

一种新的光伏发电预测模型的设计

秦俊举, 张恕远

(西华大学 机械工程与自动化学院, 成都 610039)

摘要:为了对光伏发电的输出功率进行预测,本文分析光伏发电的影响因素,提出了一种基于太阳辐射功率曲线匹配的预测模型。该模型将历史数据按时段进行分解,查找与当前时段太阳辐射功率曲线最为匹配的数据,以此构建并训练BP神经网络,来预测未来3个小时内的太阳辐射功率,能够较好的实现预测目标。实验结果表明,该模型有较高的精度,可对电网调度起到重要的指导作用。

关键词:光伏发电预测;太阳辐射曲线匹配;时段分解;BP神经网络

中图分类号:P422.1 **文献标志码:**A **文章编号:**1671-1807(2014)02-0146-05

光伏发电具有间歇性和波动性,其输出功率的剧烈变化会对要接入的电网造成巨大冲击,因此需要对光伏发电进行预测。

光伏发电预测主要分为两类。一类是利用数值天气预报和太阳辐射模型。2011年由我国气象部门研发的“太阳能光伏发电预报系统”基于中尺度数值天气预报,可以预报未来三天逐小时光伏发电量,精度达到国内领先水平。这种方法预测结果较为精确,但是计算复杂而且成本较高。第二种方法是利用历史气象资料,包括天气、太阳辐射功率等数据,通过神经元网络、模糊预测等方法进行预测,易于实现,且具有较高的精度,具有很大的实用性。本文侧重研究通过第二类方法预测光伏发电功率,文献[1]—[4]概述了该方向的相关研究成果。文献[1]将日射量、太阳能电池板的类型和温度及配线方案作为影响因素,它侧重于太阳辐射模型和光电转换模型,对天气情况的变化考虑较少。文献[2]将日太阳辐射及温度数据进行小波分解,对不同的成分采用不同的预测,但对一天中天气变化剧烈的情况仍不理想。文献[3]将气温、季节、天气类型作为考虑因素,采用神经元算法,利用了历史数据的相关性。但其过于简单地把某一天归为晴天、雨天、多云和阴天四种类型中的某一种,预测结果不准确。文献[4]通过“气候因素总体差异度”这样一个指标,考虑了太阳辐照强度、辐照时间、气温三个要素,从历史数据中找出与当前天气最为相关的数据来进行预测,但是该方法对天气变化较

大的情况辨识灵敏度不高。因此,本文提出了以太阳辐射曲线作为相似度的标准进行天气类型匹配搜索和预测。

1 基于太阳辐射曲线匹配的预测模型

1.1 影响光伏发电功率的因素

光伏发电功率的影响因素主要是光伏电池的光电转换效率和太阳辐射功率。光伏电池的光电转换效率越高,在相同的太阳辐射强度时,输出的功率就越大;而在光伏电池光电转换效率一定的情况下,太阳辐射功率越大,光伏电池的输出的功率也越大。

影响光伏电池光电转换效率的主要因素是温度,图1显示了不同温度下太阳能电池的伏安特性曲线^[5]。总体而言,随着光伏电池板温度的升高,电池板的输出功率要下降。因此,光伏电池板温度是光伏发电输出功率的重要影响因素。

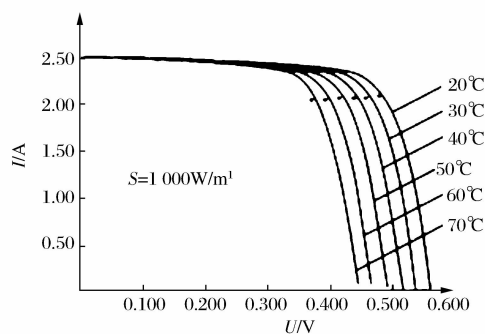


图1 光伏电池在不同温度下的伏安特性曲线

收稿日期:2013-10-31

作者简介:秦俊举(1988—),女,河南许昌人,西华大学硕士研究生,研究方向:光伏发电预测相关技术;张恕远(1964—),男,副教授,主要从事智能控制、数控技术、机电控制技术的研究。

