

## 基于时序分析的电网合并单元电平预测

张朝辉<sup>1</sup>, 罗 炜<sup>1</sup>, 林康熙<sup>1</sup>, 秦冠军<sup>2</sup>, 金岩磊<sup>2</sup>, 丁 笠<sup>3</sup>, 周 宇<sup>3</sup>

(1. 中国南方电网有限责任公司超高压输电公司广州局, 广州 510663; 2. 南京南瑞继保电气有限公司, 南京 211102; 3. 南京航空航天大学计算机科学与技术学院, 南京 211106)

**摘 要:** 合并单元设备监控依赖于现场工作人员记录、实践经验以及预设告警阈值, 缺少对系统监视数据的分析和挖掘, 不能实现设备状态预测。鉴于此, 根据监视合并单元电平数据的时序性特征, 将传统时序模型差分整合移动平均自回归 (Autoregressive integrated moving average, ARIMA) 和长短期记忆网络 (Long short-term memory, LSTM) 构建组合模型, 并采用蛙跳算法 (Shuffled frog leaping algorithm, SFLA) 进行模型优化。优化后的模型应用在合并单元激光器监视的电平数据预测分析, 将 ARIMA-LSTM 优化组合模型和单一模型进行对比, 验证了组合模型比单一模型具有更高的准确度。进一步和其他组合模型做对比实验, 实验结果表明, 组合模型经过 SFLA 优化后均优于其他组合模型, 能够更好挖掘数据中的隐藏信息和趋势, 提高时序数据预测精度和故障排查效率。将 SFLA 优化的组合 ARIMA-SVM 模型和 ARIMA-LSTM 模型对比, 实验结果表明, 所提出的 ARIMA-LSTM 模型优于 ARIMA-SVM 模型, 可以更好地分析和掌握设备状态信息, 实现对合并单元设备的电平数据预测。

**关键词:** 合并单元; 时序分析; 差分整合移动平均自回归-支持向量机; 长短期记忆网络; 蛙跳算法  
**中图分类号:** TP391      **文献标志码:** A

### Electrical Level Prediction of Power Grid Merging Unit Based on Time Series Analysis

ZHANG Zhaohui<sup>1</sup>, LUO Wei<sup>1</sup>, LIN Kangzhao<sup>1</sup>, QIN Guanjun<sup>2</sup>, JIN Yanlei<sup>2</sup>, DING Li<sup>3</sup>, ZHOU Yu<sup>3</sup>

(1. Guangzhou Bureau of EHV Transmission Company of China Southern Power Grid Co. Ltd., Guangzhou 510663, China; 2. NR Electric Co. Ltd., Nanjing 211102, China; 3. College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

**Abstract:** The equipment monitoring of the merging unit relies on on-site staff records, practical experience and preset alarm threshold, and the lack of analysis and mining of the system monitoring data makes it impossible to realize the device state prediction. In view of this, according to the timing characteristics of the level data of the monitoring merge unit, a combined model of the traditional timing model of autoregressive integrated moving average (ARIMA) and long short-term memory (LSTM) is constructed and optimized by using shuffled frog leaping algorithm (SFLA). The optimized model is applied to the level data prediction analysis of the combined unit laser monitoring. The comparison of the ARIMA-LSTM optimized combination model with the single model verifies that the former has higher accuracy than the latter. Further comparison experimental results show that the combined model is superior to the other combined models after the SFLA algorithm optimization, which can better mine the hidden

information and trend in the data, improving the accuracy of time series data prediction and the efficiency of fault troubleshooting. By comparing the combined ARIMA-SVM model and the proposed ARIMA-LSTM model, experimental results show that the proposed ARIMA-LSTM model is superior to the ARIMA-SVM model, and it can better analyze and grasp the device state information, and realize the level data prediction of the merging unit equipment.

**Key words:** merging unit; time series analysis; autoregressive integrated moving average-support vector machine (ARIMA-SVM); long short-term memory(LSTM); shuffled frog leaping algorithm(SFLA)

## 引 言

随着国内经济的迅猛增长,国民用电需求也在快速增长。与之相对应的,中国的超、特高压输变电技术也正式走向了完善和成熟。在高压输电技术日趋成熟的前提下,如何做到精益求精,对超、特高压输变电工程安全性、稳定性和智能化的要求越来越高,成为一个急需面对的严峻问题。目前直流输电正朝着高电压大容量的技术方向发展,直流输电的长距离输电等方面优势明显,但也会出现一些关键设备由于故障造成的直流系统非正常停运,对电力系统造成的影响越来越大<sup>[1]</sup>。因此,提高对直流关键设备的感知程度,对电网中关键设备存在的潜在故障隐患进行提前预判采取相应的对策,将故障消灭在萌芽阶段,对降低直流系统非正常停运风险并保证供电可靠性具有重要的意义。

