

何汉圳, 谷升阳, 秦雨松, 刘宇轩. 2026. 基于 CNN-LSTM 的临近空间温度预报研究[J]. 地球与行星物理论评 (中英文), 57(1): 118-125. doi: 10.19975/j.dqyxx.2025-015.

He H Z, Gu S Y, Qin Y S, Liu Y X. 2026. Research on temperature forecasting in near space based on CNN-LSTM algorithm[J]. Reviews of Geophysics and Planetary Physics, 57(1): 118-125 (in Chinese). doi:10.19975/j.dqyxx.2025-015.

基于 CNN-LSTM 的临近空间温度预报研究

何汉圳, 谷升阳*, 秦雨松, 刘宇轩
武汉大学 地球与空间科学技术学院, 武汉 430072

摘要: 临近空间是指地球大气层与太空之间的复杂过渡区域, 其高度介于 20~100 km, 包括平流层、中间层和低热层, 是地球大气系统的重要组成部分. 临近空间不仅受对流层小尺度扰动影响并且受到中高层大气行星波等影响, 另外在该区域中存在大气动力学过程以及各层大气之间的物理化学作用, 这使得对临近空间的研究变得困难. 国内外学者应用机器学习与深度学习算法在空间物理领域进行了大量的研究, 并取得了较好的成果. 但由于大气环境的复杂多变, 在临近空间中利用深度学习算法对环境温度进行预报的工作研究较少. 研究临近空间的温度预报对理解大气动力学过程、大气跨圈层耦合以及分析大气波动有重要的科学意义. 基于 MERRA-2 再分析数据集, 本文采用卷积神经网络-长短期记忆网络 (CNN-LSTM) 算法对临近空间温度进行窗口为 7 天的预报, 其 CNN 部分能够有效提取空间特征, LSTM 部分能够捕捉时间依赖关系. 预报结果与原始数据相比, 相对误差控制在 2% 以内. 在此基础上, 对该算法在季节、高度和纬度上的敏感性展开研究. 分析结果显示北半球夏季和秋季的预报效果更好, 且在低纬度区域预报模型的准确性也较为可靠.

关键词: 临近空间; 温度预报; 机器学习; 深度学习

doi: 10.19975/j.dqyxx.2025-015

中图分类号: P351

文献标识码: A

Research on temperature forecasting in near space based on CNN-LSTM algorithm

He Hanzhen, Gu Shengyang*, Qin Yusong, Liu Yuxuan

School of Earth and Space Science and Technology, Wuhan University, Wuhan 430072, China

Abstract: Near space refers to the complex transition region between the Earth's atmosphere and space, with its height between 20–100 km, including the stratosphere, the mesosphere and the lower thermosphere which are important components of the Earth's atmospheric system. The near space is affected not only by small-scale disturbances in the troposphere but also by planetary waves in the middle and upper atmosphere. There are also atmospheric dynamic processes in the region and physical and chemical interactions between layers of the atmosphere, which make the study of near space difficult. Scholars at home and abroad have done a lot of research on the application of machine learning and deep learning algorithms in the field of space physics, and achieved good results. However, due to the complexity and variability of atmospheric environment, there is little research on using deep learning algorithm to forecast environmental temperature in near space. The study of temperature forecasting in near space is of great scientific significance for understanding atmospheric dynamics, cross-layer coupling and analyzing atmospheric fluctuations. In this paper, CNN-LSTM algorithm is used to forecast temperature in near space

收稿日期: 2025-03-14; 录用日期: 2025-05-28

基金项目: 国家自然科学基金资助项目 (42374195, 42404168); 博士后创新人才支持计划 (BX20230273)

Supported by the National Natural Science Foundation of China (Grant Nos. 42374195, 42404168), and the Postdoctoral Innovation Talent Support Program (Grant No. BX20230273)

第一作者: 何汉圳, 在读硕士研究生, 主要从事利用机器学习预测中高层大气的研究. E-mail:

2023202120011@whu.edu.cn

*通信作者: 谷升阳, 男, 教授, 主要从事中高层大气动力学、中间层/电离层耦合等研究. E-mail:

gushengyang@whu.edu.cn



based on MERRA-2 datasets. The CNN part can effectively extract spatial features and the LSTM part can capture temporal dependencies. Its output window is 7 days, and the relative error is controlled within 2% compared with the original data. On this basis, we do some research on the sensitivity of the algorithm in season, altitude and latitude. The results show that in summer and autumn of the Northern Hemisphere, the forecast results are better, and the accuracy of the forecast model is more reliable at low latitudes.

Keywords: near space; temperature forecasting; machine learning; deep learning

0 引言

临近空间位于地表上方约 20~100 km 处, 包括平流层、中间层和低热层, 是地球大气系统的重要组成部分. 该区域不仅受对流层小尺度扰动的影响, 又受高层大气行星波、重力波等的影响. 由于空间大气动力学过程, 各层大气存在跨圈层耦合等复杂物理化学相互作用, 这使得临近空间大气环境复杂且多变. 应科学研究与实际应用需求, 临近空间环境的预报变得尤为重要, 有必要通过机器学习与深度学习算法对临近空间展开预报工作.

