

# 神经网络技术在风机 SCADA 数据分析中的应用研究

杜勉<sup>1</sup>, 易俊<sup>2</sup>, 郭剑波<sup>2</sup>, 程林<sup>1</sup>, 马士聪<sup>2</sup>, 贺庆<sup>2</sup>

(1. 清华大学电机工程与应用电力电子技术系, 北京市海淀区 100084)

2. 电网安全与节能国家重点实验室 (中国电力科学研究院), 北京市海淀区 100192)

## Research on the Application of Neural Networks on Wind Turbine SCADA Data Analysis

DU Mian<sup>1</sup>, YI Jun<sup>2</sup>, GUO Jianbo<sup>2</sup>, CHENG Lin<sup>1</sup>, MA Shicong<sup>2</sup>, HE Qing<sup>2</sup>

(1. Dept. of Electrical Engineering, Tsinghua University, Haidian District, Beijing 100084, China)

(2. State Key Laboratory of Power Grid Safety and Energy Conservation (China Electric Power Research Institute), Haidian District, Beijing 100192, China)

**ABSTRACT:** With the rapid development of offshore wind farms, the high cost of operation and maintenance is one of the important challenges for wind power development. Reducing the cost of operation and maintenance, and improving the availability of wind turbines draw widely attention from both domestic and worldwide. This paper proposed a data-driven performance evaluation method for wind turbine to improve the operation and maintenance efficiency and reduce the cost. This method combined neural network technology and stochastic process theory. It was adopted to analyze the wind turbine SCADA data, establish the wind turbine operation behavior model, and put forward the index to evaluate the fan performance. On the basis of this, the performance of the wind turbine was evaluated based on the SCADA data of 9 wind turbines. In view of the abnormal state of the wind turbine operation, the possible reasons were analyzed, and the corresponding maintenance suggestions were put forward. The results show that the method can effectively analyze the SCADA data, and the proposed indexes have important reference value for improving the operation and maintenance efficiency of the wind farm.

**KEY WORDS:** MarkovChain; Performance Evaluation; Wind Power Generation; Neural Network; Self-organizing Map

**摘要:** 随着海上风电场的快速发展, 高额的运维成本是风电发展面临的重要挑战之一, 降低运维成本, 提高风机可用性受到国内外广泛关注。本文提出了一种数据驱动的风机性能评估方法用来提高风机运维效率, 降低维护成本。该方法结合了神经网络技术和随机过程理论对风机 SCADA 数据进行分析, 建立了风机运行行为模型, 提出了评估风机运行性能的指标。在此基础上, 结合 9 台风机的 SCADA 数据, 评估了风机的运行性能。针对风机运行中出现的异常状态, 分析了可能的原因, 并提出了相应的维修建议。结果表明, 本

方法能够有效的分析 SCADA 数据, 所提指标对提高风电场运维效率具有参考价值。

**关键词:** 马尔可夫链; 性能评估; 风力发电; 神经网络; 自组织映射

**DOI:** 10.13335/j.1000-3673.pst.2016.3283

## 0 引言

风电场属于资本密集型项目, 其中建筑投资、风机制造与安装以及并网系统占比为 80%, 运行和维护成本占 20%<sup>[1]</sup>。因此, 保持风机以相对高的效率工作并制定最佳维修策略符合投资者的主要利益, 特别是在海上风电场复杂且恶劣的工作环境下, 保持风机高性能稳定运行更加重要。

降低运维成本的主要途径在于提高风机性能的和优化维护策略。为实现风机高性能稳定运行, 首先需要对风机运行性能进行评估, 然后利用评估结果优化运维策略。已有一些文章通过建立故障预测模型, 并将分析结果与故障模式进行匹配来解决该问题<sup>[2-5]</sup>。然而, 这些模型大都具一定程度的不确定性; 以及由建模过程中的简化带来的对风机运行复杂性的忽略。此外, 这些研究需要大量的高质量的历史故障数据, 这对该类方法的研究和应用带来了巨大的挑战<sup>[6-7]</sup>。

研究表明, 分析风机运行过程中采集的状态参数来是研究风机性能的有效方法<sup>[22][23]</sup>。本文将 SCADA 历史数据, 人工智能技术和随机理论结合起来, 保障研究成果具备良好的应用性的和坚实的理论基础。之所以采用风机的 SCADA 数据, 是因为它包含了反应风机组件运行状态的综合信息, 例

