

MONIKA JANASZEK¹**DARIUSZ R. MAŃKOWSKI**²**JANUSZ KOZDÓJ**³¹ Wydział Inżynierii Produkcji, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie² Zakład Nasiennictwa i Nasionoznawstwa, IHAR — PIB Radzików³ Zakład Biotechnologii i Cytogenetyki Roślin, IHAR — PIB Radzików

Sieci neuronowe typu MLP w prognozowaniu plonu jęczmienia jarego*

MLP artificial neural networks in predicting the yield of spring barley

U roślin zbożowych, poszczególne cechy plonotwórcze oraz struktura morfologiczna kłosa kształtują się w określonych fazach rozwojowych, które zachodzą w zmiennych warunkach środowiska. Interakcja genotypu z biotycznymi i abiotycznymi czynnikami środowiska w trakcie trwania okresu wegetacji roślin również wpływa na kształtowanie się wielkości plonu. W niniejszej pracy dokonano aproksymacji wielkości plonu na podstawie cech, charakteryzujących nie tylko okres wegetacji roślin, ale także fazę ich pełnej dojrzałości. Do aproksymacji wykorzystano sieć MLP o bardzo prostej topologii, wynikającej zarówno z liczby, jak i struktury dostępnych danych. Efekt wytrenowania sieci był pozytywny. Uzyskane wyniki wyraźnie wskazują, iż sieć typu MLP może służyć jako narzędzie wspomagające prognozowanie plonu jęczmienia jarego.

Słowa kluczowe: jęczmień jary, MLP, linie DH, prognozowanie plonu, sztuczne sieci neuronowe

In cereal plants, individual yielding characteristics and the morphological structure of the spike are formed in certain phases of development, which occur in varying environmental conditions. The interaction of genotype with the biotic and abiotic environmental factors during the period of vegetation also affects the formation of yield. In this paper we approximate the yield on the basis of traits that characterize not only the vegetation period, but also the phase of full maturity. For the approximation, an MLP network with a very simple topology, resulting from both the number and structure of available data, was used. The effect of network training was positive. The results obtained show clearly that the MLP network may be used as a support tool for the prediction of the yield of spring barley.

Key words: artificial neural networks, DH lines, MLP, spring barley, yield prediction

* Praca była prezentowana w ramach I Warsztatów Biometrycznych, które odbyły się w IHAR-PIB w Radzikowie w dniach 14-15 września 2010 r.

WSTĘP

Modelowanie procesu wzrostu roślin ma wartość zarówno użytkową, jak i poznawczą. Wzrost roślin jest procesem mierzalnym, a zatem dającym się ująć matematycznie. Jest to również proces nieliniowy, który kształtuje się pod wpływem wielu czynników.

Statystyczne modele procesu wzrostu roślin powstają przede wszystkim w celu wspomagania podejmowania decyzji w gospodarce rolnej, a jednym z celów modelowania jest oszacowanie produkcji roślinnej kilka tygodni, a nawet miesięcy przed zbiorem. Precyzyjne i szybkie przewidywanie wysokości plonu jest niezmiernie pomocne w planowaniu i zarządzaniu produkcją roślinną.

Analizy porównawcze roślin, przeprowadzone w dłuższym, wieloletnim okresie badań w warunkach polowych, pozwalają na ich ocenę pod względem różnych cech użytkowych oraz wielkości i jakości plonu. Tego typu analizy są przydatne, zwłaszcza w przypadku badań prowadzących do uzyskania nowych odmian o poprawionych cechach użytkowych. Ostatecznym kryterium decydującym o przydatności nowych linii jest wielkość plonu i jego stabilność w latach. Są one uzależnione od struktury morfologicznej rośliny i jej cech plonotwórczych (Kozdój i in., 2010).

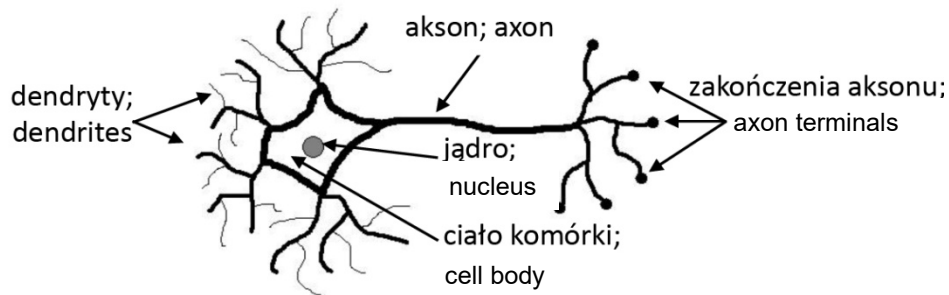
W przypadku roślin zbożowych, w tym i jęczmienia jarego, poszczególne cechy plonotwórcze oraz struktura morfologiczna kłosa kształtują się w określonych fazach rozwojowych (Klepper i in., 1998, Górny, 2004), które zachodzą w zmiennych warunkach środowiska (Łubkowski, 1968, Kozdój, 1992, 1994). Interakcja genotypu z biotycznymi i abiotycznymi czynnikami środowiska w trakcie trwania okresu wegetacji roślin uprawnych również decyduje o wielkości plonu.

Modelowanie plonu roślin nie jest zadaniem łatwym, zwłaszcza z punktu widzenia statystyki. Budowa modeli prognostycznych jest w tym wypadku często utrudniona ze względu na naturę samych danych, ich ilość, strukturę oraz występowanie interakcji między zmiennymi, które można wykorzystać do budowy tego typu modeli. Wydaje się więc uzasadnione poszukiwanie innych, bardziej wydajnych narzędzi, które sprostałyby zadaniu modelowania złożonych systemów bez jasno określonych, stałych wzorców zachowań. Praktyka pokazuje, że sieci neuronowe z powodzeniem radzą sobie z informacją zaszumioną i niekompletną, a przy tym posiadają zdolność do uogólniania zdobytej wiedzy. Z tego względu coraz częściej są one wykorzystywane również w zagadnieniach modelowania plonu roślin uprawnych (O'Neal i in., 2002; Park i in., 2005; Uno i in., 2005; Khazaei i in., 2008).

Celem niniejszej pracy był opis sztucznych sieci neuronowych jako narzędzia analizy danych i ich wykorzystania w zagadnieniach związanych z prognozowaniem plonu roślin zbożowych na podstawie informacji o podstawowych, fizjologicznych cechach plonotwórczych oraz czasie trwania poszczególnych faz rozwojowych rośliny. W pracy przedstawiono opis sztucznych sieci neuronowych — ich działania, budowy i wykorzystania. Przedstawiono również przykład możliwości wykorzystania sieci neuronowej typu MLP w prognozowaniu wysokości plonu jęczmienia jarego.

CZYM SĄ SZTUCZNE SIECI NEURONOWE?

Sztuczne sieci neuronowe naśladują swoją strukturą i działaniem układ neurologiczny człowieka. Powstały one w efekcie wieloletnich badań w dziedzinie neurofizjologii, zmierzających do opracowania matematycznego modelu działania ludzkiej komórki nerwowej. Ludzki system nerwowy zbudowany jest z przeszło 25 miliardów wzajemnie połączonych komórek nerwowych, posiadających zdolność do odbierania, przekazywania i reagowania na impulsy elektrochemiczne od blisko dziesięciu tysięcy swoich sąsiadów, a także komórek receptorowych. Uproszczony schemat budowy komórki nerwowej (neuronu) przedstawia rysunek 1.