近年来, 国内外研究学者应用机器学习与深度学习算法在空间物理领域进行了大量的研究. 在对流层区域, 通过随机森林算法预测由热带气旋引起的海平面温度变化, 取得了显著成效 (Cui et al., 2023). 基于数据驱动的天气预报使用训练模型, 如 PanguWeather、FourCastNet、GraphCast 等. 虽然这些人工智能天气模型成本低廉且能够提供与传统数值天气模式相当的准确预报, 但是预报的高度限制在对流层. 在平流层中, Wu 等 (2022) 利用深度学习算法对重力波势能进行估算, 与实际值相比误差较小. 在中层大气中, Lu 等 (2024) 证明了利用机器学习模拟参数化地形重力波阻力 (gravitational wave drag, GWD) 和一般环流模式 (general circulation model, GCM) 模拟平流层的可行性. Espinosa 等 (2022) 利用机器学习重力波参数通用化以捕获 QBO 和对增加的 CO₂ 的响应. 综上所述, 机器学习与深度学习在空间物理的应用是可行的和实用的. 传统空间天气研究通常依赖多种方法和手段, 结合观测 (地基和卫星观测)、理论模型与数值模拟来研究地球空间环境. 而机器学习相比于传统的空间天气研究方法, 无需建立复杂的物理模型, 计算资源要求更低、运行更便利以及运行成本更低. 但在临近空间中, 利用机器学习和深度学习对环境温度预报的工作研究较少. Sun 等 (2024) 利用卷积长短期记忆网络 (ConvLSTM) 和卷积门控循环单元 (ConvGRU) 构建了神经网络, 对来自

MERRA2 的十年数据集进行深度学习, 获得了临近空间大气的全球尺度短期预报模型 (每日预报). Chen 等 (2024) 提出了一种温度曲线算子方法, 融合长短期记忆网络 (long short term memory neural network, LSTM) 和其变体双向长短期记忆网络 (BiLSTM) 的能力, 以利用时间序列关联信息和深度算子网络来近似温度曲线, 从而用于平流层大气温度剖面的精细短期预报 (48 小时). 目前应用机器学习对大气温度预报的工作限于短期预报, 在长时间尺度上的预报还比较薄弱.

Adolfs 等 (2024) 使用 LSTM 成功预测了未来 24 小时电离层总电子含量 (total electron content, TEC), 与传统的基于前馈神经网络 (feedforward neural network, FNN) 的模型和 Neustrelitz TEC 模型相比较, LSTM 模型准确率更高. Luo 等 (2023) 提出基于注意力记忆机制的 ConvLSTM 模型 (SAM-ConvLSTM) 来预测未来一天全球电离层总电子含量, 与卷积神经网络 (convolutional neural network, CNN)-长短期记忆网络 (CNN-LSMT) 和 ConvLSTM 模型进行比较, SAM-ConvLSTM 模型的预测效果更加显著.

本研究将结合临近空间大气变化特征, 利用 CNN-LSMT 深度学习算法对临近空间温度进行时间尺度为 7 天的预报, 由于温度数据同时包含时间和空间上的信息, 而卷积神经网络部分能够从输入数据中有效提取局部的空间特征, 长短期记忆神经网络部分能够捕获输入数据中的时间依赖关系. 所以 CNN-LSTM 算法能够合理地应用于临近空间环境温度预报这一工作中, 从而更好地研究临近空间环境变化趋势以及其中的大气动力学机制、跨圈层耦合、大气环流效应.

1 CNN-LSTM 算法

本研究采用 CNN-LSTM 算法, CNN-LSTM 是一种结合了卷积神经网络 (CNN) 和长短时记忆网络 (LSTM) 的深度学习模型, 其优势在于能够很

好地处理包含空间特征和时间依赖两方面的数据. 其中, CNN 的主要作用是从输入数据中提取局部的空间特征, 其原理是通过卷积核扫描输入数据来提取局部特征, 同时保留主要特征, 交替使用 CNN 的卷积层和池化层能够有效提取数据局部特征并降低局部特征维度以减少特征图的尺寸. 由于权值共享的特性, 使用 CNN 可以减少权值数量、降低模型的复杂度. 一维卷积对于时间序列的特征提取的输出为:

$$Y = \sigma(W \cdot X + b) \tag{1}$$

式中, Y 为提取的特征; σ 为 sigmoid 激活函数; W 为权重矩阵; X 为时间序列; b 为偏置向量.