如风机有功功率, 齿轮箱轴承温度; 以及外部环境因素诸如风速, 环境温度等。文献[8-9]中采用 SCADA 数据来研究风机的运行行为和健康状况。

神经网络 (Neural Network, NN) 是一种被广泛使用的非线性建模技术<sup>[10-11]</sup>。近年来, 基于正常行为的偏差信号经常用于风机的故障和异常检测。文献[12]中, AP 的偏差被用于评估风机的性能。文献[13]中, GT 的偏差被认为是可能对风机健康状况有影响的风险指标。此外, 为了捕获偏差的行为模式, 本文采用自组织映射 (Self-Organizing Map, SOM) 来对偏差数据进行聚类分析。SOM 是为数据聚类设计的 NN 的变体<sup>[14]</sup>。文献[15]中, SOM 被训练用以对风功率的波动特性进行聚类分析。

基于 NN 的方法面临的重要问题之一是为基于人工神经网络的结果提供合理的解释<sup>[16]</sup>。为了解释聚类结果之间的关系, 并提供用于风机性能的信息, 本文提出了基于马尔科夫链 (Markov Chain, MC) 的性能指标及其计算模型。在风机可靠性分析中, 马尔可夫模型由于其无记忆性而被广泛使用。文献[17]中, 离散马尔可夫链被用于风机组件的可靠性建模。

本文首先对单台 SCADA 数据进行了预处理, 并在此基础上建立了基于神经网络的风机运行模型, 并计算输出功率预测值与实测值的偏差。然后应用自组织映射对偏差信号进行聚类以获取系统状态; 随后建立了基于 MC 的状态转移模型, 提出性能评价指标和计算方法; 最后将本方法应用到 9 台风机的 SCADA 数据中, 讨论了风机的运行状态, 并提出相应的维修建议, 验证了方法的有效性。

## 1 基于 SCADA 数据的风机性能评估方法

本节将通过三部分介绍本文所提方法, 各部分间逻辑关系如图 1 所示。

### 1.1 基于神经网络的风机运行行为建模

在进行建模之前, 需要对原始 SCADA 数据进行预处理。首先, 为了保持数据的一致性, 将缺省值替换为其前后数据点的平均值。排除不符合实际常识的数据点 (例如含有负值的有功功率)。其次, 为了进行对比研究, 将各个物理参数映射到统一范围。例如, 将风机运行参数中的风速 (Wind Speed, WS) 和功率 (Active Power, AP) 统一到在 [0,25] (m/s) 和 [0,2000] (kW)。其他参数范围及其相应的单位是: 转子转速 (Rotor Speed, RS) [0,20] (rpm), 环境温度 (Ambient Temperature, AT) [0,45] (°C), 齿轮箱

温度 (Gearbox Temperature, GT) [0,90] (°C)。

建立基于 NN 的风机运行行为模型。将 WS, AT, GT, 桨角 (Pitch Angle, PA) 和 RS 作为输入参数训练 NN 并预测 AP。上述参数通过采用皮尔森相关系数 [18] 作为参数选择标准。风机性能指标 “偏差 (Deviation, Dev.)” 被定义为:

$$\text{Dev.} = |\text{AP}_{\text{measured}} - \text{AP}_{\text{predicted}}| \quad (1)$$

将 Dev. 信号作为异常状态的表征量。偏差可能是由于模型本身的不准确性, 风机故障, 随机扰动或风机异常运行状态导致的。为减小误差, 当连续三个数据点中的两个在所得动态置信区间之外, 则认为该预测值有误, 作为误差处理。上述判据在质量工程中应用较多<sup>[13]</sup>。

