

## Original Article

### Improving the Technology for Managing the Training Process of Qualified Athletes

TAMARA KUTEK<sup>1</sup>, RUSTAM AKHMETOV<sup>2</sup>, VLADIMIR POTOP<sup>3</sup>, VIKTOR KOSTIUKEVYCH<sup>4</sup>,  
MYKOLA MYKULA<sup>5</sup>, INNA VOVCHENKO<sup>6</sup>, VICTOR SHAVERSKIY<sup>7</sup>, INNA ASAULYUK<sup>8</sup>, SVITLANA  
DMITRENKO<sup>9</sup>, YURIY NABOKOV<sup>10</sup>, VASILIIY TOLKACH<sup>11</sup>

<sup>1,2,6,7,10,11</sup> Zhytomyr Ivan Franko State University, UKRAIN

<sup>3</sup> Ecological University of Bucharest, ROMANIA

<sup>4,8,9</sup> Vinnytsia State Pedagogical University named after Mykhailo Kotsyubynsky, UKRAIN

<sup>5</sup> Bogomolets National Medical University, UKRAIN

Published online: November 30, 2019

(Accepted for publication: November 15, 2019)

DOI:10.7752/jpes.2019.s6330

#### Abstract

This paper discusses the issue of improving the technology of controlling the training process of qualified athletes. In the system of sports training, prediction of athletic technical and functional enhancement is of pivotal importance. The problem of employing the entire set of sports parameters to solve performance prediction tasks by the well-established mathematical methods has not been solved because of a sufficiently large number of these parameters. *The objective of the study.* The study aims at establishing the set of the most informative sports parameters for athletic performance prediction. *Study results.* The spectral analysis of correlation matrix parameters confirms the finding that the maximum number of the most informative parameters for the performance prediction is from three to six. *Conclusions.* The set of the most informative parameters depends on the age group. In this study, the choice of three to six parameters is made for reasons of high correlation with the target function and the maximum rate of repetition of the top ten eigenvectors in the correlation matrix.

**Keywords:** prediction, informative parameters, correlation ellipsoid, target function.

#### Introduction

Prediction of athletic tactical, technical and functional performance enhancement remains a cornerstone of the sports training system (Platonov, 2004; Akhmetov, 2005, 2006; Kutek, 2014). To use the entire set of sports parameters for solving athletic performance prediction tasks by well-established mathematical methods is not possible due to a considerable number of these parameters ( $P_{max} > 20$ ), while the quantity of age groups ( $S$ ) employed for prediction is typically limited to  $S_{max} = 7$  (10-16 years). This study proposes to choose among a great number of sports parameters the most informative ones with fewer dimensions  $P < S_{max}$ . Moreover, it suggests that linear prediction tasks should satisfy the required condition  $P < S - 1$  ( $S > 2$ ).

Prediction relies on the extrapolation method, which extends the findings obtained by observing one part of a studied phenomenon to other parts (Shustyn, 1995; Bulatova 1996; Bobrovnik, 2002; Gamalij, 2005; Shestakov, 2005). In sports, extrapolation allows providing predictions of performance enhancement based on the study of the relevant patterns in earlier years. The deterministic analysis of the entire set of athletic parameters (anthropometric, technical and special) carried out in (Shustyn, 1995; Akhmetov, 2005, 2006; Ohandzhanov, 2005; Kutek, 2014) reveals their physical essence and proves all of them are relevant and collectively determine sports results (target function). However, the deterministic analysis does not answer a crucial question of how to quantify the degree of influence of certain parameters or a group of parameters on the sports result. This influence is peculiar for various parameters and groups of parameters and depends on the age group. Since particular values of the parameters depend on a specific athlete, they consistently demonstrate random scattering, which can be described by mathematical statistics (Akhmetov, 2005, 2006; Kutek, 2014). In this respect, the factor analysis is significant (Akhmetov, 2006; Kutek, 2014) as it aims at determining the most informative and relevant parameters in a set of random parameters. This necessitates carrying out a statistical analysis of the entire set of sports parameters for each age group and ordering them by the degree of information content. Such a solution to the problem has not been fully considered in the available sports literature.

#### The objective of the study

This study only poses a problem. Its ultimate purpose is to outline the key issues related to the development of a modern, scientifically based methodology for establishing a set of the most informative

parameters for performance prediction of qualified high jumpers. The factor analysis and estimation of the maximum number of informative parameters for athletic performance prediction are of particular interest.

