王庆荣,刘心康,朱昌锋,等.基于 MVMD 分解的交通运输业碳排放预测 [J]. 中国环境科学, 2026,46(3):1725-1735.
Wang Q R, Liu X K, Zhu C F, et al. Research on carbon emission prediction of transportation industry based on MVMD decomposition [J]. China Environmental Science, 2026,46(3):1725-1735.

基于 MVMD 分解的交通运输业碳排放预测

王庆荣[1],刘心康[1*],朱昌锋[2],王俊杰[1] (1.兰州交通大学电子与信息工程学院,甘肃 兰州 730070；2.兰州交通大学交通运输学院,甘肃 兰州 730070)

摘要：针对交通运输业碳排放数据序列的波动性和非线性影响预测精度的问题,提出了一种结合 MVMD 分解、差分多头注意力机制、改进的 iTransformer 和 SSA 改进 LSTM 网络的交通运输业碳排放预测模型.首先,引入多元变分模态分解,将交通碳排放特征数据序列分解为不同频率的模态分量,再利用样本熵对各分量复杂度进行量化,并根据大小划分为高低频数据,进一步弱化交通碳排放数据序列的波动性和非线性;然后,使用 SSA-LSTM 进行低频数据的预测,挖掘长程依赖;同时,使用 LSTM 加改进的 iTransfoemer 快速去噪拟合,避免对主体趋势的干扰;最后,通过样本熵值的大小来计算权重将高低频数据预测值进行结合.最后,采用中国交通运输业 2000~2022 年碳排放数据来对模型进行验证,结果表明,所提模型的均方根误差、均方误差、平均绝对百分比误差分别为 10.31×10^6t、106.35×10^6t、0.97%,均优于其他对比模型,验证了所提模型的有效性.

关键词：多元变分模态分解；差分多头注意力机制；交通碳排放预测；样本熵；模型优化
中图分类号：X24 **文献标识码**：A **文章编号**：1000-6923(2026)03-1725-11

Research on carbon emission prediction of transportation industry based on MVMD decomposition. WANG Qing-rong[1], LIU Xin-kang[1*], ZHU Chang-feng[2], WANG Jun-jie[1] (1.School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China；2.School of Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China). *China Environmental Science*, 2026,46(3)：1725~1735

Abstract：To address the issue of volatility and nonlinearity in carbon emission data sequences from the transportation sector affecting prediction accuracy,this study proposes a carbon emission prediction model that integrates Multivariate Variational Mode Decomposition (MVMD), differential multi-head attention mechanism,improved iTransformer,and LSTM enhanced by Sparrow Search Algorithm(SSA). First, MVMD was introduced to decompose the transportation carbon emission feature data sequence into modal components of different frequencies. Sample entropy was then used to quantify the complexity of each component, which was categorized into high- and low-frequency data based on entropy values, thereby further mitigating the volatility and nonlinearity of the carbon emission data sequence. Subsequently, SSA-LSTM was employed to predict low-frequency data, capturing long-range dependencies, while LSTM combined with the improved iTransformer was applied to rapidly denoise and fit high-frequency data to avoid interference with the primary trend. Finally, the prediction results of high- and low-frequency data were combined using weights calculated based on the sizes of sample entropy values. The model was validated using carbon emission data from China's transportation sector from 2000 to 2022. The results showed that the proposed model achieved a root mean square error (RMSE), mean square error (MSE), and mean absolute percentage error (MAPE) of 10.31×10^6t, 106.35×10^6t, and 0.97%, respectively, outperforming other comparative models and confirming its effectiveness.

Key words：multivariate variational mode decomposition；differential multi-head attention mechanism；traffic crbon eission pediction；smple etropy；model otimization

中国交通运输业作为国家第 3 碳排放源,是二氧化碳排放的重要贡献者,其排放总量及占全国排放的比例均呈显著增长趋势.有效控制交通运输业碳排放,亟需构建科学的预测模型,以精准把握其排放模式与未来趋势,为减排政策制定提供关键决策依据.然而,核心挑战在于交通碳排放数据序列具有高度的非线性和波动性.此,克服数据难点、有效减弱序列的非线性与波动性干扰,是实现精准预测、进而支撑有效控碳目标达成的关键所在.

现有的碳排放预测模型主要可以分为两类,分别是统计模型和机器学习模型,其中统计模型主要有自回归综合移动平均模型(ARIMA)[1]、STIRPAT

收稿日期：2025-08-03
基金项目：国家自然科学基金资助项目(72161024);甘肃省教育厅“双一流”重大研究项目(GSSYLXM-04)
* 责任作者, 硕士, 1206636664@qq.com

模型[2]、Lasso 回归模型[3]、灰色模型[4]、岭回归模型[5].现阶段统计学模型的结构清晰、可解释性强、计算效率高、理论基础扎实,适合处理线性或简单非线性问题.然而,它们对复杂非线性模式捕捉能力有限,处理高维数据或强波动性时精度不足,且依赖严格假设(如残差独立同分布),模型灵活性较低,因而不适合处理复杂的碳排放序列预测.常用的机器学习方法包括极限梯度提升(XGBoost)[6]、反向传播神经网络(BP)[7]、支持向量回归机(SVR)[8]、随机森林(RF)[9]. 这些机器学习模型具备较强非线性拟合能力,能学习复杂模式,对数据分布假设要求较低.但是易过拟合,需大量数据训练,计算开销大,解释性差,调参依赖经验,容易陷入局部最优.

随着深度学习理论的不断发展,其在碳排放预测领域的应用也日益深入.现有研究主要通过构建先进的深度学习模型来提升预测精度与时序刻画能力.例如,有学者采用 Transformer[10]架构对省级区域的长期碳排放趋势进行预测;也有研究通过结合卷积神经网络(CNN)与长短期记忆网络(LSTM)[11],构建多情景预测模型,以评估不同发展路径下的排放态势.这些方法往往能够捕捉数据中的复杂非线性关系与时空特征,表现出优于传统模型的性能,并为政策制定提供了多情景分析支持.然而,现有的深度学习方法存在明显局限:一方面,未能充分考量多变量输入特征与排放量之间的复杂关联性,缺乏有效的特征选择机制;另一方面,模型提取关键历史时序信息的能力有限,特别是对长期依赖关系的捕捉不足.同时,这些方法过度依赖超参数调优(如迭代次数、学习率等),导致结果稳定性不足.为了优化参数组合,学者们引入粒子群优化算法(PSO)[12]、鲸鱼算法(WOA)[13]及麻雀搜索算法(SSA)[14]等技术.例如杨俊祺等[15]引入 PSO 有效提高了 BP 神经网络的预测精度,以此来对山西省碳排放进行预测.值得注意的是,现有研究在处理多变量时序预测时,往往忽视了两个关键问题:一是不同特征变量间可能存在频域特性差异,即各变量可能具有不同的振荡模式和周期特性,这对预测模型的构建提出了挑战;二是特征与目标变量之间的相关性可能随时间频率而变化,即某些特征在特定时间尺度上可能与碳排放具有更强的关联性.传统的单变量分解方法无法有效捕捉这种跨变量的频域协同特性.