The Encyclopedia of Science

Rys. 1. Schemat budowy ludzkiego neuronu
Fig. 1. Schematic diagram of the human neuron

Ciało pojedynczego neuronu (soma) otoczone jest wypustkami (dendrytami), które wprowadzają do niego impulsy elektrochemiczne, pochodzące od innych neuronów. Ponadto neuron posiada dokładnie jedną wypustkę wyprowadzającą impulsy (akson), dzięki której może równolegle komunikować się ze swoimi sąsiadami, jak również z otaczającymi go mięśniami. Komunikację między poszczególnymi komórkami nerwowymi zapewniają złącza nerwowe zwane synapsami. W tym obszarze akson komórki nerwowej oddzielony jest od dendrytu lub ciała innego neuronu tzw. szczeliną synaptyczną. Akson wyposażony jest w kulebki synaptyczne, które zawierają pęcherzyki wypełnione substancją o właściwościach regulujących aktywację innych komórek nerwowych (neuroprzekaźnik). Komórka nerwowa, za pośrednictwem neuroprzekaźników, odbiera od swoich sąsiadów zarówno sygnały aktywujące, jak i wyciszające, które następnie sumuje. Jeżeli wypadkowa wprowadzonych sygnałów przekroczy ustaloną wartość, zwaną progiem, wówczas neuron jest aktywowany i za pośrednictwem aksonu przekazuje sygnał innym neuronom. Aktywowane komórki nerwowe komunikują się w ten sam sposób z kolejnymi neuronami.

W podobny sposób działa sztuczna sieć neuronowa, przetwarzając dostarczone do niej sygnały wejściowe przez rzędy elementów nazywane sztucznymi neuronami. Neurony te,

podobnie jak pojedyncze komórki nerwowe, odbierają i sumują docierające do nich sygnały, przetwarzają je zgodnie z zadaną regułą i przekazują swoim sąsiadom.

Dziś sztuczne sieci neuronowe (SSN) postrzegane są jako samodzielna dziedzina wiedzy, oferująca narzędzia sterujące i decyzyjne, które realizują nie tylko zadania aproksymacji, klasyfikacji i rozpoznawania wzorców, ale również predykcji, sterowania oraz asocjacji (Osowski, 1996). SSN wykorzystywane są jako zaawansowane narzędzia analityczne, przydatne głównie w przypadku, gdy zachodzi konieczność modelowania zjawisk o silnie nieliniowym charakterze, wielowymiarowych zależności funkcyjnych, lub gdy rozpatrywany problem jest złożony i zawodzą klasyczne metody jego rozwiązania (Tompos i in., 2007). W przeciwieństwie do metod statystycznych, sieci neuronowe nie wymagają znajomości charakteru badanego zjawiska. Ich zaletą jest zdolność do samodzielnego odkrywania zależności między zmiennymi czy to przez wielokrotne prezentowanie sieci przypadków uczących, czy też na podstawie podobieństwa między przypadkami. Poprawne działanie modelu neuronowego nie wymaga wstępnych założeń dotyczących związków między zmiennymi opisującymi modelowane zjawisko, ani też znajomości rozkładu tych zmiennych (Mittal i Zhang, 2000; Seyhan i in., 2005).

OD PERCEPTRONU DO MLP

Pierwszej próby sformalizowania opisu działania i roli ludzkiego neuronu w procesie równoległego i rozproszonego przetwarzania informacji, dokonali w 1943 roku Warren McCulloch i Walter Pitts (Osowski, 2006). Działanie neuronu opisali oni prostą zależnością:

$$y = f(s), \quad (1)$$

gdzie:

$$s = w_0 + \sum_{i=0}^n x_i w_i, \quad (2)$$

w której n — liczba sygnałów wejściowych docierających do neuronu, x_i — i -ty sygnał wejściowy, w_i — waga synaptyczna i -tego sygnału wejściowego, w_0 — wartość progowa (waga, odpowiadająca progowi aktywacji $x_0=1$), f — funkcja aktywacji, s — efekt działania sumatora, odpowiadający poziomowi pobudzenia neuronu (liniowa funkcja wartości wejściowych). Zgodnie z powyższym zapisem sztuczny neuron odbiera dostarczone do niego sygnały i oblicza ich sumę ważoną. Efektem działania sumatora jest sygnał s , odzwierciedlający działanie liniowej części sztucznego neuronu. Sygnał ten jest przekształcany w wyjściowy sygnał y za pomocą funkcji aktywacji, która z reguły jest ciągłą funkcją nieliniową. Mnożenie sygnałów wejściowych przez wagi synaptyczne jest odpowiednikiem naturalnego procesu wzmacniania lub osłabiania impulsów przez synapsy. Modyfikacja wartości wag synaptycznych za pomocą odpowiedniego algorytmu umożliwi uczenie sztucznego neuronu. Funkcja aktywacji decyduje natomiast o możliwościach uczenia się neuronu lub rozpoznawania przez niego wzorców. W zależności od celu postawionego przed neuronem może ona przybrać różne postacie. W pierwszych neuronach stosowano funkcje skokowe: unipolarną (progową), przyjmującą wartości 0 lub

1 oraz funkcję bipolarną (signum), przyjmującą wartości -1 lub 1. Obecnie najczęściej wykorzystywane są funkcje:

— liniowa: $f(s) = ks \wedge k \neq 0,$ (3)

— identycznościowa: $f(s) = s,$ (4)

— tangens hiperboliczny: $f(s) = \frac{1 - \exp(-\beta s)}{1 + \exp(-\beta s)} \wedge \beta \in (0, \infty)$ (5)

— sigmoidalna: $f(s) = \frac{1}{1 + \exp(-\beta s)} \wedge \beta \in (0, \infty)$ (6)

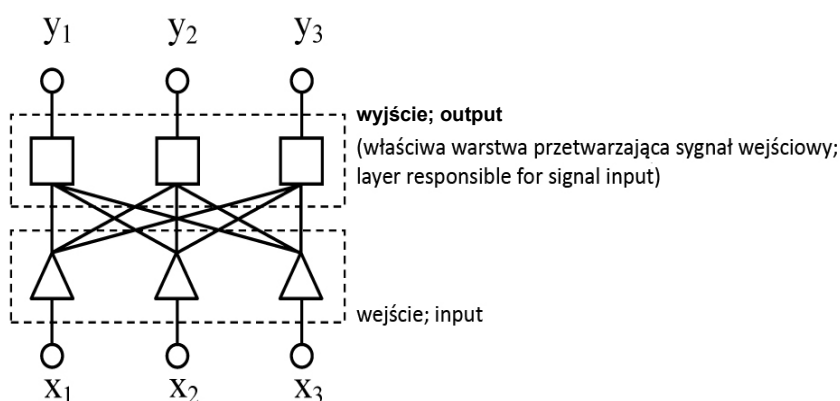
Funkcja aktywacji w modelu McCullocha i Pittsa była funkcją skokową Heaviside'a (unipolarną) w postaci:

$$H(s) = \begin{cases} 0 & \text{dla } s < 0 \\ 1 & \text{dla } s \geq 0 \end{cases} \quad (7)$$

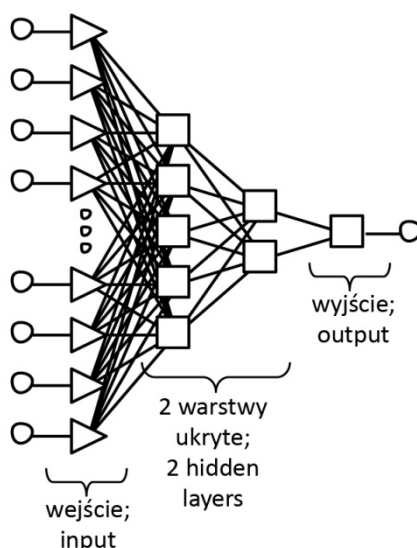
co oznaczało, że neuron był aktywowany tylko wtedy, kiedy sygnał s przyjmował wartość większą od zera. Mógł on zatem tylko rozpoznać i klasyfikować sygnały wejściowe do jednej z dwóch klas. W ogólnym przypadku, jeśli neuron posiada n wejść, wówczas przestrzeń decyzyjna jest rozdzielana za pomocą hiperpłaszczyzny o $n - 1$ wymiarach na dwie podprzestrzenie, a owa hiperpłaszczyzna nazywana jest granicą decyzyjną i opisuje ją wzór:

$$\sum_{i=1}^n w_i x_i + w_0 = 0. \quad (8)$$

Sieć neuronową o najprostszej budowie nazywamy perceptronem (Rosenblatt, 1958). Składa się ona z równolegle pracujących sztucznych neuronów, między którymi nie ma żadnych połączeń. Taka struktura bywa nazywana niekiedy perceptronem dwuwarstwowym, ponieważ dla większej przejrzystości wyodrębnia się w nim warstwę wejściową. W rzeczywistości jest to struktura jednowarstwowa, ponieważ właściwą warstwą dokonującą sumowania i obliczania wyjścia według wartości funkcji aktywacji jest tylko warstwa druga. Sygnał w tej sieci przekazywany jest tylko w jednym kierunku. Funkcja aktywacji jest taka sama dla wszystkich neuronów w warstwie, przy czym dla warstwy wejściowej jest funkcją identycznościową, a dla wyjściowej zwykle unipolarną lub bipolarną (Gatnar, 2009). Perceptron Rosenblatta może realizować zadanie klasyfikacji obiektów przez wyznaczenie liniowej funkcji dyskryminacyjnej pod warunkiem, że klasy są liniowo separowalne. O ile do klasyfikacji sygnałów (wzorców) do dwóch zbiorów wystarczy jeden neuron w warstwie wyjściowej, to klasyfikacja do k klas wymaga obecności w tej warstwie dokładnie k neuronów wyjściowych, przy czym każdy neuron reprezentuje jedną klasę (rys. 2).



Rys. 2. Perceptron, klasyfikujący obiekty charakteryzowane sygnałami x_1, x_2, x_3 do jednej z trzech klas
 Fig. 2. Perceptron, classifying objects characterized by signals x_1, x_2, x_3 to one of three classes



Rys. 3. Perceptron wielowarstwowy
 Fig. 3. Multilayer perceptron

Wzorzec jest przypisany do tej klasy, dla której neuron wyjściowy wygeneruje najsilniejszy sygnał. Siła sygnału wyjściowego zależy natomiast od zestawu wag, jaki posiada każdy neuron wyjściowy. Ograniczenia perceptronu Rosenblatta sprawiają, że do rozwiązania złożonych zadań niezbędny jest zestaw perceptronów, zorganizowanych w warstwy o różnych funkcjach aktywacji. Taką strukturę nazywamy perceptronem

wielowarstwowym (ang. multilayer perceptron) lub siecią MLP. W strukturze sieci MLP, wyraźnie wyróżnia się warstwę wejściową i wyjściową, a oprócz nich neurony ukryte, które umożliwiają modelowanie praktycznie dowolnego problemu, również nieliniowego (rys. 3).

Liczba warstw ukrytych w sieci MLP może być w zasadzie dowolna, ale nie oznacza to, że sieć z większą liczbą warstw ukrytych potrafi klasyfikować lub aproksymować w sposób bardziej dogodny lub prostszy. Udowodniono, że każda ograniczona funkcja ciągła może być aproksymowana z dowolnie małym błędem przez sieć MLP z jedną warstwą ukrytą oraz, że dowolna funkcja może być aproksymowana z dowolną dokładnością przez sieć MLP z dwoma warstwami ukrytymi (Cybenko, 1988 i 1989).

TOPOLOGIA I UCZENIE SIECI MLP

Zbudowanie i wytrenowanie sieci neuronowej nie jest zadaniem łatwym i wymaga doświadczenia. Przede wszystkim należy wybrać typ sieci neuronowej, biorąc pod uwagę charakter modelowanego zjawiska. Struktura sieci, sposób przepływu informacji między poszczególnymi jej warstwami oraz budowa i połączenie pojedynczych neuronów, a także konkretne przeznaczenie pozwalają na ogólne wyróżnienie kilku typów sieci neuronowych: jednokierunkowe (perceptrony jedno- i wielowarstwowe), rekurencyjne (sieć Hopfielda, sieć Hamminga, RTRN — ang. Real Time Recurrent Network, BAM — ang. Bidirectional Associative Memory), sieci samoorganizujące się (SOM — ang. Self Organizing Maps, sieć Heraulta-Juttana), sieci o radialnych funkcjach bazowych, sieci ART oraz sieci probabilistyczne i hybrydowe. Z uwagi na charakter poruszonego problemu w dalszej części pracy skupimy się na jednokierunkowych sieciach wielowarstwowym, czyli sieciach MLP.

Uczenie się sieci neuronowej jest zdolnością adaptacji wag zgodnie z wymaganiami, stawianymi odwzorowaniu danych wejściowych w dane wyjściowe. Proces ten zwykle przebiega cyklicznie aż do momentu osiągnięcia minimum funkcji celu. Sieci MLP są uczone w trybie nadzorowanym (z nauczycielem), co oznacza, że znane są żądane wartości wzorców, a dobór wag powinien zapewnić możliwie najlepsze dopasowanie wyjść sieci do tych wzorców. Algorytmy uczenia sieci MLP należą do grupy metod gradientowych, których stosowanie jest możliwe i efektywne jedynie w przypadku gdy funkcja aktywacji jest funkcją ciągłą. Gradientowe metody uczenia sieci neuronowych, szeroko opisane w literaturze (Tadeusiewicz, 1993; Osowski, 1996), polegają na cyklicznej aktualizacji wag sieci na podstawie informacji o gradiencie funkcji celu i wyznaczanym w każdym kroku kierunku minimalizacji. Istnieje wiele algorytmów uczenia sieci MLP, należą do nich m.in.: algorytm wstecznej propagacji błędów, algorytm największego spadku, algorytm zmiennej metryki, algorytm Levenberga-Marquardta, algorytm Gaussa-Newtona (nieliniowa metoda najmniejszych kwadratów), algorytm gradientów sprzężonych. Za najbardziej efektywne uznaje się metody oparte na algorytmie Newtona, znane pod nazwą metod zmiennej metryki lub metod quasi-Newtonowskich. Do metod tych należą: metoda BFGS (Broyden-Fletcher-Goldfarb-Shanno) oraz metoda DFP (Davidon-Fletcher-Powell). Algorytmy quasi-Newtona poszukują punktów stacjonarnych funkcji celu, przy założeniu

że może być ona aproksymowana funkcją kwadratową w otoczeniu optimum i wykorzystują pierwsze pochodne (gradient) funkcji celu do iteracyjnego przybliżania jej Hessianu (BFGS) lub jego odwrotności (DFG).

Celem uczenia sieci neuronowej jest taki dobór jej topologii i parametrów, który zapewni minimalizację błędu w wyznaczaniu wartości wyjściowej. Zbiór danych, opisujących modelowane zjawisko, powinien zostać podzielony na zbiór uczący, walidacyjny i testowy. Dobór wag odbywa się w zbiorze uczącym, ale ostatecznym celem jest minimalizacja błędu dla zbioru testowego, czyli tzw. błędu generalizacji. Na etapie uczenia sieci (w fazie „przyswajania”) wykorzystywany jest zbiór uczący oraz będący jego częścią zbiór walidacyjny, dzięki któremu możemy kontrolować przebieg uczenia przez sprawdzanie stopnia wytrenowania neuronów. W rzeczywistości uczenie obejmuje dwie fazy: dobór wag dla zbioru uczącego i testowanie wag na próbkach ze zbioru walidacyjnego. Modyfikacja wartości wag trwa tak długo aż w zbiorze uczącym nie zostanie osiągnięte kryterium aproksymacji (minimalizacja błędu aproksymacji) lub gdy błąd w zbiorze walidacyjnym nie zacznie rosnąć. W przypadku obu tych zbiorów funkcją błędu jest zazwyczaj suma kwadratów odchyleń (oznaczana jako SS lub SOS) między wartością zadaną a wyjściem sieci. Jeżeli dobór wag zostanie zakończony, sieć przechodzi w tryb „odtworzenia”, w którym wykorzystywany jest zbiór testowy. Zawiera on wartości, które wcześniej nie brały udziału w uczeniu sieci. Jeżeli struktura sieci oraz wagi zostały dobrane prawidłowo, to model powinien generować możliwie najmniejszy błąd, czyli powinno zostać osiągnięte minimum globalne funkcji celu. Walidację modelu, czyli jego ocenę pod względem zdolności odwzorowania modelowanego zjawiska, przeprowadza się na podstawie wartości globalnego błędu modelu (GE), obliczanego dla zbioru testowego:

$$GE = \sqrt{\frac{\sum_{i=1}^n (z_i - y_i)^2}{\sum_{i=1}^n z_i^2}}, \quad (9)$$

gdzie n — liczba przypadków, z — wartość zadana (wzorzec), y — odpowiedź sieci.