### Material and Methods

The entire set of parameters, including a sports result (H), is presented as N-dimensional vector  $\vec{x}_N$  (matrix column):

$$\vec{x}_N^T = (x_1, x_2, \dots, x_N),$$

where  $T$  is a matrix transposition operation,  $\vec{x}_N^T$  – a line,  $\vec{x}_N$  – a column. In this study, the analysis is limited to the case of  $N=21$ :  $x_1 = H$  is a sports result (jump height; also called the target function (TF)).

Anthropometric parameters ( $x_2 \dots x_7$ ):  $x_2$  is the length of the torso;  $x_3$  – the length of the shin;  $x_4$  – the length of the thigh;  $x_5$  – the circumference of the thigh;  $x_6$  – the circumference of the calf muscle;  $x_7$  – the body weight.

Technical parameters ( $x_8 \dots x_{14}$ ):  $x_8$  – the approach velocity before takeoff;  $x_9$  – the velocity the athlete's center of gravity (COG) at the instant of takeoff;  $x_{10}$  – the flight angle of the athlete's center of gravity (COG);  $x_{11}$  – the time of takeoff;  $x_{12}$  – the flight height of the athlete's center of gravity (COG);  $x_{13}$  – the impulse of takeoff force ( $x_{13} = x_7 x_9$ ).

Special parameters ( $x_{14} \dots x_{21}$ ):  $x_{14}$  – the degree of employing power capabilities in takeoff (%);  $x_{15}$  – the 30 meter sprint from a standing start (time, seconds);  $x_{16}$  – the sprinting speed (flying 10 m);  $x_{17}$  – the two-legged standing jump;  $x_{18}$  – the standing longjump;  $x_{19}$  – the standing triple jump;  $x_{20}$  – the single leg vertical jump;  $x_{21}$  – the jump up from a run of three steps.

The vector of sports parameters (VSP)  $\vec{x}_N$  depends on a specific athlete  $m = 1, 2, \dots, M$  in the group of  $M$  athletes (in this study  $M = 12$ ). The dependence of VSP on an athlete (his number) and on the age is presented as follows:

$$\vec{x}_N = \vec{x}_N^m(t), t = t_1, t_2, \dots, t_L, t_0; m = 1, 2, \dots, M,$$

$$t_n = 10 + (n - 1), \quad n = 1, 2, \dots, 8,$$

where  $n$  is the number of age groups (in this study  $n = 8$ );  $t_0$  – the age of leading athletes. To simplify, the dependence of VSP on time is dropped out for now and the age group is characterized by an N-dimensional set of N-dimensional VSP of athletes and presented as a rectangular matrix  $X_{NM}$ , further referred to as a group parametric matrix (GPM):

$$X_{NM} = (\vec{x}_N^1 \vec{x}_N^2 \dots \vec{x}_N^M) = (x_{nm})_{NM}, \quad x_{nm} = \vec{x}_N^m[n],$$

$$X_{NM} = \begin{pmatrix} x_{11} x_{12} \dots x_{1M} \\ x_{21} x_{22} \dots x_{2M} \\ \dots \dots \dots \\ x_{N1} x_{N2} \dots x_{NM} \end{pmatrix}, \quad (1)$$

where  $N$  is the number of lines;  $M$  – the number of columns in the matrix;  $x_{nm}$  – matrix elements (n-component (coordinate) of the vector  $\vec{x}_N^m$ ).

By singling out the sports result  $x_1 = H$ , the VSP  $\vec{x}_N$  can also be presented as a block:

$$\vec{x}_N = \begin{pmatrix} H \\ \vec{y}_{N-1} \end{pmatrix}, \quad \vec{y}_{N-1} = \begin{pmatrix} x_2 \\ x_3 \\ \cdot \\ \cdot \\ x_N \end{pmatrix}, \quad (2)$$

where  $\vec{y}_{N-1}$  is (N-1)-dimensional vector of physical parameters (VPP) of the athlete.