而 LSTM 是一种变体的循环神经网络 (recurrent neural network, RNN), 用于捕获输入数据中的长时依赖关系. 其特点是通过门机制 (输入门、遗忘门和输出门) 有效地解决了标准 RNN 的梯度消失和爆炸问题. 原理示意图如图 1 所示. LSTM 网络中加入的细胞记忆单元, 使得 LSTM 网络具有良好的记忆能力, 被广泛应用在时间序列预测当

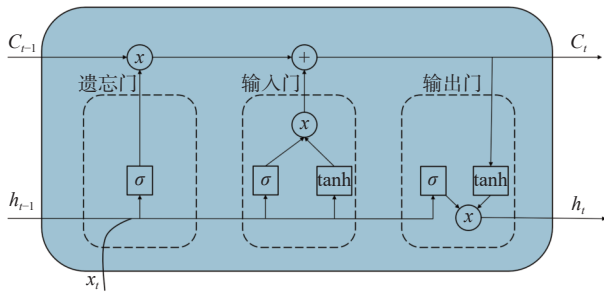


图 1 LSTM 的原理示意图

Fig. 1 The principle diagram of LSTM

中. 其具体计算公式如下所示:

$$f_t = \sigma(W_f [h_{t-1}, x_t] + b_f) \tag{2}$$

$$i_t = \sigma(W_i [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\tilde{C}_t = \tanh(W_c [h_{t-1}, x_t] + b_c) \tag{4}$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t \tag{5}$$

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = o_t \otimes \tanh(C_t) \tag{7}$$

式中, W_f 、 W_i 、 W_c 、 W_o 为权重矩阵; b_f 、 b_i 、 b_c 、 b_o 为对应的偏置向量; \tanh 为双曲正切函数; \otimes 为矩阵相乘; h_t 为上一时刻的输出; f_t 为保留程度值; C_{t-1} 为上一时刻的记忆状态; i_t 为当前时刻状态的添加程度值; \tilde{C}_t 为中间状态; C_t 为当前状态; o_t 为输出程度值; h_t 为当前时刻的输出; x_t 为当前时刻的输入.

CNN 与 LSTM 相结合则可以快速高效地同时处理包含时间和空间信息的输入数据.

CNN-LSTM 算法的流程图如图 2 所示. 该算法主要包括三个部分, 其中输入为处理好的 60 天 MERRA-2 温度数据, 模型网络搭建依次由一层长短期记忆神经网络层 (LSTM Layer)、一层卷积神经网络层 (Conv1d Layer) 和三层密集层 (Dense Layer) 组成. LSTM Layer 主要目的是提取数据的时间信息, 以便分析长时依赖关系. Conv1d Layer 的作用则是提取数据的空间分布特征, 处理局部扰动. 三层 Dense Layer 的核心功能是将输入数据通过一系列权重 (weights) 和偏置 (biases) 进

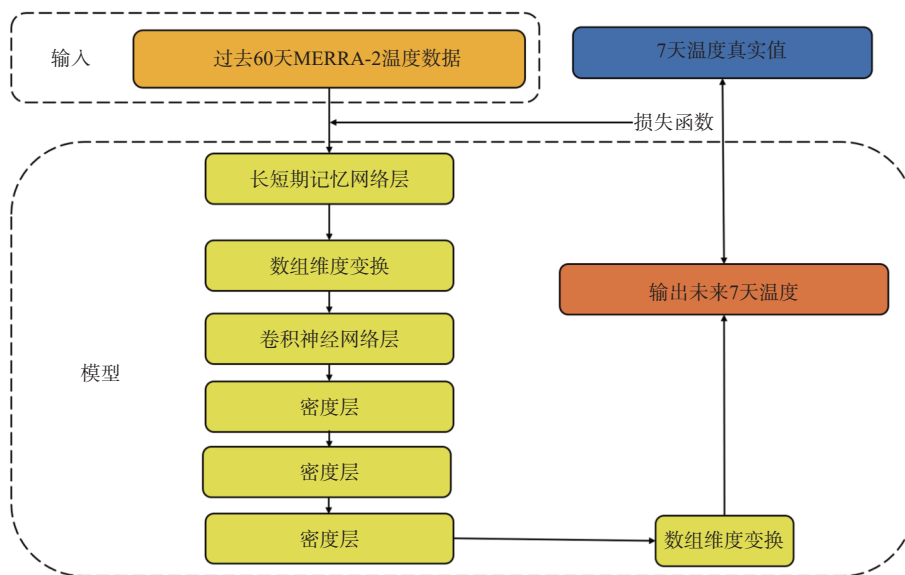


图 2 CNN-LSTM 算法流程图

Fig. 2 Logic loop of CNN-LSTM algorithm

行线性变换, 然后通常会通过一个激活函数进行非线性变换, 以便能够捕捉和学习输入数据中的复杂模式和关系. 数据顺序经过这些网络层, 输出为未来 7 天的温度数据, 通过预测值与真实值之间的 LOSS 来判断循环的次数.

