

基于 MVMD-ISSA-LSTM 融合算法的复杂系统预测

苏会强

北方民族大学数学与信息科学学院 宁夏银川

【摘要】针对复杂时序数据的多噪声、非平稳以及随机性的特点，本文提出基于多策略融合算法来实现其准确预测。首先，利用多元变分模态分解（MVMD）将复杂时序数据分解为有限个模态分量。其次，在麻雀搜索法（ISSA）的基础上加入优化后的 Tent 混沌映射使麻雀初始种群分布均匀，引入动态自适应权重调整发现者步长防止算法陷入局部最优同时加入防参数越界设置使算法搜索器更加稳定，加入高斯变异扰动提高算法跳出局部最优的能力。再次，将改进后的算法嵌入长短时记忆网络（LSTM）形成融合算法寻找最优参数从而达到准确预测复杂时序数据的目的。最后，以区域日最高气温数据为例验证融合算法的有效性及精确性。研究证实，本文所提出的融合算法对复杂时序数据的预测有较好的表现。

【关键词】复杂时序数据；多元变分模态分解；麻雀搜索算法；长短时记忆网络；降噪

【收稿日期】2023 年 11 月 6 日 **【出刊日期】**2023 年 12 月 13 日 **【DOI】**10.12208/j.aics.20230040

Prediction of complex system based on MVMD-ISSA-LSTM Fusion algorithm

Huiqiang Su

School of Mathematics and Information Science, North Minzu University, Yinchuan, Ningxia

【Abstract】Aiming at the characteristics of multi-noise, non-stationary and randomness of complex time series data, this paper proposes a multi-strategy fusion algorithm to achieve accurate prediction. Firstly, multivariate variational mode decomposition (MVMD) is used to decompose complex time series data into finite modal components. Secondly, the sparrow search method (ISSA) is improved, adding the optimized Tent chaos map to make the initial population of sparrows evenly distributed, introducing dynamic adaptive weight adjustment to the discoverer step to prevent the algorithm from falling into the local optimal, adding anti-parameter overstepping Settings to make the algorithm searcher more stable, adding Gaussian variation perturbation to improve the algorithm's ability to jump out of the local optimal. Therefore, the algorithm is embedded into the Long and short time memory network (LSTM) to form a fusion algorithm to find the best parameters and achieve the purpose of accurately predicting complex time series data. Finally, the regional daily maximum temperature data is taken as an example to verify the effectiveness and accuracy of the fusion algorithm. The research proves that the fusion algorithm proposed in this paper has a good performance in the prediction of complex time series data.

【Keywords】Complex temporal data; Multivariate variational modal decomposition; Sparrow search algorithm; Long short term memory network; Noise reduction

1 引言

随着计算机技术在气候、金融、医疗等各个领域的深入应用，复杂现实系统的时序数据规模大、纬度高、随机性强的特点越来越凸出^[1]，因此受强噪声干扰也尤为突出。为了解决复杂时间序列的噪声去除问题，近年来，研究学者提出了很多方法。

Fenghua 等人^[2]使用奇异谱分析将股票时间序列分解成不同特征的噪声，将重构后的数据采用支持向量机（Support Vector Machine, SVM）进行预测，有效提高了预测精度。Nguyen 等人^[3]使用小波变换对天然气价格时间序列降噪，将降噪的序列输入到预测模型中取得了更高的精度。但是，基于小波变换

的噪声去除方法效果取决于小波基函数的选取, 另外分解的层数也需要根据每一层的波形去选取, 这往往受限于先验知识和专业知识。经验模式分解 (Empirical Mode Decomposition, EMD) 是 Huang 等人^[4]提出的一种自适应时间序列分解方法, EMD 能将复杂时间序列分解成频率由高到低的信号分量。但是 EMD 算法有时会产生模态混叠和端点效应^[5]。变分模态分解 (Variational Mode Decomposition, VMD) 由 Drag-omiretskiy 等人^[6]提出将信号分解为具有不同频率模态分量的一类方法。该方法可有效解决 EMD 模态混叠和端点效应问题, 对噪声具有更强的鲁棒性, 但是与 EMD 一样无法处理多通道数据。Rehman 等人^[7]提出的多元变分模态分解 (Multivariate Variational Mode Decomposition, MVMD) 方法。该方法 VMD 从单通道拓展到多通道, 避免了频率不匹配问题, 同时又解决了 EMD 的模式混叠问题。

复杂时间序列的出现意味着传统的简单时间序列的模型可能不再适用, 学者们开始借助深度神经网络模型来处理复杂时序系统。深度神经网络^[8] (Deep Neural Networks, DNN)、循环神经网络^[9] (Recurrent Neural Network, RNN)、长短时记忆网络^[10] (LSTM) 是目前常用于复杂时间序列数据处理的模型。由于 DNN 和 RNN 自身结构原因, 在进行数据处理时会产生梯度爆炸等问题, 因此很难捕捉到长期的依赖关系。而 LSTM 作为 RNN 的变体则很好地解决了这些问题, 在复杂时间序列数据的预测中具有良好的准确性, 但是 LSTM 自身也有一些局限, 需要人为设置参数, 如隐含层神经元个数、学习率以及正则化系数等, 而不同的参数设置对模型预测的精确度也会产生差异, 因此选出其最优参数非常重要。

群体智能优化算法的主要思想就是寻找搜索空间中的全局最优解。研究人员通过多种智能生物的行为, 提出了大量群体智能优化算法例如麻雀搜索法^[11] (Sparrow Search Algorithm, SSA)。但是 SSA 由于初始种群分布的随机性以及迭代后期容易陷入局部最优等问题, 会造成算法收敛速度、预测精度以及平稳性出现问题。为了解决这一问题, 本文在麻雀算法 (SSA) 的初始种群加入了优化的 Tent 混沌映射, 对发现者步长进行动态权重调整以及在

迭代后期加入高斯变异对最优结果进行扰动, 从而提高算法搜索速度以及精度。

目前对于数据降噪或者优化算法改进后的复杂时间序列预测研究较为成熟, 而将数据分解降噪与优化算法改进之后融合预测的研究较少。因此, 本文构建基于多元变分模态分解 (MVMD) 和多策略改进的麻雀算法 (ISSA) 优化 LSTM 的融合算法, 该算法将复杂时间序列数据进行分解降噪后将各个分量作为输入, 利用改进的算法进行 LSTM 参数智能寻优, 从而实现更精准高效预测复杂系统的能力。

