

農業支援のための取得容易なデータに基づく
深層学習手法の検討

Investigation of a Deep Learning Prediction Model Based on Easy-to-
Measure Data for Agricultural Support

東京農工大学 大学院 生物システム応用科学府
食料エネルギーシステム科学専攻

本多誠之

Masayuki HONDA

目次

1. 緒言	4
1.1 本研究の背景	4
1.2 本研究の目的	4
1.3 本論文の構成	5
2. 関連研究	6
2.1 光合成モデリング	6
2.2 リモートセンシング	9
2.3 時系列データ予測とその応用分野	11
2.4 本研究の位置づけ	13
3. データセット	14
3.1 イネ個葉ガス交換測定データセット	14
3.1.1 データの計測方法	15
3.1.2 センサーデータの前処理	17
3.2 ドローン画像データセット	19
3.2.3 データの計測方法	19
3.2.4 アノテーション	20
4. 光合成速度データの解析	22
4.1 本章の目的	22
4.2 時系列予測の代表的な手法	22
4.2.1 再帰的ニューラルネットワーク RNN	22
4.2.2 長期-短期記憶 LSTM	23
4.2.3 時系列データ回帰予測の手法	24
4.3 実験の評価方法	26
4.3.1 ネスト化交差検証	26
4.3.2 ホールドアウト検証	27
4.3.3 実験環境	27
4.4 結果と考察	29
4.4.1 モデルの推定および予測精度	29
4.4.2 外部環境要因のみによるモデルの推定・予測精度	32
4.4.3 説明変数が学習モデルの推定・予測精度に与える影響	35
5. BLSTM-Augmented LSTM の提案	38
5.1 本章の目的	38
5.2 BALSTM の詳細	38
5.2.1 提案する予測モデルの構造	38

5.2.2 BLSTM 中間出力の評価方法	41
5.3 実験結果	42
5.3.1 BALSTM の予測精度と再現変数の関係	42
5.3.2 データ入力幅と光合成予測精度の関係	45
5.3.3 予測モデルの可用性	45
5.4 本章のまとめ	50
6. ドローン空撮画像を用いた葉面積推定	51
6.1 本章の目的	51
6.2 関連手法	51
6.2.1 一般物体検出モデル (Object Detection)	51
6.2.2 領域抽出モデル (Semantic Segmentation)	52
6.3 葉面積推定システムの構成	55
6.4 評価方法	56
6.4.1 検出評価指標：平均適合率	56
6.4.2 葉面積評価指標：Intersection Over Union (IoU)	56
6.5 実験結果と考察	57
6.5.1 作物検出・葉領域抽出モデルと作物時期の関係	57
6.5.2 検出モデルの改善	60
6.5.3 葉領域抽出モデルの改善	61
6.5.4 各モデルの組み合わせと圃場全体での葉面積推定	61
7. 結言	62
7.1 研究成果の要約	62
7.2 本研究の将来性	63
謝辞	64
参考文献	65

1. 緒言

本章では、研究の背景と目的、および論文の構成について述べる。

1.1 本研究の背景

光合成とは、緑葉が光エネルギーを利用して二酸化炭素と水から有機化合物を合成する作物生育の根幹をなす生化学的プロセスであり、光合成の速度の実測値は、チャンバーと呼ばれる密閉器を用いたガス交換測定によって測定される。これを高精度で予測可能なモデルを構築することによって、作物の収量や生産性予測が可能になる。このように、近年では様々な計測機器から取得されるデータと情報技術を応用し、農業が抱えている様々な課題に取り組む研究が実施されている。実現場においても、ロボットや人工知能、IoT 機器を駆使した農業形態はスマート農業と呼ばれ、日本においても作物品種や農業形態ごとに多種多様な方法が提案されており、その生産性や収益性について実証実験が行われている段階である。しかし高度な機器の導入や保守には高いコストがかかるため、実用の観点から取得容易なデータの活用が望まれている。

光合成の速度に関して、光量が定常状態である光環境での研究は多数存在する一方で、自然環境下での動的な光量における研究は十分になされていなかった。特に深層学習といった手法を用いて周辺環境の時系列データ、つまり前歴情報を活用したモデリングの研究は少なく、時系列特徴を利用することで予測の精度が向上するのか、また予測に重要な説明変数は何かといった点を明確にする必要があった。また、画像から光合成速度を推定する研究もあるが、これにはハイパースペクトルを用いる試みが多い。しかし、この方法では計測データの取得に高価な装置が必要であり、実用的でないという欠点がある。

また一方で、面積あたりの光合成速度の予測と同様に、作物や圃場単位での葉面積を推定することも収量や成育の予測を実現するにあたって重要である。関連技術の発達により、画像やスペクトルを用いる方法だけでも、室内、野外、定点のものから車や飛行機械に搭載する形式など様々なセンシングについての研究がなされている。特に近年、小型無人飛行機械（UAV）を用いたりモートセンシングは、機器の飛躍的な進歩や安価化、制御の簡便性を要因として活用が進んでおり、モニタリング以外にも施肥や除草などといった応用が開始されている。また、欧米圏の大規模農園を中心とした営農形態と比較して、山間部が多く農地面積の比較的狭い日本ではとりわけ利用機会が拡大している現状がある。

1.2 本研究の目的

前述の背景から、本研究では実用に即した営農支援技術を想定し、そのためにも計測が容易なセンサーデータを用いて個葉の光合成量を予測し、UAV を用いて個葉スケールから株・圃場全体へ拡張する技術確立を目的とした。光合成速度の予測モデルを構築するにあたって、4章で予測に重要な説明変数を分析し、続く5章では時系列変数を計測容易なデータから再現して予測モデ

ルに組み込む機構をもつ BALSTM の提案を行った。UAV によって撮影された RGB 空撮画像だけを用いた葉面積推定モデルの構築を行い、その課題を明確にすることを目的とした研究について 6 章で述べる。

1.3 本論文の構成

本論文は WOS を含む既報論文 2 報[1], [2]と 4 件の学会発表[3]-[6]の内容を統合したものである。本論文は 7 つの章から構成される。第 1 章（本章）では、本研究の背景と目的についてまとめ、また構成の概要を記す。第 2 章では、本研究に関連する諸分野の基礎的な知識と研究事例について述べる。第 3 章では、本研究で利用したデータセットとその構成について報告する。第 4 章では、本研究における光合成速度データに対する分析と予測モデルの構築について記し、第 5 章では計測が容易なデータだけを用いる予測モデルの提案と開発、および結果について報告をする。第 6 章ではドローン画像を用いた研究内容について報告する。最後に第 7 章では本研究成果の要約と将来性について述べる。

2. 関連研究

本章では、本論文に関連する分野である光合成のモデリング、リモートセンシング分野についての基礎的な知識及び研究事例について述べ、加えて時系列データ予測の応用事例として化学工学分野における研究動向について紹介する。

2.1 光合成モデリング

まず光合成とは、光エネルギーを利用して CO_2 を固定し、糖を生成する代謝経路である[7]。光の吸収から糖の生産までのうち、大部分を細胞内の小器官である葉緑体で行われる。葉緑体は二重の包膜で覆われており、内部にはチラコイドと呼ばれる膜構造が複数存在している。このチラコイドの内側をルーメン、外側の空間をストロマと呼ぶ。葉緑体の概観を図 1 に示す。

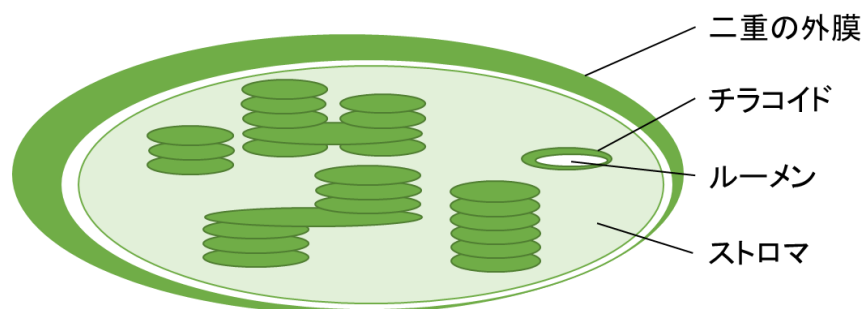


図 1 葉緑体の構造

葉緑体で行われる光合成は、ストロマ内での CO_2 を固定する化学反応群とチラコイド膜で行われる光化学反応群に分類することができる。光合成反応の模式図を図 2 に示す。ストロマでは、内部にあるルビスコ酵素が触媒となるカルボキシル化反応によって、リブローズビスリン酸 (RuBP) が CO_2 と結合し、ホスホグリセリン酸 (PGA) となる。この PGA は炭素を 3 つもつ化合物のため、最初にルビスコによって CO_2 を固定する植物は C3 植物と呼ばれる。PGA は、チラコイド膜における光化学反応の生成物であるニコチンアミドアデニンジヌクレオチドリン酸 (NADPH) による還元とアデノシントリリン酸 (ATP) によるエネルギーを用いてトリオース酸 (TP) となる。TP の一部が葉緑体外に運ばれて、細胞質におけるショ糖合成に利用され、残りの TP は、中間退社物を経て RuBP の生産に再利用されることから、一連の反応はカルビンサイクルと呼ばれている。他にもルビスコ酵素はストロマ内で O_2 を消費することによって RuBP から PGA を生成する酸素化反応も触媒している。こちらは酸素を消費し二酸化炭素を代謝するため光呼吸と呼ばれている。また、世界のより暑い地域では C4 経路と呼ばれる別の経路をもち、光合成の初期には炭素 4 つを含む化合物生成する植物が存在する。C4 経路には様々な種が存在するが、そ

のすべてが PEP カルボキシラーゼという酵素によって初期固定が起こる葉肉細胞と前述の C3 回路をもつ繊維束鞘細胞で構成されている。

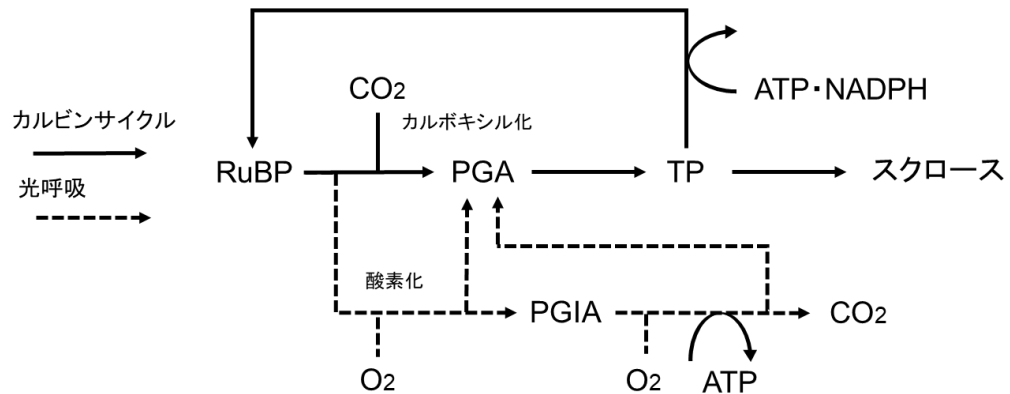


図 2 光合成の模式図

このようにルビスコが触媒する反応はカルボキシル化と酸素化の二つあるため、これらは競合してしまう。どちらの反応が起きるかという頻度、つまり速度はルビスコ周辺の CO_2 と O_2 の濃度や温度によって左右されることになる。しかし、陸上植物が固定する CO_2 のほとんどは大気から供給されるため、植物にこれを能動的に輸送する経路は無く、濃度の勾配による拡散によるのみ移動する。このため植物は葉の表皮に気孔と呼ばれる穴をもち、植物体内への CO_2 の取り入れ口となる (図 3)。取り込まれた CO_2 は葉内部の空間を拡散し、葉肉細胞の表面で水に溶け、細胞壁から、葉緑体、ストロマまで伝達していく。大気から葉緑体まで CO_2 を伝達する際に、これを抑制する段階が複数あり、これらは一般に抵抗やコンダクタンスと表現されている。一番大きな抵抗は気孔の開閉によるもので、完全に閉じた状態だと CO_2 の移動はほとんど起きなくなることが知られている。

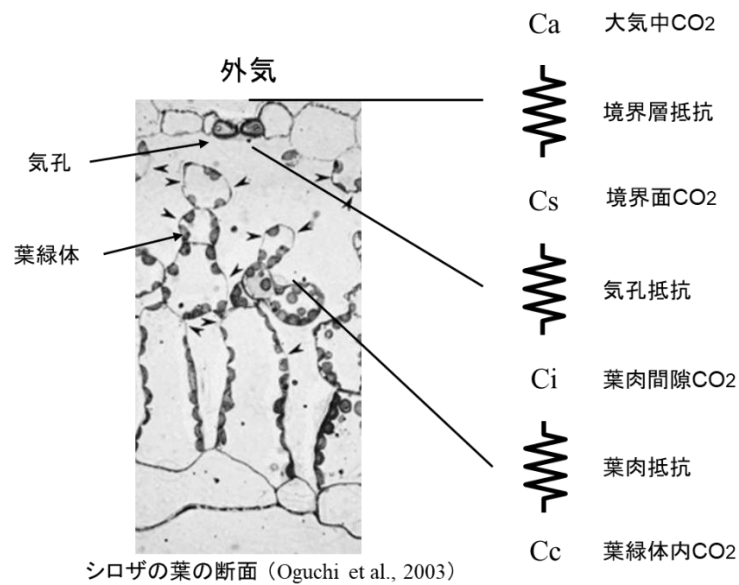


図 3 葉の構造と CO₂ の伝達における抵抗

一方で、気孔は水蒸気の出口でもある。葉肉細胞からは常に水分が蒸発（蒸散）するため、葉内部の相対湿度はほぼ 100%になっており、開口状態でいると著しく水分を失っていくことになる。そのため十分に気孔を開かないと CO₂ を取り込むことは出来ず、また開きすぎると水分を失っていくというジレンマがあり、気孔の開閉は環境に応答して厳密に制御されている。



図 4 左：個葉光合成速度測定装置 (LI-6800; LI-COR, Inc)
 右：群落光合成速度測定装置 (Photosynthesis Cell; Kyowa Co., LTD.)

光合成の速度の計測・推定は、個葉・群落（植物個体）・圃場といった対象とするスケールによって異なる計測方法が用いられている。本研究で取り扱う個葉の光合成速度であれば、チャンバーと呼ばれる空間へのガス入出力の成分濃度を分析するガス交換測定により比較的精密な計測が

できる。植物個体であれば株を密閉温室にて栽培し、ガス交換測定を行うことによって計測できる。野外圃場といった群落での光合成量、炭素同化率はガス交換による測定が不可能なため、多様なセンサー情報から近似する方法が用いられることが多い。

個葉の光合成速度を推定するための代表的なモデルとして光合成の生化学反応と光合成速度の関係を数理的に表した Farquhar モデル[8], [9]がある。このモデルは光エネルギー、炭素同化酵素であるルビスコによる CO₂ 固定反応、トリオースリン酸の再生能力をそれぞれ光合成の律速因子とし、光合成有効放射量、気温、CO₂ 濃度を外部環境要因として光合成速度を決めるものである。光合成に関するさまざまな研究の多くは、定常環境下で得られたデータに基づいて Farquhar モデル内のパラメタリゼーションを実施した後、モデルによる出力結果に基づき光合成能力を調べたものであるが、定常的な光環境は野外において存在せず、自然環境下においては絶えず変化している。

光環境が定常状態での光合成に関する研究はこれまで多数ある[10]–[16]が、自然環境下における動的光に対する光合成応答や光合成誘導反応に関する研究は未だ十分な知見は積み上げられていない。また、これらの研究は、野外環境下や強光と弱光を繰り返すパルス波形を与えた場合の気孔や光合成応答に与える影響を明らかにしようとするものであって、光合成速度を再現・予測するモデリング技術への貢献を主要な目的としたものではない。

画像ベースの推定では、小麦の葉のハイパースペクトル反射率を用いて、二次多項式ステップワイズ回帰 (QPSR)、部分最小二乗回帰、逆伝播ニューラルネットワークの 3 つの方法を採用して、光合成予測に機械学習を適用した研究がある[17]。その結果、QPSR モデルが小麦の光合成炭素同化率 A に対する最良のハイパースペクトルモデルであり、その一次導関数は小麦の葉の 750 ~ 925 nm の反射率の結果として変化することが示された。葉の反射スペクトルを用いて作物の光合成能力を予測するために、幅広い機械学習方法を評価した[18]。彼らは、部分最小二乗回帰がこの予測に最も高い予測力を持つことを観察した。Zhang らの研究[19]では、葉の表現型 (面積、長さ、幅、周囲長、葉のアスペクト比、形状因子など) を使用し、Extreme Gradient boosting (XGB)、サポートベクトル回帰、ランダムフォレスト回帰、一般化加法回帰モデルから樹木 (*Populus simonii*) の光合成速度 A を予測した報告がある。その結果、最も優れた手法は XGB であり、平均絶対誤差 (MAE) は 1.12、二乗平均平方根誤差 (RMSE) は 2.57、R² 値は 0.63 であった。

2.2 リモートセンシング

リモートセンシングとは、非接触、非破壊であることを特徴とした計測技術である。農学分野においては対象の採取や組織液の抽出といった破壊的な計測を経ずに、植生の環境情報の定量化や植物機能を推定することができるため、その重要性が認められている。つまり生存や成長を物理的な環境に依存する農作物において、環境への応答メカニズムを理解し、気候変化に対する応答を御することで将来の収穫量を改善することが可能となる。とくに農業と関係するリモートセ

ンシングは人工衛星や航空機による光学的な画像利用と解釈を主として扱われることが多い。しかし近年では通信技術や計測技術の発展を要因として、近接センシング、つまり設置型のデバイスや車載カメラによる近距離での撮影、低高度での UAV の活用など新しい時空間スケールや用途が検討されている。表 1 に代表的なリモートセンシングのプラットフォームとその特性についてまとめる。各プラットフォームの主な違いは、センサー類と時間、空間的分解能にあるが、それ以外にも利用のしやすさといった点が異なる。また工学分野においても、人の手が入りづらい工場での品質や安全管理、危険な環境下での観測といった用途を目的として様々な計測機器が用いられている。

表 1 プラットフォームと特性

	観測頻度	観測単位面積	代表的なセンサー
人工衛星	数分～10年	1m～1000km	スペクトル
航空機	1日～1月	1m～100m	スペクトル, マイクロ波, ライダー, 画像
携帯・車載	1日～1月	～1m	温・湿度, 画像
定点装置	数分～	～1m	水・窒素包有量, 画像

スペクトルを用いた計測では、近赤外域の反射率と赤色光の反射率を取得することが可能であり、これら数値から植物の分布状況や活性度を示す指標である正規化植生指数 (NDVI) が計算される。NDVI は植生と土壌などの反射応答の差を利用して環境情報を集めるものであり、多くの研究で利用されている。他にマイクロ波はその高い浸透性から、雲や小雨といった大気状態影響されづらく、さらに葉を通過して土壌情報を取得できるといった利点がある。ライダーによる計測は、主にレーザスキャナを用いることによって対象との距離を高い分解能で計測するため、観測対象の空間上の位置情報を 3 次元点群データとして生成し、より精緻に空間的な情報の抽出が可能となる利点を持つ。

指標としては人工衛星や航空機といった上空からのモニタリングでは NDVI が多くの場合で利用されてきたが、これ以外にも葉面積指数 (LAI) やクロロフィル量といった数値も植生を表す数値としてよく用いられ、バイオマス量、群落窒素量、収量といった重要な指標との関係性が研究されている。また一般に、群落に対して用いられるリモートセンシングの植生指数は、耕地ごとの場合や対象地域を任意大きさの区画に分解したうえで計測される場合が殆どである。これに対して植物個体ごとのセンシングであれば、室内栽培といった環境での定点装置による計測が主流である。

このようにスケールが大きいほど農業的に応用可能な範囲が広がる一方で、正確な値を得るま

での技術的障壁が高く、多くのリモートセンシング技術が研究されている。そこで近年、古典的な機械技術や観測手法に対し、情報技術を駆使して作物生産に関わる複雑な要因を時空間的に分析し、意思決定を支援する営農戦略に生かそうとする精密農業研究が進展している[20]–[23]。また、農業が有する様々な課題に AI 技術を適用した研究が行われ始めており、例えば、ドローンやリモートセンシングによる圃場画像からの生育状況の把握や病害虫の検知などに機械学習技術を応用した研究が多数実施されている[24]。

2.3 時系列データ予測とその応用分野

時系列データは、センサーなどによって取得される情報を特定の時間間隔で観測した集合である。とくに互いに関連する複数の時系列情報が記録されたものを多変量時系列と呼び、可視化・情報抽出・予測や制御といった目的のもと多くの解析方法が存在する。本節では関連研究として著者が従事した化学工学分野における時系列データ予測の応用事例について述べる。

