

GAO Zhixu, CHEN Yanhong, AO Xianzhi, WANG Jingjing, WANG Xin, YUE Fulu. Multi-parameter Prediction of Solar Wind Based on Deep Learning (in Chinese). *Chinese Journal of Space Science*, 2026, 46(1): 36–48. CSTR:32142.14.cjss.2025-0022. DOI:10.11728/cjss2026.01.2025-0022

基于深度学习的太阳风多参数预测*

高志旭^{1,2} 陈艳红¹ 敖先志¹
王晶晶¹ 王 昕¹ 岳甫璐³

1(中国科学院国家空间科学中心 北京 100190)

2(中国科学院大学 北京 101408)

3(宇航动力学国家重点实验室 西安 710043)

摘 要 太阳风中高速等离子体流与地球磁层相互作用会引发地磁暴等空间天气事件, 进而影响现代技术系统的稳定运行. 本文利用具有片段嵌入和交叉注意力机制的深度学习模型 TimeXer 来挖掘太阳风参数间的复杂依赖关系, 预测未来 72 h 的 4 个太阳风参数. 实验结果表明, TimeXer 仅利用历史太阳风数据及时间信息, 即可对太阳风速度、动压、质子密度、质子温度分别实现 $47.65 \text{ km}\cdot\text{s}^{-1}$, 1.00 nPa , 3.13 cm^{-3} , $4.49\times 10^4 \text{ K}$ 的预测绝对误差, 与现有的传统及深度学习方法相比, 该模型性能更优, 即使在磁暴期间, 也能较准确地捕获各太阳风参数的整体变化趋势; 基于太阳风参数依赖关系的联合建模预测优于单参数预测; 对模型交叉注意力权重的分析可反映各输入参数对不同太阳风参数预测的相对重要性.

关键词 太阳风参数, 联合预测, 注意力机制, 可解释性

中图分类号 P353

Multi-parameter Prediction of Solar Wind Based on Deep Learning

GAO Zhixu^{1,2} CHEN Yanhong¹ AO Xianzhi¹
WANG Jingjing¹ WANG Xin¹ YUE Fulu³

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

2(University of Chinese Academy of Sciences, Beijing 101408)

3(State Key Laboratory of Astronautic Dynamics, Xi'an 710043)

Abstract When interacting with the Earth's magnetosphere, high-speed plasma flows in the solar wind can trigger space weather events such as geomagnetic storms. Therefore, accurately forecasting so-

* 国家重点研发计划项目 (2024YFC2206902), 中国科学院 B 类先导专项 (XDB0560000) 和中国科学院太空探源专项 (GJ11020405) 共同资助

2025-02-16 收到原稿, 2025-06-27 收到修定稿

©The Author(s) 2026. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

lar wind parameters is critical for early warnings of space weather and the stable operation of modern technological systems. This study employs TimeXer, a deep learning model incorporating patch embedding and cross-attention mechanism, to explore the complex dependencies among solar wind speed, dynamic pressure, proton density, and proton temperature. This model can accurately predict solar wind parameters for the next 72 hours by only using historical solar wind data and time information, and it is also interpretable. Test results during low solar activity level (Year 2021) and high solar activity level (Year 2024) periods demonstrate the following conclusions. TimeXer's Root Mean Square Errors (RMSE) for solar wind speed, dynamic pressure, proton density, and proton temperature are $68.39 \text{ km}\cdot\text{s}^{-1}$, 2.12 nPa , 5.02 cm^{-3} , and $8.83\times 10^4 \text{ K}$, respectively, while the Mean Absolute Errors (MAE) are $47.65 \text{ km}\cdot\text{s}^{-1}$, 1.00 nPa , 3.13 cm^{-3} , and $4.49\times 10^4 \text{ K}$. Compared with traditional and advanced deep learning methods, TimeXer exhibits superior performance, and even can accurately capture the overall variation trends of solar wind parameters during geomagnetic storm. Optimal prediction performance is achieved with a historical input length of 336 hours (corresponding to approximate 14-day quasi-period of the solar wind). The joint modeling prediction based on the inter-parameter dependencies of solar wind parameters is significantly better than the single-parameter prediction. Cross-attention weight analysis reveals that the four solar wind parameters contribute similarly to proton temperature and solar wind speed predictions. The solar wind speed and proton temperature contribute more to the prediction of proton density, while the proton temperature, solar wind speed, and annual time information have a more substantial influence on the prediction of solar wind dynamic pressure. Moreover, the importance of time information grows with increasing scales of time information.

Key words Solar wind parameters, Joint prediction, Attention mechanism, Interpretability

0 引言

太阳风是从太阳日冕层持续向外扩散的高速等离子体流, 分布于太阳与地球等各行星之间的行星际空间, 并通过与地球空间环境的相互作用, 对卫星通信、导航系统、地面电网、载人航天等现代技术产生显著的影响^[1,2]. 例如太阳风高速流 (High-speed Solar wind Streams, HSS) 与地球磁层的相互作用会引发持续增强的亚暴活动^[3], 从而导致外辐射带中相对论电子通量的显著增加、热层密度和卫星阻力的上升^[4-6], 增加卫星遭受充电效应的风险并可能造成其性能下降或永久损伤^[7,8]. 此外, 快速太阳风追赶前方慢速太阳风所形成的共转相互作用区 (Corotating Interaction Regions, CIR), 可能会引起地磁暴, 促使电离层中形成强电流, 进而在地面高压输电系统或其他高导电性基础设施中诱发地磁感应电流, 对电网、管道及铁路系统造成危害^[9]. 另外, 由于太阳风与各类空间天气事件间存在密切联系, 其参数常作为空间环境研究与空间天气建模的关键变量. 例如 Shen 等^[10] 以太阳风参数作为三维磁流体力学模拟的背景条件, 验证了日冕物质抛射在日球层中可能发生超弹性碰撞;

Hu 等^[11] 通过将太阳风速度与密度作为广义回归神经网络的输入参数, 对紫外极光强度进行建模, 实现了 0.54 的平均结构相似度预测. 因此, 提前快速且准确地获知太阳风参数对于提高空间天气建模和预警系统能力具有重要意义.

传统的太阳风预测模型主要包括物理模型和经验模型. 其中物理模型主要是基于磁流体力学 (Magnetohydrodynamic, MHD) 原理, 通过求解太阳风等离子体的动力学方程来模拟太阳风的演化与传播. 典型的 MHD 模型包括美国国家海洋和大气的管理局 (NOAA) 空间天气预报中心使用的 ENLIL 模型^[12,13], Detman 等^[14] 开发的混合日球层建模系统 (Hybrid Heliospheric Modeling System, HHMS) 等. 经验模型则主要基于历史观测数据和统计分析, 利用太阳活动指数、日冕观测等信息建立太阳风的经验关系, 例如 Wang-Sheeley (WS) 模型^[15]、Wang-Sheeley-Argue (WSA) 模型^[16] 和冕洞边界距离模型 (Distance from the Coronal Hole Boundary, DCHB)^[17].

近年来, 随着机器学习技术的迅速发展, 数据驱动的深度神经网络模型逐渐成为太阳风预测领域的研究热点并展现出巨大的潜力^[18-21]. 相比传统的 MHD 物

理模型与经验模型, 深度学习方法能够自动从复杂动态变化的太阳风观测数据中提取特征, 挖掘传统方法难以捕捉的模式和规律, 同时具有高效并行计算的优势. 例如 Chen 等^[22] 利用 Kp 指数、 Dst 指数、 AE 指数、 $F_{10.7}$ 和太阳黑子数等地面观测数据构建了基于人工神经网络 (ANN) 的太阳风速度预测模型, 其提前 3 h 的太阳风速度预测相关系数达到 0.82. Sun 等^[23] 综合考虑太阳风速度、动压、磁场、行星际日冕物质抛射状态等多个参数, 通过太阳极紫外图像提取冕洞信息, 提出了基于图注意力机制的 GTA (Graph-Temporal-AR) 模型, 并在 2017 年的提前 96 h 太阳风速度预测中实现了 $74.82 \text{ km}\cdot\text{s}^{-1}$ 的平均绝对误差. Lin 等^[24] 利用太阳磁场和太阳风观测数据, 提出了一种基于卷积神经网络 (CNN) 和全球日冕磁场外推 (例如 PFSS) 模型的太阳风速度预测方法, 该方法在 96 h 的预测窗口上优于 WSA 等传统经验模型.

