

ZASTOSOWANIE ANALIZY SKUPIEŃ W PROCESIE NABORU DO PŁYWANIA SPORTOWEGO

Robert Rocznik¹

Podstawowym celem pracy było zweryfikowanie przydatności metod aglomeracyjnych we wspomaganie procesu naboru do pływania sportowego na dystansach sprinterskich 50m. Przeprowadzone zostały analiza skupień, analiza wariancji otrzymanych podzbiorów, analiza jednorodności wariancji dla wskaźników jakości 50m, oraz testy wielokrotnych porównań. Zbudowany dendrogram przy wykorzystaniu tylko wybranych cech charakteryzujących kandydata do szkolenia sportowego, pozwolił na wyodrębnienie trzech jednorodnych podgrup. Jakość tego podziału potwierdził wskaźnik jakości na dystansie 50m, który nie był wykorzystany przy budowie dendrogramu. Zdecydowanie najlepsze wyniki uzyskują zawodnicy z trzeciej podgrupy, najgorsze wyniki stwierdzono zaś w podgrupie pierwszej. Trafność podziału na trzy podgrupy potwierdzają również dalsze analizy, analiza wariancji ($p=0,0000002$), która potwierdza, że są istotne różnice pomiędzy średnimi w otrzymanych podgrupach, oraz testy wielokrotnych porównań. Otrzymane wyniki pozwalają z całą pewnością stwierdzić, że grupą wykazującą najlepsze wyniki po 12 miesięcznym treningu jest grupa 3.

Słowa kluczowe: *analiza skupień, analiza wariancji, optymalizacja, nabór.*

Wstęp

Współczesny trening sportowy jako długotrwały i złożony proces wymaga stałego dopływu informacji warunkującej jego efektywność. Te

¹ Katedra Analiz Systemowych w Sporcie, AWF Katowice. Opiekun naukowy: prof. dr hab. Igor Ryguła.

elementy są wykorzystywane w sterowaniu treningiem, rozumianym jako wydawanie decyzji w oparciu o informacje dotyczące stanu procesu sterowanego oraz znajomości celu sterowania [Raczek 1989]. Na efektywność procesu szkolenia wpływa wiele czynników, jednym z nich jest proces selekcji sportowej. Jej istota polega na określeniu wektora uzdolnień kandydata do szkolenia sportowego, co pozwala na zwiększenie efektywności szkolenia sportowego [Ryguła 2002].

Od wielu lat poszukuje się narzędzi, które pozwoliłyby na określenie predyspozycji osobnika do osiągania wysokich wyników sportowych [Kozioł 1984]. Na kolejnych etapach szkolenia stosuje się zróżnicowane kryteria selekcji, które mają pomóc określić szanse zawodnika na zdobywanie wysokich rezultatów [Płatonow 1997, Raczek 1989]. Coraz większe znaczenie w tej dziedzinie zajmują metody statystyczne oraz matematycznego prognozowania [Komor 1982, Mester, Perl 1999, Perl i inn. 2002]. Do nich zaliczamy wielowymiarowe techniki eksploracyjne, których zastosowanie w obszarze nauk o sporcie jest sporadyczne. Związane są one z rozpoznawaniem obiektów oraz ich grupowaniem (analiza taksonomiczna) [Witt, Burton 1996], analizą skupień, analizą czynnikową, analizą dyskryminacyjną [Ryguła 2003].

Wykorzystując wielowymiarowe techniki eksploracyjne do optymalizacji procesu naboru, mamy możliwość grupowania obiektów, bądź grupowania cech. W obydwu przypadkach znajdujemy podzbiory bardziej jednorodnych jednostek taksonomicznych. W odniesieniu do cech są to podzbiory przenoszące podobną informację o zawodnikach (Ryguła 2003).

Cel badań

Podstawowym celem pracy było zweryfikowanie przydatności metod aglomeracyjnych we wspomaganiu procesu naboru do pływania sportowego na dystansach sprinterskim 50m.

Problem badawczy:

1. Czy metody aglomeracyjne mogą wspomagać procesu naboru?
2. Czy zastosowana analiza skupień pozwoli na wyodrębnienie jednorodnych taksonomicznie podgrup zawodników bardzo dobrych, średnich oraz słabych?

Materiał, metody i narzędzia badań

Materiał do analiz stanowiły wyniki uzyskane przez grupę 40 zawodników uprawiających pływanie z terenu Makroregionu Śląskiego w wieku 12 lat.

Główne badania poprzedzone były wstępnym rocznym przygotowaniem sprawnościowym. Badani uczestniczyli w zajęciach treningowych trzy razy w tygodniu. Dwa razy na pływalni (45 min) oraz raz na sali gimnastycznej (90 min). Zajęcia miały charakter ogólny i uzupełniający. Celem zajęć na pływalni była nauka pływania.

