

複数の分光情報と深層学習の融合による 土壌理化学性の高精度推定

齋藤 嘉人

新潟大学自然科学系農学部 助教 (〒950-2181 新潟市西区五十嵐2の町 8050 番地)

E-mail: ysaito@agr.niigata-u.ac.jp

報告概要

本研究では、土壌水抽出液の励起蛍光マトリクス (EEM) を入力とした、土壌の全炭素 (TC)、全窒素 (TN)、水溶性全有機体炭素 (WS-TOC)、水溶性全窒素 (WS-TN)、および発生 CO₂ 量の推定モデル構築を目的とした。新潟大学附属農場にて採取した計 104 点の土壌を供試し、モデルの評価指標として、テストデータにおける平均二乗平方根誤差 (RMSE_{CV})、決定係数 (R^2_{CV}) および予測偏差比 (RPD_{CV}) を用いた。畳み込みニューラルネットワーク (CNN) と部分的最小二乗回帰 (PLSR) を比較した結果、CNN が優位性を示し、TC ($R^2_{CV} = 0.95$)、TN ($R^2_{CV} = 0.93$)、WS-TOC ($R^2_{CV} = 0.86$)、WS-TN ($R^2_{CV} = 0.91$) で高精度な推定が可能であった。以上の結果から、蛍光分光法と深層学習による土壌中の炭素・窒素関連成分の推定の可能性が示唆された。

1. 背景

温室効果ガス (greenhouse gas, GHG) の発生量はここ数十年で急速に増加している。人為的な活動により、主要な GHG である CO₂、CH₄、N₂O をはじめとする様々な GHG の大気中の濃度上昇を引き起こしている。農業も例外ではなく、純 GHG 排出量のうち、農業・林業・その他土地利用が 22%を占めていると報告されている¹⁾。2022 年度の環境省の統計によると、日本では CO₂、CH₄、N₂O の発生量のうちそれぞれ 4%、82%、50%を農業が占める。特に CH₄ および N₂O は地球温暖化係数がそれぞれ 27.0、273 と報告されており²⁾、農業由来の GHG は環境に対して大きな影響を与えている。GHG による環境汚染を軽減するために、GHG 排出量およびその要因となる土壌中の炭素、窒素量の正確かつ迅速な計測技術が求められている。

2023 年、日本において農業由来で発生した CH₄ のうち約 44%を稲作が、N₂O のうち約 27%が農用地の土壌が占めていると報告されている³⁾。土壌から発生する GHG の量を推定する手法として、土壌表面からのガス放出をモニタリングするチャンバー法が一般的であるが、その方法によっては、ガス濃度の上昇による非線形フラックスを示すことや、人工的な環境が必要であるといった課題が挙げられる⁴⁾。一方、GHG の発生源は土壌中の炭素量および窒素量に関係している^{5), 6)} ことに着目すると、土壌から発生する GHG 量の推定には土壌中の炭素や窒素などの成分を推定することが有用であると考えられる。また、土壌に含まれる水溶性全有機体炭素 (water-soluble total organic carbon, WS-TOC) および水溶性全窒素 (water-soluble total nitrogen, WS-TN) も土壌の様々なプロセスに対して重要な指標である。

全炭素 (total carbon, TC) および全窒素 (total nitrogen, TN) の測定方法として、乾式酸化法や湿式酸化法がよく用いられている。乾式酸化法は高温処理のための特殊な装置を必要とする。湿式酸化法は有害廃棄物となる試薬を使用する場合や有機炭素を過大評価する可能性がある場合などが指摘されており、プロセスが煩雑で不正確になりやすいといった課題も挙げられる⁷⁾。WS-TOC および WS-TN も同様に、乾式燃焼法、湿式酸化法など様々な方法で測定されている^{8), 9)}。しかし、これらの手法は実験室で作業を行う必要があり、煩雑な作業や高価な機器の購入が必要であることから、簡便な手法とは言えない。このような手法に対して、物質の分子振動に着目した近赤外分光法を用いた研究が行われている。

近赤外分光法は、迅速、非破壊、汎用的といった利点があり、可視近赤外域における分光反射特性を用いて土壌中の炭素や窒素などの含有量を推定する研究が多数報告されている¹⁰⁾。また、近赤外分光法に続き、近年蛍光分光法を用いた研究も報告されている。蛍光はある波長の光子を吸収し、より長い波長の光子を放出するという現象であり、蛍光検出が高感度かつ高い選択性を持つという特徴がある¹¹⁾。特に、励起波長・蛍光波長・蛍光強度の 3 次元データからなる励起蛍光マトリクス (excitation emission matrix, EEM) は、分光

しながら励起光を照射し、蛍光波長と蛍光強度を測定して得られるデータであり、物質の蛍光反応を網羅的に取得できる。分光データを入力とした土壌中の化学成分の推定に関しては、従来、部分的最小二乗回帰 (partial least squares regression, PLSR) や主成分回帰 (principal component regression, PCR) をはじめとした多変量解析を用いた先行研究が多く報告されている^{12), 13), 14), 15)}。PLSRは多重共線性のあるデータや多数の説明変数がある場合に有効で、目的変数を同時にモデル化する手法¹⁶⁾である。PCRはPLSRと同様に多重共線性のあるデータにも適用可能であるが、説明変数のみを用いてモデル構築をする点が特徴である¹⁷⁾。

土壌の蛍光特性を用いた研究として、2種類の紫外LEDを用いて可給態窒素、TC、TN、C/N比をPLSRによって予測した研究¹⁸⁾が報告されている。この研究では、375 nm および 392 nm のLEDで励起した際の蛍光スペクトルを用いており、TCおよびTNの推定モデルにおける決定係数はそれぞれ0.9356、0.6212と報告されている。この結果より、特にTNの推定においては、幅広い励起波長での蛍光スペクトルを取得可能なEEMを用いることで、推定精度が改善する可能性が考えられる。また、水抽出性有機物に含まれる蛍光性成分を詳細に分析することができる並列因子分析 (parallel factor analysis, PARAFAC) を用いて、EEMから取得した水抽出性有機物の蛍光性成分が土地利用や気候地域と有意な関連性を示すことを明らかにした研究¹⁹⁾や、湿地の溶存有機物 (dissolved organic matter, DOM) におけるEEM-PARAFAC分析と紫外可視分光法および機械学習を組み合わせて、土壌DOMのスペクトル特性が地理的・土壌特性よりも溶存有機炭素 (dissolved organic carbon, DOC) により影響を与えることを解明した研究²⁰⁾など、土壌EEMと機械学習を組み合わせた研究が多数報告されている。さらに、近年蛍光分光法に深層学習を用いて解析を行う研究が報告されている^{21), 22)}。深層学習は計算能力が高く膨大なデータに対応できる²³⁾ことから、様々な分野で広く用いられている。土壌のEEMに深層学習を適用した例もあり、バイオ炭施用後の土壌の水溶性有機物量を予測し、土壌活性の迅速な同定と早期予測への応用可能性を示した研究²⁴⁾などが報告されている。しかし、これらの手法は主に水溶性有機物の組成や特性の評価に焦点を当てており、土壌全体の動態を捉えるには水溶性有機物だけではなく、TCやTN、WS-TNといったより多くの成分を対象とする必要がある。これらに加えて、土壌から放出される温室効果ガスを推定することで、圃場の持続可能な栽培計画が可能になると考えられる。

そこで本研究は、EEMを用いて土壌中のTC、TN、WS-TOC、WS-TNおよび発生したCO₂の累積量 (累積CO₂) を、PLSRおよび畳み込みニューラルネットワーク (convolutional neural network, CNN) によって推定することを目的とした。畑地と水田から土壌を採取し、水抽出して作成した懸濁液と濾液からEEMを取得した。また、化学分析によりTC、TN、WS-TOC、WS-TNの4成分を計測するとともに、生土の培養実験からCO₂の発生量を計測した。取得した懸濁液および濾液のEEMデータを入力、土壌のCN関連成分を出力とする推定モデルをPLSRおよびCNNによって構築した。その後、Variable Importance in the Projection (VIP) の値を元に、PLSRモデルに対する重要波長の可視化と考察を行った。

