

機械学習を用いた工場野菜の収穫量予測モデルの開発

多田桃大† 細田侑也† 後藤仁志† 戸田清太郎† 高山弘太郎‡

豊橋技術科学大学† 愛媛大学‡

1. はじめに

気温や日射量等の生育環境を管理しながらコンテナ施設内で野菜を栽培する野菜工場では、天候に依存せずに年中安定した収穫が見込まれる。しかし、露地栽培と比較して運用費用が高いため、植物工場の運用計画を最適化する必要がある。そこで本研究では、幼苗期の生育状況から将来の収穫量を予測することを目的とする。もし収穫量を事前に把握できれば、生育不良な株の生育環境を見直すことや市場の需要に合わせて供給量を調整することでフードロス削減できる。先行研究[1]では幼苗期のレタス画像から成長モデルを構築し重回帰で収穫量を予測した。本研究では、各株の生育段階を考慮するように成長モデルを構築して、成長が早いあるいは遅い株にも有効な収穫量予測モデルを提案する。

2. 植物工場

本研究では、豊橋技術科学大学が所有するコンテナ型植物工場に栽培するフリルレタス 45 株を対象とした。図 1 は植物工場の生育棚を表す。発芽から 14 日経過した株を生育棚に植え付け、植物工場での栽培 20 日目に収穫量として根を含まない新鮮重量を測定した。このとき空調機器は 25°C 冷房で設定し、室温は 24°C から 28°C を維持した。

植物工場では、毎日 1 枚クロロフィル蛍光画像[2]を撮影した。このとき、カメラは高さ 30cm の位置で固定している。広角レンズを取り付け同時に 9 株撮影するが、レンズ歪みは事前に取得したカメラパラメータで補正している。最終的に、各株が中央に収まるように補正後の蛍光画像を 9 つに分割した。



図 1 植物工場の生育棚

3. 生育段階を考慮する収穫量予測モデル

図 2 は提案法の概要図を表す。本研究では 10 日目までの蛍光画像から 20 日目の収穫量を予測する。まず蛍光画像から生育データとして投影面積とエッジ長を抽出する。上からみた大きさを表す投影面積は大津の二値化法で、葉の複雑さを表すエッジ長は Canny 法で計測する。そして、計測値を成長モデルに適用することで、将来的な生育データを取得する。先行研究[1]では、カーブフィッティングで各株の成長モデルを構築したが、単一の株の少ない計測値だけでは不安定で株ごとに異なる生育段階も反映されていない。そこで提案法では生育段階を考慮するように複数の株の計測値から成長モデルを構築する。

まず 1 番目の株の計測値を 1 日目から順に配置して成長モデルを構築する。2 番目の株の計測値は成長モデルに合うように日数をシフトしながら配置する。このシフト量が株ごとに異なる生育段階を補正する働きをもつ。ここで、生育データ y と日数 s に対する成長モデルを $y = f(s)$ と表すと、計測値 x_1, \dots, x_{10} におけるシフト量 T は次式で与えられる。

$$T = \arg \min_{-5 \leq t \leq 5} \sum_{i=1}^{10} \{f(s+t+i) - x_i\}^2$$

本研究では、0.1 刻みで t を変動させた。そしてシフトした計測値を含めて再度成長モデルを構築する。上記の動作を繰り返して作成した成長モデルを図 3 に示す。生育段階を考慮しながら複数の株の観測値を配置して安定な成長モデルを構築している。こ

Development of A Yield Prediction Model for Plant Factories Using Machine Learning

† Tota TADA

† Yuya HOSODA(Hosoda.yuya.ho@tut.jp)

† Hitoshi GOTO

† Seitaro TODA

† Kotaro TAKAYAMA

Toyohashi University of Technology (†)

Ehime University (‡)

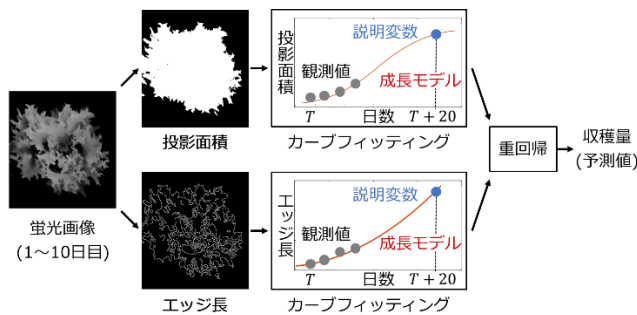


図2 提案法の概要図

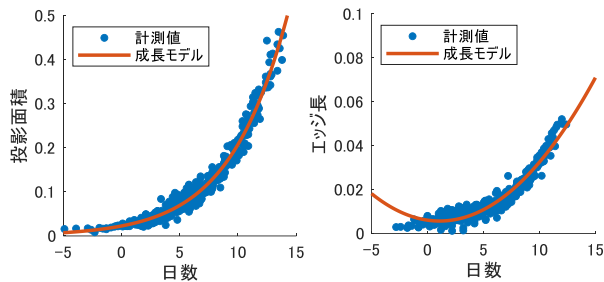


図3 成長モデル(左: 投影面積, 右: エッジ長)