尽管基于机器学习的太阳风建模和预测已取得诸多重要进展, 但该领域目前仍存在一些亟待解决的问题: 一方面, 太阳风速度、动压、温度、密度等参数作为研究太阳风变化的关键参数, 在大多数太阳风预测工作中, 仅将太阳风速度作为预测输出对象, 难以完整地描述太阳风特性; 另一方面, 几乎所有的太阳风预测模型都不同程度地依赖于 Kp 指数、太阳图像等外部观测数据, 虽然模型通过融合这些外部信息可以更准确地预测太阳风参数, 但这些外部观测数据的发布时间通常具有不同程度的延迟, 对太阳图像等数据需要复杂的预处理, 这些因素使得基于机器学习的太阳风预测模型的高效、易用性大幅下降, 制约了其在实际空间天气预报中的应用.

针对上述问题, 本文利用 1999 年 1 月至 2024 年 11 月间 1 h 分辨率的太阳风速度、动压、质子温度和质子密度数据构建数据集, 采用具有片段嵌入和交叉注意力机制的深度学习模型 TimeXer 深入分析各太阳风参数间的复杂关系, 该模型仅需要历史太阳风数据和时间信息即可实现未来 72 h 太阳风参数的高准确率预测. 为验证模型性能, 在太阳活动低年 (2021 年) 和太阳活动高年 (2024 年) 进行预测评估, 选取 2024 年 5 月 11 日 10:00 UT 至 2024 年 5 月 14 日 09:00 UT 磁暴期间的预测案例进行分析, 并与

27 天持续性模型、ARIMA、DLinear、TimesNet、Transformer 和 TimeMixer 等传统及先进的深度学习模型方法进行对比. 此外, 讨论分析不同历史数据输入长度对模型预测性能的影响、太阳风多参数联合预测的有效性、模型的参数敏感性以及各太阳风参数与时间信息在预测过程中的相对重要性.

1 数据

1.1 数据表征

太阳风的动力学特征可由以下 4 个关键参数表征^[3,25,26].

(1) 太阳风速度. 作为衡量太阳风能量与动量的重要指标, 其变化直接影响太阳风到达地球的时间及对地球磁层的作用效果.

(2) 太阳风动压. 反映太阳风对地球磁层的动态压力, 剧烈的太阳风动压脉冲会压缩磁层, 导致磁层顶电流瞬间大幅增加, 进而引发地磁急始事件.

(3) 质子温度. 该参数是研究太阳风加热机制与能量传输过程的重要参数, 不同来源的太阳风具有不同的温度特征, 即来源于日冕洞的高速太阳风往往温度较高, 而来源于日冕中开放与闭合磁场交界区域的慢速太阳风温度则较低.

(4) 质子密度. 直接影响太阳风的动压和粒子通量, 高密度太阳风可增强磁层压缩效应并提升磁暴风险.

OMNI 数据集汇集了多个近地轨道或 L1 (拉格朗日点) 航天器观测到的太阳风磁场和等离子体参数, 这里使用从 OMNIWeb^{*} 获取的每小时平均太阳风数据, 并选取太阳风速度 v_{sw} 、太阳风动压 P_{sw} 、质子温度 T_p 、质子密度 D_p 作为研究对象. 数据集的时间范围为 1999 年 1 月至 2024 年 11 月, 时间分辨率为 1 h.

1.2 数据预处理

1.2.1 时间信息选取

为增强模型对太阳风参数时序中多尺度周期性及趋势性特征的捕捉能力, 选取一天中的小时、一个月中的天数和一年中的天数等作为时间信息, 并对这些变量进行归一化处理, 将其映射至 $[-0.5, 0.5]$ 区间, 计算公式为

* <https://omniweb.gsfc.nasa.gov/form/dx1.html>

$$H_t = \frac{h_t}{23} - 0.5, h_t \in [0, 23]; \quad (1)$$

$$M_t = \frac{m_t - 1}{30} - 0.5, m_t \in [1, 31]; \quad (2)$$

$$Y_t = \frac{y_t - 1}{365} - 0.5, y_t \in [1, 366]. \quad (3)$$

其中, h_t 表示时间 t 对应的小时值, m_t 为时间 t 对应的月份中的天数, y_t 表示时间 t 对应的一年中的天数.

1.2.2 异常值处理及标准化

采用线性插值方法对数据中的缺失值进行填充, 保证输入序列的连续性, 降低数据缺失对模型建模的影响. 需要注意的是, 尽管插值能够对缺失数据进行合理估计, 但其结果与真实观测值仍存在一定误差, 因此在模型的评估阶段, 首先对插值部分进行屏蔽, 以确保评估结果的可靠性. 其次, 由于各太阳风参数的量纲与数值范围存在较大差异, 这可能会影响模型训练的收敛速度与数值稳定性. 因此, 对各太阳风参数分别进行 Z-score 标准化, 将各参数的数据分布调整至相似的尺度, 消除数值范围上的不一致性, 避免模型在训练过程中偏向某些参数. 此外, 为防止数据泄露, 使用训练集的均值与方差对整个数据集进行标准化处理.

1.2.3 数据集构建

将 1999—2020 年数据作为训练集, 2022—2023 年数据作为验证集, 将太阳活动水平较低的 2021 年和太阳活动水平较高的 2024 年数据作为测试集. 模型的输入包括过去 336 h 的太阳风参数 (v_{sw} , P_{sw} , T_p , D_p) 及对应的时间信息 (H , M , Y), 输出为未来 72 h 的太阳风参数 (v_{sw} , P_{sw} , T_p , D_p). 在样本构建方式上, 这里采用 Liu 等^[27] 在电离层 TEC 预测研究中使用的滑动窗口采样方法, 即在每个数据子集中通过滑动窗口进行采样. 滑动窗口的总长度为 408 h, 其中前 336 h 的数据作为模型的输入, 后 72 h 的数据作为预测目标. 滑窗步幅为 1 h, 通过逐步移动窗口的方式生成样本集合, 最终获得 227520 个训练样本, 17112 个验证样本, 15912 个测试样本.

2 预测方法

2.1 TimeXer 模型

研究表明, 在 MHD 过程和能量传输机制的作用下, 太阳风各参数间存在着密切的依赖关系^[2,28,29]. 这

一现象直观反映在数据层面, 如图 1 所示. 数据集中各太阳风参数间存在不同程度的相关性, 其中 T_p 与 v_{sw} 的相关系数为 0.66, D_p 与 P_{sw} 的相关系数高达 0.71. 因此, 为了充分挖掘太阳风参数间的复杂关联性, 采用基于 Transformer 架构的 TimeXer 模型^[30] 对太阳风参数进行建模, 该模型通过交叉注意力机制建模多变量相关性, 利用自注意力机制捕获变量的时序依赖, 并通过不同的嵌入策略捕捉数据的不同层次表示.

TimeXer 模型的整体结构如图 2 所示, 其建模流程如下.

