第41卷 第8期 太 阳 能 学 报 Vol. 41, No. 8
2020年8月 ACTA ENERGIAE SOLARIS SINICA Aug., 2020

文章编号:0254-0096(2020)08-0173-07

一种光伏短期出力区间预测方法

张 娜[1], 王守相[2], 葛磊蛟[2], 王志和[1]

(1. 内蒙古工业大学电力学院, 呼和浩特 010080; 2. 天津大学智能电网教育部重点实验室, 天津 300072)

摘 要: 该文建立基于相似日模糊信息粒化和Elman神经网络的光伏短期出力区间预测模型。首先对原始序列进行相似日的选取,然后将提取的样本利用模糊信息粒化进行处理,确定预测区间的上下界,并结合Elman神经网络分别预测,构建区间预测模型。仿真结果表明,所提出的区间预测方法具有较高的预测精度和实用价值。

关键词: 光伏系统; 区间预测方法; 随机性; Elman; 相似日

中图分类号: TM74 **文献标志码**: A

0 引 言

光伏发电可缓解能源危机,净化环境,但其间歇性和不确定性给电力系统带来非常大的挑战。光伏出力的准确预测,可更好地对其加以利用,并且对降低电力系统运行成本、优化系统调度具有非常重要的意义。

目前对于光伏出力预测大多采用确定性的点值预测模型,如时间序列[1]、支持向量机[2]、人工神经网络[3-4]、马尔科夫链[5]以及各种组合模型[6-7]。文献[8]利用思维进化算法优化BP神经网络,粒子群算法优化SVM和单隐层前馈网络极限学习机分别预测,方差-协方差权值动态分法来组合预测结果。但对于随机性较强的光伏输出功率而言,上述点值预测反映不了其波动性和随机性,而采用区间预测技术更为适宜。

对于光伏出力的区间预测,部分学者从概率预测技术的角度进行了研究,如文献[9-10]从各影响因素的概率出发,预测未来短期光伏发电量的概率分布。还有部分学者建立光伏出力的概率模型,如单变量的假设参数分布的光伏概率建模[11]以及基于非参数核密度估计理论的光伏输出功率概率模型[12]等。虽然光伏出力的概率模型反映了光伏出力的部分特性,但涉及参数较多,复杂度较高,将其运用到实时的预测技术中,还需进一步调整和完善。

本文从区间边界进行考虑,着重对光伏出力的变化趋势和变化范围进行预测,提出一种基于模糊信息粒化和Elman神经网络的光伏短期出力区间预测模型。利用模糊信息粒化将相似日样本处理为反映变化规律的最小值、平均值和最大值3个序列,之后采用具有适应时变特性能力的动态Elman神经网络对边界进行预测,建立区间预测模型。通过实例仿真,验证该预测模型的可行性和有效性。

1 相似日选取原则

影响光伏出力的因素有很多,如果把所有的影响因素都应用到预测建模中,所建立的模型涉及的参数过多会导致模型过于复杂难以实现,另外,有些影响因素的数据难以清晰的获得,如云量。因此本文选取较易获取的天气预报中的数据进行相似日参数提取,采用的气象数据为温度T、风速Ws、湿度H、大气压p,首先利用距离分析法中皮尔森相似系数分析光伏出力Pout与各影响因素的相关性系数,式(1)为2个$1\times n$维向量x和y的皮尔森相关系数计算公式。

$$r_{xy}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \tag{1}$$

式中,x_i、y_i——向量x和y的第i各元素;$\bar{x}$、$\bar{y}$——向量x和y的均值。其次,将相关系数进行排序,选择相关系数较大的数据作为特征向量。

再次,利用"欧式距离"确定与预测日相似的气象条

收稿日期: 2018-02-08

基金项目: 内蒙古自治区自然科学基金(2017MS0501); 内蒙古自治区高等学校科学研究项目(NJZY251)

