

# 植物工場の果菜類を対象とした 週別収量の予測

外館有希<sup>†1</sup> 大場みち子<sup>†2</sup> 高森満<sup>†3</sup>

**概要:** 頻発する天候不順や異常気象の発生, 消費者の安全・安心志向の高まりから, 植物工場の需要が世界的に拡大している. 植物工場が抱える問題の一つとして, 収量に影響を与える要因の解析や収量の予測が困難であるという問題がある. 理由として, 植物工場での栽培作物では生物学的要因が解明されていない品種が多く, 労働力等の生産工学的要因も絡んでいることが挙げられる. 特に, 果菜類に分類される果実を食用とする野菜は開花や結実という生育過程があり, 他の作物と比較して環境の影響を受ける期間が長いため, より収量予測が困難である. これらの問題に対し, 本研究では果菜類の収量に関する要因を明らかにして, 統計モデリングを用いて果菜類の複数品種の収量予測モデルを開発することを目的とする. 本稿では, ミニきゅうりを対象作物とし, 重回帰モデリングを利用した収量予測モデルの開発し, その評価結果を報告する.

**キーワード:** 収量予測, 植物工場, 果菜類, 統計モデリング

## Weekly Yield Prediction of Fruits and Vegetables in Plant Factories

YUKI TODATE<sup>†1</sup> MICHIKO OBA<sup>†2</sup> MITSURU TAKAMORI<sup>†3</sup>

**Abstract:** Demand for plant factories is expanding worldwide due to frequent inclement weather, abnormal weather, and increasing consumer's safety and security consciousness. As one of the problems in plant factories, there is a problem that analysis of factors affecting yield or prediction of yield is difficult. As a reason, there are many varieties whose biological factors have not been clarified in cultivated crops in plant factories, and production engineering factors such as labor force are also involved. Especially, it is more difficult to predict the yield of vegetables classified as fruits and vegetables because they have growth processes such as flowering and fruiting, and the period affected by the environment is longer than other crops. For these problems, the purpose of this research is to clarify the factors related to the yield of fruits and vegetables and to develop a yield prediction model for multiple varieties of fruits and vegetables using statistical modeling. In this paper, a yield prediction model using multiple regression modeling is developed for mini cucumber and its evaluation results are reported.

**Keywords:** Yield Prediction, Plant Factory, Fruits and Vegetables, Statistical Modeling

### 1. はじめに

近年, 世界的に植物工場の施設数が増加し続けており[1], 今後も増加していくと予想されている[2]. 矢野研究所の調査によると, 2020年には施設栽培作物の世界市場規模は165億円にも上ると予想されている[3]. その背景には, 頻発する天候不順や異常気象の発生, 消費者の安全・安心志向の高まりなどがある. このような状況の中, 植物工場が注目されている.

植物工場とは, 高度な環境制御と生育予測を行うことにより, 野菜等の植物の周年・計画生産が可能な栽培施設のことである[4]. 一般的に, 植物工場は, 生育環境として人工光のみを用いる人工光型植物工場と, 人工光の他に太陽光も併用する太陽光型植物工場の2つに分類される[5].

このような植物工場が抱える問題の一つとして, 栽培作物の収量予測が困難であるという問題が挙げられる. 収量予測は, 生産者が生産計画や販売計画を立てる際の大きな目安となり, 過剰生産を防ぐ重要なものである. 収量予測が

<sup>†1</sup> 公立はこだて未来大学システム情報科学研究科  
Graduate School of Systems Information Science, Future University Hakodate.

<sup>†2</sup> 公立はこだて未来大学  
Future University Hakodate

<sup>†3</sup> 株式会社アブレ  
Apure Inc.