Within the framework of statistical terminology, we will assume that each of the parameters  $x_n$  (for each age group) is a certain random variable, and the VSP is a random vector. The statistical characteristics of the VSP are determined by arithmetic averaging:

$$\bar{a}_N = \bar{\bar{x}}_N = \frac{1}{M} \sum_{m=1}^M \bar{\bar{x}}_N^m, \quad a_n = \bar{x}_n = \frac{1}{M} \sum_{m=1}^M x_{nm}, \quad (3)$$

$$D(x_n) = \sigma_n^2 = \overline{\Delta x_n^2} = \overline{x_n^2} - \overline{x_n}^2, \quad \Delta x_n = x_n - \overline{x_n}, \quad (4)$$

$$\Phi_{nk} = \overline{x_n x_k} = \frac{1}{M} \sum_{m=1}^M x_{nm} x_{km}, \quad (5)$$

$$\Psi_{nk} = \overline{\Delta x_n \Delta x_k} = \Phi_{nk} - \overline{x_n} \overline{x_k}, \quad (6)$$

$$\Psi_{nk} = \sigma_n \sigma_k \rho_{nk}, \quad \rho_{nk} = \frac{\Psi_{nk}}{\sigma_n \sigma_k}, \quad (7)$$

where  $a_n, \sigma_n^2$  – mean values and dispersions of parameters  $x_n$  ( $\sigma_n = \sqrt{D(x_n)}$  – standard deviation (SD));  $\Delta x_n$  – fluctuations of parameters as to mean values;  $\Phi_{nk}, \Psi_{nk}$  – mutual correlations and covariance of parameters  $x_n, x_k$ ;  $\rho_{nk}$  – mutual correlation coefficients ( $|\rho| \leq 1$ ).

The corresponding correlation and covariance matrices are presented as follows:

$$\Phi_{NN} = \frac{1}{M} X_{NM} X_{NM}^T = \overline{\bar{X}_N^m \bar{X}_N^{mT}} = \frac{1}{M} \sum_{m=1}^M \bar{X}_N^m \bar{X}_N^{mT}, \quad (8)$$

$$\Psi_{NN} = \overline{\Delta \bar{X}_N^m \Delta \bar{X}_N^{mT}}, \quad (9)$$

where the bar above refers to arithmetic averaging by number  $m$  ( $m = 1, 2, \dots, M$ ), namely, statistical averaging of athletes in a group with a uniform discrete probability distribution  $p_m = 1/M$ .

The input GPM  $X_{NM}$  contains information about the interrelationship between different parameters  $x_n$  and the degree of their similitude or parametric proximity in a group. It is sufficient to consider similitude or correlation of vectors  $\bar{X}_N^m$ , evaluating the scalar products of vectors (Akhmetov, 2005, 2006):

$$B_{mk} = \frac{1}{N} (\bar{X}_N^m, \bar{X}_N^k) = \frac{1}{N} \sum_{n=1}^N \bar{X}_N^m[n] \bar{X}_N^k[n]. \quad (10)$$

The matrix of scalar products (MSP) can be represented through the GPM  $X_{NM}$ :

$$B_{MM} = \frac{1}{N} X_{NM}^T X_{NM}. \quad (11)$$

The measure of parametric proximity of athletes in a group is considered as algebraic correlation of vectors  $\bar{X}_N^m$ , also called angle cosine between vectors:

$$R_{mk} = \cos \varphi_{mk} = \frac{(\bar{X}_N^m, \bar{X}_N^k)}{\|\bar{X}_N^m\| \|\bar{X}_N^k\|}, \quad (12)$$

$$\|\vec{X}_N\| = \sqrt{(\vec{X}_N, \vec{X}_N)} = \sqrt{\sum_{n=1}^N x_n^2},$$

where  $\|\vec{X}_N\|$  is the vector norm in N-dimensional Euclidean space [2].

**Multinormal distribution law and correlation ellipsoid of the vector of sports parameters. The task of the factor analysis**

Normal probability density of the VSP is standardly presented (Akhmetov, 2005, 2006) as:

$$W(\vec{X}_N / \bar{\vec{X}}_N, \Psi_{NN}) = \frac{1}{\sqrt{(2\pi)^N \det(\Psi_{NN})}} \exp\left\{-\frac{1}{2}(\Psi_{NN}^{-1} \Delta \vec{X}_N, \Delta \vec{X}_N)\right\},$$

where  $\det(\Psi_{NN})$  is the determinant of the covariance matrix.

The probability density section in the VSP space defines the so-called correlation ellipsoid:

$$W(\vec{X}_N / \cdot) = \text{const} \Rightarrow (\Psi_{NN}^{-1} \Delta \vec{X}_N, \Delta \vec{X}_N) = \text{const}' \quad (13)$$

In particular, in the case of independent parameters  $x_n$ , the equation of the correlation ellipsoid is represented as:

$$\left(\frac{x_1 - \bar{x}_1}{\sigma_1}\right)^2 + \left(\frac{x_2 - \bar{x}_2}{\sigma_2}\right)^2 + \dots + \left(\frac{x_N - \bar{x}_N}{\sigma_N}\right)^2 = \text{const}'$$

When the constant  $\text{const}'$  is properly selected, with significant probability the VSP  $\vec{X}_N$  lies inside its correlation ellipsoid. Principally, the multidimensional correlation ellipsoid is characterized by its size and orientation, resulting from correlating the quadratic form (13) to the canonical type. Figure 1 shows the correlation ellipsoid of the VSP on the plane of two parameters.