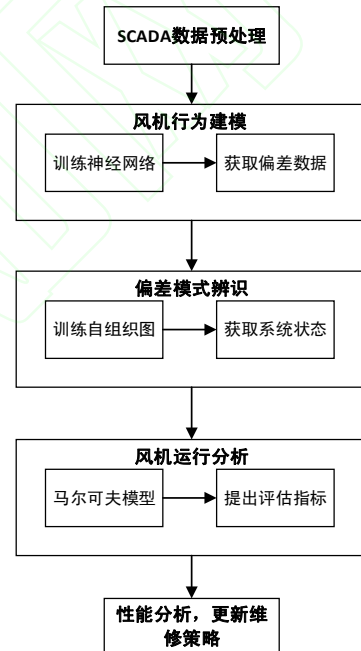


图 1 基于神经网络的风机性能评估方法框架

Fig1 Framework of the NN based wind turbine performance evaluation method

### 1.2 基于自组织映射的偏差聚类

由于所得偏差信号数量大, 模式重复度高, 采用基于无监督学习的聚类工具, 自组织映射 (SOM)<sup>[19]</sup>对偏差信号进行聚类分析。公式 (2) 表示偏差序列 Dev. 的  $t$  个数据点, 公式 (3) 表示状态空间  $S$  含有  $n$  个状态, 其中每个  $S_i$  表示聚类到状态  $i$  的偏差值。SOM 可以看作是集合  $D$  到  $S$  的一个映射过程。

$$D = \{d_1, d_2, \dots, d_t\} \quad (2)$$

$$S = \{S_1, S_2, \dots, S_n\} \quad (3)$$

以聚类结果为基础建立马尔可夫模型。状态转

移的驱动力包括是风机运行中的异常状态,故障,控制措施和采用的维修策略。

### 1.3 风机状态转移的马尔可夫模型

在这部分,通过建立有限状态空间的离散马尔可夫模型来描述风机性能状态。在风机运行在相对稳定的环境中时,马尔可夫模型可以用来对风机状态进行预测。通过统计聚类结果及其时间序列,可以计算风机状态转移概率矩阵  $T$  和初始状态概率  $I$ 。 $T$  和  $I$  分别表示为:

$$T=[p_{ij}]_{n \times n} \quad (4)$$

$$I=[q_1, q_2, \dots, q_n]_{1 \times n} \quad (5)$$

其中,  $p_{ij} = P(S_j|S_i)$ ,  $q_i = P(S_i)$ 。 $T$  表示从一个状态移动到另一个状态的概率,而  $I$  表示在开始处于每个状态的概率。Markov 模型的无记忆性可由如下公式所示:

$$P(S_n|S_{n-1}, S_{n-2}, \dots, S_1) = P(S_n|S_{n-1}) \quad (6)$$

在文献[20]中,作者定义了异常等级指数 (Anomaly Level Index, ALI),为风机运行过程各参数变化的区间阈值提供参考。为了描述和计算异常状态指数,基于马尔可夫模型,将状态转移路径  $(S_n^t, S_{n-1}^{t-1} \dots S_1^k \dots S_3^3, S_2^2, S_1^1)$  的概率定义为异常指数 (AI)。其数学定义如下

$$AI = P(S_n^t, S_{n-1}^{t-1} \dots S_1^k \dots S_3^3, S_2^2, S_1^1) \quad (7)$$

对于每个  $S_i^k$ ,上标表示时间序列,下标表示系统所处的状态。计算此概率的方法及其证明如下:

根据条件概率可以得到公式 (8):

$$P(S_n^t, S_{n-1}^{t-1} \dots S_1^k \dots S_3^3, S_2^2, S_1^1) = \quad (8)$$

$$P(S_n^t | S_{n-1}^{t-1} \dots S_1^k \dots S_3^3, S_2^2, S_1^1) \times P(S_{n-1}^{t-1} | S_{n-2}^{t-2} \dots S_1^k \dots S_3^3, S_2^2, S_1^1)$$

对于其中的每一条状态转移路径,可以按公式 (9) 进行描述:

$$\begin{aligned} &P(S_n^t, S_{n-1}^{t-1} \dots S_1^k \dots S_3^3, S_2^2, S_1^1) = \\ &P(S_n^t | S_{n-1}^{t-1} \dots S_1^k \dots S_3^3, S_2^2, S_1^1) \times \\ &P(S_{n-1}^{t-1} | S_{n-2}^{t-2} \dots S_1^k \dots S_3^3, S_2^2, S_1^1) \times \dots \times P(S_3^3 | S_2^2, S_1^1) \\ &\times P(S_2^2 | S_1^1) \times P(S_1^1) \end{aligned} \quad (9)$$

