

EKSPLORACYJNA ANALIZA I MODELOWANIE PROCESU EKSTRUZJI BŁYSKAWICZNYCH MAKARONÓW PEŁNOZIARNISTYCH

Agnieszka Wójtowicz

Katedra Inżynierii Procesowej, Uniwersytet Przyrodniczy w Lublinie

Andrzej Marciniak

Katedra Podstaw Techniki, Uniwersytet Przyrodniczy w Lublinie

Streszczenie. W pracy przedstawiono możliwości zastosowania sieci bayesowskich do analizy eksploracyjnej i modelowania procesu ekstruzji makaronów błyskawicznych. Wykrywanie i modelowanie zależności pomiędzy parametrami procesu i produktu przeprowadzono za pomocą trzech algorytmów uczenia maszynowego na danych empirycznych uzyskanych podczas procesu wytwarzania makaronów błyskawicznych: *MST*, *Taboo* oraz *SopLEQ*. Otrzymana topologia sieci była zgodna z przewidywaną strukturą zależności wewnątrzprocesowych pomiędzy parametrami procesu a cechami produktu, a oszacowane warunkowe rozkłady prawdopodobieństwa umożliwiły poprawne wnioskowanie predykcyjne i diagnostyczne.

Słowa kluczowe: ekstruzja, modelowanie, sieci bayesowskie

Wstęp

W pracy tej model procesu produkcyjnego jest sformalizowaną i zinstrumentalizowaną w postaci wykonywalnych programów komputerowych reprezentacją wiedzy o procesie. Idea oparcia procesów produkcyjnych na wiedzy i informacji zakłada, że modele procesów są równie ważnymi środkami produkcji, jak ludzie, materiały i energia. Wiedza uprzedmiotowiona w postaci modeli procesów pozwala na zarządzanie nimi w sposób powtarzalny, przewidywalny i optymalny, czyli najlepszy w określonych warunkach.

Modelowanie procesu ekstruzji środków spożywczych umożliwia zoptymalizowanie parametrów wytwarzania szerokiego asortymentu ekstrudatów spożywczych i paszowych oraz analizę wpływu parametrów procesu na właściwości fizykochemiczne i użytkowe wyrobów. W technologii ekstruzji jakość produktu gotowego można opisać wieloczynnikowym modelem zmiennych [Linko i in. 1981; Fornal 1998]:

$$J = f(S, D, W, P, T, R, F) \quad (1)$$

gdzie:

- J – jakość wyrobu,
- S – cechy i właściwości surowca,
- D – dodatki,

W	– wilgotność,
P	– ciśnienie,
T	– temperatura,
R	– czas retencji,
F	– siły ścinające.

Nie tylko parametry procesu, ale również właściwości przetwarzanego materiału, jego skład surowcowy, ilość komponentów ulegających reakcjom uplastyczniania w czasie ekstruzji, powstawanie kompleksów między składnikami produktu (interakcje białko-skrobia) oraz ich wpływ na cechy produktu będą wskazywały na dobór odpowiednich parametrów procesu [Ryu i Ng 2001; Mościcki i in. 2007].

Celem pracy jest zbudowanie komputerowego modelu procesu ekstruzji pełnoziarnistych makaronów błyskawicznych. Model ten z założenia spełnia funkcje adaptacyjnego systemu reprezentacji wiedzy procesowej opierając się na systemowej konceptualizacji relacji przyczynowo-skutkowych pomiędzy mierzalnymi i dedukowanymi wielkościami opisującymi parametry surowca (wejścia procesu), parametry procesu (przetwarzanie) oraz wielkościami wyjściowymi – parametry produktu.

