

赵瑞晓,符学龙,刘欢. 基于 MADF-DE-GRU 的瓦斯浓度动态预测[J]. 智能计算机与应用,2024,14(8):115-120. DOI:10.20169/j.issn.2095-2163.240819

基于 MADF-DE-GRU 的瓦斯浓度动态预测

赵瑞晓,符学龙,刘欢

(江苏财经职业技术学院 智能制造实训中心,江苏 淮安 223003)

摘要:为保障煤矿回采工作面安全生产,提高瓦斯浓度预测精度,本文提出一种基于门控循环单元(GRU)神经网络的瓦斯浓度动态预测模型。首先,用差分法改进的移动平均滤波器(MAF)去除原始数据中的噪声和趋势;其次,利用差分进化算法(DE)对GRU神经网络的隐藏层数、隐藏层神经元数量、时间步长和迭代次数等超参数进行寻优;最后,通过优化后的超参数搭建瓦斯浓度预测模型(MADF-DE-GRU)。通过仿真实验,并与多种预测模型进行对比,结果表明MADF-DE-GRU模型具有较高的预测精度,预测效果较好,能够应用到实际生产中。

关键词: 瓦斯浓度; 移动平均滤波器; 门控循环单元; 差分进化算法

中图分类号: TP399

文献标志码: A

文章编号: 2095-2163(2024)08-0115-06

Dynamic prediction of gas concentration based on MADF-DE-GRU

ZHAO Ruixiao, FU Xuelong, LIU Huan

(Intelligent Manufacturing Training Center, Jiangsu Polytechnic of Finance and Economics,
Huai'an 223003, Jiangsu, China)

Abstract: In order to ensure the safe production of coal mine back face and improve the accuracy of gas concentration prediction, a dynamic prediction model of gas concentration based on gated recurrent unit (GRU) neural network is proposed. Firstly, a moving average filter (MAF) improved by difference method is used to remove the noise and trend in the original data; then the hyperparameters such as the number of hidden layers, the number of neurons in the hidden layer, the time step and the number of iterations of the GRU neural network are used to find the optimal ones by using the differential evolution (DE) algorithm; and finally, a gas concentration prediction model (MADF-DE-GRU) is constructed by optimized hyperparameters, and the experimental results comparing with many kinds of prediction models, the MADF-DE-GRU model has higher prediction accuracy and better prediction effect, and can be applied to practical production.

Key words: gas concentration; moving average filter; gated cyclic unit; differential evolution algorithm

0 引言

煤矿瓦斯事故不仅影响煤矿安全生产,且破坏性强,人员伤亡多,造成的经济损失大^[1]。中国煤矿百万吨死亡率由2015年的0.162下降到2018年的0.093,首次降至0.1以下,但中国煤矿灾害防治面临的形势依然异常严峻复杂^[2]。为防范煤矿瓦斯灾害,保障煤矿的安全生产,对瓦斯浓度进行准确地预测有着重要的意义^[3-4]。近年来,随着大数据和深度学习理论的发展,已有研究人员将深度学习算法应用于煤矿瓦斯时间序列预测。相比人工神经

网络,最小二乘支持向量机等传统的时间序列预测方法,深度学习算法能够更好地挖掘出时间序列的内在联系,使得预测结果更加准确^[5-7]。程子均等^[8]利用长短期记忆网络与全连接神经网络搭建瓦斯浓度预测模型,取得了优于其它神经网络的预测效果;李伟山等^[9]利用长短期网络设计了煤矿瓦斯预测系统;张雨^[10]将深度置信网络用于煤矿瓦斯预测系统中;钱建生等^[11]提出了深度学习耦合粒子群优化支持向量机(Support Vector Machine, SVM)的瓦斯浓度预测方法,利用深度学习的相关理论学习瓦斯浓度数据特征,取得了良好的效果;Lyu

基金项目: 江苏省职业教育“双师型”教师团队培育项目。

作者简介: 符学龙(1980-),男,博士,副教授,主要研究方向:智能制造技术,先进复合材料;刘欢(1991-),男,硕士,助教,主要研究方向:智能制造。

通讯作者: 赵瑞晓(1996-),女,硕士,助教,主要研究方向:人工智能,计算机应用。Email:1978915838@qq.com

收稿日期: 2023-10-10

等^[12]提出了一种基于 LSTM (Long Short-Term Memory) 编码器/解码器的瓦斯浓度预测方法,通过与其他传感器的信息融合,取得了较好的预测效果。

