

YAN Shuainan, LI Xuebao, DONG Liang, HUANG Wengeng, WANG Jing, YAN Pengchao, LOU Hengrui, HUANG Xusheng, LI Zhe, ZHENG Yanfang. Application of $F_{10.7}$ Index Prediction Model Based on BiLSTM-attention and Chinese Autonomous Dataset (in Chinese). *Chinese Journal of Space Science*, 2024, 44(2): 251–261. DOI:10.11728/cjss2024.02.2023-0040

基于 BiLSTM-Attention 的 $F_{10.7}$ 指数预测 模型与中国自主数据集的应用*

闫帅楠¹ 李雪宝¹ 董亮² 黄文耿³
王晶³ 闫鹏朝¹ 娄恒瑞¹
黄徐胜¹ 李哲⁴ 郑艳芳¹

1(江苏科技大学 镇江 212003)

2(中国科学院云南天文台 昆明 650216)

3(中国科学院国家空间科学中心 北京 100190)

4(北方信息控制研究院集团有限公司 南京 211153)

摘要 $F_{10.7}$ 指数是太阳活动的重要指标, 准确预测 $F_{10.7}$ 指数有助于预防和缓解太阳活动对无线电通信、导航和卫星通信等领域的影响. 基于 $F_{10.7}$ 射电流量的特性, 在双向长短时记忆网络 (Bidirectional Long Short-Term Memory Network, BiLSTM) 基础上融入注意力机制 (Attention), 提出了一种基于 BiLSTM-Attention 的 $F_{10.7}$ 预报模型. 在加拿大 DRAO 数据集上其平均绝对误差 (MAE) 为 5.38, 平均绝对百分比误差 (MAPE) 控制在 5% 以内, 相关系数 (R) 高达 0.987, 与其他 RNN 模型相比拥有优越的预测性能. 针对中国廊坊 L&S 望远镜观测的 $F_{10.7}$ 数据集, 提出了一种转换平均校准 (Conversion Average Calibration, CAC) 方法进行数据预处理, 处理后的数据与 DRAO 数据集具有较高的相关性. 基于该数据集对比分析了 RNN 系列模型的预报效果, 实验结果表明, BiLSTM-Attention 和 BiLSTM 两种模型在预测 $F_{10.7}$ 指数方面具有较好的优势, 表现出较好的预测性能和稳定性.

关键词 $F_{10.7}$ 预报, 双向长短时记忆网络, 注意力机制, L&S 数据集

中图分类号 P353

Application of $F_{10.7}$ Index Prediction Model Based on BiLSTM-attention and Chinese Autonomous Dataset

YAN Shuainan¹ LI Xuebao¹ DONG Liang² HUANG Wengeng³
WANG Jing³ YAN Pengchao¹ LOU Hengrui¹
HUANG Xusheng¹ LI Zhe⁴ ZHENG Yanfang¹

* 国家自然科学基金项目 (11703009, 11803010) 和江苏省自然科学基金面上项目 (BK20201199) 共同资助

2023-03-25 收到原稿, 2023-05-10 收到修定稿

©The Author(s) 2024. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

1(Jiangsu University of Science and Technology, Zhenjiang 212003)

2(Yunnan Observatories, Chinese Academy of Sciences, Kunming 650216)

3(National Space Science Center, Chinese Academy of Sciences, Beijing 100190)

4(North Information Control Research Academy Group Co., Ltd., Nanjing 211153)

Abstract The $F_{10.7}$ index is an important indicator of solar activity. Accurate predictions of the $F_{10.7}$ index can help prevent and mitigate the effects of solar activity on areas such as radio communications, navigation and satellite communications. Based on the properties of the $F_{10.7}$ radio flux, the prediction model of $F_{10.7}$ based on BiLSTM-Attention is proposed by incorporating an Attention mechanism on the Bidirectional Long Short-Term Memory Network (BiLSTM). The Mean Absolute Error (MAE) on the Canadian DRAO dataset is 5.38, the Mean Absolute Percentage Error (MAPE) is controlled to within 5% and the correlation coefficient (R) reaches 0.987. It has superior prediction performance compared with other RNN models in both short-term and medium-term prediction. A Conversion Average Calibration (CAC) method is proposed to preprocess the $F_{10.7}$ data set observed by the Langfang L&S telescope in China. The processed data has high correlation with the DRAO dataset. Based on this dataset the forecasting effectiveness of the RNN series models is compared and analyzed. The experimental results show that both BiLSTM-Attention and BiLSTM models have significant advantages in predicting the $F_{10.7}$ index and show excellent predictive performance and good stability. The BiLSTM-Attention model has the highest prediction accuracy when forecasting future first-day data, with MAE and MAPE of 11.10 and 8.66, respectively, and the MAPE is always within 15% in the short- and medium-term forecasts. This shows that the proposed model has high generalization ability and can effectively predict the $F_{10.7}$ data set of DRAO and L&S.

Key words $F_{10.7}$ index forecasting, Bidirectional Long Short-Term Memory Network, Attention mechanism, L&S dataset

0 引言

太阳活动是影响中高层大气和电离层变化的重要因素,对通信系统及气候变化都会产生重要影响^[1]. 太阳 10.7 cm 射电流量 ($F_{10.7}$ 指数) 是日面上所有源区的总射电辐射流量值,是衡量太阳活动水平的典型参数,能够反映太阳活动的周期性变化^[2]. $F_{10.7}$ 被广泛应用于大气密度模型、电离层模型等空间环境模型,其预测的精度直接影响航天器轨道的预报精度. 因此,准确预测 $F_{10.7}$ 指数具有重要意义.

$F_{10.7}$ 指数的变化规律与太阳的周期性变化密切相关,在不同时间尺度上呈现出不同的变化特征^[2], $F_{10.7}$ 预测是典型的基于时间序列的预测回归问题. Wen 等^[3] 基于自回归方法模型,利用活动区面积、位置等参数与 $F_{10.7}$ 指数的定量关系,建立了基于活动区演化规律的中期预报模型,其准确性与传统的自回归模型相比有很大提高. Wang 等^[4] 根据太阳活动 27 天短期振荡的特点,利用过去 135 天的历史资料,

