

酒造原料米の消化性Brixの機械学習による予測に関する研究

齋藤健太* 樋口誠一* 和田健太郎* 横堀正敏*

Study on Machine Learning Prediction of Enzyme Digestibility of Rice for Sake Making

SAITO Kenta*, HIGUCHI Seiichi*, WADA Kentaro*, YOKOBORI Masatoshi*

抄録

日本酒の品質に影響を及ぼす酒造原料米の溶解性の指標となる消化性Brixを、迅速に予測するモデルの構築を目指した。精米歩合35-80%かつ水分8-15%と対象範囲を拡大することで、仕込みに使用する原料米の溶解性を把握し、最適な原料処理が可能となると期待される。消化性Brix予測値を得るため、白米水分、吸水率、メッシュ農業気象データ及び近赤外分析で得た迅速分析値を説明変数とした機械学習による回帰分析を行い、予測精度を比較した。

キーワード：米，日本酒，溶解性，消化性，機械学習

1 はじめに

酒造原料米（以下、原料米という）の溶解性は、日本酒の醸造工程や品質に影響する。原料米の溶解性は、生育時の気象条件¹⁾、品種、精米歩合²⁾及び水分²⁾によって変化する。特に近年は異常気象により、地域や生産者ごとに溶解性が異なる。この溶解性の指標は酒造用原料米全国統一分析法²⁾による消化性Brixで表される。

白米の浸漬或いは蒸米の晒しなど、原料処理を調整することで、もろみでの蒸米の溶け方を制御できる²⁾が、その条件を決めるには仕込み前に溶解性を把握する必要がある。しかし、原料米の溶解性分析には時間と手間がかかる²⁾ため、多くの酒造現場では行われていない。

酒米研究会では、全国の原料米を調査しており、品種特性と毎年の傾向を見るため、精米歩合70%かつ水分13.5%と限定された条件の原料米の消化性Brixを酒造期後に提供している。仕込みに使用する原料米がそれ以外の精米歩合や水分である場

合は直接評価できないため、溶解性は把握されないまま仕込まれ、造り手の経験則で管理されており、十分な対応ができていない。

既報³⁾では部分的最小二乗（以下、PLSという）回帰分析による消化性Brixの予測を行い、迅速に消化性Brixを把握できるようになったが、精米歩合70%かつ水分13.5%と限定された条件であった。そこで本研究では、精米歩合35-80%かつ水分8-15%と対象範囲を拡大した消化性Brix予測モデルの構築を目的とした。そのためには迅速分析が可能な白米水分、10分吸水率、20分吸水率、近赤外分析（以下、NIRという）及びメッシュ農業気象データ⁴⁾を説明変数とし、人工ニューラルネットワーク（以下、ANNという）、ランダムフォレスト（以下、RFという）及び勾配ブースティング（以下、GBという）といった機械学習による予測モデルを構築することで、最適な説明変数及び解析手法を明らかにすることを目指した。

* 食品・バイオ技術担当

表1 試料の品種		表2 試料の精米歩合		表3 試料の白米水分	
品種	調査点数	精米歩合	調査点数	白米水分	調査点数
さけ武蔵	83	30-39%	8	7.0-9.9%	16
山田錦	35	40-49%	8	10.0-11.9%	29
五百万石	8	50-59%	29	12.0-13.9%	125
その他の酒米	31	60-69%	10	14.0-15.9%	0
一般米	14	70-79%	113	16.0%以上	1
全合計	171	80%以上	3	全合計	171
		全合計	171		

2 実験方法

2.1 試料

品種、生産地、生産年、精米歩合、白米水分の異なる原料米 171 点を用いた(表 1～表 3 参照)。生産年は 2017～2023 年産であった。

2.2 原料米分析

20 分吸水率及び消化性 Brix の測定は酒造用原料米全国統一分析法²⁾に準じた。10 分吸水率は、20 分吸水率と同様の方法で、時間を 10 分間に変えた。白米水分は酒米・麴水分計(ライスタ f7、ケツト科学研究所)の酒米精米モードで測定した。

2.3 NIR

近赤外スペクトルの測定には、フーリエ変換赤外分光分析装置(Frontier MIR/NIR plus NIRA、PerkinElmer Japan)を用いた。近赤外反射アクセサリ(NIRA)の窓板を白米(米粒のまま)で覆い、セラミックディスクをリファレンスとして反射モードで測定した。637～2700 nm の範囲を 0.3～0.4 nm 間隔で、積算回数は 32 回とした。3 回繰り返して測定し、反射率の平均値を求めた。