除了温度,太阳辐射对光伏发电功率的影响可以分为两大类。一类是明确因素,包括太阳高度角、方位角、电站所处纬度、海拔等。太阳高度角、电站纬度、海拔影响太阳辐射强度;方位角影响太阳入射光线与太阳能电池板的夹角,当太阳光线垂直入射太阳能电池板表面时入射能量最强。这类因素的影响可以通过计算得到,故称明确因素。另一类不定因素则包括大气透明度、云的种类及云量、风速、风向等。这类因素影响太阳辐射到达地面的强度,而且对地面太阳辐射的综合影响难以用明确的物理公式来计算,故称不定因素。

通过与某实际光伏电站的合作,并综合考虑上述影响因素之间的相关性,本实验采用光伏电站易获得的数据类型,如太阳高度角、方位角这两个明确因素变量和电站温度、湿度、风向、风速这四个不定因素变量,以及实时的太阳辐射值来综合反映各环境变量对光伏发电的影响,力争避开复杂的物理模型,通过神经网络预测的方法对光伏发电功率进行预测。

本文主要通过时间段匹配的方法找到与当前时间段太阳辐射曲线相符的历史数据,利用历史数据构建预测网络并对待预测时段的发电功率进行预测。

1.2 太阳辐射曲线匹配模型

模型前提:将所有历史数据中 5 时到 20 时这 15 个小时的数据等间隔分为 5 段。如第 3 时间段代表 11 时、12 时、13 时。本模型基于以下假设:

- 1) 同一时间段的三个小时内,天气类型是大致相同的。
- 2) 相邻时间段,天气类型之间高度相关。
- 3) 在历史数据中,对当前三个小时的太阳辐射曲线进行匹配搜索。如果历史某一时段满足最佳匹配标准,则认为当前时间段天气类型与搜索到的历史天气类型相同。
- 4) 待预测的后三个小时的太阳辐射,可以由搜索到的历史匹配时间段的数据进行预测得到。

如图 2,首先将 $\max(\max$ 代表数据库中的总天数)天历史数据中的每一天都分为间隔三个小时的时间段。每一时间段有 1 954 个观测数据(每 5 秒钟记录一次)。每一个数据点记录着光伏电站太阳高度角 H 、太阳方位角 L 、实时的温度 T 、湿度 W 、风向 CD 、风速 CV 及太阳辐射功率 I 等观测数据,作为神经网络的输入数据类型。

通过对收集的某光伏电站一年内的气象数据及发电数据进行考查,上述假设在总体上是成立的。对于不符合该假设的个别情况,在进行数据库建立时将

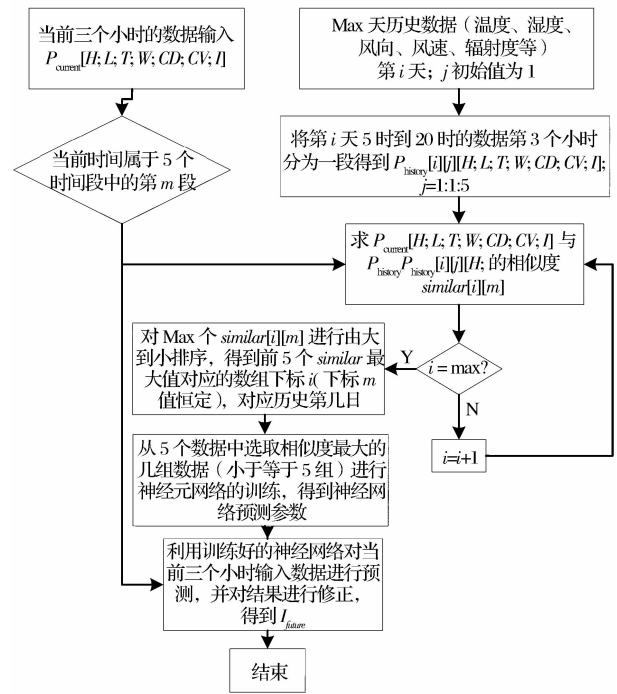


图 2 基于太阳辐射曲线匹配的光伏功率预测模型

这些极少量的数据作为“异常数据”排除。

预测步骤如下:

首先,判断当前输入数据 $P_{current} [H; L; T; W; CD; CV; I]$ 属于 5 个时间段中的第几段,即图 2-2 中 m 的取值。

其次,以 I (太阳辐射功率曲线)为对象,对 $\max$ 天的第 m 时间段历史数据进行遍历,得到历史第 i 天对应时间段 m 与当前时间段太阳辐射曲线的匹配度 $Similar[i][m]$ 。

$$Similar[i][m] = \frac{|I_{history}[i][m] - I_{current}|}{I_{current}} \quad (1)$$

其中 $I_{current}$ 为当前输入数据 $P_{current}$ 中的太阳辐射功率分量; $I_{history}[i][m]$ 表示历史第 i 天第 m 时间段内的太阳辐射功率序列。匹配度 $Similar[i][m]$ 的计算方法为:将历史第 i 天第 m 时间段太阳辐射功率 $I_{history}[i][m]$ 减去当前时间段太阳辐射功率 $I_{current}$, 然后将其绝对值与当前时段太阳辐射功率 $I_{current}$ 相除。

$$Lable[i] = \text{length}(\text{Find}(Similar[i][m] = (Similar[i][m] \leq 0.3))) \quad (2)$$

然后使用语句(2)对 $Similar[i][m]$ ($i=1, 2, 3 \dots \max$)序列中(对每一个 i 值, $Similar[i][m]$ 序列都包含所在时间段 m 内的 1 954 个点)值小于 0.3 的数据个数进行计数。计数值赋给 $Lable[i]$ 。Find 函

数功能是查找 $Similar[i][m]$ 中符合条件 $Similar[i][m] \leq 0.3$ 的数据点,而 $length$ 函数功能是计算这些数据点的个数。其中 0.3 是计算天气相符程度的阈值,选取的越小,匹配程度也越大,一般根据实验选取。

$$Sort(Lable[i]);$$

(3)

接着使用(3)对 max 个 $Lable$ 值排序。 $Sort$ 函数将 max 个 $Lable[i]$ 值由小到大排列。取 $Lable$ 最大的 5 个值所对应的下标 i ,即得到 5 个历史匹配日。

$$FN = Lable[i]/1954$$

(4)

另外,设定一个相似度衡量标准 FN 。 FN 的含义是求 5 个历史匹配日中符合 $Similar[i][m] \leq 0.3$ 条件点的个数与所在时间段内总的个数个数的比值。如果 FN 小于 0.6(根据经验选取的阈值),说明得到的 5 天虽然比其他历史日更接近当前天气,但是仍然与当前天气差别较大,则要进一步舍弃该天数据。

最后,将几组匹配数据所对应的 $[H; L; T; W; CD; CV; I]$ 分量作为输入,其后三个小时的太阳辐射分量 $I_{history-after}$ 作为输出,构建 BP 神经网络。并引入权重训练。使得匹配度最高的数据在训练中占据的权重最大。

训练完成后,即可将当前三个小时数据 $P_{current}$ $[H; L; T; W; CD; CV; I]$ 作为输入,得到后三个小时的太阳辐射输出值 I_{future} 。为了保证神经网络预测的收敛性,将历史匹配数组中下一时间段的太阳辐射值与当前预测输出值 I_{future} 进行加权平均,得到新的 I_{future} 值。可使预测更加准确。

2 预测实例

2.1 最佳匹配搜索

本文使用 MATLAB7 对某实际光伏电站一年的数据进行建模。选取 6 月 16 号作为预测日,选取该天第 3 段数据序列 $[H; L; T; W; CD; CV; I]$ 来预测第 4 段的太阳辐射输出值 I 。预测过程如下:

以 6 月 16 号第 3 段数据中的太阳辐射值 I 作为对象,在全年数据中的第 3 段数据库中寻找 I 的匹配,得到 5 组匹配度最大的数据。按相似度由大到小排列,下标分别为 180、40、146、245、302。其中最佳匹配下标为 180,表明与 I 最匹配的历史日为第 180 日,即 5 月 30 日。

5 月 30 日第 3 段数据 11 时、12 时、13 时的太阳辐射曲线如图 3 所示,而 6 月 16 号第 3 段实际的太阳辐射曲线如图 4 所示。相比之下可以看出,使用本文自定义的匹配标准搜索到的最佳匹配曲线与实际

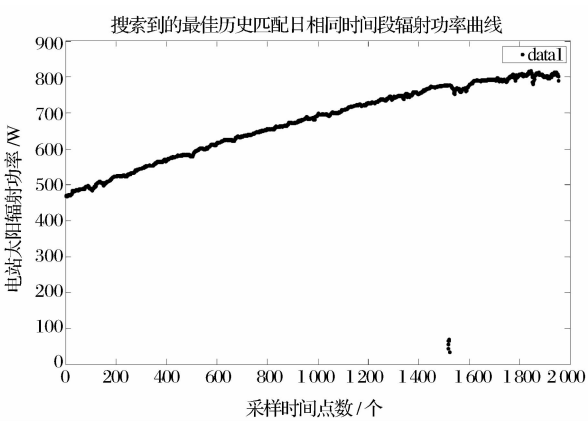


图 3 5 月 30 日第 3 段 11 时、12 时、13 时太阳辐射曲线

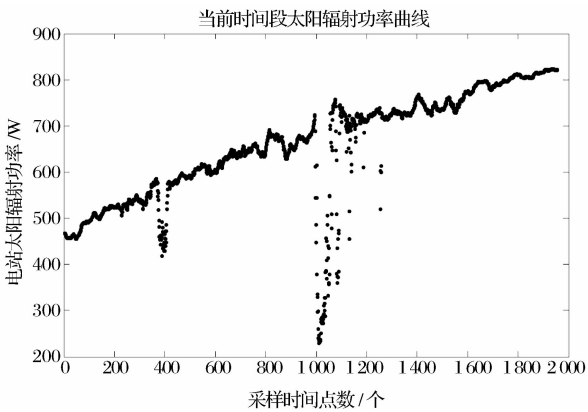


图 4 6 月 16 日第 3 段 11 时、12 时、13 时实际的太阳辐射曲线

曲线有较大的相符性。

除了最佳匹配,还找到了 4 个匹配度较高的数据组,下标为 40、146、245、302,对应 2 月 9 号、5 月 26 号、9 月 2 号、10 月 29 号。其 $Lable$ 与 FN 的值分别如表 1 所示,选取 $FN \geq 0.6$ 的 4 天数据得到历史匹配数据组。

表 1 匹配搜索结果所对应的 Lable 值与 FN 值

日期	2月9号	5月26号	5月30号	9月2号	10月29号	6月16号
Lable 值	1 386	1 291	1 542	823	1 195	1 954
FN	0.709 3	0.660 6	0.789 1	0.421 1	0.611 5	1.000 0

