

Modeling the relationship of molasses-forming mineral elements on autumn sown sugar beet (*Beta vulgaris*) using Machine Learning method in Iran

Sepinood Seyyed Alizadeh Ganji¹, Nargess Mooraki^{1*}, Masoud Honarvar¹, Parvin Azhdari²

(Received: 12 Dec. 2024 ; Accepted: 30 April. 2025)

How to cite this article:

Seyyed Alizadeh Ganji S, Mooraki M, Honarvar M, Azhdari P. Modeling the relationship of molasses-forming mineral elements on autumn sown sugar beet (*Beta vulgaris*) using Machine Learning method in Iran. 2024; (40)2: 225-236. (In Persian with English abstract). DOI: <https://doi.org/10.22092/jsb.2025.367972.1379>.

Extended Abstract

Introduction

Sugar beet (*Beta vulgaris* L.) is a vital agricultural crop, contributing approximately 30% of the world's sugar supply and ranking as the second most important source of sugar globally. This crop is predominantly grown in the autumn in suitable regions, offering advantages such as enhanced water-use efficiency, increased sugar yield, and greater profitability for growers, especially in semi-arid areas. The processing of sugar beets produces molasses, a by-product rich in sugars and minerals, where elements such as sodium, potassium, and alpha-amino nitrogen influence both sugar content and quality. Elevated concentrations of these minerals can reduce sucrose recovery by increasing sugar losses to molasses. To estimate sugar content, the Renfield formula correlates mineral concentration with molasses sugar concentration; however, it was originally designed for spring-sown sugar beets under European conditions. Applying this formula directly in Iran is challenging due to regional differences, including higher impurity levels and distinct climatic factors. This study aims to modify and calibrate the Renfield formula for autumn-sown sugar beets in Iran, focusing on samples from Khuzestan Province, by utilizing Machine Learning techniques. This approach seeks to improve the accuracy of sucrose estimation, refine beet quality assessment, and enhance overall sugar production efficiency in accordance with local soil and crop characteristics.

Materials and Methods

This study investigates the chemical and mineral composition of harvested samples of autumn-sown sugar beet from Khuzestan Province, collected over eight weeks in 2022. Samples were prepared by washing, weighing, and homogenizing root tissues, which were subsequently frozen and subjected to laboratory analyses. Key measurements included total sugar and purity levels determined via polarimetry, as well as concentrations of sodium, potassium, and alpha-amino nitrogen measured by spectrophotometry and flame photometry. Molasses sugar content was calculated using the Rinfeld formula, while alkalinity was determined through a specified relationship. A total of 382 samples were analyzed using statistical and machine learning techniques. Data normalization and normality were performed followed by one-way ANOVA. Machine Learning methods—including principal component analysis (PCA) and K-Means clustering—were employed to estimate molasses sugar content based on three independent variables (Na, K, alpha-amino nitrogen). Model validation involved comparing predicted values with actual factory data using t-tests. The results aim to enhance regional estimation and understanding of molasses sugar content based on mineral profiles, thereby facilitating improved quality assessment and process optimization in the Iranian sugar beet industry.

Results

This study examined variations in key mineral elements—sodium (Na), potassium (K), alpha-amino nitrogen (N), as well as molasses content in sugar beet samples collected over eight weeks of operation. The objective was to evaluate significant changes in these

1. Food Science and Technology Department, Faculty of Biological Science, North Tehran Branch, Islamic Azad University, Tehran, Iran.

*Corresponding author: Nargess_mooraki@yahoo.com

2. Statistics Department, Faculty of Basic Sciences, North Tehran Branch, Islamic Azad University, Tehran, Iran.

parameters during the harvest period. Analysis of variance (ANOVA) revealed statistically significant differences in Na, K, N, and molasses content across the eight-week timeframe ($p < 0.05$). The corresponding F-values indicated substantial temporal fluctuations underscoring the dynamic nature of mineral composition during harvest. These findings provide valuable insights for quality assessment and process optimization in sugar beet cultivation and processing. To refine molasses sugar estimation under regional conditions, this study utilized unsupervised Machine Learning techniques, including Principal Component Analysis (PCA) and clustering, to adapt the Renfield equation for Iran. Based on 80% of the dataset, regression models were trained using both scaled and unscaled data. The best performing model emerged from clustering analysis, yielding statistically significant regression coefficients. The resulting regression equation to estimate molasses sugar content is as follows:

$$M = 0.194 + 0.3267 \times Na + 0.3135 \times K + 0.3036 \times NM = 0.194 + 0.3267 \times Na + 0.3135 \times K + 0.3036 \times N$$

Model performance was evaluated using mean squared error, mean absolute error, and the coefficient of determination ($R^2 = 0.93$). Results indicated that linear regression outperformed more machine learning methods in predicting molasses sugar content. This suggests that the proposed model effectively captures the relationship between mineral content and molasses sugar percentage, offering a practical tool for improving processing outcomes in Iran's sugar industry. Furthermore, a statistical comparison was conducted between the molasses sugar content estimated by the Renfield equation, the newly proposed model, and the actual measured values. Using sodium, potassium, and alpha-amino nitrogen values from the samples, molasses sugar content was calculated using both equations and compared with laboratory measurements. One-sample t-tests at a significance level of $P < 0.05$ revealed a significant difference between the values predicted by the Renfield equation and the actual measurements. In contrast, the proposed model showed no significant difference from the actual measurements. Additionally, the ranges of Na, K, and N concentrations in the dataset—as the foundation for the proposed equation—are provided, further validating its applicability. These results confirmed that the proposed model reliably estimates molasses sugar content based on mineral composition and substantially outperforms the traditional Renfield equation under Iranian conditions. Recent studies have also questioned the applicability of the Renfield-based model in Iran, emphasizing the influence of climatic conditions, agricultural practices, and harvest timing on sugar beet quality. The finding of

this study support those claims through empirical comparison with actual molasses production data. Machine Learning techniques were employed to analyze the relationship between independent variables—Na, K, and N—and the dependent variable, molasses sugar percentage. Although several models were tested, regression analysis not only achieved the lowest error metrics but also provided a transparent and interpretable equation, demonstrating its superiority in modeling molasses sugar content when the number of variables is limited. While other models may offer predictive capacity, regression uniquely clarifies the direct influence of each mineral element, making it especially valuable for practical implementation.

