

植物収穫時品質や収量に関連する経時特徴量分析手法の検討

水野涼介[†] 柴田瞬^{††} 峰野博史^{†‡}

[†]静岡大学情報学部 ^{††}静岡大学大学院総合科学技術研究科 ^{†‡}静岡大学学術院情報学領域

1. はじめに

国内の農業分野において、農業従事者の持つ高品質な作物栽培技術を次世代の新規就農者へ継承することが困難となってきている。特に高糖度トマトに代表される、高品質な栽培技術は作物に対する長い年月をかけて培われた経験と勘によって行われており、継承に時間を要する。そのため、作物に対する経験と勘が十分に培われていない新規就農者への継承は難しく、高品質な作物栽培技術は喪失の危機にある。

この問題を解決するために、ICTを用いて、熟練農家の持つ高品質な作物栽培技術を形式知化する研究が進められており、気象環境が植物品質に与える影響の分析が行われている。しかし、多くの既存手法は経時特徴変化において時間帯や生育ステージが考慮されておらず、気象環境が与える経時特徴変化の分析が困難であった。

本研究では、同じ気象環境であっても、植物に与える影響が時間帯や生育ステージで異なることに着目し、経時特徴量データを生育ステージや特定の時間帯ごとに分割することで、高次元データを作成し、特徴量選択を用いて重要な特徴量を選別することで、植物の生育状態を考慮しつつ、気象環境と収穫時品質や収量との関係性を分析する手法を提案する。

2. 関連研究

植物の収穫時品質や収量と経時特徴量データに対する分析手法として、積算温度と収穫時品質との相関分析や分散分析を用いた手法[1]、収集した環境データのマイニング手法としてデータの単純化[2]がある。[1]は、気温の閾値を定め、栽培期間中の日平均気温が閾値を超えている温度差を積算することで積算温度を算出する。その後、算出した積算温度を元に収穫時品質との統計的な分析を行うことで、気温から糖度予測を行っている。ここで、積算温度算出時に日平均気温を用いているが、1日の時間帯毎に積算温度を算出し、時間帯積算温度を用いて分析を行うという方法も考えられる。[2]では、環境データを-1, 0, 1のインデックスで単純化することで、環境データ特性の把握を容易とする。また、インデックス化された値の合計値を用いることで、分析を容易としている。しかし、植物は成長とともに生育ステージが変化するため、同じ環境条件でも与える影響が異なる。

つまり、植物の収穫時品質や収量と生育・環境データとの関係を高精度に分析するには、時間帯や生育ステージの違いといった経時特徴量まで考慮する必要があると考える。

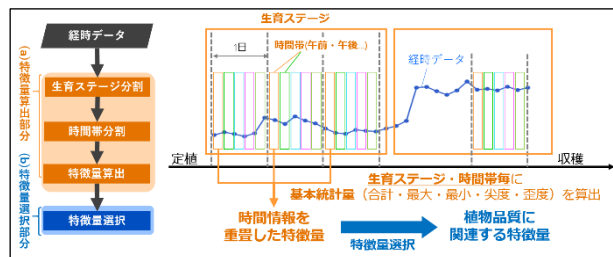


図1 提案手法概要図

3. 提案手法

本研究では、熟練農家の経験と勘による高品質な作物栽培技術を形式知化するため、植物の生育ステージや時間帯で植物に与える影響の異なる気象環境と収穫時品質や収量との関係性を分析する手法を提案する。提案手法の概要を図1に示す。

提案手法は大別して、特徴量算出部分(図1(a))と特徴量選択部分(図1(b))から構成される。特徴量算出部分においては、はじめに経時特徴量データを生育ステージ毎に分割し、新たな経時特徴量データを作成する。次に、作成した経時特徴量データに対し、1日における特定の時間帯ごとに分割し、分割したデータの基本統計量を算出することで、生育ステージや時間帯といった、時間情報を重畳した特徴量を作成する。熟練農家は一般的に、特定の日の気象環境や植物の萎れ具合などの局所的な変化を考慮した栽培だけでなく、定植日や開花日からの日数や積算温度といった、大域的な変化を考慮した栽培を実施していると言われている。そのため、算出する基本統計量には、合計と最大、最小、尖度、歪度の5つの基本統計量を用いる。