由于马尔可夫模型具有无记忆性,如公式 (6) 所示。因此公式 (9) 可以进一步简化为:

$$\begin{aligned} &P(S_n^t, S_{n-1}^{t-1} \dots S_1^k \dots S_3^3, S_2^2, S_1^1) = P(S_n^t | S_{n-1}^{t-1}) \times \\ &P(S_{n-1}^{t-1} | S_{n-2}^{t-2}) \times \dots \times P(S_3^3 | S_2^2) \times P(S_2^2 | S_1^1) \times P(S_1^1) \end{aligned} \quad (10)$$

其中  $P(S_n^t | S_{n-1}^{t-1})$  是状态  $S_n$  和  $S_{n-1}$  之间的状态转移概率。 $P(S_1^1)$  可以通过统计各状态中包含的数据点的出现频率计算。根据状态转移步长和跨越阶数可以定义多个异常指数。具体情况需要结合实际数据进

行判断。

## 2. 算例验证

将提出的方法分别用于分析 9 台风机的 SCADA 数据。其中 1-4 号风机和 5-9 号风机的数据时间跨度分别为 2 年和 5 年。采用周期均为 10 分钟。

### 2.1 神经网络模型说明

本文使用的 NN 共含有神经元 50 个,采用比例共轭梯度反向传播算法进行训练<sup>[1]</sup>。此外,以平方误差和 (SSE) 指标作为性能函数,并且最大遍历次数设置为 5000 次迭代。训练结果显示,所有 9 个 NN 模型的精确度约为 99%,这意味着它们能够以 99% 的准确度预测风机的正常行为。以图 2 为例展示风机 NN 模型的训练精度。

### 2.2 SOM 偏差聚类分析说明

文中使用的 SOM 有 24 个神经元,以均方差 (MSE) 作为其性能指数。图 3 展示了具有 24 个神经元的 SOM 的结构图。具有相同 (或相似) 颜色的神经元属于相同的簇,以具有明显区别的颜色数量作为类别数量。就风机 1 偏差的聚类结果来看,偏差信号共分为 4 类。观察图 4 中这 24 个神经元的权重曲线可得到相同的分类结果。基于聚类结果定义了风机 4 个性能状态 State1-4。这些状态数值范围如图 5 所示。