2 融合算法构建

2.1 多元变分模态分解

MVMD 可以保证复杂时序序列在分解时各个通道分量同时对其中心频率和带宽进行迭代更新。其目的是从含有 C 个数据通道的输入数据 $X(t)=[x_1(t), x_2(t), \dots, x_C(t)]$ 中提取出所包含的 K 个多元调制振荡 $U_k(t)=[u_1(t), u_2(t), \dots, u_C(t)]$, $k=1, 2, \dots, K$ 。具体步骤如下:

1) 假设 K 个多元调制振荡 $U_k(t)$, 使:

$$X(t) = \sum_{k=1}^K U_k, k=1, 2, \dots, K$$

2) 对 $U_k(t)$ 中每一个元素进行希尔伯特变换, 记为 $U_k^+(t)$, 以计算单边频谱。之后借助指数项 $e^{-j\omega_k t}$ 调整至其对应的中心频率 $\omega_k(t)$ 。利用谐波转换后 $U_k^+(t)$ 的梯度函数的 L_2 范数估计出各模态 $U_k(t)$ 的带宽。MVMD 受约束的优化问题变为:

$$\begin{aligned} \min_{\{u_{k,c}\}, \{\omega_k\}} & \left\{ \sum_{k=1}^K \sum_{c=1}^C \left\| \partial_t \left[u_{k,c}^+(t) e^{-j\omega_k t} \right] \right\|_2^2 \right\} \\ \text{s.t.} & \sum_{k=1}^K u_{k,c}(t) = x_c(t), c=1, 2, \dots, C, \end{aligned} \quad (1)$$

其中: $u_{k,c}^+(t)$ 为相应的通道 c 中模式 k 的信号表示, ∂_t 表示与时间相关的偏导数。

3) 对上述优化问题求解。构造的拉格朗日表达式为:

$$\begin{aligned} L(\{u_{k,c}\}, \{\omega_k\}, \lambda_c) &= \alpha \sum_{k=1}^K \sum_{c=1}^C \left\| \partial_t \left[u_{k,c}^+(t) e^{-j\omega_k t} \right] \right\|_2^2 \\ &+ \sum_{c=1}^C \left\| x_c(t) - \sum_{k=1}^K u_{k,c}(t) \right\|_2^2 \\ &+ \sum_{c=1}^C \left\langle \lambda_c(t), x_c(t) - \sum_{k=1}^K u_{k,c}(t) \right\rangle, \end{aligned} \quad (2)$$

其中： α 为惩罚参数， λ_c 为拉格朗日乘子； $\langle \cdot \rangle$ 为内积符号。

4) 为解决 (2) 式中的变分问题，应用乘子交替方向法 (Alternate Direction Method of Multipliers, ADMM) 求解，通过求取更新后的模式以及中心频率，以得到分解后的信号分量。得到模式更新为：

$$\hat{u}_{k,c}^{n+1}(\omega) = \frac{\hat{x}_c(\omega) - \sum_{i \neq k} \hat{u}_{i,c}(\omega) + \hat{\lambda}_c(\omega)}{1 + 2\alpha(\omega - \omega_k)^2}, \quad n=0,1,2,\dots \quad (3)$$

其中： n 为迭代次数； $\hat{x}_c$ 、 $\hat{u}_{i,c}$ 、 $\hat{\lambda}_c$ 为傅里叶变换后的信号。

中心频率更新为：

$$\omega_k^{n+1} = \frac{\sum_{c=1}^C \int_0^\infty \omega |\hat{u}_{k,c}(\omega)|^2 d\omega}{\sum_{c=1}^C \int_0^\infty |\hat{u}_{k,c}(\omega)|^2 d\omega} \quad (4)$$

2.2 麻雀搜索法及其改进

(1) 麻雀搜索法

麻雀搜索算法于 2020 年由 Xue^[11] 首次提出，该算法是模拟麻雀种群的觅食和反捕食行为来达到寻优的目的，具有寻优能力强，收敛速度快的特点。在种群中，搜寻者负责为种群寻找食物，然后再把食物的位置信息反馈给种群，位置更新如下：

当 $R_2 < ST$ 时表示周围没有危险：

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \exp(-\frac{i}{\alpha \cdot iter_{\max}}) & \text{if } R_2 < ST \\ X_{i,j}^t + Q \cdot L & \text{if } R_2 \geq ST \end{cases}, \quad (5)$$

其中： $X_{i,j}^t$ 表示的是在 t 次迭代后，麻雀 i 在第 j 维空间的位置； R_2 为警惕值； ST 为安全值； Q 为正态分布随机数； L 是 1 行 d 列的全 1 矩阵； α 在 (0, 1] 上随机生成； $iter_{\max}$ 为最大迭代数。

麻雀种群中除了搜寻者其余均为跟随者，位置更新如下：

当 $i > \frac{n}{2}$ 时，表示第 i 个跟随者食物获取不足，需要去其他地方觅食：

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp(-\frac{X_{worst}^t - X_{i,j}^t}{i^2}) & \text{if } i > \frac{n}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \cdot A^+ \cdot L & \text{if } \text{otherwise} \end{cases}, \quad (6)$$

其中： X_{worst} 为最差位置； X_p 为最优位置； $A^+ = A^T (AA^T)^{-1}$ ， A 表示 1 行 d 列的矩阵，各个元素随机被赋值为 ± 1 ； n 为种群数量。

警惕者是在麻雀种群中随机选取，若其感知到危险，种群会快速地飞往另一个安全的地方。位置更新如下：

当 $f_i > f_g$ 时，表示外围麻雀位置危险。

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t + K \left(\frac{|X_{i,j}^t - X_{worst}^t|}{(f_i - f_w) + \varepsilon} \right), f_i = f_g \\ X_{best}^t + \beta |X_{i,j}^t - X_{best}^t|, f_i > f_g \end{cases}, \quad (7)$$

其中： f_i 为麻雀 i 的适应度， f_g 为全局最佳适应度， f_w 为最差适应度； X_{best} 为最优位置； K 在 [-1, 1] 上任意取值，正负表示方向； ε 表示最小常数，避免运算错误； β 服从标准正态分布并且随机产生。