Conclusion

The final results of this study demonstrate that the independent variables—sodium, potassium, and alpha-amino nitrogen—have a significant effect on the estimation of molasses sugar content in sugar beet. The successful adaptation of the Renfield equation for Iranian conditions allowed for the development of a localized model capable of predicting molasses sugar content based on these mineral concentrations. This modified equation offers a practical and reliable tool for both researchers and growers, supporting the optimization of cultivation practices and enhancing productivity, quality assessments and decision-making in sugar beet production.

Keywords: Beet, Modeling, Molasses sugar content, nitrogen, Potassium, Python, Sodium.

References

- Amiri Z, Asgharipour MR, Moghadam EH, Kakoly E, Campbell DE. Investigating the need to replace the conventional method of sugar beet production in Lorestan province, Iran based on the arguments obtained from emergy calculations. *Ecological Modelling*. 2022; 472(4): 110091. Doi: <https://doi.org/10.1016/j.ecolmodel.2022.110091>
- Hoffmann CM, Koch HJ, Märlander B. Sugar beet. Pp. 634-672. In *Crop Physiology Case Histories for Major Crops*, Academic Press, Cambridge, MA, USA, 2021;. Doi: <https://doi.org/10.1016/B978-0-12-819194-1.00020-7>
- Rezaei J, Fasahat P. Autumn-sown sugar beet cultivation in semiarid regions. pp. 275-290. In: *Sugar Beet Cultivation, Management and Processing*, Singapore: Springer Nature Singapore. 2022. Doi: https://doi.org/10.1007/978-981-19-2730-0_14

مدل‌بندی ارتباط عناصر معدنی ملاس‌زا روی چغندر قند (*Beta vulgaris*) پاییزه به روش یادگیری ماشین در ایران[†]

سپینود سید علی‌زاده گنجی^۱، نرگس مورکی^۲، مسعود هنرور^۳، پروین اژدری^۳

تاریخ دریافت: ۱۴۰۳/۹/۲۲ تاریخ پذیرش: ۱۴۰۴/۲/۱۰

نوع مقاله: پژوهشی

DOI: 10.22092/jsb.2025.367972.1379

س. سید علی‌زاده گنجی، ن. مورکی، م. هنرور و پ. اژدری. ۱۴۰۳. مدل‌بندی ارتباط عناصر معدنی ملاس‌زا روی چغندر قند (*Beta vulgaris*) پاییزه به روش یادگیری ماشین در ایران. چغندر قند ۴۰(۲): ۲۲۵-۲۳۶

چکیده

تحقیق حاضر به بومی‌سازی معادله رینفلد برای چغندر قند پاییزه در استان خوزستان و تخمین رابطه متغیرهای سدیم، پتاسیم و نیتروژن آمینه با درصد قند ملاس با روش یادگیری ماشین پرداخته است. آماده‌سازی نمونه‌ها به این صورت انجام شد که از محموله‌های ارسالی به کارخانه قند قزوین در سال ۱۴۰۱، تعداد ۳۸۲ نمونه در طول هشت هفته بهره‌برداری (اول اردیبهشت‌ماه تا پایان خردادماه) به صورت تصادفی انتخاب شده و شستشو، توزین و خمیرگیری شدند. این خمیرها در دمای ۲۰- درجه سانتی‌گراد منجمد و برای تعیین صفات کیفی ساکارز و ناخالصی‌های سدیم، پتاسیم و نیتروژن آمینه به مؤسسه تحقیقات اصلاح و تهیه بذر چغندر قند ارسال شدند. در ابتدا معنی‌داری تغییرات داده‌های مربوط به اندازه‌گیری سدیم، پتاسیم و نیتروژن آمینه در سطح معنی‌داری پنج درصد با استفاده از مسیر تجزیه واریانس یک‌طرفه ارزیابی شد و سپس داده‌ها به نرم‌افزار پایتون منتقل و با استفاده از روش‌های یادگیری ماشین، درصد قند ملاس برآورد شد. در نهایت مدل رگرسیون به‌دست‌آمده ذیل به‌عنوان بهترین گزینه برای بومی‌سازی معادله رینفلد بر اساس چغندر قند کشت پاییزه استان خوزستان انتخاب شد: $M = 0.1940 + 0.3267 (Na) + 0.3135 (K) + 0.3036 (N)$

کلمات کلیدی: پایتون، پتاسیم، چغندر قند، درصد قند ملاس، سدیم، مدل‌سازی، نیتروژن.

[†] این مقاله مستخرج از پایان‌نامه دانشجویی دوره کارشناسی ارشد است.

۱. دانشجوی کارشناسی ارشد، گروه علوم و مهندسی صنایع غذایی، دانشکده علوم زیستی، واحد تهران شمال، دانشگاه آزاد اسلامی، تهران، ایران.

۲. دانشیار گروه علوم و مهندسی صنایع غذایی، دانشکده علوم زیستی، واحد تهران شمال، دانشگاه آزاد اسلامی، تهران، ایران.

*- نویسنده مسئول: Nargess_mooraki@yahoo.com

۳. استادیار گروه آمار، واحد تهران شمال، دانشگاه آزاد اسلامی، تهران، ایران.

