

基于神经网络的未来 3 天 Kp 指数预报 建模与可解释 AI 应用*

王听雨¹ 罗冰显^{1,2} 陈艳红¹
石育榕¹ 王晶晶¹ 刘四清^{1,2}

1(中国科学院国家空间科学中心 空间天气学国家重点实验室 北京 100190)

2(中国科学院大学 北京 100049)

摘要 当前业务中对未来 3 天 Kp 指数预报需求强烈. 但地磁暴中多参数耦合导致难以量化各预报因子对 Kp 值的贡献, 制约了预报精度提升. 本文构建了神经网络 3 天 Kp 指数预报模型, 并使用人工智能 (AI) 可解释性算法量化各因子贡献. 结果显示, 行星际磁场南向分量在提前 3 h 对 Kp 指数的贡献为 37.15%, 为主要因子, 说明模型能捕捉符合物理特征的主要预报因子. Kp 指数历史特征贡献随提前量逐渐增加, 提前 3 天总体贡献占 68.06%, 验证了对冕洞高速流引起的地磁暴事件的预报能力. 对 2015 和 2017 年特大地磁暴进行贡献分析, 模型准确捕捉了地磁暴多参数耦合的复杂特性. 研究表明, 可解释 AI 算法在一定程度上能量化各预报因子对 Kp 指数的预报贡献, 有助于改进未来 3 天 Kp 指数 AI 预报模型.

关键词 地磁暴, 未来 3 天 Kp 指数预报, 神经网络, 可解释性 AI 算法

中图分类号 P353

Modeling Next 3-day Kp Index Forecasting with Neural Networks and Exploring the Application of Explainable AI

WANG Tingyu¹ LUO Bingxian^{1,2} CHEN Yanhong¹
SHI Yurong¹ WANG Jingjing¹ LIU Siqing^{1,2}

1(State Key Laboratory of Space Weather, National Space Science Center, Chinese Academy of Sciences, Beijing 100190)

2(University of Chinese Academy of Sciences, Beijing 100049)

* 中国科学院战略性先导科技专项 (XDB0560000), 国家自然科学基金面上项目 (42074224), 中国科学院重点部署项目 (ZDRE-KT-2021-3), 中国科学院国家空间科学中心“攀登计划”青年创新课题 (E4PD40012S) 和中国科学院青年创新促进会共同资助

2023-09-27 收到原稿, 2023-11-17 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/>)

Abstract The current operational needs of space weather forecasting strongly require accurate predictions of the future 3-day Kp index. Such forecasts involve a multitude of predictors, including physical parameters observed at the Earth-Sun L1 point and historical characteristics of the Kp index. Therefore, previous research primarily relied on statistical or empirical methods for prediction. However, the complex coupling of multiple parameters during geomagnetic storm events has made it challenging to quantify the contributions of various predictors to Kp index forecasting over a 3-day timescale, hindering further improvements in forecast accuracy. This study builds a 3-day Kp index forecasting model based on neural network modeling and utilizes explainable AI (Artificial Intelligence) algorithm, specifically the integrated gradient algorithm, to quantify the contributions of individual predictor. The research results indicate that the southward interplanetary magnetic field contributes significantly to Kp index prediction, accounting for 37.15% of all factors, making it the primary contributor. Following this, solar wind speed contributes 15.73%, underscoring the model's ability to capture parameters aligned with physical characteristics as the primary predictive factors during training. The contribution of historical characteristics of Kp index (recurrence characteristics) gradually increases with the forecasting horizon and reaches 68.06% at a lead time of 3-day. This substantiates the strong predictive capabilities of the AI model in forecasting geomagnetic storm events induced by high-speed solar wind streams originating from coronal holes. Furthermore, this study conducts contribution analysis on two significant geomagnetic storm events that occurred in 2015 and 2017. It reveals that the predominant predictors contributing to each event differ. This underscores the model's capability to accurately capture the complex coupling of multiple parameters in geomagnetic storm forecasting. In conclusion, this research demonstrates that employing explainable AI algorithms can help quantify the contributions of various predictive factors to Kp index forecasting to some extent. This has the potential to enhance further research and improvements in 3-day Kp index AI forecasting models.

Key words Geomagnetic storm, 3-day Kp index forecasting, Neural network, Explainable AI algorithm

0 引言

地磁暴是指太阳爆发活动产生的高速等离子体云到达地球空间后与地球磁场相互作用引起的地球磁层扰动。这种扰动会影响空间环境,并进一步干扰人类的导航、通信、电力等系统。2022年2月8日,美国SpaceX公司发射了49颗星链卫星,其中有40颗受地磁暴影响,在大气层坠毁^[1,2]。 Kp 指数(行星性3h磁情指数)用于描述全球地磁扰动的程度,被广泛应用于地磁暴的研究和预报中,作为地磁暴预警和地磁暴持续时间的判据^[3]。当前的 Kp 指数的预报模型以短时(提前1~3h)预报为主^[4-6],但在实际业务中对未来3天的 Kp 指数预报有巨大的需求。比如,空间站的轨道预报精度十分依赖于未来3天甚至更长时间的 Kp 指数预报结果。因此,提升未来3天 Kp 指数的预报准确度有助于提高中国空间天气保障能力^[7,8]。