まず化学工学とは、物理・化学現象を利用して工業的に目標となる化合物を生産することを目指しており、その成果は燃料精製から医薬品、材料の生成など多岐にわたっている。化学工学の広範な学術領域のなかでも、大規模な化学製造プラントの設計や制御について研究を行うプロセスシステム工学分野では、プラントの高い安全性と効率性の実現を目標としている。このため、プラントにおいて流量を変化させる弁の開閉度合いや攪拌機の回転速度といった制御・入力信号と、各種センサーから得られる温度や圧力といった出力信号を説明変数として、プラントの状態を時間的にモデリングし、推定することによって制御に役立てる試みがなされている。近年、産業界においても計算機の発展と低価格化、IoT デバイスの登場により様々なデータが取得できるようになったことから、モデルベース制御 (MBC) の重要性が認識され始めている[25], [26]。モデリングの手法は、説明性の観点から以下に示す3つに分類されることが多い。

ホワイトボックスモデリング：これは第 1 原理モデリングや物理モデリングと呼ばれ、運動方程式や化学反応式などの科学原理に基づいてモデリングを行う方法である。対象システムでの現象を主に非線形・微分方程式で記述することによって挙動を再現し、計算機を用いた解析やシミュレータの構築に役立てられている。図 5 は Tennessee Eastman process (TEP)[27] と呼ばれる著名な化学プロセスモデルである。数理モデルを用いて制御変数と成分濃度や生成量の時間的な関係が厳密に表現されているため様々な研究においてシミュレーションされたデータが活用されている[28]。一方で、プラントの構造は非常に複雑である場合が殆どで、TEP のようにモデルを完全に記述し制御系の設計にそのまま利用することは難しい。また現実の設備やプロセスに合わせて調整が必要となるため、精度が必ずしも高くないといった問題がある。

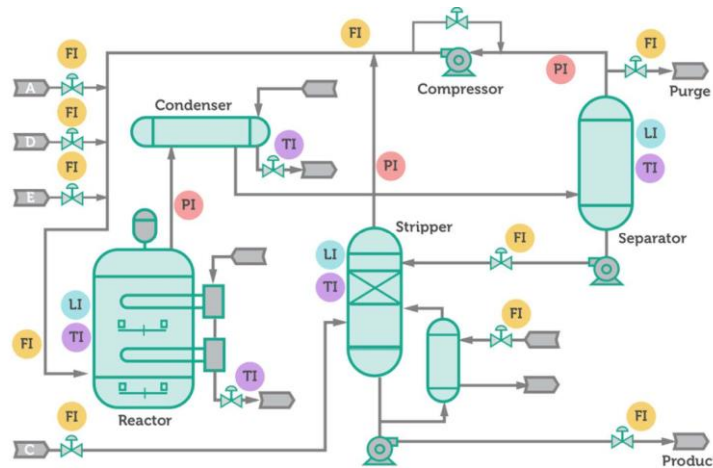


図 5 TEP の制御モデル (引用文献[27])

ブラックボックスモデリング：プロセスをブラックボックスと見なし，その制御信号やセンサーの出力データから統計的方法でモデリングを行う手法である．こちらは複雑なシステムに対しても実験データから有用な情報を抽出することによって比較的簡素かつ高精度なモデルを構築することができ，これらの手法開発は重要な課題の一つとされている[29]．しかしながら統計的手法であるため対象システムの実データの集約が必要であり，プロセスとの対応関係を明確に表現できない説明性の低さといった欠点がある．

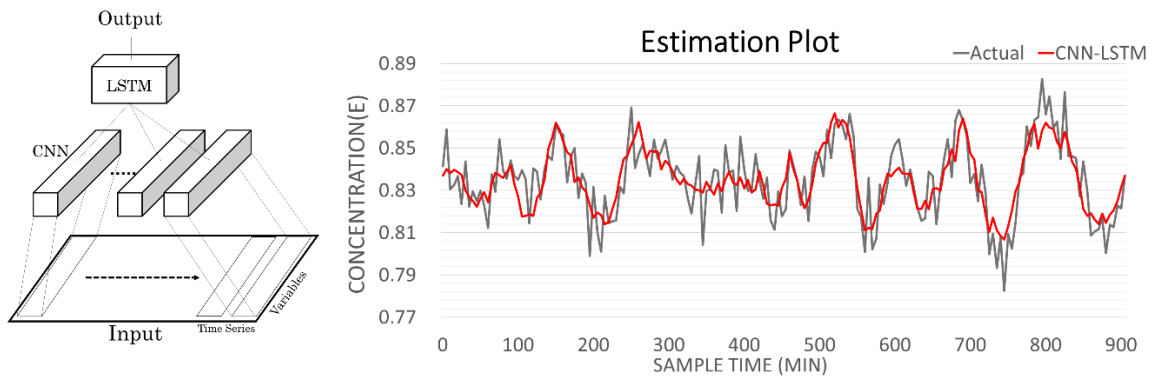


図 6 TEP データセットでの成分推定モデルと予測プロット (引用文献[3])

グレイボックスモデリング：前述したホワイトボックスモデルとブラックボックスモデルの特性を両方とりこみ，プロセスへの説明性を確保しつつ精度向上を高めようという試みが注目されている[30]．これには様々な方法が存在するが，主要な3つの形態に分類することができる．1つは物理モデルによる予測結果に対して，その誤差を実際のデータを用いた統計モデルで補正する方法である．第2に，物理モデルのパラメータを統計モデルの値で設定するという方法が存在する．これはプロセスの定性的，定量的な特性は数理モデルで表現できている一方で，表現するた

めに必要なパラメータの実測値が取得しづらいといったケースに有用である。第 3 に、実際に用いられるのは統計モデルだけであるが、その入力に際してプロセスの物理的特性を明示的に対応させることで、モデルの説明性を向上させる方法が存在する。

2.4 本研究の位置づけ

先述の関連研究を踏まえて、本研究の位置づけについて述べる。光合成速度のモデリングは、数理モデリングによる推定が主だった手法であり、時系列情報を用いたモデリング、とくに機械学習手法を用いた統計的モデリングのアプローチは無かった。よってこれに、時系列データを利用する予測手法を適用することで、まずは高い精度とその要因となる変数を特定する。しかし実用に際しては、説明変数の計測にはガス交換測定といった時間・人的コストがかかってしまうため、予測に重要な変数を比較的容易に計測が可能な環境データで再現し、光合成速度予測モデルに組み込むことで、環境データだけを用いる制約と高精度の両立を目指す。これは先述のグレイボックスモデルに近いが、変数再現と予測の両モデルに統計的手法を当てはめる点で異なっている。そして最後に、個葉単位だった光合成速度予測モデルを時空間的に拡張する技術として UAV 空撮画像を用いた葉面積推定モデルを検討する。葉面積の推定にあたっては、空撮において検討されてこなかった植物個体の検出を行うモデルと葉領域を抽出するモデルを組み合わせ、取得が容易な画像データだけを用いて実装する。検出と領域抽出の 2 段階としたのは、高精度で検出することが可能になれば、定点観測や車載カメラで実施されてきた個別の作物モニタリングを UAV による空撮で実現することが可能となり、また病理診断といったその他の画像利用に着目した手法と組みあわせることが可能となるためである。

3. データセット

3.1 イネ個葉ガス交換測定データセット

本研究では、東京農工大学付属広域都市圏フィールドサイエンス教育研究センターが所有するフィールドミュージアム府中本町水田試験圃場 (139.47° E, 35.67° N)で 2017 年に栽培したイネ (コシヒカリ)を対象として個葉のガス交換測定を実施した. 測定は植物光合成解析システム (LI-6800; LI-COR, Inc.)を用いて、6月から8月にかけて計16枚の最上位完全展開葉を対象とし、1分サンプリングでおおよそ日の出から日の入りまで行った. 表2に各個葉の測定日、測定開始・終了時刻およびログ点数を示す.

表2 本研究で実施した個葉のガス交換測定日時

Year	Cultivar	Month	Day	Start	End	Log point
2017	Koshihikari	June	26	7:01 a.m.	7:00 p.m.	719
			27	6:33 a.m.	2:16 p.m.	463
			28	1:11 p.m.	6:14 p.m.	303
			29	6:42 a.m.	6:13 p.m.	691
		July	11	7:52 a.m.	6:40 p.m.	648
			13	8:19 a.m.	6:54 p.m.	635
			14	7:07 a.m.	7:40 p.m.	753
			15	7:57 a.m.	6:39 p.m.	642
		August*	21	7:15 a.m.	7:06 p.m.	711
			22	7:41 a.m.	6:01 p.m.	620
			23	7:33 a.m.	6:25 p.m.	652
			24	7:29 a.m.	6:20 p.m.	651
		August*	21	7:17 a.m.	6:30 p.m.	682
			22	7:45 a.m.	6:50 p.m.	665
			23	7:47 a.m.	6:28 p.m.	641
			24	7:45 a.m.	6:42 p.m.	657
2019	Koshihikari	July	5	7:48 a.m.	1:48 p.m.	360
	Akitakomachi		6	7:57 a.m.	2:57 p.m.	420
	Koshihikari		17	8:05 a.m.	6:01 p.m.	571
	Akitakomachi		18	7:57 a.m.	2:45 p.m.	401
	Akitakomachi		21	7:57 a.m.	6:03 p.m.	558
	Akitakomachi	August	2	8:00 a.m.	6:01 p.m.	578
	Koshihikari		3	7:53 a.m.	6:01 p.m.	580
	Akitakomachi		17	9:59 a.m.	6:07 p.m.	467
	Koshihikari		27	11:32 a.m.	6:04 p.m.	376
	Koshihikari		29	8:31 a.m.	6:04 p.m.	550
Akitakomachi	September	1	8:39 a.m.	6:06 p.m.	547	
Takanari		16	1:33 p.m.	6:02 p.m.	259	
Takanari		17	10:02 a.m.	6:01 p.m.	461	

3.1.1 データの計測方法

図 7 に示すように測定の際は、チャンバー光源を取り外して自然光を入れることによって外部光を葉に照射させ、チャンバー内の気温・湿度は、群落近くに設置した温・湿度計 (TR-72wb; T&D 社) の値を 1 時間ごとに読み取り、手動によりその値をコンソールから入力することにより更新した。なお、本研究では、表 3 に示す光合成速度 A を目的変数、それ以外を説明変数とした。



図 7 ガス交換測定の様子

表 3 本研究で使用した目的変数および説明変数

変数名	単位	説明
A	$\mu\text{mol m}^{-2} \text{s}^{-1}$	光合成速度
E	$\text{mol m}^{-2} \text{s}^{-1}$	蒸散速度
C_a	$\mu\text{mol mol}^{-1}$	大気中（チャンバー内）の CO_2 濃度
C_i	$\mu\text{mol mol}^{-1}$	葉内 CO_2 濃度
g_{sw}	$\text{mol m}^{-2} \text{s}^{-1}$	水蒸気の気孔コンダクタンス
g_{bw}	$\text{mol m}^{-2} \text{s}^{-1}$	水蒸気の葉面境界層コンダクタンス
g_{tw}	$\text{mol m}^{-2} \text{s}^{-1}$	水蒸気のトータルコンダクタンス
g_{tc}	$\text{mol m}^{-2} \text{s}^{-1}$	二酸化炭素に対する合計コンダクタンス
SVP_{leaf}	kPa	葉温における飽和水蒸気圧
RH	%	相対湿度
VPD_{leaf}	kPa	葉温における飽差
LHF	W m^{-2}	潜熱フラックス
SHF	W m^{-2}	顕熱フラックス
Q_{in}	$\mu\text{mol m}^{-2} \text{s}^{-1}$	葉面上の光合成有効放射束密度
Q_{abs}	$\mu\text{mol m}^{-2} \text{s}^{-1}$	葉に吸収された光合成有効放射束密度
$Q_{\text{amb_in}}$	$\mu\text{mol m}^{-2} \text{s}^{-1}$	チャンバー内の光合成有効放射束密度
T_{air}	$^{\circ}\text{C}$	チャンバー内の気温

3.1.2 センサーデータの前処理

実測データ内に含まれる変数の外れ値や変数間の異なるスケールはその変数の特徴を失わせる要因となり、構築するモデルによる予測精度に悪影響を及ぼすため、モデルの入出力として用いる前に外れ値処理とスケーリングを行う必要がある。本研究では、ガス交換測定データ内の時刻 t における説明変数 x_t に対し、データ内における平均 μ 、標準偏差 σ が次の条件式を満たす値を外れ値とした。

$$|x_t - \mu| > 2\sigma \quad (1)$$

式(1)の条件で検出された外れ値を削除後、その直前の値と直後の値を用いた 1 次多項式によって線形補間を行いデータの補間を行った。機械学習技術を用いて予測モデルを構築する際の重要なテクニックの一つとして、モデルの入力値に用いる説明変数の値を基準に沿って変換することにより、尺度を揃えるスケーリングがある。スケーリングは、説明変数の値がもつ単位が異なることによって説明変数間で桁の異なる数値が存在し、このため相対的に小さな、あるいは大きな入力値が予測モデルに強く影響されてしまう影響を排除するために実施される。本研究では、次式を用いてスケーリングを行い、各説明変数が取りうる範囲によるモデルへの影響を排除した。

一般的な回帰モデルでは、説明変数の中に相関係数が高い組み合わせがあるときに多重共線性、つまり回帰係数の分散を増加させて係数が不安定になるという問題が生じる[31]。このため、相関係数が高いどちらか一方の説明変数を削除するか、変数の適切な選択を実施することが推奨されている[32]。ニューラルネットワークを用いた予測モデルの構築においても、説明変数の相関が極端に高い組み合わせは、学習の際にパターンの特徴を適切にとらえることを難しくする。また、計算コストの観点からも説明変数の適切な削減および選択が望ましい。本研究では、表 3 に示す全ての変数に対し、同時刻における各変数間の相関係数を求めた(図 8)。その結果、説明変数である $E, g_{tw}, g_{tc}, SVP_{leaf}, VPD_{leaf}, Q_{abs}, Q_{amb_in}$ がそれぞれ $LHF, g_{sw}, g_{sw}, T_{air}, RH, Q_{in}, Q_{in}$ と高い正および負の相関を有していたため、モデルに多重共線性の問題が生じないようにこれらの説明変数をモデル学習の際に用いないようにした。ここで採用した後者の変数群は、前者と比較して予測対象の光合成速度 A に対する相関係数の絶対値が大きいものである。

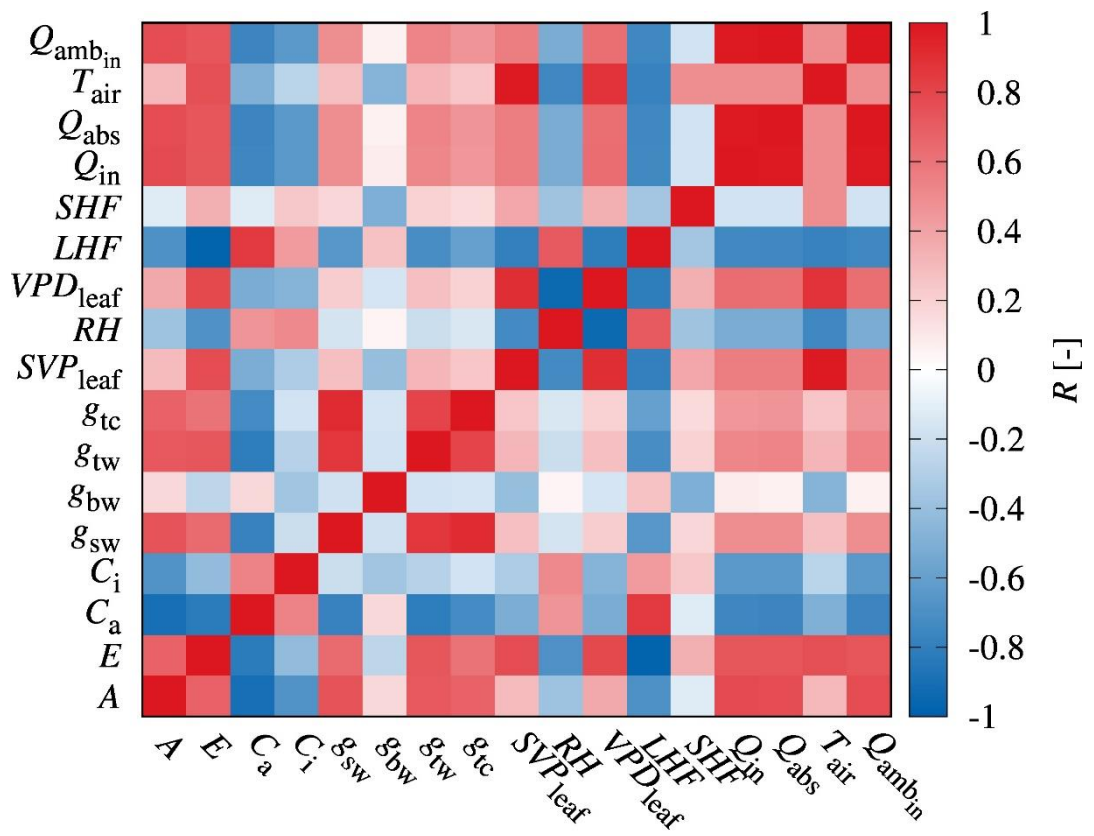


図 8 変数間の相関係数ヒートマップ

3.2 ドローン画像データセット

3.2.3 データの計測方法

実験対象とする農作物は、東京農工大学府中キャンパス内にある FM 府中の野圃場にて栽培したトマト、タロイモとした。これらに対して UAV を用いて 2020 年 5-7 月にかけて一週間おきに高度 12,18m から撮影し、取得した空撮画像(静止解像度: 5280 x 3956)から深層学習用のデータセットを作成した。このデータセットは物体検出と領域抽出の 2 種類の形式でラベル付けを行っている。

撮影日 1 日につき生成されるデータセットの内容を表 4 に示す。これが合計 10 回の撮影日分作成した。

表 4 撮影日あたりの作成したデータセット数

原画像 5280x3956 (計5枚)							
学習用							評価用
4枚							1枚
検出用モデル 1024x1024				抽出モデル 256x256			
	学習用	評価用	計		学習用	評価用	計
トマト●	68	4	80	トマト●	18	2	20
トマト○		4		トマト○	18	2	20
サトイモ		4		サトイモ	18	2	20