长短期记忆网络 (LSTM) 是目前在瓦斯浓度预测方面运用较多的深度学习方法,相比 LSTM,门控循环单元 (Gated Recurrent Unit, GRU) 能够达到相同甚至更好的效果,并且 GRU 更容易进行训练,能够很大程度上提高训练效率,因此可将 GRU 运用于煤矿瓦斯浓度预测^[13-14]。此外,人为设置并调试深度学习模型不仅耗时耗力,而且还很难确定多次调试的模型是否是最佳预测模型。为解决这一问题,利用差分进化 (Differential Evolution, DE) 算法的全局寻优能力对 GRU 神经网络的隐藏层数、隐藏层神经元数量、时间步长和迭代次数等参数进行寻优^[15]。

由于煤矿回采工作面生产环境复杂,影响瓦斯浓度变化的因素较多,矿井中采集的瓦斯浓度数据大都存在噪声,具有非稳定性,并且可能表现出不同的趋势。传统的移动平均滤波 (MAF) 可以消除噪声使数据平滑,但无法消除趋势。针对这一问题,本文提出在 MAF 中加入差分操作,以去除数据中存在的可能趋势。综上所述,本文提出了一个基于 GRU 神经网络,并集合了消除噪声、去除趋势和超参数寻优等方法的瓦斯浓度预测模型。

1 门控循环单元

循环神经网络 (Recurrent Neural Network, RNN) 是常规前馈神经网络的扩展,能够处理可变长度序列输入。RNN 通过循环隐藏状态单元来处理可变长度序列,但是 RNN 在训练时会出现梯度消失和梯度爆炸等问题,因此 RNN 很难捕获数据的长期依赖^[16]。为解决 RNN 梯度消失的难题, Hochreiter^[17]提出了 LSTM, LSTM 是一种门控循环神经网络,是 RNN 的一种特殊表现形式,能够很好地捕获数据之间的长期依赖。

GRU 也是一种门控循环神经网络,属于 RNN 的一种表现形式。GRU 和 LSTM 很相似,在 GRU 中只有两个“门”,即重置门和更新门,具体的 GRU 单元结构如图 1 所示。其中更新门 u 取代了 LSTM 中的输入门和遗忘门,决定要遗忘哪些历史信息 and 需要添加的新的输入信息,重置门 r 也是用来决定要忘记多少历史信息的门。相比 LSTM, GRU 少了一个门,因此少了很多需要控制的参数,所以处理同样的任务,GRU 的训练效率要比 LSTM 高。

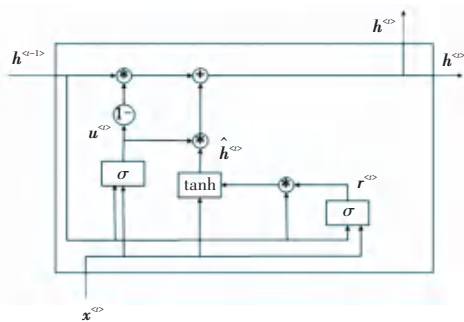


图 1 GRU 单元结构

Fig. 1 GRU unit structure

在 GRU 单元中, t 时刻更新门和重置门的值可表示为 $u^{<t>}$ 和 $r^{<t>}$, GRU 的输出值为 $h^{<t>}$, 其计算如公式(1)~公式(4)所示:

$$u^{<t>} = \sigma(W_u[h^{<t-1>}, x^{<t>}] + b_u) \quad (1)$$

$$r^{<t>} = \sigma(W_r[h^{<t-1>}, x^{<t>}] + b_r) \quad (2)$$

$$\hat{h}^{<t>} = \tanh(W_h[r^{<t>} * h^{<t-1>}, x^{<t>}] + b_h) \quad (3)$$

$$h^{<t>} = (1 - u^{<t>}) * h^{<t-1>} + u^{<t>} * \hat{h}^{<t>} \quad (4)$$

其中, W 和 b 为门的权重和偏置向量, σ 表示 sigmoid 函数。

在深度学习中,增加神经网络的深度是提高算法模型整体性能的有效方法^[18]。之前设计算法模型时采用的多为普通的三层神经网络,即输入层、隐藏层和输出层。三层结构的神经网络所涉及的参数较少,训练方便,但相应的在预测精度上可能不如具有深层次隐藏层的预测模型。图 2 所示模型就是一个具有输入层,输出层和 L 层隐藏层的深度 GRU 模型, $h_L^{<t-1>}$ 表示第 L 层 GRU 单元在 $t-1$ 时刻的输出值, $h_L^{<t>}$ 表示第 L 层 GRU 单元在 t 时刻的输出值,同时这个值也是下一层的输入值。