Materiały i metody

Eksploracyjnej analizie i modelowaniu poddano dane empiryczne dokumentujące proces ekstruzji makaronów błyskawicznych pełnoziarnistych obejmujące: wilgotność surowca (%), obroty ślimaka ($\text{obr} \cdot \text{min}^{-1}$), wydajność produktu ($\text{kg} \cdot \text{h}^{-1}$), temperaturę w sekcji I, II ekstrudera ($^{\circ}\text{C}$), temperaturę głowicy formującej ($^{\circ}\text{C}$), temperaturę produktu ($^{\circ}\text{C}$), SME - specyficzne zapotrzebowanie energii mechanicznej ($\text{kWh} \cdot \text{kg}^{-1}$), ekspandowanie produktu (-), WAI – wskaźnik absorpcji wody (%), WSI - wskaźnik rozpuszczalności w wodzie (%), WHC - zdolność pochłaniania gorącej wody (%), WB – zdolność zatrzymywania wody po hydratacji (%), straty składników po hydratacji (%), siła cięcia suchych makaronów (N), BS - stosunek siły cięcia do średnicy makaronu (-), kruchość po hydratacji (N), twardość po hydratacji (N), siłę kohezji czyli spójność materiału po hydratacji (N), adhezję po hydratacji (mJ), żujność po hydratacji (N), gumistość po hydratacji (N), sprężystość po hydratacji (-), ocenę organoleptyczną po hydratacji (w skali 1-5 pkt.) [Wójtowicz i Marciniak 2010].

Do opracowania modelu wykorzystano środowisko modelowania probabilistycznego BayesiaLab (<http://bayesialab.com>). Dane poddano dyskretyzacji metodą K-Means z ewentualną ręczną korektą przedziałów. Wykrywanie i modelowanie asocjacji pomiędzy parametrami procesu i produktu przeprowadzono za pomocą trzech algorytmów maszynowego uczenia sieci bayesowskich dostępnych w BayesiaLab: *MST* – Maksymalne Drzewo Spinające (Maximum Spanning Tree), *Taboo* oraz *SopLEQ*.

Do oceny jakości dopasowania sieci do posiadanych danych mogą być użyte dwie miary: Minimum Description Length oraz współczynnik korelacji Pearsona. W wyniku tworzone jest drzewo o niezorientowanych łukach. Aby przekształcić je w sieć bayesowską, łukom należy nadać orientację tak, aby uniknąć cykli (sieć bayesowska musi być ukierunkowanym grafem acyklicznym).

Uczenie struktury sieci zaimplementowane w algorytmie *Taboo* jest szczególnie użyteczne przy doskonaleniu sieci budowanych przez ekspertów w modelowanej dziedzinie przedmiotowej lub do uaktualniania sieci otrzymanych na innych zbiorach danych uczących. Oprócz tego, że brana jest tu pod uwagę aprioryczna wiedza reprezentowana przez sieć i odpowiadający temu zbiór przykładów uczących, punktem wyjścia w metodzie *Taboo* jest aktualna sieć, w przeciwieństwie do metody *SopLEQ*. Algorytm uczenia *EQ* oparty jest na metodzie przeszukiwania zbioru możliwych struktur z wykorzystaniem klas równoważności [Munteanu i Bendou 2002]. Jest to metoda bardzo efektywna, ponieważ nie ma ryzyka zatrzymania się algorytmu na lokalnych minimach. Podobnie, jak w algorytmie *Taboo*, punktem startu *EQ* jest aktualna sieć. Algorytm *SopLEQ* (modyfikacja *EQ*) bazuje na globalnej charakterystyce danych i zastosowaniu relacji równoważności pomiędzy strukturami sieci [Joufee 2002].