困難とされている主な要因は2つある。1つ目は、植物工場の栽培作物は、生長や収量に影響を与える生物学的要因が、解明されていない品種が多いという要因である[6][7]。2つ目は、環境制御を常に最適な状態に保つことが難しいという要因である。これは太陽光型植物工場に主に起きる問題である。その要因として、太陽光の利用や換気を行うことによる外気の影響や、従業員がいない夜間の時間帯や休日の環境制御の甘さが挙げられる。

植物工場の栽培作物の中でも、特に果菜類の収量予測が困難とされている。果菜類とは、ナスやトマト等の実がなる野菜のことである。困難とされている主な要因は2つある。1つ目は、葉菜類と比較し、果菜類は開花や結実という生育期間があるため、より環境の影響を受ける期間が長いという要因である。2つ目は、一つの苗から複数回・複数の実を収穫できるため、一つの苗からいつ・どれほど収穫できるかを推定することが困難という要因である。

そこで、本研究の目的は、植物工場に栽培される果菜類を対象として収量を予測することである。目的を達成するための目標として、果菜類の収量に関する要因を明らかにした上で、収量予測モデルの検討および評価を行い、開発した予測モデルを用いた収量予測システムを開発する。

## 2. 関連研究と課題

### 2.1 機械学習を用いた収量予測

機械学習を用いて収量予測を行なった研究は多い[8][9]が、これらの研究では、露地栽培の様々な作物を対象としたものが多く、膨大な量の学習データを用いることで、高精度な予測を実現している。しかし、植物工場では、機器の導入による環境変動や栽培品種の変動が著しいため、膨大なデータ数を用意することが難しいという問題がある。

機械学習は、処理プロセスをブラックボックス化してしまつたため、単純なモデルを除き、どのような処理プロセスを経て最終的な予測結果を算出しているのかの理解が困難であるという問題もある。農業従事者にとって、納得感のない予測モデルは受け入れ難いため、農業従事者が予測モデルの理論を理解可能なモデル開発が重要である。

機械学習を用いた多くの研究では、機械学習という技術の検証に重きを置いており、実際の農業現場で活用するための収量の要因分析や現場での検証が不十分である。

### 2.2 統計手法を用いた収量予測

統計モデリングによって、果菜類の収量予測モデルの開発をした研究として、奥野ら[10]やHoshi et. al. [11]の研究がある。奥野らの研究では、ベイジアンネットワークと重回帰分析を組み合わせた方法をアスパラガスの収量予測に適用した。その結果、予測結果に明確な根拠を示すことのできる回帰モデルを生成し、増加減少の傾向を捉えた収量推定ができたことを報告している。

Hoshi et. al.は、温室内のトマトを対象として、トポロジー事例ベースモデリング(TCBM)[12]と重回帰分析の2つの手法によって日別収量の予測をしている。その結果、TCBMによる予測が最も精度が良い結果となり、平均絶対誤差(MAE)が±31kg(26%の誤差)という結果を得たと報告している。

両研究に共通した課題は2つあると考える。1つ目として、手法の妥当性・信頼性に欠けるという課題が挙げられる。複数のモデリング手法による検討が行われておらず、その性能評価指標も少ないためである。2つ目として、実運用におけるモデルの活用方法の検討や評価まで実施していないという課題が挙げられる。これは、2.1の機械学習を用いた収量予測の研究の課題として挙げた点と同様のことである。

### 2.3 研究課題

以上の関連研究での課題より、本研究における研究課題は、次に示す通りである。

- (1) 植物工場の果菜類に適した正確な予測モデルの開発
- (2) 実運用におけるモデルの活用方法

## 3. 提案手法

2.3で述べた研究課題に対する解決アプローチは、次に示す通りである。

- (1) 予測モデルに対する説明変数の回帰構造の探索と評価が可能な統計モデリング手法の活用
  - (2) 生産者の意見を取り入れた収量予測システムの開発
- 本章では、上記の解決アプローチについて詳細に説明する。

### 3.1 統計モデリングの候補

本研究では、収量予測モデルの開発に伴い、まず植物工場の果菜類の複数品種を対象とした収量に影響を及ぼす要因を明らかにする。要因を明らかにするために、利用する統計モデリング手法の選定基準は、①予測モデルに対する説明変数の回帰構造の探索と評価が可能、②予測精度の高いモデルの構築が可能という2点である。①と②を担保できる以下の3つをモデリング手法の候補として選定する。

候補の番目は、重回帰分析(MLR: Multiple Linear Regression)である。MLRは2つ以上の説明変数を使用して連続値の目的変数を予測するための一般的な統計手法である。MLRは、収量予測において多く用いられており、果菜類の収量予測においても多く用いられている。MLRによって、ある程度の予測精度も得ていることから、本研究でもMLRを選択する。また、最も良い説明変数の組み合わせを調べるために、一般的な統計的変数選択手法である赤池情報量基準(AIC)に基づくstepwise法を実施する。

