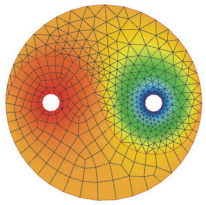
Publishing House
AKAPIT

COMPUTER METHODS IN MATERIALS SCIENCE

Informatyka w Technologii Materiałów

Vol. 6, 2006, No. 2



PORÓWNANIE METOD STATYSTYCZNYCH I WYKORZYSTUJĄCYCH SZTUCZNE SIECI NEURONOWE W ANALIZIE ZNACZENIA I WSPÓŁDZIAŁANIA PARAMETRÓW PROCESÓW TECHNOLOGICZNYCH

MARCIN PERZYK, JACEK KOZŁOWSKI

Politechnika Warszawska, Instytut Technologii Materiałowych, Narbutta 85, 02-524 Warszawa

COMPARISON OF STATISTICAL AND NEURAL NETWORKS-BASED METHODS IN ANALYSIS OF SIGNIFICANCE AND INTERACTION OF MANUFACTURING PROCESSES PARAMETERS

Abstract

Due to development of computer techniques, large amounts of data are collected and stored in many manufacturing companies, related to designs, products, equipment, materials, manufacturing processes etc. This data can be a source of a valuable information. The extracting useful knowledge from that data, using intelligent and partly automated techniques, is called data mining. Until now, data mining has been primarily used in business area. Applications to manufacturing and design problems are seldom.

Some important problems that can be solved through extracting knowledge from a recorded past data in a manufacturing company include: detection of causes of deteriorating product quality, prediction of a break-downs of machines, indication of optimal or critical process parameters and their combinations. These problems can be solved by determining relative significance factors of input variables as well as interaction coefficients between them.

Many of the data mining tasks can be performed using different methods. In general, complex problems, about which no knowledge is available, require learning systems – type models. The statistical methods can be used for less complex tasks and for those about which at least a general character of dependencies is known (e.g. reasonable assumptions about their linearity and occurrence of interactions between variables can be made). The purpose of the present work was a comparison of some statistical (ANOVA, contingency tables, polynomial approximation) and neural network based methods. The general methodology employed in this research is based on utilization of simulated data sets containing assumed hidden relationships between variables. The various types of significance factors were also evaluated for some industrial problems, related to influence of alloying components and process parameters on tensile strength of ductile cast iron.

It was found that the best performance exhibit relative significance factors for single parameters as well interaction coefficients, based on interrogation of trained neural networks. They were calculated according to special procedures, developed by the authors. The results obtained for the industrial data sets confirmed good properties of significance factor of that type. The ANOVA based factors, utilized in some commercial data mining software, essentially underestimate the significance of less important variables while the contingency based factors overestimate them. The ANOVA based factors also exhibit much higher sensitivity to the noise existing in the data.

Although the performance of some techniques discussed in the project is satisfactory, a vast further work is needed. The main goals include development of procedures and tools for cleaning and preparation of rough data, further analysis of behavior of various data mining methods and development of improved definitions of significance and interaction of variables (e.g. detection of synergetic action of more than two variables) as well as development of the software oriented at manufacturing problems.

Key words: data mining, manufacturing processes, significance of parameters, statistical methods, artificial neural networks

ISSN 1641-8581

1. WSTĘP

W ostatnich latach rozwija się intensywnie nowa dyscyplina naukowa, zajmująca się wydobywaniem przydatnych informacji ze zgromadzonych dużych zbiorów danych, zwana *eksploracją danych* (ang. *data mining*). Eksploracja danych może pomóc w rozwiązywaniu zadań różnego typu (Hand et al., 2005) i prowadzić do uzyskiwania wiedzy w różnej postaci. Najczęściej spotykanymi zadaniami są klasyfikacja obiektów oraz generowanie wiedzy w postaci reguł logicznych. Dotychczasowe obszary zastosowań eksploracji danych obejmowały głównie sfery biznesu, zarządzania i nauk społecznych, natomiast zastosowania techniczne, np. w zakresie analizy procesów technologicznych, w tym metalurgicznych, są jeszcze stosunkowo nieliczne (Braha, 2001; Demski, 2004).

Jednym z ważniejszych problemów w technice, jakie można rozwiązywać metodami eksploracji danych, jest uzyskiwanie wiedzy o procesach technologicznych w postaci informacji o znaczeniu poszczególnych ich parametrów (wielkości wejściowych) oraz wzajemnym ich współdziałaniu. Prowadzić to może do uzyskiwania tak cennych rezultatów, jak np. zdefiniowanie zmian parametrów procesu prowadzących do pogorszenia się jakości wyrobów (np. wzrostu ilości braków lub obniżenia własności materiałów), przewidywanie prawdopodobieństwa awarii urządzeń, a także identyfikacja parametrów procesów, które najefektywniej można wykorzystywać do sterowania nimi. Zadaniom tego typu poświęcano dotychczas stosunkowo niewiele uwagi; zastosowania sieci neuronowych do analizy znaczenia parametrów procesów technologicznych można znaleźć w pracach (Fujii et al., 1996; Narayan et al., 1999; Warde & Knowles, 1999; Yescas et al., 2001; Yescas, 2003).

Eksploracja danych wykorzystuje zarówno metody statystyczne, jak i metody sztucznej inteligencji. W niniejszej pracy przeprowadzono badania porównawcze niektórych metod statystycznych z metodami wykorzystującymi sztuczne sieci neuronowe, w zastosowaniu do określania względnych istotności pojedynczych parametrów (zmiennych niezależnych) oraz interakcji pomiędzy dwiema zmiennymi. Metody statystyczne obejmowały jedno i dwuczynnikową analizę wariancji (ANOVA), metody tablic wielodzielczych (test V-Cramera) oraz metody regresji (wielomiany mieszane drugiego stopnia).

Należy zauważyć, że tradycyjnie stosowane w modelowaniu procesów technologicznych (np. materiałowych) metody regresji statystycznej, w tym modele wielomianowe, wymagają założenia postaci funkcji regresji (aproksymującej), co związane jest z koniecznością posiadania pewnej wiedzy o modelowanym procesie. Statystyczne metody analizy wariancji i tablic wielodzielczych nie wymagają takiej wiedzy, jednakże ich zastosowanie jest ograniczone do analizy znaczenia i współdziałania zmiennych i nie nadają się one do *przewidywania* wartości w nowych sytuacjach. Sztuczne sieci neuronowe są z kolei świetnymi narzędziami predykcji, nie wymagającymi znajomości praw rządzących danym procesem, jednakże metody wnioskowania o znaczeniu zmiennych niezależnych i ich interakcjach nie są jednoznacznie określone i dobrze opracowane.

W niniejszej pracy metody oparte na sztucznych sieciach neuronowych wykorzystywały opracowane wcześniej przez autorów współczynniki istotności względnej pojedynczych zmiennych niezależnych (wejściowych) oraz nowo opracowaną metodę wyznaczania współczynników współdziałania dwóch zmiennych. Badania prowadzono na dwóch typach zbiorów danych: symulowanych (o znanych „ukrytych" zależnościach), a także podjęto próbę zastosowania tych metod dla danych przemysłowych, dotyczących wytopu żeliwa sferoidalnego oraz obróbki cieplnej (wytwarzania) żeliwa typu ADI.

2. METODYKA I PLAN BADAŃ

2.1. Zbiory danych

a) Zbiory symulowane

Wzorem prac innych autorów oraz wcześniejszych prac własnych, wygenerowano szereg zbiorów danych o znanych relacjach pomiędzy wielkościami wejściowymi (zmiennymi niezależnymi) a wyjściem (zmienna zależną). Liczebność wszystkich zbiorów wynosiła 1000 rekordów. Metodyka ich generowania była następująca:

- Przyjęcie dowolnego wzoru elementarnego typu Y = f(X1, X2, ..).
- Generowanie liczb losowych X1, X2, ... o rozkładzie prostokątnym i obliczanie wyjść Y dla każdego zestawu tych liczb.
- Wprowadzenie zakłóceń losowych zgodnie z rozkładem normalnym, o maksymalnej wartości ±20%, dla wygenerowanych liczb X1, X2, ...

