

GNIEWKO NIEDBAŁA¹, MAGDALENA PIEKUTOWSKA², TOMASZ PISKIER²,
MARIUSZ ADAMSKI¹, ROBERT KORZENIEWICZ³

¹Inżynierii Biosystemów
Uniwersytet Przyrodniczy w Poznaniu

²Katedra Agrobiotechnologii
Politechnika Koszalińska

³Katedra Hodowli Lasu
Uniwersytet Przyrodniczy w Poznaniu

ZASTOSOWANIE METODY REGRESJI WIELORAKIEJ DO PREDYKCJI CECH JAKOŚCIOWYCH BULW ZIEMNIAKA ODMIANY ‘INNOVATOR’

APPLICATION OF THE MULTIPLE REGRESSION METHOD
TO PREDICT THE QUALITY CHARACTERISTICS
OF POTATO TUBERS OF THE INNOVATOR VARIETY

Abstrakt

Wstęp. Ziemniak jest gatunkiem powszechnym w uprawie, szczególnie na kontynencie europejskim i azjatyckim. Szacuje się, że produkcja ziemniaka ($t \cdot ha^{-1}$) w Azji i Europie stanowi 85% produkcji światowej. Produkcja ziemniaków na frytki różni się od uprawy ziemniaków jadalnych czy skrobiowych. Jednym z najważniejszych czynników determinujących jakość surowca jest odmiana. Producenci uprawiający ziemniaki z przeznaczeniem na frytki zazwyczaj wybierają odmiany pochodzące z europejskich hodowli. Ważnym aspektem w wykonywaniu specjalistycznej produkcji rolniczej jest możliwość prognozowania spodziewanych plonów w ujęciu nie tylko ilościowym, ale także jakościowym. Na podstawie danych o tempie przyrostu suchej masy w bulwach w trakcie sezonu wegetacyjnego rolnicy mogą z wyprzedzeniem zaplanować termin defoliacji czy zbioru. Celem niniejszej pracy było opracowanie modelu regresyjnego prognozującego parametr: masę bulw pod wodą (UWW – ang. *underwater weight*) dla odmiany ziemniaka ‘Innovator’. Znając jego wartość, można łatwo określić zawartość suchej masy i skrobi w bulwach.

Materiał i metody. Do budowy modelu regresyjnego prognozującego UWW odmiany ‘Innovator’ wykorzystano dane z pól produkcyjnych gospodarstwa indywidualnego pochodzące z lat 2011–2017. Wszystkie pola były zlokalizowane w północnej Polsce, w powiatach słupskim i sławieńskim (województwo pomorskie i zachodniopomorskie). Holenderska odmiana ziemniaka ‘Innovator’ (hodowca HZPC) jest jedną z najbardziej popularnych odmian przeznaczanych do

produkcji frytek. W każdym roku uprawy w/w odmiany w omawianym gospodarstwie liczba pól oraz powierzchnia obsadzona ziemniakiem była różna. Rokiem o najmniejszej skali produkcji był rok 2012 – trzy pola o łącznej powierzchni 204 ha, z kolei najwięcej ziemniaków uprawiano w roku 2017 – pięć pól o łącznej powierzchni 344 ha. Do budowy modelu regresyjnego, będącego przedmiotem niniejszej pracy, wykorzystano dane pochodzące z ostatnich, wrześniowych próbkowań w latach 2011–2017. Łącznie użyto do badań wyniki z 82 prób pobranych z pola. Zbiór do budowy modelu regresyjnego, nazwany zbiorem I, składał się z 75 prób. Zbiór II, który tworzyło 7 losowo wybranych prób, pełnił funkcję walidacyjną i nie uczestniczył w budowie modelu. Struktura modelu bazuje na ośmiu cechach niezależnych – danych meteorologicznych i poziomach nawożenia mineralnego.

Wyniki. Wytworzony model regresyjny charakteryzuje się współczynnikiem determinacji $R^2 = 0,6623$. Na poziomie istotności statystycznej $\alpha = 0,05$ wyznaczono cztery cechy niezależne w największym stopniu wpływające na UWW. Są to: średnia temperatura powietrza od 1 kwietnia do x-dnia we wrześniu, w którym pobrano próbkę (T4-9), suma nawożenia N w roku bieżącym (N), suma nawożenia SO_3 w roku bieżącym (SO_3) oraz suma nawożenia MgO w roku bieżącym (MgO). Do walidacji modelu zastosowano cztery mierniki błędów *ex post*, tj. globalny względny błąd aproksymacji modelu (RAE = 0,0382), błąd średniokwadratowy (RMS = 22,5781), błąd średni bezwzględny (MAE = 14,5978), błąd średni bezwzględny procentowy (MAPE = 3,823).

Wnioski. Wytworzony model regresyjny charakteryzuje się małą wartością błędu średniego bezwzględnego procentowego (MAPE) na poziomie 3,823%. Oznacza to, że może on być stosowany w praktyce rolniczej. Czynnikiem o najwyższej wadze, wpływającym na suchą masę bulw ziemniaków na podstawie masy bulw pod wodą (Y_UWW), jest średnia dobowa temperatura (T4-9) w okresie do 1 kwietnia do x września.

