基于多变量灰色模型的光伏短期预测

武天雨，王 宁

（燕山大学电气工程学院，河北 秦皇岛 066004）

摘要：由于具有随机性与间歇性，属于清洁能源的光伏发电也给电网的稳定运行带来不利影响，因此需要对光伏出力进行准确预测，以便进行合理调度，保障电网安全性与经济收益。通过光伏出力的特性公式与实际物理环境挖掘影响光伏出力的影响因素，针对阴天情况将影响因素进行组合并纳入多变量灰色模型并进行预测建模，对光伏阵列在短期内的出力值进行预测。最后进行模型准确度的评估与误差定量分析，得出该方法可行性较高的结论。

关键词：光伏预测；多变量灰色模型；太阳光照强度；环境温度；风速

Short term photovoltaic prediction based on multivariable grey model

WU Tianyu, WANG Ning

(School of Electrical Engineering, Yanshan University, Qinhuangdao 066004, China)

Abstract：Due to its randomness and intermittency, photovoltaic power generation, which belongs to clean energy, also brings adverse effects to the stable operation of the power grid. Therefore, it is necessary to accurately predict the photovoltaic output, so as to carry out reasonable dispatching and ensure the security and economic benefits of the power grid. Through the characteristic formula of photovoltaic output and the actual physical environment, the influencing factors of photovoltaic output are explored. According to different weather types, the influencing factors are combined and incorporated into the multivariable grey model to predict the output value of photovoltaic array in the short term. Finally, the model accuracy evaluation and error quantitative analysis are carried out, and the conclusion that the method feasibility is higher is got.

Key words：photovoltaic prediction; multivariable grey model; sunlight intensity; ambient temperature; wind speed

0 引言

随着世界经济以及全球化的快速发展，能源也越来越不能完全满足现代人们的日常生产以及生活使用需求[1]。因此如何发展可再生清洁能源，对于各国的经济发展来说意义重大。太阳能发电作为可再生清洁能源之一，因其性能清洁环保无污染，生产成本低，同时光伏电站发电的正常运行以及维护的主要工作量较小，已成为目前科学研究的一个重点[2-3]。但是由于风力受到自然环境、气象等诸多自然因素的直接影响与电力限制，光伏风力发电机的出力仍然呈现随机性与出力间歇性[4]。由此可能会给已经接入光伏电站发电的其他电网设备造成冲击和电压波动，不利于保证电力系统的稳定正常运行。因此如何准确性地预测光伏发电出力，进而通过政策进行合理的能源调度，提升我国电网对光伏发电出力的综合消纳利用能力，成为当前学术研究的一个焦点[5]。

目前，出现很多应用气象等外界因素预测光伏出力的方法，如文献[6]运用相关性分析法分析光伏出力的影响因素，得出强影响因素；再利用强影响因素历史数据建立气象相似日特征向量，然后运用回声状态网络创建预测模型，利用相似

日历史数据训练回声状态网络，而后对待测日的输出出力值进行预测。但回声状态网络模型对样本数量要求较高，且对不良数据敏感，不适用于对模糊系统进行预测。文献[7]分析了雾霾对短期光伏出力的影响，讨论了预测过程中各类相关因素数据的预处理并加入雾霾因素，引入主成分分析、灰色关联度分析和加权相似度公式选取预测，但是该模型只在雾霾持续时间较长的情况下预测准确度高，没有给出在气象条件发生突变的情况下的预测模型。文献[8]利用发电量数据和同期气象数据进行气象因素与光伏发电量的灰色关联度和逐步回归分析。但是文中只进行了定性分析，没有对具体模型进行定量分析。文献[9]只给出了晴天天气的预测模型，没有考虑复杂天气下的预测模型。文献[10]提出了一种基于改进的灰色神经网络光伏出力预测方法。利用模糊 C 均值聚类的方法将数据分类，选择与待测日最为接近的一类作为样本，仅利用单变量灰色预测模型得到预测数据。

本文在上述文献的基础上进行扩展，同时考虑阴天的情况，建立模型并进行预测。单变量灰色模型对模糊系统的适应性较高，但准确度低，因此采用多变量灰色模型，提高预测准确度。通过基于多变量灰色模型的光伏短期预测，将光伏发电出力的天气影响因素直接考虑到综合建模当中，针对不同天气情况，将具有影响力的因素进行综合数值组合分析建模，使光伏短期发电出力的预测更加精确。