(1) 变量嵌入. 模型通过片段嵌入与序列级嵌入, 获取不同时间尺度下的变量表示. 首先, 对历史输入数据应用可逆实例归一化 (RevIN), 缓解数据中的分布漂移问题^[31]. 对历史输入时序应用片段嵌入, 提取序列的局部特征并减小计算复杂度, 具体而言, 将每个太阳风参数时序划分为不重叠的片段, 片段长度为 L_p , 然后将每个片段 s_i 投影为一个 d_v 维的时间编码单元, 该过程可表示为

$$\{s_1, s_2, \dots, s_N\} = (\mathbf{x}')_{\text{Patchify}}, \\ \mathbf{P} = (s_1, s_2, \dots, s_{N'})_{\text{PatchEmbed}}. \quad (4)$$

其中, $\mathbf{x}' \in \mathbf{R}^{L \times 4}$ 表示长度为 L_h 的 4 个历史太阳风参数序列; $\mathbf{P}$ 表示片段嵌入后的太阳风参数; $N' = L_h/L_p$ 为序列片段的总数; Patchify 表示序列分段操作; PatchEmbed 表示可训练的线性投影以及基于正弦和余弦函数的位置编码操作^[32]. 为了获得全局视角下的变量表示, 采用序列级嵌入方法, 将长度为

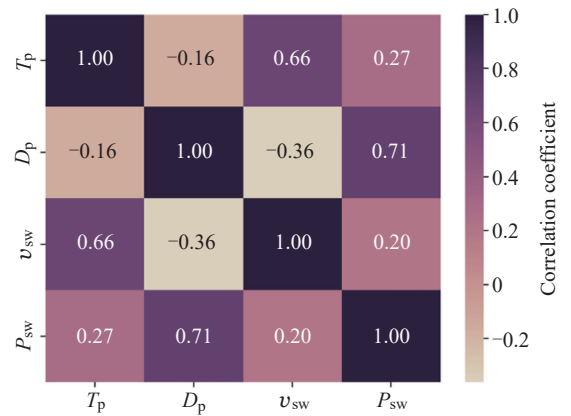


图 1 各太阳风参数间的 Pearson 相关系数

Fig. 1 Pearson correlation coefficients among solar wind parameters

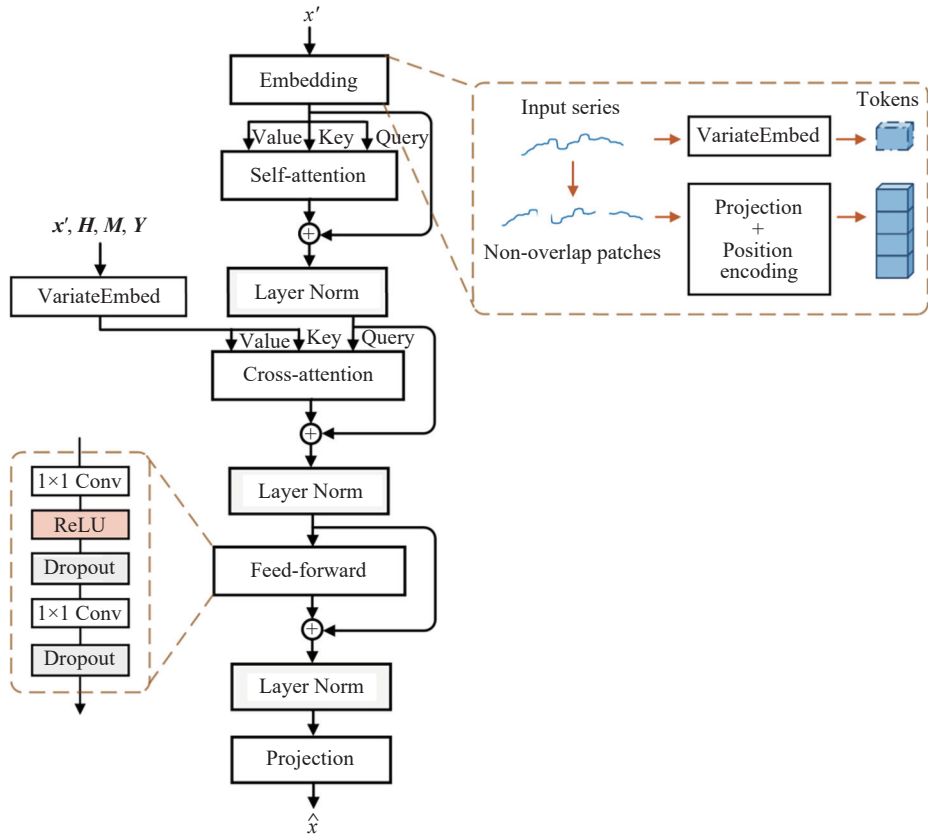


图 2 TimeXer 模型的整体结构

Fig. 2 Overall architecture of the TimeXer model

L_h 的整段序列直接映射为 d_v 维变量的全局表示 V_{em} , 该过程可表示为

$$V_{em} = (x')_{VariateEmbed}. \quad (5)$$

其中, VariateEmbed 表示可训练的嵌入矩阵. 此外, 模型将历史太阳风参数 x' 与时间信息 H, M, Y 沿变量维度进行拼接后, 通过一个线性层将长度为 L 的序列映射至 d_v 维隐空间, 该过程的表达式为

$$V_T = [x', H, M, Y]_{Linear}. \quad (6)$$

其中, $[\cdot]$ 表示拼接操作, Linear 表示线性映射操作.

(2) 变量内自注意力. 自注意力层首先将 P 与 V_{em} 沿序列维度拼接, 再通过自注意力机制捕捉变量内部的时序依赖关系, 然后进行层归一化, 得到自注意力输出 $\hat{P}$ 和 $\hat{V}_{em}$, 该过程可表示为

$$\hat{P}, \hat{V}_{em} = [P, V_{em}] + [P, V_{em}]_{Self-Attn}. \quad (7)$$

其中, Self-Attn 表示变量内部的多头自注意力操作.

(3) 变量间交叉注意力. 为了建模变量间的依赖关系, 交叉注意力层以 $\hat{V}_{em}$ 作为查询 (Query), 以 V_T 作

为键 (Key) 和值 (Value), 并通过层归一化得到更新后的全局变量表示 $\tilde{V}_{em}$, 该过程可表示为

$$\tilde{V}_{em} = \hat{V}_{em} + \left(\hat{V}_{em}, V_T, V_T \right)_{Cross-Attn}. \quad (8)$$

其中, Cross-Attn 表示变量间的多头交叉注意力操作.

(4) 预测和优化. 采用包含两个一维卷积 (核尺寸为 1×1) 的前馈网络对 $\tilde{V}_{em}$ 进行非线性变换, 通过一个线性层将其映射为预测值, 并利用 L2 损失函数衡量预测值与真实值之间的差异, 即

$$L = \sum_{i=1}^S \|x_i - \hat{x}_i\|_2^2. \quad (9)$$

其中, $x \in \mathbf{R}^{L_f \times 4}$ 为未来 L_f 个时间步的太阳风参数真实值; $\hat{x} \in \mathbf{R}^{L_f \times 4}$ 为预测值.

2.2 评价指标

为评估模型对太阳风参数的预测性能, 采用均方根误差 RMSE (δ_{RMSE}) 与平均绝对误差 MAE (δ_{MAE}) 作为评价指标, 其计算公式为

$$\delta_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2}. \quad (10)$$

$$\delta_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i|. \quad (11)$$

其中, n 表示测试样本的数量. RMSE 和 MAE 的值越小, 代表模型预测性能越好.

3 实验结果与讨论

3.1 基线模型与实验设置

这里选择以下 6 种预测模型方法作为基线与 TimeXer 进行对比.

(1) 27 天持续性模型 (27-day Persistence)^[33]. 假设当前近地太阳风条件与 27 天前相同, 对于任意时刻的太阳风参数 $\hat{x}(t)$, 直接采用 27 天前的对应值作为预测值, $\hat{x}_t = x_{t-w}$, 其中, 历史输入窗口长度 $w = 27 \times 24 = 648$ h.