(2) 多策略改进的麻雀算法

① 优化的 Tent 混沌映射

Tent 混沌映射能够搜索全局最优解，并具有较强的收敛速度^[12]。该方法可以使 SSA 初始群分布均匀，增加种群多样性。Tent 映射函数计算方式见式 (8)。

$$x_i = \begin{cases} 2x_i, 0 \leq x_i \leq 1/2 \\ 2(1-x_i), 1/2 < x_i \leq 1 \end{cases} \quad (8)$$

Tent 映射在 [0, 1] 内产生的序列具有周期性同时存在不稳定点^[13]。因此可在初始的 Tent 混沌映射函数中增加一个随机变量 $rand(0,1) \times 1/N$ 。优化后的 Tent 混沌映射函数如式 (9)。

$$x_{i+1} = \begin{cases} 2x_i + rand(0,1) \times \frac{1}{N}, 0 \leq x_i \leq 1/2 \\ 2(1-x_i) + rand(0,1) \times \frac{1}{N}, 1/2 < x_i \leq 1 \end{cases} \quad (9)$$

其中： N 是 Tent 混沌序列内的粒子个数。

② 动态自适应权重

发现者位置在受到上一步位置以及全局最优解的共同影响下，可有效防止算法陷入局部最优，因此引入上一步的全局最优解。参考惯性权重思想^[14]，在发现者的位置更新中引入权重因子 ϕ ，迭代初期时，有较大的 ϕ 值，随着迭代次数的增加， ϕ 值逐渐减小，从而提高收敛速度，公式如下：

$$\phi = \frac{e^{\frac{2(1-\frac{t}{iter_{max}})}{2(1-\frac{t}{iter_{max}})} - e^{\frac{-2(1-\frac{t}{iter_{max}})}{2(1-\frac{t}{iter_{max}})}}}{e^{\frac{2(1-\frac{t}{iter_{max}})}{2(1-\frac{t}{iter_{max}})}} + e^{\frac{-2(1-\frac{t}{iter_{max}})}{2(1-\frac{t}{iter_{max}})}}}, \quad (10)$$

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t + \phi(f_{a,b}^t - x_{i,j}^t) \cdot rand(0,1) & \text{if } R_2 < ST \\ X_{i,j}^t + Q & \text{if } R_2 \geq ST \end{cases}, \quad (11)$$

其中: $f_{a,b}^t$ 为上一次迭代的全局最优解。

③高斯变异

SSA 在迭代后期会逐渐陷入局部最优, 造成算法过早收敛, 影响预测精度。为此, 引入高斯变异策略, 选择当前适应度最好的个体进行变异, 然后比较选择较优的位置进行下一次迭代。表达式如下:

$$Lnew_{i,j}^{t+1} = X_{best}^t + Gauss(0,1) \cdot X_{best}^t, \quad (12)$$

$$X_{best} = \begin{cases} X_{i,j}^{t+1} & f(X_{i,j}^{t+1}) < f(Lnew_{i,j}^{t+1}) \\ Lnew_{i,j}^{t+1} & f(X_{i,j}^{t+1}) \geq f(Lnew_{i,j}^{t+1}) \end{cases}, \quad (13)$$

其中: $Lnew_{i,j}^{t+1}$ 表示经过高斯变异后的个体位置; X_{best}^t 为当前最优个体的位置; $Gauss(0,1)$ 为标准高斯分布的随机数。

2.3 长短期记忆网络

LSTM 网络模型在隐藏层中引入被称为“记忆细胞”的结构, 使用不同的函数去计算隐藏层的状态。在“记忆细胞”中, 使用三个门限层控制可以通过门的信息量, 对具有长期序列依赖问题的数据非常有效^[15]。

对于复杂时序序列 $x = \{x_1, x_2, x_3, \dots, x_t\}$, 通过迭代进入 LSTM 的门而更新细胞状态, 计算隐藏层 $h = \{h_1, h_2, h_3, \dots, h_t\}$, 以及输出层 $o = \{o_1, o_2, o_3, \dots, o_t\}$, 得到最终预测趋势, 具体步骤如下。

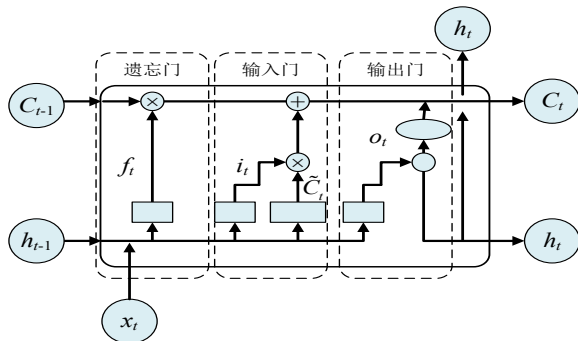


图1 LSTM 结构图

(1) 遗忘门 f 通过细胞状态决定遗忘信息, 通过 C_{t-1} 来确定能否保存到当前细胞状态 C_t 。

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f), \quad (14)$$

其中: σ 代表 *Sigmoid* 激活函数; W_f 代表遗忘门的权重系数矩阵; h_{t-1} 代表前一时刻上一个细胞输出; x_t 代表当前细胞的输出; b_f 代表遗忘门的偏置向量。

(2) 输入门 i 进行细胞状态更新, 通过当前的细胞输入 x_t 决定能否保存到当前细胞状态 C_t 。

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \quad (15)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c), \quad (16)$$

$$C_t = i_t \tilde{C}_t + f_t C_{t-1}, \quad (17)$$

其中: $\tanh$ 代表双曲正切激活函数; W_i , b_i 分别代表通过 *Sigmoid* 激活函数确定的输入门权重系数矩阵和偏置矩阵; W_c , b_c 分别代表通过双曲正切函数确定的输入门权重系数矩阵和偏置矩阵。

(3) 输出门 o 通过细胞状态确定细胞信息。

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o), \quad (18)$$

$$h_t = o_t \tanh(C_t), \quad (19)$$

其中: W_o , b_o 分别代表权重系数矩阵和偏置矩阵, o_t 代表 t 时刻的输出门。

2.4 改进麻雀算流程

对 ISSA 的改进见图 2, 具体步骤如下:

步骤 1: 算法参数初始化;

步骤 2: 利用 Tent 混沌映射初始化种群得到分布较为均匀的种群;

步骤 3: 求每只麻雀的适应度值, 确定适应度最优, 最差的麻雀的位置;

步骤 4: 按照适应度的大小来调整每只麻雀的顺序;