合并单元装置为直流控制保护系统提供真实有效的电流或电平信息,是直流关键设备,是直流测量装置的重要组成部分<sup>[2]</sup>。合并单元装置用于接收直流测量装置的数字采样信号,合并各直流测量装置的采样数据并进行组帧,然后通过光纤介质按照标准通信协议分别传输到直流控制保护等设备。同时,它还要通过供能光纤为直流测量装置的远端模块(Remote terminal unit, RTU)提供激光能量作为其工作电源。在运行时,如果出现激光器驱动电流高、激光器温度高或者接收数据电平低的情况,会导致直流测量出现偏差,造成直流非正常停运<sup>[3]</sup>。

当前,针对合并单元装置的运行状态存在一些监视措施,但大多是利用已有经验来设置合并单元设置阈值。当合并单元装置中某些数据超过或低于已设阈值时,系统发出告警信号。这种方式仅能区分系统的“健康”和“不健康”状态,对处于中间区域的“亚健康”状态无法识别并及时预警。除此之外,此方式缺少对监视数据进行充分的挖掘和分析,不能及时发现隐藏在电网时序数据中的信息和趋势。鉴于此,考虑到监视数据具有时序性,本文提出利用时序分析方法对监测数据进行时序分析并预测未来值从而达到评估合并单元设备状态的目的。而且,考虑到合并单元激光器电平是合并单元的关键参数之一,着重根据合并单元激光器驱动电平的监视数据进行合并单元设备状态的预测研究,以期创建预测电网时序数据模型,挖掘数据中的隐藏信息和趋势,降低维护成本,提高故障排查效率,分析和掌握设备状态信息,实现合并单元设备及时预警<sup>[4]</sup>。

## 1 相关工作

国内外研究学者提出很多基于时序数据的分析方法<sup>[5-8]</sup>。这些方法经历了从传统的时序模型到基于机器学习的模型,以及从单一模型到组合模型的发展,都对时间序列数据进行分析,根据时间序列体现出的趋势、规律和方向对后续过程进行扩展和总结,进而预测下一段时间内的数据水平。

传统的时间序列分析方法有很多种。1927年数学家 Yule 提出自回归(Autoregressive, AR)模型,根据之前已经存在的数据来预测未来下一时间点。1931年 Walker 提出滑动平均(Moving average, MA)模型和自回归滑动平均(Autoregressive moving average, ARMA)模型。MA模型是一种根据时序数据通过计算移项、时间序列的平均值来体现数据长期趋势的方法。相较于AR模型和MA模型,AR-

MA模型在谱估计和谱分辨率方面表现更为优异,与其相对应的参数估算也较为复杂。上述3种模型的构建均以平稳时间序列为前提,而差分整合移动平均自回归(Autoregressive integrated moving average, ARIMA)模型的优点在于可以对非平稳的时序数据建模进行时序分析预测<sup>[9-10]</sup>。

时间序列预测研究与机器学习方法中的回归分析关系密切。机器学习方法中包括了常见的支持向量机(Support vector machine, SVM)<sup>[11]</sup>、贝叶斯网络<sup>[12]</sup>等,矩阵分解等方法也常用在时间序列数据的分析预测方面,并取得了不错的效果。Kim<sup>[13]</sup>采用SVM对股票价格指数加以预测取得了很好的效果;Tay和Cao<sup>[14]</sup>也将SVM应用在金融时间序列数据方面进行预测;Das和Ghosh<sup>[15-16]</sup>提出将时空信息融合到贝叶斯网络,对气象领域的时间序列数据进行预测;在应对高维度时间序列时,一般的时序分析模型往往难以应对,因此Yu等<sup>[17]</sup>提出一种时间正则矩阵分解技术,可以很好地解决此问题,而且在包含了噪声数据的时间序列数据方面展现了良好的鲁棒性。

近年来,随着深度学习的不断深入发展以及广泛应用,它在时间序列预测方面也是有效的<sup>[18-20]</sup>,尤其是循环神经网络(Recurrent neural networks, RNN)的使用。RNN可以解决时间上的延续性问题<sup>[21]</sup>。虽然RNN将整个时间序列信息转换成一个向量,但始终存在一个无法解决的问题,就是时序数据在时序长度不断增加的情况下,RNN的记忆能力与时序长短成反比而不断减弱,最终效果也就不甚明显,由于存在如此缺陷,会导致RNN无法克服自身存在的长期依赖的问题<sup>[22]</sup>。长短期记忆网络(Long short-term memory, LSTM)由Schmidhuber在1997年提出,其目的就是为了弥补RNN的不足之处。它引入了一个上下文向量,进而在解码过程中可以选取最相关的信息,因此非常适合用于时序数据分析,构建预测模型。

但无论是传统时序分析方法中的ARIMA模型、机器学习方法中的SVM、还是深度学习中的LSTM神经网络,没有一种特定的单一模型能够统一对电网时间序列数据做出最好的预测,因此组合模型应运而生。组合模型可以弥补单个预测的缺点,受益于单个预测之间的相互作用以及各个模型的多样性,在减少单一模型使用风险的同时又能够提升预测精度。在组合预测模型中,如何分配每个单一预测模型的权重系数是提高预测准确性的关键。大多数研究采用相对简单的等权平均组合预测法和均方根误差组合预测法。等权平均组合预测法中的各个模型同等重要,即所有权重都相等,没有给预测更准确的单一模型分配更多的权重,即没有优选的概念,而均方根误差组合预测法存在权值不稳定的问题。考虑到蛙跳算法(Shuffled frog leaping algorithm, SFLA)具有很强的全局收敛性和局部搜索细致等优点<sup>[23]</sup>,因此,选取SFLA求得组合模型中各个模型的较优权重值,在各个单一模型基础上综合优化进而提高预测准确性。

