

DOI: 10.20103/j.stxb.202406251476

邓炜,刘登峰,李明亮,孟静静,黄强.机器学习方法在黄土高原沟壑区模拟二氧化碳浓度的潜力评估.生态学报,2025,45(13): - .
Deng W, Liu D F, Li M L, Meng J J, Huang Q. Evaluation of the potential of using machine learning to simulate the CO₂ concentration in the gully region of the Chinese Loess Plateau. Acta Ecologica Sinica, 2025, 45(13): - .

机器学习方法在黄土高原沟壑区模拟二氧化碳浓度的潜力评估

邓 炜^{1,2}, 刘登峰^{1,*}, 李明亮³, 孟静静¹, 黄 强¹

1 西安理工大学旱区水工程生态环境全国重点实验室, 西安 710048

2 赣江下游水文水资源监测中心, 宜春 336000

3 水利部水利水电规划设计总院, 北京 100120

摘要: 二氧化碳 (CO₂) 是主要的温室气体之一, 空气中二氧化碳浓度的变化受到多种因素的影响。现在能连续观测地面空气二氧化碳浓度的站点依旧较少, 但是常规气象数据容易观测, 在给定地点如何用常规气象数据模拟地面空气二氧化碳浓度依旧缺乏研究。基于陕西省淳化县的和家山小流域的实测气象数据和二氧化碳浓度数据, 分别采用不同位置的地面空气二氧化碳浓度或环境因子的组合驱动 MLP、LSTM、Bi-LSTM 和 RF 模型, 评价了机器学习方法在黄土高原沟壑区典型站点模拟日尺度地面空气二氧化碳浓度的潜力。研究结果表明, 无论是以同类型观测数据作为输入还是以环境因子的组合作为输入, 4 种机器学习模型都可以较好的模拟地面空气二氧化碳浓度的整体变化过程, 且都可以用于插补数据。以其他点位实测二氧化碳浓度作为输入时, MLP 模拟 2 号点的浓度 C₂ 精度最高, 在测试集的误差仅为 3.8%; 以环境因子作为输入时, RF 模拟 2 号点的浓度 C₂ 精度最高, 在测试集的误差为 6.3%。但是以环境因子作为输入时无法模拟出二氧化碳浓度剧烈的日变化过程。为了进行数据插补, 以同类型的观测数据作为输入用于插补数据的效果更好, 它可以较好地模拟出地面空气二氧化碳浓度的日变化过程。

关键词: 二氧化碳浓度; 机器学习; 黄土高原; 环境因子

Evaluation of the potential of using machine learning to simulate the CO₂ concentration in the gully region of the Chinese Loess Plateau

DENG Wei^{1,2}, LIU Dengfeng^{1,*}, LI Mingliang³, MENG Jingjing¹, HUANG Qiang¹

1 State Key Laboratory of Water Engineering Ecology and Environment in Arid Area, Xi'an University of Technology, Xi'an 710048, China

2 Hydrology and Water Resources Monitoring Center of Lower Ganjiang River, Yichun 336000, China

3 General Institute of Water Resources and Hydropower Planning and Design, Ministry of Water Resources, Beijing 100120, China

Abstract: Carbon dioxide (CO₂) is one of the major greenhouse gases, and changes in CO₂ concentration in the air is affected by many factors. At present, there are still few stations that can continuously observe ground air CO₂ concentration, but conventional meteorological data are relatively easy to obtain. There is still a lack of research on how to use conventional meteorological data to simulate ground air CO₂ concentration at a specific location. Based on the measured meteorological data and CO₂ concentration data from the Hejiashan small watershed in Chunhua County, Shaanxi Province, the MLP, LSTM, Bi-LSTM and RF models were driven by ground air CO₂ concentration or the combination of environmental factors at different locations, respectively. The potential of machine learning methods to simulate daily-scale air CO₂ concentration on the ground of typical point in the gully region of the Loess Plateau was evaluated. The results show that the four machine

基金项目: 国家重点研发计划项目 (2022YFF1302200); 国家自然科学基金 (52279025, 42071335)

收稿日期: 2024-06-25; **网络出版日期:** 2025-00-00

* 通讯作者 Corresponding author. E-mail: liudf@xaut.edu.cn

learning methods can simulate the overall change process of ground air CO₂ concentration, whether using the same type of observed data as input or combination of environmental factors as input, and can be used for data interpolating. When the measured CO₂ concentration at other points is used as input, the simulation accuracy of the CO₂ concentration at C₂ by MLP is the highest, and the error in the test set is only 3.8%. When the environmental factors are used as input, the simulation accuracy of the CO₂ concentration at C₂ by RF is the highest, and the error in the test set is 6.3%. However, when environmental factors are used as input, it is impossible to simulate the dramatic daily variation of CO₂ concentration. For data interpolation, it is better to use the same type of observed data as input, which can better simulate the daily process of ground air CO₂ concentration.

Key Words: CO₂ concentration; machine learning; Loess Plateau; environment factors

虽然影响全球气候变化的因素很多,但是化石燃料燃烧、土地利用变化等人类活动造成的大气中温室气体含量增加是近年来全球气候变化的主因^[1-2]。二氧化碳(CO₂)是主要的温室气体之一,大量温室气体的排放不仅使得全球海平面上升,还使得相关的自然灾害进一步频发,甚至造成某些生物的灭绝^[3-4]。早在 20 世纪 50 年代末,国外就已开展温室气体的观测研究,但我国在 90 年代才建立了第一个全球温室气体本底站-青海瓦里关全球大气基准站^[5]。仪器监测目前仍然是 CO₂浓度研究的主要方式。如韩炜等^[6]通过采用开路式红外分析器收集 CO₂浓度,分析了近地面 CO₂浓度与下垫面及气象要素的相关性,刘寿东等^[7]通过南京市多站点、多高度观测所得 CO₂浓度分析了南京夏季城市大气二氧化碳浓度的时空分布规律。Crawford 等^[8]使用分布在北半球的多个不同源头水流中的高频传感器的实测二氧化碳浓度数据,探索了二氧化碳分压变化的基本模式。目前仪器观测已用于城市和野外等多场景的二氧化碳浓度研究中,但仪器的铺设常具有一定的局限性,受地形等因素影响较大。因而卫星监测因其具有不受时空限制、覆盖范围广、观测稳定、时间序列长和三维观测等优点备受关注^[9]。目前专用的 CO₂探测卫星包括日本发射的温室气体观测卫星(GOSAT)^[10]、美国发射的轨道碳观测者 2(OCO-2)^[11]以及中国发射的碳卫星(TanSat),它们为广大学者提供了更好的观测方法及数据。吴迪等^[12]就对 GOSAT 卫星监测的二氧化碳浓度进行差值处理,分析了黑龙江各地市的 CO₂浓度时空分布规律。陆地生态系统中的植被呼吸、光合作用以及土壤呼吸,对地面二氧化碳浓度的动态变化具有显著影响。这些过程与当地的水热条件、有效辐射等环境变化因素又紧密相关。在植被茂盛的区域,太阳辐射较强时,植物通过气孔吸收大气中的二氧化碳进行光合作用释放氧气,近地面二氧化碳浓度降低^[13]。当水热条件适宜时,植物的生长和光合作用效率提高,进一步促进固碳。当光照不足,或者太阳辐射过强时,植被和土壤呼吸占据主导作用。植物呼吸和土壤呼吸作用增强时,近地面空气二氧化碳浓度升高。土壤含水量是土壤呼吸作用的重要控制因素,降水也可通过调节异养呼吸来调节土壤总呼吸^[14-15]。这些环境因子都会影响近地面空气二氧化碳浓度的变化。

二氧化碳浓度的数据质量直接影响研究结果。为了取得高质量的数据,更好的探究二氧化碳浓度变化规律,目前已有部分学者将机器学习方法用于二氧化碳浓度的模拟,曹世杰等^[16]以神经网络模型对室内平均二氧化碳浓度进行模拟,但由于其针对室内环境,并未考虑气候因素的影响。多种模型的耦合也同样用于二氧化碳浓度模拟研究中,尹航等^[17]融合了麻雀搜索算法和极限学习机 ELM,成功预测了羊舍的二氧化碳浓度。付子骏等^[18]采用张量分解和序列最小二乘规划对长短期记忆 LSTM 进行耦合,更好地利用了变量间的内在联系,更加准确地模拟了二氧化碳浓度。

黄土高原地区水土流失尤为突出,是我们国家乃至世界上水土流失最严重、面积最大的区域^[19]。目前大部分学者大多是对室内的二氧化碳浓度进行模拟,对于黄土高原地区沟壑区的地面空气二氧化碳浓度模拟的研究并不多,且野外实测数据也常常由于设备故障等问题影响数据的完整性^[20]。因此,本文采用机器学习方法对黄土高原沟壑区地面空气二氧化碳浓度进行模拟,对多种方法和多种输入数据的模拟效果进行了评估,

并将其用于数据插补,提高数据质量。这有助于更好的了解黄土高原沟壑区地面空气二氧化碳浓度的变化过程,从而服务于变化环境下流域生态水文过程的研究,并对黄土高原水土保持、植被恢复、节水固碳等工作的开展具有重要的应用价值和科学意义。

1 研究区与数据

1.1 研究区概况

本研究位于陕西咸阳市淳化县的和家山小流域,建设了淳化生态水文实验基地,研究区可开展地面空气二氧化碳浓度观测、水碳通量观测和常规气象观测^[21-22]。实验基地的地理位置如图 1 所示,其位于渭河北岸的黄土高原沟壑区,盛行东风。气候属温带大陆性季风气候,降雨主要集中在每年的夏季和秋季^[23]。土壤类型以黄绵土为主^[24]。研究区的植被覆盖度高,主要植被类型包括灰绿藜(*Chenopodium glaucum*)、狗尾草(*Setaria viridis*)、碱蒿(*Artemisia anethifolia*)等^[25]。