のとき, 最終的な成長モデルに対する i 番目の株のシフト量を T_i と表すと, 将来の生育データは $f(T_i + 20)$ で与えられる. ただし本研究のカーブフィッティングでは, 投影面積ではロジスティック関数を, エッジ長では二次関数を採用した.

最後に, 将来の生育データを説明変数として重回帰モデルで収穫量を予測する. 先行研究では重回帰を採用したが, 本研究では線形重回帰モデルであるRidge 回帰と Lasso 回帰, Elastic Net, サポートベクター回帰を, 非線形重回帰モデルである k-最近傍法を, アンサンブル学習モデルであるランダムフォレストと XGBoost, CatBoost を新たに導入する.

4. 実験

図4は, 重回帰モデルとして重回帰を採用した先行研究[1]と提案法の予測結果を表す. このとき学習用データとして36株と評価用データ9株をランダムに選択している. 生育段階を考慮しない先行研究では, 標準的な成長段階である80g付近の株に対しては収穫量予測が有効であるが, 60g以下の株は本来よりも重く100g以上の株は軽く予測した. 一方で生育段階を考慮するように成長モデルを構築する提案法は, 成長が早いあるいは遅い株に対しても誤差を抑制した. このとき, 先行研究に対して提案法の決定係数は, 学習用データでは0.34, 評価用データでは0.81だけ改善した.

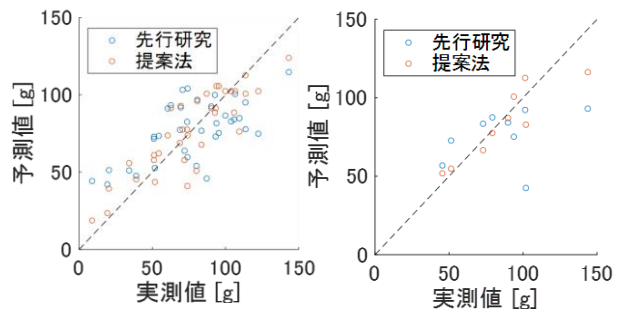


図4 収穫量予測結果(左: 学習用, 右: 評価用)

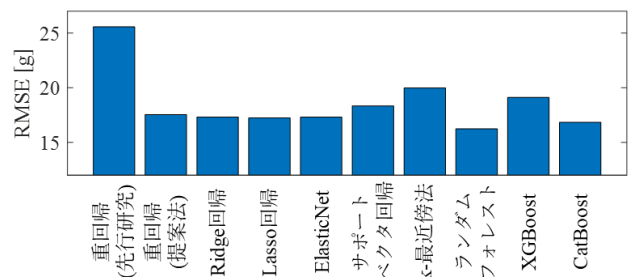


図5 各回帰モデルにおける評価値 RMSE

続いて, 図5は各回帰モデルにおける評価用データに対する評価値 Root Mean Square Error(RMSE)を表す. このとき, 汎化性能を検証するために, 交差検証($k = 5$)を実施している. 先行研究と比較すると, いずれの回帰モデルでも評価値は改善した. したがって, 生育段階を考慮しながら成長モデルを作成する手法が有効であることが示された. また, ランダムフォレストによる収穫量予測では, RMSE 16.22g で最低値を記録した. つまり, 単純な線形重回帰モデルでは表現不十分であった収穫量予測もアンサンブル学習モデルを採用することで十分に対応できることが示された.

5. おわりに

本研究では, 株ごとに異なる成長段階を考慮するように日数をシフトしながら成長モデルを構築する手法を提案した. そして, 9つの回帰モデルで将来の収穫量を予測した. その結果, ランダムフォレストを用いた提案法は, 先行研究よりも評価値 RMSE を 14.56g 改善した. 今後は, レタスの色情報や光合成情報を新たに特徴量に取り入れることを検討する.

参考文献

[1] Y. Ito, *et al.*, Plant growth prediction for lettuces using chlorophyll fluorescence information, *Proc. CIGR*, 2022.
 [2] K. Maxwell, and G. N. Johnson, Chlorophyll fluorescence—a practical guide, *J. Exp. Bot.*, 2000.