

風車の正常稼働状態に基づく教師データ選択手法の評価

安田晃久^{†1} 古澤陽子^{†1} 成末義哲^{†1} 森川博之^{†1} 飯田誠^{†1}

概要：風力発電では、風車の内部機器が故障すると停止期間が続き、大きな損害が生じてしまう。そのため、内部機器の異常を早期に検知することが重要である。これまで、筆者らは、風車の標準的な稼働監視装置である SCADA (Supervisory Control And Data Acquisition) 収集データを用いた機械学習による異常検知手法を開発しており、ここでは、モデルの異常検知性能が、過去の時系列データから選択した教師データに依存する点が課題となっている。本稿では、理想的なパワーカーブとの比較から一定期間毎の正常稼働率を算出し、この正常稼働率を基に、教師データとして適切な期間を選択するための分類手法を示し、異常検知処理における有効性を確認した。

キーワード：風車, SCADA, 機械学習, 教師データ, 正常稼働, 異常検知

Evaluation of Training Data Selection Method for Wind Turbine Degradation Assessment Based on the Normal Behavior

AKIHISA YASUDA^{†1} YOKO FURUSAWA^{†1} YOSHIAKI NARUSUE^{†1}
HIROYUKI MORIKAWA^{†1} MAKOTO IIDA^{†1}

Abstract: In wind power generation, failures of internal components of wind turbines lead to prolonged downtime, and maintenance cost will increase. It is necessary to detect incipient faults and improve availability with reducing unscheduled maintenance. We have developed an anomaly detection method by machine learning using SCADA (Supervisory Control And Data Acquisition) data which is a standard operation monitoring device for wind turbine, and anomaly detection performance is dependent on the training data selected from the past data. In this paper, we propose a training data selection method. Specifically, we calculate the normal behavior rate in comparing the ideal power curve with the normal operating rate for each fixed period, and classify it for selecting the appropriate period as the training data. The anomaly detection results showed the effectiveness of proposed method.

Keywords: Wind Turbine, SCADA, Machine Learning, Training Data, Normal Behavior, Anomaly Detection

1. はじめに

風力発電事業は世界的に堅調な成長路線を描いており、2017年の風力発電の発電出力は、再生可能エネルギーの中でも水力発電に次いで第2位の規模である[1]。今後も継続して風力発電事業を発展させるには、発電コストを下げるための施策として、故障発生時の風車ダウンタイムの短縮が有効である[2]。

ダウンタイムを短縮するには、風車の異常を早期に検知することで、交換機器をあらかじめ調達しておく必要がある[3]。風車の異常を検知するためのシステムとして CMS (Condition Monitoring System) が販売されているが、高価であり、風力発電事業活性化のためには CMS の導入が困難である事業者に向けて、初期導入コストを抑えた安価な異常検知手法が必要である。

この問題に対し、稼働監視および制御装置として、ほぼすべての風車にあらかじめ備え付けられている SCADA (Supervisory Control and Data Acquisition) 収集データを二次利用した、機械学習モデルに基づく異常検知手法が提案されており[4]-[8]、筆者らも SCADA データを二次利用し

た異常検知手法の開発に取り組んでいる[9] [10]。SCADA データを用いた異常検知のための機械学習モデルでは、教師データとして、故障時のデータではなく正常稼働データを利用することが一般的である[4]。これは、長期間のダウンタイムを要する故障は発生頻度が低く、参照可能なデータ件数が少ないからである[11]。

今後、SCADA データを用いた異常検知手法の有効性が確認された後、多様な風車に適用していく際に課題となるのは、正常稼働を示す教師データを選択することである。教師データは異常検知性能に直結する情報であるにもかかわらず、各風車は稼働年数、修理履歴が様々であり、風車毎に適切な教師データは異なる。そのため、過去の SCADA データについて、稼働状態毎に分類し、教師データとして妥当な期間を選択する手法が必要となる。

教師データ選択における参照情報として、風車メンテナンス事業者の作業日誌を利用する場合、記載内容は事業者毎に異なるため、作業者に依存してしまう点が課題である。また、選択の基準となる参照情報を用いず、SCADA データのみで教師データを自動的に選択する場合、600 s 間の荒く平均化されたデータという SCADA データの特徴を考慮すると、データに直接現れない風車の稼働状態を把握する

