

## ダム堤体で観測された地震記録の教師なし機械学習による異常検知

国土交通省国土技術政策総合研究所 正会員 ○佐藤 弘行、大谷 周、傅 斌、金銅 将史

### 1. はじめに

ダムの固有振動数等の振動特性は、ダムの地震時挙動を把握する上で重要な指標である。ダムの固有振動数は、堤体基礎部と上部（天端等）で得られた地震観測記録（加速度時刻歴データ）による伝達関数（基礎部に対する上部のフーリエ振幅スペクトル比）のランニングスペクトルの分析<sup>1)</sup>やARXモデルを用いたシステム同定<sup>2)</sup>を行った既往研究から、比較的強い地震動を受けると変化（減少）することが指摘されている。しかし、これらの手法では観測記録に含まれるノイズ等の影響により固有振動数の経時変化を精度良く評価することが困難な場合がある。固有振動数の変化はダムの構造的損傷に関わる可能性があり、精度良く把握できれば地震後の安全確認に役立てることができる。そこで本論文では、ダムで取得される地震記録のうち、堤体応答が反映される上部（天端）での観測記録をもとに、ダムの固有振動数の変化を見逃すことなく検知することを目的として、教師なし機械学習の手法による異常検知の可能性について検討した。

### 2. 検討対象と使用した地震観測記録

既往研究<sup>1), 2)</sup>により、比較的大きな地震動を受けた際に堤体の固有振動数が変化したことが指摘されている2ダム（Aダム：重力式コンクリートダム（堤高65m）、Bダム：中央コア型ロックフィルダム（堤高74.4m））を対象とし、それぞれのダムの天端で観測された加速度時刻歴データのうち、Aダムについてはダム軸方向の19データ、Bダムについては上下流方向の23データを使用した。加速度時刻歴データはいずれも100Hzのサンプリングレートで記録されている。なお、両ダムともに、地震後の臨時点検およびその後の観測データ等からダムの安全性には問題がないことが確認されている。

### 3. 教師なし機械学習による正常／異常データの分類方法

本検討では、教師なし機械学習での異常検知に用いられることの多い①One Class SVM、②Isolation Forest、③k近傍法の3手法<sup>3)</sup>を用い、固有振動数の変化が指摘されている比較的強い地震時の観測記録をテストデータ、同じダムでテストデータ以前の中小地震時に得られた観測記録を比較用データ、同じダムでのテストデータ及び比較用データ以外の観測記録を学習用データとした。なお、学習用データとして用いた観測記録については、伝達関数のランニングスペクトルあるいはARXモデルによるシステム同定の結果から、固有振動数に明確な変化がないことを別途確認している。これら各データは、機械学習の前処理として、学習用データは周波数領域における特徴量が安定して抽出できる時間間隔（5.12秒）でデータを分割した。テストデータおよび比較用データについては、機械学習による分類時のデータ数を多くするため0.5秒ずつずらしながら、学習用データと同じ時間間隔でデータを分割した。分割したそれぞれのデータについて、表-1に示す11種類の特徴量を算出したうえで、各特徴量の絶対値を揃えるため、各特徴量を平均が0、標準偏差が1となるよう標準化した。なお、表-1の特徴量のうち、動的特徴量は線形最小二乗近似の傾き、エントロピーは底を2としたパワースペクトルの対数の和に(-1)を掛けた値、パワーバンドは5から15Hzのパワースペクトルの和を表す。機械学習の各手法（①～③）について、学習用データを全て正常として学習させた後、テストデータおよび比較用データを対象に正常データと異常データの分類を試みた。

表-1 算出した特徴量

1	平均
2	分散
3	最小値
4	最大値
5	中央値
6	歪度
7	尖度
8	RMS
9	動的特徴量
10	エントロピー
11	パワーバンド

### 4. 教師なし機械学習による分類結果

テストデータと比較用データの主要動を含む時間帯におけるAダムとBダムの分類結果を図-1と図-2に示す。両図において、(a)はダム天端で記録された加速度時刻歴波形、(b)は伝達関数（基礎部の加速度に対する天端の加速度のフーリエ振幅スペクトル比のランニングスペクトル（背景））とARXモデルによるシステム同定で推定された固有振動数の経時変化（実線）、(c)はOne Class SVMによる分類結果（1：正常、-1：異常）、(d)はIsolation Forestによる分類結果（1：正常、-1：異常）、(e)はk近傍法による分類結果（1に近いほど異常）を示している。

One Class SVMによる分類結果（図-1および図-2(c)）は、テストデータでは堤体の固有振動数が減少している時間において異常と分類されているデータが多いものの、テストデータおよび比較用データにおいて堤体の固有振動数に変化が確認されない時間においても異常と誤判定されているデータが多くなっている。Isolation Forestによる分類結果（図-1および図-2(d)）は、テストデータでは堤体の固有振動数が減少している時間において異常と分類されているデータ数はOne Class SVMよりも少なくなっているものの、比較的正しく異常と分類されており、またテストデータおよび比較用データにおいて堤体の固有振動数に変化が確認されない時間においては正常と正しく分類されている。k近傍法の分類結果（図-1および図-2(e)）は、One Class SVMでの分類結果と同様の傾向となっている。これらのことから、本検討において用いたデータとその前処理方法に対しては、Isolation Forestによる分類が異常検知の上で最も高い精度が得られていると評価できる。