## 2 合并单元设备状态研究

合并单元设备状态研究与数据息息相关。考虑到合并单元激光器驱动电平是合并单元设备的一个关键参数,本文根据合并单元激光器驱动电平的监视数据进行合并单元设备状态的预测研究。

### 2.1 数据准备

本工作数据来源于中国南方电网从2017年9月到2019年8月的超高压换流站设备合并单元激光器驱动电平数据记录。

考虑到要分析数据的时序性和趋势性特征,而初始数据是散乱分布于数据库中的,因此需要以下一系列的数据预处理来构建一个干净的数据集。

(1)数据存取高效性保证:为了保证高速、有效地对数据进行划分、筛选和清洗,将所有数据从数据库中导出,并以JSON或者CSV文件格式存储。

(2)数据时序性保证:鉴于合并单元设备数量众多,需要对每台设备的监视数据进行分析,为了分析每台设备随时间变化的运行状态,需要对数据按设备进行划分,并按时间节点对每台设备的监视数

据进行排序,确保数据的有效性和时序性。

(3)冗余数据滤除:在所有监视数据中存在冗余数据,即存在因为停电、故障等原因造成的某个时间段内数据值缺失的问题,因此,需要将这些无效数据进行清洗和过滤,进而确保所收集数据的真实性和有效性。

在获得干净的数据集后,可以对监测到的合并单元激光器驱动电平数据进行趋势分析。

## 2.2 数据趋势分析

为了更好地分析合并单元激光器驱动电平数据随时间变化的趋势性和规律性,分别按天、周对合并单元激光器驱动电平数据进行随机抽样,分析其规律性和趋势变化。

### 2.2.1 一天内数据分析

对过滤后的数据集进行随机抽样,选取3天的合并单元激光器电平数据进行分析,其采集数据随时间的变化如图1所示,其中3个分图分别描绘了在一天时间内合并单元激光器驱动电平的变化趋势。

从图1可以分析出,合并单元激光器驱动电平在正常工作情况下,其值会趋于保持在一个稳定范围内,会有随机的波动,但值的波动幅度较小,且驱动电平值均在正常值范围之内变化。

### 2.2.2 一周内数据分析

对过滤后的数据集进行随机抽样,选取了3周的驱动电平数据进行分析,结果如图2所示,其中3个分图分别描绘了在3周时间内合并单元中激光器驱动电平的变化趋势。

从图2可以分析出,激光器驱动电平在正常工作情况下,其值依旧会趋于保持在一个稳定范围内,会有随机的波动,但值的波动幅度不大,且驱动电平值均在正常值范围之内变化。为进一步确定采集到的激光器驱动电平数据是否为一个平稳时间序列,可采用计算均值、标准差等方式。