مقدمه

چغندر قند (*Beta vulgaris L.*) یک گیاه صنعتی و اقتصادی بسیار مهم برای تولید قند و شکر است و تقریباً ۳۰٪ از تأمین قند و شکر جهان را به خود اختصاص می‌دهد (Lv et al. 2019). این گیاه دوساله است (Hoffman et al. 2021). و دومین منبع بزرگ تولید قند و شکر در سطح جهانی است و برای تأمین انرژی غذایی ضروری است (Nefedeva and Belopukhov 2023).

چغندر قند پاییزه در مناطق مستعد در فصل پاییز کاشت و در فصل بهار برداشت می‌شود. کشت چغندر قند در پاییز مزایای زیادی نسبت به کشت در بهار دارد و این روش در مناطق نیمه‌خشک در حال افزایش است، زیرا باعث افزایش کارایی آب و افزایش تولید شکر می‌شود و از نظر اقتصادی برای کشاورزان سود بیشتری دارد. این روش امکان استفاده از بارش‌های طبیعی را بیشتر فراهم می‌کند و نیاز به آبیاری را تا حد زیادی کاهش می‌دهد (Rezaei and Fasahat 2022). لذا زراعت چغندر قند پاییزه نسبت به چغندر قند بهاره پایدارتر است (Amiri et al. 2022). در طی فرآیند تولید قند و شکر از چغندر قند، محصولات جانبی نیز تولید می‌شود که یکی از آن‌ها ملاس نام دارد. این ماده حاوی قندها، ویتامین‌ها و مواد معدنی ضروری است و کاربردهای زیادی در صنایع مختلف از جمله صنایع غذایی دارد (Yanuartono et al. 2017).

عناصر سدیم، پتاسیم و آلفا آمینو نیتروژن، سه عنصر معدنی اصلی در ریشه چغندر قند هستند. این عناصر برای رشد گیاهان مواد مغذی مهمی هستند اما در فرآیند استخراج شکر موجب کاهش قند قابل استحصال و افزایش قند ملاس می‌شوند. هر اندازه سطوح سدیم و پتاسیم در ریشه‌های چغندر قند بالا رود، می‌تواند منجر به افزایش از دست رفتن قند در ملاس شود (El-Geddawy et al. 2023) و منجر به کاهش کیفیت محصول و کاهش درصد قند خالص در ریشه چغندر قند شود (Jafarnia et al. 2015). مقدار قند قابل استحصال از چغندر قند (ساکارز)، مهم‌ترین عامل در تعیین ارزش

این محصول است (Bahrami and Honarvar 2017). بنابراین با توجه به اهمیت بالای چغندر قند در تولید قند و شکر، ارائه راهکارهایی برای بهبود کیفیت این گیاه تأثیر مهمی در افزایش درصد قند قابل استحصال و کاهش ضایعات قندی و در نتیجه بالا رفتن میزان قند و شکر تولیدی و بالا رفتن کیفیت آن و افزایش راندمان کارخانه‌ها دارد. از جمله این راهکارها، استفاده از فرمول رینفلد و همکاران (Renfield et al. 1974; Honarvar et al. 2012). این فرمول تأثیر عناصر سدیم، پتاسیم و نیتروژن به روی درصد قند موجود در ملاس را نشان می‌دهد و به تولیدکنندگان قند و شکر کمک می‌کند تا کیفیت چغندر قند را مورد ارزیابی قرار دهند و میزان قند قابل استحصال از چغندر قند را محاسبه کنند.

استفاده از این فرمول برای کشور ایران عاری از خطا نیست و با توجه به تفاوت اساسی میان میزان ناخالصی‌ها از جمله سدیم و پتاسیم در محصول تولیدی ایران و اروپا و عدم وجود محدوده قابل راستی آزمایی (Honarvar and Alimoradi 2003) بر اساس مدل رینفلد و سایر فرمول برآورد شده توسط هافمن و همکاران (Hoffmann et al. 2021) وجود ندارد و همچنین لازم به ذکر است که فرمول‌های مزبور برای چغندر قندهای بهاره است به همین سبب تحقیق حاضر، با هدف بومی‌سازی معادله رینفلد و ایجاد یک رابطه جدید میان عناصر سدیم، پتاسیم و نیتروژن آمینه و درصد قند موجود در ملاس چغندر قند پاییزه کاشت و برداشت‌شده از استان خوزستان، با بهره‌گیری از روش‌های مربوط به حوزه یادگیری ماشین (Machine Learning) انجام گرفت و اعتبارسنجی آن بر اساس میزان ملاس تولیدی واقعی کارخانه در طی دوره بهره‌برداری برآورد شد.

مواد و روش‌ها

آماده‌سازی نمونه

در این تحقیق از نمونه‌های چغندر قند پاییزه برداشت‌شده از استان خوزستان در سال ۱۴۰۱ از اول اردیبهشت‌ماه لغایت پایان

$$ALC = (Na + K) / N \quad \text{رابطه ۲}$$

تحلیل آماری

پیش پردازش داده‌ها

برای انجام این تحقیق تعداد ۳۸۲ نمونه چغندر قند سالم برای سنجش محتوی سدیم، پتاسیم، نیتروژن و ملاس به‌طور تصادفی انتخاب شد که در این فرآیند هیچ داده گم شده و پرت وجود نداشت که این مسأله باعث عدم نیاز به بخش‌هایی از فرآیند پیش‌پردازش داده‌ها گردید. به‌منظور حصول اطمینان از صحت و دقت انجام مدل‌سازی در این مطالعه هم از داده‌های تغییر مقیاس‌شده که از طریق کم کردن از میانگین داده‌ها و تقسیم نمودن بر انحراف از معیار داده‌ها به دست آمد و هم از داده‌های بدون تغییر مقیاس استفاده شد اما در نهایت بهترین مدل تعریف شده از داده‌های بدون تغییر مقیاس به دست آمد. برای تحلیل اکتشافی داده‌ها در این تحقیق نیز در مرحله بعد از بررسی توزیع نرمال و متعاقب آن تجزیه واریانس یک‌طرفه و سپس ترسیم نمودارهای مربوطه صورت گرفت.

