

基于改进组合预测的电能质量预警研究

卢 珏¹, 孙云莲¹, 谢信霖², 郑龙武², 徐冰涵¹, 吴 莹¹

(1. 武汉大学电气与自动化学院, 湖北 武汉 430072;
2. 广东电网英德供电局, 广东 清远 513000)

摘要: 随着能源互联网的建设和分布式发电的日益增多,以及目前用电形式的逐渐多样化,产生的电能质量问题也随之受到重视。为了保证用电品质以及电网的稳定运行,对电能质量指标合理地预警具有重要意义。因此,本文提出了一种基于改进组合预测的电能质量预警模型。对于可能影响电能质量指标的因素进行整合,利用 ReliefF 算法排除相关重要性低的因素,以降低输入维数。然后,选择 Elman 神经网络、随机森林和 RBF-SVM 智能算法对相应指标预测,利用灰色关联分析计算三种方法与真实值间的关联度以确定其权重,将三种方法预测结果组合并对峰谷突变处预测值修正得到最后预测值。最后,以广东省某 10 kV 线路监测点数据进行了仿真计算,对比证明了本文所提出方法的适用性。

关键词: 电能质量指标; ReliefF 算法; Elman 神经网络; 随机森林; RBF-SVM 算法

DOI: 10. 12067/ATEEE1910026 **文章编号:** 1003-3076(2020)09-0065-09 **中图分类号:** TM726

1 引言

近十年来我国发电量不断增长,电成为人们生活必不可少的一部分,较高的用电量在保证高效便捷的社会生产生活的同时也产生了一系列相关问题。用电设备的类型越来越多,线路的结构也越来越复杂,随之而来的电能质量问题对多方造成了不同程度的影响。为此,多地搭建了电能质量监测系统,采集电能质量数据^[1,2]。而目前的研究大多集中于扰动检测^[3,4],相对而言真正对电能质量指标分析预测的不多,因此针对电能质量稳态指标,对其合理预警有重要的工程意义。

在电能质量预警研究方面,文献[5]提出利用混沌理论将具有时间特性的数据扩充到维度空间中,再通过支持向量机预测电能质量稳态数据,但其未考虑到外部环境等因素也可能产生影响;文献[6]组合了线性回归等三种模型对历史电能质量数据进行了预测,然而三种模型均是依赖时间序列的人工模型,算法精度有待提高,且是对以月为单位的数据进行研究;文献[7]利用聚类、主成分分析 KP-

CA 对输入数据进行处理,再输入神经网络进行训练预测,文献[8]则在其基础上考虑到了天气因素,但预测算法都过于单一;文献[9,10]都是在离网条件下利用预测算法进行电能质量参数预测,且依赖线路拓扑参数;其他则着重与电能质量预警阈值的分析与划定^[11-13]。

在数据预测预警上,负荷预测研究较多,并且已有了许多成果,这对电能质量的预测奠定了一定的基础,提供了理论上的参考与支持。目前的预测算法主要有神经网络^[14]、随机森林^[15]、灰色预测^[16]等。本文在此基础上,对电能质量稳态指标进行了分析研究,采用 ReliefF 算法对相关时间的电能质量指标数据、外部天气因素等相关因素进行特征选择,剔除与对应预测量关联程度低的特征量,降低算法的输入维度,以简化模型,提高预测精度。然后,采用 Elman 神经网络、随机森林、基于 RBF 函数的向量机(RBF-SVM)算法对其进行训练预测,基于灰色分析法计算每种模型与预测目标实际值关联程度,将其结果进行组合,修正数据变化曲线峰谷突变处预测结果,得到最终组合预测值。最后通过广东某

收稿日期: 2019-10-21
基金项目: 南方电网公司科技项目资助(GDKJXM20173979)
作者简介: 卢 珏(1996-),女,江西籍,硕士研究生,研究方向为电能质量分析与配网预警;
孙云莲(1962-),女,湖北籍,教授,博士,研究方向为电力信息分析。

地区 10 kV 线路监测点数据验证了本文提出方法预测精度的准确性。

2 数据关联分析与处理

为保障电能质量预警的准确性与高效性,在进行预测模型训练前,需对关联因素进行收集整理与分析。基于广东省某地区 10 kV 线路监测点每 15 min 的采集数据,筛选整理出与线路电能质量相关联的特征量,并收集统计相对应的外部天气数据。所获取的数据未经过处理前无法用于模型训练,因此需要对数据进行一系列操作。

2.1 关联数据预处理

(1)输入数据采集与整理

考虑电能质量指标在自身时间序列上存在规律,选取其在预测日前两天以及一周前对应时刻的数据作为输入的选项。此外,其可能还会受负荷、外部环境的影响。因此将一周前对应时刻的有功、无功和当时的温度、相对湿度、风力、天气状况类型、空气质量以及日期类型也考虑其中。由于人为或是系统采集交互问题,在存储数据过程中可能存在数据缺失、或是异常值的出现,因此事先对收集的数据进行整理,补充缺失值,并将异常值剔除或根据数据前后时刻特点进行修正。

(2)输入数据量化处理

对于电能质量指标本身就是数值型数据,因此不需要做量化处理,但天气状况类型和日期类型这类数据都为文字型数据,不利于模型的训练,因此需要进行量化。根据历史数据变化的情况,将晴、多云这类正常天气设作 1,阴、小雨这类天气设作 2,中雨、雷雨等恶劣天气设作 3;将工作日设作 0.8,休息日设作 1。然后,将同类型数据进行归一化以便于接下来模型的输入,而不会对预测结果的精度造成影响。最后输入量统计如表 1 所示。

