

改进 Informer 模型的苜蓿土壤湿度预测方法

王静, 刘瑞*, 杨松涛, 葛永琪
(宁夏大学信息工程学院, 宁夏银川 750021)

摘要:精准的苜蓿土壤湿度预测对于提高水资源利用率和降低智慧农业投入成本至关重要。针对传统土壤湿度预测方法在实际应用中存在预测周期短、精度低以及时空预测不足等问题,提出了一种融合快速傅里叶变换的 Informer 时空预测方法(Fast Fourier Transform and Spatio Temporal-Informer, FFT-ST-Informer)。首先,在传统 Informer 模型基础上添加了独立的时空嵌入层,从而捕获各个变量之间复杂的时空相关性。然后,根据土壤墒情与环境因素的相关性分析结果,选择降雨、灌溉量为关键环境因素,并使用快速傅里叶变换,通过提取某一周期具有先验的数据序列的频谱来表示其频域特征放入模型。此外,该模型中的 ProbSparse 自注意机制可以集中提取时空数据的重要上下文信息。FFT-ST-Informer 模型使用来自宁夏引黄灌区自采的气象和土壤数据作为输入数据。实验结果表明,FFT-ST-Informer 模型性能明显优于传统模型,比 LSTM 模型在平均绝对误差(MAE)、均方根误差(RMSE)、相关系数(R^2)等评价指标上,分别提高了56.9%,64.4%,0.12%。

关键词:苜蓿土壤湿度预测;快速傅里叶变换;空间嵌入层;ProbSparse 自注意机制;Informer 模型

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2024)06-0171-07

doi:10.20165/j.cnki.ISSN1673-629X.2024.0069

Improved Informer Model for Alfalfa Soil Moisture Prediction

WANG Jing, LIU Rui*, YANG Song-tao, GE Yong-qi

(School of Information Engineering, Ningxia University, Yinchuan 750021, China)

Abstract: Accurate prediction of alfalfa soil moisture is crucial for improving water resource utilization and reducing smart agriculture investment costs. In view of the issues of short prediction cycle, low accuracy, and insufficient spatio-temporal prediction in traditional soil moisture prediction methods, we propose a Spatio-Temporal prediction method called FFT-ST-Informer that integrates Fast Fourier Transform. Firstly, an independent spatio-temporal embedding layer is added to the traditional Informer model to capture the complex spatio-temporal correlations between variables. Then, based on the correlation analysis between soil moisture conditions and environmental factors, rainfall and irrigation volume are selected as key environmental factors. The FFT is utilized to extract the frequency domain characteristics of data sequences with prior knowledge of a certain period, which are then incorporated into the model. In addition, the ProbSparse self-attention mechanism in this model can effectively extract important contextual information from spatio-temporal data. The FFT-ST-Informer model uses meteorological and soil data collected from the self-drained Yellow River irrigation area in Ningxia as input. Experimental results show that the performance of the FFT-ST-Informer model is significantly better than that of traditional models, with improvements of 56.9%, 64.4%, and 0.12% in terms of mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2) evaluation metrics compared to the LSTM model.

Key words: alfalfa soil moisture prediction; fast Fourier transform; spatial embedding layer; ProbSparse self-attention mechanism; Informer model

0 引言

土壤湿度是指土壤中的含水量,土壤含水量作为农业生产的重要参数指标^[1],它的时空分布直接影响农作物的生长状况,快速准确地掌握土壤水分动态变

化规律,不仅对农业生产至关重要,也缓解了水资源利用率低的问题,从而制定合理的田间节水灌溉方案^[2]。土壤湿度预测,即通过采集的土壤参数数据,使用相应的模型、算法进行分析,确定未来一段时间内的土壤含

收稿日期:2023-08-18

修回日期:2023-12-21

基金项目:国家自然科学基金地区科学基金项目(62162052);宁夏回族自治区重点研发计划(2021BEB04016,2022BDE03007);宁夏自然科学基金(2021AAC03041,2022AAC03004)

