# 基于改进 KAN-iTransformer 的掘进工作面瓦斯浓度预测研究*

张晓景[1]，李新宇[1]，孔少奇[2]

（1.山西忻州神达大桥沟煤业有限公司， 山西 忻州市 034099；
2.太原理工大学， 山西 太原 030024）

**摘要**：掘进工作面是煤矿开采过程中瓦斯事故频发的区域，针对其浓度变化呈现出的强非线性及时空动态相关性特点，提出了一种融合科尔莫格罗夫-阿诺德网络（KAN）与改进 iTransformer 的掘进工作面瓦斯浓度预测模型（KAN-iTransformer）。首先，通过数据预处理和 Spearman 相关系数分析构建了瓦斯浓度预测相关特征数据集；其次，基于倒置 Transformer 结构将完整变量序列视为标记，并引入图注意力机制显式建模传感器网络的空间拓扑关联，增加了传感器之间动态依赖的有效建模；再次，利用 KAN 的可学习激活函数替代传统线性权重，提升了模型对多源监测参数间复杂非线性关系的表征能力；最后，提出了融合点预测与区间估计的分级响应机制，实现了瓦斯超限风险的超前动态分级响应。现场试验结果表明，改进 KAN-iTransformer 模型在单步预测中的 *MSE*、*MAE* 和 *MAPE* 分别为 0.000 307，0.012 921 和 2.321 373%，明显优于 iTransformer 与 Transformer 两种对比模型；在多步预测中，改进 KAN-iTransformer 模型的 *MSE* 为 0.000 913，较对比模型分别降低了 63.7%与 80.9%，展现出更好的长期预测性能，有效实现了掘进工作面瓦斯浓度的精准预测与动态预警。

**关键词**：掘进工作面；瓦斯浓度预测；深度学习；单步预测；多步预测

**中图分类号**：TD712 **文献标识码**：A **文章编号**：1005-2763(2026)07-0133-10

## Research on Gas Concentration Prediction of Tunneling Working Face Based on Improved KAN-iTransformer

*Zhang Xiaojing*[1], *Li Xinyu*[2], *Kong Shaoqi*[2]

(1.Shanxi Xinzhou Shenda Daqiaogou Coal Industry Co., Ltd., Xinzhou, Shanxi 034099, China;
2.Taiyuan University of Technology, Taiyuan, Shanxi 030024, China)

**Abstract**: The tunneling working face is an area prone to frequent gas accidents during coal mining. In view of the strong nonlinear and spatiotemporal dynamic correlation characteristics of gas concentration variations, a prediction model for gas concentration at the tunneling working face, named KAN-iTransformer, was proposed, which integrates the Kolmogorov-Arnold Networks (KAN) and an improved iTransformer (Inverted Transformer). Firstly, a relevant feature dataset for gas concentration prediction was constructed through data preprocessing and Spearman correlation coefficient analysis. Secondly, based on the inverted Transformer structure, the complete variable sequence was regarded as a token, and a graph attention mechanism was introduced to explicitly model the spatial topological relationships of the sensor network, enabling effective modeling of the dynamic dependencies among sensors. Thirdly, the learnable activation functions of KAN were used to replace the traditional linear weights, enhancing the model's ability to characterize complex nonlinear relationships among multi-source monitoring parameters. Finally, a hierarchical response mechanism combining point prediction and interval estimation was proposed to enable advanced dynamic hierarchical response to gas overrun risk. The field test results show that the *MSE*, *MAE* and *MAPE* of the improved KAN-iTransformer model in single-step prediction are 0.000 307, 0.012 921 and 2.321 373%, respectively, significantly outperforming both the iTransformer and Transformer models. In multi-step prediction, the *MSE* of the improved KAN-iTransformer model is 0.000 913, which is 63.7% and 80.9% lower than those of the comparison models, respectively, demonstrating better long-term prediction performance and effectively realizing accurate prediction and dynamic early warning of gas concentration at the tunneling working face.

**Key words**: Tunneling working face, Gas concentration prediction, Deep learning, Single-step prediction, Multi-step prediction

---

* **收稿日期**：2025-10-16

**基金项目**：山西省基础研究计划-自由探索类青年项目(202203021222099)；山西省重点研发计划项目(202402080301013)；山西省重点研发计划项目(202402080301016)

**作者简介**：张晓景(1986—)，男，山西朔州人，高级工程师，主要从事采矿工程和生产技术等方面的研究工作。E-mail：3778127414@qq.com

**通信作者**：孔少奇(1989—)，男，山西大同人，博士，副教授，主要从事矿山压力与岩层控制、充填开采与二氧化碳储存利用及封存研究。E-mail：kongshaoqi@tyut.edu.cn

## 0 引言

煤炭作为当前全球一次能源的主体,长期在能源消费结构中占据主导地位[1]。尽管近年来清洁能源发展迅速,但煤炭在电力、冶金及化工等领域仍具有不可替代性。随着煤矿开采深度和规模不断增加,煤矿安全生产问题日益突出,其中瓦斯事故仍是威胁煤矿安全生产的主要灾害[2]。掘进工作面作为煤矿生产中的关键环节,是瓦斯聚集与异常涌出的高发区,其瓦斯浓度变化具有非线性、时变性和突发性特征。因此,如何实现对掘进工作面瓦斯浓度的精准预测,成为煤矿安全开采领域亟待解决的重大问题[3]。

传统瓦斯浓度预测方法主要是基于数值模拟和机器学习方法,Bi 等[4]基于随机森林和改进回廊算法的混合机器学习模型对瓦斯涌出量相关影响因素进行筛选,实现了瓦斯浓度的精准预测。Zeng 等[5]基于多个工作面的多源数据进行研究,建立了基于灰色预测模型的瓦斯涌出量预测模型。此外,其他学者也基于极端梯度提升(eXtreme Gradient Boosting, XGBoost)[6]、支持向量机(Support Vector Machine, SVM)[7]、极限学习机(Extreme Learning Machines, ELM)[8]等经典机器学习算法实现瓦斯涌出量预测。但这类方法无法挖掘时序数据间的内部关联,且不能理解长期非平稳时间序列的非线性关系。

