

## 熱帯島嶼河川の栄養塩濃度を機械学習で予測 —沿岸生態系の保全対策への活用に期待—

### ポイント

- 機械学習を用いて、石垣島主要河川の栄養塩濃度を高精度で予測
- 栄養塩濃度に及ぼす土地利用などの流域<sup>1)</sup>特性の影響も評価可能
- サンゴ礁の衰退や赤潮発生リスク評価への活用など、沿岸生態系の保全に向けた施策立案への活用が期待

### 概要

国際農研は、機械学習を用いて、沖縄県石垣島の河川水に含まれる栄養塩（窒素、リン、ケイ素）濃度を高い精度で予測する新たなモデルを作成しました。

本研究で作成したモデルは、石垣島の上流域<sup>2)</sup>の土地利用や表層地層などの流域特性から栄養塩濃度を機械学習手法「ランダムフォレスト」<sup>3)</sup>により予測するものであり、それぞれの栄養塩に対してどのような流域特性が大きな影響を及ぼしているかについても評価することができます。また、従来のシミュレーションモデル<sup>4)</sup>の多くは、窒素とリンを対象に栄養塩濃度を推定していましたが、本モデルはこれらに加えてケイ素についても濃度の予測が可能であり、窒素、リンの過剰負荷に起因するサンゴ礁の衰退やケイ素の減少に起因する渦鞭毛藻（うずべんもうそう）など有害藻類発生リスク評価などへの活用も可能になります。

さらに、陸域から流出する栄養塩の量を推定・予測するためには、河川流量と水質の観測やシミュレーションモデルなどの高度な専門知識・技術の習得ならびに労力が必要でしたが、本モデルでは、流域の土地利用や表層地質、人口密度のデータがあれば比較的簡便な操作で栄養塩濃度を予測することが可能になります。

熱帯島嶼では、気候変動に伴う海洋環境の変化に加えて、陸域からの過剰な栄養塩の流入による、サンゴ礁などの沿岸生態系への悪影響が懸念されています。本研究の成果によって、陸域から流入する栄養塩の量を適正に管理し、健全な沿岸生態系を保全するための施策立案への活用が期待されます。

本研究成果は、「Environmental Pollution」電子版（日本時間 2022 年 11 月 5 日）に掲載されました。

### <関連情報>

予算：運営費交付金プロジェクト「熱帯島嶼における山・里・海連環による環境保全技術の開発」

## 発表論文

---

<論文著者> T Kikuchi, T Anzai, T Ouchi, K Okamoto and Y Terajima

<論文タイトル> Assessing the impact of watershed characteristics and management on nutrient concentrations in tropical rivers using a machine learning method

<雑誌> Environmental Pollution

DOI : <https://doi.org/10.1016/j.envpol.2022.120599>

## 問い合わせ先など

---

国際農研（茨城県つくば市）理事長 小山 修

研究推進責任者：国際農研 プログラムディレクター 林 慶一

研究担当者：国際農研 生産環境・畜産領域 菊地 哲郎

国際農研 熱帯・島嶼研究拠点 安西 俊彦

広報担当者：国際農研 情報広報室長 大森 圭祐

Tel : 029-838-6708 FAX : 029-838-6337

プレス用 e-mail : [koho-jircas@ml.affrc.go.jp](mailto:koho-jircas@ml.affrc.go.jp)

本資料は、農政クラブ、農林記者会、農業技術クラブ、筑波研究学園都市記者会、八重山記者クラブに配付しています。

※国際農研（こくさいのうけん）は、国立研究開発法人 国際農林水産業研究センターのコミュニケーションネームです。新聞、TV等の報道でも当センターの名称としては「国際農研」のご使用をお願い申し上げます。

## 開発の社会的背景

熱帯島嶼の沿岸生態系は、生物多様性に富み、漁業や観光業にとって重要な水産資源です。一方、気候変動に伴う海水温の上昇や海洋酸性化に加え、農業生産や土地開発に起因する陸域からの栄養塩（特に窒素、リン）の過剰な流入は、植物プランクトンの大量増殖を引き起こし、サンゴを食害するオニヒトデの大量発生などにより、サンゴ礁衰退の要因となっていると考えられています。

栄養塩の主な発生源として、下水処理場や工場など、特定の場所から流れ込む「点源」と、森林や農地などから地表面、地下水を経由して流れ込む「面源」の2つのタイプがあります。「点源」は排水処理対策などにより改善されやすい一方で、「面源」からの栄養塩の輸送経路は複雑であることから、シミュレーションモデルによる推定が一般的ですが、高度な専門知識・技術の習得が課題でした。

## 研究の経緯

国際農研では、熱帯島嶼での農業生産と環境保全の両立を目指し、山林からの土壌の流出を抑える技術開発に取り組んでいます。また、これらの技術導入によって、山林や農地からの土砂や栄養塩の流出がどの程度抑えられるのかを評価するため、河川調査やモデルによる解析を行っています。