Ponadto, zarówno na etapie uczenia, jak i testowania można, analizować wartości błędów, które stanowią kryterium jakości modelu. Do najczęściej wykorzystywanych kryteriów należą:

— błąd średni

$$ME = \frac{1}{n} \sum_{i=1}^n (z_i - y_i) \quad (10)$$

— błąd średni bezwzględny

$$MAE = \frac{1}{n} \sum_{i=1}^n |(z_i - y_i)| \quad (11)$$

— odchylenie standardowe

$$RMS = \sqrt{\frac{\sum_{i=1}^n (z_i - y_i)^2}{n-1}} \quad (12)$$

— znormalizowane odchylenie standardowe

$$nRMS = \frac{RMS}{y_{\max} - y_{\min}} \quad (13)$$

— wariancja błędu

$$MSE = \frac{1}{n-1} \sum_{i=1}^n (z_i - y_i)^2 \quad (14)$$

Dodatkowymi miernikami jakości modelu, stosowanymi w problemach regresyjnych mogą być: iloraz odchyłeń standardowych (*SDR*), obliczany jako iloraz odchylenia standardowego błędów predykcji i odchylenia standardowego zmiennej wyjściowej oraz współczynnik liniowej korelacji Pearsona (*R*), obliczany ogółem lub w poszczególnych typach zbiorów, dla odpowiedzi sieci i wartości zadanych.

Zdolność generalizacji, czyli umiejętność uogólniania wiedzy w zakresie modelowanego systemu (Trajer, 2001), jest funkcją złożoności sieci. Aby sieć uzyskała dobrą generalizację, proces uczenia powinien spełniać kilka warunków (Nghia, 2000):

- zbiór uczący powinien być reprezentatywny dla modelowanego zjawiska (procesu) i zawierać cechy, które istotnie je opisują,
- liczba przypadków uczących powinna znacznie (przynajmniej dziesięciokrotnie) przewyższać liczbę połączeń w sieci (liczbę wag sieci),
- struktura sieci powinna być możliwie najprostszą,
- proces uczenia powinien zostać zakończony w chwili osiągnięcia minimum błędu na zbiorze sprawdzającym (innym niż uczący).

Istotną rolę w zapewnianiu dobrych zdolności uogólniających sieci odgrywa wymiar Vapnika-Chervonenkisa (*VCdim*), odzwierciedlający stopień złożoności sieci. Górny i dolny zakres tej miary można oszacować na podstawie zależności (Hush i Horne, 1993):

$$2 \left\lceil \frac{K}{2} \right\rceil N \leq VC \dim \leq 2N_w (1 + \log N_n), \quad (15)$$

w której [] — część całkowita liczby, *N* — liczba wejść sieci (wymiar wektora wejściowego), *K* — liczba neuronów w warstwie ukrytej, *N_w* — całkowita liczba połączeń w sieci (odpowiada liczbie wag), *N_n* — całkowita liczba neuronów w sieci. Dolna granica tej miary jest w przybliżeniu równa liczbie połączeń między warstwą wejściową a warstwą ukrytą, a granica górna przekracza dwukrotność wszystkich połączeń sieci. W praktyce za miarę *VCdim* przyjmuje się liczbę wszystkich wag sieci. Im większa jest liczba połączeń w sieci, tym większa jest jej złożoność i wartość miary *VCdim*.

Własności uogólniające sieci MLP są zależne od stosunku liczby przypadków uczących do wartości *VCdim*. Jeśli dysponujemy małą liczbą próbek to przy ustalonej wartości wymiaru *VCdim* sieć dobrze dopasuje się do danych w zbiorze uczącym, ale nie zapewni to dobrej generalizacji. Liczba dobieranych parametrów znacznie przewyższy liczbę

dopasowywanych sygnałów wyjściowych (Osowski, 1996). Nadmiarowość parametrów sieci zapewni niemal idealne odwzorowanie danych, nie umożliwi jednak rozwiązania zadania, które te dane reprezentują, gdyż sieć nie będzie posiadać zdolności uogólniania. Oszacowanie wartości wymiaru $VCdim$ umożliwi określenie rozmiarów zbioru uczącego, który zapewni dobre zdolności uogólniające sieci. Zgodnie z wcześniejszymi wskazówkami liczba próbek w zbiorze uczącym powinna przynajmniej dziesięciokrotnie przewyższać wartość miary $VCdim$.

Zapewnienie dobrych własności uogólniających przy rozwiązywaniu złożonych problemów wymaga nie tylko ingerencji w zbiór uczący, ale również w strukturę sieci. Praktyka pokazuje, że w większości przypadków dobre rezultaty można osiągnąć stosując najprostsze struktury sieci MLP. Zgodnie z twierdzeniem Hecht-Nielsena sieć neuronowa, posiadająca N neuronów na wejściu, $(2N+1)$ neuronów ukrytych i jeden neuron wyjściowy, może z dowolną dokładnością aproksymować każdą funkcję ciągłą (Trajer, 2001).

Określenie struktury sieci MLP często sprawia trudności, zwłaszcza w zakresie doboru liczby neuronów w warstwie ukrytej. Zwiększenie liczby neuronów ukrytych, umożliwia modelowanie bardziej złożonych zjawisk, ale powoduje rozrastanie się sieci, co skutkuje spowolnieniem jej działania i trudnościami w uczeniu, zwłaszcza skłonnością do zbytniego dopasowywania się do danych uczących (modelowanie szumu). Ograniczenie stopnia złożoności sieci realizuje się zwykle w dwojaki sposób: przez budowę kilku modeli o różnych strukturach i wybór tego, który generuje najmniejszy błąd dla zbioru testowego lub przez redukcję wag, których wartość bezwzględna nie wpływa istotnie na funkcję błędu. W tym procesie skupiamy się na eliminowaniu zbędnych połączeń między neuronami warstwy wejściowej i ukrytej, nie uwzględniając możliwości ograniczenia liczby wejść sieci. Wymiar wektora wejść ustala liczba zmiennych, opisujących modelowane zjawisko, a każda z tych zmiennych wnosi do modelu różne dawki istotnych informacji. Doświadczony badacz potrafi wyeliminować niepotrzebne (mniej istotne) informacje, ale przy modelowaniu skomplikowanych zjawisk taka selekcja nie zawsze jest możliwa. W takich wypadkach można wykonać analizę wrażliwości sieci względem danych uczących. Analizę tę przeprowadza się po wstępnym nauczeniu sieci, a polega ona na wyznaczeniu sumy kwadratów odchyłeń (regresja) lub liczby błędnych klasyfikacji (klasyfikacja) po usunięciu jednej ze zmiennych. Miarą wrażliwości sieci jest iloraz błędu uzyskanego przy uruchomieniu sieci dla zbioru danych bez jednej zmiennej i błędu uzyskanego z kompletem zmiennych. Im większy błąd po odrzuceniu zmiennej, w stosunku do pierwotnego błędu, tym bardziej wrażliwa jest sieć na brak tej zmiennej. Należy jednak pamiętać, że wynik analizy wrażliwości może zostać zafałszowany jeżeli zmienne wejściowe są silnie zależne, ponieważ iloraz błędu obliczany niezależnie, dla każdej zmiennej może nie odzwierciedlać rzeczywistej sytuacji.