2. 試料および実験方法

(1) 試料

2023年4月に新潟大学附属農場の村松ステーション (新潟県五泉市、畑地、標高25 m) および新通ステーション (新潟県新潟市、水稲栽培圃場、標高-1 m) において、スコップで表層0-10 cm深から600 g程度サンプリングした (表-1)。村松ステーションの土壌は畑地から採取した黒ボク土で、1 haの圃場を10 m×10 mのグリッドに区画し、計62試料を採取した。また、新通ステーションの土壌は水田のグライ低地土で、管理条件の異なる3つの区画から土壌サンプルを採取した。具体的には、0.18 haの耕起済み土壌から10 m×15 mのグリッドを用いて12試料、0.18 haのスラグ混入土壌から5 m×15 mのグリッドを用いて12試料、0.367 haの酒米栽培土壌から10 m×15 mのグリッドを用いて18試料が採取された。採取した土について、約60 gを4 mm径のふるいにかけて土塊を細かくしたのち、10分間根取り作業を行った。根取り済みの土をTC、TN、WS-TOC、WS-TNの測定に供試した。各土壌は測定に供試するまで、4℃で冷蔵保存した。

表-1 土壌試料の採取場所と土壌の種類および試料数。

採取場所	土地利用	土壌の種類	栽培区画	試料数	試料数合計
村松ステーション	畑地	黒ボク土		62	62
			耕起済み	12	
新通ステーション	水田	グライ低地土	スラグ混入	12	42
			酒米栽培	18	

(2) 化学成分の測定

a) TC および TN の測定

各土壌サンプルから根取り後の生土を 3.00 g 計量し、それぞれについて 3 反復作成し、計 104 地点×3 反復の試料を作成した。その後 105 °C で 24 時間乾燥させてから乳棒で砕いて粒子径を細かくしたものを 1.00 g 計量し、TC および TN の測定に供試した。測定には炭素・窒素同時定量装置 (CN コーダー MT-700Mark2, ヤナコ機器開発研究所) を用い、解析には各サンプル 2 反復測定した平均値を用いた。また、標準試料には馬尿酸 (富士フィルム和光純薬) を用い、4 点で測定した後検量線の作成を行った。

b) WS-TOC および WS-TN の測定

各生土試料 3.00 g につき 30.0 mL の純水を添加し、約 20 秒間攪拌させ、15 分間静置して得た上澄み液を取り分け、2. (3) に記す EEM の測定に供試した。この液体を懸濁液と呼ぶ。EEM 測定用の懸濁液を取り分けた液体を、振とう機マルチシェーカー (MMS-3010, 東京理化学器械) を用いて 130 rpm で 24 時間振とうした。その後遠心機 (CN-2060, AS ONE) を用いて 3780 G にて 6000 rpm で 5 分間遠心分離を行った。各サンプルで 2 反復実施し、計 104 地点×2 反復の試料を作成した。水抽出で得られた上澄み液を 0.45 μm のフィルターで濾過し、細かな土粒子を除去したものを WS-TOC および WS-TN の測定に供試した。以後、濾過した液体を濾液と呼ぶ。

測定には、接触熱分解/化学発光方式の TN (全窒素) ユニット (TNM-L, 島津製作所) を取り付け窒素分析も実施可能とした、燃焼触媒酸化方式の全有機体炭素計 (TOC-L, 島津製作所) を用いた。検量線の試料には、N:1000 mg/L の窒素標準液 (富士フィルム和光純薬) および C:1000 mg/L のフタル酸水素カリウム標準液 (富士フィルム和光純薬) を用い、1, 10, 100 ppm に調製した炭素および窒素の標準液で検量線を作成した。標準液と試料の注入量は 100 μL, 通気ガス流量は 80 mL/min とした。

c) 土壌培養による CO₂ 放出量の測定

土壌試料を約 15 g 取り分け、5 分間根取りを行った後、瓶に 10 g 計量し、20 mL の蒸留水を添加した培養瓶の中に入れて蓋をした。培養開始日までは蓋をして 4 °C で冷蔵保存した。培養開始初日は純空気を流量計 5 L で 4 分間送ることで培養瓶の中の空気置換を行った後に、純空気 30 mL を培養瓶に添加して 5 回攪拌してから 30 mL 回収してサンプルとした。空気置換の後に行ったこの工程を以後サンプル回収と呼ぶ。なお、空気の置換およびサンプル回収に用いた器具は一度に 4 サンプルを処理可能であるため、空気の置換は 4 サンプルで同時に行い、これをサンプル数が終了するまで繰り返した。サンプル回収は OECD の標準試験方である TG 301B に従い、28 日間の培養期間で試料の生分解性二酸化炭素の累積量とした。ただし、CO₂ 濃度が飽和することを避けるため、培養開始初日以降は 7 日ごとにサンプル回収を行った。以後、この濃度を累積 CO₂ と表記する。培養期間中は 20 °C で保存し、測定には高性能汎用ガスクロマトグラフ (GC-2014, 島津製作所) を使用した。

(3) EEM の測定

WS-TOC および WS-TN の測定の際に作成した懸濁液と濾液から 3.0 mL 取り分け、EEM を測定した。測定には分光蛍光光度計 (FP-8350, 日本分光) を用いた。測定を行ったサンプルの様子を図-1 に示す。図-1(a) は懸濁液、図-1(b) は濾液である。土壌粒子による散乱が大きい懸濁液の EEM 測定では、励起光の減衰や蛍光の再吸収といった散乱や内部フィルター効果による影響を最小限に抑えるため、励起光をサンプル表面に照射し同じ面から蛍光を検出する Front-face 方式を採用した。これに対し、濾液は透明性が高く、散乱の影響が小さいと考えられたため、励起光がサンプル内部を透過し 90 ° の角度から蛍光を検出することで均一な蛍光情報を取得できる Right-angle 方式を用いて測定を行った。なお、予備実験の結果から、懸濁液は蛍光強度が微弱で試行による結果のばらつきが少なかったため、1 反復のみの測定とした。濾液は 2 反復測定を行い、解析には 2 反復の平均値を用いた。濾液は水のみを EEM を測定して濾液の EEM データから減算することで、レイリー散乱による影響を除去した。測定は励起波長 200–700 nm を 5 nm 間隔、蛍光波長 210–750 nm を 1 nm 間隔で行った。測定時の装置の感度を補正するため、各測定日ごとに励起波長 350 nm で水のラマン散乱を測定し、波長 371–428 nm におけるラマン散乱の積算強度を用いた²⁵⁾。

(4) CNN および PLSR を用いた予測モデルの構築

解析のフローを図-2 に示す。解析には MATLAB_R2024b (MathWorks, USA) を用いた。説明変数としてノイズを除去した懸濁液および濾液の 2 種類の EEM をそれぞれ入力し、TC, TN, WS-TOC, WS-TN, およ

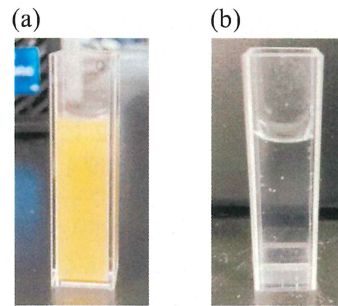


図-1 EEM 測定に使用した (a) 懸濁液 および (b) 濾液。

び累積 CO₂ を予測するモデルを CNN および PLSR によって構築した。さらに、PLSR モデルにおいて推定に寄与した波長を VIP に基づいて可視化した。

a) CNN による予測モデルの構築

CNN は第一検証と第二検証に分けて行い、グリッドサーチによるハイパーパラメータの探索を行った。CNN における学習の条件を表-2 に示す。層が比較的浅いモデルから深いモデルで精度を検証しモデルを評価するため、モデルのネットワークとして ResNet-18²⁶⁾, AlexNet²⁷⁾, EfficientNet-B0²⁸⁾の3種類を用い、それぞれについて最適化手法は、広く普及しているモーメンタム付き確率的勾配降下法 (Stochastic Gradient Descent with Momentum, SGDM)²⁹⁾と、高い学習効率で知られる適応的モーメント推定 (adaptive moment estimation, Adam³⁰⁾) を比較した。使用した3種類のネットワークは、最終層から二つ目の全結合層の出力サイズを1とするとともに、最終層の分類層を削除する処理を加えた。ミニバッチサイズと初期学習率は、それぞれ4, 8, 16, および 1.0×10^{-5} , 1.0×10^{-4} , 1.0×10^{-3} の3種類を検討した。エポック数および早期停止の停止回数はそれぞれ100, 15に固定し、CNNのモデル、最適化手法、ミニバッチサイズ、初期学習率の組み合わせを変えて計54通りの解析を行った。この解析を第一検証と呼ぶ。第一検証の結果をもとに誤差が小さかった上位3つの組み合わせを選択し、最大エポック数と早期停止の許容回数をそれぞれ50, 100, 150, および15, 30の3通りと2通りに変えて計18通りの解析を行った。この解析を第二検証と呼ぶ。第二検証の結果から、各目的変数の予測に最適な深層学習のパラメータを決定した。