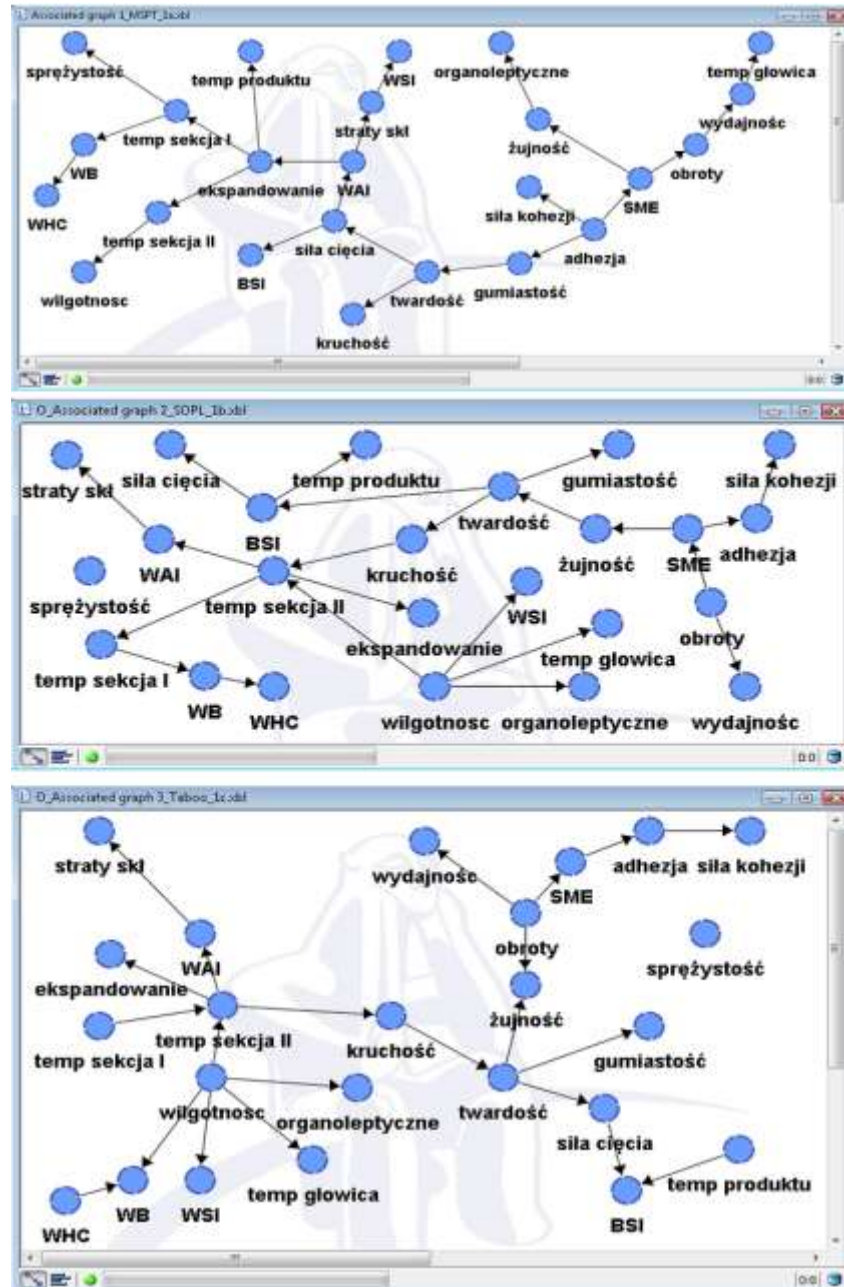
Model poddano walidacji na danych testowych wyodrębnionych losowo ze zbioru danych uczących.

