

施設栽培トマトにおける収量予測モデルを用いた農業経営改善

老林慶悟† 南野謙一†

岩手県立大学ソフトウェア情報学部†

1. はじめに

近年、センサーで計測して集めたビッグデータを解析し、収量を予測することで、効率的に栽培管理するスマート農業が進められている。施設野菜栽培では出荷量・雇用労働力の確保のため精度の良い収量予測の需要が高まっており、センサーを設置し収集データから機械学習により収量予測モデルを開発する研究が行われている。しかしながら、トマトは高温、着果負担等の原因により収量の増減が大きく、精度の良い収量予測モデルを開発するためには、センサーデータの選別、データの前処理等を行い、対象圃場に適合させなければならない。そこで本研究では、(株)いわて若江農園との共同研究により、施設で得られた環境データと収量データを使用し、機械学習により開発した実用的な収量予測モデルと、モデルの運用による農業経営改善について述べる。

2. 収量予測モデルの開発

2.1. リッジ回帰モデル

機械学習では環境データを前処理し、リッジ回帰モデルにより収量予測を行う。リッジ回帰は教師あり学習の手法の1つであり、回帰係数を推定する際、損失関数にモデルの複雑さを表す項を加え、これを含め誤差を最小化するように学習し過学習のリスクを低下させる。本研究では、説明変数に環境データの日射量と気温、目的変数に収量データをセットし、リッジ回帰モデルを開発する。

収量予測モデルには訓練データにより学習を行い、訓練データとは異なるテストデータにより評価を行う。訓練データには2018年10月から2021年5月までのデータを、テストデータには、2021年10月からの環境データを用いて2021年11月からの収量を予測する。精度は決定係数から評価する。

2.2. 説明変数と目的変数

説明変数は施設の日射量、日平均気温を前処理したものを用いる。積算日射量と収量を調査した結果、最も相関が高くなった収穫日から前4週間積算した値とする。日平均気温は一定期間の温度分布を表したHistogram関数(以下、H関数)を用いる[1]。H関数は定植から収穫開始日までの日平均気温を対象とし、1度刻みの度数分布表を作成して気温ごとの相対度数を求め、その値を用いる。



図1 リッジ回帰モデルによる収量予測

本研究では8週間の度数分布とする。

目的変数には収量データを用いる。しかし作業が休日の場合、収量は0gとなり、休日に採れるはずの収量は別の日の収量にカウントされるため、トマトが生育した環境本来の収量とは異なってしまふ。そこで1週間積算した値を取ることにする。さらに、休日の収量0gが生育本来の収量に及ぼす影響は収穫日の前1週間から後1週間にしか及ばない。そこで収穫日より前10日間、後10日間の3週の収量の移動平均を求めた値を採用する。[2]

2.3. 予測精度

2018年10月から2021年5月の訓練データで学習させたリッジ回帰モデルでは、決定係数は0.79となった。2021年11月からの収量データ(実測値)とテストデータによる予測値との適合を示す(図1)。テストデータでは決定係数は0.82となり、1年を通じて大きなずれはない。この結果より、リッジ回帰による収量予測モデルは若江農園西棟の収量予測に有用性があることが確認できた。しかしながら、収穫開始日から数日間は予測値が実測値の細かい傾向を捉えることが出来なかったため、訓練データとして採用する学習期間を分割し(冬春、夏秋トマト)、モデルを作成し実測値の細かい傾向を捉える必要がある。

3. モデルの精度向上

3.1. 着果数を説明変数に加えたモデル構築

若江農園では施設内に試験区域を設けトマトの生育状況を調査している。計測されるデータは実の大きさ、着果数、葉の大きさなどである。予測モデルの精度向上のためモデルの説明変数に着果数データを加えたモデルを開発した。

