

DOI:10.20033/j.1003-7241.(2026)01-0049-06

面向风力发电过程的多尺度无监督状态监测方法

韩涛, 姚维

(浙江大学电气工程学院, 浙江 杭州 310027)

摘要: 针对风力发电过程中快变与缓变参数非线性耦合所带来的多变化尺度问题, 提出一种新型的神经慢特征分析模型, 并将其应用于风力发电过程的状态监测任务。模型以神经网络为载体, 通过引入特征白化和慢度约束实现对传统慢特征分析方法的非线性化, 从运行参数复杂耦合关系中分解出不同变化尺度的子成分。同时, 设计一种无监督的多尺度异常检测实验方案, 无须依赖标签信息即可完成模型训练, 自动识别风机的异常运行模式。实验结果表明, 所提出的方法显著优于传统的无监督方法, 并能够达到与典型有监督方法相似的性能。

关键词: 神经慢特征分析; 多变化尺度; 无监督学习; 状态监测; 风力发电

中图分类号: TP277; TP183; TM614

文献标志码: A

文章编号: 1003-7241(2026)01-0049-06

Multi-scale method for unsupervised anomaly detection of wind power generation process

HAN Tao, YAO Wei

(College of Electrical Engineering, Zhejiang University, Hangzhou 310027, Zhejiang, China)

Abstract: Considering the multi-scale problem caused by the nonlinear coupling of fast-varying and slowly-varying parameters in the wind power generation process, this paper proposes a novel neural slow feature analysis model and applies it to the operating status monitoring of the wind power process. The method uses neural network technology and introduces feature whitening and slowness constraints to realize the nonlinear extension of the traditional slow feature analysis method, which decompose the sub-components with different varying scales from the operating parameters with complex coupling relationships. Meanwhile, an unsupervised and multi-scale experimental scheme for anomaly detection is designed to achieve model training without relying on label information, automatically identify the abnormal operating mode. Experimental results show that the proposed method can significantly outperform traditional unsupervised methods while achieving similar performance as typical supervised methods.

Keywords: neural slow feature analysis; multi-scale variation; unsupervised learning; status monitoring; wind power

风能作为一种清洁无污染的新型能源, 在“碳达峰、碳中和”目标中发挥着重要作用^[1-2]。然而, 风电机组通常位于高原、海上等恶劣环境, 运行状态随环境复杂多变, 机组异常时有发生, 影响电网的稳定。近年来, 随着大数据技术的发展, 数据驱动的风电机组监测是行业研究的热点问题之一。利用数据采集系统(SCADA)获取风机运行过程中的环境及设备运行参数, 结合数据驱动模型分析监测机组运行状态, 保障机组安全。

现有的监测方法通常分为有监督和无监督两类。有监督方法利用风机运行的历史运行数据及对应的标签信息, 建立分类监测模型。如 XGBoost^[3], 支持向量机(SVM)^[4], k 近邻(kNN)^[5], 随机森林^[6]等经典方法及长短期记忆(LSTM)^[7], 卷积神经网络(CNN)^[8]等基于深度学习的方法。

无监督方法则是利用风机正常运行数据进行模型训练, 提取表征风机运行规律的关键特征; 对于异常运行状态下的样本, 其特征将与正常数据有显著差异, 从而实现故障的自动检测。一般情况下, 故障数据难以采集, 是因为风机一旦发生故障就需要尽快进行检修, 否则容易造成设备损坏甚至引发安全事故。因此, 与有监督方法相比, 无监督方法的建模无需故障数据, 在实际应用中不仅易于操作, 而且还避免了在故障状况下收集数据的风险。李俊卿等^[9-10]提出一种基于自编码器的风机故障检测模型, 通过编解码网络提取描述风机运行的低维度特征, 并以样本点重构误差作为检测指标。Jing等^[11]使用高斯混合模型(GMM)刻画风机运行参数的概率分布, 以样本点对应的概率密度作为故障检测依据。

