

陸域から沿岸域へ流入する溶存有機物および鉄の変動要因を機械学習で解明

熱帯島嶼の河川を対象に定期的に水質調査を行い、その結果について機械学習手法を用いて解析すると、河川水中の溶存有機物量は季節変動を示すとともに、流域の土地利用や土壌特性の影響を強く受けること、特に低湿地土壌から供給される腐植物質様成分の寄与が大きいことが分かる。また、腐植物質様成分の量が多いほど、植物プランクトンや海藻類等の必須栄養素である鉄の濃度が増加することが分かる。

キーワード：機械学習、鉄、熱帯島嶼、溶存有機物、流域

背景・ねらい

熱帯島嶼の周辺には、サンゴ礁など生物多様性が高く、観光・漁業資源としても重要な沿岸生態系が発達しているが、それらは陸域からの物質負荷の影響を強く受ける。たとえば有機物は、紫外線の吸収や微生物へのエネルギーおよび養分の供給など、海洋生態系の維持に重要であるが、その過剰な流入は光の透過性を低下させ、植物プランクトンや海藻類による光合成を阻害する。また、鉄は海藻類の生育に重要な成分の一つであるほか、サンゴの高温ストレス耐性の維持に寄与していることも報告されている。一方、溶存態の有機物(DOM)は鉄と結合することにより、元来水に溶けにくい鉄の輸送・供給に大きく寄与している。

そこで、陸域から河川を介して流入する DOM の量や化学特性（分解性や鉄との結合能力など）ならびにそれらの変動要因について明らかにするために、熱帯島嶼の一つである石垣島（沖縄県）内の主要な 6 つの河川を対象に、定期的に水質調査を実施する。そして、DOM の濃度および主な構成成分に及ぼす季節および流域特性（土地利用や土壌特性など）の影響を解析する。これらの水質や流域特性のデータはしばしば非正規分布を示し、また両者の関係が非線形である場合も多いことから、解析にはこれらを含むデータセットにもそのまま適用でき、さらに各説明変数の重要度も評価できる、機械学習手法の一種であるランダムフォレスト(RF)を用いる。さらに、DOMの成分組成が河川水中の溶存鉄(DFe)濃度に及ぼす影響についても同様に解析する。

成果の内容・特徴

1. 水質調査は、計 22 地点において 2 ヶ月に 1 回の頻度で 2 年間実施する。DOM の全濃度として溶存有機態炭素(DOC)濃度を定量する。また、3 次元蛍光スペクトル(EEM)の測定データについて多変量解析の一種である parallel factor analysis (PARAFAC)を行うことにより、DOM の主要構成成分を同定する。
2. DOC および EEM-PARAFAC の各成分について、RF を用いて流域特性（各土地利用および土壌タイプの面積割合、人口密度）および水温を説明変数とするモ

デルを作成する。DOC および 3 つの腐植物質様成分(C1, C2, C3)は、予測値の実測値に対する適合度を示す Nash-Sutcliffe efficiency (NSE)が学習データに対して 0.88 ~ 0.91、検証データに対して 0.71 ~ 0.77 となり、モデルの予測精度は高いと評価できる。

3. 各説明変数の重要度から、DOC、各腐植物質様成分のいずれに対しても、水温および強グライ土（排水不良の低湿地土壌）の影響が最も大きいことが分かる（表 1）。さらに、partial dependence plot（図 1）から、水温が 26℃ 以上で DOM の量が大きく増加すること、ならびに強グライ土の面積割合が 12%以上で腐植物質様成分の量が特に大きく、同土壌から供給されるこれらの有機物の寄与が大きいことが分かる。
4. DFe 濃度について、EEM-PARAFAC の各成分および鉄の化学形態や DOM との結合に関連する水質因子を説明変数とする RF モデルを作成する。NSE が学習データに対して 0.92、検証データに対して 0.62 と高いことから、DFe 濃度は DOM の成分組成や関連する水質因子の影響を強く受けることが分かる。中でも、腐植物質様成分(C1)の重要度が大きく、これらの量が多いほど DFe 濃度が増加する（図 2）。

成果の活用面・留意点

1. 本成果は、令和 4 年度国際農林水産業研究成果情報「河川水中の栄養塩類濃度は土地利用などの流域特性から機械学習で予測できる」と併せて、陸域からの物質負荷の実態把握に活用できる。
2. 説明変数とした土地利用と土壌タイプとの間に相互依存関係が存在する場合があります。変数の重要度を評価する際はその影響について検討する必要がある。

その他

予算区分：交付金プロ [A5 熱帯島嶼環境保全（第 5 期）、A3 アジア・島嶼資源管理（第 4 期）]

研究実施期間：2020～2023 年度

研究担当者：菊地哲郎（生産環境・畜産領域）、安西俊彦（熱帯・島嶼研究拠点）

発表論文等：Kikuchi et al. (2023) *Sci. Total Environ.* 904: 166892. <https://doi.org/10.1016/j.scitotenv.2023.166892>

表 1 溶存有機態炭素(DOC)濃度および腐植物質様成分(C1, C2, C3)の各ランダムフォレスト(RF)モデルにおける各説明変数の重要度（当該変数の値をランダムに入れ替えたときの予測値の平均平方誤差の増加率(%)）

説明変数		DOC	C1	C2	C3
土地 利用	水温	16.1	20.8	21.3	17.5
	農地 ^a	10.8	9.1	11.2	10.0
	牧草地	11.2	8.8	11.5	9.9
	水田	9.7	7.5	9.3	8.1
	森林	14.1	11.1	13.2	11.9
土壌	畜舎	7.0	8.0	7.2	8.2
	赤色土	11.9	11.1	11.3	10.9
	適潤性黄色土 ^b	12.3	9.9	12.4	10.2
	その他黄色土	12.7	8.9	8.3	8.6
	細粒暗赤色土	8.1	7.0	6.3	6.1
	その他暗赤色土	7.7	6.7	6.2	5.8
	強グライ土	16.1	22.2	16.8	21.4
人口密度		11.7	9.6	8.9	8.6