第2のモデル候補は、一般化加法モデル(GAM: Generalized Additive Model)である。GAMは、一般化線形モデル(GLM: Generalized Linear Model)の各特徴量に重みをつけるだけでなく、関数とする事で複雑な現象も表現することができる非線形にも対応したモデルである。GAMは、目的変数と説

明変数の関係性が分かりやすいという線形モデルの利点を保ったまま、機械学習のモデルと同レベルでの精度の予測が可能であるため、選定する。

最後に、第3の候補として、多変量適応的回帰スプライン(MARS: Multivariate Adaptive Regression Splines)を選定した。MARSでは、CART(CART: Classification And Regression Trees)のステップ関数近似を打ち切りベキ乗基底関数に変更し、局所的に線形モデルをあてはめることで予測精度を高めたものである。GLMと比較して説明変数間の交互作用を樹木構造でティッピングポイントも含めて明示的に表現できることに強みがある[13]ため、選定する。

### 3.2 モデルの性能評価指標

回帰モデルの性能評価指標として、決定係数( $R^2$ )、二乗平均平方根誤差(RMSE)、平均絶対誤差(MAE)を利用する。 $R^2$ は、モデルの実際のデータへの適合度を評価する。RMSEは、実測値と予測値の差を測定し、外れ値を誇張して表現するものである。MAEは、誤差の単純な平均であるため、外れ値の影響が小さいという特徴を持つ。

開発したモデルに対して、 $R^2$ 、RMSE、MAEによって予測精度の評価を行い、最適なモデリング手法を選定する。

### 3.3 収量予測システムの開発

開発したモデルが実際の現場で本当に使用できるものであるかを検証するため、開発したモデルを用いて収量予測システムを開発する。システムの対象者は、植物工場の現場で生産業務を担う社員またはマネジメント業務を担う社員を想定している。提案システムの評価として、ユーザテスト、アンケート、ヒアリングによる定性的評価と、提案システムの予測結果と実測値の誤差を比較する定量的評価を実施する。

## 4. 対象の実験施設と作物

本章では、本研究の対象とする実験施設と予測対象の果菜類の作物について述べる。

### 4.1 実験施設

本研究の対象とする実験施設は、北海道函館市七飯町にある太陽光型植物工場(以下、植物工場Aと表記)である。植物工場Aでは、2つのハウスを所有しており、水耕栽培による野菜の生産・直販を行い、果菜類7種、葉菜類17種の計24種という多品目の野菜を栽培している。温度、湿度、 $CO_2$ 濃度、養液濃度等のハウス内の環境データのセンシングに加え、気温、湿度、降水量、光量等の外気象のデータのセンシングも行い、工場内の環境制御をしている。ハウス内の環境データだけでなく、外気象のデータもセンシングも行うことで、ハウス内の環境制御に役立てている。

植物工場Aでの果菜類の収量予測の必要性については、マネジメント業務を担う社員と実際に生産業務を担う社員へのヒアリングからも確認している。

## 4.2 対象とする作物

まず初めに予測モデルを開発する果菜類の作物として、ミニきゅうりのラリーノホワイトという品種を選択する。図1は、実際に栽培されているミニきゅうりの様子である。ミニきゅうりは、植物工場Aの主要の栽培作物であるため、栽培面積が大きく、収穫サイクルも早く、ほぼ毎日出荷をしている。この状態から、生産者にとって収量予測モデルを開発する需要が高いと考えたため、対象品種とした。ミニきゅうりを対象とした予測モデルの開発後、他の果菜類の予測モデルを開発していく。



図1 栽培されているミニきゅうり

## 5. 対象データと問題設定

本章では、モデリングに用いるデータと問題設定について述べる。

### 5.1 対象データ

モデリングとその評価に用いるデータは、2018年2月～2019年6月の期間の全67週分である。対象データは、環境データと生産データの大きく2種類に分けられる。

#### 5.1.1 環境データ

環境データは、ハウス内環境データと外気象データの大きく2種類に分けられる。これらのデータは、ハウス内外の数カ所に設置されたセンサーにより毎分記録され、10分毎の平均値が計測されている。表1に、計測されている環境データ一覧を示す。