Dla każdego typu zależności wykonano po 5 generowań zbiorów. Umożliwiło to późniejszą ocenę

poszczególnych metod pod kątem jednoznaczności wykrywania zależności między danymi, niezależnie od losowych zakłóceń występujących w danej ich reprezentacji (zbiorze).

b) Zbiory przemysłowe

Dla celów niniejszej pracy wykorzystano dwa zbiory przemysłowe stworzone w ramach prac wykonywanych wcześniej. Są to:

Zbiór o nazwie „ZS" opisujący zależność pomiędzy składem chemicznym żeliwa sferoidalnego a jego wytrzymałością, stworzony na podstawie danych o wytopach wykonanych w jednej z odlewni warszawskich. Szczegóły dotyczące tego zbioru znaleźć można w (Perzyk & Kochański, 2001).

Zbiór o nazwie „ADI", opisujący zależność pomiędzy składem chemicznym, parametrami obróbki cieplnej i modułem odlewu, a wytrzymałością na rozciąganie żeliwa typu Astempered Ductile Iron. Dane uzyskano w trakcie realizacji grantu badawczego KBN nr 3T08B00726. Pochodziły one zarówno z publikacji światowych, jak i badań własnych, prowadzonych wspólnie z Instytutem Odlewnictwa w Krakowie.

2.2. Definicje współczynników istotności względnej pojedynczych zmiennych wejściowych

a) Sieci neuronowe

W wyniku prac prowadzonych w latach ubiegłych (Perzyk & Kochański, 2003a; Perzyk & Kochański, 2003b) przyjęto dwa rodzaje współczynników istotności względnej zmiennych wejściowych:

Typ A. Obliczany na podstawie wzrostu błędu sieci dla danych uczących wskutek zablokowania danego wejścia na stałym poziomie (współczynnik często spotykany w literaturze).

Typ B. Obliczany na podstawie przyrostu wartości wyjścia z sieci przy zmianie danego wejścia przy pozostałych wejściach ustalanych losowo (propozycja własna).

b) Analiza wariancji

Wykorzystano wartości statystyki testowej F w jednoczynnikowej analizie wariancji obliczane dla oddziaływania danego wejścia na zmienną zależną. Podobna definicja współczynnika istotności względnej przyjmowana jest w module *Data Mining* komercyjnego pakietu *Statistica*. W niniejszej pracy wartości te były znormalizowane przez podzielenie ich przez wartość maksymalną spośród wszystkich zmiennych wejściowych. W ten sposób wartości istotności miały charakter względny i mogły zawierać się w granicach 0 – 1 (podobnie jak dla sieci neuronowych).

c) Metody tablic wielodzielczych

Do wyznaczenia współczynników istotności względnej przyjęto wartość statystyki testowej dla znanego testu tej grupy: V-Cramera. Wartości jej były normalizowane w analogiczny sposób, jak statystyki F w jednoczynnikowej analizie wariancji.

d) Regresja z wykorzystaniem wielomianów

Zastosowano wielomiany mieszane drugiego stopnia. Dzięki temu, że współczynniki wielomianów znajdowane były dla danych normalizowanych w przedziale 0 – 1, współczynniki istotności danej zmiennej niezależnej wyznaczano wprost jako sumę współczynników dla wyrazów: kwadratowego i liniowego dla tej zmiennej (z pominięciem jednak wyrazów mieszanych, które również zawierały daną zmienną). Otrzymane wartości współczynników istotności również normalizowano, analogicznie jak opisano powyżej.

2.3. Definicje współczynników interakcji pomiędzy dwiema zmiennymi wejściowymi

a) Sieci neuronowe

Definicję współczynnika interakcji oparto o zasadę odpytywania sieci przy założeniu, że wartości pozostałych zmiennych przyjmują (wielokrotnie) wartości losowe. Zasada ta, zastosowana do współczynników istotności pojedynczych zmiennych, przyniosła zdecydowanie lepsze rezultaty niż inne metody, oparte na analizie wag nauczonej sieci (Perzyk & Kochański, 2003b). Wartości współczynnika interakcji wyznaczano niejako z definicji, wg następującej procedury, opracowanej w ramach niniejszej pracy:

- Wyznaczenie znormalizowanych odpowiedzi sieci (Y) dla 9 równomiernie rozmieszczonych, znormalizowanych wartości każdej z dwóch analizowanych zmiennych wejściowych (X_i i X_j), dla 1000 losowo wybranych wartości pozostałych zmiennych.

- Obliczenie dwuwymiarowej tablicy 9 x 9 zawierającej 81 średnich wartości Y (uśrednienie dla 1000 losowań pozostałych wejść). Dalsze obliczenia wykorzystują wyłącznie wartości z tej tablicy.
- Obliczenie maksymalnych różnic Y otrzymanych przez zmianę X_j, dla wszystkich 9 poziomów X_i.
- Znalezienie maksymalnej różnicy między różnicami wyznaczonymi w poprzednim punkcie. Wartość ta (pomnożona przez 100) daje zwiększenie działania zmiennej X_j w wyniku działania zmiennej X_i, wyrażone w % zakresu wyjścia Y.
- Analogicznie znajduje się zwiększenie działania zmiennej X_i w wyniku działania zmiennej X_j, wyrażone w % zakresu wyjścia Y.
- Współczynnik interakcji oblicza się jako średnią arytmetyczną obu procentowych wartości obliczonych powyżej.
- W znajdowaniu wszystkich wartości ekstremalnych opisanych w powyższej procedurze zastosowano aproksymację wielomianem 3 stopnia (optymalizowanym dla 9 punktów). O wyborze takiej metody może świadczyć analiza kształtów wielu zależności Y(X) uzyskiwanych w praktyce z nauczonych sieci neuronowych. W przyszłości należałoby rozważyć zastosowanie bardziej uniwersalnych metod optymalizacyjnych.

b) Analiza wariancji

Wykorzystano wartości statystyki testowej F w dwuczynnikowej analizie wariancji. Wartości (nie normalizowane) współczynników interakcji zdefiniowano w niniejszej pracy jako:

$$\frac{F_{i,j}}{\left(F_i + F_j\right)/2} \tag{1}$$

gdzie F_i – wartość statystyki dla testowania istotności oddziaływania czynnika (zmiennej) i,
F_j – wartość statystyki dla testowania istotności oddziaływania czynnika (zmiennej) i,
$F_{i,j}$ – wartość statystyki dla testowania istotności wzajemnego oddziaływania (interakcji) czynników (zmiennych) i, j.

c) Regresja wielomianowa

Wartości współczynników interakcji obliczano na podstawie wartości współczynników wielomianu wg wzoru:

$$\frac{\left|b_{i,j}\right|}{\left(\left|a_i + c_i\right| + \left|a_j + c_j\right|\right)/2} \tag{2}$$

gdzie $b_{i,j}$ – współczynnik przy wyrazie mieszanym $(X_i \cdot X_j)$,
a_i, a_j – współczynniki przy wyrazach kwadratowych dla obu zmiennych, odpowiednio, X_i i X_j,
c_i, c_j – współczynniki przy wyrazach liniowych dla obu zmiennych, odpowiednio, X_i i X_j.

2.5. Metody obliczeniowe i oprogramowanie

a) Sieci neuronowe

Do uczenia, odpytywania oraz wyznaczania współczynników istotności względnych pojedynczych zmiennych i współczynników interakcji między dwiema zmiennymi wykorzystano własne oprogramowanie, w przeważającej części napisane w toku prac wykonywanych w latach ubiegłych. Opis tego programu oraz stosowanych metod uczenia sieci można znaleźć w (Perzyk & Kochański, 2002; Perzyk & Kochański, 2003). Dla celów niniejszej pracy rozbudowano to oprogramowanie o następujące elementy:

- procedurę wyznaczania współczynnika interakcji pomiędzy dwiema zmiennymi niezależnymi, opisaną wyżej
- procedurę umożliwiającą wyznaczanie współczynników istotności pojedynczych zmiennych jako średnich z wielu uczeń sieci.