Słowa kluczowe: predykcja, model regresyjny, ziemniak, 'Innovator', sucha masa

Wstęp

Ziemniak jest gatunkiem powszechnym w uprawie głównie na kontynencie europejskim i azjatyckim. Szacuje się, że jego produkcja w Azji i Europie ($t \cdot ha^{-1}$) stanowi 85% produkcji światowej, z wyraźnym jej wzrostem w Azji, a spadkiem w Europie (Dzwonkowski, 2012). Według danych FAOSTAT w 2016 roku krajem europejskim o największej produkcji ziemniaków była Rosja. Wyprodukowano tam ponad 31 mln ton ziemniaków. Polska znalazła się na czwartym miejscu z produkcją sięgającą prawie 9 mln ton (<https://www.potatopro.com/europe/potato-statistics>).

W Polsce redukcja skali uprawy ziemniaka rozpoczęła się około roku 1980 i trwała aż do wejścia naszego kraju do struktur UE. Bezpośrednią przyczyną spadku popularności ziemniaka jest mała opłacalność produkcji tego gatunku w relacji do innych roślin oraz ograniczenie jego wielostronnego użytkowania w żywieniu zwierząt czy gorzelnictwie (Nowacki, 2012). Obecnie rynek charakteryzuje pewna stabilizacja i specjalizacja. Zauważyć można dynamiczny rozwój tzw. „ziemniaczanej żywności wygodnej”, która spełnia zmieniające się wymagania konsumentów pod względem łatwości przygotowania, jakości i wartości odżywczej. Z roku na rok wzrasta także spożycie i eksport smażonych produktów ziemniaczanych, takich jak frytki i chipsy.

Produkcja ziemniaków na frytki różni się od uprawy ziemniaków jadalnych czy skrobiowych. Jednym z najważniejszych czynników determinujących powodzenie uprawy jest odmiana. Producenci uprawiający ziemniaki z przeznaczeniem na frytki

zazwyczaj wybierają odmiany pochodzące z zagranicznych hodowli. Wiele zakładów przetwórstwa ziemniaczanego kontraktuje wyłącznie te odmiany, które pozytywnie przeszły kilkuletnie doświadczenia polowe oraz próbne testy na linii produkcyjnej. Ważne jest, żeby dana odmiana była łatwa w uprawie i przechowywaniu, a zarazem dawała produkt spełniający oczekiwania konsumentów. Inne różnice wynikające głównie z stosowanej agrotechniki to: preferowana 90-centymetrowa szerokość międzyrzędzi, wysokie nawożenie mineralne, dobra ochrona plantacji przed chorobami i szkodnikami w trakcie sezonu wegetacyjnego, technologia zbioru i przechowywania (Erlichowski i in., 2017). Najważniejszym parametrem decydującym zarówno o zysku rolnika, jak i skali produkcji finalnego produktu, jest plon. Podczas realizacji dostaw surowca szczególną uwagę zwraca się także na inne cechy surowca, tj. długość ziemniaków, zawartość suchej masy, uszkodzenia zewnętrzne i wewnętrzne, obicia (Carputo i in., 2005; Geremew i in., 2007; Krzysztofik i Skonieczny, 2009; Singh i in., 2009; Spyrka, 2013; Zgórska, 2012). Zawartość suchej masy w bulwach to bardzo ważny parametr z punktu widzenia jakości otrzymywanych frytek. Jeżeli jest zbyt niska, frytki tracą właściwą strukturę oraz chrupkość i chłoną więcej oleju podczas smażenia (Gegov i in., 2007). Szacuje się, że ziemniaki przeznaczone do produkcji frytek powinny zawierać 20–22% suchej masy bulw (Zgórska, 2005). Inne źródła wskazują na szerszy zakres: od 20% do 24 % (Kabira i Lemaga, 2006).

Ważnym aspektem w wykonywaniu specjalistycznej produkcji rolniczej jest możliwość prognozowania spodziewanych plonów w ujęciu nie tylko ilościowym, ale także jakościowym. Na podstawie danych o tempie przyrostu suchej masy w bulwach w trakcie sezonu wegetacyjnego rolnicy mogą z wyprzedzeniem zaplanować termin defoliacji czy wykopków. Dzięki szacunkowym prognozom plonu bulw i zawartości suchej masy zakład przetwórczy może kontrolować własne możliwości przerobowe oraz decydować o rodzaju produktu, jaki będzie w stanie wyprodukować z danej dostawy ziemniaków. Do prognozowania zjawisk w rolnictwie wykorzystuje się liniowe i nieliniowe metody modelowania (Gonzales-Sanchez i in., 2014). Modele liniowe są z powodzeniem stosowane w praktyce rolniczej. Obecnie wykonuje się intensywne badania nad wykorzystaniem prostych modeli w codziennym użyciu (Diepenbrock, 2000). Celem niniejszej pracy było uzyskanie modelu regresyjnego prognozującego parametr masę bulw pod wodą (UWW – ang. *underwater weight*) dla odmiany ziemniaka 'Innovator'. Znając jego wartość, można łatwo określić zawartość suchej masy i skrobi w bulwach.