^{†1} 東京大学
The University of Tokyo

ことが難しく、正常稼働データを選択するのは容易でない。

従来研究では、SCADA データ項目毎に正常稼働状態を示すパラメータの範囲を設定し、風車の稼働状態を判定している[12]。しかし、メーカーや型番毎に風車の設計が異なるため、このパラメータが正常稼働状態を示す保証はない。

そこで、筆者らは、メーカーが風車の性能仕様として提供するパワーカーブに着目し、風車の稼働状態がパワーカーブに近い期間を正常稼働状態として、これを教師データ期間として選択する。ここで、本稿では、1 か月単位の SCADA データレコードのうち、正常稼働状態と判定可能なレコードが含まれる割合を正常稼働率と定義する。

パワーカーブを用いた教師データ選択において、パワーカーブで利用する入出力データは、風車の稼働状態以外の要因による影響を受けるという点が課題となる。そこで、本稿では、正常稼働率以外の情報を考慮した複合的な仕組みを構築するため、正常稼働率と共に、特異スペクトル変換 (Singular Spectrum Transformation : SST) [13]-[15]による変化点検知データと、動的時間伸縮法 (Dynamic Time Warping: DTW) [16]による他風車との比較データを用いて、時系列 SCADA データから正常稼働期間を分類し、教師データを選択する。

本稿では、風車の理想的なパワーカーブとの比較に基づく教師データ選択手法を3章に、筆者らが提案している異常検知手法を4章で示す。そして、5章では実機データと異常検知手法を用いた教師データ選択手法の評価結果を提示し、結論に繋げる。

2. 教師データ選択手法

風車メーカーは風車の設計仕様として、理想的な稼働状態を示すパワーカーブを提供している。パワーカーブとは、風速に対する発電出力を示す曲線であり、風車の入力情報である風況と、出力情報である発電出力を組み合わせた、入出力の関係情報である。理想的なパワーカーブと1 か月単位の SCADA データとを照らし合わせ、理想的な稼働状態が継続した期間を教師データとして選択する。

ここで問題となるのは、理想的なパワーカーブと風車の稼働状態との間の因果関係が強固でない点にある。つまり、パワーカーブとの比較から得られる正常稼働率は、風車の稼働状態を分類し、教師データを選択するための情報として、特徴量が不足することとなる。

本稿では、この分類のための特徴量を補足するため、過去の稼働状態との比較を示す変化点検知データと、風況から得られる入力情報を同一とする、同じ風力発電所内の他の風車の稼働状態との比較データを利用する。そして、正常稼働率データ、変化点検知データ、他の風車との比較データを基に、X-means[17]を用いて、1 か月単位の SCADA データが所属するクラスを算出する。このクラスタリング処理において、他の風車との比較データにはバイアスを

下げ、汎化性能を向上させる役割がある。

2.1 SCADA データのデータ特性

SCADA は風車の稼働監視、制御装置であり、売電対象の発電に関わるデータ以外に、風況データや温度データ、機器の稼働ステータスデータも収集する。また、遠隔からの稼働監視において、通信エラーによるデータ収集ミスを防ぐため、通信データサイズを最小限に抑える設定が一般的である。よって、SCADA データは、収集データ項目は広範囲だが、600 s 間の平均値、最大値、最小値、標準偏差という、現実の風車の稼働状態を荒く平均化したデータとなっている。

2.2 パワーカーブに基づく正常稼働率の算出

正常稼働率の算出では、まず、理想的なパワーカーブとの間の類似性を判断するための指標を作成する。理想的なパワーカーブの風速を 0.5 m/s 毎に分割し、分割した領域の発電出力毎に発電出力の 95 パーセントに該当する値を公差として求め、分割領域の発電出力の最大値に公差を加えた値、最小値に対して公差を引いた値を、当該風速における発電出力の許容範囲として設定する。

次に、時系列 SCADA データに対し、風車に負荷が掛かるレコードのみを抽出するため、風速データ項目の値に対してフィルタリング処理を行う。本稿における風車の有効風速については 9.0 - 14.0 m/s を設定する[10]。

最後に、抽出した各レコードの風速と発電出力を取り出し、理想的なパワーカーブの許容範囲内に収まるレコードを正常稼働状態レコードとし、正常稼働率を算出する。

2.3 特異スペクトル変換による正常稼働率の変化点検知

正常稼働率において時系列の変動を除外し、理想状態に近い期間を明確化するため、変化点検知を行う。変化点検知とは、スライド窓を動かしながら過去の状態データ空間と現在の状態データ空間を比較し、両者の間に相違が発生した時点を変化点として検知する手法である。