2 数据划分与预处理

为对临近空间进行温度预报, 本研究使用了现代研究和应用回顾性分析第 2 版 (MERRA-2; Gelaro et al., 2017) 的温度数据, 它的垂直覆盖范围从地面到 0.01 hPa 气压层, 包含 72 个模式层, 其空间分辨率为 $1.25^{\circ} \times 1.25^{\circ}$, 时间分辨率是 3 h. 本文选取 2000—2020 年的 MERRA-2 温度数据作为原始数据, 高度选取的范围为 100 hPa~0.01 hPa 的气压层, 对所选数据依次进行以下处理: (1) 纬向平均: 指将经度方向的分量沿着纬圈平均, 得到只含时间、纬度和高度三个分量的数据; (2) 日平均: 原始温度数据时间分辨率为 3 h, 即每天含 8 个时间分量, 将每天的 8 个时间分量相加后平均作为一天的平均气温, 得到温度日平均值; (3) 归一化: 为便于计算, 将数据映射到 0~1 范围之内, 计算公式如下所示:

$$X_{ni} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (8)$$

表 2 模型参数与预测误差

Table 2 Model parameters and prediction errors

批次大小	激活函数	迭代次数	优化器	卷积核大小	均方根误差/K	平均绝对误差/K	决定系数
32	Tanh	50	Adam	7	0.046	0.031	0.938

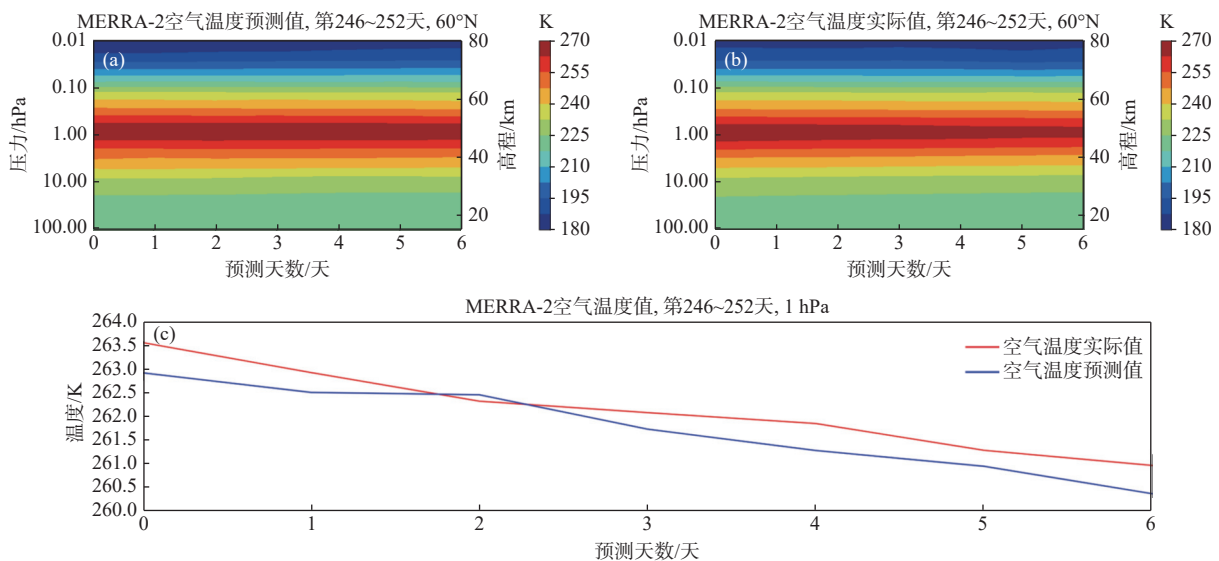


图 3 未来 7 天温度的预测

Fig. 3 Prediction of future 7 days temperature

式中, X_{ni} 为归一化后的数据值, X_i 为原始数据, X_{\min} 为数据样本最小值, X_{\max} 为数据样本最大值.

处理完成的数据按照 8 : 1 : 1 的比例划分成训练集、验证集和测试集. 数据输入的时间窗口为 60 天, 输出的时间窗口为 7 天. 各数据集的形状如表 1 所示.

表 1 各数据集形状

Table 1 The size of all datasets

	输入形状	输出形状
训练集	(6069, 60, 39)	(6069, 7, 39)
验证集	(700, 60, 39)	(700, 7, 39)
测试集	(701, 60, 39)	(701, 7, 39)

3 结果与讨论

3.1 模型参数与误差分析

将经过处理的 MERRA-2 再分析数据集输入到 CNN-LSTM 算法模型中, 模型的参数与预测的误差分析结果如表 2.

由表 2 可知, 均方根误差 (root mean square error, RMSE) 和平均绝对误差 (mean absolute error, MAE) 都较小, 分别为 0.046 K 和 0.031 K, 并且决定系数 (r^2) 为 0.938, 拟合效果较好. 接下

来, 为了定量展示 CNN-LSTM 模型的预测能力, 本文以 60°N 地区为例, 具体分析预测未来 7 天温度的效果, 如图 3 所示.