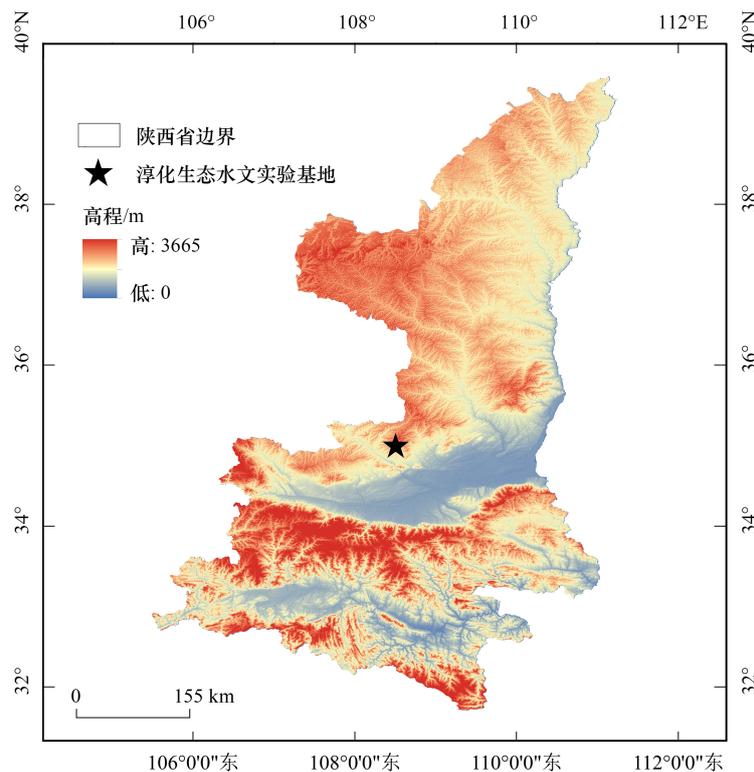


图 1 淳化生态水文实验基地位置图

Fig.1 The location of Chunhua eco-hydrological experimental base

1.2 二氧化碳浓度观测系统

二氧化碳浓度观测系统的主要设备有二氧化碳传感器(GMP252, Vaisala, Finland)和六要素气象传感器(WXA100-06, 中铭电气, China),如图 2 所示。观测系统的具体布设方案如图 3 所示,布设方案简图如图 4 所示,该套监测系统在安装初期进行了设备调试及标定,设备安装公司建议每年进行一次标定,本文仅采用数据长度为 350 的数据(未超过建议标定时间),共设置三个安装在距离地面 0.3m 高度的二氧化碳传感器用于观测地面空气二氧化碳浓度,自北向南分别命名为 1 号传感器、2 号传感器和 3 号传感器,其中 1 号传感器处的地面空气二氧化碳浓度命名为 C_1 , 2 号传感器处的地面空气二氧化碳浓度命名为 C_2 , 3 号传感器处的地面空气二氧化碳浓度命名为 C_3 。在距离地面 3m 高度处安装六要素气象传感器,用于观测空气温度、空气相对湿度、风向、风速、气压和雨量;在距地面 0.2m 深度处埋设土壤三参数传感器,用于观测土壤温度、土壤湿度和

电导率。该观测系统可对野外地面空气二氧化碳浓度、土壤状态和常规气象要素进行连续观测,可获得 10min、30min、1h 和 1 天尺度等多尺度的实测数据。



图 2 二氧化碳浓度观测系统的主要设备

Fig.2 The main equipment of carbon dioxide concentration observation system



图 3 二氧化碳浓度观测系统布设方案

Fig.3 The layout scheme of carbon dioxide concentration observation system

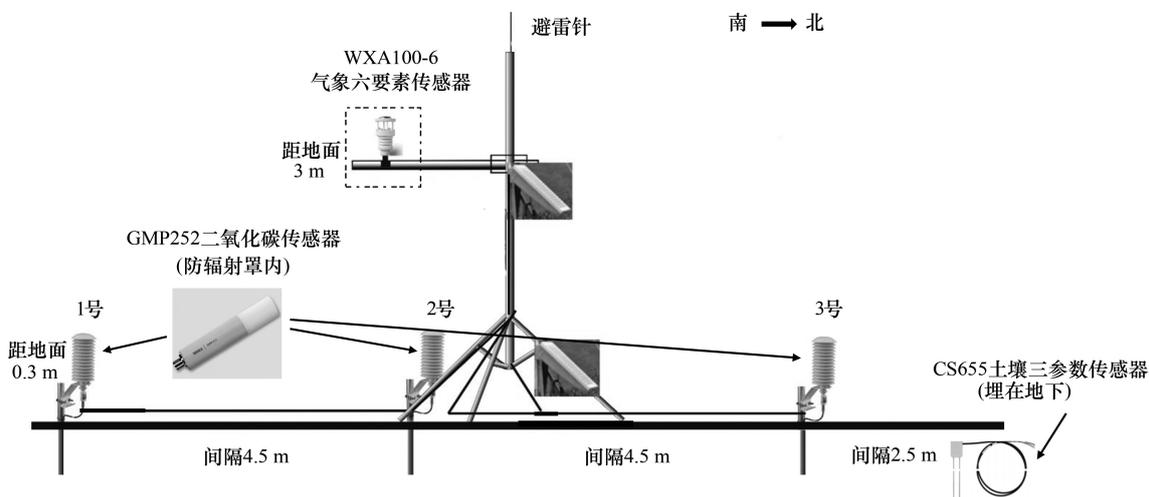


图 4 二氧化碳浓度观测系统布设方案简图

Fig.4 The layout scheme diagram of carbon dioxide concentration observation system

1.3 数据及预处理

本研究所采用的数据均为淳化生态水文实验基地的实测数据,包括自北向南三个不同位置的地面空气二氧化碳浓度(C_1 、 C_2 、 C_3)、0.2m 深度土壤温度(T_s)、空气温度(T_{air})、空气相对湿度(RH)和气压(P)。所用数

据均为日尺度数据,数据为 2021 年 2 月 1 日至 2022 年 1 月 23 日,数据长度为 350 个。作为输入变量的土壤温度、空气温度、空气相对湿度和气压插补后采用 SG 滤波^[26]进行预处理,预处理后的数据如图 5 至图 9 所示。从图 4 可以看出,二氧化碳浓度日变化十分剧烈,三个传感器所测地面空气二氧化碳浓度具有相同的变化趋势,在 1 月和 2 月时二氧化碳浓度较高,6 月至 10 月二氧化碳浓度较低。对比不同传感器的数据可知,1 号传感器于 2021 年 6 月初出现故障,1 号传感器从 2021 年 6 月份开始出现所测地面空气二氧化碳浓度偏低的情况,2021 年 6 月份之后 1 号探头的实测数据存在问题。在环境因子中,空气温度和土壤温度呈现相同的季节变化趋势,夏季温度较高,冬季温度较低。空气相对湿度和气压无明显的季节变化趋势。

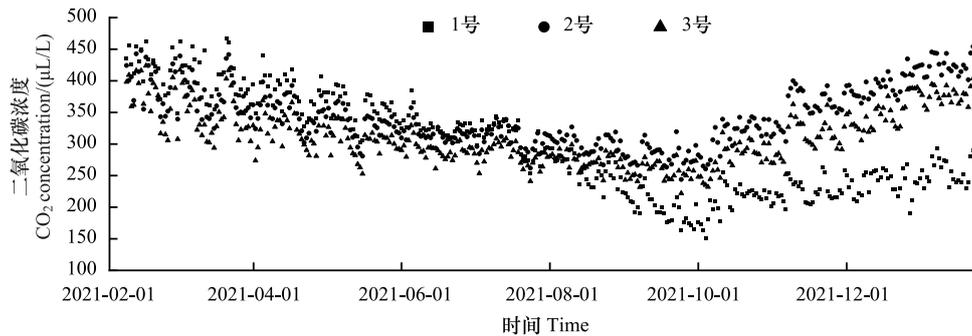


图 5 地面空气二氧化碳浓度的数据

Fig.5 The data of ground air carbon dioxide concentration

为了分析二氧化碳浓度(C_2 、 C_3),空气温度(T_{air})、土壤温度(T_s)、空气相对湿度(RH)以及气压(P)在研究时段的统计信息,本研究计算了各数据序列的平均值(x_{mean})、最大值(x_{max})、最小值(x_{min})、标准差(x_{std})、变异系数(C_v)、偏度(C_s)和峰度(C_k),计算结果见表 1。1 号传感器因发生故障,不进行统计信息计算。计算结果表明,在研究时段内,空气温度的最大值为 23.84℃,最小值为-8.34℃,平均值为 11.00℃,标准差为 7.44℃;土壤温度的最大值为 19.60℃,最小值为 0.60℃,平均值为 10.66℃,标准差为 6.17℃,空气温度和土壤温度的平均值和标准差十分接近。用于模拟的四个环境因子中

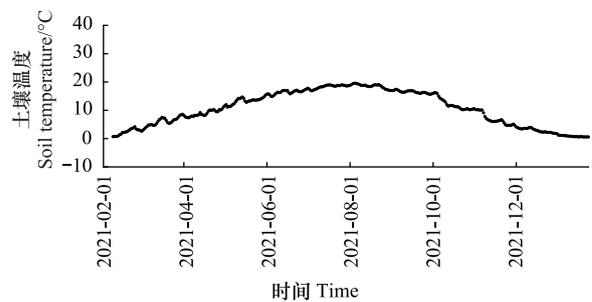


图 6 土壤温度数据

Fig.6 The data of soil temperature

只有气压为正偏,空气温度、土壤温度和相对湿度都为负偏。二氧化碳浓度的标准差较大,气压的标准差较小。气压和二氧化碳浓度的数值都较大,而气压的标准差远小于二氧化碳浓度,说明气压的数据分布较集中,