栄養塩の流出量推定を行うシミュレーションモデルの多くは、窒素とリンを対象にしており、ケイ素は扱われていませんでした。主な植物プランクトンである珪藻の生育にはケイ素も必要ですが、ケイ素に対して窒素やリンが過剰に供給されると、珪藻以外の植物プランクトンが増え、特に赤潮の原因となる渦鞭毛藻などの有害藻類の増殖に繋がる可能性も指摘されています。

そこで本研究では、沖縄県石垣島の複数の河川を対象に、ケイ素も加味した水中の栄養塩濃度を測定し、「面源」となる流域の土地利用や地質などの特徴（以下、流域特性）との関係を調べることで、栄養塩濃度に対してどのような影響を及ぼしているのか、機械学習手法の一つであるランダムフォレスト（RF）を用いて評価しました。

## 研究の内容・意義

1. 石垣島の主要な6河川（支流を含む）、計18地点（図1）において、令和2年9月から約1年間、約2ヶ月に1回の頻度で平水時<sup>5)</sup>に調査を行い、溶存無機態窒素（DIN）<sup>6)</sup>、全リン（TP）<sup>7)</sup>および溶存態ケイ素（DSi）の濃度を測定しました。
2. 流域特性として、18地点上流域の①土地利用（サトウキビ畑などの農地、牧草地、森林など）の面積割合、②表層地質（琉球石灰岩、国頭礫層など）の面積割合、③人口密度を、GISデータと統計資料により把握しました。
3. 河川調査の結果をもとに、RFを用いて、流域特性から各栄養塩濃度を予測するモデルを作成しました。季節による栄養塩濃度の変動も考慮して、「その年の1月1日から調査日までの経過日数」も栄養塩濃度を予測するパラメータに加えました。
4. データセットをランダムに5つのサブデータセットに等分し、そのうちの4つを統合したものを学習データ、残りの1つを検証データとして、計15個（5通りのサブデータセットの組み合わせ×3反復）のRFモデルを検証しました。その結果、いずれのモデルも実測値を高い精度で予測できることが示され、栄養塩濃度の変動の大部分を流域特性によって説明できることを明らかにしました。

(DIN について図 2 に表示)。河川水中の栄養塩濃度は流域特性、すなわち、流域内の土地利用や地質の影響を強く受けていることを示しています。

5. RF モデルにおける各流域特性の重要度（流域特性の値のみをランダムに入れ替えた場合の栄養塩濃度の予測値の平均平方誤差の増加率）、栄養塩濃度と各流域特性との相関性から、サトウキビ畑や畜舎などの面積割合が多い流域では、DIN および TP 濃度が高いのに対し、広葉樹林の面積割合が多い流域では DSi 濃度が高い傾向でした（図 3）。石垣島ではサトウキビ栽培や肉牛の飼育が盛んであることから、これらの農業生産活動が河川の栄養塩濃度に強い影響を与えていることが考えられ、今後、環境負荷の具体的な要因や経路を把握する等、課題解決に向けて検討していくことが必要と思われま

## 今後の予定・期待

今回作成したRFモデルは、流域の土地利用や表層地質、人口密度のデータがあれば比較的簡便な操作で栄養塩濃度の予測値を算出できます。今後、降雨直後の出水時についてもRFモデルを作成することで、年間を通じた栄養塩濃度の予測が可能になります。さらに、国土交通省が提供する水文水質データベースなどの流量観測データと組み合わせることにより、流域からの栄養塩流出量を把握することができます。

作成したRFモデルは、窒素、リンに加えてケイ素の濃度も予測することができ、ケイ素の減少に起因する有害藻類発生リスクの評価などに活用されることが期待されます。今後、行政機関などによる水質や生態系の監視、さらに沿岸生態系保全対策への活用など、様々な地域への展開が期待されます。