收稿日期: 2024-07-09; 录用日期: 2024-11-08

基金项目: 国家自然科学基金面上项目(62173295); 浙江省产学研合作协同育人项目(2023006)

作者简介: 韩涛(1980—), 男, 硕士, 工程师, 研究方向: 机器人, 智能控制与检测。

引用本文: 韩涛, 姚维. 面向风力发电过程的多尺度无监督状态监测方法[J]. 自动化技术与应用, 2026, 45(1): 49-53, 78. (HAN Tao, YAO Wei. Multi-scale method for unsupervised anomaly detection of wind power generation process[J]. Techniques of Automation and Applications, 2026, 45(1): 49-53, 78.)

事实上,风机的运行参数与外界环境参数具有不同的变化尺度,通常情况下,风速等环境因素呈现快速变动趋势,而机组内部的温度、电流等参数变动相对缓慢,需要对这种多尺度特性进行细致分析以提高监测模型的精确程度^[12],避免缓变参数中的潜在异常被快变参数中的大范围波动所掩盖,而现有方法通常假设各运行参数尺度一致,无法合理地处理这一问题。

针对上述的风机运行的多尺度特性,提出一种神经慢特征分析方法。通过缓变约束及白化操作来提取多元运行参数中具有不同变化尺度的子成分,利用神经网络实现数据的非线性变换来刻画风机参数之间的非线性耦合特性,构建静态分布指标和动态波动指标来从不同变化尺度综合地描述风机运行规律,实现无监督故障检测任务。通过对风机运行状态比较准确的监测,降低源侧出力波动,提升风电并网的稳定性,实现高比例新能源的可调度分布式异构协同的目的。

1 慢特征分析

慢特征分析(slow feature analysis, SFA)^[13-14]是一种经典的用于多变化尺度分析的无监督统计学习方法,能够用于风机运行数据的特征提取。SFA的输入数据为风机运行过程中采集的多元时间序列,定义为 $\mathbf{x}(t) = [x_1(t), \dots, x_N(t)]^T$,其中 N 为变量数目, $x_1(t), \dots, x_N(t)$ 表示风机运行过程中的重要参数在 t 时刻上的采样值,涵盖风速、对风角、环境温度等环境参数以及风机功率、发电机转速、叶片转角、叶片变桨电机温度、机舱温度、叶片直流电流等设备参数。SFA方法从上述运行数据中提取具有不同时间变化速度的特征作为多尺度时序动态性的表征。给定 $\mathbf{x}(t)$, SFA需要求解一系列线性映射 $\mathbf{w}_1, \dots, \mathbf{w}_N$,得到输出特征 $y_1(t), \dots, y_N(t)$,并使得特征的变化速度由慢到快排列,其中 $y_j(t) = \mathbf{w}_j^T \mathbf{x}(t)$ ($j = 1, 2, \dots, N$)。SFA的优化目标

$$\begin{aligned} & \min_{y_j, j=1,2,\dots,N} \langle [y_j(t) - y_j(t-1)]^2 \rangle_t \\ & \text{s. t. } \begin{cases} \langle y_j \rangle_t = 0, j = 1, 2, \dots, N \\ \langle y_j^2 \rangle_t = 1, j = 1, 2, \dots, N \\ \forall i < j, \langle y_i y_j \rangle_t = 0 \end{cases} \quad (1) \end{aligned}$$

式中, $\langle \cdot \rangle_t$ 为时间平均算子,表示对括号中成分在各时间点上的取值作平均处理。

由式(1)可知,SFA在对输出特征的变化速度做缓慢约束的同时,还要求各特征具有零均值和单位方差,并且彼此相互独立。

为满足SFA的零均值约束,首先对输入 $\mathbf{x}(t)$ 作去均值运算,即令 $\mathbf{x}(t)$ 中各变量分别减去各自的均值。随后,将 $y_j(t) = \mathbf{w}_j^T \mathbf{x}(t)$ 代入原始优化目标中,得到等价问题

$$\begin{cases} \min_{\mathbf{w}_j, j=1,2,\dots,N} \mathbf{w}_j^T \langle \dot{\mathbf{x}}(t) \dot{\mathbf{x}}(t)^T \rangle_t \mathbf{w}_j \\ \text{s. t. } \mathbf{w}_j^T \langle \mathbf{x}(t) \mathbf{x}(t)^T \rangle_t \mathbf{w}_j = 1 \end{cases} \quad (2)$$