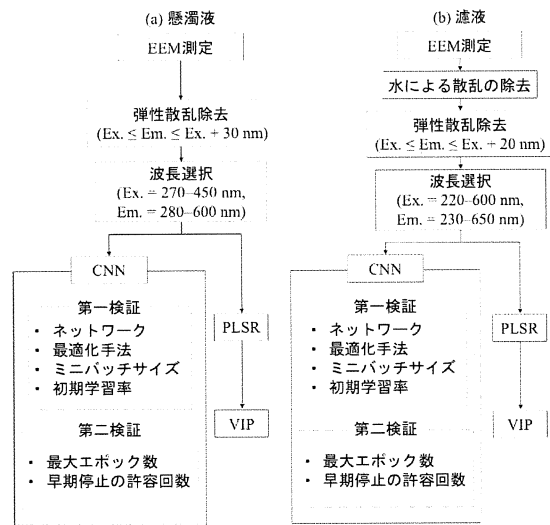


図-2 解析のフロー.

表-2 CNNの第一検証および第二検証におけるハイパーパラメータ.

ハイパーパラメータ	第一検証	第二検証
ネットワーク (Model architecture)	ResNet-18, AlexNet, Efficient-Net-B0	
最適化手法 (Optimization algorithm)	SGDM, Adam	第一検証より 選択
ミニバッチサイズ (Mini-batch size)	4, 8, 16	
初期学習率 (Initial learning rate)	1.0×10^{-3} , 1.0×10^{-4} , 1.0×10^{-5}	
最大エポック数 (Max epochs)	100	50, 100, 150
早期停止の許容回数 (Validation patience)	15	15, 30

b) PLSR による予測モデルの構築

PLSR は説明変数と目的変数の共分散が最大となる変数である潜在変数を新たに作り出して回帰を行う手法であり、説明変数と目的変数のスコア間の回帰で構成される³¹⁾。以後の説明にある各数式のサイズは以下の通りとする。なお、本研究では各目的変数に対して PLSR によるモデル構築を行っているため、目的変数側はベクトル表記とする。

- X:** サンプル数 $N \times$ 変数数 K
- y:** サンプル数 $N \times 1$
- T:** サンプル数 $N \times$ 潜在変数の個数 H
- W:** 変数数 $K \times$ 潜在変数の個数 H
- P:** 変数数 $K \times$ 潜在変数の個数 H
- u:** サンプル数 $N \times$ 潜在変数の個数 H
- q:** $1 \times$ 潜在変数の個数 H
- g:** サンプル数 $N \times 1$
- E:** サンプル数 $N \times$ 変数数 K
- f:** サンプル数 $N \times 1$

新しい変数**T**は式(1)に示すように、元の変数**X**と係数の重み**W**の線型結合として推定される。ここで、**W***はPLSRの過程で目的変数側のスコア**u**を元に最適化された重みである。

$$T = XW^* \quad (1)$$

Tにローディング**P**を乗じ、誤差**E**を加えることで説明変数**X**が適切に説明され、式(2)が成立する。

$$X = TP' + E \quad (2)$$

同様に、目的変数**y**は、**u**で表される**y**側の潜在変数にローディング**q**を乗じることで計算され、残差を**g**とすると式(3)が成立する。

$$y = uq' + g \quad (3)$$

最適化された**T**を用いて**y**を表すと式(4)の通りである。**f**は残差を表す。

$$y = Tq' + f \quad (4)$$

また、式(1)、式(4)より、式(5)に表す回帰式を得られる。ここで、 $W^*q' = b$ としている。¹⁶⁾

$$\begin{aligned} y &= XW^*q' + f \\ &= Xb + f \end{aligned} \quad (5)$$

ここで計算される残差**E**および**f**について、式(6)、式(7)が成立する³¹⁾。なお、*h*は潜在変数の計算が行われた回数を示す。

$$E_h = E_{h-1} - t_h p'_h \quad (6)$$

$$F_h = F_{h-1} - t_h q'_h \quad (7)$$

式(2)、式(4)より、1番目に計算される説明変数側と目的変数側の成分**X**₁および**y**₁は式(8)、式(9)の通りである。説明変数、目的変数側の新しい変数は、それぞれ潜在変数とローディングの積に残差を加えて計算される。

$$X_1 = t_1 p'_1 + E_1 \quad (8)$$

$$y_1 = t_1 q'_1 + f_1 \quad (9)$$

残差について式(6)、式(7)が成り立つことから、2番目に計算される**X**₂および**y**₂成分は式(10)、式(11)の通りである。

$$X_2 = E_1 = X_1 - t_1 p'_1 \quad (10)$$

$$y_2 = f_1 = y_1 - t_1 q'_1 \quad (11)$$

潜在変数の数を説明変数行列**X**のランクと等しくすれば**E**および**f**は0に近づき、元のデータをうまく再構成できる一方で、説明変数に多数の相関がある場合は過適合する可能性が高くなる。その結果訓練データへの適合度が高い一方で予測能力が低いモデルとなる。そのため、潜在変数の個数を増やすことによる予測優位性がなくなった時点で潜在変数の計算を止める必要がある¹⁶⁾。予測残差平方和 (prediction residual sum of squares, PRESS) が最小になる個数とするのが望ましいとされる³¹⁾が、用いる潜在変数の個数が増加すると誤差が0に漸近し、最小値を持たない場合がある。そのため本研究では一個抜き交差検証 (Leave-one-out cross-validation, LOOCV) を行い、予測の平均二乗誤差 (mean square error, MSE) が増加する一つ前の個数を潜在変数の個数 (the number of latent variables, LVs) として設定した。PLSRモデルの重要波長はVIPを用いて計算した。VIPスコアは潜在変数に基づいて目的変数を予測する際の変数の重要度を要約したものである。変数選択の基準として、一般的にVIPスコアが1より大きい変数が重要であるとみなされる³²⁾。

c) 交差検証

CNNとPLSRのモデル構築には5分割の交差検証を用い、全データを5つに分割したのうち、4セットを訓練、残り1セットをテストとした。全てのセットがテストになるよう訓練データとテストデータの組み合わせを変えて5通りのモデルを作成し、それらの結果の平均値をモデルの評価に用いた。モデルの精度として訓練データにおける二乗平均平方根誤差 (root mean square error of calibration, RMSE_c) と決定係数

(coefficient of determination of calibration, R^2_c) および予測偏差比 (residual prediction deviation of calibration, RPD_c) を算出した。また、テストデータにおける二乗平均平方根誤差 (out mean square error of cross-validation, RMSE_{cv}) と決定係数 (coefficient of determination of cross-validation, R^2_{cv}) および予測偏差比 (residual prediction deviation of cross-validation, RPD_{cv}) を算出し、モデルの評価に用いた。