2.2 使用匹配数据组训练神经网络模型

BP 神经网络是基于误差反向传播算法的人工神经网络,具有自我学习能力。图 5 所示为神经网络的学习算法原理图^[6]。

本次试验选用输入节点数为 7,输出节点为 1,中间层为 1 的神经网络。训练输入为 4 组历史匹配

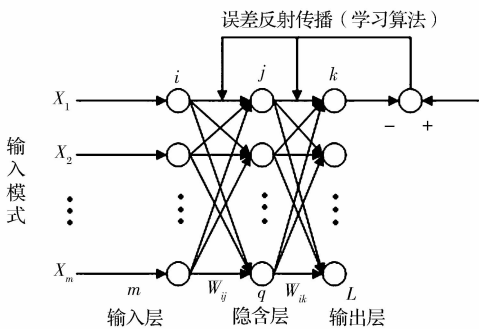


图5 神经网络的学习算法示意图

数据中第3时间段的太阳高度角、太阳方位角、温度、湿度、风向、风速及电站太阳辐射功率序列。训练输出为匹配数据中第4时间段的太阳辐射功率序列。

2.3 预测结果

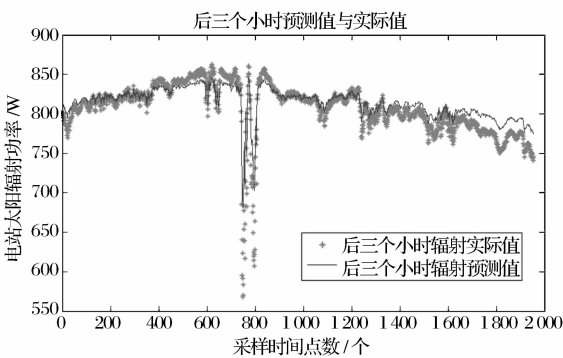


图6 由当前输入预测后三个小时的太阳辐射功率曲线

将当前三个小时的数据 $P_{current}[H;L;T;W;CD;CV;]$ 作为输入,利用训练好的神经网络,得到后三个小时的太阳辐射功率输出值 I_{future} 。6月16号14时、15时、16时太阳辐射预测值及实际曲线如图6所示。

2.4 误差分析

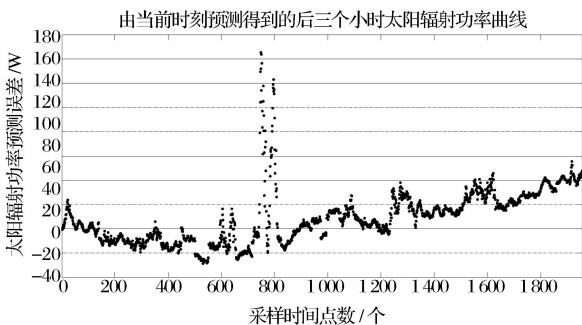


图7 后三个小时的太阳辐射功率预测误差曲线

图7是预测值减实际值所得的预测误差曲线。可以看出,在第750个采样点以内(在11时到12时之间),预测误差平稳而且较小;在第800个采样点附

近,实际辐射值变化非常大,预测机制可以快速的响应,但是缺点是响应的幅度不足。

另外值得注意的是随着预测时间的增长,误差有增大趋势。因为要预测的时间段越大,待预测数据与当前输入的相关性就越小。该不足可以通过对更长时间跨度的历史数据进行数据库的建立来改进。

为验证模型的适用性,随机对其它历史时段进行预测,预测误差计算公式如(5)所示:

$$Error = \frac{1}{1954} \times \sum_{l=1}^{1954} \left| \frac{I_{future}(l) - I_{future_real}(l)}{I_{future_real}(l)} \right| \quad (5)$$

其中, $Error$ 表示预测误差,1954表示预测时段(三个小时)内的总的的数据点数。 $I_{future}(l)$ 表示预测结果序列的第1个值,而 $I_{future_real}(l)$ 表示待预测时段真实的太阳辐射功率序列第1个值。所得的部分预测误差数据如表2所示。

从表2可以看出,对不同时段进行预测所得的预测误差是不同的。如本文实例中平均误差在5%左右,而有的情况误差则能达到30%左右。经过分析,误差大的情况是由于历史样本数据量较小,对于一些不常见的天气变化类型,所建立的分段数据库中并没有找到它的较好的匹配。但总体来说,平均误差处于20%以内的较小水平,实际为13.5%(取所有预测误差平均值),优于较新文献4中的两种预测模型(预测误差分别为20.4%和15.7%)。而且本文预测误差的计算考虑了全年中各个季节的预测结果,而非文献4中仅仅计算几天的数据,更具有普遍性。