値が大きいほど重要度が大きいと判断される。15%より大きい場合白太字で表示。
 C1：陸域由来有機物の光分解により生成
 C2：有機物の微生物分解により生成
 C3：特にリグニンの分解産物に類似
^a 水田を除く。^b 主に山林に分布

(%)
 ~ 25
 ~ 20
 ~ 15
 ~ 10

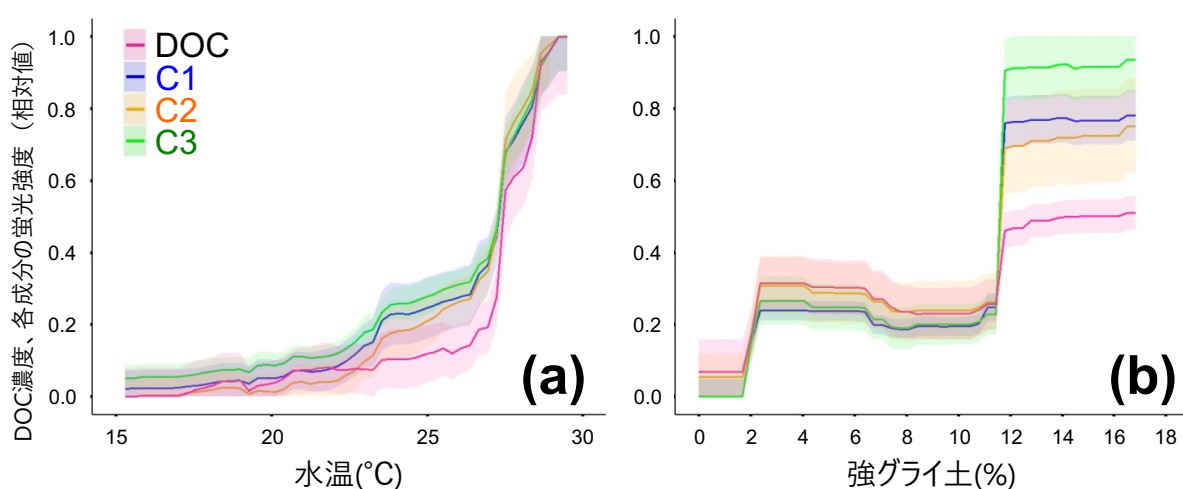


図 1 水温(a)および強グライ土の面積割合(b)に対する DOC 濃度ならびに腐植物質様成分(C1, C2, C3)の partial dependence plot (PDP)

全データセットをランダムに 5 等分し、そのうちの 4 つを統合したものを学習データとする。計 5 通りの組み合わせによる学習データのそれぞれについて RF モデルを作成する。この操作を 3 回繰り返す。実線は、得られた計 15 個の RF モデルによる計算値の平均値、網掛けは同標準偏差。比較のため、縦軸は平均値の最小値～最大値の範囲（a, b 共通）を 0～1 で標準化している。

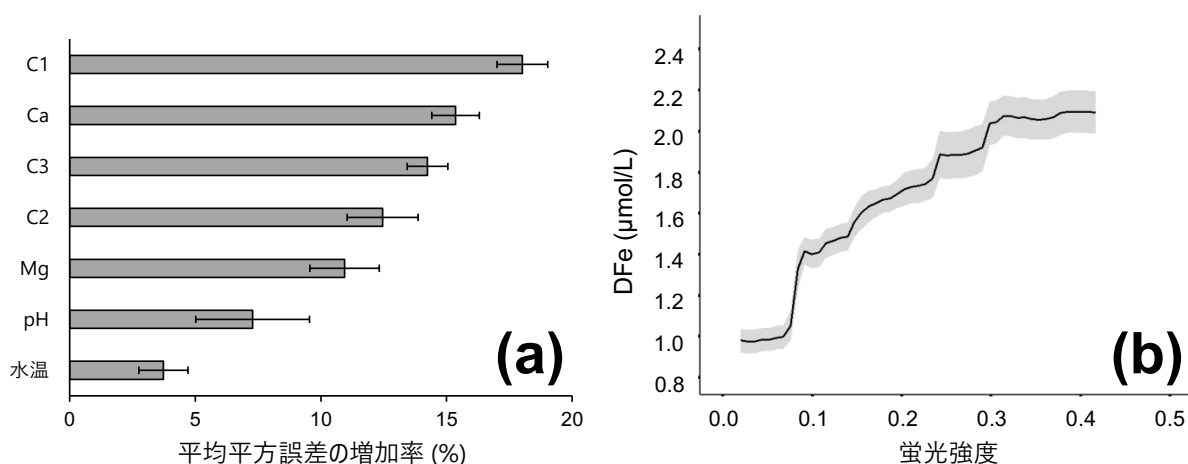


図 2 (a) 溶存鉄(DFe)濃度の RF モデルにおける各説明変数の重要度。(b) 腐植物質様成分 C1 に対する DFe 濃度の PDP

(a) 計 15 個の RF モデルの平均値（エラーバーは標準偏差）。Ca：カルシウムイオン濃度、Mg：マグネシウムイオン濃度。

(b) 実線は計 15 個の RF モデルによる計算値の平均値、網掛けは同標準偏差。

図は Kikuchi et al. (2023) © 2023 Elsevier B.V. より転載/改変して作成（転載/改変許諾済）