通信作者: 张 娜(1981—), 女, 博士、副教授, 主要从事智能配电和分布式发电出力预测方面的研究。zhangna2010337@163.com

件的日期组成时间序列。式(2)描述了第 i 天 X_i 和预测日第 j 天 X_j 之间的气象因素总体差异度的距离公式。

$$d_{ij}=\sqrt{\sum_{k=1}^{m}\left(X_{ik}-X_{jk}\right)^2} \tag{2}$$

2 模糊信息粒化

1997 年 Zadeh 教授提出信息粒化的概念,主要是把复杂信息按照对象的相近性、相似性、不可分性、函数性等进行集合的划分,也就是信息粒化的过程,信息粒化有基于模糊集理论、粗糙集理论、熵空间理论 3 种类型,本文采用模糊集形式表示的模糊信息粒。粒化及模糊过程如文献[13]所述,这里不做赘述。常用的模糊粒子隶属函数为三角型、抛物线型、高斯型、梯形,本文采用三角型模糊粒子,其隶属函数为:

$$A(x,a,m,b)=\begin{cases}0, & x<a\\ \dfrac{x-a}{m-a}, & a\leqslant x\leqslant m\\ \dfrac{b-x}{b-m}, & m<x\leqslant b\\ 0, & x>b\end{cases} \tag{3}$$

式中,x ——输入的时间序列;a、m、b ——原始数据变化的最小值、平均值、最大值。

3 ELMAN神经网络

光伏出力的数据非线性较强,传统的预测模型动态建模能力不够。而 Elman 型回归神经元网络中间层通过承接层的记忆进行自联,可增强历史数据的敏感性,增强网络的动态信息处理能力,还能以任意精度逼近任意非线性映射,对于光伏出力预测较为适用。Elman 型回归神经元网络分为 4 层:输入层、中间层(隐含层)、承接层和输出层[14],具体结构如图 1 所示。

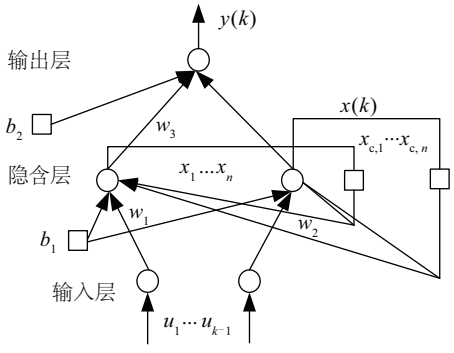


图1 Elman的结构图

Fig. 1 Structure of Elman

Elman 神经网络的隐含层的输出通过关联层的延迟与存储,自联到隐含层的输入,这种自联方式使其对历史状态的数据具有敏感性,内部反馈网络的增加提高了网络本身处理动态信息的能力,以达到动态建模的目的。同时,Elman 神经网络能够以任意精度逼近任意非线性映射,当给定系统的输入输出数据时,就可对系统进行建模。

Elman 神经网络的非线性状态空间表达式为:

$$\begin{cases}\boldsymbol{y}(k)=g\left[w_3x(k)+b_2\right]\\ \boldsymbol{x}(k)=f\left\{w_1x_c(k)+w_2\left[u(k-1)\right]+b_1\right\}\\ \boldsymbol{x}_c(k)=x(k-1)\end{cases} \tag{4}$$

式中,$\boldsymbol{y}$ —— m 维输出结点向量;$\boldsymbol{x}$ —— n 维中间层结点单元向量;$\boldsymbol{u}$ —— r 维输入向量;$\boldsymbol{x}_c$ —— n 维反馈状态向量;w_1 ——承接层到中间层的连接权值;w_2 ——输入层到中间层的连接权值;w_3 ——中间层到输出层的连接权值;b_1、b_2 ——输入层和隐含层的权值;$g(*)$ ——输出神经元的传递函数;$f(*)$ ——中间层神经元的传递函数。

Elman 神经网络采用梯度下降法对权值进行修正,学习指标函数采用误差平方和函数:

$$E(\omega)=\sum_{k=1}^{m}\left[y_k(\omega)-\tilde{y}_k(\omega)\right] \tag{5}$$

式中,$y_k(\omega)$ ——实际值;$\tilde{y}_k(\omega)$ ——预测值。

4 基于模糊信息粒化和Elman神经网络的区间预测模型

4.1 区间预测模型预测步骤

由于原始光伏出力数据具有较强的非线性和非平稳性,直接进行确定性点值建模,包含的信息较少,难以反映数据的波动性和随机性,因此本文提出基于模糊信息粒化和 Elman 神经网络的光伏短期出力区间预测模型,其流程如图 2 所示,具体步骤如下:

1)计算原始数据中各影响因素的相关系数,确定相似日的特征向量,进行相似日样本选取。

2)将提取的相似日样本采用模糊信息粒化进行处理,确定窗口的大小和预测的尺度,生成反映变化规律的最小值、平均值和最大值 3 个序列。

3)分别对模糊粒化生成的 3 个序列进行 Elman 神经网络的预测。

4)验证预测结果。

4.2 区间预测模型预测效果的评价指标

本文采用 4 项指标对区间预测的效果进行评判,即

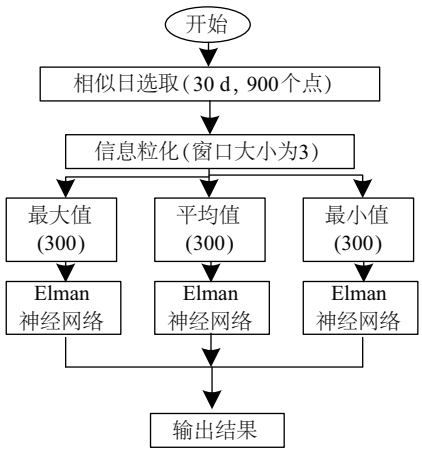


图2　预测算法流程图

Fig. 2　Flowchart of forecasting algorithm

平均区间位置误差平方和(*MSEP*)、平均区间长度误差平方和(*MSEL*)、平均区间误差平方和(*MSEI*)、平均区间相对误差和(*MRIE*)。

设某个对象的实际观察区间序列为:

$$x_t=\left[x_t^-,x_t^+\right]=\left(c_t,r_t\right),\ \ t=1,2,\cdots,n$$

预测对象的预测序列为:

$$\hat{x}_t=\left[\hat{x}_t^-,\hat{x}_t^+\right]=\left(\hat{c}_t,\hat{r}_t\right),\ \ t=1,2,\cdots,n$$

式中, x_t^- 、x_t^+ 、c_t 、r_t ——实际观察区间序列的下界、上界、区间中心、区间半径; $\hat{x}_t^-$ 、$\hat{x}_t^+$ 、$\hat{c}_t$ 、$\hat{r}_t$ ——预测区间的下界、上界、区间中心以及区间半径; n ——时间长度。

MSEP 反映了预测区间序列中点与实际区间序列中点的误差。*MSEL* 反映了预测区间序列半径与实际区间序列半径的误差。*MSEI* 为区间预测方法与实际序列的平均区间误差平方和。*MRIE* 反映的是预测区间序列与实际区间序列相离的程度,值越小,说明预测区间序列与真实序列相离的程度越小,两个区间序列越接近,预测就越有效。其表达式如式(6)~式(9)。

$$MSEP=\frac{1}{n}\sum_{t=1}^{n}\left(c_t-\hat{c}_t\right)^2 \tag{6}$$

$$MSEL=\frac{1}{n}\sum_{t=1}^{n}\left(r_t-\hat{r}_t\right)^2 \tag{7}$$

$$MSEI=\frac{1}{n}\sum_{t=1}^{n}\left(c_t-\hat{c}_t\right)^2+\frac{1}{n}\sum_{t=1}^{n}\left(r_t-\hat{r}_t\right)^2 \tag{8}$$

$$MRIE=\frac{1}{n}\sum_{t=1}^{n}\frac{\left(c_t-\hat{c}_t\right)}{\left(r_t-\hat{r}_t\right)} \tag{9}$$

5　算例及结果分析

5.1　数据来源

本研究选取自美国国家航空航天局(NASA)和澳大利亚沙漠太阳能研究中心(DKASC)[15]某一光伏发电系统 2009 年 4 月 1 日~2017 年 4 月 30 日天气与光伏出力的时间序列数据作为研究的原始数据。为避免季节对光伏出力和天气情况的影响,仅选用每年 4 月份的数据进行使用。图 3 为光伏出力的原始数据的基本情况和变化趋势,采样点间隔 20 min。