Material i metody

Do budowy modelu regresyjnego prognozującego plon bulw odmiany 'Innovator' wykorzystano dane pochodzące z pól produkcyjnych gospodarstwa indywidualnego z lat 2011–2017. Wszystkie pola były zlokalizowane w północnej Polsce, w powiatach słupskim i sławieńskim (województwo pomorskie i zachodniopomorskie). Holenderska odmiana ziemniaka 'Innovator' (hodowca HZPC) jest jedną z najbardziej popularnych odmian przeznaczanych do produkcji frytek. W każdym roku uprawy w/w odmiany w omawianym gospodarstwie liczba pól oraz powierzchnia obsadzona ziemniakiem była różna. Rokiem o najmniejszej powierzchni produkcji był rok 2012 – trzy pola

o łącznej powierzchni 204 ha, z kolei o największej – rok 2017 (pięć pól o łącznej powierzchni 344 ha).

Podstawą do określania plonu bulw ziemniaka były próby pobierane od pierwszej dekady lipca do połowy września, w 10–14-dniowych odstępach czasowych, w każdym roku uprawy i ze wszystkich pól produkcyjnych. Regularne próbkowanie umożliwiło bieżącą kontrolę przyrostu plonu, zawartości suchej masy w bulwach w okresie od końca kwitnienia do dojrzałości pełnej bulw. Plon finalny został najlepiej zobrazowany przez ostatnie próbkowanie. Pojedynczą próbę stanowiły bulwy zebrane z powierzchni 3 m², oddzielnie dla każdego z pól. Pierwszą próbę w pierwszej dekadzie lipca pobierano z wybranego miejsca na polu (min. 15 m od brzegu pola), które reprezentowało średnie warunki dla danej parceli. W kolejnych terminach próby bulw kopano w najbliższym sąsiedztwie próby pobranej w pierwszym terminie. Zdarzało się tak, że na niektórych polach za każdym razem pobieraną większą ilość prób. Wynikało to z faktu, że 25 hektarów uprawy reprezentowała jedna próba ziemniaków pobranych z konkretnego pola w dowolnym terminie. Sadzenie ziemniaków poprzedzała dokładna i głęboka uprawa gleby. W omawianym gospodarstwie praktykowano uprawę w szerokiej rozstawie wynoszącej 90 cm. Rozszerzenie międzyrzędzi w uprawie ziemniaka m.in. stymuluje wzrost bulw na długość.

Wszystkie próby ziemniaków przewożono do laboratorium oceny surowca zakładu przetwórczego, gdzie oznaczano: plon bulw o kalibrze większym niż 28 mm, procentowy udział frakcji bulw 28–40 mm, 40–50 mm, powyżej 50 mm w plonie ogólnym, długość bulw, udział defektów wewnętrznych i zewnętrznych oraz masę bulw pod wodą (UWW). Wykorzystane na potrzeby realizacji omawianego modelu regresyjnego dane dotyczące sumarycznych dawek nawożenia podstawowymi składnikami mineralnymi i wartości UWW pochodziły z bazy danych zakładu przetwórczego. Do określenia masy bulw pod wodą posłużono się metodą wagi hydrostatycznej (Fortuna, 2003), wykorzystując ziemniaki o kalibrze 50+ w jednym powtórzeniu dla każdej próby z pola.

Model regresyjny prognozujący masę bulw pod wodą wzbogacono o podstawowe dane meteorologiczne pochodzące z zasobów archiwalnych Instytutu Meteorologii i Gospodarki Wodnej – Państwowego Instytutu Badawczego dla stacji synoptycznej w Koszalinie. Regresyjny model UWW uwzględnia czynniki meteorologiczne (średnia dobową temperaturę powietrza, suma opadów) zebrane od 1 kwietnia do x września, tj. dnia ostatniego próbkowania w każdym sezonie wegetacyjnym.

Do budowy modelu regresyjnego, będącego przedmiotem niniejszej pracy, wykorzystano dane pochodzące z ostatnich, wrześniowych próbkowań w latach 2011–2017. Łącznie pracowano na wynikach z 82 prób pobranych z pola. Zbiór do budowy modelu regresyjnego, nazwany zbiorem I, składał się z 75 prób. Zbiór II, który tworzyło 7 losowo wybranych prób, pełnił funkcję walidacyjną i nie uczestniczył w budowie modelu. Szczegółowe specyfikacje danych przedstawiono w tabeli 1.

Regresja wielokrotna to metoda statystyczna, której głównym celem jest ilościowe ujęcie związków pomiędzy wieloma zmiennymi niezależnymi a zmienną zależną. Nawet jeśli nie istnieje sensowna zależność pomiędzy zmiennymi, można dążyć do powiązania ich za pomocą równania matematycznego. Równanie to może nie posiadać sensu fizycznego, ale przy pewnych założeniach pozwala na prognozowanie wielkości określonych na podstawie znajomości innych zmiennych (Trzpieciński, 2010).