随着人工智能技术的不断发展,国内外学者将深度学习算法应用于瓦斯浓度预测,解恒星等[9]、杨小彬等[10]、李鹏等[11]基于长短期记忆网络(Long Short-Term Memory, LSTM)算法,通过门控单元捕捉瓦斯浓度时序依赖关系,其预测精度和稳定性均优于传统模型。盛武等[12]、刘超等[13]采用门控循环单元(Gate Recurrent Unit, GRU),在保证训练精度的同时提升了模型训练速度。在此基础上,相关学者结合主成分分析[14-15]、注意力机制[16]、自适应优化算法[17-18]优化模型网络结构及各项超参数,进一步提高了模型的拟合能力。胡青松等[19]构建了融合时间卷积网络与时序生成对抗网络的瓦斯浓度预测模型,改善了瓦斯数据小样本过拟合的问题。

虽然当前多数研究已可实现工作面瓦斯浓度预测,但仍存在以下问题:现有研究大多基于瓦斯浓度历史数据的变化规律进行单步预测,未能深入挖掘瓦斯浓度与外部特征之间潜在的时序依赖,对多因素影响下的瓦斯浓度多步预测研究较少。针对上述问题,本文以掘进工作面多源传感器数据为基础,构建了基于改进 KAN-iTransformer 的掘进工作面瓦斯浓度预测模型。结合井下数据特性完成数据去噪与异常值处理;以倒置 Transformer(Inverted Transformer, iTransformer)为主干,融合图注意力以捕捉传感器间动态依赖,并引入更具可解释性的科尔莫格罗夫-阿诺德网络(Kolmogorov-Arnold Networks, KAN)函数逼近模块,以强化样本效率与稳定性;提出面向安全阈值管理的多步点-区间联合输出与预警策略,实现对多步预测结果的置信评估和越限风险的早期识别。最后,将本文模型与基线模型进行对比分析,以验证本文模型的预测效能。

## 1 工程背景及数据处理

### 1.1 工程背景与数据来源

本研究依托于山西忻州神达大桥沟煤业有限公司某煤与瓦斯突出矿井的掘进工作面。该矿地质条件复杂,瓦斯赋存量高,属典型的高风险矿井。试验掘进工作面位于该矿三号煤采区,工作面走向长 278 m,倾向长 2 175 m,埋藏深度为 501～590 m,煤层厚度为 2.1～3.7 m,采用压入式通风。该掘进工作面绝对瓦斯涌出量高达 27 $m^3$/min,相对瓦斯涌出量为 5 $m^3$/t,原始瓦斯含量为 11～14 $m^3$/t,煤层瓦斯压力为 1.0～1.3 MPa,具备典型的高瓦斯动力学特征。研究中收集的数据涵盖 2025 年 3 月 1 日至 2025 年 4 月 1 日期间的多源信息,每秒进行一次瓦斯数据的采集,总共采集样本 43 000 个。

该工作面的多源数据主要分为三类:第一类为生产监测数据,包含瓦斯浓度、一氧化碳浓度、温度、风速等参数;第二类为设备监测数据,包含掘进机截割电流、掘进机截割温度、开采速度等参数;第三类为瓦斯地质因素,包括煤层瓦斯抽采含量等参数。本文收集的掘进工作面现场数据类型全面,能够全面表征瓦斯赋存、通风条件、设备运行与作业工况之间的动态关系。

### 1.2 特征提取

为了揭示各监测特征之间的相互关系,本文对掘进工作面瓦斯浓度及相关特征参数进行了相关性分析。Spearman 相关系数是一种衡量变量间相似性的指标,与皮尔逊相关系数相比,该方法既不要求

变量服从正态分布，也不假设变量间存在线性关系，在变量间关系模式不明确时，能更有效地度量相关性。计算公式如式(1)所示，其中 $R(x_i)$ 和 $R(y_i)$ 是变量 $X$ 和 $Y$ 的秩统计量，$\overline{R(X)}$ 和 $\overline{R(Y)}$ 是变量 $X$ 和 $Y$ 的统计均值，$\rho$ 为两个变量之间的 Spearman 相关系数。Spearman 相关系数评估变量间相关性的判断准则与 Pearson 相关系数相同：系数的绝对值越接近 1，表明变量间的相关性越强；越接近 0，则相关性越弱；若系数为 0，则认为变量之间不存在相关性。

$$\rho=\frac{\sum_{i=1}^{n}(R(x_i)-\overline{R(X)})(R(y_i)-\overline{R(Y)})}{\sqrt{\sum_{i=1}^{n}(R(x_i)-\overline{R(X)})^2}\cdot\sqrt{\sum_{i=1}^{n}(R(y_i)-\overline{R(Y)})^2}} \tag{1}$$

Spearman 相关系数热图如图 1 所示，变量包含工作面瓦斯浓度($Y$)、掘进进尺($X_1$)、风机入风瓦斯($X_2$)、回风瓦斯($X_3$)、掘进机截割电流($X_4$)、回顺钻机下风侧瓦斯($X_5$)、风速($X_6$)、温度($X_7$)、一氧化碳($X_8$)。从图 1 可以看出，瓦斯浓度与回顺钻机下风侧瓦斯浓度、回风瓦斯、掘进机截割电流具有较强的正相关关系，表明以上参数能够在一定程度上反映工作面瓦斯浓度的波动情况。瓦斯浓度与风速和温度呈现一定的负相关关系，说明良好的通风条件和环境参数有助于降低瓦斯积聚风险。

综上所述，本文选取回风瓦斯、掘进机截割电流、回顺钻机下风侧瓦斯、风速、温度作为掘进工作面瓦斯浓度预测的特征参数。

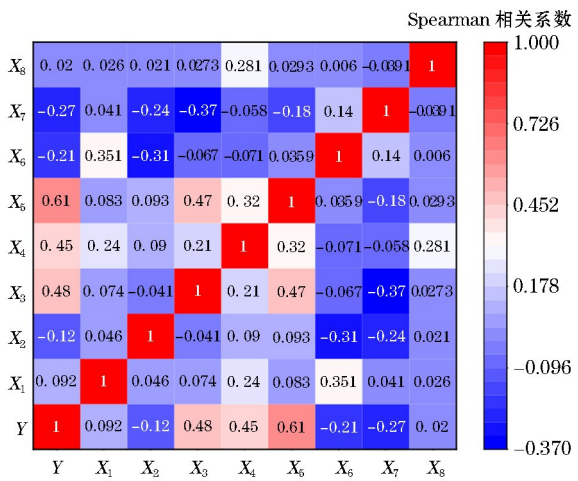



**图 1 Spearman 相关系数热图**

**Fig.1 Spearman correlation coefficient heat map**

### 1.3 数据预处理

煤矿井下环境复杂多变，原始传感器数据常受到电磁干扰、传感器校准漂移、设备故障等物理因素及传输不稳定的影响，会导致数据产生缺失值和异常值等情况，同时还存在多源异构数据采样频率不一致、量纲差异显著等问题。为确保模型训练的精度和预测结果的可靠性，需要对特征数据进行预处理。