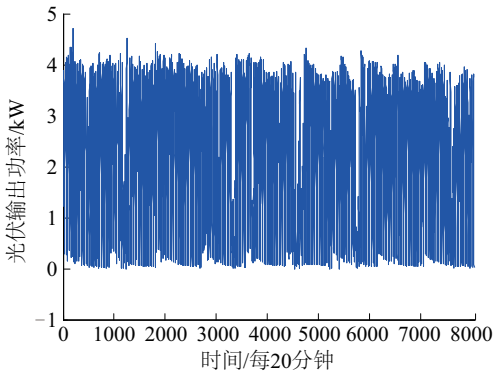


图3　光伏出力的原始数据

Fig. 3　Raw data of PV power

5.2　相似日选取

为了验证预测模型在不同天气状况下的预测效果,分别选取 2017 年 4 月 3 日(晴)、2017 年 4 月 20 日(雨)以及 2017 年 4 月 30 日(阴)为例进行预测。选取与相似日距离最近的 30 日作为预测样本。图 4 为阴天时的样本。

5.3　信息粒化

选取与预测日环境相似的相似日数据作为训练集,以每 3 个数据作为一个信息粒化窗口,将数据模糊粒化为 *Low*、*R*、*Up* 这 3 个参数,分别描述训练集的最小值、均值及最大值,并将粒化后的 *Low* 与 *Up* 作为实际观测值的上下区间。图 5 为阴天样本信息粒化后的结果。将粒化后数据的上下区间与原始数据进行对比,所

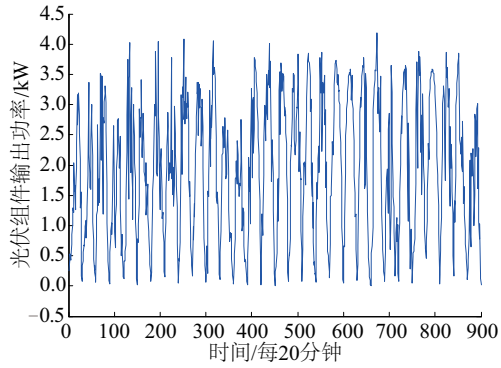


图4　2017年4月30日的相似样本

Fig. 4　Similar sample for 2017-04-30

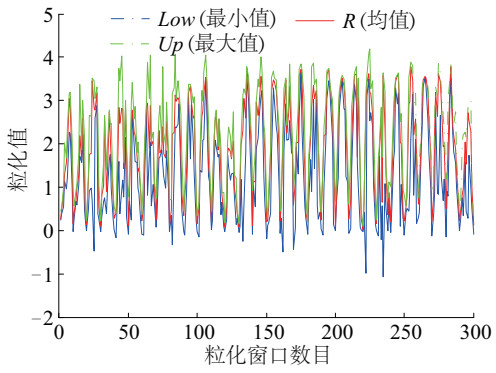


图5 2017年4月30日的粒化结果

Fig. 5 FIG results for 2017-04-30

有原始数据均在粒化区间范围内,将原始数据的均值与粒化后的均值进行比对,其相对误差为2.3473%,均方根误差为0.113。粒化后的数据区间与原始数据吻合度较高。

5.4 预测结果分析

本节建立4个模型比较说明构建相似日时间序列和FIG与Elman结合的预测模型对提升区间预测精度的作用。

模型1为传统BP神经网络与模糊信息粒化的区间预测模型,直接选取预测日前30 d的数据构成原始时间序列,将原始数据通过模糊信息粒化分解,分别对分解后的3个序列利用BP神经网络进行预测。

模型2为相似日的神经网络与模糊信息粒化的区间预测模型,采用距离公式将与预测日相似的30 d数据进行提取,对提取的数据采用FIG与BP神经网络结合的模型进行预测。

模型3为相似日的小波神经网络与模糊信息粒化的区间预测模型。

模型4为本文提出的方法。

考虑到光伏出力值应和前3~4个小时的数据关系较大,因此4个模型采用的输入数据均为预测点粒化后前4 h的数据,输出数据为粒化后预测点数据,采用滚动预测的方法进行预测。其中,传统BP神经网络结构为4,14,1;小波神经网络结构为4,6,1,小波基函数采用的是Morlet母小波基函数;Elman神经网络结构为4,11,1,1。其中BP神经网络隐含层,小波神经网络隐含层,Elman神经网络中间层中节点数量均通过枚举法进行确定,利用经验法确定枚举范围为[4,20],通过仿真实验确定各层的节点数量分别为14,6,11。