表 1 预测量影响因素

Tab.1 Influencing factors of predictor		
类别		变量名称
自身因素	历史数据	$d-1、d-2、d-7$ 天的 $t-1、t、t+1$ 时刻指标数据
	实时数据	预测当天的 $t-2、t-1$ 时刻指标数据
外部因素	温度、相对湿度、日期类型量化值、天气类型量化值、风力、空气质量指数	
其他因素	$d-7$ 天 t 时刻有功、无功	

2.2 基于 ReliefF 的输入量特征选择

当输入量维数大且不能保证任一项都与预测结

果都紧密相关时,过多的输入不仅会造成模型复杂度增加求解时间增多,而且可能由于某些相关度不强的数据导致预测结果精度的降低。因此,本文采用 ReliefF 算法对多维输入进行特征选择,留下特征权重大的变量,淘汰关联程度低的输入,以降低模型输入维度。

ReliefF 算法是由 Kononeill 提出,在 Relief^[17] 上拓展而来,相对于 Relief 只能处理两类数据的分类情况,ReliefF 可以处理多类别以及目标为连续值数据问题,模型简单且运行效率极高,因此被广泛使用于提取特征量中。其主要流程如下:

- 1) 读取训练数据库,设置样本抽样次数为 N 。
- 2) 将各维输入的特征权重初始化为 0,即 $\mathbf{W}=[w_1,w_2,\cdots,w_{n-1}],w_i=0(i=1,2,\cdots,n-1)$ 。
- 3) 每次抽样随机选择一个样本 $\mathbf{R}=[r_1,r_2,\cdots,r_n]$ 。
- 4) 在与 \mathbf{R} 同类的样本类中找到其最相近的 k 个对象,构成集合 \mathbf{NR} ,从不同类中也找出 k 个最邻近对象,构成集合 \mathbf{MR} 。计算 \mathbf{R} 与同类中其他样本间在各个输入量上的距离,即:

$$Dist(r_i,\mathbf{R},\mathbf{NR}_j)=\begin{cases}\frac{\mathbf{R}-\mathbf{NR}_j}{\max(r_i)-\min(r_i)} & r_i \text{ 为连续} \\ 0 & r_i \text{ 离散},\mathbf{R}\neq\mathbf{NR}_j \\ 1 & r_i \text{ 离散},\mathbf{R}=\mathbf{NR}_j\end{cases}\quad (1)$$

式中, \mathbf{NR}_j 是指同类中第 j 个最邻近对象; r_i 是指第 i 个输入变量,与不同类对象间的距离 $Dist(r_i,\mathbf{R},\mathbf{MR}_j)$ 也按照此定义。

- 5) 更新输入变量的权重如式(2):

$$W(i)=W(i)-\sum_{j=1}^k Dist(r_i,\mathbf{R},\mathbf{NR}_j)/(Nk)+\sum_{C\notin class(R)}\left[\frac{p(C)}{1-p(C)}\sum_{j=1}^k Dist(r_i,\mathbf{R},\mathbf{MR}_j)\right]/(Nk)\quad (2)$$

式中, $W(i)$ 是指第 i 个输入的权重; $class(R)$ 是指与 \mathbf{R} 同类的对象集合; $p(C)$ 指第 C 类样本对象出现的概率。

- 6) 考虑到样本 \mathbf{R} 每次选择的随机性,将步骤 2)~5) 重复多次取平均权重作为最后的特征权重。

权重的大小可以看作对应输入对模型输出的贡献程度。将其都除以权重总和得到最终统一化权重值,并按递减顺序排列,由于大概率事件置信水平为 95%,因此取统一化权重累计到 95% 的因素作为最后模型的输入。

3 预测模型的建立

3.1 Elman 神经网络

Elman 神经网络是种具有很强计算能力的反馈神经网络算法,它在 BP 神经网络的基本结构上,在隐含层增加了一个承接层,隐含层的输出不再是直接到输出层,而是再流过承接层^[18],相当于在层内再增加一个反馈,从而提高了网络对历史状态的敏感度,使其拥有动态记忆功能,增强了网络的全局稳定性。其数据传输计算结构如图 1 所示。

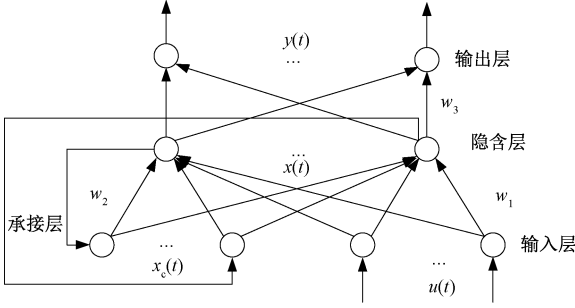


图 1 Elman 神经网络结构

Fig.1 Elman neural network structure

计算过程如下:

$$x(t) = f(w_1 u(t) + w_2 x_c(t)) \quad (3)$$

$$y(t) = g(w_3 x(t)) \quad (4)$$

式中, w_1, w_2, w_3 分别为输入与隐含层、承接层与隐含层、隐含层与输出层间权值。

相较于 BP 的静态反馈,Elman 可以实现动态系统的映射建模,在计算能力和网络稳定上都达到了更优,在预测数据方面具有优越性。

3.2 随机森林算法

随机森林^[19]是基于 bagging 框架下的决策树模型,即利用了决策树建模方法组合了 bagging 算法来输出网络。其算法流程如下:

1) 针对样本数为 N 的训练集,从总体样本集中利用重采样方式随机抽取 N 个样本作为一个训练集,重复多次,生成 T 个训练集。

2) 采用 CART 模型作为弱学习器,若样本特征维数为 M ,针对决策树的每个非叶子节点,从 M 中随机选择 m 个特征,再从这 m 个特征中抉择出最优的那个对该节点进行下一步划分。

3) 每棵树都不剪枝以保证其完整地生长。

4) 对于测试集,利用每个决策树对其进行测试,取均值作为最后强学习器的输出。

由于随机森林的随机选择特征中取优分裂思想,使得模型对新鲜样本的适应能力强且输出的方差也得到了降低,且降低了模型复杂度,无需通过降维就可获得高效结果。此外,数据集的训练是高度并行化的,并具有极好的分类能力和拟合能力,适应了大数据高速计算时代发展。

3.3 RBF-SVM 模型

支持向量机方法^[20]作为机器学习中一种重要的前馈网络型机器算法,即使在少量的样本下它也能在中和模型的复杂程度和对新样本的适应能力的过程内,以期获得最好的推广能力。应对样本不足、训练维度复杂的情况,它也表现出十分优秀的分类识别能力,并且同样适用于解决数据拟合等问题。

对于样本集 $\{(X_1, y_1), (X_2, y_2), \dots, (X_n, y_n)\}$, $X_i (i=1, 2, \dots, n)$ 为输入向量, y_i 为输出向量,其对应关系为:

$$f(X) = W^T \phi(X) + b \quad (5)$$

$$\min_{W, b} \frac{1}{2} \|W\|^2, \text{ s.t. } y_i (W^T \phi(x_i) + b) \geq 1 \quad (6)$$

式中, $\phi(X)$ 表示将 X 映射到高维空间后的对应向量。

依照 Lagrange 对偶理论,其对偶问题为:

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \phi^T(x_i) \phi(x_j) \\ \text{s.t.} \quad & \sum_{i=1}^n \alpha_i y_i = 0, \alpha_i \geq 0, i = 1, 2, \dots, m \end{aligned} \quad (7)$$

引入核函数 $K(x_i, x_j) = \phi^T(x_i) \phi(x_j)$, 求解函数得:

$$\begin{aligned} f(X) &= W^T \phi(X) + b \\ &= \sum_{i=1}^n \alpha_i y_i \phi^T(x_i) \phi(x_j) + b \\ &= \sum_{i=1}^n \alpha_i y_i K(x_i, x_j) + b \end{aligned} \quad (8)$$

RBF 函数具有非线性,在计算过程中不会因为参数的调整使得模型的复杂度发生变化,且在过去研究中表示基于 RBF 的 SVM 算法具有极好的拟合能力,因此本文选择 RBF 函数作为 SVM 的核函数,即:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (9)$$

式中, γ 为 RBF 可调节参数。

3.4 基于改进组合预测的预警模型

不同预测方法对数据的适应度不同,对每段时间内的拟合程度不同,不能在每个时刻都保证电能质量预测数据的准确性。因此,本文提出将以上三种智能机器算法有机结合,用组合后的预测模型进行仿真计算,然后对斜率突变处数据进行修正,得到最后的预测结果。

3.4.1 基于灰色关联分析的组预测

灰色关联度可反映两个系统或两组数据之间的相关程度,将变化趋势相似性变为量化值简单明了体现其间关系,且模型简单,十分适用于动态的过程分析。利用灰色分析理论将三种预测模型信息结合分析,以获取组合预测结果。

对于以上三种算法的预测数据列 $Y_1 = [y_{11}, y_{12}, \dots, y_{1n}]$ 、 $Y_2 = [y_{21}, y_{22}, \dots, y_{2n}]$ 、 $Y_3 = [y_{31}, y_{32}, \dots, y_{3n}]$, 预测日电能质量指标真实值序列为 $Opt = [o_1, o_2, \dots, o_n]$, 计算三组预测序列与真实序列的绝对差值 Abs 。

$$Abs(i, j) = |y_{ij} - o_j| \quad i = 1, 2, 3, j = 1, 2, \dots, n \quad (10)$$

各模型每个时间段数据的关联系数矩阵 RI 为

$$RI = \frac{\min(Abs) + \rho \cdot \max(Abs)}{Abs + \rho \cdot \max(Abs)} \quad (11)$$

式中, ρ 为分辨系数,且 $\rho > 0$ 。

传统灰色关联分析对某数据列每个数据点的关联系数取均值作为该组数据的关联度。但本文着重于每个时间段时间,因此并不对关联系数进行均值处理,而是直接将各个点的关联系数作为其关联度,将同一时刻的三组数据关联系数归一到 $(0, 1)$ 间,得到每个时刻每种算法预测权重,将对应预测权重与预测值相乘求和获得组合后的电能质量预测值。

$$Y = W_1 Y_1 + W_2 Y_2 + W_3 Y_3 \quad (12)$$

$$W_i = [w_{i,1}, w_{i,2}, \dots, w_{i,n}] \quad i = 1, 2, 3 \quad (13)$$

3.4.2 峰谷突变处修正

在仿真过程中发现,当数据在峰谷阶段斜率突变时,预测模型的精度会大大降低,因此本文提出在上述组合模型基础上提出改进,对峰谷突变处利用历史数据及该时间段前后数据进行修正,具体过程如下:

1) 取预测日前三天的峰谷阶段前后一个小时 T 内历史电能质量数据。

2) 计算 T 内减去峰谷突变阶段后其他时刻数据的均值 $Ave_i (i = 1, 2, 3)$, 计算峰谷阶段各个时刻

数据 D_j 对均值的倍数:

$$M_{ij} = D_j / Ave_i \quad (14)$$

式中, M_{ij} 表示预测日的前 i 天, 峰谷突变时刻 j 的数据对该日 T 时间内数据均值的比值。

3) 求得数据对三天的比率均值

$$M_j = \sum_{i=1,2,3} M_{ij} / 3 \quad (15)$$

4) 利用上述组合预测模型求出预测日的数据序列求出对应 T 内减去峰谷突变阶段后其他时刻数据的均值 $PAve$, 计算预测日峰谷阶段各个时刻修正数据

$$PD_j = M_j PAve \quad (16)$$

5) 将修正值代替预测序列中对应时刻数据, 得到改进组合预测的电能质量数据。

4 算例分析

为验证上述提出方法的有效性, 本文将 2019 年 7~10 月广东省某地区 10 kV 线路每 15 min 电能质量采集数据收集整理, 进行仿真。在进行训练前, 将所有数据进行缺失值补充、异常值修正以及归一化处理, 为预测模型的输入做好准备。由于篇幅限制, 下列只以电压偏差和 A 相电压谐波畸变率为例。

4.1 基于 ReliefF 的输入特征选择

4.1.1 电压偏差输入特征选择

将 2.1 节列出的变量按照提出顺序进行编号, 将电压偏差训练数据利用 ReliefF 算法求得各个输入变量的特征权重, 并重复 20 次, 每次 19 个输入变量权重变化如图 2 所示。各个输入权重都取 20 次的均值, 且将均值权重除以总的权重和得到最终的统一化权重如图 3 所示。将统一化权重按从大到小排列, 取其累加到 95% 的输入特征, 即丢弃特征权重小的输入。仿真得出, 当去除输入变量 14、17、18、19 时, 累加权重达到 95.27%。因此, 在进行预测训练前, 先将关联程度低的输入进行剔除, 保证预测的精度和高效性。

从图 3 以及计算得到, 当权重累加到 95.27% 时, 与预测的电压偏差直接相关的自身相似日历史数据全部保留, 即 $d-1$ 、 $d-2$ 、 $d-7$ 天的 $t-1$ 、 t 、 $t+1$ 时刻电压偏差数据, 预测日当天的 $t-2$ 、 $t-1$ 时刻电压偏差数据。其中当天 $t-2$ 、 $t-1$ 时刻数据以及前三天 t 时刻数据明显比其他时间历史数据对预测量影响更大。其次对于外部因素以及与预测量非直接相关的功率, 结果表明, 温度、天气状况对电压偏差的影响较大, 而功率和日期类型、空气质量相对而言对该

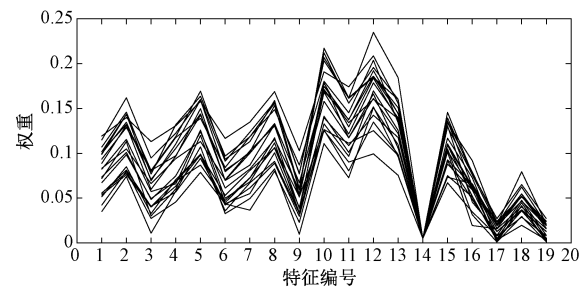


图2 20次训练中电压偏差的各个输入特征权重变化
Fig.2 Weight change of each input feature of voltage deviation in 20 trainings

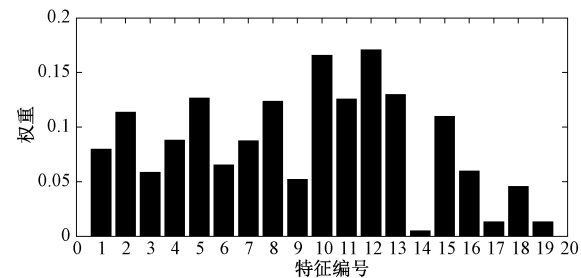


图3 电压偏差的19个输入特征统一化权重
Fig.3 Unified weight of 19 input features of voltage deviation

监测点影响较小,在现场考察时了解到该监测点主要是给一个重要化工厂供电,双休日也采取轮班不停运,因此日期类型对其影响较小。将7月20日和8月20日整天的 $d-2$ 的 t 时刻电压偏差、温度、湿度、天气状况、有功归一化值与对应电压偏差数据进行比较,如图4所示。可见这两天温度和湿度波峰波谷位置一般出现在电压偏差波峰后类似时刻,随偏差变化趋势也近似。两天的历史值偏离实时电压偏差的时刻和程度有些不同,但大体趋势可看出其间关系密切。而两天的有功变化可看出关系不大。由此说明 ReliefF 算法对于输入特征筛选的选择性和合理性。

