DOI:10.19652/j.cnki.femt.2204479

基于二次分解和GRU-attention的时间序列预测研究*

高凯悦 牟 莉
(西安工程大学计算机科学学院 西安 710600)

摘 要:针对时间序列规律难以捕捉且具有高度非平稳性特征导致的预测精度较低问题,提出了一种基于二次分解和注意力机制优化门控循环单元(GRU-attention)的时间序列预测模型。首先利用完全集合经验模态分解(complete ensemble empirical mode decomposition with adaptive noise,CEEMDAN)将时间序列分解为若干个特征互异的模态分量,并依据样本熵量化各分量复杂度。其次采用变分模态分解(variational modal decomposition,VMD)弱化高熵值分量的非平稳性特征。接着使用注意力机制优化GRU预测模型。最后对各分量建立GRU-attention模型进行预测,将各分量预测结果叠加获得最终结果。通过实验分析证明,所提出的模型与其他模型相比能够较好的捕捉序列的复杂规律、降低序列的非平稳性并且具有较高的预测性能,其平均绝对百分比误差达到了2.9%,决定系数达到了0.891。

关键词:CEEMDAN-VMD二次分解;样本熵;GRU神经网络;Attention机制;时间序列预测
中图分类号:TP393 **文献标识码**:A **国家标准学科分类代码**:520.3040

Research on time series forecasting based on quadratic decomposition and GRU-attention

Gao Kaiyue Mu Li
(School of Computer Science, Xi′an Polytechnic University, Xi′an 710600, China)

Abstract: Aiming at the problem of low prediction accuracy caused by the difficulty of capturing time series laws and the highly non-stationary characteristics, this paper proposes a time series prediction model based on quadratic decomposition and GRU-attention. First, the complete ensemble empirical mode decomposition (CEEMDAN) is used to decompose the time series into several modal components with different characteristics, and the complexity of each component is quantified according to the sample entropy. Secondly, variational mode decomposition (VMD) is used to weaken the non-stationarity characteristics of high entropy components. The GRU prediction model is then optimized using the attention mechanism. Finally, a GRU-attention model is established for each component to predict, and the prediction results of each component are added to obtain the final result. The experimental analysis proves that compared with other models, the model proposed in this paper can better capture the complex laws of the sequence, reduce the non-stationarity of the sequence, and have higher prediction performance, and its average absolute percentage error reached 2.9%, the coefficient of determination reached 0.891.

Keywords: CEEMDAN-VMD secondary decomposition; sample entropy; GRU neural networks; attention mechanism; time series forecast

0 引 言

时间序列是指按相同时间间隔顺序对某变量进行观测并记录得到的一组序列值。时间序列预测奠定着数据挖掘和决策的基础,可以有效提高管理者与决策者判断的正确性,在金融、气象、医学等多个领域中具有广泛应用。然而随着通信、计算机技术的成熟,时间序列的数据量成倍增加,导致数据变化规律逐渐复杂、非平稳性逐渐增高,使时间序列预测变得十分困难。

收稿日期:2022-11-06
* **基金项目**:陕西省2020年技术创新引导专项(基金)计划(2020CGXNG-026)项目资助

在早期研究中，国内外学者多数基于传统统计学方法预测时间序列。王英伟等[1]使用差分整合移动平均自回归模型(autoregressive integrated moving average model，ARIMA)实现了对时间序列的预测、顾海燕[2]基于自回归滑动平均模型(autoregressive moving average model，ARMA)实现了对中国大陆人口的预测，但这些传统统计学方法的优势只限于线性关系模型，对数据有较高限制。基于机器学习的方法可以良好改善传统统计学方法的缺陷，王毅等[3]利用人工神经网络(artificial neural network，ANN)实现了对河西走廊地区生态安全的预测、刘斌等[4]采用支持向量机(support vector machine，SVM)实现了对变电站蓄电池退化趋势的预测，但由于机器学习方法对高频序列学习效果不佳，所以在预测复杂时间序列时学习效果未有显著提升。近年来，基于深度学习的方法在预测时间序列的领域中得到了广泛应用，吴琼等[5]使用BP神经网络实现了对锂离子电池容量进行预测的预测，该模型的出现提高了时间序列预测的准确性和实时性，但其存在易陷入局部最优值且无法获取时间序列记忆特征的缺陷。程换新等[6]采用循环神经网络(recurrent neural networks，RNN)实现了对短期电力负荷的预测，证明了该模型在时序预测方面的有效性，但其在预测较长序列时易产生梯度爆炸的现象。陈安民等[7]使用长短期记忆网络(long short-term memory，LSTM)实现了对无人机认知无线电频谱的预测，门控单元的加入使模型能够更完整有效的学习时间序列中的变化规律，减缓了RNN神经网络梯度爆炸问题。崔立志等[8]采用门控循环单元(gated recurrent unit，GRU)实现了对在线学生绩效的预测，该模型通过对LSTM模型门控结构的简化，使模型在保留LSTM模型原预测效果的同时变得更加简易。