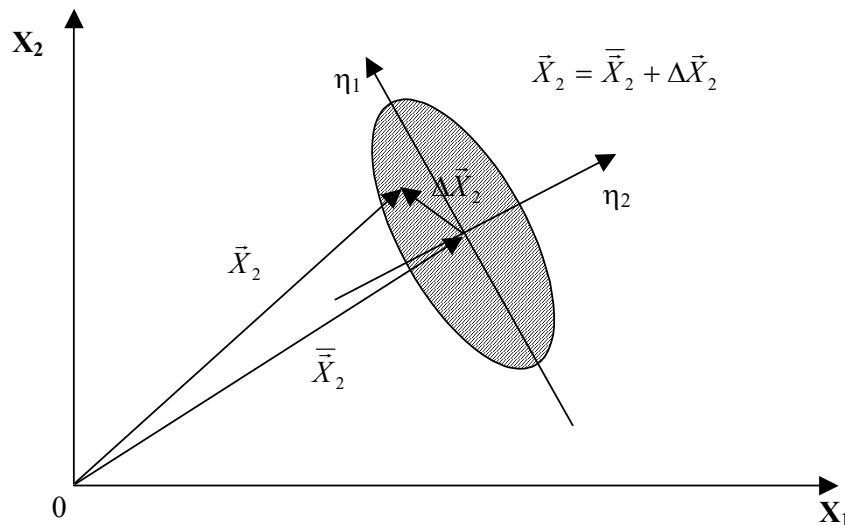


Fig. 1. Correlation ellipsoid (ellipse) on the plane of two ( $N = 2$ ) parameters ( $X_1, X_2$ )

In contrast with most existing researches, in this paper the factor analysis is considered from the standpoint of analyzing the orientation and the size of the multidimensional correlation ellipsoid of the full vector of sports parameters (VSP). Hereby, the so-called localization principle of the VSP in limited subspaces of smaller dimensionality is distinguished, when the dimensions of the correlation ellipsoid in some of the main directions become negligible. It is necessary, however, to emphasize one specific feature of the statistical processing of parameters in a small group of athletes. It lies in the fundamental limitation of the number of athletes in a group ( $M = 12$ ), which can lead to big ratio errors in equidistributed estimates of indeterminate statistical averages (at  $M = 12$ , they are 30–47% (Akhmetov, 2006)). Accordingly, it is of primary importance to clarify the main objective of estimating the group statistical parameters and whether “arithmetic” statistical

characteristics make any sense. In this paper, the main aim is to solve the problem of performance prediction on the basis of a certain set of informative parameters of athletes, depending on the training method. Therefore, at the first stage of the research the influence of ratio errors of arithmetic values for statistical characteristics is not taken into account, and arithmetic averaging is considered as a special case of statistical averaging (with a uniform probability distribution) for solving the issues of localization and factorial analysis of the VSP. The rationale and criterion for the usefulness of such an approach are conditioned by its acceptability for the practice of solving the ultimate task of performance prediction.

### ***Singular numbers of the GMP and maximum number of the most informative athletic parameters***

In most cases, the number of the analyzed physical parameters exceeds the number of athletes in a group:  $N > M$ .

In this case, the ranks of symmetric matrices  $F_{NN}$  and  $B_{MM}$  coincide at the level of  $M$ :

$$\text{Rank} \Phi_{NN} = \text{Rank} B_{MM} = M.$$

This follows from the fact that line and column ranks in arbitrary matrices are the same. Moreover, it can be shown that non-zero eigenvalues of matrices  $(X_{NM} X_{NM}^T)_{NN}$  and  $(X_{NM}^T X_{NM})_{MM}$  coincide and equal the squares of singular numbers in the GMF  $X_{NM}$  (Akhmetov, 2006).

Thus, in the case of  $N = 21$  and  $M = 12$  among twenty physical parameters it is possible to distinguish by methods of mathematical statistics no more than twelve informative parameters for prediction tasks. In our future research, it seems appropriate to bring together a group of athletes from several autonomous groups to ensure the inequality  $M > N$ . Then, all  $N$  parameters can be used to predict athletic performance.