(2) ARIMA 模型^[34]. 广泛用于时间序列分析和预测的统计中, 由自回归 (AR)、差分 (I) 和移动平均 (MA) 三部分组成.

(3) 基于线性方法的 DLinear 模型^[35]. 将时间序列分解为趋势成分和残差成分, 分别用单层线性网络进行建模, 最后将两部分结果相加得到预测值, 模型结构简单、计算高效且具有较强的可解释性.

(4) 基于 CNN 的 TimesNet 模型^[36]. 通过将时间序列转换到时域和频域, 并引入二维卷积处理时间序列的时频表示, 使其能够同时建模时间序列的时间依赖性和周期性特征.

(5) 经典的 Transformer 模型^[32]. 以自注意力机制为核心, 通过编码器层捕获各变量之间以及时间步

之间的关联, 然后利用解码器生成未来时刻的预测值.

(6) 基于线性方法的 TimeMixer 模型^[37]. 利用过去可分解混合模块提取历史序列中不同尺度的季节性和趋势性信息, 并通过未来多预测器混合模块融合多尺度信息进行未来序列的预测, 在一系列广泛的基准测试中展现出卓越的长期和短期预测性能与效率.

在实验设置上, 27 天持续性模型的历史太阳风数据输入长度为 648 h, 其他模型均为 336 h, 预测输出长度均为 72 h. 针对每个模型, 分别调优各项参数以获得最佳性能. 对于 TimeXer, 设置了批量大小为 48, 学习率为 0.001, 片段长度 $L_p = 8$, 嵌入维度 $d_v = 36$. 由于多头注意力机制允许模型在不同的表示子空间中并行地捕捉不同尺度和类型的依赖关系, 因此, 设置注意力头数为 8, 以更全面地捕捉多元时序数据中的复杂依赖关系. 此外, 在模型训练过程中应用早停机制, 若验证损失连续 7 次迭代未改善, 则提前终止训练以防止模型过拟合.

3.2 不同模型对未来 72 h 太阳风参数的预测

利用 TimeXer 模型以及 3.1 节中描述的 6 种基线方法对未来 72 h 的太阳风速度 v_{sw} 、动压 P_{sw} 、质子温度 T_p 、质子密度 D_p 进行预测, 在测试集上的预测结果列于表 1. 由表 1 可知, 所有深度学习方法在 72 h 太阳风参数预测任务上的表现均显著优于简单的 27 天持续性模型和 ARIMA 统计模型, 这表明深度学习方法在捕捉太阳风参数的非线性变化和复杂时变特征方面具有明显优势. 27 天持续性模型忽略了太阳活动区域的变化, 而 ARIMA 模型受限于对数据线性特征的假设, 难以适应太阳风参数的非平稳特

表 1 7 个模型在测试集上的 72 h 预测性能
Table 1 72-hour prediction performance of seven models on the test set

Model	$v_{\text{sw}}/(\text{km}\cdot\text{s}^{-1})$		P_{sw}/nPa		$T_p (\times 10^4)/\text{K}$		D_p/cm^{-3}	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
27-day persistence	106.09	75.72	3.06	1.45	12.22	6.20	7.28	4.68
ARIMA	87.57	60.41	5.61	1.48	118.83	7.31	7.40	4.59
DLinear	69.97	52.08	2.14	<u>1.03</u>	<u>8.87</u>	4.82	5.29	3.37
TimesNet	<u>68.53</u>	<u>48.98</u>	2.15	1.05	8.93	<u>4.62</u>	<u>5.04</u>	<u>3.17</u>
Transformer	69.75	50.23	<u>2.13</u>	1.04	8.97	4.90	5.16	3.28
TimeMixer	72.63	51.64	2.15	1.07	8.95	4.71	5.23	3.40
TimeXer	68.39	47.65	2.12	1.00	8.83	4.49	5.02	3.13

注 加黑部分为最优结果, 加下划线部分为次优结果.

性, 因此预测精度较低. 值得注意的是, TimeXer 在所有太阳风参数的预测中均取得最佳结果, 其预测误差分别为: 太阳风速度 RMSE 为 $68.39 \text{ km}\cdot\text{s}^{-1}$, MAE 为 $47.65 \text{ km}\cdot\text{s}^{-1}$; 太阳风动压 RMSE 为 2.12 nPa , MAE 为 1.00 nPa ; 质子温度 RMSE 为 $8.83 \times 10^4 \text{ K}$, MAE 为 $4.49 \times 10^4 \text{ K}$; 质子密度 RMSE 为 5.02 cm^{-3} , MAE 为 3.13 cm^{-3} . 相比其他深度学习模型, TimeXer 模型在太阳风参数预测任务中表现出更优异的预测性能, 这主要归因于以下几个方面: 有效的嵌入策略, TimeXer 中同时采用片段嵌入和序列级嵌入, 以捕捉时间序列数据的不同层次表示, 从而使模型能够更好地学习太阳风参数中的局部动态与全局变化趋势; 变量内自注意力机制, 太阳风具有多尺度的时间变化特征 (日变化、27 天周期等), TimeXer 通过变量内多头自注意力来学习时序内的多尺度时间模式, 使模型既可以关注细粒度的局部变化, 又能够提取全局趋势与周期性模式; 变量间交叉注意力机制, 各太阳风参数间存在复杂的动态关联, TimeXer 通过变量间多头交叉注意力对变量间的复杂依赖关系进行建模, 实现跨变量的信息融合, 提升模型的预测精度.

为直观比较不同模型对 4 个太阳风参数的预测结果, 图 3 给出了 7 个模型在 2021 年测试集上的逐 72 h 预测 (每间隔 72 h 进行一次预测) 结果. 由图 3 可知, 相比其他基线模型, TimeXer 均能更准确地捕捉到不同太阳风参数的演变趋势. 尤其是在太阳风速度预测中, 其预测值与实际观测值高度吻合; 而在太阳风动压、质子温度和质子密度等参数的预测中, 由于这些参数在演变过程中存在急剧上升和下降的尖峰, 这在没有物理先验知识约束的情况下, 增加了预测的难度, 使得模型对高值区域的预测误差较大, 但是总体上仍能较好地反映参数的整体变化趋势. 上述结果表明, TimeXer 模型在处理多变量太阳风数据时, 能够有效提取和利用各参数之间的内在关联性, 从而显著提高预测精度.

为进一步分析 TimeXer 模型在磁暴期间的预测效果, 选取 2024 年 5 月 11 日 10:00 UT 至 2024 年 5 月 14 日 09:00 UT 的测试数据作为分析案例, 在此期间, 最大 Kp 指数达到 9, 最小 Dst 指数为 -326 nT . 根据图 4 中的预测结果可以发现, 相比 27 天持续性模型和 ARIMA 等传统模型, TimeXer 可以更准确地捕获各太阳风参数在磁暴期间的恢复下降趋势, 而与

其他先进的深度学习模型相比, TimeXer 模型可以更细致地刻画各太阳风参数的变化细节, 但是由于模型缺乏物理先验知识的约束和更多观测信息的支持, TimeXer 对于磁暴期间的太阳风参数中快速上升至高值区域的预测精度仍有待进一步提升.