1 光伏功率的影响因素

光伏发电出力的影响因素主要分为天气因素和非天气因素，其中天气因素包含空气湿度、大气云层厚度、风速和风向。而空气湿度、大气云层厚度、风速和风向的影响可以统一归结为晴天、阴天对光伏出力的影响[11-13]。另外，雨天情况下光伏出力几乎为零[14]，因此不予考虑雨天。

根据

$$P=\eta SI\left[1-\frac{1}{200}(t_0+25)\right] \tag{1}$$

式中，η 为光电转换效率；S 为太阳能电池板面积；I 为太阳光照强度；t_0 为大气温度。

由于光电转换效率、太阳能电池板面积可通过人为进行优化改善，将其考虑作为建模时的影响因素意义不大，且运算复杂烦琐，因此不予考虑。本文最终考虑到的影响因素包括太阳光照强度、环境温度以及天气类型。

根据天气类型的不同，选择光照强度或者环境温度单独作为影响因素或者二者同时作为灰色模型的输入量。例如，在晴天状况下，光伏出力受天气影响较小，而在阴天，仅凭单变量灰色模型进行预测，可能会导致最终结果产生较大误差。

2 多变量灰色模型的建模

灰色系统，顾名思义系统的数据不完整，系统结构不固定，灰色系统中常用的工具之一灰色模型因此也具有同样特点，可以对部分数据已知而大部分数据未知的非线性系统进行预测[15]，而光伏出力的特性恰好是非线性特性。同时灰色模型中的多变量灰色模型包含多种变量，各个变量之间互相影响互相作用，而光伏出力的特性同样也受到诸多因素的影响，符合多变量灰色模型建模的要求。因此可以通过建立 n 元一阶常微分方程组来进行光伏出力的建模，预测模型中的参数会随着变量定量的变化而变化。反映到实际应用，即出力特性曲线的拟合准确度，多变量灰色模型不同于多个单变量模型的组合，该模型可以结合系统的目标因素和关联因素，定性与定量分析系统的特征，其步骤如下。

2.1 灰色预测模型建立

(1) 生成一次累加序列

$$x_i^{(1)}(j)=\sum_{k=1}^{j}x_i^{(0)}(k)\quad i=1,2,3;\ j=1,2,\cdots,n \tag{2}$$

式中，$x_1^{(0)}(k)$为光伏出力功率序列；$x_2^{(0)}(k)$为光照强度序列；$x_3^{(0)}(k)$为环境温度序列。

这一步是将离散的数据线性化，从而达到构建线性常微分方程组的条件[16]。

(2) 建立 n 元一阶常微分方程组

$$\begin{cases}\dfrac{\mathrm{d}x_1^{(1)}}{\mathrm{d}t}=a_{11}x_1^{(1)}+a_{12}x_2^{(1)}+a_{1n}x_n^{(1)}+b_1\\ \dfrac{\mathrm{d}x_2^{(1)}}{\mathrm{d}t}=a_{21}x_1^{(1)}+a_{22}x_2^{(1)}+a_{2n}x_n^{(1)}+b_2\\ \dfrac{\mathrm{d}x_3^{(1)}}{\mathrm{d}t}=a_{31}x_1^{(1)}+a_{32}x_2^{(1)}+a_{3n}x_n^{(1)}+b_3\end{cases} \tag{3}$$

记

$$\boldsymbol{A}=\begin{bmatrix}a_{11}&a_{12}&a_{1n}\\a_{21}&a_{22}&a_{2n}\\a_{31}&a_{32}&a_{3n}\end{bmatrix},\boldsymbol{B}=\left(b_1,b_2,b_3\right)^{\mathrm{T}}$$

则方程组可改写为

$$\frac{\mathrm{d}X^{(1)}}{\mathrm{d}t}=\boldsymbol{A}X^{(1)}+\boldsymbol{B} \tag{4}$$

(3) 对方程组进行离散化处理

$$x_i^{(0)}(k)=\sum_{j=1}^{n}\frac{1}{2}a_{ij}\left(x_j^{(1)}(k)+x_j^{(1)}(k-1)\right)+b$$

$$i=1,2,3;\ k=2,3,\ldots,n \tag{5}$$

(4) 参数辨识

$$\boldsymbol{H}=\begin{bmatrix}\tilde{a}_{11}&\tilde{a}_{21}&\tilde{a}_{31}\\\tilde{a}_{12}&\tilde{a}_{22}&\tilde{a}_{32}\\\tilde{a}_{1n}&\tilde{a}_{2n}&\tilde{a}_{3n}\\\tilde{b}_1&\tilde{b}_2&\tilde{b}_3\end{bmatrix}$$