Niedbała, G., Piekutowska, M., Piskier, T., Adamski, M., Korzeniewicz, R. (2018). Zastosowanie metody regresji wielorakiej do predykcji cech jakościowych bulw ziemniaka odmiany 'Innovator'. *Nauka Przyr. Technol.*, 12, 3, 285–296. <http://dx.doi.org/10.17306/J.NPT.00260>

Tabela 1. Struktura danych w predykcyjnym modelu regresyjnym

Symbol	Jednostka	Nazwa cechy	Zakres danych
T4-9	°C	średnia temperatura powietrza od 1 kwietnia do dnia września, w którym pobrano próbkę	14,07–15,31
R4-9	mm	suma opadów od 1 kwietnia do dnia września, w którym pobrano próbkę	288,3–497,6
I4-9	h	suma nasłonecznienia od 1 kwietnia do dnia września, w którym pobrano próbkę	1012,5–1371,6
N	kg·ha ⁻¹	suma nawożenia N w roku bieżącym	147–244
P ₂ O ₅	kg·ha ⁻¹	suma nawożenia P ₂ O ₅ w roku bieżącym	138–161
K ₂ O	kg·ha ⁻¹	suma nawożenia K ₂ O w roku bieżącym	210–300
SO ₃	kg·ha ⁻¹	suma nawożenia SO ₃ w roku bieżącym	252–336
MgO	kg·ha ⁻¹	suma nawożenia MgO w roku bieżącym	60–123
Y_UWW	g	masa bulw pod wodą (UWW)	327–444

Regresję wieloraką poprzedza się badaniem współczynnika determinacji R^2 dla badanych cech. Służy on do oceny stopnia wyjaśnienia całkowitej zmienności zmiennej zależnej przez zmienną niezależną. Jest on równy kwadratowi współczynnika korelacji wielorakiej między badanymi cechami. Kontynuacją analizy regresji jest określenie współczynnika prawdopodobieństwa dla statystyki bezwzględnego t , weryfikowane na poziomie istotności $\alpha = 0,05$ (różnica statystycznie istotna). W końcowej fazie tego etapu buduje się równanie regresji w postaci:

$$Y = a + b1X1 + b2X2 + \dots + bpXp \quad (1)$$

gdzie:

Y – zmienna zależna (badana cecha),

a – wyraz wolny,

Xp – wartość zmiennej niezależnej,

bp – współczynnik regresji.

Równanie (1) przedstawia model regresyjny dla prognozowanej cechy – masy bulw pod wodą (UWW).

Metodyka oceny wytworzonego modelu

Oceny zdolności predykcyjnej wytworzonego modelu dokonuje się za pomocą mierników błędu prognozy (*ex post*), porównując dane ze zbioru II do wyników predykcji wytworzonych na bazie zbioru I. Błędy te cechują się tym, że są obliczane na

podstawie danych historycznych, a więc informacji o prognozach już wygasłych i odpowiadającej im realizacji zmiennej prognozowanej. Błąd prognozy stanowi różnicę między realizacją zmiennej prognozowanej w czasie a prognozą realizowaną na ten sam okres (Stańko, 2013).

Walidację wytworzonych modeli wykonano, bazując na danych z lat 2011–2017, z siedmiu losowo wybranych pól. Danych tych nie brano pod uwagę podczas budowy modelu. Przy ocenie jakości prognoz zastosowano metodykę szeroko opisaną w literaturze (Emamgholizadeh i in., 2015; Kantanantha i in., 2010; Li i in., 2016; Logan i in., 2016; Niażian i in., 2018; Stańko, 2013)

- RAE – globalny względny błąd aproksymacji modelu:

$$\text{RAE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i)^2}} \quad (2)$$

- RMSE – błąd średniokwadratowy:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{n}} \quad (3)$$

- MAE – błąd średni bezwzględny

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \bar{y}_i| \quad (4)$$

- MAPE – błąd średni bezwzględny procentowy:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \bar{y}_i}{y_i} \right| \cdot 100\% \quad (5)$$

gdzie:

- n – liczba obserwacji,
- y_i – wartości rzeczywiste uzyskane podczas badań,
- \hat{y}_i – wartości wyznaczone za pomocą modelu.

W kolejnym kroku tworzy się wykres obrazujący wzajemne relacje oraz wyznaczone jest równanie liniowe. Zabiegi te obrazują relacje pomiędzy UWW rzeczywistym a prognozowanym przez model.

Wyniki i dyskusja

Wytworzony model regresyjny bazował na ośmiu cechach niezależnych, które zostały przedstawione w tabeli 1. Cechą zależną była masa bulw pod wodą (UWW). Szczegółowe wyniki analizy regresji wielorakiej dla prezentowanych cech niezależnych i cechy zależnej przedstawiono w tabeli 2.

Tabela 2. Współczynniki regresji, błędy standardowe oraz poziomy prawdopodobieństwa dla wytworzonego modelu

Cecha	UWW: R = 0,8138, R ² = 0,6623, Wyraz wolny = 1228,393		
	b	p	istotność
T4-9	-47,038	0,00018	*
R4-9	0,116	0,07433	–
I4-9	-0,018	0,64336	–
N	1,067	0,00259	*
P ₂ O ₅	-0,895	0,39527	–
K ₂ O	-0,078	0,55164	–
SO ₃	-1,041	0,00000	*
MgO	0,702	0,00096	*

Oznaczenie poziomu istotności statystycznej: „–” – nieistotne, * – istotne dla $\alpha = 0,05$.

