



数据驱动的粮食产能组合预测模型

摘要

针对长短期记忆网络(LSTM)在粮食产能预测上存在超参数众多、长时序列信息丢失以及难以区分主次特征的问题,提出一种数据驱动的粮食产能组合预测模型.在超参数部分,通过引入动态权重和拉普拉斯变异的秃鹰算法(WLBES)对LSTM进行超参数寻优,避免了手动调参的过程.在预测部分,利用岭回归(RR)对预测结果进行残差修正,弥补LSTM数据丢失的缺陷;同时加入注意力机制,以权重大小区分主次特征,提升粮食产能相关性较大特征的重要性.研究表明,WLBES-LSTM-RR组合模型与LSTM模型和WLBES-LSTM模型相比,均方根误差(RMSE)分别下降了75%、19%,相较于其他优化LSTM的组合模型,RMSE大幅下降,该组合模型在粮食产能预测上具有更高的预测精度.

关键词

粮食产能预测;秃鹰优化算法;长短期记忆网络;拉普拉斯变异;注意力机制;残差修正

中图分类号 TP183

文献标志码 A

收稿日期 2023-04-24

资助项目 湖北省教育厅科技项目(B2020061)

作者简介

张岳,男,硕士生,研究方向为智能预测建模及应用.980224622@qq.com

陈为真(通信作者),女,博士,副教授,研究方向为智能预测建模及应用、智能计算与决策.38306783@qq.com

¹ 武汉轻工大学 电气与电子工程学院,武汉,430023

0 引言

联合国粮农组织发布的《2022 全球粮食危机报告》指出:在 2021 年有 53 个国家或地区约 1.93 亿人经历了粮食危机.在百年未有之大变局下,准确预测粮食产能有利于保障粮食安全,对国民经济、国家规划以及安全稳定有重要意义.

目前,关于粮食产能的预测大多采用统计模型^[1]、灰色理论^[2]、回归模型^[3]等.Verma 等^[4]采用基于天气参数的统计模型预测甘蔗产量,通过加权天气指数降低预测误差;Bouras 等^[5]开发了将遥感干旱指数与气候和天气变量相结合的早期机器学习预测模型,提高了预测性能;王秀^[6]通过粮食产量的主要因子,建立灰色 GM(1, N) 预测模型,更好地拟合了粮食产量的发展趋势;戎陆庆等^[7]使用 BP 神经网络对广西省粮食产量进行预测,证明了 BP 神经网络的预测可行性.这些算法对长时间数据具有一定的依赖性,使其预测能力受到限制.长短期记忆神经网络(Long Short-Term Memory, LSTM)克服了上述缺点^[8].郭佳丽等^[9]采用改进的 LSTM 对流量进行预测;戴秉佑等^[10]建立了多维高速公路路标温度的 LSTM 预测模型.单一模型面临超参数众多以及迟滞等问题,直接影响模型的预测性能,所以众多学者提出了组合预测模型^[11-15].高超等^[16]采用黑猩猩算法寻找 LSTM 的最优参数,提高短期负荷的预测精度;许美贤等^[17]利用粒子群算法优化 BP 神经网络和支持向量机,有效地提升了抗乳腺癌药物性质的预测性能;韩莹等^[18]融合了 LSTM 与宽度学习系统使模型在运算效率未降低的情况下提升了预测性能;黄琦兰等^[19]通过差分进化算法优化最小二乘支持向量机的主要参数,建立了粮食产量的预测模型.

上述预测方法未考虑到不同特征对预测结果有不同程度的影响,难以区分主次特征,也未考虑到 LSTM 网络层数的影响以及会造成信息丢失的情况.因此,本文在 LSTM 中加入注意力机制(Soft Attention Mechanism, SAM),来提升粮食产能相关性较大特征的重要性,运用改进后的秃鹰算法(Dynamic Weights and Laplacian variation of Bald Eagle Search Optimization Algorithm, WLBES)优化 LSTM 的超参数,对粮食产能进行预测,然后通过岭回归(Ridge Regression, RR)对预测结果进行残差修正,弥补 LSTM 本身信息丢失的情况.研究表明,与其他模型相比,WLBES-LSTM-RR 组合模型均方根误差(RMSE)、平均绝对误差(MAE)以及平均绝对百分比误差(MAPE)均为最优,说明

其具有更高的预测精度,有效地解决了粮食产能预测上难以区分主次特征和 LSTM 本身的问题。

1 基础框架

1.1 秃鹰优化算法(BES)

秃鹰优化算法是 Alsattar 等^[20]通过观察秃鹰的捕食过程提出的一种模拟秃鹰狩猎的群智能优化算法,算法分为 3 个阶段:

1) 选择阶段

秃鹰在选定的搜索空间内通过猎物的数量来确定搜索的位置,以全局最优位置、种群平均位置以及个体位置作为先验信息.该阶段的数学描述为

$$P_{i,\text{new}} = P_{\text{best}} + \alpha r (P_{\text{mean}} - P_i). \quad (1)$$

式中: α 是控制位置变化的参数,取值范围是 [1.5, 2]; r 是 (0,1) 的随机数; P_{best} 是到目前为止全局最优的位置; P_{mean} 是种群内所有个体的平均位置; P_i 是第 i 只秃鹰目前所在的位置。

2) 搜索阶段

秃鹰会在选定的搜索区域内加快搜索进程,以螺旋的方式朝不同的方向移动来搜索猎物,同时为了寻找最佳的俯冲捕捉位置,螺旋飞行的数学模型通过极坐标方程的方式进行位置更新.该阶段的数学描述为

$$P_{i,\text{new}} = P_i + y(i) \times (P_i - P_{i+1}) + x(i) \times (P_i - P_{\text{mean}}), \quad (2)$$

$$x(i) = \frac{xr(i)}{\max(|xr|)}, y(i) = \frac{yr(i)}{\max(|yr|)}, \quad (3)$$

$$\begin{cases} xr(i) = r(i) \times \sin(\theta(i)), \\ yr(i) = r(i) \times \cos(\theta(i)), \end{cases} \quad (4)$$

$$\begin{cases} \theta(i) = a \times \pi \times \text{rand}, \\ r(i) = \theta(i) + R \times \text{rand}. \end{cases} \quad (5)$$

式中: a 是控制螺旋飞行的角度参数,取值范围是 [5, 10]; R 是控制螺旋飞行的圈数参数,取值范围是 [0.5, 2]; $x(i)$ 和 $y(i)$ 是秃鹰在极坐标下的位置; rand 是 (0,1) 的随机数。

