

# 一ツ瀬ダムにおけるアオコ発生 of 機械学習を用いた要因分析

長崎県立大学 情報システム学部 情報システム学科

平岡 透

# 一ツ瀬ダムにおけるアオコ発生 of 機械学習を用いた要因分析

平岡 透

## 1. はじめに

ダム湖は、治水に限らず水源確保の役割も担っており、その水質の維持や保全に関する対策は重視されている。しかしながら、近年、ダム湖において有害藻類（アオコ）の異常発生が問題として挙げられている[1][2]。アオコは異臭を放つことや生態系への悪影響をもたらすこと、景観を悪化させること、利水に悪影響をもたらすことから対策が急務とされている。また、アオコが大量発生した際には水質確保に多くの資金や労力がかかってしまう。このようなアオコの異常発生は、宮崎県西都市の一ツ瀬ダムでも起こっており、アオコの早期発見や発生の予測を行わなければならない。

そこで、本研究では、一ツ瀬ダムにおけるアオコ（ミクロキスティス）の発生に関わると考えられる要因を基にミクロキスティスを機械学習によって推定することを目的とする。

## 2. 方法と実験

本研究では、大きく三つの研究を行う。一つ目がサポートベクトル回帰分析とランダムフォレストを用いてミクロキスティスを推定する研究、二つ目が類似度を用いてミクロキスティスを推定する研究、三つ目がミクロキスティスの発生予測と発生要因を分析する研究である。

### 2.1 使用データ

本研究では、一ツ瀬ダムで2015年から2020年の1月、4月、5月、6月、7月、8月、9月、10月、11月に合計54回観測されたデータを用いた。データ項目は、ミクロキスティス、水温、pH、渦鞭毛藻類、全リン、全窒素、硝酸、流入量を用いた。アオコの発生には、水温が関係しており、夏場に発生しやすくなる[3]。ミクロキスティスの最大繁殖が得られるpHは10付近であり、ミクロキスティスがアオコ場外となっている湖沼のpHは9前後が多い[3]。筆者らが調査した結果では、ミクロキスティスが異常発生する数か月前に渦鞭毛藻類が多く発生していた。アオコの成長に必要な栄養源としてリンや窒素などがあり、水中に窒素があることにより硝酸なども発生し、それらも栄養源の一部とされている[3][4]。台風などの急激な流量の増加でもアオコの発生量が変化する[5]。また、観測地点は、ダム湖における一ツ瀬川の上流地点（以下、上流）および銀鏡川の流入口地点（以下、銀鏡川）とした。

### 2.2 サポートベクトル回帰分析とランダムフォレストを用いたミクロキスティスの推定

2015年から2020年のデータに対してサポートベクトル回帰分析とランダムフォレスト

を用いて実験を行った。このとき、マイクロキシティスの量が 100 以上の場合を発生とした場合、マイクロキシティスの量が 1000 以上の場合を発生とした場合、マイクロキシティスの値を推定した場合の三つの場合に分けて分析を行った。

実験で得られた決定係数を表 1 から表 3 に示す。マイクロキシティスの値をランダムフォレストで予測した場合に上流および銀鏡川において、決定係数がそれぞれ 0.581 と 0.629 と比較的大きくなった。上流よりも銀鏡川の決定係数が大きくなった要因として、マイクロキシティスの発生回数に関係していると考えられる。上流においてマイクロキシティスの発生回数がサンプル数 54 のうち 10 回、銀鏡川において 12 回であった。銀鏡川において、決定係数が 0.6 を上回っていることから、ランダムフォレストにおいてデータ分析の意味をなしているといえる。

表 1 ミクロキシティスの値が 100 以上の場合を発生とした場合の決定係数

	上流	銀鏡川
サポートベクトル回帰分析	-0.021	-0.019
ランダムフォレスト	0.384	0.36

表 2 ミクロキシティスの値が 1000 以上の場合を発生とした場合の決定係数

	上流	銀鏡川
サポートベクトル回帰分析	0.002	-0.008
ランダムフォレスト	0.487	0.389

表 3 ミクロキシティスの値を予測した場合の決定係数

	上流	銀鏡川
サポートベクトル回帰分析	0.021	0.104
ランダムフォレスト	0.581	0.629

### 2.3 類似度を用いたマイクロキシティスの推定

類似度[6]を用いて 2015 年から 2019 年の観測データから 2020 年のマイクロキシティスの量を推定する方法を示す。 $I$ 個の測定日があり、 $i$ 番目のマイクロキシティスを $p_i$ とする。また、 $J$ 項目の観測データがあり、 $i$ 番目の観測日の $j$ 番目の観測データを $f_{i,j}$ とする。予測する 2020 年の観測日の $j$ 番目のデータを $\hat{f}_j$ とする。