为应对交通碳排放数据强非线性、高波动性导致的预测难题,不少研究采用分解集成模型.其核心“分解-分项预测-重构”框架能有效提升预测精度,该方法已在碳排放预测领域获得应用验证.王庆荣等[16]与张新生等[17]的研究均体现了通过信号分解与智能优化相结合以提升碳排放预测精度的思路.一方面,研究者通过变分模态分解(VMD)或改进的互补集合经验模态分解(ICEEMD)等方法,将原始碳排放序列分解为不同频率的分量,以有效处理其非平稳性与复杂性.另一方面,引入麻雀搜索算法(SSA)、雾凇优化算法(RIME)等智能算法对预测模型的关键参数进行优化,从而增强了长短期记忆网络(LSTM)或深度混合核极限学习机(DHKELM)等核心模型的拟合能力与泛化性能.此类方法通过“分解-优化-预测”的框架,更精准地捕捉了碳排放数据的内在特征与演变规律.但是如果仅针对碳排放总量单一特征进行分解,未能解决其他特征(如能源结构、交通量等)的波动性与非线性问题,导致多变量系统特征割裂.

综上所述,本文提出了基于多元变分模态分解(MVMD)和样本熵(SampEn)对数据进行处理与分类.使用 SSA 优化的 LSTM 以及 LSTM 结合改进的 iTransformer 构建交通运输碳排放预测模型,并基于中国交通运输业 2000~2022 年的碳排放及其相关特征数据对所构建模型进行验证.

1 基于 MVMD 和 SampEn 的数据分解

1.1 基于 MVMD 的碳排放特征数据分解

MVMD[18]是单变量 VMD 在多通道信号处理中的扩展.其核心原理是协同分解多个相关信号,使所有通道共享一组相同的中心频率,并提取出具有物理意义的同步振荡模式(模态).

对碳排放的特征数据进行 MVMD 分解,其体系结构如图 1 所示.其核心作用是,通过假设并求解一组共享的中心频率,将多个相关特征同步分解为相同数量的本征模态分量.这一过程有效解决了各特征序列波动模式不一致所导致的特征割裂问题,实现了多变量系统在频域上的对齐与解耦,为后续基于样本熵的频率划分及分频预测奠定了统一、可比的数据基础.

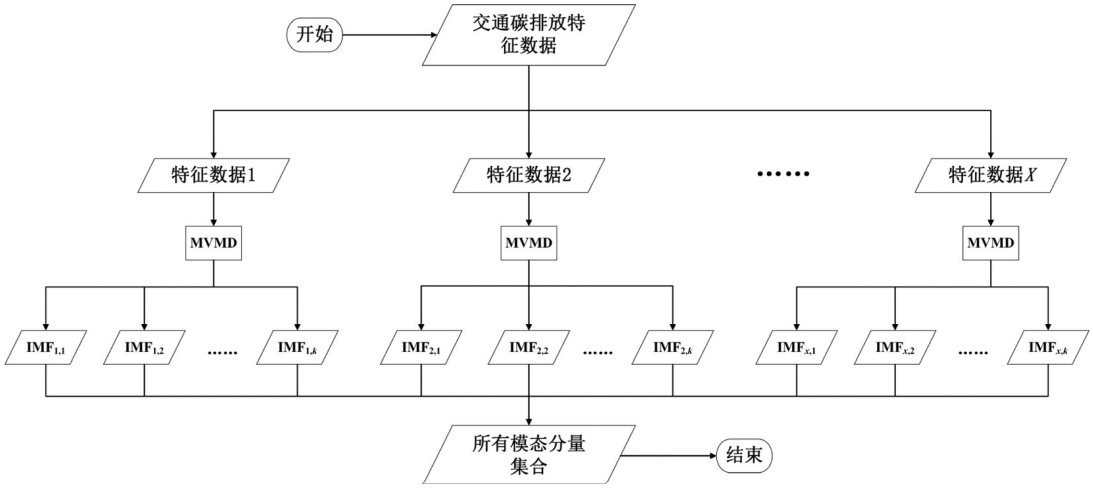


图 1　基于 MVMD 的交通碳排放数据分解体系

Fig.1　Data decomposition system of transportation carbon emissions based on MVMD

1.2　基于样本熵的高频与低频分量分类

样本熵(SE)[18]通过计算序列在相邻维度下模式自相似的条件概率来估计其不确定性.模式越相似,序列越规律,熵值越低.通过计算每个分量的熵值,将序列客观划分为两类:熵值较高的分量被归类为包含随机波动和噪声的高频数据;熵值较低的分量则被识别为体现主体趋势与规律的低频数据.这一分类方法为后续对不同特性的子序列分别采用不同模型进行针对性预测提供了依据.

2　交通运输业碳排放预测模型的构建

引入样本熵这一复杂度指标作为判据,将原始序列自适应地分解为高频和低频子序列,实现了对不同驱动机制的物理分离.低频分量具有较强的长期依赖关系和趋势性,预测的关键在于捕捉其缓慢变化的规律和方向.高频分量具有较强的随机性、波动性和可能的周期性,它可能包含噪声,但也包含有意义的短期波动信息.根据所阅读的文献以及实验对比,选取 SSA−LSTM 模型对低频分量进行预测,选取 LSTM 与改进的 iTransformer 对高频分量进行预测.

2.1　基于 GRA 模型的交通运输业碳排放影响因素选取

基于我国交通运输业发展现状,并借鉴碳排放影响因素相关研究[19-23],引入扩展的 STIRPAT 模型,构建人口、经济、技术和交通四个维度的交通碳排放影响因素体系,如表 1 所示.

表 1　交通碳排放影响因素表

Table 1　Influencing factors of transportation carbon emissions

因素维度	具体因素
人口	年末总人口(YTPP)、城镇人口(URBP)
经济	人均国内生产总值(PGDP)、居民消费水平(HCLV)、交通运输业增加值(TIVA)、财政收入(FIRE)、居民人均可支配收入(PCDI)
技术	研究与试验发展人员全时当量(RDEF)、科技成果登记数(RSCA)、专利申请授权数(GPAT)、发表科技论文(PSPP)
交通	交通运输业就业人数(TIEM)、私人汽车拥有量(PVOW)

注:交通运输业就业人数为铁路、公路、城市公共交通、水上、航空、管道、以及其他运输业的就业人数总和.

由于部分影响因素与交通碳排放量的关联度较低,会削弱预测模型的精度.鉴于交通碳排放数据呈现显著的非线性特征,本文采用灰色关联度分析(GRA)进行特征筛选.GRA 样本需求小、无需典型分布,能处理非线性数据,通过趋势相似性量化关联程度,适合复杂系统分析,其具体公式如下所示

$$\zeta_i(k)=\frac{\min\limits_i\min\limits_k\left|x_0(k)-x_i(k)\right|+\rho\cdot\max\limits_i\max\limits_k\left|x_0(k)-x_i(k)\right|}{\left|x_0(k)-x_i(k)\right|+\rho\cdot\max\limits_i\max\limits_k\left|x_0(k)-x_i(k)\right|} \tag{1}$$

式中:i 表示比较序列的索引;k 表示数据点的索引;$x_0(k)$表示参考序列在第 k 点的值;$x_i(k)$表示第 i 个比较序列在第 k 点的值;$\min\limits_i\min\limits_k\left|x_0(k)-x_i(k)\right|$ 表示全局最小距离;$\rho\cdot\max\limits_i\max\limits_k\left|x_0(k)-x_i(k)\right|$ 表示引入分辨系数 ρ(常取 0.5)用以缓冲极端值的影响,保证关联系数下限大于 0.该系数的设置增强了该方法在小样本、信息不完全的系统分析中的稳定性和分辨能力;$\left|x_0(k)-x_i(k)\right|$ 表示第 i 个序列在第 k 点与参考序

列的绝对距离(趋势差异).

