文章编号：1003-1421(2025)12-0171-12
引用格式：曹瑞，帅斌，高鹏飞.基于CEEMDAN-Attention-BiLSTM-GWO的煤炭进口量预测混合模型研究[J].铁道运输与经济，2025，47(12)：171-182.
CAO Rui，SHUAI Bin，GAO Pengfei. Hybrid Model for Predicting Mongolian Coal Import Volume Based on CEEMDAN-Attention-BiLSTM-GWO[J]. Railway Transport and Economy，2025，47(12)：171-182.

基于CEEMDAN-Attention-BiLSTM-GWO的煤炭进口量预测混合模型研究

曹 瑞[1]，帅 斌[2,3,4,5]，高鹏飞[2,3,4,5]

(1.国能包神铁路集团有限责任公司，内蒙古 包头 014000；2.西南交通大学 交通运输与物流学院，四川 成都 611756；3.西南交通大学 系统科学与系统工程研究所，四川 成都 611756；4.西南交通大学 综合交通运输智能化国家地方联合工程实验室，四川 成都 611756；5.西南交通大学 综合交通大数据应用技术国家工程实验室，四川 成都 611756)

摘 要：煤炭是我国能源结构的重要组成部分，预测煤炭短期进口量有助于保障煤炭供应链稳定与高效运行，然而既有的先进混合模型中对于影响因素特征考虑不足，因此，本研究提出一种基于CEEMDAN-Attention-BiLSTM-GWO的煤炭进口量预测混合模型。该方法使用CEEMDAN，将具备复杂、随机和非线性特点的煤炭进口量时间序列信号进行分解得到9个表示不同频率信号的模态，使用格兰杰因果分析筛选与煤炭进口量呈现显著时间关联的影响因素，将分解后的模态和影响因素均作为特征输入至Attention-BiLSTM模型中，最后使用GWO寻找最优超参数。实证表明，该模型优于既有基线模型，并且具备模型结构上的合理性与完整性，取得了较好的预测效果。该预测方法可以为政策制定者、能源企业及相关产业提供重要的参考，帮助优化资源配置、减少预测误差。

关键词：煤炭进口量预测；自注意力机制；时间序列分解；混合模型；格兰杰因果分析
中图分类号：U294.1+3 **文献标识码：**A **DOI：**10.16668/j.cnki.issn.1003-1421.20241107001

Hybrid Model for Predicting Mongolian Coal Import Volume Based on CEEMDAN-Attention-BiLSTM-GWO

CAO Rui[1]，SHUAI Bin[2,3,4,5]，GAO Pengfei[2,3,4,5]

(1. China Energy Investment Baoshen Railway Group Co., Ltd., Baotou 014000, Inner Mongolia, China; 2. School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 611756, Sichuan, China; 3. Institute of System Science and Engineering, Southwest Jiaotong University, Chengdu 611756, Sichuan, China; 4. National United Engineering Laboratory of Integrated and Intelligent Transportation, Southwest Jiaotong University, Chengdu 611756, Sichuan, China; 5. National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Southwest Jiaotong University, Chengdu 611756, Sichuan, China)

Abstract: Coal is an essential component of China's future energy structure. Predicting short-term coal imports helps ensure the stability and efficient operation of the coal supply chain. However, existing advanced hybrid models lack sufficient consideration of influencing factors. Therefore, this

收稿日期：2024-11-07 **修订日期：**2025-10-22
通信作者：高鹏飞(2000—)，男，山东潍坊人，西南交通大学交通运输与物流学院硕士研究生。
基金项目：国家自然科学基金项目(71173177)；国家铁路局科研项目(KFJF2024-063)

paper proposed a CEEMDAN-Attention-BiLSTM-GWO-based hybrid model for predicting Mongolian coal imports. This method used complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) to decompose the coal import time series, which exhibited complex, random, and nonlinear characteristics, into 9 intrinsic mode functions representing different frequency signals. Granger causality analysis was applied to select the influencing factors that showed significant temporal correlations with coal imports, and both the decomposed modes and influencing factors were used as features in the Attention-BiLSTM model. Finally, the Grey Wolf Optimizer (GWO) was employed to find the optimal hyperparameters. Empirical results show that this model outperforms existing baseline models and demonstrates structural and functional validity, achieving better prediction performance. This prediction method can provide important references for policymakers, energy companies, and related industries, help optimize resource allocation, and reduce prediction errors.

Keywords: Coal Import Volume Prediction; Self-Attention Mechanism; Time Series Decomposition; Hybrid Model; Granger Causality Analysis

0　引言

煤炭作为我国的主要能源，在经济发展和能源安全中发挥重要作用[1-2]。尽管我国积极推进能源转型和可持续发展，煤炭仍然是未来能源结构的重要组成部分。研究煤炭进口量短期预测，能够为口岸调度与物流规划提供数据支持，保障煤炭供应链的稳定与高效运行。

煤炭进口量短期预测的研究方法与客货运量、需求量、气象状态等时间序列对象预测的方法相似，国内外学者多应用统计回归方法或机器学习方法进行研究。曾佑新等[3]梳理了我国铁路煤炭运量预测的4种方法：一元线性回归、直线趋势、自回归和二次指数平滑，并对4种预测模型的误差进行比较。张翠云[4]基于灰色马尔可夫模型，预测铁路煤炭运量。万洁[5]运用自回归差分移动平均模型(ARIMA)和三参数指数平滑法(Holt-Winters)2种时间序列模型，分别对铁路旅客周转量进行短期预测并进行对比分析。随着机器学习与深度学习等先进预测方法的兴起，学者们也将其应用到客货运量和需求量等对象的预测研究中，颜保凡等[6]构建基于径向基(RBF)神经网络的预测模型，预测铁路旅客周转量。杨丁红等[7]应用粒子群算法优化的支持向量机，建立江海联运运量预测模型。由于各种模型具备自己独特的优势，近年来涌现了大量的组合预测方法用于集成多个模型。魏姝瑶等[8]基于季节自回归积分移动平均-长短期记忆网络(SARIMA-LSTM)组合模型，预测新冠疫情这类突发性公共卫生事件影响下的铁路短时客流。龚映梅等[9]将灰色预测、时间序列ARIMA及二次平滑指数法结合，预测冷链物流需求。考虑到组合预测中机器学习与深度学习模型对于超参数的敏感程度极高，将优化算法与深度学习算法进行结合，寻找最优超参数，成为一种常见手段。例如，张赟宁等[10]提出一种基于随机森林算法(RF)特征选择和灰狼优化算法(GWO)优化高斯过程回归的组合预测模型。由于时间序列具有复杂性、随机性和非线性的特点，许多研究尝试从源头上对时间序列进行分解以提高预测性能。例如，臧海祥等[11]提出一种基于改进的自适应噪声完备集合经验模态分解(ICEEMDAN)-LSTM和残差注意力的短期太阳辐照度预测方法，其中利用ICEEMDAN将原始辐射序列分解为多尺度模态分量。王艺涵等[12]结合完备集合经验模态分解(CEEMDAN)与经粒子群优化(PSO)的变分模态分解(VMD)技术，并使用LSTM对股价进行预测。Bommidi等[13]提出一种名为ICEEMDAN-处理框架(Informer)-GWO的新颖混合框架，结合了3个组件来提高风速预测的准确性。