目前,未来3天 Kp 指数预报相关的研究^[9-11]和数据^[12,13]仍然较少,且研究主要的是利用统计、经验或者人工智能(AI)进行建模^[14]。然而,由于各个预报因子与 Kp 指数之间具有非线性关系,而且受制于AI模型的“黑盒”性质,因此很难定量其对预测结果的数量贡献以及这种贡献随时间的变化情况。在不同的预测提前期内,各预测因素对预测值的影响差异显著,但由于缺乏量化的贡献研究,因子的筛选变得更加困难,阻碍了模型的改进以及预测准确性的进一步提高。为了在AI模型中明确各种预报因子对预报结果的定量贡献,可以考虑借鉴天气和气候领域的已有方法。

当前,已经有一部分天气、气候预报研究借助可解释AI(Explainable Artificial Intelligence, XAI)来评估预报因子对预报值的贡献程度以及可预报性的变化^[15-18]。可解释AI是近几年发展起来的一个研究方向,其关键在于寻求对模型工作机理的理解,通过

算法为 AI 的决策提供人类可以理解的解释^[19]. 可解释 AI 有多种实现形式, 本文侧重于研究预报因子 (日地 L1 点观测的太阳风参数、 Kp 指数的历史特征) 对预报结果的定量贡献, 即通过 AI 可解释性算法, 给出每一个输入模型的参数对预报的贡献占比, 确保 AI 的预报决策符合基本物理规律. 同时, 给出各种观测参数与 Kp 指数之间的统计定量关系. 总之, 发展可解释 AI 将有助于空间天气预报的决策、复盘以及准确度的提升.

本文利用神经网络模型对未来 3 天 Kp 指数进行预报建模, 并对预报结果进行评估. 借助可解释性算法, 研究 L1 点行星际观测数据以及历史 Kp 指数等多种特征对 Kp 指数预报的定量贡献, 并给出其贡献占比, 以达到探索可解释 AI 算法在空间天气领域应用的目的, 并为基于 AI 的未来 3 天 Kp 指数预报结果提供一定的理论支撑.

1 神经网络模型、数据与方法

1.1 前馈神经网络

使用前馈神经网络 (Feed forward neural network) 对行星际-地磁扰动的过程进行建模. 该模型具有 1 个隐藏层, 隐藏层的神经元节点设置为 20 个. 选择此模型一方面是因为 Shprits 等^[10]的工作验证了该模型在 3 天 Kp 指数预测中的可行性, 并将其应用在德国地球科学研究中心 (German Research Centre for Geosciences, GFZ) 3 天预报数据中. 另一方面, 由于该类模型的可解释性较强^[20], 可以利用 AI 可解释性算法对输入神经网络中的每一个特征进行贡献度计算, 进而可以挖掘每一个特征在实际的预报过程中产生的定量贡献. 本研究用较为简单的模型验证 AI 可解释性算法的有效性, 将有助于未来将该算法广泛应用在空间天气 AI 建模预报的研究中.

所使用的超参数与训练策略见表 1. 将预报的时间分辨率设定为 3 h, 对每个提前时间 (提前 3, 6, 9, ..., 72 h) 分别训练模型, 便于使用 AI 可解释性算法从多个提前期模型中提取各预报因子的贡献值.

1.2 数据获取

研究使用的数据包括 L1 点的太阳风监测数据和地磁 Kp 指数数据. 太阳风监测数据来自 OMNI-Web 的 1 min 高分辨率监测数据^[21], Kp 指数来自于 GFZ 发布的结果^[3]. 用于训练和验证的时间段为

表 1 训练模型所用到的超参数以及训练策略
Table 1 Hyperparameters and training strategies used for training the model in this study

参数名	参数值/名称
Mini batch	1
Epoch	10
Optimizer	AdamW
Loss function	MSE
Learning rate	0.004
Momentum	0.6
Scheduler	StepLR

2011 年 4 月至 2022 年 6 月, 共计 32872 个样本. 本文选择十折交叉验证的训练方式, 针对每一个验证时段分别训练建模, 避免评估结果受到训练集影响.

在 3 天 Kp 指数的预报中, 将起报时间定为 T_0 时刻, 对日地 L1 点-地球上的几个主要观测数据进行预处理, 输入模型的特征如下.

(1) Kp 指数的历史特征: T_0 时刻之前的 3, 6, 9 h 的 Kp 指数 (共 3 个特征).

(2) Kp 指数的历史特征 (重现性): 考虑到太阳的自转周期为准 27 天, 选择 T_0 时刻之前的第 26, 27, 28, 53, 54, 55 这 6 天中的 6, 12, 24 h 内对应时间的平均值、最大值和最小值 (共 54 个特征). 例如起报时间为 2021 年 3 月 1 日 09:00 UT, 则预报因子为该时间之前 26, 27, 28, 53, 54, 55 这 6 天中, 06:00 UT—12:00 UT, 03:00 UT—15:00 UT、前一天 21:00 UT—21:00 UT 这 3 个时间窗口中的平均值、最大值和最小值.

(3) 太阳风参数的历史特征: 从多个太阳风参数中选择太阳风速度 (v) 行星际磁场南向分量 (B_z)、总磁场 (B)、质子密度 (n) 共 4 个观测数据, T_0 时刻之前的 0~3, 3~6 和 6~9 h 内观测到的平均值、最大值和最小值 (共 36 个特征).

预报值选择提前 3 h 至 3 天的 24 个 Kp 指数. 考虑到这些特征之间的量纲不同, 所以统一对所有数据进行 0~1 标准化之后, 再统一输入前馈神经网络之中.