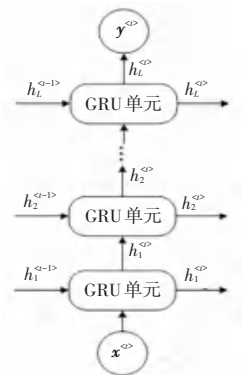


图 2 深度 GRU 模型

Fig. 2 Deep GRU model

2 差分进化算法

差分进化(DE)算法是一种随机的启发式并行搜索算法,简单易用,有较强的鲁棒性和全局寻优能力^[19]。Kachitvichyanukul^[20]通过实验对比了 DE 算法、遗传算法和粒子群算法 3 个进化算法的优化性能,DE 算法性能最优,且收敛速度快、算法稳定,多次寻优均能收敛到同一个解。

DE 算法的执行步骤主要包括初始化、变异、交叉和选择。初始化时,需要定义种群规模 NP , 种群个体向量的维度 D 。每个个体可以表示为

$$x_{i,G}(i = 1, 2, \dots, NP) \quad (5)$$

其中, G 为进化代数, $x_{i,G}$ 表示第 G 代种群中序列为 i 的个体。

种群的随机生成则可表示为

$$x_{ij,0} = \text{rand}[0, 1] \times (x^U - x^L) + x^L \quad (6)$$

其中, $i = 1, 2, \dots, NP; j = 1, 2, \dots, D; x_{ij,0}$ 表示初始种群中第 i 个体在维度 j 上的值; $\text{rand}[0, 1]$ 表示随机产生 $[0, 1]$ 之间的数; x^U, x^L 分别表示参数变量的上下界。

对于每个目标个体 $x_{i,G}(i = 1, 2, \dots, NP)$, 变异个体的产生方式可表示为

$$v_{i,G+1} = x_{r1,G} + F \times (x_{r2,G} - x_{r3,G}) \quad (7)$$

其中, $r1, r2, r3 \in [1, 2, \dots, i - 1, i + 1, \dots,$

$NP]$, $r1, r2, r3$ 为随机选取且互不相同; F 表示变异算子, $F \in [0, 2]$ 。

变异完成后,将变异个体与目标个体交叉得到实验个体,得到的实验个体即为交叉个体,交叉个体的确定方式为

$$u_{ij,G+1} = \begin{cases} v_{ij,G+1}, & \text{randb}(j) \leq CR \text{ 或 } j = \text{rnbr}(i) \\ x_{ij,G+1}, & \text{其它} \end{cases} \quad (8)$$

其中, CR 表示交叉算子,且 $CR \in [0, 1]$; $\text{randb}(j)$ 表示产生 $[0, 1]$ 之间随机数发生器的第 j 个估计值; $\text{rnbr}(i)$ 表示从 $(1, 2, \dots, D)$ 随机取一个值,用来确定 $u_{i,G+1}$ 至少从 $v_{i,G+1}$ 获得一个参数。

交叉个体产生后,利用目标评估函数评估其函数值,若函数值比目标个体的函数值小,则实验个体替代目标个体加入种群成为下一代目标个体:

$$x_{i,G+1} = \begin{cases} u_{i,G+1}, & f(u_{i,G+1}) \leq f(x_{i,G}) \\ x_{i,G}, & \text{其它} \end{cases} \quad (9)$$