تجزیه واریانس یک‌طرفه

داده‌های حاصل از سنجش میزان سدیم، پتاسیم، نیتروژن میلی‌اکی‌والان در ۱۰۰ گرم خمیر ریشه و درصد قند ملاس با استفاده از آزمون Kolmogorov-Smirnov در سطح معنی‌داری پنج درصد از نقطه‌نظر نرمال بودن پراکنش داده‌ها مورد بررسی قرار گرفت که مشخص شد داده‌ها دارای توزیع نرمال بودند و از این‌رو از تجزیه واریانس یک‌طرفه و مقایسه میانگین به روش چند دامنه‌ای دانکن در سطح احتمال پنج درصد در طی هشت هفته بهره‌برداری با استفاده از نرم‌افزار SPSS انجام گرفت.

مدل‌سازی

پس از انجام اندازه‌گیری‌ها، نمونه‌ای ۳۸۲ تایی به‌طور تصادفی انتخاب و به نرم‌افزار پایتون منتقل شد و با استفاده از روش‌های یادگیری ماشین، ابتدا به دلیل نامشخص بودن درصد

خردادماه استفاده شد. از مجموع محموله‌های ارسالی به کارخانه قند قزوین، به‌صورت روزانه و تصادفی نمونه‌برداری و ریشه‌ها در بخش عیارسنج کارخانه مطابق دستورالعمل مؤسسه اصلاح و تهیه بذر چغندر قند آماده‌سازی شدند. به این صورت که ریشه‌ها شستشو و توزین شدند و سپس خمیر ریشه تهیه شد. این خمیرها در کارخانه به‌صورت یکنواخت درآمدند و سریعاً در دمای ۲۰- درجه سانتی‌گراد منجمد شدند. این آزمایش‌ها در مدت‌زمان هشت هفته انجام شدند و برای هر هفته تعداد ۴۸ نمونه خمیر ریشه تهیه و منجمد شد. سپس این خمیرها به مؤسسه تحقیقات اصلاح و تهیه بذر چغندر قند کرج ارسال و تحت آزمایش قرار گرفتند.

آزمون‌های شیمیایی

برای انجام آزمایش‌ها، از خمیر ریشه‌ها استفاده گردید. نمونه‌های منجمد شده، یخ‌زدایی شدند و سپس مورد تجزیه شیمیایی قرار گرفتند. برای شفاف‌سازی نمونه‌ها، از استات سرب قلیائی استفاده شد. تمام نمونه خمیرها در طی برداشت و همچنین در طی نگهداری به‌صورت یک هفته‌ای، بدون تأخیر مورد آزمایش قرار گرفتند. درصد قند ناخالص، به روش پلاریمتری با استفاده از دستگاه پلاریمتر مدل P3000 متعلق به برند کروز کشور آلمان، اندازه‌گیری شد. میزان عناصر معدنی سدیم و پتاسیم، با دستگاه فلیم فتومتر مدل AF100 Biotech ساخت کشور انگلستان، اندازه‌گیری شدند. میزان عنصر نیتروژن آمینه، به روش عدد آبی با استفاده از دستگاه اسپکتروفتومتر مدل D-3016 ساخت انگلستان، اندازه‌گیری شد. درصد قند ملاس، بر اساس میزان سدیم و پتاسیم و نیتروژن آمینه و از طریق فرمول رینفولد (رابطه ۱) محاسبه شد.

$$SM = 0.343 (Na + K) + 0.094 N + 0.029 \quad \text{رابطه ۱}$$

در این رابطه SM قند موجود در ملاس و Na، K و N به ترتیب میزان سدیم، پتاسیم و نیتروژن آمینه را نشان می‌دهند. درصد قند خالص، از طریق تفاضل درصد قند ناخالص و درصد قند ملاس محاسبه شد. قلیائیت نیز بر اساس رابطه (۲) محاسبه شد.

بر مبنای مقایسه مقدار قند واقعی ملاس حاصل از گزارش نهائی بهره‌برداری کارخانه قند قزوین با مقادیر به‌دست‌آمده از رابطه پیشنهادی و رابطه رینفلد توسط آزمون t تک نمونه‌ای بررسی گردید.

نتایج

مقایسه مقدار عناصر اندازه‌گیری شده در نمونه‌ها طی

هشت هفته بهره‌برداری

همانطور که عنوان شد در نخستین مرحله لازم است مقادیر عناصر سدیم، پتاسیم و نیتروژن آمینه در طی ۸ هفته بهره‌برداری مورد مقایسه قرار گیرد و بازه تغییرات مشخص شود که نتایج در جداول ۱ و ۶ و شکل‌های ۱، ۲، ۳ و ۴ ارائه شده است. همچنین نتایج تجزیه واریانس یک طرفه داده‌های مربوط به عناصر سدیم، پتاسیم و نیتروژن آمینه و همچنین ملاس در طی هشت هفته بهره‌برداری در جدول ۱ ارائه شده است.