RPD は標準偏差を予測の RMSE で割った値で、モデルの信頼性を評価する手法である。RPD の基準値は多数報告されている^{12), 33)}が統計的な根拠がなく、明確な指標は統一されていない^{34), 35)}。一般的には RPD の値が大きいほどモデルの信頼性が高いことを意味する。本研究では土壌を対象にした Chang ら (2021)の研究に則り、RPD ≥ 2.0 を優れたモデルとする。

RMSE, R^2 , RPD の計算式をそれぞれ式(12), (13), (14)に示す。nはサンプル数, yは実測値, \hat{y} は予測値, \bar{y} は平均値, σ は実測値の標準偏差を表す。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (13)$$

$$RPD = \frac{\sigma}{RMSE} \quad (14)$$

3. 結果および考察

(1) 励起蛍光マトリクス (EEM) の結果

畑地土壌と水田土壌における懸濁液を用いて測定した EEM の代表的な結果を図-3(a)から図-3(d)に、濾液を用いて測定した EEM の代表的な結果を図-3(e)から図-3(h)に示す。図-3の(a)から(d), (e)から(h)はそれぞれ順に畑地土壌, 耕起済み水田土壌, スラグ入り水田土壌, 酒米栽培水田土壌のサンプルを表す。なお、レイリー散乱および蛍光物質に由来しないノイズ領域を除去し、懸濁液で励起波長 (Ex.) 270–450 nm, 蛍光波長 (Em.) 280–600 nm, 濾液で Ex. = 220–600 nm, Em. = 230–650 nm の範囲を示している。縦軸は励起波長, 横軸は蛍光波長, 色は蛍光強度を示す。なお、各蛍光ピーク間のスケール差が 100 倍など大きく異なる場合が見られなかったことに加え、ノイズ除去済みデータを使用していることから、特定の蛍光領域がモデルに過度な影響を与えるリスクは低いと判断し、本研究においてパリト化は実施していない。

図-3(a)から図-3(d)より、Ex. = 330–360 nm, Em. = 420–470 nm 付近 (Peak A) に現れているのはフミン様物質³⁶⁾と考えられる。さらに、Ex. = 280–320 nm, Em. = 300–360 nm 付近 (Peak B) に見られるアミノ酸様物質およびタンパク質様物質³⁶⁾と考えられるピークは水田土壌にのみ見られ、畑地土壌には見られなかった。水田は湛水期には嫌気的な条件になることから、特定の微生物等の活動が活発になり、タンパク質の分解お

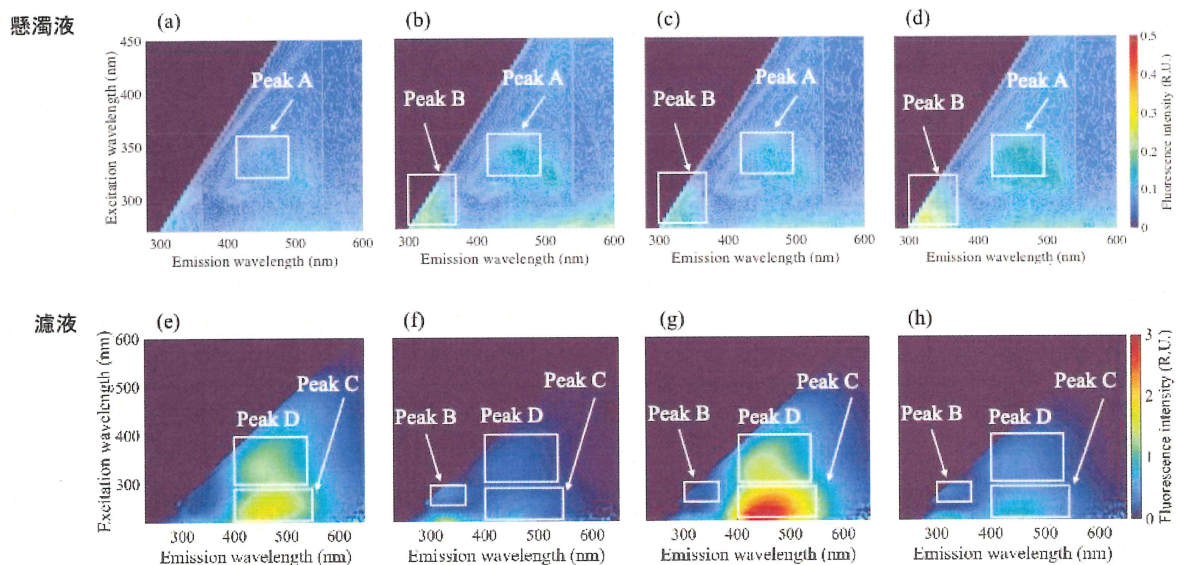


図-3 懸濁液および濾液における(a), (e)畑地土壌, (b), (f)耕起済みの水田土壌, (c), (g)スラグ混入水田土壌, (d), (h)酒米栽培水田土壌の EEM.

よび生成プロセスが生じたため、特有の物質が蓄積して蛍光ピークとして現れた可能性がある。

図-3(e)から図-3(h)より、濾液では Ex. = 280–300 nm, Em. = 300–360 nm 付近 (Peak B) に懸濁液と同様のアミノ様物質およびタンパク質様物質³⁶⁾, Ex. = 220–290 nm, Em. = 400–550 nm 付近 (Peak C) にフルボ酸³⁷⁾, Ex. = 300–400 nm, Em. = 400–540 nm 付近 (Peak D) にフミン酸³⁷⁾由来と考えられる蛍光ピークが存在する。

Peak C, Peak D に現れたと考えられるフルボ酸およびフミン酸は腐植物質と呼ばれ、土壌の肥沃度や植物栄養において重要な物質である³⁸⁾。水田土壌の濾液の中で、スラグが混ざっている場合は Peak C および Peak D が強いという特徴が見られた。この原因として、スラグ中に含まれる CaO が湛水条件により水と反応し、強アルカリである Ca(OH)₂ に変化した可能性が考えられる。フミン酸は酸に不溶である一方アルカリには可溶³⁹⁾であることから、生成された強アルカリ成分により、多くのフミン酸が溶解し蓄積した結果、スラグ入りサンプルの EEM でフミン酸由来と考えられる Peak D が強く現れた可能性がある。また、フルボ酸は酸にもアルカリにも可溶³⁹⁾であるが、フミン酸と同様に腐植物質であることから、土壌中の複雑な化学的変化を通じてフミン酸と相互に影響し合っている可能性がある。このため、スラグ入りのサンプルでは溶液の液性に変化し、Peak C および Peak D の蛍光強度が強くなったと考えられる。

これを検証するため、濾液の pH をコンパクト pH メータ LAQUATwin (HPRIBA) で測定した。各サンプルで 2 回測定し、その平均値を用いて解析した結果、スラグ混入土壌から取得した 12 試料のうち 10 試料が pH 7.0 以上を示した。一方、スラグ混入土壌を除く 92 試料では、全て pH 7.0 未満を示した。このことから、スラグ入りのサンプルで溶液がアルカリ性に寄っていたことが示された。

また、図-3(a)から図-3(d)と図-3(e)から図-3(h)を比較すると、懸濁液と濾液では蛍光強度のスケールが大きく異なることがわかる。懸濁液では土壌粒子密度が増加し、蛍光物質による蛍光強度の増加よりも土壌粒子によって励起光や蛍光が吸収または散乱される影響、即ち内部遮蔽効果が発生し、蛍光強度が減衰したためと考えられる。

(2) CNN および PLSR による予測モデルの結果

CNN と PLSR による予測モデルの精度を表-3 に示す。CNN は第一検証と第二検証の結果から、濾液および懸濁液を用いた場合の各目的変数でテストデータにおける R^2_{cv} が最も大きかったハイパーパラメータの組み合わせと、その時の RMSE_{cv}, R^2_{cv} , RPD_{cv} を示している。精度は試行ごとに多少ばらつきが生じたが、5 分割の交差検証を実施しているため、データ分割には影響を受けていないとみなす。また、表-3 中の PLSR モデルの結果として、用いた潜在変数の個数とテストデータにおける RMSE_{cv}, R^2_{cv} , RPD_{cv} を示している。