本文采用直接删除法处理单个孤立缺失值，使用平滑指数法处理连续性缺失值，在监测到缺失值时跳过该点的误差计算，但继续使用模型的平滑方程来递推计算下一个时间点的水平分量和趋势分量，从而自然地产生一个对缺失点的估计值，并以该估计值完成缺失值填补。基于箱线图法判别数据样本中的异常值，首先，获取第一四分位数(第 25 百分位数)$Q_1$ 和第三四分位数(第 75 百分位数)$Q_2$；然后，通过计算数据样本四分位数间的差值确定四分位距 $IQR$；最后，基于该区间设定上限 $U_B$ 和下限 $L_B$ 以识别异常值。下限计算原理为 $L_B=Q_1-1.5\times IQR$，上限计算原理为 $U_B=Q_2+1.5\times IQR$，不满足此赋值区间的值判定为异常值。

为进一步消除传感器数据采集过程中的随机噪声和测量误差，本研究采用卡尔曼滤波算法对原始数据进行平滑处理。卡尔曼滤波基于状态空间模型，通过递归预测和更新两个步骤实现最优估计。本文基于网格搜索方法进行参数寻优，通过计算滤波前后信号噪声比选定最优参数组合。其中过程噪声协方差和观测噪声协方差分别设定为 0.01 和 0.1，设状态转移矩阵和观测矩阵均为 1。预测状态和预测误差协方差计算公式见式(2)和式(3)，其中 $\hat{x}_k$ 表示 $k$ 时刻的预测状态，$P_k^-$ 为 $k$ 时刻的预测误差协方差，$\hat{x}_{k-1}$ 表示 $k-1$ 时刻最优后验状态估计值，$P_{k-1}$ 表示 $k-1$ 时刻后验误差协方差，$Q$ 为过程噪声协方差。

$$\hat{x}_k=\hat{x}_{k-1} \tag{2}$$

$$P_k^-=P_{k-1}+Q \tag{3}$$

在更新过程中，通过观测数据 $z_k$ 对预测状态进行修正，重新计算新的误差协方差。$K_k$ 表示卡尔曼增益，用于衡量预测值与观测值之间的权重关系，$R$ 表示观测噪声协方差，$K_k$ 的计算公式见式(4)。根据预测状态和观测值给出修正后的状态估计 $\hat{x}_k$，见式(5)。公式(6)表示更新后的误差协方差 $P_k$。

得到的数据预处理效果如图 2 所示。从图 2 中可以看出,经过上述预处理过程后能够明显抑制原始掘进机电流信号中的高频噪声和数据采集过程中的异常值,同时完整保留了原始信号的全部宏观动态趋势与关键特征。

$$K_k=\frac{P_k^-}{P_k^-+R} \tag{4}$$

$$\hat{x}_k=\hat{x}_k^-+K_k(z_k-\hat{x}_k^-) \tag{5}$$

$$P_k=(1-K_k)P_k^- \tag{6}$$

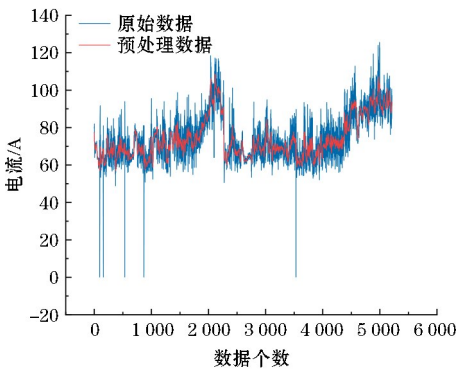



**图 2 数据预处理效果**

**Fig.2 Data preprocessing effect**

## 2 基于 KAN-iTransformer 的瓦斯浓度预测模型

### 2.1 模型总体流程

构建瓦斯浓度预测模型的整体框架,如图 3 所示,包含数据处理、模型训练与模型评估 3 个核心阶段。首先通过对原始数据集进行预处理,以保障后续模型构建的可靠性。然后根据采集频率对特征变量进行重采样,将所有特征数据与工作面瓦斯浓度数据进行时间对齐,按照 8∶2 的比例将数据分为训练集和测试集。

模型训练阶段,首先将训练集数据导入模型,完成模型参数初始化,然后计算模型输出预测值和输入真实值的损失,通过梯度计算与损失函数优化模型性能,并依据评估指标不断调整模型超参数,直至模型训练效果达到最优。模型评估阶段,利用测试集完成预测,经反归一化处理后将预测结果转化为实际的瓦斯浓度预测值,并计算均方误差(MSE)、平均绝对误差(MAE)和平均绝对百分比误差(MAPE)多项性能指标,以全面评估模型精度与泛化能力。