### **Conclusions**

In summary, the current results and the results of the earlier study conducted in 8 age groups, 12 athletes each (including a group of masters of sports of international class), allow us to derive the following conclusions:

The groups are homogeneous in terms of parametric similitude of the athletes in the group. Physical parameters turn to be more determinate with small dispersion, which causes their parametric similitude. The latter places increased demands on the accuracy of the spectral algebraic analysis of the correlation matrices of the parameters ( $\epsilon < 10^{-12}$ ).

The essential task of the factor analysis in identifying the most informative athletic parameters lies in defining the localized area of the vector of physical parameters (VPP) in a certain limited subspace of the full multidimensional Euclidean parameter space. Concurrently, the subspace basis is the set of the first "significant" eigenvectors of the covariance matrix of the VPP, which determine the orientation of the correlation ellipsoid of the VPP. The eigenvalues of the covariance matrix of the VPP determine the size of the correlation ellipsoid where the WFT is localized.

The spectral analysis of the correlation matrixes of parameters confirms the theoretical finding about the maximum number of informative parameters. This number is equal to the number of athletes in the group ( $M = 12$ ). Moreover, there is a sharp decline in the eigenvalues of the matrices, starting with numbers 4–7. This implies that for the tasks of the TF prediction at the first stage it is sufficient to confine to three or six most informative parameters:  $x_{12}$  (the flight height of the athlete's center of gravity (COG));  $x_9$  (the velocity the athlete's center of gravity (COG) at the instant of takeoff);  $x_{21}$  (the jump up from a run of three steps);  $x_8$  (the approach velocity before takeoff);  $x_{15}$  (the 30 meter sprint from a standing start);  $x_{14}$  (the degree of employing power capabilities in takeoff).

The set of the most informative combinations of parameters depends on the age group. Therefore, the question of a unique set of the most informative parameters for all groups remains open and further study of the issue is necessary. In this paper, the choice of three to six dimensional sets is made from considerations of significant correlation with the TF and the maximum repetition rate of the correlation matrix in the top ten eigenvectors.

### **Conflict of interests**

We believe that there is no conflict of interests.

## References

- Akhmetov R.F. (2005) *Theoretical and methodological foundations for the long-term training management of qualified high jumpers.* (in Ukraine)
- Akhmetov R. F. (2006) *Theoretical and methodological foundations for the system of long-term training management of speed and power athletes.* (in Ukraine)
- Bobrovnik V. I. (2002) Rational system of the high jumps training organization at the stages of maximum realization of individual capacities and preservation of achievements. *Theory and methods of physical education and sports.1*, 3-11. (in Ukraine)
- Bulatova M. M. (1996) *Theoretical and methodological foundations for the implementation of functional reserves of athletes in training and sports activities.*(in Ukraine)
- Gamaliy V. V. (2005) Modeling the technique of motor actions in sport. *Science in Olympic sport*, 2, 108-116. (in Ukraine)
- Gamaliy V. V. (2017) *Biomechanical aspects of the motor action technique in sports.* (in Ukraine)
- Kutek T. B. (2014) *Theoretical and methodical foundations for the long-term training management of qualified women athletes specializing in athletic jumps.* (in Ukraine)
- Ohandzhanov A. L. (2005) *Training management of qualified athletes-jumpers.* (in Russia)
- Platonov V. (2004) Biomechanical ergogenic means in modern sports. *Science in Olympic sport*, 2, 96-100. (in Ukraine)
- Platonov V. N. (2004) *The system of training athletes in Olympic sport. General theory and its practical applications.* (in Ukraine)
- Stryzhak A. P. (1992) *Scientific and methodological basis for the training process management of highly qualified athletes-jumpers.* (in Russia)
- Khmelnitska I. V. (2010) The software for biomechanical video and computer analysis of sports actions. *Olympic sports and sports for everyone*, 568. (in Ukraine)
- Shestakov M. P. (2000) Biomechanical aspects of training high-class jumpers and sprinters. *IAAF Bulletin*, 2, 156-170. (in Russia)
- Shestakov M. P. (2005) Technical training management in athletics on the basis of computer modeling. *Science in Olympic sport*, 2, 187-196. (in Ukraine)
- Shustyn B. N. (1995) Modeling and prediction in the system of sports training. *Modern system of sports training*, 226-237. (in Russia)