其中, f 为目标评估函数,可根据不同的求解问题定义不同的函数。

3 基于 MADF 和 DE 算法优化的 GRU 预测模型

本文采用 DE 优化后的深度 GRU 模型,如图 3 所示。模型的主体包括输入层的数据预处理和 DE 对 GRU 进行超参数寻优两部分。

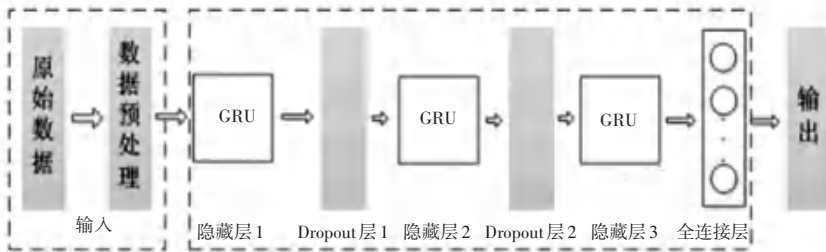


图 3 深度 GRU 预测模型框架

Fig. 3 Deep GRU prediction model framework

3.1 数据预处理

本文使用的数据是煤矿回采工作面的原始瓦斯监测数据,这些监测数据通常可能包含噪声。此外,瓦斯浓度时间序列数据集具有非稳定性,并且可能表现出不同的趋势,这意味着数据的统计特征,如方差和均值会随时间交替变化,所以在预测模型的训练中不宜使用原始监测数据。因此,本文将对数据进行处理。

(1) 消除噪声,稳定趋势。移动平均滤波器是数字信号处理中常用的数据处理方法,能够使原始监测数据平滑并消除任何可能的噪声。瓦斯浓度数据是非

稳定数据,在一定程度上可能会表现出特定趋势,但稳定数据更容易建模,并且很有可能使预测更加精确。

本文设计的移动平均差分滤波器 (Moving Average Difference Filter, MADF) 提供时间序列数据中过去数据点的加权平均,以生成时间序列的平滑估计,删除了数据中的趋势属性。在之后的预测中,再将趋势重新添加到数据中,以便将预测数值回归到原始尺度,并计算可比较的预测误差。移动平均差分滤波器的计算为

$$\text{diff}[t] = \frac{1}{N} \left(\sum_{i=1}^N x_{t-i+1} - \sum_{i=1}^N x_{t-d-i+1} \right) \quad (10)$$

其中, N 表示时间序列中每次采样数据的长度, d 为序列的时间延迟。

(2) 数据归一化。在深度学习中, 对数据集进行归一化操作是最常用也是非常必要的预处理步骤, 归一化操作不仅可以消除量纲, 将数据统一到相同的尺度上, 还可以提升模型的收敛速度和预测精度。本文利用归一化将相空间域数据集和原始混沌时间序列数据集都统一到 $(0, 1)$ 范围之内。归一化函数可表示为

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (11)$$

(3) 数据划分。将经过上述步骤处理好的数据按照简单交叉验证法划分为训练集和测试集, 取前 90% 为训练集, 后 10% 为测试集。

3.2 搭建 DE-GRU 模型

不同于经验法和试错法等常规的神经网络搭建方法, DE 算法优化 GRU 的过程是寻找最佳 GRU 网络模型的过程, 在寻优的过程中 DE 算法会确定 GRU 网络模型的隐藏层数、每一层隐藏层的隐藏单元数量、时间步长以及训练数据集在网络中的迭代次数。在整个求解过程中, 最优参数的评价标准是使得模型在训练集和验证集上的误差最小, 因此本文定义 DE 算法的目标评估函数为

$$f = \frac{1}{2} \left[\sum_{i=1}^N (y_i - \hat{y}_i)^2 + \sum_{j=1}^M (y_j - \hat{y}_j)^2 \right] \quad (12)$$

其中, N 表示训练集长度; \hat{y}_i 表示训练集中第 i 个预测值; y_i 表示对应的真实值; M 表示验证集长度; \hat{y}_j 表示验证集中第 j 个预测值; y_j 表示对应的真实值。

DE 优化 GRU 算法流程如下:

(1) 初始化 DE 参数。确定种群规模 NP , 个体向量维度 D , 变异算子 F , 交叉算子 CR , 最大进化迭代次数 G ;

(2) 种群随机初始化。每一个目标个体可表示为 $x_{i,0}(h_1, h_2, h_3, t, n)$, h_1, h_2, h_3 分别为第一、第二和第三层隐藏层隐藏单元数量, t 为时间步长, n 为迭代次数;

(3) 评价初始种群。利用式 (10) 计算初始种群中每一个目标个体的评估函数值;

(4) 判断是否达到最大进化代数。若是, 则 DE 算法寻优终止, 输出最佳个体, 获得模型最优解; 否则, 继续下一步操作;