1.3 模型评估方法

在实际的应用中, 仅使用相关系数作为评估指标难以体现 Kp 指数预报的实际效果. 对 Kp 指数的预报评估需考虑多个方面, 例如地磁暴的发生时间、未来 3 天是否将会发生地磁暴, 以及 3 天中的地磁暴时

间序列以及强度变化. 因此, 借助如下多种评估指标对模型进行评估:

$$E_{\text{MAE}}(X, Y) = \frac{1}{N} \sum_{i=1}^N |X_i - Y_i|, \quad (1)$$

$$E_{\text{RMSE}}(X, Y) = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - Y_i)^2}, \quad (2)$$

$$E_r(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X] \text{Var}[Y]}}. \quad (3)$$

其中, X 代表实测数据, Y 代表模型的预测数据. 借助多种评估指标, 可以更全面地展现 3 天 Kp 指数模型的预报能力.

1.4 AI 可解释性算法—积分梯度算法

使用积分梯度算法对训练完毕的 AI 模型进行可解释性分析^[18], 该算法的一个优势在于不受模型结构的限制, 可以应用在各种模型中. 该算法介绍如下: 给定一个预报特征组合的输入, 令其为 $\mathbf{X}$. 以函数 $F(\mathbf{X})$ 表示前馈神经网络模型, 由于该模型具有非线性激活函数 ReLU, 因此 $F(\mathbf{X})$ 代表一个非线性函数. 使用一阶泰勒展开将 $F(\mathbf{X})$ 近似成一个线性函数, 可以表示为

$$F(\mathbf{X}) \approx \omega^T \mathbf{X} + b. \quad (4)$$

对于预报参数组合 $\mathbf{X}$ 中的任意一个参数 x_i ($i=1, 2, \dots, 93$), 其变化都可能会使 $F(x_i)$ 发生变化, 而 $F(x_i)$ 随 x_i 的变化可以用梯度 ω 表示, 即

$$\omega_i = \frac{\partial F(x_i)}{\partial x_i}. \quad (5)$$

令输入积分梯度公式中的基准值为 $\mathbf{X}$ (一般是各参数的最小值或者平均值, 本文使用最小值作为基准值), 输入模型中的起报时间的真实观测 (真值) 为 $\mathbf{X}'$. 则每一个参数的积分梯度可以通过计算从基准值到真值之间的一条直线的路径积分得到. 将这一直线路径函数 $\gamma(\alpha)$ 定义为 $\gamma(\alpha) = \mathbf{X} + \alpha \times (\mathbf{X}' - \mathbf{X})$. 因此, 对参数 x_i 的积分梯度贡献值 G_i 就可以用如下改进后的积分梯度公式表示:

$$G_i = \left| (x'_i - x_i) \int \frac{\partial F[x_i + \alpha(x'_i - x_i)]}{\partial x_i} d\alpha \right|. \quad (6)$$

通过积分梯度算法计算每个物理参数对预报结果的贡献是具有意义的. AI 建模构建了预报因子和预报值之间的非线性关系. 因此, 当模型训练完毕后, 输入任意一组预报因子, 都会根据模型的非线性函数预测未来时刻的 Kp 指数. 假设某个预报因子对预报

结果具有贡献, 那么此预报因子从基准值变化到真值的过程中, $F(x_i)$ 的数值也会发生变化, 这些变化表现在函数的梯度值 (ω_i) 中. 对梯度进行路径积分可以理解为, 量化某个预报因子在变化的整个过程中, 对预报值造成变化的大小. 通过对比不同预报因子的贡献值, 以及某个因子在所有预报因子中贡献值所占比例, 即可量化该预报因子对于预报的贡献程度.

2 预报结果评估

评估 2011—2022 年 Kp 指数预测与实测数据之间 RMSE, MAE 与相关系数 r (correlation coefficient), 绘制为评估随预报提前量变化图 (见图 1).

由于本研究侧重于对输入模型的各参量贡献度进行量化研究, 因此使用的模型结构较为简单, 并且在所有的提前时间中都使用了同样的输入特征, 但通过图 1 的评估结果可知, 当前训练的模型仍然具有一定的预报能力, 对于提前 3 h 的预报准确度: RMSE=0.9, MAE=0.70, $r=0.70$; 并且随着预报时长的增加, 模型的预报准确度表现为逐渐下降.

同时, 选择 2015 年 6 月和 2017 年 9 月的两次地磁暴个例进行预测评估, 如图 2 所示. 通过图 2 可知, 模型在提前 3 h 预报两次地磁暴事件时, 虽然强度无法达到和实际结果一致, 但从预报时间序列的趋势分析, 模型都能成功预测出两次特大地磁暴的演变趋势, 说明该模型已经具备从输入参数中获取地磁暴事件关键前兆信号的能力.

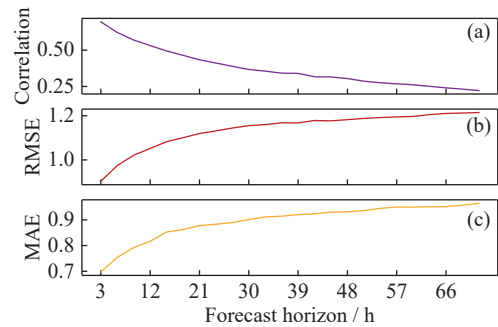


图 1 未来 3 天 Kp 指数预报模型评估结果 (2011 年 4 月至 2022 年 6 月). 曲线为评估指标随预报时长 (3 h 分辨率) 变化的函数

Fig. 1 Evaluation results (April 2011 to June 2022) of the future 3 day Kp index forecast. Model depict as a function of the forecast horizon (at a 3 h resolution)