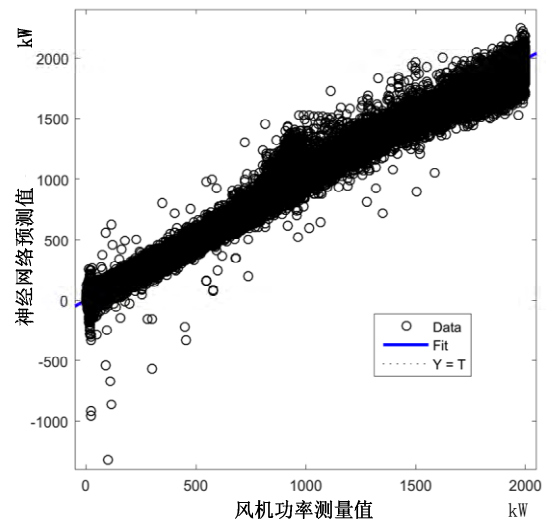


图 2 1 号风机的神经网络模型训练精度

Fig2 Performance of NN model for wind turbine 1





图 3 1 号风机 SOM 中神经元分布及聚类结果

Fig3 SOM for wind turbine 1 with Neurons positions and clusters

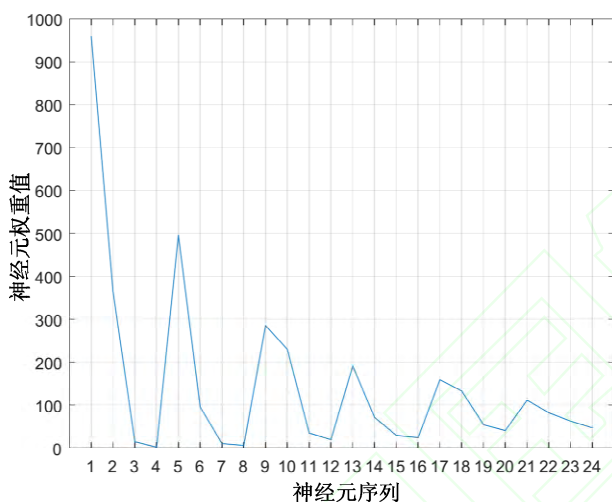


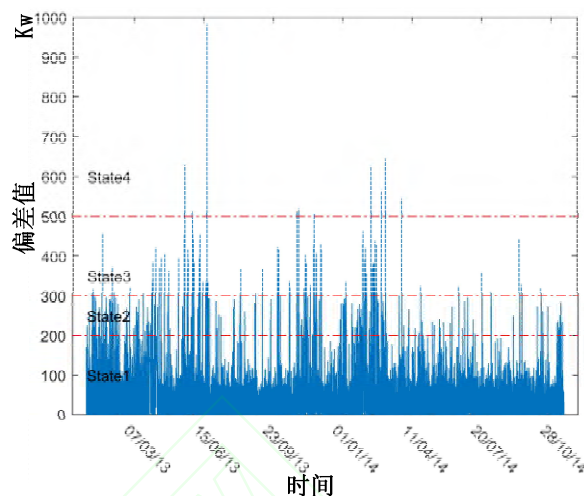
图 4 1 号风机 SOM 中各神经元的权重值分布

Fig4 Distribution of weights of neurons in SOM for wind turbine 1

在得到聚类状态后,调查各状态中数据对应输入参数是否存在异常来明确各状态的含义。为此,从每个状态抽取几个样本,对应实际运行中各参数的值及故障记录来检查是否存在异常。抽查结果表明, *State1* 和 *State2* 中选择的样本没有显示任何物理问题; *State3* 中极少部分数据, *State4* 几乎所有数据均在实际运行记录中找到对应的异常记录。

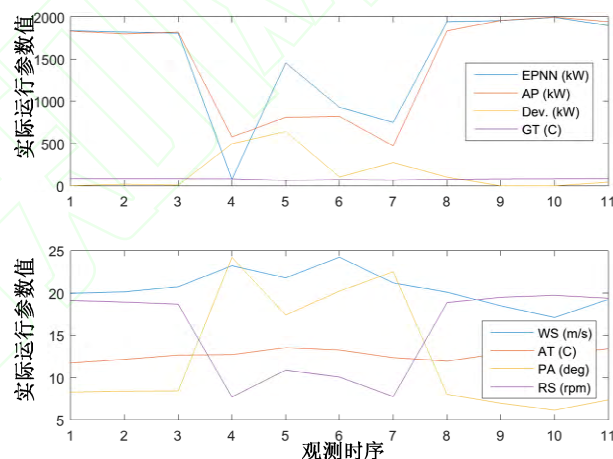
图 7 中显示了选自 *State4* 的样本数据,该样本包含了大约两个小时的操作时段,11 个连续数据点。其中(Expected Power by NN, EPNN)和 Dev.分别代表通过 NN 预测值和偏差信号。从图中可以看出,该时段 WS 稳定保持在较高水平。然而 AP 却出现明显下降。观察到当 WS 接近切入速度时导致 PA 行为异常。在正常运行情况下,变桨系统应该是小幅度的调节,保持输出平滑功率。图 7 中 PA 异常行为导致了约 60%的生产损失。因此,基于 NN 的

模型和偏差信号可以有效的描述异常行为。



5 1 号风机偏差信号及各状态区间

Fig5 Deviation signal and states for wind turbine 1

图 7 1 号风机 *State4* 中 2 小时运行过程中各参数变化Fig7 Deviation sample from *State4* of wind turbine 1

## 2.3 马尔可夫模型

表 1 给出了 1 号风机的状态转移矩阵。

表 11 号风机状态转移矩阵

Table2 Transition Possibility Matrix for wind turbine 1

状态	<i>State1</i>	<i>State2</i>	<i>State3</i>	<i>State4</i>
<i>State1</i>	0.99400	0.00497	0.00101	0.00002
<i>State2</i>	0.62901	0.27755	0.09066	0.00279
<i>State3</i>	0.33865	0.27092	0.36255	0.02789
<i>State4</i>	0.35714	0.14286	0.28571	0.21429