针对时间序列存在的复杂性、非平稳性等特征，数据分解与深度学习结合的方法逐渐应用到时间序列预测中。早期的数据分解方法有小波变换(wavelet transform，WT)[9]，该方法可消除噪声对序列的干扰，但分解时需人工选择小波基，易导致预测误差。李孟威等[10]使用经验模态分解(empirical mode decomposition，EMD)实现了对受热面积灰的预测，该方法避免了小波分解的缺陷，但它存在模态混叠和边界效应的问题。集成经验模态分解(ensemble empirical mode decomposition，EEMD)[11]是一种叠加高斯白噪声的多次EMD分解，该方法有效降低了模态混叠对预测结果造成的影响，但不适合的高斯白噪声幅值会造成重建误差的现象。完全集合经验模态分解(complete ensemble empirical mode decomposition with adaptive noise，CEEMDAN)[12]通过减少EEMD的迭代次数，改善了其重建误差的缺陷，显著提高了后续建模分析的效率。

鉴于数据分解结合深度学习的优势，本文提出一种基于二次分解的GRU-Attention时间序列预测模型。首先，将时间序列通过CEEMDAN分解为若干模态分量，并计算各分量的样本熵。其次采用变分模态分解(variational modal decomposition，VMD)对高频分量二次分解，削弱其非线性、不平稳性特征。最后对各分量作为经Attention机制优化的GRU神经网络的输入序列进行预测，叠加各分量预测值获取最终预测结果。通过算例分析，证明本文提出的基于二次分解和GRU-Attention的时间序列预测模型较其他模型具有更高的精确度。

1 模态分解

1.1 CEEMDAN

CEEMDAN可以自适应分解出规律较强且在不同时间尺度上频率从高到低的本征模态分量和残差余量[13]供后续模型更高效率的提取序列特征，以提高模型预测精度。通过加入正负白噪声重构信号，不仅可以较清晰的区分出相近模态频率，解决EMD的模态混叠问题，还可以在分解过程中相互抵消噪声，减少重构误差[14]。设时间序列为 $x(t)$，则CEEMDAN的算法步骤如下。

1)将时间序列 $x(t)$ 添加 I 次自适应白噪声，则得到第 i 次信号 $x_i(t)$：

$$x_i(t)=x(t)+\varepsilon\omega_i(t)\quad i=1,2,3,\cdots,I \tag{1}$$

式中：ε 为高斯白噪声幅值系数；$\omega_i(t)$ 为第 i 次实验处理时的高斯白噪声；I 为实验次数。

2)对 $x_i(t)$ 采用EMD分解，得到CEEMDAN分解后初次的模态分量 $IMF_1(t)$ 和余量信号 $r_1(t)$：

$$IMF_1(t)=\frac{1}{I}\sum_{i=1}^{I}IMF_{1i}(t) \tag{2}$$

$$r_1(t)=x(t)-IMF_1(t) \tag{3}$$

3)将式(3)得到的 $r_1(t)$ 再次通过 $r_1(t)+\varepsilon_1E_1(\omega_i(t))$ 计算，得到 $IMF_2(t)$、$r_2(t)$，其中 $E_1(\cdot)$ 为 $IMF_1(t)$ 的算子。

4)重复步骤2)和3)，若 $r_i(t)$ 达到最大分解数量则停止迭代，最终的时序信号可表示为：

$$x(t)=\sum_{i=1}^{I}IMF_i(t)+r_i(t) \tag{4}$$

1.2 VMD

VMD在处理非平稳信号上具备显著优势[15]。对复杂程度较高的分量进行VMD分解，可减弱时间序列的复杂程度及非平稳性特征[16]。设 $f(t)$ 表示时间序列信号，分解步骤如下。

1)采用Hillbert得到每个模态分量的解析信号及单边品牌，再引入 $e^{-j\omega_k t}$ 各信号频谱调整至对应的基带上。

2)使用H1高斯平滑估算每个模态 $u_k(t)$ 的带宽：

$$\min_{\{u_k\},\{\omega_k\}}\left\{\sum_{k=1}^{K}\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)u_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right\} \tag{5}$$

$$\text{s.t.}\quad\sum_{k=1}^{K}u_k(t)=f(t) \tag{6}$$

式中：$\{u_k\}$ 为时序序列分解后的模态；$\{\omega_k\}$ 为模态对应的中心频率；K 为分解个数。

3)将式(5)、(6)转化为非约束的变分问题：

$$L(\{u_k\},\{\omega_k\},\lambda)=\alpha\sum_{k=1}^{K}\left\|\partial_t\left[\left(\delta(t)+\frac{\mathrm{j}}{\pi t}\right)u_k(t)\right]\mathrm{e}^{-\mathrm{j}\omega_k t}\right\|_2^2+a\left\|f(t)-\sum_{k=1}^{K}u_k\right\|_2^2+\left\langle\lambda(t),f(t)-\sum_{k=1}^{K}u_k(t)\right\rangle \tag{7}$$

4)对式(7)求解得到 u_k^{n+1}、ω_k^{n+1}：

$$\hat{u}_k^{n+1}(\omega)=\frac{\hat{f}(\omega)-\sum_{i\neq k}\hat{u}_i^n(\omega)+\frac{\hat{\lambda}^n(\omega)}{2}}{1+2\alpha(\omega-\omega_k^n)^2} \tag{8}$$

$$\omega_k^{n+1}=\frac{\int_0^\infty\omega|\hat{u}_k^{n+1}(\omega)|^2\mathrm{d}\omega}{\int_0^\infty|\hat{u}_k^{n+1}(\omega)|^2\mathrm{d}\omega} \tag{9}$$

