



# 基于注意力时间卷积网络的农产品期货分解集成预测

## 摘要

针对农产品期货时间序列数据受多方面因素影响,非线性、非平稳数据特征难以提取而导致预测准确性不高的问题,基于“分解-集成”的预测思想,本文提出一种基于自适应噪声完备经验模态分解(CEEMDAN)与Transformer-Encoder-TCN的农产品期货预测方法。首先,使用CEEMDAN将时间序列分解为多尺度多频率的本征模态分量(IMF)与残差,降低了序列建模复杂度;其次,使用融合多阶段自注意力单元Transformer-Encoder的时间卷积网络(TCN)对各个分量序列进行特征提取与预测,优化了序列显著特征建模权重;最后,将各个子序列预测值线性相加集成得到最终预测结果。以南华期货公司农产品指数中的大豆期货指数为研究对象,采用时序交叉验证与参数迁移的方式进行模型重训练,消融和对比实验结果表明,提出的新模型在RMSE、MAE和DS三个评价指标上具有良好的效果,验证了该模型对农产品期货预测的有效性。

## 关键词

农产品期货;自适应噪声完备经验模态分解;自注意力机制;Transformer-Encoder;时间卷积网络

中图分类号 F830.91;TP183

文献标志码 A

收稿日期 2023-08-22

资助项目 国家自然科学基金面上项目(71971089);国家自然科学基金青年项目(72001083);广东省自然科学基金面上项目(2022A1515011612)

## 作者简介

张大斌,男,博士,教授,研究方向为信息预测理论与方法.zdbff@aliyun.com

林锐斌(通信作者),男,硕士生,研究方向为智能预测建模及应用.1103290907@qq.com

## 0 引言

农产品作为关系国计民生的大宗商品之一,是农业经济发展中重要的组成部分,其价格剧烈波动将阻碍农产品经济的可持续健康发展。而期货市场具备套期保值与价格发现功能,稳定高效的农产品期货交易市场预测机制在稳定农产品供应量、粮食现货价格预警方面具有重要意义,也能为政府决策者制定农业经济政策提供重要参考<sup>[1]</sup>。

传统的农产品时间序列预测建模方法主要基于计量统计模型,该类模型具有较强的线性拟合能力,包括ARIMA、ETS、GARCH等<sup>[2]</sup>。然而,以上传统计量统计模型非线性建模能力较差,随着市场影响因素的增多,模型泛化能力也随之下降。因此,不少学者将分解集成策略与预测模型相结合以提高农产品时间序列预测精度<sup>[3]</sup>。分解集成预测方法是一种分而治之的策略,常见的分解方法包括小波分解(WD)<sup>[4]</sup>、经验模态分解(EMD)<sup>[5]</sup>以及自适应噪声完备经验模态分解(Complementary Ensemble Empirical Mode Decomposition with Adaptive Noise,CEEMDAN)<sup>[6]</sup>等,这些分解方法在时序预测任务中起到重要作用。例如:唐振鹏等<sup>[7]</sup>利用VMD与EMD进行粮食期货价格序列二次分解,再进行集成学习预测;Zhou等<sup>[8]</sup>利用CEEMDAN分解原油价格序列后使用XGBoost算法对其进行集成预测。

随着人工智能算法的发展,基于神经网络的深度学习预测模型被应用于农产品时间序列预测领域,常见的算法包括长短期记忆网络(LSTM)<sup>[9]</sup>、时间卷积网络(Temporal Convolution Network,TCN)<sup>[10]</sup>等。以往的研究表明,LSTM具备记忆时序特征能力,TCN则具有较强的局部时序相关性建模能力,上述模型在农产品时间序列短期预测领域具有较强优势<sup>[11-12]</sup>。此外,这些模型在各类农产品时间序列预测任务上呈现优越的非线性建模能力。Wang等<sup>[13]</sup>利用LSTM对大连期货交易所的大豆期货价格进行预测,验证了LSTM模型具有更好的预测精度;Chuluunsaikhan等<sup>[14]</sup>将主题文本建模技术与深度学习模型CNN、LSTM相结合,预测了韩国市场猪肉每日零售价格;张杰等<sup>[15]</sup>构建了融合传递熵的图神经网络,有效地提高了大豆期货价格走势预测精度。

然而,上述深度学习模型对于长序列的全局建模仍具有局限性,因此有学者开始尝试将全局建模能力极佳的注意力机制与深度学习

1 华南农业大学 数学与信息学院,广州,510642

模型融合并应用到农产品时间序列预测领域<sup>[16]</sup>.

注意力机制最早应用于计算机视觉中<sup>[17]</sup>.2017年,谷歌机器翻译团队提出了经典的 Transformer 模型<sup>[18]</sup>,该模型采用多头自注意力机制(Multi-head Self-Attention)自动捕捉输入序列不同位置处的相对关联来进行机器翻译任务,并且取得了良好的效果;Niu 等<sup>[19]</sup>利用 CEEMDAN 分解能源类时间序列,再利用融合注意力机制的 GRU 模型进行预测,预测精度得到显著提高;贾宁等<sup>[20]</sup>设计了 LSTM-DA 模型用于蔬菜类农产品价格预测,先利用 CNN 层提取局部特征,再利用 LSTM 提取序列特征,两层均融合了注意力机制,有效地降低了预测误差;Yin 等<sup>[21]</sup>利用 STL 分解策略将蔬菜价格分解为趋势成分、季节性成分和残差成分,并使用融合注意力机制的 LSTM 进行预测,解决了因季节性影响而产生的预测滞后问题.

本文结合传统复杂时间序列分解集成建模策略与最新人工智能理论研究各自的优势,提出一种分解集成预测策略下融合多阶段自注意力单元 Transformer-Encoder 的时间卷积网络模型,并应用在农产品期货收益率序列预测任务中.本文设计的 CEEMDAN-Transformer-Encoder-TCN 方法主要的创新点和贡献点如下:

1)在 TCN 建模过程中融合了多阶段自注意力单元 Transformer-Encoder,注意力机制与 TCN 模型的因果卷积机制两者形成互补,有效地增强了模型局部与全局建模能力,自动优化农产品期货显著性时间点与季节性成分的建模权重,构建了一种新颖的时间卷积神经网络结构.