图 8 为风机 1 的马尔可夫模型,其中数值代表各状态间的转移概率。考虑到前面的分析,绝大部分 *State4* 中的观察值都对应了风机的异常装态。从图 8 中可以得到从其他状态转移到 *State4* 的概率和路径。并且可以看出从 *State3* 到 *State4* 的转移概率最高,这意味着 *State4* 中的大多数观察值是从 *State3*

转移过来的。因此，监测过程中当偏差值到进入 State3 的范围时，需要密切注意风机运行状况以避免故障的发生。本文中也为其他 8 台风机创建了相应的 MC 模型。

对风机 1 而言，PA 系统的异常是导致其偏差信号进入 State4 的主要原因之一。实际运行中需要对 PA 系统进行故障排查。基于 MC 模型提供的信息，图 9 指出 PA 系统的问题平均每 20 天发生一次。值得注意的是，该问题每出现一次将带来至少 500kW 的功率损失。这不仅导致利润损失，而且当风电场达到一定规模后还会给电网带来稳定性问题<sup>[21]</sup>。

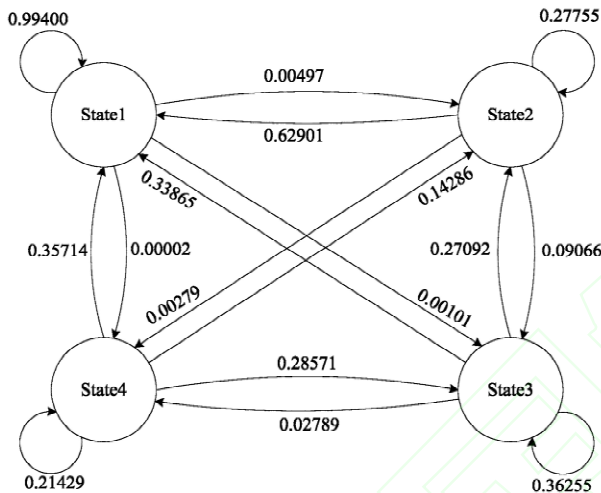


图 8 1 号风机的 4 状态马尔可夫模型

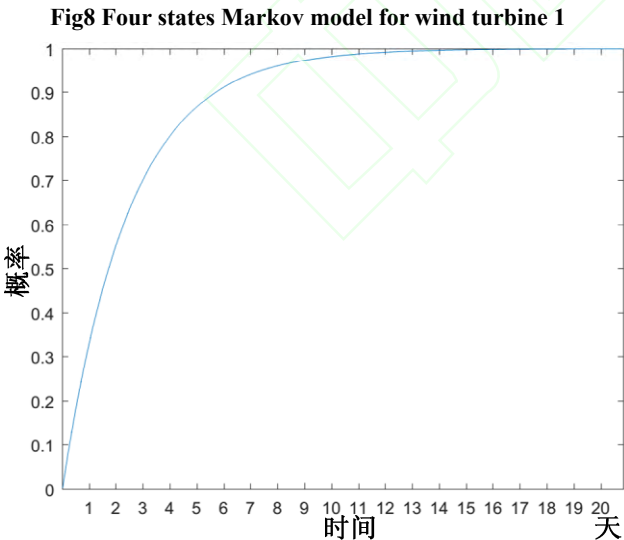


图 9 1 号风机变桨系统故障发生频率

Fig9. Frequency of pitch system failure for wind turbine 1

根据研究结果，定义当风机在从 State3 变为 State4 时，风机的性能存在微小的异常，即异常指数 1 (AI1)。定义当风机在 State1 停留至少两个时

间步长然后直接转移到 State4 时，风机存在严重异常状态，即异常指数 2 (AI2)。这两个指标对应风机实际运行中两个不同的安全级别，不同安全级别对应了采取措施的优先排序，可以为决策提供参考依据。

根据公式 (7)，可以 AI1 表示为：

$$P(x_2=S_4,x_1=S_3)=P(S_4^2|S_3^1)\times P(S_3^1) \tag{11}$$

AI2 可以表示为：

$$P(x_3=S_4,x_2=S_1,x_1=S_1)=P(S_4^3|S_1^2)\times P(S_1^2|S_1^1)\times P(S_1^1)\tag{12}$$

