

# データマイニングを用いた 水稻栽培スケジュール予測について

菅井 友駿<sup>1</sup> 神部 公佑<sup>3</sup> 西内 俊策<sup>2</sup> 北野 英己<sup>2</sup> 北 栄輔<sup>1</sup>

概要：日本の農業において新規就農者の増加が急務であり、そのために作業の効率化および収入の増加を図ることが重要である。そこで本研究では、水稻栽培において作業効率化と収入最大化に繋がる出穂日と刈取日を予測することを目的とする。本手法において出穂日予測は田植えから出穂までの名古屋市の気象データを用い、田植え日から出穂日までの日数を重回帰分析によりモデルを作成し予測した。刈取日予測は出穂日から毎日の平均気温を積算し、地域、品種ごとの刈取りの基準となる積算温度に達した日を刈取日とし予測を行った。予測の結果、熊本県と新潟県のコシヒカリについて出穂日、刈取日の誤差は共に約2日以内であった。

キーワード：水稻栽培，出穂日予測，刈取日予測，重回帰分析

## Rice Cultivation Schedule Prediction Using Data Mining

TOMOTAKA SUGAI<sup>1</sup> KOUSUKE KANBE<sup>3</sup> SYUNSAKU NISHIUCHI<sup>2</sup> HIDEMI KITANO<sup>2</sup> EISUKE KITA<sup>1</sup>

### 1. はじめに

日本における農業従事者は高齢化が深刻化している。それに伴い経験豊富な農業従事者の離農が増加し、新規就農者の増加が急務である。そのためには経験豊富な従事者の知識やノウハウをデータマイニングや機械学習で見える化し、未熟者が効率的に営農できるよう支援する必要がある。作業の効率化及び収入の増加を図ることが重要である。作業の効率化のためには作業工程をあらかじめ予測し、最適なスケジュールを決定しなければならない。また、収入増加のためには収量、品質が落ちてしまわないように刈取適期を予測する必要がある。そこで、本研究では水稻の栽培スケジュールを決定するにあたり必要な出穂日予測と水稻の最適な刈取日を決定するための刈取日予測について述

べる。

### 2. 水稻の栽培ステージ

水稻は一般的に発芽から成熟までに3~6ヶ月を必要とする[1]。この期間に発芽から幼穂分化までの栄養生長期と幼穂分化から出穂までの生殖生長期の二つの生長段階を完了し、生殖生長期はさらに出穂前と、出穂後の登熟期に分類される。そして、登熟期を経て出穂した日からの平均気温の積算値が基準を超えた日に刈取を行う。これは地域、品種で異なるが1000°C前後が基準となっている。

### 3. 提案アルゴリズム

#### 3.1 出穂日予測アルゴリズム

水稻の出穂日予測における目的変数は、田植日から出穂日までにかかる日数の出穂日数である。説明変数は、田植日を0日として出穂日までの毎日の気象データである。気象データは、日平均、日最高、日最低の各気温、降水量、日照時間であり、これらに加えて日最高気温から日最低気温を引いた気温日較差の6種類を用いる。これらのデータ

<sup>1</sup> 名古屋大学大学院情報科学研究科  
Graduate School of Information Science, Nagoya University

<sup>2</sup> 名古屋大学大学院生命農学研究科  
Graduate School of Bioagricultural Sciences, Nagoya University

<sup>3</sup> 名古屋大学情報文化学部  
School of Informatics and Sciences, Nagoya University

を用いて以下の順に出穂日予測モデルを決定する。

- 1 全ての説明変数のうち最適な変数のみを採用するため相関分析を行う。また、同時に無相関検定を行う。
- 2 説明変数および目的変数を正規化し、説明変数間の多重共線性を排除するために主成分分析を行う。
- 3 採用する主成分のスコアと出穂日数で重回帰分析 [2] を行い以下の予測式を生成する。なお、 $y$  が目的変数、 $x_i$  が説明変数である。

$$y = f + \varepsilon = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + \varepsilon$$

### 3.2 刈取日予測アルゴリズム

次に刈取日予測について以下に記す。

まず出穂日  $D_0$  を与える。ここで、過去の  $N$  年の出穂日  $D_0$  から 10 月末までの平均気温データから各年の気温データ  $n$  における出穂日  $D_0$  からの日数  $d$  の気温を  $t_n^d$  とする。また収穫に適した積算温度を  $T_h$  とする。 $N$  個すべてに対し、 $t_n^d$  を  $d = 1$  からの総和を計算し、はじめて  $\sum_{d=1}^{D_n} t_n^d > T_h$  となるときの  $D_n$  を求め、得られた  $D_n$  の最小値～最大値が刈取日の予測範囲となる。