作者简介:王静(1998-),女,硕士,研究方向为深度学习应用、智慧农业;通信作者:刘瑞(1983-),女,博士,副教授,研究方向为智能决策支持系统。

水量变化,对进行精准灌溉预报有重要意义^[3]。

目前,有关气象站台和土壤质量监测站在内的监测工具在各地广泛部署。由于该设备收集到的大量气象数据及土壤相关数据为土壤湿度预测和气象研究提供大量的数据支持,因此,通过数据驱动方法构建的土壤湿度预测模型越来越受到学者的关注。传统的土壤湿度预测方法主要包括经验公式法^[4]、水量平衡法^[5]、消退指数法^[6]、线性回归法^[7]等。然而,这些模型往往过于简化土壤湿度不同监测站点之间的空间动态变化关系,无法充分提取空间的复杂相关性。近年来,神经网络法^[8]被越来越多地应用到土壤湿度预测上,尤其是以 LSTM^[9] (Long Short - Term Memory, LSTM) 为代表的神经网络在处理上常用在不同时间跨度的土壤湿度预测,并且具有良好的自适应性和鲁棒性。肖天云等人^[10]选取贝叶斯优化算法对 LSTM 模型进行超参数调优,发现 BOA (Bayesian Optimization Algorithm) 迭代次数较少且考虑之前的参数信息,效率上 BOA 对 LSTM 模型有很大提升。在 LSTM 基础上,王学智提出基于编码器-解码器的长短期记忆网络,以及融合迁移学习的土壤湿度预测时空模型,改善了深度学习模型预测土壤湿度的性能^[11]。考虑到各种因素会影响到土壤含水量,在构建预测模型时需要考虑数据的长期时间依赖性和动态空间相关性。例如,结合卷积神经网络和长短期记忆网络,对输入数据进行卷积操作,从而提取空间特征,来捕捉时空序列数据的动态变化和长期依赖性^[12-13]。上述模型存在卷积滤波器的空间局限性,LSTM 网络的预测周期短的问题。一些学者随即提出加入注意力机制来解决这些问题。李清亮等人^[14]针对深度学习方法过度参数化的黑匣子性质以及对物理知识和可解释性的忽视,提出了一个注意感知的 LSTM(ILSTM_Soil),用于土壤湿度和土壤温度预测,提高了土壤墒情的预测精

度和效率。并且,利用大量的时间序列对过去行为的数据进行长期预测,即长序列时间序列预测^[15] (Long Sequence Time-series Forecasting, LSTF)。但与此同时,注意力机制固有特性限制了其对长时空序列的预测能力。与 LSTM 模型相比,Transformer^[16]模型在捕获远程依赖关系方面表现出了优越的性能。

然而,针对 Transformer 在长时间序列预测上存在计算复杂且内存耗费大等问题,利用时空嵌入和 ProbSparse 自注意机制构建的算法对原始 Transformer 网络进行优化,Informer 模型就是建立在这一趋势的基础上^[17]。但原有 Informer 模型忽略了变量之间的空间相关性, Ma 等人^[18]增加了独立的时空嵌入层,捕获输入信息之间复杂的动态时空相关性。并且,模型中的 ProbSparse 自注意机制可以专注于提取时空数据的重要上下文信息。

为了克服由于传感数据的复杂性造成的注意力机制无效和预测趋势不明显问题中所存在的不足^[19-20], 该文提出了一种融合快速傅里叶变换的 ST-Informer 预测模型,主要贡献有:

- (1) 提出一个独立的空间嵌入层,捕获输入信息之间复杂的动态空间相关性。
- (2) 设计了 FFT 特征提取层,利用 FFT 分解数据,提取先验因子的频域特性。
- (3) 使用来自宁夏引黄灌区的两个站点的数据集作为输入数据进行对比实验。

