

水稲の品質予測モデルの研究について

柘植 綾香¹ 吉田 詠梨¹ 西内 俊策² 北野 英己² 北 栄輔¹

概要: 本研究では、日本の農業従事者にとって重要な農作物である水稲栽培において、重回帰分析 (MRA) やニューラルネットワーク (NN) 分析を用いて、水稲の品質予測をする方法について述べる。評価期間は 2010 年から 2014 年、学習期間は 1985 年から各評価年の前年までとし、予測結果を比較した。2010 年と 2013 年の予測値は実測値に近い値を取っているが、それ以外の年度については改善の余地がある結果となった。

キーワード: 水稲, 品質予測, 重回帰分析 (MRA), ニューラルネットワーク (NN)

Quality Prediction Model of Paddy Rice

AYAKA TSUGE¹ ERI YOSHIDA¹ SHUNSAKU NISHIUCHI² HIDEMI KITANO² EISUKE KITA¹

1. はじめに

我が国の農業は単位面積当たりの生産性が高い一方で、新規就労者数は減少の一途をたどっている。今後、新規就労者が安心して農業へ従事するためには、安定して収入を得ることが必要である。このためには、農作物による収入の安定化が重要である。本研究では、日本の農業従事者にとって重要な農作物である水稲の栽培において、水稲の品質予測をする方法について述べる。

品質予測モデルは線形関数やニューラルネットワーク (NN) モデルを用いた重回帰分析 (MRA) によって定義する。このとき、過去の品質データを目的変数、栽培地域の過去の気象データ (平均・最高・最低気温, 日照時間, 降水量) を説明変数として採用する。定義した品質予測モデルに評価年の気象データを入力して予測値を算出し、実測値との比較によって予測精度を検証する。

2. 水稲の生育

水稲は一般に、発芽してから成熟するまでに 3~6 ヶ月を必要とする [1]。この期間に水稲は、栄養生長 (vegetative growth) と生殖生長 (reproductive growth) の二つの成長段階 (growth stage) を経由する。生殖生長期はさらに、出穂前と出穂後の 2 期間に区分され、出穂後の期間は登熟期間 (ripening period) と呼ばれる。ここで、出穂とは水稲の開花のことであり、出穂日は水稲が開花した日のことを示す。作物の収量の潜在的な大きさは出穂前期に決定され、最終収量は主として出穂後の期間に決定される。そこで、説明変数の候補として出穂日を基点とした出穂日前 40 日から出穂日後 20 日の合計 61 日間の日毎の気象情報を採用する。

3. 予測モデル

3.1 重回帰分析

説明変数 x_i , 説明変数の総数を N とすると、説明変数ベクトルは $\vec{x} = \{x_1, x_2, \dots, x_N\}$ と定義される。重回帰分析では、目的変数 Y を説明変数 \vec{x} の関数として定義する。

$$Y = f(\vec{x}) \quad (1)$$

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

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

3.2 線形関数を用いる場合

線形関数で近似する場合、次式となる。

$$Y = a_0 + \sum_{i=1}^N a_i x_i \quad (2)$$

説明変数 x_i と目的変数 Y の相関分析を行い、回帰分析に採用する説明変数を選択する。そして、最小自乗近似によってモデルパラメータを定める。

3.3 ニューラルネットワークモデルを用いる場合

ニューラルネットワーク (NN) は脳の神経回路の仕組みを模したモデルで、非線形回帰分析、非線形判別分析の有力な機械学習法である [2], [3]。本研究では多層パーセプトロンを用いる。