3 太阳风及历史 Kp 指数对未来 3 天 Kp 预报的定量化结果

3.1 预报因子的总体贡献

量化不同输入参数在 2011 年 4 月至 2022 年 6 月的总体贡献. 选择提前 3 h、提前 1 天 (24 h)、提前 2 天 (48 h) 和提前 3 天 (72 h) 共 4 个时间节点, 通过积分梯度算法分别计算 93 个预报因子在

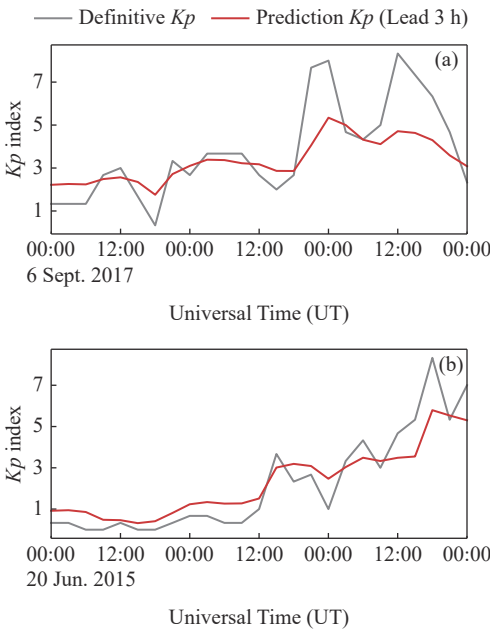


图 2 2017 年与 2015 年两次特大地磁暴时实际结果与预报结果对比

Fig. 2 Comparison between actual and forecasted Kp index for two extreme geomagnetic storms in 2017 and 2015

32872 个样本中的贡献值, 以此得到 32872×93 的贡献值矩阵, 同样使用十折交叉中的验证集部分用于计算和分析. 计算所有样本贡献值之和, 可以得到 93 个贡献值, 代表 93 个特征各自的总贡献. 将 93 个特征按照表 2 分成 7 个大类.

对 7 大类的预报特征计算后, 得到如图 3 所示的各特征预报贡献百分比柱状图.

从图 3 的贡献占比变化中可以看出, 在提前 3 h 预报时, 贡献最多的一个变量是 B_z , 占比 37.15%, 说明其是对地磁暴预报相当关键的要素, 这与以往的研究结论一致^[22]. 随预报提前量的延长, 到 3 天时 B_z 的贡献仅占 13.62%, 说明模型在没有引入任何先验知识的情况下, 从数据中学习到了预报中的关键物理规律. 同时也要注意, B_z 贡献占比随时间逐渐减小; 其他例如 B , v , n 这 3 个特征则在一开始具有一定的贡献, 但每个变量的贡献占比都小于 B_z , 且也随预报提前量变长存在下降趋势. 仅有 Kp 指数的两个重现性特征是随预报提前量延长而表现出贡献增加, 在提前 3 h 贡献总共占比 18.17%, 提前 3 天贡献占比提高到 68.06%, 成为最主要的预报要素. 这可能是因为这两组特征能对重现性地磁暴的预报具有较好的指示作用. 其中前一个太阳自转周期的历史 Kp 指数贡献最大, 其次是前两个周期的历史 Kp 指数, 这是因为冕洞在一个自转周期内的特征变化要显著少于两个周期内的变化.

虽然 Kp 的重现性特征随预报提前量增加, 贡献占比逐渐增加, 但并不意味着重现性对预报结果发挥的作用增加. 实际上, 所有特征随着预报提前量延长,

表 2 用于输入模型的 93 个特征分类 (按照物理参数或者物理特征进行归类)

Table 2 Classification of 93 features inputting into model (categorized by physical parameters or characteristics)

特征类型	特征因子	特征数
B_z	行星际磁场南向分量(B_z)在 T_0 时刻之前的 0~3 h, 3~6 h 和 6~9 h 内观测到的平均值、最大值和最小值	9
B	总磁场(B)在 T_0 时刻之前的 0~3 h, 3~6 h 和 6~9 h 内观测到的平均值、最大值和最小值	9
v	太阳风速度(v) 在 T_0 时刻之前的 0~3 h, 3~6 h 和 6~9 h 内观测到的平均值、最大值和最小值	9
n	质子密度(n) 在 T_0 时刻之前的 0~3 h, 3~6 h 和 6~9 h 内观测到的平均值、最大值和最小值	9
H [Kp (Hist)]	T_0 时刻之前的 3, 6 和 9 h 的 Kp 指数	3
R_1 [Kp (Rec 1)]	T_0 时刻之前的 26, 27, 28 天(前一个太阳自转周期)的 Kp 指数在 6, 12, 24 h 内的平均值、最大值和最小值	27
R_2 [Kp (Rec 2)]	T_0 时刻之前的 53, 54, 55 天(前两个太阳自转周期)的 Kp 指数在 6, 12, 24 h 内的平均值、最大值和最小值	27