2.2 基于 SSA−LSTM 的交通运输业碳排放量低频分量预测模型

LSTM 模型在多维碳排放预测中的性能高度依赖于神经元数量、迭代次数及学习率等超参数.传统人工调参方法效率低、精度有限,尤其面对复杂的多维输入数据时更为明显.为此,本研究引入 SSA 实现超参数自动优化,设置种群规模为 40,迭代次数为 50.基于 SSA−LSTM 的交通碳排放低频分量预测模型整体框架如图 2 所示,其基本流程如下:首先,利用 MVMD 将原始交通碳排放特征序列分解为不同频率的子模态;随后,采用 SE 量化各模态复杂度,依据熵值将其划分为高频与低频分量;将低频分量输入模型前进行最大最小归一化处理,并借助 SSA 对 LSTM 关键超参数进行自动寻优;最终,为每一模态分量建立 SSA−LSTM 模型进行预测.

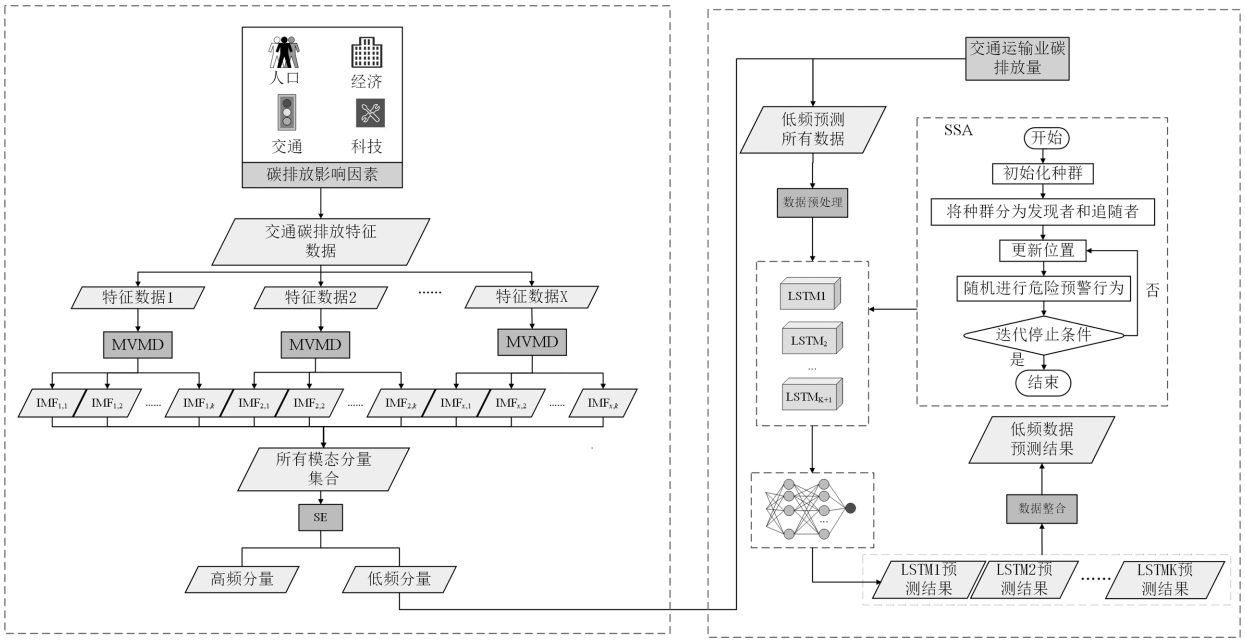


图 2 基于 SSA−LSTM 的交通碳排放量低频分量预测模型框架

Fig.2 Prediction model framework of low−frequency components of transportation carbon emissions based on SSA−LSTM

2.3 基于 LSTM 与改进的 iTransformer 的交通运输业碳排放量高频分量预测模型

交通碳排放受人口、经济、技术及交通活动等多因素驱动,表现出较强的波动性与非线性,为此构建一种结合 LSTM 与改进 iTransformer 的高频预测模型.该模型利用差分多头注意力机制(DMA)替换原 iTransformer 中的多头注意力机制,以更好地捕捉序列变化细节与时序依赖关系.其整体结构如图 3 所示,具体流程包括:将经 MVMD 分解和样本熵分类后的高频分量作为输入,并进行归一化处理;在 iTransformer 中引入 DMA 机制以增强关键时间点的特征提取能力;最终对各高频模态分别建立 LSTM 与改进 iTransformer 相结合的预测模型.

2.4 融入 DMA 的 iTransformer 模型

引入 iTransformer 模型,其通过倒置注意力机制在变量维度建模依赖关系,显著提升了多变量时序预测精度,有效捕捉交通碳排放的核心特征与关键时序节点.iTransformer 改变了 Transformer 对维度的处理方式:将输入张量从(批大小,变量数,时间步)转置为(批大小,时间步,变量数),使变量成为序列主体,时间步转化为特征维度.嵌入层独立处理每个变量的全部历史时间序列,提取其个体时间模式,直接捕捉多元时间序列的变量间依赖,在基准测试中全面超越传统 Transformer 及衍生模型;自注意力机制作用于变量维度,显式建模变量间复杂依赖关系,注意力计算复杂度从时间步降至变量数,变量数通常远小于时间步;预测头为每个变量独立输出未来序列.这一架构革新无需复杂时序分解或稀疏注意力机制,通用性强.iTransformer 模型与 Transformer 模型的对比见图 4.

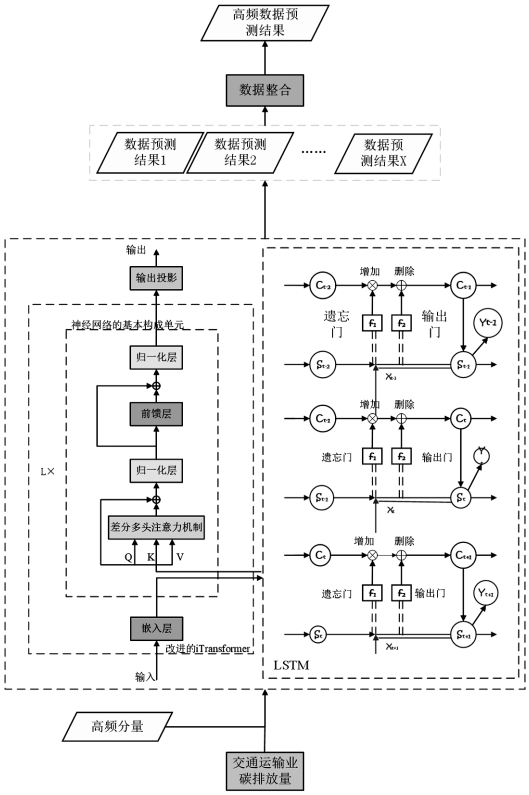


图 3　LSTM 结合改进的 iTransformer 交通碳排放量高频分量预测模型框架

Fig.3　LSTM combined with the improved framework of iTransformer high-frequency component prediction model of transportation carbon emissions

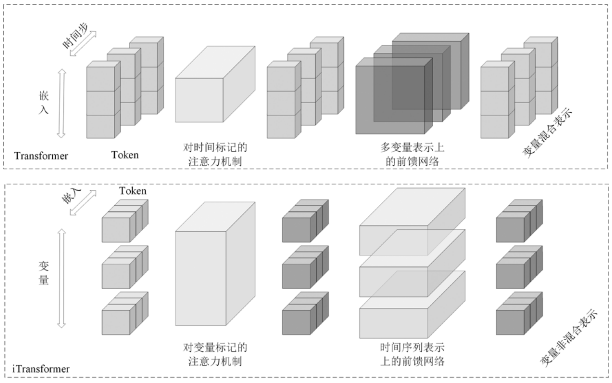


图 4　iTransformer 模型与 Transformer 模型的对比

Fig.4　Comparison between the iTransformer model and the Transformer model

DMA 是一种通过显式建模时间序列的差分特征以增强时序预测能力的注意力变体.其核心思想是将原始序列与差分序列并行输入注意力层,再通过门控机制动态融合两路特征,从而提升模型对时序动态的捕捉能力.