$$\boldsymbol{Y}=\begin{bmatrix}x_1^{(0)}(2)&x_2^{(0)}(2)&x_3^{(0)}(2)\\x_1^{(0)}(3)&x_2^{(0)}(3)&x_3^{(0)}(3)\\x_1^{(0)}(n)&x_2^{(0)}(n)&x_3^{(0)}(n)\end{bmatrix}$$

$$\boldsymbol{L}=\begin{bmatrix}\overline{x}_1(2)&\overline{x}_2(2)&\overline{x}_3(2)&1\\\overline{x}_1(3)&\overline{x}_2(3)&\overline{x}_3(3)&1\\\overline{x}_1(n)&\overline{x}_2(n)&\overline{x}_3(n)&1\end{bmatrix}$$

式中，$\overline{x}_n^{(1)}(3)=\dfrac{1}{2}\left(x_n^{(1)}(3)+x_n^{(1)}(2)\right)$

运用最小二乘法计算 $\boldsymbol{H}$，即

$$\boldsymbol{H}=(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{Y} \tag{6}$$

得到参数 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的辨识值 $\tilde{A}$和 $\tilde{B}$。

值得注意的是，因为最小二乘法的本质是 $\boldsymbol{AX}=\boldsymbol{B}$ 的线性方程组求解，可能由于数据本身的原因，形成病态方程组。所谓病态方程组是指系数 $\boldsymbol{A}$ 的微小数量变化导致解 X 改变很大的方程组，$\boldsymbol{A}$ 作为系数矩阵被称为病态矩阵。而由 $\boldsymbol{AX}=\boldsymbol{B}$ 构成的方程组成为病态方程组，如果不考虑因为数据的关系而导致出现的该种情况，那么很可能导致预测出现巨大偏差甚至出现错误数据。

$\boldsymbol{AX}=\boldsymbol{B}$ 有解的条件是使 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的行数相等，即 $\boldsymbol{L}^{\mathrm{T}}\boldsymbol{Y}$ 与$(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})^{-1}$ 的行数一致。但是在实际运算当中发现，由于数据值本身的问题，很容易导致 $\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}$ 接近奇异值，因此采用求伪逆矩阵来求得参数 $\boldsymbol{H}$。

由伪逆矩阵的定理及条件可知，若 $\boldsymbol{L}^{\mathrm{T}}\boldsymbol{Y}\in R^{(m+1)m}$，$\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}\in R^{(m+1)(m+1)}$，则 $\boldsymbol{H}=(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})^{-1}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{Y}$ 有解的充要条件是

$$(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})\ (\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})^{+}\boldsymbol{H}=\boldsymbol{L}^{\mathrm{T}}\boldsymbol{Y} \tag{7}$$

此时可求出方程的通解为

$$\boldsymbol{H}=(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})^{+}\boldsymbol{L}^{\mathrm{T}}\boldsymbol{Y}+\left(\boldsymbol{I}-(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})(\boldsymbol{L}^{\mathrm{T}}\boldsymbol{L})^{+}\right)y \tag{8}$$

将式（3）进行离散化变形得

$$\tilde{X}^{(1)}(k+1)=\mathrm{e}^{\tilde{A}k}(X^{(1)}(1)+\tilde{A}^{-1}\tilde{B}\ )-\tilde{A}^{-1}\tilde{B} \tag{9}$$

其中

$$\mathrm{e}^{Ak}=I+\sum_{k=1}^{\infty}\frac{A^k}{k!}t^k\ \ t\in(1,n)$$

$$\tilde{X}^{(1)}(1)=\ X^{(0)}(1)$$

(5) 累减还原

$$\tilde{X}^{(0)}(k)=\tilde{X}^{(1)}(k)-\tilde{X}^{(1)}(k-1)\ \ k=2,3,\cdots \tag{10}$$

其中

$$\tilde{X}^{(0)}(1)=\tilde{X}^{(1)}(1)=x^{(1)}\ (1)$$

最后代入训练样本求出预测值。

2.2 误差估计

本文选用光伏预测常用评估方法中的均方根误差 RMSE 和平均绝对误差 MAE 评估预测的模型准确度。

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(X'(i)-X(i)\right)^2} \tag{11}$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}\left|X'(i)-X(i)\right| \tag{12}$$