W trakcie badań pomiarom poddano następujące cechy:

Zmienne stanu: X1 - Skok w dal z miejsca [cm], X2 - Krok pływacki [ilość pełnych cykli na odcinku 25 m], X3 - Wysokość ciała [cm], X4 - Ciężar ciała [kg], X5 - Pojemność życiowa płuc [cm³]. Do obliczeń wykorzystano pomiary wykonane przy naborze do pływania sportowego po rocznym wstępnym przygotowaniu sprawnościowym. Wskaźnik jakości: rezultat na dystansie 50m stylem dowolnym. Czas przepłynięcia przeliczona na punkty według tabel wielobojowych. Uwzględniono pomiary wykonane po 12 miesiącach sportowego treningu pływackiego.

Prowadzone badania miały charakter empiryczny oraz eksploracyjny o charakterze przekrojowym, zatem podstawową metodą badawczą była obserwacja bezpośrednia-uczestnicząca, która była uzupełniana poprzez sondaż diagnostyczny.

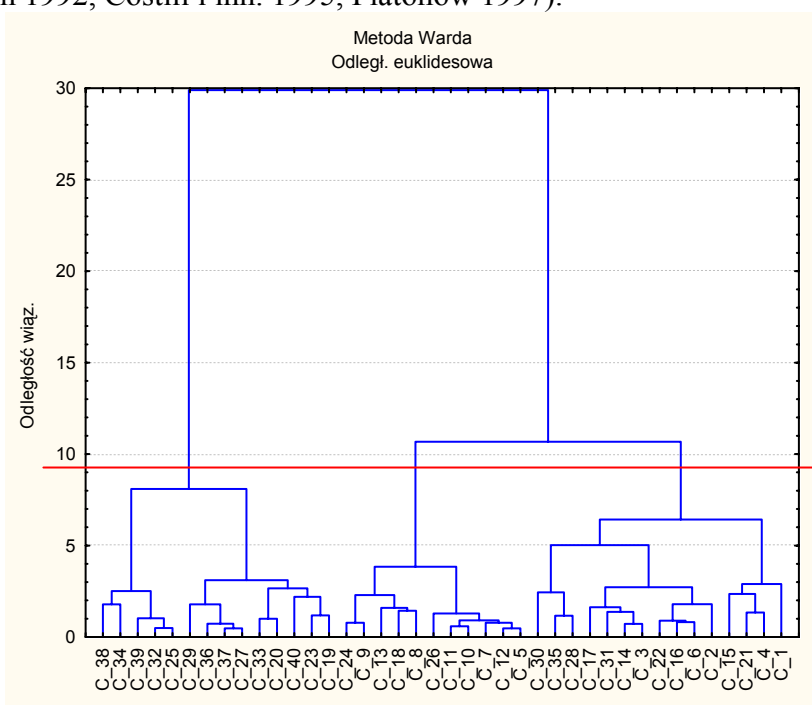
Wyniki badań

W oparciu o dane pomiarowe, dla których zastosowano klasyczną formułę standaryzacji przeprowadzona została analiza skupień, wynikiem której powstał dendrogram uzyskany metodą Warda. Ponieważ zmienne oryginalne miały charakter zmiennych ciągłych, zastosowano odległość euklidesową jako miarę odległości w wykonywanej analizie:

$$\text{odległość euklidesowa } (x,y) = \{\sum_i (x_i - y_i)^2\}^{1/2}$$

Przeprowadzone zostały również: analiza wariancji otrzymanych podzbiorów, analiza jednorodności wariancji dla wskaźników jakości 50m, oraz testy wielokrotnych porównań.

Skonstruowany model nie uwzględniał wpływu treningu sportowego na uzyskany poziom rozwoju. Oceniając potencjalne możliwości kandydata przyjęto założenie o swoistości treningu sportowego pływaków prowadzonego według metodyki opisanej w literaturze (Colwin 1992, Costill i inn. 1995, Płatonow 1997).