研究了 $F_{10.7}$ 中期预报方法,并采用傅里叶级数方法评价 $F_{10.7}$ 预测的效果. 由于太阳活动区的出现或者衰减, $F_{10.7}$ 指数在震荡时间范围内的幅度和周期不断变化,传统的统计回归方法缺乏高质量且稳定的预测能力,对于预测性能的提高十分有限^[5]. 机器学习因具有出色的数据处理能力而被逐步应用到太阳活动研究中. Valdés 等^[6] 利用无监督和有监督机器学习技术,研究了 2006—2018 年的磁图结构及其与 10.7 cm 太阳射电通量的关系,结果表明机器学习在空间天气领域具有巨大的应用潜力. Huang 等^[7] 使用 SVR (Support Vector Regression) 进行短期 $F_{10.7}$ 预测. Luo 等^[8] 提出一种将反向传播神经网络与经验模态分解 (Empirical Mode Decomposition, EMD) 相结合的方法,预测未来 27 天的 $F_{10.7}$ 日值. 深度学习神经网络 (DNNs) 作为机器学习的一个新分支,已成为解决天文学和其他科学分支中大规模学习问题的一种高度可靠的技术^[9]. 循环神经网络 (Recurrent Neural Network, RNN) 由于隐藏层结点互相连接而具有良

好的时间序列处理能力. Yang 等^[10] 使用长短时记忆网络 (Long Short-Term Memory, LSTM) 构建了预测未来 27 天高低年 $F_{10.7}$ 指数太阳活动的预测模型, 平均相对误差最优可达 10% 以内. Luo 等^[11] 提出了一种 EMD-LSTM 混合预测方法来预测未来 27 天的 $F_{10.7}$ 日值. 与其他孤立算法和混合方法相比, 该方法具有较高的精度.

RNN 在处理时序问题中具有优越性, 但当输入序列过长时容易丢失序列信息^[12], 进而影响预测准确度, 因此有研究提出将注意力机制应用于 RNN, 对关键信息分配足够的关注以提高模型的预测能力. Liu 等^[13] 在 LSTM 中引入注意力机制进行太阳耀斑预报并取得不错的进展. 因此, 在 LSTM 中引入注意力机制能否提高 $F_{10.7}$ 预报性能是值得探究的. 本文采用国际 $F_{10.7}$ 数据集进行模型训练、测试和验证. 通过分析 $F_{10.7}$ 预测问题的研究现状和特点, 在 LSTM 强时序处理能力基础上增加了双向机制, 构建双向长短时记忆网络 (Bidirectional Long Short-Term Memory Network, BiLSTM) 并引入注意力机制, 提出了一种基于 BiLSTM-Attention 的 $F_{10.7}$ 预测方法. 目前关于 $F_{10.7}$ 预报研究中^[10,11,13,14] 使用的 $F_{10.7}$ 数据均由加拿大哥伦比亚彭带克顿市的射电天文台 DRAO (Dominion Radio Astrophysical Observatory) 观测得到, 没有相关研究使用其他 $F_{10.7}$ 数据源进行预测分析. 中国廊坊 L&S 多波段太阳精密流量射电望远镜 (L&S 望远镜) 包含了 8 个 L 波段内的频点和 $F_{10.7}$ 频点, $F_{10.7}$ 数据经与日面黑子面积数据比对, 相关性较高, 是中国当时唯一一台 $F_{10.7}$ 标准测量望远镜. 为探究算法性能的优越性, 处理了 L&S 望远镜观测得到的 $F_{10.7}$ 数据并进行测试.

1 数据与处理方法

所使用数据集来自空间标准和创新中心^{*[15]}. 其中与 $F_{10.7}$ 指数相关的数据包括观测值、调整值和 D 系列值. 在本文工作中使用原始观测得到的 $F_{10.7}$ 指数值来训练预测模型. 望远镜每天测量三次 $F_{10.7}$ 流量值, $F_{10.7}$ 日值取为 20:00 LT 测量值. 选择了 1957—2022 年的 $F_{10.7}$ 数据, 共包含 6 个太阳活动周期. 由于时间序列数据的相关性, 不能使用随机划

分的策略划分数据集^[16]. 为确保各个子集中包含完整的太阳活动周期, 依据活动周期划分训练集、测试集和验证集. 图 1 显示了处理后的数据, 将 1957—1989 年的数据作为训练集, 1989—2000 年的数据作为验证集, 2000—2022 年的数据作为测试集.

使用中国 L&S 望远镜观测得到的数据, $F_{10.7}$ 频段数据由 2840 MHz 频段接收机得到. 由于原数据格式不适用于模型训练测试, 提出了一种转换平均校准 (Conversion Average Calibration, CAC) 数据处理方法, 该方法主要由转换、均值化和校准三部分组成. 数据处理过程如图 2 所示.

为了计算 $F_{10.7}$ 指数, 需要有天线对准冷空时的数据 (由天线偏开太阳 2° 时采集得到), 以及跟踪太阳

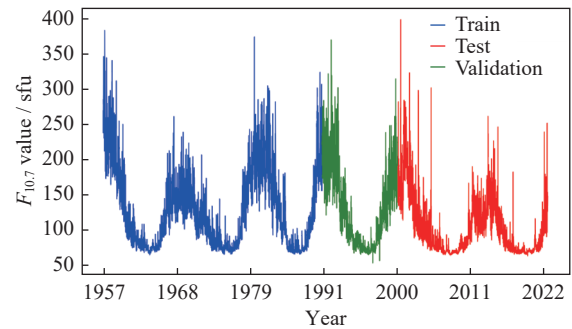


图 1 1957—2022 年 $F_{10.7}$ 日值

Fig. 1 Daily value of $F_{10.7}$ from 1957 to 2022

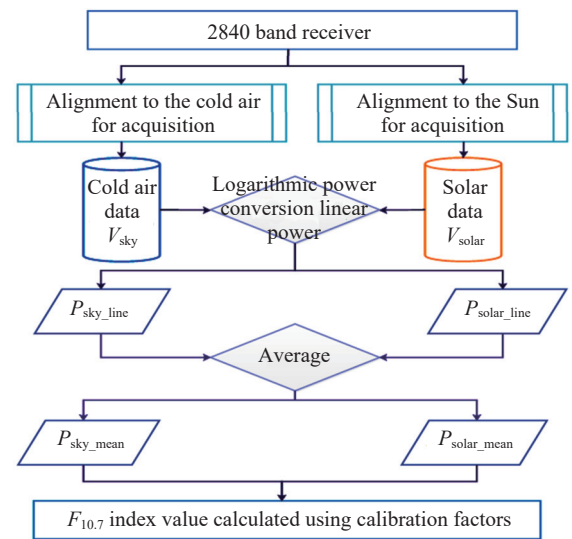


图 2 L&S 望远镜 $F_{10.7}$ 数据处理

Fig. 2 L&S telescope $F_{10.7}$ data processing

* <http://www.celstrak.com/SpaceData/SW-All.txt>

观测得到的数据, 即天线跟踪太阳采集得到的太阳射电辐射功率. 天线对准冷空时, 获得接收机对天空背景温度的响应, 该数据是稳定值. 通过每天中午观测得到的太阳射电流量数据与冷空数据的对比, 可以消除接收机在长期工作中产生的增益变化影响. 对准太阳时, 采集得到电压数组 $V_{\text{solar}}[i]$; 对准冷空时, 采集得到电压数组 $V_{\text{sky}}[i]$. 由于采用的是对数检波器, 需要逐点对对数功率转换为线性功率, 分别得到对准太阳时的线性功率