式中，$X'(i)$ 为预测值；$X(i)$为实测值；i 为序列号。

3 算例设计

当太阳辐照度到达一定数值时，光伏阵列才输出功率，因此将数据的采样时间初步定为每天8:00～18:00，各整点为采样点，利用灰色模型预测待测日这 12 个整点的出力值。预测时选取连续多个样本日的数据作为灰色模型的输入量，太阳光照强度选择在整点时刻的值，环境温度取整点时的温度值。

另外，本文将太阳光照强度、环境温度和风速分别单独作为影响因素进行建模，对比三者作为影响因素时的误差，确定最终的预测模型。

本文数据来源于石家庄某光伏电站单位光伏阵列 2017 年 11 月的数据，包含当月 7:00～16:00 逐时的温度数据、光照强度数据以及光伏出力数据。

采用 MATLAB 2019b 作为实验平台，利用多变量灰色模型进行次日的光伏出力预测，计算出相应的预测值，并利用均方根误差 RMSE 和平均绝对误差 MAE 进行模型准确度的评估。

分析当年气象报告，2017 年 11 月 1～9 日为雾霾天气，7:00 的光伏出力为零，且白天的天气状况以阴天为主，因此选用 11 月 1～9 日 8:00～16:00 的数据作为样本数据，预测 10 日 8:00～16:00 的光伏实际出力情况，具体数据如表 1～表 3 所示。

表 1 光伏出力实测值

Tab. 1 Measured value of photovoltaic output (单位：W)

时间	1 日	2 日	3 日	4 日	5 日	6 日	7 日	8 日	9 日	10 日
8:00	106.2	64.9	109.6	103.8	103.9	110.3	75.3	77.1	64.0	73.3
9:00	218.4	211.2	211.5	198.5	223.6	218.9	197.7	196.3	203.7	150.0
10:00	334.5	327.4	278.7	278.7	301.2	286.5	294.2	287.4	280.8	239.9
11:00	244.4	245.9	342.1	286.4	224.5	340.5	348.5	248.7	234.1	320.9
12:00	361.3	334.5	354.7	381.7	373.3	362.0	361.5	355.7	333.5	318.8
13:00	324.7	282.5	285.7	335.4	314.0	312.2	318.0	327.8	286.3	308.0
14:00	139.0	132.5	243.2	164.4	150.4	206.1	187.8	143.5	127.8	151.1
15:00	130.6	124.7	128.2	111.2	116.6	108.7	101.8	101.6	113.3	103.5
16:00	34.3	36.52	36.0	33.98	32.38	29.92	31.46	33.4	29.68	27.4

表 2 光照强度数据

Tab. 2 Light intensity data (单位：W/m^2)

时间	1 日	2 日	3 日	4 日	5 日	6 日	7 日	8 日	9 日	10 日
8:00	133	81	137	127	127	135	97	95	79	92
9:00	275	269	268	247	278	273	260	247	255	189
10:00	430	426	356	349	380	361	395	367	356	304
11:00	319	328	443	365	286	433	469	320	300	410
12:00	476	452	461	487	479	465	485	462	428	409
13:00	430	387	371	433	406	403	430	426	368	398
14:00	185	181	317	212	195	269	256	187	165	195
15:00	174	169	166	143	151	141	137	132	145	133
16:00	45	49	46	43	41	38	42	43	38	35

表 3 温度数据

Tab. 3 Temperature data (单位：℃)

时间	1 日	2 日	3 日	4 日	5 日	6 日	7 日	8 日	9 日	10 日
8:00	9	8	8	5	5	5	13	7	7	9
9:00	9	11	10	7	7	7	16	9	8	9
10:00	12	14	11	8	9	9	19	11	10	10
11:00	15	19	14	11	11	11	20	13	12	12
12:00	16	20	14	11	12	12	19	14	12	12
13:00	17	22	14	13	13	13	20	14	12	13
14:00	18	22	15	14	14	15	22	15	13	14
15:00	18	21	14	13	14	14	20	14	12	13
16:00	17	20	13	13	13	13	19	13	12	12

这里着重给出含光照与温度的多变量灰色模型，能更精准地预测模型的仿真参数以及仿真波形，具体如表 4 所示。

表 4 仿真参数

Tab. 4 Simulation parameters

	1 日	2 日	3 日	4 日	5 日	6 日	7 日	8 日	9 日
光伏 a_{1n}	1.94	1.95	1.89	2.02	1.97	2.05	1.97	1.79	1.84
光照 a_{2n}	1.61	1.63	1.61	1.66	1.65	1.75	1.60	1.48	1.53
温度 a_{3n}	−0.66	−2.32	−3.95	−2.56	−4.28	−4.69	−1.30	−1.06	−0.34
常数 b	0	0	0	0	0	0	0	0	0