式中, $\dot{\mathbf{x}}(t)$ 为 $\mathbf{x}(t)$ 的时间点导数,针对离散系统,可以由一阶差分来近似,即 $\dot{\mathbf{x}}(t) = \mathbf{x}(t) - \mathbf{x}(t-1)$ 。

对于上述有约束优化问题,通常采用拉格朗日乘子法进行求解。将拉格朗日乘子记作 λ_j ($j = 1, 2, \dots, N$),则拉格朗日函数 L 计算。

$$L = \mathbf{w}_j^T \langle \dot{\mathbf{x}}(t) \dot{\mathbf{x}}(t)^T \rangle_t \mathbf{w}_j - \lambda_j (\mathbf{w}_j^T \langle \mathbf{x}(t) \mathbf{x}(t)^T \rangle_t \mathbf{w}_j - 1) \quad (3)$$

根据优化理论可知,极值点处的偏导数为0,因此

$$\frac{\partial L}{\partial \mathbf{w}_j} = 0 \Rightarrow \langle \dot{\mathbf{x}}(t) \dot{\mathbf{x}}(t)^T \rangle_t \mathbf{w}_j = \lambda_j \langle \mathbf{x}(t) \mathbf{x}(t)^T \rangle_t \mathbf{w}_j \quad (4)$$

从而,SFA模型的求解转化为一个广义特征值分解问题,所得到的特征向量即为所求线性映射 $\mathbf{w}_1, \dots, \mathbf{w}_N$ 。

将式(4)代入到式(1)和式(2)中,可以得到如下结论。

$$\begin{aligned} \langle [y_j(t) - y_j(t-1)]^2 \rangle_t &= \frac{\mathbf{w}_j^T \langle \dot{\mathbf{x}}(t) \dot{\mathbf{x}}(t)^T \rangle_t \mathbf{w}_j}{\mathbf{w}_j^T \langle \mathbf{x}(t) \mathbf{x}(t)^T \rangle_t \mathbf{w}_j} = \\ &= \frac{\lambda_j \mathbf{w}_j^T \langle \mathbf{x}(t) \mathbf{x}(t)^T \rangle_t \mathbf{w}_j}{\mathbf{w}_j^T \langle \mathbf{x}(t) \mathbf{x}(t)^T \rangle_t \mathbf{w}_j} = \lambda_j \quad (5) \end{aligned}$$

因此SFA所求得的不同线性映射对应的特征值 λ_j 即反映了相应输出特征 $y_j(t)$ 的变化快慢。将 λ_j 按升序排列,表示所对应的输出特征会随之由慢到快排列,从而反映出风机运行过程中蕴含的不同变化尺度信息。

2 神经慢特征分析模型

SFA方法能够有效地提取原始数据中具有不同变化尺度的子成分,然而其只能对数据进行线性变换,无法处理具有显著非线性特性的风机运行数据。如,风机的功率会随着外界风速的增加在一定范围内近似线性的增长,而当风机运行到满功率状态后,其功率将趋于平稳,因此风机功率与风速间整体上呈现非线性关系。神经网络由于其结构灵活、能够端到端运算等优点成为处理数据非线性的重要方法之一。在一般情况下可以使用神经网络代替线性变换以解决非线性问题,通过构建损失函数并利用梯度反向传播完成神经网络模型优化求解。然而,在使用SFA算法提取慢特征的过程中需要引入特征值分解操作,这一操作是不可微的,即梯度不存在,因此无法直接利用神经网络替换SFA中的线性变换来解决非线性问题。为此,本研究提出一种神经网络驱动的非线性SFA方法,称为神经慢特征分析(Neural SFA, NSFA),通过将SFA的缓变优化目标、零均值、单位方差和去相关约束嵌入神经网络结构,提取风机时序采样信号中的多尺度特性。同时,NSFA引入幂法^[15]完成SFA中的特征值分解运算,以保证模型可以通过梯度反向传播进行优化求解。