Wyniki

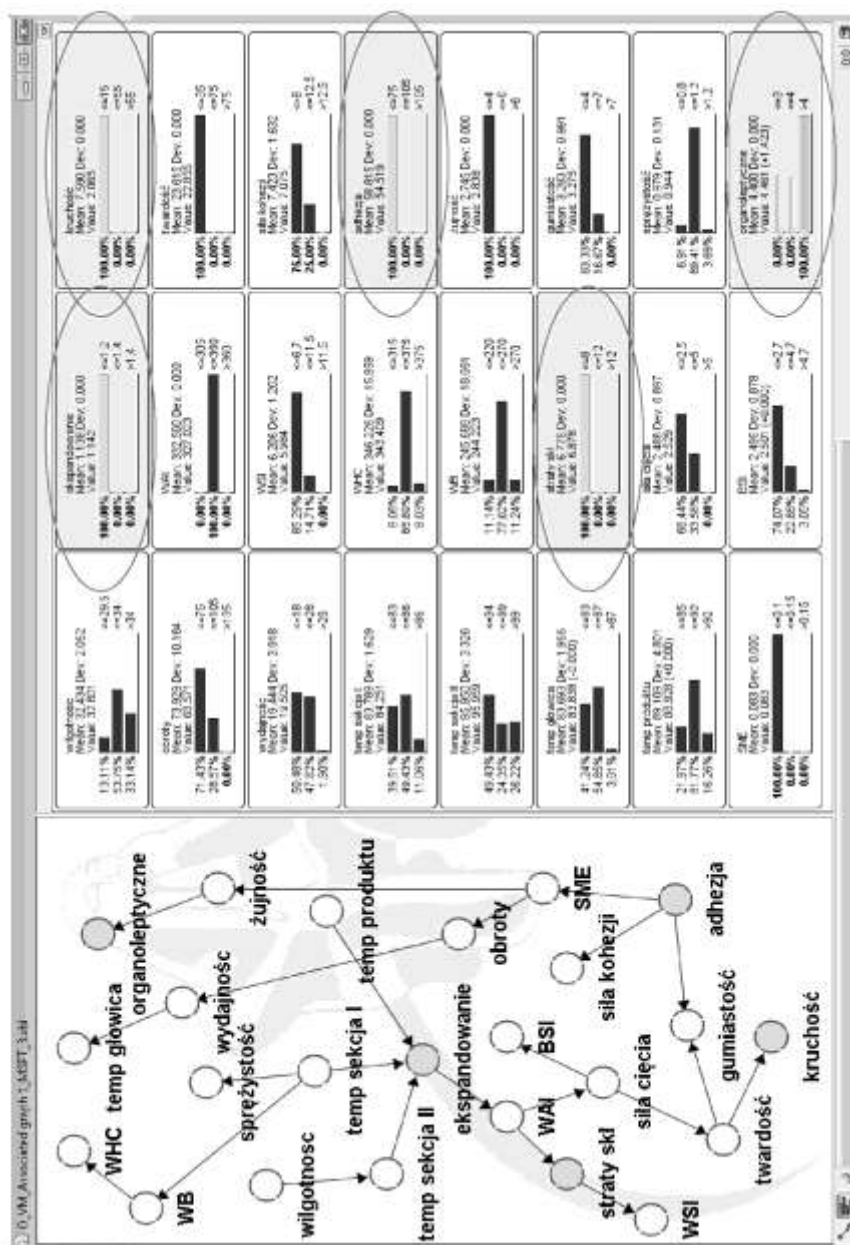
W badaniach zastosowano modelowanie probabilistyczne, ponieważ występuje tu zarówno niepewność stochastyczna związana z losowymi fluktuacjami parametrów surowca i procesu, a także niepewność epistemiczna, wynikająca z uproszczeń przestrzeni informacyjnej, w której opisywany jest badany proces.

W modelowaniu interesująca jest nie tyle sformalizowana postać zależności międzyprocesowych dla statystycznego procesu, ale procesu konkretnego, przebiegającego w określonych warunkach [Kusz i Marciniak 2006]. Adaptacyjność modelu oznacza automatyczne dopasowywanie się do warunków wytwarzania, zarówno organizacyjnych, jak i technologicznych w konkretnej firmie [Ghallab i Alaoui 1989]. Wyniki modelowania przedstawiono w postaci acyklicznych grafów skierowanych (A, B, C) na rys. 1. Na podstawie wiedzy o procesie ekstruzji makaronów błyskawicznych i wzajemnych relacjach pomiędzy parametrami procesu i parametrami produktu stwierdzono, że najbardziej kompleksowo i dokładnie zależności opisuje topologia sieci A uzyskana w wyniku uczenia maszynowego na przykładach z zastosowaniem metody *Maximum Spanning Tree*. Metoda uczenia sieci *MST* pozwala na opracowanie empirycznych warunkowych rozkładów prawdopodobieństwa nad poszczególnymi parametrami procesu i produktu.