そこで、Isolation Forestによる分類結果への各特徴量の影響を調べるため、固有振動数の変化が比較的大きいBダムの学習用データおよびテストデータを対象として、各特徴量が分類結果に及ぼす寄与度を表す指標であるShap値<sup>4)</sup>を算出した結果を図-3に示す。図-3において、赤色のプロットは3.で述べた時間間隔で分割した個々のデータから算出した特徴量が大きな値であることを示し、青色のプロットは特徴量が小さいことを表している。図-3では、分類への寄与度が高い（Shap値の絶対値の平均が大きい）ほど上位となるように特徴量を並べるとともに、Shap値のマイナスの絶対値が大きくなるほど異常へ分類されやすいことを示している。図-3の学習用データとテストデータのShap値の分布を比較すると、学習用データおよびテストデータいずれにおいてもパワーバンド（図-3の赤実線で囲まれたプロット）および分散（図-3の赤点線で囲まれたプロット）の寄与度が高くなっている。学習用データではパワーバンドおよび分散のShap値は正のプロットが多くなっているものの、テストデータではそれらのShap値は全て負の値となっており、学習用データとテストデータにおいてパワーバンドと分散のShap値の分布は大きく変化している。また、学習用データではエントロピーおよび歪度の寄与度の順位が低くなっているが、テストデータではこれらの寄与度の順位が高くなっていることから、これらの特徴量の影響によりテストデータが異常と分類されやすくなった

キーワード：ダム、地震動、固有振動数、機械学習、異常検知

連絡先：〒305-0804 茨城県つくば市旭1 国土技術政策総合研究所 TEL:029-864-2211

可能性がある。そこで、Isolation Forest によりテストデータが異常と分類された要因をさらに推定するために、B ダムの学習用データ、テストデータおよび比較用データについて、パワーバンド、分散、エントロピーおよび歪度の経時変化を図-4 に示す。同図より、Isolation Forest によりテストデータの中で異常と分類された時間帯において、テストデータのパワーバンドおよび分散は学習用データおよび比較用データのそれらの値よりも大きくなっている（同図の緑囲み箇所）。このような学習用データおよび比較用データの特徴量の経時変化とは異なるテストデータの特徴量の経時変化の傾向が、テストデータにおいて異常として分類された時間帯が存在する要因の一つと考えられる。一方、図-4 からは、テストデータのエントロピーおよび歪度の経時変化については、学習用データおよび比較用データのそれらの経時変化との明確な差異は確認されなかった。

5. まとめ

本検討では、地震によるダム堤体の固有振動数の変化を異常として検知することを目的に、ダム天端で記録された加速度時刻歴データを対象として、教師なし機械学習による正常・異常の分類を試みた。その結果、本検討での前処理方法によった場合には、Isolation Forest により比較的精度良く異常検知が可能となる結果が得られた。今後は、異なる特徴量を用いた場合にその特徴量が異常検知に及ぼす影響の検討を行うなど、引き続きダムの固有振動数の変化から詳細点検を必要とする異常を見逃すことなく検知する手法について検討を行う予定である。

参考文献

- 1)金銅将史・小堀俊秀・佐々木隆:地震動がコンクリートダムの振動特性に及ぼす影響、ダム工学、27(4)、pp.265-278、2017.
- 2)尾茂淳平・佐藤弘行・金銅将史:ARXモデルによるシステム同定を用いたダム堤体の振動パラメータの高精度推定、ダム工学研究発表会、2019.
- 3)例えば、福井健一:識別・予測・異常検知、148p、2018.
- 4)S.M. Lundberg and S.I. Lee : A unified approach to interpreting model predictions, 31st Conference on Neural Information Processing Systems, pp. 4765-4774, 2017.