$$P_{\text{solar_line}}[i] = 10^{\frac{V_{\text{solar}}[i]}{0.021}}, \quad (1)$$

对准冷空时的线性功率

$$P_{\text{sky_line}}[i] = 10^{\frac{V_{\text{sky}}[i]}{0.021}}. \quad (2)$$

将 $P_{\text{solar_line}}[i]$ 和 $P_{\text{sky_line}}[i]$ 取平均值可以得到 $P_{\text{solar_mean}}$ 和 $P_{\text{sky_mean}}$, 防止出现将一些线性动态增量压缩或者畸变的情况. 利用校准因子 S_0 计算得到当日的 $F_{10.7}$ 指数值, 即

$$F_{10.7} = S_0 \frac{P_{\text{solar_mean}}}{P_{\text{sky_mean}}}. \quad (3)$$

由于 $F_{10.7}$ 指数震荡幅度和周期不断变化, 为了提高校准后数据的准确度, 校准因子 S_0 通过前一天的 $F_{10.7}$ 指数值与国际 $F_{10.7}$ 指数值校准得到.

使用 CAC 方法进行数据处理后, 需要对数据进行质量检验和处理, 剔除异常值与坏点. 经过质量检

验后方可使用模型进行训练与评估. 具体而言, 如果校准后的数据出现明显抖动, 波动值较大时, 需要配合 X-ray 数据进行检查. 若抖动较大的数据点当天 X-ray 辐射较大, 则该点正常, 不予修正和剔除; 若当天没有发生太阳耀斑, X-ray 流量值正常, 则表示该点异常, 需要进一步修正, 就近采取前一天的数据作为当天数据. 将处理后的 L&S 望远镜数据与 DRAO 数据进行对比, 时间范围为 2022 年 1 月 28 日至 2022 年 9 月 25 日, 如图 3 所示. 由图 3 可知, 校准后的廊坊数据曲线和国际 $F_{10.7}$ 数据曲线的变化趋势基本一致, 二者相关性较高.

2 预报方法

LSTM 网络的基本单元中包含遗忘门、输入门和输出门, 如图 4 所示. 遗忘门主要对上一时刻信息和新输入信息进行筛选, 控制信息是否保留, 遗忘门中信息可以表达为

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f). \quad (4)$$

输入门主要确定需要在当前细胞状态中存储的新信息, 新信息由更新值和新的候选变量两部分组成, 结合这两部分信息得到更新后的细胞状态 C_t , 输入门可表示为

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \quad (5)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c), \quad (6)$$

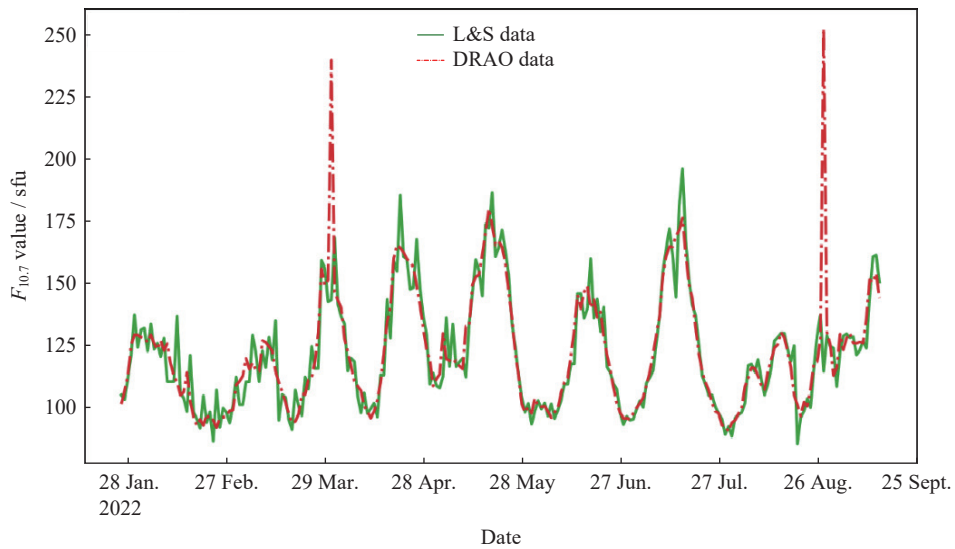


图 3 L&S 望远镜数据与 DRAO 数据对比

Fig. 3 Comparison of L&S telescope data with DRAO data

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t. \quad (7)$$

最后在时间步 t 时刻, 基于当前 LSTM 单元的新状态 C_t , 输出该单元的隐藏状态 h_t , 输出门可表示为^[17]:

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o); \quad (8)$$

$$h_t = o_t \odot \tanh(C_t). \quad (9)$$

式中, W_f, W_i, W_c, W_o 为相应门与输入和中间输出相乘的矩阵权重, b_f, b_i, b_c, b_o 为对应的偏置项, σ 为 sigmoid 函数变化.

在 $F_{10.7}$ 指数预测问题中, 在时刻 t , 将 l 个历史长度的 $F_{10.7}$ 指数数据输入 LSTM 模型, 作为输入序列 $x_{t-l+1}, x_{t-l+2}, \dots, x_t$, 其中 x_t 代表第 t 个时刻的 $F_{10.7}$ 指数值, 如图 5 所示. LSTM 根据输入序列和前一个时刻的状态 h_{t-1} 计算当前时刻的输出 h_t , 并更新记忆单元 C_t , 预测未来某一时刻的 $F_{10.7}$ 射电流量, 即

$$\hat{x}_{t+h} = f(x_{t-l+1}, x_{t-l+2}, \dots, x_t). \quad (10)$$

其中, h 为预测步长, f 为预测模型的隐函数.

在以往研究^[10,11,13,14] LSTM 模型的基础上增加

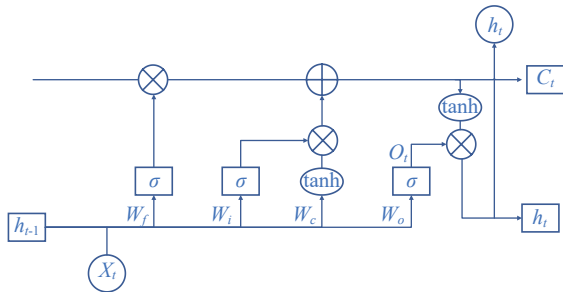


图 4 LSTM 结构

Fig. 4 LSTM network structure

双向结构, 双向 LSTM (BiLSTM, Bidirectional Long Short-Term Memory Network) 让网络在序列的某点处不仅可以获取该时间之前的信息, 还可获取未来的信息, 能够发现仅单向 LSTM 网络无法找到的关于时间序列的其他规律^[18], 如图 6 中 BiLSTM 所示.

