

雲被り衛星画像に対するクロロフィル a 濃度推定のための U-Net モデルの検討

A Study of U-Net-Based Chlorophyll-a Estimation from Cloud-Contaminated Satellite Images

徳永 穂乃香

松岡 諒

Honoka Tokunaga

Ryo Matsuoka

北九州市立大学

The University of Kitakyushu

1 はじめに

クロロフィル a (Chl-a) は、水質や藻類の分布を示す重要な生物指標であり、海洋や湖沼環境のモニタリングにおいて不可欠なパラメータである。近年、衛星リモートセンシング技術の進展により、広域かつ高頻度な Chl-a 濃度推定が可能となり、赤潮の早期検出や環境保全への応用が期待されている。

従来の Chl-a 濃度推定には、可視および近赤外域の複数のスペクトルバンドを用いた経験的な回帰モデル (例: OCx 系列) が広く用いられてきた [1]。しかし、光学センサによる観測は雲の影響を強く受けやすく、雲被り領域ではスペクトル情報が欠損するため、Chl-a の空間的な連続性や推定精度の低下が課題となっている。

この課題を解決する手法として、近年では深層学習による画像補完技術が注目されており、U-Net や GAN といったアーキテクチャを用いた手法が提案されている [2, 3, 4]。これらの多くは、視覚的な自然さや構造的整合性に着目して補完を行うが、再構成されたスペクトルが物理的に妥当かどうか、また Chl-a 濃度推定への影響については十分な検証がなされていない。

一方で、雲の影響を受けにくい合成開口レーダ (SAR) や、条件によって取得可能な熱赤外 (TIR) ・地表面温度 (LST) など、異なる観測原理をもつ複数のセンサデータを組み合わせたマルチモーダルな情報統合により、欠損した光学情報を補完し、より安定した推定を実現できる可能性がある。本研究では、このようなマルチモーダルデータの活用によって、雲被り下でも物理的整合性を保った Chl-a 濃度推定が可能になると考え、その実現に向けたモデル構築を行う。

具体的には、雲により欠損のある Sentinel-2 の光学バンドを対象とし、Sentinel-1 の SAR や Landsat に由

来する LST といったマルチモーダルセンサデータを統合的に活用することで、Chl-a 濃度推定の精度を維持・向上させることを目的とする。深層学習モデルとしては、空間的特徴を保持しながら高精度な補完が可能な U-Net モデルを採用する。本手法により、雲被り領域においても物理的に整合性のあるスペクトルを再構成し、実環境の状態をより正確に反映した Chl-a 濃度推定を実現することを目指す。

2 関連研究

海洋における Chl-a 濃度の推定に関しては、従来よりさまざまな手法が提案されてきた。代表的なものとしては、衛星の可視バンドを用いた経験的な回帰モデルである OCx (Ocean Color x) 系列があり、広く利用されている。しかし、この手法は大気や雲の影響を強く受けやすく、特に雲被り時における推定精度の低下が課題となっている。

近年では、深層学習を用いたアプローチが注目されており、とりわけ画像のセグメンテーションや再構成に優れる U-Net アーキテクチャは、衛星画像の補完や回帰タスクへの応用でも成果を上げている。本研究では、衛星画像に含まれる雲の影響を軽減しつつ、Chl-a 濃度を高精度に推定するために、U-Net に基づく深層学習モデルを構築する。OCx モデルは、海洋表面の反射率から Chl-a 濃度を推定するための経験的な多項式回帰モデルであり、NASA の SeaWiFS や MODIS などのセンサデータに基づいて開発されてきた [1]。特に、OC3 や OC4 といったモデルが良く用いられており、可視光バンド (例: 青・緑・赤) の反射率比の対数変換を用いて濃度を推定する。

OCx モデルは、計算が簡便で処理速度が速いという利点を持つ一方で、雲や大気の影響によるデータ欠損に弱く、センサ間の特性の違いや空間的な文脈情報を考慮できないという制限がある。また、局所的な変動を反映しにくい点も課題である。

3 提案手法

本章では、提案する雲除去および Chl-a 濃度推定手法について述べる。

3.1 雲マスクの推定

本研究では、雲の影響を受けた領域を特定するため、Sentinel-2 衛星画像の RGB バンドに基づく簡易な手法を用いて雲マスクを推定する。具体的には、各バンドの画素値を正規化した上で、各画素における平均輝度 $A(x, y) = \frac{R(x, y) + G(x, y) + B(x, y)}{3}$ を計算する。平均輝度が所定の閾値 τ 以上の画素を雲と判定し、以下のように雲マスク $M(x, y)$ を生成する。

$$M(x, y) = \begin{cases} 0 & \text{if } A(x, y) \geq \tau \\ 1 & \text{otherwise} \end{cases} \quad (1)$$