変化点検知手法として、本稿では SST を適用する。SST は時系列データに対する特異値分解[18]からパターンの特徴を抽出する手法であるが、時系列データ毎の特徴に合わせた補正操作や修正操作を必要としない点を利点としており、季節変動を伴う気候の時系列変化に対する解析などで利用されている[19]。

SST による変化点検知処理を説明する。

時系列データ $X = [x_1 \ \cdots \ x_t \ x_{t+1} \ \cdots \ x_{t+w-1} \ \cdots \ x_T]$ において、 $[x_t \ x_{t+1} \ \cdots \ x_{t+w-1}]$ をウィンドウ幅 w の部分列とし、これを列ベクトル y_t として表す。

$$y_t = [x_t \ x_{t+1} \ \cdots \ x_{t+w-1}]^T \quad (1)$$

この列ベクトル y_t を、左から時系列順に k 個縦に並べた行列 $H(t)$ を履歴行列とする。

$$H(t) = \begin{bmatrix} \mathbf{x}_t & \mathbf{x}_{t+1} & \cdots & \mathbf{x}_{t+k-1} \\ \mathbf{x}_{t+1} & \mathbf{x}_{t+2} & \cdots & \mathbf{x}_{t+k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{t+w-1} & \mathbf{x}_{t+w} & \cdots & \mathbf{x}_{t+w+k-2} \end{bmatrix} \quad (2)$$

$H(t)$ からパターンの特徴を抽出するため、特異値分解を行う。この特異値分解で得られた特異値 u_i を降順で並べ、上位 m 個から構築した左特異ベクトル $U_m(t)$ を過去の代表パターンとする。

$$U_m(t) = [\mathbf{u}_1(t) \quad \mathbf{u}_2(t) \quad \cdots \quad \mathbf{u}_m(t)] \quad (3)$$

ここで、時系列上での履歴行列とテスト行列の間の相互位置の距離をラグ L と呼ぶ。

比較対象のテスト行列についても、履歴行列と同様に時系列データから抽出し、特異値分解を行い、得られた特異値から構築したテストベクトルを $Q_m(t)$ と表す。

$$Q_m(t) = [\mathbf{q}_1(t) \quad \mathbf{q}_2(t) \quad \cdots \quad \mathbf{q}_m(t)] \quad (4)$$

履歴行列に対するテスト行列の変化度として異常度 $a(t)$ を算出する。

$$a(t) = 1 - \|U_m(t)^T Q_m(t)\|_2^2 \quad (5)$$

ここで、 J 次元のベクトル $\mathbf{x}$ に対する行列 p ノルムの定義は以下となる。

$$\|\mathbf{x}\|_p = (|\mathbf{x}_1|^p + |\mathbf{x}_2|^p + \cdots + |\mathbf{x}_J|^p)^{1/p} \quad (6)$$

任意の点 t における異常度 $a(t)$ が1に近づく場合、履歴行列とテスト行列の間の相関関係が弱く、相違が大きいことを意味する。つまり、当該点 t が変化点を示すこととなる。

2.4 風力発電所内他風車との比較データの算出

ここまで、単一の風車の時系列 SCADA データを対象に、1 か月単位の正常稼働率と当該正常稼働率の変化点を求めた。しかし、単一の風車の SCADA データから算出しているため、例えば計測された SCADA データの発電量の変動が激しい場合、それが単一の風車固有に発生した事象なのか、周囲の風況の乱れによって生じた結果なのか、判定が困難である。

そこで、同一風力発電所内の他の風車の SCADA データから得られた値と比較することで、単一の風車固有の稼働状態と周囲環境から生じた稼働状態を同時に考慮し、クラスタリングの汎化性能を向上させる。比較対象のデータは風車の入力情報である風速データと、出力情報から得た正常稼働率データである。

具体的には、同一風力発電所内に存在する全ての風車に

対し、1 か月単位の有効風速レコード数の割合を算出する。そして、1 か月単位の有効風速レコード数の割合を平均化し、時系列の風速比較データとする。同様に、正常稼働率データについても全ての風車で平均化し、正常稼働率比較データとする。