養液濃度と照度は、説明変数の対象から除外した。養液濃度は、常に一定に設定されており、実際にデータの分散を確認しても、大きな変化が見られなかったからである。照度は、照射時間を朝4時～夜7時までの15時間と定めており、時間や日毎に違いが見られないため、除外した。変化が非常に少ないデータを説明変数に組み込んでも、重回帰モデルにおける説明変数の重要度は極めて低くなり、ノイズになる可能性が高いと判断したためである。

表1 計測されている環境データ一覧

ハウス内データ	外気象データ
● ハウス内温度	● 気温
● ハウス内湿度	● 外気湿度
● CO <sup>2</sup> 濃度	● 光量
● ほうさ	
● 養液濃度	
● 照度	

### 5.1.2 生産データ

生産データには、播種日、定植日、収穫日、収量、収穫本数、等級(A 級品, B 級品)がある。等級は、各野菜の商品規格に基づき、商品として出荷可能かの判断や価格の決定に用いられる。ミニきゅうりの場合は、曲がり度や傷の有無によって決まる。A 級品, B 級品は価格や販売先は異なるが、どちらも出荷可能という点では同じであるため、目的変数となる週別収量には、A 級品, B 級品の両方の収量を含めることにする。

### 5.2 問題設定

植物工場のマネジメント社員と生産社員へのヒアリングにより、目的変数として、収量を予測する日(以下、予測日)の次の日から 1 週間分の収量を設定した。予測日は毎週土曜日と固定し、それに合わせたデータセットを作成する。毎週土曜日とするのは、植物工場 A からのニーズとして、次の週の販売計画を立てるための会議が毎週土曜日にあり、その会議で来週の収量予測のデータを用いたいというニーズがあるためである。目的変数である予測日の次の日から 1 週間分の収量(以下、週別収量)を、予測日当日までの環境データや生産データを使って予測することを目指す。

## 6. 予備調査

予備調査の目的は、収量を正確に予測するために、ミニきゅうりの収量に関係する要因を調査することである。実施する予備調査の項目は、以下に示す通りである。

- (1) 週別収量の推移の傾向分析
- (2) ハウス内環境データの分析
- (3) 目的変数との相関分析

### 6.1 週別収量の推移の傾向分析

週別収量の推移の傾向と収量増減の要因を分析した。図 2 は、モデリングに用いる期間の収量の推移をグラフにしたものである。基本的な統計量を算出したところ、週別収量の平均は 135kg(4500 本)、日別収量の平均は 19.3kg(642 本)であることがわかった。

まず、時系列データに現れるデータの傾向パターンについて分析した。収量のデータは時系列データの一種である。時系列データに現れるデータの傾向パターンとしては主に 4 つあり、傾向変動、循環変動、季節変動、不規則変動があ

る。分析の結果、傾向変動、循環変動、季節変動についての傾向は確認できず、不規則変動の傾向パターンであることがわかった。

次に、収量が大きく変動している時期とその要因について分析した。収量が大きく増加しているのは 2018 年 5 月末から 10 月初旬にかけての期間である。この増加が株数の増加によるものなのかを株数のデータで調べたところ、季節によって播種する株数を特に大きく変えているわけではないため、株数が要因でないことがわかった。そこで、増加は環境による要因が大きいと考え、目的変数と環境データの相関を分析することにした。この分析については 6.3 で述べる。以上の分析より、日付データ(収穫日、播種日、定植日)と株数は、収量との関係性が見られないため、説明変数から外すこととした。

