

有明海における深層学習を活用した赤潮発生予測手法の検討 A Study on Red Tide Occurrence Prediction Method using Deep Learning in the Ariake Sea

○田畑俊範*, 杉江秀香**, 龔含遠***, 平松和昭*, 原田昌佳*

Toshinori Tabata, Hideka Sugie, Hanyen Gong, Kazuaki Hiramatsu, Masayoshi Harada

1. はじめに 有明海では陸域由来の栄養塩類を起源とする赤潮の発生が漁業被害を引き起こし、赤潮発生予測が喫緊の課題である。人工知能技術を援用した赤潮発生予測の研究が進められているが、その手法は未だ確立しておらず、現状ではデータ収集のコスト面や予測のタイミングという点で課題が残る。本研究では、階層型ニューラルネットワークモデル（HNN）および再帰型ニューラルネットワークモデル（RNN）に着目し、入手が容易なデータを利用した新たな赤潮発生予測手法の確立を目指した。

2. 構築モデル概要 赤潮は陸域由来の栄養塩類と海域環境が複雑に関係した非線形現象である。本研究では、非線形現象に対しても強いパターン認識能力を持つ HNN および RNN を活用した赤潮発生予測手法の確立を目指した。今回は、メッシュ型 HNN (M-HNN)、県別型 HNN (P-HNN)、県別型 RNN (P-RNN) の3つのモデルを構築した。

3. M-HNN まずは Fig.1 に示す有明海を含む流入河川の流域全体を 402 個の 5 km メッシュに分割し、海域メッシュごとの発生の有無を予測可能な HNN の構築を目指した。入力変数は、5つの土地利用面積（水田、畑地、山地、都市、水域）、4つの水文気象データ（日平均気温、日降水量、日照時間、日平均風速）の赤潮発生日前日までの14日積算値とした。なお、水文気象データは Fig.1 の AMeDAS 観測地点を Thiessen 法によりメッシュ毎に割り当てた。一方で出力変数は、陸域を除く 81 個の海域メッシュに、4つの藻類綱別（渦鞭毛藻、珪藻、ラフィド藻、その他）赤潮の発生の有無をメッシュ毎に [1, 0] で代入した。すなわち、入力層、中間層、出力層のニューロン数を、3618 個（ $=402 \times 9$ ）、1809 個、324 個（ $=81 \times 4$ ）とした。学習には 1997 年から 2019 年の間の全ての赤潮発生イベントのデータ（581 件）を用いた。学習した赤潮イベントの一例を Fig.2 に示す。発生領域を含め良好に出力できているが、再現できないイベントも多数確認された。Table 1 に本モデルによる学習結果を示す。なお、出力結果のうち 0.5 以上を発生、0.5 未満を非発生と判別した。また、評価指標として Precision, Recall, Accuracy についても計算した。Precision, Accuracy はそれぞれ 0.898, 0.985 と高い値を示したが、Recall が 0.514 と低いことが確認できた。漁業被害の対策という面から

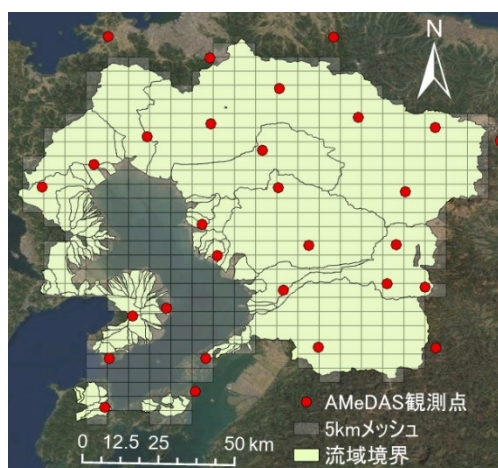


Fig.1 M-HNN のメッシュ分割
および AMeDAS 観測地点

*九州大学大学院農学研究院 / Faculty of Agriculture, Kyushu University **九州大学大学院生物資源環境科学府 / Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University ***東京大学大学院新領域創成科学研究科 / Graduate School of Frontier Sciences, the University of Tokyo

キーワード：水環境、有明海、赤潮、ニューラルネットワークモデル

Recall は重要な評価指標で、学習精度の向上が求められる。これはメッシュ型としたことで非発生 (= 0) となるニューロンが過剰となり、学習精度の低下に繋がったと考えられる。

