为0.181、0.179、0.176 g。可见，该模型不受输入样本的影响，拟合置信度较高，具有较好的鲁棒性。

4 结论

本研究搭建了柑橘幼苗环境与生长信息采集系统，采集了柑橘幼苗的生长环境数据以及蒸发量数据，应用上述数据建立了LSTM预测模型；该模型能预测柑橘幼苗的蒸发量，且预测效果较优，模型的MAE、RMSE、 R 分别为0.011 3 g、0.015 5 g、0.993 9，优于RNN、GRU等模型。可见，相比于RNN、GRU，基于LSTM的柑橘幼苗蒸发量预测模型提升了柑橘幼苗蒸发量预测精度，但不足之处的是采集的数据信息量较少，模型还需要在实际的灌溉工程中进一步验证分析。

参考文献：

[1] 吴宏霞. 基于BP神经网络的参考作物蒸发蒸腾量预测研究[D]. 南京：河海大学，2006.

[2] 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，2007，18(2)：317–321.

[3] NYOLEI D, DIELS J, MBILINYI B, et al. Evapotranspiration simulation from a sparsely vegetated agricultural field in a semi-arid agro-ecosystem using Penman-Monteith models[J]. Agricultural and Forest Meteorology, 2021, 303: 108370.

[4] 刘天凤. 基于作物蒸腾量智能预测的农田优化灌溉系统[D]. 上海：东华大学，2018.

[5] 李灶鹏，李文竹，刘婧然，等. 基于LoRa与GA-BP的规模化大棚节水灌溉系统[J]. 人民长江，2020，51(7)：225–230.

[6] 冯培存，魏正英，张育斌，等. 基于思维进化算法优化BP神经网络的温室甜瓜 ET_0 预测研究[J]. 节水灌溉，2019(9)：36–39.

[7] 张庆玉，董晓华，葛亮，等. 基于人工神经网络的柑橘树日间蒸腾速率精细模拟与预测[J]. 植物生理学报，2020，56(3)：489–500.

[8] MAJHI B, NAIDU D, MISHRA A P, et al. Improved prediction of daily pan evaporation using Deep-LSTM model[J]. Neural Computing and Applications, 2020, 32(12): 7823–7838.

[9] LAKMINI PRARTHANA JAYASINGHE W J M, DEO R C, GHARAMANI A, et al. Development and evaluation of hybrid deep learning long short-term memory network model for pan evaporation estimation trained with satellite and ground-based data[J]. Journal of Hydrology, 2022, 607: 127534.

[10] 高鹏，谢家兴，孙道宗，等. 基于物联网和LSTM的柑橘园土壤含水量和电导率预测模型[J]. 华南农业大学学报，2020，41(6)：134–144.

[11] 李珊珊，赵春江，吴华瑞，等. 基于CNN-BiLSTM融合神经网络的温室甘蓝潜在蒸散量预测[J/OL]. 吉林农业大学学报，1–10(2021–11–16). <http://kns.cnki.net/kcms/detail/22.1100.S.20211116.1432.002.html>.

[12] 李莉，李文军，马德新，等. 基于LSTM的温室番茄蒸腾量预测模型研究[J]. 农业机械学报，2021，52(10)：369–376.

[13] 武捷. 不同基质及肥料配比对酸柚苗生长和生理的影响[D]. 海口：海南大学，2018.

[14] 李荣华. 积优株筛选及柑橘育苗基质和容器的改良[D]. 长沙：湖南农业大学，2014.

[15] 李建明，任瑞丹，范兵华，等. 灌溉量对温室全有机营养液栽培甜瓜根际环境和茎流的影响[J]. 农业工程学报，2020，36(7)：217–224.

[16] 彭致功. 日光温室滴灌条件下小气候变化和植株蒸腾规律的研究[D]. 北京：中国农业科学院，2002.

[17] 庞新生. 缺失数据插补处理方法的比较研究[J]. 统计与决策，2012(24)：18–22.

[18] 高鹏，杨明欣，周平，等. 基于物联网和Deep-LSTM的茶树净光合速率动态预测模型[J]. 农业工程学报，2022，38(4)：159–168.

责任编辑：邹慧玲

英文编辑：柳正