表-3 より、CNN モデルにおける最適なサンプルは TC で懸濁液、TN, WS-TOC, WS-TN, 累積 CO₂ で濾液だった。累積 CO₂ を除くすべての目的変数で R^2_{cv} が 0.8 以上かつ、RPD_{cv} が 2.0 以上となったことから優れたモデルと判断でき、高精度に推定できる結果となった。

累積 CO₂ は CNN モデルで懸濁液、濾液のいずれを用いた場合も R^2_{cv} が 0.45 程度を示し、他の成分と比較すると推定が難しい結果となった。累積 CO₂ の培養実験の結果から、同一土壌タイプでも標準偏差が大きい場合があり、測定自体が適切に行えず真値を反映しきれなかった可能性がある。また、一般的に深層学習のネットワークの深さが深いほど多くの特徴量を取得できるため、層が深い EfficientNet-B0 が選択されている累積 CO₂ では特徴量は多く取得されていると考えられる。推定精度が低くなった原因として、正解データのばらつきが大きかったため、EEM データと CO₂ の有用な関連性を学習できず、適切な特徴量を取得できなかった可能性がある。

TC, TN および累積 CO₂ は懸濁液を用いた場合も濾液を用いた場合もテストデータにおける精度に大きな変化がなかったことから、これらの目的変数はサンプルの性質よりもモデルのネットワークに依存して精度が変動すると考えられる。一方、WS-TOC および WS-TN に関して、懸濁液の EEM を用いて推定した場合の R^2_{cv} はそれぞれ 0.68, 0.50 だったのに対し、濾液では 0.86, 0.91 を示した。これらは懸濁液では予測が難しく、土壌粒子を除去することで推定に有効な蛍光特性を取得できたためと考えられる。

ハイパーパラメータとして、ミニバッチサイズ、初期学習率は TC, TN, WS-TOC, WS-TN でそれぞれ $16, 1.0 \times 10^{-3}$ が選択された。検証した中では最も大きいものが選択されたことから、効率的に学習できたと考えられる。一方、累積 CO₂ は小さい初期学習率が選択されており、分光データと化学成分の乖離が大きかったため学習が進まなかったと考えられる。

表-3 において PLSR モデルによる結果に注目すると、TC の懸濁液および濾液、TN の濾液で R^2_{cv} が 0.9 以上の結果となり、TN の懸濁液でも R^2_{cv} が 0.8 以上を示した。ここで、表-4 に TC, TN, WS-TOC, WS-TN 含有量および累積 CO₂ 量と、懸濁液における Ex. = 320 nm, Em. = 450 nm および濾液における Ex. = 350 nm, Em. = 450 nm の相関係数を示す。なお、TC および TN の測定時に 1.00 g を計量したが、スラグ入りのサンプルは土壌部分が 1.00 g より少なくなったことから、表-4 の相関係数はスラグ入り土壌の結果を除いている。表-4 に示す通り、TC および TN の含有量と懸濁液、濾液の蛍光強度に線形の相関関係があったため、線形モデルである PLSR でも比較的高い精度を得られたと考えられる。特に濾液を使った場合は TC, TN でそれぞ

表-3 R^2_{cv} が最も高かったハイパーパラメータによる CNN モデルの精度および PLSR モデルの精度. 略語は略語リストを参照.
(※1 単位は, TC および TN が mg/g-ds, WS-TOC, WS-TN および累積 CO_2 が ppm.)

Objective variables	Data	Model	Optimization algorithm	Mini-batch size	Max epochs	Initial learning rate	Validation patience	LVs	RMSE _{cv} (※1)	R^2_{cv}	RPD _{cv}
TC	Sus.	Efficient-Net-B0	Adam	16	150	1.0×10^{-3}	30	-	4.54	0.95	4.62
	Filt.	ResNet-18	SGDM	8	50	1.0×10^{-3}	30	-	4.94	0.93	4.49
	Sus.	PLSR	-	-	-	-	-	6	5.65	0.92	3.74
	Filt.	PLSR	-	-	-	-	-	7	6.25	0.90	3.25
TN	Sus.	Efficient-Net-B0	Adam	16	150	1.0×10^{-3}	15	-	0.34	0.90	3.36
	Filt.	Efficient-Net-B0	Adam	16	50	1.0×10^{-3}	30	-	0.28	0.93	4.27
	Sus.	PLSR	-	-	-	-	-	6	0.42	0.84	2.86
	Filt.	PLSR	-	-	-	-	-	9	0.35	0.90	3.29
WS-TOC	Sus.	AlexNet	SGDM	16	150	1.0×10^{-3}	30	-	1.19	0.68	1.93
	Filt.	AlexNet	SGDM	16	50	1.0×10^{-3}	15	-	0.80	0.86	2.86
WS-TN	Sus.	AlexNet	Adam	16	100	1.0×10^{-3}	15	-	0.59	0.50	1.56
	Filt.	ResNet-18	Adam	16	100	1.0×10^{-3}	15	-	0.23	0.91	3.57
	Sus.	PLSR	-	-	-	-	-	3	0.67	0.32	1.32
	Filt.	PLSR	-	-	-	-	-	9	0.30	0.85	2.89
累積 CO_2	Sus.	Efficient-Net-B0	SGDM	4	50	1.0×10^{-5}	15	-	1408	0.45	1.39
	Filt.	Efficient-Net-B0	SGDM	4	50	1.0×10^{-5}	15	-	1396	0.46	1.43
	Sus.	PLSR	-	-	-	-	-	2	1452	0.40	1.35
	Filt.	PLSR	-	-	-	-	-	4	1554	0.29	1.25

略語リスト

- ・ TC: total carbon (全炭素)
- ・ TN: total nitrogen (全窒素)
- ・ WS-TOC: water-soluble total organic carbon (水溶性全有機体炭素)
- ・ WS-TN: water-soluble total nitrogen (水溶性全窒素)
- ・ Sus.: suspension (懸濁液)
- ・ Filt.: filtrate (濾液)
- ・ SGDM: stochastic gradient descent with momentum (モーメント付き確率的勾配降下法)
- ・ Adam: adaptive moment estimation (適応的モーメント推定)
- ・ PLSR: partial least squares regression (部分的最小二乗回帰)
- ・ LVs: the number of latent variables (潜在変数の個数)
- ・ RMSE_{cv}: root mean square error of cross-validation (交差検証の二乗平均平方根誤差)
- ・ R^2_{cv} : coefficient of determination of cross-validation (交差検証の決定係数)
- ・ RPD_{cv}: residual prediction deviation of cross-validation (交差検証の予測偏差比)

れ 0.910, 0.930 の相関係数を示し, 蛍光強度と化学成分に強い正の相関が見られる. 一方, 懸濁液は TC, TN でそれぞれ-0.744, -0.740 と中程度の負の相関だったが, PLSR の R^2_{cv} は TC, TN とともに 0.8 以上となった.

WS-TOC および WS-TN における相関係数は, 懸濁液でそれぞれ-0.408, -0.509, 濾液でそれぞれ 0.763, 0.611 を示した. 絶対値で比較すると濾液では TC および TN と懸濁液と同程度の相関係数を示したが, 濾液の EEM を用いた WS-TOC の PLSR モデルの R^2_{cv} は 0.6 程度と, 高精度な推定は難しい結果となっ

た。一方、WS-TNを濾液で推定した場合の R^2_{cv} は0.85となり、線形関係が強くない場合でもPLSRモデルで高精度に推定できる結果も見られた。累積CO₂は懸濁液と中程度の正の相関、濾液と弱い負の相関が見られた。PLSRモデルによる R^2_{cv} はそれぞれ0.40, 0.29で、懸濁液の精度はCNNモデルによる精度と同程度の結果だった。一方、濾液を用いた場合の精度は著しく低く、濾液中の蛍光成分とCO₂との関連性を線形モデルでは捉えきれなかったため、推定精度が低下したと考えられる。