2)由于农产品期货收益率具有非平稳、非线性特征,基于 CEEMDAN 将收益率序列分解为多个频率的 IMF(Intrinsic Mode Function)分量,降低了序列复杂度.同时,采用时序交叉验证与参数迁移的方式构建的模型训练流程有效地挖掘了子序列特征,改善了预测效果.

3)基于真实大豆期货收益率数据集进行实验,探索了一种新的农产品期货收益率预测方法,未来可扩展运用到多种农产品期货预测任务上.

## 1 研究方法

### 1.1 自适应噪声完备经验模态分解(CEEMDAN)

CEEMDAN 是一种有效的信号序列分解算法,由 Torres 等<sup>[22]</sup>在 2011 年提出,其对残差加入高斯白

噪声并继续分解计算 IMF 分量均值,有效地解决了白噪声从高频到低频的转移传递问题.CEEMDAN 算法过程描述如下:

步骤 1.将待分解信号  $x(t)$  添加  $K$  次均值为 0 的高斯白噪声,构造  $K$  个待分解序列  $x_i(t)$ :

$$x_i(t) = x(t) + \varepsilon \delta_i(t). \quad (1)$$

式中: $\varepsilon$  为高斯白噪声权值系数, $\delta_i(t)$  为第  $i$  次处理时产生的高斯白噪声; $i = 1, 2, 3, \dots, K$ .

步骤 2.对步骤 1 得到的序列  $x_i(t)$  进行 EMD 分解,得到多个第 1 阶模态分量并求总体平均作为 CEEMDAN 分解得到的第 1 个 IMF 分量:

$$f_{\text{IMF},1}(t) = \frac{1}{K} \sum_{i=1}^K f_{\text{IMF},1}^i(t). \quad (2)$$

步骤 3.计算去除第 1 个模态分量后的残差:

$$r_j(t) = x(t) - f_{\text{IMF},1}(t). \quad (3)$$

步骤 4.将分解后得到的第  $j$  阶段余量信号序列添加特定噪声后,继续进行 EMD 分解:

$$f_{\text{IMF},j}(t) = \frac{1}{K} \sum_{i=1}^K E_1(r_j(t) + \varepsilon_j E_j(\delta_i(t))), \quad (4)$$

$$r_{j+1}(t) = r_j(t) - f_{\text{IMF},j}(t). \quad (5)$$

式中: $f_{\text{IMF},j}(t)$  表示 CEEMDAN 分解得到的第  $j$  个模态分量; $E_j(\cdot)$  表示对序列进行 EMD 分解后的第  $j$  个 IMF 分量; $\varepsilon_j$  表示 CEEMDAN 对第  $j$  阶段余量信号加入噪声的权值系数; $r_{j+1}(t)$  表示第  $j+1$  阶段残差信号序列.

步骤 5.重复上述步骤,如果满足 EMD 停止条件,则迭代停止,CEEMDAN 算法分解结束,原始信号序列  $x(t)$  被分解为

$$x(t) = \sum_{j=1}^N f_{\text{IMF},j} + r_N. \quad (6)$$

### 1.2 Transformer-Encoder 模型

人类大脑在处理接收到的画面信息时无法同时仔细观察眼前的一切,会选择性地聚焦某些重要且较有价值的区域,这种机制被称为注意力机制.本文将作为神经网络组件的一部分以有效地提取全局特征权重信息,其公式如下:

$$A(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \quad (7)$$

式(7)把自注意力机制描述成查询(Query,  $Q$ )和键值对(Key-Value,  $K-V$ )之间的映射关系.当 Query 和 Key 的输入维度都是  $d_k$  维即  $Q=K$  时,则为自注意力,Value 是  $d_v$  维,计算  $Q$  和  $K$  的点积结果,除以  $\sqrt{d_k}$  以简化特征维度,再使用 softmax 归一化运算求

得权重, 权重乘以  $V$  可得到当前时间步最终优化后的特征向量.

将  $n$  个自注意力模块并联组合, 能让每个模块去优化序列中不同特征部分, 运算公式如下:

$$h_i = A(QW_i^Q, KW_i^K, VW_i^V), \quad (8)$$

$$h_{\text{Mult}}(Q, K, V) = \text{concat}(h_i, \dots, h_n) \cdot W. \quad (9)$$

式中,  $W$  为需要训练的权重值.

Transformer 模型包含有 Transformer-Encoder 与 Transformer-Decoder 两个部分. 本文将编码器端引入预测模型之中作为神经网络的一层, 编码器端用于进行进一步的高层次特征提取优化, 为下一层网络提供有效的权重信息. Transformer-Encoder 架构如图 1 所示.

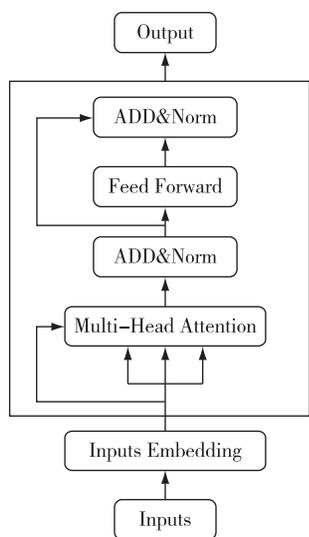


图 1 Transformer-Encoder 架构  
Fig. 1 Transformer-Encoder architecture

### 1.3 时间卷积网络 (TCN)

在基于神经网络模型处理时序问题的研究中, 通常采用经典的 RNN 结构, 但最新的研究发现 TCN 在多种任务上表现不输于 RNN. TCN 本质上是针对一维空间进行全维度卷积以处理时序问题的神经网络, 具有捕捉长期时序关系的能力, 其优点如下: 不同于 RNN 的顺序计算, TCN 的计算速度更快, 所需内存量更少, 可以进行大规模的并行处理从而缩短了神经网络训练和验证的时间; TCN 为改变卷积核感受域大小提供了更多灵活性, 通过堆叠更多的卷积层、使用更大的膨胀系数及增大滤波器大小等更好地控制模型时间记忆的长短; TCN 运用了残差连接, 避免了 RNN 中经常出现的梯度爆炸或梯度消失问题.