RNN 在处理时序问题中具有优越性, 但当输入序列过长时容易丢失序列信息, 进而影响预测准确度. 为解决该问题, 在 BiLSTM 中添加与 Liu 等^[13] 在太阳耀斑预测中使用的注意力机制, 该注意力机制通过计算输入数据序列的加权和来映射并捕获输入与输出之间的关系, 注意力层输出的注意力上下文向量 A_i 计算过程如下:

$$S_{i,j} = V \times \tanh(W'(S_i, H_j)), \quad (11)$$

$$W_{i,j} = \frac{e^{S_{i,j}}}{\sum_{k=1}^m e^{S_{i,k}}}, \quad (12)$$

$$A_i = \sum_{j=1}^m W_{i,j} H_j. \quad (13)$$

其中, $\tanh$ 为双曲正切激活函数, S_i 为输出与输出相应的输出状态, V 和 W 为神经网络学习的权重矩阵. m 为输入序列长度, H_j 为对应于输入数据样本 x_j 的隐藏状态, W 表示隐藏状态的权重.

3 结果评估

3.1 模型参数配置

为构建基于 BiLSTM-Attention 的 $F_{10.7}$ 预测模型, 需要确定模型的输入层时间步数、隐藏层个数及神经元数等参数. 输入层时间步数等于用来进行

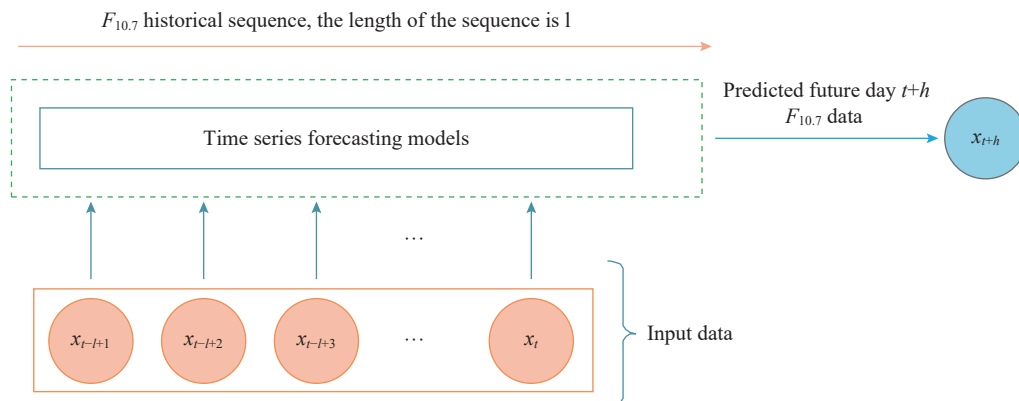


图 5 $F_{10.7}$ 预测基本思路

Fig. 5 Basic idea for $F_{10.7}$ prediction

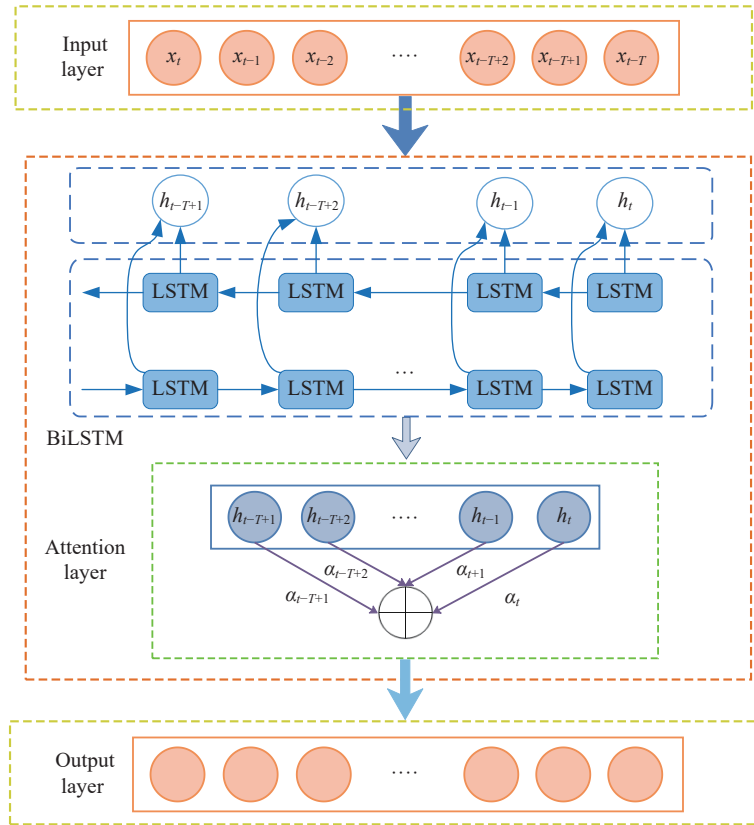


图6 BiLSTM-Attention 模型结构

Fig. 6 Structure of BiLSTM-Attention model

$F_{10.7}$ 流量预测的时间序列长度, 即图 5 中的序列长度 l . 考虑到模型训练的有效性和预测精度, 通过试探将时间序列长度设置为 27, 即输入前 27 天的 $F_{10.7}$ 数据进行预测. 将 BiLSTM 设置为 3 层, 神经元数量分别为 16, 16, 8. 训练模型所需轮次设置、优化器以及损失函数等详细参数列于表 1.