通过Matlab软件进行程序仿真,其预测结果分别如图6~图8所示。预测误差如表1所示。

模型1与模型2~模型4比较,可明显看出相似日的选取对于预测结果的精度提高有一定的作用。将不同类型的神经网络进行预测,即模型2~模型4的比较,可清楚地看出模型4具有较高的预测精度,特别是针对晴天情况时(4月3日),区间的预测中点、半径以及与原始区间相离程度都高度一致(见图6)。对于平稳的雨天情况的预测(4月20日),区间预测的结果可以预测出变化趋势和变化的范围,模型2的预测结果稍有延迟,模型3与模型4稍有改善(见图7)。针对于(4月30日)变化较快的阴晴天气,几种模型的预测结果都需要一定程度的改善,预测区间半径较大,预测区间的中心与实际区间的中心偏差较大,区间相离程度最小为0.1278。但是对比于4个模型,模型4基本上可跟随其变化趋势,其区间中心的预测精度较高,说明Elman网络本身处理动态信息的能力较强,对于随机性较强的光伏输出功率预测有一定的适应性(见图8)。

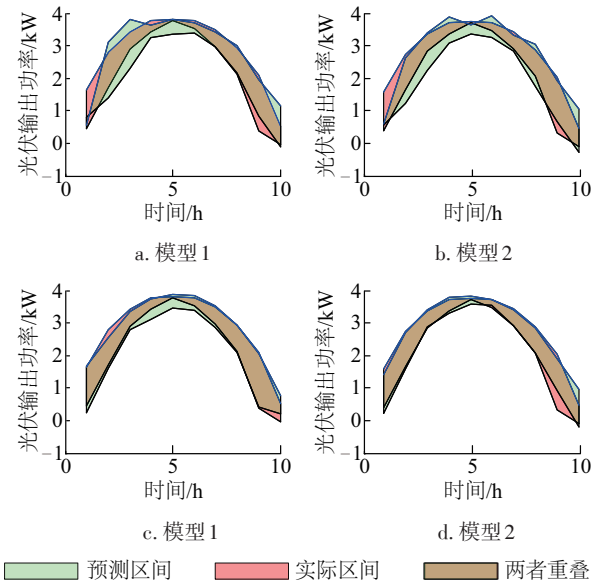
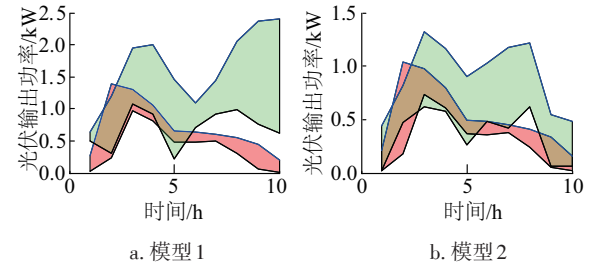
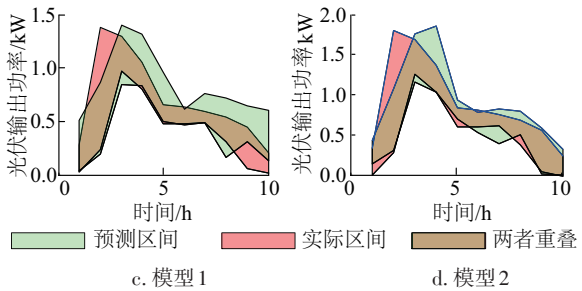


