

DOI: 10.19595/j.cnki.1000-6753.tces.161746

# 基于深度自编码网络模型的风电机组 齿轮箱故障检测

刘辉海<sup>1</sup> 赵星宇<sup>2</sup> 赵洪山<sup>1</sup> 宋 鹏<sup>3</sup> 邓 春<sup>3</sup>

(1. 华北电力大学电气与电子工程学院 保定 071003

2. 中国科学院大学物理科学学院 北京 110116

3. 国网冀北电力有限公司电力科学研究院 北京 100045)

**摘要** 为了实现风机齿轮箱的故障检测分析,提出一种基于风电机组齿轮箱的数据采集与监视控制(SCADA)数据和振动信号的深度自编码网络模型。该模型作为一种典型的深度学习方法,通过逐层智能学习初始样本特征,可以获取数据蕴含的规则与分布特征形成更加抽象的高层表示。首先,利用限制性玻尔兹曼机对网络参数进行预训练和反向传播算法对参数进行调优,建立深度自编码网络模型。然后,通过对齿轮箱的状态变量进行编码和解码,计算重构误差并将其作为齿轮箱的状态检测量。为了有效检测重构误差的趋势变化,选用自适应阈值作为风机齿轮箱故障检测的决策准则。最后,利用对齿轮箱故障前、后记录的数据进行仿真分析,结果验证了深度自编码网络学习方法对齿轮箱故障检测的有效性。

**关键词:** 风电机组 齿轮箱 故障检测 深度自编码网络 自适应阈值

**中图分类号:** TM614

## Fault Detection of Wind Turbine Gearbox Based on Deep Autoencoder Network

Liu Huihai<sup>1</sup> Zhao Xingyu<sup>2</sup> Zhao Hongshan<sup>1</sup> Song Peng<sup>3</sup> Deng Chun<sup>3</sup>

(1. School of Electrical and Electronic Engineering North China Electric Power University  
Baoding 071003 China

2. School of Physics Science University of Chinese Academy of Sciences Beijing 110116 China

3. State Grid Jibei Electric Power Research Institute Beijing 100045 China)

**Abstract** In order to achieve the fault detection of wind turbine gearbox, a deep autoencoder network model from deep learning method based on supervisory control and data acquisition (SCADA) data and vibration signals of wind turbine gearbox is proposed in this paper. The deep autoencoder network, as one of the typical deep learning methods, can obtain the underlying rules and distribution characteristics of the data through learning features of original sample by layer-wise intelligent learning to form a more abstract and high-level representation. Firstly, restricted boltzmann machine was used to pre-train parameters and the back-propagation algorithm was used to optimize these parameters to build the deep autoencoder model in this paper. Then through encoding and decoding condition variables of gearbox, reconstruction error was computed as the gearbox condition monitoring variable. In order to monitor the trend change of reconstruction error effectively,

the adaptive threshold was chosen as the decision criterion of gearbox fault. Finally, by utilizing the record data before and after fault to simulation, results showed the validity of deep autoencoder model on gearbox fault detection.

**Keywords:** Wind turbine, gearbox, fault detection, deep autoencoder, adaptive threshold

## 0 引言

风电作为一种清洁、高效的新型能源,随着其装机容量的持续增长,风机的故障问题也变得更加突出。齿轮箱做为风电机组的关键部件,由于承受复杂的载荷及其特殊的工作环境,造成风电机组停机时间最长<sup>[1-3]</sup>。因此,利用数据采集与监视控制(Supervisory Control and Data Acquisition, SCADA)数据或状态监测信息分析研究齿轮箱故障检测,能够有效降低齿轮箱的维修成本,提高经济效益。

通常风电机组齿轮箱处于早期故障阶段时,由于故障产生的冲击成分比较微弱,并且受到环境噪声的干扰,使得齿轮箱早期故障特征难以提取。目前故障检测方法一般分为基于数据的研究方法和基于模型的分析方法。基于模型的分析方法通常是建立部件的物理模型,根据模型参数的变化实现故障的检测。文献[4]通过构建齿轮箱模型,分析其故障物理机制,从而提取不同故障状况下的故障特征。

