Journal of Physical Education and Sport ® (JPES), Vol 20 (Supplement issue 6), Art 441 pp 3253 – 3259, 2020
online ISSN: 2247 - 806X; p-ISSN: 2247 – 8051; ISSN - L = 2247 - 8051 © JPES

Original Article

Classification of athletic decathlon using methods of hierarchical analysis

JAROSLAV BROĎÁNI[1], NATÁLIA DVOŘÁČKOVÁ[2], MONIKA CZAKOVÁ[3]
[1,2,3,]Department of Physical Education and Sport, Faculty of Education, Constantine the Philosopher University in Nitra, SLOVAKIA

Published online: November 30, 2020
(Accepted for publication: November 22, 2020)
DOI:10.7752/jpes.2020.s6441

Abstract
Authors deals with the problematics of group classification of athletics disciplines, which influence the sports performance in the men's decathlon. For the group identification, the indicators of the best world's performance in decathlon above the 8000 points according to the data from IAAF (1986-2019) were used. From the classification methods of clustering the hierarchical models as the Average linkage (Between & Within-group), Single Linkage - Nearest neighbor, Complete Linkage - Farthest neighbor, Centroid linkage, Median clustering, and Ward´s method were used. The substructure of the clusters is differentiated according to the relations of used methods. All seven clustering methods agreed in four groups of disciplines in 4 cluster [100 meters, 400 meters, Long jump, 110 meters hurdles, Pole vault] [High jump] [1500 meters]. Stability test with a substructure of decathlon of 4th. cluster is identified in 85.71% cases. Hierarchical models allow identifying groups of athletics disciplines that influence the sports performance in men's decathlon. Understanding the structure of sports performance contributes to the streamlining of the training process and determining the typology of top-class decathletes.
Keywords: decathlon, men, cluster analysis

Introduction

Decathlon requires great motor and technique skills of every single individual. The individuals should have a relatively tall body disposition with great speed-strength and endurance assumptions and feeling for the movement coordination (IhringandHoráček, 1978;Ryba 2002). The essential tasks of decathlon training include the development of special factors of motoric abilities and improvement of the technique of individual decathlon disciplines (Koukal, 1985; Vindušková and Koukal, 2003). The division of decathlon disciplines plays an important role here. It is well known from the experiences that after a 100 meters and Long jump it is difficult to compete in Shot put. There are considerable problems with the 1500 meters as the final endurance discipline in the decathlon. Individuals with an innate disposition for the long-term cyclic work have an advantage here. Decathlons on the lower performance level indicate the low versatility of the athletes (Broďáni, 2006). It is well known that in the lower performance level, there is a preference for running and jumping disciplines over technical-throwing disciplines.

Decathlon is a complex athletics discipline in which is necessary to respect the requirements of individual disciplines (Vinduškováand Koukal, 1987). The top-class decathlon is by its specificity and structure share of individual disciplines characterized by versatility but also by unilaterality of some disciplines (Cox and Dunn, 2002; Dawkins et al. 1994; Shen and Huang, 2012). Complex relationships and connections in the performance structure of decathlon point to the different focus and content of training units at different levels of sports performance (Schomaker and Heumann, 2011; Zhang and Lu 2014). Knowing the share of individual disciplines in the decathlon at different levels of sports performance creates space for a better understanding of the complexity of the disciplines and to better understand the difficulty in the training process. The clarification of the structure of performance by individual decathlon disciplines allows trainers and athletics coaches to be better oriented in the intricate structure of relationships and connections, and it also allows them to understand the share of preparedness in each discipline (Kenny et al., 2005; Park and Zatsiorsky 2011;Zatsiorsky and Godik, 1962; Woolf et al., 2007). The isolation of every common in the character of decathlete´s trainability, determining the hierarchy of requirements for the training process in addition to observing the general rules, it is necessary to create adequate space for respecting the athlete's personality and his advantages and disadvantages (Pavlovic andIdrizovic, 2017;Etcheverry 1995;Bilic et al., 2015;Koštiaľ, 1993; Walker andCaddigan 2015; Wang and Lu, 2007).

Aim

The aim of this article was to identify groups of athletic disciplines that affect the sport performance in men´s decathlon with an average performance of 8282.75 points. The secondary aim was to explain the typology of top-class combined event athletes and the focus on the training process.