1 FFT-ST-Informer 模型的首蓿土壤湿度预测

1.1 总体架构

该文提出了一种将快速傅里叶变换与 Informer 时空预测模型相结合的方法,如图 1 所示。

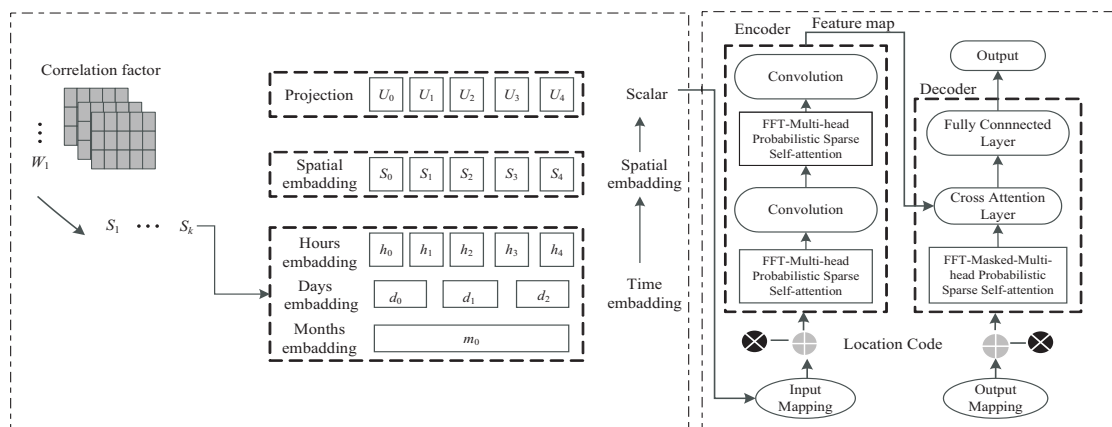


图 1 FFT-ST-Informer 模型框架

该模型核心改进部分为两个模块:(a)时空嵌入层:对具有空间相关性的输入数据进行时空嵌入;(b)

特征提取层:利用 FFT 函数对具有强先验性的张量进行特征提取,使用多头 ProbSparse self-attention 和一维

卷积网络进行进一步特征提取,而后进行特征融合,并通过全连接层输出。

1.2 时空嵌入层

在数据输入部分添加了嵌入层,以确保模型能够捕获输入序列的正确顺序。其中包括时间嵌入层、空间嵌入层以及融合层三部分。

1.2.1 时间嵌入层

本次实验历经 4 年,划分成不同年份,不同月份,不同天数甚至精确到分钟的时间信息。由于数据在时间上的依赖很复杂,在嵌入数据序列的时间特征的同时,还需要给出在输入序列的位置关系。位置编码在考虑嵌入数据的时间特征以外,也考虑时间序列的周期性,如公式 1:

$$\begin{cases} PE(\text{pos}, 2i) = \sin\left(\frac{\text{pos}}{1000(2i/d_{\text{model}})}\right) \\ PE(\text{pos}, 2i + 1) = \cos\left(\frac{\text{pos}}{1000(2i/d_{\text{model}})}\right) \end{cases} \quad (1)$$

其中, d_{model} 指的是通过输入层映射后的数据维度;pos 是指输入数据中的一行数据中的映射特征的位置; $2i$ 和 $2i + 1$ 表示数据行数。等式 1 把时间向量的特征拓展到新的时间戳的矢量化索引中,可以生成输入到最终的时间嵌入层。

1.2.2 空间嵌入层

由于降雨等气象因素会影响土壤湿度变化,再加上人为灌溉的影响,在不同水氮梯度下的土壤含水量互不相同。因此,通过划分不同水氮梯度区域,提取数据之间的空间相关性。

为了获取空间相关性元素,使用 Pearson 相关性系

数。它是一种统计指标^[21],用于衡量两个变量之间的线性相关性,而后进行性状及类别的比较。可以知道, Pearson 的取值范围在 -1 到 1 之间,它的绝对值越接近 1,表示两个变量之间的相关性越强。计算公式如下:

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X\sigma_Y} = \frac{E((X - \mu_X)(Y - \mu_Y))}{\sqrt{E(X^2) - E^2(X)} \sqrt{E(Y^2) - E^2(Y)}} \quad (2)$$

其中, X 和 Y 表示两个变量, $\text{cov}(X, Y)$ 表示两个参数间的协方差, $\sigma_X\sigma_Y$ 表示两个变量的标准差。

在灌溉量的作用下,同一站点不同水氮区域之间存在较强的空间相关性。图 2(a) 显示了每种土壤数据的不同水氮梯度区域间的相关性系数热图。从热图中看出,不同深度下的土壤湿度相较于 N0 区域, N2 区域的相关性更强,并且,随着水梯度的上升,土壤湿度的相关性越来越强。不同深度下的土壤湿度与各个水氮区域的相关性系数均大于 0.50。

另外,实验分别在宁夏引黄灌区下相距 10 km 的两个站点进行,考虑到不同站点会受不同的局部降雨、太阳辐射等因素的影响,分别采集了来自两个区域的数据集,并针对不同站点的气象数据和土壤湿度数据进行了相关性分析。见图 2(b),数据集 1 和数据集 2 在相同土壤深度下的土湿受不同气象因素的影响各不相同,数据集 1 下的土湿在各个气象因素的相关系数普遍不高。因此,土壤湿度随着空间位置的差距会有动态的差异,同时,会对模型训练造成影响。