有効風速の割合と正常稼働率の比較については DTW を用いて風車全体の平均との間の類似度を算出する。DTW は時系列データ間の類似性を示す指標として利用される手法である。

DTW は 2 つの時系列データの類似度を求めるため、総当たりで両者の距離を計算し、最小値となる場合を両者の関係性を示す値とする。本稿では距離算出のスライド窓幅 W を設定し、 W を1 か月単位で動かしながら距離を求める。

2.5 X-means による正常稼働期間の分類

教師なしクラスタリング手法に k-means がある。k-means では、分類対象の N 個の多次元データ $\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$ において、予め分類するクラスタ数 K を決定し、各クラスタ k の重心位置 μ_k を変更しながら、それぞれのデータ点から最も距離に近い μ_k との二乗距離の総和が最小となる μ_k を求めることで、クラスタと所属データの割り当てを行う。k-means は各データが必ず1つのクラスタに所属するハードクラスタリングであり、広範囲に適用可能だが、 K の事前決定が必要であるという点が課題である。

この課題を解決する手法が X-means である。X-means では、k-means の処理に加え、ベイズ情報量基準 (Bayesian Information Criterion : BIC) [18]を用いて妥当なクラスタ数を自動的に決定する。具体的には、最初に十分小さなクラスタ数から開始し、各クラスタを2分割する処理を繰り返しながら、分割処理毎に BIC の値を算出する。ここで、分割前後の BIC の値を比較し、分割後に BIC の値が大きくなる場合、分割後のクラスタが妥当でないと判断し、分割処理を終了する。

X-means における BIC は、 μ_k の周囲にデータが正規分布として存在するという仮定に基づいて算出する。

$$\text{BIC} = -2 \log L(\hat{\theta}_i) + q \log n_i \quad (7)$$

ここで、 i はクラスタの識別子、 $\hat{\theta}_i$ は多次元正規分布の最尤推定値、 $L(\hat{\theta}_i)$ は尤度関数、 q は次元数、 n_i はクラスタ i に含まれるデータ数を示す。

X-means によって複数のクラスタに分割した1 か月毎の風車の稼働状態に対し、規準化した正常稼働率の値を用いて教師データの判定を行う。判定処理では、正常稼働率が最も大きい1 か月を含むクラスタを教師データとし、当該クラスタの所属期間を教師データ期間と判定する。

3. 稼働状態監視モデルに基づく異常検知手法

教師データ選択手法を用いて分類した教師データに対し、

筆者らが提案している稼働状態監視モデルを評価器として適用し、異常検知による評価を行う。本章では、筆者らが提案している異常検知手法を説明する。

3.1 稼働状態監視モデル

稼働状態監視モデルとは、1 機の風車全体の健全性評価結果を異常度として出力する機械学習モデルであり、風車ドライブトレインに関わる複数の SCADA データ項目を利用する。稼働状態監視モデルによる異常検知では、大局的な風車の健全性評価を行うため、時系列データの 4 つの変動要因のうち、傾向変動要因を SCADA データから抽出する。そのため、フィルタリング処理を行い、傾向変動要因を抽出すると同時に、季節変動要因と不規則変動要因を除外する。また、循環変動要因についてはフィルタリング処理による除外が困難であるため、モデル構築後、テストデータを用いて算出した異常度に対し、指数加重移動平均を適用して循環変動要因を除外する。

稼働状態監視モデルは、機械学習アルゴリズムとして、混合正規分布 (Gaussian Mixture Model : GMM) を用いて異常検知処理を実行する。GMM はテストデータ $\mathbf{x}$ と教師データとの類似度を示す確率分布 $p(\mathbf{x}|\Theta)$ を、正規分布 $\mathcal{N}(\mathbf{x}|\mu_k, \Sigma_k)$ を用いて算出する。

$$\begin{cases} p(\mathbf{x}|\Theta) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x}|\mu_k, \Sigma_k) \\ 0 \leq \pi_k \leq 1 \\ \sum_{k=1}^K \pi_k = 1 \end{cases} \quad (8)$$

ここで、 π_k は正規分布 k の重み係数、 μ_k は正規分布 k の平均値、 Σ_k は正規分布 k の分散値、 Θ は正規分布毎の π_k 、 μ_k 、 Σ_k の 3 つのパラメータの集合を示す。