Na podstawie powyższych wyników, zbudowano równanie regresji wielorakiej, które przyjęło postać:

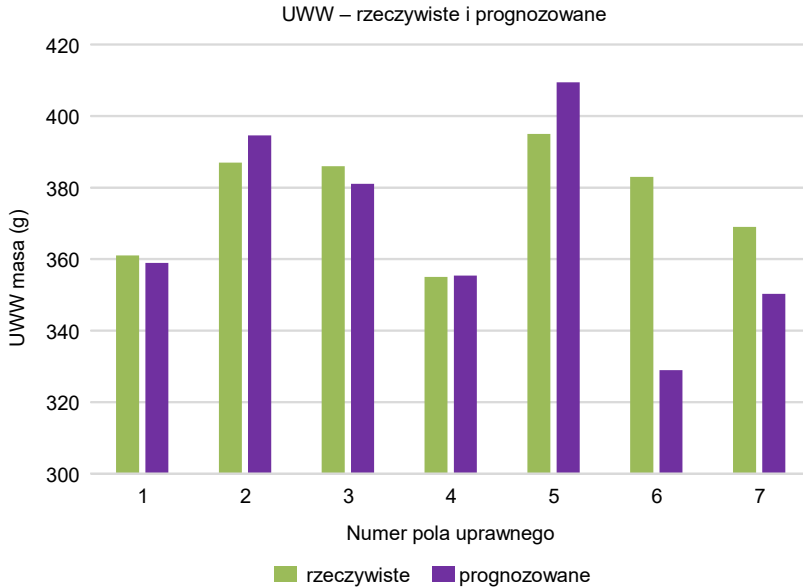
$$UWW = 1228,393 - 47,038 \cdot T4-9 + 0,116 \cdot R4-9 - 0,018 \cdot I4-9 + 1,067 \cdot N - 0,895 \cdot P_{2O_5} - 0,078 \cdot K_2O - 1,041 \cdot SO_3 + 0,702 \cdot MgO \quad (6)$$

W celu określenia jakości prognozy, obliczenia zastosowane dla metod *ex post* zostały przeprowadzone przy użyciu wzorów (2–5), a wyniki przedstawiono w tabeli 3.

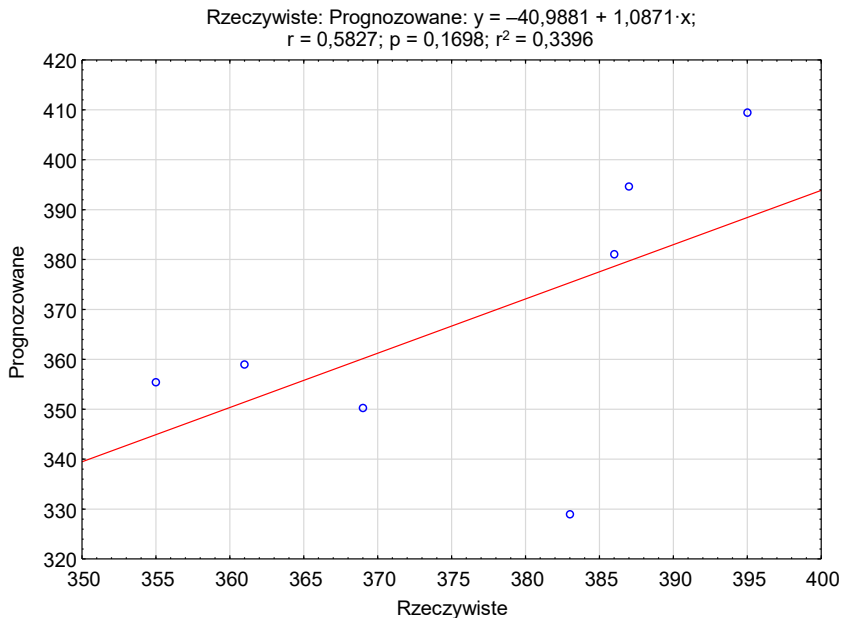
Tabela 3. Mierniki predykcyjne *ex post* w analizowanym modelu regresyjnym

RAE (–)	RMS (t·ha ⁻¹)	MAE (t·ha ⁻¹)	MAPE (%)
0,0382	22,5781	14,5978	3,823

W kolejnym kroku utworzono wykres zależności między obserwowanymi i prognozowanymi przez model wartościami UWW oraz na podstawie równania liniowego (6) wyznaczono relacje pomiędzy uzyskanymi wynikami (ryc. 1 i 2).