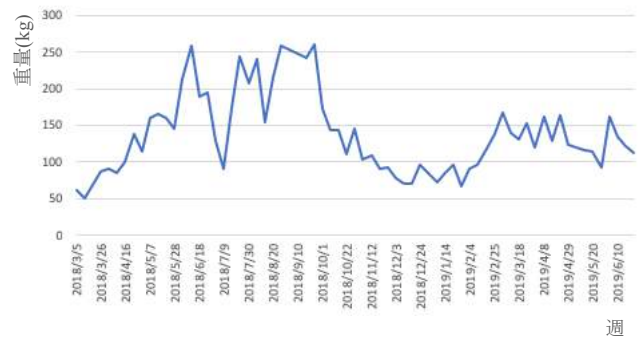


図2 ミニきゅうりの週別収量の推移

### 6.2 ハウス内環境データの分析

温度、湿度、ほうさ、CO<sup>2</sup>濃度、光量の 1 日の平均、最高値、最低値、最高値と最低値の差という複数の統計量を算出し、グラフで可視化することによって、各環境データの日単位、年単位での推移とデータの分布を分析した。本稿では、その分析結果の一例として湿度を取り上げて説明する。図 3 に、年単位での湿度の推移を示す。これより、年単位と日単位共に変動が大きいことがわかる。この傾向は、湿度以外の他の環境データにおいても類似した結果となった。以上の結果より、データに差異が見られたため、温度、湿度、ほうさ、CO<sup>2</sup>濃度、光量を説明変数として含めることにした。



図3 年単位・日単位での湿度の推移

### 6.3 目的変数との相関分析

#### 6.3.1 分析内容与方法

この分析では、収量に影響を与える要因と環境データの積算期間で最も収量に影響を及ぼす期間を調査した。データの前処理として、これまでの分析結果を参考にし、環境データの積算期間を複数パターンによって説明変数を作成した。分析は、目的変数との単相関係数の算出と散布図の作成によって行なった。単相関係数の算出には、Spearman の順位相関係数を用いた。

#### 6.3.2 前処理

環境データの積算期間を、基準日(予測日前日)午前までの積算に関して、7日前から、14日前から、28日前からの積算の3パターンで作成した。7日前としたのは、ミニきゅうりの開花から収穫までのおよその日数であり、このあたりの期間の環境要因が最も影響するという報告が多かった[7][9]からである。しかし、定植後40日までの環境要因がきゅうりの収量に関係するという報告もあり[14]、開花以前の時期の環境が最も影響を及ぼしている可能性もあるため、14日前と28日前での積算期間の説明変数も作成した。

これらを踏まえて作成した説明変数の一覧を表2に示す。各環境データは、基本的に対象期間の平均値を使用した。ハウス内の温度、湿度、気温の3つについては、関連研究[9]より、平均値の他に、最高値、最低値、最高値と最低値の差の3つも追加して、説明変数を作成した。

表2 作成した説明変数の一覧

ハウス内データ	外気象データ	生産データ
● 温度	● 気温	● 1週間前の週別収量
● 湿度	● 湿度	● 2週間前の週別収量
● CO <sub>2</sub> 濃度	● 光量	● 3週間前の週別収量
● ほうさ		● 4週間前の週別収量

#### 6.3.3 結果と考察

週別収量との単相関が絶対値 0.5 以上のものを抽出した結果が表3である。なお、互いに強い相関(絶対値 0.8 以上)が認

められた項目については相関関係が弱い方を除去している。

週別収量との相関関係が強いものを抽出した結果、外気象データ、ハウス内環境データ、生産データの3カテゴリ全体的に変数が抽出された。外気温最低(積算期間が収穫日の14/28日までの)単相関係数が0.662となり、環境データの中では最も相関関係が強く見られた。図4は外気温最低と週別収量の散布図であり、散布図からも正の相関が確認できる。外気温最低と光量との相関が高かったことから、太陽光型植物工場では外気象の影響を多く受けていることが確認できた。最低気温の相関関係が強くなった理由として、ミニきゅうりの栽培適温は、およそ18~25度といわれており、それよりも低すぎたり高すぎたりするとミニきゅうりの生育が阻害されるという傾向が現れたと考える。

次に、最も収量に影響を及ぼす期間についての分析結果として、多くの環境変数の場合で、収穫日の28日前からの積算が最も収量との関係性が高いことがわかった。28日前は、ミニきゅうりの栄養成長の最終段階の時期と一致する。

1週間前の週別収量の相関係数が0.774となり、説明変数の中で最も相関関係が強いという結果となった。この結果は、関連研究[11]の環境データよりも生産工学的要因に関わるデータの方が収量への影響度が高くなるという結果と一致した。

これらの結果について、生産者に確認したところ、生産者の感覚とほぼ一致することを確認した。

表3 週別収量との単相関

変数	積算期間	相関係数
外気温最低	収穫日の14/28日まで	0.662
光量	収穫日の28日まで	0.508
温度(ハウス内)	収穫日の28日まで	0.657
CO <sub>2</sub> 濃度(ハウス内)	収穫日の28日まで	-0.597
1週間前の週別収量	—	0.774
2週間前の週別収量	—	0.719
3週間前の週別収量	—	0.658