また、表-4に示したように、懸濁液における蛍光強度と化学成分との相関係数はTC, TN, WS-TOC, WS-TNで負の値を示した。これは、土壌粒子による内部遮蔽効果が発生し、蛍光強度が減衰したためと考えられる。即ち、例えばTCが多い場合はその中に存在する蛍光物質も多くなる一方で、励起光を吸収および蛍光を再吸収する物質も多くなるため、本来現れるはずの蛍光が減衰して検出される。反対にTCが少ない場合は蛍光物質の量が少なく、吸光性物質も少ないため内部遮蔽効果の影響を受けにくくなる。一方、累積CO₂でのみ相関係数の正負がその他の成分とは反対の結果を示した。土粒子に吸着した有機物と二酸化炭素の発生量とに相関があった可能性が考えられる。

表-4から、TCおよびTNは濾液のEEMと強い相関があった一方、WS-TOCおよびWS-TNは化学計測に用いたサンプルとEEMの測定に用いたサンプルが同じだったにもかかわらず中程度の相関しか示さなかった。これは、WS-TOCおよびWS-TNの化学測定が濾液中の全ての水溶性有機物および窒素化合物を測定している一方で、EEMは蛍光性の物質の情報のみを捉えていることに起因すると考えられる。一方、TCおよびTNは土壌中の全ての有機物および窒素化合物を測定しているため、土壌中の蛍光性物質と、より関連性が高い可能性がある。

表-3においてCNNモデルとPLSRモデルの結果を比較すると、PLSRよりもCNNを使った場合に精度が高いことから、CNNが予測に効果的である可能性が示された。EEMデータは「励起波長 × 蛍光波長」の二次元行列として得られ、各画素値が蛍光強度を表す。ピークは座標(Ex, Em)に位置づけられるため、ピーク同士の相対距離や強度勾配、およびピーク成分の広がりといった空間的構造そのものが化学的意味を持つ。しかし、PLSRではこの二次元行列を一次元ベクトルへ平坦化するため、隣接波長間の局所パターン、複数ピークの共存・相対配置、強度勾配やピーク幅といった空間的・非線形特徴は失われやすく、非線形的な相互作用を十分に捉えられない。また、同一波長帯の強度のみを比較するため、ピークが1-2ピクセル移動しただけで寄与度が大きく変動する可能性がある。

対照的にCNNは5 × 5などの大きさの畳み込みカーネルを画像全体に対してスライドさせ、局所勾配やピーク形状に関する情報を抽出し、他の特徴との相対的な配置がおおよそ維持されている限り、その絶対的な位置情報は重要視されない⁴⁰⁾。このとき、プーリング層はピーク位置の微小な波長シフトに対し平行移動不変性を付与するため、試料間でピークがわずかにずれても特徴量が安定して保持されると考えられる。したがって、EEM画像に固有の二次元空間構造を保ったまま局所特徴とその相対配置を階層的に学習できるCNNを利用することで、線形モデルであるPLSRよりも化学成分の推定において高い予測精度を得ることができたと考えられる。特にWS-TOCは濾液を用いたCNNによるモデル構築の場合にのみ RPD_{cv} が2.0を超えたことから、PLSRモデルでは捉えられなかった特徴量をうまく学習できていたと考えられる。

この結果から、説明変数と目的変数に線形の強い相関関係がある場合はPLSRでも十分に予測可能である一方、線形の相関が強くないなど、複雑なモデルが必要な場合はCNNが適していると考えられる。累積CO₂のように説明変数と目的変数の乖離が大きく予測自体が難しい場合は、モデルの違いは精度に大きく影響を与えず、一様に低い精度を示した。

本研究ではハイパーパラメータ探索にグリッドサーチを採用した。学習率・ミニバッチサイズ・最大エポック数・最適化アルゴリズムなどの主要パラメータについてあらかじめ離散的な候補集合を設定し、全組合せを総当たりで評価した。グリッドサーチは実装が容易で再現性も高い反面、データセットの次元が高い場合やハイパーパラメータの組み合わせが多い場合は探索に時間を要する⁴¹⁾。さらに、探索空間が増えると計算負荷が指数関数的に増大する⁴²⁾ため、連続値を細かく最適化する場合には探索効率が低下するという制約がある。

将来的には、より効率的な探索手法としてベイズ最適化^{43), 44)}やランダムサーチ⁴⁵⁾を利用することが望ましい。ベイズ最適化は既存の評価結果から確率モデルを構築し、高精度が期待される領域を優先的にサンプリングすることで試行回数を大幅に削減できる。Itakuraら(2019)による研究では、ベイズ最適化によるハイパーパラメータの探索が精度の向上に貢献したことが報告されている²¹⁾。同様に、ランダムサーチがグリッドサーチよりも効率的にハイパーパラメータの最適化を行えることも報告されている⁴⁵⁾。本研究の今後の課題として、ベイズ最適化やランダムサーチなどの自動的な最適化手法を用いることで、より効率的かつ網羅

表-4 各目的変数と蛍光強度の相関係数。

目的変数	懸濁液	濾液
	(Ex./Em. = 320/450 nm)	(Ex./Em. = 350/450 nm)
TC	-0.744	0.910
TN	-0.740	0.930
WS-TOC	-0.408	0.763
WS-TN	-0.509	0.611
累積CO ₂	0.567	-0.455

的なハイパーパラメータの選択を行うことが必要である。

また、本研究では交差検証によるモデル性能の議論および重要波長の検証を目的としているため、完全外部検証を実施できていない。今後は本研究で構築したモデルを、全く異なる水田や畑で取得した土壌に適用することで外部検証を行い、性能を評価する必要がある。

(3) VIPスコアに基づくPLSRモデルにおける重要波長の可視化

懸濁液と濾液のEEMデータを用いてPLSRを行い、VIPスコアが1以上の重要波長を可視化した結果をそれぞれ図-4(a)から図-4(e)、および図-4(f)から図-4(j)に示す。図-4の(a)から(e)、(f)から(j)はそれぞれ順に、TC、TN、WS-TOC、WS-TN、累積CO₂の結果を表す。なお、EEMに見られたPeak AからPeak Dのおおよその位置を図中に示している。図-4より、VIPスコアはEEM全体の蛍光ピークに重複する部分で高い値を示し、EEMの結果を強く反映していることがわかる。

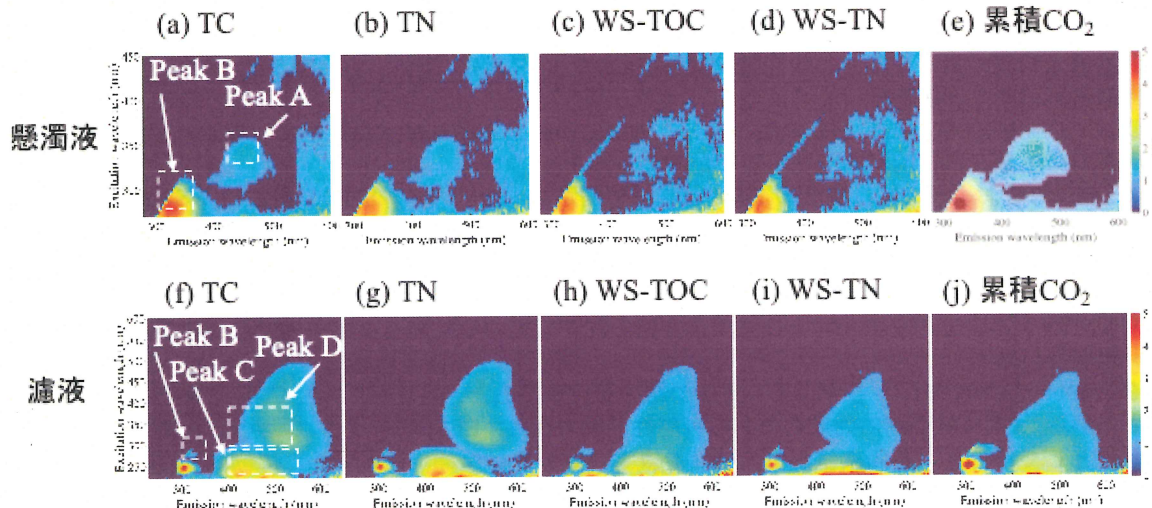


図-4 懸濁液および濾液のEEMデータによるPLSRを用いた(a), (f) TC, (b), (g) TN, (c), (h) WS-TOC, (d), (i) WS-TN, および(e), (j) 累積CO₂の予測におけるVIPスコア