表 2 给出了 9 台风机初始状态概率的计算结果；表 3 给出了基于上述两个假设提出的异常判断指标应用到 9 台风机的计算结果。

表 2 9 台风机的初始状态概率

Table3 Initial state possibility of all 9 wind turbines				
风机	$P(S_1^1)$	$P(S_2^1)$	$P(S_3^1)$	$P(S_4^1)$
风机 1	0.9892	0.0079	0.0028	0.00015
风机 2	0.9510	0.0382	0.0079	0.00290
风机 3	0.9494	0.0449	0.0052	0.00048
风机 4	0.9870	0.0120	0.0008	0.00021
风机 5	0.9900	0.0079	0.0019	0.00018
风机 6	0.9699	0.0282	0.0018	0.00005
风机 7	0.9557	0.0400	0.0037	0.00056
风机 8	0.9943	0.0038	0.0015	0.00045
风机 9	0.9860	0.0122	0.0014	0.00037

表 3 9 台风机的异常指标计算结果

Table4 Results of anomaly indexes for 9 wind turbines		
异常指标	AI1	AI2
状态转移路径	S3-S4	S1-S1-S4
风机 1	0.00008	0.00002
风机 2	0.00071	0.00042
风机 3	0.00010	0.00006
风机 4	0.00002	0.00007
风机 5	0.00005	0.00006
风机 6	0.00002	0.00001
风机 7	0.00012	0.00016
风机 8	0.00004	0.00006
风机 9	0.00004	0.00010

3 结果分析

根据 AI2 比较各风机的性能状态。在风机 1-风机 4 之间，风机 2 具有更高的轻微异常的风险；而在风机 5-风机 9 之间，风机 7 的 AI2 值最大，相应面临的风险就越高。根据 AI1 来看，风机 2 和风

机 7 同样排在高风险的前列。从中可以得出风机 2 和风机 7 相比其他风机的运行载荷更高。

对表 3 的另一种解读,可以通过比较每台风机两个指数的值来推断当前风机的性能状态。例如,如果 AI2 的值大于 AI1 的值,则风机在运行过程的风险将会更高。因此,在实际运行中,对于出现这种情况的风机应进行详细的潜在故障排查,以提前维修,降低故障风险。例如,风机 7 的当前预防性维护策略是每三个月的计划维护,由于风机 7 的 AI2 值大于 AI1 值,建议结合未来的风资源情况,风机出力安排,提前进行计划维修或安排额外的故障检查以排除隐患,提高运行可靠性。

为了进一步利用本文的研究成果,结合表 5 给出的风机 5-风机 9 在研究期间报警和维护数据。从表 5 中可以看出,风机 7 在这段时间内具有最高的报警次数和维修次数。这也从另一方面验证了上面的分析,既风机 7 在研究期间面临更高的运行风险。

## 4 结论

本文结合神经网络技术和随机过程理论分析了风机 SCADA 数据,提出了一种数据驱动的风机性能评估方法,其结果表明:

1) 神经网络技术能够有效的提取风机 SCADA 数据中包含的运行模式,对于建立风机系统级运行行为模型有效。

2) 基于状态转移路径的风机异常状态指标在故障数据稀少,无法确定判据阈值的情况下,能够为异常状态检测提供依据。但是,状态转移路径中所跨阶数与异常程度的关联尚未明确,需要在后续工作中,结合更多的数据样本进一步研究。