式中：$\hat{u}_k^{n+1}(\omega)$ 为 $\hat{f}(\omega)-\sum_{i\neq k}\hat{u}_i^n(\omega)$ 的维纳滤波；$\hat{u}_i(\omega)$、$\hat{f}(\omega)$、$\hat{\lambda}(\omega)$、$\hat{u}_k^{n+1}(\omega)$ 为 $u_i(t)$、$f(t)$、$\lambda(t)$、$u_k^{n+1}(t)$ 傅里叶变换后的值；ω_k^{n+1} 为当前模态函数的功率谱重心；n 为迭代求解次数。

1.3 样本熵

当样本熵值越大，表示信号自我相似越低且分量复杂程度越高[17-18]。

1)采集的时间序列包含 N 个数据 $\{x(n)\}=x(1),x(2),x(3),\cdots,x(N)$，对数据进行 m 维的向量序列重构排序，即：

$$\boldsymbol{X}_m(i)=[x(i),x(i+1),\cdots,x(i+m-1)] \tag{10}$$

2)当 $1\leqslant i\leqslant N-m+1$ 时，计算向量 $\boldsymbol{X}_m(i)$ 与 $\boldsymbol{X}_m(j)$ 各元素绝对值的差。记为时序序列中绝对值相差最大的距离为 $d[\boldsymbol{X}_m(i),\boldsymbol{X}_m(j)]$：

$$d[\boldsymbol{X}_m(i),\boldsymbol{X}_m(j)]=\max(|x(i+k)-x(j+k)|) \tag{11}$$

3)统计满足条件阈值 $r\geqslant d[\boldsymbol{X}_m(i),\boldsymbol{X}_m(j)]$ 的向量个数 B_i 和 $N-m$ 的比值 $B_i^m(r)$，再求取均值 $B^m(r)$：

$$B_i^m(r)=\frac{B_i}{N-m} \tag{12}$$

$$B^m(r)=\frac{\sum_{i=1}^{N-m+1}B_i^m(r)}{N-m+1} \tag{13}$$

4)增加维度到 $m+1$，统计满足条件阈值 $r\geqslant d[\boldsymbol{X}_{m+1}(i),\boldsymbol{X}_{m+1}(j)]$ 的个数 A_i 和 $N-m$ 的比值 $A_i^m(r)$，再求取均值 $A^m(r)$：

$$A_i^m(r)=\frac{A_i}{N-m} \tag{14}$$

$$A^m(r)=\frac{\sum_{i=1}^{N-m+1}A_i^m(r)}{N-m+1} \tag{15}$$

5)计算样本熵：

$$SampEn(m,r)=\lim_{N\to\infty}\left[-\ln\frac{A^m(r)}{B^m(r)}\right] \tag{16}$$

样本熵值和嵌入维数 m、阈值 r 两个参数相关，一般情况下 $m=1$ 或 2，$r=0.1SD\sim0.25SD$（SD 为时间序列 $x(N)$）。

2 预测模型

2.1 GRU 神经网络

传统预测模型在预测时往往忽略了时间序列具有较强的时序性和周期性的特征。深度学习模型中 LSTM 神经网络能够良好的对较长序列进行长期记忆并解决普通神经网络易出现的梯度爆炸问题，但其内部结构较为复杂，预测时间较长。GRU 神经网络在 LSTM 神经网络的基础上引入了重置门和更新门，与 LSTM 神经网络相比其训练参数更少，结构更加简单[19]，其具体结构如图 1 所示。

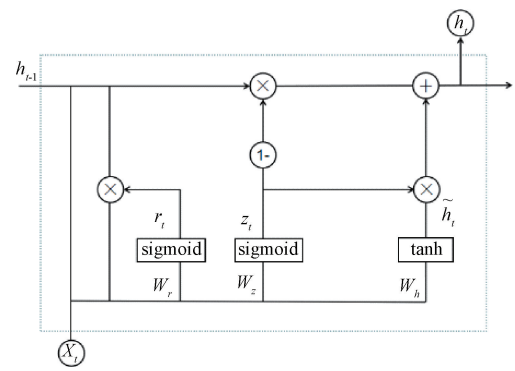

图 1 GRU 神经网络结构

GRU 神经网络单元的运行原理首先是通过式(17)和(18)计算重置门的值以及候选状态。接着通过式(19)使更新门控制特征信息需要传递多少信息至当前隐藏状态，实现了单元的长期记忆功能。最后通过式(20)计算出当前隐藏状态的最终值。

$$r_t=\mathrm{sigmoid}(\boldsymbol{W}_r\cdot[h_{t-1},X_t]+b_r) \tag{17}$$

$$\tilde{h}_t=\tanh(\boldsymbol{W}_h\cdot[r\cdot h_{t-1},X_t]+b_h) \tag{18}$$

$$z_t=\mathrm{sigmoid}(\boldsymbol{W}_z\cdot[h_{t-1},X_t]+b_z) \tag{19}$$

$$h_t=(1-z_t)\cdot h_{t-1}+z_t\cdot\tilde{h}_t \tag{20}$$

式中：$\boldsymbol{W}_r$、$\boldsymbol{W}_h$、$\boldsymbol{W}_z$ 和 b_r、b_h、b_z 为各门结构对应的权重矩阵和偏置；X_t 代表当前时刻的输入；h_{t-1} 代表上一时刻 GRU 单元的隐藏状态；h_t 代表当前 GRU 单元计算出的隐藏状态；z_t 为更新门；r_t 为重置门。