3.3 不同历史数据输入长度对模型预测性能的影响

理论上分析, 增加历史回溯窗口的长度可以扩大模型的感受野, 从而提升模型的预测性能^[35,38], 但是太阳活动在不同时间尺度上的周期性可能会对模型的预测效果产生影响. 因此, 为进一步分析历史数据输入长度对模型预测性能的影响, 设置历史回溯窗口长度 $L_h = \{24, 48, 96, 192, 336, 648, 1296\}$, 预测输出长度 $L_f=72$, 图 5 给出了 TimeXer 模型在不同历史回溯窗口长度下的预测结果, 图 5 中 RMSE 为 4 个太阳风参数标准化后计算的平均 RMSE. 由图 5 可知, 随着 L_h 的增加, 模型整体的预测误差 (RMSE) 迅速降低, 在 $L_h=336$ 时达到最小值, 随后 RMSE 开始逐渐上升. 背景太阳风依赖于日面上冕洞的存在情况, 而日冕物质抛射 (CME) 会引起太阳风的剧烈变化, 多个研究表明太阳风存在明显的 9 天、14 天左右和 27 天的周期性变化^[39,40], 本文研究结果表明, 在 14 天 (336 h) 的回溯窗口中, 模型实现了最佳性能, 这可能是综合日面上冕洞个数及 CME 爆发的结果, 而更长的历史回溯窗口不仅有效信息增量有限, 还可能引入额外的干扰噪声.

3.4 太阳风参数联合预测的有效性

为验证太阳风参数联合预测方式的有效性, 对比了太阳风参数的多变量预测多变量 (Multivariate predict Multivariate, M-M) 与单变量预测单变量 (Univariate predict Univariate, U-U) 两种策略. 前者以 4 个太阳风参数的历史数据作为输入, 预测输出其未来值; 后者每次只输入单个太阳风参数的历史数据, 预测输出其未来值. 实验结果列于表 2. 由表 2 可知, M-M 策略对于所有太阳风参数的预测均实现不同程度的性能提升, 特别是质子密度的预测 MAE 降低了 6.29%, 这表明不同太阳风参数间存在多变量相关性, 通过 TimeXer 可以有效捕获太阳风数据中的复杂依赖关系, 从而提升多个太阳风参数的预测精度.

3.5 模型的参数敏感性

通过实验分析片段长度 L_p 和嵌入维度 d_v 这两个关键超参数对模型预测性能的影响. 具体地, 在实验

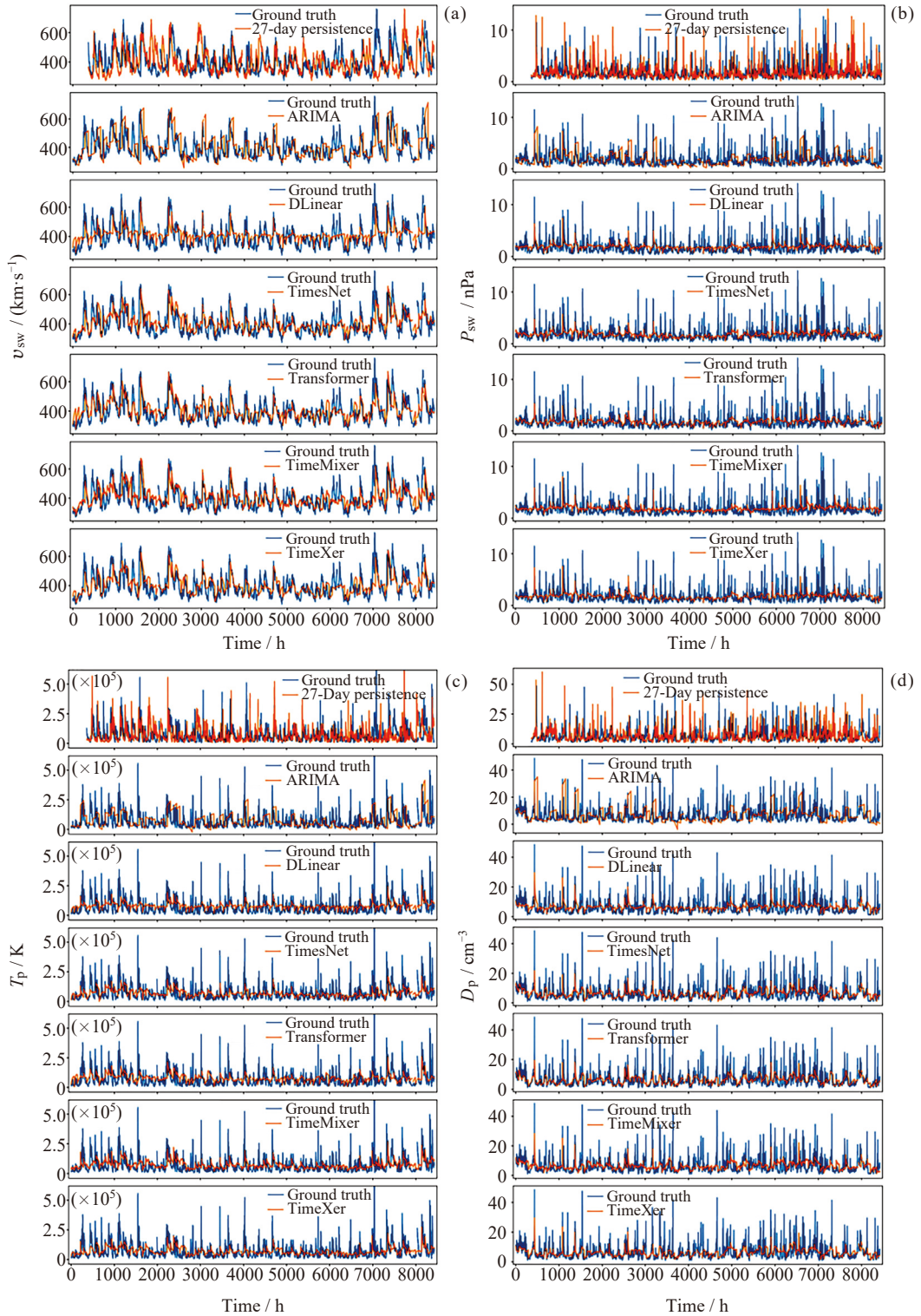


图 3 不同模型在 2021 年测试集上对太阳风速度、动压、质子温度、质子密度的 72 h 预测结果
 Fig. 3 72 h prediction results of solar wind speed, dynamic pressure, proton temperature, and proton density by different models on the test set in 2021

中设置片段长度 $L_p = \{2, 4, 8, 16, 32, 64\}$, 嵌入维度 $d_v = \{8, 16, 36, 72, 144\}$, 图 6 给出了 TimeXer 在不