图 2 对 TCN 整体结构进行了说明, 图 3 则表示 TCN 中每一层的残差结构. TCN 主要包含 3 种基本结构: 因果卷积、膨胀卷积与残差连接. TCN 的每一个隐层之间使用膨胀卷积进行连接, 每一个隐层即为一个因果卷积残差模块, TCN 由多个因果卷积残差模块组成.

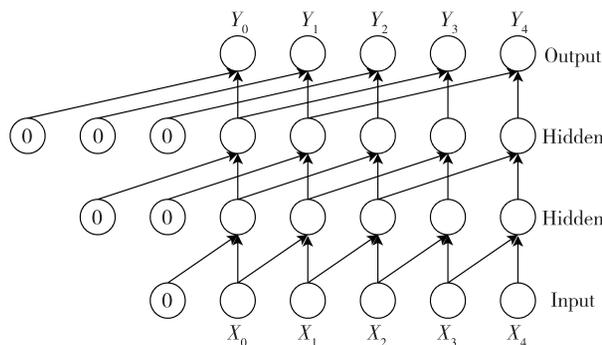


图 2 TCN 架构

Fig. 2 TCN architecture

### 1.4 Transformer-Encoder-TCN 模型

本文设计了一种新的时间卷积网络结构, 设膨胀系数为 4, 构建包含有 4 层残差块的 TCN 网络, 并在每一层因果残差块之间引入由自注意力单元构成的 Transformer-Encoder, 将每一层经因果卷积编码后的时间步特征作为其输入, 通过自注意力机制进行特征权重优化后继续作为下一层因果残差块的输入, 最后拼接一层全连接层作为预测输出, 模型如图 4 所示.

Transformer-Encoder-TCN 结构设计思路如下:

1) 在时间步特征中融合位置编码能为模型提供当前时间步的前后顺序信息. 最新的研究表明, 模型层数越深、感受域越大, 卷积操作可以在一定程度上编码目标的更多绝对位置信息<sup>[23]</sup>. 在本文提出的模型中, TCN 的因果卷积残差模块为了保证输入与输出的神经元数量一致, 在时间步中添加了 Zero-padding 进行时间步补全, 这一机制同样为下一层 Transformer-Encoder 层隐式建模了位置信息, 有助于更深层次的时序信息建模.

2) 从互补性角度看, 卷积是一种模板匹配式的局部操作运算, 不同位置采用相同的卷积核进行滤波, 一个卷积层通常只会建模邻域数据之间的关系. 而自注意力机制则属于自适应滤波式的全局操作运算, 一个 Transformer-Encoder 层具有并行化建模数据内部全局关系的能力, 但其缺少对于局部信息的先验知识, 因此卷积操作和自注意力机制可以形成

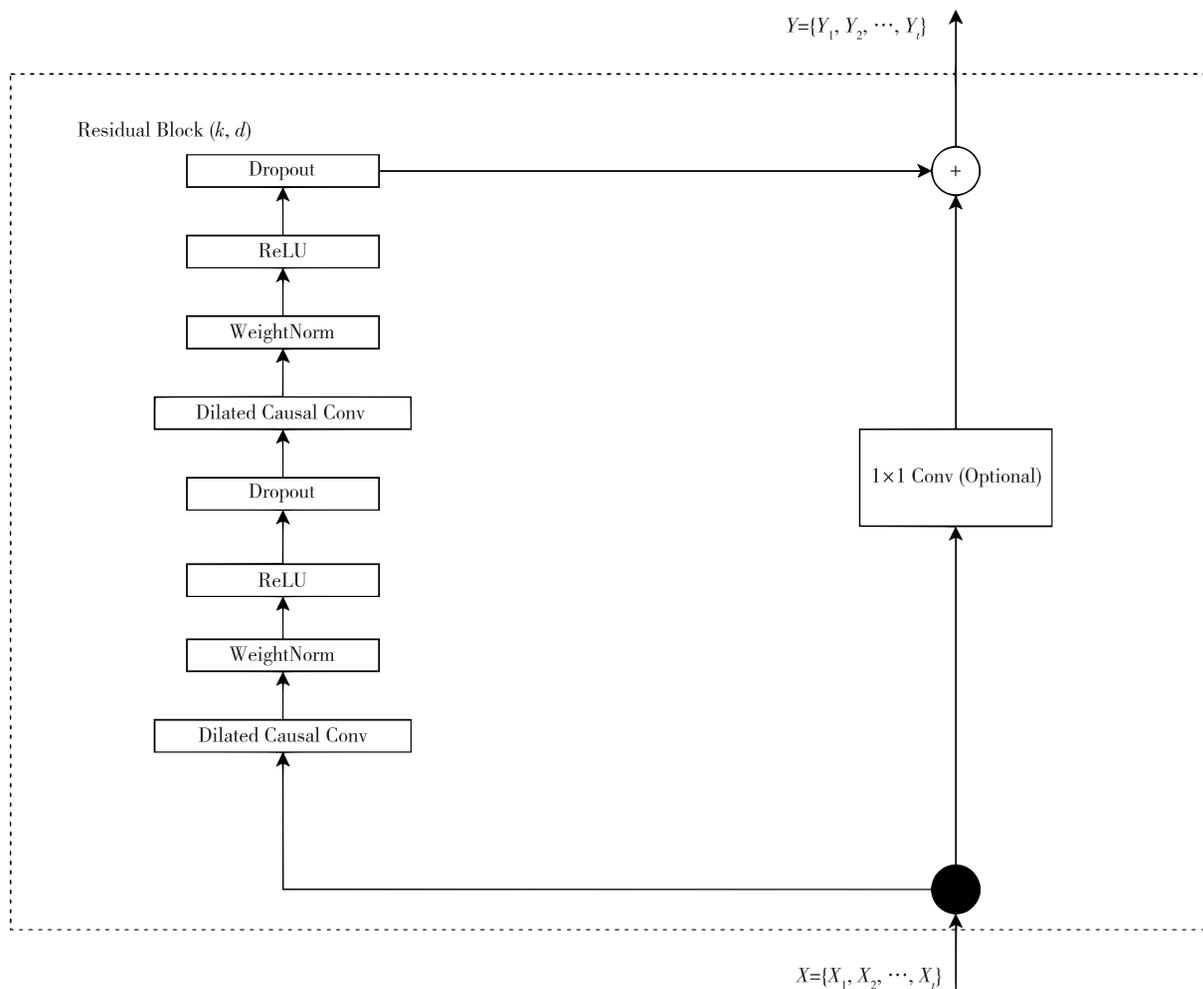


图3 因果卷积残差模块

Fig. 3 Causal convolution residual module

良好的互补<sup>[24]</sup>.