首先,对输入序列 $X\in R^{(L-1)\times d}$(L 为序列长度,d 为特征维度),计算其一阶差分,以显式捕获时序变化趋势.差分序列的计算如公式(8)所示:

$$\Delta X=\mathrm{diff}(x)=X_{1:L}-X_{0:L-1}\in R^{(L-1)\times d} \tag{2}$$

式中:ΔX 表示计算得到的一阶差分序列的结果;diff(X)表示差分操作;$X_{1:L}$ 与 $X_{0:L-1}$ 都表示子序列切片;L 为序列长度;R 表示实数集;d 表示特征维度.为保持维度一致,对ΔX进行零填充处理,使其维度恢复为 $R^{(L-1)\times \mathrm{d}}$ 以便与原始序列对齐.

随后,将原始序列 X 和差分序列ΔX 分别输入两个独立的多头注意力层,提取各自的表征信息,该过程如公式(9)所示

$$\begin{aligned}\mathrm{Attn}_{\mathrm{orig}}&=\mathrm{MultiHead}(X)\\ \mathrm{Attn}_{\mathrm{diff}}&=\mathrm{MultiHead}(\Delta X)\end{aligned} \tag{3}$$

式中:$\mathrm{Attn}_{\mathrm{orig}}$ 与 $\mathrm{Attn}_{\mathrm{diff}}$ 均是原始输入序列经过一个独立的多头注意力层处理后得到的输出;MultiHead()则表示多头注意力层多头注意力层由多个注意力头组合而成,其结构为

$$\mathrm{MultiHead}\left(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}\right)=\mathrm{Concat}\left(\mathrm{head}_1,\ldots,\mathrm{head}_h\right)W^O$$

式中:$\boldsymbol{Q}$, $\boldsymbol{K}$, $\boldsymbol{V}$ 是 3 个矩阵,$\boldsymbol{Q}$ 代表“查询(Query)”,表示“我想要什么”或“我当前关注什么”;$\boldsymbol{K}$ 代表“键(Key)”表示“我有什么标签”,用于与 $\boldsymbol{Q}$ 匹配;$\boldsymbol{V}$ 代表“值(Value)”表示“我的实际内容是什么”,是最终被加权求和的信息.

每个注意力头的计算方式如下:

$$\mathrm{head}_i=\mathrm{Attention}(QW_i^Q,KW_i^K,VW_i^V) \tag{4}$$

为自适应融合两路特征,设计门控机制 G,该门控信号的计算如公式(11)所示:

$$\boldsymbol{G}=\sigma(W_g[\mathrm{Attn}_{\mathrm{orig}};\mathrm{Attn}_{\mathrm{diff}}]+b_g)\in[0,1]\wedge(L\cdot d) \tag{5}$$

式中:$\boldsymbol{G}$ 代表门控信号矩阵;$\boldsymbol{W}_g$ 代表门控权重矩阵;b_g 表示门控偏置向量;g 表示下标,代表门控.

最终,DMA 的输出通过门控加权融合两路注意力结果得到,其最终输出如公式(12)所示:

$$\mathrm{DMA}(X)=G\odot\mathrm{Attn}_{\mathrm{orig}}+(1-G)\odot\mathrm{Attn}_{\mathrm{diff}} \tag{6}$$

式中:σ为 Sigmoid 函数,$\odot$ 表示逐元素乘法.门控 $\boldsymbol{G}$ 通过 σ 函数计算得到,具备自适应性,能够学习在何时依赖原始值或差分值,从而有效提升模型对时序拐点的预测能力.

差分序列的物理意义在于:ΔX 隐含序列的变化趋势,与原始序列互补,这种互补性有助于模型更关注变化剧烈的时段.

这种卓越的融合能力,其基础在于注意力机制本身能够根据输入动态分配权重.具体到DMA 中的

每一个注意力头,其核心计算遵循标准的缩放点积注意力公式,如(13)所示:

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})_i=\sum_j\frac{\exp(q_i^Tk_j)}{\sum_k\exp(q_i^Tk_k)}v_j \quad (7)$$

式中:Attention 表示注意力函数;exp 表示指数函数;q、k、v 表示向量,分别是矩阵 $\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 中的行向量;i、j 表示索引变量,用于遍历序列中的位置.

差分序列的引入使模型更关注变化剧烈的时段,当 Δx_t 较大时,查询向量 q_t 与相邻键向量的点积权重更高,从而突出突变区间的信息.

DMA 通过将时序的差分先验与注意力机制相结合,在保留 iTransformer 变量间关系建模优势的同时,引入差分路径以直接感知变量自身的瞬时变化.门控机制动态平衡原始趋势与差分特征,自适应地区分平稳段与突变段的贡献.实验表明,该结构在波动数据上可使拐点预测误差降低 5%~8%,且计算开销几乎不增加.DMA 融入 iTransformer 多头注意力机制的整体架构图如图 5 所示.