Corresponding Author: MONIKA CZAKOVÁ, E-mail: monika.czakova@ukf.sk

Material &Methods

The accumulated empirical material is a collection of the sports performance parameters in the decathlon of N = 255 world's best performances based on available data from IAAF to December 31. 2019 (1986 - 2019), with the average sports performance of M = 8282.75 points, with the standard deviation SD = 232.74 (Max = 9126 points, Min = 8001 points) see table 1. For the analysis, the point values of all disciplines from each day were used. From the first day it was data from 100 meters (R100m), Long jump (LJ), Shot put (SP), High jump (HJ), 400 meters (R40m), and from the second day 110 meters hurdles (R110mh), Discus throw (DT), Pole vault (PV), Javelin throw (JT), and 1500 meters (R1500m).

Table 1: Characteristics of sports performance of men decathletes with the average sports performance of 8282.74 points

	M	SD	Min	Max
Marks (points)	**8282.75**	**232.74**	**8001.00**	**9126.00**
Run 100 m (s)	10.90	0.25	10.23	11.73
Long jump (m)	7.43	0.26	6.72	8.11
Shot put (m)	14.66	0.92	12.08	17.32
High jump (m)	2.01	0.07	1.76	2.21
Run 400 m (s)	48.88	1.07	45.00	53.01
Run 110 m hurdles (s)	14.47	0.38	13.47	16.11
Discus throw (m)	44.77	3.48	35.58	55.22
Pole vault (m)	4.90	0.25	4.30	5.70
Javelin throw (m)	61.02	5.60	46.75	79.05
Run 1500 m (m:ss.00)	4:34.87	0:11.39	3:59.13	5:05.41
Run 100 m (points)	884.40	56.78	705.00	1040.00
Long jump (points)	919.01	63.18	748.00	1089.00
Shot put (points)	769.20	56.48	611.00	933.00
High jump (points)	817.58	68.19	593.00	1002.00
Run 400 m (points)	867.57	50.68	681.00	1060.00
Run 110 m hurdles (points)	914.51	47.73	720.00	1044.00
Discus throw (points)	762.45	71.70	575.00	980.00
Pole vault (points)	879.82	76.80	702.00	1132.00
Javelin throw (points)	753.52	84.36	541.00	1028.00
Run 1500 m (points)	714.70	72.75	529.00	960.00

The clustering algorithms for the classification of groups of athletics disciplines in the structure of decathlon were used. A cluster hypothesis is a suitable way to express the quality of clustering by the degree of separation of the described distribution (TrebuňaandBeres, 2010). Hierarchical clustering methods are based on individual objects (each object creates a primary cluster). By their connecting, the number of clusters decreases gradually at each step until all the clusters eventually come together in one unit (however, the procedure can also be the opposite: from the initial-all object involving a cluster-smaller and smaller clusters are gradually divided up to the individual objects). Hierarchical methods leading to the hierarchic (tree) structure which is graphically displayed as a tree diagram - dendrogram (fig. 1). The objects in the dendrogram are ordered in order to follow the gradual joining of objects into clusters. The classification of the clusters in groups of athletics disciplines and sports performance in combine athletics events are also discussed in works of Bilić (2015), BilićandSmajlović (2012), Cox and Dunn (2002), Dziadek (2018), Fröhlich et al. (2015), Gassmann, Fröhlich andEmrich (2016), Pavlovič (2017), SchomakerandHeumann (2011), Trevor et al. (2002), Woolf, Ansley and Bidgood et al. (2007).