## 3 合并单元设备状态分析方法

对处理好的合并单元激光器电平数据,选取、确定一些时序性分析方法,尝试建立预测

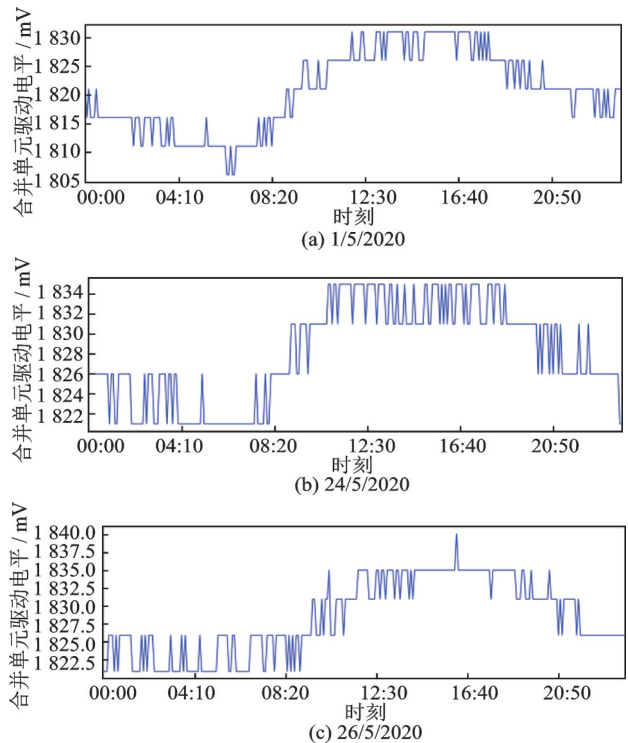


图1 随机3天的激光器驱动电平变化

Fig.1 Random three-day laser drive level changes

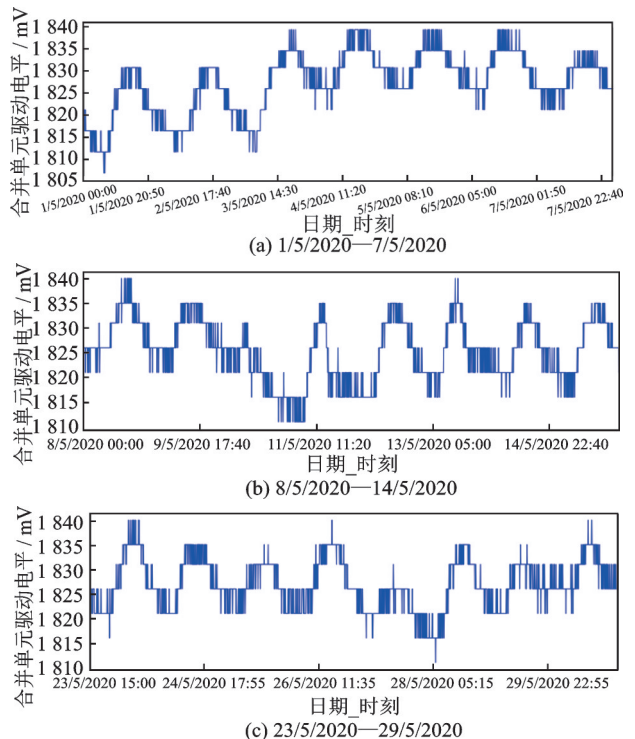


图2 随机3周的激光器驱动电平变化

Fig.2 Random three-week laser drive level changes



模型,对合并单元设备状态进行组合模型预测分析。

### 3.1 组合模型预测优化原理

组合模型的原理是分别对单一时序预测模型预测结果进行再处理,首先采取一定的方法计算后得到适当的权重,然后对各个模型的预测结果加权平均计算,以最终得到的结果作为最终预测结果。组合模型可以在多个单一模型预测结果的基础上综合优化,降低单一模型不适用的风险,提高预测精度。本文选择将传统的预测模型分别和SVM、LSTM神经网络构建两个组合预测模型对电网数据进行实验,分析挖掘数据内在的规律,组合模型的基本原理为

$$y^p = \sum_{j=1}^k w_j y_j^p \quad (1)$$

式中: $y_i(i=1,2,\dots,n)$ 为真实电量时序数据; $y_j^p(j=1,2,\dots,k)$ 为第 $j$ 个模型的预测值; $W_j$ 为第 $j$ 个模型权重的估计值,满足 $\sum_{j=1}^k w_j = 1$ 和 $w_j \geq 0$ ;  $y^p$ 为组合预测值。

本方法进行预测的基本原理是基于误差平方和最小的固定权系数进行组合预测,即

$$\min E = \sum \left( y - \sum_{j=1}^k w_j y_j^p \right)^2 \quad (2)$$

本文采用SFLA算法对式(2)进行优化,即误差平方和取得最小值时,得出相应权重系数,并赋值给式(1),从而获得最优组合预测值。

### 3.2 组合模型预测优化过程

(1) ARMA模型。ARMA模型进行时序数据分析的首要条件是数据平稳性,它是由AR模型与MA模型两个传统时序模型改进结合两者优点而构成的模型。假设一个随机过程在时间的维度上,通过计算其均值和方差值,如果两者的计算结果都是固定不变的数值,满足这样条件的随机时间序列即为平稳的。ARMA模型可表示为

$$Y_t = \sum_{j=1}^p a_j Y_{t-j} + \sum_{j=0}^q b_j \epsilon_{t-j} \quad (3)$$

式中: $p$ 、 $q$ 分别代表了ARMA模型的阶数,简记为ARMA( $p, q$ ),即时序预测值 $Y_t$ 是现在和过去的误差或冲击值 $\epsilon_t$ 以及在此之前的时间序列 $Y_{t-1}$ 的线性组合。

(2) ARIMA模型。ARIMA模型是针对不满足平稳性条件的时间序列建模。通过差分将非平稳的时间序列进行平稳化。ARIMA模型本质上可以也能看作AR模型与MA模型的组合,但和ARMA模型有所区别。ARIMA模型可表示为

$$\left( 1 - \sum_{j=1}^p \phi_j L^j \right) (1 - L)^d X_t = \left( 1 + \sum_{j=1}^q \theta_j L^j \right) \epsilon_t \quad (4)$$

式中: $L$ 代表滞后算子; $d$ 表示对非平稳时序数据经过 $d$ 次差分后得到的平稳的时序数据。

(3) SVM模型。SVM模型是以统计学理论为基础的学习方法,在高维数据处理方面有较强优势,而且可以在参数少、样本小的条件下具有很好的泛化能力。从理论上分析只存在唯一的全局最优解,其泛化能力在小样本集合条件下有更好的体现。SVM模型的性能由误差惩罚因子 $C$ 和核函数中的参数决定。误差惩罚因子的作用在于确定数据子空间时调节置信区间范围。本文选用径向基函数作为核函数,其参数 $\sigma$ 为径向基半径。在构建支持向量机模型时,需要综合衡量之后合理地选择支持向量机的参数,提高模型的学习和泛化能力。