قند ملاس (متغیر وابسته) در ایران، با استفاده از دو روش غیر نظارتی تجزیه به مؤلفه اصلی (PCA) و خوشه‌بندی (K -Means)، درصد قند ملاس محاسبه شد. در روش تجزیه به مؤلفه اصلی، از روی سه متغیر مستقل (x) سدیم، پتاسیم و نیتروژن آمینه، متغیر وابسته درصد قند ملاس (y) به دست آمد. این روش یک‌بار برای داده‌های تغییر مقیاس داده شده و یک‌بار برای داده‌های اصلی و بدون اعمال تغییر پیاده شد و متغیر وابسته محاسبه گردید که به ترتیب با نمادهای $yorgpca$ و $ypca$ مشخص شدند. سپس از روش خوشه‌بندی استفاده شد و برای دقت بیشتر تعداد خوشه‌ها با استفاده از متد elbow به تعداد ۳۵ خوشه در نظر گرفته شد و میانگین داده‌های هر سه متغیری که در یک خوشه بودند، محاسبه شد. متغیر وابسته حاصل از این روش با نماد $ycluster$ مشخص شد.

بررسی صحت مدل

پس از مدل‌سازی رابطه بومی‌شده محتوی قند ملاس بر اساس سه عنصر سدیم، پتاسیم و نیتروژن آمینه، صحت مدل

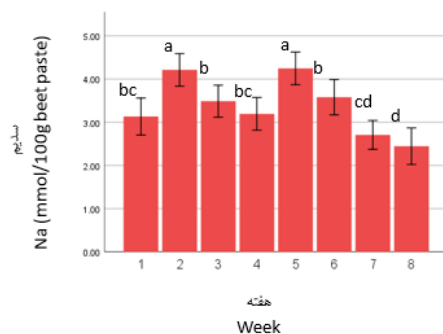
جدول ۱ تجزیه واریانس مقایسه مقدار عناصر در طی هشت هفته بهره‌برداری

Table. 1: Analysis of variance comparing the amount of elements over eight weeks of operation

		Sum of Squares	df	Mean Square	F	Sig.
Na	Between Groups	125.669	7	17.953	10.128	.000
	Within Groups	655.831	370	1.773		
	Total	781.500	377			
K	Between Groups	16.640	7	2.377	3.525	.001
	Within Groups	249.530	370	.674		
	Total	266.170	377			
N	Between Groups	20.520	7	2.931	10.846	.000
	Within Groups	100.004	370	.270		
	Total	120.524	377			
Molase	Between Groups	16.998	7	2.428	8.529	.000
	Within Groups	105.346	370	.285		
	Total	122.345	377			

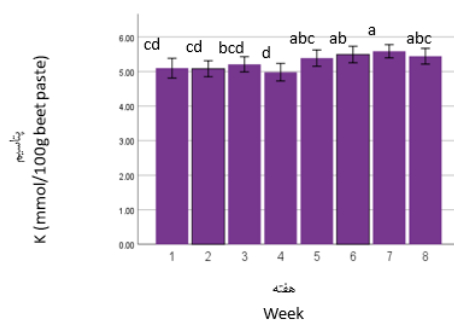
مفهوم در شکل‌های ۱ الی ۴ به تفکیک ۸ هفته بهره‌برداری از اول تا پایان خردادماه ۱۴۰۱ نمایش داده شده است.

از جدول ۱ مشخص است که مقادیر اندازه‌گیری‌شده در طول هشت هفته دارای اختلاف معنی‌داری ($P < 0.05$) بودند که این



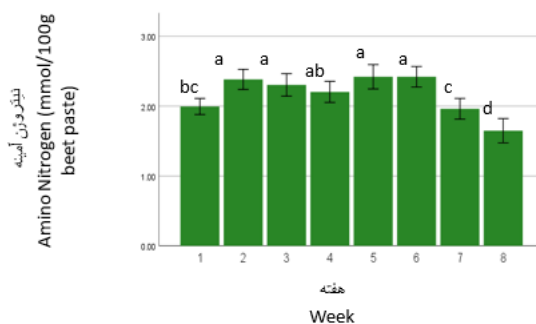
شکل ۱ مقایسه محتوی سدیم (میلی مول در صد گرم چغندر قند) نمونه‌های چغندر قند پاییزه در طی هشت هفته بهره‌برداری سال ۱۴۰۱ از اول اردیبهشت ماه-پایان خرداد ماه

Fig. 1 Comparison of sodium content (mmol/100SB) in autumn beet samples over eight-week period (May-June) of operation in 2022



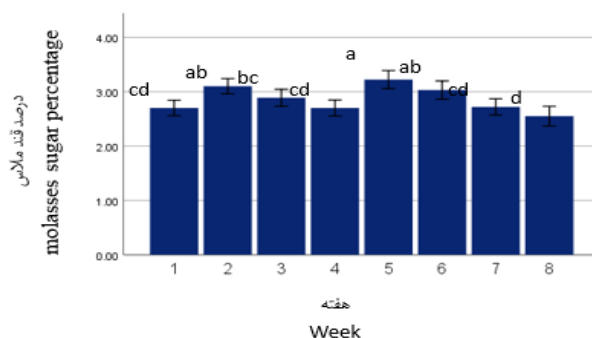
شکل ۲ مقایسه محتوی پتاسیم (میلی مول در صد گرم چغندر قند) نمونه‌های چغندر قند پاییزه در طی هشت هفته بهره‌برداری سال ۱۴۰۱ از اول اردیبهشت ماه-پایان خرداد ماه

Fig. 2 Comparison of potassium content (mmol/100 g sugar beet) in autumn beet samples over eight weeks of operation (May-June) in 2022



شکل ۳ مقایسه محتوی نیتروژن آمینه (میلی مول در صد گرم چغندر قند) نمونه‌های چغندر قند پاییزه در طی هشت هفته بهره‌برداری سال ۱۴۰۱ از اول اردیبهشت ماه-پایان خرداد ماه

Fig. 3 Comparison of nitrogen content (mmol/100 g sugar beet) in autumn beet samples over eight weeks of operation (May-June) in 2022



شکل ۴ مقایسه محتوی قند ملاس نمونه‌های چغندر قند پاییزه در طی هشت هفته بهره‌برداری سال ۱۴۰۱ از اول اردیبهشت ماه-پایان خرداد ماه
Fig. 4 Comparison of molasses content in autumn beet samples over eight weeks of operation (May-June) in 2022.