2.2 Attention 机制优化的 GRU 预测模型

在时间序列预测过程中，离预测点时间间隔较小的特征信息更具参考性，对预测点的值影响更大，反之间隔较远的特征信息随时间推移影响逐渐递减，因此时间序列不同时间的输入特征对预测值的影响力不同[20]。传统 GRU 神经网络认为其所有时间点的输入的特征影响力相同，并给予相等权重值，忽略了输入特征值对预测值的影响程度。本文提出基于 Attention 注意力机制优化传统 GRU 神经网络，Attention 机制基于评分函数为影响力较高的

因素赋予更高权重，将模型注意力聚合到关键位置，降低了非关键位置的注意，减少了对资源的耗费，能够有效提高了模型预测的精度[21]。Attention机制优化的GRU预测模型具体结构如图2所示。

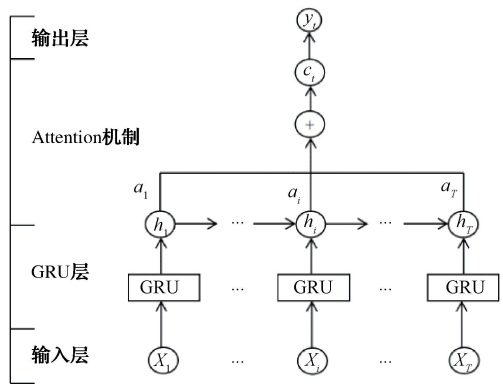

图2 GRU-Attention模型层级图

1)输入层，将分解信号作为GRU神经网络的输入特征，在t时刻中的输入序列具体为：

$$X_t=[x_1,\cdots,x_i,\cdots,x_T] \tag{21}$$

式中：X_t为当前t时刻分解的信号；T为时间步数。

2)GRU层，将X_t作为GRU的输入序列，计算该层第i个输出的隐藏状态h_i：

$$h_i=GRU(X_{i-1},X_i) \tag{22}$$

$$h_t=[h_1,\cdots,h_i,\cdots,h_T] \tag{23}$$

式中：h_i为t时刻中第i个GRU层的隐层状态；h_t为最终在t时刻的GRU层隐藏状态。

3)Attention机制，将GRU层输出的隐层状态h_t作为Attention机制的输入，计算得到每个隐层单元的得分e_i、归一化后得分a_i和最终Attention层输出c_t：

$$e_i=\tanh(Wh_i+b) \tag{24}$$

$$a_i=\frac{\exp(e_i)}{\sum_{j=1}^{T}\exp(e_j)} \tag{25}$$

$$c_t=\sum_{i=1}^{T}a_ih_i \tag{26}$$

式中：W、b为权重系数和偏置系数；h_i为GRU神经网络中t时刻中输出的第i个隐藏单元状态值。

4)输出层，选择$sigmoid$函数作为激活全连接层的函数，输出最终t时刻最终预测值记作y_t：

$$y_t=\text{sigmoid}(\omega c_t+b) \tag{27}$$

2.3 时间序列预测模型的建立

为了充分学习时间序列的复杂变化规律，弱化其非平稳性和非线性的特点，本文提出一种基于二次分解和GRU-Attention的预测模型，其流程如图3所示，具体步骤分为如下5步。

1)使用CEEMDAN分解原始时间序列，得到一个余量信号Res以及$n-1$个IMF模态分量。

2)采用样本熵衡量各模态分量的复杂程度。

3)将复杂度较高的IMF模态分量进行VMD二次分解，进一步降低时间序列的非平稳性，其中每个经过VMD分解后的模态信号可被再次分解为m个Model分量。

4)对获得IMF与Model分量分别建立利用Attention机制优化GRU模型进行预测。

5)将各分量预测结果叠加获得最终结果。

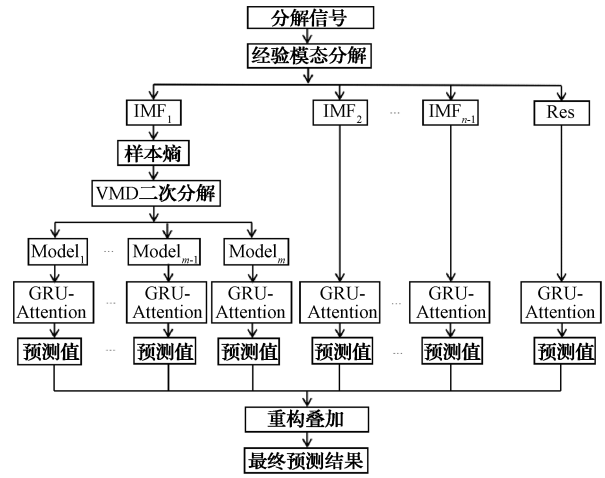

图3 基于二次分解和GRU-Attention的预测模型流程

3 算例分析

3.1 数据来源

本文选取平安银行从2006年1月5日～2019年6月5日所有交易日的收盘价共3 337条数据作为实验数据集。如图4所示，该数据集大且全面，具有复杂规律及较强的非平稳性特征。本文利用数据集的前80%条数据为训练集，后20%条数据为测试集，验证提出模型的准确性。