図 9 撮影時の原画像

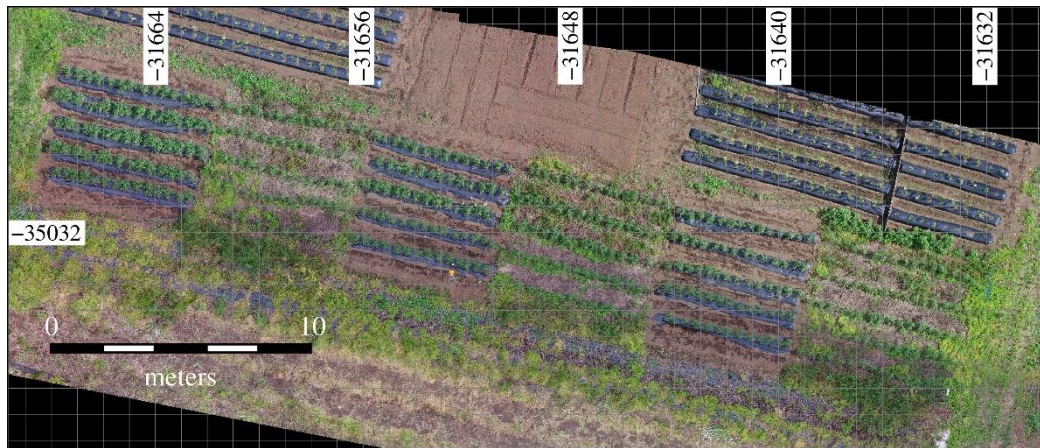


図 10 実験圃場の全体のオルソ画像

3.2.4 アノテーション

オープンソースのアノテーションツールである CVAT を用いて、元画像に対する検出領域・葉領域に対するラベル付けを実施した。

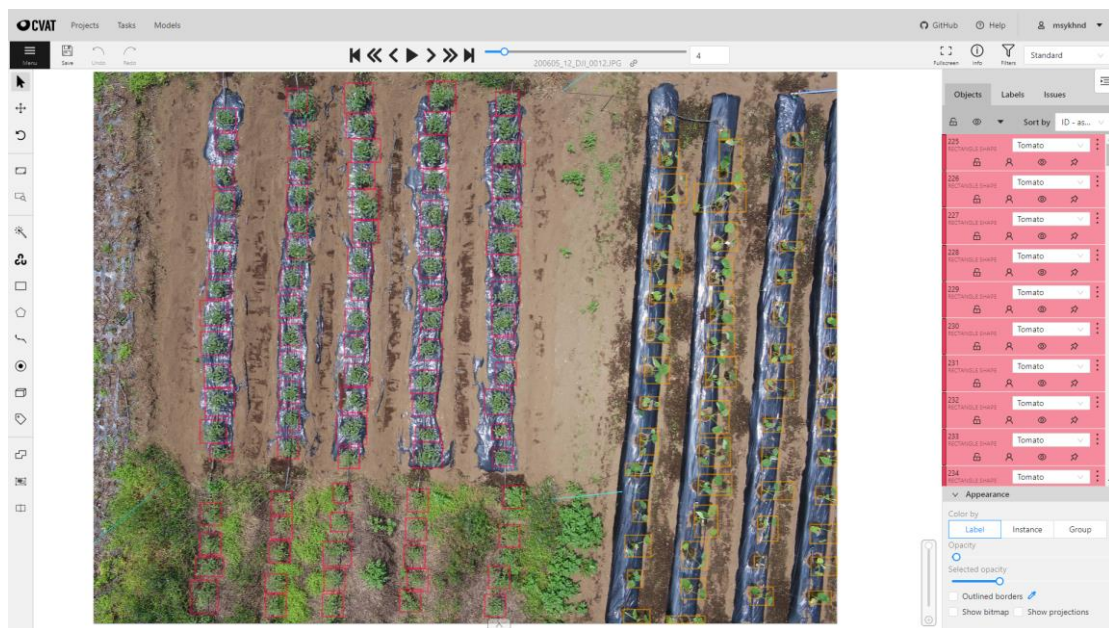


図 11 検出データセット用ラベリング

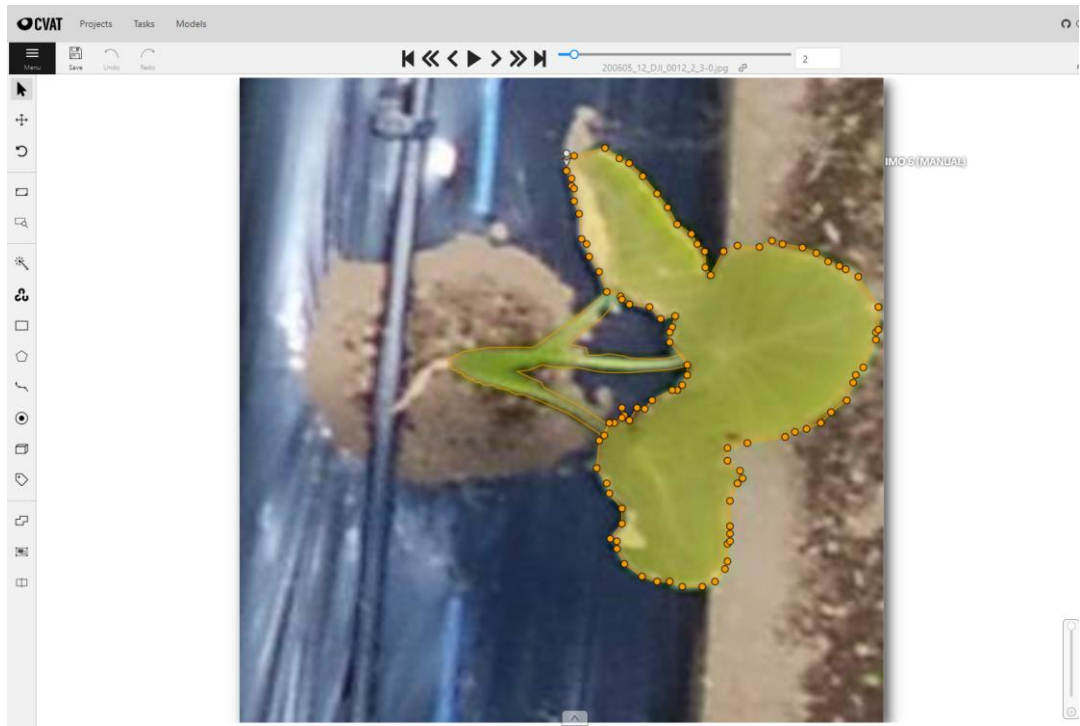


図 12 抽出データセットのラベリング

4. 光合成速度データの解析

4.1 本章の目的

先の章で紹介したように光合成研究と時空間分析に関わる技術は機械学習手法を含め進展を遂げているが、非定常環境下での個葉の光合成速度の予測に機械学習技術を適用した研究はほとんどなかった。また、その予測に必要な情報を探索し、課題を明らかにすることは、作物生長モデル、センシング、営農モデルなどの情報技術を軸とする意思決定支援ツールエンジンを構築するための一つのモジュールとして有用な知見となる。本章では、時間とともに変動する光環境下における光合成速度に対して、機械学習による予測技術に着目し、リカレントニューラルネットワークの一種である Long Short-Term Memory Neural Network (LSTM) を光合成速度の律速要因となる説明変数の時系列データを入力として適用した。データは野外圃場においてガス交換測定を実施して得られた実測値を利用する。そして、外部気象環境および葉内環境の長期情報、つまりこれらの前歴を学習に用いることが予測精度の向上に結び付くのか、十分な予測精度を保つことができる予測期間はどの程度なのか、機械学習による光合成速度の予測モデルを使用する上で重要な説明変数や課題は何かを同定することを目的とした。

4.2 時系列予測の代表的な手法

ニューラルネットワークによるイネ個葉の光合成速度の予測モデルには、再起的な仕組みを持たない一般的な3層の全結合型ニューラルネットワーク (FC: Fully Connected Neural Network)、代表的な再帰的ニューラルネットワークであり自己回帰型の構造をもつニューラルネットワークである Recurrent Neural Network (RNN) および RNN が長期間の時系列情報の反映が難しいという欠点を補い、説明変数の長い時系列的変化特徴を捉えることができる LSTM の3種を用いた[32]。さらに、両ニューラルネットワークモデルとの比較のため、主成分回帰 (PCR: Principal Components Regression) [33]、ランダムフォレスト回帰 (RFR: Random Forest Regression) [34]を用いた。これらの回帰モデルはニューラルネットワークベースの手法と異なり、時系列データの長期時間依存性やその変動などの動的特徴を捉えることができない。光合成速度がどの説明変数あるいはその前歴情報に基づいて主に決定されるかを調べるために、これらの回帰型モデルを光合成速度 A の推定・予測精度の比較対象として用いた。以下にそれぞれのモデルについて概略を述べる。

4.2.1 再帰的ニューラルネットワーク RNN

図2にRNNの概要を示す。通常のニューラルネットワークでは、入力層においてすべての値を入力データとして受け付け、中間層にて人間のニューロンに相当するノード間で重み計算を行

い、出力層において結果を出力する。RNN は前回の出力を現在の入力に追加するものであり、時系列データを 1 サンプルずつ受け付ける点が特徴である。RNN では時刻 t での説明変数 X_t と過去のノード出力である重み行列 H_t を受け付け、重み計算を実施する。ここで計算ノードの出力を次の時刻における入力の一部として再度扱う。つまり RNN の中間層のノードは、現在のサンプルからの入力と、過去の隠れ層の状態を表現する重み行列 H_{t-1} を受け取る。以上のように、RNN は自己ループを備えたニューラルネットワークであり、自己ループを時間軸方向に展開することによって過去の情報を基に予測させることができる。一方で、同じ重みを過去から現在の入力にかけて適用するため、系列が長くなると重みがかかる回数が多くなり、勾配の消失/爆発の問題が発生するため、時系列値の長期依存を学習できなくなるという課題が生じる。

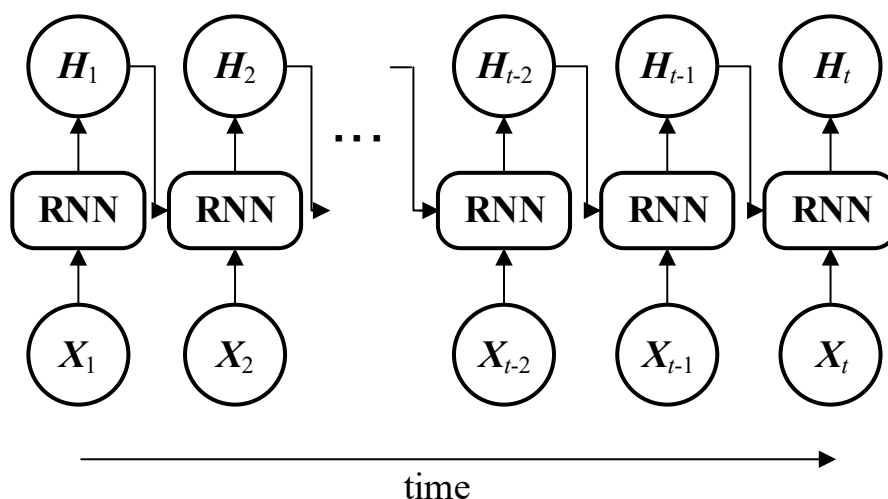


図 13 RNN の基本構造

4.2.2 長期-短期記憶 LSTM

長期記憶の困難性や勾配消失が起りやすいなどの問題点を解決するため、長い時系列データに対して学習可能なモデルとして LSTM[35]がある。LSTM は RNN と同様過去から現在にかけての情報伝達構造を有しているが、伝達する重み形式が異なる。その最大の特徴は過去のノード出力である重み行列 H_t に加え、記憶セル C_t と呼ばれる状態保存のための隠れ層を入れることであり、長期記憶や勾配消失を避ける役割を果たしている (図 3)。LSTM の内部には、ネットワーク内の情報伝搬のコントローラとして、記憶を保つ内容を選別するための忘却ゲート f_t 、各時刻においてどの程度新しいデータを反映するか、または出力するかについて制御するための入力ゲート (i_t, g_t) および出力ゲート o_t が存在し、接続されたノードの状態に基づいて重みを直接上書きする単純な RNN とは異なる構造を有している。これらゲートは特殊な構造で相互作用する 4 つの計算ユニットで表現されており、一連の重みと活性化関数が含まれている。 f_t, i_t, g_t, o_t はそれぞれ以下のように求める。

$$\left\{ \begin{array}{l}
 f_t = \text{sigmoid}(X_t W_{X,f} + H_{t-1} W_{H,f} + b_f) \\
 i_t = \text{sigmoid}(X_t W_{X,i} + H_{t-1} W_{H,i} + b_i) \\
 g_t = \text{sigmoid}(X_t W_{X,g} + H_{t-1} W_{H,g} + b_g) \\
 o_t = \text{sigmoid}(X_t W_{X,o} + H_{t-1} W_{H,o} + b_o) \\
 C_t = (C_{t-1} \odot f_t) \oplus (i_t \odot g_t) \\
 H_t = o_t \odot \tanh(C_t)
 \end{array} \right. \quad (2)$$

ここで、sigmoid はシグモイド関数、 W は学習する重み行列のベクトル型パラメータ、 b はバイアスのベクトル項である。 \oplus は記憶セルへの追加記号、 \odot は要素積である。

f_t, i_t, g_t により定義される C_t は、直前の記憶セルに対して不要な情報を f_t で処理後、選別された情報を追加している。これら一連の記憶の制御構造によって、LSTM ユニット自身が記憶能力を持ち、中間層に使用することで長期の情報を制御することにより、長い時系列データの特徴を考慮することができる。

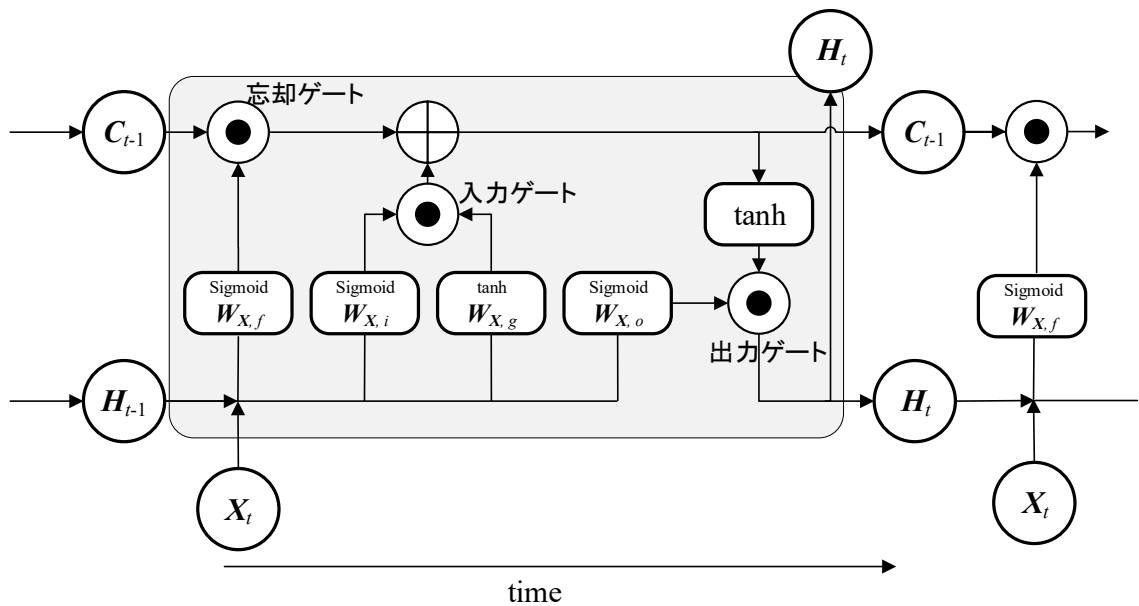


図 14 LSTM の基本構造

4.2.3 時系列データ回帰予測の手法

今回扱う光合成速度を決定づける時系列データは、直前の値に対する相関係数や説明変数間の相関係数が高い組み合わせがあるため、回帰式において適合する超平面が求められなくなる多重共線性問題が生じる[31]。これを緩和することを目的とした回帰分析手法として PCR が挙げられ

る。PCR は回帰式を求める前に、回帰式を求めやすい成分へと説明変数を変換する過程を持っており、説明変数の分散が最大となる処理がそれぞれ行われる。また、RFR は線形モデルを用いないため、多重共線性による問題が起きにくく、時系列データを用いた予測に利用されている。なお、PCR、RFR による回帰分析では、ある 1 時点の説明変数だけを用いた瞬時値による予測と、RNN や LSTM の入力のように連続する時系列データを入力とする予測を行った。

4.3 実験の評価方法

4.3.1 ネスト化交差検証

本研究で使用した予測モデルの作成に利用できたログは計 10,133 点存在するが、計測期間が約 3 か月と長く、気温や光環境も多様であるため、光合成特性が大きく異なる。したがって、検証データを特定の日に測定された個葉データに固定すると、学習や調整に用いた個葉データの影響により予測モデルが適切に評価できなくなる。そこで本研究では、予測モデル評価にはネスト化交差検証[36]を用いた。ネスト化交差検証は、分割したデータセットの役割を学習用、調整用、検証用としてそれぞれ遷移させることによって偏りをなくした予測モデルの検証を行う手法である(図 4)。

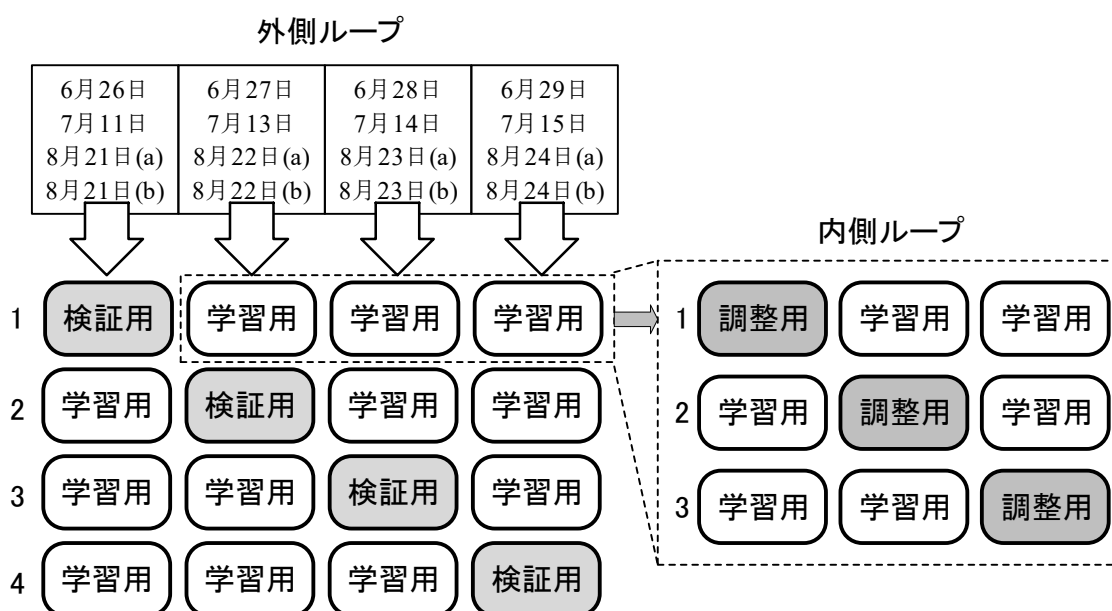


図 15 ネスト化交差検証

最初に、ガス交換測定で得られた全データを各月の個葉データが均等になるように配分し、検証データと学習データに分離する。学習データ内は、さらに学習用と調整用データセットに分割する。学習用データを用いて学習させ、調整用データにより学習誤差が最小となるハイパーパラメータの最適値の探索および学習の過学習検知と早期終了判断を行う。以上によって作成された予測モデルを検証データによって評価するが、学習と調整データの偏りによる影響を排除するため、学習データと調整データの役割を遷移させて予測モデルを構築する。次に、検証用データに含まれる偏りによる影響を排除するため、検証用データを遷移させ、得られた検証用データごとの評価の平均を最終的なモデルの評価値とした。本研究では、モデル精度検証に用いた時系列データの長さを 5 分間とし、1 分ずつずらしていくことにより、各時系列データにおける推定値お

よび予測値を求めた (図 16). なお, モデルの精度評価に R^2 (決定係数) と MAE (平均絶対誤差) を用いた.

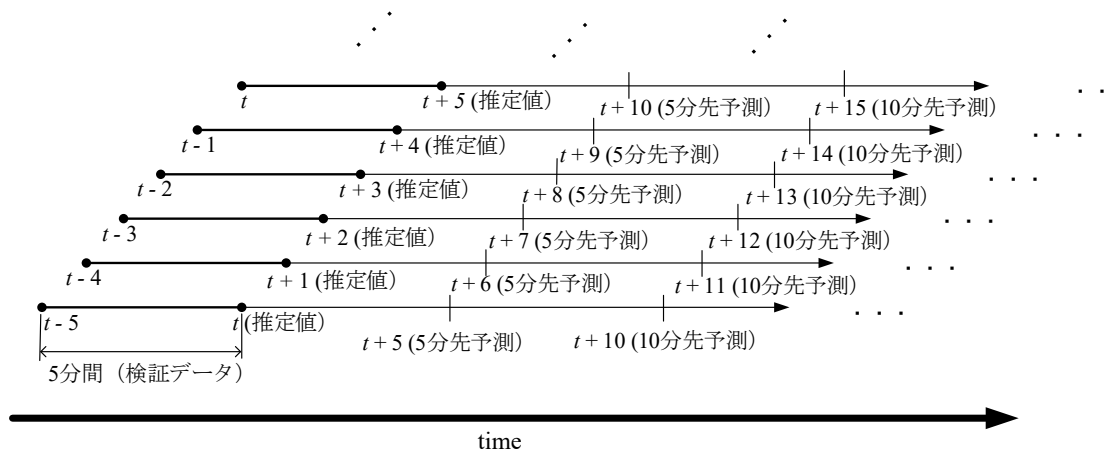


図 16 モデルによる推定値と予測値の関係

4.3.2 ホールドアウト検証

実験では一般的にホールドアウト検証と呼ばれる学習用データ, 調整用データ, 検証用データを固定のものとする検証も並行して行った. これは予測モデルが実際にどの程度予測できるかを図示するために用いた. 前述したネスト化交差検証では, 個葉データの影響を排して予測モデルの評価を定量的に実施することが可能であるが, 複数の予測モデルの生成を必要とするために, 実際の予測がどのような結果になるかを可視化することができない. しかし, ホールドアウト検証では各データを遷移させないため検証を通して生成される予測モデルは一つであり, 固定の検証用データに対して推定結果が得られるため, 予測モデルの挙動を直感的に比較, 解釈しやすいという利点がある.