图 3a 和 3b 分别表示 MEERA-2 再分析数据集在 2020 年 60°N 第 246~252 天 (8 月底 9 月初) 的模型预测空气温度与实际空气温度, 预测窗口长度为 7 天. 从图中, 明显可以看出模型预测值与实际值相差无几, 这是 0.01 hPa~100 hPa 压力层之间的结果, 说明 CNN-LSTM 对空气温度的空间信息的捕捉能力较强, 在不同压力层的预测效果都较好. 单独选取 1 hPa 气压层展示如图 3c 所示, 红线表示空气温度实际值, 蓝线表示空气温度预测值, 可发现两者温度差距始终未超过 1 K, 由此看出, CNN-LSTM 模型对空气温度的时序变化能够有效提取并且正确地预报, 模型的预测效果是较好的.

3.2 季节敏感性与纬度敏感性分析

为了进一步分析该 CNN-LSTM 算法对不同高度和季节的敏感性, 本文将每一次输出的 7 天时间窗口中的第一天滑动提取, 同样以 60°N 为例, 得到 2020 年的温度预测图如图 4 所示. 图 4a 和 4b 分

别表示模型预测值与 MERRA-2 再分析数据集值, 两者整体趋势是吻合的, 但在极大值和极小值处存在偏差, 为探寻这些偏差的分布情况, 将预测值与 MERRA-2 再分析数据集值做差, 得到图 4c 和 4d, 分别表示绝对误差和相对误差, 最大误差值为±10 K, 主要分布在春冬两季, 且在年初 (1—3 月) 中间层呈负向误差, 平流层呈正向误差. 可能是因为该时间段中间层和平流层的大气扰动较为强烈导致模型预测相对不佳, 同时夏季和秋季在 0.1 hPa 气压高度也存在正向偏差, 大小在 3 K, 其余区域在时间和高度上的预测偏差均不超过 3 K, 效果较好. 为更直观地观察这种误差, 取气压高度为 1 hPa 的剖面得到图 4e, 其中红线代表 MERRA-2 再分析数据集值, 蓝线代表模型预测值. 我们能够发现在春冬两季的温度变化较为剧烈, 小尺度扰动居多, 故而模型对极大值和极小值的预测不能很好地体现, 但在夏季和秋季气温变化比较平缓的情况下, 模型能够比较好地贴合实际值. 既然模型对季节变化较为敏感, 我们不妨计算出不同季节的预测相对误差得到图 4f, 红线代表夏季和秋季的温度预测的相对误差, 蓝线代表春季和冬季的温度预测