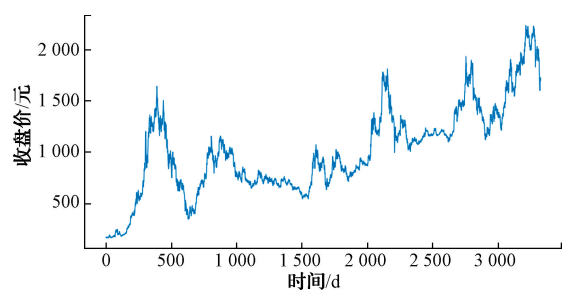

图4 实验数据集

3.2 基于CEEMDAN分解的时间序列

CEEMDAN通过引入了有限次自适应白噪声避免了传统EMD分解基础存在模态混叠现象，同样也减少了EEMD迭代次数改善了重建误差的问题。基于CEEMDAN分解的结果如图5所示。

求解各模态分量的样本熵，图6所示为样本熵变化趋势，表1为样本熵计算结果。

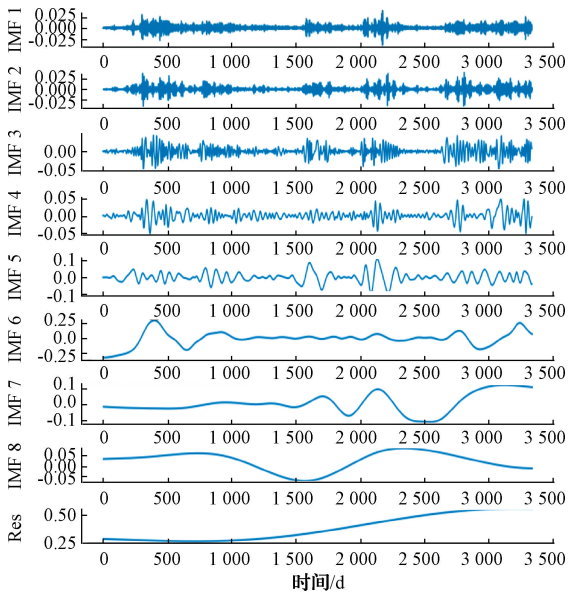

图 5 CEEMDAN 分解结果

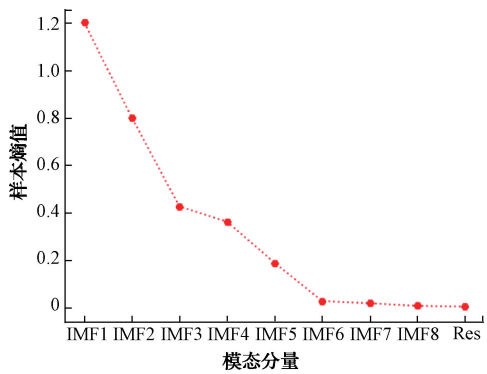

图 6 样本熵趋势

表 1 样本熵计算结果

Signal	样本熵
IMF1	1.202
IMF2	0.798
IMF3	0.424
IMF4	0.359
IMF5	0.186
IMF6	0.024
IMF7	0.015
IMF8	0.003
Res	0.001

从图 6 可以观察到，各个模态分量的样本熵值呈现整体递减的趋势，即复杂度逐渐降低。由表 1 可知，其中 IMF1 分量的样本熵远高于其他分量达到了 1.202，故 IMF1 分量的波动程度较为复杂，受一种或多种趋势性、季节性、周期性等成分的影响较大。其中趋势性是指数据在随时间变化的同时呈现持续升高、下降或短暂不变的变化幅度，季节性是指数据在一年之中由周期性因素的影响而重复出现的规律性周期变动，周期性是指除去季节性外的影响因素，它能够使时间序列产生一种波浪或震动形式的变动。这些成分使序列复杂程度提高，样本熵值增大，对后续预测的准确性有较大影响。

3.3 VMD 二次分解

因 IMF1 复杂度较高，本文考虑使用 VMD 对其进行再次分解[22]。由文献[23]可知，可通过中心频率法判断最佳的模态分解个数 K 的取值，即当其他 K 值下每个分量之间中心频率相减所得到的值相较于当前 K 值是否明显过大，若出现此情况则取最终 K 为 $K-1$，表 2 为不同 K 值下的中心频率。

表 2 不同 K 值下的中心频率

K	中心频率/Hz						
2	0.248 2	0.325 7	—	—	—	—	—
3	0.172 1	0.251 4	0.328 8	—	—	—	—
4	0.173 6	0.250 0	0.320 1	0.395 4	—	—	—
5	0.172 2	0.248 7	0.312 9	0.367 5	0.450 2	—	—
6	0.159 4	0.212 3	0.255 2	0.316 5	0.369 9	0.450 9	—
7	0.116 5	0.182 8	0.242 3	0.243 4	0.324 8	0.377 9	0.452 4

由表 2 可知，$K=7$ 时出现过分解，即各模态分量中心频率的差值仅为 11×10^{-4}，故最终取 $K=6$，最终 VMD 分解结构如图 7 所示。

原始时间序列基于二次分解后，得到 Model1，Model2，…，Modal6 与 IMF2，IMF3，…，Res 共 14 个分量，其样本熵趋势如图 8 所示。由图 8 能够观察到，IMF1 模态分量的复杂程度经 VMD 分解显著降低。