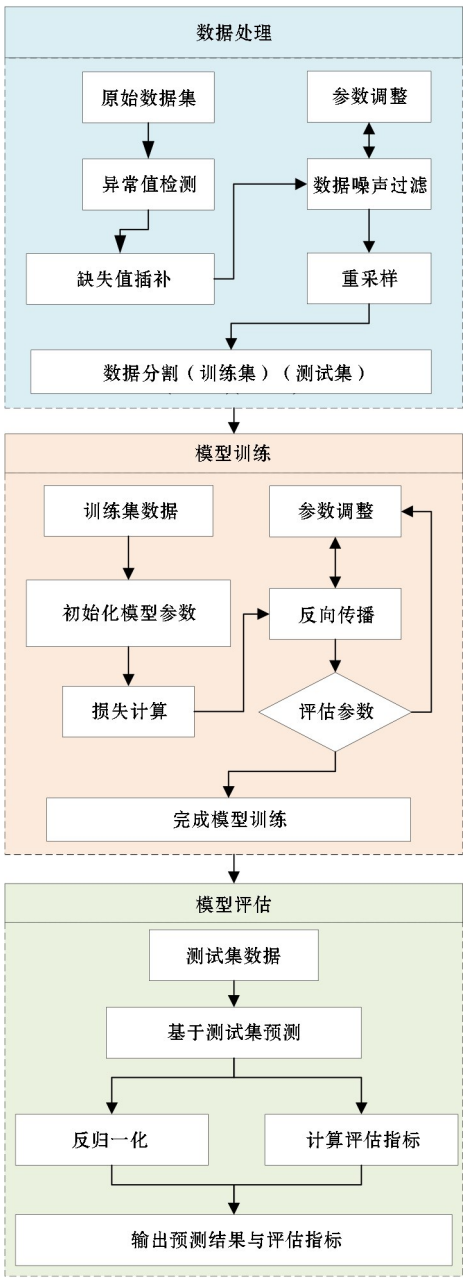



**图 3 模型总体框架**

**Fig.3 The overall framework of the model**

### 2.2 KAN 模型架构

KAN 是基于科尔莫戈罗夫-阿诺德表示定理提出的新型神经网络结构[20]。该定理指出任何多元函数都可以表示为有限个连续单变量函数与加法操作的组合。KAN 将线性权重替换为可学习的一维函数,通常采用 $B$ 样条(B-spline)进行参数化。网络节点仅执行求和运算,非线性建模完全由连接边上的函数实现。给定第 $l$ 层的输入向量 $x^{(l)}\in\boldsymbol{R}^{nl}$,输出为 $x^{(l+1)}\in\boldsymbol{R}^{nl+1}$,每条连接边对应一个单变量函数 $\varphi_{j,i}^{(l)}$,则第 $l+1$ 层的输出见式(7),其中

$\varphi(x)$ 的计算公式见式(8)，由残差基函数与 $B$ 样条线性组合构成，其中 $w_b$ 和 $w_s$ 表示可学习的权重，$silu(x)$ 表示预先定义的非线性函数，$c_k$ 为可学习系数，$B_k(x)$ 为 B-spline 基函数。

$$x_j^{(l+1)}=\sum_{i=1}^{nl}\varphi_{j,i}^{(l)}(x_i^{(l)}),j=1,\cdots,n_{l+1} \tag{7}$$

$$\varphi(x)=w_b\cdot silu(x)+w_s\sum_k c_k B_k(x) \tag{8}$$

### 2.3　改进 iTransformer 模型架构

瓦斯浓度预测模型的核心是倒置 Transformer，其重新定义了多变量时间序列的分词方案[21]。倒置 Transformer 不再将每个时间步视为独立分词[22]，而是将每个变量的完整历史序列视为一个变量分词。这种设计将自注意力机制的关注点从时间维度转移到变量维度，从而能够显式建模变量间的关联关系。模型架构如图 4 所示，该架构包含嵌入层、倒置 Transformer 模块、预测头 3 个主要组件。

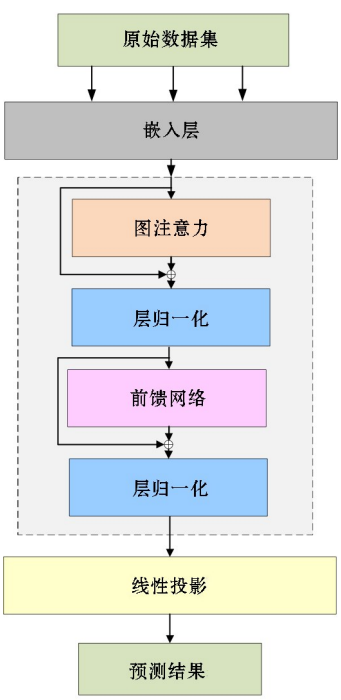



**图 4　改进 iTransformer 模型架构**

**Fig.4　Improved iTransformer model architecture**

#### 2.3.1　嵌入层

在多变量时间序列预测任务中，历史监测数据可表示为 $\boldsymbol{X}=[x_1,\cdots,x_T]\in\mathbb{R}^{T\times N}$，其中 $\mathbb{R}$ 为实数域，$\boldsymbol{X}$ 为时间 $T$ 内的多变量世界序列数据矩阵，$N$ 为变量个数。与时间维度上定义标记的标准 Transformer 不同，倒置 Transformer 在变量维度上重新定义了标记。每个变量的完整历史序列 $\boldsymbol{X}_{:,n}\in\mathbb{R}^T$ 均被视为单个变量标记。嵌入过程可表示为式(9)，生成的嵌入向量被堆叠为式(10)。通过将每个变量的完整时间序列独立编码为变量标记，避免了传统方法中将同一时间步的所有变量混合嵌入为时间标记所造成的时间错位及相关性丢失问题。

$$h_n^{(0)}=\dot{f}_{\text{emb}}(X_{:,n}),n=1,\cdots,N \tag{9}$$

$$\boldsymbol{H}^{(0)}=[h_1^{(0)},h_2^{(0)},\cdots,h_N^{(0)}]\in\boldsymbol{R}^{N\times D} \tag{10}$$

式中：$h_n^{(0)}$ 为第 $n$ 个变量的嵌入值；$\dot{f}_{\text{emb}}(\cdot)$ 为嵌入函数；$\boldsymbol{H}^{(0)}$ 为模型输入层的嵌入矩阵；$D$ 为嵌入维度。

#### 2.3.2　改进倒置 Transformer 模块

每个倒置 Transformer 模块由层归一化、前馈网络、图注意力 3 个主要部件组成。

为应对瓦斯监测变量的非平稳性和尺度异质性，在每个倒置 Transformer 模块内部应用了时间层归一化(TLN)。与基础 Transformer 模型在每个时间步跨特征维度进行归一化不同，TLN 沿着每个变量序列的时间维度进行归一化，能够进一步减小多个变量间的时间错位问题。具体而言，对于变量 $n$ 的归一化表示计算见式(11)至式(13)，其中 $\mu_n$ 和 $\sigma_n$ 为沿变量 $n$ 的时间轴计算的均值与标准差，$\tilde{h}_{n,t}$ 为变量 $n$ 在时间步 $t$ 归一化后的特征值，$h_{n,t}$ 为变量 $n$ 在时间步 $t$ 的原始特征值。

$$\tilde{h}_{n,t}=\frac{h_{n,t}-\mu_n}{\sigma_n} \tag{11}$$

$$\mu_n=\frac{1}{T}\sum_{t=1}^{T}h_{n,t} \tag{12}$$

$$\sigma_n^2=\frac{1}{T}\sum_{t=1}^{T}(h_{n,t}-\mu_n)^2 \tag{13}$$

在注意力机制和层归一化处理后，每个变量标记都会经过时间方向的前馈网络(FFN)，以优化其时间表征结果。与基础 Transformer 模型中 FFN 作用于每一时间点的标记不同，倒置 Transformer 模型中每个变量标记都包含了该变量整个历史窗口的信息，因此这里的 FFN 只用于建模单个变量内部的整个时间的依赖关系。该结构通过提取变量内部的复杂非线性关系，实现对变量内部季节性、周期性和瞬态冲击等共性时间模式的协同学习。