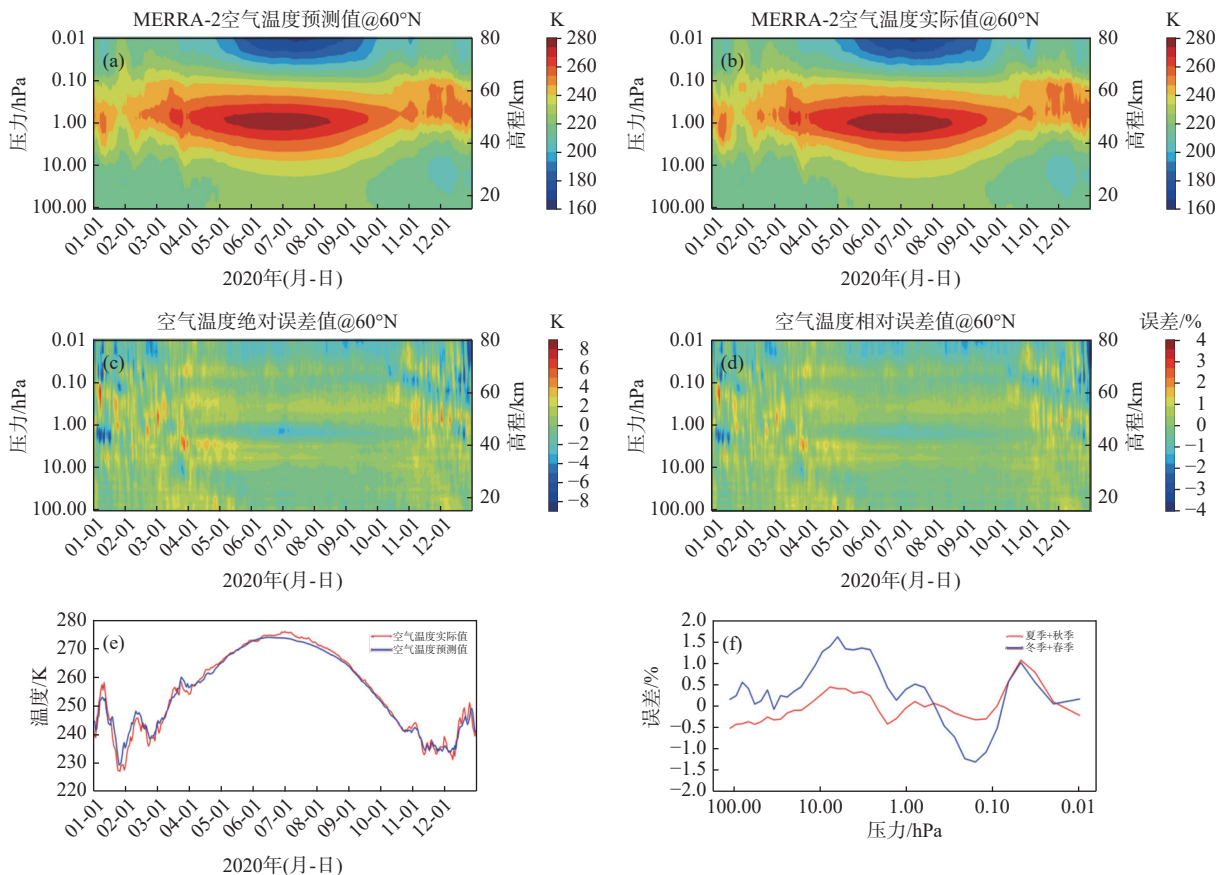


图 4 温度预测的季节和高度敏感性分析

Fig. 4 Seasonal and altitudinal sensitivity analysis of temperature prediction

的相对误差. 其中, 红线在 100 hPa~0.01 hPa 气压高度区间的相对误差值都未超过 1%, 整体曲线平缓, 但在 0.1 hPa~0.01 hPa 气压高度区间 (中间层) 出现明显跳动. 蓝线在平流层下部区域 (100 hPa~10 hPa) 相对误差较小, 在 $\pm 0.5\%$ 左右, 但在

10 hPa~0.01 hPa 气压高度区间大致呈正弦函数分布, 最大相对误差接近 2%.