3.2 異常度の定義

GMM で求めた教師データとテストデータとの類似度 $p(\mathbf{x}|\Theta)$ を用いて、テストデータの異常度を負の対数尤度 (Negative Log Likelihood Probability : NLLP) で定義する。

$$\text{NLLP} = -\ln p(\mathbf{x}|\Theta) \quad (9)$$

NLLP はデータ間の情報量変化を表しており、教師データとして与えられる正常稼働状態と、テストデータとして与えられる SCADA データとの間の類似度の差が大きいほど、NLLP は大きな値を示す。

(8)式より求めた NLLP に対し、時系列データの変動要因である循環変動要因を除外するため、指数加重移動平均 W_t を適用する。

$$W_t = (1 - \alpha) \cdot W_{t-1} + \alpha \cdot \text{NLLP}_t \quad (10)$$

ここで、 t はテストデータの時系列インデックス、 α は (0,1) の平滑化定数を示す。

3.3 異常度の閾値による判定

教師データは正常稼働状態であるという想定より、構築した GMM 自体に再びこの教師データを適用し、NLLP を算出することで、正常稼働状態の異常度を得ることができる。得られた異常度の 99 パーセンタイルに該当する値を正常稼働状態と異常状態とを分ける閾値として設定し、テストデータから算出した NLLP がこの閾値を超過し続けることで、風車が異常状態であると判定する。

4. 実機 SCADA データを用いた評価

4.1 評価対象データ

教師データ選択手法の評価で利用する SCADA データは、風力発電所で実際に稼働している出力 2.0 MW 以下の風車 10 機 A~J において、600 s 間隔で収集されたデータを使用する。これら風車のうち、A, H, I において内部機器の交換作業が発生しており、選択した教師データから、この交換作業を行った期間を適切に分類することが可能となることで、分類操作の性能を評価する。同時に、分類操作における過学習を防止する操作の有効性を評価する。

また、教師データ期間を分類し、選択したデータを用いて、稼働状態監視モデルによる異常検知が正常に実行できることで、選択した教師データの妥当性を評価する。

4.2 機器交換事例による教師データ選択手法の評価

風車 A, H, I に対し、SCADA データから取得した 1 か月単位の正常稼働率、変化点検知、有効風速と正常稼働率の DTW 距離の 4 種類のデータを説明変数とし、1 か月毎に複数のクラスタに分類した結果を図 1 - 図 3 に示す。各グラフにおいて、左軸は分類結果であるクラスタ、右軸は説明変数である 4 種類のデータである。ここで、季節変動