在原始倒置 Transformer 模型中，自注意力机制作用于变量维度，每个变量标记都对应传感器测量的历史序列。该设计虽然能有效建模变量间的依赖关系，但在地下煤矿中，监测传感器并非随机分

布,而是沿掘进巷道、通风路径和工作面开采设备按照标准布设,瓦斯浓度的动态演化受气流拓扑结构、巷道连通性及传感器空间邻接性的显著影响。

在改进倒置 Transformer 模块中,采用图注意力机制替换原来的自注意力[23],利用各个传感器的空间拓扑结构,引入空间先验知识,注意力聚合仅发生在具有物理或统计意义的相邻节点间,避免了远距离或不相关变量间产生虚假依赖关系的问题,大大提高了模型的归纳能力,使其可用于解决多传感器瓦斯浓度预测的特定挑战。

每个变量标记 $h_i \in \mathbb{R}^D$ 被视为图 $G=(V,E)$ 中的一个节点,其中边集 $E$ 可编码以下 3 种关系之一:(1) 物理网络拓扑结构;(2) 传感器的空间邻近性;(3) 变量间具有统计显著相关性的关联。

对于每个相连节点 $(i,j) \in E$,对节点特征的更新过程见式(14),其中 $\boldsymbol{W}$ 为可学习的权重矩阵,$\sigma(\cdot)$ 为激活函数,$E(i)$ 表示节点 $i$ 的邻接节点集合,$\alpha_{ij}^{(l)}$ 为注意力系数,其计算公式如式(15) 所示,其中 $\boldsymbol{a}$ 为可学习的注意力向量,$\|$ 表示向量拼接操作,LeakyReLU(·) 表示带泄露的 ReLU 激活函数。

$$h_i^{(l+1)}=\sigma\Big(\sum_{j\in E(i)}\alpha_{ij}^{(l)}\boldsymbol{W}h_j^{(l)}\Big) \tag{14}$$

$$\alpha_{ij}^{(l)}=\frac{\exp(\mathrm{LeakyReLU}(\boldsymbol{a}^{\mathrm{T}}[\boldsymbol{W}h_i^{(l)}\,\|\,\boldsymbol{W}h_j^{(l)}]))}{\sum_{k\in E(i)}\exp(\mathrm{LeakyReLU}(\boldsymbol{a}^{\mathrm{T}}[\boldsymbol{W}h_i^{(l)}\,\|\,\boldsymbol{W}h_k^{(l)}]))} \tag{15}$$

为了提升模型表达能力,引入多头注意力机制,节点特征更新过程见式(16)。其中 $M$ 为注意力头的数量,$\boldsymbol{W}^{(m)}$ 和 $\alpha^{(m)}$ 为各个注意力头独立的参数。

$$h_i^{(l+1)}=\|_{m=1}^{M}\sigma\Big(\sum_{j\in N(i)}\alpha_{ij}^{(l,m)}\boldsymbol{W}^{(m)}h_j^{(l)}\Big) \tag{16}$$

在图注意力更新后,输出结果仍输入至前馈网络,该改进模型在保留倒置 Transformer 基础架构的基础上,引入了基于图的约束,使跨变量依赖关系的学习更加符合传感器的空间拓扑与物理规律。通过嵌入煤矿传感器的拓扑信息,改进模型能够同时捕捉局部空间内变量间的复杂相关性与全局依赖关系,从而提升井下瓦斯浓度预测的物理一致性与鲁棒性。

2.3.3 预测头

原倒置 Transformer 模型中最终标记 $h_n^{(l)} \in \mathbb{R}^N$ 通过回归网络被投影至未来预测时段,在原始设计中,这一投影由多层感知机实现。为了提升模型对煤矿井下瓦斯浓度这类高度非线性和非平稳数据的动态表征能力,将多层感知机替换为 KAN,利用样条基函数展开增强模型的非线性表达能力。结合 2.2 小节,最终得到预测值计算公式见式(17),其中 $S$ 为待预测的时间步长,$\hat{y}_s$ 为第 $s$ 个待预测时间步的预测值。

$$\hat{y}_s=\sum_{i=1}^{D}\varphi_i^{(s)}(h_{n,i}^{(l)}),s=1,2,\cdots,S \tag{17}$$

## 3 试验结果与分析

### 3.1 模型评估及试验配置

基于改进 KAN-iTrnsformer 瓦斯浓度预测模型的训练采用 Intel(R)Core-i7 11850H,主频 32.50 GHz,内存 32 G,显卡 NVIDIA T600。该模型采用均方误差(*MSE*) 作为损失函数进行模型训练,其中 $\hat{y}_i$ 为模型预测值,$y_i$ 为瓦斯浓度真实值。采用 Adam 优化算法进行训练,学习率设置为 0.001,权重衰减系数设置为 0.000 01,并启用早停机制。

将数据集中前 80% 的数据作为训练集,后 20% 的数据作为测试集,把完成特征构造的数据输入瓦斯浓度预测模型,以 1 min 为时间单位对试验工作面未来瓦斯浓度进行多步预测,最终通过 *MSE*、*MAE* 和 *MAPE* 评估模型的预测效果。