NSFA方法的输入数据与SFA相同,为风机运行过程中的重要参数,包含环境参数与设备参数,其涵盖4个运算步骤,具体如下。

1) 非线性特征提取

将风机运行数据 $\mathbf{x}(t) = [x_1(t), \dots, x_N(t)]^T$ 输入到多层全连接神经网络,进行非线性特征提取。将每层全连接神经网络的映射函数记为 f , 每层网络的输入 \mathbf{x}_{in} , 其输出 \mathbf{x}_{out} 的表达式为

$$\mathbf{x}_{out} = f(\mathbf{x}_{in}) = \delta(\mathbf{W}\mathbf{x}_{in} + \mathbf{b}) \quad (6)$$

式中, \mathbf{W} 和 \mathbf{b} 为可训练的权重矩阵和偏置向量; δ 为非线性激活函数, 在方案中选取的激活函数为 Relu, 其计算方式为

$$\text{Relu}(x) = \max(0, x)。$$

通过堆叠多个全连接层, 来提取深层次非线性特征 $\mathbf{h}(t)$, 即

$$\mathbf{h}(t) = f(f(\dots f(f(x(t))))))。$$

2) 特征白化

SFA 模型中利用特征一阶差分的模长来表示其变化速率, 并通过最小化该速度提取风机运行参数的动态特性, 然而在神经网络框架中无法直接对特征施加这一优化目标。具体来讲, 由于神经网络的模型拟合能力较强, 可能会得到平凡解。极端情况下, 模型可能会学习到常值特征, 从而使得特征变化速率为零以实现优化目标的最小化, 而常值特征无法携带任何能够反映风机参数变化信息。为克服上述问题, 在 NSFA 方法中设计了一种特征白化机制来满足零均值, 单位方差, 以及去相关约束, 使特征需要满足单位方差约束, 避免上述的平凡解问题, 使得模型能够在神经网络架构下提取非线性特征。

首先对步骤 1) 得到的非线性特征 $\mathbf{h}(t)$ 进行去均值操作。随后, 计算 $\mathbf{h}(t)$ 的协方差矩阵 \mathbf{C} , 再对其进行正交分解, 计算公式为

$$\mathbf{C} = \frac{1}{n} \mathbf{h}(t) \mathbf{h}^T(t) = \mathbf{U} \mathbf{D} \mathbf{U}^T \quad (7)$$

式中, n 为样本个数, 矩阵 \mathbf{U} 和 \mathbf{D} 分别为 \mathbf{C} 的特征向量及特征值构成的矩阵。

值得注意的是, 在上述正交分解过程中需要计算矩阵 \mathbf{C} 的特征值和特征向量。如前文所述, 在一般的神经网络框架中, 该操作的梯度不存在, 不可直接进行求解, 为解决该问题, 此处引入幂法进行迭代求解。

首先, 随机生成一个向量 $\mathbf{u}_1^{(0)}$ 作为矩阵 \mathbf{C} 绝对值最大的特征值 λ_1 所对应的特征向量初始值。随后, 按照如下方式迭代运算直至收敛, 得到最终的特征向量 \mathbf{u}_1 , 计算公式为

$$\mathbf{u}_1^{(i)} = \mathbf{C} \mathbf{u}_1^{(i-1)}, \quad i = 1, 2, \dots \quad (8)$$

由此, 对应的特征值 λ_1 计算为

$$\lambda_1 = \mathbf{u}_1^T \mathbf{C} \mathbf{u}_1 \quad (9)$$

此后, 将矩阵 \mathbf{C} 按如下方式更新, 并重复上述步骤依次计算出 \mathbf{C} 的其他特征值及特征向量。

$$\mathbf{C} \leftarrow \mathbf{C} - \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T \quad (10)$$