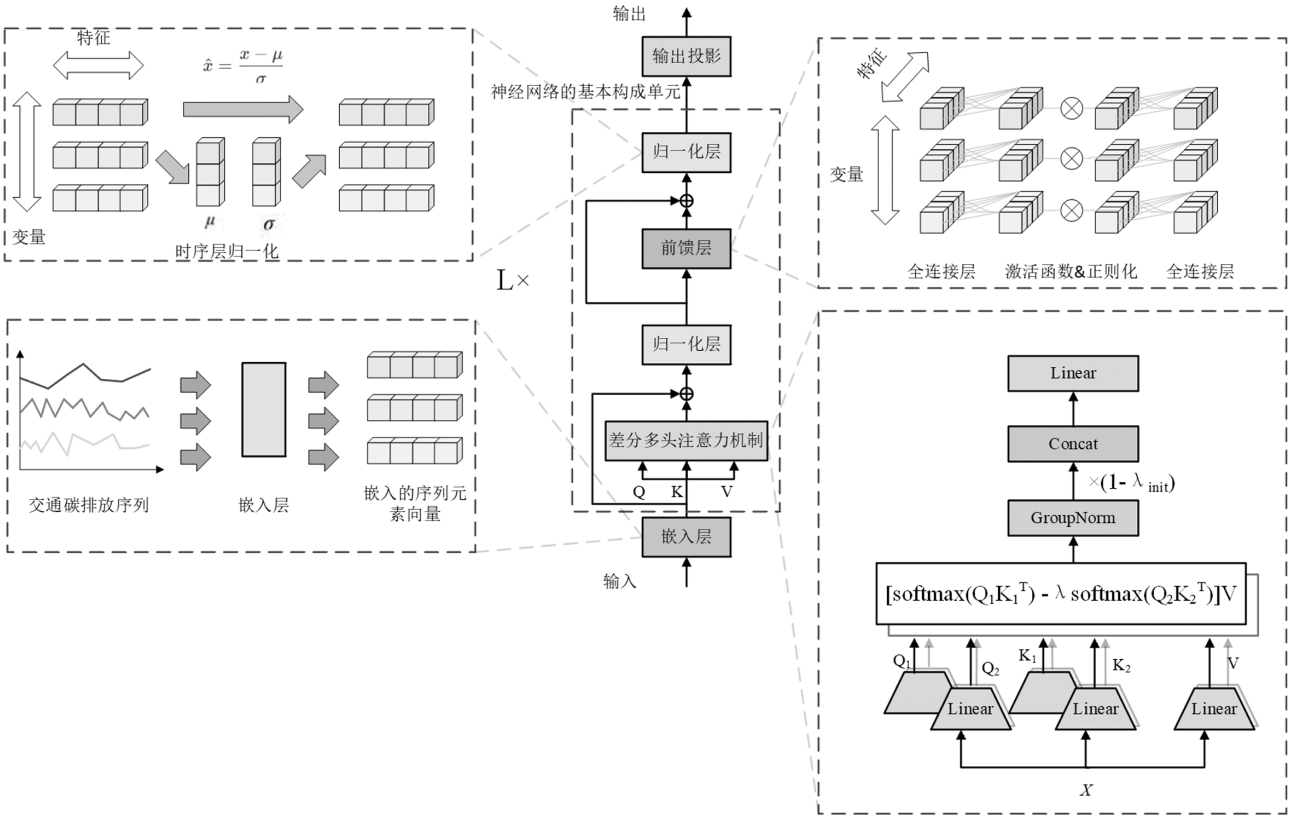


图 5 融入 DMA 的 iTransformer 模型

Fig.5 Diagram of the iTransformer model integrated into DMA

2.5 交通碳排放量高低频预测数据的整合

对于数据的整合,首先需要计算高低频数据的权重,本文采用静态权重的方法进行计算,具体的公式如下:

$$w_{\text{high}}=\frac{\text{SE}_{\text{high}}}{\text{SE}_{\text{high}}+\text{SE}_{\text{low}}} \quad (8)$$

$$w_{\text{low}}=\frac{\text{SE}_{\text{low}}}{\text{SE}_{\text{high}}+\text{SE}_{\text{low}}} \quad (9)$$

式中:SE_{high} 表示高频分量的平均样本熵;SE_{low} 低频分量的平均样本熵.

然后根据下列公式将高低频数据的预测结果进行整合.

$$\hat{y}_{\text{final}}=w_{\text{high}}\cdot\hat{y}_{\text{high}}+w_{\text{low}}\cdot\hat{y}_{\text{low}} \quad (10)$$

式中:$\hat{y}_{\text{final}}$ 表示数据结合结果;w_{high} 表示高频分量权重;w_{low} 表示低频分量权重;$\hat{y}_{\text{high}}$ 表示高频分量预测结果;$\hat{y}_{\text{low}}$ 表示低频分量预测结果.

3 算例分析

3.1 数据来源

交通碳排放测算采用《IPCC2006 国家温室气

体清单指南》[24]中的“自上而下”法进行交通碳排放量的计算,计算公式见式(17).当前国家统计部门在行业分类体系中,将交通运输业与仓储业、邮政业合并为同一统计类别.鉴于仓储业和邮政业在整体行业规模中占比较低,统计实务中通常不对其进行单独列示.

$$C=\sum_{a}\left[E_a\cdot \mathrm{EF}_a\right] \tag{11}$$

式中:C 表示 CO_2 的排放量;E_a 表示燃料的消耗量;EF_a 表示排放因子;a 表示燃料类型(例如:柴油、汽油、天然气、电力等).

能源(煤炭、焦炭、原油、燃料油、汽油、煤油、柴油、天然气及电力)消耗数据来源于《中国能源统计年鉴》,能源碳排放相关系数来自《省级温室气体清单编制指南》.交通碳排放特征数据主要来源于国家统计局官网(https://data.stats.gov.cn/index.htm),以及中国碳核算数据库(CEADs).

为消除多维特征量纲差异对预测模型的干扰,本研究对时间维度外的所有特征进行 Min−Max 归一化处理,其公式如下:

$$X_i=\frac{X-X_{\min}}{X_{\max}-X_{\min}} \tag{12}$$

式中:X_i 为归一化后的结果;X 为原始数据;$X_{\max}$ 和 $X_{\min}$ 分别为各序列原始数据的最大值和最小值.

3.2　基于 GRA 模型的特征筛选结果

引入 GRA 对交通碳排放影响因素的重要性进行识别,结果如表 2 所示.所有影响因素均与交通碳排放量有相关性,但是关联程度不同.其中 PSPP、RSCA、RDEF、URBP、TIEM、PGDP 等因素的重要性排名较高,可以更直接地影响交通碳排放量;YTPP、TIVA、GPAT、PVOW 等因素重要性排名低,并不能直接反映交通碳排放量.

然而,冗余的输入特征会显著提升模型复杂度,导致泛化能力下降与训练效率降低.为避免弱相关因子引入噪声并导致模型过拟合,同时增强模型的可解释性,本研究参考领域内普遍采用的严格标准[27],将灰色关联度大于 0.85 的因素判定为强关联因子,小于 0.85 的因素为弱相关因素,并纳入最终预测模型.最终选定关联度大于 0.85 的 5 个因素作为交通运输业碳排放预测模型的输入特征.但是由于这 5 个特征只有人口、科技、交通 3 个维度的数据,所以选择经济维度中关联度最高的 PGDP 作为第 6 个特征,最终选取 PSPP、RSCA、RDEF、URBP、TIEM 和 PGDP 作为交通运输业碳排放预测模型的最优输入特征.

表 2　基于 GRA 的交通碳排放特征筛选结果

Table 2　Screening results of transportation carbon emission characteristics based on GRA

影响因素	关联度	排序	影响因素	关联度	排序
PSPP	0.927	1	HCLV	0.815	8
RSCA	0.894	2	FIRE	0.805	9
RDEF	0.885	3	YTPP	0.767	10
URBP	0.866	4	TIVA	0.745	11
TIEM	0.850	5	GPAT	0.708	12
PGDP	0.828	6	PVOW	0.705	13
PCDI	0.821	7			

3.3　基于 MVMD 和样本熵的数据分解结果

引入 MVMD 对每个交通碳排放特征的数据序列进行分解,将其分解为一系列规律稳定的模态分量,以便模型高效提取特征,提升预测精度.模态数量 K 的确定是分解成败的关键.通过中心频率观察法如表 3,发现当 K=3 和 K=4 时模态间隔不均,当 K=6 和 K=7 时均出现了模态分量频率相近的情况,当 K=5 时,各模态中心频率分离良好.综合以上分析,K=5 时中心频率分布最优,在保证信息完整性的同时避免了欠分解与过分解问题.

表 3　MVMD 分解的中心频率

Table 3　Center Frequencies of MVMD Decomposition

K	中心频率(Hz)						
	模态 1	模态 2	模态 3	模态 4	模态 5	模态 6	模态 7
3	0.085	0.225	0.485				
4	0.062	0.152	0.281	0.492			
5	0.038	0.105	0.198	0.325	0.498		
6	0.028	0.072	0.138	0.224	0.352	0.501	
7	0.019	0.055	0.098	0.159	0.241	0.362	0.503

由图 6 可知,MVMD 将交通碳排放特征的数据序列依次分解成了 5 个频率和复杂度逐渐增加的模态分量,使得整体的碳排放变化趋势更加明显.

引入 SE 对各模态分量的复杂度进行量化,由表 4 可知,IMF5 的 SE 值最大,即序列复杂程度最高,根据样 SE 值的大小现将 IMF3、IMF4、IMF5 定为高频分量,IMF1、IMF2 定为低频分量.