از آنجا که انتظار می‌رفت متغیرهای سدیم، پتاسیم و نیتروژن

آمین به روی درصد قند ملاس تأثیر مثبت داشته باشند یعنی ضرایب رگرسیونی مثبت شود، این انتظار در مورد متغیر پاسخ به دست آمده به روش تجزیه به مؤلفه اصلی چه در حالتی که داده‌ها تغییر مقیاس دارند (جدول ۲) و چه در حالتی که تغییر مقیاس ندارند (جدول ۳) برآورده نشد، بنابراین تنها متغیر پاسخ حاصل از روش خوشه‌بندی مناسب بود که ضرایب رگرسیونی مطابق جدول ۴ است.

روش‌های غیر نظارتی تحلیل مؤلفه اصلی و خوشه‌بندی برای بومی‌سازی معادله رینفیلد در ایران در مرحله بعد، برای یافتن رابطه میان سه متغیر مستقل و متغیر وابسته و ایجاد معادله (مدل)، به صورت تصادفی از ۸۰ درصد داده‌ها به عنوان داده آموزشی استفاده شد و روش نظارتی رگرسیون روی این داده‌ها برای هر سه متغیر وابسته yorgpca و ycluster و yzca برآزش داده شد.

جدول ۲ خروجی پایتون برای ضرایب رگرسیونی مربوط به متغیر وابسته به دست آمده از روش تحلیل مؤلفه اصلی برای داده‌های تغییر مقیاس شده

Table. 2 Python output of the regression coefficients for the dependent variable obtained using the principal component analysis method on the given rescaled data.

	coefficient	standard error	t	P> t	[0.025	0.975]
const	-1.239e-10	1.16e-10	1.069	0.286	-3.52e-10	1.04e-10
Na	0.9842	1.2e-11	8.17e+10	0.000	0.984	0.984
K	-0.0709	2e-11	-3.54e+09	0.000	-0.071	-0.071
N	0.1622	3.12e-11	5.2e+09	0.000	0.162	0.162

جدول ۳ خروجی پایتون برای ضرایب رگرسیونی مربوط به متغیر وابسته به دست آمده از روش تحلیل مؤلفه اصلی برای داده‌های بدون تغییر مقیاس

Table. 3 Python output of the regression coefficients for the dependent variable obtained using the principal component analysis method on the unscaled data.

	coefficient	standard error	t	P> t	[0.025	0.975]
const	22.7973	3.853	5.917	0.000	15.216	30.379
Na	-0.0491	0.427	-0.115	0.909	-0.890	0.792
K	-0.2999	0.686	-0.437	0.662	-1.650	1.050
N	-2.4105	1.087	-2.218	0.027	-4.549	-0.272

جدول ۴ خروجی پایتون برای ضرایب رگرسیونی مربوط به متغیر وابسته به دست آمده از روش خوشه‌بندی

Table. 4 Python output of the regression coefficients associated with the dependent variable obtained using a clustering method

	coefficient	standard error	t	P> t	[0.025	0.975]
const	0.1940	0.056	3.462	0.001	0.084	0.304
Na	0.3267	0.006	54.390	0.000	0.315	0.339
K	0.3135	0.010	31.237	0.000	0.294	0.333
N	0.3036	0.016	19.564	0.000	0.273	0.334

رابطه ۳ $M = 0.1940 + 0.3267 (Na) + 0.3135 (K) + 0.3036 (N)$ مقدار احتمال‌ها در ستون $|t| > P$ ، برای همه مقادیر صفر شده است که یعنی متغیرهای سدیم، پتاسیم و نیتروژن آمینه حضورشان در مدل معنی‌دار است.

معیارهای میانگین توان دوم خطاها و میانگین قدر مطلق خطاها دقت پیش‌بینی با کدهای تعریف شده در نرم‌افزار پایتون در این چهار روش نظارتی یادگیری ماشین مطابق جدول ۵ مقایسه گردیده است. همچنین مقدار R^2 نیز در جدول آمده است. لازم به ذکر است روشی که دارای خطای کمتر یعنی میانگین توان دوم خطاها و میانگین قدر مطلق خطاها کمتر باشد، بهتر است؛ روشی که مقدار R^2 بیشتری دارد مناسب‌تر است. ملاحظه می‌شود که روش رگرسیون بهتر از سایر روش‌ها عمل می‌کند زیرا میانگین توان دوم خطاها و میانگین قدر مطلق خطاهای کمتری دارد. مقدار ضریب تبیین به‌دست آمده ($R^2 = 0.93$)، نشان‌دهنده‌ی این است که این روش می‌تواند درصد زیادی از تغییرات متغیر وابسته توسط متغیرهای مستقل را نشان دهد.

به همین علت برای بقیه تحقیقات تنها متغیر پاسخ به‌دست‌آمده از روش خوشه‌بندی به کار گرفته شد. سایر روش‌های یادگیری ماشین مانند رگرسیون بردار پشتیبان، درخت تصمیم، جنگل تصادفی و بگینگ روی داده‌های آموزشی برازش داده شد که مطابق کدهای ضمیمه است.