将基于二次分解后得到的 14 个 IMF 分量输入到经 Attention 机制优化的 GRU 神经网络中，模型采用 Adam 优化算法对模型中的权重进行迭代更新。模型各参数选用经验数据且通过多次实验进行调整，最终确定 GRU 神经网络层的隐层节点个数为 64 个，迭代次数为 100 次，batch-size 大小为 64，滑动窗口长度为 5，学习率为 0.01，其他网络均做相同设置。

3.4 时间序列实验结果分析

本文选用评估指标平均绝对百分比（mean absolute

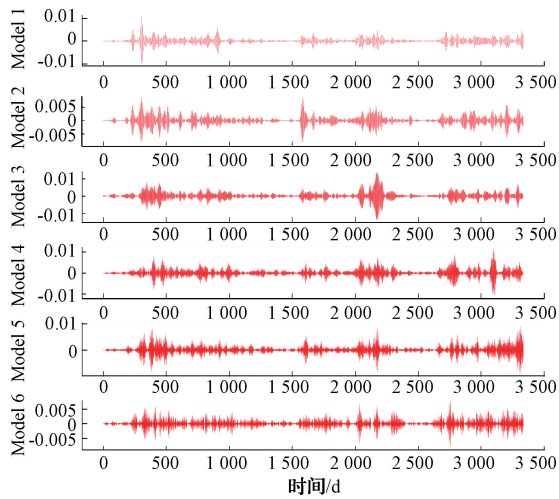

图 7　VMD 分解 IMF1 结果

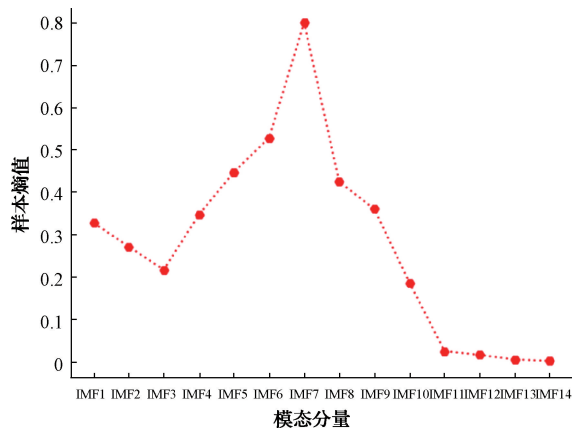

图 8　二次分解后各模态分量样本熵趋势

percentage，MAPE）与决定系数对时间序列模型进行误差分析，其定义公式为：

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_{true}(i)-y_{pre}(i)}{y_{true}(i)}\right|\times 100\% \tag{28}$$

$$\text{决定系数} = 1-\frac{\sum_{i=1}^{n}(y_{true}(i)-y_{pre}(i))^2}{\sum_{i=1}^{n}\left(y_{true}(i)-\frac{1}{n}\sum_{i=1}^{n}y_{true}(i)\right)^2} \tag{29}$$

式中：n 为样本个数；$y_{true}(i)$ 为 i 时刻的实际值；$y_{pre}(i)$ 为 i 时刻的预测值。

为了直观体现本文提出的基于二次分解和 GRU-Attention 时间序列模型的预测优势，基于上述评价指标，将基于二次分解和 GRU-Attention 时间序列预测模型与多个模型进行对比实验，观测各模型对测试集的拟合效果。由于神经网络具有随机性的特点，如随机初始化权重等，这种随机性会造成相同数据使用同样的神经网络但计算结果不同的现象。为减少实验中随机性对各模型预测精度带来的影响，本文对每种模型进行了多次实验并记录最佳结果，不同模型预测结果对比如表 3 所示。

表 3　不同模型预测结果对比

预测模型	MAPE/%	决定系数
LSTM	7.946	0.698
GRU	7.853	0.758
EMD-GRU	6.289	0.809
CEEMDAN-GRU	5.276	0.826
CEEMDAN-VMD-GRU	4.877	0.863
CEEMDAN-VMD-GRU-Attention	2.986	0.891

从表 3 可观察到，本文提出的 CEEMDAN-VMD-GRU-Attention 模型时间序列在 MAPE 与决定系数两种评估指标中的计算值分别为 2.986%、0.891，将 CEEMDAN-VMD-GRU-Attention 模型与 LSTM、GRU、EMD-GRU、CEEMDAN-GRU、CEEMDAN-VMD-GRU、CEEMDAN-VMD-GRU-Attention 进行对比，其 MAPE 指标的值分别下降了 4.960%、4.867%、3.303%、2.290%、1.891%，决定系数指标的值分别提高了 0.193、0.133、0.082、0.065、0.028，由此可以验证 CEEMDAN-VMD-GRU-Attention 模型可以有效提高对时间序列预测的精度。另外，将 CEEMDAN-GRU 与 EMD-GRU 相比，其 MAPE 指标的值下降了 1.013%，决定系数值提高了 0.017，验证了 CEEMDAN 有效避免了 EMD 的模态混叠问题，有效提高了后续模型的预测效果。将 CEEMDAN-VMD-GRU 与 CEEMDAN-GRU 相比，其 MAPE 指标的值下降了 0.399%，决定系数值提高了 0.037，验证了复杂程度较高的模态分量经 VMD 分解后，其非平稳性能够被充分弱化。将 CEEMDAN-VMD-GRU-Attention 与 CEEMDAN-VMD-GRU 相比，其 MAPE 指标的值下降了 1.891%，决定系数值提高了 0.028，验证了 Attention 机制可有效优化 GRU 神经网络，进一步捕捉关键时刻的信息。