3) 俯冲阶段

秃鹰会从搜索阶段的最佳位置快速俯冲向猎物,并且所有的个体都会向全局最优的位置移动,而在俯冲阶段同样采用极坐标方程描述秃鹰在俯冲时的运动.该阶段的数学描述为

$$P_{i,\text{new}} = \text{rand} \times P_{\text{best}} + x_1(i) \times (P_i - c_1 P_{\text{mean}}) + y_1(i) \times (P_i - c_2 P_{\text{best}}), \quad (6)$$

$$x_1 = \frac{xr(i)}{\max(|xr|)}, y_1(i) = \frac{yr(i)}{\max(|yr|)}, \quad (7)$$

$$\begin{cases} xr(i) = r(i) \times \sinh[\theta(i)], \\ yr(i) = r(i) \times \cosh[\theta(i)], \end{cases} \quad (8)$$

$$\theta(i) = a \times \pi \times \text{rand}, r(i) = \theta(i), \quad (9)$$

式中: c_1 和 c_2 分别是秃鹰向全局最优位置和中心点的运动强度,取值范围是 [1, 2]。

综上所述,BES 在每一阶段结束均会进行适应度值的计算,并更新全局最优,其总共进行 3 次适应度函数的计算,所以相对于其他优化算法会拥有比较好的收敛速度以及优化性能.但是该算法本身存在全局搜索能力较差和容易陷入局部最优的问题。

1.2 长短期记忆网络(LSTM)

LSTM 是由 Hochreiter^[21]提出的针对循环神经网络(RNN)难以在长序列中传递远程信息而导致的梯度消失和梯度爆炸问题的一种变体.其核心概念在于细胞状态以及各种‘门’控结构,主要包含遗忘门、输入门以及输出门 3 个门控单元.LSTM 基本结构如图 1 所示。

细胞状态通过 3 个门控单元进行递归传递信息,从而实现细胞状态信息的更新以及信息的长期保存.遗忘门决定遗弃前一层无用的信息;输入门决定新输入的信息中有用的会被保留;输出门将经过遗忘门和输入门更新后的细胞状态以及隐藏状态整合作为当前时刻的输出.细胞状态以及各门的表达式如下:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (10)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (11)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \quad (12)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \quad (13)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (14)$$

$$h_t = o_t * \tanh(C_t), \quad (15)$$

式中: * 表示哈达马积。

LSTM 因为细胞状态的存在,能够在其他单元传递信息的过程中保持长时序特征或信息不变,解决了 RNN 的信息传递会随时间间隔太长而不断衰减的问题,因此,LSTM 更适合处理时间序列数据.但是 LSTM 对较长时序的数据仍存在丢失数据的情况。

2 WLBES-LSTM-RR 组合预测模型

组合预测模型的基本思路是将模型的主要影响因素和其他影响因素分为两部分预测,由 WLBES-LSTM 模型计算主要影响因素与粮食产能的映射关系,通过 RR 模型预测前一个模型中受其他影响因素影响的残差序列,将两部分的预测结果组合完成

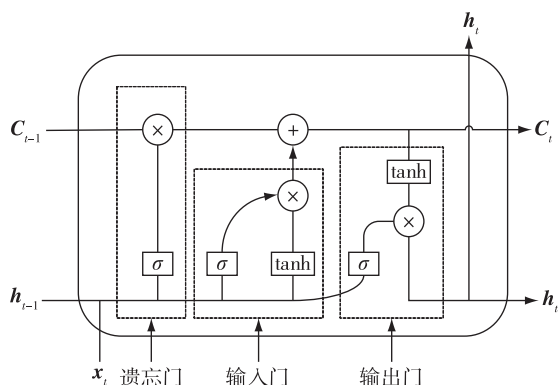


图1 LSTM 结构

Fig. 1 Structure diagram of LSTM

对粮食产能的预测。

2.1 超参数优化部分

LSTM 的超参数设置对模型的性能至关重要,所以提出一种 WLBS 算法对模型的超参数进行寻优,从而提高模型的预测精度. WLBS 算法相较于传统的 BES 算法有以下改进:1)搜索阶段加入 Levy 飞行来提高算法的全局搜索能力;2)俯冲阶段对全局最优位置融入拉普拉斯变异提升跳出局部最优的能力;3)通过拉普拉斯和高斯混合动态权重平衡全局搜索和局部搜索,加快收敛速度。

2.1.1 Levy 飞行

Levy 飞行^[22]以法国数学家保罗·莱维命名,指的是步长的概率分布为重尾分布的随机行走,主要是以小步移动为主,在其中会有大跨步的移动,这就使得随机行走时不会重复一个地方搜索,能够加大搜索的范围.尽管它的运动方向是随机性的,但是其运动的步长是按照幂次律分布的.搜索阶段位置式(2)更新为

$$P_{\text{new},i} = \alpha \otimes \text{Levy}(\beta) \times P_i + x(i) \times (P_i - P_{\text{mean}}) + \gamma(i) \times (P_i - P_{i+1}). \quad (16)$$

式中: α 是步长的控制量,此处设置为 0.01; $\text{Levy}(\beta)$ 是莱维飞行步长,其计算公式如下:

$$\text{Levy}(\beta) \sim \frac{u}{|v|^{-\beta}}, \quad (17)$$

式中: u, v 满足正态分布, $u \sim N(0, \sigma_u^2)$, $v \sim N(0, \sigma_v^2)$. σ_u, σ_v 定义如下:

$$\sigma_u = \left[\frac{\Gamma(1 + \beta) \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left(\frac{1 + \beta}{2}\right) \beta \times 2^{\frac{\beta-1}{2}}}\right]^{\frac{1}{\beta}}, \quad \sigma_v = 1, \quad (18)$$

式中: β 的取值范围是(0,2),此处取值为 1.5.

2.1.2 动态权重以及拉普拉斯变异

拉普拉斯分布^[23]是一种连续的概率分布,其中中心点两边近似由两个指数分布背靠背拼在一起,所以它又被称为双指数分布.它能够设置任意一点概率分布的峰值,概率密度函数公式如下:

$$\text{Lap}(x | \mu, \lambda) = \frac{1}{2\lambda} e^{-\frac{|x-\mu|}{\lambda}}, \quad (19)$$

式中: μ 是拉普拉斯的期望,并且在这一点达到极大值.