Analogicznie jak w poprzednich pracach (Perzyk & Kochański, 2003a; Perzyk & Kochański, 2003b) zastosowano sieci typu MLP z jedną warstwą ukrytą. Ponieważ nie zaobserwowano istotnego wpływu zwiększania liczby neuronów ukrytych powyżej wartości równej liczbie neuronów w warstwie wejściowej na przewidywania sieci, ostatecznie zastosowano tę właśnie liczbę z tym jednak, że w żadnym przypadku nie była ona mniejsza od 5. Użyto funkcję aktywacji neuronów typu sigmoidalnego. Przyjęto generalną zasadę dziesięciokrotnego uczenia sieci.

b) Analiza wariancji

Obliczenia jednoczynnikowej analizy wariancji, potrzebne do wyznaczania współczynników istotności względnej zaprogramowano w postaci procedury dołączonej do programu realizującego obliczenia dla sieci neuronowej. Dzięki temu może ona być łatwo wykorzystana jako narzędzie wstępnej oceny istotności

zmiennych, co może w pewnych przypadkach ułatwić decyzję o redukcji liczby zmiennych niezależnych dla sieci neuronowej (podobne rozwiązanie zastosowano w module *Data mining* pakietu *Statistica*).

Obliczenia dwuczynnikowej analizy wariancji, potrzebne dla wyznaczenia statystycznych współczynników interakcji między dwiema zmiennymi, wykonano z użyciem modułu *Zaawansowane modele liniowe i nieliniowe* pakietu *Statistica*. Należy zwrócić uwagę, że obliczenia z wykorzystaniem analizy wariancji wymagają, aby zmienne niezależne były typu nominalnego lub porządkowego, podczas gdy ich wartości w wygenerowanych zbiorach symulowanych, jak i przemysłowych, były typu ciągłego. Zamiany zmiennych rzeczywistych (ciągłych) na porządkowe wykonano z wykorzystaniem metod podobnych do opisanych w (Perzyk & Biernacki, 2004; Perzyk et al., 2005), jednak z zachowaniem stałej liczby klas równej 10.

c) Test V-Cramera

Niezbędne obliczenia, potrzebne do wyznaczania współczynników istotności względnej, zaprogramowano również w postaci procedury dołączonej do programu realizującego obliczenia dla sieci neuronowej, podobnie jak w przypadku jednoczynnikowej analizy wariancji. Zamiany zmiennych rzeczywistych (ciągłych) na porządkowe wykonano z wykorzystaniem metod zbliżonych do podanych wyżej, jednak z zastosowaniem pewnej optymalizacji liczby przedziałów.

d) Regresja wielomianowa

Wyznaczenie współczynników wielomianów wykonano z zastosowaniem własnego oprogramowania stworzonego wcześniej przez autorów dla celów innej pracy. Należy zaznaczyć, że ogólnie dostępne w popularnych programach obliczeniowych narzędzia służące do tego celu (np. *MS Excel*) nie mogły tu być zastosowane z uwagi na dużą liczbę wyrazów wielomianu (np. w przypadku zbioru przemysłowego „ADI" wynosiła ona aż 105). Zastosowano dwie metody optymalizacyjne. Najpierw z wykorzystaniem metody symulowanego wyżarzania procedura optymalizacyjna znajduje punkt leżący w pobliżu minimum globalnego, a następnie wykorzystując metodę gradientową z modułu *Solver* zawartego w *MS Excel* znajduje rozwiązanie końcowe, bardziej dokładne.

3. WYNIKI OBLICZEŃ

3.1. Wyniki dla danych symulowanych

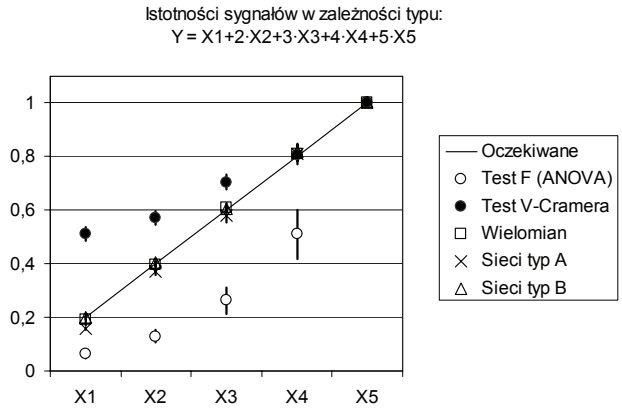
Najważniejsze z uzyskanych wyników przedstawiono w postaci wykresów zbiorczych. Pokazane wartości przedstawiają średnie otrzymane z 5 generowań zbiorów (każdy z zakłóceniami losowymi), zaś zaznaczone słupki błędów oznaczają 95% przedziały ufności dla tych średnich. Wartości współczynników istotności pojedynczych zmiennych dla poszczególnych generowań wyznaczano jako średnie z 10 uczeń sieci, natomiast współczynniki interakcji dla jednego, wybranego uczenia, dającego przewidywania sieci najbliższe średnim ze wszystkich 10 uczeń.

2.1.1. Współczynniki istotności względnej pojedynczych zmiennych

Na rysunku 1 pokazano zestawienie wartości współczynników istotności względnej dla zbioru wygenerowanego (z zakłóceniami) wg zależności $Y=X1+2\cdot X2+3\cdot X3+4\cdot X5+5\cdot X5$, czyli o istotnościach zmiennych liniowo narastających. Widoczne jest, że najlepszą, praktycznie idealną zgodność z oczekiwaniami, dały współczynniki istotności oparte na odpytywaniu sieci neuronowej, typu B oraz wyliczone ze współczynników wielomianu, zaś nieco gorszą – oparte na odpytywaniu sieci neuronowej, typu A. Interesujące są rezultaty uzyskane dla obu metod statystycznych. Obie one odzwierciedlają wprawdzie poprawnie tendencje w wartościach istotności (narastanie od zmiennej X1 do X5), jednakże ich wartości są dość odległe od oczekiwań. Współczynniki oparte na jednoczynnikowej analizie wariancji wykazują wyraźną tendencję do zaniżania istotności dla mniej istotnych zmiennych (w sposób nieliniowy), tj. większego eksponowania różnic pomiędzy tymi istotnościami. Charakteryzują się także większymi rozrzutami wynikającymi z zakłóceń wprowadzanych przy generowaniu poszczególnych zbiorów, a więc są stosunkowo czułe na losowe zakłócenia występujące w „ukrytej" zależności. Współczynniki oparte na teście V-Cramera wykazują tendencję odwrotną, tj. do zmniejszania różnic pomiędzy istotnościami zmiennych, jak również wykazują stosunkowo mniejszy rozrzut.

Na rysunku 2 pokazano zestawienie zakresów otrzymywanych wartości współczynników istotności względnych dla zbiorów wygenerowanych na podstawie wzoru typu $Y=X1+X2+X3+X4+X5$, tj. za-

kładającego identyczne wartości istotności wszystkich zmiennych. Wyniki te dobitnie potwierdzają dużą wrażliwość metody opartej na teście jednoczynnikowej analizy wariancji na losowe zakłócenia w zbiorze, a także najlepsze przewidywania metody wykorzystującej nauczoną sieć neuronową, typu B.