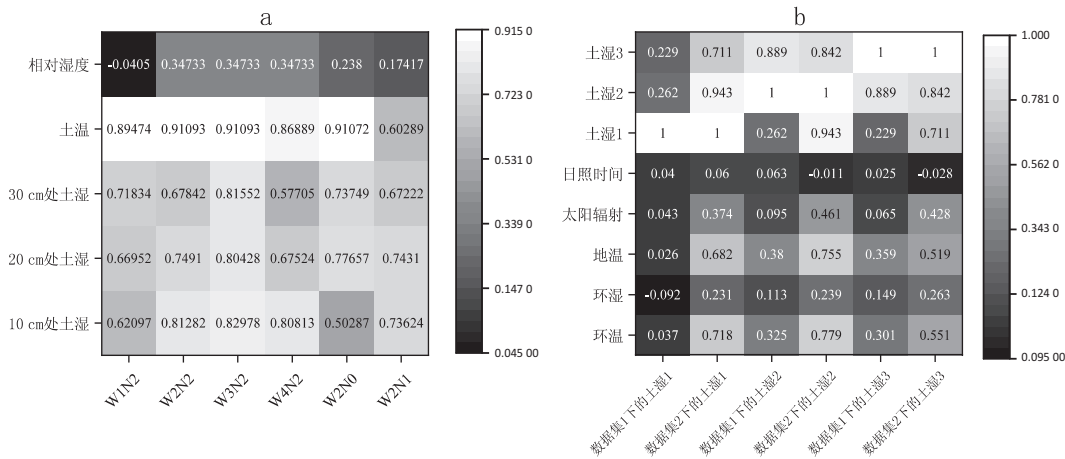


图 2 有关土壤数据的不同水氮区域之间的相关性热图(a)和有关土壤数据的不同数据集的相关性热图(b)

通过前馈神经网络将输入数据投影到高维子空间。

$$Z_t = w * X_t + b \quad (3)$$

其中, Z_t 代表 t 时刻的高维表示。使用上述方程,可以使用互协方差函数计算位点 i 和位点 j 的相关性强度 $S_{t,i,j}$,公式如下:

$$S_{t,i,j} = \text{xcorr}(Z_{t,i}, Z_{t,j}) \quad (4)$$

S_t 是 $N_s \times N_v$ 的平方矩阵,表示 t 时刻两个位置之间的相关性强度,给定时刻下空间嵌入层输出 E_s 的值,计算公式如下:

$$E_s = w * S_t * Z_t + b \quad (5)$$

1.2.3 融合层

融合模块通过使用一维卷积(内核宽度=3,步长=1)将每个时间戳的输入值扩展到一个高维空间 E_V , 并通过学习权重来调整这些嵌入值的相关性,使模型更好地捕捉到不同时间点之间的变化和关联性。通过连接空间、时间和值嵌入层数据来表示完全嵌入特征数据。

$$X_{ST} = E_T \oplus E_S \oplus E_V \quad (6)$$

1.3 FFT 特征提取层

1.3.1 FFT 的特征提取

提取输入序列的特征有利于提高机器学习算法的性能。快速傅里叶变换是一种常用于预处理时间序列的特征提取方法^[22],可以选择性地去除或保留某些频率成分,实现信号的滤波操作。由于原始数据的波动性,实验使用低通滤波频谱处理,提取最大振幅及其相位作为突出的数据特征。接着,对处理后的频域信号应用快速傅里叶逆变换(Inverse Fast Fourier Transform,IFFT),将频谱空间带回物理空间。

设 Z 为长度为 N 的时间序列。使用滑动窗口, Z 被划分成 m 个数据段。其离散傅里叶变换(Discrete Fourier Transform,DFT)为:

$$X(n) = \sum_{k=0}^{N-1} C_k e^{-j\frac{2\pi}{N}nk} \quad (7)$$

通过 FFT 提取幅度和相位作为特征。通过三角函数将原始数据和提取的特征拟合为等式 7,向量 A_m 表示通过三角函数拟合获得的特征信息(最大振幅 a_{m1} 、相位 a_{m2} 和移动 a_{m3})。

$$A_m = [a_{m1}, a_{m2}, a_{m3}] \quad (8)$$

$$\rho(x_i) = a_{m1} \sin(a_{m2} \frac{X_{\Pi}}{n/2}) + a_{m3} \quad (9)$$

原始结构状态数据 Z 被转换成特征向量 A 。方程 9 用来拟合窗口数据,其中 a_{m1} 和 a_{m2} 由最大振幅和相位固定。作为特征向量公式 8 中的参数向量 A_m 是模型的输入。