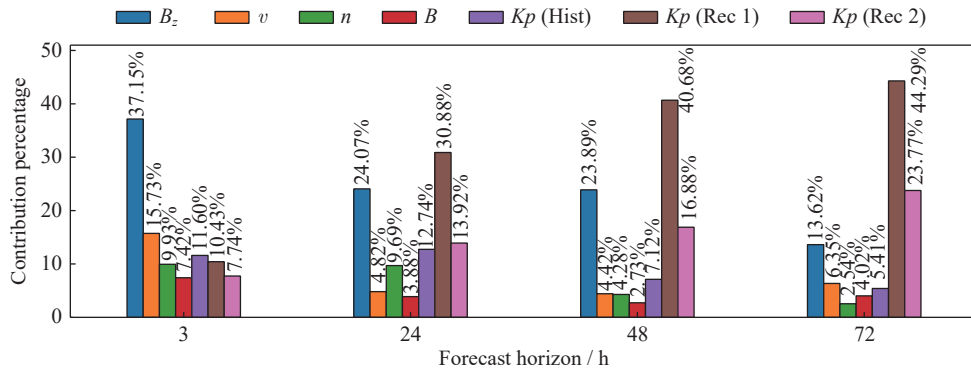
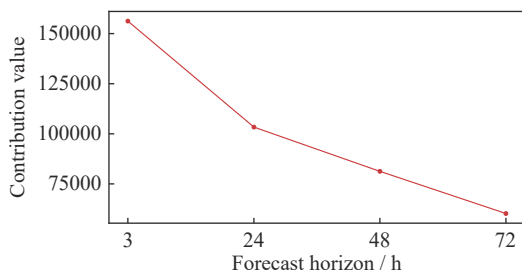
图3 提前 3, 24, 48, 72 h 预报 K_p 指数时 7 类特征贡献值占比Fig. 3 Contribution percentages of the 7 categories of features forecasting the K_p index 3, 24, 48, 72 h in advance

图4 93 个特征贡献度之和随预报时间的变化 (提前 3, 24, 48, 72 h)

Fig. 4 Sum of contributions from 93 features as a function of the forecast horizon (3, 24, 48, 72 h ahead)

总体贡献值是逐渐减少的. 从图 4 的贡献值随时间变化的曲线就能看出, 93 个特征因子在提前 3 h 能提供的贡献值是提前 3 天的 2.59 倍. 因此, 也可以从预报因子的贡献度来解释预报准确度的下降, 即 93 个预报因子能在提前 3 h 为 AI 模型提供更多的贡献, 因此预报结果也更为准确, 而提前 3 天时, 各预报因子为模型提供的贡献显著降低, 导致了预报准确度的下降. 这也说明, 计算贡献值不仅可以用于解释模型、定量不同参量的贡献占比, 也可能用于解释预报信息随提前期的变化. 因为根据基本的物理规律, 提前预报的时间越长, 初始条件的参量中可用于预报的信息量就越少, 对预报结果的贡献也会变小.

3.2 预报因子贡献度的个案分析

利用积分梯度算法提取各特征因子在预报时的统计贡献值, 此算法可以对单一样本进行积分运算, 所以也可以用于对单次地磁暴事件中各物理参量的贡献进行度量, 该算法的特性将有利于对地磁暴个例的成因进行复盘和研究. 挑选 2015 年 6 月 22 日

18:00 UT^[23] 和 2017 年 9 月 8 日 12:00 UT^[24,25] 这两次具有代表性的事件 (两次事件 K_p 指数都达到 8⁺ 级别), 对该算法的个例分析能力进行初步的探究.

选取两次事件发生的前 3 h 作为起报时间, 计算输入模型中的各个特征对两次预报的贡献. 虽然在统计上, 提前 3 h 的预报中 B_z 是一个最关键的要素, 但这两次事件中, 关键要素却各不相同. 其中对 2017 年的地磁暴的预报的贡献较大的前 3 个特征分别是 $v = 23.89\%$, $B_z = 23.00\%$, $H = 18.99\%$, 如图 5(a) 所示, 说明在这次事件中, 太阳风速度的贡献达到了和南向磁场相同的水平. 而 2015 年的事件中前三个特征是 $B_z = 29.60\%$, $H = 20.20\%$, $n = 17.59\%$, 如图 5(b) 所示, 说明在该次地磁暴事件中, 对预报其主要贡献的仍是行星际磁场南向分量, 但质子密度对预报同样具有较大作用. 通过对比两次事件的贡献占比可知, 对单次地磁暴事件具有主要贡献的特征参数并不相同, 这也反映出地磁暴产生机制中多参数耦合的复杂特性.

3.3 模型拟合能力对定量化贡献度的影响

积分梯度作为一种可解释 AI 算法, 对预报的贡献进行量化的准确度依赖于前序建模模型的拟合能力强弱. 所以有必要对不同拟合能力的模型在可解释性分析中的结果进行对比, 确认不同拟合能力的模型如何影响贡献度分析的准确性. 考虑到积分梯度算法的算力消耗较大, 将模型的拟合能力使用 3 种情况表示, 即欠拟合、拟合能力正常及过拟合. 其中欠拟合与过拟合属于模型精度下降的两种情况.