Ryc. 1. Rzeczywiste i prognozowane przez model wartości UWW



Ryc. 2. Relacja pomiędzy UWW rzeczywistym i prognozowanym wraz z równaniem liniowym

Czynnikami wpływającymi na gromadzenie suchej masy i skrobi w bulwach są odmiana, agrotechnika oraz lokalne czynniki siedliskowo-glebowe (Kołodziejczyk, 2014). Ponadto istotną rolę odgrywają także warunki meteorologiczne panujące w danym sezonie wegetacji ziemniaka, które kształtują cechy jakościowe plonu (Bombik i in., 2007). Wyniki analiz zawarte w tabeli 2 dla modelu regresyjnego zbudowanego na podstawie danych empirycznych ze zbioru I, częściowo potwierdzają wyniki uzyskane w innych opracowaniach. W niniejszej pracy czynnikiem o najwyższej wadze jest średnia dobową temperatura (T4-9) w okresie 1 kwietnia–x września (data pobrania ostatniej próby bulw z pola). Gromadzeniu suchej masy w bulwach sprzyja ciepła, słoneczna pogoda przypadająca na końcowy etap wegetacji ziemniaków (Wierzbička, 2012). Optymalna temperatura do uprawy ziemniaka waha się między 15–25°C w ciągu dnia i nie przekracza 12°C w nocy. Zwiększona długość dnia w połączeniu z wysokimi temperaturami powietrza warunkują wzrost zawartości cukrów redukujących w bulwach – cechy niepożądaną w przetwórstwie ziemniaczanym.

Innymi cechami niezależnymi o znacznym wpływie na wyjaśnienie zmienności UWW bulw odmiany 'Innovator' jest nawożenie składnikami mineralnymi. W przedstawionych w pracy analizach wykazano istotność statystyczną azotu (N), siarki (SO₃) i magnezu (MgO). Jednak pomimo istotności statystycznej tych cech na poziomie $\alpha = 0,05$, nie wniosły one do równania regresji (6) istotnego wkładu. Ich współczynnik „b” wahał się w zakresie od –1,041 do 1,067. W porównaniu do współczynnika „b” dla średniej dobowej temperatury (T4-9), który wyniósł –47,038, różnica wartości jest znacząca.

Model regresyjny został poddany walidacji na podstawie losowo wybranych danych polowych z siedmiu lat. Zastosowano cztery mierniki błędów *ex post*, tj. globalny względny błąd aproksymacji modelu (RAE), błąd średniokwadratowy (RMS), błąd średni bezwzględny (MAE), błąd średni bezwzględny procentowy (MAPE). Mierniki te miały zastosowanie w określeniu jakości modeli i wyznaczenia błędów prognozy suchej masy bulw ziemniaków na podstawie masy bulw pod wodą (Y_UWW). W tabeli 3 przedstawiono wartości błędów *ex post* dla wytworzonego modelu. Do najczęściej używanych wskaźników charakteryzujących wartości błędów predykcji zalicza się MAPE (Kantanantha i in., 2010; Niedbała i in., 2007; Zhang i in., 1998). Wartość błędu prognozy MAPE dla modelu regresyjnego wyniosła 3,823%. Mając na uwadze krytyczny poziom błędu MAPE na poziomie do 10%, w przypadkach, które pozostają pod znacznym wpływem warunków losowych (Stańko, 2013), uzyskany wynik jest bardzo dobry.

Na rycinie 1 przedstawiono graficzną formę relacji pomiędzy wartościami UWW rzeczywistymi i prognozowanymi przez model. Porównaniem objęto siedem pól, które w jedynym przypadku pola numer 6 znacznie odbiega od innych wartości. Pole to obejmowało dane z roku 2016, na którym zanotowano znaczną ilość opadów atmosferycznych oraz wysokie dawki nawożenia siarką i magnezem. Prawdopodobnie z tych przyczyn wartości prognozowanego UWW są zdecydowanie niższe niż wartości rzeczywiste. Z kolei na rycinie 2 relacje pomiędzy obydwojma zbiorami zostały dodatkowo opisane równaniem liniowym, oraz został wyznaczony współczynnik determinacji $R^2 = 0,3396$. Tak niska wartość współczynnika R^2 jest nieadekwatna do bardzo niskiej wartości MAPE na poziomie 3,823%. Z danych przedstawionych na rycinie 2 można zaobserwować, że tak duże różnice związane są także z parametrami pogodowo-nawozowymi występującymi na polu numer 6.

Wnioski

1. Wytworzony model regresyjny charakteryzuje się niską wartością błędu średniego bezwzględnego procentowego (MAPE) na poziomie 3,823%. Oznacza to, że może on być stosowany w praktyce rolniczej.

2. Czynnikiem o najwyższej wadze wpływająca na suchą masę bulw ziemniaków na podstawie masy bulw pod wodą (Y_{UWW}) jest średnia dobowa temperatura ($T4-9$) w okresie od 1 kwietnia do x września.

3. Należy podjąć dalsze badania poszerzone o większą liczbę pól w celu zwiększenia dokładności modelu i ustaleniu istotności wagowej poszczególnych cech niezależnych.