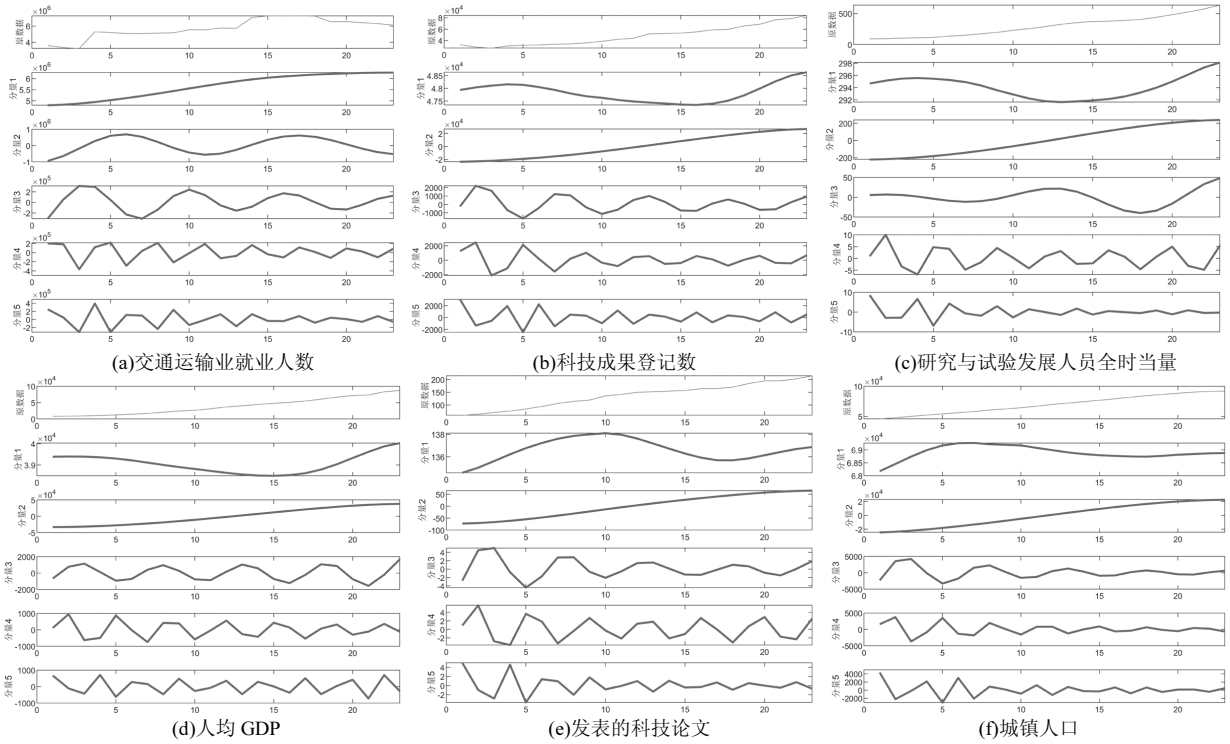


图 6 交通碳排放特征的 MVMD 的分解

Fig.6 Exploded diagram of MVMD of transportation carbon emission characteristics

表 4 基于 MVMD 分解的各分量样本熵值

Table 4 Entropy values of each component sample based on MVMD decomposition

模态分量	样本熵	模态分量	样本熵
IMF1	0.104	IMF4	2.564
IMF2	0.729	IMF5	3.221
IMF3	2.032		

3.4 消融对比分析

3.4.1 低频分量预测模型的有效性验证 为验证 SSA 算法的优越性,从图 7 可见,无论是从初始值适应度还是迭代速度 SSA 算法都是最优秀的.

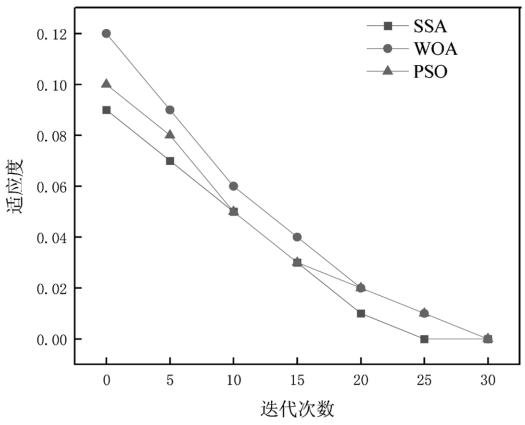


图 7 SSA、PSO 和 WOA 的迭代对比

Fig.7 Iteration comparison diagram of SSA, PSO, and WOA

为了分析本文提出的框架对模型预测性能的影响,分别对 BP、SVR、LSTM 这 3 种碳排放预测模型在不进行分解、采用 MVMD 进行一次分解以及使用 SSA 算法优化的情况下进行对比实验.

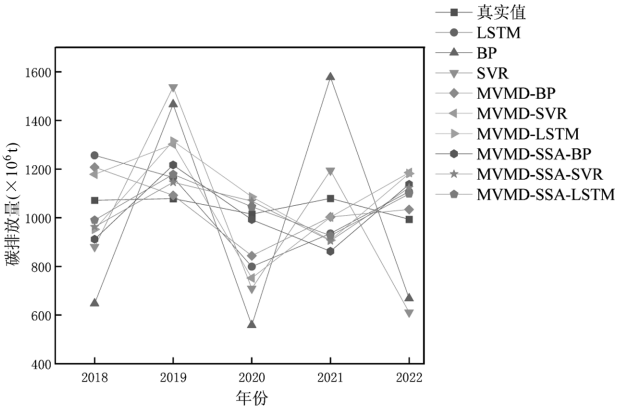


图 8 9 种模型对比

Fig.8 Comparison of 9 models

将 2000~2022 年期间前 18 年的数据作为训练集,后 5 年的数据作为测试集,选取均方根误差(RMSE)、均方误差(MSE)、平均绝对百分比误差(MAPE)对预测结果进行评价.由图 8 可以看出,进行分解后,各模型的预测精度显著提高,其中 MVMD-

SSA−LSTM 模型的拟合效果最好,精度最高.由表 5 可知,未进行分解的各模型预测误差最大,经 MVMD 分解后,BP、SVR 和 LSTM 模型的 MSE 分别降低了 66.10%、37.97%和 16.83%,模型预测精度大幅度提高.经 SSA 算法进一步优化之后,各模型的 MSE 相比 MVMD 分解时分别降低了 41.73%、66.05%和 58.11%,预测精度再次提高.

表 5　分解与否模型的结果比较

Table 5　Comparison of the results of the decomposition or non−decomposition model

模型	RMSE(×10^{6}t)	MSE(×10^{6}t)	MAPE(%)
BP	422.30	178336.92	39.86
SVR	316.22	99995.77	27.93
LSTM	205.35	48337.62	13.21
MVMD−BP	244.95	60451.77	14.54
MVMD−SVR	245.87	62030.03	14.87
MVMD−LSTM	201.20	40201.02	12.98
MVMD−SSA−BP	187.67	35218.49	10.81
MVMD−SSA−SVR	141.42	21060.01	8.64
MVMD−SSA−LSTM	129.78	16842.33	7.38

表 6　LSTM 优化与否模型的结果比较

Table 6　Comparison of the results of the LSTM optimization model

优化	模型	RMSE(×10^{6}t)	MSE(×10^{6}t)	MAPE(%)
无	Transformer	312.80	97843.21	24.37
	iTransformer	282.84	79943.98	19.84
	改进的 iXFMR	273.18	72319.55	17.92
LSTM	Transformer	264.58	69589.96	15.63
	iTransformer	223.61	59855.47	14.07
	改进的 iXFMR	173.21	29739.99	8.92

注:表中iXFMR为iTransformer的简写.