(4) LSTM神经网络。LSTM模型具有记忆长短期信息的能力,可以从过去的数据中挖掘有用信息来分析未来变化的规律。LSTM网络可以解决梯度消失不能进行很好反馈的问题。根本原因在于LSTM网络采用记忆块取代传统的RNN的隐含节点,从而能够更好地学习时序数据内在长期依赖关系。LSTM神经网络模型在训练时涉及到许多超参数的设定,设神经元数量为 $m$ ,时间步长 $T$ ,批数据大小batch size,迭代次数epoch轮。神经元个数决定了神经网络对时序数据的拟合程度,时间步长和批数据大小决定了模型训练的结果,选择适应性矩估计(Adaptive moment estimation, ADAM)优化器对参数进行优化调整。

(5) SFLA算法。基本原理是:在 $D$ 维搜索空间内,随机产生一只青蛙种群。 $X_i=(x_{i1}, x_{i2}, \dots, x_{id})$ 表示种群的第 $i$ 只青蛙个体,并按照适应度进行非递增排序,然后将种群按照一定方法分为 $s$ 个子群。其方法为第一只青蛙分配到第一个子群,第二只青蛙分配到第二个子群,第 $s$ 只青蛙分配到第 $s$ 个子群,值得注意的是第 $s+1$ 只青蛙会重新分配到第一个子群,以此类推,直到青蛙分配完毕。

在每个子群,适应度最优的青蛙个体设为 $X_{\text{best}}$ ,适应度最差的青蛙个体设为 $X_{\text{worst}}$ 。对整个青蛙群体而言,一定存在适应度最优的青蛙个体,记为 $X_g$ 。对从第一个子群开始对每个子群进行局部搜索,最终搜到第 $s$ 个子群,在这个过程中,按照以下更新策略迭代每个子群中最差的个体。

$$D_i = \text{rand}() \times (X_{\text{best}} - X_{\text{worst}}) \quad (5)$$

$$X_{\text{worst}} = X_{\text{worst}} + D_i \quad (-D_{\text{max}} \leq D_i \leq D_{\text{max}}) \quad (6)$$

式中: $\text{rand}()$ 表示范围大于零小于1的随机数, $D_{\text{max}}$ 为青蛙个体可以移动的最大步长。如果经过更新迭代后,计算出的最新的 $X_{\text{worst}}$ 个体优于原来的 $X_{\text{worst}}$ 个体,则取代原子种群中最差的青蛙个体,否则就用 $X_g$ 代替 $X_{\text{best}}$ 重新进行更新,如果其适应度仍没有改进,则随机产生一个 $X'_{\text{worst}}$ 取代原有的 $X_{\text{worst}}$ 。不断执行上述步骤,直到完成所有更新。当子群体全部经过局部搜索后,将青蛙个体进行混合运算,重新划分子群,进行局部搜索,直到迭代次数完成为止。

## 4 实验评估

### 4.1 评估度量指标

(1) 算术平均值,表示为

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (7)$$

(2) 方差,表示为

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (8)$$

(3) 均方根误差,表示为

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_{\text{real},i} - X_{\text{predict},i})^2} \quad (9)$$

### 4.2 实验结果分析

随机选取一台合并单元设备从2020年5月1日到5月30日的数据进行实验分析评估,电平数据采样的时间间隔为5 min,数据的趋势变化如图3所示。从图3可以看出,合并单元在正常工作状态下,激光器驱动电平趋于平稳,波动幅度较小,其算术平均值趋近于一个

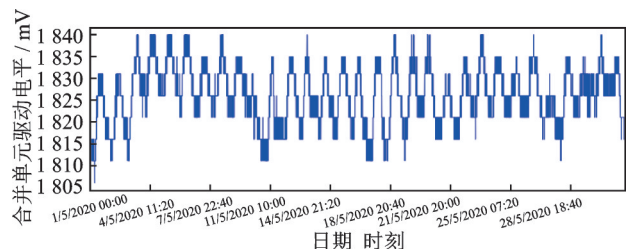


图3 激光器驱动电平一个月内数据变化图

Fig.3 Data change diagram of laser drive level within one month

常量,方差值亦趋近于0。此外,运用迪基-福勒检验进行单位根校验<sup>[25]</sup>,计算得到 $p$ 值小于0.01,说明此时间序列在超过99%的置信水平下是显著的,即数据是时间平稳序列,不需要经过差分运算就可以建模。

本文方法中的参数设置:基于激光器驱动电平的数据是时间平稳序列,先建立ARIMA(1,0,2)模型,确定模型参数,进行预测。将惩罚因子 $C$ 设为10,用径向基函数(Radial basis function, RBF)作为进行预测。优化误差惩罚因子 $C$ 和核函数能够很好提升SVM的性能,LSTM神经网络的神经元数量 $m=128$ ,批数据大小 batch size=32,迭代次数 epoch=10。SFLA算法的群体数为150,子群数 $s=5$ ,子群迭代次数设为50,混合迭代总数1000。