4. Dalsze badania należy skierować także na inne odmiany uprawiane na cele frytkowe w celu głębszej analizy danych.

Bibliografia

- Bombik, A., Rymusza, K., Markowska, M., Stankiewicz, C. (2007). Variability analysis of selected quantitative characteristics in edible potato varieties. *Acta Sci. Pol. Agric.*, 6(3), 5–15.
- Carputo, D., Aversano, R., Frusciante, L. (2005). Breeding potato for quality traits. *Acta Hort.*, 684, 55–64.
- Diepenbrock, W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L.): A review. *F. Crop. Res.*, 67, 35–49.
- Dzwonkowski, W. (2012). Międzynarodowy rynek a eksport ziemniaków z Polski. W: J. Chotkowski (red.), *Produkcja i rynek ziemniaka* (s. 7–20). Warszawa: Wieś Jutra.
- Emamgholizadeh, S., Parsaeian, M., Baradaran, M. (2015). Seed yield prediction of sesame using artificial neural network. *Eur. J. of Agron.*, 68, 89–96. <https://doi.org/10.1016/j.eja.2015.04.010>
- Erlichowski, T., Urbanowicz, J., Osowski, J. (2017). Intensywna produkcja ziemniaków na różne kierunki użytkowania na przykładzie dwóch specjalistycznych gospodarstw w woj. pomorskim. *Ziemniak Polski*, 3, 49–56.
- Fortuna, T. (2003). *Podstawy analizy żywności*. Kraków: Wyd. Akad. Roln.
- Gegov, Y., Pevicharova, G., Nacheva, E., Slavchev, V. (2007). Potato breeding lines suitable for production of frozen french fries. *Bulg. J. Agric. Sci.*, 13, 15–29.
- Geremew, E. B., Steyn, J. M., Annandale, J. G. (2007). Evaluation of growth performance and dry matter partitioning of four processing potato (*Solanum tuberosum*) cultivars. *New Zeal. J. Crop Hort. Sci.*, 35: 385–393.
- Gonzales-Sanchez, A., Frausto-Solis, J., Ojeda-Bustamante, W. (2014). Attribute selection impact on linear and nonlinear regression models for crop yield prediction. *Sci. World J.* <http://dx.doi.org/10.1155/2014/509429>
- Kabira, J., Lemaga, L. (2006). Potato processing quality evaluation procedure for research and food industry applications in east and central Africa. *Kenya Agric. Res. Insti. Nairobi.*, 10.
- Kantanatha, N., Serban, N., Griffin, P. (2010). Yield and price forecasting for stochastic crop decision planning. *J. Agr. Biol. and Envir. St.*, 15(3), 362–380. <https://doi.org/10.1007/s13253-010-0025-7>
- Kołodziejczyk, M. (2014). Wpływ warunków opadowo-termicznych na skład chemiczny oraz wybrane parametry jakości bulw średnio późnych i późnych odmian ziemniaka jadalnego. *Ann. UMCS- Polonia, Sect. E*, 69(3), 1–10.
- Krzysztofik, B., Skonieczny, P. (2009). Wpływ kształtu bulw ziemniaka na wydajność frytki surowej. *Inż. Roln.*, 5, 114, 131–138.

Niedbała, G., Piekutowska, M., Piskier, T., Adamski, M., Korzeniewicz, R. (2018). Zastosowanie metody regresji wielorakiej do predykcji cech jakościowych bulw ziemniaka odmiany 'Innovator'. *Nauka Przyr. Technol.*, 12, 3, 285–296. <http://dx.doi.org/10.17306/J.NPT.00260>

- Li, F., Qiao, J., Han, H., Yang, C. (2016). A self-organizing cascade neural network with random weights for nonlinear system modeling. *Appl. Soft Comput.*, 42, 184–193. <http://dx.doi.org/10.1016/j.asoc.2016.01.028>
- Logan, T. M., McLeod, S. i Guikema, S. (2016). Predictive models in horticulture: A case study with Royal Gala apples. *Sci. Hortic.*, 209, 201–213. <https://doi.org/10.1016/j.scienta.2016.06.033>
- Niazian, M., Sadat-Noori, S. A. i Abdipour, M. (2018). Artificial neural network and multiple regression analysis models to predict essential oil content of ajowan (*Carum copticum* L.). *J. Appl. Res. Med. Arom. Plants*, 9 (April), 124–131. <https://doi.org/10.1016/j.jarmap.2018.04.001>
- Niedbała, G., Przybył, J., Sęk, T. (2007). Prognosis of the content of sugar in the roots of sugar-beet with utilization of the regression and neural techniques. *Agric. Eng.*, 2(90), 225–234.
- Nowacki W. (2012). O kierunkach zmian w uprawie ziemniaka w Polsce. *Biuletyn Instytutu Hodowli i Aklimatyzacji Roślin*, 266, 21–35.
- Singh, S. V., Marwaha, R. S., Kumar, D., Kumar, P., Pandey, S. K. (2009). Suitability of potato varieties grown in north-eastern Indian plains for processing. *Potato J.*, 36, 25–34.
- Spyrka, B. W. (2013). Wpływ warunków siedliskowych na powstawanie i wielkość uszkodzeń mechanicznych bulw ziemniaka dla przetwórstwa spożywczego. Praca doktorska. Wrocław: UP.
- Stańko, S. (2013). Prognozowanie w agrobiznesie. Teoria i przykłady zastosowania. Wyd. I. Warszawa: Wydawnictwo SGGW.
- Trzpieciński, T. (2010). Zastosowanie regresji wielokrotnej i sieci neuronowej do modelowania zjawiska tarcia. *Zesz. Nauk. WSI*, 9(3), 31–43.
- Wierzbicka, A. (2012). Wpływ odmiany, nawożenia azotem i terminu zbioru na zawartość suchej masy i skrobi w bulwach ziemniaków wczesnych. *Fragm. Agron.*, 29(2), 134–142
- Zgórska, K. (2005). Kryteria oceny ziemniaka i jego przetworów w programie „PDZ – poznaj dobrą żywność”. Warszawa: Min. Rol. i Rozwoju Wsi, Produkcja Roślinna. Ziemniak. Poznaj Dobrą Żywność, [www.minrol.gov.pl].
- Zgórska, K. (2012). Ziemniak- surowiec do produkcji żywności wygodnej i minimalnie przetworzonej. W: J. Chotkowski (red.). *Produkcja i rynek ziemniaka* (s. 324–333). Warszawa: Wieś Jutra.
- Zhang, G. P., Patuwo, E. B., Michael, Y. H. (1998). Forecasting with artificial neural networks: The state of the art. *Int. J. Forecasting*, 14(1), 35–62. [https://doi.org/10.1016/S0169-2070\(97\)00044-7](https://doi.org/10.1016/S0169-2070(97)00044-7)