图6 2017年4月3日预测结果

Fig. 6 Forecasting results for 2017-04-03

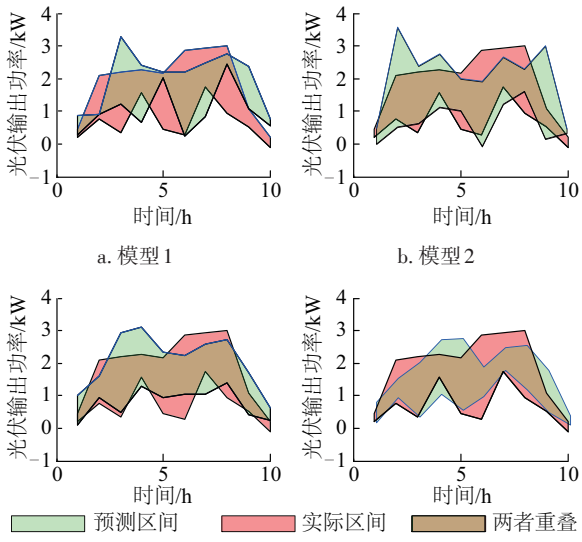




c. 模型1　　d. 模型2

图7　2017年4月20日预测结果

Fig. 7　Forecasting results for 2017-04-20



a. 模型1　　b. 模型2

c. 模型1　　d. 模型2

图8　2017年4月30日预测结果

Fig. 8　Forecasting results for 2017-04-30

将区间预测的中心作为点值预测结果与原始序列进行对比，从表1中*MSEI*的结果可看出本文提出的方法在各种天气情况下其预测精度也较高，因此也可将区间预测的中心作为点值预测的结果进行使用。

表1　不同模型预测结果比较(2017年4月)

Table 1　Comparison of effect among different forecasting models(April 2017)

日期	模型	*MSEP*	*MSEL*	*MSEI*	*MRIE*
3日	1	0.137930	0.452030	0.589960	0.10517
	2	0.030588	0.022383	0.052970	0.056696
	3	0.026723	0.012829	0.039552	0.062857
	4	0.009107	0.012557	0.024663	0.016375
20日	1	0.198940	0.459160	0.658100	0.237850
	2	0.047601	0.052636	0.100240	0.184280
	3	0.016804	0.026162	0.042966	0.125510
	4	0.013629	0.012763	0.026393	0.064282
30日	1	0.386850	0.385970	0.772830	0.236050
	2	0.292410	0.171290	0.463700	0.200290
	3	0.141380	0.095077	0.236460	0.142990
	4	0.147030	0.042460	0.189490	0.127800

6　结　论

本文提出的基于相似日模糊信息粒化和Elman神经网络的区间预测模型可实现不同天气情况的光伏出力预测，与传统的神经网络、模糊信息粒化神经网络、模糊信息粒化小波神经网络的预测模型相比较，预测区间覆盖度高，评价指标均有所改善，可为消纳光伏或计划安排等提供数据支持。另外根据*MSEI*的结果，区间预测的中心作为点值预测的结果进行使用。

致　谢：感谢天津大学智能电网教育部重点实验室对本研究的支持。

[参考文献]

[1] REIKARD G. Predicting solar radiation at high resolutions: A comparison of time series forecasts[J]. Solar energy, 2009, 83(3): 342-349

[2] 朱永强，田军. 最小二乘支持向量机在光伏功率预测中的应用[J]. 电网技术，2011，35(7)：54-59.
ZHU Y Q, TIAN J. Application of least square support vector machine in photovoltaic power forecasting[J]. Power system technology, 2011, 35(7): 54-59

[3] 茆美琴，丁杰，周海，等. 地基云图结合径向基函数人工神经网络的光伏功率超短期预测模型[J]. 中国电机工程学报，2015，35(3)：561-567.
MAO M Q, DING J, ZHOU H, et al. Short-term photovoltaic generation forecasting based on EEMD-SVM combined method[J]. Proceedings of the CSEE, 2013, 35(3):561-567.

[4] 张程熠，唐雅洁，李永杰. 适用于小样本的神经网络光伏预测方法[J]. 电力自动化设备，2017，37(1)：101-111.
ZHANG C Y, TANG Y J, LI Y J, et al. PV power forecast based on neural network with a small number of samples[J]. Electric power automation equipment, 2017, 37(1): 101-111.

[5] 丁明，徐宁舟. 基于马尔可夫链的光伏发电系统输出功率短期预测方法[J]. 电网技术，2011，35(1)：152-157.
DING M, XU N Z. A method to forecast short-term output power of photovoltaic generation system based on

Markov chain[J]. Power system technology, 2011, 35(1): 152-157.