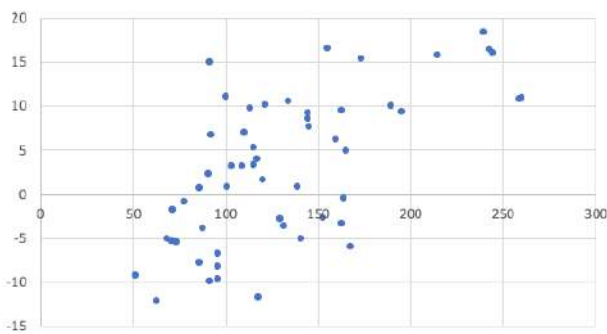


図4 最低気温と週別収量との単相関

## 7. 重回帰分析における収量予測

本章では、重回帰分析によるモデルの開発について述べる。

### 7.1 目的と方法

実験目的は、重回帰分析における収量予測の精度を確かめることである。説明変数の選択法として stepwise 法を利用し、AIC による評価値が最も高くなる変数の組み合わせを調べる。分散拡大要因(VIF: Variance Inflation Factor)を求めることで、多重共線性が発生していないかどうかを確認する。

### 7.2 評価方法

対象データの 4/5 を学習データに、1/5 をテストデータに分割して、Hold-out 法によって評価をする。提案手法でも述べたように、回帰モデルの評価には、 $R^2$ 、RMSE、MAE の 3 つを性能評価指標として用いる。

### 7.3 評価結果

AIC による stepwise 法の結果、1 週間前の週別収量とハウス内の温度(収穫日の 14, 28 日前から予測日の午後まで積算)の説明変数の組み合わせが AIC の値が最も小さくなり、重回帰モデル式(1)  $Weekly\ Yield = 0.626 X_1 + 19.44 X_2 - 368.205$  を得た。  $X_1$  = 予測日の 1 週間前の量,  $X_2$  = ハウス内温度の平均(収穫日の 14/28 日前からの積算)である。学習データとテストデータにおける 3 つの性能評価指標による評価結果を表 4 に示す。予測モデルの各説明変数についての統計量をまとめたものが表 5 である。  $p < 0.05$  となったことから、選ばれた説明変数は有意であることが確認できた。図 5 の週別収量の実測値と予測値の比較結果では、実際の収量(青線)と予測結果(赤線)を確認すると、増加・減少の傾向をとらえた予測ができていることが分かる。しかし、収量が大きく変動している 2018 年 5 月末から 10 月初旬では、誤差が大きくなる傾向があることがわかった。

6.3 で単相関係数を調べたときに、過去の収量とハウス内温度との相関係数が 0.590 と相関が高かったため、多重共線性が発生していないかを調査した。VIF を求めた結果、1.734( $< 10$ )であったことから、式(1)における多重共線性の疑いは小さいと考える。

表 4 各性能指標による評価結果

評価指標	学習	テスト
$R^2$	0.681	0.746
RMSE	29.4	36.7
MAE	23.4	8.79

表 5 予測モデルの各係数の値と有意性

項目	1 週間前の週別収量	ハウス内温度
係数	0.690	12.9
標準誤差	0.105	5.604
t 値	6.550	2.301
p 値	2.795E-08 ( $p < 0.05$ )	0.0255 ( $p < 0.05$ )