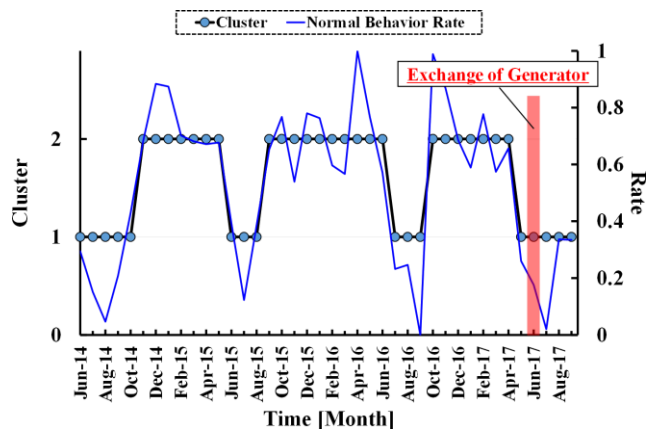


図 1 風車 A の教師データ分類結果

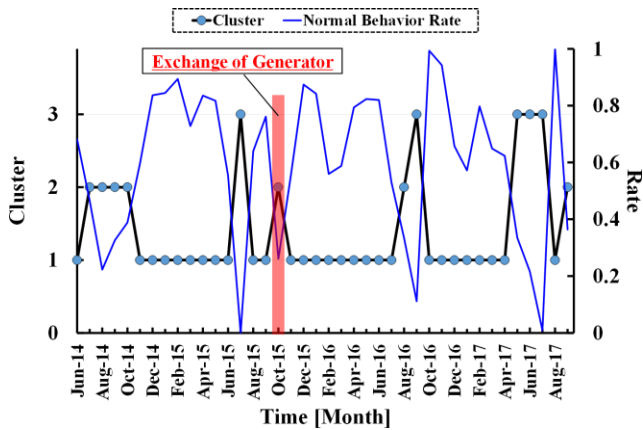


図 2 風車 H の教師データ分類結果

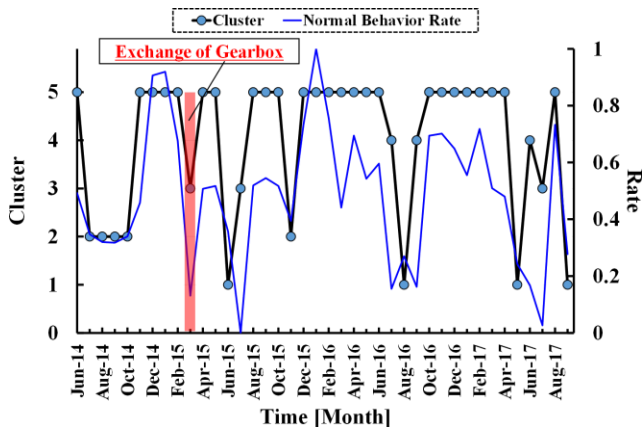


図 3 風車 I の教師データ分類結果

性を考慮した SST パラメータとして $w = 4$, $m = 2$, $k = 2$, $L = 3$ を適用した。また, DTW パラメータとして $W = 6$ を適用した。

図 1 - 図 3 中の赤棒は内部機器の交換作業が発生した年月を示すが, 赤棒に割り当てられたクラスタと, 正常稼働率が低下した年月のクラスタが一致した。よって, 教師データ選択手法では, 規準化した正常稼働率の値が高い年月に割り当てられたクラスタを選択することで, 教師データとして適切な年月を選択することができる。

4.3 教師データ選択手法の汎化性能向上策の評価

同一発電所内の他の風車との比較による DTW 距離情報を考慮した結果を a, DTW 距離情報を考慮しない結果を b とし, DTW 距離情報がクラスタ数に与える影響を図 4 に示す。

a と b を比較すると, 風車 A, B, C, G, I, J ではクラスタ数が異なり, 特に, A, B, C では b におけるクラスタ数が a の 2 倍以上となった。これは, 周囲の他の風車との間の比較を行うと過学習を防止し, 正常稼働率の変動に従って, 多数のクラスタに分類されることを示している。

分類結果から教師データを選択する際, b のように多数のクラスタに分類されると, 選択処理の判断基準として, 各クラスタの特徴を明確化しなければならず, 選択が困難となる。よって, 汎化性能を向上させ, 選択を容易に実行

するには, DTW 距離情報の考慮が妥当である。

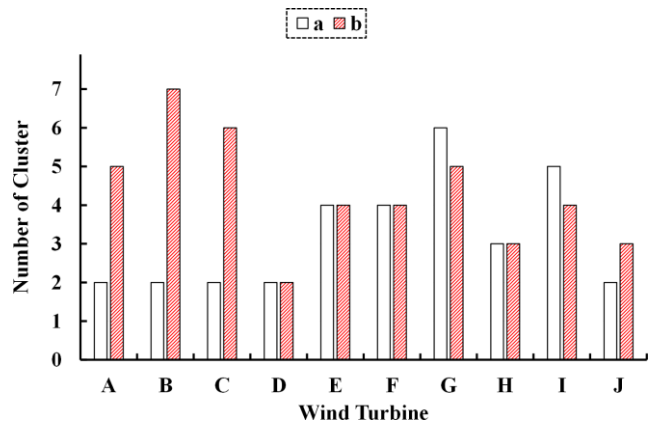


図 4 他風車との比較の有無による分類クラスタ数の変化

4.4 異常検知手法を用いた教師データ選択の妥当性評価

内部機器の交換作業が発生した風車 A, H, I に対し, 教師データ選択手法を用いて分類した教師データ期間の SCADA データを用いて稼働状態監視モデルを構築し, 異常検知を行った結果を図 5 - 図 7 に示す。

風車 A について, 教師データ期間として 2014/11/01 - 2015/05/01 を選択し, この期間の SCADA データを用いて異常検知を行った。その結果, 図 5 において, 発電機を交換した 2017 年 5 月の約 4 か月前である 2017 年 1 月に異常