尽管既有研究已经通过结合模态分解算法、应用先进的深度学习模型及组合模型或使用优化算法等方式提高了时间序列预测的性能，但大多数研究忽视了影响因素中蕴含丰富信息，仅对原始时间序列信号进行处理，即使引入了影响因素也往往缺乏严格的统计关联检验。此外，不同的预测任务应寻求和设计最佳的混合模型组合形式。因此，本研究构建基于CEEMDAN-注意力机制(Attention)-双向长短时记忆网络(BiLSTM)-GWO的煤炭进口量预测混合模型，该方法集成煤炭时间序列分解模态特征和经过严格筛选的影响因素特征，并进行了一系列实

验，验证了该模型预测煤炭进口量的有效性。

1　煤炭进口量与影响因素分析

利用自相关性(Autocorrelation)分析和偏自相关性(Partial Autocorrelation)分析，揭示煤炭进口量的时间序列特征；通过格兰杰因果关系检验(Granger Causality Test)，筛选出对煤炭进口量具备预测能力的影响因素。

1.1　煤炭进口量时间规律分析

本研究使用的数据为甘其毛都口岸进口的蒙古国煤炭(以下简称“蒙煤”)量。甘其毛都口岸位于巴彦淖尔市乌拉特中旗，是中蒙贸易的关键枢纽，也是内蒙古自治区通关量和贸易额最大的公路口岸。尽管中蒙煤炭贸易发展态势良好，但伴随着国际煤炭市场环境动荡、国际局势变动等影响，两国煤炭市场贸易发展也出现了一些波动[14]。随着我国碳达峰和碳中和目标的推进，煤炭进口面临新挑战，研究甘其毛都口岸煤炭进口量趋势具有重要价值。

2017年8月16日—2024年6月30日蒙煤甘其毛都口岸进口量趋势如图1所示，在2018—2021年呈现出年周期性变化，在2021—2022年进口量处于较低水平，2022年至今进口量虽有所波动但总体处于上升水平。

自相关性分析：在时间序列分析中，自相关性用于衡量一个序列在不同时期之间的相关程度。自相关函数(ACF)用于确定某个时间序列在不同滞后期(Lags)之间的相关性，揭示了数据的周期性和模式。ACF的值在[-1，1]之间，反映了在不同时期的值之间的相关性程度。对于时间序列X_t，其ACF定义为

$$\rho_k=\frac{\sum_{t=1}^{n-k}\left(X_t-\bar{X}\right)\left(X_{t+k}-\bar{X}\right)}{\sum_{t=1}^{n}\left(X_t-\bar{X}\right)^2} \tag{1}$$

式中：ρ_k为滞后k的自相关系数；X_t为时间序列在时间t的值；$\bar{X}$为时间序列的均值；n为时间序列的长度。

偏自相关性分析：偏自相关性用于测量一个时间序列在滞后k时的纯相关性，即排除其他滞后期影响的情况下的相关性。偏自相关函数(PACF)在蒙煤进口量预测中可以确定每个滞后期的纯相关性，从而判断在模型中引入多少滞后项是合理的。通过分析PACF图，可以识别出在滞后期达到多少时，相关性开始显著下降，从而确定适当的模型阶数。本研究中使用尤尔-沃克(Yule-Walker)方程，估计模型的偏自相关性系数定义为

$$\phi_{kk}=\frac{\rho_k-\sum_{j=1}^{k=1}\phi_{j,\,k-1}\rho_{k-j}}{1-\sum_{j-1}^{k-1}\phi_{j,\,k-1}\rho_j} \tag{2}$$

式中：ϕ_{kk}为偏自相关性系数；$\phi_{j,\,k-1}$为滞后$k-1$阶的偏自相关系数；ρ_{k-j}为滞后$k-j$的自相关系数；ρ_j为滞后j的自相关系数。

蒙煤甘其毛都口岸进口量的自相关性分析和偏自相关性分析结果如图2所示。结果表明蒙煤进口量在30个滞后期内总体上均存在显著正相关性，且相关性随着滞后期的增加而减弱。而在控制了所有中间滞后效应后，偏自相关性结果表明蒙煤进口量在短期(7个滞后期以内)内存在显著的正相关性，随着时间的推移(特别是超过14个滞后期)，这种相关性逐渐减弱，甚至消失。这表明蒙煤进口量的预测模型应重点关注前14个滞后期内的时间序列信

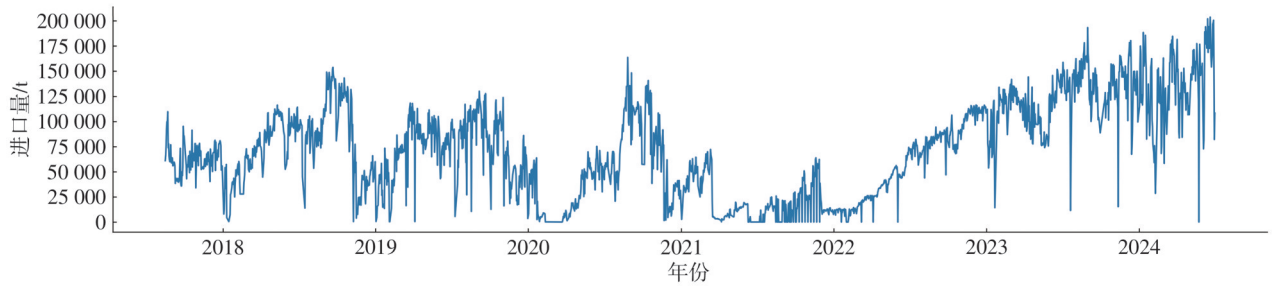

图1　蒙煤甘其毛都口岸进口量趋势

Fig.1　Trend of Mongolian coal imports at Ganqimaodu Port

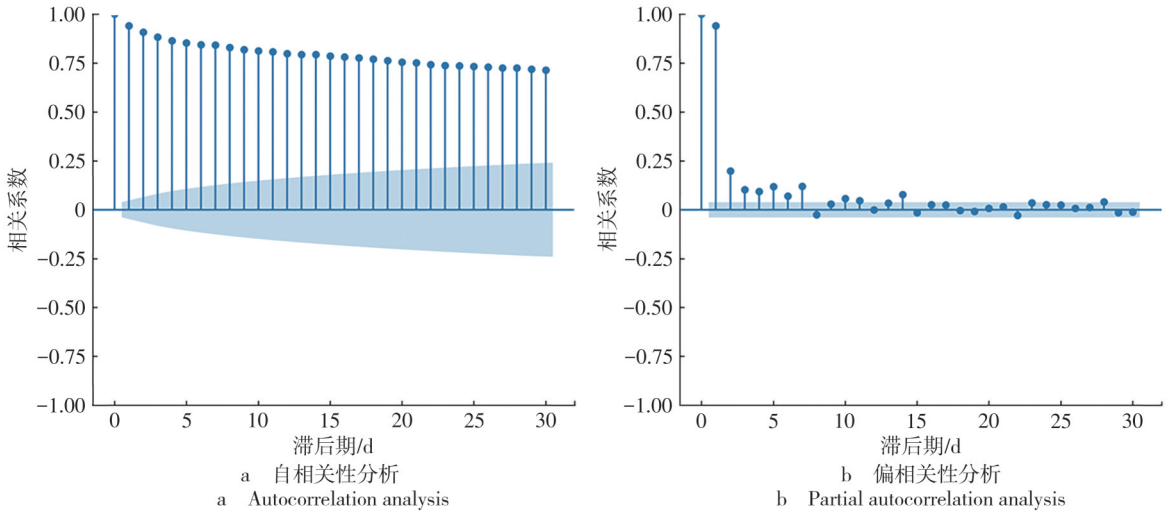

图2　蒙煤甘其毛都口岸进口量的自相关性分析与偏相关性分析结果

Fig.2　Autocorrelation and partial autocorrelation analysis of Mongolian coal imports at Ganqimaodu Port