步骤 5: 根据全局适应度变化情况, 调整发现者和跟随者的比例;

步骤 6: 更新发现者、跟随者和警惕者的位置;

步骤 7: 对最优位置进行高斯扰动, 若是最优则更新最优位置, 若不是则更新种群位置重复步骤 3

到步骤 7, 直到最大迭代次数。

2.5 ISSA 的算法性能测试

为了测试 ISSA 的收敛精度、稳定性以及收敛速度, 本文利用 9 种测试函数, 分别为 4 个单峰函数、5 个多峰函数对改进之后的算法进行测试并与 SSA 以及 PSO 进行比较, 为公平的验证算法性能, 实验需要满足其运行环境、迭代次数 1000 算法独立运行次数以及种群规模 100 均保持相同, 本节的仿真实验均基于 Matlab R2019a 的仿真软件。

测试函数如表 1 所示, 表中 f_1 - f_4 为单峰测试函数, 单峰函数只有一个全局最优解, 因此可用来测试算法的收敛速度, f_5 - f_9 为多峰测试函数, 多峰函数有多个局部最优解, 会让算法陷入局部最优, 可用来测试算法跳出局部最优的能力。

由于仅运行一次算法, 其结果会有偶然性, 为测试 ISSA 算法的性能, 取算法独立运行 30 次的结果的平均值作为最后结果, 以每个测试函数的平均值以及最优值作为评价算法的性能指标结果见表 2。

从表 2 中结果, 在算法的收敛精度方面, ISSA 和 SSA 在 f_1 , f_2 , f_3 , f_7 的最优值和平均值都为 0 达到了已知最优解, 即在这几个测试函数基础上两种算法寻优能力和收敛速度都很优秀, 而在 f_3 , f_4 , f_6 , f_8 , f_9 为基础的测试函数中, 虽然 ISSA 没有达到最优解, 但是相较于 SSA、PSO, ISSA 的最优值和平均值的精度要更优。在算法收敛速度方面列举了单峰 f_1 , 多峰 f_8 的迭代图像如图 3, 图 4 所示。图 3 中虽然 ISSA 和 SSA 收敛速度相似, 但 ISSA 收敛更加光滑且在迭代 100 次左右收敛完成, 而 SSA 的收敛情况则比较曲折, PSO 收敛最差。图 4 中三种优化算法的收敛情况则更为明显, ISSA 最优, SSA 其次, PSO 最后。

综上, 综合多策略改进的 ISSA 与其余算法相比, 其收敛速度和稳定性都更加优越, 具有良好的寻优能力。

3 基于 MVMD-ISSA-LSTM 融合算法的复杂时序模型

MVMD-ISSA-LSTM 融合算法综合了 VMD 上改进的自适应性、多通道分量频率一致性, ISSA 收敛更快、精度更高以及 LSTM 长期记忆的优势。模型的建立包含以下步骤:

步骤 1: MVMD 序列分解。MVMD 在这个过程

中, 各通道的 IMF 分量同时迭代更新其中心频率和带宽, 从而自适应地得到 k 个频率一致的 IMF 分量。

步骤 2: ISSA 参数寻优。经过上述检验 ISSA 拥有较好的寻优能力。利用综合多策略的 ISSA 对 LSTM 参数进行优化, 使得隐含层节点数 (h), 学习率 (α), 正则化系数达到最优。

步骤 3: 训练 LSTM 网络模型。在寻得最优参数的基础上建立预测模型, 对每个归一化处理后的 IMF 子序列进行模型训练。复杂时序数据是非平稳的, 归一化处理可避免训练过程中出现数据较大波动。

步骤 4: 模型预测。在训练好的 LSTM 网络的基础上对每个 IMF 子序列进行预测, 各个 IMF 的预测结果的叠加为最终预测结果。

4 实例分析

本文的样本数据集为 2021 年 1 月到 2022 年 9 月银川市的日最高气温数据 (数据来源为银川市气象自动站的地面观测资料)。数据集按照 3: 1 的结构进行训练、测试划分, 日最高气温趋势见图 6。

MVMD 算法中有两个重要的参数需要提前设置, 分别为分解个数 k 以及惩罚参数 α 。若 k 值设置较小, 就会导致分析信号分解不充分; 若所设置的 k 值大于被分析信号中所包含的有用分量信号的个数, 则会导致分析信号过分解。合理的选择 k 值是 MVMD 算法的关键步骤之一。此外, MVMD 算法中的惩罚参数 α 用以保障信号的重构精度。最高气温复杂时序数据经过 MVMD 分解后可有效降低数据非平稳带来的噪声影响从而达到数据降噪的目的。根据文献^[16], 最终选定惩罚参数为 $\alpha=2000$, 模态分量个数为 $k=6$, 日最高气温数据经过 MVMD 分解得到的分量如图 7 所示, 数据模态分解前后的信号情况如图 8 所示。

通过 MVMD 分解以后, 重构后的数据要更加平稳。对于上述的每一个 IMF 分量分别建立 ISSA-LSTM 的预测模型, 各个子序列预测结果的叠加即为最终的预测数据。在对每个子序列进行预测时, 经过多次测试统一将每个子序列的延时步长设置为 15 得到每个分量的最优参数见表 3。

4.1 预测结果对比分析

为了验证 MVMD-ISSA-LSTM 的可靠性, 建立 VMD-ISSA-LSTM, ISSA-LSTM 预测日最高气温,

SSA-LSTM 预测最高气温与 MVMD-ISSA-LSTM 进行结果比较得到预测结果对比见图 9, 各个预测方法的相对误差对比结果见图 10。

通过对比分析可得: 优化后的 LSTM 模型明显比优化之前的预测效果要好; 改进得到的 ISSA-LSTM 与改进前的 SSA-LSTM 对比可以直观看出 SSA-LSTM 在进行模型改进以后, 预测结果对比 SSA-LSTM 波动趋势平稳, 整体上预测效果更接近于真实值; 而 ISSA-LSTM 在加入多元模态分解后预测情况比 VMD-ISSA-LSTM、ISSA-LSTM 更加接近真实值且波动更加趋于平稳, 说明 MVMD 算法相比于 VMD 起到了很好的降噪作用。为了进一步增