4.1.2 电压谐波畸变率输入特征选择

同理将 A 相电压谐波畸变率利用 ReliefF 算法求得各个输入变量的特征权重,并重复 20 次,每次 19 个输入变量权重变化如图 5 所示。可以看出各个特征权重变化趋势基本统一,因此 ReliefF 算法对于电压畸变率的特征选择也具有适用性。各个输入都取 20 次的均值,且将均值权重除以总的权重和得到最终的统一化权重如图 6 所示。将统一化权重按从大到小排列,取其累加到 95% 的输入特征。

从图 5、图 6 以及计算可看出,对于电压谐波畸变率而言,前 11 个输入量的权重大小相差较小,当

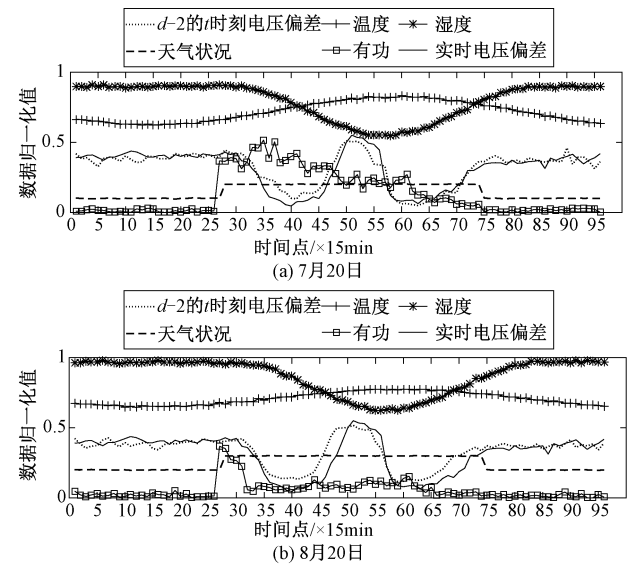


图4 两天实际数据比较
Fig.4 Comparison between actual data of two days

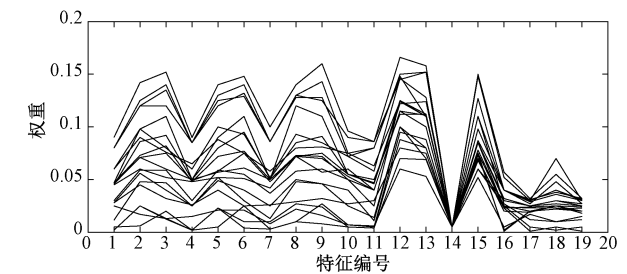


图5 20次训练中电压谐波畸变率的各个输入特征权重变化
Fig.5 Weight change of each input feature of voltage harmonic distortion rate in 20 trainings

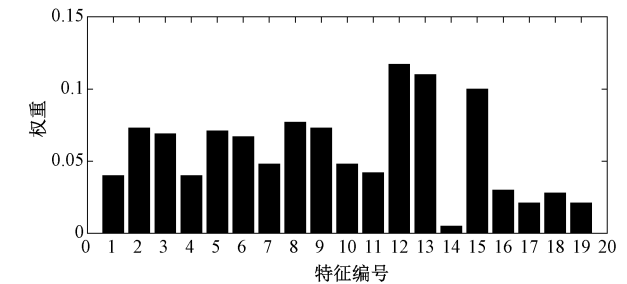


图6 电压谐波畸变率的19个输入特征统一化权重
Fig.6 Unified weight of 19 input features of voltage harmonic distortion rate

累加权重达到 96.08% 时,剔除输入量 14、17、19 即日期类型量化值、空气质量指数和无功。与电压谐波畸变率直接相关的自身历史数据也全部保留下来,其中 $d-7$ 天的数据相较于前三天的数据对预测量影响相对大一些。此外外部因素中温度和湿度对

畸变率的影响也明显大于其他因素,有功功率也对电压谐波畸变率有一定程度上影响。由此说明 Relief 算法的选择性和合理性。

将 7 月 20 日和 8 月 20 日整天的 $d-2$ 的 t 时刻电压谐波畸变率、温度、湿度、天气状况、有功归一化值与对应电压谐波畸变率数据进行比较,如图 7 所示。可见,历史谐波畸变率、温度、湿度、天气状况变化趋势与实时畸变率也是密切相关。

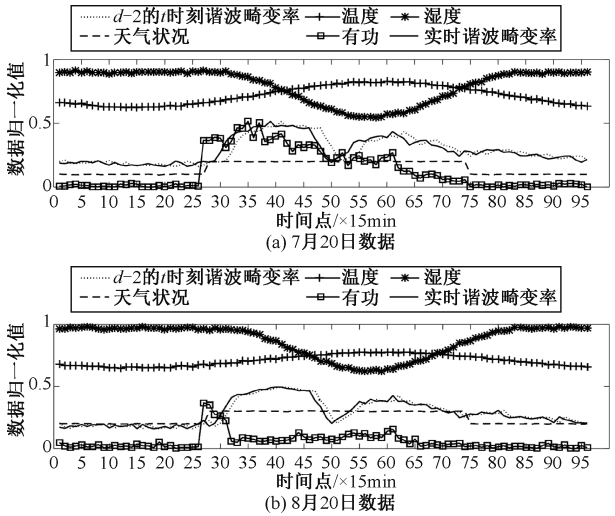


图 7 两天实际数据比较

Fig.7 Comparison between actual data of two days

4.2 预测结果分析

将电压偏差数据和对应输入变量数据分别输入 Elman 神经网络、随机森林、RBF-SVM 预测模型中训练预测,并将三种预测模型结果利用灰色理论组合且进行峰谷突变阶段数据修正。四种模型输出的预测数据及其相对误差如图 8 和图 9 所示,修正前后模型误差如图 10 所示。