4.3.3 実験環境

本研究で用いたモデルには, 学習させる前に設定が必要なハイパーパラメータが存在する. FC, RNN, LSTM であれば中間層や学習係数, PCR では主成分分析時に抽出する次元数, RFR であれば決定木の数が挙げられる. これらのハイパーパラメータの最適値を求めることは非常に難しいため, 事前いくつか用意したハイパーパラメータの組み合わせ候補から最も性能の良い組み合わせを選出する手法が一般的である[37]. つまり前述したネスト化交差検証での内側ループにおいて, ハイパーパラメータの組み合わせを変化させながら調整用データに対する平均二乗誤差スコアの平均を求め, 比較することによって最適に近いハイパーパラメータの組み合わせを決定し

た。その後、内側ループで用いた全データと求めたハイパーパラメータを用いて学習を行い、外側ループにて検証用として隔離していたデータを使って評価を行い、これを外側ループで繰り返すことによって最終的な評価値を求めた。ニューラルネットモデルの場合には過学習検出のためのデータを確保する必要があるため、監視のための調整用データを遷移させた複数の検証用データへの平均を内側ループの評価値とした。

本実験はプログラミング言語 Python を用いて実装し、機械学習フレームワークとして Tensorflow + Keras, 機械学習ライブラリとして主に Scikitlearn を利用し、ニューラルネットの学習には NVIDIA 製 GPU RTX 2080 および GTX 1080 を使用した。

4.4 結果と考察

4.4.1 モデルの推定および予測精度

すべての学習用説明変数データを用いた学習済みの FC, RNN, LSTM, PCR および RFR の各モデルに対し、検証用データを用いた光合成速度 A の推定精度結果を表 3 に示す。表より、 R^2 と MAE の両指標において、LSTM による A の推定精度が今回検証したモデルの中で最も高い結果となった。一方、回帰モデル間において時系列値と瞬時値を使った推定精度には大きな差は見られなかったが、RFR が PCR よりも優れた再現性を示した。以上より、 A の日トレンドの再現および実測値との絶対誤差の点において全アルゴリズムの中では LSTM が優位であることが確認できる。

表 5 モデルによる推定精度

Model	MAE	R^2
FC	1.320	0.930
RNN	1.002	0.952
LSTM	0.516	0.977
PCR (瞬時値 ¹)	1.854	0.910
PCR (時系列値 ²)	2.046	0.913
RFR (瞬時値 ¹)	1.557	0.940
RFR (時系列値 ²)	1.561	0.939

¹ 入力値に瞬時値を利用

² 入力値に時系列値を利用

図 6 に 60 分先までの A の予測精度結果を示す。本図は 5 分間の説明変数の値を学習済みのモデルに入力して、その時刻からリードタイム 60 分までの値を 5 分間隔で予測することをタイムステップ 1 分として繰り返し、実測値と予測値を比較することで、モデルの光合成速度の予測精度がどのように遷移するかを調べたものである。最も高い推定精度があった LSTM は、 R^2 指標で 5 分先までの予測はモデル間で最も優位な予測精度を有するが、10 分先以降の予測は瞬時値および時系列値を用いた PCR よりも低くなる。また、MAE 指標では 20 分先の予測までは有意に予測誤差が小さいが、それ以降先の予測誤差は時系列データを用いた PCR による予測結果と同程度と

なる。さらに LSTM は FC と比較して常に優位な精度を保っており、時系列特性を考慮する再帰的なモデル構造の効果が確認できる。RFR は PCR より高い推定精度が得られた(表 3)が、予測精度は LSTM や PCR と比較して低く、 A の予測には劣位であることが確認できた。

以上のことから、少なくとも 20 分先以降の A の予測には、LSTM の説明変数の時系列特性を考慮する利点が減少し、説明変数から予測に貢献する主成分を抽出することが可能な PCR の特性が重要になっていることが示唆された。これらの結果から、ある時刻 A の値はその時刻の 20 分前までの説明変数の情報、つまり外部気象や葉内環境の前歴が予測に大きく影響していることが考えられる。

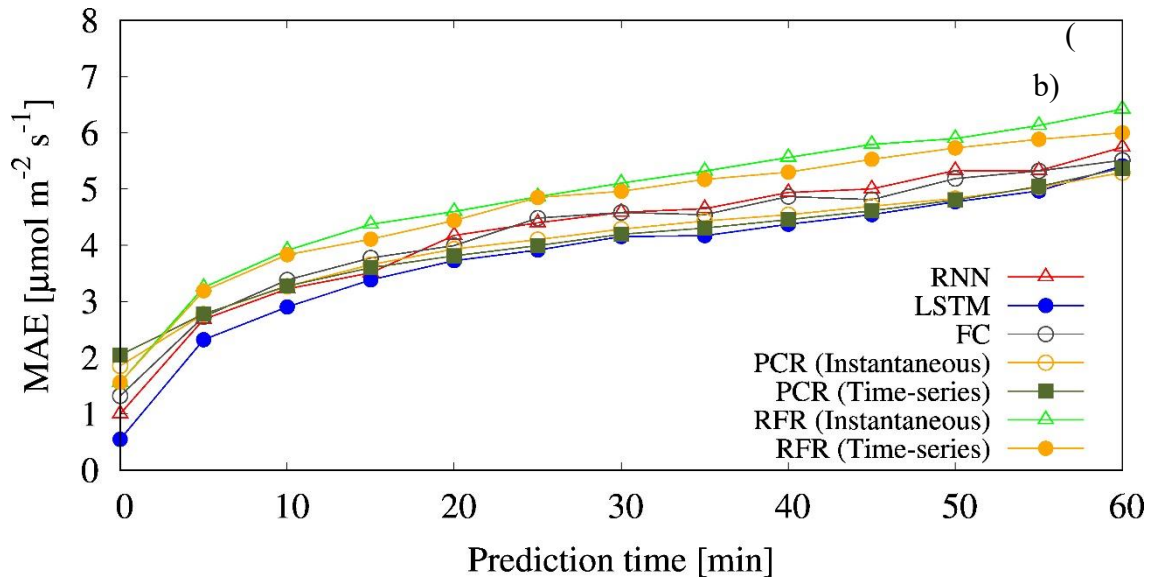
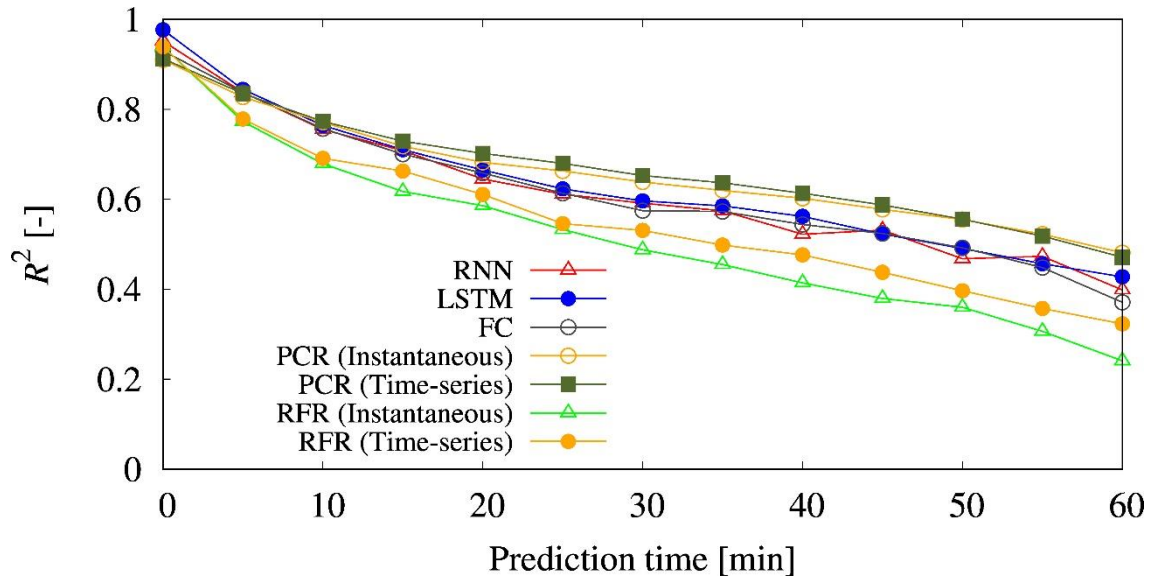


図 17 各モデルによる予測時間と予測精度の関係 ((a) R^2 , (b) MAE)

4.4.2 外部環境要因のみによるモデルの推定・予測精度

前節では、外部環境要因に加え、取得に人的・物的コストが必要な葉内 CO_2 濃度 C_i 、 CO_2 や水蒸気に対する葉内コンダクタンス g を説明変数に入れてモデルを構築した。また、光合成速度は前 20 分程度の説明変数が予測に貢献していることを示したが、外部の気象環境と葉内環境のどちらが予測にとってより重要であるのかが不明であった。本節では、モデルの応用可能性も考え、比較的計測が容易でデータの入手可能性が高い Q_{in} 、 T_{air} 、 C_a のみを説明変数として構築した学習モデルの推定・予測精度を評価した(図 7)。 R^2 と MAE の両指標にみる再現・予測精度は全説明変数を考慮した結果と比較して低下するが、現在値の推定において LSTM が最も優れている一方、5 分先以降の予測精度は時系列値を用いた PCR が LSTM を上回ることがわかる (図 7(a), 7(b))。このことから、前節の結果も含めて勘案すると、20 分程度先までの LSTM による A の予測には、葉の内部における情報が A の決定に強く寄与していることが示唆され、その情報を考慮しなくなったことが予測精度の低下の原因であると考えられる。

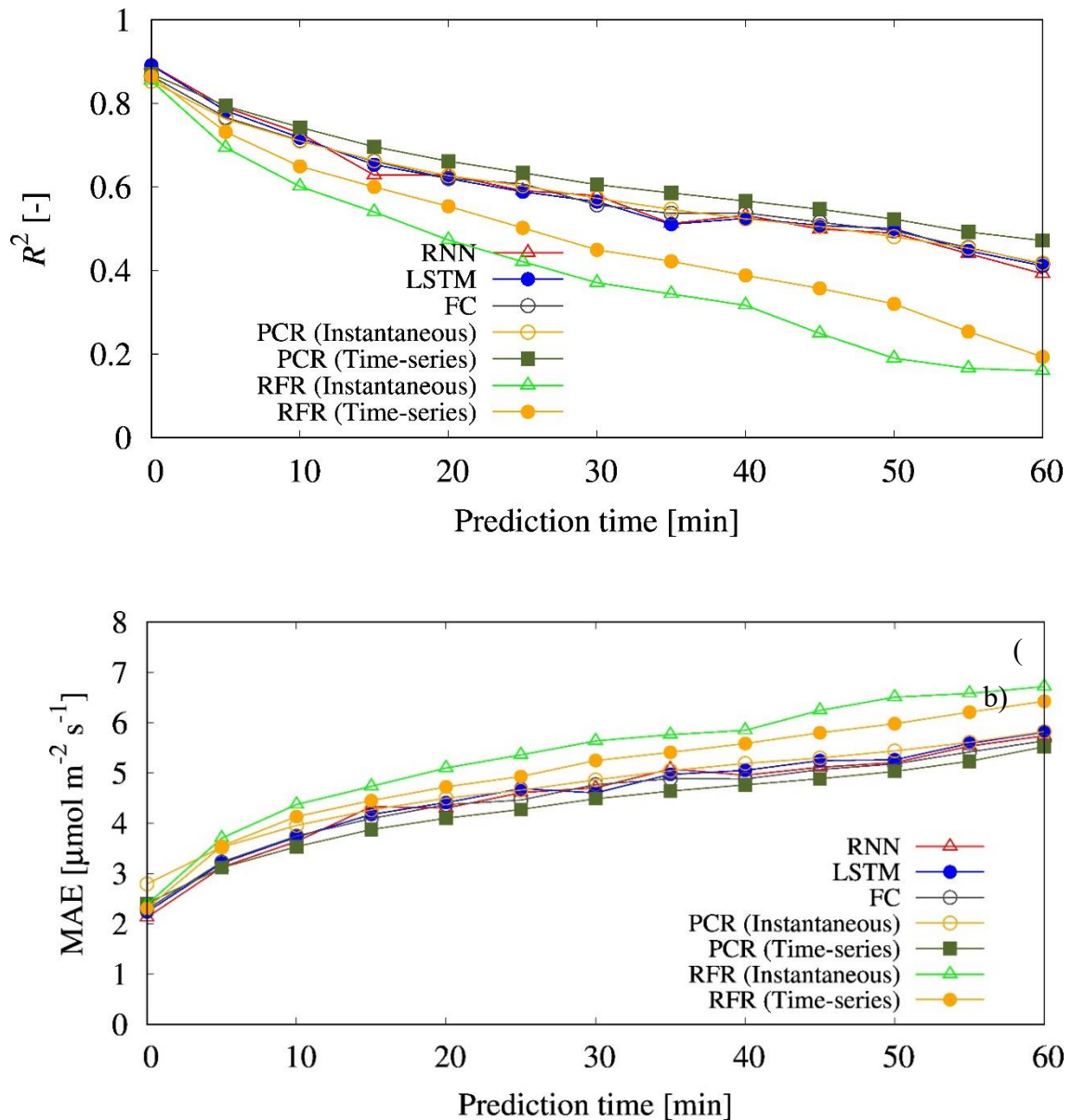


図 18 Q_{in} , T_{air} , C_a を説明変数とした各モデルによる予測時間と予測精度の関係 ((a) R^2 , (b)MAE)

図 19 に学習済みの LSTM による 8 月(b)における A の推定結果を示す. なお, 本図は 6 月と 7 月のデータを学習用, 8 月(a)のデータを調整用, 8 月(b)のデータを検証用とするホールドアウト検証によるものである. なお, ネスト化交差検証の時とは異なり, 計測月ごとにデータセットを分割したのは, 個体の偏りによる影響を排除できないので, 時間的な整合性のある分割の方が解釈しやすいものと考えたからである. 全説明変数を用いて学習したモデルと比較して, Q_{in} , T_{air} , C_a のみを用いた場合の推定誤差は大きく, 全説明変数を用いた場合の R^2 は 0.956, 外部環境要因の

みを説明変数に用いた場合の R^2 は 0.806 であった。外部環境要因のみを用いた推定では、8月21日(b)における結果は実測値との誤差は少ない一方、8月24日(b)におけるモデルによる予測値は実測値を過大評価する傾向を示していることがわかる。 A が外部気象環境因子に加えて、葉室素濃度や葉の光合成ポテンシャルなど他因子の影響を受けることを示唆するものであるが、条件を満たせば外部環境要因のみによる学習モデルでも十分な推定精度を得ることができることを示している。今後の予測モデルを構築する際は、作物の生育ステージや葉室素濃度を分類して学習モデルを作成し、外部気象環境のみを説明変数とした場合に正確に予測できる条件を詳細に分析する必要があると考えている。

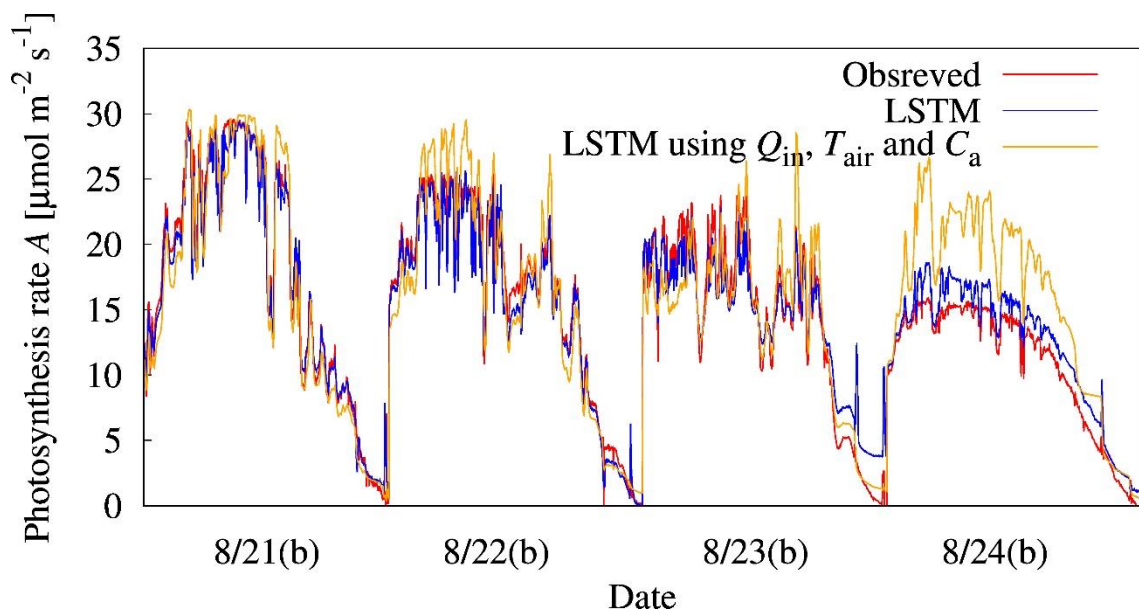


図 19 実測値と予測結果のプロット

4.4.3 説明変数が学習モデルの推定・予測精度に与える影響

ニューラルネットを用いた予測や分類は目覚ましい成果を残す一方で、その本質が統計的なパターンマッチングである以上、学習済モデルはブラックボックスである。そのため予測や分類のプロセスが人間にはわからないという問題を抱えている。機械学習が出した答えがどのような思考により出されたものなのかを定量的に明らかにするための研究がなされているが、現在において画一的な解は得られていない。機械学習の解釈性を提供するアプローチとして、画像分類の分野では、画像領域をヒートマップで表示する手法 [38]が、自然言語処理分野では単語同士の関係性を視覚化することのできる機能が注目されている[39]。しかしながら、時系列予測モデリングにおける統計的アプローチにおいては、時系列値前後の相関の強さや画像分野のような直感的な視覚化が困難なことから、学習モデルの結果の定量的解釈に関する研究はほとんどなされていない。

本研究の学習モデルの予測精度評価結果から、 A を律速する要因に C_i や g の時間変動が強く影響することが示唆された。ここでは、学習に用いた各説明変数が A の予測に与える影響を調べるために、予測モデル作成の際に、説明変数を時系列ごと削除することによってその影響をヒートマップで可視化する。これは予測モデル作成において重要な説明変数を削除した場合に予測精度が低下する仮説に基づくものである。

上記を調べるための予備実験として、単独の説明変数のみで LSTM による予測モデルを作成し A の推定精度を評価した (図 9)。その結果、 Q_{in} , C_a は単体でも R^2 が 0.8 程度となり、相対的に高い推定精度が得られた。

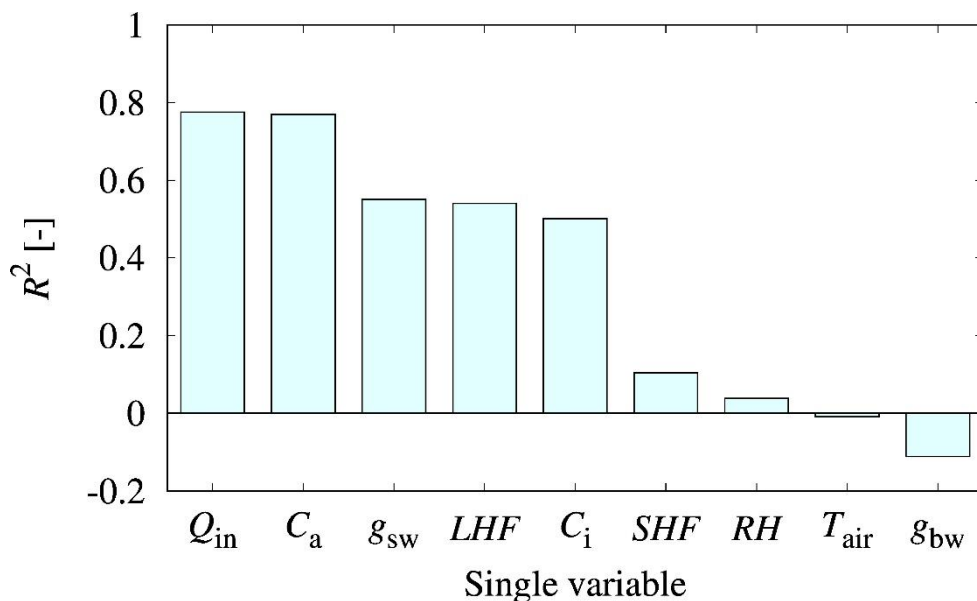


図 20 単変数による光合成速度の再現精度

図 10 は、説明変数を LSTM による学習モデル構築の際に取り除いた際の推定・予測精度への

影響を示したものである。縦軸の変数名は削除した変数、横軸の数字は 60 分先までの A の予測結果を 5 分先ごとに示している。変数削除による影響は、縦軸にある説明変数を削除した残りの変数群で予測モデルを作成して R^2 スコアを計算し、得られたスコアは列ごとに平均 0, 分散 1 になるように標準化して数値化した。つまり、変数削除の予測精度への影響が大きいものは青色の領域で重要度が高い変数とみなすことができ、赤色の領域は相対的に予測に重要でない変数であることを示す。