得られたマスクにおいて、値が 0 の領域を雲と判定し、学習および推論時の処理における欠損領域として扱う。

3.2 U-Net モデルによる学習

提案手法では、雲除去と Chl-a 濃度推定を同時に実現するために、エンコーダ-デコーダ構造を有する U-Net モデルを採用する。モデルへの入力には、以下の複数のモダリティから構成される衛星データをチャンネル方向に連結したテンソルを用いる。すなわち、Sentinel-2 のマルチスペクトルバンド、Landsat 由来の地表面温度 (LST) 画像、Sentinel-1 による SAR 画像 (VV/VH 偏波)、および対応する雲マスクである。

本モデルは、雲が除去された多バンド再構成画像と、Chl-a 濃度マップの 2 種類の出力を同時に生成する。学習においては、入力画像を一定サイズのパッチに分割し、それぞれのパッチに対して雲の影響を受けていない有効画素のみに基づいて損失を計算することで、観測データに準拠した教師あり学習を実現する。

ここで、Chl-a 濃度の教師データには、Sentinel-3 の海洋観測ミッションから提供される Chl-a 濃度プロダクト (OC4ME) を用いる。これは空間分解能こそやや

低いものの、広域かつ高信頼な Chl-a 濃度推定値を提供することから、本研究では教師データとして適していると判断した。

提案手法では、残差ブロックとデュアル出力を備えた U-Net ベースのエンコーダ-デコーダ構造を採用している。エンコーダは、3 段の畳み込み層、インスタンス正規化、Leaky ReLU 活性化、および残差ブロックで構成され、各段の後にマックスプーリングが適用される。ボトルネックでは高次特徴を抽出するための残差ブロックを用いる。デコーダでは、転置畳み込みによるアップサンプリングとエンコーダからのスキップ接続により特徴を統合し、残差ブロックで精緻化を行う。最終的に、13 チャンネルの MSI 再構成出力と Chl-a 濃度予測の 1 チャンネル出力という 2 つの出力を同時に得る構成である。残差ブロックには汎化性能向上のためのドロップアウトも導入している。

提案する U-Net モデルを次式のように定義する。

$$\{\hat{X}, \hat{Y}\} = \text{U-Net}_{\theta}([X, M]) \quad (2)$$

ここで、 $[\cdot]$ はチャンネル方向の結合を表す。U-Net モデルは、入力多バンド画像 X とその対応マスク M (二値マスク、1 が有効画素) の各パッチを入力としその再構成 \hat{X} と Chl-a 濃度画像 \hat{Y} を出力する。

このモデルのパラメータ θ は、次式で定義する損失関数を最小化することで求める。

$$\mathcal{L}_{\theta} = \lambda_{\text{MS}} \text{MSE}_M(\hat{X}, X) + \lambda_{\text{Chl-a}} \text{MSE}_M(\hat{Y}, Y) \quad (3)$$

ここで、 Y は Sentinel-3 由来の参照 Chl-a 濃度画像である。 λ_{MS} および $\lambda_{\text{Chl-a}}$ は、それぞれの損失項に対するスケール係数である。

マスク付き平均二乗誤差 $\text{MSE}_M(\cdot)$ は、マスク M に基づき、各パッチ内の有効画素上でのみ誤差を計算し、以下のように定義される。

$$\text{MSE}_M(\hat{Z}, Z) = \frac{1}{\sum_{i,j} M_{i,j}} \sum_{i,j} M_{i,j} \cdot (\hat{Z}_{i,j} - Z_{i,j})^2 \quad (4)$$

ここで、 \hat{Z} および Z は同じ次元を持つ予測値と教師データを、 M は対応する雲マスクを表す。この定義により、雲のかかっていないパッチ内の画素のみを対象として損失が評価される。

3.3 反復的な画像再構成と濃度推定

提案手法では、入力画像を一定サイズのパッチに分割し、雲の有無を考慮しながら、画像の再構成を反復

的に実施する。雲領域が完全に補完された後、最終的な再構成画像を用いて Chl-a 濃度の推定を行う。

ステップ t において、前回の再構成結果 $\hat{X}^{(t)}$ と雲マスク $M^{(t)}$ を用いて、U-Net モデルによる補完処理を次のように定義する。

$$\hat{X}^{(t+1)} = \text{U-Net}_{\theta}([\hat{X}^{(t)}, M^{(t)}]) \quad (5)$$

ここで、 $[\cdot]$ はチャンネル方向の結合を表す。U-Net はパッチ単位で動作するため、各パッチ出力をスライディングウィンドウにより重ね合わせ、全体画像として統合した再構成結果を得る。