除此之外,由于农产品期货时间序列具有显著的季节性,自注意力机制在季节性特征建模中能发挥有效作用.自注意力机制能增强保留重要特征、减弱冗余特征,而季节性是一种周期性的重复特征,其往往在时序预测任务中发挥重要作用,运用自注意力机制能自动有效地增强季节性成分的建模权重,省略繁琐的季节性调整步骤;同时,由于其优越的时间维度全局加权能力,不同周期的季节性成分加权后能获得较为平稳显著的季节性序列特征.

### 1.5 本文预测模型

本文的模型框架主要包括3个步骤:1)使用CEEMDAN分解方法对时间序列进行多尺度自适应分解,得到多个高频到低频排序的IMF分量与残差;2)分解子序列的预测,使用Transformer-Encoder-TCN模型对各个IMF分量序列进行特征抽取;3)将子序

列预测值线性相加集成.本文预测模型结构如图5所示.

## 2 实验结果与分析

### 2.1 数据材料

为验证所提出模型的有效性,本文选取大豆期货作为实证研究对象,数据集爬取于中国南华期货公司的多类商品指数历史走势网站(<http://www.nanhua.net/nhzc/varietytrend.html>),该样本涵盖了2011年1月4日至2021年1月4日期间的大豆价格指数日度数据,共2428个时间观测值,并对其进行对数处理以获得2011年1月5日至2021年1月4日期间大豆期货对数收益率序列(对数收益率 $r_t = \ln \frac{x_{t+1}}{x_t}$ ).大豆对数收益率指数如图6所示,大豆价格

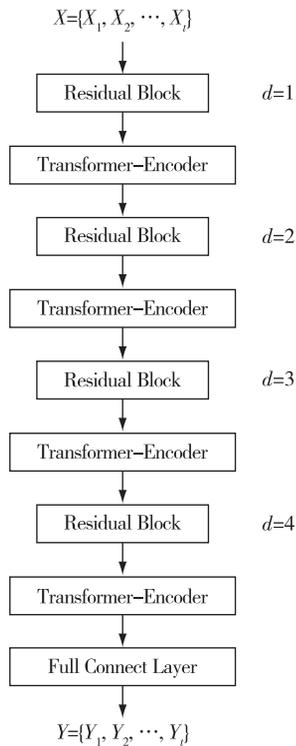


图 4 Transformer-Encoder-TCN 架构  
Fig. 4 Transformer-Encoder-TCN architecture

和大豆对数收益率指数的常规描述性统计如表 1 所示,可以观察到对数收益率指数均值明显围绕 0 附近上下波动.

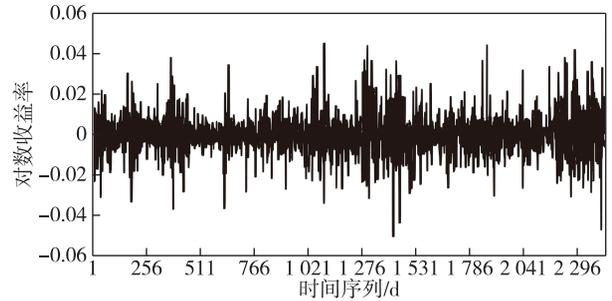


图 6 大豆对数收益率指数

Fig. 6 Logarithmic rate of return for soybeans

从经济学意义的角度上讲,由于对数收益率具备可加性,其均值能正确反映出该投资物的真实收益率,收益率的预测更具重要意义.而从计算建模的角度看,对数收益率易于统计处理的性质给模型构建带来了巨大的便利性,取对数可以让数据更加平稳,削弱了原始价格变量的异方差与共线性.由图 7 可知,对数收益率序列围绕着 0 均值上下波动,数据

表 1 大豆期货序列描述性统计结果

Table 1 Descriptive statistical results of soybean future series

	最大值	最小值	均值	标准差	偏度	峰度
价格	5 766.916 8 元	3 250.485 7 元	4 163.531 6 元	481.164 8 元	0.068 5	-0.666 5
对数收益率	4.521 3%	-5.030 6%	0.009 8%	0.922 9%	0.221 5	3.510 2