将提出的基于SFLA算法优化的ARIMA(1,0,2)和SVM组合模型方法与等权组合模型方法、ARIMA及SVM的预测结果进行比较,验证SFLA算法优化后组合模型的有效性。从图4可以看出,采用SFLA算法优化组合模型所预测的曲线与ARIMA(1,0,2)和SVM预测曲线相比,在合并单元驱动电平数据发生变化上更加接近真实电平曲线。

表1通过均方根误差指标评估组合模型质量,SFLA算法优化组合模型的均方根误差为3.12,优于单一的ARIMA模型和SVM模型。说明基于SFLA算法优化组合预测模型比单一模型更加适合应用在电网中电平预测,获得的预测效果更好。

将提出的基于SFLA算法优化的ARIMA(1,0,2)和LSTM神经网络组合模型方法与等权组合模型方法,ARIMA(1,0,2)及LSTM神经网络的预测结果进行比较,验证优化后组合模型的有效性。从图5中可以看出,采用SFLA算法优化组合模型所预测的曲线比ARIMA(1,2)和LSTM预测曲线相比更加接近真实电平曲线。ARMA(1,0,2)可以很好拟合真实电平在运行过程中变化的曲线,但SFLA算法优化组合模型可以在ARIMA(1,0,2)模型的基础上进一步提升拟合精度。

表2说明,经过SFLA算法优化的ARIMA-LSTM组合在预测模型的均方误差和

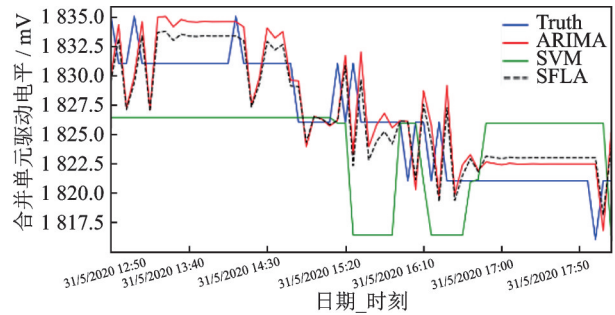


图4 ARIMA、SVM、SFLA预测模型的预测结果

Fig.4 Prediction results of ARIMA, SVM and SFLA models

表1 SFLA优化组合模型与ARIMA、SVM模型的性能对比

Table 1 Performance comparison among ARIMA, SVM models and the combined model optimized by SFLA

评估度量指标	ARIMA	SVM	SFLA优化组合
MSE	11.72	31.91	9.75
RMSE	3.42	5.64	3.12

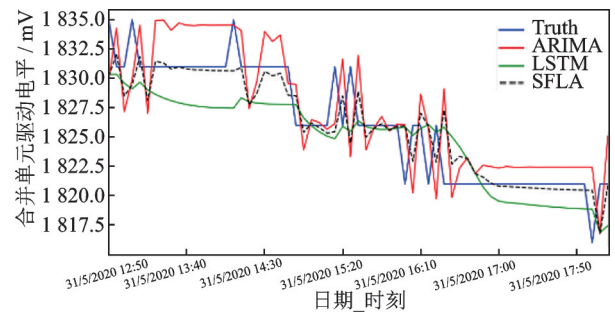


图5 ARIMA、LSTM、SFLA预测模型的预测结果

Fig.5 Prediction results of ARIMA, LSTM and SFLA models

表2 SFLA优化组合模型与ARIMA、LSTM模型的性能对比

Table 2 Performance comparison among ARIMA, LSTM models and the combined model optimized by SFLA

评估度量指标	ARIMA	LSTM	SFLA优化组合
MSE	11.72	8.42	5.72
RMSE	3.42	2.90	2.39

均方根误差上均优于单一模型,组合模型可以结合两个模型的优点提高预测精度,在电平数据预测中获得更好的效果。

为了比较 SFLA 算法优化后的 ARMA-SVM 组合模型和 SFLA 算法优化后的 ARMA-LSTM 组合模型,确定哪个组合模型更适合电网数据预测,将其进行单独对比,结果如图 6 所示。从图 6 可以看出,经过 SFLA 算法优化的 ARIMA-LSTM 组合模型明显优于 ARIMA-SVM 模型。

表 3 进一步说明 ARIMA-LSTM 组合模型适合电网中的电量预测,主要原因是 LSTM 神经网络属于深度学习方法,和浅层机器学习方法算法相比具有更好的非线性能力和泛化能力。SVM 在训练集样本学习到拟合效果较好的映射关系,在实际应用中面对新数据,其预测效果并不理想,其原因是浅层机器学习模型的泛化能力较差。而深度学习则通过多层连接和权值共享的网络结构增强了模型的泛化性能。

为了进一步比较 SFLA 算法对组合模型优化的性能,将经过 SFLA 算法优化后的 ARIMA-LSTM 组合模型与基于 ARIMA 和 LSTM 两个单一模型形成的等权组合模型和方差-协方差组合模型进行对比,结果如图 7 所示。从图 7 中可以发现,3 组组合模型都能很好拟合电平的变化趋势。采用 SFLA 算法优化组合模型所预测的曲线与另外两个组合模型形成的曲线比较接近,是因为通过 SFLA 算法经过一系列计算之后,计算出的权值和另外两个组合模型的权值比较接近。经局部 SFLA 算法优化后的模型能够更好地反应真实数据。