図-4(a), 図-4(b), 図-4(e)より、TC、TN、累積CO₂の推定には懸濁液の微小なPeak AとPeak Bが両方も明確に選択されている。このことから、Peak Aの由来と考えられるフミン様物質はTC、TN、累積CO₂量と関連が深く、土壌粒子に付着していたフミン様物質が3つの成分を推定する上で重要な役割を果たしていることが予想される。さらに、図-4(c), 図-4(d)より、WS-TOCおよびWS-TNのVIPではPeak Aと比較してPeak Bがより明確に選択されたことから、水溶性成分のうち蛍光性の物質と関連が深いのはPeak Bの由来と考えられるアミノ酸様物質およびタンパク質様物質であると考えられる。さらに、懸濁液による推定では累積CO₂を除く4つの成分で、蛍光ピークとは異なる領域でモザイク状に波長が選択されている。これは、これら4つのCN関連成分が懸濁液のEEMの全体的な傾向と関連している可能性を示唆している。一方、累積CO₂ではノイズの領域はほとんど選択されておらず、CO₂濃度は懸濁液のEEM全体ではなく特定の蛍光ピークとの関連性が強かったと考えられる。

図-4(f)から図-4(j)より、濾液のEEMを用いた予測ではEEMの蛍光ピークと類似した領域でVIPスコアが高くなっていることがわかる。特に、全ての成分の予測において、フルボ酸と考えられるPeak Cが強く現れている。フルボ酸は腐植物質の中でも分子量が小さく水溶性が高い性質を持つ³⁸⁾ことから、水に溶け出したフルボ酸はCN関連成分の短期的な特性を示すことが考えられる。また、図-4(j)より、積算CO₂の予測にはフミン酸と考えられるPeak Dの領域が明確に寄与していることがわかる。フミン酸はフルボ酸と比較して分子量が大きく³⁸⁾安定した構造をしていることから、土壌中にとどまる腐植物質が温室効果ガスの量と関連している可能性が考えられる。一方、EEMでピークとして強く現れなかったEx. = 250 nm, Em. = 300 nm付近でVIPスコアが高くなっている領域が見られる。この領域で蛍光反応を示すチロシン様物質やタンパク質様物質の蛍光がわずかに現れていた可能性がある³⁶⁾。これらのVIPの結果より、全ての目的変数に対して、EEMで見られる蛍光ピークと共通の領域を選択している一方で、EEMに明瞭に現れなかった蛍光領域も捉えていることが示された。

VIPは個々の説明変数がPLSモデルに与える影響を要約しており、抽出された各潜在変数で説明された目的変数を考慮に入れる重み w^* を用いて計算する⁴⁶⁾。入力したすべてのサンプルの説明変数を読み込むことから、VIPはデータセット全体に対して等しく重要度を評価しており、EEMにおける蛍光ピークのパターン

を反映したスコアマップとなりやすいと考えられる。

4. 結論

本研究では、畳み込みニューラルネットワーク (CNN) と部分的最小二乗回帰 (PLSR) を用いて、土壌から取得した懸濁液および濾液の励起蛍光マトリクス (EEM) から全炭素 (TC)、全窒素 (TN)、水溶性全有機体炭素 (WS-TOC)、水溶性全窒素 (WS-TN) 含量、および発生した CO₂ 量 (累積 CO₂) を推定するモデルを構築した。TC および TN の量は蛍光強度と強い線形の相関があったことから、モデルを問わず懸濁液と濾液の両方でテストデータにおける決定係数 R^2_{cv} が 0.8 以上となった。一方、WS-TOC および WS-TN では濾液の EEM に CNN を用いると高精度に予測できる結果となり、線形回帰を目的とする PLSR では捉えきれなかった特徴量を学習できたと考えられる。累積 CO₂ は分光データと測定データの乖離が大きかったため、CNN を用いた場合でも推定が困難だった。

今後の展望として、CNN を実装する際のハイパーパラメータの探索方法として、より効率的な手法であるベイズ最適化やランダムサーチを用いることが求められる。さらに、深層学習によるモデルの実応用に向けては、判断根拠を把握しながらモデルを評価することが必要である。そのため、Local Interpretable Model-agnostic Explanation (LIME)⁴⁷⁾ や Gradient-weighted Class Activation Mapping (Grad-CAM)⁴⁸⁾ などの説明可能 AI を用いた判断根拠の可視化を導入し、解釈性を一層高める予定である。

将来的には、本研究をもとに最適な波長の選択や外部検証用データを使用したモデルの性能評価を実施する必要がある。これらの検証を経て、最終的には可搬型分光計に加え、Edge-CNN のような少量データで学習できる軽いモデルを内蔵した小型センサーを実装し、フィールドで使用可能な簡易土壌診断センサーの構築を目指す。

参考文献

- 1) IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Lee, H., and Romero, J. (eds.)], Section 2, pp. 41-66, 2023.
- 2) Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D. J., Mauritsen, T., Palmer, M. D., Watanabe, M., Wild, M., and Zhang, H.: The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity [Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B. (eds.)], Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, pp. 923-1054, 2021.
- 3) 日本国温室効果ガスイベントリ報告書 2025 年, 温室効果ガスイベントリオフィス編, 環境省地球環境局総務課脱炭素社会移行推進室監修, 地球環境研究センター: https://www.nies.go.jp/gio/archive/nir/ua88o20000099s22-att/NID-JPN-2025-v3.0_J_gioweb.pdf (最終閲覧日 2025 年 6 月 24 日)
- 4) Mumu, N. J., Ferdous, J., Müller, C., Ding, W., Zaman, M. and Jahangir, M. M. R.: Methodological progress in the measurement of agricultural greenhouse gases, *Carbon Management*, Vol.15, No.1, pp.2366527, 2024.
- 5) Le Mer, J. and Roger, P.: Production, oxidation, emission and consumption of methane by soils: a review, *European journal of soil biology*, Vol.37, No.1, pp.25-50, 2001.
- 6) Butterbach-Bahl, K., Baggs, E. M., Dannenmann, M., Kiese, R. and Zechmeister-Boltenstern, S.: Nitrous oxide emissions from soils: how well do we understand the processes and their controls?, *Phil. Trans. R. Soc. B*, Vol.368, No.1621, pp.20130122, 2013.
- 7) Provin, T.: Total carbon and nitrogen and organic carbon via thermal combustion analyses, *Soil test methods from the southeastern United States*, pp.149-154, 2014.
- 8) Bisutti, I., Hilke, I. and Raessler, M.: Determination of total organic carbon – an overview of current methods, *TrAC Trends in Analytical Chemistry*, Vol.23, No.10-11, pp.716-726, 2004.
- 9) 大槻晃: アルカリ性ペルオキシ二硫酸カリウム分解を用いる環境水中の全窒素測定法における硝酸イオンの紫外吸光度定量法の応用, *分析化学*, Vol.30, No.10, pp.688-689, 1981.
- 10) Ahmadi, A., Emami, M., Daccache, A. and He, L.: Soil properties prediction for precision agriculture using visible and near-infrared spectroscopy: A systematic review and meta-analysis, *Agronomy*, Vol.11, No.3, pp.433, 2021.
- 11) Camarca, A., Varriale, A., Capo, A., Pennacchio, A., Calabrese, A., Giannattasio, C., Almuzara, C. M., D'Auria, S. and Staiano, M.: Emergent Biosensing Technologies Based on Fluorescence Spectroscopy and Surface Plasmon Resonance, *Sensors*, Vol.21, No.3, pp.906, 2021.
- 12) Chang, C.-W., Laird, D. A., Mausbach, M. J. and Hurburgh, C. R.: Near - Infrared Reflectance Spectroscopy - Principal Components Regression Analyses of Soil Properties, *Soil Science Soc of Amer J*, Vol.65, No.2, pp.480-490, 2001.
- 13) Yanli, L., Youlu, B., Liping, Y. and Hongjuan, W.: Hyperspectral extraction of soil organic matter content based on principal