データ処理は、最小二乗法を基にした平滑化と数値微分を行い、915, 1200, 1355, 1460, 1580, 1708, 1776, 1825, 1936, 2058, 2106, 2174, 2282, 2320, 2430 nm の各波長スペクトルにおける 2 次微分ピーク面積値を求めた。

2.4 メッシュ農業気象データ

国立研究開発法人農業・食品産業技術総合研究機構のメッシュ農業気象データシステム⁴⁾の平均気温より、出穂後 20 日間の平均気温を求めた。圃場位置が特定できない米については、県単位の平均気温とした。

2.5 回帰分析

消化性 Brix への単回帰分析での相関係数を比較した。

機械学習では R (version 4.3.0)⁵⁾のパッケージ tidymodels を用いた。説明変数には白米水分、10 分吸水率、20 分吸水率、出穂後 20 日間の平均気温、915, 1270, 1355, 1460, 1580, 1708, 1776, 1825, 1936, 2058, 2106, 2174, 2282, 2320, 2430 nm の各波長スペクトルにおける 2 次微分ピーク面積値及び 1930-1933 nm の反射率を用いた。動作エンジンには、ANN で nnet パッケージ、RF で ranger パッケージ、GB で XGboost を指定した。試料はランダムに区分し、全体の 80%に当たる 135 点を学習データ、残りの 20%に当たる 36 点を検証データとした。10 分割交差検証法を用いた。目的変数となる消化性 Brix には対数変換を行い、全ての変数を平均 0、分散 1 となるように標準化した。最適なハイパーパラメータの選択には、二乗平均平方根誤差(以下、RMSE)を指標とした。

2.6 回帰分析の予測精度

予測精度は、許容誤差 0.3 以内に含まれる試料の割合（以下、正解率という）、平均絶対誤差（以下、MAE という）、RMSE、決定係数（以下、 R^2 という）で評価した。酒造用原料米全国統一分析法²⁾に定められるばらつき許容範囲は 5% である。既報³⁾より消化性 Brix 予測値 0.5 の差があると、カブロン酸エチル濃度に差がみられたことから、許容誤差を 0.3 とした。

3 結果及び考察

3.1 各種分析値

一例として、さけ武蔵（2017 年産、埼玉県川越市鴨田）の NIR の反射率及び 2 次微分値を図 1 に示した。各種分析値の平均値、最大値、最小値の一部を表 4 に示す。

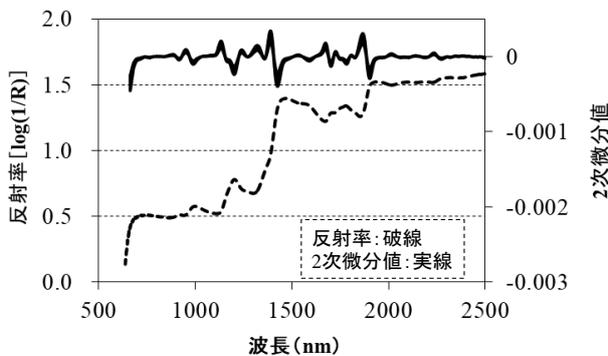


図 1 NIR の反射率及び 2 次微分値

表 4 各種分析値の平均値、最大値、最小値

変数	平均値±SD	最大値	最小値
消化性Brix (%)	10.3±0.8	12.1	7.8
精米歩合 (%)	63.7±10.8	80.0	35.0
白米水分 (%)	12.1±1.3	17.6	7.2
10分吸水率 (%)	20.6±4.5	34.2	7.4
20分吸水率 (%)	30.7±4.6	49.8	10.6
出穂後 20日間の 平均気温 (°C)	25.9±1.8	29.6	20.5

n=171

3.2 消化性 Brix への影響

消化性 Brix への各説明変数の単回帰分析による相関係数及び R^2 の一部を表 5 に示す。10 分吸水率の R^2 が 0.22 で最も高かった。

3.3 回帰分析の予測精度

3 つの機械学習による回帰分析の予測精度を表 6 に示す。正解率は、ANN で 19.4%、RF で 30.6%、GB で 41.7% であった。MAE と RMSE は、ANN が最も高く、GB で最も低かった。 R^2 は ANN が最も低く、RF 及び GB は 0.68 と 0.66 であった。表 5 の単回帰分析による R^2 よりも表 6 の機械学習による R^2 の方が 1 に近かったことから、機械学習による消化性 Brix の予測が有用な手法であると考えられる。3 つの機械学習による消化性 Brix 予測値及び実測値を図 2~図 4 に示す。ANN 及び RF の予測値が 9-11 付近に集中しているのに対し、GB による予測値は 8-11.5 と幅広く予測していた。以上のことから、現状では予測精度は不十分であるが、広範囲の予測が可能な GB による予測が適していると考えられる。