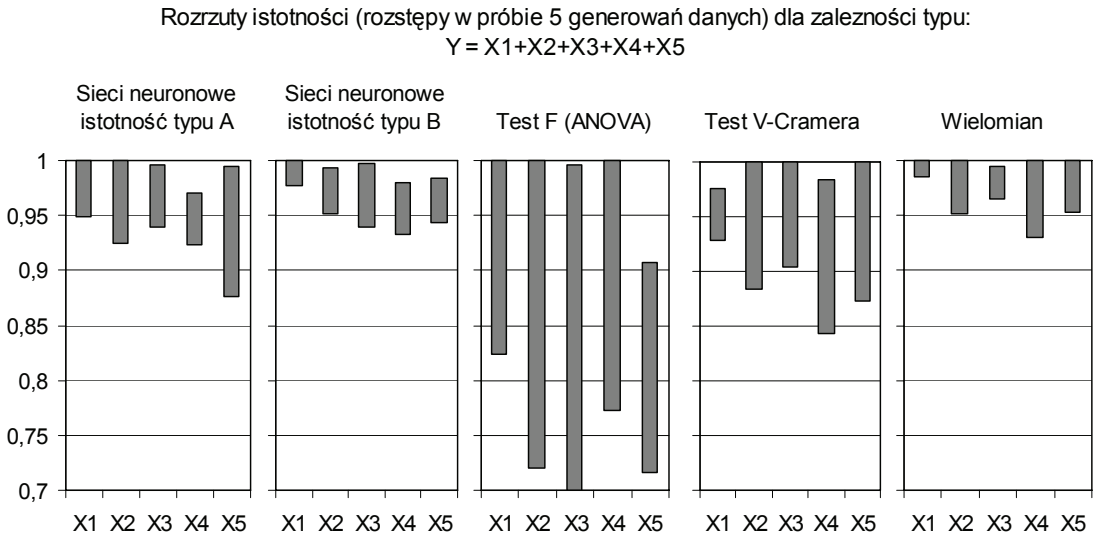


Rysunek 1. *Porównanie istotności względnych zmiennych niezależnych, wyznaczanych różnymi metodami*

Figure 1. *Comparison of various types of relative significance factors of independent variables, obtained from 5 generations of simulated data sets with imposed noise; vertical bars denote 95% confidence intervals*

czynnikach wielomianu daje bardzo duże rozrzuty wyników. Również średnie są niezgodne z przewidywaniami, zwłaszcza w przypadku prostej zależności typu Y=X1·X2. Niewątpliwie wynikało to z przyjętej definicji, w której pomijano wyrazy mieszane, zawierające również daną zmienną. Najlepsze wyniki uzyskano dla metody opartej na teście V-Cramera (jednakowe istotności dla obu zmiennych, bardzo małe rozrzuty). Na drugim miejscu jest metoda oparta na sieciach neuronowych, typu B.

Analizując wykresy pokazane na rysunku 4 należy zauważyć, że wszystkie metody przewidują mniejsze znaczenie tych zmiennych, które występują tylko w postaci iloczynu (X1 i X2 na rysunku 4a oraz X1, X2 i X3 na rysunku 4b), a nie występują niezależnie (X3, X4 i X5 na rysunku 4a oraz X34 i X5 na rysunku 4b). Jest to generalnie zgodne z oczekiwaniami: znaczenie takiej zmiennej jest mniejsze, gdyż „potrzebuje" ona działania innej zmiennej. Należy zauważyć, że istotności oparte na odpytywaniu sieci neuronowej, typu B, dają wartości istotności zmiennych występujących w iloczynach najbardziej zbliżone do oczekiwanych intuicyjnie: ok. 1/2 istotności zmiennych występujących niezależnie w przypadku wzoru typu Y=X1·X2+X3+X4+X5

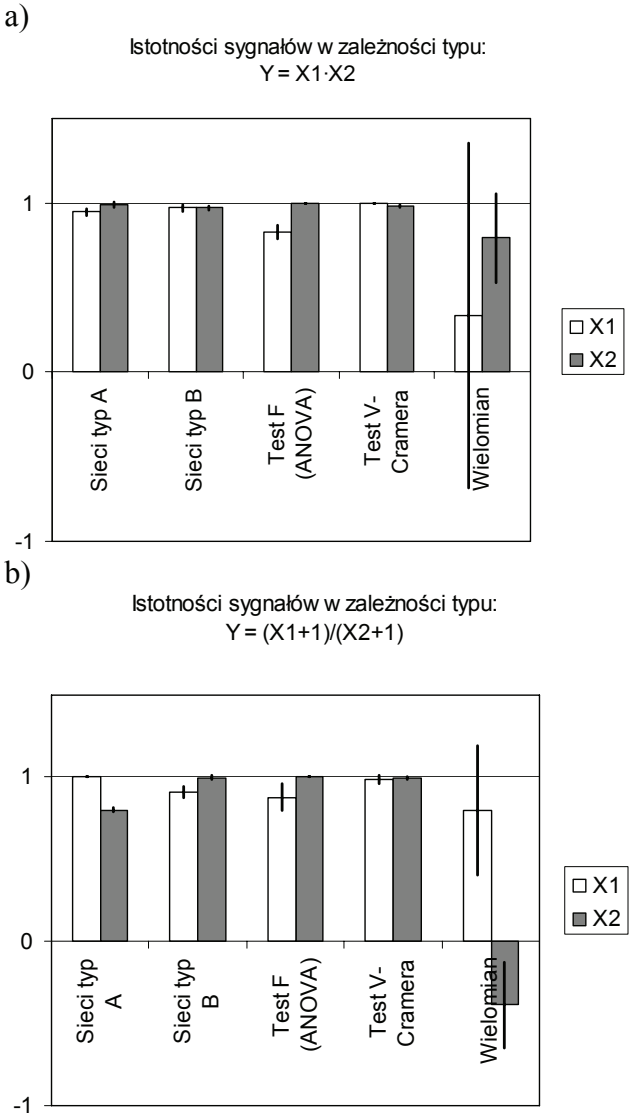


Rysunek 2. *Porównanie rozrzutów (rozstępów) istotności względnych zmiennych niezależnych, wyznaczanych różnymi metodami, wynikających z losowych zakłóceń dla 5 generowań zbiorów*

Figure 2. *Comparison of dispersions (ranges) of various types of relative significance factors of independent variables, obtained from 5 generations of simulated data sets with imposed noise*

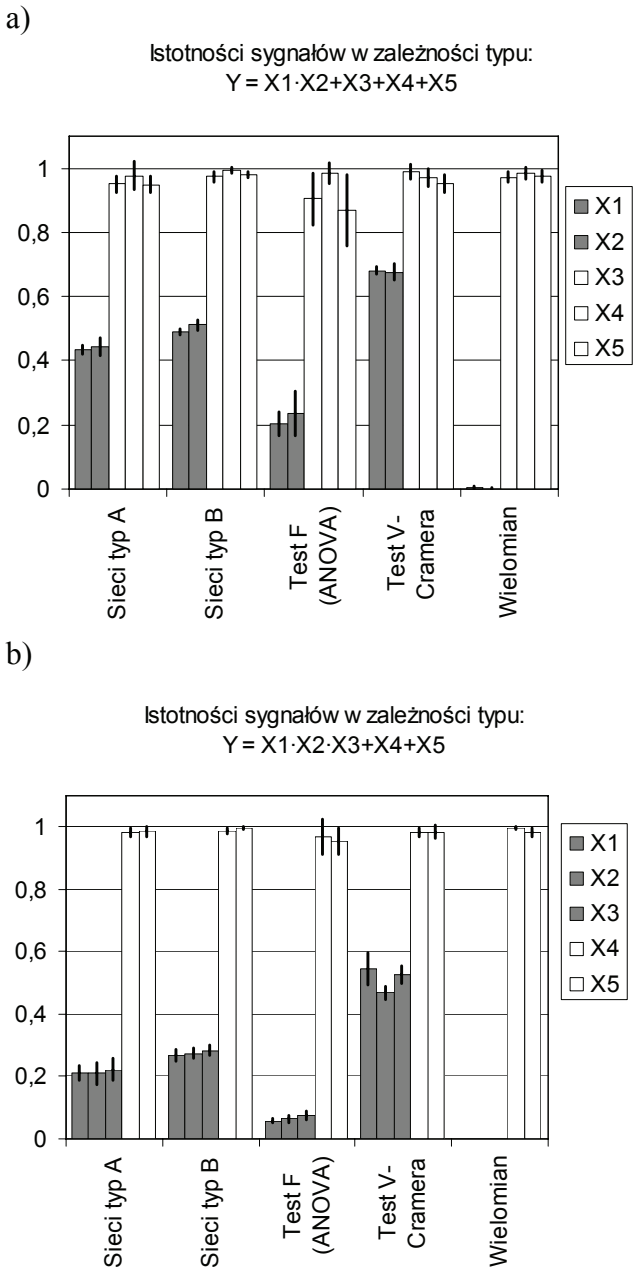
Na rysunkach od 3 do 6 pokazano wykresy istotności względnych otrzymane dla innych, wybranych zbiorów danych symulowanych. Z porównania istotności dla dwóch typów wzorów (rysunek 3), w których występowały jedynie 2 zmienne X1 i X2 z interakcjami wynika, że metoda oparta na współ- oraz ok. 1/3 w przypadku zależności typu Y=X1·X2·X3+ X4+X5. Charakteryzują się one ponadto najmniejszymi rozrzutami oraz zróżnicowaniem pomiędzy średnimi dla zmiennych o tym samym charakterze. Podobnie jak na wykresie przedstawionym na rysunku 1, widoczna jest tendencja do