以上步骤通过多步可微的迭代操作替代了矩阵的特征值分解, 能够在神经网络框架下进行运算。

根据 \mathbf{C} 的分解结果, 可以计算白化矩阵 $\mathbf{\Omega}$, 计算公式为

$$\mathbf{\Omega} = \mathbf{D}^{-1/2} \mathbf{U}^T \quad (11)$$

最终的特征白化结果为

$$\mathbf{y}(t) = \mathbf{\Omega} \mathbf{h}(t) = \mathbf{D}^{-1/2} \mathbf{U}^T \mathbf{h}(t) \quad (12)$$

$\mathbf{y}(t)$ 的协方差矩阵 $\text{cov}[\mathbf{y}(t)]$ 计算过程为

$$\begin{aligned} \text{cov}[\mathbf{y}(t)] &= \frac{1}{n} \mathbf{D}^{-1/2} \mathbf{U}^T \mathbf{h}(t) \mathbf{h}^T(t) \mathbf{U} \mathbf{D}^{-1/2} = \\ &= \frac{1}{n} \mathbf{D}^{-1/2} \mathbf{U}^T \mathbf{U} \mathbf{D} \mathbf{U}^T \mathbf{U} \mathbf{D}^{-1/2} \end{aligned} \quad (13)$$

由于 \mathbf{U} 由协方差矩阵 \mathbf{C} 的特征向量构成, 因此有 $\mathbf{U}^T \mathbf{U} = \mathbf{I}$, 其中 \mathbf{I} 为单位阵, 从而得到

$$\text{cov}[\mathbf{y}(t)] = \frac{1}{n} \mathbf{D}^{-1/2} \mathbf{U}^T \mathbf{U} \mathbf{D} \mathbf{U}^T \mathbf{U} \mathbf{D}^{-1/2} = \frac{1}{n} \mathbf{D}^{-1/2} \mathbf{D} \mathbf{D}^{-1/2} = \mathbf{I} \quad (14)$$

$\mathbf{y}(t)$ 的协方差矩阵为单位阵, 即 $\mathbf{y}(t)$ 满足 SFA 的零均值、单位方差以及去相关约束, 从而避免了前述的平凡解问题。

3) 计算慢度损失

慢度损失由最小化特征的一阶差分的二范数来表征, 基于此, NSFA 中设计的慢度损失函数 L 为

$$L = \frac{1}{n-1} \sum_{t=2}^n \|\mathbf{y}(t) - \mathbf{y}(t-1)\|_2^2 \quad (15)$$

于是, NSFA 的优化目标为最小化慢度损失函数 L 。

4) 梯度反向传播

NSFA 通过梯度信息更新非线性特征提取中各神经网络层的参数, 实现损失函数 L 的最小。通过反向传播算法^[16] 可以计算 L 对各参数的梯度, 并使用梯度下降算法进行优化, 以便得到最终的模型参数并完成模型训练。

综上, NSFA 通过多层神经网络提取原始数据的深层次非线性特征, 再利用特征白化机制使数据满足 SFA 的三方面约束以避免平凡解, 同时引入幂法来避免梯度不存在的特征值分解运算, 最后构建慢度损失函数 L 并通过梯度反向传播来进行模型训练。通过这种方式, NSFA 在传统 SFA 模型的基础上兼顾数据的非线性特性, 更利于风机数据的多尺度特性提取与分析。

3 状态监测模型构建

基于上述 NSFA 方法, 建立相应的无监督风机运行状态监测模型, 其离线建模与在线监测框架如图 1 所示。

在离线建模阶段, 利用风机在正常工况下的历史运行数据训练 NSFA 模型, 提取表征风机运行参数的动态变化规律的特征 $\mathbf{y}(t)$ 。按照变化速度由慢到快的顺序将 $\mathbf{y}(t)$ 中各子信号进行排列, 得到 $\mathbf{y}_1(t), \dots, \mathbf{y}_N(t)$, 其中, 变化速度可用式(1)中的优化目标函数来衡量, 其取值越小说明变化越缓慢。根据各特征变化速度的快慢, 将其分为缓变成分和快变成分, 从不同的变化尺度表述风机的运行特性。设置划分阈值 K (可取作变量个数 N 的 50%), 利用变