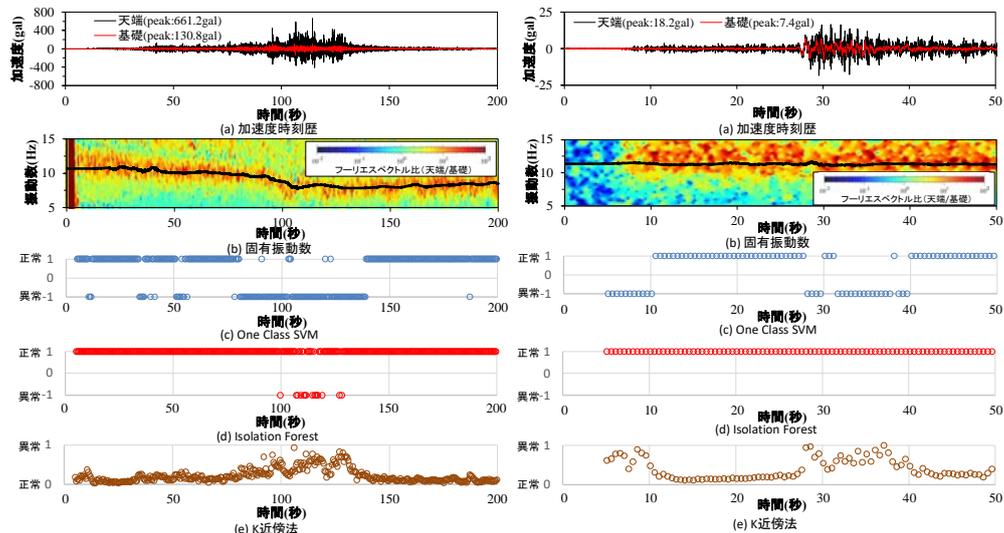


図-1 Aダムの分類結果(左:テストデータ、右:比較用データ)

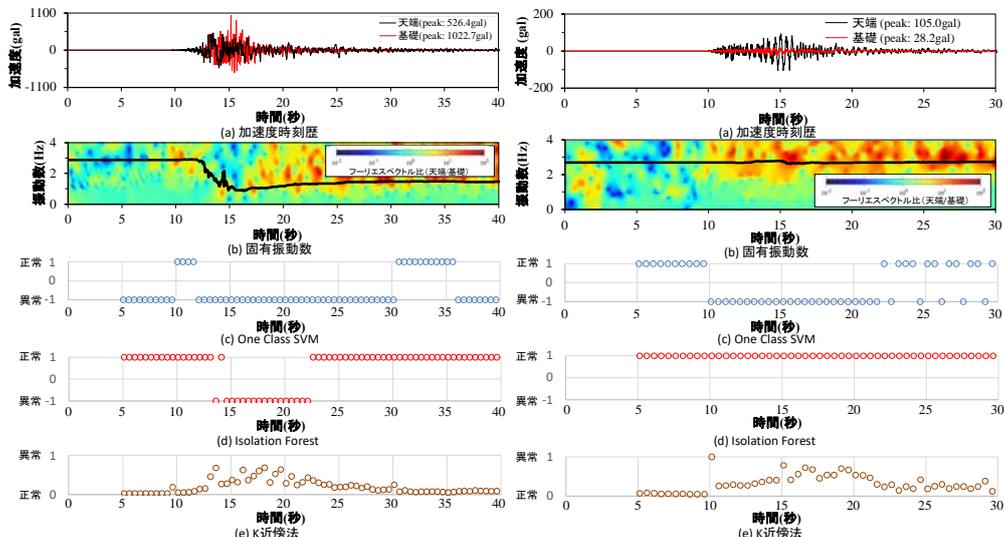


図-2 Bダムの分類結果(左:テストデータ、右:比較用データ)

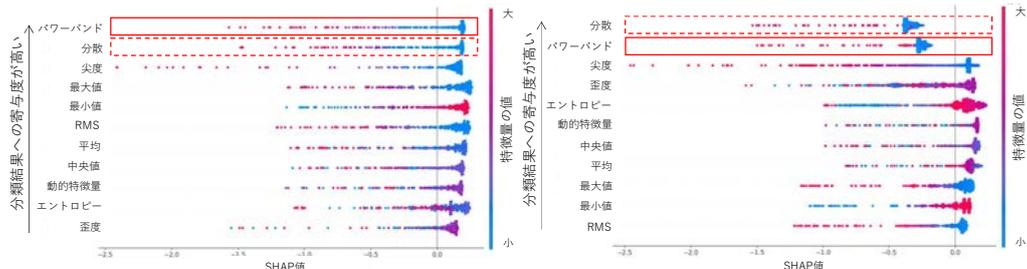


図-3 分類に及ぼす各特徴量の寄与度(Bダム、左:学習用データ、右:テストデータ)

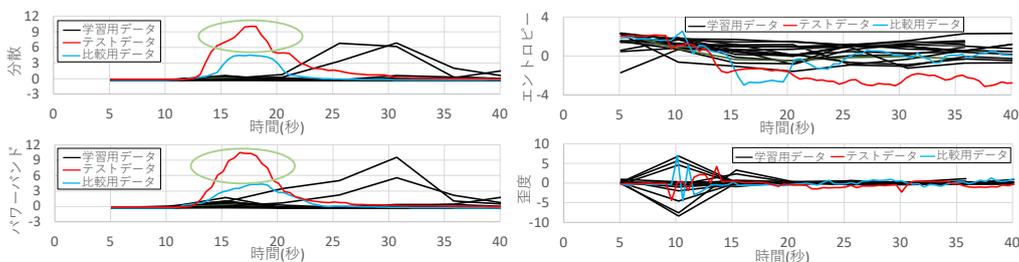


図-4 標準化した特徴量の経時変化 (Bダム、左上:分散、左下:パワーバンド、右上:エントロピー、右下:歪度)