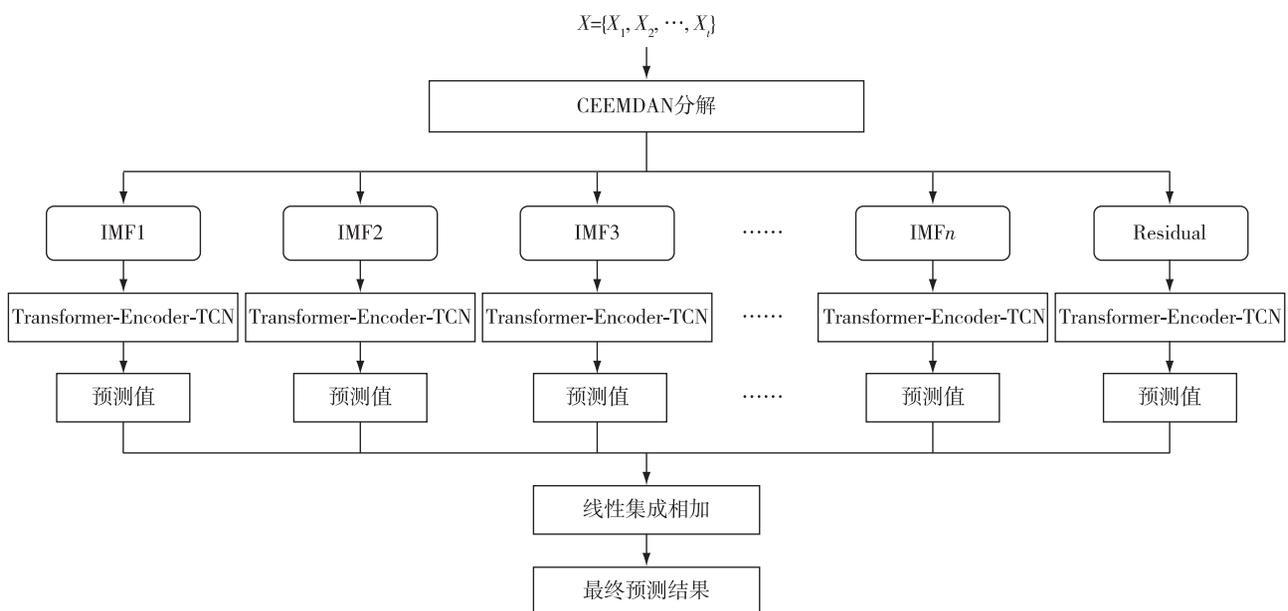


图 5 分解集成预测模型结构

Fig. 5 Structure of the proposed decomposition and ensemble prediction model

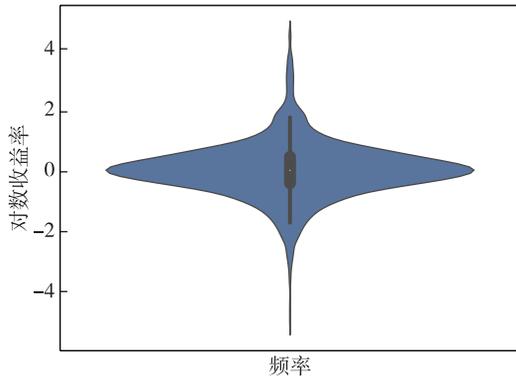


图7 对数收益率序列小提琴图

Fig. 7 Violin plot of logarithmic rate of return series

呈正态分布,具有波动聚集性.

此外,本文通过单位根检验(Augmented Dickey-Fuller test, ADF)进行平稳性检验,判断序列是否存在单位根:如果序列平稳,则单位根不存在;反之,序列不平稳即存在单位根.原假设存在单位根,如果得到的显著性检验统计量小于5%显著性水平下的临界值,则对应有95%的把握来拒绝原假设,说明序列平稳具有显著性.检验结果如表2所示,对数收益率序列在显著性水平为5%的情况下平稳性远高于价格序列.

表2 ADF 平稳性检验结果

Table 2 ADF stationarity test results

	统计量	5%临界值	结论
大豆价格	-0.716 567	-2.862 732	非平稳
大豆对数收益率	-49.084 466		平稳

## 2.2 预测过程

为提高预测模型的泛化能力,缓解时序预测中常出现的过拟合与概念漂移问题,本文采用时间序列交叉验证中的前向滚动窗口方法<sup>[25]</sup>,通过分段预测<sup>[26]</sup>并求误差均值来测试模型的整体预测效果.为了保证模型能学习到较长时间段内的规律,在每一个预测周期内,用过去3年的数据来构造输入序列,将时序排序的前90%数据作为训练集,并使用CEEMDAN进行序列逐级分解,得到当前预测周期由高频到低频排序的多尺度分量.

序列分解结果如图8所示,红色序列为原收益率序列,蓝色序列为IMF分量与残差,可以观察到序列被分解为高中低三类频级的分量.残差代表着收益率的长期发展趋势,低频项波动幅度代表重大事件对收益率的影响期限长短,期限越长期期货市场受

影响时间越久,高频项则代表短期期货市场受突发事件和供需事件引起的收益剧烈波动.

本文选择迁移学习式的训练策略,模型依次对当前周期得到的多尺度分量进行特征趋势提取,当上一个频率分量训练结束后,模型在上一个频率分量训练得到的参数基础上不断进行重训练,这种在预训练模型上进行参数微调(fine-tune)的训练方式有利于更好地捕捉不同时期的时序规律,加快模型训练收敛速度.而过去3年里最后10%数据则作为验证集,并将在验证集上预测效果最好的模型权重保存,接着使用未来6个月数据构造的测试集来检验模型预测效果.

在下一轮周期验证中,将构造的训练集、验证集与测试集整体往后推移6个月,以此类推,直至覆盖全部10a样本,最后将分段训练得到的模型预测误差求平均.交叉验证分段预测过程如图9所示.

## 2.3 参数设置

超参数的合理设置对于模型效果具有至关重要的影响.由于本文的关键改进在神经网络模型结构创新,因此CEEMDAN分解方法仅采用py-EMD库中的内置默认参数设置,经过多次实验,TCN中因果卷积残差模块层数设置为4,各层卷积核数量分别设置为20、15、10、5,卷积核大小设置为2,膨胀因子则为 $2^n$ , $n$ 为残差块层数,而Transformer-Encoder中blocks设置为1,heads设置为5,即五头自注意力.

为了构建适合TCN输入格式的数据集,本文将时间滑动窗口设置为10,即使用过去10天的信息来预测下一个交易日的收益率,批大小(batch\_size)设置为128,使用均方误差(MSE)作为损失函数,选择Adam算法进行梯度训练,初始学习速率设置为0.001,训练过程中使用余弦退火算法进行学习率调节,最大训练轮数设置为5000,当连续5轮验证集误差不再下降时使用早停机制停止训练.

## 2.4 评价指标

本文对模型的评价指标从两个维度来选取:第1个维度采用均方根误差(Root Mean Squared Error, RMSE)和平均绝对误差(Mean Absolute Error, MAE)来衡量真实值和预测值之间的误差,其值越小越好;第2个维度是方向性命中率(Directional Symmetry, DS),这一指标能衡量模型预测走势的精确度,取值范围为0至1,其值越接近1越好.

$$\sigma_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{i+1} - \hat{x}_{i+1})^2}, \quad (10)$$