前文均以 60°N 区域为例展示 CNN-LSTM 模型预测效果, 接下来, 针对该模型的纬度敏感性进行分析. 本文选取了三个不同高度 (10、1、

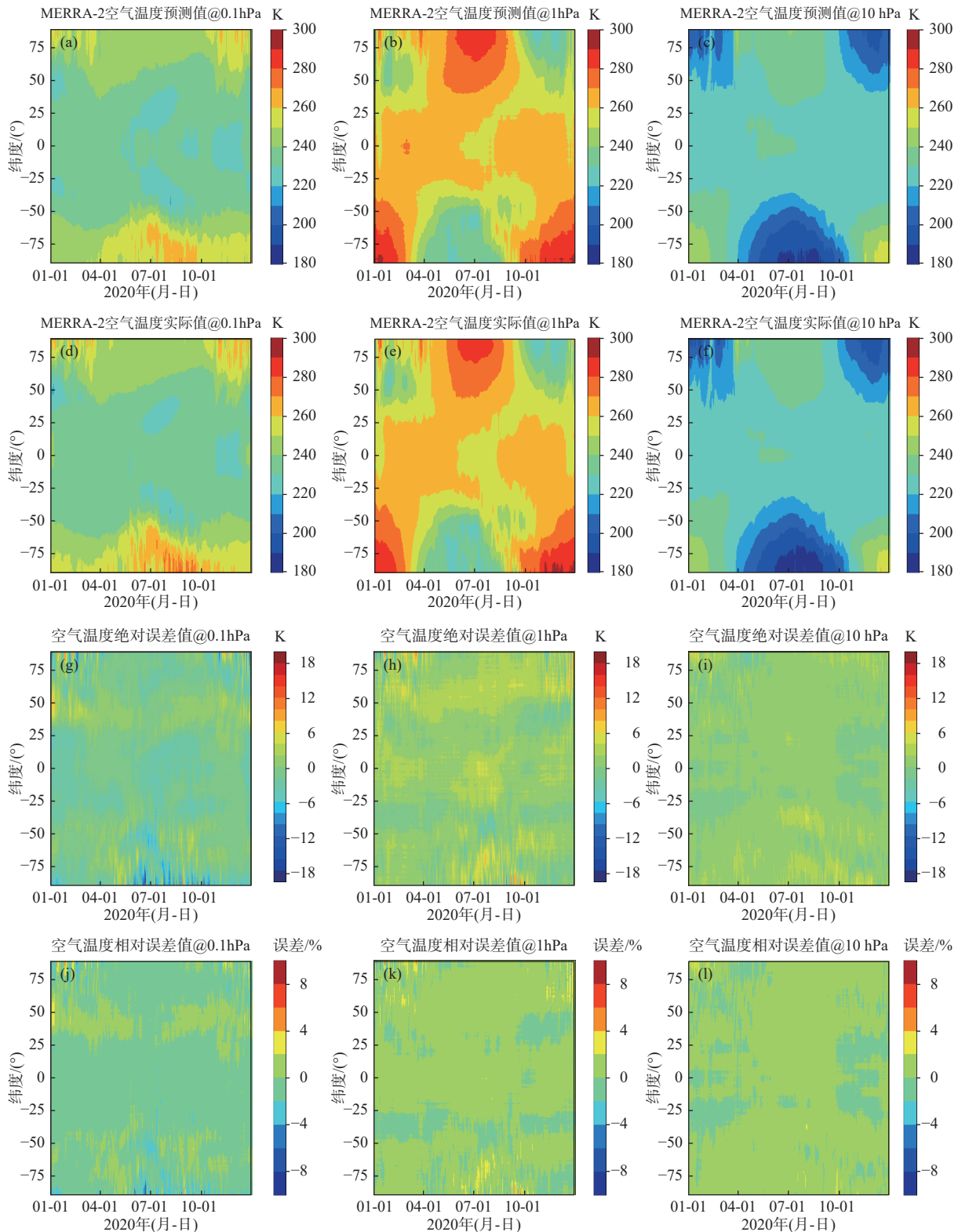


图 5 温度预测的纬度敏感性分析

Fig. 5 Latitudinal sensitivity analysis of temperature prediction

0.1 hPa), 得到 2020 年温度随时间和纬度变化的图如图 5 所示. 左侧一列为 0.1 hPa 气压高度的温度变化图, 从图上可以看出, 低纬度和中纬度区域预测误差较小, 误差主要集中在 12 月—5 月的北半球高纬度区域以及 6 月—11 月的南半球高纬度区域, 且为负向误差, 峰值达 -18 K, 相对误差为 -8%. 中间一列为 1 hPa 气压高度的温度变化图, 与 0.1 hPa 气压高度温度变化图类似, 低纬度和中纬度区域预测误差较小. 误差主要集中在 12 月—5 月的北半球高纬度区域以及 6 月—11 月的南半球高纬度区域, 但此时误差为正向误差, 峰值达 18 K, 相对误差为 8%. 右侧一列为 10 hPa 气压高度的温度变化图, 各时间和纬度上的整体误差均较小, 数值在 ±6 K 以内, 相对误差绝对值不超过 4%, 效果较好. 综上所述, 中间层上部区域 (0.1 hPa) 温度预测存在负向误差, 且主要集中在高纬度区域. 平流层顶区域 (1 hPa) 温度预测存在正向误差, 也集中在高纬度区域. 平流层下部区域 (10 hPa) 温度预测在时间和纬度上的效果均较好, 这是因为此高度的大气是层结稳定大气, 大气变化较为缓慢, 预测算法能够有效地提取到温度的时空变化信息.

为进一步具体分析温度预测的纬度敏感性, 本

文选取三个参数: 均方根误差 (RMSE)、平均绝对误差 (MAE) 和相关系数 (r^2) 作为预测效果的评判指标. 为比较 CNN-LSTM 模型与单一 LSTM 模型的性能差异, 图 6 展示了不同纬度下的 RMSE、MAE 和 r^2 . 如图 6a 和 6b 所示, 红线表示 CNN-LSTM 模型的误差曲线, 蓝线表示 LSTM 模型的误差曲线, 由图可知, RMSE 和 MAE 的曲线趋势类似, 呈 “U” 字形, 在低纬度 (30°S~30°N) 区域误差较小, 随着纬度升高, RMSE 和 MAE 数值均随之增大, 但南北半球存在差异, 北半球预测误差略大于南半球. 并且 CNN-LSTM 模型的 RMSE 和 MAE 均小于单一的 LSTM 模型. 而从图 6c 中可以看到, CNN-LSTM 模型与 LSTM 模型的相关系数曲线类似, 但在各个纬度, CNN-LSTM 模型的相关系数始终大于 LSTM 模型. CNN-LSTM 模型在低纬度区域相关系数较小, 数值均未超过 0.9, 拟合效果不佳, 但在高纬度区域 (60°~80°), 相关系数大于 0.9, 拟合效果较好, 但越靠近极区, 拟合效果就越差. 而 LSTM 模型仅在南半球高纬度区域拟合效果较好, 相关系数在 0.8~0.9 之间. 综上所述, 由于增加了 CNN 模型对空间信息的捕捉能力, CNN-LSTM 模型的整体预测效果优于单一 LSTM 模型.