从表 4 中计算的均方根误差和均方误差说明,经过 SFLA 算法优化的 ARIMA-LSTM 组合预测模型各项指标均优于 ARIMA-LSTM 等权组合模型和方差-协方差组合模型,SFLA 算法对组合模型的优化效果更好。

## 5 结束语

采用 SFLA 算法对组合模型进行优化,综合改进组合模型 ARIMA-SVM 和 ARIMA-LSTM 中各自的优点,在单一模型的基础上进一步提升预测精度。更好地处理时序序列数据预测的随机性因素,与

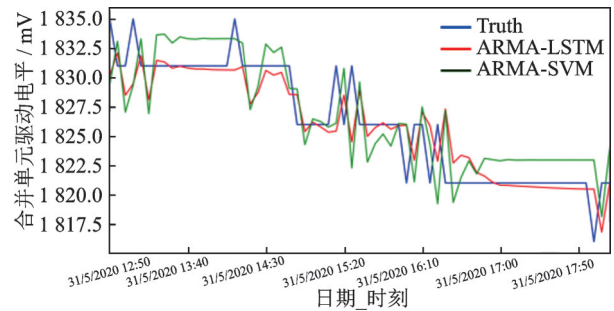


图 6 SFLA 优化后不同组合预测模型的预测结果

Fig.6 Prediction results of different combination prediction models after SFLA algorithm optimization

表 3 两种 SFLA 优化组合模型性能对比

Table 3 Performance comparison between two combined models optimized by SFLA

评估度量 指标	经过 SFLA 优化的 ARIMA-LSTM	经过 SFLA 优化的 ARIMA-SVM
MSE	5.72	9.75
RMSE	2.39	3.12

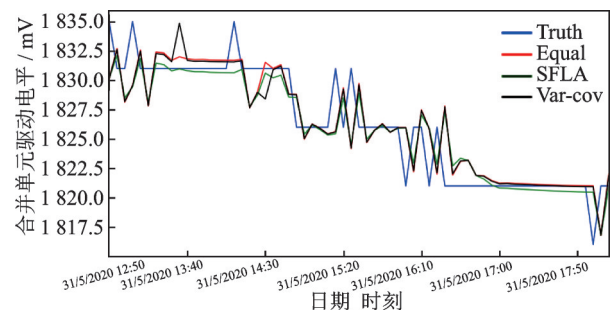


图 7 不同组合模型预测结果

Fig.7 Prediction results of different prediction models

表 4 不同优化方法的组合模型性能对比

Table 4 Performance comparison among different optimized methods for combined models

评估度量 指标	SFLA 优化 组合	等权平均组合	方差-协方差 组合
MSE	5.72	6.34	6.53
RMSE	2.39	2.57	2.55



单一模型相比,在降低模型选择风险的同时提高预测精度。

SFLA算法具有很强的全局搜索性和局部搜索仔细的特点。采用SFLA算法确定组合预测模型中每个模型的权重系数,相比较等权组合模型中各个模型权重都相等,没有优选的概念,而均方根误差组合预测法存在权值不稳定的问题,能够进一步提高预测精度。