加预测结果的可靠性, 本文用 RMSE、MAE、MAPE 作为各预测模型的误差评判结果, 如表 4 所示。

各模型的指标结果对比如图 11 所示。通过模型结果, MVMD-ISSA-LSTM 模型相比于 VMD-ISSA-LSTM 模型, RMSE 降低了 8.0%, MAE 降低了 6.5%, MAPE 降低了 6.3%; MVMD-ISSA-LSTM 模型相比于 ISSA-LSTM 模型, RMSE 降低了 13.9%, MAE 降低了 14.6%, MAPE 降低了 21.8%; ISSA-LSTM 模型相比于 SSA-LSTM 模型, RMSE 降低了 18.5%, MAE 降低了 21.0%, MAPE 降低了 24.9%; 综上, 本文所提出的 MVMD-ISSA-LSTM 模型的预测性能是较好的。

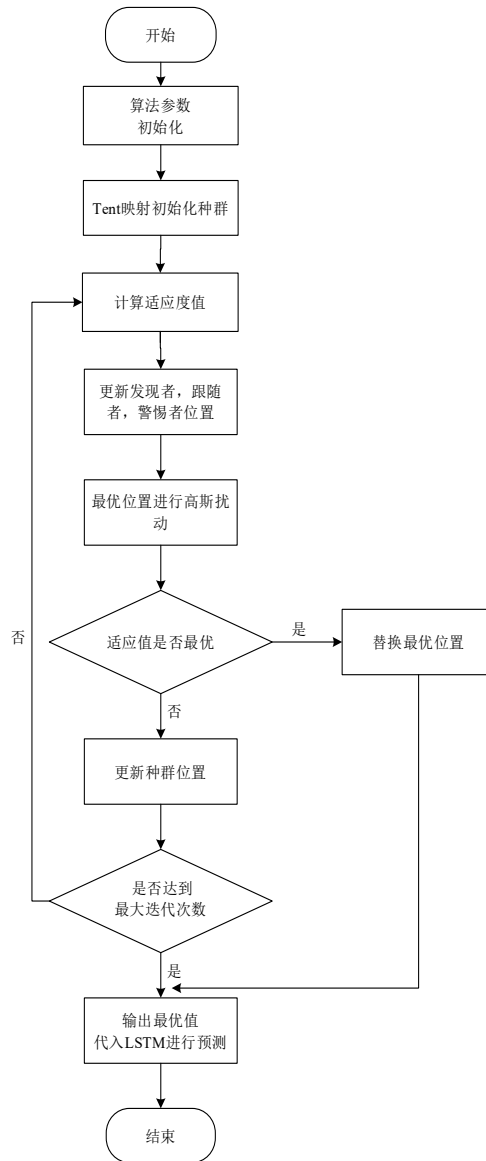


图2 ISSA 算法流程

表 1 测试函数信息

测试函数	维度	搜索空间	最优值
$f_1(x) = \sum_{i=1}^n x_i^2$	30	[-100, 100]	0
$f_2(x) = \sum_{i=1}^n (\sum_{j=1}^i x_j)^2$	30	[-100, 100]	0
$f_3(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	30	[-30, 30]	0
$f_4(x) = \sum_{i=1}^n ix^4 + random[0,1)$	30	[-1.28, 1.28]	0
$f_5(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	30	[-5.12, 5.12]	0
$f_6(x) = -20 \exp(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}) - \exp(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)) + 20 + e$	30	[-32, 32]	0
$f_7(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos(\frac{x_i}{\sqrt{i}}) + 1$	30	[-600, 600]	0
$f_8(x) = \frac{\pi}{n} \{10 \sin(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})] + (y_n - 1)^2\} + \sum_{i=1}^n u(x_i, 10, 100, 4)$ $y_i = 1 + \frac{x_i + 1}{4}$ $u(x_i, a, k, m) = \begin{cases} k(-x_i - a)^m & x_i < -a \\ 0 & -a < x_i < a \\ k(x_i - a)^m & x_i > a \end{cases}$	30	[-50, 50]	0
$f_9(x) = 0.1 \{ \sin^2(3\pi x_i) + \sum_{i=1}^n (x_i - 1)^2 [1 + \sin^2(3\pi x_i + 1)] + (x_n - 1)^2 [1 + \sin^2(2\pi x_n)] \} + \sum_{i=1}^n u(x_i, 5, 100, 4)$	30	[-50, 50]	0

表 2 测试函数寻优结果

函数	最优解	ISSA		SSA		PSO	
		最优值	平均值	最优值	平均值	最优值	平均值
f ₁	0	0	0	0	0	4.3343E+01	7.5614E+01
f ₂	0	0	0	0	0	3.0452E+02	9.7134E+02
f ₃	0	3.0418E-11	5.7012E-09	3.0171E-09	6.6998E-08	6.3097E+02	8.0095E+02
f ₄	0	4.0423E-06	1.3574E-05	3.4976E-05	5.2334E-05	1.3984E-01	2.2232E-01
f ₅	0	0	0	0	0	4.0196E+01	5.5386E+01
f ₆	0	8.8818E-16	8.8818E-16	8.8818E-16	8.8818E-16	6.3029E+00	6.3029E+00
f ₇	0	0	0	0	0	1.2047E+00	1.6368E+00
f ₈	0	2.8286E-13	1.7982E-11	1.2553E-11	8.0086E-11	6.7177E+00	9.9049E+00
f ₉	0	4.4861E-12	1.3636E-11	2.1736E-11	1.1318E-10	5.2703E+00	5.7679E+00