3.2 评估指标

模型预测性能评价指标采用了均方根误差 (RMSE, E_{RMSE})、平均绝对误差 (MAE, E_{MAE})、平均绝对百分比误差 (MAPE, E_{MAPE})、决定系数 (R^2 , E_R^2) 和相关系数 (R , E_R). RMSE 和 MAE 表示预测值与真实值之间的误差, 反映模型的预测性能. 误差越小, 则预测效果越好. MAPE 常用作时间序列预测中表示预测误差的评价指标, 其值越小, 说明预测模型具有越高的精确度^[14]. R 表示预测值与真实值之间的相关性, 相关系数越高, 表明预测值与真实值同步增加或减少, 预测效果越好. R^2 表示决定系数, 反映模型拟合数据的准确程度, 其值越接近 1, 表明模型对数据拟合程度越好. 具体公式如下:

表 1 参数配置信息

Table 1 Parameter configuration information

参数配置项	值或信息
训练轮次(epoch)	80
学习率	0.001
损失函数	MSE
batch_size	8
优化器	RMSprop

$$E_{\text{RMSE}} = \sqrt{\frac{\sum_{t=1}^T (\hat{y}_t - y_t)^2}{T}}, \quad (14)$$

$$E_{\text{MAE}} = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|, \quad (15)$$

$$E_{\text{MAPE}} = \frac{1}{T} \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{\hat{y}_t} \right|, \quad (16)$$

$$E_R = \frac{\sum_{t=1}^T y_t \hat{y}_t}{\sqrt{\sum_{t=1}^T \hat{y}_t^2} \sqrt{\sum_{t=1}^T y_t^2}}, \quad (17)$$

$$E_{R^2} = 1 - \frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{\sum_{t=1}^T (y_t - \bar{y})^2}. \quad (18)$$

其中, T 表示样本数量, $\hat{y}_t$ 和 y_t 分别表示预测值和真实值, $\bar{y}$ 表示 $y_t (t = 1, 2, \dots, T)$ 的平均值.

3.3 基于 DRAO 数据集模型评估结果

构建的 BiLSTM-Attention 模型利用加拿大 DRAO 观测的 $F_{10.7}$ 数据集所得性能指标列于表 2. 基于同一数据集, 构建了 RNN 系列预测模型, 即 LSTM 和 BiLSTM 模型. 其中 LSTM 模型已被广泛应用于 $F_{10.7}$ 预测问题中, 并取得了较好的预测效果^[10,11,13,14]. BiLSTM 在 LSTM 强时序处理能力基础上增加了双向机制, 能够发现 LSTM 无法发现的时序规律. 为探究双向结构能否提高 $F_{10.7}$ 预测效果, 构建了 BiLSTM 预测模型. 值得一提的是, 本文在 Luo 等^[15] 构建的 CNN-LSTM 基础上进行改进, 建立了 CNN-BiLSTM 模型. 表 2 为各算法模型在 DRAO 数据集上进行 $F_{10.7}$ 短期和中短期预报的性能比较, 各模型在 DRAO 数据集上的 4~26 天预测结果详见表 2. 由表 2 可知, 在短期预报中, 各模型在第一天的预测效果最好, 其中, BiLSTM-Attention 模型各项指标最佳, RMSE 和 MAE 误差值分别为 17.39 和 5.38, MAPE 为 4.19%, 预测值与真实值的误差最小, 预测

精度最高, 明显优于其他 RNN 系列模型. R^2 值为 0.830, 相关系数 R 高达 0.987, 表明预测值与观测值密切相关.

通过对比表 2 中各个模型在不同预测天数的 MAE 和 MAPE 可知, BiLSTM-Attention 在 DRAO 数据集上的短期预测 (第 1~3 天) 和中短期预测 (第 27 天) 性能表现最佳, 与其他模型相比, 其预测误差更小, 精度更高, 并且与目标值之间的相关系数较高, 表明该模型无论是在短期预测还是在中短期预测中均有良好的表现. 由表 2 中 MAE, MAPE 与相关系数的变化可知, 随着预测天数的增加, 模型的 MAE 逐渐上升, MAPE 值由 4.19% 增长至 10.92%, 预报误差逐渐增大. 相反, 相关系数下降了 0.01, 模型的预测能力有所下降, 但整体预测性能变化幅度不大. 由 LSTM 与 BiLSTM 的结果对比可以看出, 增加双向结构的模型预测性能要优于单向结构, 预测误差和相关系数都有一定幅度的减小和提高, 这与双向结构能够双向提取时序特征, 可以发现单向网络无法找到的时序规律密切相关. 由 BiLSTM 与 BiLSTM-Attention 结果可知, 在时序模型上增加注意力机制对于模型整体性能有一定幅度的提升, 其主要原因是模型在增加注意力机制后, 可以自适应地获取历史 $F_{10.7}$ 时间序列上的关键时间点信息和关键特征信息,

表 2 各模型在 DRAO 数据集上多天预测性能比较
Table 2 Comparison of multi-day forecast performance of each model on DRAO dataset

预测天数	模型	Metrics				
		RMSE /sfu	MAE/sfu	MAPE/ (%)	R^2	R
1	LSTM	18.50	7.55	7.16	0.808	0.986
	BiLSTM	17.46	5.97	5.16	0.828	0.987
	BiLSTM-Attention	17.39	5.38	4.19	0.830	0.987
	CNN-BiLSTM	21.60	8.70	7.13	0.738	0.981
2	LSTM	19.47	9.82	9.79	0.787	0.985
	BiLSTM	19.69	9.27	8.81	0.782	0.985
	BiLSTM-Attention	19.23	7.41	6.08	0.792	0.985
	CNN-BiLSTM	23.27	11.72	11.24	0.696	0.979
3	LSTM	19.98	8.80	7.80	0.775	0.983
	BiLSTM	19.82	8.23	6.80	0.779	0.984
	BiLSTM-Attention	20.14	8.04	6.46	0.772	0.983
	CNN-BiLSTM	23.15	10.81	9.58	0.699	0.978
27	LSTM	25.40	14.54	13.33	0.635	0.974
	BiLSTM	25.92	14.54	12.88	0.620	0.973
	BiLSTM-Attention	25.69	13.18	10.92	0.627	0.973
	CNN-BiLSTM	26.00	13.75	12.25	0.618	0.972

从而指导模型预测, 提升预测效果. 经实验结果可知, 在 $F_{10.7}$ 日值预测问题中, BiLSTM-Attention 拥有较好的预测性能. 为更加直观地展示模型预测性能的优劣, 使用 BiLSTM-Attention 模型对 2021 年 10 月 1 日至 2022 年 10 月 1 日时间段的 $F_{10.7}$ 日值进行测试. 图 7 展示了 BiLSTM-Attention 模型在该时间段 $F_{10.7}$ 数据的预测效果. 对于每个测试点, 通过输入测试点的前 27 天历史数据来预测未来一天的数据. 由图可知, 预测曲线与真实值曲线变化趋势基本一致, 说明预测值与真实值之间误差较小, 预测精度更高.

此外, 预测曲线的总体趋势和局部变化与真实值曲线变化趋势基本相同, 证明了模型对 $F_{10.7}$ 预测的有效性和优越性.