まず、観測データ $f_{i,j}$ と $\hat{f}_j$ を式(1)から式(4)で正規化し、それぞれ $F_{i,j}$ と $\hat{F}_j$ とする。

$$\bar{f}_j = \frac{\sum_{i=1}^I f_{i,j}}{I} \quad (1)$$

$$\bar{\bar{f}}_j = \sqrt{\frac{\sum_{i=1}^I (f_{i,j} - \bar{f}_j)^2}{I}} \quad (2)$$

$$F_{i,j} = \frac{f_{i,j} - \bar{f}_j}{\bar{f}_j} \quad (3)$$

$$\hat{F}_j = \frac{\hat{f}_j - \bar{f}_j}{\bar{f}_j} \quad (4)$$

同様に、マイクロキスティス  $p_i$  を正規化し、 $P_i$  とする。

次に、観測データ  $\hat{F}_j$  と  $i$  番目の観測日の観測データ  $F_{i,j}$  の類似度  $S_i$  を式(5)で求める。

$$S_i = e^{-\sum_{j=1}^L \alpha_j |\hat{F}_j - F_{i,j}|} \quad (5)$$

類似度  $S_i$  は、0 から 1 の値をとり、1 に近いほどマイクロキスティスを推定する観測日の観測データ項目  $\hat{F}_j$  と  $i$  番目の観測日の観測データ項目  $F_{i,j}$  が類似していることを示している。 $\alpha_j$  は、正定数である。マイクロキスティスの値を推定する場合、すべての類似度を使うと精度が落ちると考えられる。そのため、類似度  $S_i$  が大きいものから  $L$  個抽出し、そのとき  $i$  の集合を  $U$  とする。推定するマイクロキスティス  $\hat{p}$  を式(6)で求める。

$$\hat{p} = \frac{\sum_{i \in U} P_i S_i}{\sum_{i \in U} S_i} \quad (6)$$

最後に、正規化された値のマイクロキスティス  $\hat{p}$  をマイクロキスティス  $p_i$  と同じ単位の値に変換する。

実験の結果、マイクロキスティスの観測値と推定値に大きな隔たりがあり、類似度を用いた方法では良好な結果が得られないことがわかった。

## 2.4 ミクロキスティスの発生予測と発生要因分析

観測された生の水質データと、特徴量エンジニアリングで得られたデータから、予測モデルの構築と発生要因分析を行う。水質データは、予測対象日の3ヶ月前から12ヶ月前までの観測値である。特徴量エンジニアリングによるデータは、予測対象日の3ヶ月前から12ヶ月前までの1年間の水質の統計量である。これらのデータから、マイクロキスティスの発生と関係性の高い特徴集合を、特徴選択アルゴリズムによって獲得する。獲得した特徴集合を入力とした機械学習モデルによって、マイクロキスティスの発生予測を行う。また、選ばれた特徴集合から、発生要因分析を行う。

特徴量エンジニアリングと特徴選択アルゴリズムを適用した結果、テストデータにおける accuracy が 0.950、F 値が 0.863 であった。これは、ほとんどの従来手法で行われている生データのみ解析と比べて、accuracy が 0.0108point、F 値が 0.691point と高い値である。このことから、マイクロキスティスの発生予測に対する特徴量エンジニアリングと特徴選択アルゴリズムの有効性が確認された。

## 3. まとめ

本研究では、大きく三つの研究を行った。一つ目がサポートベクトル回帰分析 (SVR) と

ランダムフォレストを用いてマイクロキスティスを推定する研究，二つ目が類似度を用いてマイクロキスティスを推定する研究，三つ目がマイクロキスティスの発生予測と発生要因を分析する研究である．実験の結果，水温，pH，渦鞭毛藻類，全リン，全窒素，硝酸，流入量の要因からマイクロキスティスを推定するにあたり，ランダムフォレストを用いる方法が有効であることがわかった．また，マイクロキスティスの発生予測において，特徴量エンジニアリングと特徴選択アルゴリズムの有効性が確認できた．今後の課題は，2021年以降の観測データの数を増やして実験を行うことである．

## 参考文献

- [1] 梅田信，古里栄一，浅枝隆，“富栄養化したダム湖におけるアオコ発生指標として水温成層安定性”，ダム工学，Vol.16, No.4, pp.269-281, 2006.
- [2] 増木新吾，矢島啓，管原庄吾，“アオコ発生時のダム湖における集中観測から得られた水質の変動と栄養塩の関係”，土木学会論文集 B1(水工学)，Vol.74, No.4, pp.I\_367-I\_372, 2018.
- [3] 近藤竜二，鍵谷豪，幸保孝，廣石伸互，渡辺眞之，“アオコの分類”，日本水産学会誌，Vol.64, No.2, pp.291-292, 1998.
- [4] 佐藤紗知子，大城等，馬庭章，管原庄吾，神谷宏，大谷修司，“宍道湖におけるアオコ発生の環境要因とその事前判別”，陸水学雑誌，Vol.76, No.3, pp.217-223, 2015.
- [5] 芹沢浩，雨宮隆，伊藤公紀，“相模湖と津久井湖におけるアオコ異常発生現象の数理モデル解析”，技術マネジメント研究，No.9, pp.1-14, 2010.
- [6] 西村詩央里，平岡透，熊野稔，“九州地方における道の駅の類似度を用いた利用者数の推定”，産業応用工学会論文誌，Vol.8, No.2, pp.198-202, 2020.