(5) 变异、交叉和选择。进行变异和交叉操作, 得到临时种群, 并利用式 (10) 计算临时种群中每个

个体的评估函数值, 对临时种群和原始种群的相对对应个体进行比较选择, 产生新的种群;

(6) 进化代数, 转步骤 (4);

(7) 在 DE 算法完成模型寻优之后, 将最佳 GRU 模型输出, 此时将测试数据输入预测模型, 即可进行瓦斯浓度时间序列预测。为评价预测误差, 除了采用常用的均方根误差 ($RMSE$)、平均绝对误差 (MAE) 和平均绝对百分比误差 ($MAPE$) 外, 再引入一个在评价时间序列预测中常用的均方根百分比误差 ($RMSPE$), 如公式 (13) ~ 公式 (16) 所示:

$$RMSE = \sqrt{\frac{1}{I} \sum_{i=1}^I (y_i - \hat{y}_i)^2} \quad (13)$$

$$MAE = \frac{1}{I} \sum_{i=1}^I |y_i - \hat{y}_i| \quad (14)$$

$$MAPE = \frac{1}{I} \sum_{i=1}^I \frac{|\hat{y}_i - y_i|}{y_i} \times 100\% \quad (15)$$

$$RMSPE = \sqrt{\frac{1}{I} \sum_{i=1}^I \left| \frac{y_i - \hat{y}_i}{y_i} \right|^2} \times 100\% \quad (16)$$

其中, I 为预测样本数, y_i 和 \hat{y}_i 分别为预测值和真实值。

3.3 模型其它设置

在实际的工程运用中, 对瓦斯浓度进行动态预测更符合煤矿安全生产的需求, 在瓦斯浓度监测系统中, MADF-DE-GRU 可以根据实时的监测数据来预测未来一段时间内的瓦斯浓度变化。在动态预测中, 采取单步预测方式, 在每个时间步更新预测模型, 并从测试数据中插入新的观察值。简而言之, 动态方案使用因变量的前一个预测值来计算下一个。

在深度 GRU 模型中, 为了防止模型在训练时出现过拟合, 提升模型的鲁棒性和泛化能力, 在第一层和第二层 GRU 隐藏层后加入 Dropout 机制。

4 仿真实验及分析

4.1 实验数据

实验数据来源于邢台矿某回采工作面实际监测数据, 共计 1 200 组瓦斯浓度样本数据。该工作面属 II 类自燃煤层, 煤尘具有爆炸危险性, 且爆炸指数较高。瓦斯浓度的变化对工作面的采掘进度影响较大, 需加强通风管理, 防止瓦斯积聚。对瓦斯浓度进行预测, 可以有效的了解瓦斯浓度的变化趋势, 为工作面采掘提供预警。

4.2 基于 RNN 和门控 RNN 的预测对比

RNN 和门控 RNN (包括 LSTM 和 GRU) 在时间

序列预测方面都有着较好的表现,为了验证 GRU 神经网络在瓦斯浓度预测方面的性能,用 DE 算法对 RNN、LSTM 和 GRU 进行同样的超参数寻优操作,3

种模型所用数据皆用 MADF 进行预处理,具体对比结果见表 1。

表 1 GRU, LSTM 和 RNN 预测对比

Table 1 GRU, LSTM and RNN prediction comparison

模型	隐藏层数	各层隐藏单元	迭代次数	时间步长	MAE	RMSE	MAPE	RMSPE
GRU	3	[12, 35, 22]	1 883	3	0.033 4	0.046 7	0.103 8	0.129 8
LSTM	3	[24, 22, 38]	1 897	2	0.034 1	0.047 1	0.116 6	0.144 2
RNN	2	[26, 32]	1 286	6	0.034 5	0.047 6	0.118 4	0.145 5

从表 1 中可以看出,在经过 DE 算法寻优后,3 种循环神经网络的超参数各不相同。从寻优结果分析,3 种模型皆为深度模型,而非常见的、简单的三层神经网络结构,表明深度学习模型在一定程度上可以提高预测精度;而合理的设置隐藏神经元数量可以减少不必要的运算,提升模型的训练的预测速度;在不同的模型结构下,迭代次数和时间步长对模型的拟合能力影响较大,不同的时间步在提取时间序列的规律时,需要进行迭代训练的次数也不尽相同。