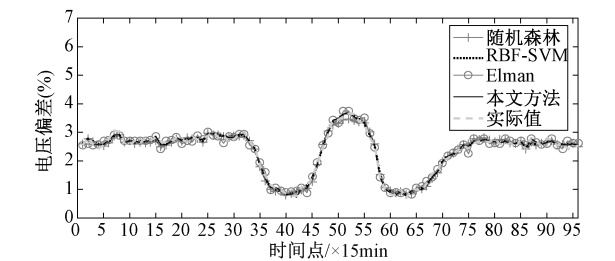


图 8 四种预测方法预测值与实际值比较

Fig.8 Comparison between predicted value and actual value of four prediction methods

利用均方根误差 $\varepsilon_{\text{RMSE}}$ 以及平均绝对误差百分比 $\varepsilon_{\text{MAPE}}$ 比较四种模型的预测效果, $\varepsilon_{\text{RMSE}}$ 计算如式 (17), $\varepsilon_{\text{MAPE}}$ 计算如式 (18),结果如表 2 所示。

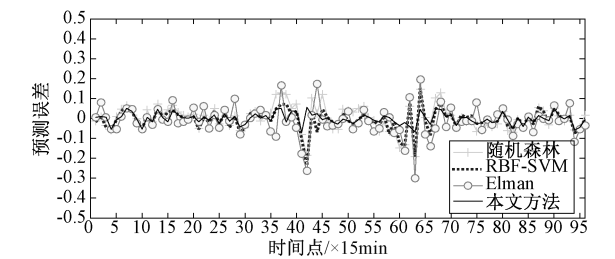


图 9 四种预测方法相对误差比较

Fig.9 Comparison of relative errors of four prediction methods

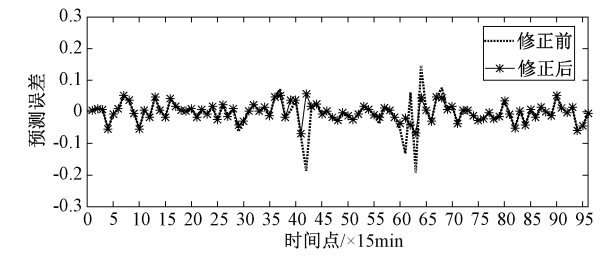


图 10 修正前后误差比较

Fig.10 Comparison of error before and after correction

$$\varepsilon_{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2}$$
 (17)

$$\varepsilon_{\text{MAPE}} = \frac{1}{N} \sum_{i=1}^N |y_i - y'_i| \times 100\%$$
 (18)

式中, N 为样本数量; y_i 为第 i 个样本的真实值; y'_i 为第 i 个样本预测值。

表 2 几种方法效果评价比较

Tab.2 Comparison of effect evaluation of several methods		
预测方法	$\varepsilon_{\text{RMSE}}$	$\varepsilon_{\text{MAPE}} (\%)$
Elman	0.123 9	5.26
随机森林	0.101 8	4.39
RBF-SVM	0.103 9	4.34
组合预测	0.081 2	3.66
改进组合预测	0.060 0	2.46

通过图 9 可看出,本文方法无论是在平稳变化阶段还是峰谷值阶段,预测相对误差都维持在 0.1 以内,可见本文方法的优越性。表 2 的评价指标可看出,本文提出的方法比均方误差最小的随机森林在 $\varepsilon_{\text{RMSE}}$ 上低了 0.041 8,比平均绝对误差百分比最低的 RBF-SVM 在 $\varepsilon_{\text{MAPE}}$ 上也低了 1.88%,修正后比修正前两个参数各降低了 0.018 8 和 1.20%。由上可知,改进组合预测在电能质量指标预测上具有优秀的准确度。

由于篇幅限制,A 相电压谐波畸变率只给出最

后仿真结果,如图 11~图 13 所示。几种方法效果评价比较见表 3。

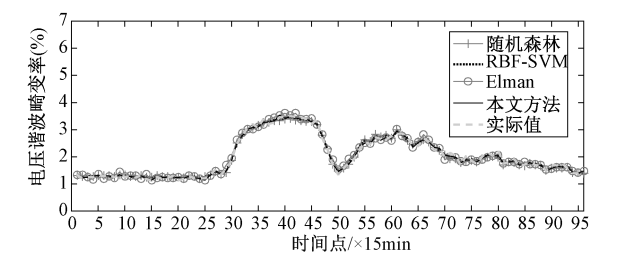


图 11 A 相电压谐波畸变率预测值与实际值对比
Fig.11 Comparison between predicted value and actual value of harmonic distortion rate of phase A voltage

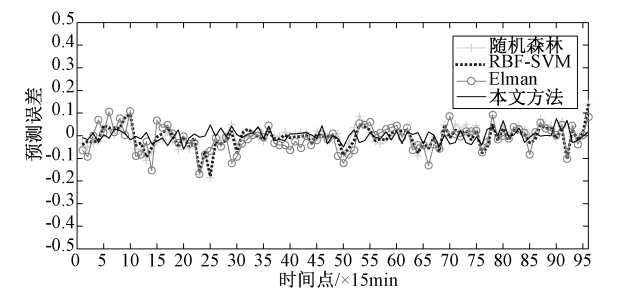


图 12 电压谐波畸变率相对误差对比
Fig.12 Comparison of relative error of voltage harmonic distortion rate