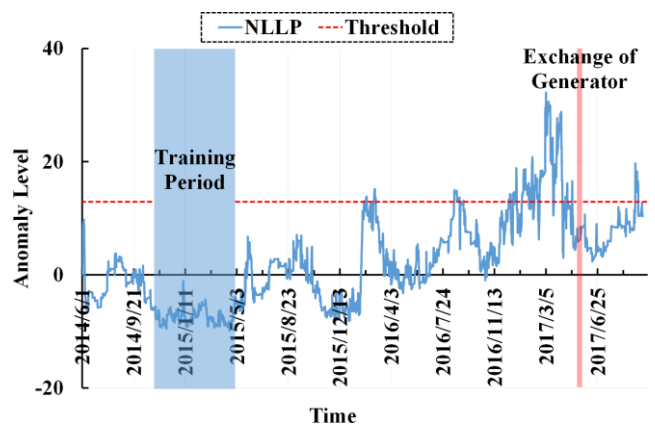


図 5 風車 A の異常検知結果

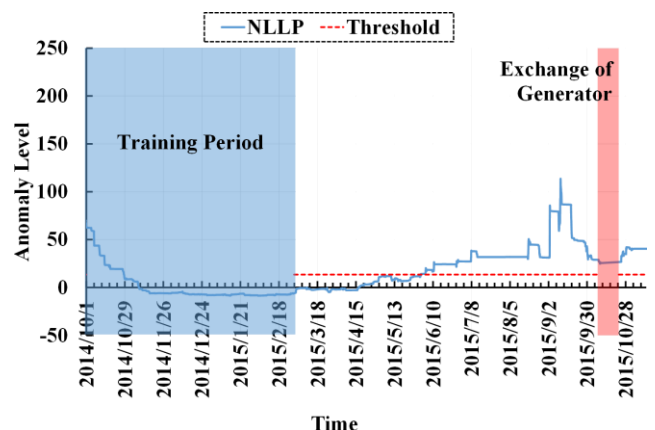


図 6 風車 H の異常検知結果

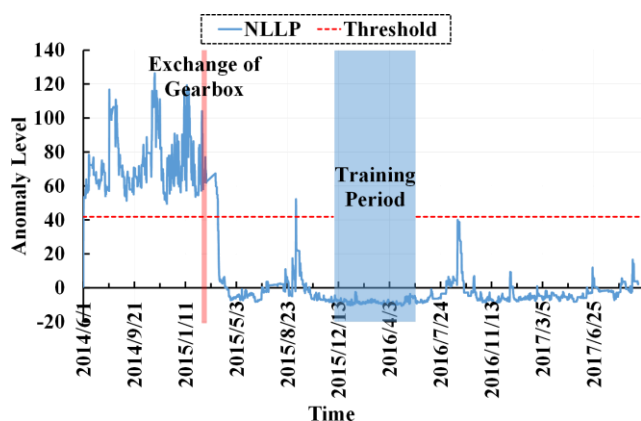


図 7 風車 I の異常検知結果

度が閾値を超過した。よって、風車 A の異常検知結果から、教師データ選択手法の分類処理が妥当であることが分かった。

風車 H について、教師データ期間として 2014/11/01 - 2015/03/01 を選択し、異常検知を行った結果、図 6 において、2015 年 6 月には異常度が閾値を超過し、発電機交換の約 4 か月前に閾値を超過した。風車 H では教師データ期間から発電機交換までの間は 7 か月と短い、このような異常検知処理にとって不利な条件においても、教師データ選択手法で分類した教師データを利用すれば異常検知が可能であり、教師データ選択手法の分類結果が適切であることを示している。

風車 I について、教師データ期間として 2015/12/01 - 2016/06/01 を選択した。これは、図 3 におけるクラスタ 5 のうち、十分な連続データ期間として 6 か月を選択することが可能であったからである。この教師データを用いて異常検知処理を行った結果、発電機交換前後で閾値を超過したことが分かった。また、教師データ期間以降の異常度は閾値を超過することなく、誤検知は発生していない。これは、教師データ期間における風車の稼働状態と、発電機交換以前の風車の稼働状態が明示的に異なり、教師データとして分類したクラスタの妥当性を示している。