寻优的 MADF-DE-GRU 预测模型在各项预测指标上的表现都很优秀,提升较为明显,拥有更高的预测精度。对比 MADF-DE-GRU 和 DE-GRU 两种模型,可以发现 MADF 数据预处理过程能够显著提升模型的预测精度,说明该数据预处理过程在一定程度上消除了数据中存在的噪声,去除了数据的相关趋势,使得数据更加平滑与稳定;与 MADF-GRU 对比, MADF-DE-GRU 的各项预测指标值亦有所降低,说明经过 DE 算法寻优而建立的预测模型比较合理,能够提升模型的预测精度;通过将 DE-GRU、MADF-GRU 和单一的 GRU 3 种模型进行对比,可以发现 DE-GRU 和 MADF-GRU 预测精度明显好于单一的 GRU 模型,进一步说明了本文提出的 MADF 数据预处理过程与 DE 算法优化 GRU 神经网络超参数能够提升预测精度,使模型具有更好的预测效果。

在预测指标 MAE 和 RMSE 上,GRU 的表现略好于 LSTM 和 RNN;从 MAPE 和 RMSPE 方面看,GRU 对比 RNN 和 GRU 提升效果显著;相对于 LSTM,GRU 的 MAPE 值和 RMSPE 值分别降低了 11% 和 10%,而较 RNN 而言,两项预测指标则分别降低了 12% 和 11%。因此,在本文提出的模型思想下,GRU 的预测效果优于 LSTM 和 RNN。

4.4 传统预测模型对比

4.3 变体预测模型对比

为进一步验证本文提出的 MADF-DE-GRU 模型在瓦斯浓度时间序列上的预测能力,将其与传统的预测方法进行对比。对比了 MADF-DE-GRU、最小二乘向量机(LSSVM)、径向基神经网络(RBF)、BP 神经网络和 ARIMA 在相同数据集上的预测误差,结果见表 3。通过对比,可以发现 MADF-DE-GRU 的各项预测误差指标的值均为最小,模型预测能力明显优于其它传统预测方法。

本文提出的预测模型包含 MAPE 数据预处理方法和 GRU 神经网络的超参数寻优两部分,为了验证这两个部分对预测模型性能的提升,对比了未经 MADF 数据预处理的 DE-GRU 模型、经过数据预处理但未经过超参数寻优的 MADF-GRU 模型以及单一的 GRU 模型。其中 MADF-GRU 和单一 GRU 模型的超参数按照经验设置,具体为隐藏层数 3,各层隐藏单元数均为 50,迭代次数 2 000,时间步长 5,对比结果见表 2。

表 3 传统预测模型对比

Table 3 Comparison of traditional prediction models

模型	MAE	RMSE	MAPE	RMSPE
MADF-DE-GRU	0.033 4	0.046 7	0.103 8	0.129 8
LSSVM	0.043 9	0.056 2	0.132 6	0.148 9
RBF	0.048 5	0.059 3	0.134 5	0.164 4
BP	0.066 5	0.068 4	0.157 5	0.175 9
ARIMA	0.059 6	0.645 2	0.163 5	0.172 1

表 2 变体预测模型对比

Table 2 Comparison of variant prediction models

模型	MAE	RMSE	MAPE	RMSPE
MADF-DE-GRU	0.033 4	0.046 7	0.103 8	0.129 8
DE-GRU	0.048 4	0.059 8	0.116 1	0.147 8
MADF-GRU	0.040 1	0.053 1	0.124 3	0.149 2
GRU	0.054 1	0.069 2	0.165 2	0.170 4