Rys. 1. Diagram drzewa

Tabela 1

Wskaźniki jakości

n	50 m [pkt.]	n	50 m [pkt.]
1	230	21	264
2	233	22	215
3	200	23	140
4	177	24	46
5	174	25	38
6	214	26	231
7	108	27	139
8	123	28	214
9	78	29	83
10	157	30	309
11	158	31	227
12	175	32	103
13	62	33	111
14	248	34	57
15	102	35	239
16	218	36	99
17	133	37	128
18	75	38	73
19	145	39	84
20	88	40	99

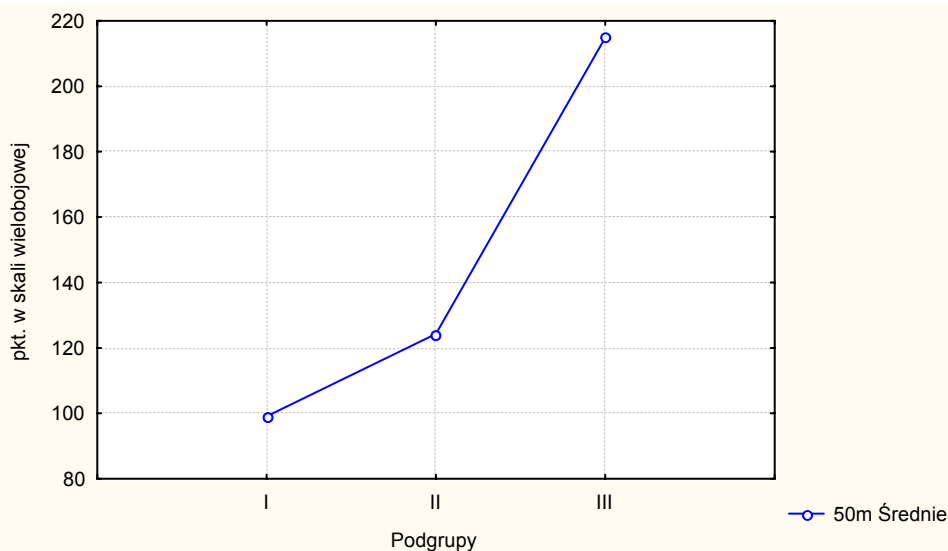
Tabela 2

Porównanie trzech grup zawodników uzyskanych w wyniku zastosowania analizy skupień

Cecha	Jednostka miary	Średnia Grupy I	Średnia Grupy II	Średnia Grupy III	p Levene'a $\alpha=0,05$	p ANOVA $\alpha=0,05$
50 m	pkt.	99,07	124,27	214,87	0,20	0,0000002

Analiza skupień została wyliczona w oparciu o zmienne stanu, których pomiary wykonano przy naborze do pływania sportowego. Na podstawie tych danych został stworzony dendrogram metodą Warda. Na podstawie oceny struktury dendrogramu uznano, że najodpowiedniejszą

odległością aglomeracyjną poziomu odcinającego będzie 9. Tworzą się wówczas trzy podgrupy zawodników. Patrząc od lewej strony dendrogramu: pierwsza z nich obejmuje 14 zawodników, druga 11 zawodników i trzecia 15 zawodników. Dla powstałych trzech podgrup do dalszych obliczeń wykorzystane zostały wskaźniki jakości, które uzyskano po 12 miesięcznym okresie treningu (wyniki na 50m). Wyniki te nie były wykorzystywane przy tworzeniu dendrogramu (weryfikacja metody aglomeracyjnej) i dla tych wyników przeprowadzone zostały analizy statystyczne dotyczące jednorodności wariancji a następnie analizy polegające na testowaniu równości w otrzymanych podgrupach. Dane te stanowią podstawę do weryfikacji hipotezy, że średnie w podgrupach są jednakowe $H_0: m_1 = m_2 = m_3$, wobec hipotezy alternatywnej: **H_1 : co najmniej dwie średnie różnią się między sobą.** Brak podstaw do odrzucenia hipotezy zerowej prowadzi do stwierdzenia, że metody aglomeracyjne nie są skuteczne jako narzędzie wspomagające proces naboru. Po przeprowadzeniu analiz stwierdzających jednorodność wariancji (test Levene'a) nie ma podstaw do odrzucenia hipotezy zerowej o równości wariancji $p = 0,2$, przeprowadzono jednoczynnikową analizę wariancji, która pozwoliła na odrzucenie hipotezy zerowej i przyjęcie hipotezy alternatywnej zarówno dla wyników otrzymanych na dystansie 50m. Po odrzuceniu hipotezy zerowej przeprowadzono dalsze analizy statystyczne wykorzystujące testy post-hoc (testy wielokrotnych porównań) pozwoliły na stwierdzenie, że istotne różnice występują pomiędzy grupami I a III oraz II a III, nie otrzymano natomiast istotnych statystycznie różnic pomiędzy grupami I a II.