着果数は一定期間積算した値とする。積算期間は1日毎の積算着果数と収量を調査した結果、最も相関が高くなった収穫日から前30日間積算した



図 2 着果数を説明変数に加えた予測モデル



図 3 特定の期間に限定した予測モデル

値とする。モデルは西棟の 2018 年 10 月から 2021 年 5 月の期間を学習させ 2021 年 11 月から 8 月までの収量を予測した。(図 2)テストデータの決定係数は 0.87 となり、着果数を考慮しないモデルと比べて精度が上がった。

3.2. 栽培期間の分割による複数モデル構築

モデル開発において 1 シーズンを通しての訓練データを使用したものは実測値の細かい傾向まで捉えることが出来なかった。そこで、訓練データとして採用する学習期間を冬春、夏秋トマトの 2 つに分割しモデルを作成して実測値の細かい傾向を捉える修正を行った。テストデータは 2022 年 11 月から 6 月までとした。図 3 に示すように収穫開始日から数日は誤差を少なく予測することが出来た。この結果より実測値の細かい傾向まで捉えられるようになったが、予測期間の後半は実測値の変動があり傾向を捉えられなかった期間もあった。

4. 収量予測モデルを活用した農業経営改善

4.1. 週単位の収量予測の活用

若江農園へのインタビュー結果より、収量予測から農園の経営を効率的に行えることが分かった。若江農園では収穫の際、収量の目標値を日毎に設定している。予測モデルを運用する以前は昨年度の 1 か月毎の収量を日数で割り、1 日の目標値を設定していたが、トマト本来の生育状況を考慮した目標値とはなっていなかった。収量予測モデルを運用することで 1 週間先の収量を予測することが出来るため、1 週あたりで生育状況を考慮した目標値を設定することが出来るようになり、出荷の検討や人員調整など効率的な経営方法に改善できた。特に、出荷先へ収量の目安を伝えながら出荷作業を行うため、以前までは経験に基づき大まかな目安を伝えていたものが、収量予測モデルを運用す



図 4 平年値環境データによる収量予測

ることで収量の増減が予測でき、出荷先へ客観的なデータに基づく目安を伝えられるようになり過不足なく取引ができる。

4.2. 年単位の収量予測の活用

若江農園では 2018 年からデータを取得している。図 4 は西棟の気温と日射量の 2018 年からの平均値(平年値)を作成し、予測モデルで計算した値をグラフ化したものである。平年値の利用により西棟の 1 シーズン収量の予測ができた。若江農園では 1 年を通して栽培計画を立てるが、例年の収量と経験を基にした作成であった。予測値を基に時期ごとの販売計画作成や毎日の収量目標値、人員調整計画など農業経営の様々な分野への利用が可能になり以前と比べ客観的なデータに基づく作成へ改善できた。特に収量のピークを予測することで人員調整を効率的に行える。次年度の 1 シーズンを通しての収量の長期見通しが立てられ、若江農園の従業員も含め共有することで人員調整の理解も深められる。なお収量モデルは毎年更新し最新のものを使用し、1 シーズンの収量予測をする。

5. まとめ

本研究では施設栽培トマトを対象に、2018 年 10 月から 2021 年 5 月の訓練データを用いて収量予測モデルを開発し、2021 年 11 月からのテストデータで評価した。1 シーズンのテストでは大きな誤差はなく、モデルの有用性を確認できた。またこのモデルに栽培現場のデータである着果数を説明変数として加えると精度が向上することが明らかになった。また学習させる時期を分割することで細かい傾向も捉えることが可能となった。さらに平年値データを使用したモデルを作成することで 1 シーズンを通しての長期予測が可能であり、人員調整や販売計画などの農業経営に利用できるようになった。今後の課題として、実の重さや収穫速度を考慮したモデルを開発する。

参考文献

- [1] 遠藤隆也・割澤伸一・山田一郎, “トマト低段密植栽培における収穫日・収量の予測手法(計算モデル)の検討”, 日本生物環境工学会大会講演要旨, 2017.
- [2] 久枝和昇・仁科弘重, “大規模トマト生産温室における生産性向上に関する研究—積算日射量に基づいたトマトの出荷量予測—”, 植物環境工学(J.SHITA)19(1):11-18.2007.