Głównym celem niniejszej pracy było sprawdzenie możliwości wykorzystania jednokierunkowych sieci wielowarstwowych MLP jako narzędzia wspomagającego prognozowanie plonu jęczmienia jarego. Realizacja celu pracy wymagała również wyboru optymalnej struktury (topologii) sieci neuronowej, wskazania wektora zmiennych, istotnie determinujących plon jęczmienia jarego oraz oceny wpływu redukcji tego wektora na zdolności predykcyjne sieci.

MATERIAŁ I METODY

Przedmiotem badań były rośliny linii podwojonych haploidów (DH) jęczmienia jarego (*Hordeum vulgare* L.) oraz rośliny odmiany Scarlett. Badane linie DH uzyskano na drodze androgenezy w kulturze *in vitro* z izolowanych mikrospor jęczmienia jarego odmiany Scarlett.

Doświadczenie założono na glebie zbielicowanej, czarnej ziemi, klasy bonitacyjnej IIIb/IVa, kompleksu glebowego żytniego bardzo dobrego. Siewu ziarniaków linii DH kolejnej generacji i odmiany Scarlett wykonano: 14-04-2004 (10 linii DH), 29-03-2007 (4 linie DH), oraz 2-04-2008 (3 linie DH). Przedsięwzięcie wykonano nawożenie mineralne wieloskładnikowym nawozem mineralnym Polifoska 6 w ilości 300 kg/ha, pogłównie — nawożenie azotowe w postaci saletry amonowej (34% N) w ilości 100 kg/ha. Zastosowano chemiczne odchwaszczanie preparatem Mustang w fazie krzewienia.

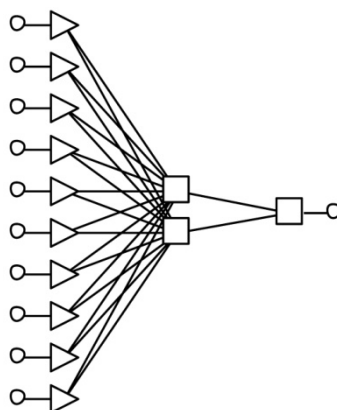
W okresie wegetacji roślin wykonano opisy stadiów rozwojowych roślin wg dziesiątej skali Zadoksa (Zadoks i in., 1974). Fazy rozwojowe rośliny wyodrębniono na podstawie kryteriów, którymi była struktura morfologiczna rośliny wyrażona jej pokrojem, czyli liczbą liści w siewce, liczbą pędów w roślinie, liczbą węzłów na pędzie głównym, pojawieniem się kłosa ponad liściem flagowym, pyleniem. Fazy dojrzewania ziarniaków wyodrębniono na podstawie jego konsystencji, barwy i stopnia twardości.

W fazie dojrzałości pełnej, z każdego poletka pobierano losowo po 15 roślin. Tak więc dla każdej badanej linii DH i dla odmiany Scarlett, w każdym roku badań uzyskano po 15 obserwacji analizowanych cech, traktowanych dalej jako powtórzenia. Łącznie przeanalizowano 315 roślin, w tym 165 (2004 r.), 60 (2007 r.) i 90 roślin (2008 r.). Dokonano pomiaru długości najdłuższego pędu w roślinie, określano całkowitą liczbę pędów w roślinie z uwzględnieniem pędów produktywnych (in. kłosonośnych), niedogonów (niedojrzałe), w tym suchych (ontogenetycznie najmłodsze pędy w roślinie, niedorozwinięte i silnie zaschnięte podczas zbioru roślin). Przeprowadzono pomiary długości kłosów, policzono ogólną liczbę kłosków w kłosie z uwzględnieniem liczby kłosków sterylnych i płodnych, liczbę ziarniaków w kłosie, zważono ziarniaki ze wszystkich kłosów z jednej rośliny. Średnią masę ziarniaków z kłosa stanowił iloraz masy ziarniaków z rośliny i liczby kłosów z rośliny. Jednostkową masę ziarniaka stanowił iloraz masy ziarniaków z rośliny i liczby ziarniaków z rośliny. Średnią masę ziarniaków z kłosa i jednostkową masę ziarniaka wyliczono na podstawie danych uzyskanych z 15 roślin z trzech powtórzeń zarówno dla linii DH jak i odmiany wzorcowej.

Do budowy modelu prognostycznego posłużył zbiór danych, który zawierał 315 przypadków (traktowanych na równi obserwacji dla linii DH i odmiany Scarlett) oraz 11 zmiennych, określających: długość trwania faz: krzewienia, strzelania w źdźbło, kłoszenia i pylenia oraz łączny czas trwania faz dojrzałości mleczej, woskowej i pełnej, a także długość pędu, długość kłosa, liczbę kłosów z rośliny, liczbę kłosków sterylnych, liczbę kłosków płodnych, masę ziarna z kłosa oraz plon ziarna z rośliny. Zbiór danych podzielono na zbiór testowy, zawierający 78 (25% danych) losowo wybranych próbek oraz na zbiór uczący, który zawierał 237 (75% danych) pozostałych próbek, z czego 24 (10% danych) z nich weszło do zbioru walidacyjnego.

W celu sprawdzenia, czy do prognozowania wartości plonu jęczmienia jarego, na podstawie analizowanych danych, wystarczy zastosowanie statystycznych modeli liniowych, przeprowadzono analizę regresji wielokrotnej plonu ziarna z rośliny względem pozostałych analizowanych cech plonotwórczych. Analizę regresji wielokrotnej wraz z badaniem uzyskanego modelu przeprowadzono w Systemie SAS[®] w wersji 9.2 (SAS Institute Inc., 2009).

Neuronowy model prognostyczny w postaci sieci MLP zaimplementowano w programie Statistica 9.0. Biorąc pod uwagę liczbę próbek uczących oraz liczbę wejść sieci ustalono, że powinna ona mieć strukturę 10:2:1 (10 wejść, 2 neurony w jednej warstwie ukrytej, 1 wyjście) aby posiadać zdolność uogólniania (rys. 4).



Rys. 4. Struktura sieci MLP, wykorzystanej do prognozowania plonu
Fig. 4. The structure of the MLP network used for field prediction

Wartość miary $VCdim$ dla tego modelu wynosiła 23, a więc liczba próbek uczących (237) przekraczała dziesięciokrotnie liczbę połączeń sieci. Wyjście sieci stanowiła zmienna, określająca plon ziarna z rośliny, na wejścia wprowadzono pozostałych 10 zmiennych. Sieć uczono przez 200 epok (cykli) z wykorzystaniem algorytmu quasi-Newtonowskiego BFGS. Wagi w pierwszym cyklu uczenia zainicjalizowano w sposób losowy i w trakcie uczenia nie stosowano ich redukcji. Oceny stopnia dopasowania modelu do rozpatrywanego problemu, dokonano głównie na podstawie wartości błędu globalnego, współczynnika korelacji liniowej Pearsona oraz ilorazu odchyłeń standardowych. Wpływ redukcji wektora zmiennych wejściowych na zdolności predykcyjne sieci zbadano na podstawie wyników, uzyskanych w analizie wrażliwości modelu. Na jej podstawie wybrano również optymalną strukturę sieci MLP.

OMÓWIENIE WYNIKÓW

W celu sprawdzenia, czy do prognozowania plonu ziarna z rośliny jęczmienia jarego wystarczy prosty model liniowy, przeprowadzono analizę funkcji regresji wielokrotnej dla

plonu ziarna z rośliny względem pozostałych analizowanych cech plonotwórczych. Do wyboru finalnego modelu regresji wielokrotnej, zawierającego wyłącznie cechy plonotwórcze powiązane w sposób statystycznie istotny z plonem ziarna z rośliny, wykorzystano dwukierunkową metodę krokową doboru zmiennych do modelu (ang. stepwise selection). Finalny model zawierał 5 zmiennych przyczynowych (tab. 1). Współczynnik determinacji dla tego modelu wynosił $R^2 = 0,9837$, a współczynnik determinacji poprawiony względem stopni swobody wynosił $R_{adj}^2 = 0,9835$.