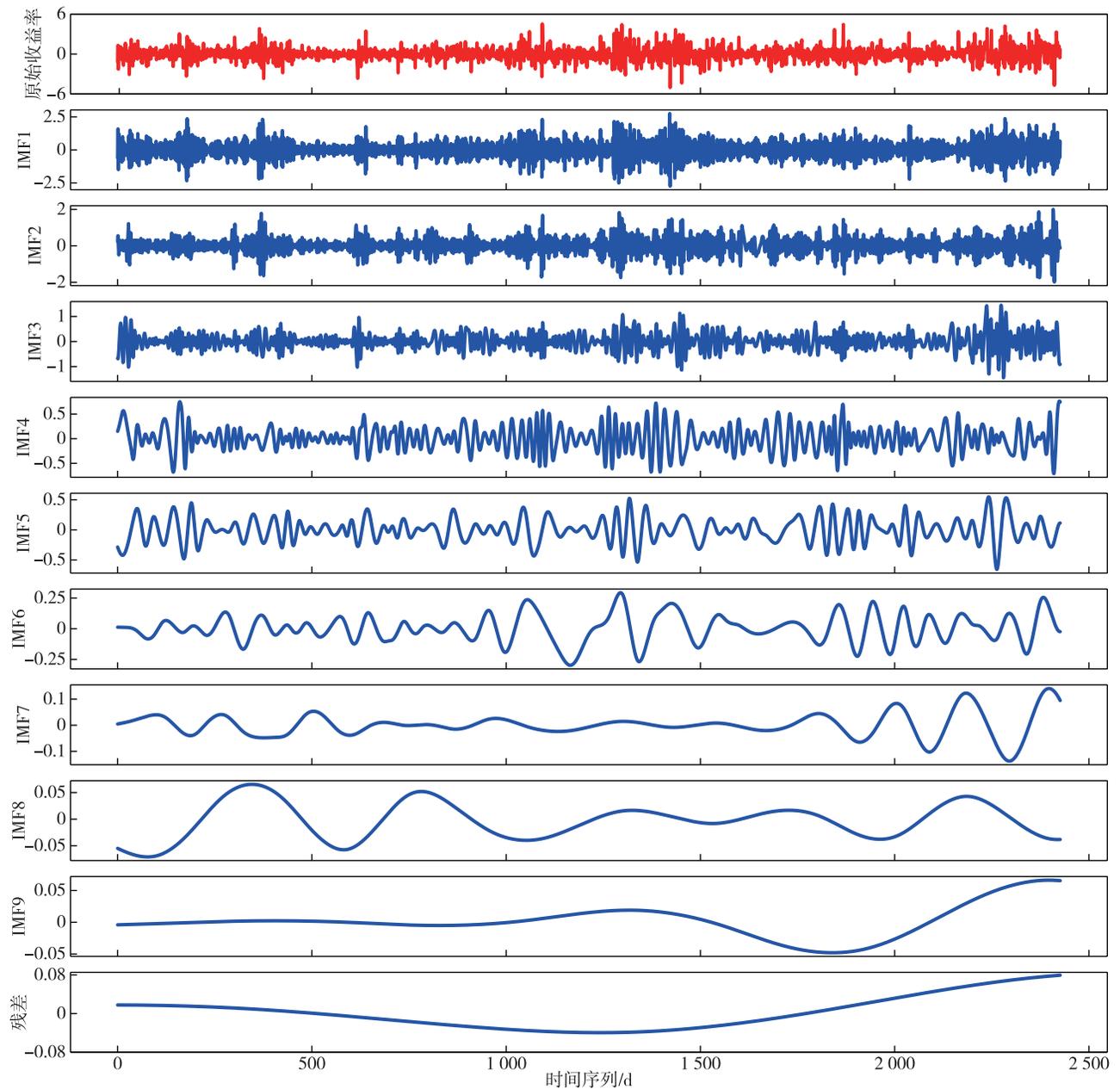


图 8 收益率序列分解

Fig. 8 Exploded plots of the rate of return series

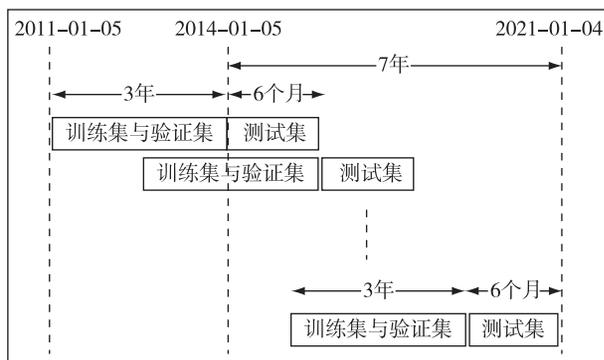


图 9 交叉验证分段预测过程

Fig. 9 Cross-validation segmented forecasting process

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

$$S_D = \frac{1}{n} \sum_i u_i,$$

$$u_i = \begin{cases} 1, & (x_{t+1} - x_t)(\hat{x}_{t+1} - \hat{x}_t) > 0; \\ 0, & (x_{t+1} - x_t)(\hat{x}_{t+1} - \hat{x}_t) \leq 0. \end{cases} \quad (12)$$

其中:  $n$  为样本数量;  $x_t$  和  $\hat{x}_t$  分别为第  $t$  天的真实值和预测值;  $x_{t+1}$  和  $\hat{x}_{t+1}$  分别为第  $t+1$  天的真实值和预测值.

## 2.5 消融实验

本文模型主要由 CEEMDAN、TransformerEncoder 与 TCN 3 个核心组件构造.为了验证本文模型的有效性,对其进行消融实验,设计了如下 4 种组合方式:

- 1) TCN, 记为模型 1;
- 2) Transformer-Encoder-TCN, 记为模型 2;
- 3) CEEMDAN-TCN, 记为模型 3;
- 4) CEEMDAN-Transformer-Encoder-TCN, 记为本文模型.

由表 3 消融实验的结果可知,本文所提出的模型通过多个周期交叉验证后在 3 个评价指标上均取得了最佳预测效果.模型 1 作为基线模型,预测精度较低.模型 2 由于添加了自注意力单元模块,建模了全局时序信息后在 RMSE 与 MAE 两个指标上预测精度比起单一的 TCN 模型略有提高,但受限于数据复杂的全局波动性,DS 精度有所降低.而模型 3 由于采用了先分解再集成的预测策略,依据分而治之的思想降低了预测复杂度,比起单一 TCN 模型预测效果有了非常显著的提升.本文模型整合了分解集成策略与自注意力单元各自的优势,通过分解算法得到多尺度分量后,自注意力单元更容易聚焦于不同频率的显著特征,比起单一 TCN 模型, RMSE 误差指标降低了 33.7%, MAE 误差指标降低了 31.5%, DS 精度提升了 43.4%.