息，以提高预测的准确性。

1.2　煤炭进口量影响因素时间关联分析

引入煤炭进口量影响因素作为特征有利于提高预测准确性，即使深度学习模型可以自动地进行特征工程，但输入到模型中的特征仍然需具备适当的逻辑关系，本研究使用格兰杰因果分析判断影响因素与煤炭进口量之间是否存在显著时间关联关系，从而验证影响因素的时间序列是否可以用来预测煤炭进口量，假设时间序列X和Y，首先建立包含Y自身滞后项和X滞后项的回归模型为

$$Y_t = \alpha_0 + \sum_{i=1}^{p}\alpha_i Y_{i-1} + \sum_{j=1}^{q}\beta_j X_{t-j} + \epsilon_t \tag{3}$$

式中：Y_{i-1}为Y的滞后项；X_{t-j}为X的滞后项；ϵ_t为误差项；α_0为常数项；α_i和β_j分别为Y的滞后项和X的滞后项的系数。

使用F检验比较包含X滞后项和不包含X滞后项的模型，若包含X滞后项的模型显著优于不包含的模型，则认为X是Y的格兰杰原因，即X在该滞后期上与Y存在显著的时间关联，纳入X对于预测Y有积极作用。

本研究收集到的潜在影响因素涉及下游港口煤炭报价、国内市场煤炭价格、蒙煤甘其毛都口岸进口价格及人民币汇率。下游港口煤炭报价包括曹妃甸港、秦皇岛港、天津港沫煤平仓价格。国内市场煤炭价格包括晋中市、临汾市、吕梁市和唐山市焦煤出厂价格，平顶山市、宿州市和太原市焦煤车板价格，沈阳市肥煤出厂价格和延安市气煤出厂价格。蒙煤甘其毛都口岸进口价格包括5#原煤和5#精煤进口价格。影响因素格兰杰因果分析P值分布结果如图3所示，设置显著性水平为90%，经过筛选，在30个滞后期内与蒙煤进口量存在显著关联的影响因素包括人民币汇率、临汾市焦煤出厂价格、沈阳市肥煤出厂价格、宿州市焦煤车板价格和太原市焦煤车板价格。其中，人民币汇率和宿州市焦煤车板价格之间的关联是即时性的，随着滞后期增加而不再产生关联关系。临汾市焦煤出厂价格在第5和第6滞后期上存在关联，太原市焦煤车板价格在16~21个滞后期上与蒙煤进口量显著关联，沈阳市肥煤出厂价格在第1个滞后期上呈现显著关联，此外，在第9个滞后期以后均与蒙煤进口量存在显著关联。考虑到蒙煤进口量在14个滞后期内存在纯自相关性，因此去除太原市焦煤车板价格，共计得到4个用于预测的影响因素，蒙煤进口量影响因素描述性统计指标如表1所示。

2　预测方法

首先介绍用于煤炭进口量预测的CEEMDAN-Attention-BiLSTM-GWO预测方法思路，之后分别介绍混合模型的各个关键组成部分，包括CEEMDAN，Attention-BiLSTM和GWO。

2.1　CEEMDAN-Attention-BiLSTM-GWO预测方法思路

CEEMDAN-Attention-BiLSTM-GWO预测方法

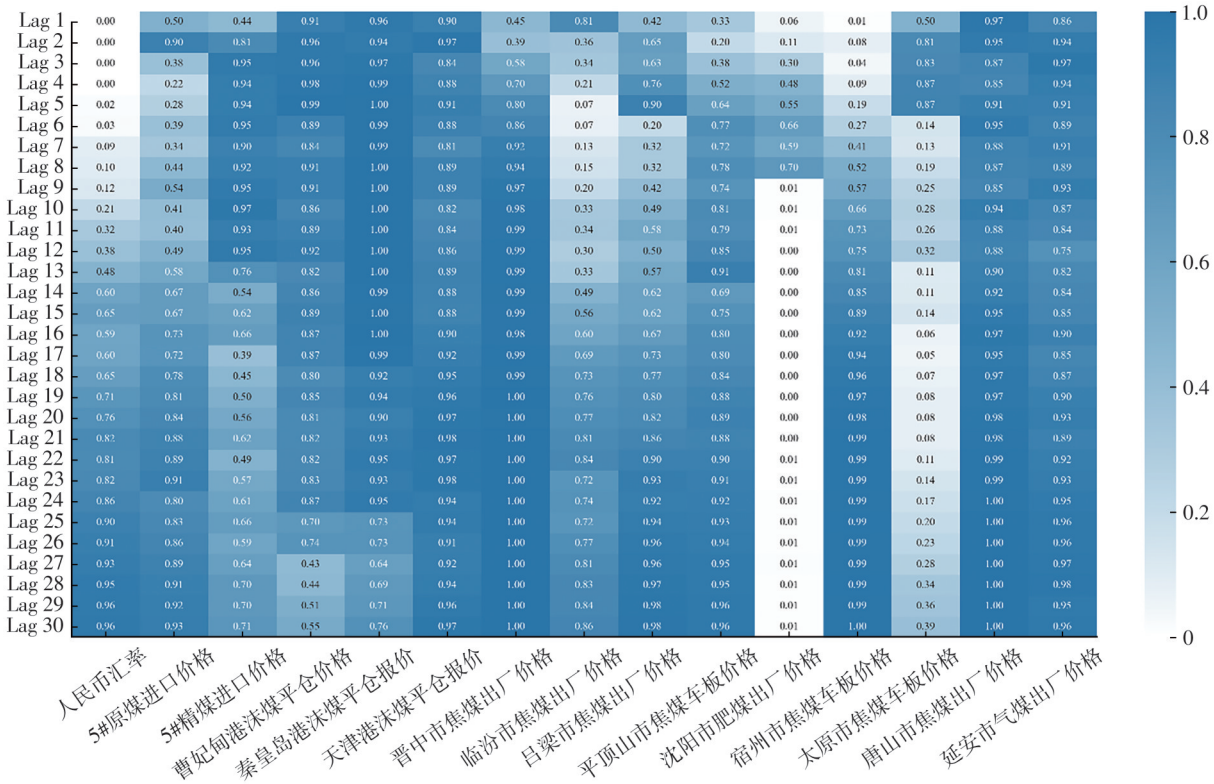

图3　影响因素格兰杰因果分析P值分布结果

Fig.3　*P*-value distribution of Granger causality analysis for influencing factors

表1　蒙煤进口量影响因素描述性统计指标

Tab.1　Descriptive statistics of influencing factors for Mongolian coal imports

影响因素	均值	标准差
蒙煤进口量/t	70 807.774 2	46 377.463 5
人民币汇率/(人民币/美元)	6.786 4	0.276 2
临汾市焦煤出厂价格/元	1 956.103 1	635.015 3
沈阳市肥煤出厂价格/元	1 948.188 0	544.096 5
宿州市焦煤车板价格/元	1 611.708 5	458.055 0