其中前三行对应 $\boldsymbol{H}$ 矩阵中的参数 A，第四行对应参数 B。

将太阳光照强度与温度数据单独作为变量以及共同作为变量代入到灰色模型中进行运算，得出 10 日的光伏出力数据结果如表 5 所示。

表 5 预测值

Tab. 5 Predicted value (单位：W)

时间	1	2	3	4	5
8:00	73.3	200.9	73.1	76.3	73.3
9:00	172	225.3	146.9	182.8	150.0
10:00	250	286.3	236.3	264.6	239.9
11:00	261	292.0	321.2	289.6	320.9
12:00	301	329.9	327.5	340.8	318.8
13:00	268	296.5	273.0	313.8	308.0
14:00	141	158.6	129.3	150.7	151.1
15:00	96.5	102.5	124.6	101.3	103.5
16:00	27.3	28.5	36.5	28.6	27.4
平均误差	18.5	34.5	11.4	13.7	0

其中 1 组为光照组，2 组为温度组，3 组为组合组，4 组为传统单变量灰色模型组，5 组为实测组。分析误差结果可知：在阴天时，含光照与温度的多变量灰色模型能更精准地预测短期的光伏出力情况。而该模型的 RMSE 值为 4.092 7，MAE 值为 2.970 0，模型误差符合标准[17]。

光伏出力、预测值和误差的仿真波形如图 1～图 3 所示。从仿真波形来看，实测值与预测值的分布趋势一致，且近似呈正态分布特性，二者拟合度较高。同时针对每天 10:00～13:00，光伏单位出力能稳定在 200 W 以上，可稳定向电网输送功率，因此可以在此时段对其进行有功调度，调节电网潮流。其次比较实测值与预测值曲线间的偏差，可以发现同一时刻，10 日内的误差变化幅度小，说明该预测模型不仅预测准确度高，而且稳定。同时，单条曲线上误差的增长率也保持较低水平，说明该模型不因气象条件突变而受到个别偏差数据的影响。其次待测日 10 日 10:00 之前误差虽在允许范围内，但相比其他时刻，误差仍较大，分析可能的原因是受当日云团移动等未知因素的影响而造成的。