使用 BiLSTM-Attention 模型对 27 天 (2022 年 8 月 19 日至 2022 年 9 月 14 日) 的 DRAO 数据进行预测, 并与真实值进行比较. 模型在 DRAO 数据集上得到的预测曲线与真实值曲线如图 8 所示. 图中红色曲线为 8 月 18 日预测的 8 月 19 日至 9 月 14 日的 $F_{10.7}$ 指数, 蓝色曲线表示真实结果. 由图 8 可以看出, 预测曲线的总体趋势和局部变化与真实值曲线变化

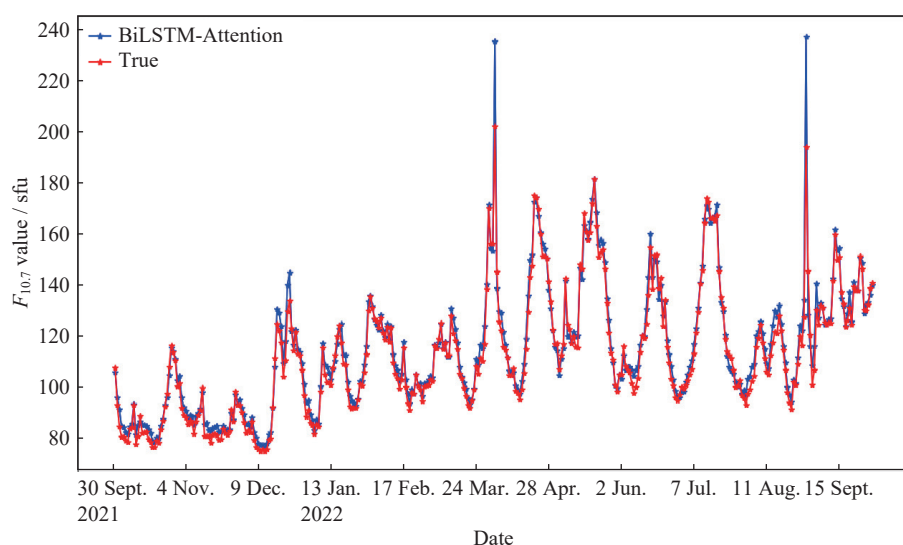


图 7 基于 BiLSTM-Attention 模型的 DRAO 数据预测值与真实值对比

Fig. 7 Comparison of predicted and true values of DRAO data based on BiLSTM-Attention model

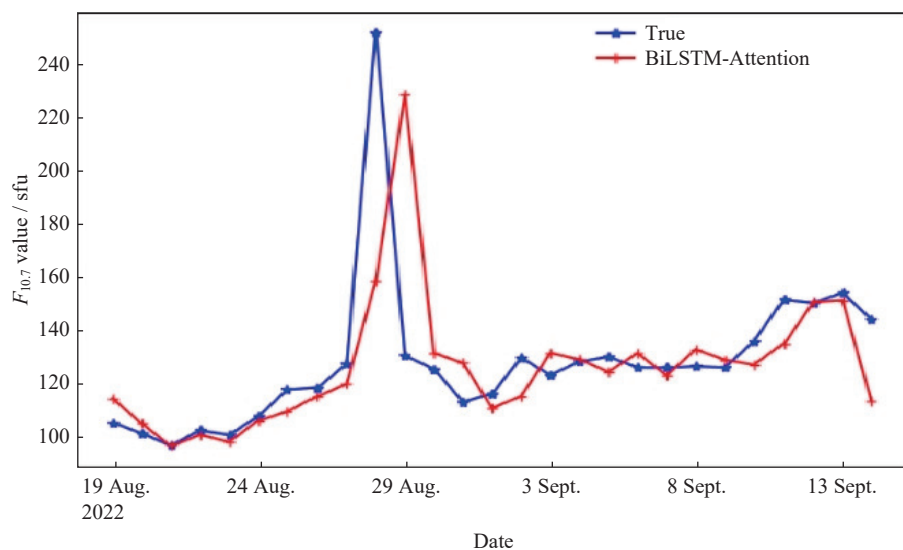


图 8 基于 BiLSTM-Attention 模型的未来 27 天 DRAO 数据 $F_{10.7}$ 指数预测

Fig. 8 $F_{10.7}$ index forecast for the next 27 days based on BiLSTM-Attention model

趋势基本相同, 当 $F_{10.7}$ 的数值急剧变化时, 模型很难进行有效预测, 导致一定的延迟. 但从两个曲线的整体变化情况来看, 模型能够有效实现未来 27 天 $F_{10.7}$ 指数预测.

3.4 基于 L&S 望远镜数据模型评估结果

为了研究 RNN 系列预测模型基于中国自主数据的预测能力, 将处理后的 L&S 数据集作为测试集进行测试. 表 3 为各算法模型在 L&S 数据集上 $F_{10.7}$ 短期和中短期预报的性能比较, 各模型在 L&S 数据集上的 4~26 天预测结果详见表 3. 由表 3 可知, 在短期预测 (1~3 天) 中, 各模型预测性能均有较好的表现, 尤其是在预测未来第一天数据时, 预测精度最高, MAE 和 MAPE 分别为 11.10 和 8.66. 通过对比短期预报和中短期预报 (27 天) 的 MAE 和 MAPE 结果可知, 随着预测天数的增加, 模型的预测性能呈现逐渐下降的趋势, 与在 DRAO 数据集上得到的结论一致, 但 BiLSTM-Attention 模型的 MAE 和 MAPE 相比其他模型较小, 预测精度最高, MAPE 始终在 15% 以内. 通过对比模型在两个数据集上的预测结果, 可以发现模型在 L&S 数据集上的预测精度较低, 而且在中短期预测中 R^2 较低, 这可能是由于 L&S 数据集中的样本较少以及数据处理过程中可能存在误差等因素导致的. 总体而言, BiLSTM 和 BiLSTM-

Attention 在 L&S 数据集上具有较优秀的预测能力, 同时验证了模型的泛化能力较强, 基于国内外 $F_{10.7}$ 数据均能进行有效的预测.

基于 2022 年 1 月 28 日至 2022 年 9 月 25 日的 L&S 数据, 对 BiLSTM-Attention 模型进行了测试, 结果如图 9 所示. 对于每个测试点, 通过输入测试点的前 27 天历史数据来预测未来一天的数据. 蓝色曲线表示真实结果, 红色曲线表示模型预测结果. 由图 9 可知, 预测曲线与真实值曲线变化趋势基本相同, 虽然在某些时间点的预测中准确度较低, 这可能是因为数据处理过程中存在误差导致, 但总体上模型的预测表现良好. 值得注意的是, 模型在预测时出现了一定的时间偏移 (见图 9), 推测是由于 L&S 望远镜数据缺失导致. 当前望远镜数据在预处理过程中使用的校准参数不一致, 且望远镜在试运行阶段出现多次调整, 导致数据缺失和坏点较多, 使得校准后的数据存在一定的误差. 此外, L&S 数据量较少也可能是出现该现象的主要原因之一. 尽管如此, 综合来看 BiLSTM-Attention 预测模型能够在中国数据中适配, 预测精度表现良好.