基于数据的分析方法又可以分为基于状态监测数据和基于 SCADA 运行数据两种方法。齿轮箱状态监测信息一般用于监测设备的健康状况(如振动信号和油分析),利用状态监测信息可以实现对故障特征准确提取。文献[5]运用多尺度线性调频小波分解法研究了齿轮箱变工况条件下的振动信号,实现了故障诊断。为了分析振动信号的频域特征,文献[6]将非平稳的时域信号转换为平稳或准平稳的频域信号,并采用 K 邻近度异常检测方法,提取了齿轮箱的故障信息。文献[7, 8]运用 EMD 方法分解齿轮箱振动信号,然后对其模态函数建模分析,从而判断出齿轮箱轴承的性能退化程度。文献[9]利用小波包分析提取信号特征向量,从而提高信号的分辨效果,然后对特征向量进行降维处理,实现对故障诊断时间的提高,但该方法易受强噪声的影响。文献[10, 11]通过对齿轮箱油液采样分析,根据油液粒子成分对比反映齿轮箱的故障。SCADA 信息一般指的是设备的运行数据,可利用这些数据的趋势变化来实现故障检测。文献[12, 13]针对齿轮箱油温和轴承温度的实际值与估计值,运用统计过程控制法分析温度的残差趋势,实现了齿轮箱异常状态的检测。文

献[14, 15]通过采用内存批处理技术和 Storm 实时流数据处理方法构建风电机组在线预警和故障诊断模型,能够实现对齿轮箱故障的判定。风机在运行过程中产生的 SCADA 数据蕴含着特定的规则与结构,这使得数据挖掘方法在齿轮箱的故障研究中得到应用。文献[16]分析了风机 SCADA 数据变量间的相关性,从而实现对风机不同运行状况下的健康状态的定量评估。

风电机组 SCADA 数据与状态监测数据具有变量多、数据量大、类型复杂的特点。深度学习方法能够深度学习数据内在的结构特征,并将学习到的特征信息融入模型的建立过程中,从而减少了人为设计特征的不充分性和传统特征提取所带来的复杂性。深度学习算法与传统故障诊断方法相比,克服了传统方法对诊断经验的依赖性和大数据下模型诊断能力与泛化能力的不足<sup>[17, 18]</sup>。本文基于风电机组 SCADA 数据和振动信号,提出深度自编码网络(Deep Autoencoder, DA)模型,对齿轮箱的状态进行整体分析,深度学习数据规则,挖掘其蕴含的分布式特征,从而提取齿轮箱状态检测量,实现故障检测。

DA 网络模型的分析过程包括:首先采集正常状态下风机齿轮箱的 SCADA 数据和振动数据创建训练数据集;基于训练数据集建立 DA 模型,然后运用建立的 DA 模型计算风机齿轮箱的重构误差  $R_e$  序列;根据将要发生故障的齿轮箱的  $R_e$  的趋势会发生变化,运用自适应阈值法检测  $R_e$  的趋势,在风电机组齿轮箱发生故障但未失效之前,检测出即将发生的故障,实现齿轮箱的故障预测。

## 1 DA 模型

DA 模型是一种由多个限制玻尔兹曼机(Restricted Boltzmann Machine, RBM)连接构成的深度学习网络,其底层表示原始数据的浅层特征,高层表示数据类别或属性。该模型从低层到高层逐层抽象,因此可以深度挖掘和识别数据内部的本质特征<sup>[19, 20]</sup>。每个 RBM 的输出作为更高一级 RBM 的输入,实现对学习结果的逐层传递,最终在高层获取比低层更具特征识别能力且更抽象的输出编码

$y$ , 该过程称为编码。然后通过 RBM 按照编码的方式反向进行解码, 由原来的高层输入  $y$  得到重构的  $\hat{x}$ , 编码和解码过程组成了 DA 模型的逐层学习过程, 该过程如图 1 所示。DA 模型对样本数据的学习过程包括由低层到高层的 RBM 预训练和由高层到低层的参数优化两部分。

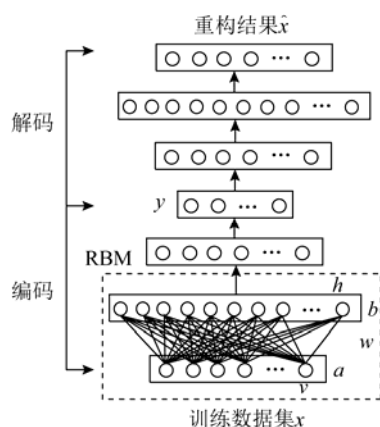


图 1 DA 模型的逐层学习过程

Fig.1 The Layer-wise learning process of DA model

### 1.1 预训练

预训练的实质是采用逐层无监督特征优化算法对网络参数进行初始化。首先需要初始化的参数为层与层之间的连接权值及各层的偏置值。RBM 的结构如图 1 中所示。RBM 包含一个可见层  $v$  和一个隐含层  $h$ , 层间各单元没有连接, 层与层之间的单元全连接。

为了能够尽可能地描述风机齿轮箱的运行状态, 选择齿轮箱的 SCADA 数据与振动信号作为模型可见层的输入。齿轮箱振动传感器的采样频率为 25.6 kHz, 对振动信号每秒求取一次峭度特征值, 形成峭度值序列。齿轮箱状态变量见表 1。