思路如图4所示。首先利用CEEMDAN分解煤炭进口量时间序列信号，得到高频、中频、低频和趋势项，将其与经过格兰杰因果分析筛选后的影响因素一并作为预测特征，经过极值归一化后划分为训练集、验证集和测试集。之后将特征与标签数据输入Attention-BiLSTM模型中，使用GWO将验证集上的表现作为目标函数，寻找Attention-BiLSTM的最优超参数组合，最终利用最优超参数训练模型得到预测输出。

2.2　CEEMDAN

CEEMDAN算法是经验模态分解(EMD)的改进版本[15]。其通过引入自适应噪声逐步消除模态混叠问题，并得到更为稳定和精确的本征模态函数(IMF)。煤炭进口量数据表现出显著的非平稳性和非线性特征，传统的分解方法难以准确捕捉其复杂性。CEEMDAN通过对原始序列进行多次加噪分解，能够有效分离出具有物理意义的IMF，有助于捕捉长期趋势、季节性变化和随机波动，提升模型对多样性特征的表达能力。此外，通过分解减少时间序列的复杂性，便于后续模型对不同频率成分进行针对性建模，避免直接对原始数据建模可能导致的特征混杂问题。以下是CEEMDAN算法的详细流程和计算公式。

第1步：对原始信号添加白噪声，生成多个含噪信号$X_i(t)$，计算公式为

$$X_i(t) = X(t) + W_i(t) \tag{4}$$

式中：$X(t)$为原始信号；$W_i(t)$为第i次添加的白噪声。

对每个含噪信号$X_i(t)$进行EMD，得到第一层的每个IMF，$IMF_{1i}(t)$计算公式为

$$IMF_{1i}(t) = X_i(t) + r_{i1}(t) \tag{5}$$

式中：$r_{i1}(t)$为剩余的信号。

第2步：计算所有含噪信号在第一层每个IMF

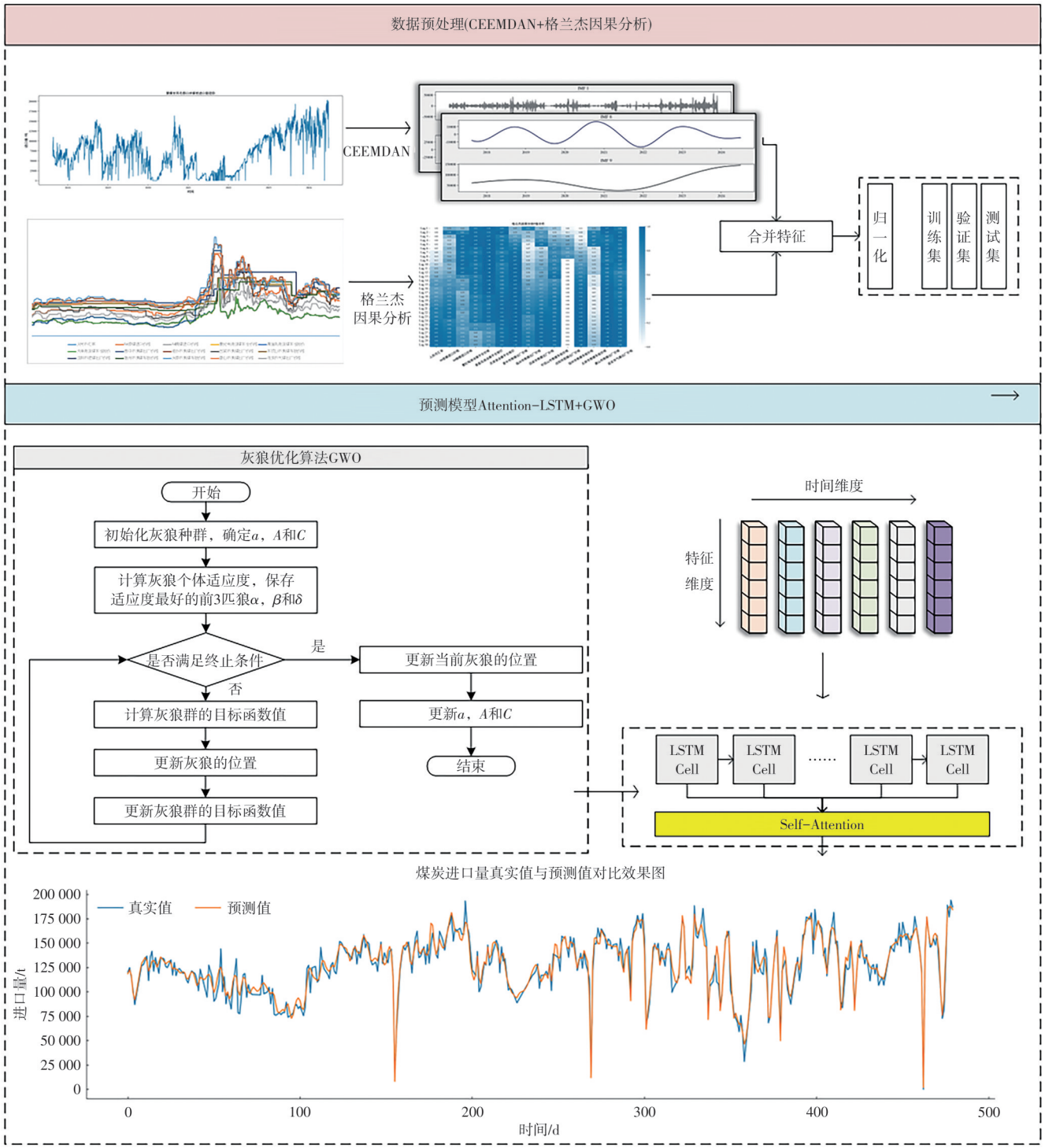

图4　CEEMDAN-Attention-BiLSTM-GWO预测方法思路

Fig.4　CEEMDAN-Attention-BiLSTM-GWO predicting method

上的均值，得到第一层IMF，$IMF_1(t)$计算公式为

$$IMF_1(t)=\frac{1}{M}\sum_{i=1}^{M}IMF_{1i}(t) \tag{6}$$

式中：M为添加白噪声的次数。

第3步：计算残差信号$r_1(t)$，$r_1(t)$是原始信号减去第一层IMF的结果，计算公式为

$$r_1(t)=X(t)-IMF_1(t) \tag{7}$$

对残差信号$r_1(t)$再次添加白噪声$W_i(t)$，并进行EMD以提取下一个IMF。

第4步：重复上述步骤，直到所有IMF都被提

取出来。每次迭代提取IMF后，都会计算残差信号，并添加噪声进行下一层的IMF分解。此过程继续进行，直到残差信号变为单调或无法进一步分解为止。

2.3　Attention-BiLSTM模型

LSTM模型是一种常用于处理和预测序列数据的深度学习模型，特别适用于具有长期依赖关系的序列数据。相比传统的循环神经网络(RNN)模型，LSTM模型能够更好地解决梯度消失和梯度爆炸等问题，从而更好地捕捉序列数据中的依赖关系。LSTM的核心在于其独特的结构，包含输入门、遗忘门和输出门3个门，以及一个细胞状态(Cell State)。遗忘门决定了前一时刻的细胞状态有多少部分需要被遗忘。输入门决定了当前输入对细胞状态的影响，输入门包括2个部分：一个决定更新部分的值，另一个决定更新部分的比例。而细胞状态根据遗忘门和输入门的结果进行更新。输出门决定了当前细胞状态对隐状态的影响，并产生输出。LSTM的前向传播方式为