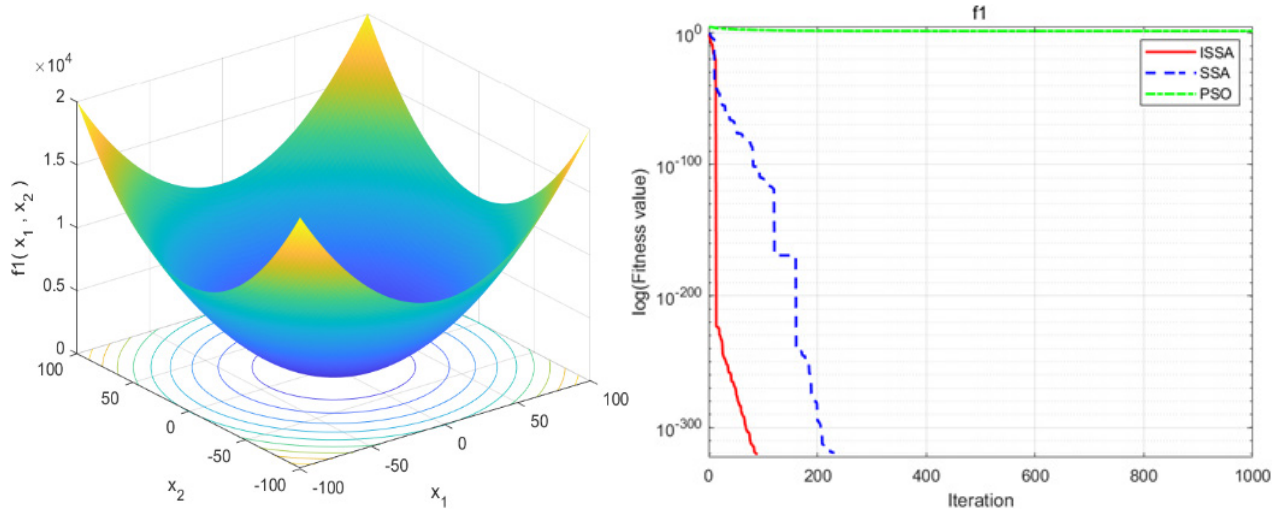
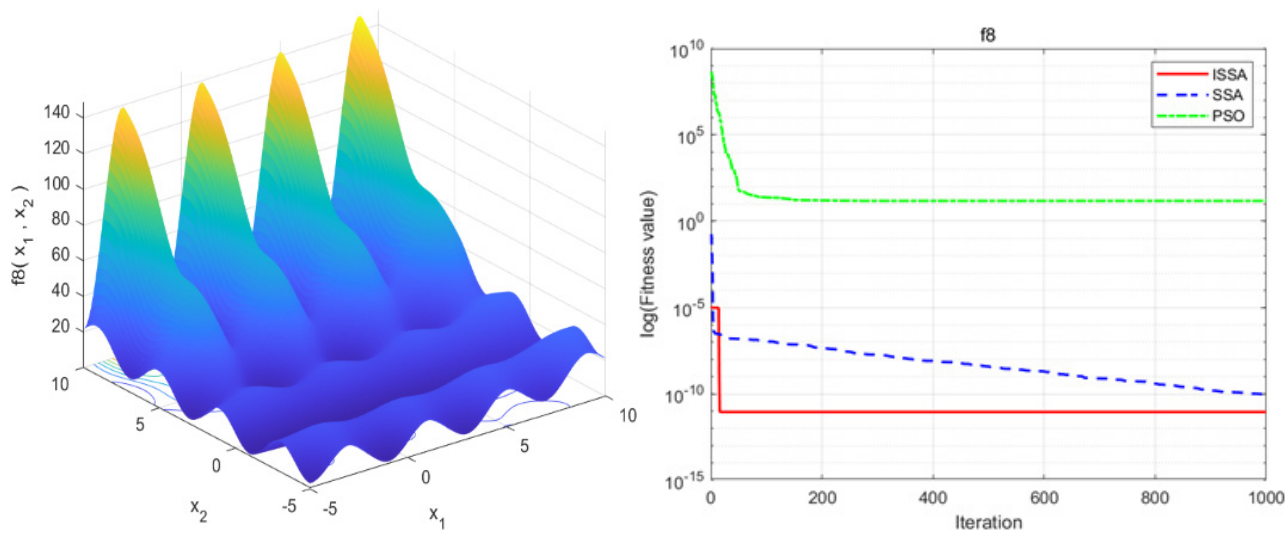
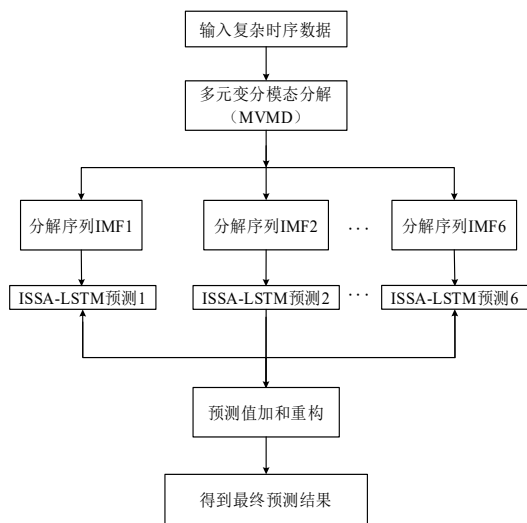
图3 f_1 三维图和适应度曲线图4 f_8 三维图和适应度曲线