化最慢的 K 个特征信号作为风机运行状态的慢尺度表征, 记作 $\mathbf{y}_s(t) = [y_1(t), \dots, y_K(t)]^T$, 其余特征信号则用于构成运行状态的快尺度表征, 记作 $\mathbf{y}_f(t) = [y_{K+1}(t), \dots, y_N(t)]^T$ 。参考 Shag 等^[17]从分布特性和变化特性两个角度构建风力发电过程缓变和快变两个尺度的监测统计量, 计算公式为

$$T_s^2 = \mathbf{y}_s^T(t) \mathbf{y}_s(t) \quad (16)$$

$$T_f^2 = \mathbf{y}_f^T(t) \mathbf{y}_f(t) \quad (17)$$

$$S_s = \|\mathbf{y}_s(t) - \mathbf{y}_s(t-1)\|_2^2 \quad (18)$$

$$S_f = \|\mathbf{y}_f(t) - \mathbf{y}_f(t-1)\|_2^2 \quad (19)$$

式中, T_s^2 和 T_f^2 分别为慢尺度和快尺度的分布特性统计

量, 其数学意义为特征向量到整体分布的马氏距离的平方, 其取值越小说明当前样本越贴近正常范围, 由于 NSFA 中的特征满足零均值和单位方差约束, 因此呈现为特征模的平方形式; S_s 和 S_f 分别为慢尺度和快尺度的变化特性统计量, 度量了特征的变化速度, 由于 NSFA 要求特征的变化速度尽可能慢, 因此变化特性统计量取值越小则说明当前样本的变化特性越趋近于正常模式。两类统计量从不同视角综合地描述了风机运行状态, 且细粒度地区分了不同变化尺度的运行特性, 能够显著提高监测的准确性和灵敏性。

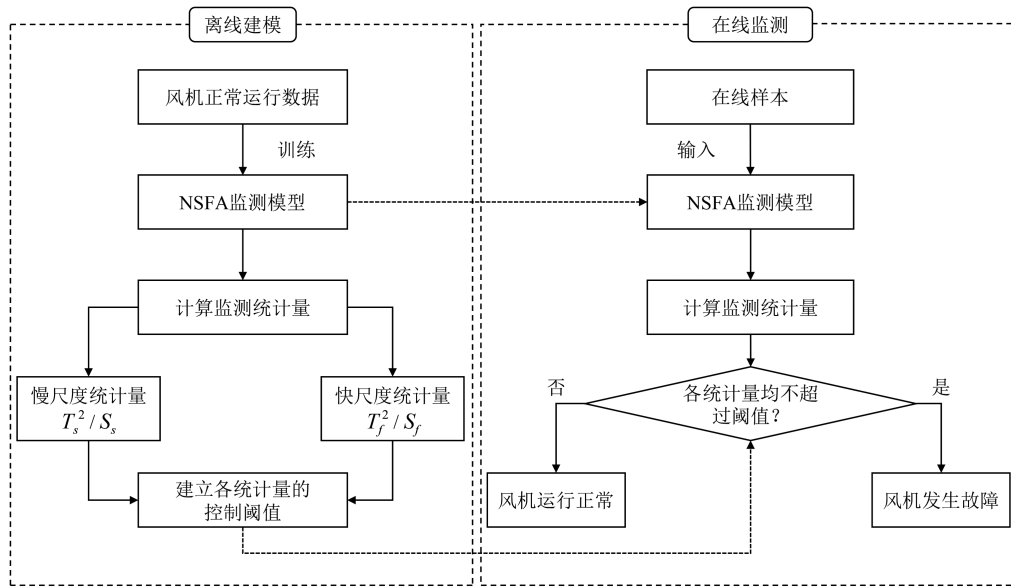


图1 基于 NSFA 的无监督风机运行状态监测模型

Fig. 1 Unsupervised wind turbine operation state monitoring model based on NSFA

为了确定风力发电过程参数的合理取值范围, 使用正常数据训练 NSFA 模型, 并获取每个正常样本点的 4 个统计量, 各统计量的控制阈值计算方式如下: 对于任一统计量 α , 其控制阈值 $C(\alpha)$ 表达式为