$$MSE=\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2 \tag{18}$$

$$MAE=\frac{1}{N}\sum_{i=1}^{N}|y_i-\hat{y}_i| \tag{19}$$

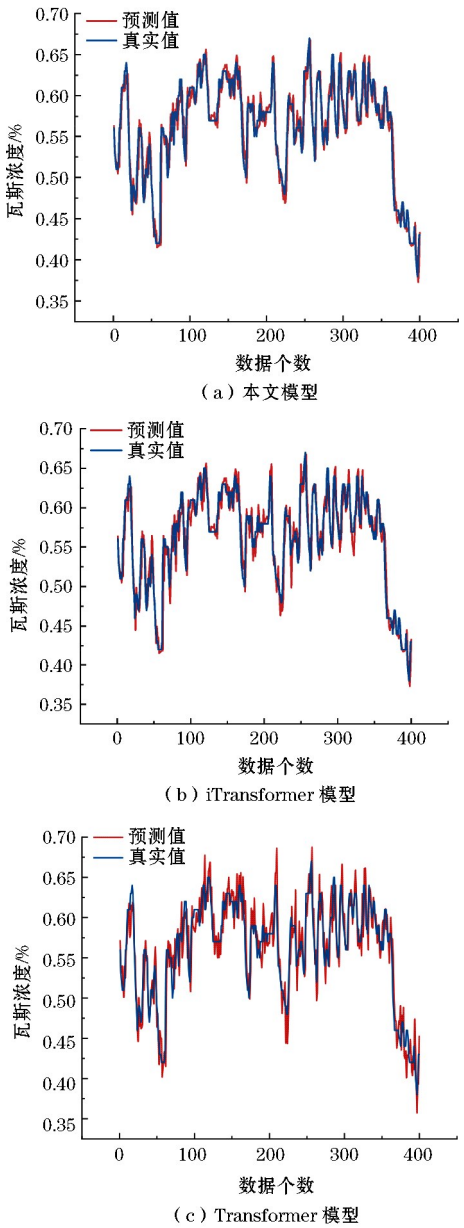
$$MAPE=\frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \tag{20}$$

### 3.2 模型单步预测效果对比分析

图 5 为模型预测值与真实值的对比曲线。由图 5 可知,本文提出的掘进工作面瓦斯浓度预测模型在 400 个时间步长(采样间隔 1 min)的瓦斯浓度预测中展现出优异性能,预测值和真实值高度吻合,能够准确刻画瓦斯浓度在不同时间点的波动趋势。在瓦斯浓度快速变化阶段,模型不仅能够正确预测波动的起伏趋势,还能准确拟合峰值与谷值的变化幅度。在瓦斯浓度慢速变化阶段,模型预测呈现周期性小幅波动,不存在明显偏移。当真实值表现出明显的快速下降趋势,即从 0.58 下降至 0.40 左右时,预测曲线在该阶段保持同步下降的变化趋势。

通过对比分析发现,当瓦斯浓度在短时间内发生较大波动时,iTransformer 模型虽然能够反映瓦斯浓度真实变化,但会产生更多的瞬时尖峰,无法精

准预测数据的变化幅度。Transformer 模型在预测过程中出现了较为明显的偏差和滞后性，当真实值已经开始上升或下降时，模型的预测值仍停留在之前的水平，当真实值没有发生较明显波动时，预测值依然会发生局部波动，偏离真实值，这表明 Transformer 模型无法有效捕捉数据中的动态时序依赖性。




（a）本文模型

（b）iTransformer 模型

（c）Transformer 模型

**图 5　不同模型单步预测结果对比**

**Fig.5　Comparison of single-step prediction results of different models**

各个模型单步预测的性能指标对比见表 1。由表 1 可知，本文模型的所有指标上均优于 iTransformer 和 Transformer 模型，本文模型的 *MSE* 为 0.000 307，分别比二者降低了 14.2% 和 45.4%；本文模型的 *MAE* 为 0.012 921，分别比二者降低了 7.8% 和 29.8%。*MSE* 对误差更为敏感，其显著降低表明本文模型有效减少了预测中出现大幅偏差的情况，增强了模型预测的稳定性。*MAE* 的降低则表明模型的整体平均绝对误差更小，预测精度更高。本文模型的 *MAPE* 为 2.321 373%，该值较 iTransformer 和 Transformer 模型分别降低了 7.6%和 30.1%。本文模型的 $R^2$ 高达 0.916 450，较另外两个模型提升了 3.8%和 8.2%，具有更好的拟合能力。

这是因为本文模型通过图注意力和 KAN 加强了模型对各个变量间真实依赖关系和非线性特征的动态表征，使得本文模型在预测精度、稳定性和拟合能力方面均优于对比模型。

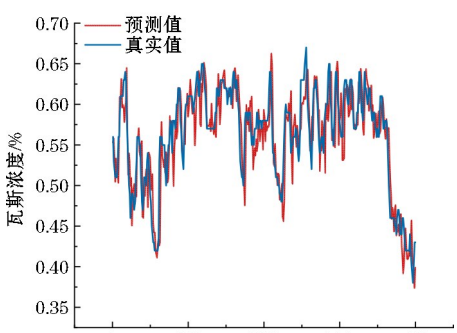
**表 1　不同模型单步预测效果对比**

**Table 1　Comparison of the single-step prediction effect of different models**

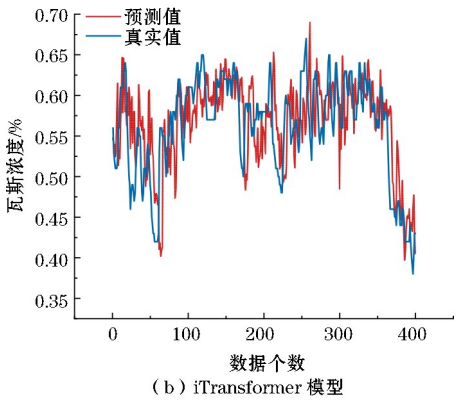
| 模型 | *MSE* | *MAE* | *MAPE*/% | $R^2$ |
|---|---|---|---|---|
| 本文模型 | 0.000 307 | 0.012 921 | 2.321 373 | 0.916 450 |
| iTransformer 模型 | 0.000 358 | 0.014 020 | 2.511 322 | 0.882 508 |
| Transformer 模型 | 0.000 562 | 0.018 397 | 3.320 037 | 0.846 822 |

### 3.3　模型多步预测效果对比分析

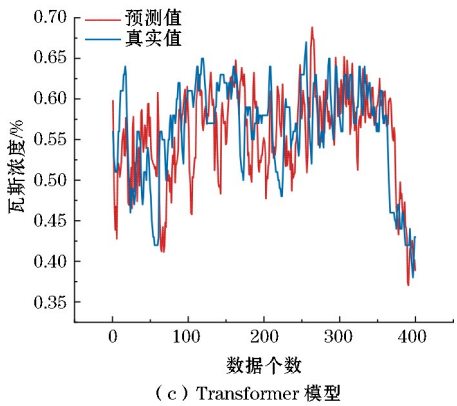
为了进一步验证本文模型在瓦斯浓度多步预测的性能，设置预测步长为 10 进行模型多步预测效果分析，对比结果如图 6 所示。从图 6 可以看出，本文模型相较单步预测结果存在不同幅度的波动和少量偏差，但在数据变化趋势上仍然能体现出较高的拟合能力，能够准确预测出瓦斯浓度变化的整体演化趋势。另外两种对比模型的整体预测曲线与真实值存在较大幅度的系统性偏差和波动衰减现象，当真实值平稳变化时，模型预测值能够大致反映真实情况，当真实值发生急剧上升和下降时，预测曲线与真实值出现分离，滞后现象明显。