图5 算法融合流程

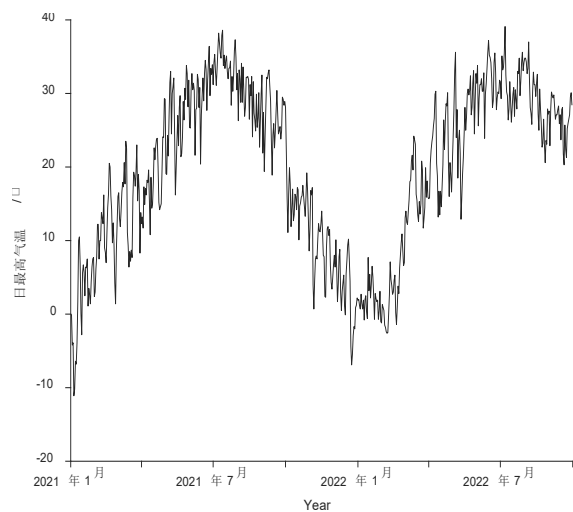


图 6 气温变化趋势

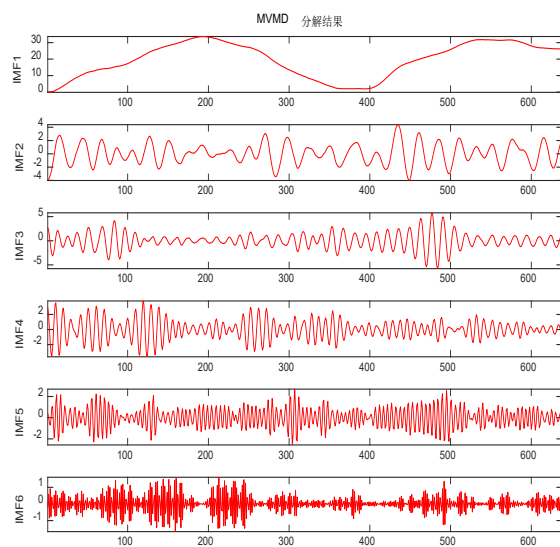


图 7 MVMD 分解各分量结果

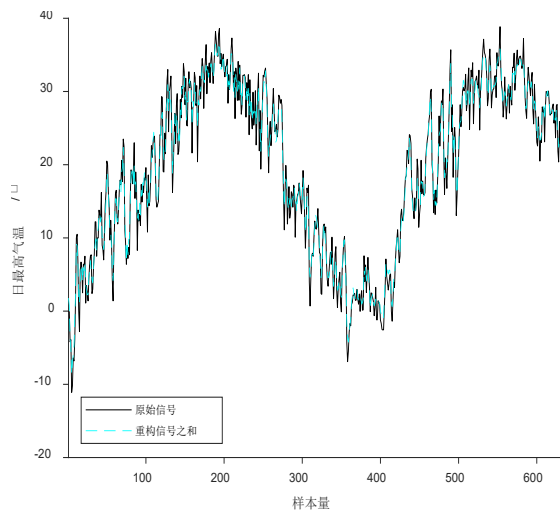


图 8 分解前后信号对比

表 3 各分量的对应的参数寻优结果

IMF	h	α	正则化系数
IMF1	159	0.0031	1E-04
IMF2	200	0.0024	1E-04
IMF3	109	0.0050	0.0049
IMF4	151	0.0056	1E-04
IMF5	115	0.0066	0.0034
IMF6	45	0.0054	1E-04

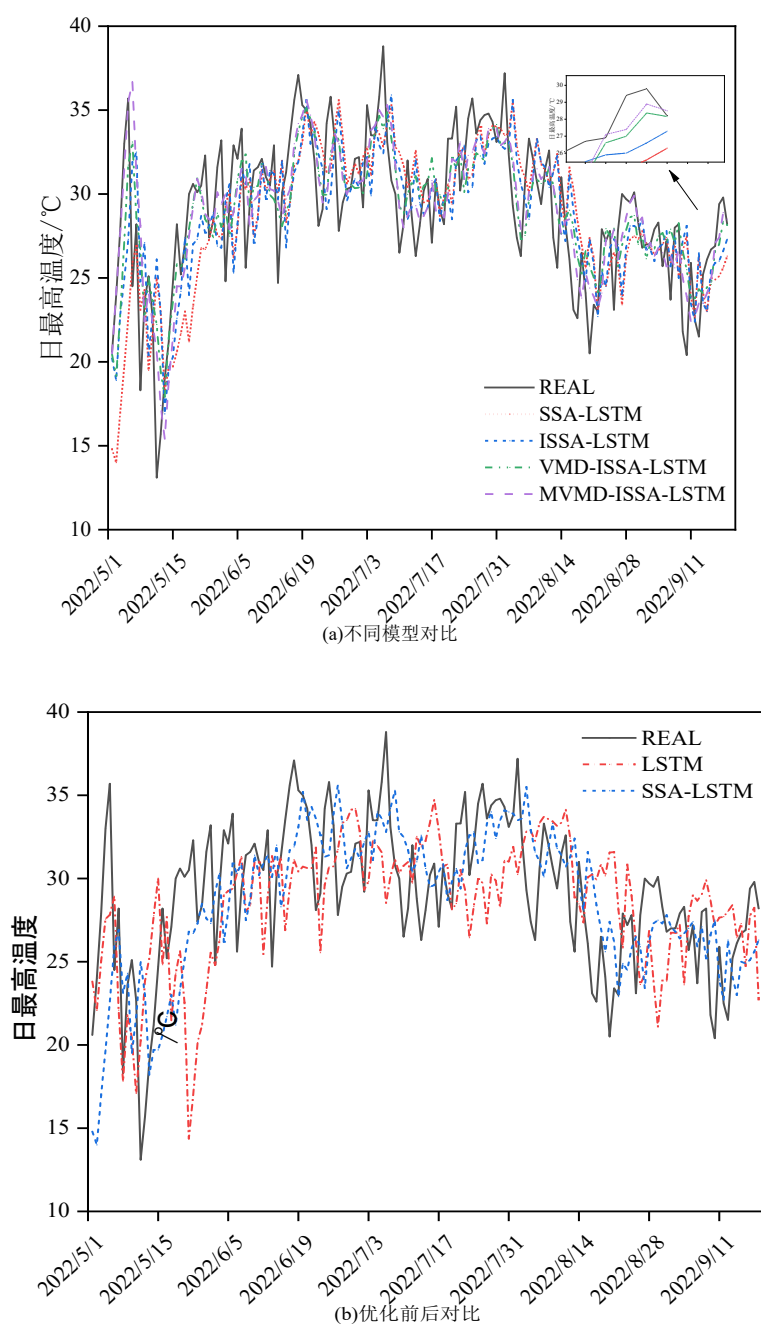


图 9 各模型预测曲线

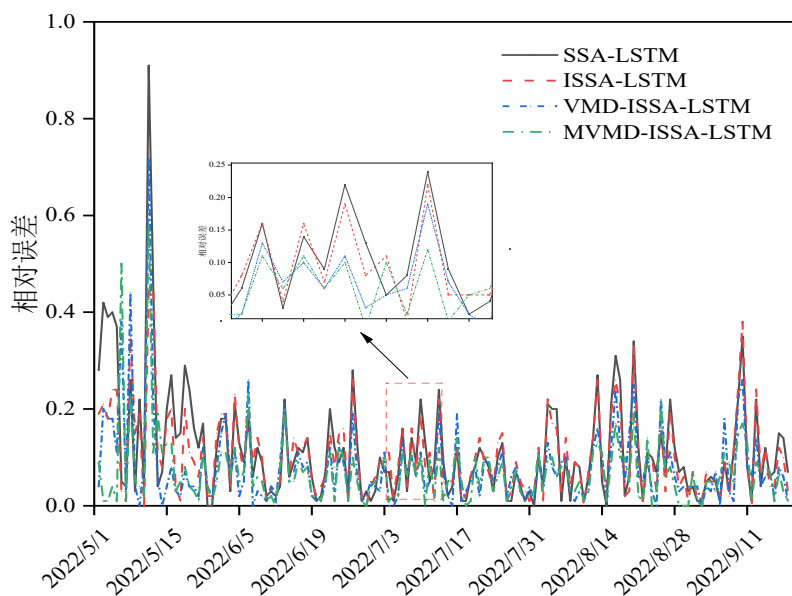


图 10 各模型相对误差结果

表 4 指标对比

算法	RMSE	MAE	MAPE (%)
MVMD-ISSA-LSTM	2.55	1.99	7.43%
VMD-ISSA-LSTM	2.78	2.13	7.90%
ISSA-LSTM	2.96	2.33	9.50%
SSA-LSTM	3.13	2.52	9.90%
LSTM	3.90	3.09	11.01%
CNN	4.02	3.25	11.51%