中間変数を u_j 、その総数を M とする。入力変数と中間変数、中間変数と出力変数の関係は次式で与えられる、

$$u_j = g \left(\sum_{i=1}^N w_{ij} x_i - \theta_j \right) \quad (3)$$

$$Y = g \left(\sum_{j=1}^M v_{kj} x_j - \Theta_k \right) \quad (4)$$

ここで、関数 g は次式で与えられる。

$$g(v) = \frac{1}{1 + e^{-v}} \quad (5)$$

ここで、 $w_{ij}, v_{kj}, \theta_j, \Theta_k$ は、重み変数と閾値である。

階層型ネットワークにおける結合の重みと閾値は、誤差逆伝播学習法を用いて定める。誤差逆伝播学習法は 1986 年に Rumelhart らが提案した学習法であり、ある学習データが与えられたとき、ネットワークの出力が学習データと一致するように各層間の結合係数を修正する [4], [5]。誤差逆伝播法では、入力パターンが与えられたとき、その出力と教師信号との誤差を減らすように、二乗誤差和を評価関数として誤差を逆伝播させながら結合荷重を修正することにより学習を進める。

4. 予測実験

水稻の品質指標は、その数値が小さい程品質が高く、5 以下ならば最上位の一等米に分類される。また気象データについては、最高・最低・平均気温、気温日較差、降水量、日照時間を用い、出穂日の 40 日前から 20 日後までの期間について品質と相関分析 [6] を行い、算出した相関係数が一定以上であった気象データを説明変数として採用する。本研究ではコシヒカリの実験についてのみ示す。

本研究では、説明変数に気象データの実測値を用い、線形重回帰モデルとニューラルネットワークモデルで学習を行う。評価期間は 2010 年から 2014 年、学習期間は 1985 年から各評価年の前年までとする。

予測結果と品質データの実測値の誤差が最小となったの

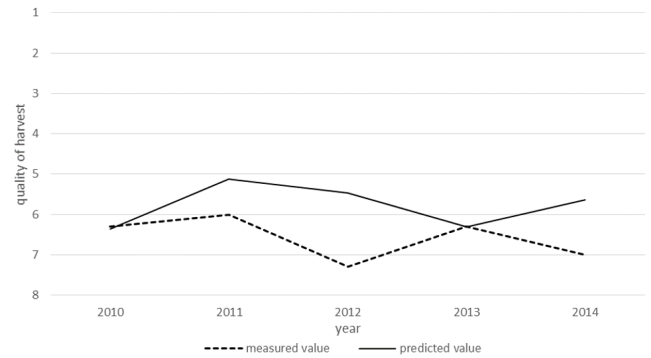


図 1 コシヒカリの品質予測

Fig. 1 Quality Prediction of Koshihikari

は、相関係数に閾値を設けず、有意である項目をすべて説明変数に選び、MRA を用いて予測を行った場合であった。その結果を図 1 に示す。図において、横軸は評価年、縦軸は品質の値を表している。また、破線は実測値、実線は予測値を示す。これから分かるように、2010 年と 2013 年の予測値は特に実測値と良く一致しており、その絶対誤差は 2010 年が 0.05、2013 年が 0.003 となっている。その他の年の絶対誤差は、2011 年が 0.873、2012 年が 1.84、2014 年が 1.37 となっている。

5. 結論

本研究では、水稻の品質予測モデルについての研究結果を示した。目的変数に水稻の品質指数を、説明変数に日毎の気象データを用いた。MRA と NN を比較したところ、MRA のほうが良い精度を示した。予測精度は年ごとにばらつきがみられたので、今後はどの年でも精度良い予測を示すように改善するとともに、他の地域における栽培データでの精度検定を行う。

6. 謝辞

本研究の一部には、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。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, 吉田よし子, 長谷川周一, 末永一博: 稲作科学の基礎, 博友社 (1986).
- [2] 石川博, 新見礼彦, 白石陽, 横山昌平: データマイニングと集合知, 共立出版 (2012).
- [3] 後藤正幸, 小林学: 入門パターン認識と機械学習, 共立出版 (1994).
- [4] D.E.Rumelhart, G.E.Hinton, R.J.Williams. Learning representations by back-propagating errors, Nature, 323, pp.533-536 (1986).
- [5] D.E.Rumelhart, J.L.McClelland, the Reserch Group. Parallel distributed processing, MIT Press (1986).
- [6] 長畑秀和: 多変量解析へのステップ, 共立出版 (2001).