## 4. 解析方法

出穂日予測はアルゴリズムに則り予測モデルを構築する。採用する説明変数は 0.3 以上、0.4 以上、0.5 以上、0.6 以上とし、各相関係数で生成される予測モデルで直近 5 年間の平均誤差を比較し、最良のモデルを提案モデルとする。刈取日予測は過去 5 年以前のデータから  $D_n$  の最小値～最大値を求め、直近の 5 年間で予測を行い精度評価を行う。予測対象は愛知県コシヒカリ、愛知県あいちのかおり、岐阜県飛騨コシヒカリ、新潟県コシヒカリ、熊本県コシヒカリである。また、それぞれの地域・品種で記録されている期間が異なる。したがって、愛知県コシヒカリ 1985～2014 年、愛知県あいちのかおり 1987～2014 年、岐阜県飛騨コシヒカリ 1988～2013 年、新潟県コシヒカリ 1980～2013 年、熊本県コシヒカリ 1981～2013 年というようにデータを整形し使用する。

## 5. 予測結果

### 5.1 出穂日予測

予測対象の 5 種類について出穂日予測の結果として、最良の予測モデルの平均絶対誤差日数とそのときの相関係数を表 1 に示す。表 1 より、愛知県では約 3 日、その他の県では約 2 日以内の誤差で予測ができています。

### 5.2 刈取日予測

刈り取りに適した積算温度は愛知県コシヒカリ 1000°C、岐阜県飛騨地方コシヒカリ 950°C、熊本県コシヒカリ 850°C、新潟県コシヒカリ 1000°C、愛知県あいちのかおり 1100°C である。これらから使用データの直近 5 年において予測を

表 1 出穂日予測結果のまとめ

地域名	相関係数	平均絶対誤差日数 (日)
愛知県コシヒカリ	0.6	3.04
愛知県あいちのかおり	0.4	2.74
岐阜県飛騨コシヒカリ	0.5	2.45
新潟県コシヒカリ	0.4	1.29
熊本県コシヒカリ	0.5	0.95

行う。予測対象の 5 種類について  $D_n$  の最小値～最大値、平均絶対誤差日数を表 2 に示す。表 2 より愛知県、熊本県、新潟県のコシヒカリについては約 2 日以内の誤差日数であり精度良く予測ができていますといえる。しかし、愛知県あいちのかおり、岐阜飛騨のコシヒカリについては誤差が大きい。岐阜飛騨について、品質と積算温度、収量と積算温度のそれぞれの相関係数は、0.11、0.32、0.02 であり相関なしと判断され、積算温度が変化しても品質、収量に影響が小さいと考えられる。目視で初めの黄化率 85～90%あるいは、青化率 10～15%の時に収穫適期とされているため黄化率、青化率を確認し収穫日を決定することも必要と考えられる。または、積算温度が湿度や高温期、低温期などを加味して変動するモデルにより精度を高めることができる可能性がある。

表 2 刈取日予測結果のまとめ

地域名	最小値～最大値	平均絶対誤差日数 (日)
愛知県コシヒカリ	34～41	1.5
愛知県あいちのかおり	41～53	3.8
岐阜県飛騨コシヒカリ	38～52	4.2
新潟県コシヒカリ	35～48	1.2
熊本県コシヒカリ	41～53	2.2

## 6. 結論

本研究では、統計的手法を用いて出穂日を予測することと、平均気温の積算値を用いて刈取日予測を行った。その結果、出穂日予測に関しては熊本県と新潟県のコシヒカリ、刈取日予測に関しては愛知県、熊本県、新潟県のコシヒカリについて 2 日以内の誤差であった。

謝辞 本研究の一部は、農林水産省革新的技術事業創造促進事業（異分野融合研究）の助成を受けたものである。また、愛知県農業総合試験場からデータの提供を受けた。

### 参考文献

- [1] 村山 登 (1986) 『稲作科学の基礎』 株式会社博友社
- [2] 多変量解析へのステップ  
長畑秀和 共立出版株式会社 pp29-127