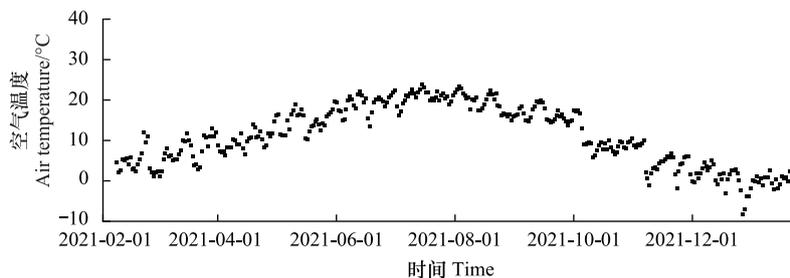


图 7 空气温度数据

Fig.7 The data of air temperature

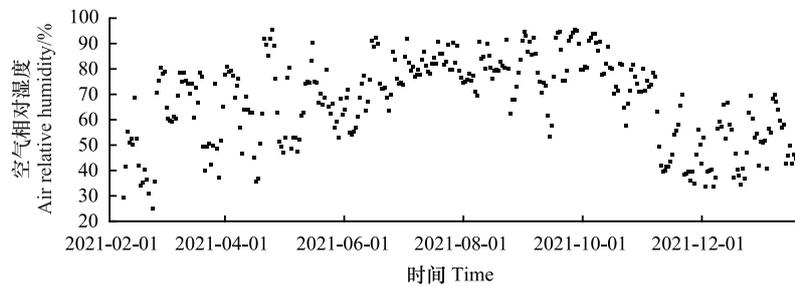


图8 空气相对湿度数据

Fig.8 The data of air relative humidity

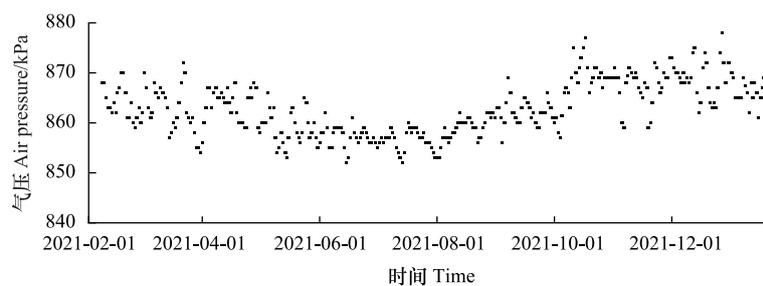


图9 气压数据

Fig.9 The data of air pressure

变化较平稳。此外,二氧化碳浓度、空气温度、土壤温度、相对湿度和气压的峰度都为负数,说明它们的分布峰值较平。对比不同传感器实测的二氧化碳浓度来看,2号传感器和3号传感器处的地面空气二氧化碳浓度的最大值、最小值、均值和方差相差不大,仅相差10%左右。

表1 二氧化碳浓度和环境因子的简要统计参数

Table 1 The brief statistical parameters of carbon dioxide concentration and environmental factors

数据 Data	统计参数 Statistical parameters						
	x_{mean}	x_{max}	x_{min}	x_{std}	C_v	C_s	C_k
2号处二氧化碳浓度/($\mu\text{L/L}$) 度 CO_2 value at the second sensor location	343.34	454.15	243.96	47.05	0.14	0.38	-0.69
3号处二氧化碳浓度/($\mu\text{L/L}$) CO_2 value at the third sensor location	310.26	423.85	222.00	43.46	0.14	0.55	-0.42
空气温度 Air temperature/ $^{\circ}\text{C}$	11.00	23.83	-8.34	7.44	0.68	-0.14	-1.11
土壤温度 Soil temperature/ $^{\circ}\text{C}$	10.66	19.60	0.60	6.17	0.58	-0.16	-1.41
空气相对湿度 Air relative humidity/%	67.37	95.50	18.66	16.88	0.25	-0.42	-0.73
气压 Air pressure /kPa	862.73	878.00	852.00	5.34	0.01	0.26	-0.69

综合来看,二氧化碳浓度和温度、相对湿度、气压数据在数值上的差距都较大。因此采用温度、相对湿度、气压数据对二氧化碳浓度时间序列模拟时,需要对输入变量进行归一化处理。为了减少绝对尺度对模拟精度的影响,本研究对模型所需的所有变量都按式(1)进行归一化处理。

$$x_{\text{normal}} = \frac{x - x_{\text{mean}}}{x_{\text{std}}} \quad (1)$$

式中, x_{normal} 为经过归一化处理的变量序列, x 为原始观测的变量序列, x_{mean} 为相应变量序列的平均值, x_{std} 为相应变量序列的标准差。

2 研究方法

采用多种机器学习方法,分别以其他位置二氧化碳浓度或环境因子的组合作为输入,评价采用机器学习用于模拟黄土高原沟壑区地面空气二氧化碳浓度的潜力,并遴选出最优机器学习方法对问题数据进行插补,提高数据质量。

2.1 多层感知机

感知机最早由 Frank 提出用于解决分类问题^[27]。多个感知机连接即为多层感知机^[28] (Multilayer Perceptrons, MLP)。MLP 是一种具有较好非线性全局作用 and 高度并行能力,前向反馈的人工神经网络,可用于解决分类问题和回归问题^[29-30]。其基本结构由输入层、隐藏层和输出层构成^[31],如图 10 所示。一层隐藏层的 MLP 可用下式表示:

$$y = J_1(b^{(2)} + W^{(2)}(J_2(b^{(1)} + W^{(1)}x))) \quad (2)$$

式中: y 是输出的模拟值; J_1 和 J_2 是激活函数; $W^{(1)}$ 和 $b^{(1)}$ 是隐藏层的权值系数和偏置项; $W^{(2)}$ 和 $b^{(2)}$ 是输出层的权值系数和偏置项。

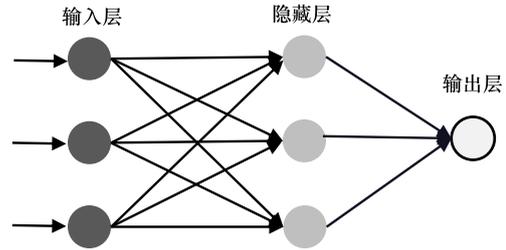


图 10 多层感知机的结构

Fig.10 The structure of MLP

2.2 长短期记忆神经网络

长短期记忆神经网络 (Long Short Term Memory, LSTM) 最早由 Hochreiter 和 Schmidhuber 提出^[32]。目前 LSTM 在水文模拟中应用十分广泛^[33],它主要通过设置遗忘门、输入门和输出门来解决传统 RNN 循环神经网络存在的长期记忆能力不足、梯度爆炸和梯度消失等问题^[34]。其常规的单元结构见图 11。它的运行过程可以用公式(3)表示:

$$\begin{aligned}
i_t &= \sigma(W_{hi}h_{t-1} + W_{xi}x_t + W_{ci}c_{t-1} + b_i) \\
f_t &= \sigma(W_{hf}h_{t-1} + W_{xf}x_t + W_{cf}c_{t-1} + b_f) \\
c_t &= f_t \otimes c_{t-1} + i_t \otimes \tanh(W_{hc}h_{t-1} + W_{xc}x_t + b_c) \\
o_t &= \sigma(W_{ho}h_{t-1} + W_{xo}x_t + W_{co}c_{t-1} + b_o) \\
h_t &= o_t \otimes \tanh(c_t) \\
y_t &= W_{hy}h_t + b_o
\end{aligned} \quad (3)$$

式中: x_t 和 y_t 是 LSTM 在 t 时刻的输入和输出; i_t , f_t , c_t 和 o_t 是在 t 时刻的输入门、遗忘门、记忆单元状态和输出

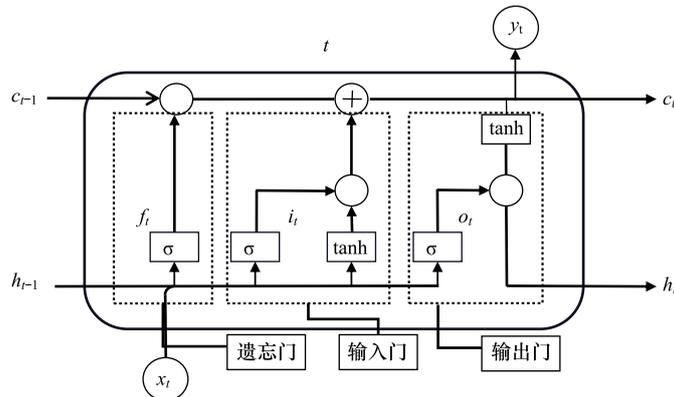


图 11 长短期记忆神经网络的单元结构

Fig.11 The structure of LSTM

门; W 和 b 是对应时刻和对应门的权重系数矩阵和偏置项; h_t 是在 t 时刻的递归输入; σ 是 sigmoid 激活函数; \tanh 是双曲正切切激活函数。

2.3 双向长短期记忆神经网络

双向长短期记忆神经网络 (Bidirectional Long Short-term Memory, Bi-LSTM) 是基于 LSTM 产生的。一般 LSTM 都是前向模拟预测, 而 Bi-LSTM 可以理解为一个前向 LSTM 和一个后向 LSTM 组合而成, 如图 12 所示, 预测结果由两个 LSTM 共同决定^[35]。双向长短期记忆可以更好地利用过去和未来的数据^[36]。