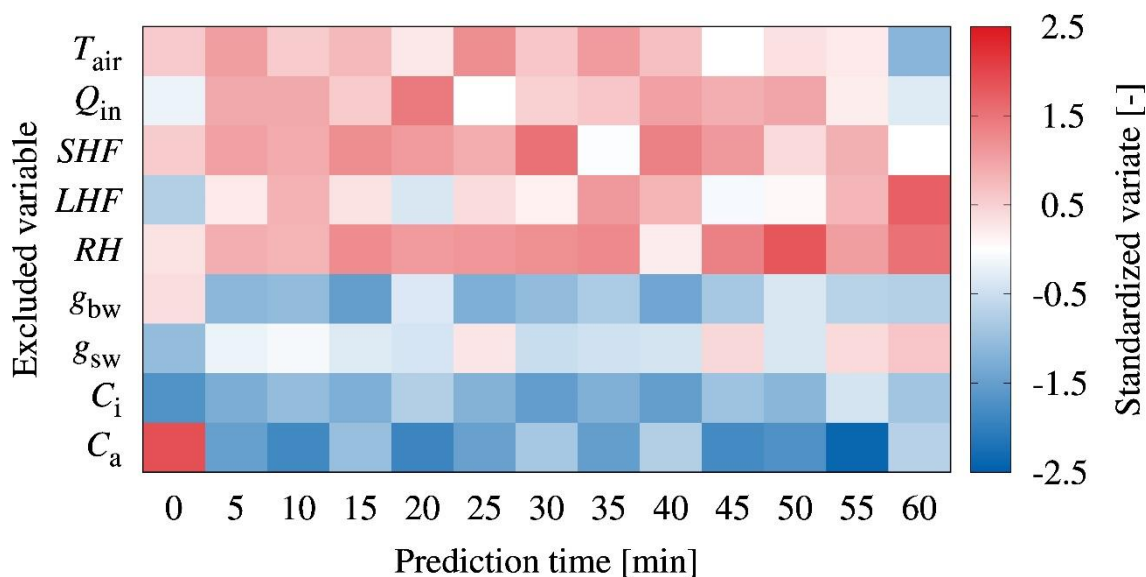


図 21 LSTM モデルにおいて予測時間ごとの説明変数が予測精度に与える影響

C_a , C_i , g_{sw} , g_{bw} は少なくとも 60 分先までの予測精度にとって重要な変数であることがわかる。また、現在値の推定においては、 C_i が最も A の値を決める上で重要な変数であることがわかる。予備計算から、 C_i のみを使った LSTM モデルによる再現精度は Q_{in} , C_a をそれぞれ用いた場合と比較して低いことから、 C_i の変動と他の変数が影響して高再現精度を実現することがわかる。このことは、光合成速度が葉内外の CO_2 濃度差に比例することからも説明が可能である。一方、予備計算において最も重要な説明変数であった Q_{in} は、全説明変数を対象にした LSTM モデルではその重要度は低い結果となった。これは Q_{in} の影響が他の説明変数に反映された結果、 Q_{in} の影響が相対的に小さくなったことが原因であると考えられる。また、 C_a と C_i の時系列値が A の 60 分先までの予測精度に強い影響を及ぼしつづけていることがわかる。

同様に、PCR について各説明変数を削除した場合の再現・予測精度への影響を図 11 に示す。なお、図 11 の上図が説明変数に瞬時値を使った場合、下図が時系列値を使った結果である。PCR モデルの推定・予測精度への影響を見ると、説明変数に瞬時値および時系列値を使った場合のいずれもヒートマップの特徴が似ており、 C_a が A の予測に重要な因子として扱われている。また、予測先の時間が長くなると g_{bw} の重要度が増すことがわかる。一方、図 11(b)を見る

と、25分先以降の予測では Q_{in} の重要度が相対的に低下し、 T_{air} が C_a の次に A の予測精度に欠かせない説明変数となることがわかる。したがって、時系列データにより構築されたPCRモデルによる25分先までの予測には、 C_a, g_{bw}, T_{air} が相対的に重要であることがわかる。また、25分以降の予測では Q_{in} の重要度が大きく低下しており、予測期間が長い場合は、光の影響よりは気温の影響が卓越することが読み取れ、その閾値はおおよそ25分程度であることがわかる。

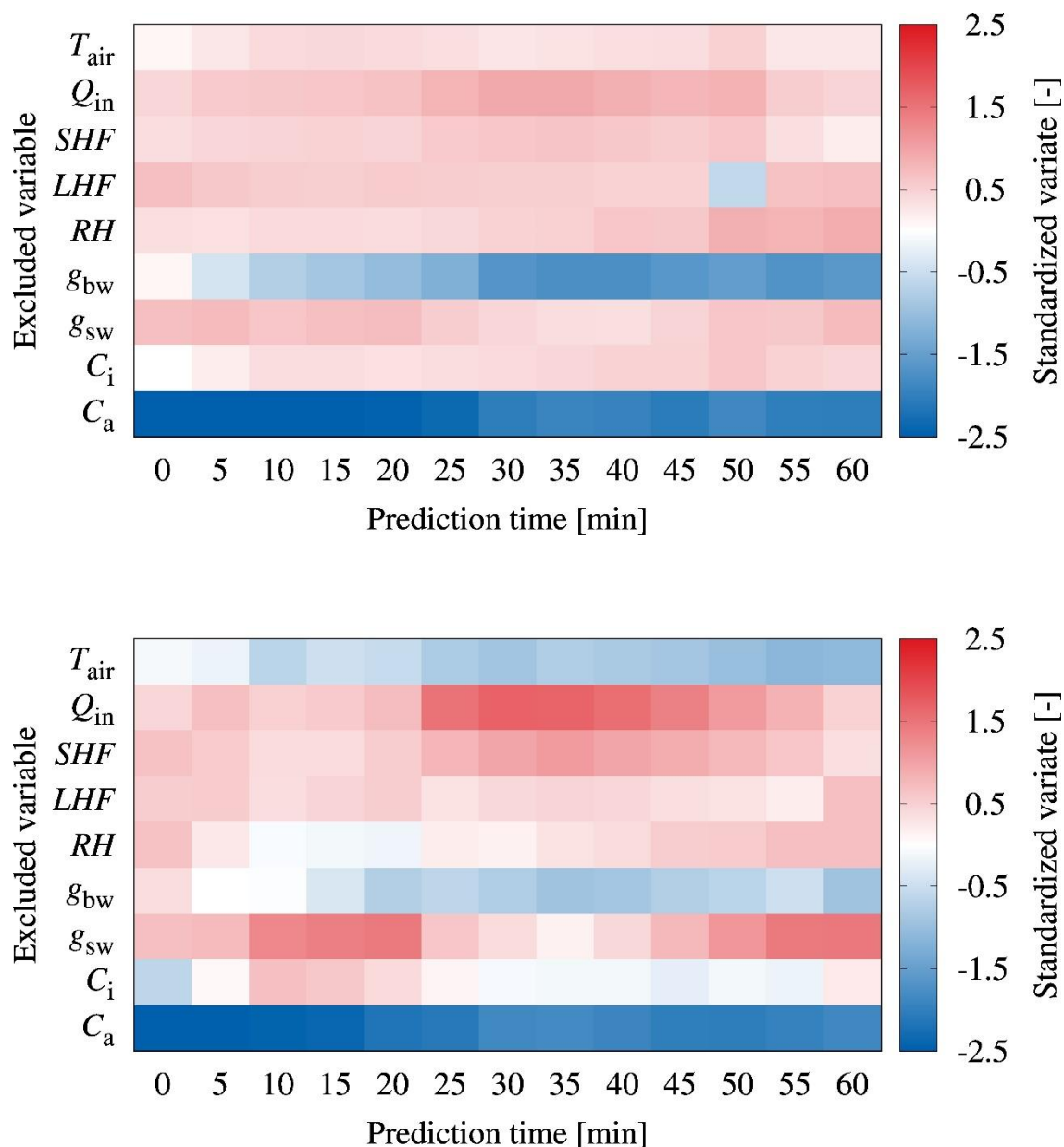


図 22 PCRモデルによる予測時間ごとの説明変数が予測精度に与える影響 ((a)瞬時値, (b)時系列値をそれぞれ入力値に使用)

5. BLSTM-Augmented LSTM の提案

5.1 本章の目的

先の章では、圃場で栽培したイネ (*Oryza sativa* L.) を対象にガス交換計測を実施し、予測モデルの構築と分析の結果について述べた。そこで本章では、得られた分析結果から葉内部の説明変数を A の速度決定因子と仮定して、それを双方向 LSTM (BLSTM)[40]–[42]で再現し予測モデルに組み込む構造をもつ BLSTM Augmented LSTM を提案する。本章では、葉内部の変数がどの程度再現できるのか、再現は予測に貢献するのか、そしてモデルが適用できる範囲について明確にすることを目的とした、

5.2 BALSTM の詳細

5.2.1 提案する予測モデルの構造

モデルで使用する説明変数は、データ取得コストが低く、容易に測定可能であることが望ましい。一方、4.4.1 節で示したように、表 3 に示した外部環境因子 Q_{in} , C_a , T_{air} のみを用いた場合の A の予測誤差は、すべての説明変数を用いて学習したモデルの結果よりも相対的に大きくなることがわかった。さらに、各説明変数のスクリーニングテストでは、 A を決定する際に C_i の変動が重要であることが示唆されている (4.4.3 節)。そこで、メソフィル説明変数の再現には計算の容易な外部環境変数をオグメントアーキテクチャで用い、外部環境因子のみを用いた A の再現にはその値を LSTM に明示的に組み込むことで予測精度を向上させることを目的とした。

図 23 は BLSTM モデル[42]の概要を描いたものである。通常 LSTM などの RNN は時刻 $t-1$ の隠れ状態を時刻 t の入力に対する追加入力として扱うが、BLSTM は中間層として LSTM を追加することで時刻 $t+1$ の状態を時刻 t の入力に対する追加入力として扱う逆方向の動作をさせる。この構造により未来への順伝播と過去への逆伝播の両方で予測ことが可能となる。BLSTM はある時点の未来値を用いて予測を行うため、その応用分野は機械翻訳[43]、音声[44]、手書き文字認識[45]、[46]などに限定されている。本研究では BLSTM モジュールの生成変数は、LSTM モジュールの予測時刻 A から相対的に全て過去の時系列であるため、モデルのリアルタイム性が損なわれないことためこの手法の利点を活用できる。

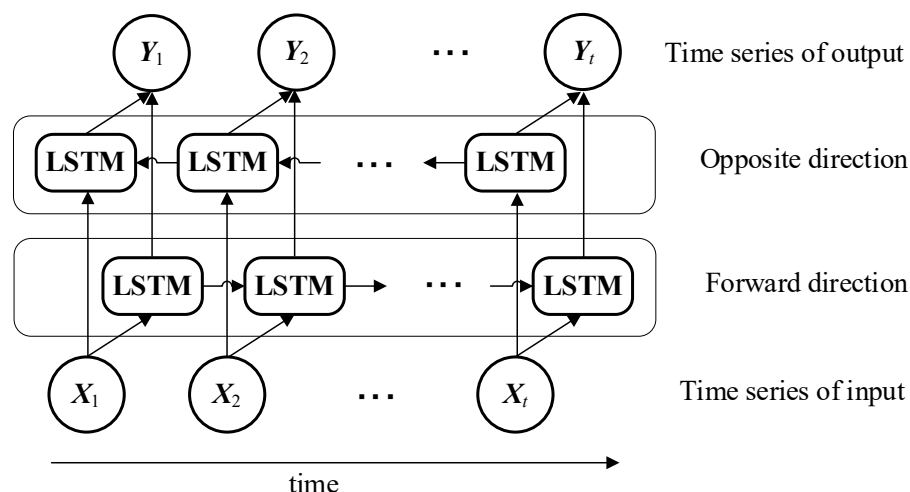


図 23 BLSTM の構造

図 24 は BALSTM モデルの概要を示したもので、メソフィル変数の逐次データを再現するモジュールと、再現されたメソフィル変数と外部環境要因の逐次データから光合成炭素同化率を予測するモジュールの 2 つから構成されている。前者は、生成モデルとして BLSTM を用い、外部環境変数の逐次データからメソフィル変数の逐次データを再生成するものである。表 6 に示すように、中間生成を行うニューラルネットワークモデルをいくつか比較した結果、BLSTM が最有力候補であることがわかった。なお、入力時系列変数と生成時系列変数は同時刻上のデータとなっている。すなわち、LSTM モデルで説明変数として入力される逐次メソフィルデータは、BLSTM モデルによる時系列出力となる。

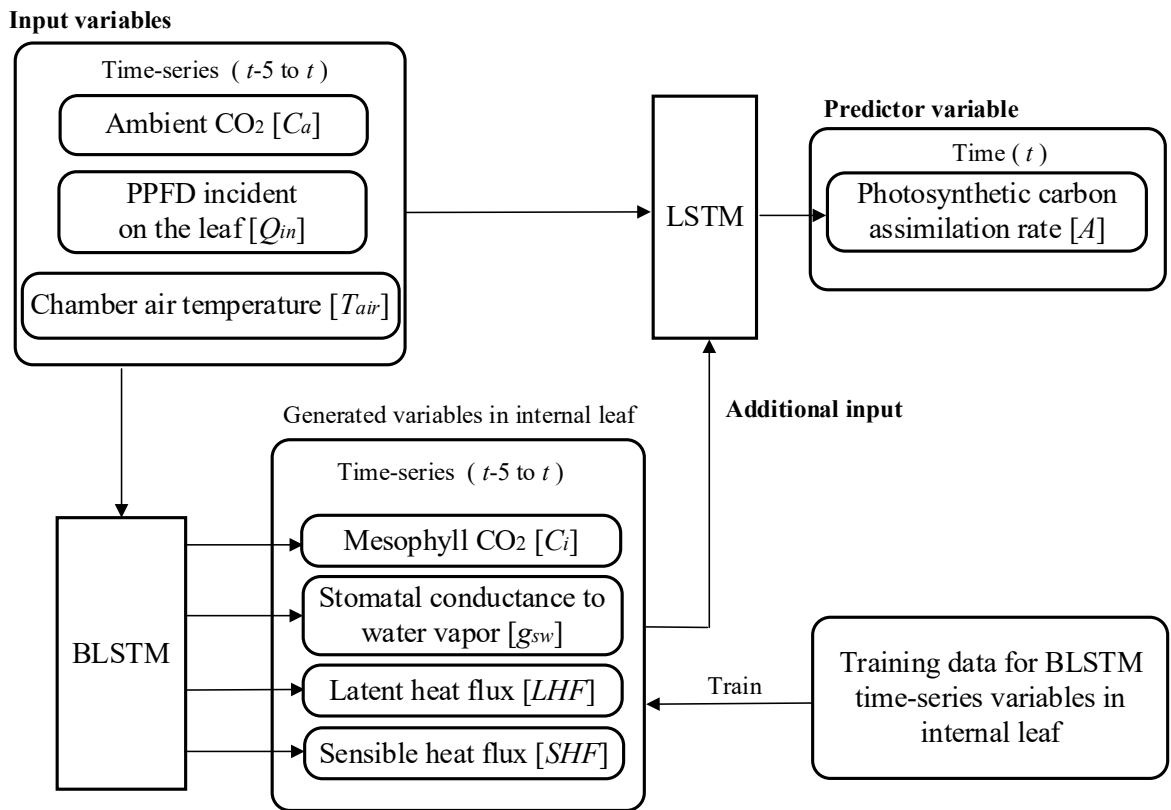


図 24 BALSTM の基本構造

表 6 葉内部の変数を生成するモデルの比較

Augmentation Model	Intermediate variable (C_i)		
	TMAE	MAPE(0)	MAPE(4)
Full Connected Net	3.32	7.0	8.9
LSTM	2.59	7.2	4.4
GRU	2.82	7.2	4.7
BLSTM	1.21	2.8	4.2

BALSTM モデルは 2 段階で学習される。まず、外部環境要因の時系列変数を説明変数とし、葉内部の時系列変数を目的変数として、中間変数の生成をになう BLSTM モデルを学習させる。次のステップでは、学習した BLSTM モデルを用いて予測した葉内部の時系列変数と外部環境要因の時系列変数を説明変数として結合し、目的変数である光合成速度 A を LSTM によって学習させる。

5.2.2 BLSTM 中間出力の評価方法

再現された逐次データは、BALSTM 中間ステージで再現されたメソフィル変数に関連する逐次データ予測への影響を調べるために、以下の 2 つの方法で評価された。まず、再現された時系列の総合的な性能について次式で表される時系列平均絶対誤差 (TMAE) と、再現された時系列の各要素の評価に用いられる平均絶対割合誤差 (MAPE) を式 (3) ~ (5) のように算出した。

$$\text{MAE} = \frac{\sum_{i=1}^n |O_i - P_i|}{n} \quad (3)$$

$$\text{TMAE} = \frac{\sum_{i=1}^m \text{MAE}}{m} \quad (4)$$

$$\text{MAPE}(t) = \frac{100}{m} \sum_{i=1}^m \left| \frac{O_{t,i} - P_{t,i}}{O_{t,i}} \right| \quad (5)$$

ここで、 O はガス交換測定によって得られたメソフィル変数の測定値、 P はモデルによって計算された値、 n は逐次データの入力幅、 m は評価データにおける全サンプルサイズ、 t は逐次データの入力幅の t 番目の要素を示す。

5.3 実験結果

5.3.1 BALSTM の予測精度と再現変数の関係

外部環境変数を除くと、LSTM による A の現在値予測に対して、 C_i , g_{sw} , LHF , RH の順に予測モデルへの影響が大きい (4.4.3 項)。まず、すべての説明変数を用いて LSTM の予測精度の値を算出した。次に、外部環境変数 Q_{in} , C_a , T_{air} のみで LSTM (Q_{in} , C_a , T_{air}) の値を予測した。同じ外部環境変数によって算出された BALSTM (Q_{in} , C_a , T_{air}) は、BLSTM を介してメソフィル変数 C_i , g_{sw} , LHF , RH を再現し利用するものである。なお、再現された変数は添え字で表現している。それぞれの手法の精度を比較した結果、全変数を用いた LSTM を除くと、 C_i と g_{sw} を生成して用いた BALSTM $_{C_i,g_{sw}}$ (Q_{in} , C_a , T_{air}) が最も予測精度が高いことが分かった (表 7)。一方、 LHF と RH を加えて 3 変数以上を生成して得た BALSTM $_{C_i,g_{sw},LHF}$, BALSTM $_{C_i,g_{sw},LHF,RH}$ の再生精度値は BALSTM $_{C_i,g_{sw}}$ と比較して向上していない。これは、予測対象の増加に伴い学習パラメータが増加し、モデル全体の学習が困難となったためと考えられる。さらに、重要度が低く精度の低いデータが LSTM に入力されることで、 A をノイズとして予測していた。しかし、より広範囲の学習データを用いてメソフィル変数を高精度に再現できれば、 A の予測精度を向上させることができる可能性があり、今後モデルの改良が期待される。さらに、メソフィル変数の TMAE 値は以下の通りであった。 $C_i = 1.21-1.32 \mu mol mol^{-1}$, $g_{sw} = 0.055-0.077 mol m^{-2} s^{-1}$ であった。MAPE(0)については、次のような値であった。 $C_i = 2.8\%-3.2\%$, $g_{sw} = 2.2\%-3.3\%$ となった。MAPE(4)では、次のような値であった。 $C_i = 4.2\%-7.1\%$, $g_{sw} = 6.3\%-7.6\%$ (表 7)であった。入力時系列の初期値 (MAPE(0)) が、それ以降の予測点 (MAPE(4)) よりも高い精度で再現できたのは、BLSTM の特性上、再現する時系列の初期値を相対的に未来の値で生成しているためと考えられる (図 23)。この特性の優位性は、表 7 に示した補強モデルの比較からも明らかである。

表 7 生成する中間変数の違いによる炭素同化率 (A)の予測比較

Model	Intermediate variable	Intermediate variable			A	
		TMAE	MAPE(0)	MAPE(4)	R ²	MAE
LSTM					0.977	0.516
LSTM (Q_{in}, C_a, T_{air})					0.891	2.24
BALSTM _{Ci} (Q_{in}, C_a, T_{air})	C_i	1.21	2.8	4.2	0.903	2.08
BALSTM _{gsw} (Q_{in}, C_a, T_{air})	g_{sw}	0.060	2.2	6.3	0.907	2.03
BALSTM _{Ci,gsw} (Q_{in}, C_a, T_{air})	C_i	1.33	3.0	5.1	0.909	2.01
	g_{sw}	0.055	3.2	7.6		
BALSTM _{Ci,gsw} (Q_{in}, C_a, T_{air})	C_i	1.32	3.1	5.1	0.902	2.06
	g_{sw}	0.061	2.7	7.3		
	LHF	3.34	11.7	13.2		
BALSTM _{Ci,gsw,LHF,RH} (Q_{in}, C_a, T_{air})	C_i	1.25	3.2	7.1	0.899	2.36
	g_{sw}	0.071	3.3	6.7		
	LHF	3.26	10.7	15.6		
BALSTM _{gsw,LHF,RH} (Q_{in}, C_a, T_{air})	RH	2.77	5.9	11.0	0.892	2.26
	g_{sw}	0.077	3.1	7.6		
	LHF	3.01	9.5	10.9		
	RH	2.98	7.1	10.2		

図 25 は、60 分までの A の予測精度を示したものである。この図は、学習済みモデルに 5 分間の説明変数の値を入力することで、その時刻からリードタイム 60 分までを 5 分間隔で予測し、A の予測精度がどのように変化するかを示している (図 16)。さらに、それを 1min のタイムステップで繰り返し、実測値と予測値を比較する。MAE 指標に基づく 60 分前予測では、LSTM (Q_{in}, C_a, T_{air})、BALSTM_{Ci,gsw} と比較して、LSTM が最も予測精度が高いことがわかった。一方、BALSTM_{Ci,gsw} は、MAE 指数では LSTM(Q_{in}, C_a, T_{air})よりも予測誤差が小さかった。また、R2 指標では、BALSTM_{Ci,gsw} による 20 分前の予測は LSTM による予測よりも精度が低かったが、20 分後の予測値は LSTM の予測値よりも良好であることがわかった。