## 用語の解説

- 1) 流域：河川や海に流れ込む雨水（雪や氷を含む）が降り集まる地域。「集水域」とも言います。
- 2) 上流域：調査河川の上流から調査地点までの間に雨水が流れ込む地域です。
- 3) ランダムフォレスト：2001年にLeo Breiman（米）によって提唱された機械学習アルゴリズムの一種です。データセットの中から重複を許して繰り返しデータを抽出し、それぞれを学習データとして決定木（樹形図によってデータを分析し、予測値を算出するもの）を作成し、各決定木の予測値の平均値を最終的な予測値とします。
- 4) シミュレーションモデル：ここでは、流域（地下部を含む）における水の移動、土壌や栄養塩など物質の移動、反応に関わる各プロセスを数式化し、それらの時間的・空間的な変化量を一元的に解くことにより、河川の流量や物質の流出量を計算するモデルのことを指します。
- 5) 平水時：降雨などによって、河川の流量や水質の顕著な変化が起きていない、平常の状態を言います。
- 6) 溶存無機態窒素（DIN）：Dissolved Inorganic Nitrogenの略で、硝酸態窒素（NO<sub>3</sub>-N）、亜硝酸態窒素（NO<sub>2</sub>-N）、アンモニア態窒素（NH<sub>4</sub>-N）の総称です。本研究におけるDIN濃度とは、これらの形態の窒素の総濃度のことを指します。なお、本研究では、河川水中の物質を孔隙の大きさが0.7 μm（DINの場合）または0.22 μm（DSiの場合）のフィルターでろ過し、そのろ過水中に溶けている成分を溶存態としました。

- 7) 全リン (TP) : Total Phosphorus の略です。水中に存在する全ての形態のリン (溶解態、土砂などの懸濁粒子に含まれる、あるいは吸着しているもの) を指します。