$$f_t=\sigma\left(W_f\left[h_{t-1},\ x_t\right]+b_f\right) \tag{8}$$

$$i_t=\sigma\left(W_i\left[h_{t-1},\ x_t\right]+b_i\right) \tag{9}$$

$$\tilde{C}_t=\tanh\left(W_C\left[h_{t-1},\ x_t\right]+b_C\right) \tag{10}$$

$$C_t=f_t\cdot C_{t-1}+i_t\cdot\tilde{C}_t \tag{11}$$

$$o_t=\sigma\left(W_o\left[h_{t-1},\ x_t\right]+b_o\right) \tag{12}$$

$$h_t=o_t\cdot\tanh\left(C_t\right) \tag{13}$$

式中：f_t为遗忘门的激活值；W_f为遗忘门的权重；$\left[h_{t-1},\ x_t\right]$为上一个隐藏状态$h_{t-1}$和当前输入$x_t$的拼接；$b_f$为遗忘门的偏置项；$i_t$为输入门的激活值；$\tilde{C}_t$为候选的细胞状态值；$W_i$和$W_C$为对应的权重；$b_i$和$b_C$为偏置项；$C_t$为更新后的细胞状态；$C_{t-1}$为上一个状态的细胞状态；$o_t$为输出门的激活值；$W_o$为输出门的权重；$b_o$为偏置项；$h_t$为当前状态的隐状态(输出)。

自注意力(Self-Attention)是神经网络中一种机制，用于计算输入序列中各个元素之间的相关性，并根据计算结果对输入信息进行加权，从而突出重要信息。在蒙煤进口量预测中，采用Self-Attention机制来动态调整LSTM的输出。该机制能根据输入序列，为隐藏状态的不同位置分配重要性权重，有助于组合预测模型更有效地捕获输入序列中的关键信息，并提高其在序列建模任务中的性能，计算过程如下所示。

$$h\times\left(W_Q,\ W_K,\ W_V\right)=\left(\boldsymbol{Q},\ \boldsymbol{K},\ \boldsymbol{V}\right) \tag{14}$$

$$S\left(\boldsymbol{Q},\ \boldsymbol{K}\right)=\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\left|\left|\boldsymbol{Q}\right|\right|\times\left|\left|\boldsymbol{K}\right|\right|} \tag{15}$$

$$W_A=\mathrm{softmax}\left(S\left(\boldsymbol{Q},\ \boldsymbol{K}\right)\right) \tag{16}$$

$$O=W_A\boldsymbol{V} \tag{17}$$

式中：h为由BiLSTM每一步输出拼接成的张量；$\left\{W_Q,\ W_K,\ W_V\right\}$为生成查询向量$\boldsymbol{Q}$、键向量$\boldsymbol{K}$和值向量$\boldsymbol{V}$的权重；$S\left(\boldsymbol{Q},\ \boldsymbol{K}\right)$为查询向量$\boldsymbol{Q}$和键向量$\boldsymbol{K}$之间的余弦相似度；$W_A$为经softmax操作处理后的注意力权重；$O$为经注意力权重调整后的输出。

在本研究中，BiLSTM为在LSTM上设置双向操作，BiLSTM通过正向和反向的双向循环单元，能够同时捕捉蒙煤进口量时间序列的前向和后向信息，适应具有长期依赖和复杂动态变化的预测任务。Self-Attention机制能够根据输入特征的重要性动态分配权重，突出对预测结果贡献较大的CEEMDAN分解模态和影响因素，增强模型对关键信息的关注能力，提升预测精度。

2.4　灰狼优化算法

GWO是一种基于灰狼社会行为的群体智能优化算法，由Mirjalili等于2014年提出[16]。该算法模拟了灰狼在自然界中的捕猎行为，主要分为3个阶段：搜索猎物、包围猎物和攻击猎物。由于其具备强大的全局搜索与局部开发能力，适用于复杂、高维参数空间的优化，能够有效调整Attention-BiLSTM模型的超参数。在蒙煤进口量预测问题中，GWO能够根据模型性能指标(如损失函数)动态调整搜索策略，因此更能提升模型在复杂、多尺度数据环境下的适应性。

灰狼优化算法中，灰狼分为4种等级：α狼、β狼、δ狼和ω狼。α狼是领导者，β狼和δ狼辅助δ狼进行决策，ω狼则跟随其他灰狼行动。算法通过模拟灰狼围捕猎物的行为更新狼群的位置，并寻找全局最优解。主要步骤如下。

第1步：初始化灰狼群体的位置，包括α狼、β狼、δ狼和ω狼的位置。通常，位置是随机生成的。

第2步：计算每只灰狼的位置对应的适应度值，并根据适应度值确定α狼、β狼、δ狼。

第3步：根据α狼、β狼、δ狼的位置，更新其他灰狼的位置。

第4步：重复适应度评估和位置更新步骤，直到满足终止条件，如达到最大迭代次数或满足预设精度。

位置更新依靠的计算公式为

$$\vec{D}_\alpha = \left|\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}\right| \tag{18}$$

$$\vec{D}_\beta = \left|\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}\right| \tag{19}$$

$$\vec{D}_\delta = \left|\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}\right| \tag{20}$$

式中：$\vec{D}_\alpha$，$\vec{D}_\beta$，$\vec{D}_\delta$为ω狼与α狼、β狼、δ狼各自的相对距离；$\vec{X}$为当前灰狼的位置；$\vec{X}_\alpha$，$\vec{X}_\beta$和$\vec{X}_\delta$分别为α狼、β狼、δ狼的位置；$\vec{C}_1$，$\vec{C}_2$，$\vec{C}_3$为随机变量。

进一步更新位置为

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 \cdot \vec{D}_\alpha \tag{21}$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2 \cdot \vec{D}_\beta \tag{22}$$

$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3 \cdot \vec{D}_\delta \tag{23}$$

$$\vec{X}(t-1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{24}$$

式中：$\vec{X}_1$，$\vec{X}_2$，$\vec{X}_3$为α狼、β狼、δ狼分别建议ω狼移动到的位置；$\vec{A}_1$，$\vec{A}_2$和$\vec{A}_3$为随机变量。

$\vec{A}$和$\vec{C}$的更新计算公式为

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{25}$$

$$\vec{C} = 2 \cdot \vec{r} \tag{26}$$

式中：$\vec{a}$为从2线性递减到0；$\vec{r}$为[0，1]的随机变量。

3　实验分析

3.1　实验设置

3.1.1　实验输入与参数说明

输入至模型中的数据格式为[批次大小、时间序列长度、特征维度]，批次大小表示每次输入至模型的特征时间序列个数，时间序列长度表示用多长时间的历史特征预测未来的煤炭进口量，特征维度则包括模态特征和影响因素特征。本研究中预测模型的批次大小为32；时间序列长度为14，即选取14 d的特征预测未来1 d的蒙煤进口量；特征维度为14，包括9个模态分解特征及5个影响因素特征。