zawyżania różnic w istotnościach obliczanych z wykorzystaniem jednoczynnikowej analizy wariancji oraz zaniżania tych istotności przez metodę opartą na teście V-Cramera. Obliczenia oparte na współczynnikach wielomianu dały nienaturalnie bardzo niskie istotności dla zmiennych występujących w iloczynach. Podobnie jak dla zależności przedstawionych na rysunku 3, oczywistym powodem tego był fakt, że do ich wyliczenia brano współczynniki przy wyrazach zawierających pojedyncze zmienne (suma współczynników wyrazu kwadratowego i liniowego), które dla omawianych wzorów otrzymuje się jako zasadniczo równe zeru.



Rysunek 3. *Porównanie istotności względnych zmiennych niezależnych, wyznaczanych różnymi metodami, w zbiorach zawierających tylko dwie takie zmienne*

Figure 3. *Comparison of various types of relative significance factors of independent variables, obtained from 5 generations of simulated data sets containing two such variables only, with imposed noise; vertical bars denote 95% confidence intervals*



Rysunek 4. *Porównanie istotności względnych zmiennych niezależnych, wyznaczanych różnymi metodami, w zbiorach ze zmiennymi wejściowymi z interakcjami i bez*

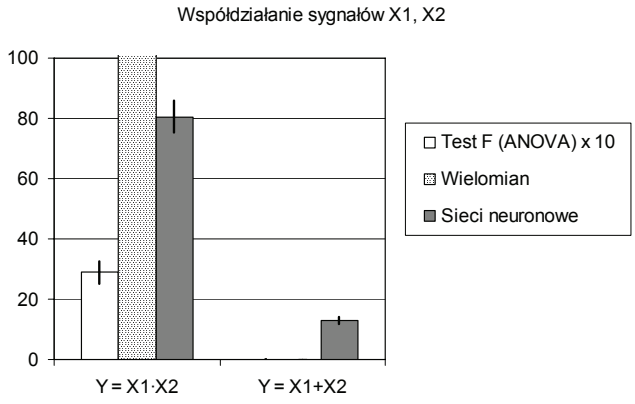
Figure 4. *Comparison of various types of relative significance factors of independent variables, obtained from 5 generations of simulated data sets containing input variables, with - and without - interactions, with imposed noise; vertical bars denote 95% confidence intervals*

2.1.2. *Współczynniki interakcji dwóch zmiennych*

Na rysunku 5 pokazano porównanie współczynników interakcji dla dwóch najprostszych wzorów, z interakcją i bez. Generalnie wszystkie trzy metody wyznaczania tych współczynników dały wyniki zgodne z przewidywaniami: w przypadku wzoru typu Y = X1·X2 są one istotnie wyższe, niż dla wzoru typu Y = X1+X2. Dla metody opartej na dwu-

czynnikowej analizie wariancji oraz współczynnikach wielomianu uzyskano wartości dla drugiego wzoru pomijalnie małe, co jest w pełni zgodne z oczekiwaniami.

Na rysunku 6 pokazano porównania współdziałania dwóch zmiennych w zbiorach z dwiema zmiennymi z najprostszym modelem interakcji, tj. typu Y = X1·X2, ze współdziałaniem uzyskanym dla wzorów, w których występują także inne składniki: te same zmienne, lecz jako wyrazy wolne (rysunek 6a) oraz inne zmienne, także jako wyrazy wolne (rysunek 6b). Widoczne jest wyraźne zmniejszenie się współczynników interakcji wskutek obecności wolnych wyrazów, przy czym dla przypadku pokazanego na rysunku 6a wielkość tego zmniejszenia jest bardziej racjonalna dla metody opartej na sieci neuronowej niż w przypadku pozostałych dwóch metod.



Rysunek 5. *Współczynniki interakcji między dwiema zmiennymi niezależnymi, uzyskane różnymi metodami, dla najprostszych przypadków*

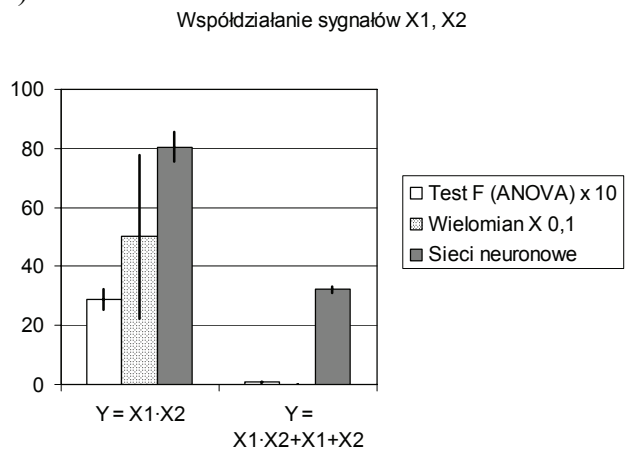
Figure 5. *Interaction coefficients between two independent variables, obtained by different methods, for the simplest cases*

Na rysunku 7 pokazano porównanie współczynników interakcji otrzymanych dla różnych par zmiennych, występujących w trzech typach wzorów zawierających po 5 zmiennych. Dla wzoru typu Y=X1·X2+X3+X4+X5 otrzymano wartości generalnie zgodne z oczekiwaniami z tym, że oczekiwane zerowe interakcje dla par zmiennych X1, X5 oraz X4, X5 przewiduje tylko metoda oparta na współczynnikach wielomianu.

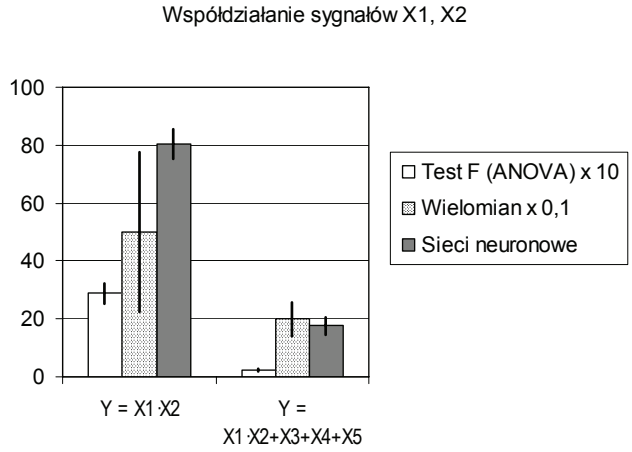
Wyniki zgodne z oczekiwaniami uzyskano także dla wzoru typu Y=X1+X2+X3+X4+X5, nie zawierającego interakcji, gdzie dla wszystkich zmiennych wartości współczynników interakcji są w przybliżeniu równe sobie i niewielkie (w stosunku do wartości otrzymanych dla pary X1 i X2 na rysunku 7a).