3) 针对异常状态的原因分析和提出的维修建议,应考虑实际的运行环境,并结合运行记录给出合理结论。

## 参考文献

- [1] Martin J, Ramsey D. The Economics of Wind Energy[J]. Journal of Applied Corporate Finance, 2009, 21(2):100-109.
- [2] Marugán A P, Márquez F P G, Pérez J M P. Optimal Maintenance Management of Offshore Wind Farms[J]. Energies, 2016, 9(1):46.
- [3] 刘璐洁,符杨,马世伟,赵华. 基于运行状态监测与预测的海上风机维护策略[J]. 电网技术,2015,11: 3292-3297.  
LIU Lujie, FU Yang, MA Shiwei, ZHAO Hua. Maintenance Strategy for Offshore Wind Turbine Based on Condition Monitoring and Prediction[J]. Power System Technology, 2015,11:3292-3297.
- [4] Nie M, Wang L. Review of Condition Monitoring and Fault Diagnosis Technologies for Wind Turbine Gearbox[J]. Procedia CIRP, 2013, 11:287-290.
- [5] Nie M, Wang L. Review of Condition Monitoring and Fault Diagnosis Technologies for Wind Turbine Gearbox[J]. Procedia CIRP, 2013, 11:287-290.
- [6] Sinha Y, Steel J A. A progressive study into offshore wind farm maintenance optimization using risk based failure analysis[J]. Renewable & Sustainable Energy Reviews, 2015, 42:735-742.
- [7] Y. Sinha and John A. Steel, Failure prognostic schemes and database design of a software tool for efficient management of wind turbine maintenance[J]. Wind Engineering, 39:453-478, 2015
- [8] Yan Y. Nacelle orientation based health indicator for wind turbines[J] Prognostics and Health Management. IEEE, 2015:1-7.
- [9] Castellani F, Astolfi D, Sdringola P, et al. Analyzing wind turbine directional behavior: SCADA data mining techniques for efficiency and power assessment[J]. Applied Energy, 2015.
- [10] Pelletier F, Masson C, Tahan A. Wind turbine power curve modelling using artificial neural network[J]. Renewable Energy, 2016, 89:207-214.
- [11] Rubio J D J. Analytic neural network model of a wind turbine[J]. Soft Computing, 2015, 19(12):3455-3463.
- [12] De A V R J, Sanz-Bobi M A, Kato S. Wind turbine condition assessment based on changes observed in its power curve[C]. International Conference on Renewable Energy Research and Applications. 2013:31-36.
- [13] Vieira R J D A, Sanz-Bobi M A. Failure Risk Indicators for a Maintenance Model Based on Observable Life of Industrial Components With an Application to Wind Turbines[J]. IEEE Transactions on Reliability, 2013, 62(6):569-582.
- [14] 刘绚,文俊,刘天琪. 基于自组织神经网络的模糊聚类同调机群识别[J]. 电网技术,2010,07: 98-102.  
LIU Xuan, WEN Jun, LIU Tian-qi, A Fuzzy Clustering Method to Recognize Coherent Generator Groups Based on Self-Organizing Neural Network[J]. Power System Technology, 2010,07:98-102.
- [15] 李驰,刘纯,黄越辉,王伟胜. 基于波动特性的风电出力时间序列建模方法研究[J]. 电网技术,2015,01: 208-214.  
LI Chi, LIU Chun, HUANG Yuehui, WANG Weisheng Study on the Modeling Method of Wind Power Time Series Based on Fluctuation Characteristics[J]. Power System Technology, 2015,01:208-214.
- [16] Tu J V. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes[J]. Journal of Clinical Epidemiology, 1996, 49(11):1225-1231.
- [17] G. Wilson. Quantifying the relationship between wind turbine component failure rates and wind speed.[2015].<http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.665229>
- [18] K Pearson, F Galton Pearson product-moment correlation coefficient. [2016.11.20].[https://en.wikipedia.org/wiki/Pearson\\_product\\_moment\\_correlation\\_coefficient](https://en.wikipedia.org/wiki/Pearson_product_moment_correlation_coefficient)
- [19] Ultsch A, Siemon H P. Kohonen Self Organizing Feature Maps for Exploratory Data Analysis[C]. Proceedings international neural Network Conference, 1990.
- [20] Sun P, Li J, Wang C, et al. A generalized model for wind turbine anomaly identification based on SCADA data[J]. Applied Energy, 2016, 168:550-567.
- [21] Kling W L. Impacts of Wind Power on Power System Stability[J]. Wind Power in Power Systems, 2012.
- [22] Du, Mian, Shichong Ma, Qing He. A SCADA data based anomaly detection method for wind turbines[C]. 2016 China International Conference on Electricity Distribution (CICED). IEEE, 2016.
- [23] M. Du, S.C. Ma, Q. He, et.al. A SOM based Anomaly Detection Method for Wind Turbines Health Management through SCADA Data[J].International Journal of Prognostics and Health Management,2016, 29:1-13.

收稿日期: 2016-12-20。

作者简介:

杜勉(1989), 男, 博士研究生, 研究方向为数据驱动的风电机故障检测运行维护, E-mail: dm13@mails.tsinghua.edu.cn;

郭剑波(1960), 男, 教授级高级工程师, 从事电力系统规划、运行和电网可靠性等方面的研究;



杜勉