通过增加模型的学习率到 0.8, 使模型达到欠拟合的效果; 减小学习率到 0.0002, 使模型达到过拟合的效果. 将这两个模型分别与之前学习率 0.04 的模

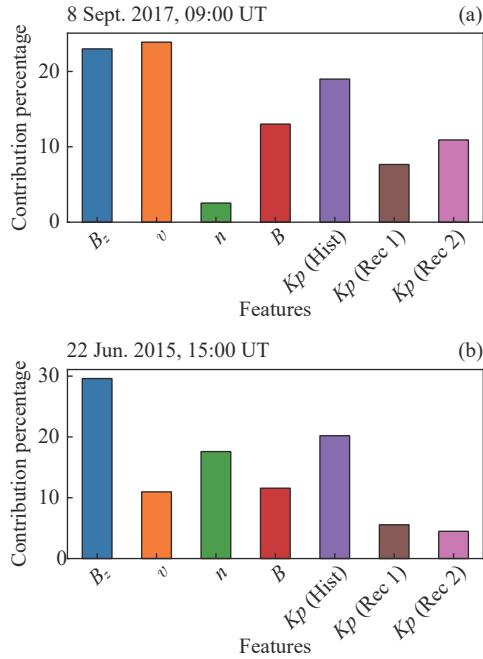


图 5 2017 年与 2015 年两次特大地磁暴时 7 类特征对预报的贡献占比 (提前 3 h 预报)

Fig. 5 Contribution percentages of the 7 feature categories to the forecast during the two extreme geomagnetic storms in 2017 and 2015 (3 h lead time)

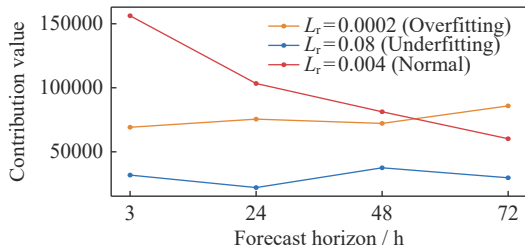


图 6 三种拟合能力不同的模型度量的总贡献度随预报时间的变化 (提前 3, 24, 48, 72 h)

Fig. 6 Total contribution measured by three models with different fitting abilities varies over forecast horizon (3, 24, 48, 72 h in advance)

型结果进行对比 (见图 6), 可以发现仅有正常模型能够展现出提前预报的时间越长, 初始条件的参量中可用于预报的信息量越少的特征. 过拟合模型出现随提前量增加各特征总贡献量增加的特征, 说明模型从输入参数中提取了许多不符合预报规律的特征, 以此试图提高训练集的拟合能力, 导致了模型泛化能力的下降. 而欠拟合模型从各时段提取的总贡献远小于其他两组模型, 说明一旦模型的拟合能力不足, 从参数中

提取对预报有贡献的信息的能力也会显著降低.

同时, 过拟合与欠拟合也会导致模型无法从大数据中获取真实有效的知识. 图 7(a) 显示出在多个提前量中, B_z 都具有主要贡献, 这是由于在训练集中的过度拟合, 使模型学习到了错误的知识, 夸大了 B_z 在各个提前量中的贡献. 图 7(b) 的结果则是由欠拟合所致, 模型无法学习到任何对预报有贡献的知识. 因此, 模型的模拟能力的好坏对于后续的量化贡献具有至关重要的作用, 需要在建模过程中保证模型的精度在合理的范围内, 并且模型不出现过拟合和欠拟合, 才能进一步用于可解释性 AI 分析.

4 结论与讨论

使用神经网络构建了未来 3 天的 Kp 指数预报模型, 并对其进行了预报能力的评估. 由于本文更加侧重对 AI 建模进行可解释性研究, 因此模型架构选择了可解释性更强, 且训练更为简单的前馈神经网络. 利用积分梯度算法, 从 2011—2022 年的所有样本中提取了每个预报特征对预报结果的贡献度数值, 验证了模型在建模时能准确抓住 L1 点观测数据中的关键要素 B_z , 因此具有较高的预报准确度. 同时本文给出了 B_z 在预报中的贡献占比的变化, 从提前 3 h 的 37.15% 下降到了提前 3 天的 13.62%, 这一结论为 1~3 天的 Kp 指数预报提供了参考. 同时, 模型中也需要历史 Kp 指数等参数的配合, 引入重现性特征, 才能在提前 1 天以上的 Kp 指数预报中取得一定的预报效果, 因为模型会随着预报提前量的延长, 而调整不同参数的贡献比例, 在提前 1 天乃至 3 天的预报中, 模型选择历史 Kp 指数中的重现性特征作为关键要素来指导定量预报结果. 本文也试图将 AI 可解释性算法应用在地磁暴个例的分析中, 发现即使是同等级的地磁暴, 对预报起关键作用的参数也并不一致, 直观证明了 AI 模型能够针对地磁暴事件进行具有物理意义的建模. 通过构建的一个过拟合模型与欠拟合模型与正常模型进行对比, 指出需要让模型的拟合能力处在合理的范围, 才能使可解释性 AI 算法发挥作用, 以更好地度量预报因子的量化贡献.

过去基于神经网络的 Kp 指数预报研究更重视深度学习模型的建模过程, 但是在天气学和气候学领域, 针对臭氧层污染^[16]、降水^[26]、厄尔尼诺预报^[15,17]已经进行了大量 AI 可解释性研究, 而空间天气领域