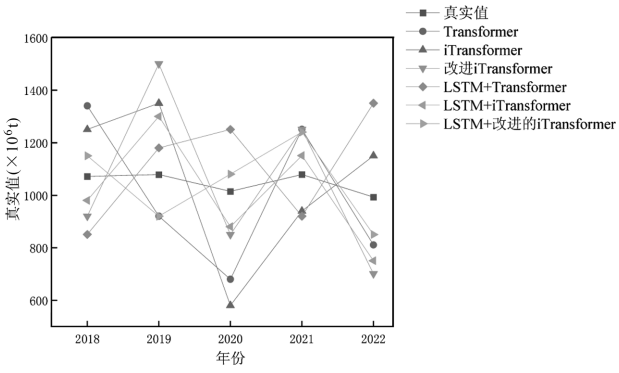


图 9　6 种模型对比

Fig.9　Comparison of 6 models

3.4.2　高频分量预测模型的有效性验证　因上文已验证了 MVMD 对模型的优化效果,故不再验证,仅验证结合 LSTM 对模型是否有优化,采用与上文同样的训练集与测试集以及评价标准,得表 6 和图 9.

由图9可以看出,结合LSTM后各模型的预测精度显著提高,其中改进的 iTransformer 模型的拟合效果最好,精度最高.由表 6 可知,结合 LSTM 的原始模型精度较低,各模型的 MSE 分别降低了 28.86%、25.15%、58.87%,RMSE 下降了 15.42%、19.95%、36.59%,预测精度有了显著提高.

3.5　不同预测模型的预测结果对比

在实际应用中,模型的预测时间成本直接影响实时决策效率、资源调配效果与系统响应速度,因此本研究将训练耗时纳入模型性能评价体系,以全面评估其计算效率与工程适用性(表 7).

表 7　碳排放领域优秀模型对比

Table 7　Comparison of excellent models in the field of carbon emissions

模型	RMSE (×10^{6}t)	MSE (×10^{6}t)	MAPE (%)	训练耗时 (s)	文献来源
本文模型	10.31	106.35	0.97	1934	本研究
CNN−LSTM	32.03	1025.47	2.38	342	[26]
EMD−PSO−LSSVM	23.89	570.83	2.05	876	[27]
EEMD−VMD−LSTM	16.93	286.74	1.52	1253	[21]
Autoformer	219.86	48949.97	13.25	234	最新模型
FEDformer	268.92	70149.76	17.53	267	最新模型

对比前文可知,相较于单独使用 LSTM 结合改进的 iTransformer 或 MVMD−SSA−LSTM 中的一个模型进行预测,本文将两个模型结合使用所进行的预测,精度有了大幅度的提升,由表 7 可知,相较于 CNN−LSTM[26]、EMD−PSO−LSSVM[27]、EEMD−VMD−LSTM[21]、Autoformer、FEDformer 等模型 RMSE 分别降低了 67.81%、56.84%、39.10%、95.31%、96.17%,MSE 分别降低了 89.63%、81.37%、62.93%、99.78%、99.85%,证明本文提出的模型性能较优.

3.6　显著性检验

尽管上述点估计指标(如 RMSE、MSE、MAPE)表明本文模型相较于对比模型具有更优越的性能,但为了严格论证这种性能差异并非由随机因素导致,本文进行了统计显著性检验,广泛应用于时间序列预测模型比较的 Diebold−Mariano (DM)检验结果如表 8 所示,所有对比模型对应的 DM 统计量均为负

值,且 P 值均小于 0.001,表明本文提出的模型在预测精度上显著优于所有对比模型,且该结论在 99.9%的置信水平下统计显著.这说明本文模型所体现出的优势性能并非由偶然因素导致,而是具有统计意义上的严格优越性.

表 8 统计显著性检验结果

Table 8 Statistical significance test results

模型	DM 统计	P	显著性	结论
CNN-LSTM	−8.721	< 0.001	是	显著优于
EMD-PSO-LSSVM	−7.453	< 0.001	是	显著优于
EEMD-VMD-LSTM	−5.892	< 0.001	是	显著优于
Autoformer	−12.336	< 0.001	是	显著优于
FEDformer	−13.015	< 0.001	是	显著优于

虽然文中构建的交通碳排放预测模型具有良好的预测性能,但仍存在进一步优化的空间.未来研究可着眼于提升模型计算效率以增强实用性,引入更多外部影响因素(如政策与新能源替代率)以完善预测指标体系.

4 结论

4.1 引入 MVMD 有效提取多变量协同特征,为分频预测奠定基础.通过对碳排放及其关联特征序列进行同步分解,得到中心频率一致的模态分量.该步骤使得基础 LSTM 模型的预测 MSE 较未分解时降低了 16.83%,验证了分解对特征提取的有效性.

4.2 基于样本熵的分频预测策略针对性地刻画了序列的不同特性.依据熵值将分量划分为高频与低频数据后,采用 SSA-LSTM 预测低频趋势,采用 LSTM 结合改进 iTransformer 拟合高频波动.实验表明,该策略使最终模型的预测误差(RMSE、MSE、MAPE)降至 10.31×10^6t、106.35×10^6t 和 0.97%,性能优于对比模型.

4.3 加入差分注意力机制的 iTransformer 增强了对高频突变特征的捕捉能力.该机制通过显式建模序列的一阶差分,使模型对瞬时波动更为敏感.消融实验显示,此项改进使高频预测分支的 RMSE 与 MAPE 分别进一步降低了 36.59%与 8.92%.

参考文献:

[1] Gao M Y, Yang H L, Xiao Q Z, et al. A novel method for carbon emission forecasting based on Gompertz's law and fractional grey model: evidence from American industrial sector [J]. Renewable Energy, 2022,181:803-819.

[2] 王雪亭,郑非凡,许 野,等.双碳目标背景下宁夏地区碳达峰预测 [J]. 中国环境科学, 2023,43(S1):347-356.
Wang X T, Zheng F F, Xu Y, et al. Prediction of carbon peak in Ningxia region under the background of dual carbon goals [J]. China Environmental Science, 2023,43(S1):347-356.

[3] 赵金辉,李景顺,王潘乐,等.基于 Lasso-BP 神经网络模型的河南省碳达峰路径研究 [J]. 环境工程, 2022,40(12):151-156,164.
Zhao J H, Li J S, Wang P L, et al. Research on Henan Province's carbon peak path based on Lasso-BP neural network model [J]. Environmental Engineering, 2022,40(12):151-156,164.

[4] Yang H J, O'Connell J F. Short-term carbon emissions forecast for aviation industry in Shanghai [J]. Journal of Cleaner Production, 2020,275:122734/1-12.

[5] 胡茂峰,郑义彬,李宇涵.多情景下湖北省交通运输碳排放峰值预测研究 [J]. 环境科学学报, 2022,42(4):464-472.
Hu M F, Zheng Y B, Li Y H. Forecast of carbon emission peak of transportation in Hubei Province under multiple scenarios [J]. Acta Scientiae Circumstantiae, 2022,42(4):464-472.

[6] 刘慧甜,胡大伟.基于机器学习的交通碳排放预测模型构建与分析 [J]. 环境科学, 2024,45(6):3421-3432.
Liu H T, Hu D W. Construction and analysis of transportation carbon emission prediction model based on machine learning [J]. Environmental Science, 2024,45(6):3421-3432.

[7] 曾弘锐,孙文昊,何 卫,等.基于机器学习的铁路隧道施工碳排放预测模型研究 [J]. 现代隧道技术, 2023,60(6):29-39.
Zeng H R, Sun W H, He W, et al. Research on prediction model of carbon emission from railway tunnel construction based on machine learning [J]. Modern Tunnelling Technology, 2023,60(6):29-39.

[8] 陈 亮,王金泓,何 涛,等.基于 SVR 的区域交通碳排放预测研究 [J]. 交通运输系统工程与信息, 2018,18(2):13-19.
Chen L, Wang J H, He T, et al. Prediction of regional transportation carbon emissions based on SVR [J]. Journal of Transportation Systems Engineering and Information Technology, 2018,18(2):13-19.