表 5 各説明変数の相関係数及び R^2

変数	白米水分	10分吸水率	20分吸水率	出穂後 20日間の 平均気温
相関係数	-0.45	0.47	0.37	-0.19
R^2	0.20	0.22	0.14	0.04

n=171

表 6 機械学習による回帰分析の予測精度

回帰分析	ANN	RF	GB
正解率	19.4	30.6	41.7
MAE	0.57	0.47	0.41
RMSE	0.63	0.57	0.52
R^2	0.49	0.68	0.66

n=36

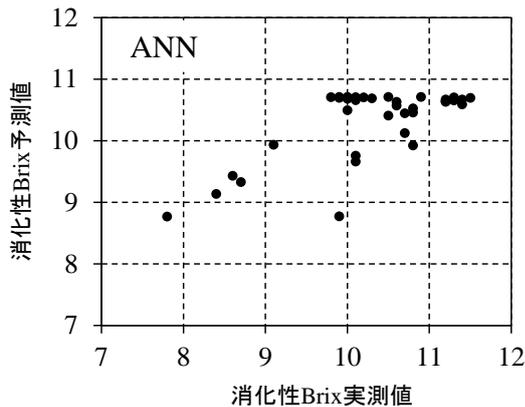


図2 ANNによる消化性Brix予測値及び実測値

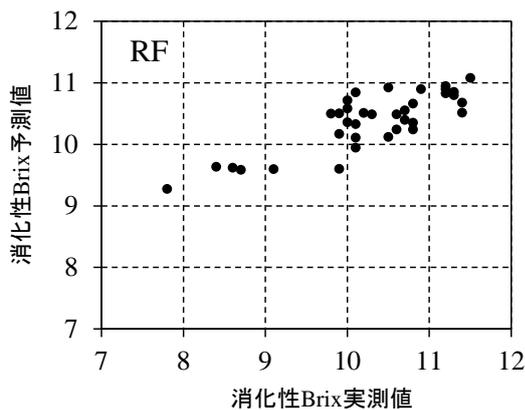


図3 RFによる消化性Brix予測値及び実測値

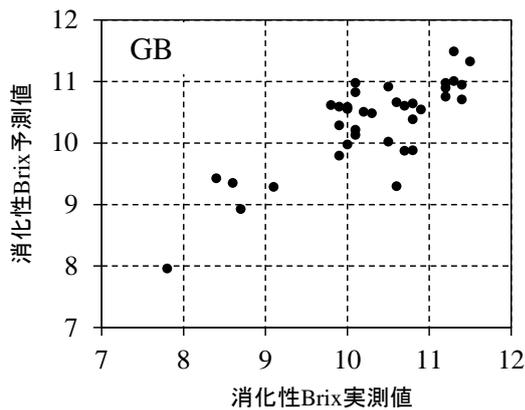


図4 GBによる消化性Brix予測値及び実測値

4 まとめ

迅速分析が可能な白米水分、10分吸水率、20分吸水率、NIR及び平均気温を説明変数として、機械学習による消化性Brix予測モデルを構築した。原料米の対象範囲を精米歩合35-80%かつ水分8-15%と拡大した。これらの原料米に対してANN、RF及びGBにより消化性Brixを予測した。現状ではGBの正解率41.7%と最も高かった。

今後は、実際に現場で使用される白米水分7-11.9%の試料を増やし、メッシュ農業気象データから得られる最高気温、最低気温、降水量、日射量、日照時間を説明変数に加えることで、予測精度の向上を目指す。正解率80%以上、MAE0.1以下、RMSE0.1以下、 R^2 0.8以上を目標とする。さらに重回帰分析、PLS回帰分析及びサポートベクターマシンとの比較を行う予定である。

謝辞

メッシュ農業気象データシステムを御提供いただきました国立研究開発法人農業・食品産業技術総合研究機構、原料米を御提供いただきました酒造会社及び農家の皆様に感謝の意を表します。

参考文献

- 1) 奥田将生：米デンプン構造と醸造特性・気象条件との関係，生物工学会誌，**90**，5(2012)227
- 2) 酒米研究会編：酒造用原料米全国統一分析法，酒米研究会，(2023)
- 3) 齋藤健太，横堀正敏，樋口誠一，和田健太郎：酒造原料米の溶解性予測システムの確立に関する研究—酒造原料米の消化性Brix予測値の活用—，埼玉県産業技術総合センター研究報告，**18**，(2020)35
- 4) 農研機構メッシュ農業気象データシステム，https://amu.rd.naro.go.jp/wiki_open/doku.php?id=start2，2025.1.22
- 5) The R Project for Statistical Computing，<https://www.r-project.org>，2025.1.22