同参数下的预测性能, 图中 RMSE 为 4 个太阳风参数标准化后计算的平均 RMSE. 由图 6 可知, 模型的

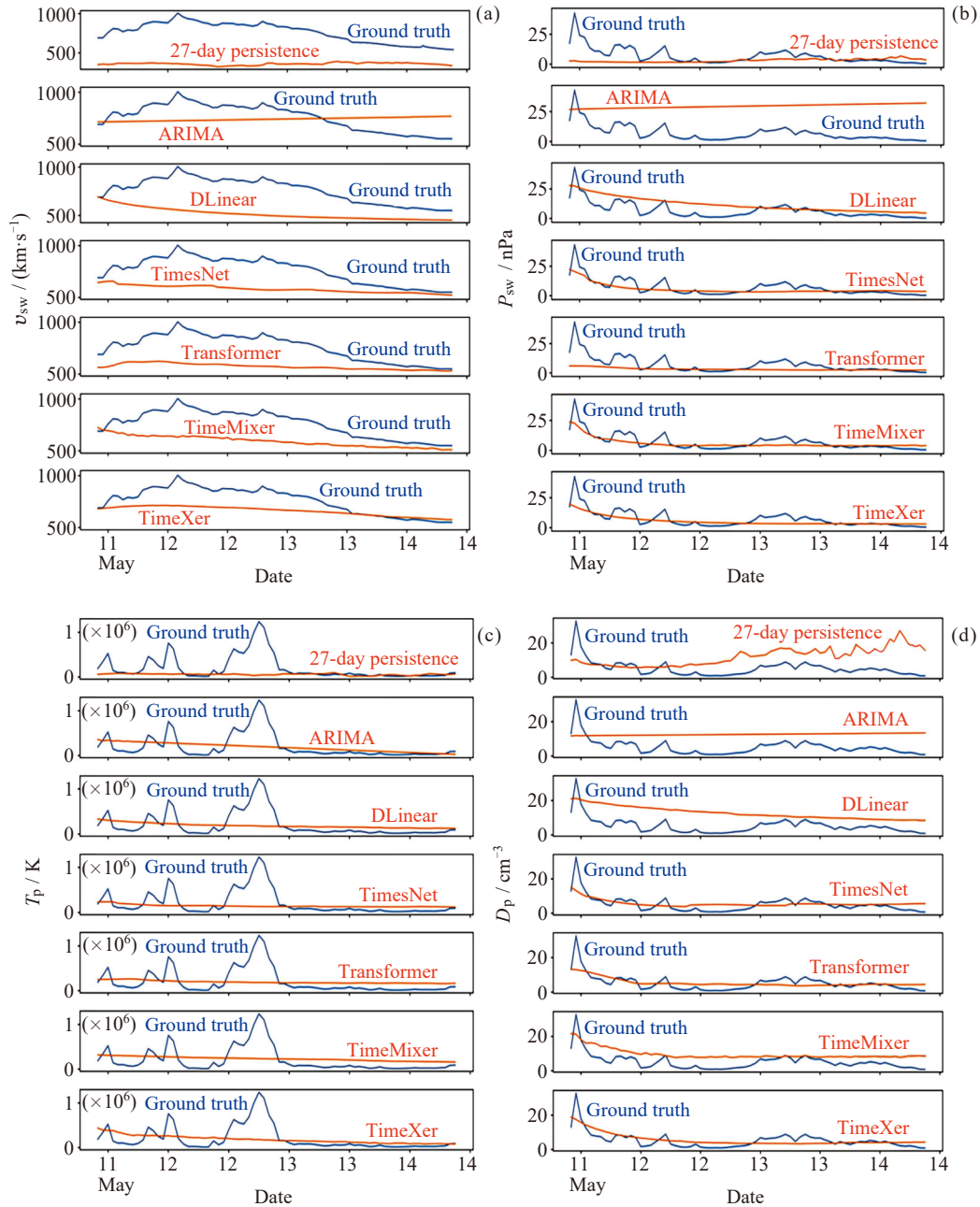


图 4 2024 年 5 月 11 日 10:00 UT 至 2024 年 5 月 14 日 09:00 UT 磁暴期间不同模型的 72 h 预测结果对比

Fig. 4 Comparison of the 72 h prediction results by different models during the geomagnetic storm from 10:00 UT on 11 May 2024 to 09:00 UT on 14 May 2024

预测 RMSE 均随着 L_p 和 d_v 的逐渐增大而先减小后增加, 并分别在 $L_p = 8$, $d_v = 36$ 时达到最小值. 这可能是因为当片段长度较小时, 模型获得的局部信息不足, 难以捕捉到足够的时间依赖和序列动态变化; 当片段长度过长时, 冗余信息和噪声可能干扰模型对关键特征的提取, 从而导致预测性能下降. 同样, 较低的嵌入维度限制了模型表达数据复杂变化的能力, 而过

高的嵌入维度则可能导致参数冗余和过拟合问题, 削弱了模型的泛化能力. 因此, 在实验中, $L_p = 8$, $d_v = 36$ 这组参数在模型的表达能力和噪声控制方面达到较好的平衡, 使得模型的预测性能达到最优.

3.6 模型的可解释性

为进一步理解模型的预测行为, 对 TimeXer 学习到的交叉注意力权重进行详细分析. 图 7 给出了交

又注意力层中的全局注意力权重矩阵, 其中第 i 行向量代表第 i 个太阳风参数对所有变量的关注分布, 权重值越大, 表示对应变量对该参数预测的重要性越高, 图中权重为 8 个注意力头的平均权重. 从图 7 中可以看出, 4 个太阳风参数在预测质子温度和太阳风速度时均起到关键作用. 相比其他参数, 太阳风速度和质子温度对质子密度的预测更为重要, 特别是太阳

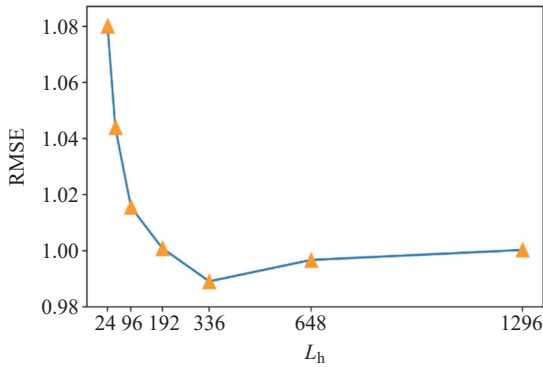


图 5 模型预测性能 (RMSE) 随历史回溯窗口长度的变化

Fig. 5 Variation of model prediction performance (RMSE) with historical look-back window length

风速度的注意力权重达到了较高值 0.21, 这是因为根据 Parker 太阳风模型^[41,42], 太阳风在传播过程中满足质量守恒方程 $4\pi r^2 D_p v_{sw} = \text{Const}$, 其中 r 表示日心距离, 这意味着质子密度与太阳风速度成反比关系, 并且随距离的平方衰减; 太阳风在近似绝热膨胀条件下, 其温度与密度之间满足 $T_p \propto D_p^{\gamma-1}$, 其中 γ 为绝热指数. 在预测太阳风动压时, 质子温度、太阳风速度与年时间信息显示出较强的影响力. 近期研究表明, 由于日心距离及日球纬度的周期性变化, 太阳风动压具有明显的年变化特征^[43,44]: 两极冕洞产生的高速低密度太阳风在每年春秋更容易到达地球附近, 导致动压在这些月份偏高; 夏季由于地球位于远日点附近, 太阳风以慢速成分为主, 动压偏低. 因此, 年时间信息对太阳风动压的预测具有较为显著的影响. 此外, 随着时间信息尺度的增加, 各太阳风参数对其关注程度也显著上升, 对年时间信息 Y 的注意力权重最大, 这可能是因为太阳风的中长期变化过程中, 其长期的季节性周期变化比短期的日变化更为显著^[40,45]. 这一结果表明, 对于太阳风参数的中长期预测, 利用不同尺度的时间信息有助于模型捕捉太阳风多尺度的变化特征.

表 2 TimeXer 对太阳风参数的多变量预测与单变量预测结果

Table 2 Multivariate and univariate forecasting results of solar wind parameters by TimeXer

Forecasting paradigms	$v_{sw}/(\text{km}\cdot\text{s}^{-1})$		P_{sw}/nPa		$T_p (\times 10^4)/\text{K}$		D_p/cm^{-3}	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
U-U	69.92	48.40	2.13	1.02	8.85	4.50	5.21	3.34
M-M	68.39	47.65	2.12	1.00	8.83	4.49	5.02	3.13
Improve	2.19%	1.55%	0.47%	1.96%	0.23%	0.22%	3.65%	6.29%