از ۲۰ درصد باقی‌مانده داده‌ها به عنوان داده‌های آزمایشی (تست) استفاده شد و عملکرد مدل رگرسیونی و سایر روش‌های یادگیری ماشین به این صورت که متغیر وابسته برآورد شده (y_{pred}) با مقدار واقعی داده‌های تست (y_{test}) که از روش‌های مختلف به دست آمده‌اند مقایسه شدند. از سه معیار میانگین توان دوم خطاها، میانگین قدر مطلق خطاها و R^2 برای این مقایسه استفاده شد.

مطابق جدول ۴ ضرایب رگرسیونی معادله بومی‌سازی شده (رابطه ۳) به صورت زیر که در آن M درصد قند موجود در ملاس و Na ، K و N به ترتیب میزان سدیم، پتاسیم و نیتروژن آمینه را نشان می‌دهند، به دست آمد:

جدول ۵ مقایسه خطا و R^2 در روش‌های رگرسیونی، رگرسیون بردار پشتیبان، درخت تصمیم، جنگل تصادفی و بگینگ

Table.5 Comparison of error and R^2 in regression methods, support vector regression, decision tree, random forest and bagging

Methods	متدها	MSE	MAE	R^2
Regression	رگرسیون	0.016	0.098	0.093
SVR	رگرسیون بردار پشتیبان	0.027	0.13	0.926
Decision Tree	درخت تصمیم	0.037	0.0909	0.9
Random Forest	جنگل تصادفی	0.031	0.11	0.915
Bagging	بگینگ	0.03	0.11	0.919

پیشنهادی ارائه شده قرار داده و با مقدار ملاس واقعی به‌دست‌آمده از نمونه‌های چغندر قند مطابق جدول زیر مقایسه انجام شد و همان‌طور که در جدول ۶ نمایش داده شده است با استفاده از آزمون t تک‌نمونه‌ای در سطح معنی‌داری مشخص شد ($P < 0.05$) که مقادیر قند موجود در ملاس به‌دست‌آمده از طریق فرمول رینفلد با مقادیر واقعی دارای اختلاف معنی‌دار بود اما مقدار واقعی

مقایسه آماری نتایج محتوی قند ملاس تخمینی به‌دست‌آمده از فرمول رینفلد، فرمول ارائه شده در تحقیق حاضر با مقدار واقعی

در این قسمت برای مقایسه عملکرد رابطه پیشنهادی مقادیر سدیم، پتاسیم و نیتروژن آمینه مربوط به نمونه‌های حاضر را در فرمول (Renfield et al. 1974)، همچنین این مقادیر در رابطه رگرسیونی

با مقادیر به‌دست‌آمده از رابطه ارائه‌شده اختلاف معنی‌دار نداشت. همچنین در جدول ۶ بازه تغییرات سه عنصر سدیم، پتاسیم و آلفا آمینو نیتروژن که فرمول پیشنهادی بر اساس آن تبیین شده ارائه گردیده است.

جدول ۶ مقایسه عملکرد رابطه پیشنهادی با رابطه رینفیلد و مقدار واقعی ملاس به‌دست‌آمده و دامنه عملکرد فرمول پیشنهادی بر اساس سه متغیر سدیم، پتاسیم و نیتروژن آمینه

Table. 6 Comparison of the performance of the proposed equation with the Reinfield equation and the actual molasses content, along with the performance range of the proposed equation based on Na, K, α -amino N contents.

عناصر Elements	مقدار واقعی actual amount				رابطه پیشنهادی Proposed equation				رابطه رینفیلد Renfield equation			
	نیتروژن آمینه α -amino N				پتاسیم K				سدیم Na			
	انحراف معیار Stdev.	ماکزیمم Max.	مینیمم Min.	میانگین Mean	انحراف معیار Stdev.	ماکزیمم Max.	مینیمم Min.	میانگین Mean	انحراف معیار Stdev.	ماکزیمم Max.	مینیمم Min.	میانگین Mean
	0.58	3.9	0.65	2.18	0.84	9.12	3.2	5.27	1.43	7.39	0.71	3.4

^a حروف کوچک متمایز بیانگر وجود اختلاف معنی‌دار در سطح احتمال پنج درصد برای مقادیر میانگین قند ملاس محاسبه‌شده با استفاده از رابطه پیشنهادی تحقیق حاضر، رابطه رینفیلد و مقدار واقعی می‌باشد.

^a Distinctive lowercase letters indicate significant difference at the 5% probability level among the average molasses sugar values calculated using the proposed equation in this study, the Reinfield equation, and the actual values.

روش بهترین عملکرد را در پیش‌بینی متغیر وابسته داشته است.

شاید علت عملکرد بهتر روش رگرسیون از روش‌های دیگر که بسیار پیچیده‌تر هستند این است که تعداد متغیرها کم است و نیاز به این همه پیچیدگی نیست. ضمن این که تنها روش رگرسیون است که یک معادله برای تأثیر سدیم، پتاسیم و نیتروژن روی درصد قند ملاس می‌دهد. سایر روش‌ها برای پیش‌بینی مناسب هستند نه ارائه رابطه بین متغیرهای مستقل و وابسته. هرچند که در این تحقیق به دلیل کمتر بودن خطا در رگرسیون این اتفاق نمی‌افتد.