表 1 风机齿轮箱状态参数描述

Tab.1 Description of wind turbine gearbox condition variables

变量名称	变量符号
齿轮箱油温度/℃	$T_o$
齿轮箱轴承温度/℃	$T_g$
第一级垂直振动信号峭度	$V_1$
第二级垂直振动信号峭度	$V_2$
箱平行轴水平振动信号峭度	$V_3$
高速轴垂直振动信号峭度	$V_4$
高速轴轴向振动信号峭度	$V_5$

在训练过程中, 这些变量选为风电机组正常运行状态下一段时间内的监测数据, 而且均为无标签数据, 将其作为预训练样本。根据表 1, DA 网络模型的输入  $x$  表示为

$$x = [T_o, T_g, V_1, V_2, V_3, V_4, V_5]$$

考虑到风电机组齿轮箱不同类型变量的数值差异性较大, 为了减小数据的差值, 降低计算误差, 并且保证原始数据结构相对不变, 采用式 (1) 对各变量进行归一化处理。

$$\bar{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

式中,  $\bar{x}_i$  为归一化后的各变量的值;  $x_i$  为各变量原始数据;  $x_{\min}$  为训练集中该类变量的最小值;  $x_{\max}$  为训练集中该类变量的最大值。

RBM 网络作为一种无监督学习的方法, 其目的是尽可能地表达输入数据的规则和特征。为此引入能量模型<sup>[21]</sup>, RBM 作为一个系统所具备的能量定义如式 (2) 所示。

$$E_{\theta}(v, h) = -\sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i w_{ij} h_j \quad (2)$$

式中,  $\theta$  为 RBM 的参数,  $\theta = \{w, b, a\}$ ;  $w_{ij}$  为可见层单元  $i$  与隐含层单元  $j$  之间的权值;  $a_i$  为可见层单元  $i$  的偏置;  $b_j$  为隐含层单元  $j$  的偏置;  $h_j$  为第  $j$  个隐含层单元的状态;  $v_i$  为第  $i$  个可见单元的状态;  $n$  为可见层单元数量;  $m$  为隐含层单元数量。基于该能量函数, 可以得到  $(v, h)$  的联合概率分布为

$$P_{\theta}(v, h) = \frac{e^{-E_{\theta}(v, h)}}{Z(\theta)} \quad (3)$$

式中,  $Z(\theta)$  为归一化因子,  $Z(\theta) = \sum_{v, h} e^{-E_{\theta}(v, h)}$ 。通过对可见层与隐含层积分可以得到  $P_{\theta}(v)$  和  $P_{\theta}(h)$  分布, 进而得出隐含层与可见层的条件概率为

$$P_{\theta}(v) = \frac{\sum_h e^{-E_{\theta}(v, h)}}{\sum_{v, h} e^{-E_{\theta}(v, h)}} \quad (4)$$

$$P_{\theta}(h) = \frac{\sum_v e^{-E_{\theta}(v, h)}}{\sum_{v, h} e^{-E_{\theta}(v, h)}} \quad (5)$$

由于隐含层内部和可见层内部不存在连接, 由式 (4)、式 (5) 可进一步得到激活函数为

$$P_{\theta}(h_j = 1 | v) = \frac{1}{1 + \exp(-a_j - \sum_i v_i w_{ij})} \quad (6)$$

$$P_{\theta}(v_i = 1 | h) = \frac{1}{1 + \exp(-b_i - \sum_j h_j w_{ij})} \quad (7)$$

RBM 的学习结果就是求出参数  $\theta$  的值,以最大可能拟合训练数据。用梯度上升法构造可见层  $v$  的似然函数如式 (8) 所示,达到求出参数  $\theta$  的目的。

$$L(\theta; v) = \prod_v P_\theta(v) = \prod_v \frac{\sum_h e^{-E_\theta(v, h)}}{\sum_{v, h} e^{-E_\theta(v, h)}} \quad (8)$$

然后,对式 (8) 取对数,求导数

$$\begin{aligned} \frac{\partial \ln(L(\theta; v))}{\partial \theta} &= \frac{\partial \sum \ln P_\theta(v)}{\partial \theta} \\ &= \sum \left\{ E_{P_\theta(h|v)} \left[ -\frac{\partial E_\theta(v, h)}{\partial \theta} \right] - E_{P_\theta(v, h)} \left[ -\frac{\partial E_\theta(v, h)}{\partial \theta} \right] \right\} \end{aligned} \quad (9)$$

式中,  $E_{P_\theta(h|v)}[-\partial E_\theta(v, h)/\partial \theta]$  表示函数  $-\partial E_\theta(v, h)/\partial \theta$  在概率  $P(h|v)$  下的期望。在  $v$  描述似然函数的原始数据的可视状态的情况下,  $E_{P_\theta(h|v)}[-\partial E_\theta(v, h)/\partial \theta]$  可以表示原始数据的能量函数的偏导数在其自身分布下的期望值;  $E_{P_\theta(v, h)}[-\partial E_\theta(v, h)/\partial \theta]$  为函数  $-\partial E_\theta(v, h)/\partial \theta$  在概率  $P(v, h)$  下的期望,理解为能量函数的偏导在 RBM 网络表示样本分布下的期望值<sup>[22]</sup>。