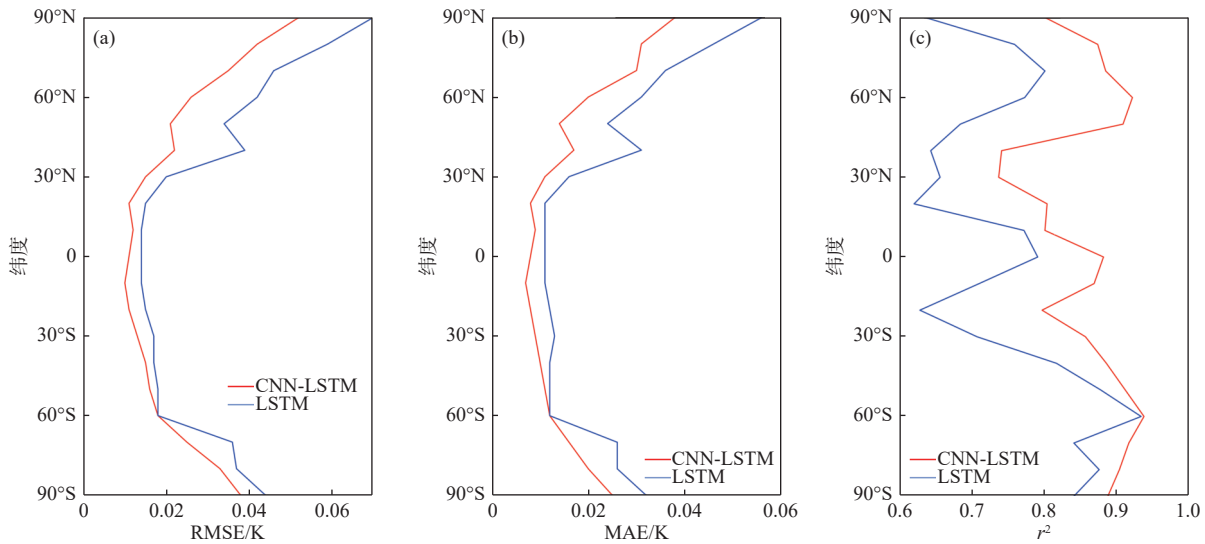


图 6 纬度预测误差与相关系数
Fig. 6 Latitudinal prediction error and correlation coefficient

4 结 论

本研究基于 CNN-LSTM 深度学习算法模型对临近空间的温度进行了 7 天预报. 并分析了其在季节、高度和纬度上的敏感性. 夏季和秋季的预测效果最佳, 在 100 hPa~0.01 hPa 气压高度上的相对

误差值均小于 1%, 稳定性较好. 但在春冬两季中间层较高层 (0.1 hPa~0.01 hPa) 相对误差较大, 接近 2%, 预测效果相对较差. 而关于纬度敏感性的特征比较突出, RMSE 和 MAE 的曲线趋势类似, 呈 “U” 字形, 在低纬度 (30°S~30°N) 区域误差较小, 随着纬度升高, RMSE 和 MAE 数值均随之增

大, 整体误差变化较为平滑. 本文目前只对温度这一变量进行了预报, 未来可能拓展到多变量预测(温度、风场、位势高度、水汽等物理量), 另一方面, 模型结构还有优化的空间, 提高预报的精度和增大预报的尺度(大于 7 天)也是未来需要进行的研究.

数据与来源

本研究使用的现代研究和应用回顾性分析第 2 版(MERRA-2)来源于 Global Modeling and Assimilation Office (<https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2>).

References

- Adolfs M, Hoque M M, Shprits Y Y. 2024. Forecasting 24-Hr total electron content with long short-term memory neural network[J]. *Journal of Geophysical Research: Machine Learning and Computation*, 1(2): e2024JH000123.
- Chen B, Sheng Z, Cui F. 2024. Refined short-term forecasting atmospheric temperature profiles in the stratosphere based on operators learning of neural networks[J]. *Earth and Space Science*, 11(4): e2024EA003509.
- Cui H, Tang D, Mei W, et al. 2023. Predicting tropical cyclone-induced sea surface temperature responses using machine learning[J]. *Geophysical Research Letters*, 50(18): e2023GL104171.
- Espinosa Z I, Sheshadri A, Cain G R, et al. 2022. Machine learning gravity wave parameterization generalizes to capture the QBO and response to increased CO₂[J]. *Geophysical Research Letters*, 49(8): e2022GL098174.
- Gelaro R, McCarty W, Suárez M J, et al. 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2)[J]. *Journal of Climate*, 30(14): 5419-5454.
- Lu Y, Xu X, Wang L, et al. 2024. Machine learning emulation of sub-grid-scale orographic gravity wave drag in a general circulation model with middle atmosphere extension[J]. *Journal of Advances in Modeling Earth Systems*, 16(3): e2023MS003611.
- Luo H, Gong Y, Chen S, et al. 2023. Prediction of global ionospheric total electron content (TEC) based on SAM-ConvLSTM model[J]. *Space Weather*, 21(12): e2023SW003707.
- Sun X, Zhou C, Feng J, et al. 2024. Research on short-term forecasting model of global atmospheric temperature and wind in the near space based on deep learning[J]. *Atmosphere*, 15(9): 1069.
- Wu Y, Sheng Z, Zuo X, et al. 2022. Application of deep learning to estimate stratospheric gravity wave potential energy[J]. *Earth and Planetary Physics*, 6(1): 70-82.