高斯分布^[24]又称正态分布,它与拉普拉斯一样都是一种连续的概率分布,并且它的图像分布与拉普拉斯近似,都是由中心点向两边下降,不过它是一种平滑的曲线,概率密度函数公式如下:

$$N(x | \mu, \sigma) = \sqrt{\frac{1}{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (20)$$

式中: μ 是高斯的期望,此点是分布的均值.

相比较高斯分布,拉普拉斯分布的随机变量出现极端大的值的概率要远远大于正态分布,借此选取拉普拉斯分布对全局最优位置进行变异,使之跳出局部最优更为有效,其式(6)中全局最优更新如下:

$$P_{\text{new,best}} = P_{\text{best}} + P_{\text{best}} \times \text{Lap}(0,1). \quad (21)$$

为了平衡前期全局搜索和后期局部搜索的能力,将高斯和拉普拉斯混合策略作为动态权重引入到俯冲阶段,结合式(21),俯冲阶段位置式(6)更新公式为

$$P_{\text{new},i} = w \times \text{rand} \times (P_{\text{best}} + P_{\text{best}} \times \text{Lap}(0,1)) + x_1(i) \times (P_i - c_1 P_{\text{mean}}) + y_1(i) \times (P_i - c_2 P_{\text{best}}), \quad (22)$$

$$w = \left(1 - \frac{t^3}{t_{\text{max}}^3}\right) \times \text{Lap}(1,2) + \frac{t^3}{t_{\text{max}}^3} \times \text{Gauss}(1,2) + 1, \quad (23)$$

式中: t 是当前迭代次数; t_{max} 是最大迭代数.因为拉普拉斯随机数相对高斯随机数会有更大的随机范围,有大概率出现极端大值,所以在迭代前期选取较大的权重系数在空间中探索更多的未知解以及更优解;而在迭代后期,拉普拉斯随机数权重系数变小,高斯随机数权重系数变大,能够在最优解附近搜索更优解,以增强秃鹰优化算法后期的局部搜索能力,并且对后期收敛速度影响比较小。

2.2 预测部分

粮食产能每个特征对预测结果的影响都是不同的,所以为了聚焦主要特征,降低次要特征的影响,在 LSTM 网络中加入注意力机制.除此之外, LSTM

网络本身对较长时序的数据仍会存在信息丢失的情况,利用岭回归的残差修正弥补 LSTM 模型的缺陷,同时还能够抑制残差带有的共线性问题。

2.2.1 注意力机制(SAM)

注意力机制广泛应用于自然语言处理、图像检测等方面,其本质是实现信息处理资源的高效分配。以高权重聚焦主要特征,以低权重降低次要特征的影响,使其能够在不同的情况下区分主次特征,从而提高预测的精度。SAM 计算公式^[25]如下:

$$\text{attention} = \text{softmax}(\mathbf{Q}\mathbf{K}^T)V, \quad (24)$$

式中: \mathbf{Q} 是查询向量, \mathbf{K} 是序列数据的地址, V 是 \mathbf{K} 中的数据值。

SAM 的计算步骤^[26]如下所示:

1) 根据 \mathbf{Q} 和 \mathbf{K} 计算两者相似度

$$S_{ij} = F(Q_i, k_j). \quad (25)$$

式中: S_{ij} 是数据 i, j 之间的相似度; F 是计算相似度的方法公式, 此处采用小型神经网络拟合相似度计算并结合步骤 2) 进行归一化。

2) 用 softmax 函数进行归一化

$$\alpha_{ij} = \text{softmax}(S_{ij}), \quad (26)$$

式中: α_{ij} 是进行归一化后的相似度; $\text{softmax}()$ 是归一化函数。

3) 加权求和

$$w_i = \sum \alpha_{ij}v_i, \quad (27)$$

式中: w_i 是经过加权求和后的权重。

步骤 1) 中计算 \mathbf{Q} 和 \mathbf{K} 的相似度所得到的数值取值范围不一样,就需要在步骤 2) 中进行 softmax 函数归一化,将原始计算的相似度归一化成权重和为 1 的概率分布,通过 softmax 函数的内在机制突出所需要的主要特征的权重,最后将所得到归一化后的相似度与数据值进行加权求和即可得到注意力权重。由此就可使主要特征得到高权重,将低权重分配给次要特征。

2.2.2 岭回归残差修正

岭回归^[27]实际上是一种改良的最小二乘法,它放弃了最小二乘法的无偏估计,损失部分信息,导致其方程的 R^2 通常会稍低于普通的最小二乘法,但其估计的偏回归系数就会更接近真实情况,从而提高回归模型的稳定性和可靠性,对于病态数据具有较好的效果。而且岭回归主要用于解决多重共线性,有助于提升对残差序列的预测精度。

一般的线性回归数学模型比较简单,在拟合上有比较广泛的应用,在求解线性回归问题中的最优

拟合参数是由最小二乘法决定,又称为最小二乘法回归模型。但是当它的输入矩阵不是列满秩或者一些列之间线性相关性比较大时,就会导致误差很大。传统的最小二乘法缺乏稳定性和可靠性,使其在实际的应用当中受到一定的限制。

$$\theta = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}, \quad (28)$$

式中: \mathbf{X} 为自变量, \mathbf{Y} 为 \mathbf{X} 对应的因变量。

岭回归就是在式(28)基础上增加一个 L2 正则化,解决最小二乘法的缺陷,加入岭回归系数对角矩阵,使矩阵的特征值变大,从而有效地提高稳定性,使其更加接近真实情况。岭回归的最优拟合参数为

$$\theta = (\mathbf{X}^T\mathbf{X} + \alpha\mathbf{I})^{-1}\mathbf{X}^T\mathbf{Y}, \quad (29)$$

式中: α 为岭回归系数, \mathbf{I} 为单位矩阵。当 $\alpha = 0$ 时就是最小二乘法。

针对残差序列中具有一定的共线性,通过岭回归对其共线性进行抑制,提升残差序列的预测精度,从而提升整体模型的预测精度,对 WLBES-LSTM 预测模型进行残差修正。

2.3 组合模型建立

组合模型能够利用不同的预测模型的优点弥补其他模型存在的缺陷, LSTM 网络具有良好的处理时间序列的能力,但是其众多超参数需手动设置,而 WLBES 算法具有较好的收敛速度以及寻优性能,用其对 LSTM 进行参数优化,有效地解决手动试参的过程。随后通过岭回归抑制 LSTM 网络预测的残差序列中存在一定程度共线性问题和弥补信息丢失的情况,实现模型之间的互补。WLBES-LSTM-RR 预测模型的流程如图 2 所示。

基于 WLBES-LSTM-RR 组合预测模型流程如下:

1) WLBES 超参数寻优。根据数据量的不同, LSTM 网络中 LSTM 层、Dense 层以及神经元的个数对模型结果具有较大的影响,所以将其作为 WLBES 算法的寻优超参数,通过不断的迭代计算出最佳超参数。

2) LSTM 训练与预测。选取与粮食产能相关性较强的特征作为输入,划分训练集与预测集;通过注意力机制以高权重聚焦主要特征,降低次要特征的影响;利用 WLBES 算法得到的最优超参数进行训练,分别得到训练集的预测值与实际值的残差以及预测集的预测值。

3) 岭回归残差修正。将通过 LSTM 网络计算得到的残差序列作为输入,划分训练集和预测集,然后采

用交叉验证法评估出最优的岭回归系数,建立岭回归预测,得到 LSTM 训练集残差的预测值。

4) 计算组合模型预测结果.组合模型的预测结果为 LSTM 预测结果与岭回归残差预测结果之和,并将 RMSE、MAE、MAPE 作为评价指标。

3 研究结果与分析

3.1 标准测试函数实验

为了验证 WLBES 的性能,选取表 1 所示的 10 个测试函数作为适应度函数进行测试,其中: $f_1 \sim f_5$ 为单峰函数,用来测试算法的收敛速度和收敛的精度; $f_6 \sim f_{10}$ 为多峰函数,用来测试算法的全局搜索和跳出局部最优的能力.为了使实验的结果具有客观性和公正性,将种群数设置为 30,最大迭代次数为 1 000,上下界根据测试函数的区间确定.选取鲸鱼优化算法 (Whale Optimization Algorithm, WOA)、粒子群优化算法 (Particle Swarm Optimization Algorithm, PSO)、正余弦优化算法 (Sine Cosine Algorithm,

SCA)、麻雀搜索算法 (Sparrow Search Algorithm, SSA)、灰狼优化算法 (Grey Wolf Optimizer, GWO) 以及 BES 算法与 WLBES 算法进行比较验证.所有算法的共有参数均保持一致,其他非共有参数如表 2 所示。

表 3 整理了表 1 中的测试函数经过 30 次独立运算后与其他经典智能优化算法的结果.结果表明:在单峰函数中,对于函数 $f_1 \sim f_5$, WLBES 均能够在 200 次以内收敛,寻到了函数最优值,平均值和标准差均为 0,尤其 f_1 和 f_3 , WLBES 在 100 次以内就能够收敛,而 SCA、GWO、WOA、SSA、PSO 均未能收敛寻到最优值, WLBES 表现更优;相对于 BES 算法, WLBES 算法在收敛速度上有了明显的提升;在 f_5 中, BES 算法出现未收敛的情况,而 WLBES 仍旧寻到了最优值,提升了 BES 算法的全局搜索能力.在所有的单峰函数中, WLBES 均能够收敛,且收敛次数最低,证明了 Levy 飞行提高全局搜索能力的有效性,极大地提升了 WLBES 的收敛速度和收敛精度。

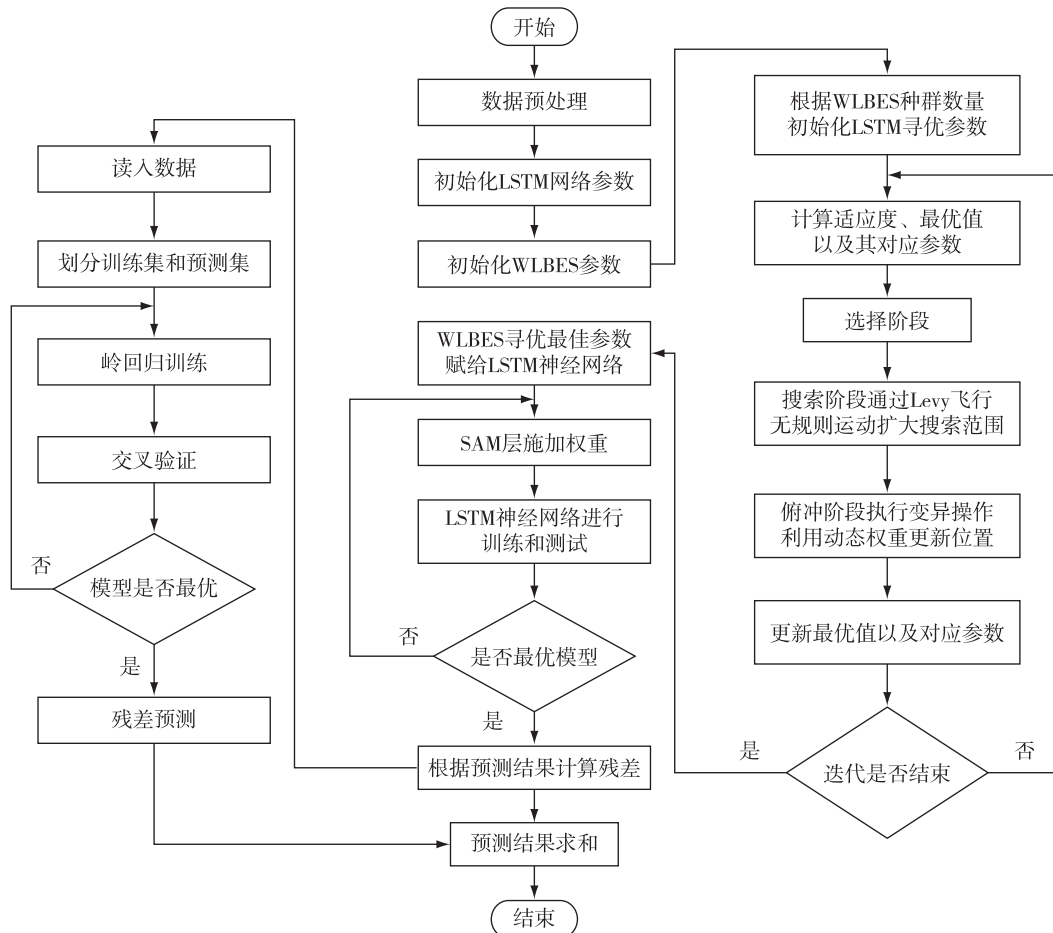


图 2 WLBES-LSTM-RR 预测模型流程

Fig. 2 Flow chart of WLBES-LSTM-RR forecasting model

表 1 测试函数
Table 1 Test functions

序号	函数	维度	区间	最优值
f_1	Sphere Function	30	[-100, 100]	0
f_2	Schwefel's Problem 2. 22	30	[-10, 10]	0
f_3	Schwefel's Problem 1. 2	30	[-100, 100]	0
f_4	Schwefel's Problem 2. 21	30	[-100, 100]	0
f_5	Step Function	30	[-100, 100]	0
f_6	Generalized Schwefel's Problem 2. 26	30	[-500, 500]	-12 569. 5
f_7	Eneralized Rastrigin's Function	30	[-5. 12, 5. 12]	0
f_8	Ackley's Function	30	[-600, 600]	0
f_9	Generalized Griewank's Function	30	[-50, 50]	0
f_{10}	Generalized Penalized Function 2	30	[-50, 50]	0