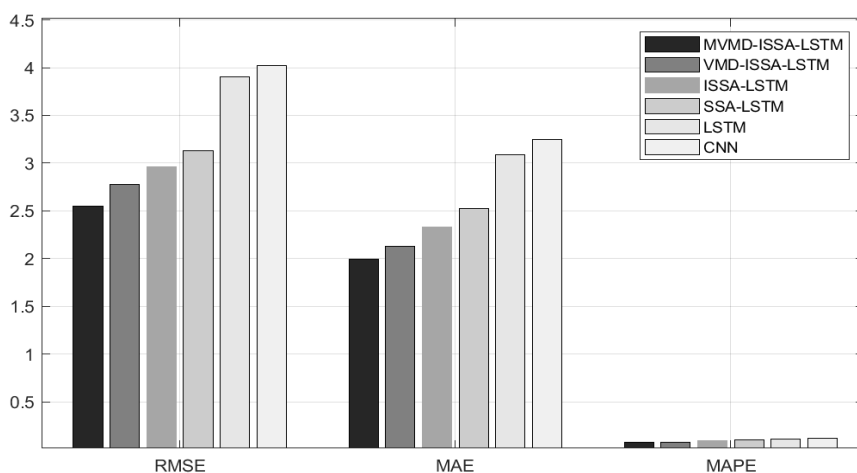


图 11 模型结果对比

5 结束语

针对复杂时间序列数据的非线性以及多噪声的特点, 本文提出了融合算法 MVMD-ISSA-LSTM。

通过测试函数对 ISSA 进行性能测试, 同时与未改进的 SSA、PSO 进行对比; 为了验证该融合算法的适用性, 以区域日最高气温复杂时序数据为实例, 与

VMD-ISSA-LSTM、ISSA-LSTM、SSA-LSTM、LSTM 以及 CNN 模型预测结果对比, 得出以下结论:

(1) ISSA 在经过测试函数的检验后与 SSA、PSO 进行对比, 其收敛速度更快, 模型准确性更高, 稳定性以及跳出局部最优能力都为最优。

(2) 与未降噪模型相比, 经过数据分解降噪的预测模型的预测精度明显提高, 分解后模型的优势更加显著; 同时 MVMD 作为改进的 VMD 方法分解数据后的预测效果比 VMD 误差更小。因此, 信号分解方法减弱了数据噪声, 使预测模型更好地学习复杂时序数据中的周期、趋势等本质特征, 提高预测精度。

从本文的实例中不难看出, 改进后的算法拥有更良好的性能, 并且不同的信号降噪方法对于模型的预测准确率也是有影响的。由此, 进一步研究更适合的算法改进策略、灵活选择去噪工具, 用以增强模型对于复杂时序数据有用信息的学习能力, 是未来提升预测准确度的重要着手点。

参考文献

- [1] BAUM L E, PETRIE T. Statistical Inference for Probabilistic Functions of Finite State Markov Chains[J]. The Annals of Mathematical Statistics, 1966, 37(6): 1554-1563.
- [2] WEN F H, XIAO J H, HE Z F, GONG X. Stock Price Prediction Based on Ssa and Svm[J]. Procedia Computer Science, 2014, 31: 625-631.
- [3] NGUYEN H T, NABNEY I T. Short-Term Electricity Demand and Gas Price Forecasts Using Wavelet Transforms and Adaptive Models[J]. Energy, 2010, 35(9): 3674-3685.
- [4] HUANG N E, SHEN Z, LONG S R, et al. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis[J]. Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, 1998, 454(1971): 903-995.
- [5] 李军宁, 罗文广, 陈武阁. 面向振动信号的滚动轴承故障诊断算法综述[J]. 西安工业大学学报, 2022, 42(2): 105-122.
- [6] DRAGOMIRETSKIY K, ZOZZO D. Variational Mode Decomposition[J]. IEEE Transactions on Signal Processing, 2014, 62(3): 531-544.
- [7] NAVEED U R, HANIA A. Multivariate Variational Mode Decomposition[J]. IEEE Transactions on Signal Processing, 2019, 67(23): 6039-6052.
- [8] ALAA S, MOSTAFA K. Time Series Forecasting of Petroleum Production Using Deep LSTM Recurrent Networks[J]. Neurocomputing, 2018, 323: 203-213.
- [9] ZHANG J S, XIAO X C. Predicting Chaotic Time Series Using Recurrent Neural Network[J]. Chinese Physics Letters, 2000, 17(2): 88-90.
- [10] QING X Y, NIU Y G. Hourly Day-Ahead Solar Irradiance Prediction Using Weather Forecasts By LSTM[J]. Energy, 2018, 148: 461-468.
- [11] XUE J k, SHEN B. A Novel Swarm Intelligence Optimization Approach: Sparrow Search Algorithm[J]. Systems Science & Control Engineering, 2020, 8(1): 22-34.
- [12] 单梁, 强浩, 李军, 等. 基于 Tent 映射的混沌优化算法[J]. 控制与决策, 2005(02): 179-182.
- [13] 黄敬宇. 融合 t 分布和 Tent 混沌映射的麻雀搜索算法研究[D]. 兰州大学, 2021: 27-32.
- [14] 王永贵, 曲彤彤, 李爽. 基于指数衰减惯性权重的分裂粒子群优化算法[J]. 计算机应用研究, 2020, 37(04): 1020-1024.
- [15] 武大硕, 张传雷, 陈佳等. 基于遗传算法改进 LSTM 神经网络股指预测分析[J]. 计算机应用研究, 2020, 37(S1): 86-87+107.
- [16] 魏勇召. 基于变分模态分解的机车轴承故障诊断[D]. 北京交通大学, 2018: 25-60.