为此,利用式 (8) 分别对  $w$ 、 $a$ 、 $b$  求偏导,运用对比差异度 (Contrastive Divergence) 快速算法实现  $P(v, h)$  的有效估计<sup>[23]</sup>。然后根据式 (7) 计算可视单元的条件概率,再次用 Gibbs 抽样<sup>[24]</sup>确定可见层状态,这实现了对可见层单元的一个重构。因此,最大似然化式 (8),使用快速梯度上升法确定参数  $\theta = \{w, b, a\}$  的更新准则为  $\Delta\theta = \{\Delta w, \Delta b, \Delta a\}$ ,即

$$\Delta w_{ij} = \varepsilon \left\{ \langle v_i h_j \rangle_{P_\theta(h|v)} - \langle v_i h_j \rangle_{\text{recon}} \right\} \quad (10)$$

$$\Delta b_i = \varepsilon \left\{ \langle v_i \rangle_{P_\theta(h|v)} - \langle v_i \rangle_{\text{recon}} \right\} \quad (11)$$

$$\Delta a_j = \varepsilon \left\{ \langle h_j \rangle_{P_\theta(h|v)} - \langle h_j \rangle_{\text{recon}} \right\} \quad (12)$$

式中,  $\varepsilon$  为学习率;  $\langle \cdot \rangle_{P_\theta(h|v)}$  为偏导数在  $P_\theta(h|v)$  分布下的期望;  $\langle \cdot \rangle_{\text{recon}}$  为偏导数在重构模型分布下的期望。

## 1.2 调优

在 DA 模型的预训练结束后,每层 RBM 获得一个初始参数,这构成了齿轮箱 DA 模型的初步结构。但由于 DA 模型的预训练过程是对样本数据集的无监督学习,因此预训练得到的参数可以作为 DA 模型的监督学习的先验值,模型的训练结果可以通过运用标签数据集进行监督学习实现优化,使得网络的识别性能更好。

调优训练采用风机齿轮箱长期正常运行状态下的标签数据集实现有监督学习,运用 BP 算法对

网络参数进行微调,最终使网络模型的识别性能达到整体最优。该过程训练得到的参数比单纯 BP 算法训练效果好,训练速度更快,而且收敛所需要的时间明显缩短<sup>[25]</sup>。

## 1.3 重构误差

在训练过程中,DA 模型用到了隐含层的激活信息,这种激活信息正是 DA 模型运用 RBM 训练与调优过程中学习到的输入数据的特征,以可见层与隐含层的权值和偏置值来表示。另外,DA 模型在训练过程中能够使信息损失最小,而且能够准确保留抽象和深层的特征信息。这是选择 DA 模型提取齿轮箱状态深层特征,分析其运行状态的原因。

DA 模型利用齿轮箱 7 个变量作为输入进行多层深度学习,训练过程中对原始数据进行编码和解码,并以非监督方式学习数据的固有结构和特征。在正常状态下,齿轮箱数据间存在的特征规则保持相对稳定,当齿轮箱故障时这种规则发生改变,反映重构误差  $R_e$  的趋势发生变化。在测试中,利用建立的 DA 模型计算新数据集的重构误差  $R_e$ ,将其作为齿轮箱的状态趋势的检测量。由重构值  $\hat{x}$  与原始输入  $x$  计算得到  $R_e$ ,如式 (13) 所示。

$$R_e = \|\hat{x} - x\|^2 \quad (13)$$

## 2 自适应阈值设定

通过分析  $R_e$  变化趋势或突变程度,实现对齿轮箱的故障检测。考虑到风机齿轮箱运行状态的波动性,并且是一个非平稳过程,计算得到的  $R_e$  序列仍然具有非平稳性,对其设定恒定阈值可能会出现误报警。因此引入  $R_e$  的自适应阈值作为判断故障预警的决策条件,其原理如图 2 所示。结合自适应阈值的故障检测方法将会减少干扰对风机齿轮箱故障检测的影响。假设故障发生在  $t_B$  时刻,若设定恒定阈值,将会在  $t_A$  时刻产生误报警,而且无法检测到发生在  $t_B$  时刻的故障;若采用随  $R_e$  变化的自适应阈值,在避免误报警的情况下能够检测到  $t_B$  时刻发生的故障。