$$C(\alpha) = 0.95 \times \max \alpha \quad (20)$$

即取统计量在所有样本点上最大值的 0.95 倍作为控制阈值, 统计量超过该阈值则将该样本判定为异常。设置 0.95 作为系数是为了防止正常数据中存在离群点。

在线监测时, 获取风机运行时的采样点输入到离线建模时训练好的 NSFA 模型中, 获取该时刻的特征并计算上述统计量, 如果任一统计量超出了控制阈值, 则说明风机的运行发生了异常。

值得注意的是, 所提出 NSFA 模型为无监督方法, 仅需要正常数据即可完成建模。相比于同时需要充分正常数据和故障数据的有监督方法, 所提方法对数据的要求较低, 在实际应用中避免了在故障状况下采集数据的风险, 然而由于其建模过程中未对故障数据进行分析与利用, 无法如有监督方法一样明确掌握故障数据的分布或模式, 因此当数据中的故障征兆不明显时可能存在故障误报与漏

报情况。

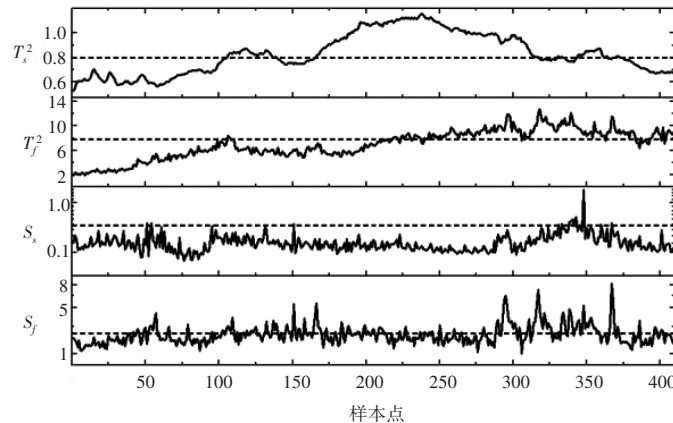
4 实验验证

为验证所提出 NSFA 模型的有效性, 选取某风电场 SCADA 的实测数据进行实验。数据中包含风速、外界温度、风向等环境参数, 以及风机叶片转角、机舱温度、电流、功率等设备参数, 采样周期为 30 s。选取 2 000 个正常工况下采集的历史数据用于模型训练, 同时用故障数据集进行测试, 故障数据集中包含 410 个样本, 其中前 50 个样本点为正常数据, 后 360 个样本中发生了叶片结冰故障。

图 2 中可视化了所提出 NSFA 模型各统计量在测试数据上的变化曲线, 分别描述风力发电过程中不同的特性。统计量 S_f 在故障发生后迅速超过阈值, 说明其成功捕捉到数据中异常的快变特性。由于风力发电并不是一个稳定的过程, 而是会产生复杂的波动, 这导致 S_f 始终在控制阈值上下波动, 在某些时刻将故障样本误判为正常。而 NSFA 具有 4 个统计量, 能够从不同角度对运行状态进行描述, 可以发现统计量 T_s^2 在第 100 个样本点附近超出控制阈值, 并能够在大多数故障样本点上保持在阈值上