Natomiast wyniki pokazane na rysunku 7c, uzyskane ze wzoru również nie zawierającego interakcji, typu Y=X1+2·X2+3·X3+4·X5+5·X5, są dość nieoczekiwane. O ile wielomian przewiduje prawidłowo brak interakcji, to obie pozostałe metody wskazują na ich istnienie, choć w odmiennych przypadkach. Metoda wykorzystująca dwuczynnikową analizę wariancji wskazała na istnienie istotnej interakcji dla pary zmiennych najmniej znaczących, natomiast metoda wykorzystująca sieci neuronowe - dla pary zmiennych najbardziej znaczących. Z drugiej strony, wartości współczynników interakcji dla zmiennych rzeczywiście ją posiadających (X1 i X2 na rysunkach 6a i b) są istotnie większe od pokazanych na rysunku 7c, (zwłaszcza dla metody opartej na sieciach neuronowych).

a)



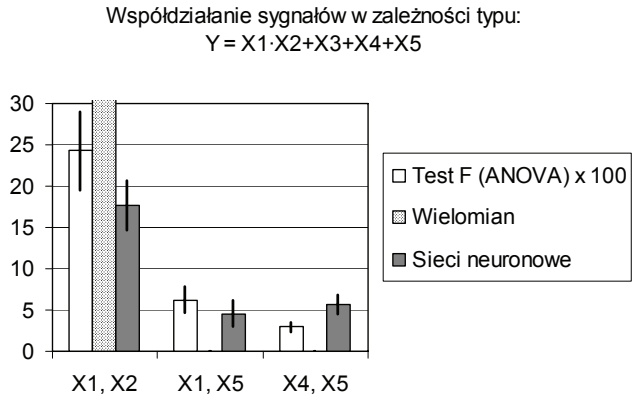
b)



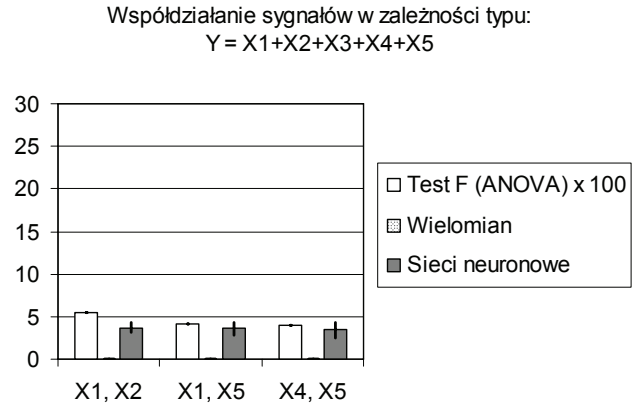
Rysunek 6. *Współczynniki interakcji między dwiema zmiennymi niezależnymi, uzyskane różnymi metodami w zależnościach, w których występują dodatkowe wyrazy zawierające: a) te same dwie zmienne, b) inne zmienne; dla porównania pokazano również wyniki dla najprostszej zależności, typu Y = X1·X2*

Figure 6. *Interaction coefficients between two independent variables, obtained by different methods, in the relationships in which additional terms appear including: a) the same two variables, b) other variables; the results for the simplest relationship of the type Y = X1·X2 are also shown for comparison*

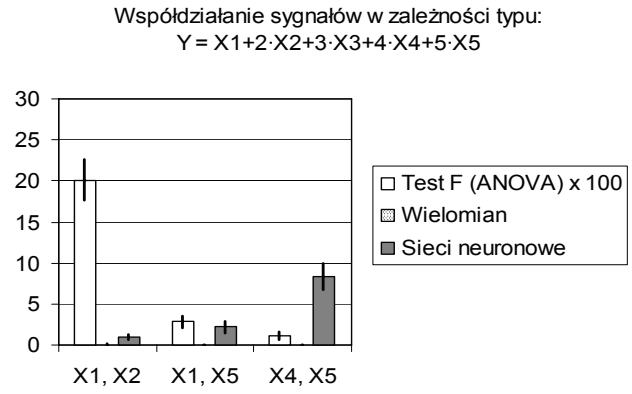
a)



b)



c)

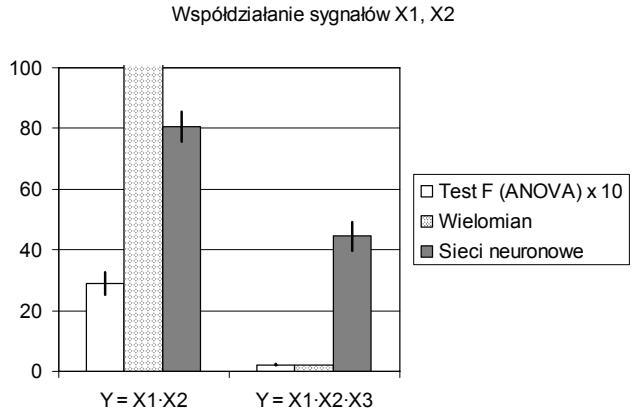


Rysunek 7. *Współczynniki interakcji między dwiema zmiennymi niezależnymi, uzyskane różnymi metodami: a) dla przypadku z interakcją zmiennych X1 i X2, b) i c) bez interakcji*

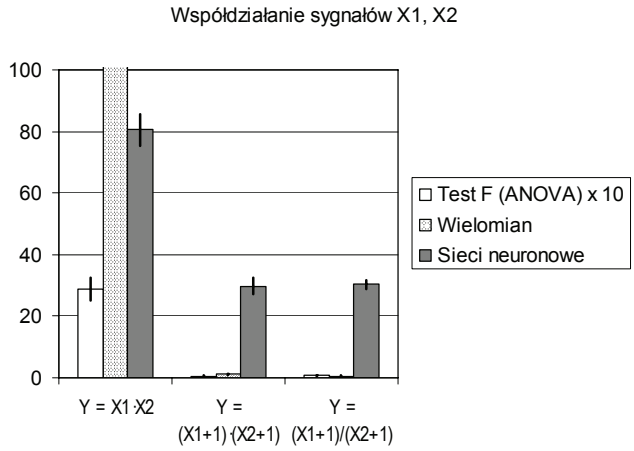
Figure 7. *Interaction coefficients between two independent variables, obtained by different methods: a) for the case of interaction between variables X1 and X2, b) and c) without interactions*

Podsumowując omówienie wykresów przedstawionych na rysunkach 6 i 7 należy stwierdzić, że obecność niezależnych innych zmiennych o zróżnicowanej istotności może wpływać na zafałszowanie wyników dla pary zmiennych bez interakcji.

a)



b)



Rysunek 8. *Współczynniki interakcji między dwiema zmiennymi niezależnymi uzyskane różnymi metodami w wybranych zależnościach zawierających: a) trzecią zmienną będącą w interakcji z tymi dwiema, b) tylko te dwie zmienne ale w złożonych zależnościach*

Figure 8. *Interaction coefficients between two independent variables obtained by different methods in selected relationships including: a) a third variable, being in an interaction with those two, b) the two variables only, but in more complex relationships*

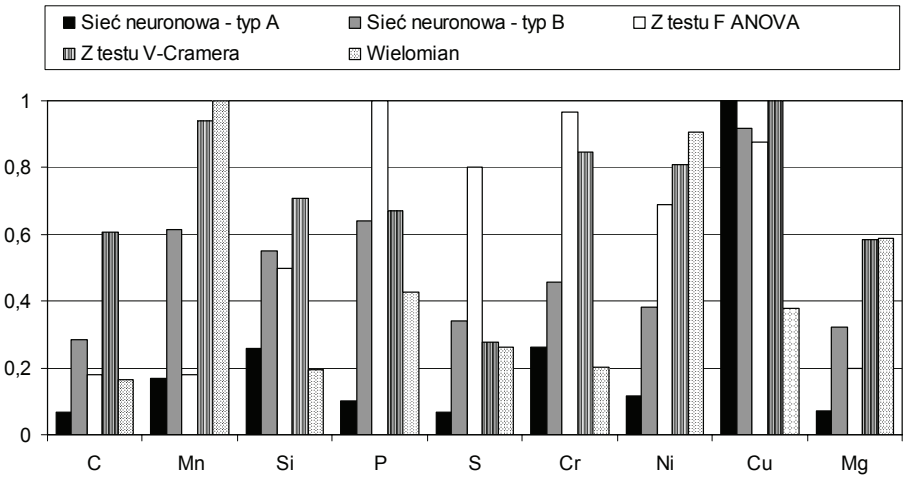
Na rysunku 8 pokazano dalsze porównania współdziałania dwóch zmiennych w zbiorach z dwiema zmiennymi z najprostszym modelem interakcji, tj. typu Y = X1·X2, ze współdziałaniem uzyskanym dla wzorów, w których występują także inne składniki. Na rysunku 8a pokazano sytuację, gdy rozważane dwie zmienne współdziałają jednocześnie z trzecią, natomiast na rysunku 8b porównanie tego najprostszego przypadku z dwoma bardziej złożonymi. We wszystkich tych porównaniach najbardziej racjonalne wartości wykazuje metoda wykorzystująca sieć neuronową. Obie pozostałe metody wskazują na nieracjonalne małe interakcje w porównaniu do wzoru typu Y = X1·X2; np. metoda oparta na dwuczynnikowej analizie wariancji daje w tych przypadkach wartości porównywalne lub

niższe od tych, jakie przewidywała przy rzeczywistym braku interakcji w sytuacjach pokazanych na rysunku 7.