在划分数据集时，首先将2017年8月16日—2024年6月30日的特征时间序列处理成76个[32，14，13]维度的批次输入，按照时间序列顺序将70%的输入作为训练集，10%的输入作为验证集，20%的输入作为测试集，因此训练集、验证集及测试集的批次个数分别为54，7，15。

训练的轮数设置为1 000，经过GWO得到的最优超参数中批次大小为32，BiLSTM层和Self-Attention的隐藏单元数皆为18，Self-Attention的头数为1，BiLSTM的层数为5，学习率设置为0.004，使用验证集上表现最好的模型作为最终的预测模型，并在测试集上测试效果。算法计算流程与数据流动过程如图5所示。

3.1.2　评价指标与损失函数

使用均方误差(MSE)、均方根误差(RMSE)和平均绝对值误差(MAE)作为评价指标，MSE也被作为本研究的损失函数，计算公式为

$$A_{\mathrm{MSE}} = \frac{1}{N}\sum_{I=1}^{N}\left(y_t - \hat{y}_t\right)^2 \tag{27}$$

$$A_{\mathrm{RMSE}} = \sqrt{\frac{1}{N}\sum_{I=1}^{N}\left(y_t - \hat{y}_t\right)^2} \tag{28}$$

$$A_{\mathrm{MAE}} = \frac{1}{N}\sum_{i=1}^{N}\left|y_t - \hat{y}_t\right| \tag{29}$$

式中：A_{MSE}，A_{RMSE}，A_{MAE}分别为MSE，RMSE，MAE；N为数据集中样本的数量；y_t为煤炭进口量真实值，t；$\hat{y}_t$为煤炭进口量的预测值，t。

3.1.3　基线模型

使用的基线模型涉及常见的时间序列预测深度学习模型，包括LSTM，门控循环单元(GRU)，BiLSTM，以及时序卷积网络(TCN)。为了验证CEEMDAN算法有效性，基于CEEMDAN-Attention-BiLSTM-GWO框架，构建EMD-Attention-BiLSTM-GWO和集合经验模态分解(EEMD)-Attention-BiLSTM-GWO进行对比。此外，将模型的Attention-BiLSTM部分替换为LSTM，GRU和TCN，得到CEEMDAN-LSTM-GWO，CEEMDAN-GRU-GWO和CEEMDAN-TCN-GWO，共计得到9个基线模型。为了充分发挥模型的最佳性能，所有的基线模型均配合GWO寻找最优超参数。

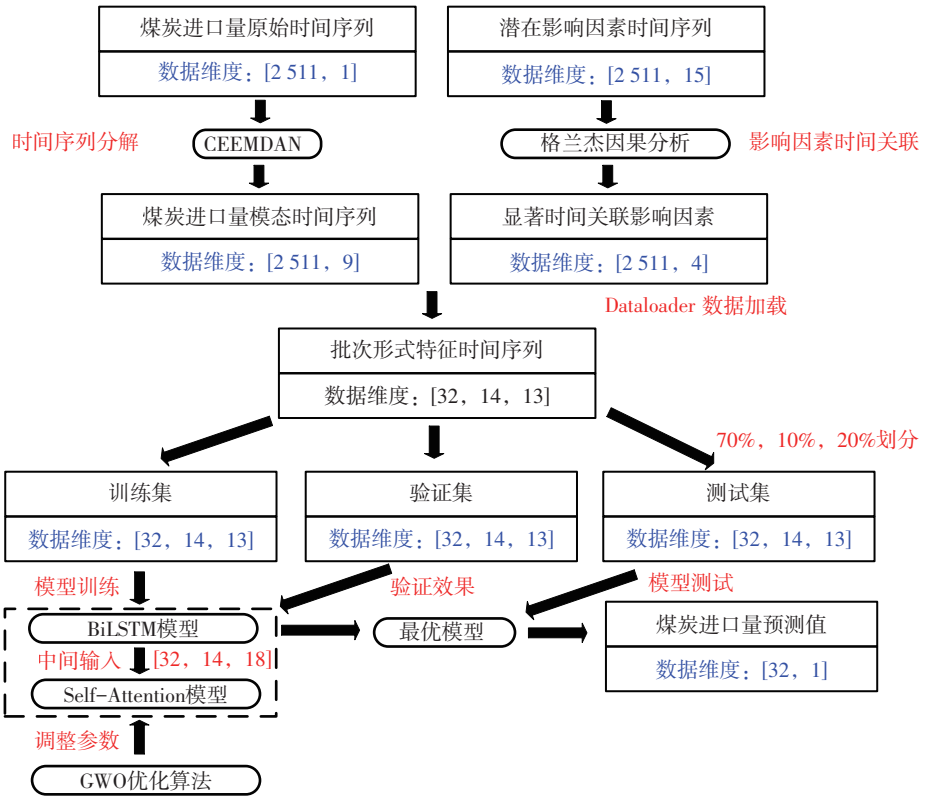

图5　算法计算流程与数据流动过程

Fig.5　Algorithm calculation process and data flow

3.2　实验结果分析

3.2.1　煤炭进口量时间序列分解结果

煤炭进口量时间序列经过CEEMDAN分解后得到多个IMF，煤炭进口量时间序列CEEMDAN分解结果如图6所示，这些IMF代表了原始信号中不同频率成分的振荡模式。高频IMF(IMF1，IMF2)表示煤炭进口量的噪声部分或短期波动。中频IMF(IMF3，IMF4)对应信号中的中等频率成分，与周期性或半周期性趋势有关。低频IMF(IMF5及之后的分量)则表示信号的长周期趋势或较慢的变化成分，与信号的整体趋势或季节性变化有关。在低频IMF中，IMF9是趋势项，捕捉了煤炭进口量中的长期变化趋势，与煤炭进口量趋势保持一致，在2021—2022年迎来谷底之后不断上升。

3.2.2　模型对比实验结果

在测试集的特征时间序列上利用模型计算得到预测值，与对应的真实值趋势进行对比，煤炭进口量预测值与真实值对比如图7所示。从图7看出，真实值与预测值的整体趋势基本一致，表明提出的方法CEEMDAN-Attention-BiLSTM-GWO，能基于历史一段时间内煤炭进口量及相关影响因素的变化趋势，实现对未来煤炭进口量的准确预测，尤其是在波动较大的时期依然能跟随主趋势变化。从整个时间序列来看，模型在多数时间段内表现稳定，未出现大幅度的漂移或预测失效，说明模型在时间上的泛化性较好，也具备一定的鲁棒性。

将CEEMDAN-Attention-BiLSTM-GWO与其他各种基线模型进行对比，模型对比实验结果如表2所示。LSTM，GRU，TCN及BiLSTM由于未使用CEEMDAN算法进行模态分解，导致原始煤炭进口量时间序列数据中包含较多噪声，因而均表现出较差的性能，其中BiLSTM由于设置了双向结构相比LSTM模型效果较好，GRU由于减少了门控机制的复杂性相比LSTM表现也相对较好。在使用EMD方法分解了煤炭进口量时间序列之后，模型的MSE，RMSE，MAE均大幅度上升。EEMD通过减轻模式混叠并提高分解稳定性，显著提升了分解效果，相较于EMD方法更为有效。在将Attention-BiLSTM替换为LSTM，GRU，TCN之后，由于CEEMDAN的应用，相较于不使用CEEMDAN的LSTM，GRU和