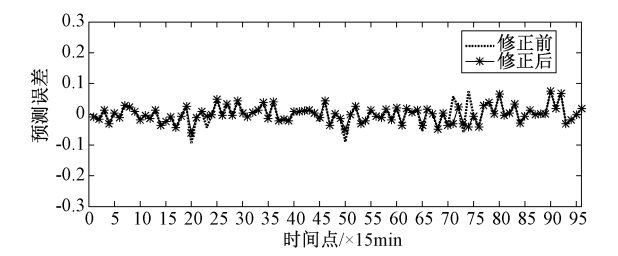


图 13 修正前后误差比较
Fig.13 Comparison of error before and after correction

表 3 几种方法效果评价比较		
Tab.3 Effect Evaluation Comparison of several methods		
预测方法	$\varepsilon_{\text{RMSE}}$	$\varepsilon_{\text{MAPE}}(\%)$
Elman	0.100 7	4.48
随机森林	0.096 8	3.74
RBF-SVM	0.092 6	3.40
组合预测	0.060 2	2.59
改进组合预测	0.051 0	2.12

4.3 预警阈值设定

根据国标 GB/T12325-2008 规定,10 kV 电压偏差允许范围为 $\pm 7\%$,电压谐波畸变率限值为 4.0%,因此对电压偏差和电压谐波畸变率不同预警

等级的阈值作出设定,如表 4 和表 5,按照所属等级来进行报警。

表 4 电压偏差预警等级划定	
Fig.4 Early warning level of voltage deviation	
预警等级	电压偏差绝对值范围
1	7%~7.5%
2	7.5%~8.5%
3	8.5%以上

表 5 电压谐波畸变率预警等级划定	
Fig.5 Early warning level of voltage harmonic distortion rate	
预警等级	畸变率绝对值范围
1	4%~4.5%
2	4.5%~5.5%
3	5.5%以上

5 结论

能源互联网的建设使得电能质量问题的关注度不断上升,本文提出了一种改进组合预测的电能质量指标预警方法。利用 ReliefF 算法求得统一化特征权重,减少输入维度,降低预测模型的复杂程度;利用主流的三组预测模型分别对电压偏差和电压谐波畸变率进行了预测,基于灰色关联理论有机组合模型预测结果,并对斜率突变的峰谷处进行预测值的修正,获得改进组合后的预测日数据;然后,根据国标规定进行不同等级预警。通过仿真计算说明,本文提出的方法在电能质量数据预测上有一定的优越性。最后,针对预警阈值的设定本文并未进行深入研究,不同监测点对数据的敏感程度不同,预警阈值也应不同,今后将在这方面进一步学习。

参考文献 (References):

[1] Wilsun X, Jing Y. Power disturbance data analytics - New application of power quality monitoring data [J]. Chinese Society for Electrical Engineering, 2013, 33 (19): 93-101.

[2] 林顺富, 谢潮, 汤波, 等 (Lin Shunfu, Xie Chao, Tang Bo, et al.). 数据挖掘在电能质量监测数据分析中的应用 (The data mining application in the power quality monitoring data analysis) [J]. 电测与仪表 (Electrical Measurement & Instrumentation), 2017, 54 (9): 46-51.

[3] Mohammadi Mohammad, Afrasiabi Mousa, Afrasiabi Shahabodin, et al. Detection and classification of multiple power quality disturbances based on temporal deep learning [A]. 2019 IEEE International Conference on Envi-