表 2 各算法参数
Table 2 Parameters for each algorithm

算法	参数
SCA	$a=2$
GWO	a 线性递减 2→0
WOA	a 线性递减 2→0, $b=1$
SSA	PR=20%, SD=10%, ST=0. 8
PSO	$c_1=c_2=2$, 自适应惯性权重 $w_{\max}=0. 8, w_{\min}=0. 4$

在多峰函数中,从收敛次数来看:函数 f_7 和 f_9 , WLBES均能够在3次以内收敛,寻到最优值,并且平均值和标准差均为0,SCA、GWO、WOA、SSA、PSO收敛速度与WLBES有较大的差距或未能收敛,BES能够在6次以内收敛,但仍与WLBES有一定的差距;对于函数 f_6 ,WLBES、BES和SSA相比其他3种算法能够寻到更优值;对于函数 f_8 ,WLBES、BES、WOA最优值、平均值以及标准差均相同,SSA最优值相同,但是具有较大的波动性,其他算法表现较差;对于函数 f_{10} ,WLBES领先SCA、GWO、WOA、SSA、PSO算法10个量级以上,相对于BES算法,寻优结果相同,但是其波动性更小更稳定,提升了BES算法跳出局部最优的能力.在所有的多峰函数中,WLBES在函数 f_7 、 f_9 以及 f_{10} 寻优中具有极大的优势,在 f_6 和 f_8 中优势较小,足以证明混合动态权重能够平衡全局搜索和局部搜索的能力和拉普拉斯变异的有效性,使WLBES能够更快地跳出局部最优.

WLBES在5个单峰函数中均表现出其优秀的寻优能力,证明该算法有极好的收敛速度以及收敛精度,能够以较短的次数寻到函数最优值;而在5个多峰函数中,其表现没有在单峰函数中突出,但是相比其他算法有着更好的全局搜索能力和跳出局部最

优的能力,从而证明了Levy飞行、拉普拉斯变异以及混合动态权重的有效性.

3.2 预测模型实验

3.2.1 输入变量与实验设置

实验测试环境基于Windows 10,运行内存8 GB,处理器为i5-8300H CPU @ 2.30 GHz,用Python3.8以及Pycharm2022.2.1进行仿真实验.

因为人均粮食产量能够体现出我国粮食的生产能力,所以选取人均粮食产量(kg/人)作为预测目标,将粮食总产量(万t)、总人口(万人)、农作物播种总面积(10^3 hm^2)、粮食单位面积产量(kg/(10^3 hm^2))、粮食作物播种面积(10^3 hm^2)、受灾面积(10^3 hm^2)以及成灾面积(10^3 hm^2)作为其影响因素,数据源自国家统计局1950—2021年的统计年鉴,其中:1950—2010年数据作为训练集,2011—2021年数据作为预测集.MAPE、RMSE以及MAE作为评价指标,公式如下:

$$e_{\text{MAPE}} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|, \quad (30)$$

$$e_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \quad (31)$$

$$e_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|. \quad (32)$$

因为数据集为小样本数据,所以LSTM层数和Dense层数分别设置为1~3层,神经元的个数范围为(1,256),利用WLBES进行寻优.通过多次实验取最优值,LSTM网络的迭代训练设置为100,批次大小设置为16,学习率采取自适应下降学习率.为了测试加入SAM层WLBES-LSTM预测模型的预测精度,同时验证WLBES对LSTM的寻优性能,选取BES-LSTM、WOA-LSTM、PSO-LSTM以及GWO-LSTM进行比较.

3.2.2 WLBES-LSTM 时序预测

对粮食相关数据进行处理后,通过各模型进行寻优得到预测结果(表4),可以看出GWO-LSTM、WOA-LSTM以及PSO-LSTM模型对粮食产能的预测结果皆不太理想,而WLBES-LSTM预测模型的RMSE、MAE、MAPE分别为23.798、23.293 kg/人和5%,其各项评价指标均为最优,但是误差还是较大,相对于BES-LSTM模型有较大的提升,这表明在相同的实验条件下,WLBES-LSTM模型对整体变化趋势反应更加灵敏,预测精度高于其他模型.