本文对ARMA和SVM组合模型进行优化,将其应用在电网中的电平预测,精度比单一模型和等权组合模型有所提高,为电网中的电平预测提供了一种行之有效的预测方法。

#### 参考文献:

- [1] 崔宗清,孙仁龙,邹荣盛.换流站在线监测系统的电源方案[J].集成电路应用,2019,36(10): 46-48.  
CUI Zongqing, SUN Renlong, ZOU Rongsheng. Power supply scheme for on-line monitoring system of converter station[J]. Integrated Circuit Application, 2019, 36(10): 46-48.
- [2] 赵艳君,茹作伟,陈宇.浅谈合并单元应用及发展[J].电子世界,2018(20): 82-83.  
ZHAO Yanjun, RU Zuwei, CHEN Yu. A brief discussion on the application and development of consolidated units[J]. Electronic World, 2018(20): 82-83.
- [3] 吴怀宇.时间序列分析与综合[M].武汉:武汉大学出版社,2004.  
WU Huaiyu. Time series analysis and synthesis[M]. Wuhan: Wuhan University Press, 2004.
- [4] 林子雨,江弋,赖永炫.一种新的时间序列延迟相关性分析算法——三点预测探查法[J].计算机研究与发展,2012(12): 147-157.  
LIN Ziyu, JIANG Yi, LAI Yongxuan. A new time series delay correlation analysis algorithm: Three-point predictive probe method[J]. Computer Research and Development, 2012(12): 147-157.
- [5] 吴志华,韩燮,张永梅,等.基于知识发现的时序数据分析方法[J].测试技术学报,2002(Z1): 571-575.  
WU Zhihua, HAN Xie, ZHANG Yongmei, et al. Temporal data analysis method based on knowledge discovery[J]. Journal of Testing Technology, 2002(Z1): 571-575.
- [6] 张美英,何杰.时间序列预测模型研究综述[J].数学的实践与认识,2011,41(18): 189-195.  
ZHANG Meiyong, HE Jie. Review of time series prediction model research[J]. Mathematics Practice and Understanding, 2011, 41(18): 189-195.
- [7] 杨海民,潘志松,白玮.时间序列预测方法综述[J].计算机科学,2019,46(1): 21-28.  
YANG Haimin, PAN Zhisong, BAI Wei. A review of time series prediction methods[J]. Computer Science, 2019, 46(1): 21-28.
- [8] 王丽娜,肖冬荣.基于ARMA模型的经济非平稳时间序列的预测分析[J].武汉理工大学学报(交通科学与工程版),2004,28(1): 133-136.  
WANG Lina, XIAO Dongrong. Forecasting analysis of economic non-stationary time series based on ARMA model[J]. Journal of Wuhan University of Technology (Transportation Science and Engineering Edition), 2004, 28(1): 133-136.
- [9] JENKINS G M, RIESEL G C. Time series analysis: Forecasting and control[M]. NY: Prentice Hall Inc., 1994.
- [10] SHUMWAY R H, STOFFER D S. Time series analysis and its applications[M]. New York: Springer Science Business Media, 2006: 79-99.
- [11] 张学工.关于统计学习理论与支持向量机[J].自动化学报,2000,26(1): 32-42.  
ZHANG Xuegong. Statistical learning theory and support vector machine[J]. Acta Automatica Sinica, 2000, 26(1): 32-42.
- [12] PEARL J. Fusion, propagation, and structuring in belief networks[J]. Artificial Intelligence, 1986, 2(3): 241-288.
- [13] KIM K. Financial time series forecasting using support vector machines[J]. Neurocomputing, 2003, 55(1): 307-319.
- [14] TAY F E H, CAO L. Application of support vector machines in financial time series forecasting[J]. Omega, 2001, 29(4): 309-317.
- [15] DAS M, GHOSH S K. A probabilistic approach for weather forecast using spatio-temporal inter-relationships among climate variables[C]//Proceedings of International Conference on Industrial & Information Systems. [S.l.]: IEEE, 2015.
- [16] DAS M, GHOSH S K. semBnet: A semantic Bayesian network for multivariate prediction of meteorological time series data [J]. Pattern Recognition Letters, 2017, 93: 192-201.

- [17] YU H F, RAO N, DHILLON I S. Temporal regularized matrix factorization for high-dimensional time series prediction[C]// Proceedings of the 30 th International Conference on Neural Information Processing Systems. [S.l.]: [s.n.], 2016: 847-855.
- [18] LÄNGKVIST M, KARLSSON L, LOUTFI A. A review of unsupervised feature learning and deep learning for time-series modeling[J]. Pattern Recognition Letters, 2014, 42(1): 11-24.
- [19] LV Y, DUAN Y, KANG W, et al. Traffic flow prediction with big data: A deep learning approach[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(2): 865-873.
- [20] ROMEU P, ZAMORA-MARTÍNEZ F, BOTELLA-ROCAMORA P, et al. Time-series forecasting of indoor temperature using pre-trained deep neural networks[M]. Berlin: Springer, 2013.
- [21] 高茂庭, 徐彬源. 基于循环神经网络的推荐算法[J]. 计算机工程, 2019, 45(8): 198-202.  
GAO Maoting, XU Binyuan. Recommendation algorithm based on cyclic neural networks[J]. Computer Engineering, 2019, 45(8): 198-202.
- [22] 翟毅, 徐丽燕, 季学纯, 等. 基于长短期记忆神经网络的短期负荷预测[J]. 信息技术, 2019(10): 27-31.  
ZHAI Yi, XU Liyan, JI Xuechun, et al. Short-term load prediction based on long and short-term memory neural network[J]. Information Technology, 2019(10): 27-31.
- [23] TANG Deyu, LIU Zhen, YANG Jin, et al. Memetic frog leaping algorithm for global optimization[J]. Soft Computing, 2019, 23(21): 11077-11105.
- [24] DOLADO J J, GONZALO J, MAYORAL L. A fractional Dickey-Fuller test for unit roots[J]. Econometrica, 2002, 70(5): 1963-2006.

#### 作者简介:



张朝辉(1989-),男,高级工程师,研究方向:高压直流输电运维方向, E-mail: zzhsddxdq@163.com。



罗伟(1984-),男,硕士,高级工程师,研究方向:电网设备运行管理方向。



林康熙(1986-),男,本科,工程师,研究方向:高压直流输电设备运行维护方向。



秦冠军(1982-),男,硕士,高级工程师,研究方向:新能源数据分析、电力监控系统数据分析。



金岩磊(1978-),男,硕士,工程师,研究方向:电力系统监控研发。



丁笠(1992-),男,硕士研究生,研究方向:时序分析。



周宇(1981-),通信作者,男,教授,研究方向:智能化软件技术、大数据与云计算、软件演化分析, E-mail: zhouyu@nuaa.edu.cn。

(编辑:王静)