- ronment and Electrical Engineering [C]. Genoa, Italy, 2019.
- [4] 程志友, 吴吉, 刘荡荡, 等 (Cheng Zhiyou, Wu Ji, Liu Dangdang, et al.). 基于 AMIST 的复合电能质量扰动检测 (Detection of multiple power quality disturbance based on adaptive modified incomplete S-transform) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2020, 39 (5): 59-65.
- [5] 林顺富, 汤继开, 汤波, 等 (Lin Shunfu, Tang Jikai, Tang Bo, et al.). 典型电能质量稳态指标预测模型研究 (Study on forecasting model of typical power quality steady state indices) [J]. 电网技术 (Power System Technology), 2018, 42 (2): 614-620.
- [6] 丁泽俊, 刘平, 欧阳森, 等 (Ding Zejun, Liu Ping, Ouyang Sen, et al.). 电能质量预测与预警机制及其应用 (Mechanism of power quality forecast and early warning and their application) [J]. 电力系统及其自动化学报 (Proceedings of the CSU-EPSC), 2015, 27 (10): 87-92.
- [7] 王知芳, 杨秀, 潘爱强, 等 (Wang Zhifang, Yang Xiu, Pan Aiqiang, et al.). 基于改进集成聚类和 BP 神经网络的电压偏差预测 (Voltage deviation forecasting based on improved ensemble clustering and BP neural network) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2018, 37 (5): 73-80.
- [8] 于西西, 杨秀, 王知芳, 等 (Yu Xixi, Yang Xiu, Wang Zhifang, et al.). 基于改进 KPCA 和 GA-BP 神经网络的电能质量稳态指标预测 (Power quality steady state indices forecasting based on improved KPCA and GA-BP neural network) [J]. 水电能源科学 (Water Resources and Power), 2019, 37 (5): 189-193.
- [9] Vantuch Tomas, Misak Stanislav, Jezowicz Tomas, et al. The power quality forecasting model for off-grid system supported by multiobjective optimization [J]. IEEE Transactions on Industrial Electronics, 2017, 64 (12): 9507-9516.
- [10] Stuchly Jindrich, Misak Stanislav, Vantuch Tomas, et al. A power quality forecasting model as an integrate part of active demand side management using artificial intelligence technique - Multilayer neural network with back-propagation learning algorithm [A]. 2015 IEEE 15th International Conference on Environment and Electrical Engineering [C]. Rome, Italy, 2015. 611-616.
- [11] 刘建华, 刘艳梅, 冯纯纯, 等 (Liu Jianhua, Liu Yanmei, Feng Chunqun, et al.). 基于 k 中心点聚类的稳态电能质量预警阈值研究 (Early warning thresholding of steady-state power quality based on k-medoids clustering) [J]. 电测与仪表 (Electrical Measurement & Instrumentation), 2018, 55 (23): 41-45.
- [12] 卢彦锋, 郜慢慢, 钟红梅, 等 (Lu Yanfeng, Gao Manman, Zhong Hongmei, et al.). 考虑数据聚类特性的电能质量分级预警阈值设置方法 (Graded early warning threshold setting method of power quality considering the data clustering features) [J]. 电力电容器与无功补偿 (Power Capacitor & Reactive Power Compensation), 2016, 37 (3): 88-94.
- [13] 柏晶晶, 袁晓冬, 张帅, 等 (Bai Jingjing, Yuan Xiaodong, Zhang Shuai, et al.). 基于组合赋值法的稳态电能质量预警阈值研究 (Combinatorial weighting method for early warning thresholding of power quality) [J]. 电测与仪表 (Electrical Measurement & Instrumentation), 2014, 51 (12): 70-74.
- [14] 管鑫, 刘会家, 张振, 等 (Guan Xin, Liu Huijia, Zhang Zhen, et al.). 主动配电网短期负荷预测研究 (Research on short-term load forecasting of active distribution network) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2019, 38 (1): 31-38.
- [15] Lahouar A, Ben H S J. Day-ahead load forecast using random forest and expert input selection [J]. Energy Conversion and Management, 2015, 103: 1040-1051.
- [16] 马星河, 闫炳耀, 唐云峰, 等 (Ma Xinghe, Yan Bingyao, Tang Yunfeng, et al.). 基于优选组合预测技术的中长期负荷预测 (Medium and long-term load forecasting based on optimized combination forecast technology) [J]. 电力系统及其自动化学报 (Proceedings of the CSU-EPSC), 2015, 27 (6): 62-67.
- [17] 刘科研, 吴心忠, 石琛, 等 (Liu Keyan, Wu Xinzhong, Shi Chen, et al.). 基于数据挖掘的配电网故障风险预警 (Fault risk early warning of distribution network based on data mining) [J]. 电力自动化设备 (Electric Power Automation Equipment), 2018, 38 (5): 148-153.
- [18] 于道林, 张智晟, 韩少晓, 等 (Yu Daolin, Zhang Zhisheng, Han Shaoxiao, et al.). 计及需求响应的 Elman-NN 短期负荷预测模型研究 (Study of short-term load forecasting model based on Elman-NN considering demand response) [J]. 电工电能新技术 (Advanced Technology of Electrical Engineering and Energy), 2017, 36 (4): 59-65.
- [19] 霍娟, 孙晓伟, 张明杰 (Huo Juan, Sun Xiaowei, Zhang Mingjie). 电力负荷预测算法比较-随机森林与支持向量机 (Comparison between power load forecasting algorithms based on random forest and support vector machine) [J]. 电力系统及其自动化学报 (Proceedings

of the CSU-EPSC), 2019, 31 (7): 129-134.

[20] 王东, 史晓霞, 尹交英 (Wang Dong, Shi Xiaoxia, Yin Jiaoying). 不同核函数的支持向量机用于空调负荷预测的对比研究 (Prediction on hourly load of air conditioning by RBF support vector machine) [J]. 电工技术学报 (Transactions of China Electrotechnical Society), 2015, 30 (Sup. 1): 531-535.

Research on power quality early warning based on improved combined forecasting

LU Jue¹, SUN Yun-lian¹, XIE Xin-lin², ZHENG Long-wu², XU Bing-han¹, WU Ying¹

- (1. School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China;
2. Yingde Power Supply Bureau, Guangdong Power Grid Co. Ltd., Qingyuan 513000, China)

Abstract: With the construction of energy Internet and the increasing of distributed generation, as well as the gradual diversification of current power consumption forms, much attention has also been paid to the power quality problems. In order to ensure the power quality and the stable operation of the power grid, it is of great significance to have reasonable early warning for the power quality indicators. Therefore, an early warning model of power quality based on improved combined forecasting is proposed. In order to reduce the input dimension, the factors that may affect the power quality index are integrated and the ReliefF algorithm is used to eliminate the factors of low importance. Then, the random forest, Elman neural network and RBF-SVM intelligent algorithm are selected to predict the corresponding index, and the grey correlation analysis is used to calculate the correlation degree between the three methods and the real value to determine its weight. The three methods are combined to modify the prediction value at the peak and valley mutation to get the final prediction value. Finally, the simulation calculation is carried out with the data of a 10 kV line monitoring point in Guangdong Province, and the comparison proves the applicability of the method proposed in this paper.

Key words: power quality index; ReliefF algorithm; Elman neural network; random forest; RBF-SVM algorithm