- component regression, *New Zealand Journal of Agricultural Research*, Vol.50, No.5, pp.1169-1175, 2007.
- 14) Kodaira, M. and Shibusawa, S.: Mobile Proximal Sensing with Visible and Near Infrared Spectroscopy for Digital Soil Mapping, *Soil Systems*, Vol.4, No.3, pp.40, 2020.
 - 15) 小平正和, 澁澤栄: リアルタイム土壌分光計測と精密農業, レーザー研究, Vol.49, No.10, pp.556-561, 2021.
 - 16) Wold, S., Sjöström, M. and Eriksson, L.: PLS-regression: a basic tool of chemometrics, *Chemometrics and Intelligent Laboratory Systems*, Vol.58, No.2, pp.109-130, 2001.
 - 17) Vigneau, E., Devaux, M. F., Qannari, E. M. and Robert, P.: Principal component regression, ridge regression and ridge principal component regression in spectroscopy calibration, *J. Chemometrics*, Vol.11, No.3, pp.239-249, 1997.
 - 18) 織井孝治, 井上直人: 紫外LED励起蛍光分析による畑作土壌の可給態窒素, 全炭素, 全窒素, C/N比の推定, 日本土壌肥科学雑誌, Vol.90, No.2, pp.116-122, 2019.
 - 19) Rinot, O., Borisover, M., Levy, G. J. and Eshel, G.: Fluorescence spectroscopy: A sensitive tool for identifying land-use and climatic region effects on the characteristics of water-extractable soil organic matter, *Ecological Indicators*, Vol.121, pp.107103, 2021.
 - 20) Wenliang, J., Dengke, M., Meilian, C., Jordi, S., Biao, Z., Ji, L., Ji, C., Pen uelas, J. and Xiang, L.: Soil-derived dissolved organic matter in inland wetlands along a temperate river: Insights from spectroscopic characteristics coupled with machine learning methods, *Water Research X*, Vol. 29, pp. 100339, 2025.
 - 21) Itakura, K., Saito, Y., Suzuki, T., Kondo, N. and Hosoi, F.: Estimation of Citrus Maturity with Fluorescence Spectroscopy Using Deep Learning, *Horticulturae*, Vol.5, No.1, pp.2, 2018.
 - 22) Li, Y., Liu, X., Zhang, X., Gu, X., Yu, L., Cai, H. and Peng, X.: Using solar-induced chlorophyll fluorescence to predict winter wheat actual evapotranspiration through machine learning and deep learning methods, *Agricultural Water Management*, Vol.309, pp. 109322, 2025.
 - 23) Sun, C., Shrivastava, A., Singh, S. and Gupta, A.: Revisiting unreasonable effectiveness of data in deep learning era, *Proc. of the IEEE international conference on computer vision*, pp.843-852, 2017.
 - 24) Liang, J., Wei, D., Yin, D., Zou, G., Li, Y., Zhang, Y., Ding, J., Wang, L., Liang, L., Sun, L., Wang, W., Shen, H., Wang, Y. and Xu, J.: Prediction of Soil Water - Soluble Organic Matter by Continuous Use of Corn Biochar Using Three - Dimensional Fluorescence Spectra and Deep Learning, *Computational Intelligence and Neuroscience*, Vol.2023, No.1, pp.7535594, 2023.
 - 25) Lawaetz, A. J. and Stedmon, C. A.: Fluorescence Intensity Calibration Using the Raman Scatter Peak of Water, *Appl Spectrosc*, Vol.63, No.8, pp.936-940, 2009.
 - 26) He, K., Zhang, X., Ren, S. and Sun, J.: Deep residual learning for image recognition, *Proc. of the IEEE conference on computer vision and pattern recognition*, pp.770-778, 2016.
 - 27) Krizhevsky, A., Sutskever, I. and Hinton, G. E.: Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, Vol.25, pp. 1097-1105, 2012.
 - 28) Tan, M. and Le, Q.: Efficientnet: Rethinking model scaling for convolutional neural networks, *Proc. of International conference on machine learning*, pp.6105-6114, PMLR, 2019.
 - 29) Qian, N.: On the momentum term in gradient descent learning algorithms, *Neural Networks*, Vol.12, No.1, pp.145-151, 1999.
 - 30) Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, arXiv:1412.6980, 2017.
 - 31) Geladi, P. and Kowalski, B. R.: Partial least-squares regression: a tutorial, *Analytica chimica acta*, Vol.185, pp.1-17, 1986.
 - 32) Chong, I.-G. and Jun, C.-H.: Performance of some variable selection methods when multicollinearity is present, *Chemom. Intell. Lab. Syst.*, Vol.78, No.1-2, pp.103-112, 2005.
 - 33) Saeys, W., Mouazen, A. M. and Ramon, H.: Potential for Onsite and Online Analysis of Pig Manure using Visible and Near Infrared Reflectance Spectroscopy, *Biosystems Engineering*, Vol.91, No.4, pp.393-402, 2005.
 - 34) Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.-M. and McBratney, A.: Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy, *TrAC Trends in Analytical Chemistry*, Vol.29, No.9, pp.1073-1081, 2010.
 - 35) Minasny, B. and McBratney, A.: Why you don't need to use RPD, *Pedometron*, Vol.33, No.600, pp.14-15, 2013.
 - 36) Coble, P. G.: Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy, *Marine Chemistry*, Vol.51, No.4, pp.325-346, 1996.
 - 37) Mounier, S., Patel, N., Quilici, L., Benaim, J. Y. and Benamou, C.: FLUORESCENCE 3D DE LA MATIÈRE ORGANIQUE DISSOLUÈ DU FLEUVE AMAZONE (THREE-DIMENSIONAL FLUORESCENCE OF THE DISSOLVED ORGANIC CARBON IN THE AMAZON RIVER), *Wat. Res.*, Vol.33, No.6, pp.1523-1533, 1999.
 - 38) Pettit, R. E.: Organic matter, humus, humate, humic acid, fulvic acid and humin: their importance in soil fertility and plant health, *CTI Research*, Vol.10, pp.1-7, 2004.
 - 39) Hayes, M. H. and Swift, R. S.: The chemistry of soil organic colloids, *The chemistry of soil constituents*, pp.179-320, 1978.
 - 40) LeCun, Y. and Bengio, Y.: Convolutional networks for images, speech, and time series, *The handbook of brain theory and neural networks*, Vol.3361, No.10, pp.1995, 1995.
 - 41) Syarif, I., Prugel-Bennett, A. and Wills, G.: SVM parameter optimization using grid search and genetic algorithm to improve classification performance, *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, Vol.14, No.4, pp.1502-1509, 2016.
 - 42) Yu, T. and Zhu, H.: Hyper-Parameter Optimization: A Review of Algorithms and Applications, arXiv:2003.05689, 2020.
 - 43) Jones, D. R., Schonlau, M., and Welch, W. J.: Efficient global optimization of expensive black-box functions, *Journal of Global*

Optimization, Vol.13, No.4, pp.455-492, 1998.

- 44) Snoek, J., Larochelle, H. and Adams, R. P.: Practical bayesian optimization of machine learning algorithms, *Advances in neural information processing systems*, Vol.25, 2012.
 - 45) Bergstra, J. and Bengio, Y.: Random search for hyper-parameter optimization, *The journal of machine learning research*, Vol.13, No.1, pp.281-305, 2012.
 - 46) Farrés, M., Platikanov, S., Tsakovski, S. and Tauler, R.: Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation, *Journal of Chemometrics*, Vol.29, No.10, pp.528-536, 2015.
 - 47) Ribeiro, M. T., Singh, S. and Guestrin, C.: 'Why Should I Trust You?': Explaining the Predictions of Any Classifier, *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.1135-1144, San Francisco, 2016.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D.: Grad-cam: Visual explanations from deep networks via gradient-based localization, *Proc. of the IEEE international conference on computer vision*, pp.618-626, 2017.