特徴量選択部分では、Lasso (Least Absolute Shrinkage and Selection Operator[3])を用いた特徴量選択を行う。Lassoは回帰モデルの損失関数に対し、 L_1 正則化項を加えた損失関数を最小化する、パラメータ推定手法である。 L_1 正則化は L_1 ノルムに基づいた正則化を行うため、特徴量選択を可能としており、特徴量選択にLassoを採用する。特徴量算出部分により、算出された時間情報を重畳した特徴量と収穫時糖度、収量との相関分析、特徴量選択部分において選択された特徴量に対する重要度を分析する。

4. 基礎評価

4.1. 実験方法

植物収穫時品質や収量に関連する経時特徴量分析の基礎実験として、提案手法を用いて算出された特徴量と収穫時糖度、収量との相関分析、提案手法を用いて算出された特徴量に対する特徴量選択を行った。本実験は、静岡大学農学部のトマト低段密植栽培実験現場内の12か所の栽培処理区で温度と飽差、散乱光を無線散乱光センサノード[4]で、茎径をレーザラインセンサ(HL-T1)で、

Proposal of timeseries feature analysis method for harvest quality and yield

Ryosuke Mizuno[†], Shun Shibata^{††}, Hiroshi Mineno^{†‡}

[†] Faculty of Informatics, Shizuoka University

^{††} Graduate School of Integrated Science and Technology, Shizuoka University

^{†‡} College of Informatics, Shizuoka University

表 1 収穫時糖度に対する特徴量選択結果

| 重要度 | 生育ステージ | 時間帯 | 特徴量 |
|-------|--------|-----|---------|
| 0.18 | 2 | I | 蒸発散量_最大 |
| 0.08 | 3 | III | 蒸発散量_尖度 |
| 0.075 | 3 | V | 温度_歪度 |
| 0.065 | 6 | III | 温度_歪度 |

株重量を重量計 (EK-3000i) で計測し、2017年5月18日から2017年7月19日の間、1分間隔で収集したデータを用いて実施した。糖度、収量データは収穫時にそれぞれ糖度計 (PAL-1)、重量計 (GF-600) を用いて計測した。また、株重量と灌水量の差分をとることで蒸発散量を、現在までの茎径最大値と現在の茎径との差である DSD (Difference in Stem Diameter) を算出した。

生育ステージは、1. 定植日から開花日、2. 開花日から開花後10日、3. 開花後10日から開花後20日、4. 開花後20日から開花後30日、5. 開花後30日から開花後40日、6. 開花後40日から収穫日の6ステージで区分した。時間帯区分は、I. 日の出から11時、II. 11時から14時、III. 14時から日没、IV. 日没から日没5時間後、V. 日没5時間後から日の出の5段階で区分し、日の出時刻は5時、日没時刻は19時とした。

4.2. 実験結果

表 1 に収穫時糖度を目的変数として特徴量選択を行った結果における重要度の高い5個の特徴量を示す。表1から最も重要度の高い特徴量は、生育ステージ 2 の日の出から 11 時における蒸発散量の最大値であった。蒸発散量は毎日の天候に応じて変化し、晴天時に PPF (Photosynthetic Photon Flux Density: 光合成光量子束密度) と共に多くなる傾向がある[6]。そのため、糖度の高い処理区ほど生育ステージ 2 の日の出から 11 時における時間帯は日射量が多く、光合成が行われたと考えられる。また、蒸発散量が給水量を上回ったことで水分ストレスがかかり、糖度が高まったと考えられる。したがって、生育ステージ 2 の日の出から 11 時における蒸発散量の最大値が最も重要度が高くなったと考える。表 2 に特徴量算出部分において算出された特徴量と収穫時糖度との相関分析結果において相関の高い5個の特徴量を示す。表2から収穫時糖度と最も相関の高い特徴量は、生育ステージ 2 の日の出から 11 時における蒸発散量の最大値であり、特徴量選択した結果と同様であった。以上の結果から、生育ステージ 2 の日の出から 11 時における蒸発散量を高める栽培を行うことで収穫時糖度を高められると考える。