Rys. 2. Średnie arytmetyczne wskaźników jakości

Zbudowany dendrogram przy wykorzystaniu tylko wybranych cech charakteryzujących kandydata do szkolenia sportowego, pozwolił na wyodrębnienie trzech jednorodnych podgrup. Jakość tego podziału potwierdziły wskaźniki jakości (wyniki na dystansach 50), które nie były wykorzystywane przy budowie dendrogramu. Zdecydowanie najlepsze wyniki uzyskują zawodnicy z trzeciej podgrupy, najgorsze wyniki stwierdzono zaś w podgrupie pierwszej.

Trafność podziału na trzy podgrupy potwierdzają również dalsze analizy, analiza wariancji, która potwierdza, że są istotne różnice pomiędzy średnimi w otrzymanych podgrupach, czyli potwierdza skuteczność zastosowanej metody aglomeracji. Dalsze analizy pozwalają na dokładne stwierdzenie, że istotnie różnią się grupy I i III oraz II i III, oraz nie ma podstaw do stwierdzenia istotności różnic pomiędzy grupą I i II. Otrzymane wyniki pozwalają z całą pewnością stwierdzić, że grupą wykazującą najlepsze wyniki po 12 miesięcznym treningu jest grupa 3. Przy następnym naborze wykorzystując wielkości charakteryzujących budowę ciała kandydata oraz wyników testów sprawnościowych

zawodnicy, którzy znajdą się w tej grupie będą dawali większe nadzieje na uzyskiwanie dobrych wyników sportowych.

Dużą trudność stanowi porównanie uzyskanych wyników z doniesieniami innych autorów ze względu na brak publikacji poruszających praktyczne zastosowanie metod analizy skupień w procesie selekcji. Jediną pracą w tym zakresie są analizy przeprowadzone przez Rygułę [2002], w której omawiane metody zastosowano w ocenie zawodniczek trenujących piłkę ręczną. Autor wykazał duże znaczenie tych metod w naukach o sporcie, jednakże w pracy nie określono wyraźnie wskaźnika jakości do którego należałoby odnieść uzyskane wyniki badań.

Podsumowując przeprowadzone pomiary oraz obliczenia statystyczne można stwierdzić, że wielowymiarowe techniki eksploracyjne mogą być wykorzystywane do wspomagania procesu naboru do pływania sportowego zwiększając jego efektywność.

Piśmiennictwo

1. Colwin C.M. 1992. *Swimming into 21-st Century*. Leisure Press, Champaign, Illinois.
2. Costil D.L., Maglisco E.W., Richardson A.B. 1995. *Swimming*. Blackwell Scienceetific Publicatory, Oxford.
3. Komor J.A. 1982. *Zastosowanie metod modelowania w sporcie*. Instytut Sportu, Warszawa.
4. Kozioł R. 1984. *Podjęcie systemowo-cybernetyczne w treningu i walce sportowej*. Zeszyty Naukowe, AWF, Kraków, nr 49.
5. Mester J., Perl J. 1999. *Unconventional simulation and empirical evaluation of biological response to complex high training loads*. [In:] Parisi P., Pigozzi F., Prinzi G. (Eds.): *Sport Science'99 in Europe*. Rome.
6. Perl J., Lames M., Glitsch U. 2002. *Modellbildung in der Sportwissenschaft. Beiträge zur Lehr und Forchung im Sport*, Bd 132, Schorndorf, Hofmann.
7. Płatonow W.N. 1997. *Trening wyczynowy w pływaniu*. Biblioteka Trenera. Warszawa.

8. Raczek J. 1989.: Szkolenie młodzieży w systemie sportu wyczynowego. AWF Katowice.
9. Ryguła I.(red) 2002. Elementy teorii, metodyki, diagnostyki i optymalizacji treningu sportowego. AWF Katowice.
10. Ryguła I. 2003. Proces badawczy w naukach o sporcie. AWF Katowice.
11. Wit A., Barton G. 1996. Sztuczne sieci neuronowe. Sport Wyczynowy nr 9.

Summary

A constructed dendrogram with chosen features characteristic of a candidate for sports training allowed to select three homogenous sub-groups. The quality of that division was confirmed by the quality indicator results (at distance of 50m) which were not used during the construction of dendrogram. The best results were achieved in the third sub-group, the worst results were in the first group.

Accuracy of the division for 3 sub-groups are certified by later analysis of variation, which confirm that there are essential differences between the groups, so it the efficiency of applied agglomeration method. Further analysis justifies point to essential differences is in groups I and III as well as II and III. Achieved results allow to state that the group with best results after 12 months of training is group 3.

It is difficult to compare the achieved results with other authors because of the lack of data dealing with the practical application of the methods of analysis of concentration in the process of selection. The only work is presented by Ryguła (2002) in which discussed methods were used for the assessment of participants practicing handball. The author showed the importance of these methods in sport sciences, but in there is no clear quality factor to which the achieved results should be applied.