Źródła internetowe:

<https://www.potatopro.com/europe/potato-statistics>, dostęp: 22.12.2018, 21:00

APPLICATION OF THE MULTIPLE REGRESSION METHOD TO PREDICT THE QUALITY CHARACTERISTICS OF POTATO TUBERS OF THE INNOVATOR VARIETY

Abstract

Introduction. The potato is a common species in cultivation, mainly on the European and Asian continents. It is estimated that total production in Asia and Europe accounts for 85% of world production. Potato production for French fries differs from the cultivation of ware potatoes or starch potatoes. One of the most important factors determining the success of cultivation is variety. Potato growers who grow potatoes for French fries usually choose foreign varieties. An important aspect of specialist agricultural production is the possibility of forecasting the expected

yields in terms of quantity, but also quality. On the basis of data on the rate of dry matter growth in tubers during the growing season, farmers can plan in advance the date of defoliation or excavation. The aim of this study was to obtain a regression model predicting the parameter of tuber weight underwater (UWW – underwater weight) for Innovator potato variety. Knowing its value it is easy to determine the dry matter and starch content of tubers.

Materials and methods. For the construction of the regression model forecasting UWW Innovator cultivar, data from the production fields of the individual farm from 2011–2017 were used. All fields were located in the northern Poland, in the Słupski and Sławieński districts (Pomeranian and West Pomeranian Voivodeships). The Dutch potato variety Innovator (HZPC cultivator) is one of the most popular varieties intended for French fries. In each year of cultivation of the above mentioned cultivar, the number of fields and the area planted with potatoes was different in the farm in question. The year with the smallest scale of production was the year 2012 – 3 fields, 204 ha, while the highest number of potatoes was cultivated in 2017 – 5 fields, 344 ha. In order to build a regression model, which is the subject of this study, data from the last, September samples in the years 2011–2017 were used. In total, results from 82 samples taken from the field were worked on. The set for the construction of the regression model, called set I, consisted of 75 samples. Set II, which consisted of 7 randomly selected samples, performed a validation function and did not participate in the construction of the model. The structure of the model is based on 8 independent features – meteorological data and mineral fertilization levels.

Results. The produced regression model is characterized by a determination coefficient $R^2 = 0.6623$. At the level of statistical significance $\alpha = 0.05$ four independent features influencing UWW to the greatest extent were determined. These are: average air temperature from 1st April to the day in September when the sample was taken (T4-9), sum of N fertilization in the current year (N), sum of SO_3 fertilization in the current year (SO_3) and sum of MgO fertilization in the current year (MgO). Four *ex-post* error measures were used for model validation, i.e. global relative approximation error of the model (RAE = 0.0382), root mean square error (RMS = 22.5781), mean absolute error (MAE = 14.5978), mean absolute percentage error (MAPE = 3.823).

Conclusions. The produced regression model is characterized by a low value of the mean absolute percentage error (MAPE) of 3.823%. This means that it can be used in agricultural practice. The factor with the highest weight affecting the dry weight of potato tubers on the basis of the weight of tubers under water (Y_{UWW}) is the average daily temperature (T4-9) in the months from 1st April to x September.

Key words: prediction, regression model, potato, Innovator, dry matter

Adres do korespondencji – Corresponding address:

Gniewko Niedbała, Instytut Inżynierii Biosystemów, Wydział Rolnictwa i Bioinżynierii, Uniwersytet Przyrodniczy w Poznaniu, ul. Wojska Polskiego 50, 60-627 Poznań, e-mail: gniewko@up.poznan.pl

Zaakceptowano do opublikowania – Accepted for publication:

06.08.2018

Do cytowania – For citation:

Niedbała, G., Piekutowska, M., Piskier, T., Adamski, M., Korzeniewicz, R. (2018). Zastosowanie metody regresji wielorakiej do predykcji cech jakościowych bulw ziemniaka odmiany 'Innovator'. *Nauka Przyr. Technol.*, 12, 3, 285–296. <http://dx.doi.org/10.17306/J.NPT.00260>