これらの結果から、BALSTM_{Ci,gsw} は、特に 20 分前の A 予測において、メソフィル変数の順次性を考慮することにより、予測精度の低下を防ぐ重要な役割を担っていることが示唆された。また、これらの結果から、BLSTM によって再現されたメソフィル変数の時系列値は、20 分先までの予測時間においても、A の予測精度の向上に寄与することが示された。これらの結果は、メ

ソフィル環境の事前情報が A の予測に大きな影響を与えることを示唆している。

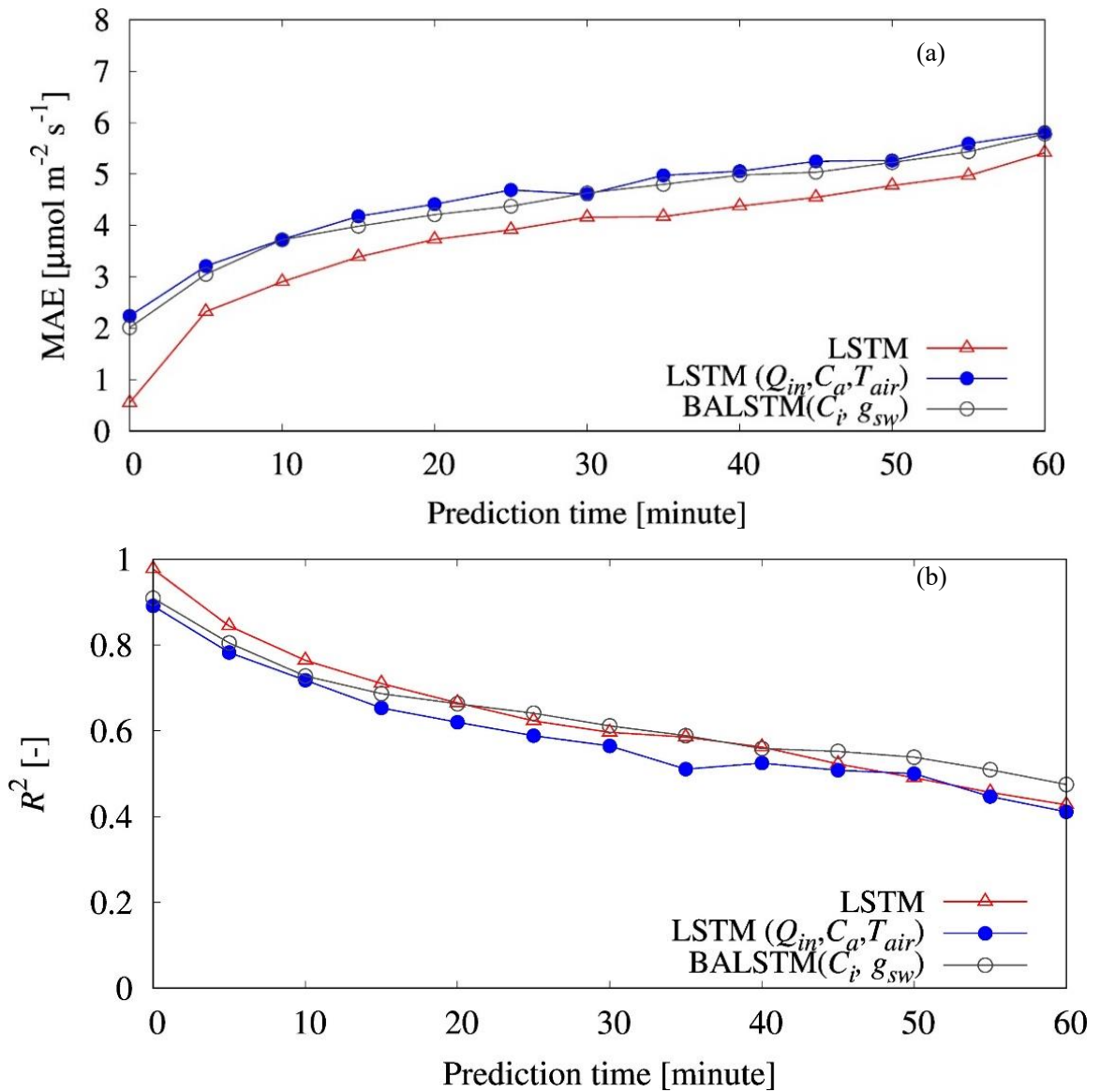


図 25 各モデルによる予測時間と A の予測精度の関係

(a) MAE and (b) R^2

5.3.2 データ入力幅と光合成予測精度の関係

本研究の予測モデルで使用した入力データの時系列長は 5 分と設定している。これは、LSTM で A の現在値を予測する際に、時系列データのウィンドウサイズを 5 分以上に設定しなくても、高精度な予測が可能であったためである (表 8)。さらに、外部環境の事前履歴が予測精度に影響するかどうかを、LSTM (Q_{in}, C_a, T_{air}) と BALSTM $_{Ci,gsw}$ について検討した。その結果、LSTM (Q_{in}, C_a, T_{air})、BALSTM $_{Ci,gsw}$ とともに時系列長が 5 分のときに最も精度が高く、入力幅が大きくなると精度が低下することがわかった。一方、時系列の窓サイズを 3 分にすると、精度の低下はわずかであった。以上の結果から、 A の予測には最新値に重要な情報が集中しており、LSTM、LSTM (Q_{in}, C_a, T_{air})、BALSTM $_{Ci,gsw}$ (Q_{in}, C_a, T_{air}) を用いて外部変数の時系列情報から A の動的特性を捉えるのに適した窓サイズを示すことができた。一方、ウィンドウサイズを大きくすると、モデルの複雑さが増し、予測精度が低下することがわかった。

表 8 入力幅と予測精度の関係

Window size	LSTM		LSTM(Q_{in}, C_a, T_{air})		BALSTM $_{Ci,gsw}$ (Q_{in}, C_a, T_{air})	
	MAE	R^2	MAE	R^2	MAE	R^2
3	0.72	0.963	2.25	0.877	2.10	0.900
5	0.52	0.977	2.24	0.890	2.01	0.909
10	0.78	0.964	2.25	0.881	2.20	0.898
15	0.81	0.961	2.28	0.867	2.23	0.889
20	0.98	0.957	2.40	0.852	2.29	0.870

5.3.3 予測モデルの可用性

図 26 と表 9 に、2019 年度の葉の個体ごとのガス交換測定値のデータセット (表 2) に対して、2017 年度のデータセットで学習したモデルを適用して A を予測した結果を示す。この結果によれば、LSTM (Q_{in}, C_a, T_{air}) に対して、BALSTM $_{Ci,gsw}$ が C_i と g_{sw} を中間積変数として組み込んだことによって誤差が大幅に減少した場合 (図 26 (a))、両モデルで誤差はほぼ変わらず、予測精度は概ね満足できる場合 (図 26 (b))、両モデルで絶対誤差は大きく、予測精度も低い場合 (図 26 (c)) に分類されることが判明した。

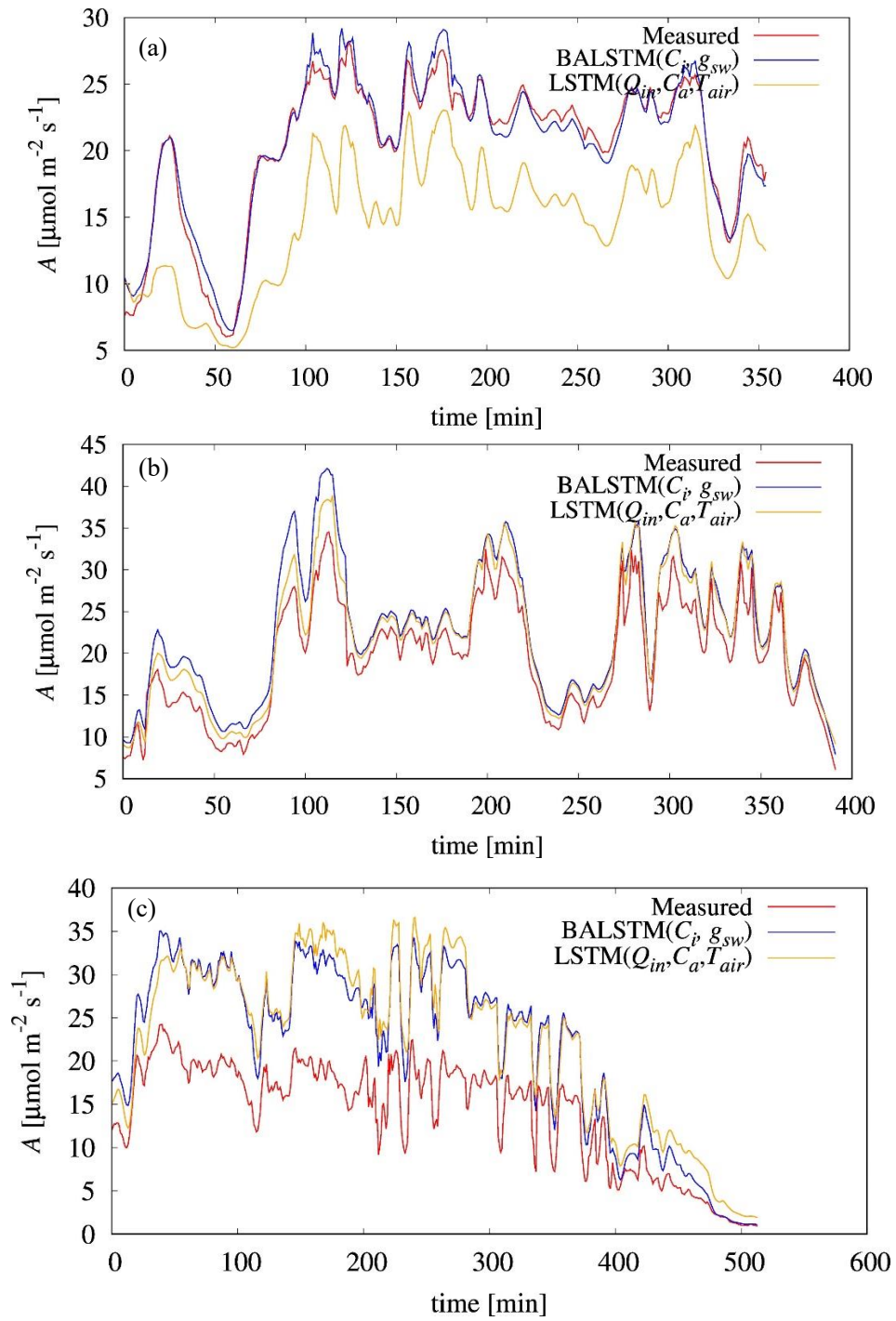


図 26 BALSTM による実測値と予測値

(a) July 5, 2019, (b) July 18, 2019, and (c) July 17, 2019

表 9 イネ (2019 年) の各葉のテストデータを用いたモデルの予測評価

Month	Day	Cultivar	MAE		Category
			LSTM(Q_{in}, C_a, T_{air})	BALSTM $_{Ci,gs}$ (Q_{in}, C_a, T_{air})	
July	5	Koshihikari	5.80	0.73	A
	6	Akitakomachi	4.31	1.98	A
	17	Koshihikari	8.35	7.73	C
	18	Akitakomachi	2.45	3.39	B
	21	Akitakomachi	1.72	1.94	B
August	2	Akitakomachi	4.52	4.54	B
	3	Koshihikari	6.09	6.01	C
	17	Akitakomachi	6.64	5.99	C
	27	Koshihikari	6.80	6.07	C
	29	Koshihikari	2.99	2.56	B
September	1	Akitakomachi	7.96	7.95	C
	16	Takanari	4.07	6.53	C
	17	Takanari	6.64	6.91	C

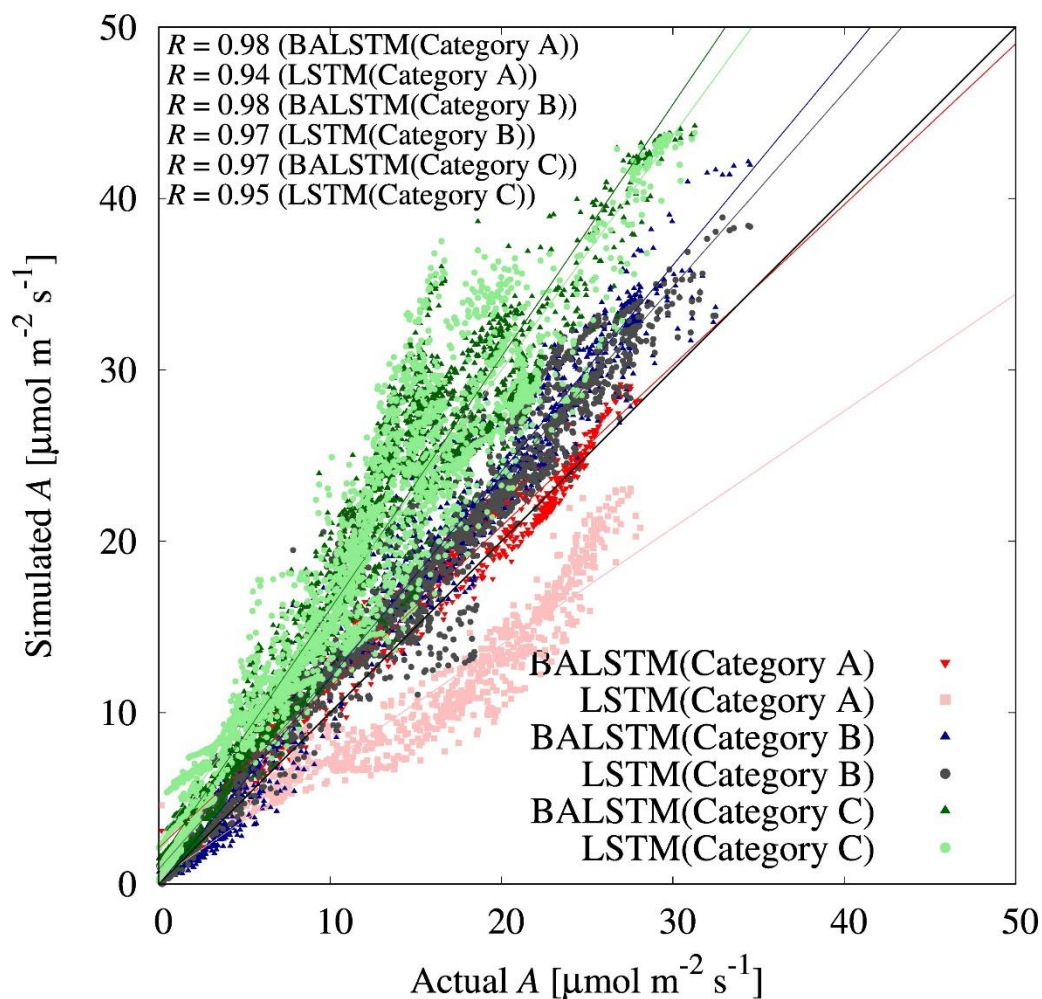


図 27 表 6 の各カテゴリーにおける同化率(A)の残差プロット

BALSTM_{C_i,g_{sw}}を用いることで、 C_i と g_{sw} の再現性と、 A の予測精度が大きく向上したサンプル(表 9 のカテゴリーA)を確認することができる。これらの結果は、外部環境変数だけでは十分な精度が得られない試料において、メソフィル変数をBALSTMで再現し、それを予測の説明変数とすることで精度が向上したことを示唆している。さらに、表 9 のサンプルのうち半数は、カテゴリーAとBを比較すると、LSTM(Q_{in} , C_a , T_{air}), BALSTM_{C_i,g_{sw}}(Q_{in} , C_a , T_{air})ともに予測精度が比較的低く、表 9 および図 27 で示したようにカテゴリーCと分類されたサンプルの外部環境変数のデータ分布は、夏季の光合成光量子束と気温が学習データより高いことが分かる(図 28)。また、モデルの学習に使用しなかった「コシヒカリ」や「あきたこまち」とは系統が異なる「タカナリ」でも予測精度が低かった。したがって、時間、品種、生育ステージなど学習データの数値範囲外の入力データを用いた予測や、学習に用いていない品種のモデルへの適用には、未解決の課題が残されている。一方、メソフィルの C_i と g_{sw} は、BALSTM_{C_i,g_{sw}}を用いれば、 A の予測に加えて再現することができる。したがって、今後、機械学習モデルの学習データとして、幅広いガス交換測定データを利用することによって、 C_i , g_{sw} ともに再現が可能になり、様々な外部環境シ

ナリオに基づいた A 予測が可能になると考えられる。

本研究で用いた密閉型チャンバー内の気象環境は、光環境を除き、野外圃場環境と異なっている。また、学習に用いた対象品種は「コシヒカリ」であり、葉位置の測定には、最も発達した葉のみを用いた。そのため、提案モデルをより汎用的なものにするためには、品種、土壌条件、施肥条件、生育ステージ、葉の位置など、さまざまな条件下での測定値が必要となる。さらに、本技術を圃場における群落規模の光合成予測に応用するためには、気象環境が個々の葉の光合成炭素同化率 A に与える影響について、追加の検討を行う必要がある。また、圃場周辺の鉛直方向の光量といった、ミクロな気象の精密な推定モデルとの接続など通じて、実験および理論研究による知見の蓄積が必要である。これらの問題を解決することで、樹冠の成長や乾物生産の正確な予測が可能になるとともに、予測モデルを用いて成長に影響する作物の表現型データを大量に取得可能になると考えられる。こういったデータを利用することによって、効率やスピードを追求した育種を行う際にどの表現型をターゲットにすべきか、という判断のための情報を提供することができると思われる。

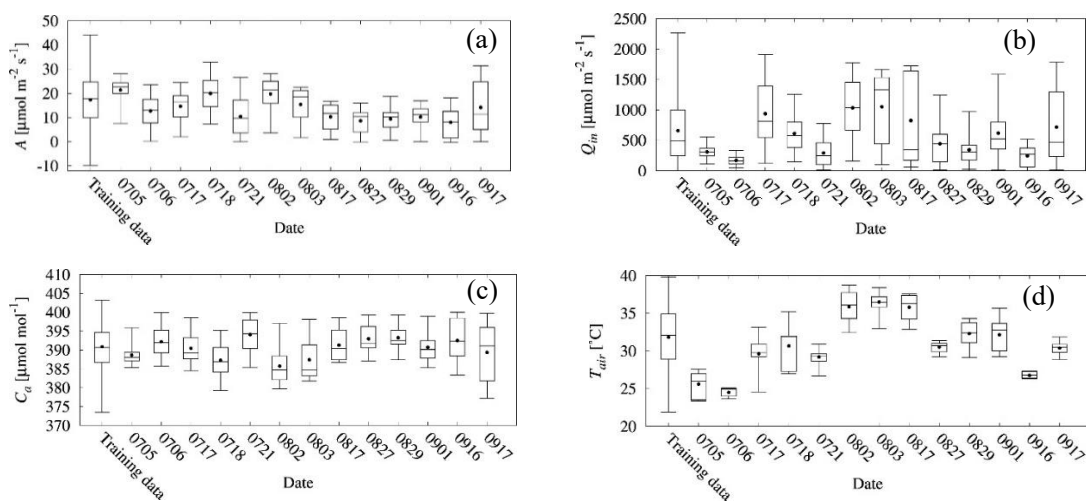


図 28 2019 年に測定した学習データと各葉の変数値の箱ひげ図：(a) 同化率 (A), (b) 葉に入射する PPFD (Q_{in}), (c) 周囲の CO_2 (C_a), (d) チャンバー空気温度 (T_{air})

5.4 本章のまとめ

本章では、イネ葉の光合成炭素同化率 A を個々に予測することに焦点を当てた。BALSTM を用いて A の予測モデルを構築し、その精度を検証した。BALSTM の最大の特徴は、測定しやすい環境変数である Q_{in} , C_a , T_{air} のみを用いて、 A に加え、 C_i と g_{sw} を中間生成物として再現するハイブリッドなアーキテクチャであることである。そして、得られた出力は学習モデルに加えらる。その結果、以下のことがわかった。

- (1) $BALSTM_{C_i, g_{sw}}(Q_{in}, C_a, T_{air})$ は $LSTM(Q_{in}, C_a, T_{air})$ よりも有意に高い予測精度を示す。
- (2) 中間生成物として C_i と g_{sw} を用いたハイブリッドモデルは他の中間生成変数の組み合わせよりも高い A の再現精度を示した。
- (3) C_i と g_{sw} の再現における MAPE 値は 5.1%であった。
- (4) メソフィル変数を再現する機構を持つことで、外部環境変数のみを用いた場合と比較して、 A の再現性・予測精度が向上することが確認された。

今後は、より広範囲の外部環境変数を含むガス交換測定を行い、モデル学習に必要なデータ量を増やすことでより頑強なモデルを構築し、モデルの予測精度をさらに向上させることが可能になると考えられる。また、比較的計測が容易な外部環境を説明変数とする BALSTM は、リモートセンシングデータなどの空間的広域情報を組み合わせることで、リアルタイムで A を予測するシステムを構築できると考えられ、ひいては現場での農作業の意思決定に適用できる情報基盤の構築を目指すことができる。