使用 BiLSTM-Attention 模型对 2022 年 8 月 19 日至 9 月 14 日 27 天的 L&S 数据进行预测, 并与真实值进行比较. 图中红色曲线为 8 月 18 日预测的

表 3 各模型在 L&S 数据集上多天预测性能比较

Table 3 Comparison of multi-day forecast performance of each model on L&S dataset

预测天数	模型	Metrics				
		RMSE /sfu	MAE/sfu	MAPE/ (%)	R^2	R
1	LSTM	14.29	11.26	9.04	0.625	0.993
	BiLSTM	14.23	11.02	8.71	0.628	0.993
	BiLSTM-Attention	14.75	11.10	8.66	0.600	0.993
	CNN-BiLSTM	16.56	13.64	11.45	0.496	0.991
2	LSTM	17.95	13.97	11.30	0.405	0.989
	BiLSTM	16.68	12.49	9.92	0.486	0.991
	BiLSTM-Attention	16.79	13.04	10.46	0.479	0.990
	CNN-BiLSTM	19.23	15.95	13.71	0.316	0.990
3	LSTM	19.14	15.11	12.07	0.325	0.985
	BiLSTM	17.43	13.99	11.39	0.440	0.990
	BiLSTM-Attention	18.16	14.88	12.19	0.392	0.989
	CNN-BiLSTM	20.21	16.40	13.63	0.247	0.987
27	LSTM	24.52	19.84	15.30	-0.065	0.982
	BiLSTM	24.11	19.40	15.26	-0.030	0.982
	BiLSTM-Attention	24.19	19.14	14.75	-0.037	0.982
	CNN-BiLSTM	23.85	19.64	15.87	-0.008	0.982

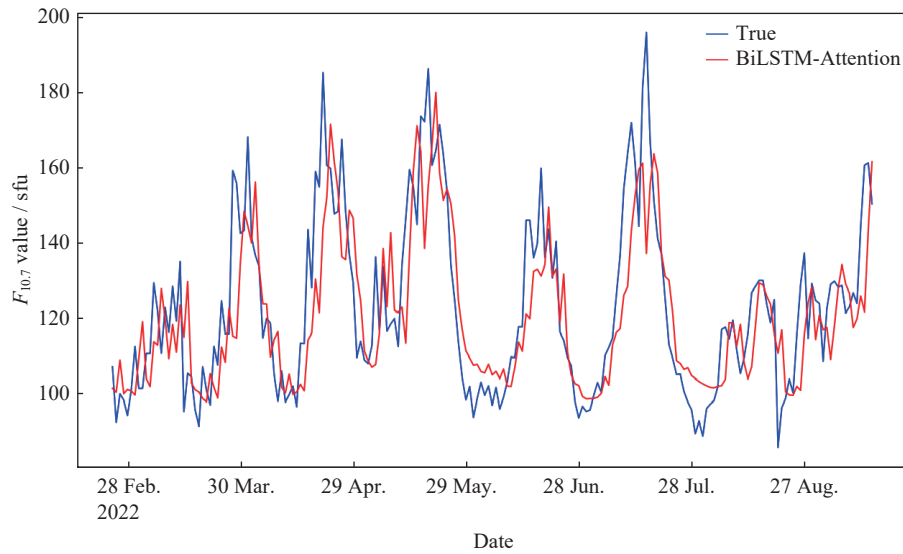
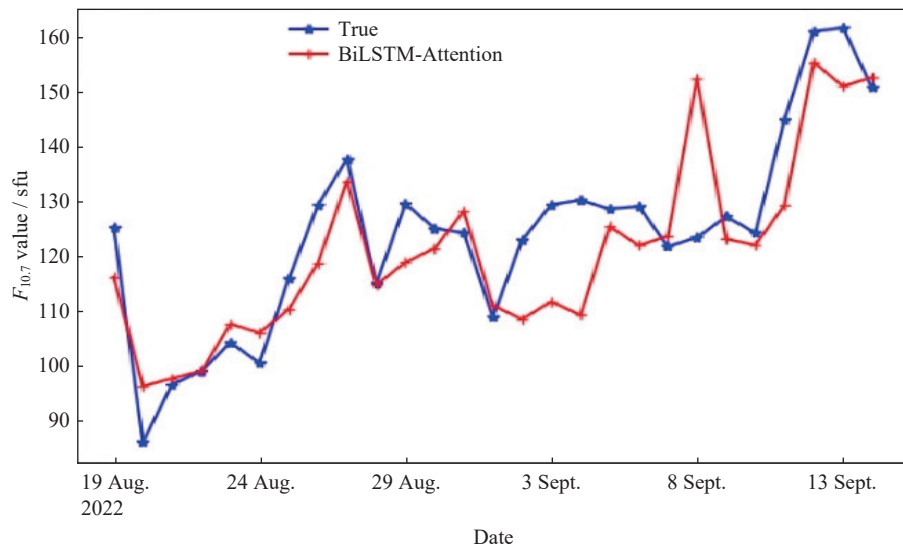


图9 基于 BiLSTM-Attention 模型的 L&S 数据预测值和真实值对比

Fig. 9 Comparison of predicted and true values of L&S data based on BiLSTM-Attention model

图10 基于 BiLSTM-Attention 模型的未来 27 天 L&S 数据 $F_{10.7}$ 指数预测曲线Fig. 10 $F_{10.7}$ Index forecast for the next 27 days based on BiLSTM-Attention model

8月19日至9月14日的结果,蓝色曲线为L&S数据真实结果.在L&S数据集上的预测曲线与真实曲线如图10所示.由图10可以看出,在短期预测中,真实值曲线与预测值曲线吻合程度较好,随着预测天数增加,预测能力有所下降,但整体上表现良好.

4 结论

针对空间天气预报业务中对 $F_{10.7}$ 射电流量准确预测的需求,在分析 $F_{10.7}$ 数据和 LSTM 特点基础上,

提出一种基于 BiLSTM-Attention 的 $F_{10.7}$ 指数预测模型,并构建了 LSTM, BiLSTM, CNN-BiLSTM 一系列预测模型进行对比.在加拿大 DRAO 观测的 $F_{10.7}$ 数据集上,与其他模型对比, BiLSTM-Attention 模型的 RMSE, MAE, MAPE 误差均最小,预测精度最高, R^2 值为 0.830, 相关系数 R 高达 0.987, 预测值与观测值密切相关.针对中国 L&S 望远镜观测的 $F_{10.7}$ 数据,提出了一种 CAC 数据处理方法,经过校准分析,处理后的 L&S 望远镜数据和 DRAO 数据

集有较高的相关性. 基于中国自主数据, 分析和研究了 BiLSTM-Attention 模型和其他 RNN 模型的预测能力, 经对比分析, BiLSTM 和 BiLSTM-Attention 在 L&S 数据集上具有较优秀的预测能力, 同时说明所提模型泛化能力较强, 对于国内外 $F_{10.7}$ 数据源均能进行有效的预测.