各模型预测结果如图 9～14 所示，可观察到 LSTM 与 GRU 的预测结果与真实值的拟合程度较差，说明直接使用单模型预测时间序列性能较低。EMD-GRU、CEEMDAN-GRU 相比于 LSTM、GRU 两个单模型预测结果同真实值的拟合效果较好，说明单次分解能够简化序列的复

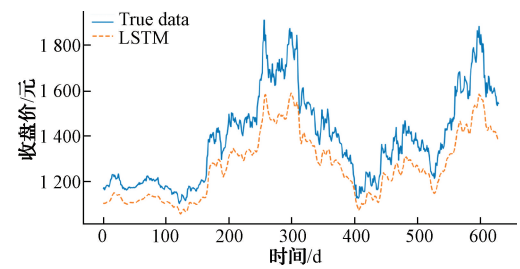

图 9　LSTM 模型预测结果

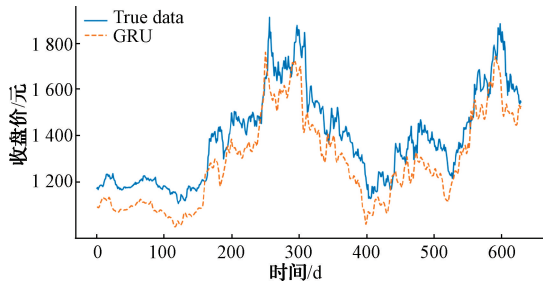

图 10　GRU 模型预测结果

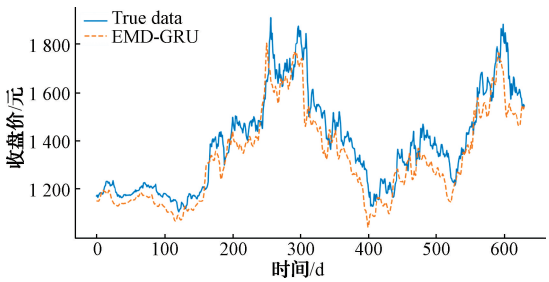

图 11　EMD-GRU 模型预测结果

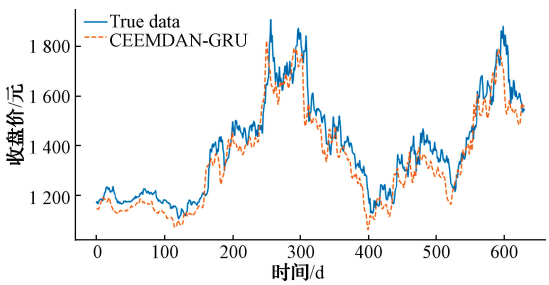

图 12　CEEMDAN-GRU 模型预测结果

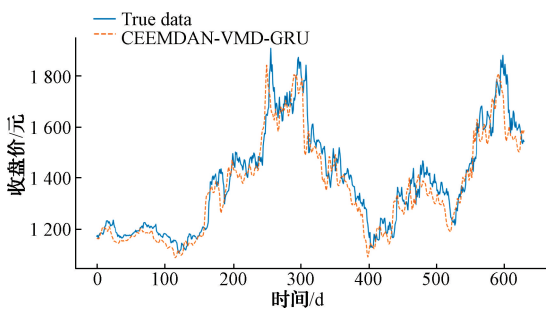

图 13　CEEMDAN-VMD-GRU 模型预测结果

杂度，便于后续模型的建模分析。CEEMDAN-VMD-GRU、CEEMDAN-VMD-GRU-Attention 模型与其他模型的预测结果相比，还原真实变化趋势的能力最佳，并且本文提出的预测方法拟合程度最高，说明经过二次分解的模型能够在简化序列复杂度的基础上降低高频子序列的非平稳性，减少高频子序列对模型预测误差的影响，并且 Attention 机制的加入能够使模型精准捕获序列复杂的变化规律，进一步提高模型预测精度。