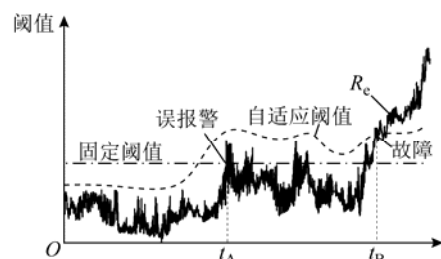


图 2 自适应阈值原理

Fig.2 Principle of adaptive threshold



将统计学中参数置信区间的思想应用于自适应阈值的设计,  $R_e$  的均值和方差求取方法分别为

$$\mu(R_{ej}, t_k) = \frac{1}{n} \sum_{i=1}^n r_i(t_k) \Big|_{R_{ej}} \quad (14)$$

$$\sigma^2(R_{ej}, t_k) = \frac{1}{n-1} \sum_{i=1}^n [r_i(t_k) - \mu(R_{ej}, t_k)]^2 \Big|_{R_{ej}} \quad (15)$$

式中,  $R_{ej}$  为不同时刻对应的  $R_e$ 。

置信度为  $(1-\alpha)$  的均值的置信区间可表示为

$$P\{\bar{\mu} - z\alpha < \mu < \bar{\mu} + z\alpha\} = 1 - \alpha \quad (16)$$

式中,  $\alpha$  为置信水平;  $z$  为与置信水平相关的系数。在实际应用中, 置信度  $(1-\alpha)$  通常选为 95%~99%。本文取置信度为 96%, 则  $z$  为 2.06, 由式 (16) 求得阈值为

$$J_{th} = \mu(R_{ej}, t_k) \pm 2.06\sigma^2(R_{ej}, t_k) \quad (17)$$

从而得到风机齿轮箱故障检测的决策准则为

$$\begin{cases} J > J_{th} & \text{故障预警} \\ J \leq J_{th} & \text{正常} \end{cases}$$

### 3 齿轮箱的 DA 模型分析步骤

正常状态下, 该  $R_e$  维持在其自适应阈值范围内变化, 当  $R_e$  变化越过阈值并保持在阈值之上, 则可以判定齿轮箱发生了故障。基于 DA 模型的风机齿轮箱故障检测应用步骤为:

1) 选取齿轮箱正常状态下的监测数据并建立变量数据集, 对数据进行归一化处理后按一定比例分为训练集和测试集。

2) 对 DA 模型参数初始化, 采用训练集中的无标签样本, 通过对比散度算法对底部 RBM 进行预训练。

3) 选用齿轮箱训练数据集中的标签样本, 运用 BP 算法对预训练之后的网络参数进行调优。

4) 保存网络参数, 建立关于齿轮箱的 DA 网络模型, 运用测试数据集计算  $R_e$ , 并进行测试。

5) 引入  $R_e$  的自适应阈值检测其趋势变化, 当  $R_e$  超过阈值, 判定齿轮箱发生故障。

## 4 算例分析

### 4.1 数据分析与参数设定

本文选取某风场 A13 号机组和 A22 号机组齿轮箱在 2014 年 1 月 4 日~9 月 14 日正常状态下的 SCADA 数据和振动监测数据建立齿轮箱的 DA 模型, 并进行参数训练和模型测试。通过异常数据处

理之后, 训练数据集的数据量分别为  $984\,000 \times 7$  和  $823\,000 \times 7$ , 总时间为 3 个月。其中 A22 号风机训练数据分为 121 批次, A13 号风机训练数据分为 101 批次, 每批次的数据量均为  $8\,125 \times 7$ , 时间为一天。

DA 模型设定 4 个隐含层, 隐含层的单元数分别为 1 000、500、250、50。数据集按 9:1 分为训练集和测试集。在训练过程中, 用训练集来学习每个隐含层的权值和偏置值, 然后利用 DA 模型计算测试集的  $R_e$ 。DA 模型的参数  $w$ 、 $a$ 、 $b$  初始化为服从高斯分布的随机较小值, 初始学习率设为 0.1, 网络更新速率设为 0.001。

### 4.2 DA 模型的重构

在预训练完成之后, 参数调优过程可以通过齿轮箱  $R_e$  的数值反映特征提取的质量和调优的性能。在 DA 模型计算重构值需要隐含层的权值与偏置值做为特征传递。因此, 隐含层的参数包含了输入数据的特征集。随着调优过程中迭代次数  $k$  的增加, 参数调优性能会越好, DA 模型隐含层对数据学习到的特征越充分, 反映到  $R_e$  的数值越小。图 3 为 A13 号风机齿轮箱 DA 模型选取 3 个月的数据作为训练数据集, 且预训练条件相同时随迭代周期变化的  $R_e$  比较。根据  $R_e$  越小模型越优, 综合考虑迭代周期越大运行时间越长的因素, 本文选择参数调优的迭代周期为 200。