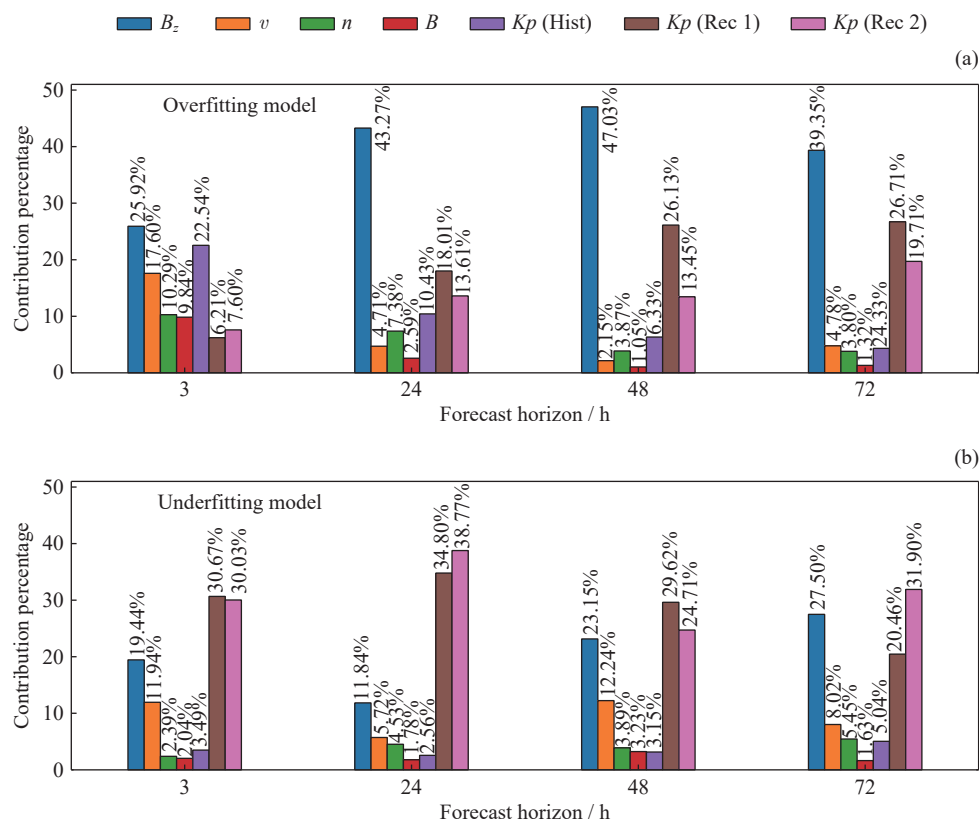


图7 提前3, 24, 48, 72 h 预报 K_p 指数时7类特征贡献值占比

Fig. 7 Contribution percentages of the 7 categories of features forecasting the K_p index 3, 24, 48, 72 h in advance

目前仍缺乏对未来3天的 K_p 指数预报建立可解释AI的分析框架, 因此本文针对该问题进行了可解释性研究. 未来的研究方向应是在此基础上发展出更为完善且合理的分析框架, 完善的可解释AI分析框架可能在以下几个方面推进 K_p 指数的预报研究.

(1) 通过对用于输入模型各个特征参量进行贡献度的评估, 针对不同预报时间范围, 筛选最为合适的参量组合, 避免过多的特征输入导致模型欠拟合. 过去基于神经网络的预报建模研究只能通过控制变量的方式, 了解其在预报过程中的贡献程度, 即首先需要筛选输入模型中的参数组合, 然后通过增加/减少参数, 对比模型的预报效果, 筛选出对预报最有贡献的参数组合. 这种方法只能对每个参数进行间接评估, 无法准确量化其贡献. 但是借助AI可解释性算法, 可以将所有重要的参数引入模型, 并计算每个参数的贡献度, 通过排除一些贡献较少的参数, 达到快速优化模型预报效果的目的.

(2) 对复杂的空间天气AI模型进行可解释性分析, 使预报员能够理解模型预报决策的来源, 可提高

模型预报的可靠性.

(3) 成熟的AI可解释性框架配合AI的大数据挖掘能力, 在辅助科研人员探索不同地磁暴个例的发生发展机制也具有很广阔的前景.

(4) 图4中出现的各预报因子随预报提前量延长总贡献度减小的情况, 是由于L1点与地球位于日地因果链的末端, 缺乏因果链中的源头—太阳观测信息, 未来的研究应该选择合适的太阳观测参量引入模型, 并去掉模型中一些贡献较低的预报因子, 以提高未来3天 K_p 指数的预报准确度.

参考文献

- [1] FANG T W, KUBARYK A, GOLDSTEIN D, *et al.* Space weather environment during the SpaceX starlink satellite loss in February 2022[J]. *Space Weather*, 2022, **20**(11): e2022SW003193
- [2] LUO Bingxian, WANG Ronglan, LIU Wei, *et al.* Analysis of the space environment in the event of the destruction of the Starlink satellite by a geomagnetic storm[J]. *Space International*, 2022(5): 35-39 (罗冰显, 王荣兰, 刘卫, 等. “星链”卫星毁于地磁暴事件空间环境分析[J]. *国际太空*,