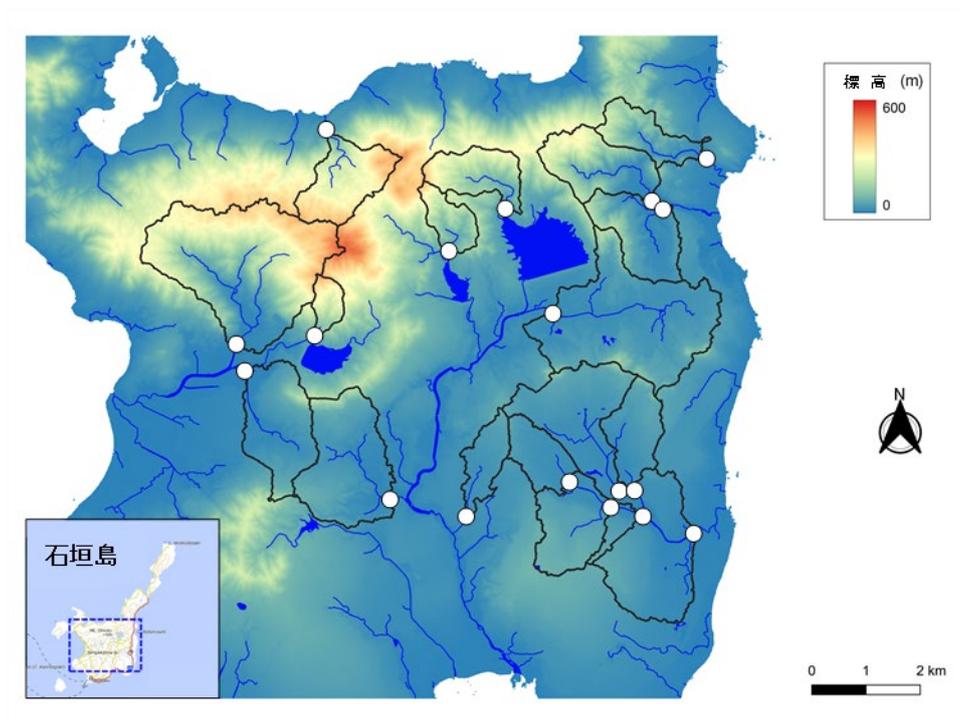


図 1. 調査対象地域

地図上の白色○印は調査地点 (18 箇所) を、黒の実線で囲まれた領域は各調査地点における河川流域をそれぞれ表します。

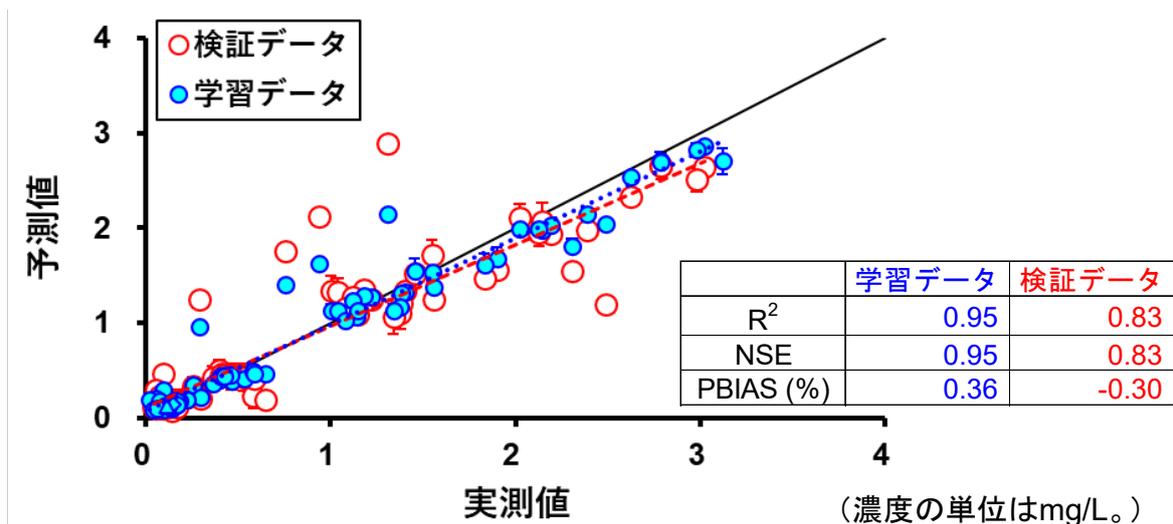


図 2. 無機態窒素 (DIN) の実測値と RF モデルによる予測値との関係

予測値は、15 個の RF モデルによる各予測値の平均値 (エラーバーは標準偏差) を表します。黒色線は実測値=予測値のライン、青色点線は学習データに対する回帰直線、赤色破線は検証データに対する回帰直線をそれぞれ表します。Moriyas et al. (2015) によるモデルの予測精度評価基準では、 $R^2$  が 0.70、NSE (Nash-Sutcliffe Efficiency) が 0.65 よりそれぞれ大きく、PBIAS (percent bias) が  $\pm 15\%$  より小さい場合、予測精度は「非常に良い」と評価できるとしています。

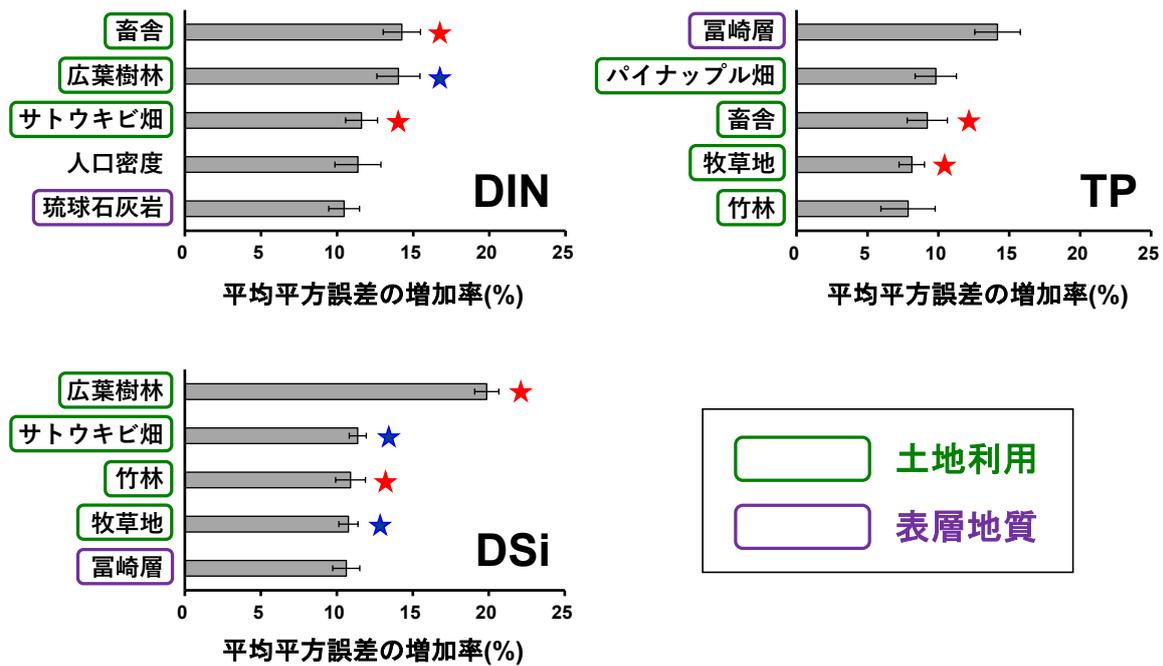


図 3. 栄養塩濃度に対する流域特性の重要度（上位 5 特性のみ表示）

重要度は、各流域特性の値のみをランダムに入れ替えた場合の、各栄養塩濃度の予測値の平均平方誤差の増加率（%）

- ★は、各栄養塩濃度と強い正の相関（相関係数が 0.7 以上）が認められる特性
- ★は、各栄養塩濃度と強い負の相関（相関係数が-0.7 以下）が認められる特性

琉球石灰岩：更新世（約 258 万年前～約 1 万年前）にサンゴ礁によって形成された石灰岩の地層。南西諸島に広く分布し、石垣島では主に南部に分布。

富崎層：中生代（約 2 億 5217 万年前～約 6600 万年前）に堆積した礫岩、チャート、砂岩、石灰岩などよりなる層。石垣島では主に南西部に分布。