MADF-DE-GRU 模型预测对比图如图 4 所示,可以看出 MADF-DE-GRU 的拟合能力较好,不仅预测值与真实的误差较小,且能够较好的预测瓦斯浓

从表 2 中可以看出,经过数据预处理和超参数

度的变化趋势。模型预测误差如图5所示,可见模型的整体预测误差在4%以内,能够满足煤矿实际生产对瓦斯浓度的预测精度要求。

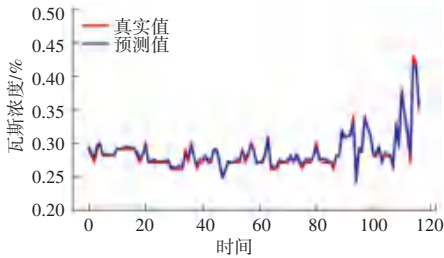


图4 MADF-DE-GRU 预测对比

Fig. 4 MADF-DE-GRU forecast comparison

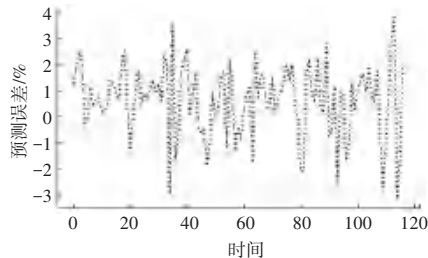


图5 预测误差

Fig. 5 Prediction error

5 结束语

本文提出一种基于GRU神经网络的瓦斯浓度动态预测模型,针对瓦斯数据存在噪声和不稳定性的问题,提出一种在消除噪声的同时和去除趋势的数据预处理方法MADF,并采用DE算法优化的GRU神经网络超参数构建瓦斯浓度预测模型,对瓦斯浓度进行动态预测。仿真实例验证表明,与同类型的神经网络预测模型、变体预测模型和传统预测模型对比,本文提出的MADF-DE-GRU模型表现出较好的预测性能,具有较高的预测精度和较好的泛化能力,为煤矿瓦斯浓度预测提供了一种可行、有效的方法。

参考文献

[1] 王建国,傅文,刘颖. 2012—2016年我国煤矿较大以上瓦斯事故发生规律分析研究[J]. 矿业安全与环保,2018,45(6):108-111,116.

[2] 中国煤炭地质总局. 加强煤矿冲击地压等重大灾害防治[EB/OL]. [2020-5-26]. http://www.ccgcc.cn/art/2019/3/14/art_63_210053.html.

[3] 林立堂,王山峰,王清彦. 煤矿瓦斯防治技术的现状与存在的问题[J]. 内蒙古煤炭经济,2023(16):181-183.

[4] 梁志荣,梁伟. 煤矿瓦斯时间序列特性及模式研究进展[J]. 煤矿现代化,2017(5):61-64.

[5] 张以文,郭海帅,涂辉,等. 基于随机隐层神经网络神经网络的瓦斯浓度预测[J]. 计算机工程与科学,2019,41(4):699-707.

[6] 付华,丰盛成,刘晶,等. 基于DE-EDA-SVM的瓦斯浓度预测建模仿真研究[J]. 传感技术学报,2016,29(2):285-289.

[7] 沈旭东. 基于深度学习的时间序列算法综述[J]. 信息技术与信息化,2019(1):71-76.

[8] 程子均,马六章,张翼翔. 基于LSTM-FC的瓦斯浓度时空分布预测[J]. 计算机工程与应用,2020,56(16):7.

[9] 李伟山,王琳,卫晨. LSTM在煤矿瓦斯预测预警系统中的应用与设计[J]. 西安科技大学学报,2018,38(6):1027-1035.

[10] 张雨. 基于深度学习的井下瓦斯浓度预测系统设计与实现[D]. 徐州:中国矿业大学,2019.

[11] 钱建生,邱春荣,李紫阳,等. 深度学习耦合粒子群优化SVM的瓦斯浓度预测[J]. 煤矿安全,2016,47(11):173-176.

[12] LYU P, CHEN N, MAO S. LSTM based encoder-decoder for short-term predictions of gas concentration using multi-sensor fusion[J]. Process Safety and Environmental Protection, 2020, 137:93-105.

[13] CHUNG J, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. arXiv preprint arXiv:1412.3555, 2014.

[14] CHO K, VAN MERRIENBOER B, BAHDANAU D, et al. On the properties of neural machine translation: Encoder-decoder approaches[C]//Proceedings of Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation. IEEE, 2014: 103.

[15] RAINER S, KENNETH P. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces[J]. Journal of Global Optimization, 1997, 11(4):341-359.

[16] PASCANU R, MIKOLOV T, BENGIO Y. On the difficulty of training recurrent neural networks[C]//Proceedings of 30th International Conference on Machine Learning. Atlanta, USA: IEEE, 2013: 2347-2355.

[17] HOCHREITER S, SCHMIDHUBER J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):1735-1780.

[18] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[19] 丁青锋,尹晓宇. 差分进化算法综述[J]. 智能系统学报,2017, 12(4):431-442.

[20] KACHITVICHYANUKUL V. Comparison of three evolutionary algorithms: GA, PSO, and DE[J]. Journal of Electrical Engineering & Technology, 2012, 11(3):215-223.