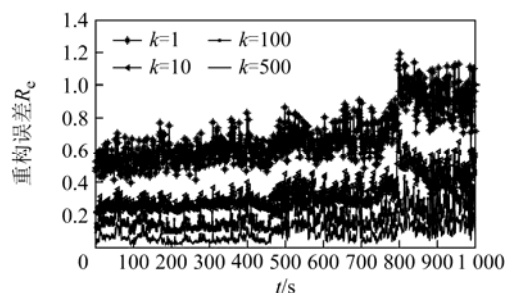
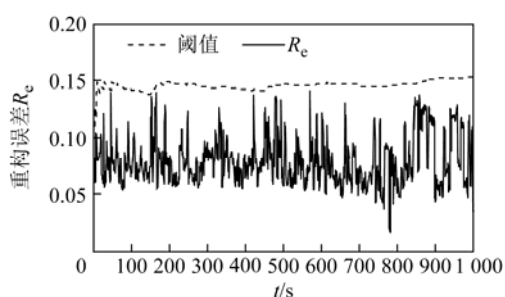
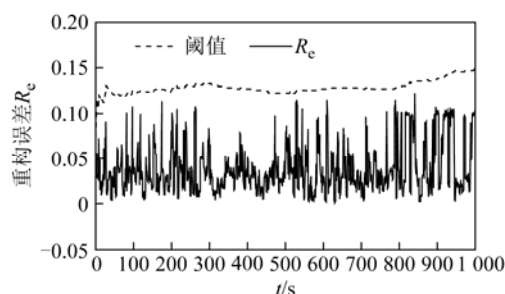


图 3 A13 号风机齿轮箱 DA 模型在预训练条件相同时不同迭代周期下的  $R_e$

Fig.3  $R_e$  of No.A13 wind turbine gearbox DA model at same training condition and different iterations

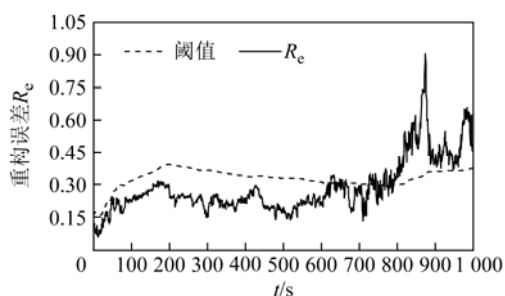
### 4.3 正常状态下的齿轮箱 DA 模型测试

运用齿轮箱的状态变量和振动信号数据集建立 DA 模型后, 对齿轮箱正常状态下的数据进行测试,  $R_e$  由式 (13) 求取。在自适应阈值条件下, 由齿轮箱的 DA 模型计算得到齿轮箱  $R_e$  的变化趋势如图 4 和图 5 所示。可以观察到正常状态下 A13 和 A22 机组齿轮箱的  $R_e$  的波动一直处于其自适应阈值范围内, 并且波动幅度较小。

图 4 A13 号风机齿轮箱正常状态下  $R_e$  自适应阈值控制Fig.4 The  $R_e$  adaptive threshold chart of wind turbine gearbox No. A13 in normal condition图 5 A22 号风机齿轮箱正常状态下  $R_e$  自适应阈值控制Fig.5 The  $R_e$  adaptive threshold chart of turbine No. A22 in normal condition

#### 4.4 故障状态下的齿轮箱 DA 模型测试

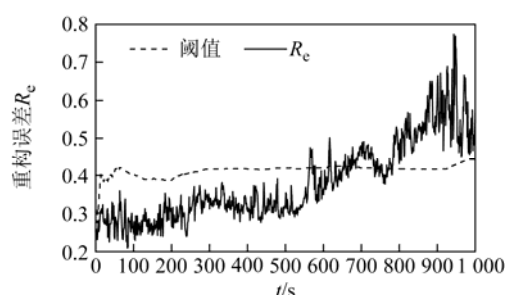
利用 A13 号机组齿轮箱故障发生前、后记录的数据,验证模型对故障的检测有效性。运用已经训练好的 DA 网络模型,计算故障状态下齿轮箱数据集的  $R_e$ ,图 6 给出了 A13 机组齿轮箱故障状态下的  $R_e$  结果。

图 6 A13 号风机齿轮箱的  $R_e$  自适应阈值控制Fig.6 The  $R_e$  adaptive threshold chart of wind turbine gearbox No. A13

仿真结果显示,在齿轮箱发生故障前, $R_e$ 一直处于其自适应阈值范围内,并且变化波动不大。然而,在 620 s 时刻齿轮箱  $R_e$  变化趋势开始上升超过阈值,并一直处于自适应阈值之上,从而判定齿轮箱出现故障,发出故障预警,实现对故障的预测。