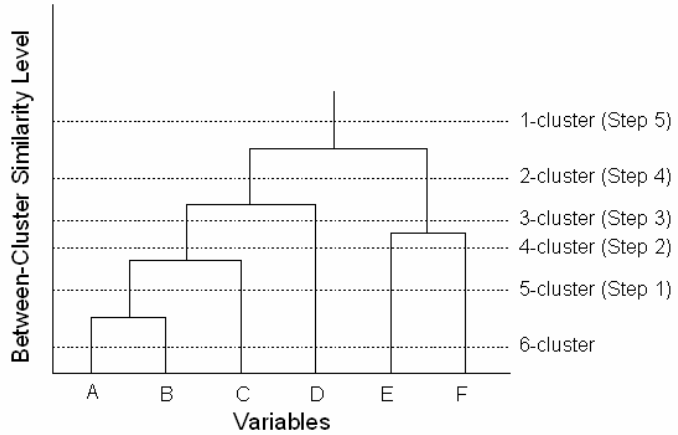

Figure 1:Tree diagram – dendrogram

There are several clustering methods that differ in specific procedures and can lead to different results, but they are mathematically statistically equivalent (therefore, it is recommended to test several methods and select the optimal one according to the reached results). From the clustering classification methods were used in our work hierarchical models as Average linkage (Between & Within groups), Single Linkage - Nearest neighbor, Complete Linkage - Farthest neighbor, Centroid linkage, Median clustering, Ward´s method. To determine the distance between variables, we chose the square Euclidean distance recommended by Wolf et al. (2007) because of greater consistency than in the methods based on Pearson correlation. The results were processed in MS Excel and IBM SPSS. Findings and conclusions are formulated based on the logical evaluation of the obtained results.

Results

By finding a match between methods and comparing dendrograms at all levels, we found that all seven methods created the same solutions in the 4th cluster (tab. 2, fig. 3-6). Hierarchical methods created four clusters of disciplines [100m, 400m, LJ, 110mh, PV] [SP, DT, JT] [HJ] [1500m]. The largest cluster consists of three sprint disciplines (100 meters, 400 meters, 110 meters hurdles) and two jumping disciplines (Long jump and Pole vault). All disciplines are conditioned by speed and speed-strength abilities. The second cluster consists of three throwing disciplines (Shot put, Discus throw, Javelin throw). All methods set aside High jump and 1500 meters into the single clusters. In the 2. cluster only 4 methods were identical, assigning a High jump and 1500 meters to the second cluster of disciplines [100 m, 400m, LJ, 110mh, PV] [SP, DT, JT, HJ 1500m]. We have identified five hierarchical methods in the 3rd and 5th clusters. In 3rd cluster was the High jump assigned to the second group of disciplines, were 3 clusters of disciplines were created [100m, 400m, LJ, 110mh, PV] [SP, DT, JT, HJ] [1500m]. In the first step of clustering (5th cluster) was set aside the Pole vault, Javelin throw, and 1500 meters into the single clusters whilst the sprint disciplines and Long jump create the largest groups along with the throwing disciplines [100m, 400m, LJ, 110mh] [PV] [SP, DT, JT] [HJ] [1500m].

Table 2: Clusters of decathlon disciplines and matching hierarchical models

Cluster	Clusters of disciplines	Methods
2	[100m, 400m, LJ, 110mh, PV] [SP, DT, JT, HJ 1500m]	1, 5, 6, 7
3	[100m, 400m, LJ, 110mh, PV] [SP, DT, JT, HJ] [1500m]	1, 3, 5, 6, 7
4	**[100m, 400m, LJ, 110mh, PV] [SP, DT, JT] [HJ] [1500m]**	**1, 2, 3, 4, 5, 6, 7**
5	[100m, 400m, LJ, 110mh] [PV] [SP, DT, JT] [HJ] [1500m]	1, 3, 5, 6, 7

Method codes: 1. Average linkage (Between groups), 2. Average linkage (Within groups), 3. Single Linkage - Nearest neigbor, 4. Complete Linkage - Farthest neigbor,5. Centroid linkage, 6. Median clustering, 7. Ward´s method.

The classic scaling graph displays groups of decathlon disciplines on the sides of graph and presenting the clustering options of 2-5 cluster (fig. 2)

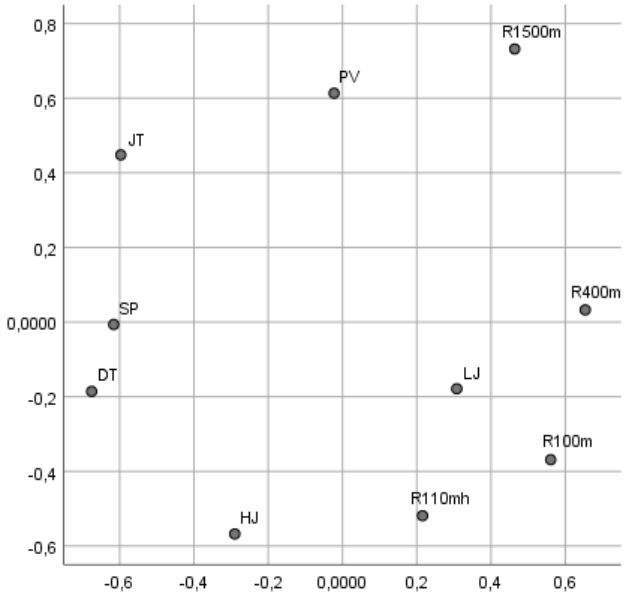

Figure 2:Classic scaling graph of decathlon disciplines from the 255 best performances of the world.