通过实验验证,结合表4和表5可看出加入SAM

表 3 WLBES 与其他优化算法结果比较

Table 3 Comparison between WLBES and other optimization algorithms

函数	算法	最优值	平均值	标准差	收敛数	函数	算法	最优值	平均值	标准差	收敛数
f_1	WLBES	0.00E+00	0.00E+00	0.00E+00	64.67	f_6	WLBES	-4 189.83	-3 046.68	803.29	
	BES	0.00E+00	0.00E+00	0.00E+00	93.80		BES	-4 189.83	-3 074.06	883.80	
	WOA	2.74E-96	5.68E-84	2.93E-83			WOA	-3 003.11	-2 597.22	135.82	
	PSO	3.97E-07	1.34E-02	5.40E-02			PSO	-2 083.12	-1 261.87	295.89	
	SCA	1.29E-37	2.37E-24	1.27E-23			SCA	-2 508.67	-2 225.15	139.35	
	SSA	0.00E+00	1.12E-14	5.92E-14			SSA	-4175.34	-3 044.44	210.00	
	GWO	8.89E-125	3.78E-118	6.80E-118			GWO	-3380.46	-2826.10	325.74	
f_2	WLBES	0.00E+00	0.00E+00	0.00E+00	129.37	f_7	WLBES	0.00E+00	0.00E+00	0.00E+00	2.53
	BES	0.00E+00	0.00E+00	0.00E+00	189.63		BES	0.00E+00	0.00E+00	0.00E+00	4.33
	WOA	2.66E-58	8.63E-53	4.27E-52			WOA	0.00E+00	0.00E+00	0.00E+00	464.03
	PSO	2.54E-03	1.10E-01	1.51E-01			PSO	8.24E-07	1.57E-01	3.72E-01	
	SCA	1.95E-23	3.80E-19	7.47E-19			SCA	0.00E+00	7.40E-01	2.76E+00	
	SSA	2.36E-204	1.45E-06	4.65E-06			SSA	0.00E+00	0.00E+00	0.00E+00	360.77
	GWO	1.29E-68	1.13E-66	3.04E-66			GWO	0.00E+00	0.00E+00	0.00E+00	155.43
f_3	WLBES	0.00E+00	0.00E+00	0.00E+00	64.50	f_8	WLBES	4.44E-16	4.44E-16	9.86E-32	
	BES	0.00E+00	0.00E+00	0.00E+00	94.97		BES	4.44E-16	4.44E-16	9.86E-32	
	WOA	1.79E-38	6.81E-26	3.63E-25			WOA	4.44E-16	4.44E-16	9.86E-32	
	PSO	8.04E-06	1.40E-02	3.04E-02			PSO	1.69E-04	8.17E-03	9.02E-03	
	SCA	1.03E-16	4.72E-10	2.18E-09			SCA	4.00E-15	2.44E-14	4.81E-14	
	SSA	0.00E+00	3.83E-11	2.00E-10			SSA	4.44E-16	2.39E-07	7.55E-07	
	GWO	9.53E-63	2.70E-52	1.45E-51			GWO	4.00E-15	4.35E-15	1.07E-15	
f_4	WLBES	0.00E+00	0.00E+00	0.00E+00	129.30	f_9	WLBES	0.00E+00	0.00E+00	0.00E+00	2.97
	BES	0.00E+00	0.00E+00	0.00E+00	188.03		BES	0.00E+00	0.00E+00	0.00E+00	5.07
	WOA	6.58E-25	2.78E-20	1.14E-19			WOA	0.00E+00	0.00E+00	0.00E+00	492.1
	PSO	7.06E-03	5.72E-02	5.99E-02			PSO	1.05E-07	6.76E-05	2.34E-04	
	SCA	3.96E-12	2.06E-07	5.68E-07			SCA	0.00E+00	4.03E-02	1.10E-01	
	SSA	0.00E+00	1.06E-07	5.51E-07			SSA	0.00E+00	6.85E-16	3.11E-15	
	GWO	1.58E-40	5.08E-37	1.12E-36			GWO	0.00E+00	1.06E-02	1.12E-02	
f_5	WLBES	0.00E+00	0.00E+00	0.00E+00	106.50	f_{10}	WLBES	1.35E-32	2.23E-21	1.19E-20	
	BES	0.00E+00	3.59E+01	3.61E+01			BES	1.35E-32	1.96E-01	3.31E-01	
	WOA	8.96E-06	8.33E-02	1.07E-01			WOA	6.18E-06	6.32E-02	6.31E-02	
	PSO	1.94E-07	2.18E-04	6.28E-04			PSO	2.10E-02	6.89E-02	3.61E-02	
	SCA	1.70E-01	3.70E-01	1.06E-01			SCA	1.80E-02	1.09E-01	6.34E-02	
	SSA	4.96E-12	3.33E-08	1.05E-07			SSA	3.50E-13	7.80E-10	1.16E-09	
	GWO	3.43E-07	8.37E-03	4.51E-02			GWO	2.06E-07	9.76E-07	4.74E-07	

层对不同的特征施加不同的权重后各个模型的预测精度明显有所提升, WLBES-LSTM 预测模型的 RMSE、MAE 以及 MAPE 分别减小了 15.776、16.646 kg/人和 3.6 个百分点, RMSE 和 MAE 减到了 10 kg/人以内,而其他模型均在 10 kg/人以上.说明通过 SAM 层对粮食产能主要特征施加高权重,能够有效提升该预测模型的预测精度,降低了次要特征对预测结果的影响.

表 4 基础 LSTM 模型预测结果评价价值

Table 4 Evaluation index values for base LSTM model prediction results

模型	RMSE/(kg/人)	MAE/(kg/人)	MAPE
WLBES-LSTM	23.798	23.293	0.050
BES-LSTM	38.021	37.772	0.081
GWO-LSTM	56.216	55.686	0.119
WOA-LSTM	44.786	44.219	0.094
PSO-LSTM	45.685	45.239	0.096

表 5 加入 SAM 层预测结果评价

Table 5 Evaluation index values of prediction results after adding SAM layer

模型	RMSE/(kg/人)	MAE/(kg/人)	MAPE
WLBES-LSTM	9.594	6.704	0.014
BES-LSTM	13.819	11.846	0.026
GWO-LSTM	19.702	19.119	0.041
WOA-LSTM	23.313	22.671	0.048
PSO-LSTM	15.759	12.882	0.028

3.2.3 组合模型预测结果分析

通过计算 WLBES-LSTM 预测模型的预测值与岭回归残差预测的结果之和,得到 2011—2021 年的粮食产能预测值.与此同时,利用支持向量回归模型 (Support Vector Regression, SVR)、差分自回归移动平均模型 (Autoregressive Integrated Moving Average Model, ARIMA)、LSTM 模型和 WLBES-LSTM 模型与其进行对比,以上 3 种 LSTM 模型中均加入了 SAM 层,从而得到 5 种预测模型的预测评价指标和平均运行时间.

表 6 5 种预测模型结果比较

Table 6 Performance comparison between the proposed model and 4 prediction models

预测模型	RMSE/(kg/人)	MAE/(kg/人)	MAPE	平均时间/s
SVR	32.07	31.47	0.067	1.982
ARIMA	30.94	29.97	0.064	8.658
LSTM	25.94	25.37	0.054	7.821
WLBES-LSTM	9.59	6.70	0.014	10.027
WLBES-LSTM-RR	6.47	5.73	0.012	10.083

由表 6 可以看出, WLBES-LSTM-RR 组合模型的各项评价指标均表现最优,从侧面反映出其拟合效果最好,精确度最高,而单一模型 SVR、ARIMA 和 LSTM 与之相比有较大差距,误差较大,充分证明了该组合模型的预测效果比单一模型和 WLBES-LSTM 模型更优. WLBES-LSTM-RR 组合模型的 RMSE 和 MAE 相比较 WLBES-LSTM 分别降低了 19%、13%,说明通过岭回归残差修正能够有效地降低误差,弥补 LSTM 信息丢失的情况,提高了组合预测模型的预测精度.同时,可以看出 WLBES-LSTM-RR 组合模型的平均运行时间最长,这是因为在寻优过程中 LSTM 的层数不同会对运行时间造成影响,但是极大地提升了预测精度,在粮食产能预测上,预测精度优先级较高,所以该运行速度在可接受范围内.