方。此外,统计量 T_f^2 在第 250 个样本点附近超出阈值,并稳定在阈值以上,进一步弥补 T_s^2 统计量在最后回落到阈

值以下所造成的漏检测。因此,不同统计量对风力发电过程的多尺度描述可以显著地提高故障检测准确率。



实线为监测统计量,虚线为各统计量的控制阈值

图 2 NSFA 模型监测结果

Fig. 2 NSFA model monitoring results

为了给出定量的评价标准,设定误报率 (false alarm rate, FAR) 和检测精度 (detection accuracy, DA) 两个指标来评估所提出方法的监测性能,其中 FAR 指的是测试数据中被误判为异常的正常样本占有所有正常样本的比例,取值越低说明模型性能越强;DA 指测试数据集中被模型成功检测到的故障样本占有所有故障样本的比例,其取值越高则说明模型对故障数据更灵敏,性能越好。另外,选取自编码器、GMM、SFA、XGBoost、SVM 5 种方法作为对比,其中前 3 种方法为无监督方法,而后两种方法为有监督方法。对于两种有监督方法,额外加入 500 个结冰故障样本进行建模。各模型的监测性能指标如表 1 所示。

表 1 各模型监测性能指标

Tab. 1 Monitoring performance indicators of each model %

模型	FAR	DA
NSFA	12	94.72
自编码器	14	79.44
GMM	12	61.94
SFA	12	86.66
Xgboost	8	96.39
SVM	10	92.78

可见 3 种无监督模型的 FAR 指标大多相近,均略高于 10%,根据图 2 中 S_f 的变化曲线可以看出,其取值在故障开始之前的较短时间内就已经超出阈值,这可能并不是由于模型性能不佳而产生的误报,而是模型提前检测到了数据中潜在的异常趋势。对于 DA 指标,所提出的 NSFA 显著高于其他对比的无监督方法,这体现出其分别提取快慢变化尺度特性以及多指标监测的优势。此外,NSFA 的 DA 明显高于 SFA,这也说明 NSFA 对 SFA 的非线性化有助于模型更精确地捕捉风力发电过程中的故障特征,提高监测性能。此外,两种有监督对比方法的监测性能略好于所提出的 NSFA 方法,由于这两种方法引入了较多的故障

样本进行建模,能够直接对正常和故障运行模式进行区分,因此这一结果符合预期。然而,即便在额外引入大量故障样本的情况下,两种有监督对比方法的性能也仅仅是略好于 NSFA,这说明所提出 NSFA 即便在不使用故障数据时,也能够通过挖掘风机运行中不同变化尺度的特征来达到与有监督方法相近的监测性能,这进一步凸显了所提出方法的优势和有效性。

5 结论

本文面向风力发电过程的智能运维,提出了一种无监督学习范式下的多变化尺度、非线性特征提取与状态监测模型。针对风力发电中不同参数变化尺度不同,且参数之间存在复杂非线性耦合的问题,所提出的 NSFA 模型通过对特征的缓变约束提取出具有不同变化尺度的子成分,并采用神经网络模块和特征白化机制实现对传统 SFA 模型的非线性化。此外,所提出的监测模型针对不同的变化尺度信息,分别从分布特性与变化特性两个角度对数据特性进行精细描述,并构建四个统计量综合地评估风力发电的过程状态。在实际风机数据上的实验结果表明,所提出方法不仅无需标签信息即可检测到故障,且相比于其他的常用无监督方法具有显著的精度提升,并能够达到与典型有监督方法相近的监测性能。

参考文献

[1] 余德启. 大规模风电场集电线路路径优化与数字地图显示[J]. 风能, 2022(3):62-65.
 [2] 黄强, 郭怿, 江建华, 等. “双碳”目标下中国清洁电力发展路径[J]. 上海交通大学学报, 2021, 55(12):1499-1509.
 [3] 苏连成, 朱娇娇, 郭高鑫, 等. 基于 XGBoost 和 Wasserstein 距离的风电机组塔架振动监测研究[J]. 太阳能学报, 2023, 44(1):306-312.
 [4] 韩涛, 姚维. 基于 K 近邻与支持向量机协同训练的风机叶片结冰早期检测[J]. 实验室研究与探索, 2021, 40(9):52-56.

(下转第 78 页)