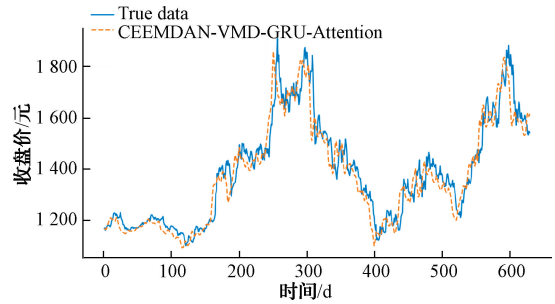

图 14　CEEMDAN-VMD-GRU-Attention 模型预测结果

4　结　论

本文提出一种基于 CEEMDAN-VMD-GRU-Attention 的时间序列预测模型。针对时间序列高度的非平稳特征，首先采用 CEEMDAN 将时间序列分解为若干规律较平稳的 IMF 模态分量，求出各模态分量的样本熵，其次将样本值较高的 IMF 采用 VMD 二次分解，弱化样本熵值较高的模态分量，最后建立经 Attention 机制优化的 GRU 模型，使模型在预测过程中具有精准捕获关键时刻信息的能力，最后将各分量预测结果相加获得最终结果。将本文提出模型与 LSTM、GRU、EMD-GRU、CEEMDAN-GRU、CEEMDAN-VMD-GRU 五种模型对比，从实验的评估指标看，本文提出模型在 MAPE 和决定系数两种指标上均具有显著优势，分别达到了 2.9%与 0.891。从实验预测结果与真实值的拟合程度来看，本文提出模型具有更高的拟合性，能较好还原序列复杂的变化趋势，更适用于时间序列的预测。

参 考 文 献

[1] 王英伟，马树才. 基于 ARIMA 和 LSTM 混合模型的时间序列预测[J]. 计算机应用与软件，2021，38(2)：291-298.

[2] 顾海燕. 时间序列分析在人口预测问题中的应用[J]. 黑龙江工程学院学报，2007(3)：69-71.

[3] 王毅，魏江超，孙启元，等. 基于 ARIMA-ANN 模型的生态安全评价及预测——以河西走廊城市群为例[J]. 生态学杂志，2020，39(1)：326-336.

[4] 刘斌，毕小熊，党军朋，等. 基于支持向量回归的变电站蓄电池退化趋势预测[J]. 电源学报，2020，18(6)：207-214.

[5] 吴琼，徐锐良，杨晴霞，等. 基于 PCA 和 GA-BP 神经网络的锂电池容量估算方法[J]. 电子测量技术，2022，45(6)：66-71.

[6] 程换新，黄震. 基于改进 PSO 优化 RNN 的短期电力负荷预测模型[J]. 电子测量技术，2019，42(20)：94-98.

[7] 陈安民，张春元，张泽林. 基于长短期记忆网络的无人机认知无线电频谱预测[J]. 国外电子测量技术，

2021,40(1):37-43.
[8] 崔立志,何泽彬,李璇.基于注意力的R-GCN-GRU的在线学生绩效预测[J].电子测量技术,2021,44(19):69-75.
[9] 王耀庆,孙建平,李冰,等.基于小波变换和LSTM的短期风速预测研究[J].计算机仿真,2021,38(2):438-443.
[10] 李孟威,史元浩,杨彦茹,等.融合EMD和LSTM的受热面积灰预测研究[J].电子测量与仪器学报,2020,34(11):166-172.
[11] HASAN N. A methodological approach for predicting COVID-19 epidemic using EEMD-ANN hybrid model[J]. Internet of Things,2020,11:100228.
[12] JIANG X, WEI P, LUO Y, et al. Air pollutant concentration prediction based on a CEEMDAN-FE-BiLSTM model[J]. Atmosphere,2021,12(11):1452.
[13] ZHANG Y, CHEN Y. Application of hybrid model based on CEEMDAN, SVD, PSO to wind energy prediction[J]. Environmental Science and Pollution Research, 2022,29(15):22661-22674.
[14] LIN Y, YAN Y, XU J, et al. Forecasting stock index price using the CEEMDAN-LSTM model[J]. The North American Journal of Economics and Finance, 2021,57:101421.
[15] 陈妮,覃玉荣,熊艳婷,等.基于变分模态分解的脑电锁相刺激方法[J].仪器仪表学报,2020,41(5):205-213.
[16] ZHANG Y, PAN G, CHEN B, et al. Short-term wind speed prediction model based on GA-ANN improved by VMD[J]. Renewable Energy, 2020, 156: 1373-1388.
[17] 魏乐,李思莹.基于MEEMD-LSSVM的风电功率超短期预测研究[J].智慧电力,2020,48(5):21-26.
[18] 余向阳,赵怡茗,杨宁宁,等.基于VMD-SE-LSSVM和迭代误差修正的光伏发电功率预测[J].太阳能学报,2020,41(2):310-318.
[19] WANG Y, LIAO W, CHANG Y. Gated recurrent unit network-based short-term photovoltaic forecasting[J]. Energies, 2018, 11(8): 2163.
[20] LIANG Y, XU G, REHMAN S U. Multi-scale attention-based deep neural network for brain disease diagnosis [J]. CMC-Computers Materials & Continua, 2022, 72(3): 4645-4661.
[21] FU L, ZHAO F. Prediction of hot topics of agricultural public opinion based on attention mechanism LSTM model[J]. International Journal of Agricultural and Environmental Information Systems, 2021, 12(4): 1-16.
[22] 陈妮,覃玉荣,熊艳婷,等.基于变分模态分解的脑电锁相刺激方法[J].仪器仪表学报,2020,41(5):205-213.
[23] 陶志勇,胡启振,任晓奎.基于二层分解技术和改进神经网络的河流溶解氧预测研究[J].云南大学学报(自然科学版),2022,44(2):262-270.

作者简介

高凯悦,硕士研究生,主要研究方向为时间序列预测方法研究。
E-mail:903119649@qq.com

牟莉,硕士,副教授,研究生导师,主要研究方向为智能化信息系统、嵌入式系统应用、图像处理与识别等。
E-mail:626470481@qq.com