（a）本文模型




(b) iTransformer 模型




(c) Transformer 模型

**图 6 不同模型多步预测结果对比**

**Fig.6 Comparison of multi-step prediction results of different models**

各个模型多步预测的性能指标对比见表 2。由表 2 可知,本文模型各指标均优于对比模型,本文模型的 *MSE* 为 0.000 913,相比两个对比模型分别降低了 63.7%和 80.9%,表明本文模型的多步预测能力准确性明显优于对比模型。本文模型的 $R^2$ 高达 0.851 230,相较单步预测仅降低了 0.065 22,这表明本文模型的拟合能力并未发生大幅退化,能够同时保证短期预测的精确性和长期预测的稳定性。

**表 2 不同模型多步预测效果对比**

**Table 2 Comparison of the multi-step prediction effect of different models**

| 模型 | *MSE* | *MAE* | *MAPE*/% | $R^2$ |
|---|---|---|---|---|
| 本文模型 | 0.000 913 | 0.022 867 | 4.117 202 | 0.851 230 |
| iTransformer 模型 | 0.002 517 | 0.039 033 | 7.183 084 | 0.706 521 |
| Transformer | 0.004 784 | 0.052 639 | 9.741 005 | 0.303 680 |

### 3.4 模型应急响应机制

在煤矿掘进工作面的复杂环境中,瓦斯浓度多步预测存在不可避免的误差与不确定性,若仅依赖点预测值与固定阈值进行超限判断,容易出现误报与漏报。本文在多步预测基础上引入预测区间,通过计算预测残差的分布特征来估计模型预测的不确定性,以 95%置信水平为例,第 $k$ 步预测的区间形式见公式(21)。其中 $\hat{y}_{t+k}^{(L)}$ 和 $\hat{y}_{t+k}^{(U)}$ 分别为上下界置信区间,置信水平设为 95%。

$$\hat{y}_{t+k}^{(L)} \leqslant y_{t+k} \leqslant \hat{y}_{t+k}^{(U)} \tag{21}$$

为瓦斯浓度超限设置置信阈值,可得应急响应分级效果,如图 7 所示,当预测均值低于阈值,且置信区间整体位于安全范围内时,判定为绿色等级,无需进行任何应急措施。当预测上限超过阈值,但单点预测值仍处于安全范围内时,判定为蓝色等级,此时处于潜在风险阶段,建议加强该区域瓦斯浓度监测,增设便携式瓦斯检测定点抽检。当预测均值已超过阈值,但区间下限仍低于阈值时,判定为橙色等级,此时处于一般风险阶段,需要采取预防性措施提高瓦斯抽采效率。当预测区间整体超过阈值,即置信下界也超过阈值时,判定为红色等级,此时处于高风险阶段,必须检查各个传感器是否正常,做好设备紧急断电准备并启动人员撤离等强制性应急措施。

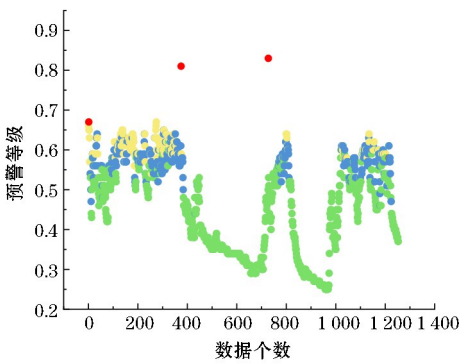



**图 7 应急响应分级效果**

**Fig.7 Emergency response classification effect**

## 4 结论

本文针对煤矿掘进工作面瓦斯浓度预测中存在的强非线性、多源耦合与时空动态特性等挑战,提出了一种融合 KAN 网络与改进倒置 Transformer 的瓦斯浓度预测模型(KAN-iTransformer)。通过多源传感器数据的特征提取与预处理,结合长序列时序建模与空间拓扑感知的图注意力机制,实现了掘进工作面瓦斯浓度的精准预测。得到的主要结论如下。

(1) 构建了基于 KAN 和改进倒置 Transformer 的掘进工作面瓦斯浓度预测模型,其中 KAN 网络

采用可学习核函数替代传统线性权重，增强了模型对非线性关系的表征能力；改进的倒置 Transformer 通过倒置序列编码与图注意力机制，有效捕捉了多变量间的长期依赖关系。

（2）本文模型在单步预测中的 *MSE* 为 0.000 307，*MAE* 为 0.012 921，*MAPE* 为 2.321 373%，$R^2$ 高达 0.916 450，相比倒置 Transformer 和 Transformer 模型，误差指标显著降低（*MSE* 降幅分别为 14.2% 和 45.4%），表明本文模型在捕捉瞬态变化与峰值特征方面具有优越性；在多步预测中，本文模型的 *MSE* 为 0.000 913，虽较单步预测有所上升，但显著低于对比模型（降幅分别为 63.7% 和 80.9%），且 $R^2$ 为 0.851 230，表明本文模型在长期预测中仍能保持较高的趋势一致性和稳定性，有效缓解了传统 Transformer 类模型在长序列预测中普遍存在的滞后性和累积误差问题。

（3）通过引入预测区间估计与分级预警机制，使模型在多步预测中能够量化预测不确定性，依托置信区间实现瓦斯超限风险动态分级管控，可为煤矿智能应急决策提供可靠依据。

（4）由于本文试验数据仅取自山西忻州某单一高瓦斯煤矿，模型在不同地质构造、瓦斯压力及采矿工艺下的泛化能力尚未得到充分验证。后续将采集多矿井、多地质条件与采掘工艺的实测数据构建综合性数据集，对预训练模型进行微调，进一步提升模型泛化能力。

**参考文献**(References)：


[1] 王国法，刘峰，庞义辉，等.煤矿智能化：煤炭工业高质量发展的核心技术支撑[J].煤炭学报，2019，44(2)：349-357.
Wang Guofa，Liu Feng，Pang Yihui，et al. Coal mine inellectualization：the core technology of high quality development[J]. Journal of China Coal Society，2019，44(2)：349-357.

[2] 杨科，郭鹏慧，袁亮，等.深部开采煤矿典型动力灾害孕灾主控因素与机制研究进展[J].煤炭学报，2025，50(7)：3466-3487.
Yang Ke，Guo Penghui，Yuan Liang，et al. Research progress on key controlling factors and mechanisms of typical dynamic hazards in deep coal mining.[J]. Journal of China Coal Society，2025，50(7)：3466-3487.