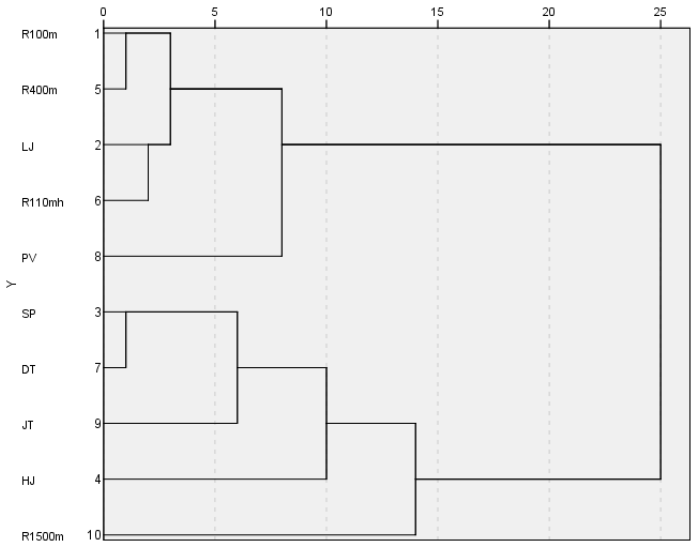

Figure 3: Dendrogram presenting the same representation by methods of Average linkage (Between Groups), Median linkage and Ward linkage

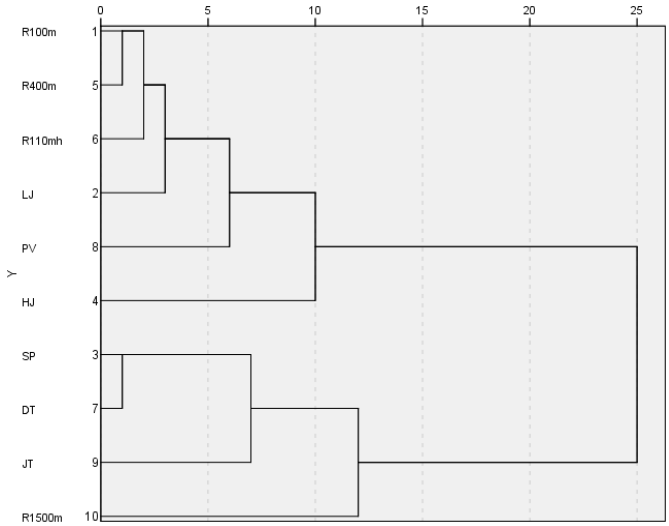

Figure 4:Dendrogram by the method of Average linkage (Within Groups)