1.3.2 多头 ProbSparse 自注意机制

另外模型引入了多头 ProbSparse 自注意机制,提取主要特征进行压缩。它解决了自注意力点积计算的二次时间复杂度,通过对每个单独的 Query 点进行采样分析,得到公式:

$$A(Q, K, V) = \text{Softmax}(\frac{\bar{Q}K^T}{\sqrt{d}})V$$

$$\left(Q \in R^{L_q \times d}, K \in R^{L_k \times d}, V \in R^{L_v \times d}, \text{Softmax}(t) = \frac{1}{1 + e^{-t}} \right) \quad (10)$$

其中, Q, K 和 V 分别是通过特征变量的线性变换获得的三个相同大小的矩阵, d 是输入维度, \bar{Q} 是通过概率

稀疏化从 Q 中获得的, Softmax 是激活函数。

1.3.3 FFT-ProbSparse 自注意力

考虑到灌溉、降雨等气象因素的不确定性,通过 Pearson 相关性分析发现(如公式 2),不同深度下的土壤墒情值与降雨、灌溉呈正相关性,且大于其他参数相关性。这表明随着人为灌溉量的增多和降雨量的突发变化,土壤墒情也会随之变化。并将得到的强先验性的数据进行 FFT 的特征提取,然后通过快速傅里叶变换与 ProbSparse 自注意集成,将分析得到的降雨、灌溉的频域信息和其他特征信息融合,称作 FFT-ProbSparse 自注意力。

1.3.4 distilling 操作

作为 FFT-ProbSparse 自注意机制的结果,Encoder 的特征映射存在值的冗余组合。为提取长序列的远程依赖性,改进的 Informer 引入了卷积、激活和最大池化操作,已将前一层的输入序列长度减半,从而解决长输入序列的内存占用过多问题。从 j 到 $j+1$ 层的 distilling 操作的过程如下:

$$X_{j+1}^t = \text{MaxPool}(\text{ELU}(\text{Conv1d}([\ X_j^t]_{AB}))) \quad (11)$$

其中, $[\cdot]_{AB}$ 包含了多头 ProbSparse self-attention 以及在注意力块中的关键操作。Conv1d 表示时间序列上的一维卷积操作,并通过 ELU 作为激活函数。

1.3.5 输出层

模型使用批量生成预测直接输出多步预测结果,从而提高长序列的预测速度。方程式如下:

$$X_{de}^t = \text{Concat}(X_{token}^t, X_o^t) \in \mathbb{R}^{(L_{token} + L_o) \times d_{model}} \quad (12)$$

其中, $X_{token}^t \in \mathbb{R}^{L_{token} \times d_{model}}$ 是起始 token, $X_o^t \in \mathbb{R}^{L_o \times d_{model}}$ 是目标序列的占位符(将标量设置为 0)。通过将 masked 点积设置为 $-\infty$,将 masked 多头注意力应用于 ProbSparse 自注意计算。它防止每个位置都关注即将到来的位置信息,从而避免了自回归。全连接层获得最终输出,它的超大规模取决于是进行单变量预测还是多变量预测。

2 实验

2.1 数据集

2.1.1 数据来源

提供了在宁夏引黄灌区连续多年不同试验区域下的苜蓿多层土壤墒情田间数据,整理了 19 763 条多土层土壤墒情数据 ($>0 \sim 10$ cm, $>10 \sim 20$ cm, $>20 \sim 30$ cm),同时包括其同步的气象数据、作物生长数据和田间管理数据。

试验区位于银川市农垦茂盛草业有限公司第一草场试验点,地处西北地区农牧交错带,属温带大陆性气候。根据宁夏引黄灌区灌溉实际,设计了漫灌和地下点源滴灌方式下不同水氮施灌量的田间试验,设计 4

个灌溉水平 (W1, 5 250 m³/hm²)、(W2, 6 000 m³/hm²)、(W3, 6 750 m³/hm²)和(W4, 7 500 m³/hm²)和4个氮水平 (N0, 0)、(N1, 60 kg/hm²)、(N2, 120 kg/hm²)、(N3, 180 kg/hm²),共16个处理区域。

2.1.2 数据处理

在使用数据集之前,需要进行数据预处理。其中包括缺失值填充和数据归一化。

(1)缺失值填充:由于2017年和2018年设备损耗造成的部分数据缺失,选择用历史同期的数据填充。2022,2023年田间实验设备崭新完善,采集到的数据缺失较少,缺失的部分通过前后数据的平均值来填充。