5. おわりに

風車メーカーが提供するパワーカーブを基に、SCADA データから抽出した正常稼働率を用いて、風車の異常検知で利用可能な教師データ選択手法について示した。また、実機 SCADA データと、筆者らが提案する異常検知手法を用いて教師データ選択手法の妥当性を評価し、風車の異常検知に適用可能であることを示す結果を得た。

謝辞

本研究は NEDO 事業「風力等自然エネルギー技術研究開発／風力発電高度実用化研究開発／スマートメンテナンス技術研究開発(分析)」による支援を受けた。分析にあたり、

本事業へご協力くださった関係各位に感謝申し上げます。

参考文献

- [1] C. Lins, et al.. RENEWABLES 2017 GLOBAL STATUS REPORT, http://www.ren21.net/gsr_2017_full_report_en, アクセス 2018 年 7 月 25 日.
- [2] 風力発電競争力強化研究会報告書. http://www.meti.go.jp/committee/kenkyukai/energy_environment/furyoku/pdf/report_01_01.pdf, アクセス 2018 年 7 月 25 日.
- [3] Milborrow, D.. Operation and maintenance costs compared and revealed. Windstats Newsletter. 2006, vol. 19, no. 1-3.
- [4] Tautz-Weinert, J., & Watson, S. J.. Using SCADA data for wind turbine condition monitoring—a review. IET Renewable Power Generation. 2016, vol. 11, no. 4, p. 382-394.
- [5] Marvuglia, A. and Messineo, A.. Monitoring of wind farms' power curves using machine learning techniques. Applied Energy. 2012, vol. 98, p. 574-583.
- [6] Lapira, E., Brisset, D., Ardakani, H. D., Siegel, D. and Lee, J.. Wind turbine performance assessment using multi-regime modeling approach. Renewable Energy. 2012, vol. 45, p. 86-95.
- [7] Sun, P., Li, J., Wang, C., and Lei, X. A generalized model for wind turbine anomaly identification based on SCADA data. Applied Energy. 2016, vol. 168, p. 550-567.
- [8] Zaher, A. S. A. E., McArthur, S. D. J., Infield, D. G. and Patel, Y.. Online wind turbine fault detection through automated SCADA data analysis. Wind Energy. 2009, vol. 12, no. 6, p. 574-593.
- [9] Yasuda, A., Ogata, J., Murakawa, M., Morikawa, H., and Iida, M.. A Fundamental Study and Factor Analysis on System Health Monitoring of Wind Turbines Using SCADA. WVEC 2016 TOKYO. 2016, no. C-1-1.
- [10] A. Yasuda, J. Ogata, M. Murakawa, H. Morikawa and M. Iida. System Health Monitoring of Wind Turbines Using SCADA Data and Gaussian Mixture Models. PHMAP17. 2017, P-088.
- [11] Skrimpas, G. A., Ursin, T., Sweeney, C., Marhadi, K., Mijatovic, N., and Holboell, J.. Residual signal feature extraction for gearbox planetary stage fault detection. Wind Energy. 2017, vol. 20, no. 8, p. 1389-1404.
- [12] Schlechtingen, Meik, Ilmar Ferreira Santos, and Sofiane Achiche. Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: System description. Applied Soft Computing. 2013, 13.1, p. 259-270.
- [13] Idé, Tsuyoshi, and Keisuke Inoue. Knowledge discovery from heterogeneous dynamic systems using change-point correlations. Proceedings of the 2005 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics. 2005.
- [14] Idé T, and Inoue K. Knowledge discovery from time-series data using nonlinear transformations. In: Proceedings of the fourth data mining workshop (The Japan Society for Software Science and Technology, Tokyo, 2004). 2004, no. 29, p. 1-8.
- [15] Idé, Tsuyoshi. Speeding up change-point detection using matrix compression. IBIS. 2006.
- [16] Senin, Pavel. Dynamic time warping algorithm review. Information and Computer Science Department University of Hawaii at Manoa Honolulu, USA. 2008, 855, 1-23.
- [17] Pelleg, Dan, and Andrew W. Moore. X-means: Extending k-means with efficient estimation of the number of clusters. Icml. 2000, vol. 1, p. 727-734.
- [18] C.M. ビショップ, パターン認識と機械学習. シュブリンガー・ジャパン, 2008.
- [19] Itoh, Naoki, and Jürgen Kurths. Change-point detection of climate time series by nonparametric method. Proceedings of the world congress on engineering and computer science, 2010, vol. 1.