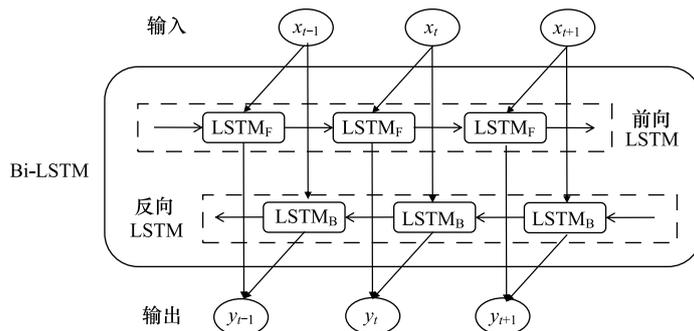


图 12 双向长短期记忆神经网络的结构

Fig.12 The structure of Bi-LSTM

2.4 随机森林

随机森林 (Random Forests, RF) 是由 Breiman 在 2001 基于分类和回归树与随机子空间等方法提出的一种机器学习算法^[37]。随机森林由多个决策树构成, 其中的每个回归树都根据数据子集和解释变量子集进行训练, 共同决定预测值^[38-39]。随机森林的构造过程如图 13 所示。随机森林因其具有较高的稳定性和特征鲁棒性, 能够有效的减少过拟合的风险^[40], 目前已广泛应用于随机分类和随机回归。

2.5 机器学习模型的参数设置

在 MLP, LSTM, Bi-LSTM 和 RF 的训练阶段都利用 Python 中 Sklearn 工具库的随机搜索和五折交叉验证确定参数。MLP 的超参数设置包括隐藏层数量, 神经元数量, 激活函数, 学习和遗忘率; LSTM 和 Bi-LSTM 的超参数设置包括神经元数量, 激活函数, 遗忘率。RF 的超参数设置包括树的数量, 节点可分得的最小样本数, 叶子结点含有的最少样本数, 树的深度等。由于 1 号传感器正常监测的二氧化碳浓度数据约占整个研究时段的 40%, 本研究中 4 种机器学习中训练集所占比例均设为 40%。

2.6 模型评价指标的选择

为了评价对二氧化碳浓度模拟的效果好坏, 本研究采用了决定系数 (Coefficient of determination, R^2), 平均绝对误差 (Mean absolute error, MAE), 均方根误差 (Root mean square error, RMSE), 以及 KGE 系数 (Kling-Gupta efficiency coefficient, KGE) 作为评价指标。4 个评价指标的计算公式如下:

$$R^2 = \left[\frac{\sum_1^N (\hat{y}_i - \bar{y}_i) (y_i - \bar{y})}{\sqrt{\sum_1^N (\hat{y}_i - \bar{y}_i)^2} \sqrt{\sum_1^N (y_i - \bar{y})^2}} \right]^2$$

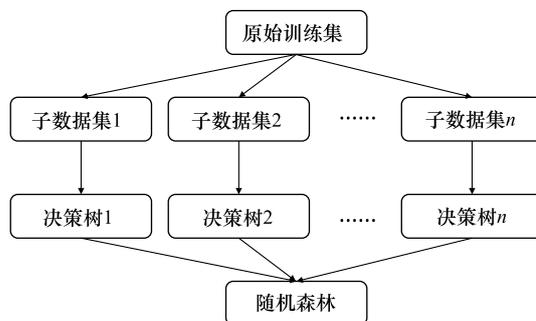


图 13 随机森林的构造过程

Fig.13 The construction process of RF

$$\begin{aligned} \text{MAE} &= \frac{\sum_1^N |y_i - \hat{y}_i|}{N} \\ \text{RMSE} &= \sqrt{\frac{\sum_1^N (y_i - \hat{y}_i)^2}{N}} \\ \text{KGE} &= 1 - \sqrt{(r - 1)^2 + \left(\frac{\mu_s}{\mu_0} - 1\right)^2 + \left(\frac{\sigma_s/\mu_s}{\sigma_0/\mu_0} - 1\right)^2} \end{aligned} \quad (4)$$

式中, \hat{y}_i 和 y_i 是模拟值和实测值, \bar{y}_i 是模拟值系列的均值, \bar{y} 是实测值系列的均值; N 是样本总数; μ_s, μ_0 是模拟值和实测值系列的均值, σ_s, σ_0 是模拟值和实测值系列的均方差, r 是模拟值和实测值系列之间的线性相关系数。四个评价指标中 MAE 和 RMSE 越小模拟效果越好, R^2 和 KGE 越接近 1 模拟效果越好。

3 结果与分析

3.1 采用 C_3 作为输入对 C_2 的模拟结果

对比 2 号传感器处的地面空气二氧化碳浓度 C_2 和 3 号传感器处的地面空气二氧化碳浓度 C_3 可知, 在同一观测系统中的不同位置的地面空气二氧化碳浓度存在一定的差异, 但是它们具有相同的变化规律。因此以 C_3 作为输入, 对 C_2 进行模拟。采用 MLP、LSTM、Bi-LSTM 和 RF 模拟的评价指标计算结果见表 2。测试集的评价指标计算结果表明, 以其他位置的地面空气二氧化碳浓度作为输入, MLP 和 RF 更为适合用于模拟, 其中 MLP 的性能较好, RF 的性能次之。地面空气二氧化碳浓度的日变化本就十分剧烈, LSTM 和 Bi-LSTM 的强记忆性反倒使得模拟精度远低于 MLP 和 RF。评价指标的对比显示, 以其他位置实测二氧化碳浓度作为输入, MLP 模拟 2 号传感器处的地面空气二氧化碳浓度 (C_2) 性能最好, 且其在测试集的误差仅为 3.8%, 模拟精度很高。

表 2 采用不同机器学习模型 (以 C_3 作为输入) 对 2 号传感器处地面空气二氧化碳浓度 (C_2) 模拟的评价指标

Table 2 The different machine learning models (C_3 as input) were used to evaluate the simulation of ground air carbon dioxide concentration at sensor 2 (C_2)

模型 Models	训练集 Training dataset				测试集 Testing dataset			
	MAE/ ($\mu\text{L/L}$)	RMSE/ ($\mu\text{L/L}$)	R^2	KGE	MAE/ ($\mu\text{L/L}$)	RMSE/ ($\mu\text{L/L}$)	R^2	KGE
多层感知机 MLP	6.650	9.429	0.938	0.940	13.060	15.767	0.905	0.804
长短期记忆 LSTM	22.572	28.772	0.407	0.429	26.115	32.337	0.606	0.464
双向长短期记忆 Bi-LSTM	19.890	25.534	0.533	0.618	21.346	26.504	0.735	0.625
随机森林 RF	4.794	6.507	0.971	0.974	13.671	16.935	0.891	0.801

MLP: 多层感知机 Multilayer perceptrons; LSTM: 长短期记忆 Long short term memory; Bi-LSTM: 双向长短期记忆 Bidirectional long short-term memory; RF: 随机森林 Random forests

不同机器学习模型的模拟值与实测值的散点图如图 14 所示, 不同机器学习模型的模拟值与 2 号传感器处的地面空气二氧化碳浓度实测值 C_2 的对比图如图 15 所示。从散点图来看, 在测试集和训练集, 不同机器学习模型所得模拟值与实测值的皮尔逊相关系数 r 均大于 0.7, 模拟效果是可接受的。对比模拟值和实测值来看, 以其他位置实测二氧化碳浓度作为输入, 可以较好的模拟出地面空气二氧化碳浓度的整体变化过程。MLP 还可以较好的模拟地面空气二氧化碳浓度高值和低值, 明显优于 LSTM、Bi-LSTM 和 RF。

3.2 用于模拟 C_2 的输入变量组合

一般情况下输入参数的选择应该基于较简单的关系实现高准确性的模拟。因此本研究绘制了目标变量 (C_2) 和各计算了各变量之间的散点图, 并计算了他们之间的皮尔逊相关系数, 散点图如图 16, 相关系数的计算结果见表 3。从散点图来看, 空气温度、土壤温度和二氧化碳浓度的散点分布较集中, 其他两个输入变量