(2)数据归一化:对数据进行归一化^[23],以提高预测精度。归一化表达式如下:

$$X_i = \frac{x - \min}{\max - \min} \quad (13)$$

2.2 评估准则

实验按照训练集:测试集:验证集以7:2:1来划分数据。采用平均绝对误差(MAE)、均方误差(MSE)、相关系数(R²)评价模型预测的效果。

(1)MAE是绝对误差的平均值,能更好地反映预测值误差的实际情况。

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (14)$$

(2)MSE用于衡量预测值与真实值之间的平均方差,能很好地反映模型的预测精度,误差越大,权重越高,拟合程度越差。

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (15)$$

(3)R²用来衡量模型对数据的拟合优度,R²越接近1,表示模型能够解释因变量的变异程度越高,拟合

效果越好;R²越接近0,表示模型对观测值的变异程度解释能力较弱,拟合效果较差。

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{pred} - \hat{y}_i)^2}{\sum_{i=1}^n (y_{act} - \hat{y}_i)^2} \quad (16)$$

2.3 实验参数设置

实验代码使用Python语言及第三方库,深度学习环境采用Pytorch框架,并在GPU上进行训练,模型训练过程中,使用的部分参数如表1所示。

表1 模型训练参数设置

参数	取值范围
Pred_len	2,4,8,12,24
Train_epochs	10,15,20
Batch_size	32,64
Learning rate	1e-2,1e-3,1e-4
Dropout	0.05
Num of heads	8,16

2.4 实验结果与分析

2.4.1 定性分析

在模型训练完毕之后,通过将自采的数据集输入到5个模型中进行预测,如图3所示。考虑数据完整性选择2017年的数据(数据集2),图3中时间为60天,时间单位是1小时。同时,添加LSTM^[24]、ILSTM^[14]、Resnet^[25]、Informer^[17]网络作为比较算法,以验证FFT-ST-Informer的性能。实验结果表明FFT-ST-Informer模型拟合程度明显高于其他模型,而Informer模型和改进后的Informer模型的拟合曲线接近,这表明Informer模型能更好地反映时间序列模型的预测性能。

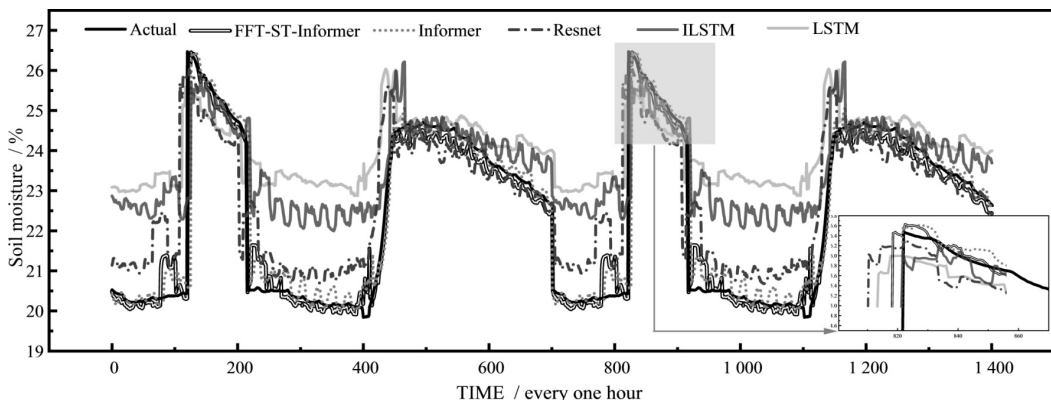


图3 不同模型在数据集上的预测结果

2.4.2 定量分析

为验证该模型在长期预测的表现,选择不同预测时间进行土壤墒情预测。通过在未来2小时,4小时,8小时,16小时,20小时和24小时的预测分别对5个模型进行评估,针对表2所示的评估指标计算得出结

果,发现FFT-ST-Informer模型评估效果均优于其他4个模型。结果表明,随着预测时间的推移,模型的拟合会慢慢变差,但均保持在0.3以上。同时,随着预测时间的增长,模型的误差越大,预测效果也越差。对比发现,经过改进的Informer模型,即FFT-ST-Informer模