続いて、統合された再構成画像に基づいて雲領域を再判定し、次のステップで用いる新たな雲マスク $M^{(t+1)}$ を更新する。

パッチの処理は、雲の占有率があらかじめ定めた閾値(例: 50%) 以下であるものに限定する。ただし、パッチを重ねてスライドさせながら処理することで、雲の少ない周辺領域を含むパッチがスキップされた領域の再構成にも間接的に寄与する。このようにして、雲の少ない領域から順に再構成とマスク更新を繰り返すことで、初期には補完が困難であった広範な雲被覆領域も、反復的な推定により段階的に再構成されていくことが期待される。反復処理は、雲マスク $M^{(t)}$ の全画素がゼロ(すなわち全域が雲除去された状態)になるまで継続される。

雲領域が完全に補完された再構成画像 $\hat{X}^{(T)}$ に対して、U-Net モデルを適用し Chl-a 濃度画像 \hat{Y} を得る。

4 実験

本研究では、人工的に雲を付加して実験を行うため、雲の影響が少ない晴天時の衛星画像を選定し、実験データとして使用した。衛星画像は、静岡県沿岸域(東経約 138.3°, 北緯約 34.8°)を対象として、2023 年 8 月 20 日付近に取得されたものである。

入力データとして、Sentinel-2 SR 画像の 10 バンド(B2~B4, B5~B8A, B11, B12)に加え、Landsat-8/9 の地表面温度(LST) データおよび Sentinel-1 SAR データ(VV 偏波, VH 偏波)を使用し、これら 13 バンドを 1 つの入力テンソルとして統合した。

Chl-a 濃度の教師データは Sentinel-3B 衛星から取得したものであり、欧州宇宙機関(ESA)のコペルニクスプログラムのデータポータルよりダウンロードし、Sentinel-2 画像と同一日時・同一座標のものを使用した。

Chl-a 濃度の教師データを除く全てのデータは、Google Earth Engine 上で雲マスク、正規化、空間的リ

スケーリングの処理を行い、空間解像度を 10 m に統一して使用した。一方、教師データは、Sentinel-2 と空間的に整合するように投影系を変換し、補間処理を行った。すべての画像データは、上述した経度緯度を中心とし上下左右に 500 ピクセルずつ切り出し、1000×1000 ピクセルの領域を使用した。

各バンドに対して以下の正規化を行った。

- Sentinel-2 の反射率バンド: 14 ビット (0~16384) を基準に 0~1 の範囲へスケーリング
- SAR データ (VV, VH): dB 単位で -35~30 dB の範囲を 0~1 の範囲へスケーリング
- LST データ: -40 °C~40 °C を 0~1 の範囲へスケーリング

雲除去モデルの訓練および評価用データとして、晴天画像に人工的な雲を付加したデータを作成した。雲の生成には、ゼロ行列上にランダムな位置・大きさ・回転角を持つ複数の楕円を重ね合わせ、これにガウシアンフィルタを適用することで、現実の雲に類似した滑らかな雲形状を生成した。各楕円の長軸・短軸の半径は所定の範囲内で一様乱数により決定され、回転角も 0~180 度の範囲でランダムに設定された。さらに、生成されたプロブには輝度の高い値を加えることで、雲の輝度分布を再現した。図 2 に、原画像、作成した雲マスク、雲合成画像および推定した雲マスクを示す。

雲の面積は、楕円の数・大きさ・ぼかし処理を考慮したマスク画像上のピクセル比に基づいて確認を行い、画像全体の約 20~25% となるように調整した。また、RGB バンドから計算される平均輝度に基づき閾値処理を行い、雲マスクを生成して実験に使用した。本研究では、この雲マスクを用いて雲領域外のみ学習対象とするため、Sentinel-2 の画像にのみ人工的な雲を付与している。

U-Net モデルの入力は 64×64 ピクセルのパッチ単位であり、元画像(1000×1000 ピクセル)からハーフオーバーラップさせてパッチを抽出した。学習時には雲の影響がない領域のパッチのみを使用し、雲がかかっているパッチは除外した。

1 画像から得られるすべてのパッチを 1 エポックとし、100 エポック繰り返してモデルを学習した。最適化には Adam を用い、初期学習率は $1.0 \cdot 10^{-3}$ 、重み減衰は $1.0 \cdot 10^{-5}$ とした。

評価には、予測された Chl-a 濃度と実測値の誤差の大きさを示す指標として、二乗平均平方根誤差(RMSE)を用いた。RMSE は、推定値と実測値の差の 2 乗平均の平方根であり、値が小さいほど推定精度が高いことを意