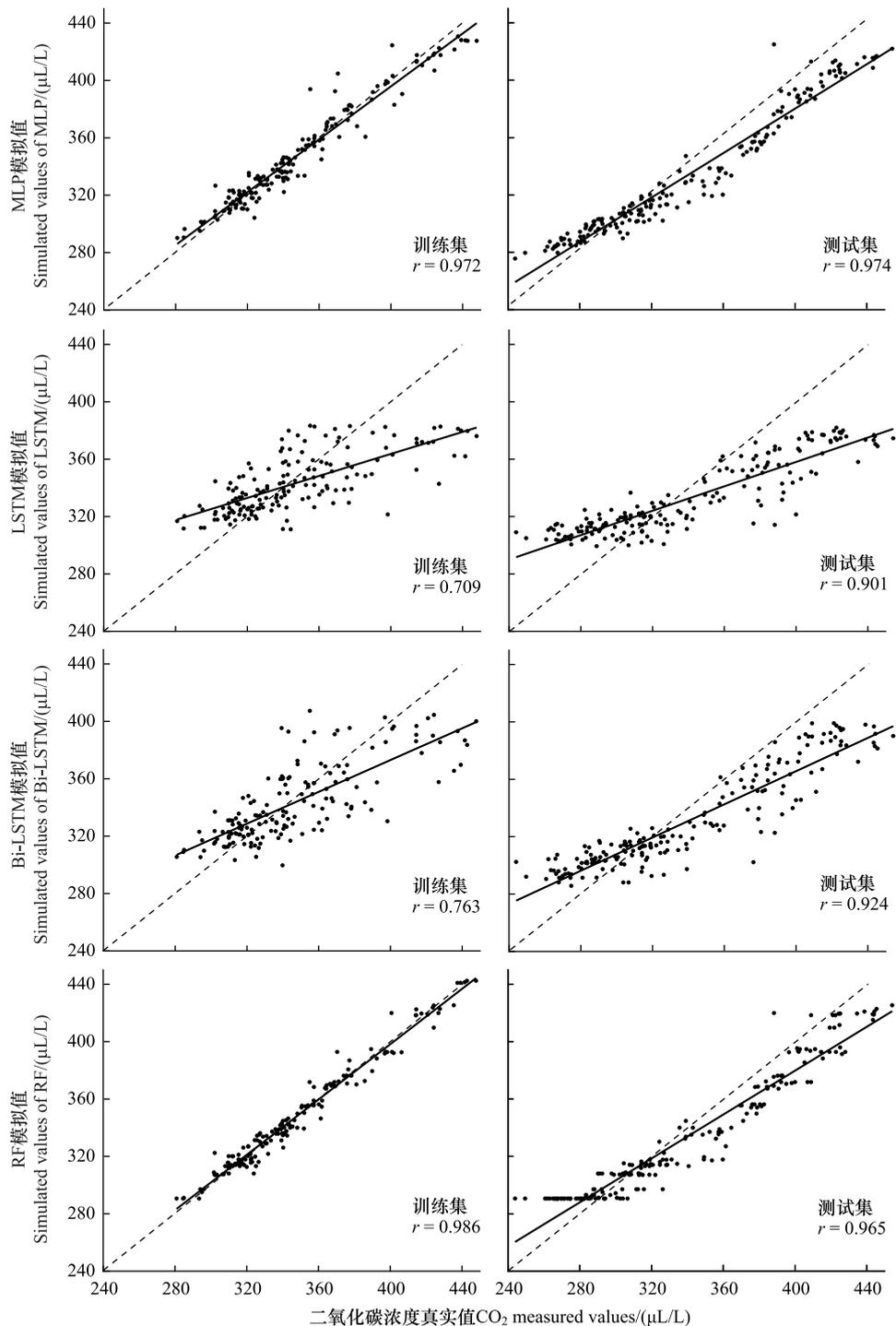


图 14 不同机器学习模型模拟值(以 C_3 作为输入)与 C_2 实测值散点图

Fig.14 The scatter plot of simulated values of different machine learning models (C_3 as input) and measured values of C_2

(相对湿度和气压)和二氧化碳浓度的散点分布较分散。从相关性的计算结果来看,土壤温度和二氧化碳浓度的线性相关系数为 -0.870 ,空气温度和二氧化碳浓度的线性相关系数为 -0.806 ,相对湿度和二氧化碳浓度的线性相关系数为 -0.671 ,在研究时段中二氧化碳浓度和土壤温度、空气温度都呈明显的负相关,二氧化碳浓度高时土壤温度、空气温度都较低,二氧化碳浓度低时土壤温度、空气温度都较高。气压和二氧化碳浓度的相关系数最低。

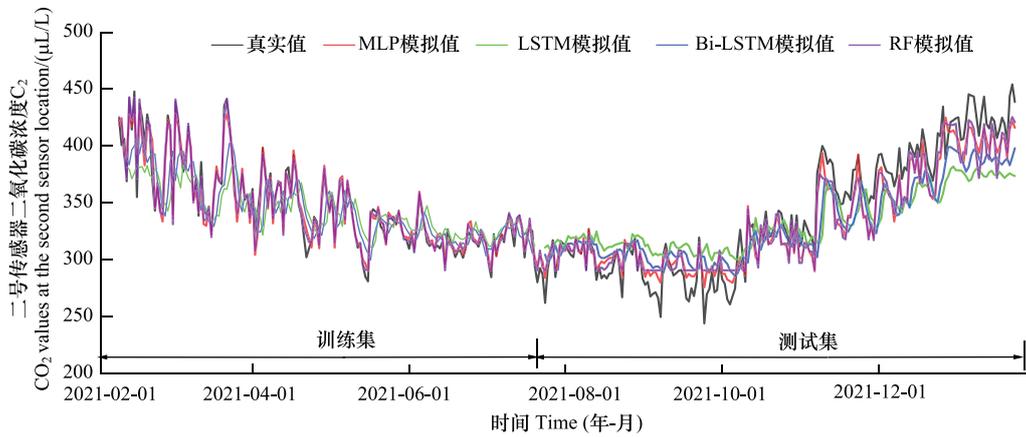


图 15 不同机器学习模型模拟值(以 C_3 作为输入)与 C_2 实测值对比图

Fig.15 The comparison of simulated values of different machine learning models (C_3 as input) and measured values of C_2

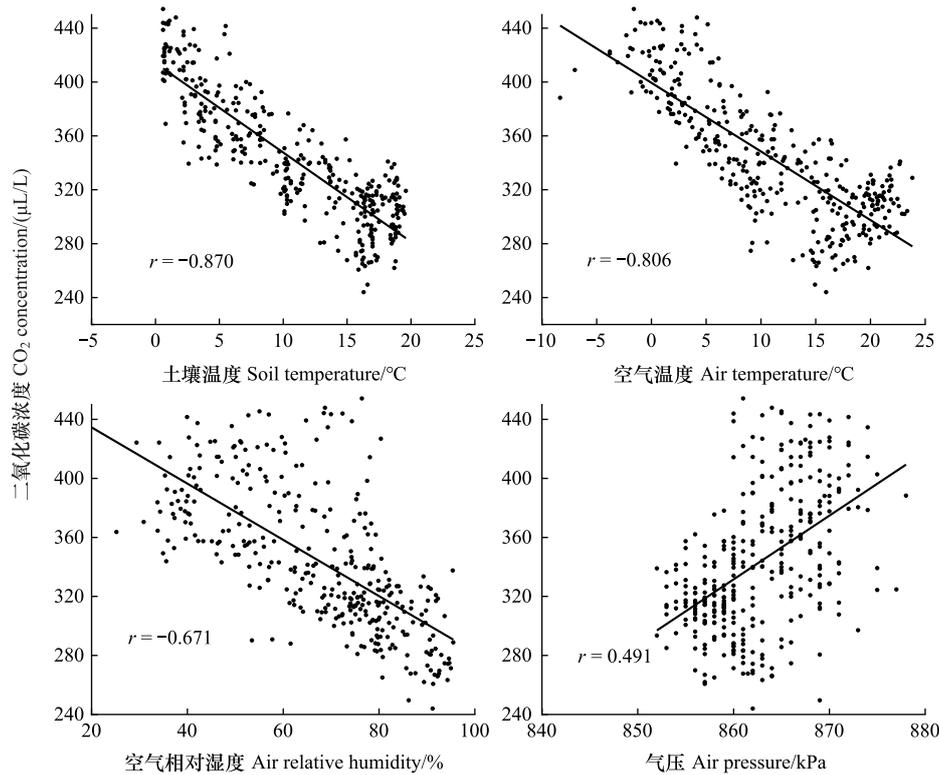


图 16 输入变量和 2 号传感器处地面空气二氧化碳浓度(C_2)的散点图

Fig.16 The scatter plot of input variables and ground air carbon dioxide concentration at sensor 2 (C_2)

表 3 输入变量和 2 号传感器处地面空气二氧化碳浓度(C_2)的皮尔逊相关系数

Table 3 The pearson correlation coefficient between input variables and ground air carbon dioxide concentration at sensor 2 (C_2)

数据 Data	2 号处二氧化碳浓度 CO ₂ value at the second sensor location/(µL/L)
20cm 深度土壤温度 Soil temperature at depth of 20cm /°C	-0.870
空气温度 Air temperature/°C	-0.806
空气相对湿度 Air relative humidity/%	-0.671
气压 Air pressure /kPa	0.491

根据相关性分析的结果,确定 6 种用于二氧化碳浓度模拟的输入组合,不同输入组合的环境因子组成见表 4。在设置的 6 种输入组合中,输入组合 1 由土壤温度单独作为输入;输入组合 2 由空气温度单独作为输入;输入组合 3 由土壤温度和空气温度共同作为输入;输入组合 4 由土壤温度、空气温度和相对湿度共同作为输入;输入组合 5 由土壤温度、空气温度和气压共同作为输入;输入组合 6 由全部环境因子(土壤温度、空气温度、相对湿度和气压)共同作为输入。

表 4 用于模拟 2 号传感器处地面空气二氧化碳浓度(C_2)的输入组合

Table 4 The input combinations for simulating ground air carbon dioxide concentration at sensor 2 (C_2)

编号 Number	输入组合 Input combination	编号 Number	输入组合 Input combination
1 NO.1	T_{s20cm}	4 NO.4	T_{s20cm} T_{air} RH
2 NO.2	T_{air}	5 NO.5	T_{s20cm} T_{air} P
3 NO.3	T_{s20cm} T_{air}	6 NO.6	T_{s20cm} T_{air} RH P

: T_{s20cm} 表示 20cm 深度的土壤温度, T_{air} 表示空气温度, RH 表示空气相对湿度, P 表示气压

3.3 评价采用环境因子作为输入对 C_2 的模拟结果

以环境因子构造的多个输入组合作为输入,对 2 号传感器处的地面空气二氧化碳浓度(C_2)采用 MLP、LSTM、Bi-LSTM 和 RF 模拟的评价指标计算结果见表 5。评价指标计算结果表明,虽然空气温度、土壤温度、气压和相对湿度都会影响二氧化碳浓度,但是不同的模型的性能有所差异,它们的最优输入组合也不尽相同。综合多个评价指标在测试集的结果来看,MLP、Bi-LSTM 和 RF 的最优输入组合均为输入组合 1,LSTM 的最优输入组合为输入组合 4。值得强调的是,以土壤温度(输入组合 1)作为输入时,4 种机器学习方法模拟的平均绝对误差 MAE 最小。