6. ドローン空撮画像を用いた葉面積推定

6.1 本章の目的

野外圃場において農作物の生育情報や環境情報をモニタリングすることは営農者の意思決定において重要な役割を果たす。そして近年、小型無人飛行機械（UAV）はリモートセンシングを始め施肥や農薬散布といった実用化が進んでいる[47]。先の章で述べた外部環境要因から個葉の光合成速度を予測するようなモデルを株単位や圃場規模に拡大し適用するためにも、葉面積（LAI）といった植生指数を推定する技術が必要となる。しかし多くの空撮画像を用いて植生指数を求める研究では、圃場単位あるいは区画単位で平均化された指標を用いるのが主流であり、個株に注目するような方法は定点装置や車載センサーを用いる場合が多い。UAVによる近接センシングでは、近年3次元点群情報として仮想的に再現することによる精緻な計測が可能となったが、機器コスト・計算コストが高いという欠点が存在している。

一方で、画像処理分野においては計算機性能の進歩や深層学習技術の発展と大規模なデータセットの整備[48]–[51]が進んだことから、物体の位置とカテゴリーを識別する物体検出（Object Detection）、および画素毎に物体を識別し領域を抽出する領域抽出（Semantic Segmentation）といったアルゴリズムが高い精度で実装され、画像を扱う多くの分野で活用が進んでいる。以上の背景から、本章ではUAVを用いて取得される高解像度の空撮画像から農作物を検出し、その作物領域を抽出するモデルの実装を試みた。これは空撮画像から作物領域を高い精度で抽出し、農作物の生育状態を定量的に表現するにあたっての問題点を明確化することを目的としている。

6.2 関連手法

6.2.1 一般物体検出モデル（Object Detection）

物体検出は1枚の画像から複数物体のクラスと位置情報を抽出するタスクであり、かつては画像から領域情報を抽出しサポートベクトルマシンなどを用いてクラス分類とバウンディングボックスへの回帰が用いられてきた。近年の深層学習手法の発展や計算機の高速化と畳み込み層（CNN）の利用やRegionOfInterest（RoI）構造によって、Fast-RCNN[52]、Faster-RCNN[53]といった手法は物体領域の抽出を大幅に高速化したことでよく知られている。これら著名な二つの手法は、物体領域を特定する段階と、特定された領域ごとに物体のクラスを特定する2段階構造となっている。これに対して、YOLO[54]では画像をグリッドごとに分割し、またバウンディングボックスの検出、クラス認識、物体である確率計算を同時に実行する1段階構造の物体検出が実現された。この手法は2016年に提案されてから2020年に至るまでバージョンアップ[55]–[57]を繰り返している。CNNベースの手法は高速かつ高精度で認識できるため、リアルタイム性が重視されるケースで多

く利用されている。さらに近年では、自然言語処理で利用される Transformer 構造を画像認識タスクに取り込んだモデルとして Vision Transformer (ViT) が様々な課題で成果を出している。このモデルにおける Self-Attention 構造から計算量を削減した Swin Transformer[58]が提案され、その汎用性の高さから注目を浴びている。

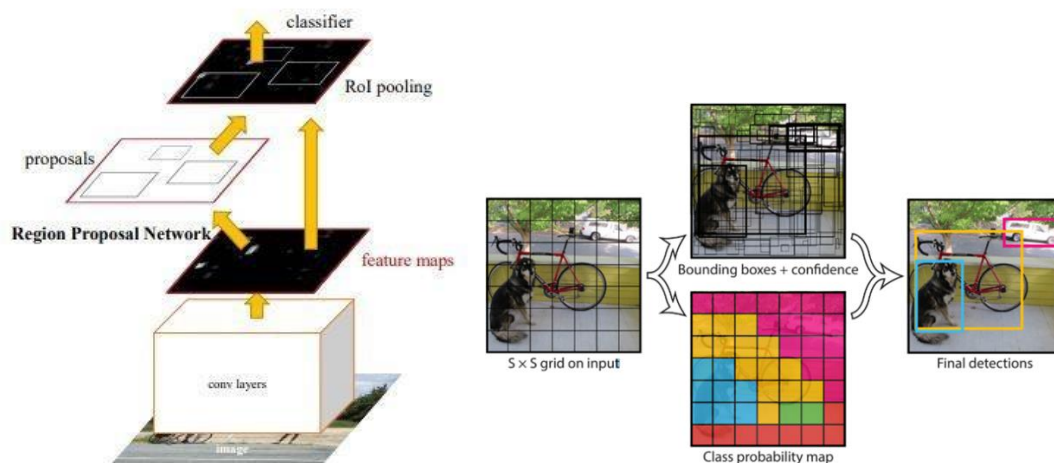


図 29 モデル概要 (左 : Faster RCNN[53], 右 : YOLO[54])

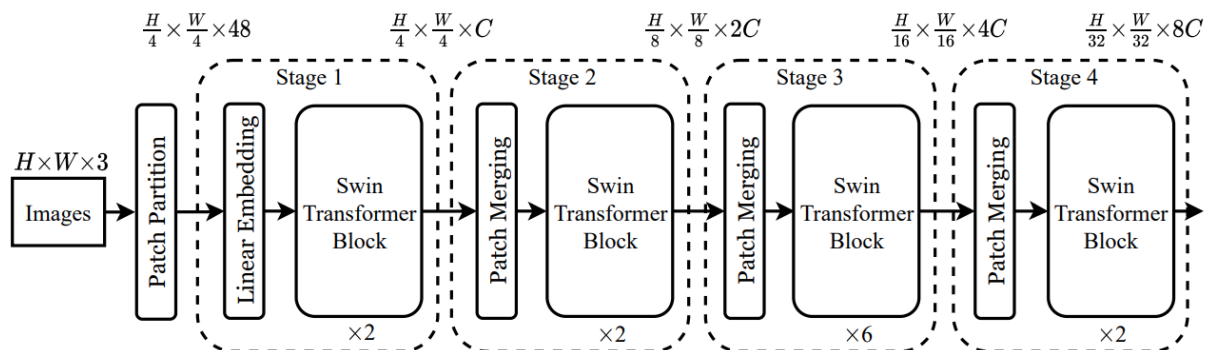


図 30 モデル概要 SwinTransformer[58]

6.2.2 領域抽出モデル (Semantic Segmentation)

セマンティックセグメンテーションは1枚の画像内のピクセルごとに物体クラスを判別するタスクであり、深層学習手法の進歩とともにCNNを用いた手法が提案されてきた。中でも代表的なのが Fully Convolutional Network (FCN) [59]であり、これはエンコーダデコーダ構造によって、CNNによる画像特徴の抽出とピクセル単位のクラス分類を実現している。このエンコーダデコーダ構造には、特徴抽出の段階で情報圧縮が起きるため、局所情報を損失してしまうという欠点がある。そこで、それら構造の間をバイパスする機構を持つ U-NET[60]やピラミッド型の画像分割

により局所から大域までの情報を利用する PSPNet[61]が提案されている。また物体検出モデルである Faster R-CNN を拡張した Mask R-CNN[62]では検出領域ごとに物体マスクを同時に推定することで、高速かつ高精度なセマンティックセグメンテーションを実現している。

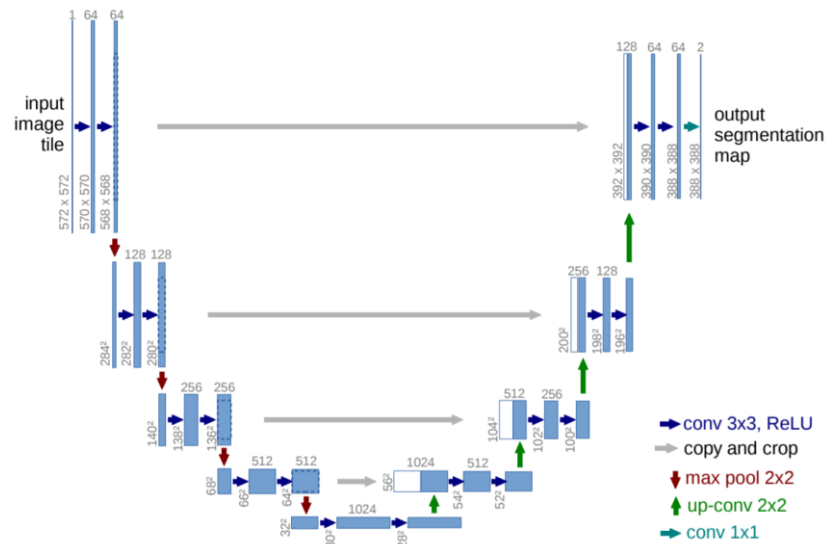


図 31 モデル概要 U-NET[60]

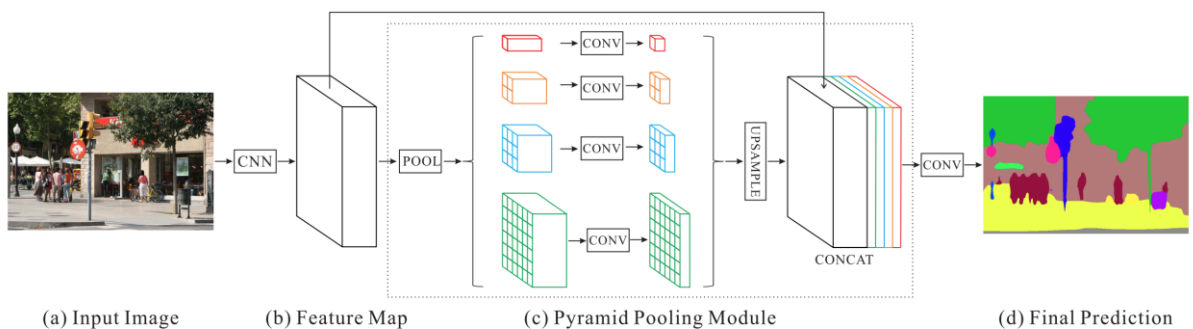


図 32 モデル概要 PSPNet[61]

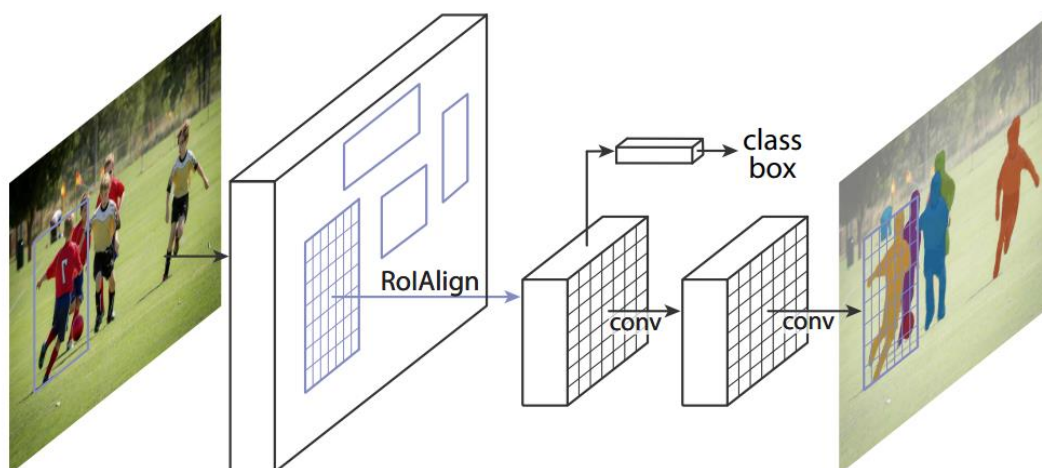


図 33 モデル概要 Mask R-CNN[62]

6.3 葉面積推定システムの構成

提案する葉面積推定システムの概要を図 34 に示す。モデルは 4 つのモジュールから構成されている。1. 画像前処理では、UAV によって空撮された原画像が高解像度(5280x3956)これを 1024x1024 の正方系に切り分ける処理である。これは後に続く作物検出モデルの入力としては画像サイズが大きすぎるためである。続く 2. 作物検出では、一般物体認識モデルを用いて作物の位置を検出する。3. 葉領域抽出部では、検出された作物位置の中心座標をもとに 256x256 の領域をさらに切り取り、セマンティックセグメンテーションモデルを用いた葉領域の抽出を行う。最後に 4. 復元部では抽出された葉領域を原画像サイズに戻し、ラベルデータとの比較と評価を行う。各モデルの学習において、ImageNet データベース[63]を利用した事前学習を実施し、本研究のデータセットを用いて転移学習を実施した。

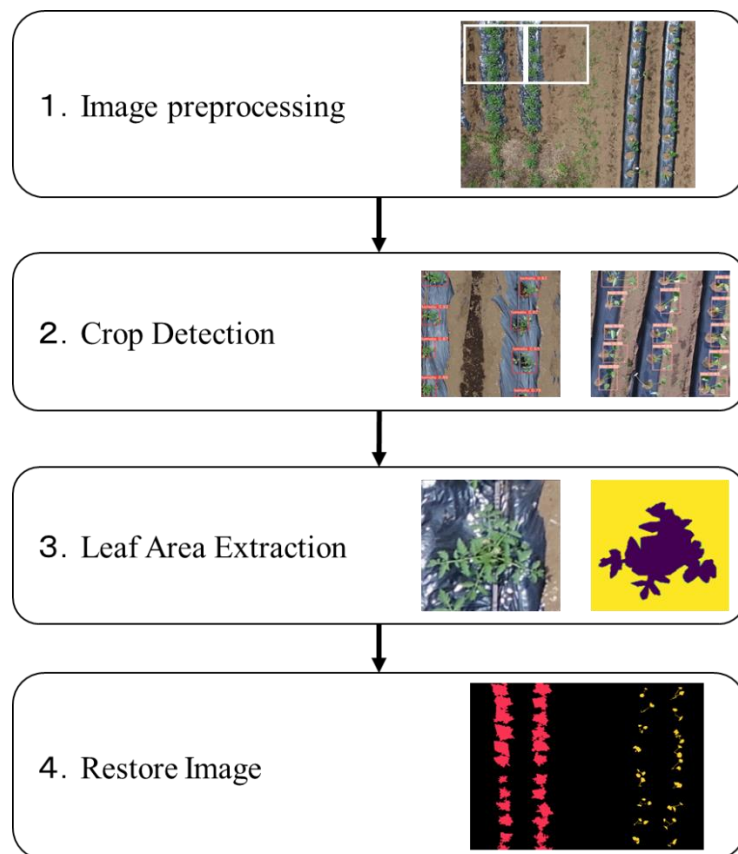


図 34 ドローン空撮による葉面積推定システムの概要

6.4 評価方法

鉛直方向から見た葉面積をモデル評価の対象とするため、入力画像に対するターゲットと予測領域の重なり度合いを示す Intersection Over Union (IoU)を評価指標とする。また個株を検出するモジュールについては、検出の一般的な指標である Average Precision (AP)を評価値とし、葉領域抽出モジュールの評価はシステム全体と同様に IoU を使用する。

6.4.1 検出評価指標：平均適合率

以下の式で表現される適合率と再現率を用いて求めるクラス毎の平均適合(AP)が物体検出器の一般的な評価指標として用いられている。

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (6)$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \quad (7)$$

$$\text{平均適合率(AP)} = \text{Precision-Recall グラフの面積} \quad (8)$$

本研究では、検出モデルによる検出結果の信頼度スコアが 0.25 を超えた場合に正の予測結果 (Positive)として出力する閾値設定とし、作物枠の予測と実際の枠ラベルの IoU が 0.45 を超える場合に予測結果が正(True Positive)となる閾値設定とした。

表 10

		ラベルデータ Ground Truth	
		作物枠内 : True	枠外 : False
認識結果	作物枠内 : Positive	True Positive	False Positive
	枠外 : Negative	False Negative	True Negative

6.4.2 葉面積評価指標：Intersection Over Union (IoU)

葉面積抽出の結果は Intersection over Union(IoU)を用いて評価する。IoU は、次式のように計算できる。これは Ground Truth と認識結果の和集合の中の積集合の割合を表しており、本研究では実際の葉領域が Target 領域、予測結果の葉領域が Prediction 領域である。

$$IoU = \frac{Target \cap Prediction}{Target \cup Prediction} \quad (9)$$

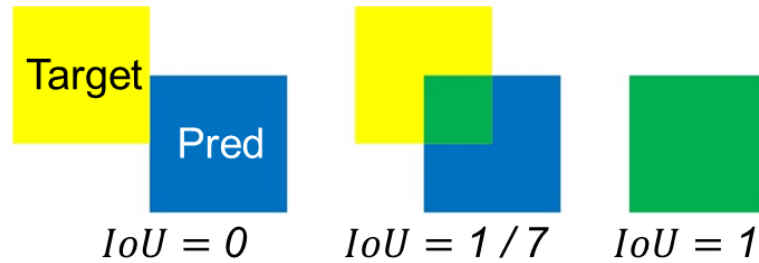


図 35 IoU の概要

6.5 実験結果と考察

6.5.1 作物検出・葉領域抽出モデルと作物時期の関係

撮影日毎の画像データそれぞれを評価用として、各モジュールの性能と作物時期の影響について精度評価実験を行った。計算速度の関係から本実験では、検出モデルとして YOLO を、葉領域検出モデルとして U-Net を利用した。

システム全体の葉面積推定の結果を図 36 に示す。また作物検出部と葉領域抽出部の結果を図 37 に示す。葉面積推定システムの結果から、生育初期～中期にかけては高い精度で、鉛直方向からの葉面積の推定が実現できている。トマトについて最初期の精度が低いのは、マルチが無い場合に、初期の芽と雑草の区別が難しいためと考えられる。また、一方で両作物において 7 月上旬から著しい精度の低下が発生している。これは 7 月上旬から作物検出部が適切に検出できないケースが増え、葉領域抽出部に作物枠がフィードされないためと考えられる。この現象は、夏の繁茂期に入ることによって作物同士の葉が触れ合うことで輪郭が曖昧になり、作物の検出難易度が上昇するためと考えられる (図 37, 図 38, 図 39)。

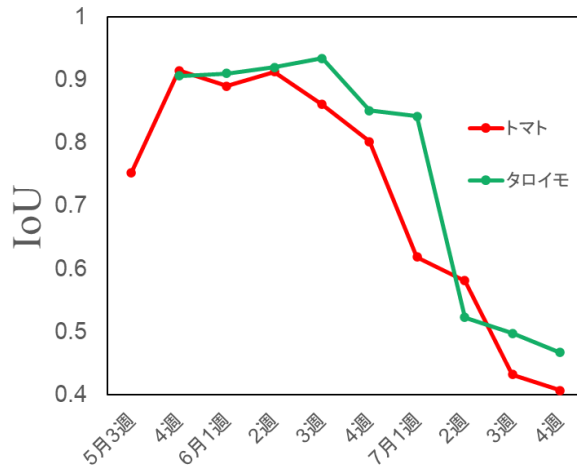


図 36 葉面積推定システムの精度

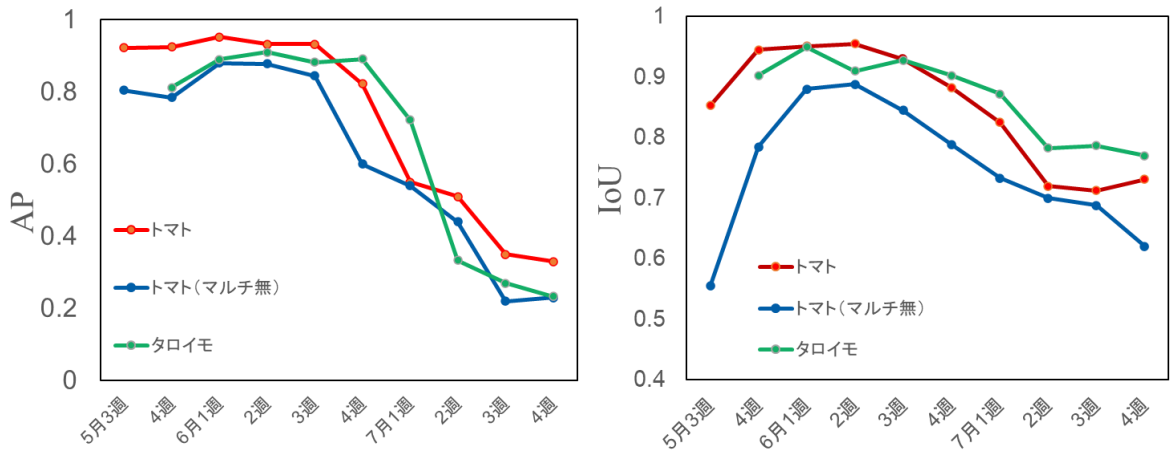


図 37 作物検出モジュールの精度 (左) と葉領域抽出モジュールの精度 (右)



図 38 6月上中旬の検出結果 (左: マルチ有, 右: マルチ無)