Wykryte zależności pomiędzy parametrami procesu a parametrami produktu przedstawiono graficznie w postaci empirycznych warunkowych rozkładów prawdopodobieństwa nad poszczególnymi parametrami procesu i produktu. Wykorzystanie otrzymanej sieci probabilistycznej (model A) do wnioskowania o parametrach procesu umożliwiających uzyskanie zadanych parametrów produktu pokazano na rys. 2.



Rys. 1. Topologia sieci uzyskana w wyniku uczenia maszynowego na przykładach z użyciem trzech metod: A – MST, B – Taboo, C – SopLEQ
 Fig. 1. Network topology learned using three methods: A – MST, B – Taboo, C – SopLEQ



Rys. 2. Warunkowe rozkłady prawdopodobieństwa i wnioskowanie objaśniające możliwości uzyskania wymaganych parametrów produktu w metodzie uczenia *MST*
 Fig. 2. Conditional probability distributions and explanation inference of product with required parameters achievement with learning method *MST*

W przypadku badanego procesu ekstruzji makaronów błyskawicznych pełnoziarnistych parametrami krytycznymi produktu, warunkującymi uzyskanie wyrobów o akceptowalnej jakości były: wskaźnik ekspandowania promieniowego, którego odpowiednio niska wartość umożliwia skrócenie czasu przygotowania makaronu do spożycia przez hydratację w gorącej wodzie, niskie straty składników przechodzących do wody w czasie hydratacji świadczące o dobrym zintegrowaniu i przetworzeniu składników podczas procesu, niska twardość makaronów oraz adhezyjność po hydratacji, odpowiadające za wrażenia sensoryczne podczas konsumpcji, a także jak najwyższe wskaźniki oceny cech organoleptycznych makaronów. Wymagane wartości tych parametrów zadano jako cel wnioskowania (obramowania na rys. 2). W wyniku tego wnioskowania stwierdzono, że jeżeli wilgotność surowców wprowadzanych do procesu nie będzie przekraczała 34% a zastosowane obroty ślimaka podczas ekstruzji nie będą większe niż 105 obr min^{-1} , to z dużym prawdopodobieństwem uzyskamy produkt o wskaźniku ekspandowania 1,2-1,4, wskaźniku WSI poniżej 6,7%, wskaźniku WHC w granicach 315-375%, zdolności wiązania wody pomiędzy 220 a 270%, straty składników podczas hydratacji nie przekroczą 8%, twardość wyrobów będzie niższa niż 35 N, sprężystość makaronu po hydratacji wyniesie od 0,8 do 1,2 N, zaś cechy organoleptyczne osiągną wartości 4,0 i powyżej, co świadczy o dobrej jakości produktów.

Zastosowanie wnioskowania objaśniającego umożliwia wykrycie najbardziej istotnych parametrów procesu ekstruzji oraz wskazanie tych cech produktów, które przy zadanych ograniczeniach można otrzymać, ze wskazaniem, z jakim prawdopodobieństwem cechy te zostaną uzyskane. Innym zastosowaniem mechanizmu wnioskowania objaśniającego w połączeniu z wnioskowaniem predykcyjnym jest wykrywanie anomalii w procesie produkcyjnym odpowiedzialnych za zaobserwowane odchylenia parametrów produktu od wymaganych wartości.

Podsumowanie

Zbudowany model procesu ekstruzji makaronu posiada wymagane cechy funkcjonalne, takie jak poprawność rzeczową wnioskowania predykcyjnego i objaśniającego oraz racjonalny, zgodny z wiedzą przedmiotową mechanizm reprezentacji zależności wewnątrzprocesowych i uaktualniania niepewności.

Z eksperymentów przeprowadzonych na modelu wynika, że parametry produktu są współzależne. Uwzględnienie tych zależności może się odbywać poprzez iteracyjne włączanie poszczególnych parametrów produktu w odpowiedniej kolejności do zbioru zmiennych objaśniających.