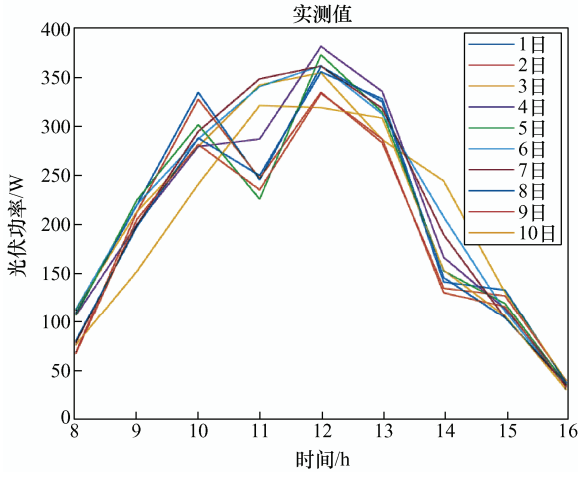

图 1　光伏出力仿真波形

Fig. 1 Photovoltaic output simulation waveform

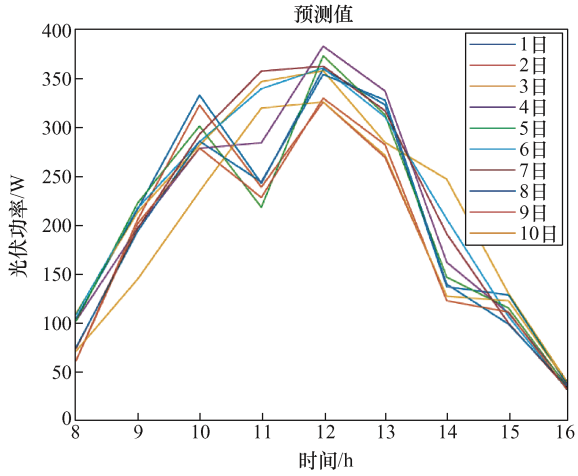

图 2　预测值仿真波形

Fig. 2 Prediction simulation waveform

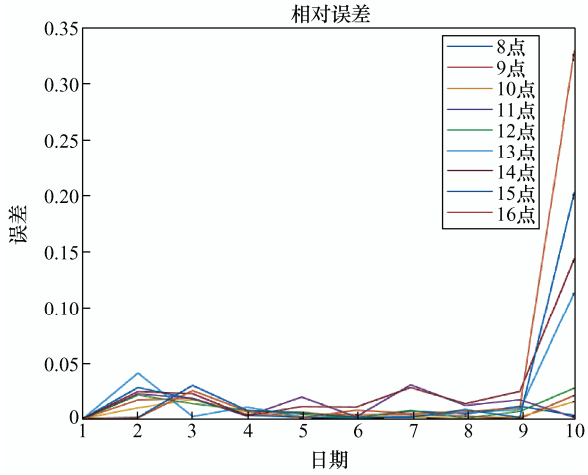

图 3　误差仿真波形

Fig. 3 Error simulation waveform

4　结束语

本文考虑了阴天情况下单变量灰色模型预测准确度不高的情况，提出了一种基于多变量灰色模型的短期光伏出力预测方法。利用光照强度和温度对光伏出力特性进行约束，将二者与光伏出力特性有机地集合在一起，提高了模型的预测准确度，一定程度上改进了模型的实用性能。

1) 该模型的 RMSE 值与 MAE 值符合误差标准。

2) 本文中用到了小时级的数据样本，数据跨度较大，模糊程度较高，但从实验结果来看，误

差变化幅度小，对模糊系统的适应性较高。

3）本文重点考虑的是太阳辐射度和温度，而没有考虑云量、空气湿度等因素的影响，因此，后期研究中，应综合考虑天气信息和各种影响因素，以提高不同天气状况的预测准确度。EA

参考文献

[1] 管霖，赵琦，周保荣，等. 基于多尺度聚类分析的光伏功率特性建模及预测应用[J]. 电力系统自动化，2018，42(15)：24-30.

[2] 朱想，居蓉蓉，程序，等. 组合数值天气预报与地基云图的光伏超短期功率预测模型[J]. 电力系统自动化，2015，39(6)：4-10.

[3] 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.

[4] 刘卫亮，刘长良，林永君，等. 计及雾霾影响因素的光伏发电超短期功率预测[J]. 中国电机工程学报，2018，38(14)：4086-4095.

[5] 赵书强，张婷婷，李志伟，等. 基于数值特性聚类的日前光伏出力预测误差分布模型[J]. 电力系统自动化，2019，43(13)：36-45.

[6] 安鹏跃，孙堃. 基于相似日和回声状态网络的光伏发电功率预测[J]. 智慧电力，2020，48(8)：38-43.

[7] 杨文海，程华新，高亚静，等. 计及雾霾影响的短期光伏出力预测和负荷预测[J]. 中国电力，2016，49(增刊 1)：148-153.

[8] 吕学梅，朱虹，王金东，等. 气象因素对光伏发电量的影响分析[J]. 可再生能源，2014，32(10)：1423-1428.

[9] 杨阳，练冲，马超. 晴天光伏发电功率的日内变化规律及预测方法研究[J]. 天津大学学报(自然科学与工程技术版)，2020，53(6)：565-572.

[10] 孙成胜，张红民，王艳，等. 一种改进的灰色 BP 神经网络光伏出力预测方法[J]. 重庆理工大学学报(自然科学版)，2020，34(2)：217-221.

[11] 袁晓玲，施俊华，徐杰彦. 计及天气类型指数的光伏发电短期出力预测[J]. 中国电机工程学报，2013，33(34)：57-64，76.

[12] 王飞，米增强，甄钊，等. 基于天气状态模式识别的光伏电站发电功率分类预测方法[J]. 中国电机工程学报，2013，33(34)：75-82，96.

[13] 焦润海，苏辰隽，林碧英，等. 基于气象信息因素修正的灰色短期负荷预测模型[J]. 电网技术，2013，37(3)：720-725.

[14] 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.

[15] 仇伟杰，刘思峰. GM(1，N)模型的离散化结构解[J]. 系统工程与电子技术，2006(11)：1679-1681，1699.

[16] 王允平，黄殿勋，熊浩清，等. 智能电网环境下采用关联分析和多变量灰色模型的用电量预测[J]. 电力系统保护与控制，2012，40(1)：96-100.

作者简介：

武天雨

武天雨(1994—)，男，硕士研究生，研究方向为光伏功率的中长期预测。

(收稿日期：2020-09-23)