表 3 消融实验结果

Table 3 Ablation experiment results

模型	RMSE	MAE	DS
模型 1	0.996 7	0.736 2	0.517 7
模型 2	0.936 9	0.681 4	0.419 7
模型 3	0.802 8	0.617 3	0.662 9
本文模型	0.660 5	0.504 1	0.742 6

为了更直观地观测本文模型的预测效果,由于篇幅限制,仅取多折交叉验证中最后一折的测试集作出预测,结果如图 10 所示,黑色实线代表真实值,灰色虚线代表预测值,本文模型在一定程度上能较好地拟合真实值,但更倾向于保守预测,少数异常波动收益率点的预测结果具有滞后性,这符合期货投资收益率中宁可少赚也不多亏的原则.以上实验结果表明,本文提出的预测模型改进方式具有一定可行性.

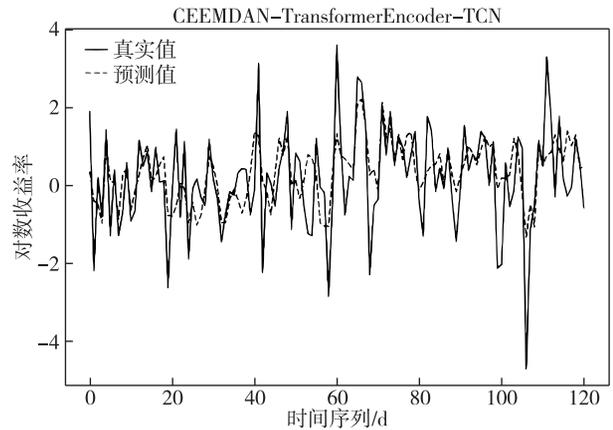


图 10 部分预测结果

Fig. 10 Some prediction results

## 2.6 对比实验

为了进一步验证本文所提出模型的优越性,将其与 LSTM、CEEMDAN-LSTM、SVR、CEEMDAN-SVR 模型进行对比实验.为了保证实验公平性,对比模型预测流程与本文模型一致.其中:LSTM 隐藏层神经元数量根据经验与滑动窗口长度设置为 10,隐藏层数设置为 1,隐藏层激活函数设置为 ReLU,其余训练参数如学习速率等则与本文模型一致;SVR 模型使用 sklearn 库中的默认参数,核函数为 RBF.对比实验结果如表 4 所示.

表 4 对比实验结果

Table 4 Comparison of experiment results

模型	RMSE	MAE	DS
LSTM	0.937 5	0.679 5	0.498 2
CEEMDAN-LSTM	0.676 6	0.532 8	0.702 4
SVR	0.960 9	0.709 6	0.497 6
CEEMDAN-SVR	0.741 9	0.567 6	0.702 4
本文模型	0.660 5	0.504 1	0.742 6

由表 4 可知,LSTM 作为深度学习模型的一种,能建模序列时间尺度上的依赖关系,其独特的门控记忆机制使得其在序列建模任务上比起传统机器学习模型 SVR 更具优势,而两者在应用分解集成策略将复杂序列转换为简单子序列后,从 3 个评价指标结果上看预测精度均获得显著提升,验证了分解策略的有效性.

4 种对比模型中,LSTM 模型容易陷入梯度弥散或爆炸的困境,SVR 模型则无法很好地捕捉时间尺度上的依赖关系,而本文模型在应用分解集成策略的基础上同时融合了局部信息与全局信息进行建

模,避免了梯度消失或爆炸问题,因果卷积的 Zero-padding 机制与自注意力单元形成互补,较好地建模了农产品期货季节性成分,在误差精度和方向性精度上均展示出较强优势.本文模型改进了大豆期货收益率短期预测效果,在农产品期货预测上具有一定潜力.

### 3 结论

针对农产品期货预测精度低的问题,本文提出一种基于 CEEMDAN 与 Transformer-Encoder-TCN 模型的分解集成预测方法,并在真实大豆期货收益率数据集上进行实验.CEEMDAN 通过挖掘序列多尺度特征,降低了序列建模复杂度,TCN 模型的 Zero-padding 机制为自注意力单元抽象表达了时序数据中的绝对位置信息,而自注意力单元 Transformer-Encoder 则有效地增强了 TCN 模型的全局关联建模能力,优化了显著特征的建模权重.实验通过时间序列交叉验证与参数迁移训练的方式进行,研究结果表明本文模型在农产品期货收益率预测研究中具备一定可行性.

本文仍存在一些不足之处,CEEMDAN 分解后的子序列数量不确定,其中高频子序列仍存在较大复杂性,这增加了建模难度.另外,由于 CEEMDAN 的计算时间较长,因此本文模型更侧重于离线计算场景,对于众多农产品期货在线高频交易场景的需求无法满足,未来研究将重点将针对其他农产品期货高频交易场景下的预测问题进行扩展研究.

### 参考文献

#### References

- [ 1 ] 刘洋,罗其友,周振亚,等.我国主要农产品供需分析与预测[J].中国工程科学,2018,20(5):120-127  
LIU Yang, LUO Qiyu, ZHOU Zhenya, et al. Analysis and prediction of the supply and demand of China's major agricultural products[J]. Strategic Study of CAE, 2018, 20(5):120-127
- [ 2 ] Han Z Y, Zhao J, Leung H, et al. A review of deep learning models for time series prediction[J]. IEEE Sensors Journal, 2021, 21(6):7833-7848
- [ 3 ] 张永安,颜斌斌.一种股票市场的深度学习复合预测模型[J].计算机科学,2020,47(11):255-267  
ZHANG Yongan, YAN Binbin. Deep learning hybrid forecasting model for stock market[J]. Computer Science, 2020, 47(11):255-267
- [ 4 ] 赵娜,孙红,黎铨祺,等.基于小波分解的时序预测模型 mWDLNet 及其应用研究[J].小型微型计算机系统,2022,43(3):561-567  
ZHAO Na, SUN Hong, LI Quanqi, et al. Research on ap-