表 3 に収穫時収量との相関分析結果において相関の高い 10 個の特徴量を示す。表 3 から生育ステージ 6 の各時間帯における飽差の最小値が強い正の相関を持っていることがわかる。飽差はある温度と湿度において水蒸気が飽和するために必要な水蒸気量であり、飽差が小さいと気孔が開き、蒸発散が行われる。そのため、生育ステージ 6 において蒸発散が活発に行われ、果実の成長が促進されたため、生育ステージ 6 の各時間帯における飽差の最小値が強い正の相関を持っていると考える。また、表 3 において飽差の最小値だけでなく、生育ステージ 6 が強い相関を持っているため、生育ステージ区分を変化させることで、気象環境と収穫時品質や収量との関係性に対し、より詳細な分析が期待できる。

表 2 各特徴量と収穫時糖度相関分析結果

| 生育ステージ | 時間帯 | 特徴量 | 相関係数 | P 値 |
|--------|-----|---------|-------|-------|
| 2 | I | 蒸発散量_最大 | 0.894 | 0.000 |
| 4 | I | DSD_尖度 | 0.774 | 0.003 |
| 3 | V | 温度_歪度 | 0.764 | 0.004 |
| 1 | V | 蒸発散量_最大 | 0.762 | 0.004 |
| 2 | III | 飽差_歪度 | 0.756 | 0.004 |

表 3 各特徴量と収穫時収量相関分析結果

| 生育ステージ | 時間帯 | 特徴量 | 相関係数 | P 値 |
|--------|-----|---------|--------|-------|
| 6 | IV | 飽差_最小 | 0.821 | 0.001 |
| 6 | I | 飽差_最小 | 0.819 | 0.001 |
| 2 | I | 蒸発散量_最大 | -0.818 | 0.001 |
| 6 | II | 飽差_最小 | 0.818 | 0.001 |
| 6 | V | 飽差_最小 | 0.808 | 0.001 |
| 6 | III | 飽差_最小 | 0.803 | 0.002 |
| 2 | III | 温度_最大 | 0.768 | 0.004 |
| 6 | V | 飽差_最大 | 0.762 | 0.004 |
| 4 | V | 蒸発散量_最大 | -0.748 | 0.005 |
| 6 | V | 飽差_合計 | 0.738 | 0.006 |

5. おわりに

本研究では、植物の生育ステージや時間帯で植物に与える影響の異なる気象環境と収穫時品質や収量との関係性を分析する手法を提案した。提案手法によって、時間帯や生育ステージを考慮した収穫時品質や収量に対する経時特徴量分析を実現した。今後、提案手法を用いて作成した時間帯や生育ステージなど、時間情報を重畳した特徴量を用いた収穫時品質推定モデル作成を検討する。また、植物生理学の知見も参考に農業 AI の実現性を模索していく。

謝辞

本研究の一部は JST さきがけの支援を受け実施された。

参考文献

- [1] 横塚弘毅: 山梨県における積算温度に基づいたブドウ糖度の予測, J.ASEV Jpn., Vol.17, No.1. 7-13(2006).
- [2] 山下勇輝, ほか: トマト栽培実験による環境データ収集と環境要因の評価, 情報処理学会第 77 回全国大会 (2015).
- [3] Robert Tibshirani: Regression Shrinkage and Selection via the Lasso, J.R.Statist.Soc.B, 58, No.1, pp.267-288(1996).
- [4] Ibayashi, H., et al.: A Reliable Wireless Control System for Tomato Hydroponics, MDPI Sensors, Vol.16(5), 664 (2016).
- [5] 若森和昌, ほか: 草姿画像を用いた植物萎れ具合高精度推定, 情報処理学会第 79 回全国大会(2017).
- [6] 大石直記: トマトの養液栽培における水分ストレスに応じた給液制御システムの開発(1)ー茎径変化による水分ストレスの非破壊評価ー, 生物環境調節, Vol.40(1), pp.81-89 (2002).