- 2022(5): 35-39)
- [3] MATZKA J, STOLLE C, YAMAZAKI Y, *et al.* The geomagnetic Kp index and derived indices of geomagnetic activity[J]. *Space Weather*, 2021, **19**(5): e2020SW002641
 - [4] WING S, JOHNSON J R, JEN J, *et al.* Kp forecast models[J]. *Journal of Geophysical Research: Space Physics*, 2005, **110**(A4): A04203. DOI: [10.1029/2004JA010500](https://doi.org/10.1029/2004JA010500)
 - [5] WINTOFT P, WIK M, MATZKA J, *et al.* Forecasting Kp from solar wind data: input parameter study using 3-hour averages and 3-hour range values[J]. *Journal of Space Weather and Space Climate*, 2017, **7**: A29
 - [6] TAN Y, HU Q H, WANG Z, *et al.* Geomagnetic index Kp forecasting with LSTM[J]. *Space Weather*, 2018, **16**(4): 406-416
 - [7] XIAO Jianjun, GONG Jiancun. Foreign space weather support capability building and the enlightenments to our country[J]. *Spacecraft Environment Engineering*, 2015, **32**(1): 9-13 (肖建军, 龚建村. 国外空间天气保障能力建设及对我国的启示 [J]. *航天器环境工程*, 2015, **32**(1): 9-13)
 - [8] LI Rong, QUAN Lin, LI Ling, *et al.* Review on status and development of space weather services[J/OL]. *Progress in Geophysics*, (2023-06-27)[2023-09-15]. <https://kns.cnki.net/kcms2/article/abstract?v=3uoq1hG8C45S0n9fL2suRadTyEV12pW9UrhTDCdPD66mi5bQfuHu-RxLnEeZBXUV7crkn5fWLWA2JP2zYD9YfcusDT2ApoMe&uniplat-form=NZKPT>
 - [9] WANG J J, LUO B X, LIU S Q, *et al.* A machine learning-based model for the next 3-day geomagnetic index (Kp) forecast[J]. *Frontiers in Astronomy and Space Sciences*, 2023, **10**: 1082737
 - [10] SHPRITS Y Y, VASILE R, ZHELAVSKAYA I S. Nowcasting and predicting the Kp index using historical values and real-time observations[J]. *Space Weather*, 2019, **17**(8): 1219-1229
 - [11] BERNOUX G, BRUNET A, BUCHLIN É, *et al.* Forecasting the geomagnetic activity several days in advance using neural networks driven by solar EUV imaging[J]. *Journal of Geophysical Research: Space Physics*, 2022, **127**(10): e2022JA030868
 - [12] 3-Day forecast | NOAA / NWS Space Weather Prediction Center[EB/OL]. [2023-09-15]. <https://www.swpc.noaa.gov/products/3-day-forecast>
 - [13] Forecast Kp index-space weather[EB/OL]. [2023-09-15]. <https://spaceweather.gfz-potsdam.de/products-data/forecasts/Kp-index-forecast>
 - [14] HU Xiaoyan, XU Jiyao, ZOU Ziming. Preliminary study on paradigm shift in space weather research driven by big data and artificial intelligence[J]. *Frontiers of Data & Computing*, 2023, **5**(2): 24-36 (胡晓彦, 徐寄遥, 邹自明. “大数据&人工智能”驱动的空间天气科研范式变革初步探索 [J]. *数据与计算发展前沿*, 2023, **5**(2): 24-36)
 - [15] HAM Y G, KIM J H, LUO J J. Deep learning for multi-year ENSO forecasts[J]. *Nature*, 2019, **573**(7775): 568-572
 - [16] ZHANG A X, FU T M, FENG X, *et al.* Deep learning-based ensemble forecasts and predictability assessments for surface ozone pollution[J]. *Geophysical Research Letters*, 2023, **50**(8): e2022GL102611
 - [17] WANG T Y, HUANG P. Superiority of a convolutional neural network model over dynamical models in predicting central pacific ENSO[J]. *Advances in Atmospheric Sciences*, 2024, **41**(1): 141-154
 - [18] SUNDARARAJAN M, TALY A, YAN Q Q. Axiomatic attribution for deep networks[C]//Proceedings of the 34th International Conference on Machine Learning. Sydney: JMLR.org, 2017: 3319-3328
 - [19] DAS A, RAD P. Opportunities and challenges in explainable artificial intelligence (XAI): a survey[OL]. arXiv preprint arXiv: 2006.11371, 2020
 - [20] LU Z, PU H M, WANG F C, *et al.* The expressive power of neural networks: a view from the width[OL]. arXiv preprint arXiv: 1709.02540, 2017
 - [21] PAPITASHVILI N E, KING J H. OMNI 1-min data set[DS/OL]. NASA Space Physics Data Facility, 2020[2023-09-25]. <https://hpde.io/NASA/NumericalData/OMNI/HighResolutionObservations/Version1/PT1M>
 - [22] KNIPP D J, GROSS N A. Understanding space weather and the physics behind it[C]//Proceedings of the American Geophysical Union, Fall Meeting 2011. Washington D C: AGU, 2011
 - [23] AUGUSTO C R A, NAVIA C E, DE OLIVEIRA M N, *et al.* The 2015 summer solstice storm: one of the major geomagnetic storms of solar cycle 24 observed at ground level[J]. *Solar Physics*, 2018, **293**(5): 84
 - [24] HAJRA R, TSURUTANI B T, LAKHINA G S. The complex space weather events of 2017 September[J]. *The Astrophysical Journal*, 2020, **899**(1): 3
 - [25] BADRUDDIN B, ASLAM O P M, DEROUICH M, *et al.* Forbush decreases and geomagnetic storms during a highly disturbed solar and interplanetary period, 4–10 September 2017[J]. *Space Weather*, 2019, **17**(3): 487-496
 - [26] HAM Y G, KIM J H, MIN S K, *et al.* Anthropogenic fingerprints in daily precipitation revealed by deep learning [J]. *Nature*, 2023, **622**(7982): 301-307

作者简介



王听雨 男, 1996 年 7 月出生于湖南省, 现为中国科学院空间科学中心特别研究助理, 主要研究方向为空间天气预报建模、人工智能与空间天气预报交叉。
E-mail: wangtingyu@nssc.ac.cn



罗冰显 (通信作者) 男, 1981 年 11 月出生, 现为中国科学院国家空间科学中心研究员、博士生导师, 国家“万人计划”青年拔尖人才获得者, 中国科学院大学岗位教授, 长期从事空间天气和空间环境预报和应用研究。
E-mail: luobx@nssc.ac.cn