表 5 采用不同机器学习模型(以环境因子组合作为输入)对 2 号传感器处地面空气二氧化碳浓度(C_2)模拟的评价指标

Table 5 The evaluation indexes of ground air carbon dioxide concentration at sensor 2 (C_2) simulation using different machine learning models (the combination of environmental factors as input)

模型 Model	输入 Input	训练集 Training dataset				测试集 Testing dataset			
		MAE/ ($\mu\text{L/L}$)	RMSE/ ($\mu\text{L/L}$)	R^2	KGE	MAE/ ($\mu\text{L/L}$)	RMSE/ ($\mu\text{L/L}$)	R^2	KGE
多层感知机 MLP	1	17.383	22.036	0.698	0.747	22.191	26.498	0.760	0.592
	2	18.498	23.362	0.661	0.721	28.668	34.507	0.592	0.535
	3	17.258	21.779	0.705	0.760	23.624	27.856	0.734	0.584
	4	16.376	21.178	0.721	0.657	26.559	31.072	0.669	0.534
	5	17.059	21.482	0.713	0.750	28.252	33.250	0.621	0.517
	6	14.819	19.056	0.774	0.821	26.696	31.093	0.669	0.632
长短期记忆 LSTM	1	19.543	25.710	0.526	0.503	24.100	29.758	0.666	0.532
	2	18.985	24.913	0.555	0.534	27.448	33.094	0.587	0.538
	3	18.629	24.055	0.585	0.572	24.809	30.011	0.660	0.574
	4	18.506	24.035	0.586	0.573	24.010	29.124	0.680	0.591
	5	18.684	24.518	0.569	0.560	27.240	32.785	0.595	0.536
	6	18.795	24.823	0.559	0.548	26.798	32.239	0.608	0.543
双向长短期记忆 Bi-LSTM	1	18.420	23.597	0.601	0.636	20.574	25.300	0.759	0.665
	2	17.189	22.501	0.637	0.650	26.274	31.841	0.618	0.710
	3	17.266	22.325	0.643	0.661	22.673	27.941	0.706	0.708
	4	16.856	21.783	0.660	0.681	21.936	26.799	0.729	0.743
	5	16.741	21.752	0.661	0.686	24.786	30.420	0.651	0.691
	6	16.597	21.531	0.668	0.687	23.849	28.894	0.685	0.710
随机森林 RF	1	14.289	19.109	0.747	0.775	19.818	24.697	0.767	0.708
	2	13.427	18.266	0.768	0.797	27.482	34.061	0.557	0.696
	3	13.701	18.515	0.762	0.775	20.701	26.497	0.732	0.741
	4	9.096	12.233	0.896	0.867	21.729	26.772	0.726	0.798
	5	13.101	17.033	0.799	0.787	27.694	34.179	0.554	0.681
	6	8.634	11.852	0.903	0.878	26.646	31.478	0.622	0.791

以最优输入组合作为输入,不同机器学习模型的模拟值与实测值的散点图如图 17 所示,不同机器学习模型的模拟值与 C_2 实测值的对比图如图 18 所示。从散点图可知,4 种机器学习方法在训练集中和测试集中模拟值与实测值的皮尔逊相关系数都较高,其中在训练集的最小值为 0.786,在测试集的最小值为 0.896,均大于 0.7,说明模拟效果是可接受的。从模拟值与实测值的对比图来看,4 种机器学习模型都可以较好的模拟出二

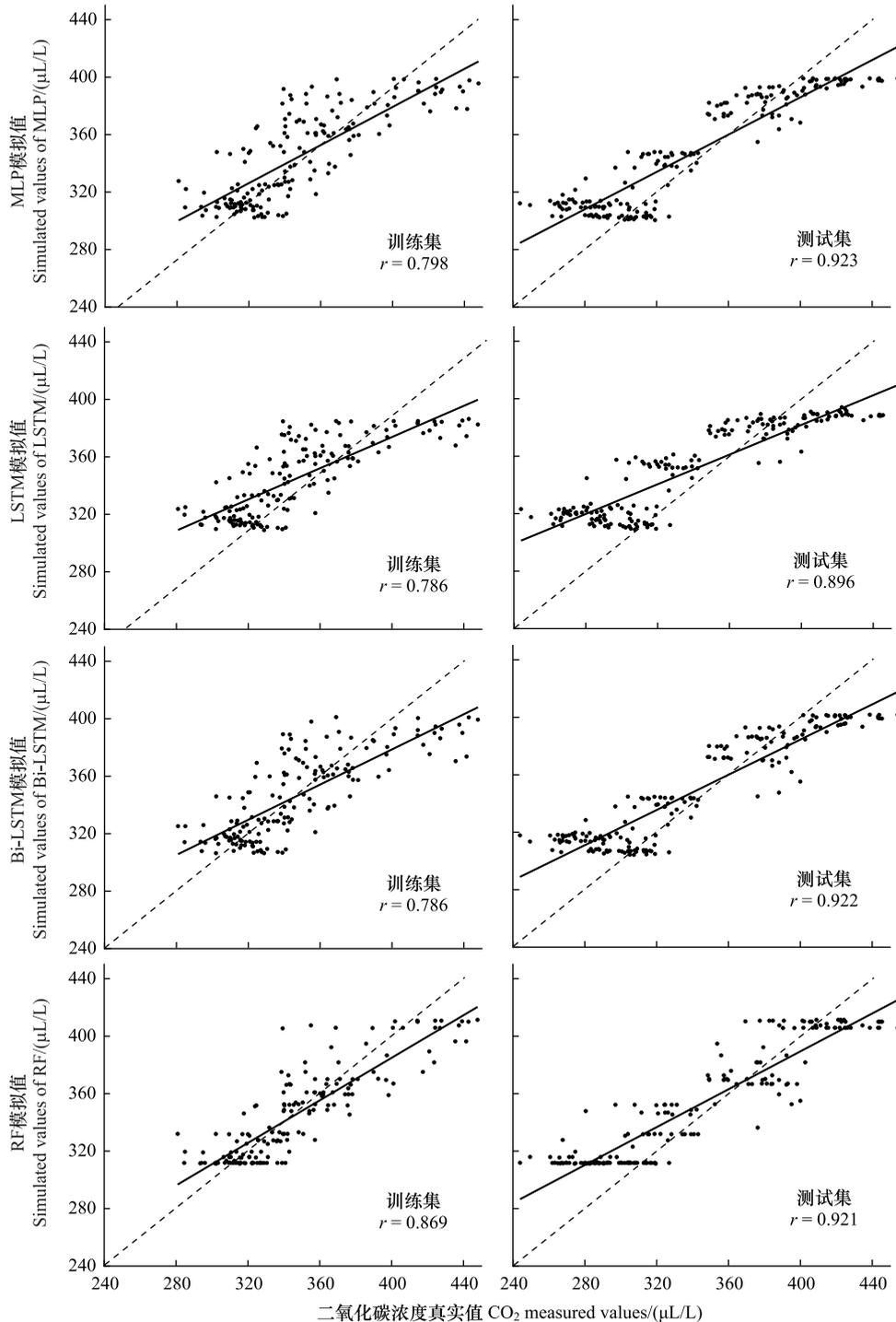


图 17 不同机器学习模型模拟值(以环境因子组合作为输入)与 C_2 实测值散点图

Fig.17 The scatter plot of simulated values of different machine learning models (the combination of environmental factors as input) and measured values of C_2

氧化碳浓度的整体变化过程,但是在以环境因子作为输入时,它们都无法模拟出日尺度地面空气二氧化碳浓度剧烈的变化。4 种机器学习方法中 LSTM 对冬季二氧化碳浓度的低估最为突出,主要原因是由于 LSTM 的最优输入组合与 MLP 和 Bi-LSTM 的有所不同。LSTM 的最优输入组合是以土壤温度、空气温度和相对湿度共同作为输入,其中空气温度比土壤温度的变化要剧烈的多,空气温度在冬季又大部分时间都要低于土壤温度,再加之相对湿度不具有明显的变化规律,共同导致了 LSTM 对地面空气二氧化碳浓度的低估。

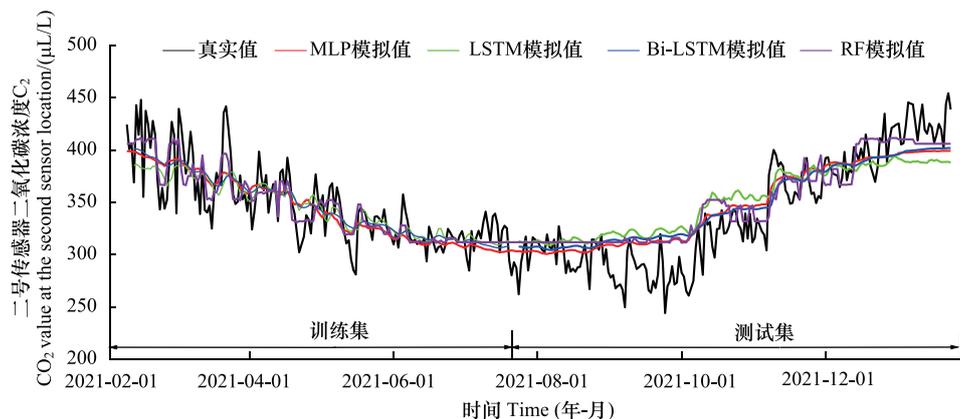


图 18 不同机器学习模型模拟值(以环境因子组合作为输入)与 C_2 实测值对比图

Fig.18 The comparison of simulated values of different machine learning models (the combination of environmental factors as input) and measured values of C_2