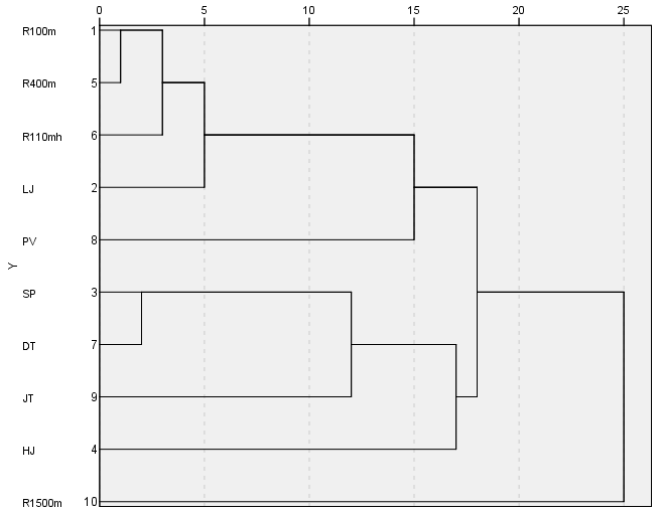

Figure 5: Dendrogram by the method of Single Linkage - Nearest neighbor

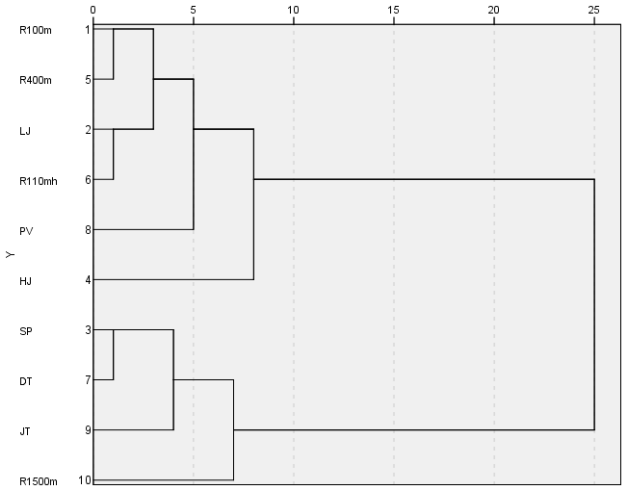

Figure 6: Dendrogram by the method of Complete Linkage - Farthest neighbor

Stability test

We examined the stability of the 4th cluster calculation by the series of tests by using all seven methods. Decathlon sample was divided into two groups containing the first 156 data (8001-8295 points) and the remaining 99 data (8302-9126 points). The dendrogram was obtained from each subgroup. All subgroups reproduced the same solutions in four clusters of disciplines [100m, 400m, LJ, 110mh, PV] [SP, DT, JT] [HJ] [1500m] by the methods of Average linkage (Between groups), Complete Linkage - Farthest neighbor, Median clustering. The consistency showed tests of Single Linkage - Nearest neighbor and Centroid linkage in the subgroup of 8001-8295 points, and tests of Average linkage (Within Groups) and Ward´s method in the subgroup of 8302-9126 points (fig. 7). The remaining methods did not have consistency with the above structure of clusters in the 4th cluster. The stability test with the substructure of decathlon in the 4th cluster is 85,71%.

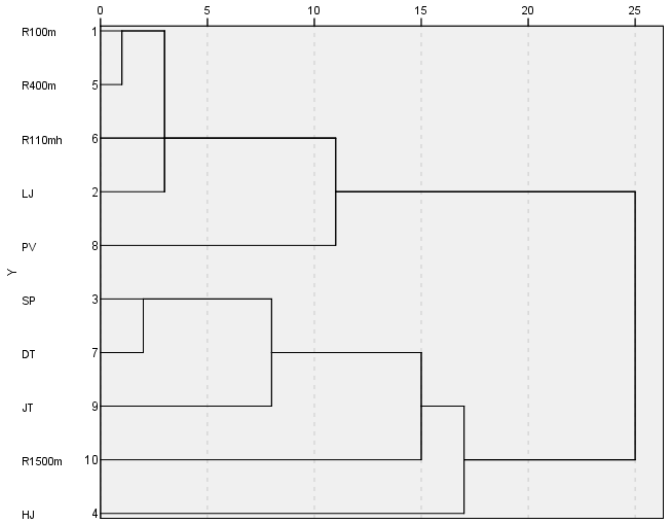

Figure7:Dendrogram representing the equal display of stability tests by the methods of Average linkage (Between Groups), Median linkage a Ward linkagein two differentiated subgroups (8001-8295 points and 8302-9126 points)

Discussion

Our results are following the results of authors Woolf, Ansley and Bidgood (2007), who confirmed through the six hierarchic methods the dominancy of running and jumping disciplines in 173 top-class decathletes from the year 1986 to 2005. At the level of three clusters, they presented clusters of disciplines: 3-cluster ([100m 400m 110mH LJ PV HJ] [SP DT JT] [1500m]), 4-cluster ([100m 400m 110mH LJ PV] [SP DT JT] [HJ] [1500m]) and 5-cluster ([100m 400m 110 mH LJ] [SP DT JT] PV] [HJ] [1500m]).

On the other hand, Cox and Dunn (2002) claims that dominant athletes in decathlon are those who are successful in the throwing and jumping disciplines. Similar results were reached in the study of Broďáni et al. (2020) where the focus was on the versatility of athletes with an emphasis on the throwing and jumping disciplines in average performance over 8200 points [JT 13.34 %, DT 12.85 %, LJ 13.25 %, HJ 10.73 %, SP 10.57 %, 110mH 9.64 %,

100m 9.55 %, PV 9.10 %, 400 m 8.38 %]. According to Koštiaľ (1993), in the top ten decathletes with an average performance of 8490.28 points, two disciplines from the first day participated in their final performance LJ 15.24 %, SP 13.41 %, DT 12.82 %, 110mH 12.63 %, PV 11.13 %). Decathlon with average performance at a lower level around 7219.91 points refer to a low versatility of decathletes and preference of running and jumping disciplines (100 m 17.62 %, LJ 15.94 %, PV 14.21 %, 110mH 13.85 %) over technical-throwing disciplines (Broďáni, 2006).