نتیجه‌گیری

در نهایت، نتایج این تحقیق نشان می‌دهد که متغیرهای مستقل سدیم، پتاسیم و نیتروژن تأثیر قابل توجهی بر تعیین درصد قند ملاس دارند و بومی‌سازی معادله رینفیلد برای ارزیابی کیفیت چغندر قند در ایران با موفقیت انجام شد و امکان پیش‌بینی درصد قند ملاس با در دست داشتن مقدار سدیم، پتاسیم و نیتروژن آمینه با معادله بومی‌سازی شده مهیا گردید. این معادله

بحث

بر اساس مطالعات عده‌ای انجام‌شده طی سالیان اخیر (Amiri *et al.* 2022; Bahadorbaigy *et al.* 2016; Bahrami and Honarvar 2017; Honarvar *et al.* 2012; Jafarnia *et al.* 2105)، مدل ارزیابی کیفیت تکنولوژیکی چغندر قند طبق رابطه رینفیلد برای شرایط کشور ایران مناسب نیست و ارزش تکنولوژیکی چغندر قند به بسیاری از پارامترها از جمله شرایط اقلیمی، چگونگی استفاده از نهاده‌های کشاورزی و زمان برداشت بستگی دارد لذا این تحقیق با توجه به مقایسه با مقدار ملاس تولیدی واقعی طی یک دوره بهره‌برداری از چغندر پاییزه مؤید این ادعا می‌باشد.

تجزیه و تحلیل داده‌ها با استفاده از روش یادگیری ماشین، نتایج قابل توجهی را در ارتباط با متغیرهای وابسته (درصد قند ملاس) نشان داد. این نتایج تأثیر متغیرهای مستقل سدیم (Na)، پتاسیم (K) و نیتروژن آمینه (N) بر متغیرهای وابسته را نشان می‌دهد. معیارهای خطا مانند میانگین توان دوم خطاها و میانگین قدرمطلق خطاها برای روش رگرسیون نشان می‌دهد که این

می‌تواند به کشاورزان و محققان در بهینه‌سازی کشت و تولید
چغندر قند کمک شایانی نماید.

سیاسگزاری

از کارخانه قند قزوین و همچنین مؤسسه اصلاح و تهیه بذر
چغندر قند به خاطر فراهم آوردن امکانات مورد نیاز برای انجام
تحقیق حاضر نهایت تشکر را داریم.

تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان بیان نشده
است.

References

منابع مورد استفاده

- Amiri Z, Asgharipour MR, Moghadam EH, Kakolv E, Campbell DE. Investigating the need to replace the conventional method of sugar beet production in Lorestan province, Iran based on the arguments obtained from emergy calculations. *Ecological Modelling*. 2022; 472(4): 110091. Doi: <https://doi.org/10.1016/j.ecolmodel.2022.110091>.
- Bahadorbaigy A, Honarvar M, Ansari K, Bahrami ME. Determining the area of sugar beet using digital image processing method and its correlation with the technological factors. *Research and Innovation in Food Science and Technology*. 2016; 5(1): 15-26. Doi: <https://doi.org/10.22101/JRIFST.2016.06.01.512>
- Bahrami ME, Honarvar M. The Effect of environmental conditions, planting and harvest time on the quantity and technological values of beet sugars cultivated in west part of Iran. *Journal of Food Technology and Nutrition*. 2017; 2(14): 49-62. Doi: <https://sanad.iau.ir/en/Journal/jftn/Article/832527>
- El-Geddawy MA, Mustafa OA, El-Syiad SI, Mohamed ESGI. The potential impact of the quality of sugar Beet roots (*Beta vulgaris L.*) on sugar loss in the beet molasses. *Assiut Journal of Agricultural Sciences*. 2023; 54(4): 108-122. Doi: <https://doi.org/10.21608/AJAS.2023.232654.1289>
- Hoffmann CM, Koch HJ, Märlander B. Sugar beet. Pp. 634-672. In *Crop Physiology Case Histories for Major Crops*, Academic Press, Cambridge, MA, USA, 2021;. Doi: <https://doi.org/10.1016/B978-0-12-819194-1.00020-7>
- Honarvar M, Kalbasi Ashtari A, Karimi K. Estimation of sugar loss in the produced molasses in the sugar beet factory based on the technological quality of sugar beet. *Journal of Food Technology and Nutrition*. 2012; 9(35): 31-38. Doi: <https://sid.ir/paper/143406/en>
- Honarvar M, Alimoradi I. Factors affecting quality of sugar beet at semi-arid areas of Isfahan. In: *Proceedings of the 1st Joint IIRB-ASSBT Congress; 2003 February 26-March 1; San Antonio, US; 2003; P: 597-601*.
- Jafarnia B, Zare Fayzabadi A, Ghorbani R, Rezvani Moghadam P. The effect of plant density and nitrogen and biological fertilizers on the qualitative performance of sugar beet in two regions of Mashhad and Torbat Jam. *Iranian Journal of Field Crops Research*. 2015; 13(2): 278-286. Doi: <https://doi.org/10.22067/gsc.v13i2.26511>
- Lv X, Chen S, Wang Y. Advances in understanding the physiological and molecular responses of sugar beet to salt stress. *frontiers in plant science*. 2019; 10: 14-31. Doi: <https://doi.org/10.3389/fpls.2019.01431>
- Nefedeva E, Belopukhov S, Kravchenko N. Application of a universal herbicide for beets: an overview. *Agroekoinfo*. 2023; 3(57): 19-19. Doi: <https://doi.org/10.51419/202133319>.

- Reinefeld E, Emmerich A, Baumgarten G, Winner C, Beiss U. Zur voraussage des melassezuckers aus rubenanalysen. Zucker. 1974; 27: 2-15
- Rezaei J, Fasahat P. Autumn-sown sugar beet cultivation in semiarid regions. pp. 275-290. In: Sugar Beet Cultivation, Management and Processing, Singapore: Springer Nature Singapore. 2022. Doi: https://doi.org/10.1007/978-981-19-2730-0_14
- Yanuartono Y, Nururrozi A, Indar-Julianto S, Purnamaningsih H, Rahardjo S. Molasses: dampak negatif pada ruminansia. Jurnal ilmu-ilmu peternakan. 2017; 27(2): 25-34. Doi: <https://doi.org/10.21776/UB.JIIP.2017.027.02.04>