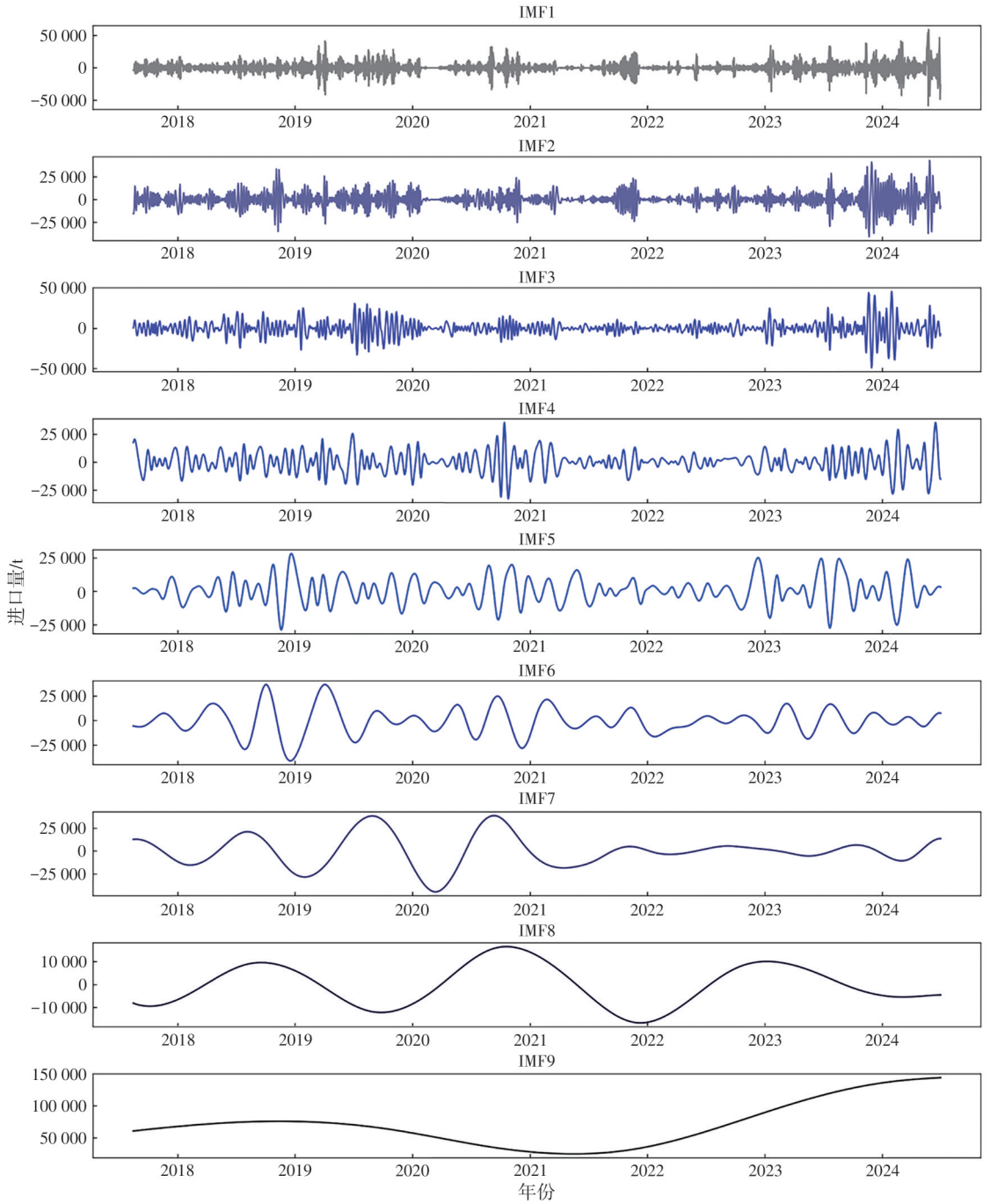

图6　煤炭进口量时间序列CEEMDAN分解结果

Fig.6　CEEMDAN decomposition results of Mongolian coal import time series

TCN结果存在较大区别，MSE下降了60%左右，RMSE下降了35%~40%，MAE下降了28%~34%。

相较于使用了EMD和EEMD的方法，本研究提出的方法CEEMDAN-Attention-BiLSTM-GWO依托CEEMDAN可以引入自适应噪声，精确估计每个IMF的残差并且能够处理具有复杂时变特性的非平稳信号的能力，在各个指标上均大幅度领先其他使用模态分解算法的模型。而相较于CEEMDAN-LSTM-GWO，CEEMDAN-GRU-GWO和CEEMDAN-TCN-GWO，由于Attention-BiLSTM的更加高级和深层次的模型结构，仍然表现出卓越的性能。

3.2.3　模型消融实验结果

为了进一步验证CEEMDAN-Attention-BiLSTM-GWO的完整性和有效性，还进行了一系列消融实

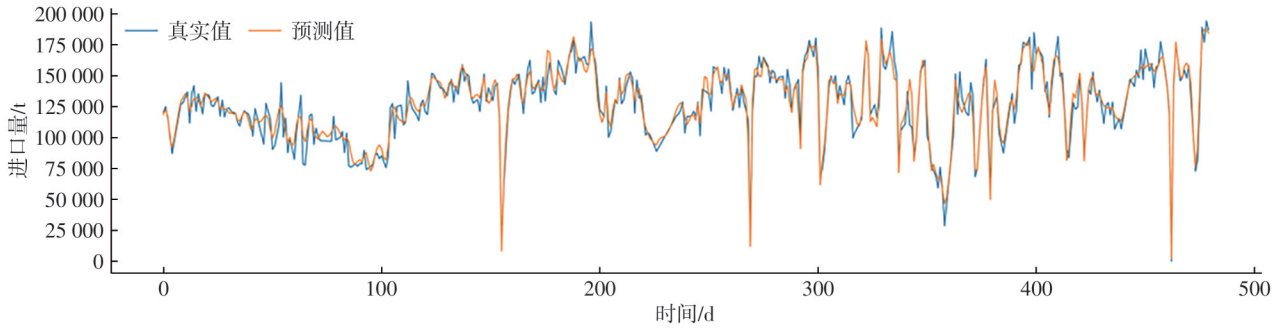

图7　煤炭进口量预测值与真实值对比

Fig.7　Comparison of predicted and actual values of Mongolian coal imports

表2　模型对比实验结果

Tab.2　Model comparison experiment results

模型	MSE	RMSE	MAE
LSTM-GWO	0.006 12	0.078 23	0.048 38
GRU-GWO	0.004 69	0.068 48	0.042 50
TCN-GWO	0.006 12	0.078 25	0.049 24
BiLSTM-GWO	0.004 60	0.067 82	0.041 65
EMD-Attention-BiLSTM-GWO	0.002 52	0.050 18	0.034 22
EEMD-Attention-BiLSTM-GWO	0.001 62	0.040 20	0.027 34
CEEMDAN-LSTM-GWO	0.002 28	0.047 72	0.031 61
CEEMDAN-GRU-GWO	0.001 95	0.044 17	0.030 27
CEEMDAN-TCN-GWO	0.002 22	0.047 09	0.031 32
CEEMDAN-Attention-BiLSTM-GWO	0.001 43	0.037 84	0.023 40