- plication of time series forecast model mWDLNet based on wavelet decomposition [J]. Journal of Chinese Computer Systems, 2022, 43(3):561-567
- [ 5 ] Wang S, Mu L, Liu D R. A hybrid approach for El Niño prediction based on empirical mode decomposition and convolutional LSTM encoder-decoder [J]. Computers & Geosciences, 2021, 149:104695
- [ 6 ] Wang Y N, Yuan Z, Liu H Q, et al. A new scheme for probabilistic forecasting with an ensemble model based on CEEMDAN and AM-MCMC and its application in precipitation forecasting [J]. Expert Systems with Applications, 2022, 187:115872
- [ 7 ] 唐振鹏,吴俊传,张婷婷,等.基于二次分解和集成学习的粮食期货价格预测研究[J].系统工程理论与实践,2021,41(11):2837-2849  
TANG Zhenpeng, WU Junchuan, ZHANG Tingting, et al. Research on grain futures price forecasting based on secondary decomposition and ensemble learning[J]. Systems Engineering Theory & Practice, 2021, 41(11):2837-2849
- [ 8 ] Zhou Y R, Li T Y, Shi J Y, et al. A CEEMDAN and XG-BOOST-based approach to forecast crude oil prices [J]. Complexity, 2019, 2019:1-15
- [ 9 ] 韩金磊,熊萍萍,孙继红.基于 LSTM 和灰色模型的股价时间序列预测研究[J].南京信息工程大学学报(自然科学版),2023,15(6):631-642  
HAN Jinlei, XIONG Pingping, SUN Jihong. Stock price time series prediction based on LSTM and grey model [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(6):631-642
- [ 10 ] Bai S, Kolter J Z, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling [J]. arXiv e-print, 2018, arXiv:1803.01271
- [ 11 ] Li G Q, Chen W, Li D H, et al. Comparative study of short-term forecasting methods for soybean oil futures based on LSTM, SVR, ES and wavelet transformation [J]. Journal of Physics: Conference Series, 2020, 1682(1):012007
- [ 12 ] Cheng W, Wang Y, Peng Z, et al. High-efficiency chaotic time series prediction based on time convolution neural network [J]. Chaos, Solitons & Fractals, 2021, 152:111304
- [ 13 ] Wang C H, Gao Q. High and low prices prediction of soybean futures with LSTM neural network [C]//2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS). November 23-25, 2018, Beijing, China. IEEE, 2019:140-143
- [ 14 ] Chuluunsaikhan T, Ryu G A, Yoo K H, et al. Incorporating deep learning and news topic modeling for forecasting pork prices: the case of the republic Korea [J]. Agriculture, 2020, 10(11):513
- [ 15 ] 张杰,甄柳琳,徐硕,等.融合传递熵的图神经网络农产品期货预测模型[J].计算机工程与应用,2023,59(2):321-328  
ZHANG Jie, ZHEN Liulin, XU Shuo, et al. Graph neural network model based on transfer entropy for agricultural futures forecasting [J]. Computer Engineering and Appli-

- cations, 2023, 59(2):321-328
- [16] Wei X L, Lei B B, Ouyang H B, et al. Stock index prices prediction via temporal pattern attention and long-short-term memory [J]. *Advances in Multimedia*, 2020, 2020: 1-7
- [17] Mnih V, Heess N, Graves A, et al. Recurrent models of visual attention [C]// *Advances in Neural Information Processing Systems*, 2014: 2204-2212
- [18] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [C]// *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017: 6000-6010
- [19] Niu H L, Xu K L, Liu C. A decomposition-ensemble model with regrouping method and attention-based gated recurrent unit network for energy price prediction [J]. *Energy*, 2021, 231: 120941
- [20] 贾宁, 郑纯军. 基于 LSTM-DA 神经网络的农产品价格指数短期预测模型 [J]. *计算机科学*, 2019, 46(增刊 2): 62-65, 71  
JIA Ning, ZHENG Chunjun. Short-term forecasting model of agricultural product price index based on LSTM-DA neural network [J]. *Computer Science*, 2019, 46(sup2): 62-65, 71
- [21] Yin H L, Jin D, Gu Y H, et al. STL-ATTLSTM; vegetable price forecasting using STL and attention mechanism-based LSTM [J]. *Agriculture*, 2020, 10(12): 612
- [22] Torres M E, Colominas M A, Schlotthauer G, et al. A complete ensemble empirical mode decomposition with adaptive noise [C]// *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. May 22 - 27, 2011, Prague, Czech Republic. IEEE, 2011: 4144-4147
- [23] Islam M A, Jia S, Bruce N D B. How much position information do convolutional neural networks encode? [J]. *arXiv e-print*, 2020, arXiv: 2001.08248
- [24] Wang X L, Girshick R, Gupta A, et al. Non-local neural networks [C]// *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18 - 23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7794-7803
- [25] Nair B B, Mohandas V P, Sakthivel N R. A decision tree-rough set hybrid system for stock market trend prediction [J]. *International Journal of Computer Applications*, 2010, 6(9): 1-6
- [26] M'Ng J C P, Mehralizadeh M. Forecasting East Asian indices futures via a novel hybrid of wavelet-PCA denoising and artificial neural network models [J]. *PLoS One*, 2016, 11(6): e0156338

## Forecasting agricultural commodity futures with decomposition and ensemble strategy based on attentional temporal convolution network

ZHANG Dabin<sup>1</sup> HUANG Junjie<sup>1</sup> LING Liwen<sup>1</sup> LIN Ruibin<sup>1</sup>

<sup>1</sup> College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China

**Abstract** To address the low prediction accuracy in agricultural commodity futures due to their nonlinear and non-smooth features resulting from various influencing factors, this paper proposes a decomposition and ensemble forecasting approach based on CEEMDAN and Transformer-Encoder-TCN. First, the Complementary Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) is used to decompose the time series into multiscale Intrinsic Mode Function (IMF) and residuals, reducing the complexity of series modeling. Second, each subseries is predicted via Temporal Convolutional Network (TCN) incorporating multi-stage self-attention unit (Transformer-Encoder), which optimizes the modeling weights of significant features. Finally, the prediction results of each subseries are linearly summed and integrated to obtain the final prediction results. The soybean futures revenue index in the agricultural commodity index of South China Futures Company is used as the research object. The model is retrained by time-series cross-validation and parameter transfer. The ablation and comparison experimental results show that the proposed model has superiority in RMSE, MAE and DS, verifying its effectiveness in predicting agricultural commodity futures.

**Key words** agricultural commodity futures; complementary ensemble empirical mode decomposition with adaptive noise (CEEMDAN); self-attention; Transformer-Encoder; temporal convolution network (TCN)