Sieć bayesowska modelująca zależności wewnątrzprocesowe, zaimplementowana w informatycznej infrastrukturze procesu produkcji może funkcjonować jako adaptacyjna baza wiedzy, automatycznie aktualizująca się i dostrajająca do warunków produkcji w konkretnej firmie w miarę napływu nowych danych gromadzonych w związku z obligatoryjnym rejestrowaniem i dokumentowaniem procesów produkcyjnych dla potrzeb zarządzania jakością. Mechanizmy automatycznego wnioskowania predykcyjnego i objaśniającego mogą być wykorzystywane do szybkiego wykrywania błędów w procesie produkcyjnym, dostrajania parametrów procesu do zmieniających się parametrów surowców oraz rozwiązywania innych podobnych problemów tak, aby proces i produkt były porównywalne i powtarzalne w zmieniającym się otoczeniu.

Bibliografia

- Fornal L.** 1998. Ekstruzja produktów skrobiowych – nowe wyroby. *Pasze Przemysłowe*. Nr 3. s. 7-14.
- Ghallab M., Alaoui A. M.** 1989. Managing efficiently temporal relations through indexed spanning trees. **Proceedings of the 11th international joint conference on Artificial intelligence – Volume 2**. Knowledge Representation. Detroit, Michigan. Morgan Kaufmann Publishers Inc. San Francisco. USA. s. 1297-1303.
- Jouffe L.** 2002. Nouvelle classe de méthodes d'apprentissage de réseaux bayésiens. *Journées Francophones d'Extraction et Gestion des Connaissances (EGC 2002)*. Montpellier. Francja.
- Linko P., Colonna P., Mercier C.** 1981. High-temperature, short-time extrusion cooking. w: *Advances in Cereal Science and Technology*. Volume 4 (red. Pomeranz Y.) AACC. Minnesota. USA. s. 145-235.
- Lu H., Ravi R.** 1998. Approximating maximum leaf spanning trees in almost linear time. *Journal of Algorithms*. 29. s. 132-141.
- Munteanu P., Bendou M.** 2002. The EQ framework for learning equivalence classes of bayesian network. *First IEEE International Conference on Data Mining (IEEE ICDM)*. San José. USA. s. 417-424.
- Kusz A., Marciniak A. W.** 2006. Dynamiczne sieci probabilistyczne jako system reprezentacji wiedzy. *Inżynieria Rolnicza*. Nr 12 (87). s. 285-294.
- Mościcki L., Mitrus M., Wójtowicz A.** 2007. *Technika ekstruzji w przemyśle rolno-spożywczym*. PWRiL. Warszawa. ISBN 9788309010272.
- Ryu G., Ng P.** 2001. Effects of selected process parameters on expansion and mechanical properties of wheat flour and whole cornmeal extrudates. *Starch/Starke*. Nr 53. s. 147-154.
- Wójtowicz A., Marciniak A.** 2010. Eksploracyjna analiza i modelowanie procesu ekstruzji błyskawicznych makaronów pełnoziarnistych. *Materiały konferencyjne XIII Ogólnopolskiej Konferencji Naukowej „Zastosowanie technologii informacyjnych w rolnictwie”*. Kazimierz Dolny. s. 91.

*Praca naukowa finansowana ze środków na naukę w latach 2008-2010
jako projekt badawczy N N312 162334*

EXPLORATORY ANALYSIS AND MODELING OF EXTRUSION-COOKING PROCESS OF PRECOOKED WHOLEWHEAT PASTA PRODUCTS

Abstract. The paper presents application of Bayesian Network to exploratory analysis and modeling of extrusion-cooking process of precooked wholewheat pasta products. For knowledge discovery in extrusion process data and modeling interdependencies of process and product parameters there were used machine learning methods available in BayesiaLab BN modeling system: *MST*, *Taboo* and *SopLEQ*. Resulted BN topology and conditional probability distributions assured satisfied accuracy of both predictive and diagnostic reasoning.

Key words: extrusion-cooking, modeling, Bayesian Networks

Adres do korespondencji:

Agnieszka Wójtowicz; e-mail: agnieszka.wojtowicz@up.lublin.pl

Katedra Inżynierii Procesowej

Uniwersytet Przyrodniczy w Lublinie

ul. Doświadczalna 44

20-236 Lublin