Pavlović and Idrizović (2017) in their study analyzed world record holders between 1984 and 2016. Through the factor analysis, they defined the typology of decathletes as a jumper-thrower-runner [PV, HJ, DT, SP, 400m, 110mH]. By the cluster analysis of top-class decathletes and their performances, Pavlović (2017) classified in the first group the runner-jumper type of decathlete through disciplines [100m, 110mH, 400m, PV, LJ]. In a similar study Park andZatsiorkij (2011) identified the prediction factors utilizing factor analysis. According to them, the performance in the decathlon was mainly influenced by disciplines [100m, 400m, 110mH, LJ]. In a complex relationship structure between disciplines of the decathlon, we can characterize today´s athletes as versatile with high performance in running and jumping disciplines conditioned by a quality performance in throwing disciplines.

Conclusion

Hierarchical models allow identifying the groups of athletics disciplines affecting the sports performance with an average of 8282,75 points in men's decathlon.

Utilized hierarchical clustering methods selected four groups of disciplines on the level of 4th cluster [100 meters, 400meters, Long jump, 110 meters hurdles, Pole vault] [Shot put, Discus throw, Javelin throw] [High jump] [1500 meters].

The first group is made of sprint-jumping disciplines with high speed and technical potential. Throwing disciplines made the second group. Individual groups were created by High jump and 1500 meters. The stability of the 4th cluster calculation is examined by a series of tests. The stability test was realized and compared through all seven methods. The stability test with the substructure of decathlon in the 4th cluster is 85.71%.

Understanding the structure of sports performance contributes to streamlining the training process and determining the world's top-class decathletes typology. The structure of athletics groups reflects the structure of the training process, which is based on the versatility of decathletes with a dominance of sprint-jumping and throwing character.

References

Bilić, M.(2015). Determination of taxonomic type structures of top decathlon athletes. Acta Kinesiologica**9**(S1): 20-23.

Bilić, M., Smajlovic, N., Balić, A. (2015). Contribution to discipline decathlon total score results in relation to decathlon age and result-level. Acta Kinesiologica, **9**(1): 66-69.

Bilić. M. &Smajlović, N. (2012). Model for longitudinal analysis of an individual all-rounder athlete's potential. HOMO SPORTICUS Scientific Journal of Sport and Physical Education. **14**(1):44-48.

Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C.J. (1994). Classification and Regression trees. Wadsworth, Belmont CA.

Broďáni, J. (2006). Decathlon with an Average Performance of 7219.91 Points. Journal Of Coimbra Network On Excercise Sciences. **3**(1): 63-66.

Broďáni, J., Czaková, M., Kováčová, N. (2020). Share of atletic disciplines on average performance in decathlon of men with a value of 8282.74 points. Physical education, sport and olympism - Proceedingsreport ofinternational scientific conference. Ružomberok : VERBUM vydavateľstvo KU, 2020. [In press].

Camp, N., Slattery, M. (2002). Classification tree analysis: a statistical tool to investigate risk factor interactions with an example for colon cancer. Cancer Causes & Control. **13**: 813-823, DOI: 10.1023/A:1020611416907.

Cohen, J. S. (1988). Statistical power analysis for the behavioral sciences. 2nd ed. New York: Lawrence Erlbaum Associates.

Cox, T. F., Dunn, R. T. (2002). An analysis of decathlon data. Journal of the Royal Statistical Society Series D: The Statistician, **51**(2):179-187.

Brian P. Dawkins, Peter M. Andreae, Paul M. O'connor. (1994). Analysis of Olympic Heptathlon Data, Journal of the American Statistical Association, **89**:427, 1100-1106, DOI: 10.1080/01621459.1994.10476848.

Dziadek, B., Iskra, J., Przednowek, K. (2018). Principal Component Analysis in the Study of Structure of the Best Polish Decathlon Competitors from the Period between 1985–2015. Central European Journal of Sport Sciences and Medicine, **3**(23): 77–87. DOI: 10.18276/cej.2018.3-08.

Etcheverry, S. G. (1995). Profile of the decathlete. New Studies in Athletics, **10**(3): 51-55.