表2 不同历史时段下的预测误差

待预测日期及时间段	预测误差	待预测日期及时间段	预测误差
3月12号第1段 预测第2段	20.25%	第3段预测第4段	6.95%
第2段预测第3段	12.59%	第4段预测第5段	10.87%
第3段预测第4段	9.93%	9月12号第1段 预测第2段	30.72%
第4段预测第5段	17.26%	第2段预测第3段	24.92%
5月12号第1段 预测第2段	17.43%	第3段预测第4段	14.06%
第2段预测第3段	11.18%	第4段预测第5段	27.61%
第3段预测第4段	8.24%	11月12号第1段 预测第2段	18.38%
第4段预测第5段	15.37%	第2段预测第3段	12.90%
7月12号第1段 预测第2段	9.2%	第3段预测第4段	10.39%
第2段预测第3段	6.75%	第4段预测第5段	19.23%

3 结论

本文采用基于太阳辐射曲线匹配的光伏预测模型,并运用某实际光伏电站的历史数据进行验证,具有较高的预测精度,优于同类型的预测模型。证明该方法具有一定的实用性,值得进一步研究。由于本文中使用的数据时间跨度不大,最佳匹配搜索得到的数据在实际上并不一定与当前类型完全相符,是预测误差的主要原因。该误差可以通过增加历史数据样本而减小。

参考文献

[1] 杨超. 太阳能光伏发电系统发电量的预测方法[J]. 太阳能

光伏发电技术,2011,5(2):29-34.
[2] 谢宏,陈志业,牛东晓,等. 基于小波分解与气象因素影响的电力系统日负荷预测模型研究[J]. 中国电机工程学报, 2001,21(5):5-10.
[3] 陈昌松,段善旭,殷进军. 基于神经网络的光伏阵列发电预测模型的设计[J]. 电工技术学报,2009,24(9):153-158.
[4] 李建红,陈国平,葛鹏江,等. 基于相似日理论的光伏发电系统输出功率预测[J]. 华东电力,2012,40(01):153-157.
[5] 卢静,翟海青,刘纯,王晓蓉. 光伏发电功率预测统计方法研究[J]. 华东电力,2010,38(4):0563-0567.
[6] 韩力群. 人工神经网络理论、设计及应用[M]. 北京:化学工业出版社,2001.

A Model Design for the Forecasting of Photovoltaic System Power Generation Power

QIN Jun-ju, ZHANG Shu-yuan

(The College of Mechanical Engineering and Automation,Xihua University,Chengdu 610039,China)

Abstract: To predict the power of photovoltaic power generation, we analyzed the factors of photovoltaic power generation and proposed a method of predicting photovoltaic power generation power based on the curve matching of solar radiation. By decomposing the historical data, searching the solar radiation curve, using the BP neural network, this method could predict photovoltaic power generation power within the next three hours and have a good performance. Experimental results show that this model can have a small forecasting error, which means that it's important in electric network management.

Key words: photovoltaic power prediction; solar radiation curve matching; period decomposition; BP neural network

(上接第 112 页)

The Scale of Investment,Consumption and Economy Growth

——An empirical analysis based on the time series data of Xinjiang

GU Cheng-jun

(Urumqi Central Sub-branch of People's Bank of China,Urumqi 830002,China)

Abstract: The results of empirical analysis show that: the relationship between Xinjiang's economy growth and its investment in a long-term equilibrium is a parabola with downward opening, and the relationship between Xinjiang's economy growth and its consumption in the long-run equilibrium is a parabola with upward opening. The highest level of output of Xinjiang's economy growth driven by its investment has not appeared, and the maximum driven-effect of Xinjiang's economy growth will reach when the scale of investment reached 4. 64666 trillion Yuan. The lowest level of output of Xinjiang's economy growth driven by its consumption appeared between 1979 and 1980, and the minimum driven-effect of Xinjiang's economy growth had reached when the consumption reached 31. 58051 billion Yuan, and its contribution to GDP only accounted for 1/5 to 1/4 to corresponding year. The investment scale of Xinjiang is still in the interval with the positive promoting effect on economy growth, and its consumption is in the interval with the positive promoting effect on economy growth, too. However, they are certain constraints on Xinjiang investment scale and its consumption level to enhance more.

Key words: the scale of investment;the consumption level;the economy growth;the co-integration test