[3] 张志刚，张庆华，刘军.我国煤与瓦斯突出及复合动力灾害预警系统研究进展及展望[J].煤炭学报，2024，49(增刊 2)：911-923.
Zhang Zhigang，Zhang Qinghua，Liu Jun，et al. Research progress and prospects of coal and gas outburst and composite dynamic disaster warning systems in China[J]. Journal of China Coal Society，2024，49(Suppl.2)：911-923.

[4] Bi Shengbao，Shao Liangshan，Qi Zihan，et al. Prediction of coal mine gas emission based on hybrid machine learning model[J]. Earth Science Informatics，2023，16(1)：501-513.

[5] Zeng Jun，Li Qinshen. Research on prediction accuracy of coal mine gas emission based on grey prediction model[J]. Processes，2021，9(7)：1147.

[6] Song Weihua，Han Xiaowei，Qi Jifei. Prediction of gas emission in the working face based on LASSO-WOA-XGBoost[J]. Atmosphere，2023，14(11)：14111628.

[7] 杨宝琳，刘杰，娄来宗.基于粒子群算法优化 BP 神经网络及 SVM 的瓦斯超限预测预警方法[J].科学技术与工程，2025，25(26)：11098-11106.
Yang Baolin，Liu Jie，Lou Laizong. Prediction and early warning method for gas overrun based on PSO optimized BP neural network and SVM.[J]. Science Technology and Engineering，2025，25(26)：11098-11106.

[8] 张科学，李中旭，陈学习，等.基于 PCA-HPO-ELM 的智能化矿井瓦斯涌出量预测研究[J].矿业科学学报，2025，10(5)：879-889.
Zhang Kexue，Li Zhongxu，Chen Xuexi，et al. Gas emission prediction of intelligent mines based on PCA-HPO-ELM.[J]. Journal of Mining Science and Technology，2025，10(5)：879-889.

[9] 解恒星，张雄，董锦洋，等.基于 CNN-BiLSTM 的矿井瓦斯涌出量预测模型[J].中国安全生产科学技术，2024，20(11)：53-59.
Xie Hengxing，Zhang Xiong，Dong Jinyang，et al. Prediction model of mine gas emission based on CNN-BiLSTM.[J]. Journal of Safety Science and Technology，2024，20(11)：53-59.

[10] 杨小彬，韩超，胡慢谷，等.考虑时空特征的瓦斯涌出量预测模型及其可解释性分析[J].安全与环境学报，2025，25(10)：3794-3802.
Yang Xiaobin，Han Chao，Hu Mangu，et al. A spatiotemporal feature-based prediction model for gas emission：an interpretability analysis.[J]. Journal of Safety and Environment，2025，25(10)：3794-3802.

[11] 李鹏，辛诗雨，闫凡壮，等.基于深度学习的煤巷掘进工作面瓦斯涌出量预测研究[J].煤炭工程，2024，56(12)：115-124.
Li Peng，Xin Shiyu，Yan Fanzhuang，et al. Prediction of gas emission from coal roadway heading face based on deep learning theory[J]. Coal Engineering，2024，56(12)：115-124.

[12] 盛武，樊斌斌.EMD-GRU 组合模型在煤矿瓦斯体积分数预测中的应用[J].安全与环境学报，2025，25(4)：1339-1348.
Sheng Wu，Fan Binbin. Application of the EMD-GRU combined model for predicting gas volume fraction in coal mines.[J]. Journal of Safety and Environment，2025，25(4)：1339-1348.

[13] 刘超,雷晨,李树刚,等.基于 CNN-GRU 的瓦斯浓度预测模型及应用[J].中国安全生产科学技术,2022,18(9):62-68.
Liu Chao, Lei Chen, Li Shugang, et al. Prediction model of gas concentration based on CNN-GRU and its application. [J]. Journal of Safety Science and Technolog,2022,18(9):62-68.

[14] 李海宏.基于 KPCA-PSO-LSTM 模型的瓦斯涌出量时序预测[J].能源与环保,2025,47(9):41-46+52.
Li Haihong. Time series prediction of gas emission based on KPCA-PSO-LSTM model. [J]. China Energy and Environmental Protection,2025,47(9):41-46+52.

[15] Lin Haifei, Li Wenjing, Li Shugang, et al. Coal mine gas emission prediction based on multifactor time series method [J]. Reliability Engineering and System Safety, 2024, 252:110443.

[16] 王宝会,高瞻,徐林,等.基于深度学习的矿井瓦斯浓度预测算法研究与实现[J].计算机科学,2025,52(增刊 1):626-632.
Wang Baohui, Gao Zhan, Xu Lin, et al. Research and implementation of mine gas concentration prediction algorithm based on deep learning[J]. Computer Science, 2025,52(Suppl.1):626-632.

[17] 祁云,白晨浩,代连朋,等.改进双向长短期记忆神经网络的瓦斯涌出量预测[J].安全与环境学报,2024,24(12):4630-4637.
Qi Yun, Bai Chenhao, Dai Lianpeng, et al. Enhanced Bi-directional Long ShortTerm Memory neural network for gas emission forecasting[J]. Journal of Safety and Environment, 2024,24(12):4630-4637.

[18] Zhang Jingzhao, Cui Yuxin, Yan Zhenguo, et al. Time series prediction of gas emission in coal mining face based on optimized Variational mode decomposition and SSA-LSTM [J]. Sensors,2024,24(19):6454.

[19] 胡青松,郑硕,李世银,等.基于改进 TCN-TimeGAN 的矿井瓦斯浓度智能预测方法[J].煤炭科学技术,2024,52(增刊 2):321-330.
Hu Qingsong, Zheng Shuo, Li Shiyin, et al. An intelligent prediction method of gas concentration in coal mines based on improved TCN-TimeGAN. [J]. Coal Science and Technology,2024,52(Suppl.2):321-330.

[20] Liu Ziming, Wang Yixuan, Vaidya Sachin, et al. Kan: Kolmogorov-arnold networks[J]. arXiv,2024.

[21] Liu Yong, Hu Tengge, Zhang Haoran, et al. iTransformer: inverted transformers are effective for time series forecasting [J]. arXiv,2023.

[22] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. arXiv,2017.

[23] Veličković P, Cucurull G, Casanova A, et al. Graph attention networks[J]. arXiv,2017.