Tabela 1

Model regresji wielokrotnej Multiple regression model				
Analiza wariancji dla modelu regresji wielokrotnej; Analysis of variance for multiple regression model				
Źródło; Source	Stopnie swobody; Degrees of freedom	Suma kwadratów; Sum of squares	Średni kwadrat; Mean square	Statystyka F; F statistic
Model; Model	5	7855,676	1571,1352	3738,73**
Reszty; Residual	309	129,852	0,4202	
Razem; Total	314	7985,528		
Oszacowanie parametrów modelu regresji wielokrotnej; Parameter estimation for multiple regression model				
Parametr; Parameter	Oszacowanie; estimation	Błąd standardowy; Standard error	Statystyka t t statistic	
Stała; Intercept	46,4619	7,4879	6,20**	
Czas trwania fazy kłoszenia i pylenia; Duration of ear emergence and anthesis stages	-1,4728	0,2247	-6,56**	
Liczba kłosów z rośliny; Spikes no. per plant	1,1661	0,0151	77,27**	
Długość kłosa Length of spike	-0,1410	0,0539	-2,62**	
Masa ziarniaków z kłosa Grain mass per spike	7,4685	0,2490	30,00**	
Czas trwania fazy dojrzewania; Duration of maturation phases	-1,1845	0,1645	-7,20**	

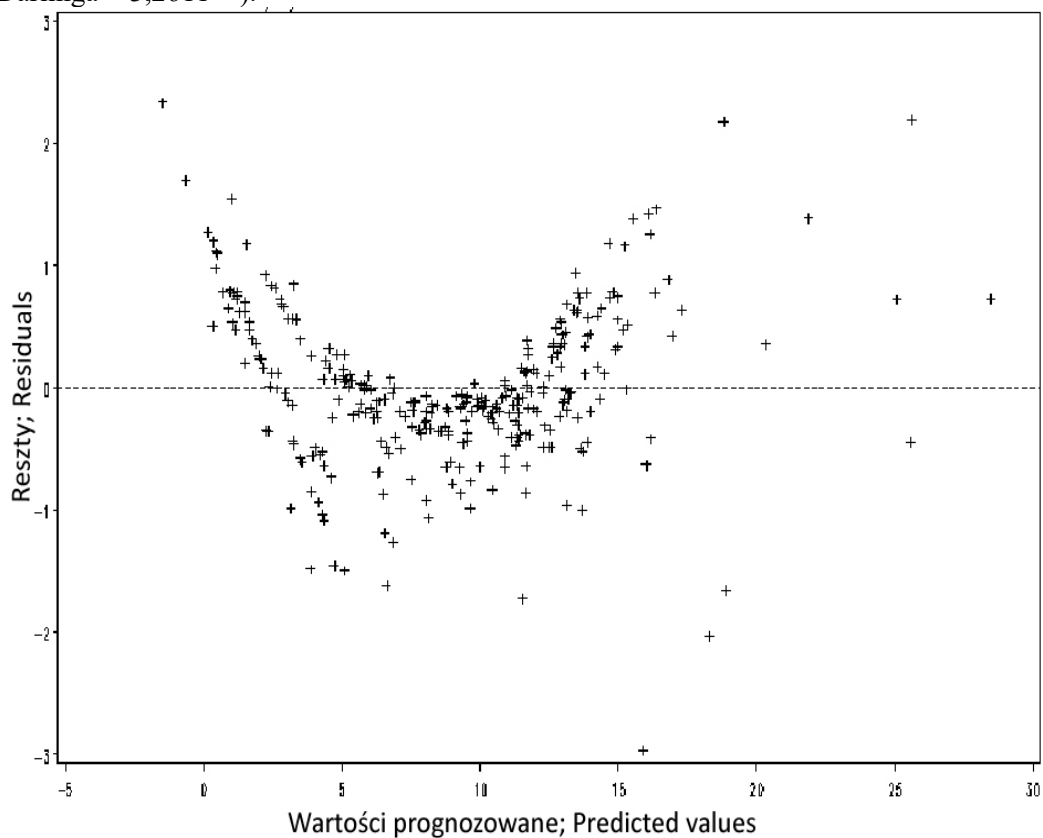
** — Istotne przy $\alpha = 0,01$

** — Significant at $\alpha = 0.01$

Mogłoby się więc wydawać, że do prognozowania wartości plonu jęczmienia jarego, w tym przypadku wystarczyłoby prosty, statystyczny model regresji wielokrotnej. Nie wolno jednak zapominać, że zastosowanie modelu liniowego obłożone jest ścisłymi założeniami co do zmiennych i cech samego modelu. Pierwszy warunek mówi o tym, że zmienna zależna, tu: plon ziarna z rośliny, musi mieć rozkład normalny. Ponieważ plon jest rodzajem średniej (łączny zbiór podzielony przez jednostkę pomiarową), to zgodnie z twierdzeniem Lindeberga-Lévy'ego, zwanym Centralnym Twierdzeniem Granicznym (Wójcik i Laudański, 1989), ma on z założenia rozkład normalny. Ten warunek jest więc spełniony. Drugie, fundamentalne założenie modelu liniowego regresji wielokrotnej dotyczy reszt (ang. residuals), czyli różnic pomiędzy wartościami uzyskanymi z modelu

regresyjnego, a rzeczywistymi wartościami zmiennej zależnej. Założenie to mówi, że reszty muszą być losowe oraz ich rozkład musi być rozkładem normalnym (Rawlings i in., 2001). W rozpatrywanym modelu reszty zdecydowanie nie miały charakteru losowego (rys. 5), gdyż nie tworzyły losowej „chmury” punktów na wykresie ich rozrzutu.

Dodatkowo analiza zgodności rozkładu reszt z rozkładem normalnym z zastosowaniem czterech różnych testów statystycznych wykazała, że rozkład reszt nie jest nawet zbliżony do rozkładu normalnego (statystyka Shapiro-Wilka = 0,9636**, statystyka Kołmogorova-Smirnova = 0,0800**, statystyka Cramera-von Misesa = 0,6096**, statystyka Andersona-Darlinga = 3,2611**).



Rys. 5. Wykres rozrzutu reszt względem wartości przewidywanych
Fig. 5. Residuals vs. predicted scatterplot

Miary dopasowania modelu
Measures of fit of the model

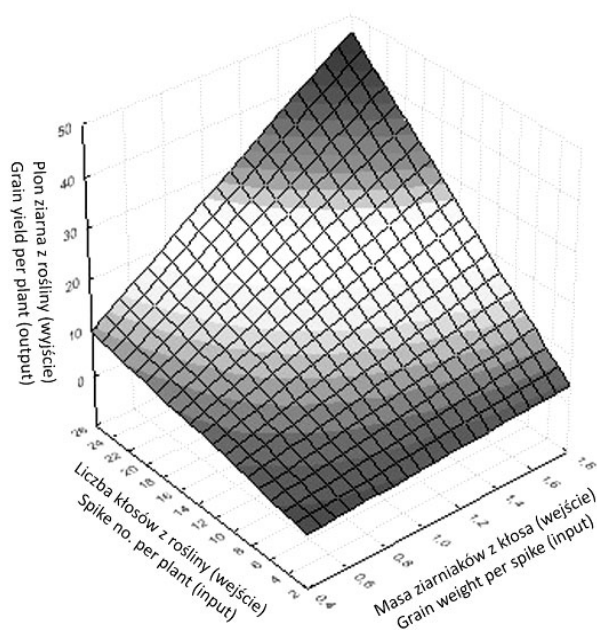
Kryterium Criterion	Aproksymacja (zbiór uczący) Approximation (training set)	Walidacja (zbiór testowy) Validation (test set)
SS	0,120855	1,413112
MAE	0,001446	0,014710
MSE	0,000512	0,018352
RMS	0,065094	0,13547
R	0,999990	0,999606
SDR	0,004395	0,028444
GE		0,012771

Niespełnienie założeń liniowego modelu regresji wielokrotnej co do rozkładu reszt pozwoliło na wyciągnięcie wniosku, że model liniowy, a w szczególności model regresji wielokrotnej, nie jest właściwym modelem do prognozowania wielkości plonu jęczmienia jarego. Do tego celu należałoby wykorzystać modele nieliniowe. Mogą to być modele statystyczne (wielocechowe modele nieliniowe uważa się za skomplikowane i stwarzające problemy w interpretacji) oraz modele matematyczne. Do tej drugiej grupy zaliczyć można sztuczne sieci neuronowe.