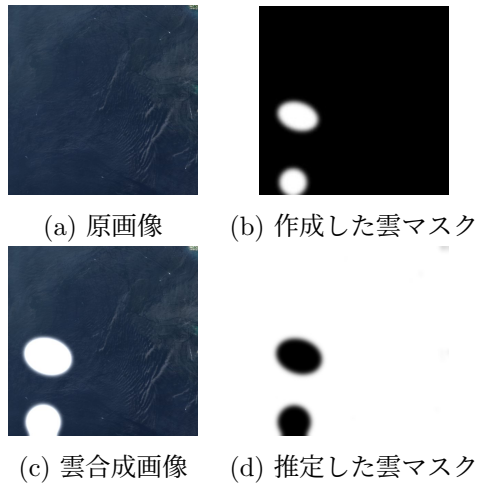


図 1: 衛星画像の原画像と雲マスクの例

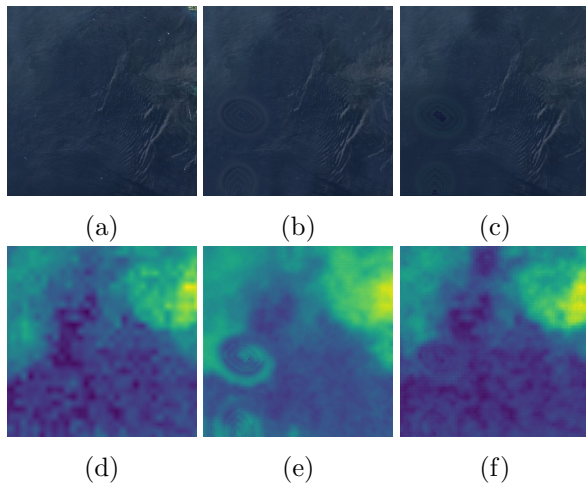


図 2: 雲除去および Chl-a 濃度推定の結果: (a, d) 原画像および Chl-a 濃度の正解データ, (b, e) SAR なしモデルによる再構成画像と推定結果, (c, f) SAR ありモデルによる再構成画像と推定結果

味する。表 1 に示すように、雲の有無および SAR データの有無によって領域を分けて評価を行った。表 1 から、SAR データを用いない構成と比較して、SAR データを併用した構成ではすべての領域において RMSE が低下していることが分かる。特に雲有り領域では、RMSE が約 43% 低減しており、SAR の活用が雲被り領域における推定精度向上に大きく寄与していることが示されている。また、雲の有無にかかわらず、SAR あり構成が一貫して最も低い誤差を示しており、本手法の有効性を裏付ける結果となっている。

図 2 に、提案する U-Net モデルによって推定されたマルチスペクトル画像 (RGB バンド) および Chl-a 濃度画像の結果を示す。ここでは、SAR データの有無による影響を検証するため、SAR を使用しない場合の結

表 1: Chl-a 濃度推定誤差 (RMSE)

モデル構成	雲無し領域	雲有り領域	画像全体
SAR 無し	0.0369	0.1196	0.0647
SAR 有り	0.0316	0.0676	0.0421

果も併せて比較している。これらの結果から、提案手法は雲被覆下においても高精度に Chl-a 濃度を推定できることが確認された。特に SAR データを併用することで欠損領域の補完精度が向上し、推定誤差 (RMSE) の低減に加えて、空間的に自然な分布の再現が可能となった。これにより、雲の影響下においても高い信頼性を維持した推定が可能となり、広域的な水質モニタリングへの応用において有効であると考えられる。

5 むすび

本研究では、雲被りのある光学衛星画像においても高精度な Chl-a 濃度推定を実現する手法を提案した。U-Net を用いた雲領域の逐次的な補完と再構成に加え、SAR および LST データを統合することで、雲の影響を効果的に緩和しつつ推定精度の向上を図った。実験により、特に SAR データの活用が雲被り領域での精度向上に有効であることが示された。

謝辞

本研究の一部は、JSPS 科研費 24K15075 の助成を受けた。

参考文献

- [1] John E O'Reilly, Stéphane Maritorena, B Greg Mitchell, David A Siegel, Kendall L Carder, Sara A Garver, Mati Kahru, and Charles McClain. Ocean color chlorophyll algorithms for seawifs. *J. Geophys. Res. Oceans*, 103(C11):24937–24953, 1998.
- [2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Proc. MICCAI*, pages 234–241. Springer, 2015.
- [3] Jiahao Zheng, Xiao-Yang Liu, and Xiaodong Wang. Single image cloud removal using u-net and generative adversarial networks. *IEEE Trans. Geosci. Remote Sens.*, 59(8):6371–6385, 2020.
- [4] Yue Zi, Fengying Xie, Ning Zhang, Zhiguo Jiang, Wentao Zhu, and Haopeng Zhang. Thin cloud removal for multispectral remote sensing images using convolutional neural networks combined with an imaging model. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 14:3811–3823, 2021.