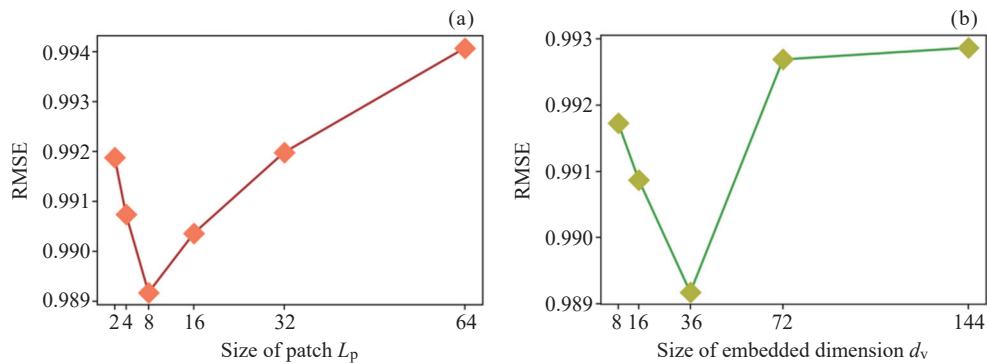


图 6 TimeXer 的参数敏感性

Fig. 6 Parameter sensitivity of TimeXer

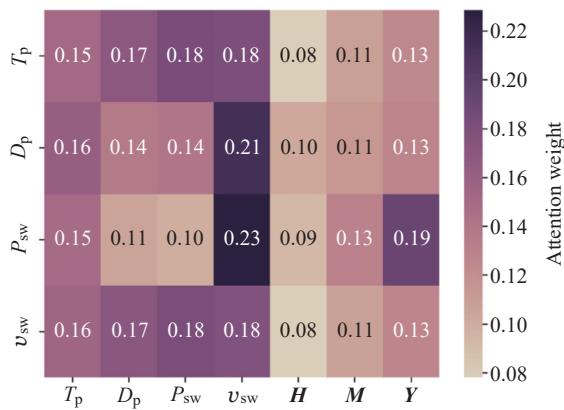


图 7 交叉注意力层中的全局注意力权重

Fig. 7 Global attention weights in the cross-attention layer

4 结论

基于深度学习方法 (TimeXer 模型), 实现了对太阳风速度、动压、质子温度与质子密度等多参数的联合预测. 模型利用过去 336 h 内 4 个太阳风参数及其对应的时间信息, 预测未来 72 h 的太阳风参数值. 为全面评估模型性能, 与经典的 27 天持续性模型、统计模型 ARIMA 以及近年来先进的深度学习模型 DLinear、TimesNet、Transformer 和 TimeMixer 等进行了对比. 此外, 还对不同历史数据输入长度对模型预测性能的影响、太阳风多参数联合预测的有效性以及各太阳风参数与时间信息在预测过程中的相对重要性进行了深入分析.

经过各项实验验证, 得出以下主要结论.

(1) TimeXer 模型在预测未来 72 h 的 4 个太阳风参数时, 其预测精度均优于 27 天持续性模型、ARIMA、DLinear、TimesNet、Transformer 和 TimeMixer 等传统及先进的深度学习方法, 并在不同太阳风活动状态下表现出较高的鲁棒性与准确性, 即使在磁暴期间, 也能较准确地捕获各太阳风参数的整体变化趋势, 但是对于太阳风参数中快速上升至高值区域的预测精度仍有待进一步提升.

(2) 当历史回溯窗口长度设为 $L_n=336$ 时, 模型性能达到最佳, 这可能与太阳风约 14 天的准周期性特征有关.

(3) 由于太阳风各参数之间存在明显的多变量相关性, 基于其依赖关系的联合预测明显优于单参数预测, 预测误差显著降低.

(4) 模型的参数敏感性分析表明, 在实验中, $L_p = 8$, $d_v = 36$ 这组参数可以在模型的表达能力与噪声控制方面达到较好的平衡, 从而使模型的预测性能达到最优.

(5) 通过交叉注意力权重的可视化分析发现, 4 个太阳风参数在对质子温度和太阳风速度的预测中均起到关键作用, 太阳风速度和质子温度对质子密度的预测贡献较大, 质子温度、太阳风速度与年时间信息对太阳风动压的预测具有较强的影响力; 此外, 随着时间信息尺度的增加, 各太阳风参数对其关注程度也显著上升.

未来可进一步在多源观测数据融合、物理先验知识约束等方面进行模型改进, 提升模型在中长期太阳风预测任务中的稳定性及精度, 推动构建高实时性和准确性的空间天气预报系统.

利益冲突 所有作者均声明不存在利益冲突.

参考文献

- [1] CRANMER S R, GIBSON S E, RILEY P. Origins of the ambient solar wind: implications for space weather[J]. *Space Science Reviews*, 2017, **212**(3): 1345-1384
- [2] VIDOTTO A A. The evolution of the solar wind[J]. *Living Reviews in Solar Physics*, 2021, **18**(1): 3
- [3] MEREDITH N P, HORNE R B, LAM M M, et al. Energetic electron precipitation during high-speed solar wind stream driven storms[J]. *Journal of Geophysical Research: Space Physics*, 2011, **116**(A5): A05223
- [4] WANG X, MIAO J, AA E, et al. Statistical analysis of Joule heating and thermosphere response during geomagnetic storms of different magnitudes[J]. *Journal of Geophysical Research: Space Physics*, 2020, **125**(8): 2020JA-027966
- [5] WANG X, MIAO J, LU X, et al. Latitudinal impacts of joule heating on the high-latitude thermospheric density enhancement during geomagnetic storms[J]. *Journal of Geophysical Research: Space Physics*, 2021, **126**(7): 2020JA-028747
- [6] WANG X, REN T, WANG R L, et al. Estimates of spherical satellite drag coefficients in the upper thermosphere during different geomagnetic conditions[J]. *Space Weather*, 2024, **22**(11): 2024SW003974
- [7] BOROVSKY J E, DENTON M H. Differences between CME-driven storms and CIR-driven storms[J]. *Journal of Geophysical Research: Space Physics*, 2006, **111**(A7): A07S08
- [8] CHEN G M, XU J Y, WANG W B, et al. A comparison of