3.1. Wyniki dla danych przemysłowych

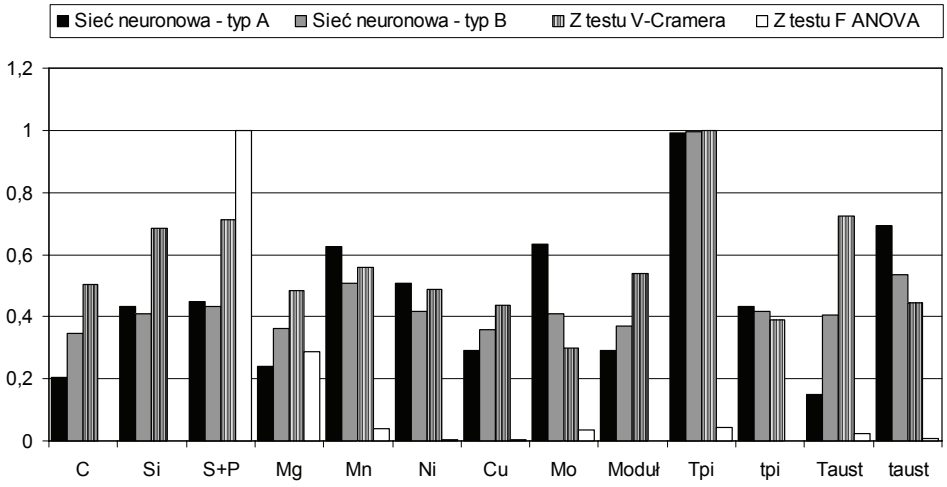
Na rysunku 9 pokazano wartości współczynników istotności względnej uzyskane dla zbioru przemysłowego „ZS". Obie metody oparte na sieciach neuronowych oraz metoda oparta na teście V-Cramera wskazały jednoznacznie na wyróżniające się znaczenie miedzi, która jest pierwiastkiem stosowanym w praktyce produkcyjnej w celu uzyskania struktury perlitycznej i tym samym zwiększenia wytrzymałości żeliwa sferoidalnego. Zbliżony rezultat uzyskano także dla metody wykorzystującej jednoczynnikową analizę wariancji, chociaż w tym przypadku wskazała ona na dwa inne pierwiastki, jako mające zbliżone, nawet nieco większe, znaczenie.

Zupełnie nietrafione wskazania otrzymano natomiast dla istotności obliczanych na podstawie współczynników wielomianu. Wydaje się jasne, że wielomian 2 stopnia nie „poradził sobie" ze złożonymi zależnościami, które zachodziły w badanym procesie. Jedną z przyczyn mogła być definicja współczynników istotności pojedynczych zmiennych, oparta tylko na wyrazach zawierających tylko pojedyncze zmienne; możliwe, że inne rezultaty uzyskano by z odpytywania wielomianu, przy którym uwzględniane byłyby także wyrazy mieszane.



Rysunek 9. *Porównanie istotności względnych składników chemicznych żeliwa sferoidalnego z punktu widzenia jego wytrzymałości na rozciąganie, wyznaczanych różnymi metodami*

Figure 9. *Significance factors of alloying components calculated by different methods, for tensile strength of ductile cast iron*



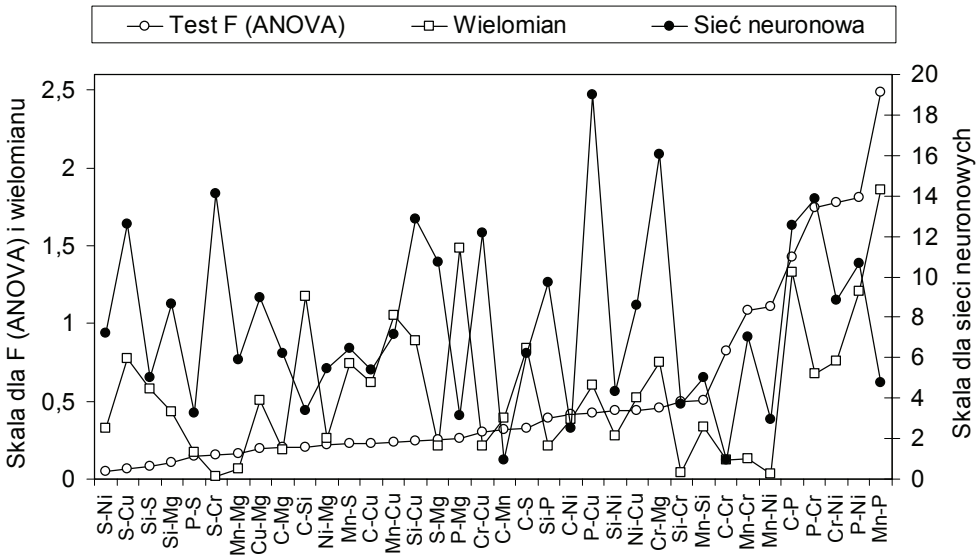
Rysunek 10. *Porównanie istotności względnych parametrów mających wpływ na wytrzymałość żeliwa typu ADI, wyznaczanych różnymi metodami*

Figure 10. *Significance factors of parameters influencing tensile strength of ADI type cast iron, calculated by different methods*

Na rysunku 10 pokazano wartości współczynników istotności względnej uzyskane dla zbioru przemysłowego „ADI". Tu także obie metody oparte na sieciach neuronowych oraz metoda oparta na teście V-Cramera wskazały jednoznacznie na wyróżniające się znaczenie tej samej wielkości, którą jest temperatura przesycania izotermicznego Tpi.

zie wariancji i współczynnikach wielomianu nie potwierdziły tego, wskazując na dwie inne, różne zmienne.

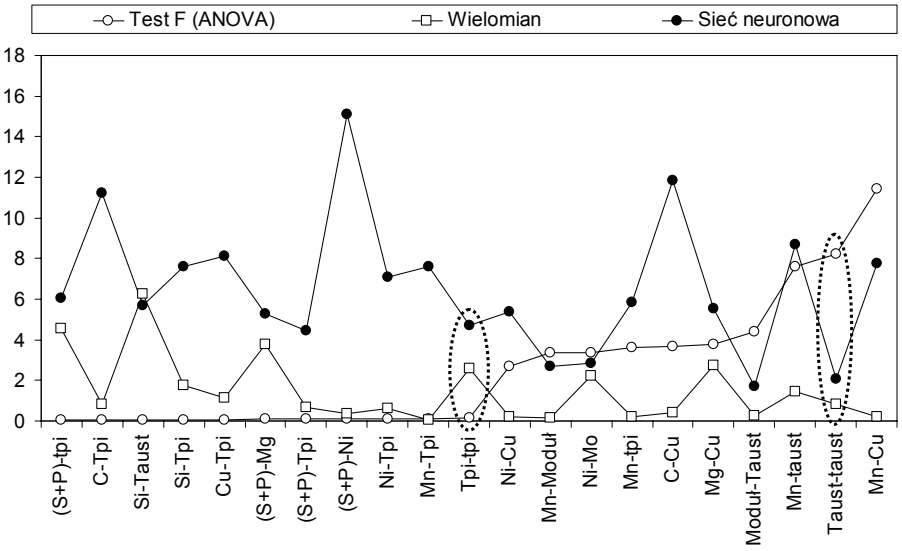
Na rysunku 11 pokazano wartości współczynników interakcji między wszystkimi możliwymi parami pierwiastków w żeliwie sferoidalnym, w postaci wykresu, na którym uszeregowano ich wartości