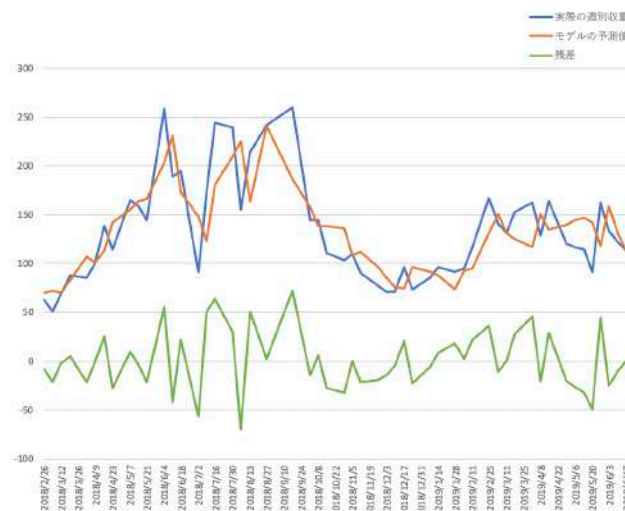


図 5 週別収量の実測値と予測値の結果

### 7.4 考察

表 4 の学習データとテストデータの評価結果の比較から、テストデータでの精度が学習データの精度より大きく下回らないことが確認された。よって、過学習は発生していない可能性が高いことがわかるため、今回開発したモデルは有効である。そこで、テストデータの結果を用いて各性能評価指標による評価結果の考察をする。

まず  $R^2$  は、一般にモデルの実際のデータへの適合度があるといえる 0.7 以上となり予測値が実測値と同じ挙動を示すため、モデルが意味を持つことがわかる。関連研究[15]での精度  $R^2 = 0.743$ (促成栽培)と比較しても高い値となった。

次に、外れ値を誇張して表現する RMSE では、RMSE 36.7 となり、週別収量平均値 135kg に対して、27.2%の誤差となった。関連研究[9]の RMSE の値 10.5(果菜類のみ)と比較すると、予測誤差が大きいという結果となった。RMSE の値が大きくなった要因として、収量の増減を正確に捉えられて

いないことが挙げられる。そのため、収量が大きく増減する要因を今後分析していく必要がある。

最後に、予測値の精度を表す MAE では、MAE 8.79 となり、週別収量平均値 135kg に対して 6.51%の誤差となった。関連研究[11][15]と比較して予測精度が高いことがわかった。

本実験より、重回帰分析によって、一定の予測精度を示すモデルを開発できることが示唆できた。しかし、誤差の面ではまだ改善の余地があることがわかった。

## 8. 収量予測システムの開発

### 8.1 システム概要

今後開発した各種予測モデルを元に、1週間分の収量の予測結果を表示する Web システム(以下、提案システムと記す)を開発する。提案システム概要を図 6 に示す。利用ユーザは、植物工場でのマネジメント社員と生産社員を対象とする。ユーザが収量予測システムを利用することで、正確な生産・出荷計画の立案に役立てることを想定している。

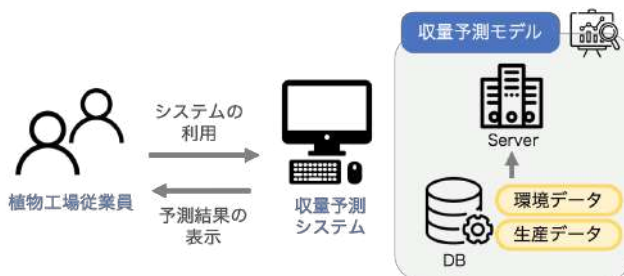


図 6 提案システム概要

### 8.2 システム評価

提案システムの評価は次の 2 つを予定している。

- (1) ユーザテスト、アンケート、ヒアリングによる定性的評価

ユーザテストでは、植物工場のマネジメント社員と生産社員の数名に、1週間程度、業務の中で予測システムを利用してもらう。利用後、使いやすさ、分かりやすさなどについて、リカード尺度を用いたアンケートで評価する。アンケート後に、ヒアリングを実施することにより、アンケートだけでは拾いきれなかった利用者の評価を加味して総合的な評価をする。

- (2) 予測モデルの予測結果と実測値の誤差を比べる予測精度の定量的評価

予測精度の評価では、提案システムで予測した結果と、実際の収量を比較し、どの程度正確に収量を予測できたかを評価する。

## 9. まとめと今後の課題

### 9.1 まとめ

本研究の目的は、植物工場で栽培している果菜類を対象

とした収量予測モデルを開発することである。本稿では、単相関係数を用いて目的変数との相関関係が強い要素の抽出を行い、重回帰分析を利用して収量予測モデルを開発した。実験により、過去の週別収量とハウス内の温度が最も良い説明変数の組み合わせとして選ばれ、増加・減少の傾向をとらえた予測が可能であることを確認した。