4. P-HNN 構築した M-HNN の過剰なニューロン数を減らすため、入出力変数を選別したモデルを構築した。Fig.1 に示す AMeDAS 観測地点のうち有明海沿岸の 12 地点を選別し、その 4 つの水文気象データの予測日前日までの 21 日積算値を入力変数とし、有明海を沿岸 4 県（福岡県、佐賀県、熊本県、長崎県）で領域分けし、各領域での赤潮発生の有無を [1, 0] の出力変数とした。すなわち、入力層、中間層、出力層のニューロン数を、48 個 (=12×4)、24 個、4 個とする ANN を構築した。学習には 1997 年から 2019 年の間の全ての赤潮イベント発生期間の日別データ (4299 日) を用いた。Table 2 に本モデルによる学習結果を示す。Precision, Accuracy, Recall はそれぞれ 0.934, 0.914, 0.933 とすべて高い値を示し、学習精度が高いことが確認できた。

5. P-RNN P-HNN の構築では、水文気象データの過去の履歴を反映させるために積算値を使用した。それに代わり、HNN の中間層を連結させることで時系列データの学習に特化した RNN への改良を目指した。入出力変数は、P-ANN で使用したデータを積算値でなく時系列データとして入力し、さらに熊本港の最大潮差を加えた。すなわち、入力層、中間層、出力層のニューロン数をそれぞれ 49 個 (=12×4+1)、85 個、4 個とする RNN モデルを構築した。中間層の間の連結数（シーケンス長）は 5 と設定した。学習には 1997 年から 2020 年の間の全日数データ (8036 日) を用い、ここからランダムに 900 個のデータセットを作成し学習データとして用いた。Table 3 に本モデルによる学習結果を示す。Precision, Accuracy, Recall はそれぞれ 0.404, 0.424, 0.629 といずれも十分な再現性は示されなかった。これは、RNN へと改良したことで層が深くなったため、深層学習でしばしば問題となる勾配消失が発生したことが原因と考えられる。

6. おわりに 本研究では、赤潮発生予測手法の確立に向け HNN および RNN を活用した 3 つのモデルを構築した。今後の方針として、勾配消失問題の解消を可能とする RNN モデルを改良した長・短期記憶（Long Short-Term Memory）を用いることで、より効率的な学習を目指す。さらには、入力変数として衛星画像を利用した畳み込みニューラルネットワークモデルも導入することで、学習精度を上げていく。

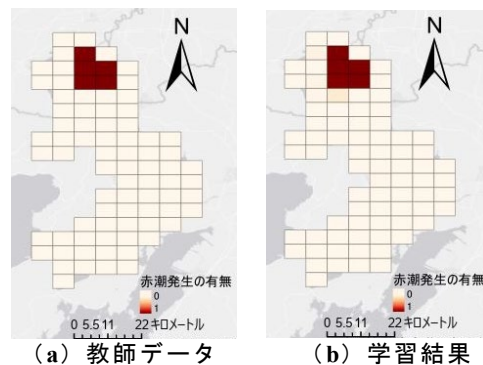


Fig.2 M-HNN で良好に学習できた赤潮イベント

Table 1 M-HNN の学習結果

		出力結果	
		発生 (≥0.5)	非発生 (<0.5)
教師 データ	発生 (= 1)	2751 (TP)	2602 (FN)
	非発生 (= 0)	312 (FP)	182579 (TN)
Precision :		0.898	Recall : 0.514
Accuracy :		0.985	

Table 2 P-HNN の学習結果

		出力結果	
		発生 (≥0.5)	非発生 (<0.5)
教師 データ	発生 (= 1)	6973 (TP)	660 (FN)
	非発生 (= 0)	495 (FP)	9,068 (TN)
Precision :		0.934	Recall : 0.914
Accuracy :		0.933	

Table 3 P-RNN の学習結果

		出力結果	
		発生 (≥0.5)	非発生 (<0.5)
教師 データ	発生 (= 1)	4197 (TP)	5708 (FN)
	非発生 (= 0)	6203 (FP)	16018 (TN)
Precision :		0.404	Recall : 0.424
Accuracy :		0.629	