Rysunek 11. *Porównanie współczynników interakcji składników chemicznych żeliwa sferoidalnego z punktu widzenia jego wytrzymałości na rozciąganie, wyznaczanych różnymi metodami*

Figure 11. *Comparison of interactions coefficients between alloying components, calculated by different methods, influencing tensile strength of ductile cast iron*

Wskazanie to jest także zgodne z przewidywaniami, uzasadnionymi praktyką przemysłową i analizą procesu obróbki cieplnej żeliwa ADI. Pozostałe dwie metody, tj. oparte na jednoczynnikowej analirosnąco dla metody opartej na dwuczynnikowej analizie wariancji. Widoczny jest zupełny brak korelacji pomiędzy przewidywaniami uzyskanymi dla wszystkich trzech metod. Ponadto, dla żadnego



Rysunek 12. *Porównanie współczynników interakcji parametrów mających wpływ na wytrzymałość żeliwa typu ADI, wyznaczanych różnymi metodami*

Figure 12. *Comparison of interactions coefficients between parameters influencing tensile strength of ADI type cast iron, calculated by different methods*

wskazania o dużej wartości trudno jest mówić o jego fizycznym (metalurgicznym) uzasadnieniu.

Na rysunku 12 pokazano wartości współczynników interakcji między wybranymi parami parametrów wpływających na wytrzymałość żeliwa typu ADI, w postaci wykresu, na którym uszeregowano ich wartości rosnąco dla metody opartej na dwuczynnikowej analizie wariancji. Z uwagi na dużą liczbę kombinacji ograniczono się do 20 par (10 największych i 10 najmniejszych wartości otrzymanych metodą opartą na analizie wariancji). Podobnie jak dla zbioru „ZS" widoczny jest zupełny brak korelacji pomiędzy przewidywaniami uzyskanymi dla wszystkich trzech metod. Należy zwrócić uwagę, że pary parametrów obróbki cieplnej typu „temperatura – czas" powinny w zasadzie wykazywać znaczące współdziałanie. Na wykresie pary te zaznaczono elipsami (parametry przesycania izotermicznego: Tpi – tpi oraz parametry austenityzacji: Taust – taust). Jak widać, nie zostały one wyróżnione przez żadną z zastosowanych metod.

4. PODSUMOWANIE I WNIOSKI

W rezultacie przeprowadzonych obliczeń i analiz można sformułować następujące oceny i wnioski odnośnie zaproponowanych metod obliczania współczynników istotności względnej zmiennych niezależnych (parametrów procesu) oraz współczynników interakcji między dwiema zmiennymi:

- Obliczenia wykonane dla zbiorów danych symulowanych wyraźnie wskazały, że najlepszym współczynnikiem istotności pojedynczych zmiennych jest zaproponowany wcześniej przez autorów, wykorzystujący nauczoną sieć neuronową (typu B).
- Wyniki obliczeń dla dwóch zbiorów przemysłowych potwierdziły lepsze, w porównaniu do innych metod, właściwości współczynnika tego typu. Należy podkreślić, że rezultat ten uzyskano dla współczynników istotności wyznaczanych jako średnie z wielu uczeń, przy występowaniu znacznych rozrzutów wartości, podobnych zresztą do uzyskanych w innych pracach, np. (Yescas et al., 2001, Yescas, 2003).
- Współczynniki istotności oparte na jednoczynnikowej analizie wariancji (wykorzystywane m.in. w komercyjnym pakiecie *Statistica*), a także na teście V-Cramera, wymagają ulepszeń, zmierzających w kierunku lepszego ilościowego odzwierciedlenia istotności zmiennych.
- Najlepsze wyniki dla wykrywania interakcji zmiennych w zbiorach symulowanych, o znanych zależnościach, uzyskano w przypadku współczynnika opartego na odpytywaniu sztucznej sieci neuronowej. Zastosowanie wszystkich trzech typów współczynnika interakcji do zbiorów danych przemysłowych nie przyniosło pozytywnych wniosków.
- Współczynniki istotności i interakcji oparte na regresji wielomianowej w obecnej wersji wykazały, jak można było oczekiwać, istotne ograniczenia co do zdolności wykrywania prawidłowości w bardziej złożonych zależnościach między zmiennymi.

Potrzebne jest prowadzenie dalszych badań, mających m.in. na celu:

- ulepszenie definicji współczynników istotności opartych na metodach statystycznych,
- przedefiniowanie współczynników interakcji wykorzystujących dwuczynnikową analizę wariancji,
- ulepszenie definicji współczynników interakcji wyznaczanych z wykorzystaniem sieci neuronowych w taki sposób, ażeby zredukować wpływ innych zmiennych,
- sformułowanie definicji i ocena współczynników interakcji pomiędzy większą liczbą zmiennych.

LITERATURA

Braha, D. (ed.), 2001, *Data Mining for Design and Manufacturing – Methods and Applications*, Kluwer Academic Publ., Dordrecht, Boston, London.

Demski, T, 2004, *Data mining w przemyśle: projektowanie, udoskonalanie, wytwarzanie, Statystyka i data mining w praktyce*, StatSoft, Warszawa – Kraków 2004, 207–221.

Fujii, H., MacKay, D.J.C., Bhadeshia, H.K.D.H., 1996, Bayesian Neural Network Analysis of Fatigue Crack Growth Rate in Nickel Base Superalloys, *ISIJ Int.*, 36, 1373–1382.

Hand, D., Mannila, H. Smyth P., 2005, *Eksploracja danych*, WNT Warszawa.

Narayan, V., Abad, R., Lopez, B., Bhadeshia, H.K.D.H., MacKay D.J.C., 1999, Estimation of hot torsion stress strain curves in iron alloys using a neural network analysis, *ISIJ International*, 39, 999–1005.

Perzyk, M., Kochański A., 2001, Prediction of ductile cast iron quality by artificial neural networks, *J. Mat. Proc. Techn.*, 109/3, 305–307.

Perzyk, M., Kochański A., 2003, Detection of causes of casting defects assisted by artificial neural networks, *J. Eng. Manuf.*, 217B, 1279–1284.

Perzyk, M., Kochański A., 2003, Istotność względna sygnałów wejściowych sieci neuronowej, *Informatyka w Technologii Materiałów*, 3, 125–132.

Perzyk, M., Kochański A., 2002, System komputerowy wspomagający projektowanie procesów metalurgicznych

wykorzystujący modelowanie sztucznymi sieciami neuronowymi, w monografii *Polska metalurgia w latach 1998 – 2002*, red. K. Świątkowski, tom 2, wyd. Komitet Metalurgii PAN, Kraków, 236–242.

Perzyk, M., Biernacki, R., 2004, Modelowanie procesów produkcyjnych za pomocą naiwnego klasyfikatora Bayesa i sztucznych sieci neuronowych, *Informatyka w Technologii Materiałów*, 4, 98–104.

Perzyk, M., Biernacki, R., Kochański, A., 2005, Modeling of manufacturing processes by learning systems: the naive Bayesian classifier versus artificial neural networks, *J. Mat. Proc. Techn*., 164–165, 1439–1435.

Warde, J., Knowles, D.M., 1999, Application of Neural Networks to Mechanical Property Determination of Ni-base Superalloys, *ISIJ Int*., 39, 1006–1014.

Yescas, M.A., Bhadeshia, H.K.D.H., MacKay D.J.C., 2001, Estimation of the amount of retained austenite in austempered ductile irons using neural networks, *Mat. Sci. Eng*., A311, 162–173.

Yescas, M.A., 2003, Prediction of the Vickers hardness in austempered ductile irons using neural networks, *Int. J. Cast Metals*, 15, 513–521.

Received: May 28, 2006
Received in a revised form: September 16, 2006
Accepted: October 14, 2006