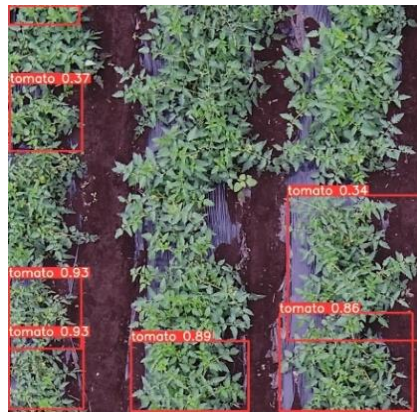


図 39 7月上旬の検出例 (マルチ有)

6.5.2 検出モデルの改善

前項の実験結果から、システム全体の精度の低下は繁茂期にかけて検出モジュールの精度低下に起因することが判明した。そこでこの問題を解決するため、近年提案された検出モデルを実装評価する実験を行った。評価のためのデータとして、テストデータの内、成育初期の5月4週の画像と7月4週の画像を利用する。結果を以下表11に示す。

表 11 作物検出モデルの精度比較

モデル	5月4週			7月4週		
	トマト○	トマト●	タロイモ	トマト○	トマト●	タロイモ
YOLOv4	0.788	0.921	0.803	0.255	0.350	0.227
Faster R-CNN	0.800	0.905	0.784	0.155	0.281	0.198
Swin Transformer	0.852	0.915	0.822	0.280	0.374	0.235

○マルチなし ●マルチあり

結果から Swin Transformer モデルは、評価用データのうち複数の項目で過去に採用していた YOLO モデルより優れた結果を示した。しかし、繁茂期での検出能は依然として低く、その改善は限定的であった。よって今後の研究で検出能を向上させるには、葉の形状や輪郭に注目したより明示的な検出方法の導入や学習データを拡張する手法を導入、あるいは繁茂期には画像以外のデータを用いた検出モデルを用いるといったアプローチが必要があると考えられる。

6.5.3 葉領域抽出モデルの改善

前項と同様に葉領域の抽出についても新規に複数のモデルを実装し比較実験を行った。結果を表 1 に示す。

表 12 葉領域抽出モデルの精度比較

モデル	5月4週			7月4週		
	トマト○	トマト●	タロイモ	トマト○	トマト●	タロイモ
U-NET	0.761	0.940	0.891	0.606	0.725	0.775
PSP Net	0.763	0.924	0.900	0.651	0.732	0.767
Mask R-CNN	0.772	0.939	0.911	0.666	0.745	0.782

○マルチなし ●マルチあり

結果から葉領域抽出モジュールにおいては、Mask R-CNN が顕著に優れた結果を示した。このように、セマンティックセグメンテーションタスクで研究・提案されているモデルの利用により、本研究における葉領域抽出の精度が改善することが明らかとなった。

6.5.4 各モデルの組み合わせと圃場全体での葉面積推定

表 13 に、技術的課題である繁茂期 7 月 4 週のテスト画像における作物全体の IoU を評価指標として、作物検出モデル・葉領域抽出モデルを組みあわせて評価実験を行った結果を示す。

表 13 モデルの組み合わせとテスト画像 IoU 関係

モデル	7月4週	
	トマト	タロイモ
YOLO + U-Net	0.401	0.472
YOLO + Mask RCNN	0.421	0.477
Swin Trans + Mask RCNN	0.489	0.525
Swin Trans + U-Net	0.472	0.507

結果から、葉領域抽出モジュールの精度向上の効果は限定的である一方で、検出モジュールを改善した場合にシステム全体での葉面積推定性能が顕著に向上することが判明した。

7. 結言

本章では、計測容易なデータを用いた光合成速度予測モデルと UAV 画像を用いた葉面積推定に関する研究の成果の要約と、本研究の将来性について述べる。

7.1 研究成果の要約

- (1) 光合成速度の時系列特徴を用いた予測には、葉内部変数のうち、特に葉内 CO₂ 濃度および気孔の水蒸気コンダクタンスがモデルにとって特に重要な変数であることが判明した。
 - ・すべての説明変数を使用できる場合に LSTM 手法は最も優れた性能を示した。
- (2) 葉内部の時系情報を用いない場合においては、既往の回帰予測手法と比較して LSTM を始めとするニューラルネット手法に明瞭な優位性は無いことがわかった。
- (3) (1)の変数群を生成する機構を持たせる BALSTM の構造は、外部環境要因・計測容易なデータだけを用いた場合にモデルの予測精度が改善することが確認された。
 - ・今後、これらの内部数値をより高い角度で推定できるようになれば、さらに A の再現性能が上がる可能性がある。
 - ・さらに適用可能な範囲を広げるには、より広範な外部環境変数を含む測定を行い、また異なる生育ステージや品種へのデータ量を増やす必要がある。
- (4) 空撮画像だけを用いた葉面積推定モデルは運用する時期によっては高い精度で実現できることが示された。
 - ・本研究でのモデルは、鉛直方向からだけ抽出した葉面積指標のため、実際の葉面積、実運用のためには更なる空間情報の加味が必要である。
- (5) 作物枠検出と葉領域抽出葉 2 段階構造にした葉面積推定モデルは、繁茂期に検出部分に課題があることが示された。
 - ・高解像度画像をダウンサンプリングせず運用するための試みであった。
 - ・作物検出段階を挟むことで、画像に基づく病理診断といった他の用途と連携させることも目指したものであった。
- (6) 近年提案された検出モデル・葉領域抽出モデルを使用する事による性能改善が確認された。
 - ・繁茂期の検出能を改善することが今後の焦点となる
 - ・とくに作物検出モデルの影響が大きいいため、本研究に適したモデル開発の必要がある

7.2 本研究の将来性

本研究により、光合成速度を取得容易なデータから再現するためのモデルの開発および提案と、空撮画像だけから個別の作物および圃場の葉面積を求めるための深層学習を用いた手法の検討を行うことができた。しかし前者は生育ステージや品種の違いによる影響についての検証が必要であり、後者においては繁茂期における課題点が明確になった。対象とする作物が異なるため、本研究におけるこれら手法を単純に掛け合わせることは出来ないが、取得容易な計測データによる光合成のモデリングとリモートセンシングを用いた時空間的な情報技術の組み合わせは、今後圃場単位でのリアルタイムかつ高精度な光合成の予測、つまり生育管理や営農者の意思決定補助に結びつくと考えられる。よって本研究を試金石として、計測容易なデータ利用を中心としたアプローチの研究と開発が期待される。

謝辞

本研究を進めるにあたり多くの方のご指導，ご協力を賜りました。

本研究の一部は JSPS 科研費(課題番号:16KK0169)及び JST さきがけ(課題番号:JPMJPR1603)の助成を受けたものです。

著者の元指導教官である中川正樹特任教授には，退官後にも関わらず研究内容に関する様々なご提案・ご指導を賜りました。ここに心から御礼を申し上げます。

山下善之教授と辰己賢一准教授には，食料エネルギーシステム科学専攻の課程として多くのご協力ならびに研究のご指導を賜りました。とくに辰己賢一准教授には本研究の具体的な内容について格別のご指導を賜りました。ここに心から御礼を申し上げます。

金子敬一教授には，後任の研究指導教官として多くのご協力を賜りました。秋澤淳教授には著者の教育指導教官として多くのご助言を賜りました。藤波香織教授，清水郁子准教授には，本論文の副査としてご助言を頂くとともに本論文の細部にわたりご指導をいただきました。深く御礼申し上げます。

研究を行う上で，元中川研究室の Nguyen Tuan Cuong 特任准教授，Nguyen Tuan Hung 特任助教，Truong Thanh Nghia 特任助教ならびに，本学の同期及び先輩後輩の皆様ほか多数の方々からご協力をいただきました。ここに深く感謝申し上げます。

最後に，博士後期課程進学へ理解を示してくれた父隆之，母恵子，そして祖父母と親類の皆様

に深く感謝いたします。

参考文献

- [1] 誠之本多, 賢一辰己, and 正樹中川, “Lstm による光環境変動下におけるイネ個葉の光合成速度の予測,” *農業情報研究*, vol. 30, no. 2, pp. 96–108, 2021, doi: 10.3173/air.30.96.
- [2] M. Honda, K. Tatsumi, and M. Nakagawa, “Prediction of Photosynthetic Carbon Assimilation Rate of Individual Rice Leaves under Changes in Light Environment Using BLSTM-Augmented LSTM,” *Comput. Model. Eng. Sci.*, vol. 133, no. 3, pp. 557–577, 2022, doi: 10.32604/cmescs.2022.020623.
- [3] M. Honda and Y. Yamashita, “Time Series Estimation of Product Quality Based on Neural Net Algorithms,” presented at the The 18th Asian Pacific Confederation of Chemical Engineering Congress, Sapporo, 2019.
- [4] M. Honda and K. Tatsumi, “機械学習を用いたイネ個葉の光合成速度の予測,” presented at the 日本農業気象学会 2020 年全国大会, 大阪, Mar. 2020.
- [5] M. Honda and K. Tatsumi, “機械学習を用いたイネ個葉の光合成速度の予測,” presented at the 日本農業気象学会 2021 年全国大会, online, Mar. 2021.
- [6] M. Honda and K. Tatsumi, “深層学習を用いた画像の作物検出およびその領域抽出に関する研究,” presented at the 第 79 回農業食料工学会年次大会大会, Sep. 2021.
- [7] D. O. Hall, “Photosynthesis — A Practical Energy Source?,” in *Research in Photobiology*, A. Castellani, Ed. Boston, MA: Springer US, 1977, pp. 347–359. doi: 10.1007/978-1-4613-4160-4_36.
- [8] G. D. Farquhar, S. von Caemmerer, and J. A. Berry, “A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species,” *Planta*, vol. 149, no. 1, pp. 78–90, Jun. 1980, doi: 10.1007/BF00386231.
- [9] S. Caemmerer, *Biochemical Models of Leaf Photosynthesis*. 2000. doi: 10.1071/9780643103405.
- [10] E. Naumburg and D. S. Ellsworth, “Short-term light and leaf photosynthetic dynamics affect estimates of daily understory photosynthesis in four tree species,” *Tree Physiol.*, vol. 22, no. 6, pp. 393–401, Apr. 2002, doi: 10.1093/treephys/22.6.393.
- [11] T. Lawson, D. M. Kramer, and C. A. Raines, “Improving yield by exploiting mechanisms underlying natural variation of photosynthesis,” *Curr. Opin. Biotechnol.*, vol. 23, no. 2, pp. 215–220, Apr. 2012, doi: 10.1016/j.copbio.2011.12.012.
- [12] E. Carmo-Silva, J. C. Scales, P. J. Madgwick, and M. A. J. Parry, “Optimizing Rubisco and its regulation for greater resource use efficiency,” *Plant Cell Environ.*, vol. 38, no. 9, pp. 1817–1832, Sep. 2015, doi: 10.1111/pce.12425.
- [13] H. Poorter *et al.*, “Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field,” *New Phytol.*, vol. 212, no. 4, pp. 838–855, 2016, doi: 10.1111/nph.14243.
- [14] J. Sun, M. Ye, S. Peng, and Y. Li, “Nitrogen can improve the rapid response of photosynthesis to

changing irradiance in rice (*Oryza sativa* L.) plants,” *Sci. Rep.*, vol. 6, p. 31305, Aug. 2016, doi: 10.1038/srep31305.

[15]S. Violet-Chabrand, J. S. A. Matthews, A. J. Simkin, C. A. Raines, and T. Lawson, “Importance of Fluctuations in Light on Plant Photosynthetic Acclimation,” *Plant Physiol.*, vol. 173, no. 4, pp. 2163–2179, Apr. 2017, doi: 10.1104/pp.16.01767.

[16]R. A. Slattery, B. J. Walker, A. P. M. Weber, and D. R. Ort, “The Impacts of Fluctuating Light on Crop Performance1[OPEN],” *Plant Physiol.*, vol. 176, no. 2, pp. 990–1003, Feb. 2018, doi: 10.1104/pp.17.01234.

[17]LÜ Wei, LI Yu-huan, MAO Wei-bing, GONG Xue, and CHEN Shi-geng, “Comparison of Estimation Methods for Net Photosynthetic Rate of Wheat’s Flag Leaves Based on Hyperspectrum,” *J. Agric. Resour. Environ.*, vol. 34, no. 6, pp. 582–586, Jan. 2017, doi: 10.13254/j.jare.2017.0173.

[18]D. Heckmann, U. Schlüter, and A. P. M. Weber, “Machine Learning Techniques for Predicting Crop Photosynthetic Capacity from Leaf Reflectance Spectra,” *Mol. Plant*, vol. 10, no. 6, pp. 878–890, Jun. 2017, doi: 10.1016/j.molp.2017.04.009.

[19]X.-Y. Zhang *et al.*, “Machine learning models for net photosynthetic rate prediction using poplar leaf phenotype data,” *PloS One*, vol. 15, no. 2, p. e0228645, 2020, doi: 10.1371/journal.pone.0228645.

[20]R. C. Andrew, R. Malekian, and D. C. Bogatinoska, “IoT solutions for precision agriculture,” in *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, May 2018, pp. 0345–0349. doi: 10.23919/MIPRO.2018.8400066.

[21]L. García, L. Parra, J. M. Jimenez, J. Lloret, and P. Lorenz, “IoT-Based Smart Irrigation Systems: An Overview on the Recent Trends on Sensors and IoT Systems for Irrigation in Precision Agriculture,” *Sensors*, vol. 20, no. 4, Art. no. 4, Jan. 2020, doi: 10.3390/s20041042.

[22]K. Jha, A. Doshi, P. Patel, and M. Shah, “A comprehensive review on automation in agriculture using artificial intelligence,” *Artif. Intell. Agric.*, vol. 2, pp. 1–12, Jun. 2019, doi: 10.1016/j.aiia.2019.05.004.

[23]A. G. Mohapatra, B. Keswani, and S. K. Lenka, “ICT specific technological changes in precision agriculture environment,” *Int J Comput Sci Mob Appl*, vol. 6, pp. 1–16, 2018.

[24]K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, “Machine Learning in Agriculture: A Review,” *Sensors*, vol. 18, no. 8, Art. no. 8, Aug. 2018, doi: 10.3390/s18082674.

[25]A. J. Isaksson, I. Harjunkoski, and G. Sand, “The impact of digitalization on the future of control and operations,” *Comput. Chem. Eng.*, vol. 114, pp. 122–129, Jun. 2018, doi: 10.1016/j.compchemeng.2017.10.037.

[26]W. Hu, S. L. Shah, and T. Chen, “Framework for a smart data analytics platform towards process monitoring and alarm management,” *Comput. Chem. Eng.*, vol. 114, pp. 225–244, Jun. 2018, doi: 10.1016/j.compchemeng.2017.10.010.

- [27] Rieth, Cory A. (Pacific Science And Engineering Group, Inc); Amsel, Ben D. (Pacific Science And Engineering Group, Inc); Tran, Randy (Pacific Science And Engineering Group, Inc); Cook, Maia B. (Pacific Science And Engineering Group, Inc), "Additional Tennessee Eastman Process Simulation Data for Anomaly Detection Evaluation." Harvard Dataverse, 2017. doi: 10.7910/DVN/6C3JR1.
- [28] G. Lee, C. Han, and E. S. Yoon, "Multiple-Fault Diagnosis of the Tennessee Eastman Process Based on System Decomposition and Dynamic PLS," *Ind. Eng. Chem. Res.*, vol. 43, no. 25, pp. 8037–8048, Dec. 2004, doi: 10.1021/ie049624u.
- [29] J. H. Lee, J. Shin, and M. J. Realff, "Machine learning: Overview of the recent progresses and implications for the process systems engineering field," *Comput. Chem. Eng.*, vol. 114, pp. 111–121, Jun. 2018, doi: 10.1016/j.compchemeng.2017.10.008.
- [30] N. Asprion *et al.*, "Gray-Box Modeling for the Optimization of Chemical Processes," *Chem. Ing. Tech.*, vol. 91, pp. 305–313, Nov. 2018, doi: 10.1002/cite.201800086.
- [31] M. Kuhn and K. Johnson, *Applied predictive modeling*, vol. 26. Springer, 2013.
- [32] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [33] Y.-L. Xie and J. H. Kalivas, "Evaluation of principal component selection methods to form a global prediction model by principal component regression," *Anal. Chim. Acta*, vol. 348, no. 1, pp. 19–27, Aug. 1997, doi: 10.1016/S0003-2670(97)00035-4.
- [34] C. Chen and L. Breiman, "Using Random Forest to Learn Imbalanced Data," *Univ. Calif. Berkeley*, Jan. 2004.
- [35] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.
- [36] S. Raschka, "Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning," *ArXiv181112808 Cs Stat*, Nov. 2020, Accessed: Apr. 20, 2021. [Online]. Available: <http://arxiv.org/abs/1811.12808>
- [37] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 281–305, 2012.
- [38] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.
- [39] A. Vaswani *et al.*, "Attention Is All You Need," *ArXiv170603762 Cs*, Dec. 2017, Accessed: Oct. 13, 2020. [Online]. Available: <http://arxiv.org/abs/1706.03762>
- [40] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Netw.*, vol. 18, no. 5–6, pp. 602–610, Jul. 2005, doi:

10.1016/j.neunet.2005.06.042.

[41]T. Hayashi, S. Watanabe, T. Toda, T. Hori, J. Le Roux, and K. Takeda, “BLSTM-HMM hybrid system combined with sound activity detection network for polyphonic Sound Event Detection,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2017, pp. 766–770. doi: 10.1109/ICASSP.2017.7952259.

[42]M. Schuster and K. Paliwal, “Bidirectional recurrent neural networks,” *Signal Process. IEEE Trans. On*, vol. 45, pp. 2673–2681, Dec. 1997, doi: 10.1109/78.650093.

[43]D. Bahdanau, K. Cho, and Y. Bengio, “Neural Machine Translation by Jointly Learning to Align and Translate,” *ArXiv14090473 Cs Stat*, May 2016, Accessed: Dec. 02, 2021. [Online]. Available: <http://arxiv.org/abs/1409.0473>

[44]A. Graves, A. Mohamed, and G. Hinton, “Speech Recognition with Deep Recurrent Neural Networks,” *ICASSP IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 38, Mar. 2013, doi: 10.1109/ICASSP.2013.6638947.

[45]N. T. Ly, C. T. Nguyen, and M. Nakagawa, “An attention-based row-column encoder-decoder model for text recognition in Japanese historical documents,” *Pattern Recognit. Lett.*, vol. 136, pp. 134–141, Aug. 2020, doi: 10.1016/j.patrec.2020.05.026.

[46]H. T. Nguyen, C. T. Nguyen, P. T. Bao, and M. Nakagawa, “A database of unconstrained Vietnamese online handwriting and recognition experiments by recurrent neural networks,” *Pattern Recognit.*, vol. 78, pp. 291–306, Jun. 2018, doi: 10.1016/j.patcog.2018.01.013.

[47]“令和3年度農業分野におけるドローンの活用状況,” 農林水産省農産局技術普及課, Aug. 2021. [Online]. Available: <https://www.maff.go.jp/j/kanbo/smart/attach/pdf/drone-176.pdf>

[48]M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The Pascal Visual Object Classes (VOC) Challenge,” *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, Jun. 2010, doi: 10.1007/s11263-009-0275-4.

[49]T.-Y. Lin *et al.*, “Microsoft COCO: Common Objects in Context,” *ArXiv14050312 Cs*, Feb. 2015, Accessed: Jul. 19, 2021. [Online]. Available: <http://arxiv.org/abs/1405.0312>

[50]B. Zhou *et al.*, “Semantic understanding of scenes through the ade20k dataset,” *Int. J. Comput. Vis.*, vol. 127, no. 3, pp. 302–321, 2019.

[51]M. Cordts *et al.*, “The cityscapes dataset for semantic urban scene understanding,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 3213–3223.

[52]R. Girshick, “Fast R-CNN.” arXiv, Sep. 27, 2015. doi: 10.48550/arXiv.1504.08083.

[53]S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” Jun. 2015, doi: 10.48550/arXiv.1506.01497.

[54]J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” *ArXiv150602640 Cs*, May 2016, Accessed: Jul. 19, 2021. [Online]. Available:

<http://arxiv.org/abs/1506.02640>

[55]J. Redmon and A. Farhadi, “YOLO9000: Better, Faster, Stronger.” arXiv, Dec. 25, 2016. doi: 10.48550/arXiv.1612.08242.

[56]J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement.” arXiv, Apr. 08, 2018. doi: 10.48550/arXiv.1804.02767.

[57]A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection.” arXiv, Apr. 22, 2020. doi: 10.48550/arXiv.2004.10934.

[58]Z. Liu *et al.*, “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows.” arXiv, Aug. 17, 2021. doi: 10.48550/arXiv.2103.14030.

[59]J. Long, E. Shelhamer, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” Nov. 2014, doi: 10.48550/arXiv.1411.4038.

[60]O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” *ArXiv150504597 Cs*, May 2015, Accessed: Jul. 19, 2021. [Online]. Available: <http://arxiv.org/abs/1505.04597>

[61]H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid Scene Parsing Network.” arXiv, Apr. 27, 2017. doi: 10.48550/arXiv.1612.01105.

[62]K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN.” arXiv, Jan. 24, 2018. doi: 10.48550/arXiv.1703.06870.

[63]J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848.