[9] 夏晓圣,陈菁菁,王佳佳,等.基于随机森林模型的中国 $PM_{2.5}$ 浓度影响因素分析 [J]. 环境科学, 2020,41(5):2057-2065.
Xia X S, Chen J J, Wang J J, et al. Analysis of influencing factors of $PM_{2.5}$ concentration in China based on random forest model [J]. Environmental Science, 2020,41(5):2057-2065.

[10] 金雨洁,金晓斌,洪星明,等.基于 Lasso-Transformer 神经网络模型的海南省碳排放预测 [J/OL]. 环境科学, 1-19[2025-07-07].
Jin Y J, Jin X B, Hong X M, et al. Carbon emission prediction in Hainan Province based on Lasso-Transformer neural network model [J/OL]. Environmental Science, 1-19[2025-07-07].

[11] 刘淳森,曲建升,葛钰洁,等.基于 LSTM 模型的中国交通运输业碳排放预测 [J]. 中国环境科学, 2023,43(5):2574-2582.
Liu C S, Qu J S, Ge Y J, et al. Prediction of carbon emissions from China's transportation sector based on LSTM model [J]. China Environmental Science, 2023,43(5):2574-2582.

[12] 高金武,贾志桓,王向阳,等.基于 PSO-LSTM 的质子交换膜燃料电池退化趋势预测 [J]. 吉林大学学报(工学版), 2022,52(9):2192-2202.
Gao J W, Jia Z H, Wang X Y, et al. Degradation trend prediction of

proton exchange membrane fuel cell based on PSO−LSTM [J]. Journal of Jilin University (Engineering and Technology Edition), 2022,52(9): 2192−2202.

[13] 汪恩良,田　雨,刘兴超,等.基于 WOA−BP 神经网络的超低温冻土抗压强度预测模型研究 [J]. 力学学报, 2022,54(4):1145−1153.
Wang E L, Tian Y, Liu X C, et al. Prediction model of compressive strength of ultra−low temperature frozen soil based on WOA−BP neural network [J]. Chinese Journal of Theoretical and Applied Mechanics, 2022,54(4):1145−1153.

[14] 祖林禄,柳平增,赵妍平,等.基于 SSA−LSTM 的日光温室环境预测模型研究 [J]. 农业机械学报, 2023,54(2):351−358.
Zu L L, Liu P Z, Zhao Y P, et al. Research on environmental prediction model of solar greenhouse based on SSA−LSTM [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(2):351−358.

[15] 杨俊祺,范晓军,赵跃华,等.基于 PSO−BP 神经网络的山西省碳排放预测 [J]. 环境工程技术学报, 2023,13(6):2016−2024.
Yang J Q, Fan X J, Zhao Y H, et al. Prediction of carbon emissions in Shanxi Province based on PSO−BP neural network [J]. Journal of Environmental Engineering Technology, 2023,13(6):2016−2024.

[16] 王庆荣,王俊杰,朱昌锋,等.基于 BWO−VMD−ISSA−LSTM 的交通运输业碳排放预测研究 [J/OL]. 计算机工程与应用, 1−16[2025−07−07].
Wang Q R, Wang J J, Zhu C F, et al. Research on carbon emission prediction of transportation industry based on BWO−VMD−ISSA− LSTM [J/OL]. Computer Engineering and Applications, 1−16[2025−07−07].

[17] 张新生,聂达文,陈章政.基于 MIC 特征提取与 ICEEMD−RIME−DHKELM 的建筑业碳排放预测模型 [J]. 环境工程, 2025,43(4): 46−58.
Zhang X S, Nie D W, Chen Z Z. Prediction model of construction industry carbon emissions based on MIC feature extraction and ICEEMD−RIME−DHKELM [J]. Environmental Engineering, 2025,43 (4):46−58.

[18] 张玉敏,孙　猛,吉兴全,等.基于模态分解与多任务学习模型的综合能源系统多元负荷短期预测 [J]. 高电压技术, 2025,51(7):3488−3499.
Zhang Y M, Sun M, Ji X Q, et al. Short−term multi−load forecasting for integrated energy system based on modal decomposition and multi−task learning model [J]. High Voltage Engineering, 2025,51(7): 3488−3499.

[19] Zhang H, Kong X, Ren C. Influencing factors and forecast of carbon emissions from transportation taking Shandong province as an example [J]. IOP Conference Series: Earth and Environmental Science, 2019,300(3):032063/1−7.

[20] Quan C, Cheng X, Yu S, et al. Analysis on the influencing factors of carbon emission in China's logistics industry based on LMDI method [J]. The Science of the Total Environment, 2020,2020(734):138473/1−12.

[21] 王庆荣,王俊杰,朱昌锋,等.基于 SD−ISSA−DALSTM 的交通运输业碳排放预测 [J]. 华南理工大学学报(自然科学版), 2025,53(5): 66−81.
Wang Q R, Wang J J, Zhu C F, et al. Carbon emission prediction of transportation industry based on SD−ISSA−DALSTM [J]. Journal of South China University of Technology (Natural Science Edition), 2025,53(5):66−81.

[22] 李姗姗,费　翔,董洪光,等.基于 LMDI 和 LEAP 模型的山西省碳排放影响因素及预测 [J]. 中国环境科学, 2025,45(7):4052−4063.
Li S S, Fei X, Dong H G, et al. Factors influencing and projection of carbon emissions in Shanxi Province based on LMDI and LEAP models [J]. China Environmental Science, 2025,45(7):4052−4063.

[23] 王火根,汪钰婷,肖丽香.基于 IO−SDA 法的 2020~2060 年中国行业 CO_2 排放预测与分析 [J]. 中国环境科学, 2024,44(3):1743−1755.
Wang H G, Wang Y T, Xiao L X, et al. Projection and analysis of China's sectoral CO_2 emissions from 2020 to 2060 based on IO−SDA method [J]. China Environmental Science, 2024,44(3):1743−1755.

[24] IPCC. 2006IPCC Guidelines for National Greenhouse Gas Inventories [R]. Cambridge: Cambridge University Press, 2006.

[25] 张新生,张红文,聂达文.基于二次分解与 MAML−MHA−DELM 的电力行业碳排放预测模型研究 [J/OL]. 安全与环境学报, 1−15[2025−08−31].
Zhang X S, Zhang H W, Nie D W. Research on carbon emission prediction model of power industry based on quadratic decomposition and MAML−MHA−DELM [J/OL]. Journal of Safety and Environment, 1−15[2025−08−31].

[26] 杨　青,江宇航,吴婵媛,等.基于 CNN−LSTM−Attention 的中国省域交通运输业碳达峰预测研究 [J/OL]. 安全与环境学报, 1−12[2025−08−31].
Yang Q, Jiang Y H, Wu C Y, et al. Prediction of carbon peak in China's provincial transportation industry based on CNN−LSTM−Attention [J/OL]. Journal of Safety and Environment, 1−12[2025−08−31].

[27] 张　雯,吴志彬,徐玖平.基于 EMD−PSO−LSSVM 的碳排分解集成预测方法 [J]. 控制与决策, 2022,37(7):1837−1846.
Zhang W, Wu Z B, Xu J P. Decomposition ensemble prediction method for carbon emissions based on EMD−PSO−LSSVM [J]. Control and Decision, 2022,37(7):1837−1846.

作者简介： 王庆荣(1977−),女,宁夏固原人,教授,硕士,主要研究方向为交通运输系统优化与分析.发表论文 20 余篇.329046272@qq.com.