在 A22 机组的 DA 模型建立之后,选取该机组

2014 年 10 月~12 月的监测数据,运用 DA 模型对该机组数据集进行重构,得到  $R_e$  的结果如图 7 所示。结果显示,该机组在 560 s 时刻  $R_e$  超过阈值,并处于阈值之上,从而判定齿轮箱即将发生故障。通常风电场风电机组 SCADA 系统采用监测某变量的幅值变化,如齿轮箱油温,通过温度越过告警阈值发出告警信号。利用仿真分析 A22 机组的运行数据,得到 A22 齿轮箱的故障信息,与现场故障信息比较一致。根据对比该机组齿轮箱的实际状态信息,验证了 DA 模型对齿轮箱故障检测的有效性,实现了对齿轮箱故障预测的目的。

图 7 A22 号风机齿轮箱的  $R_e$  自适应阈值控制Fig. 7 The  $R_e$  adaptive threshold chart of wind turbine gearbox No. A22

通过风机齿轮箱的仿真结果,可以看出 DA 网络模型  $R_e$  检测法能够有效检测齿轮箱的状态变化趋势,实现尽可能早地进行故障检测和预警,从而可以对将要发生故障的齿轮箱采取有效措施,降低损失。

## 5 结论

针对齿轮箱故障检测方法的研究,本文提出了基于风电机组齿轮箱 SCADA 数据与振动信号相结合的齿轮箱 DA 网络模型故障检测方法,实现了齿轮箱的故障检测。通过运用正常状态下齿轮箱的状态数据建立 DA 模型,使其深度学习齿轮箱运行数据的内在特征。选取重构误差  $R_e$  做为齿轮箱故障检测变量,并引入  $R_e$  的自适应阈值做为齿轮箱故障预警的决策准则,仿真结果验证了该方法能够基于 SCADA 数据与振动信号对齿轮箱故障实现有效检测。

## 参考文献

- [1] 赵洪山,郭伟,邵玲,等. 基于子空间方法的风机齿轮箱故障预测算法[J]. 电力自动化设备, 2015, 35(3): 27-32.  
Zhao Hongshan, Guo Wei, Shao Ling, et al. Gearbox

- fault prediction algorithm based on subspace method for wind turbine[J]. *Electric Power Automation Equipment*, 2015, 35(3): 27-32.
- [2] 杭俊, 张建忠, 程明, 等. 风力发电系统状态监测和故障诊断技术综述[J]. *电工技术学报*, 2013, 28(4): 262-271.
- Hang Jun, Zhang Jianzhong, Cheng Ming, et al. An overview of condition monitoring and fault diagnostic for wind energy conversion system[J]. *Transactions of China Electrotechnical Society*, 2013, 28(4): 261-271.
- [3] 杨明, 董传洋, 徐殿国. 基于电机驱动系统的齿轮故障诊断方法研究综述[J]. *电工技术学报*, 2016, 31(4): 58-63.
- Yang Ming, Dong Chuanyang, Xu Dianguo. Review of gear fault diagnosis methods based on motor drive system[J]. *Transactions of China Electrotechnical Society*, 2016, 31(4): 58-63.
- [4] 孙娟. 基于模型的风电机齿轮箱故障检测[D]. 南京: 南京理工大学, 2015.
- [5] 冯志鹏, 范寅夕, Liang Ming, 等. 行星齿轮箱故障诊断的非平稳振动信号分析方法[J]. *中国电机工程学报*, 2013, 33(17): 105-110.
- Feng Zhipeng, Fan Yinxi, Liang Ming, et al. A nonstationary vibration signal analysis method for fault diagnosis of planetary gearboxes[J]. *Proceedings of the CSEE*, 2013, 33(17): 105-110.
- [6] 顾煜炯, 宋磊, 徐天金, 等. 变工况条件下的风电机组齿轮箱故障预警方法[J]. *中国机械工程*, 2014, 25(10): 1346-1452.
- Gu Yujiong, Song Lei, Xu Tianjin, et al. Research on wind turbine gearbox fault warning method under variable working condition[J]. *China Mechanical Engineering*, 2014, 25(10): 1346-1452.
- [7] Liu B, Riemenschneider S, Xu Y. Gearbox fault diagnosis using empirical mode decomposition and Hilbert spectrum[J]. *Mechanical Systems and Signal Processing*, 2006, 20(3): 718-734.
- [8] 杨茂, 陈郁林. 基于 EMD 分解和集对分析的风电功率实时预测[J]. *电工技术学报*, 2016, 31(21): 86-93.
- Yang Mao, Chen Yulin. Real-time prediction for wind power based on EMD and set pair analysis[J]. *Transactions of China Electrotechnical Society*, 2016, 31(21): 86-93.
- [9] 于生宝, 何建龙, 王睿家, 等. 基于小波包分析和概率神经网络的电磁法三电平变换器故障诊断方法[J]. *电工技术学报*, 2016, 31(17): 102-112.
- Yu Shengbao, He Jianlong, Wang Ruijia, et al. Fault diagnosis of electromagnetic three-level inverter based on wavelet packet analysis and probabilistic neural networks[J]. *Transactions of China Electrotechnical Society*, 2016, 31(17): 102-112.
- [10] 陈学峰, 梁培钧. 油液分析技术在齿轮箱故障诊断中的应用[J]. *机械工程与自动化*, 2011(1): 114-115.
- Chen Xuefeng, Liang Peijun. Application of oil analysis technology in gearbox fault diagnosis[J]. *Mechanical Engineering & Automation*, 2011(1): 114-115.
- [11] 杨明, 董传洋, 徐殿国. 基于电机驱动系统的齿轮故障诊断方法综述[J]. *电工技术学报*, 2016, 31(4): 132-140.
- Yang Ming, Dong Chuanyang, Xu Dianguo. Review of gear fault diagnosis methods based on motor drive system[J]. *Transactions of China Electrotechnical Society*, 2016, 31(4): 132-140.
- [12] 赵洪山, 胡庆春, 李志为. 基于统计过程控制的风机齿轮箱故障预测[J]. *电力系统保护与控制*, 2012, 40(13): 67-73.
- Zhao Hongshan, Hu Qingchun, Li Zhiwei. Failure prediction of wind turbine gearbox based on statistical process control[J]. *Power System Protection and Control*, 2012, 40(13): 67-73.
- [13] 郭鹏, David Infield, 杨锡运. 风电机组齿轮箱温度趋势状态监测及分析方法[J]. *中国电机工程学报*, 2011, 31(32): 129-136.
- Guo Peng, David Infield, Yang Xiyun. Wind turbine gearbox condition monitoring using temperature trend analysis[J]. *Proceedings of the CSEE*, 2011, 31(32): 129-136.
- [14] 周国亮, 朱永利, 王桂兰, 等. 实时大数据处理技术在状态监测领域中的应用[J]. *电工技术学报*, 2014, 29(1): 432-437.
- Zhou Guoliang, Zhu Yongli, Wang Guilan, et al. Real-time big data processing technology application in the field of state monitoring[J]. *Transactions of China Electrotechnical Society*, 2014, 29(1): 432-437.
- [15] 张少敏, 毛东, 王保义. 大数据处理技术在风电机组齿轮箱故障诊断与预警中的应用[J]. *电力系统*