### 9.2 今後の課題

モデルの精度向上を目的として、労働時間等の生産工学的要因を説明変数として追加することを検討している。重回帰モデルによる精度改善を行なった後、複数のモデリング手法による比較を行い、最も良い精度となる手法と説明変数の組み合わせを探し出す。その後、開発したモデルを用いた収量予測システムを開発し、その有効性を評価していく。

### 謝辞

本研究を進めるにあたり、予測モデルの開発に用いた植物工場内の環境データやミニきゅうりの生産データは、株式会社アプレの協力によるものである。ここに深く感謝の意を表す。

### 参考文献

- [1]伊地知宏:NAPA リサーチ・レポート2018第3章 植物工場ビジネス - 植物工場経営の現状と将来展望,  
[https://www.nomuraholdings.com/jp/company/group/napa/data/20180219\\_03.pdf](https://www.nomuraholdings.com/jp/company/group/napa/data/20180219_03.pdf), (参照 2019-12-20).
- [2]一般社団法人日本施設園芸協会, “大規模施設園芸・植物工場実態調査・事例集”,  
<https://www.maff.go.jp/j/seisan/ryutu/engei/sisetsu/pdf/daikibo.pdf>, (参照 2019-12-20).
- [3]矢野研究所: 高機能・高付加価値型「次世代植物工場」市場に関する調査を実施(2018年) [https://www.yano.co.jp/press-release/show/press\\_id/1990](https://www.yano.co.jp/press-release/show/press_id/1990), (参照2020-01-07).
- [4]農林水産省: 植物工場の説明 一般財団法人・社会開発研究センター, <http://www.maff.go.jp/heya/sodan/1308/01.html>, (参照2019-07-15).
- [5]古在 豊樹: 図解でよくわかる植物工場のきほん: 設備投資・生産コストから、養液栽培の技術、流通、販売、経営まで, 誠文堂新光社, 2010, (参照2014-11-17).
- [6]一般社団法人日本施設園芸協会: 大規模施設園芸・植物工場実態調査・事例集,  
<https://www.maff.go.jp/j/seisan/ryutu/engei/sisetsu/pdf/daikibo.pdf>, (参照 2019-12-20).
- [7]博通伊藤, 律也山下: 植物工場におけるレタス成長の解析, 植物工場学会誌, 11巻,1号, pp.50-58,(1999).
- [8]A. Chlingaryan, S. Sukkarieh and B. Whelan: Machine learning approaches for crop yield prediction and nitrogen status estimation

- in precision agriculture: A review, *Comput.Electron.Agric*, Vol.151, pp. 61–69, (2018).
- [9]González-Sánchez, Alberto Frausto-Solis, Juan & Ojeda, Waldo, Predictive ability of machine learning methods for massive crop yield prediction, *SPANISH JOURNAL OF AGRICULTURAL RESEAR*, (2014).
- [10]奥野源, 新谷俊了, 機械学習と統計モデリングを組み合わせた手法によるアスパラガスの収量推定, *日本ソーシャルデータサイエンス論文誌*, Vol.2, No.1, pp.14-18, (2018).
- [11] Takehiko Hoshi, Taro Sasaki, Hiroaki Tsutsui: A daily harvest prediction model of cherry tomatoes by mining from past averaging data and using topological case-based modeling, *Computers and Electronics in Agriculture*, Vol. 29, No. 1, pp. 149-160, (2000).
- [12]宏明筒井, 淳黒崎, 友彦佐藤: 履歴データを事例として使用する非線形モデリング技術 TCBM: Topological Case Based Modeling, *計測自動制御学会論文集*, Vol. 33, No. 9, pp. 947–954, (1997).
- [13]孝典松井, 徹宇賀田, 尚町村: 機械学習アルゴリズムによる耕作放棄の要因分析および予測モデルの開発, *土木学会論文集(環境)*, Vol. 70, No. 6, II\_131-II\_139, (2014).
- [14]鈴木真美, 松尾誠治, 梅田大樹, 岩崎泰永: CO<sub>2</sub>施用時の高い相対湿度がキュウリの生育, 光合成速度, 窒素含量に及ぼす影響, *日本冷凍空調学会論文集*, Vol.31, No.3, pp.331-337, (2012).
- [15]大河内理貴, 青野雅樹, 川嶋和子: ノンパラメトリック回帰を利用した施設園芸作物の生産性の推定, *DEIM Forum*, E5-2, (2011).