Fröhlich, M., Gassmann, F., Emrich, E. (2015). Werist der besteMehrkämpfer der Leichtathletik? Eine analytischeBetrachtungim Sieben- und Zehnkampf. In M. Fröhlich,F. Gassmann, & E. Emrich (Eds.), ZurStrukturanalyse des Mehrkampfs in der Leichtathletik. Eine empirischeStudiezumZusammenhang von

Leistung und ErfoLJimSiebenkampf der Frauen und Zehnkampf der Männer. Saarbrücken: Universaar. 65-100.

Gassmann, F. Fröhlich M., Emrich, E. (2016). Structural Analysis of Women's Heptathlon. Sports 2016, **4**(1): 12. https://doi.org/10.3390/sports4010012.

IAAF. (2019). Decathlon Men. Available on: https://www.worldathletics.org/.

Ihring, P., Horáček, L. (1978). Combined Events in athletics. In LACZO, E. Athletics. Textbooks for training trainers II. Triedy. Bratislava: Šport 1978, 477-506.

Jayal, A., Mcrobert, A., Oatley, G., O'donoghue, P. (2018). Sports Analytics: Analysis, Visualisation and Decision Making in Sports Performance. Routledge : New York. 272 p.

Kenny, I. C., Sprevak, D., Sharp, C., Boreham, C. (2005). Determinants of success in the olympic decathlon: Some statistical evidence. Journal of Quantitative Analysis in Sports, **1**(1): 1-6.

Koštiaľ, J. (1993). A Structure of Performance in decathlon. Physical Education and Sport, **3**(4): 19-22.

Koukal, J. (1985). Decathlon. Praha: ÚV ČSTV VMO, 1985.

Koukal, J., Vindušková, J. (1987). Combined Events. In Kuchen, A. et al. Theory and didactics of athletics. Bratislava: SPN. 321-337.

Lehmann, E.L. (1975). Nonparametrics: Statistical Methods Based on Ranks. Holden-Day, San Francisco.

Pavlović, R.(2017). Discipline homogeneity based on the most successful decathlon scoring placement. Turk J Kin. **3**(1): 6-11.

Pavlović, R.,Idrizović, K.(2017). Factor analysis of world record holders in athletic decathlon. Sport Science**10**(1): 109-116.

Park, J., Zatsiorsky, V. M. (2011). Multivariate statistical analysis of decathlon performance results in olympic athletes (1988–2008). World Acad. Sci. Eng. Technol, **5**: 985-988.

Ryba. J. a kol. 2002. Combined Events in athletics. Praha: Olympia.

Schomaker, M., Heumann, C.(2011). Model averaging in factor analysis: An analysis of olympic decathlon data. Journal of Quantitative Analysis in Sports**7**(1): 1-15.

Shen, W., Huang, C.(2012). Grey relational analysis on the performances of decathlon world record holder Roman SebrleProceedings. ICIDT 2012, 8th International Conference on Information Science and Digital Content Technology **1**: pp. 1-58.Jeju; South Korea; 26 June 2012 through 28 June 2012.

Trebuňa, P. Beres, M. (2010). Klasifikáciametódzhlukovaniaaoblasti ichvyužitia. Transfer inovácií 16/2010, 31-34.

Trevor F., Ryan T. Cox Dunn (2002). An analysis of decathlon data Journal of the Royal Statistical Society. **51**(2), 179-187.

Vindušková, J., Koukal, J. (2003). Training Combined Events. In Vindušková, J. Alphabet of athletic trainer. Praha, Olympia 2003,269-283.

Zatsiorsky, V. M., Godik, M. A. (1962). Mathematics and decathlon. Track and Field, **10**: pp. 28-30.

Zhang, P., Lu, J.(2014). Chinese and foreign men's decathlon performance comparison and structural factor correlation test based on SPSS regression model. BioTechnology: An Indian Journal**10**(3): 441-449.

Walker, J.A., Caddigan, S.P.(2015). Performance trade-offs and individual quality in decathletes. Journal of Experimental Biology**218**(22): 3647-3657.

Wang, Z., Lu, G. (2007). The Czech Phenomenon of Men's Decathlon development. International. Journal of Sports Science and Engineering, **1**(3):209-214.

Woolf, A., Ansley, L., Bidgood, P. (2007). Grouping of decathlon disciplines Journal of Quantitative Analysis in Sports, **3**(4):1-13.