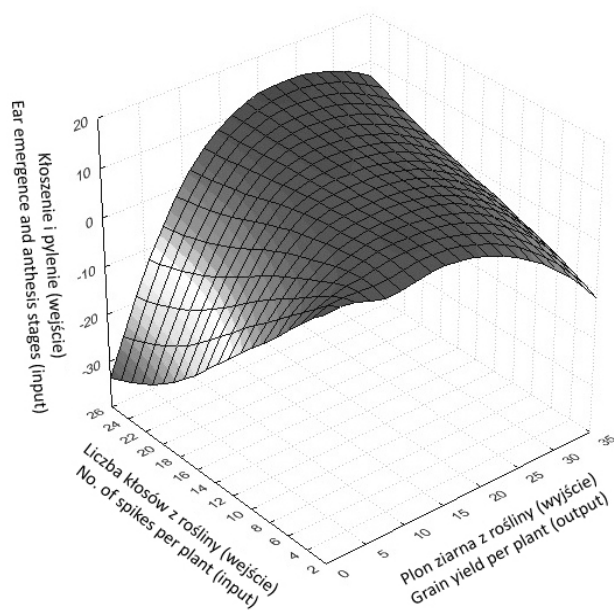
Model neuronowy o strukturze 10:2:1 osiągnął kryterium aproksymacji w 179 cyklu, co oznacza, że adaptacja wag przez algorytm uczący została dokonana 179 razy (jest to algorytm szybko zbieżny, stąd tak krótki czas uczenia). Jakość dopasowania modelu prezentuje tabela 2. Współczynniki korelacji liniowej Pearsona zarówno dla zbioru uczącego, jak i testowego były wysokie, co oznacza, że odpowiedź modelu jest mocno zbliżona do wartości zadanych, a sieć potrafi prawidłowo odwzorować zależności charakterystyczne dla modelowanego zjawiska. Przykładowe odwzorowania plonu w funkcji dwóch zmiennych wejściowych prezentuje rysunek 6.

Ostatecznej oceny jakości dopasowania modelu dokonano na podstawie wartości błędu globalnego, który wskazuje, że proponowana sieć popelnia przy prognozowaniu plonu błąd rzędu 0,013, a więc odwzorowanie jest dość dokładne. Łatwo zauważyć, że wartości błędów generowanych dla zbioru testowego są nieco wyższe niż w zbiorze uczącym. Jest to pośredni efekt prawidłowego doboru topologii sieci, zapewniającej zdolności uogólniające modelu, przy określonej na wstępie liczbie wejść. Sieć przewiduje wartości wyjściowe z większym błędem, ale mimo wszystko jest to błąd dopuszczalny.

W celu ograniczenia liczby zmiennych wejściowych, co jest równoznaczne z uproszczeniem struktury sieci, przeprowadzono analizę wrażliwości sieci. Jej wyniki wyraźnie wskazały cztery zmienne, które w porównaniu do pozostałych zmiennych wejściowych nie wносиły do modelu na tyle istotnych informacji, żeby nie można było z nich zrezygnować (tab. 3). Wartości ilorazów błędów uzyskanych przy uruchomieniu sieci dla zbioru danych bez tych zmiennych i błędu uzyskanego z kompletem zmiennych były zbliżone do jedności, co oznacza, że usunięcie którejkolwiek z nich praktycznie nie ma wpływu na jakość modelu.



a



b

Rys. 6. Przykładowe odwzorowania zmiennej zależnej przez sieć MLP
Fig. 6. The examples of mapping of the dependent variable by the MLP network

Tabela 3

Analiza wrażliwości Sensitivity analysis	
Zmienne; Variables	Sieć — MLP 10:2:1; Network — MLP 10:2:1
Czas trwania fazy krzewienia; Duration of tillering stage	7,01
Czas trwania fazy strzelania w źdźbło; Duration of shooting phase	26,64
Czas trwania fazy kłoszenia i pylenia; Duration of ear emergence and anthesis stages	97,47
Czas trwania fazy dojrzewania; Duration of maturation phases	19,61
Długość pędu; Length of stem	1,13
Liczba kłosów z rośliny; Spikes no. per plant	193,56
Długość kłosa; Length of spike	1,04
Liczba kłosków sterylnych; Sterile spikelet no.	1,13
Liczba kłosków płodnych; Fertile spikelet no.	1,10
Masa ziarniaków z kłosa; Grain weight per spike	54,70

Tabela 4

**Miary dopasowania modeli z ograniczoną liczbą zmiennych
Fit measures of models with a limited number of variables**

Sieć Network	MLP 6:2:1		MLP 6:3:1	
	aproksymacja (zbiór uczący) approximation (training set)	walidacja (zbiór testowy) validation (test set)	aproksymacja (zbiór uczący) approximation (training set)	walidacja (zbiór testowy) validation (test set)
SS	0,122071	1,435406	0,143788	1,449856
MAE	0,000152	0,016502	0,001276	0,014389
MSE	0,000517	0,018642	0,000609	0,018829
RMS	0,022743	0,136534	0,024683	0,137220
R	0,999990	0,999647	0,999988	0,999641
SDR	0,004514	0,026843	0,004893	0,027028
GE		0,013928		0,013998

A zatem z modelu bazowego usunięto zmienne: liczba kłosków sterylnych, długość pędu, liczba kłosków płodnych, długość kłosa. Ograniczenie liczby wejść sieci dało podstawy do ponownego rozpatrzenia jej struktury. Przy sześciu zmiennych wejściowych i jednej zmiennej wyjściowej należało rozpatrzyć modele MLP o strukturach 6:2:1 ($VCdim = 14$) oraz 6:3:1 ($VCdim = 21$). Obie struktury spełniały warunek, dotyczący minimalnej liczby przypadków uczących, a zatem uczono i testowano oba modele w celu ustalenia, który z nich uzyska lepsze parametry jakościowe. Wartości błędów globalnych oraz

pozostałych kryteriów oceny obu sieci były zbliżone (tab. 4), zatem jako optymalną wybrano sieć o najprostszej strukturze, czyli 6:2:1.

Różnica w dokładności odwzorowania wartości zmiennej zależnej między modelem z ograniczoną liczbą zmiennych a modelem bazowym wynosiła zaledwie 0,001, a zatem uproszczenie struktury modelu nie miało wpływu na jego jakość.

PODSUMOWANIE I DYSKUSJA

Metody statystyczne wykorzystywane w badaniach rolniczych opierają się najczęściej na modelach liniowych. Wielocechowe modele nieliniowe są stosowane rzadziej, prawdopodobnie ze względu na utarte twierdzenie, że są trudne w zastosowaniu i interpretacji wyników. Jak wspomniano we wstępie, sieci neuronowe typu MLP z łatwością radzą sobie z modelowaniem zjawisk i procesów przyrodniczych, również o charakterze nieliniowym, przez co mogą stanowić alternatywę dla modeli statystycznych.