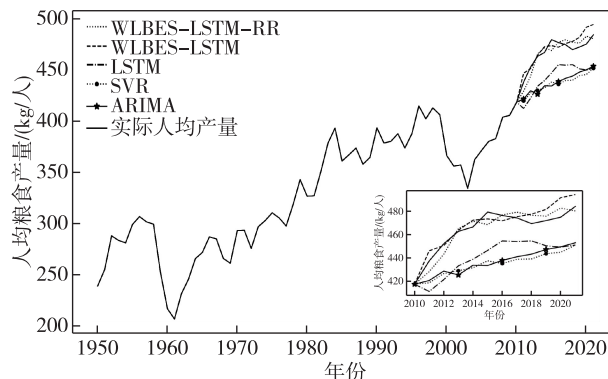


图 3 各模型预测结果

Fig. 3 Grain yields predicted by various models

为了让实验结果对比更明显,图 3 展示了各模型的整体预测结果,可以看出 WLBES-LSTM-RR 组合模型预测结果整体更加接近实际值,与实际值的折线吻合程度最好;而单一模型均在实际值的折线下方,不在实际值的区间范围,拟合程度与之相比具有一定的差距.由以上分析可知改进的秃鹰算法具有很好的寻优效果,加入的注意力机制能够充分地聚焦主要特征,而岭回归残差修正有效地弥补了 LSTM 模型本身的缺陷和对残差共线性的抑制.因此, WLBES-LSTM-RR 组合预测模型在粮食产能预测上具有更高的预测精度和更稳定的预测性能.

4 结论

本文提出一种数据驱动的 WLBES-LSTM-RR 组合预测模型对粮食产能进行预测.通过引入动态权重和拉普拉斯变异提升 BES 算法的寻优性能;针对预测粮食产能难以区分主次特征和 LSTM 本身存在的问题,在 LSTM 中加入注意力机制以权重大小区分主次特征;将改进的秃鹰算法 (WLBES) 用于 LSTM 网络参数寻优,最后通过岭回归残差修正进一步提升预测精度,从而实现粮食产能预测.研究表明, WLBES-LSTM-RR 组合预测模型对粮食产能的预测与单一模型和 WLBES-LSTM 相比, RMSE、MAE 以及 MAPE 最低,而且与实际值的折线吻合程度最好.表明本文模型具有更高的预测精度和更稳定的预测性能,并且预测误差最小.但是在模型训练、计算的过程中,如果时间跨度较大,在寻优的过程中选取的网络层数较多,就会导致计算所需时间较长,因此在今后的研究需要做更深入的探讨.

参考文献

References

- [1] 贾梦琦,蔡振江,胡建,等.基于机器学习的粮食产量预测模型研究[J].河北农业大学学报,2021,44(3):103-108
JIA Mengqi, CAI Zhenjiang, HU Jian, et al. Research on forecast model of grain yield based on machine learning [J]. Journal of Hebei Agricultural University, 2021, 44(3):103-108
- [2] 陈振坤,贾积身.基于分数阶灰色模型的河南省粮食产量预测[J].中国农机化学报,2022,43(6):135-141
CHEN Zhenkun, JIA Jishen. Prediction of grain yield in Henan province based on fractional grey model [J]. Journal of Chinese Agricultural Mechanization, 2022, 43(6):135-141
- [3] 郭晓婷.基于多元回归分析的安徽省粮食产量预测方法研究[D].合肥:安徽农业大学,2017
GUO Xiaoting. Prediction of grain yield in Anhui province based on multiple regression analysis method study [D]. Hefei: Anhui Agricultural University, 2017
- [4] Verma A K, Garg P K, Prasad K S H, et al. Sugarcane yield forecasting model based on weather parameters [J]. Sugar Tech, 2021, 23(1):158-166
- [5] Bouras E H, Jarlan L, Er-Raki S, et al. Cereal yield forecasting with satellite drought-based indices, weather data and regional climate indices using machine learning in Morocco [J]. Remote Sensing, 2021, 13(16):3101
- [6] 王秀.灰色预测模型粮食产量预测比较研究[J].农机化研究,2011,33(9):78-80
WANG Xiu. Comparison and research of grain production forecasting based on grey system forecasting model [J]. Journal of Agricultural Mechanization Research, 2011, 33(9):78-80
- [7] 戎陆庆,陈飞,欧阳浩.基于 GRA&BPNN 的广西粮食产量预测研究[J].中国农业资源与区划,2017,38(2):105-111
RONG Luqing, CHEN Fei, OUYANG Hao. Predicting grain yield of Guangxi province based on GRA & BPNN [J]. Chinese Journal of Agricultural Resources and Regional Planning, 2017, 38(2):105-111
- [8] 李洁,彭其渊,文超.基于 LSTM 深度神经网络的高速铁路短期客流预测研究[J].系统工程理论与实践,2021,41(10):2669-2682
LI Jie, PENG Qiyuan, WEN Chao. Short term passenger flow prediction of high speed railway based on LSTM deep neural network [J]. Systems Engineering—Theory & Practice, 2021, 41(10):2669-2682
- [9] 郭佳丽,邢双云,栾昊,等.基于改进的 LSTM 算法的时间序列流量预测[J].南京信息工程大学学报(自然科学版),2021,13(5):571-575
GUO Jiali, XING Shuangyun, LUAN Hao, et al. Prediction of time series traffic based on improved LSTM algorithm [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2021, 13(5):571-575
- [10] 戴秉佑,杨文臣,戢晓峰,等.冬季高速公路逐时路表温度 LSTM 预测模型[J].中国安全科学学报,2023,33(1):136-144
DAI Bingyou, YANG Wenchen, JI Xiaofeng, et al. Hourly road surface temperature LSTM prediction model of expressway in winter [J]. China Safety Science Journal, 2023, 33(1):136-144
- [11] 李嘉雨,廖如超,李钰楷.不平衡数据下基于 PSO-BP 算法的输电线路弧垂预测[J].南京信息工程大学学报(自然科学版),2021,13(5):576-581
LI Jiayu, LIAO Ruchao, LI Yukai. Transmission line sag prediction based on PSO-BP neural network with unbalanced data [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2021, 13(5):576-581
- [12] Tuexun W, Xu C, Guo H Y, et al. A wind power forecasting model using LSTM optimized by the modified bald eagle search algorithm [J]. Energies, 2022, 15(6):2031
- [13] Gundu V, Simon S P. PSO-LSTM for short term forecast of heterogeneous time series electricity price signals [J]. Journal of Ambient Intelligence and Humanized Computing, 2021, 12(2):2375-2385
- [14] 杨智勇,叶玉玺,周瑜.基于 BiLSTM-SA-TCN 时间序列模型在股票预测中的应用[J].南京信息工程大学学报(自然科学版),2023,15(6):643-651
YANG Zhiyong, YE Yuxi, ZHOU Yu. Application of BiLSTM-SA-TCN time series model in stock price prediction [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(6):643-651
- [15] 廖志豪,陈志钦,王长龙.基于 GA-BP 神经网络的广东淮山产量预测分析[J].农机化研究,2023,45(8):183-187
LIAO Zhihao, CHEN Zhiqin, WANG Changlong. Prediction and analysis of Guangdong Chinese yam yield based on GA-BP neural network [J]. Journal of Agricultural Mechanization Research, 2023, 45(8):183-187
- [16] 高超,孙谊嫒,赵洪峰,等.基于 ICOA-LSTM 的短期负荷预测研究[J].电子测量技术,2022,45(13):88-95
GAO Chao, SUN Yiqian, ZHAO Hongfeng, et al. Research on short-term load forecasting based on ICOA-LSTM [J]. Electronic Measurement Technology, 2022, 45(13):88-95
- [17] 许美贤,郑琰,李炎举,等.基于 PSO-BP 神经网络与 PSO-SVM 的抗乳腺癌药物性质预测[J].南京信息工程大学学报(自然科学版),2023,15(1):51-65
XU Meixian, ZHENG Yan, LI Yanju, et al. Prediction of properties of anti-breast cancer drugs based on PSO-BP neural network and PSO-SVM [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(1):51-65
- [18] 韩莹,管健,曹允重,等.LSTM-WBLS 模型在日降水量预测中的应用[J].南京信息工程大学学报(自然科学版),2023,15(2):180-186
HAN Ying, GUAN Jian, CAO Yunzhong, et al. Application of improved LSTM-WBLS model in daily precipitation forecast [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(2):180-186