[6] 王守相, 张娜. 基于灰色神经网络组合模型的光伏短期出力预测[J]. 电力系统自动化, 2012, 36(19): 37-41.
WANG S X, ZHANG N. Short-term output power forecast of photovoltaic based on a grey and neural network hybrid model[J]. Automation of electric power systems, 2012, 36(19): 37-41.

[7] 李乐, 刘天琪. 基于近邻传播聚类和回声状态网络的光伏预测[J]. 电力自动化设备, 2016, 36(7): 41-46.
LI L, LIU T Q. PV power forecasting based on AP-ESN[J]. Electric power automation equipment, 2016, 36(7): 41-46.

[8] 单英浩, 付青, 耿炫, 等. 基于改进BP-SVM-ELM与粒子化SOM-LSF的微电网光伏发电组合预测方法[J]. 中国电机工程学报, 2016, 36(12): 3334-3342.

[9] SHAN Y H, FU Q, GENG X, et al. Combined forecasting of photovoltaic power generation in microgrid based on the improved BP-SVM-ELM and SOM-LSF with particlization[J]. Proceedings of the CSEE, 2016, 36(12): 3334-3342.

[9] BRACALE A, CARAMIA P, CARPINELLI G, et al. A Bayesian method for short-term probabilistic forecasting of photovoltaic generation in smart grid operation and control[J]. Energies, 2013(6): 733-747.

[10] 董雷, 周文萍, 张沛, 等. 基于动态贝叶斯网络的光伏发电短期概率预测[J]. 中国电机工程学报, 2013, 33: 38-45.

[10] DONG L, ZHOU W P, ZHANG P, et al. Short-term photovoltaic output forecast based on dynamic Bayesian network theory[J]. Proceedings of the CSEE, 2013, 33: 38-45.

[13] RUIZ-RODRIGUEZ F J, HERNANDEZ J C, JURADO F. Probabilistic load flow for radial distribution networks with photovoltaic generators[J]. IET renewable power generation, 2011, 3(6): 110-121.

[12] 颜伟, 任洲洋, 赵霞, 等. 光伏电源输出功率的非参数核密度估计模型[J]. 电力系统自动化, 2013, 37(10): 35-40.
YAN W, REN Z Y, ZHAO X, et al. Probabilistic photovoltaic power modeling based on nonparametric kernel density estimation[J]. Automation of electric power systems, 2013, 37(10): 35-40.

[13] ZADEL L A. Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic[J]. Fuzzy sets and system, 1997, 90(2): 111-127.

[14] 时小虎, 梁艳春, 徐旭. 改进的Elman模型与递归反传控制神经网络[J]. 软件学报, 2003, 14(6): 1110-1119.
SHI X H, LIANG Y C, XU X. An improved Elman model and recurrent back-propagation control neural networks[J]. Journal of software, 2003, 14(6): 1110-1119.

[15] DKASC. BP Solar, poly-Si, fixed, roof mounted[EB/OL]. http://dkasolarcentre.com.au/locations/alice-spring.

A PHOTOVOLTAIC SYSTEM SHORT-TERM POWER INTERVAL FORECASTING METHOD

Zhang Na[1], Wang Shouxiang[2], Ge Leijiao[2], Wang Zhihe[1]

(1. *College of Electric Power*, *Inner Mongolia University of Technology*, *Hohhot* 010080, *China*;
2. *Key Laboratory of Smart Grid of Ministry of Education*, *Tianjin University*, *Tianjin* 300072, *China*)

Abstract: At present, the prediction of PV output power usually adopts point predicting model. For the strong randomness and volatility of the PV output power, the interval forecast method is more meaningfully. In this paper, an interval forecast model for PV short-term out put power based on similar day fuzzy information granulation and FIG-Elman is proposed. Firstly, the similar days are selected from the original sequence, and then the data are decomposed into the upper and lower bounds of the prediction interval by fuzzy information granulation. At last, the Elman neural network is used to forecast the upper and the lower bounds, respectively. Finally, an interval prediction model is established by combining Elman. The interval forecasting method is built. Simulation results show that the proposed method has higher prediction accuracy and the practical value.

Keywords: photovoltaic system; interval forecasting method; randomness; Elman; similar day