验，实验分别从输入特征与模型结构入手，具体包括：不使用CEEMDAN分解方法、移除模型中的Self-Attention机制、在预测特征中排除影响因素。模型消融实验结果如表3所示。和模型对比实验中引入模态分解对于基线模型LSTM、GRU和TCN的提升一样，使用CEEMDAN会显著提升煤炭进口量预测性能。在去除Self-Attention机制之后，模型性能同样会较大幅度下降，这是由于Self-Attention机制允许模型在处理输入序列时，动态地关注重要的部分，而忽略不太相关的部分。这样，模型可以更有效地从长序列中提取有用信息，提高对复杂任务的处理能力，去除之后则会导致信息提取的失真，进而影响预测性能。去除影响因素特征之后相当于抛弃了一部分宝贵信息，减弱了信息的丰富度，因而导致模型性能下降。根据结果可以得知，模型包括特征中组成部分的重要度排序为CEEMDAN>影响因素特征>Self-Attention机制。

表3　模型消融实验结果

Tab.3　Model ablation experiment results

模型	MSE	RMSE	MAE
Attention-BiLSTM-GWO	0.005 12	0.071 52	0.047 36
CEEMDAN-BiLSTM-GWO	0.002 14	0.046 31	0.031 98
不使用影响因素特征	0.003 26	0.057 13	0.036 81
CEEMDAN-Attention-BiLSTM-GWO	0.001 43	0.037 84	0.023 40

4　结束语

基于CEEMDAN-Attention-BiLSTM-GWO的煤炭进口量预测混合模型，使用CEEMDAN将煤炭进口量时间序列进行分解得到高频、中频、低频和趋势项，使用格兰杰因果分析筛选与煤炭进口量呈现显著关联的影响因素，将两者作为特征输入到由Self-Attention机制和BiLSTM组合形成的Attention-BiLSTM框架[17-18]，最后使用GWO算法寻找最优超参数。对比实验结果表明提出的模型优于基线模型，消融实验结果表明该模型具备结构上的合理性和完整性。尽管模型对于煤炭进口量预测取得了较好的效果，但选取的影响因素尚未涵盖国际形势变动、国际上其他煤炭来源的价格因素及更多的下游港口多种类煤炭报价的动态变化。未来，需要进一步探讨这些因素与煤炭短期进口量之间的复杂关系，并尝试将其量化后纳入预测维度，进一步提升模型的预测精度和稳定性。

参考文献：

[1] LI B B，LIANG Q M，WANG J C. A Comparative Study on Prediction Methods for China's Medium-and Long-Term Coal Demand[J]. Energy，2015，93：1671-1683.

[2] LIN B Q，LIU J H. Estimating Coal Production Peak and Trends of Coal Imports in China[J]. Energy Policy，2010，38(1)：512-519.

[3] 曾佑新，刘海燕，周尚超. 我国铁路煤炭运量预测模型的研究[J]. 华东交通大学学报，2000，17(1)：83-88.

[4] 张翠云. 基于灰色马尔可夫链的铁路煤炭运量预测研究[J]. 煤炭经济研究，2007，27(12)：41-43.

[5] 万　洁. ARIMA模型与Holt-Winters模型在铁路旅客周转量预测中的比较[J]. 铁道运输与经济，2010，32(2)：91-94.

[6] 颜保凡，郭垂江. 基于RBF神经网络的铁路旅客周转量预测研究[J]. 铁道运输与经济，2017，39(12)：65-69.
YAN Baofan，GUO Chuijiang. A Study on the Prediction of Railway Passenger Volumes Based on RBF Neural Network[J]. Railway Transport and Economy，2017，39(12)：65-69.

[7] 杨丁红，郑彭军. 基于因素分析和SVM模型的江海联运运量预测方法[J]. 宁波大学学报(理工版)，2020，33(2)：86-92.
YANG Dinghong，ZHENG Pengjun. Method for Prediction of River-Sea Combined Transport Volume Based on Factor Analysis and SVM Model[J]. Journal of Ningbo University (Natural Science & Engineering Edition)，2020，33(2)：86-92.

[8] 魏姝瑶，张　瑾. SARIMA-LSTM组合模型在铁路疫情短时客流的预测研究[J]. 重庆理工大学学报(自然科学)，2023，37(3)：204-211.
WEI Shuyao，ZHANG Jin. Prediction Research of SARIMA-LSTM Combination Model in the Short-Term Railway Passenger Flow during Epidemics[J]. Journal of Chongqing University of Technology (Natural Science)，2023，37(3)：204-211.

[9] 龚映梅，王　宁. 基于组合预测法的云南省生鲜农产品冷链物流需求预测[J]. 江苏商论，2024(5)：33-37.
GONG Yingmei，WANG Ning. Prediction of Cold Chain Logistics Demand for Fresh Agricultural Products in Yunnan Province Based on Combination Prediction Method[J]. Jiangsu Commercial Forum，2024(5)：33-37.

[10] 张赟宁，魏广军. 考虑特征选择的短期光伏功率组合预测模型[J]. 电力系统及其自动化学报，2024，36(8)：122-132.
ZHANG Yunning，WEI Guangjun. Combined Prediction Model for Short-Term Photovoltaic Power Considering Feature Selection[J]. Proceedings of the CSU-EPSA，2024，36(8)：122-132.

[11] 臧海祥，张　越，程礼临，等. 基于ICEEMDAN-LSTM和残差注意力的短期太阳辐照度预测[J]. 太阳能学报，2023，44(12)：175-181.
ZANG Haixiang，ZHANG Yue，CHENG Lilin，et al. Short Term Solar Radiation Forecasting Based on ICEEMDAN-LSTM and Residual Attention[J]. Acta Energiae Solaris Sinica，2023，44(12)：175-181.

[12] 王艺涵，路　翀，龙　洁，等. 基于优化VMD再分解的LSTM下的股价预测[J]. 信息技术与信息化，2024(7)：4-11.

[13] BOMMIDI B S，TEEPARTHI K，DULLA MALLESHAM V K. ICEEMDAN-Informer-GWO：A Hybrid Model for Accurate Wind Speed Prediction[J]. Environmental Science and Pollution Research，2024，31(23)：34056-34081.

[14] 乌齐卡. 中蒙煤炭贸易发展问题研究[D]. 哈尔滨：黑龙江大学，2021.

[15] TORRES M E，COLOMINAS M A，SCHLOTTHAUER G，et al. A Complete Ensemble Empirical Mode Decomposition with Adaptive Noise[EB/OL]. (2011-12-31) [2024-10-12]. https：//ieeexplore. ieee. org/abstract/document/5947265.

[16] MIRJALILI S，MIRJALILI S M，LEWIS A. Grey Wolf Optimizer[J]. Advances in Engineering Software，2014，69：46-61.

[17] 江路路，高天泽. 基于组合模型的黄大铁路货运量预测研究[J]. 铁路物流，2025，43(9)：58-64.
JIANG Lulu，GAO Tianze. Research on Freight Volume Forecasting of Huanghua-Dajiawa Railway Based on a Combined Model[J]. Railway Logistics，2025，43(9)：58-64.

[18] 张海山，王文斌，周　瑾. 基于异质时空图注意力网络的铁路车站货运量预测[J]. 铁道货运，2024，42(6)：52-59.
ZHANG Haishan，WANG Wenbin，ZHOU Jin. Freight Volume Prediction of Railway Station Based on Heterogeneous Spatial-Temporal Graph Attention Network[J]. Railway Freight Transport，2024，42(6)：52-59.

(责任编辑　李丹丹)