为了更加直观的对比不同机器学习模型的性能差异,绘制了不同机器学习模型(最优输入组合)评价指标对比图,如图 19 所示。从对比图可知,在训练集和测试集上,RF 的所有评价指标都是最优的,其中测试集的 MAE 为 $14.289\mu\text{L/L}$, RMSE 为 $19.109\mu\text{L/L}$, R^2 为 0.747, KGE 为 0.775;训练集的 MAE 为 $19.818\mu\text{L/L}$, RMSE 为 $24.697\mu\text{L/L}$, R^2 为 0.767, KGE 为 0.708。评价指标的对比显示,以环境因子作为输入,RF 模拟二号传感器处的地面空气二氧化碳浓度(C_2)性能最好,且其在测试集的误差为 6.3%,模拟精度较高。

3.4 采用机器学习方法模拟 C_1

1 号传感器处的地面空气二氧化碳浓度 C_1 数据部分存在明显低估,现采用多种机器学习方法,以 3 号传感器处的地面空气二氧化碳浓度 C_3 和环境因子作为输入,对进行 C_1 模拟。以 3 号传感器处的地面空气二氧化碳浓度 C_3 作为输入时,机器学习方法采用 MLP;以环境因子(土壤温度)作为输入时,机器学习方法采用 RF。模拟值与实测值的对比图如图 20 所示,在测试集上的模拟值不再有明显的低估情况。插补数据无法通过计算评价指标判断其数据质量,因此计算了插补后数据的简要统计参数见表 6。与实测的不同传感器实测二氧化碳浓度的简要统计参数来看,插补后的 1 号传感器和 2 号、3 号传感器处的地面空气二氧化碳浓度的最大值、最小值、均值和方差同样相差不大,插补后的数据质量有了明显提升。

表 6 1 号传感器处地面空气二氧化碳浓度和环境因子的简要统计参数

Table 6 The brief statistical parameters of ground air carbon dioxide concentration at sensor 1 (C_1) and environmental factors

数据 Data	统计参数 Statistical parameters						
	x_{mean}	x_{max}	x_{min}	x_{std}	C_v	C_s	C_k
多层感知机模拟值 MLP simulation value	363.21	457.78	295.79	42.38	0.12	0.57	0.80
随机森林模拟值 RF simulation value	368.74	437.87	319.42	42.48	0.12	0.23	1.38

4 结论

本文采用陕西省淳化生态水文实验基地二氧化碳浓度观测系统实测的二氧化碳浓度、空气温度、土壤温

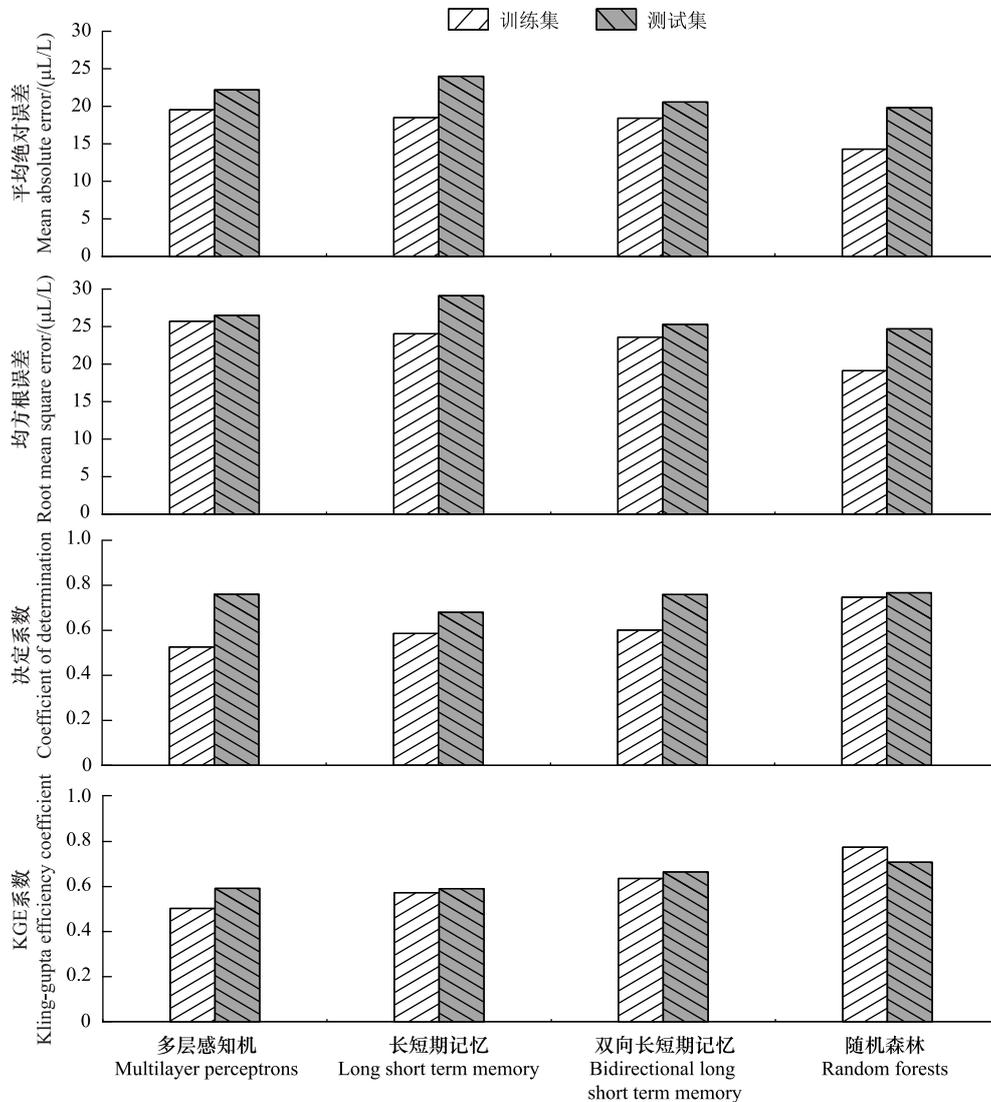


图 19 不同机器学习模型(最优输入组合)评价指标对比图

Fig.19 The comparison of evaluation indexes of different machine learning models (optimal input combination)

度相对湿度和气压数据,以环境因子的组合和其他位置二氧化碳浓度作为输入,对比研究了不同机器学习方法对二氧化碳浓度模拟的性能差异,评价了不同输入组合对二氧化碳浓度模拟的影响,验证了采用环境因子对二氧化碳浓度模拟的可行性,确定了二氧化碳浓度模拟的最优机器学习模型,并将最优机器学习方法用于插补数据。研究结果一方面可提高野外监测数据的质量,另一方面有助于更好了解黄土高原沟壑区地面空气二氧化碳浓度的变化过程,从而服务于变化环境下流域生态水文过程的研究,并对黄土高原水土保持、植被恢复、节水固碳等工作的开展具有重要的应用价值和科学意义。

研究发现,(1)采用机器学习方法,以系统内其他传感器实测的二氧化碳浓度作为输入模拟地面空气二氧化碳浓度是可行的,4种机器学习方法中MLP可以较好的模拟地面空气二氧化碳浓度高值和低值,明显优于LSTM、Bi-LSTM和RF。(2)采用机器学习方法,以环境因子的组合作为输入模拟地面空气二氧化碳浓度是可行的,4种机器学习方法中RF模拟地面空气二氧化碳浓度性能最好,但需要注意的是,在以环境因子作为输入时,它们都无法模拟出日尺度地面空气二氧化碳浓度剧烈的变化。(3)采用机器学习方法,以环境因子和同类型数据用于插补地面空气二氧化碳浓度实测数据都是可行的。其中以同类型的数据作为输入用于插补数据的效果要明显优于以环境因子作为输入,以同类型数据作为输入可以较好地模拟出地面空气二氧化碳

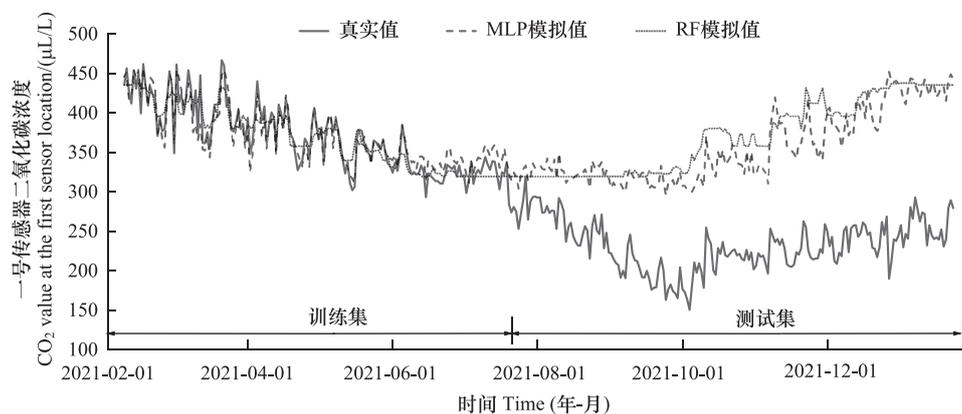


图 20 不同机器学习模型模拟值与 C_1 实测值对比图

Fig.20 The comparison of simulated values of different machine learning models with measured values of C_1