未来随着中国数据源的不断扩展, 将继续尝试使用深度学习等方法基于中国自主观测的不同 $F_{10.7}$ 数据源进行预测研究, 例如来自国家卫星气象中心的石岛观测站、国家天文台的正镶白旗观测站以及未来建设的新设备等的的数据.

参考文献

- [1] MA Ruiping, JI Qiao, XU Jiyao. Wavelet analysis of quasi-27-day oscillations in the solar index $F_{10.7}$ [J]. *Chinese Journal of Space Science*, 2007, **27**(2): 89-95 (马瑞平, 纪巧, 徐寄遥. 太阳 $F_{10.7}$ 指数准 27 天振荡的小波分析 [J]. *空间科学学报*, 2007, **27**(2): 89-95)
- [2] LIU Siqing, ZHONG Qiuqin, WEN Jing, et al. Modeling research of 10.7 cm solar radio flux 27-day forecast (I)[J]. *Chinese Journal of Space Science*, 2010, **30**(1): 1-8 (刘四清, 钟秋珍, 温靖, 等. 太阳 10.7 cm 射电流量中期预报模型研究 (I)[J]. *空间科学学报*, 2010, **30**(1): 1-8)
- [3] WEN Jing, ZHONG Qiuqin, LIU Siqing. Model research of 10.7cm solar radio flux 27-day forecast (II)[J]. *Chinese Journal of Space Science*, 2010, **30**(3): 198-204 (温靖, 钟秋珍, 刘四清. 太阳 10.7 cm 射电流量中期预报模型研究 (II)[J]. *空间科学学报*, 2010, **30**(3): 198-204)
- [4] WANG H B, XIONG J N, ZHAO C Y. The mid-term forecast method of solar radiation index[J]. *Chinese Astronomy and Astrophysics*, 2015, **39**(2): 198-211
- [5] GAO Yang, LV Jianyong, WANG Ming, et al. Short-term prediction of solar $F_{10.7}$ radiation flux based on deep learning[J]. *Acta Astronomica Sinica*, 2022, **63**(1): 11 (高扬, 吕建永, 王明, 等. 基于深度学习的太阳 $F_{10.7}$ 辐射通量的短期预报研究 [J]. *天文学报*, 2022, **63**(1): 11)
- [6] VALDÉS J J, NIKOLIĆ L, TAPPING K. Machine learning approaches for predicting the 10.7 cm radio flux from solar magnetogram data[C]//2019 International Joint Conference on Neural Networks (IJCNN). Budapest, Hungary: IEEE, 2019
- [7] HUANG C, LIU D D, WANG J S. Forecast daily indices of solar activity, $F_{10.7}$, using support vector regression method[J]. *Research in Astronomy and Astrophysics*, 2009, **9**(6): 694
- [8] LUO J Q, ZHU H B, JIANG Y, et al. The 10.7-cm radio flux multistep forecasting based on empirical mode decomposition and back propagation neural network[J]. *IEEE Transactions on Electrical and Electronic Engineering*, 2020, **15**(4): 584-592
- [9] ABRAHAM S, ANIYAN A K, KEMBHAVI A K, et al. Detection of bars in galaxies using a deep convolutional neural network[J]. *Monthly Notices of the Royal Astronomical Society*, 2018, **477**(1): 894-903
- [10] YANG Xu, ZHU Yaguang, YANG Shenggao, et al. Application of LSTM neural network in $F_{10.7}$ solar radio flux mid-term forecast[J]. *Chinese Journal of Space Science*, 2020, **40**(2): 176-185 (杨旭, 朱亚光, 杨升高, 等. LSTM 神经网络在太阳 $F_{10.7}$ 射电流量中期预报中的应用 [J]. *空间科学学报*, 2020, **40**(2): 176-185)
- [11] LUO J Q, ZHU L C, ZHU H B, et al. A new approach for the 10.7-cm solar radio flux forecasting: based on empirical mode decomposition and LSTM[J]. *International Journal of Computational Intelligence Systems*, 2021, **14**(1): 1742-1752
- [12] CINAR Y G, MIRISAEI H, GOSWAMI P, et al. Period-aware content attention RNNs for time series forecasting with missing values[J]. *Neurocomputing*, 2018, **312**: 177-186
- [13] LIU H, LIU C, WANG J T L, et al. Predicting solar flares using a long short-term memory network[J]. *The Astrophysical Journal*, 2019, **877**(2): 121
- [14] ZHANG W T, ZHAO X H, FENG X S, et al. Predicting the daily 10.7-cm solar radio flux using the long short-term memory method[J]. *Universe*, 2022, **8**(1): 30
- [15] LUO J Q, ZHU L C, ZHANG K L, et al. Forecasting the 10.7-cm solar radio flux using deep CNN-LSTM neural networks[J]. *Processes*, 2022, **10**(2): 262
- [16] STEVENSON E, RODRIGUEZ-FERNANDEZ V, MINISCI E, et al. A deep learning approach to solar radio flux forecasting[J]. *Acta Astronautica*, 2022, **193**: 595-606
- [17] ZHENG Y F, LI X B, YAN S N, et al. Multiclass solar flare forecasting models with different deep learning algorithms[J]. *Monthly Notices of the Royal Astronomical Society*, 2023, **521**(4): 5384-5399
- [18] SCHUSTER M, PALIWAL K K. Bidirectional recurrent neural networks[J]. *IEEE Transactions on Signal Processing*, 1997, **45**(11): 2673-2681

作者简介



闫帅楠 男, 2001 年 10 月出生于江苏省徐州市, 2023 年 7 月本科毕业于江苏科技大学. 现为中国科学院大学硕士研究生在读, 主要研究方向为空间智能信息处理技术, 人工智能与空间灾害性天气事件预报.
E-mail: shuainanyan2021@163.com



李雪宝 (通信作者) 男, 1986 年 1 月出生于江苏省淮安市. 现为江苏科技大学计算机学院副教授, 硕士生导师, 主要研究方向为人工智能与太阳活动预报、空间灾害性天气预报等.
E-mail: zyf062856@163.com