- 自动化, 2016, 40(14): 129-134.
- Zhang Shaomin, Mao Dong, Wang Baoyi. Application of big data processing technology in fault diagnosis and early warning of wind turbine gearbox[J]. Automation of Electric Power Systems, 2016, 40(14): 129-134.
- [16] Yang Wenxian, Court R, Jiang Jiesheng. Wind turbine condition monitoring by the approach of SCADA data analysis[J]. Renewable Energy, 2013, 53(9): 365-376.
- [17] 赵光权, 葛强强, 刘小勇, 等. 基于 DBN 的故障特征提取及诊断方法研究[J]. 仪器仪表学报, 2016, 39(7): 1946-1953.
- Zhao Guangquan, Ge Qiangqiang, Liu Xiaoyong, et al. Fault feature extraction and diagnosis method based on deep belief network[J]. Chinese Journal of Scientific Instrument, 2016, 39(7): 1946-1953.
- [18] 雷亚国, 贾峰, 周昕, 等. 基于深度学习理论的机械装备大数据健康监测方法[J]. 机械工程学报, 2015, 51(21): 49-56.
- Lei Yaguo, Jia Feng, Zhou Xin, et al. A deep learning-based method for machinery health monitoring[J]. Journal of Mechanical Engineering, 2015, 51(21): 49-56.
- [19] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [20] Chen Yushi, Lin Zhouhan, Zhao Xing, et al. Deep learning-based classification of hyperspectral data[J]. IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing, 2014, 7(6): 2094-2107.
- [21] Hinton G, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [22] Salakhutdinov R, Murray I. On the quantitative analysis of deep belief networks[C]//Proceedings of the 25th International Conference on Machine Learning, New York, USA, 2008: 872-879.
- [23] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(1): 504-507.
- [24] Bengio Y, Lamblin P, Popovici D, et al. Greedy layer-wise training of deep network[J]. Advances in Neural Information Processing Systems, 2007, 19: 153-160.
- [25] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6008): 533-536.
- 
- 作者简介**
- 刘辉海** 男, 1991 年生, 硕士研究生, 研究方向为风电机组传动链系统故障检测。  
E-mail: hhliu3@126.com (通信作者)
- 赵星宇** 男, 1995 年生, 本科生, 研究方向为大数据处理与算法。  
E-mail: asteroid@126.com
- (编辑 于玲玲)