浓度的日尺度变化。

观测基地位于黄土高原,黄土高原不同范围的生态因素存在差异,不同的植被覆盖状态对碳循环有不同的影响^[41]。三个观测点的植被覆盖度有所差异,其中 1 号观测点处植被覆盖度最低,2 号观测点处植被覆盖度中等,3 号观测点处植被覆盖度最高。这是在观测场能做到植被覆盖度方面的生态因素差异,从某种程度上可以反应黄土高原沟壑区的不同植被覆盖度。基于设备布设方案,植被作为近地面二氧化碳浓度的关键影响因素^[42]。结合研究区 2021 年雨量数据和土壤湿度来看,在 8 月该区域降雨量仅 40mm 左右,土壤湿度不足 20%,而该地区 9 月份降雨量为 230mm 左右,土壤湿度在 25%–35%。9 月份植被生长更茂盛,9 月地面二氧化碳浓度低于 8 月。地面二氧化碳浓度的低值往往出现在降雨时期附近,随着降雨量的增加,土壤呼吸受到抑制,释放的 CO_2 有所减少^[43]。地面二氧化碳浓度与空气温度、土壤温度、降雨量、土壤湿度、风速和水汽压等都有关系。对于绝大多数机器学习方法,采用与目标变量具有相同变化趋势的环境因子作为输入,模拟精度会更高。但对于日尺度变化相对剧烈的地面二氧化碳浓度,过多的环境因子作为输入反而使得模拟精度有所降低。

实测数据常由于设备故障等问题而有一定的缺失,本研究探究了不同输入数据组合对二氧化碳浓度模拟精度的影响,验证了机器学习方法可用于插补二氧化碳浓度实测数据。应该注意的是,对于插补后的数据质量还需要进一步定量分析。本研究仅使用了淳化生态水文实验基地的数据,是机器学习方法在黄土高原沟壑区模拟二氧化碳浓度的一次尝试。虽然不同观测点的植被覆盖度有所差异,但相对于整个黄土高原沟壑区的空间变异性还是不够,未来需要使用更多站点数据对其潜在推广性进行进一步验证。

参考文献 (References):

- [1] 刘维,曹龙. 气候系统对 CO_2 强迫和太阳辐射强迫在不同时间尺度的响应[J]. 气候变化研究进展,2017,13(03):231-242.
- [2] 李家琪,姜振蛟,戴鑫,汪钊,武丽文,谢月清. 河源区土壤与河流二氧化碳浓度变化特征及相关性[J]. 中国环境科学,2023,43(12):6667-6676.
- [3] 许静静,龚威,张劲,张豪伟,马昕,韩舸. 2009-2020 年基于 GOSAT 卫星的全球中低纬二氧化碳柱浓度数据集[J]. 中国科学数据(中英文网络版),2023,8(03):468-477.
- [4] 张远辉,陈立奇. 南沙珊瑚礁对大气 CO_2 含量上升的响应[J]. 台湾海峡,2006,(01):68-76.
- [5] 陈昆,沈竞,陈鲍发. 景德镇地区大气 CO_2 浓度变化特征[J]. 气象与减灾研究,2019,42(02):113-118.
- [6] 蔡寅潮,韩炜,管文轲,费兵强,马霄华,党亚玲. 乌鲁木齐夏季近地面 CO_2 浓度空间分布特征[J]. 环境科学与技术,2018,41(6):22-27.
- [7] 高韵秋,刘寿东,胡凝,王淑敏,邓力琛,于洲,张圳,李旭辉. 南京夏季城市冠层大气 CO_2 浓度时空分布规律的观测[J]. 环境科学,2015,36(7):2367-2373.
- [8] Crawford J T, Stanley E H, Dornblaser M M, Striegl R G. CO_2 time series patterns in contrasting headwater streams of North America[J]. Aquatic Sciences,2017,79(3):1-14.

- [9] 相如. 基于 SRCNN 的卫星 CO₂ 观测数据超分辨率重建研究[D]. 中国矿业大学, 2022.
- [10] Yokota T, Yoshida Y, Eguchi N, Ota Y, Tanaka T, Watanabe H, Maksyutov S. Global concentrations of CO₂ and CH₄ retrieved from GOSAT: first preliminary results[J]. *Sola*, 2009, 5: 160-163.
- [11] Crisp D, Pollock H R, Rosenberg R, Wunch D. The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products[J]. *Atmospheric Measurement Techniques*, 2017, 10(1): 1-45.
- [12] 吴迪, 杨爱玲, 周源. 基于 GOSAT 卫星监测二氧化碳浓度时空变化[J]. *测绘与空间地理信息*, 2023, 46(S1): 1-3.
- [13] Xu L, Baldocchi D D. Seasonal variation in carbon dioxide exchange over a Mediterranean annual grassland in California[J]. *Agricultural and Forest Meteorology*, 2004, 123(1-2): 79-96.
- [14] 张东秋, 石培礼, 张宪洲. 土壤呼吸主要影响因素的研究进展[J]. *地球科学进展*, 2005, (07): 778-785.
- [15] 原樱其, 朱仁超, 杨宇, 余爱华. 不同生态系统土壤呼吸影响因素研究进展[J]. *世界林业研究*, 2023, 36(04): 15-21.
- [16] Cao S J, Ding J, Ren C. Sensor deployment strategy using cluster analysis of Fuzzy C-Means Algorithm: Towards online control of indoor environment's safety and health[J]. *Sustainable Cities and Society*, 2020.
- [17] 尹航, 吕佳威, 陈耀聪, 岑红蕾, 李景彬, 刘双印. 基于 LightGBM-SSA-ELM 的新疆羊舍 CO₂ 浓度预测[J]. *农业机械学报*, 2022, 53(1): 261-270.
- [18] 付子骏, 吴永明, 徐计. TD-LSTM-S 模型在二氧化碳浓度预测中的应用[J]. *重庆理工大学学报(自然科学)*, 2023, 37(4): 192-199
- [19] Fu B. Soil erosion and its control in the Loess Plateau of China[J]. *Soil Use & Management*, 2010, 5(2): 76-82.
- [20] 雷未, 王建, 吉同元, 李鹏飞. 基于深度学习框架的长序列大坝监测缺失数据插补模型[J]. *水利水电科技进展*, 2023, 43(06): 82-88.
- [21] Guo F, Liu D, Mo S, Huang Q, Ma L, Xie S, Deng W, Ming G, Fan J. Estimation of daily evapotranspiration in gully area scrub ecosystems on Loess Plateau of China based on multisource observation data[J]. *Ecological Indicators*, 2023, 154: 110671.
- [22] Zhang K, Liu D, Liu H, Lei M, Guo F, Xie S, Meng X, Huang Q. Energy flux observation in a shrub ecosystem of a Gully Region of the Chinese Loess Plateau[J]. *Ecohydrology & Hydrobiology*, 2022, 22(2): 323-336.
- [23] 张奎月, 刘登峰, 刘慧, 赵笑雨, 郭凤年, 孟宪萌, 黄强. 黄土高原灌丛生态系统土壤呼吸特征及其影响因素[J]. *人民珠江*, 2022, 43(04): 83-94.
- [24] 高海东, 庞国伟, 李占斌, 程圣东. 黄土高原植被恢复潜力研究[J]. *地理学报*, 2017, 72(05): 863-874.
- [25] 张奎月. 基于涡度相关的黄土高原沟壑区灌丛的水碳通量研究[D]. 西安理工大学, 2022.
- [26] Tsai F, Philpot W. Derivative analysis of hyperspectral data[J]. *Remote Sensing of Environment*, 1998, 66(1): 41-51.
- [27] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain[J]. *Psychological Review*, 1958, 65: 386-408.
- [28] 冯斌. 基于 MLP 的铝合金凸凹成形减薄率预测[D]. 湖南大学, 2020.
- [29] Murtagh F. Multilayer perceptrons for classification and regression[J]. *Neurocomputing*, 1991, 2(5-6): 183-197.
- [30] 程连虎. 基于机器学习的复杂场景图像质量评价与目标识别算法研究[D]. 南昌大学, 2023.
- [31] Sanzana M R, Abdulrazic M O M, Wong J Y, Maul T, Yip C. Charging water load prediction for a thermal-energy-storage air-conditioner of a commercial building with a multilayer perceptron[J]. *Journal of Building Engineering*, 2023, 75: 107016.
- [32] Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [33] Rahmani F, Shen C, Oliver S, Lawson K, Appling A. Deep learning approaches for improving prediction of daily stream temperature in data-scarce, unmonitored, and dammed basins[J]. *Hydrological Processes*, 2021, 11.
- [34] Yu B, Liu J, Lyu T, Li Z, Wang M, Yang. A new detection method to assess the influence of human activities and climate change of CO₂ emissions in coal field[J]. *Ecological Indicators*, 2022, 143: 109417.
- [35] Hu X, Yuan S, Xu F, Leng Y, Yuan K, Yuan Q. Scalp EEG classification using deep Bi-LSTM network for seizure detection[J]. *Computers in Biology and Medicine*, 2020, 124: 103919.
- [36] Zhou P, Shi W, Tian J, Qi Z, Li B, Hao H, Xu B. Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification[J]. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 2016.
- [37] Ho T K. The random subspace method for constructing decision forests[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1998, 20(8): 832-844.
- [38] 方匡南, 吴益彬, 朱建平, 谢邦昌. 随机森林方法研究综述[J]. *统计与信息论坛*, 2011, 26(03): 32-38.
- [39] Isles P D F. A random forest approach to improve estimates of tributary nutrient loading[J]. *Water Research*, 2024, 248.
- [40] Han T, Jiang D, Zhao Q, Wang L, Yin K. Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery[J]. *Transactions of the Institute of Measurement and Control*, 2017.
- [41] 黑哲. 黄土高原植被覆盖变化过程及评价方法差异性研究[D]. 中国科学院教育部水土保持与生态环境研究中心, 2020.
- [42] 柴华, 钟尚志, 崔海莹, 李杰, 孙伟. 植物呼吸释放 CO₂ 碳同位素变化研究进展[J]. *生态学报*, 2018, 38(08): 2616-2624.
- [43] 李新鸽, 韩广轩, 朱连奇, 陈超男. 降雨引起的干湿交替对土壤呼吸的影响: 进展与展望[J]. *生态学杂志*, 2019, 38(02): 567-575.