W niniejszej pracy dokonano aproksymacji wielkości plonu na podstawie cech, charakteryzujących nie tylko okres wegetacji roślin, ale także fazę ich pełnej dojrzałości. Do aproksymacji wykorzystano sieć MLP o bardzo prostej topologii, wynikającej zarówno z liczby, jak i struktury dostępnych danych. Efekt wytrenowania sieci był pozytywny, a proponowany model podawał prognozowaną wielkość plonu z błędem rzędu 0,014 dla danych testowych. Uzyskane wyniki wyraźnie wskazują, iż sieć typu MLP może służyć jako narzędzie wspomagające prognozowanie plonu jęczmienia jarego. Należy jednak pamiętać, iż trafność prognoz wykonywanych z pomocą tego narzędzia jest ściśle uzależniona od charakteru danych, na podstawie których zostało ono wytrenowane. Przyszła eksploatacja takiego modelu polega na wprowadzaniu do niego zestawu informacji i odczytywaniu prognozowanej wartości plonu. Do zbioru informacji wejściowych dla proponowanego modelu należały m.in. liczba kłosów z rośliny oraz masa ziaren z kłosa, a więc informacje, które można zdobyć dopiero po wejściu roślin w fazę dojrzałości pełnej. Zatem prognoza wykonywana przez ten model nie jest czasowo zbyt odległa od momentu zbioru. W celu wykonywania prognoz długoterminowych należałoby się zastanowić nad zbiorem informacji wejściowych, który nie może zawierać danych niemożliwych do uzyskania w momencie wykonywania prognozy, a zatem ilość i charakter informacji wprowadzanej do sieci neuronowej w postaci wektora zmiennych wejściowych jest ściśle uzależniona od przyszłego przeznaczenia modelu. Wydłużanie terminu prognozy jest równoznaczne ze zwiększeniem błędu prognozy. Z drugiej strony można model prognostyczny potraktować jak pewnego rodzaju symulację i wówczas oszacować część danych na długo przed uzyskaniem ich faktycznych wartości. Trafność prognoz wykonanych przez taki model również będzie mniejsza, ze względu na błędy szacunkowe.

Dokładność działania sieci MLP zależy przede wszystkim od jakości i ilości wprowadzanych do niej danych, co ściśle determinuje strukturę i własności sieci. W przypadku modeli prognostycznych istotne jest zachowanie równowagi między ich zdolnością do aproksymacji i generalizacji, a więc topologia sieci nie może być w całości dziełem przypadku. Prezentowany przykład pokazał, że zastosowanie znanych z literatury zaleceń oraz kryteriów umożliwił już na wstępie prawidłowy dobór struktury sieci, co zapewniło

zdolność modelu do aproksymacji, przejawiającą się w możliwie najdokładniejszym odwzorowaniu wpływu czynników plonotwórczych na wysokość plonu oraz zdolność do generalizacji, umożliwiającą wykonywanie trafnych prognoz dla nowych, nieznanych przypadków.

Jakość prognozowania plonu przez proponowany model przekonuje, że sześć zmiennych, pozostałych po analizie wrażliwości (tab. 3), stanowi najważniejsze cechy plonotwórcze jęczmienia jarego. Jednak model ten nie opisuje struktury ani siły zależności między tymi składowymi, a jedynie odwzorowuje efekt ich wspólnego wpływu na wielkość plonu.

LITERATURA

- Cybenko G. 1988. Continuous valued neural networks with two hidden layers are sufficient. Technical Report. Department of Computer science, Tufts University, Medford.
- Cybenko G. 1989. Approximation by superposition's of a sigmoidal function. *Mathematics of Control, Signals, and Systems* 2: 303 — 314.
- Gatnar E. 2009. Analiza dyskryminacyjna. W: *Statystyczna analiza danych z wykorzystaniem programu R*, Waleśiak M., Gatnar E. (red.). PWN, Warszawa.
- Górny A. G. 2004. Zarys genetyki jęczmienia (*Hordeum vulgare* L.). W: *Zarys genetyki zbóż*. Praca zbiorowa pod red. A. G. Górnego. t. 1: 15 — 80.
- Hush D., Horne B. 1993. Progress in supervised neural networks. *IEEE Signal Processing Magazine*, January: 8 — 39.
- Khazaei J., Naghavi M. R., Jahansouz M. R., Salimi-Khorshidi G. 2008. Yield estimation and clustering of chickpea genotypes using soft computing techniques. *Agron. J.* 100 (4): 1077 — 1087.
- Klepper B., Rickman R.W., Waldman S., Chevalier P. 1998. The physiological life cycle of wheat: Its use in breeding and crop management. *Euphytica* 100: 341 — 347.
- Kozdój J. 1992. Wpływ wybranych czynników środowiska na morfogenezę kłosa i potencjał plonotwórczy zbóż. *Biul. IHAR* 183: 59 — 71.
- Kozdój J. 1994. Wzrost i rozwój rośliny zbożowej — badania botaniczne a praktyka rolnicza. *Biul. IHAR* 192: 3 — 21.
- Kozdój J., Mańkowski D. R., Oleszczuk S. 2010. Analiza potencjału plonotwórczego linii podwojonych haploidów jęczmienia jarego (*Hordeum vulgare* L.) otrzymanych na drodze androgenyzy. *Biul. IHAR* 256: 97 — 116.
- Łubkowski Z. 1968. Jęczmień, Wyd. 2. PWRiL, Warszawa.
- Mittal G. S., Zhang J. 2000. Prediction of temperature and moisture content of frankfurters during thermal processing using neural network. *Meat Sci.* 55: 13—24.
- Nghia D. D. 2000. Sieci neuronowe w zastosowaniu do rozpoznawania i klasyfikacji wzorców. Rozprawa doktorska. Wydział Elektryczny, Politechnika Warszawska.
- O'Neal M. R., Engel B. A., Ess D. R., Frankenberger J. R. 2002. Neural network prediction of maize field using alternative data coding algorithms. *Biosyst. Eng.* 83: 31 — 45.
- Osowski S. 1996. Sieci neuronowe w ujęciu algorytmicznym. WNT, Warszawa.
- Osowski S. 2006. Sieci neuronowe do przetwarzania informacji. WPW, Warszawa.
- Park S. J., Hwang C. S., Vlek P. L. G. 2005. Comparison of adaptive techniques to predict crop yield response under varying soil and land management conditions. *Agric. Syst.* 85: 59 — 81.
- Rawlings J. O., Pantula S. G., Dickey D. A. 2001. *Applied Regression Analysis — a Research Tool*. Second Edition. New York, USA: Springer-Verlag Inc.
- Rosenblatt F. 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65 (6): 386 — 408.
- SAS Institute Inc. 2009. *SAS/STAT 9.2 User's Guide*, Second Edition. Cary, NC, USA: SAS Publishing, SAS Institute Inc.

- Seyhan A. T., Tayfur G., Karakurt M., Tangolu M. 2005. Artificial neural network (ANN) prediction of compressive strength of VARTM processed polymer composites. *Comput. Mater. Sci.* 34: 99 — 105.
- Tadeusiewicz R. 1993. *Sieci neuronowe*, Wyd. 2. Akademicka Oficyna Wydawnicza, RM., Warszawa.
- Tompos A., Margitfalvi J. L., Tfirst E., Heberger K. 2007. Predictive performance of “highly complex” artificial neural networks. *Appl. Catal. A* 324: 90 — 93.
- Trajner J. 2001. *Modelowanie procesu przechowywania warzyw w wybranych jego aspektach*. Rozprawy naukowe i monografie. Wydawnictwo SGGW, Warszawa.
- Uno Y., Prasher S. O., Lacroix R., Goel P. K., Karimi Y., Viau A., Patel R. M. 2005. Artificial neural networks to predict corn yield from Compact Airborn Spectrographic Imager data. *Comput. Electron. Agric.* 47: 149 — 161.
- Wójcik A. R., Ludański Z. 1989. *Planowanie i wnioskowanie statystyczne w doświadczeniach*. Warszawa: PWN.
- Zadoks J. C., Chang T. T., Konzak C. F. 1974. A decimal code for the growth stages of cereals. *EUCARPIA Bulletin*. Vol. 7: 42 — 52.