- the effects of CIR- and CME-induced geomagnetic activity on thermospheric densities and spacecraft orbits: case studies[J]. *Journal of Geophysical Research: Space Physics*, 2012, **117**(A8): 2012JA017782
- [9] EASTWOOD J P, BIFFIS E, HAPGOOD M A, *et al.* The economic impact of space weather: where do we stand[J]. *Risk Analysis*, 2017, **37**(2): 206-218
- [10] SHEN F, SHEN C L, WANG Y M, *et al.* Could the collision of CMEs in the heliosphere be super-elastic/Validation through three-dimensional simulations[J]. *Geophysical Research Letters*, 2013, **40**(8): 1457-1461
- [11] HU Z J, HAN B, LIAN H F. Modeling of ultraviolet auroal intensity based on generalized regression neural network associated with IMF/solar wind and geomagnetic parameters[J]. *Chinese Journal of Geophysics*, 2020, **63**(5): 1738-1750 (胡泽俊, 韩冰, 连慧芳. 基于广义回归神经网络的行星际/太阳风参数和地磁指数的紫外极光强度建模 [J]. *地球物理学报*, 2020, **63**(5): 1738-1750)
- [12] ODSTRCIL D. Modeling 3D solar wind structure[J]. *Advances in Space Research*, 2003, **32**(4): 497-506
- [13] ODSTRCIL D, PIZZO V J, LINKER J A, *et al.* Initial coupling of coronal and heliospheric numerical magnetohydrodynamic codes[J]. *Journal of Atmospheric and Solar-Terrestrial Physics*, 2004, **66**(15/16): 1311-1320
- [14] DETMAN T, SMITH Z, DRYER M, *et al.* A hybrid heliospheric modeling system: background solar wind[J]. *Journal of Geophysical Research: Space Physics*, 2006, **111**(A7): 2005JA07102
- [15] WANG Y M, SHEELEY JR N R. Solar wind speed and coronal flux-tube expansion[J]. *Astrophysical Journal*, 1990, **355**(2): 726-732
- [16] RILEY P, LINKER J A, MIKIĆ Z. An empirically-driven global MHD model of the solar corona and inner heliosphere[J]. *Journal of Geophysical Research: Space Physics*, 2001, **106**(A8): 15889-15901
- [17] C N, ODSTRCIL D, PIZZO V J, *et al.* Improved method for specifying solar wind speed near the sun[J]. *AIP Conference Proceedings*, 2003, **679**(1): 190-193
- [18] YANG Y, SHEN F, YANG Z C, *et al.* Prediction of solar wind speed at 1 AU using an artificial neural network[J]. *Space Weather*, 2018, **16**(9): 1227-1244
- [19] UPENDRAN V, CHEUNG M C M, HANASOGE S, *et al.* Solar wind prediction using deep learning[J]. *Space Weather*, 2020, **18**(9): e2020SW002478
- [20] SUN Y R, XIE Z X, CHEN Y H, *et al.* Solar wind speed prediction with Two-Dimensional attention mechanism[J]. *Space Weather*, 2021, **19**(7): e2020SW002707
- [21] BROWN E J E, SVOBODA F, MEREDITH N P, *et al.* Attention-based machine vision models and techniques for solar wind speed forecasting using solar EUV images[J]. *Space Weather*, 2022, **20**(3): e2021SW002976
- [22] CHEN W X, FU S, GUO D Y, *et al.* Reconstruction and prediction on solar wind speed at L1 by artificial neural network and random forest methods[J]. *Chinese Journal of Geophysics*, 2023, **66**(12): 4850-4862 (陈伟鑫, 付松, 郭德宇, 等. 基于人工神经网络和随机森林方法的 L1 点处太阳风速度重构与预测 [J]. *地球物理学报*, 2023, **66**(12): 4850-4862)
- [23] SUN Y R, XIE Z X, WANG H C, *et al.* Solar wind speed prediction via graph attention network[J]. *Space Weather*, 2022, **20**(7): e2022SW003128
- [24] LIN R, LUO Z K, HE J S, *et al.* Prediction of solar wind speed through machine learning from extrapolated solar coronal magnetic field[J]. *Space Weather*, 2024, **22**(6): e2023SW003561
- [25] SHEN X F, NI B B, GU X D, *et al.* A statistical analysis of solar wind parameters and geomagnetic indices for the Solar Cycle 23[J]. *Chinese Journal of Geophysics*, 2015, **58**(2): 362-370 (沈晓飞, 倪彬彬, 顾旭东, 等. 第 23 太阳活动周期太阳风参数及地磁指数的统计分析 [J]. *地球物理学报*, 2015, **58**(2): 362-370)
- [26] LI Y X, XIE Y Q. Relationship between dynamic pressure pulse and geomagnetic sudden commencement[J]. *Chinese Journal of Space Science*, 2017, **37**(6): 659-666 (李玉鑫, 解妍琼. 太阳风动压脉冲结构与地磁急始关系研究 [J]. *空间科学学报*, 2017, **37**(6): 659-666)
- [27] LEI H J, LEI D X, YUAN J, *et al.* Ionospheric TEC prediction based on Attention-LSTM[J]. *Chinese Journal of Geophysics*, 2024, **67**(2): 439-451 (刘海军, 雷东兴, 袁静, 等. 基于注意力机制 LSTM 的电离层 TEC 预测 [J]. *地球物理学报*, 2024, **67**(2): 439-451)
- [28] ELLIOTT H A, HENNEY C J, MCCOMAS D J, *et al.* Temporal and radial variation of the solar wind temperature-speed relationship[J]. *Journal of Geophysical Research: Space Physics*, 2012, **117**(A9): 2011JA017125
- [29] SAMSONOV A A, BOGDANOVA Y V, BRANDUARDI-RAYMONT G, *et al.* Long-term variations in solar wind parameters, magnetopause location, and geomagnetic activity over the last five solar cycles[J]. *Journal of Geophysical Research: Space Physics*, 2019, **124**(6): 4049-4063
- [30] WANG Y X, WU H X, DONG J X, *et al.* TimeXer: empowering transformers for time series forecasting with exogenous variables[OL]. arXiv preprint arXiv: 2402.19072, 2024
- [31] KIM T, KIM J, TAE Y, *et al.* Reversible instance normalization for accurate time-series forecasting against distribution shift[C]//Proceedings of the 10th International Conference on Learning Representations. Virtual: ICLR, 2022
- [32] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017
- [33] OWENS M J, CHALLEN R, METHVEN J, *et al.* A 27

- day persistence model of near-Earth solar wind conditions: a long lead-time forecast and a benchmark for dynamical models[J]. *Space Weather*, 2013, **11**(5): 225-236
- [34] BOX G E P, PIERCE D A. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models[J]. *Journal of the American Statistical Association*, 1970, **65**(332): 1509-1526
- [35] ZENG A L, CHEN M X, ZHANG L, *et al.* Are transformers effective for time series forecasting[C]//Proceedings of the 37th AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence. Washington: AAAI, 2023: 11121-11128
- [36] WU H X, HU T G, LIU Y, *et al.* TimesNet: Temporal 2D-Variation Modeling for General Time Series Analysis[EB/OL]. (2023-04-12)[2025-01-10]. <https://arxiv.org/abs/2210.02186>
- [37] WANG S Y, WU H X, SHI X M, *et al.* TimeMixer: Decomposable Multiscale Mixing for Time Series Forecasting[EB/OL]. (2024-05-23)[2025-01-10]. <https://arxiv.org/abs/2405.14616>
- [38] LIU Y, HU T G, ZHANG H R, *et al.* iTransformer: Inverted Transformers are Effective for Time Series Forecasting[EB/OL]. (2024-03-14)[2025-01-10]. <https://arxiv.org/abs/2310.06625>
- [39] MURSULA K, ZIEGER B. The 13.5-day periodicity in the Sun, solar wind, and geomagnetic activity: the last three solar cycles[J]. *Journal of Geophysical Research: Space Physics*, 1996, **101**(A12): 27077-27090
- [40] KATSAVRIAS C, PREKA-PAPADEMA P, MOUSSAS X. Wavelet analysis on solar wind parameters and geomagnetic indices[J]. *Solar Physics*, 2012, **280**(2): 623-640
- [41] PARKER E N. Dynamics of the interplanetary gas and magnetic fields[J]. *Astrophysical Journal*, 1958, **128**: 664-676
- [42] ECHIM M M, LEMAIRE J, LIE-SVENDSEN O. A review on solar wind modeling: kinetic and fluid aspects[J]. *Surveys in Geophysics*, 2011, **32**(1): 1-70
- [43] MURSULA K, HOLAPPA L, LUKIANOVA R. Seasonal solar wind speeds for the last 100 years: unique coronal hole structures during the peak and demise of the grand modern maximum[J]. *Geophysical Research Letters*, 2017, **44**(1): 30-36
- [44] OWENS M J, LOCKWOOD M, BARNARD L A, *et al.* Annual variations in the near-Earth solar wind[J]. *Solar Physics*, 2023, **298**(9): 111
- [45] CHOWDHURY P, CHOUDHARY D P, GOSAIN S, *et al.* Short-term periodicities in interplanetary, geomagnetic and solar phenomena during solar cycle 24[J]. *Astrophysics and Space Science*, 2015, **356**(1): 7-18

作者简介



高志旭 男, 1999年9月出生于云南省弥勒市, 现为中国科学院国家空间科学中心硕士研究生, 主要研究方向为机器学习在空间环境预报中的应用。

E-mail: gaozhixu22@mailsucas.ac.cn



陈艳红 (通信作者) 女, 1977年11月出生于湖北省孝感市, 现为中国科学院国家空间科学中心研究员, 博士生导师, 主要研究方向为电离层扰动特征分析、预报建模和人工智能空间环境预报研究。

E-mail: chenyh@nssc.ac.cn