- tion), 2023, 15(2): 180-186
- [19] 黄琦兰, 彭正昌. 基于 ADE-LSSVM 算法的粮食产量预测研究[J]. 粮食与油脂, 2021, 34(11): 36-40
HUANG Qilan, PENG Zhengchang. Research on grain yield prediction based on ADE-LSSVM algorithm[J]. Cereals & Oils, 2021, 34(11): 36-40
- [20] Alsattar H A, Zaidan A A, Zaidan B B. Novel meta-heuristic bald eagle search optimisation algorithm [J]. Artificial Intelligence Review, 2020, 53(3): 2237-2264
- [21] Hochreiter S, Schmidhuber J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780
- [22] Li J J, Wu Y H. Improved sparrow search algorithm with the extreme learning machine and its application for prediction [J]. Neural Processing Letters, 2022, 54(5): 4189-4209
- [23] 翟志波, 周鹏鹏, 闫汝山. 基于拉普拉斯分布与鲍德温效应的教与学算法及其应用[J]. 计算机集成制造系统, 2023, 29(8): 2611-2621
ZHAI Zhibo, ZHOU Pengpeng, YAN Rushan. Teaching and learning optimization algorithm based on Laplace distribution and Baldwin learning effect and its application [J]. Computer Integrated Manufacturing Systems, 2023, 29(8): 2611-2621
- [24] 曾争, 陈红芳, 黄友朋, 等. 基于正态分布优化的计量准确性方案研究[J]. 电测与仪表, 2019, 56(15): 127-131
ZENG Zheng, CHEN Hongfang, HUANG Youpeng, et al. Research on measurement accuracy scheme based on normal distribution optimization [J]. Electrical Measurement & Instrumentation, 2019, 56(15): 127-131
- [25] 胡梦, 李慧敏, 唐轶, 等. 基于双向长短时记忆网络和注意力机制的 RNA m5C 甲基化位点预测[J]. 中国生物化学与分子生物学报, 2023, 39(2): 303-310
HU Meng, LI Huimin, TANG Yi, et al. Prediction of RNA m5C methylation sites based on bi-directional long short-term memory and attention mechanism [J]. Chinese Journal of Biochemistry and Molecular Biology, 2023, 39(2): 303-310
- [26] 吉兴全, 曾若梅, 张玉敏, 等. 基于注意力机制的 CNN-LSTM 短期电价预测[J]. 电力系统保护与控制, 2022, 50(17): 125-132
JI Xingquan, ZENG Ruomei, ZHANG Yumin, et al. CNN-LSTM short-term electricity price prediction based on an attention mechanism [J]. Power System Protection and Control, 2022, 50(17): 125-132
- [27] 郑宪宝, 王宏伟, 苗志国, 等. 低渗油藏注入水构成岭回归量化评价方法[J]. 特种油气藏, 2022, 29(4): 128-134
ZHENG Xianbao, WANG Hongwei, MIAO Zhiguo, et al. Quantitative evaluation method with ridge regression for injected water composition in low-permeability reservoirs [J]. Special Oil & Gas Reservoirs, 2022, 29(4): 128-134

Data-driven grain productivity forecasting model

ZHANG Yue¹ CHEN Weizhen¹ CHEN Mengjiao¹

¹ School of Electrical and Electronic Engineering, Wuhan Polytechnic University, Wuhan 430023, China

Abstract To address the problem of numerous hyperparameters, loss of long time series information and difficulty in distinguishing primary and secondary features in Long Short-Term Memory network (LSTM) for grain yield capacity prediction, this paper proposes a combined data-driven grain capacity forecasting model. In the hyperparameter part, the proposed model performs hyperparameter search optimization for LSTM by introducing Dynamic Weights and Laplacian variation of Bald Eagle Search Optimization Algorithm (WLBES), to avoid the process of manual parameter adjustment. In the prediction part, the proposed model uses Ridge Regression (RR) to correct the residuals of the prediction results to make up for the deficiency of LSTM data loss, and adds an attention mechanism to distinguish primary and secondary features by weight size to enhance the importance of features with greater relevance to grain production. The results show that the combined WLBES-LSTM-RR model decreases the root mean square error (RMSE) by 75% and 19% compared with the LSTM and WLBES-LSTM models, respectively, and substantially decreases the RMSE compared with other combined models of optimized LSTM. This combined model has higher prediction accuracy in grain yield capacity prediction.

Key words grain production capacity forecast; bald eagle search optimization algorithm; long short-term memory (LSTM); Laplacian variation; attentional mechanism; residual correction