型拟合优度高于其他 4 个模型,相较于 LSTM、ILSTM、Resnet、Informer 高了 19% ,24% ,12% ,10% ,并且误差率较 Informer 模型低了 17% 。

表 2 4 种模型的评估比较分析

模型	MAE/($\mu\text{g}/\text{m}^3$)	MSE/($\mu\text{g}/\text{m}^3$)	R^2
LSTM	11.2 ± 3.72	13.4 ± 2.40	0.61 ± 0.20
ILSTM	14.6 ± 5.31	18.51 ± 2.85	0.56 ± 0.19
Resnet	9.1 ± 2.82	9.32 ± 3.52	0.68 ± 0.12
Informer	7.11 ± 2.22	7.15 ± 3.01	0.70 ± 0.18
FFT-ST-Informer	5.71 ± 2.52	5.44 ± 2.92	0.80 ± 0.13

2.4.3 消融实验

为验证改进后模型的 FFT 特征融合的效果,进行消融实验,比较 Informer 模型在有无 FFT 的情况下的评估结果,如表 3 所示,表明带有 FFT 的 Informer 模型在不同预测时间下的性能高于 Informer 模型,在预测 2 小时的土壤墒情上可达 93%。这证明了在模型中添加 FFT 模块有利于提高模型的精准预测性。

表 3 两种模型在不同预测时数的对比评估结果

预测时数	模型	MAE	MSE	R^2
2	Informer *	0.31	0.34	0.93
	Informer	0.49	0.43	0.88
4	Informer *	0.49	0.36	0.89
	Informer	0.68	0.55	0.84
8	Informer *	0.52	0.44	0.79
	Informer	0.78	0.46	0.70
12	Informer *	0.7	0.50	0.75
	Informer	0.82	0.72	0.68
24	Informer *	0.82	0.83	0.59
	Informer	0.92	1.1	0.45

注: * 代表融合 FFT 方法的 Informer 结构。

表 4 模型在不同数据集下的对比评估结果

Datesets		Batch_size	Learning rate	Adam	Multi-head	MAE	MSE	R^2
数据集 1	数据集 2							
√		32	0.000 1	GLUE	32	7.85	1.32	0.52
	√					0.39	0.41	0.91

表 5 模型在多头注意力下的对比评估结果

FFT-ST-Informer	Multi-head					MSE	MAE	R^2
	head=2	head=4	head=8	head=16	head=32			
√	√					0.41	0.36	0.90
√		√				0.30	0.39	0.92
√			√			0.25	0.31	0.94
√				√		0.24	0.34	0.95
√					√	0.27	0.34	0.93

实验分别在两个站点进行,会存在空间的限制,将改进后的 Informer 模型在这两个数据集上分别进行评估,见表 4。

实验发现,模型在不同数据集上的效果存在偏移,相较于数据集 1,数据集 2 的拟合效果明显提高了 39%。这证明模型可以普遍适用不同数据集,然而,在考虑不同空间位置的情况下,且通过不同的气象数据影响,使得网络模型存在数据偏移现象,导致模型的误差值变大,训练效果变差。

考虑模型中的 Multi-head 的注意力机制,不同于卷积神经网络的多个卷积层中的卷积核,它是将输入的特征值拆分成更加细碎的小块,对每个小块赋值一个单独的可训练权重参数,然后共用一个隐藏层输出结果。

本次实验将 head 分解成 2,4,8,16,32 个不同的注意力,可以从表 5 看出,随着设置的 head 越高,模型性能水平相对提高,当 head 值为 16 时,模型训练效果最优,误差更小。由此可知,应当设置合适的 head 值,才更有利于 FFT-ST-Informer 模型。

3 结束语

针对土壤墒情预测模型的精度不高以及时空预测不足等问题,使用了包括 FFT 特征提取和 ProbSparse 注意力机制混合的方法降低了空间复杂度,并在原始 Informer 方法上对嵌入层进行修改,加入空间嵌入,弥补了空间预测的不足。该工作在自采土壤墒情数据集上实现了更精准的预测。与 LSTM、ILSTM、Resnet、Informer 模型进行对比,结果表明采用改进后的 Informer 土壤墒情预测模型所得到的结果与实际的土壤墒情变化趋势更贴近,模型的预测性能更好。

参考文献:

- [1] 薛明,韦波,李娟,等.基于改进 BP 神经网络与支持向量机的土壤墒情预测方法研究[J].土壤通报,2021,52(4):793-800.
- [2] ZHU S,LI H,XU W,et al. Dynamic process and model simulation of soil water content during citrus growth period[J]. Acta Geophys,2023,71:1525-1537.
- [3] 王铁英,王仰仁,战国隆,等.基于实时含水率数据的土壤墒情动态建模及预测[J].中国农业气象,2021,42(1):13-23.
- [4] 吴建龙.基于 LSTM 神经网络的土壤墒情预测方法的研究与应用[D].北京:北京化工大学,2023.
- [5] 落全富,杨宁.基于土壤水量平衡法的山区小流域地下侧向流时空特征分析[J].水力发电,2023,49(7):16-23.
- [6] 韩红亮,王雪梅.两种土壤墒情预报模型在沧州地区的应用及评价[J].吉林水利,2022(5):47-51.
- [7] HOSSEINT,ALI-AKBAR S,MOHAMMAD A. Comparison of artificial neural network and multivariate linear regression methods for estimation of daily soil temperature in an arid region[J]. Meteorology and Atmospheric Physics,2011,110:135-142.
- [8] MENG Xiangjin,MAO Kebiao,MENG Fei,et al. A fine-resolution soil moisture dataset for China in 2002 - 2018[J]. Articles,2021,13(7):3239-3261.
- [9] HAN H,CHOI C,KIM J,et al. Multiple-depth soil moisture estimates using artificial neural network and long short-term memory models[J]. Water,2021,13(18):2584.
- [10] 肖天云,张子晨,魏佳妹,等.基于 BOA-LSTM 模型的土壤湿度预测[J].华北理工大学学报:自然科学版,2023,45(3):65-73.
- [11] 王学智.基于深度学习的土壤温湿度预测相关技术研究[D].长春:吉林大学,2022.
- [12] SUN J,DI L P,SUN Z H,et al. County-level soybean yield prediction using deep CNN-LSTM model[J]. Sensors,2019,19(20):4363.
- [13] HUANG C J,KUO P H. A deep CNN-LSTM model for particulate matter (PM_{2.5}) forecasting in smart cities[J]. Sensors,2018,18(7):2220.
- [14] LI Q,ZHU Y,SHANGGUAN W. An attention-aware LSTM model for soil moisture and soil temperature prediction[J]. Geoderma,2022,409:115651.
- [15] LI Xinjuan,YU Jie,XU Lingyu,et al. Time series classification with deep neural networks based on hurst exponent analysis [C]//Lecture notes in artificial intelligence. Guangzhou: [s. n.],2017:194-204.
- [16] WU N,GREEN B,BEN X,et al. Deep transformer models for time series forecasting:the influenza prevalence case[J]. arXiv:2001.08317,2020.
- [17] ZHOU Haoyi,ZHANG Shanghang,PENG Jieqi,et al. Informer:beyond efficient transformer for long sequence time-series forecasting[C]//Proceedings of the AAAI conference on artificial intelligence. Palo Alto:AAAI,2021:11106-11115.
- [18] MA Zhanfei,LUO Wenli,JIANG Jing,et al. Spatial and temporal characteristics analysis and prediction model of PM_{2.5} concentration based on spatio-temporal-informer model[J]. PloS One,2023,18:6.
- [19] LI Zuohua,HE Jingbo,LIU Diankun,et al. Influence of uniaxial stress on the shear-wave spectrum propagating in steel members[J]. Sensors,2019,19:492.
- [20] CHEN Zhicheng,LI Hui,BAO Yuequan. Analyzing and modeling inter-sensor relationships for strain monitoring data and missing data imputation;a copula and functional data-analytic approach[J]. Structural Health Monitoring,2019,18(4):1168-1188.
- [21] JEBLI I,BELOUADHA F Z,KABBAJ M I,et al. Prediction of solar energy guided by pearson correlation using machine learning[J]. Energy,2021,224(1):120109.
- [22] MA Jihao,DAN Jingpei. Long-term structural state trend forecasting based on an FFT-informer model[J]. Applied Sciences,2023,13:2553.
- [23] 杨寒雨,赵晓永,王磊.数据归一化方法综述[J].计算机工程与应用,2023,59(3):13-22.
- [24] YU Yong,SI Xiaosheng,HU Changhua,et al. A review of recurrent neural networks:LSTM cells and network architectures [J]. Neural Computation,2019,31(7):1235-1270.
- [25] SHAFIQ M,GU Zhaoquan. Deep residual learning for image recognition;a survey[J]. Appl. Sci,2022,12(18):8972.