HOW TO CONSTRUCT MATHEMATICAL MODEL OF SPORT TRAINING

by

IGOR RYGUŁA*

The author demonstrates that the base of training optimization is precise determination of all factors materially influencing sport result in given contest or discipline. The measuring of variables influencing score is not enough, because it is necessary to accurately describe the plasticity of separated features that are highly susceptible to training stimuli such as: strength and muscle mass as well as features with strong genetic determination, almost totally unsusceptible to endogenous and exogenous elements of training. Another aspect of training process optimization is finding such load values (intensity, volume, length of rest periods) that give most favorable course of adaptive changes in the body of given athlete, taking into account his or her genetic potential. The most important step of this process is finding relations and cross-relations between observed factors and applied loads and mathematical description of the dynamics of continuous process on the basis of data representing observations results [Ryguła and Wyderka 1993].

Key words: mathematical model, optimal control methods, optimal training loads.

Introduction

The problem of training process optimization and maximization of the score is very complex, because we are dealing with an attempt of controlling multiple factors, more or less plastic, subject to material variation in time. The key problem here is the number of variables influencing given parameter (value of strength, speed, oxygen ceiling, etc.). Another important problem is time-limited effect of given stimulus action, connected with the process of exercise adaptation [Maas and Mester 1996]. To solve this problem most often a deterministic model is constructed, including a limited number of independent variables, describing the behavior of one or more dependent variables (score),

* Assist. Prof., Dep. of Theory of Sports, Acad. of Phys. Educ., Katowice; Correspondence address: 72a Mikolowska Str. 40-065 Katowice, Poland

or their behavior is predicted [Mader 1998]. Infrequently are found criteria containing the action of training stimulus and changes of adaptation processes as function of time [Perl 1997]. One possible solution of the above problem is construction of non-stereotype meta-models together with analysis of individual cases and investigation of time series. Meta-models are abstract mathematical models, generally created with the help of genetic algorithms. They contain basic logical concepts enabling explanation of the problem of structural similarities of the investigated mechanisms (biological or training systems) [Mester and Perl 1999].

In solving the structure of training loads another approach is also possible, namely mathematical model may be used in form of a set of differential equations [Ryguła and Wyderka 1990]. Basic problem of today competition sport seem to be finding and verifying tools that would enable non-verbal control of training. For this reason this work aims at the construction of such mathematical model, that would enable calculation of optimal controls of the training. The implementation of this aim is connected with obtaining answers to the following investigation problems:

1. How runs the course of construction of mathematical model of sport training?
2. How to determine parameters of training loads?

To answer the above questions it is necessary first define successive steps of the construction of mathematical model of sport training, and then, on the basis of this model to calculate optimal controls.

Material and methods of the investigation

Model foredesign

The construction of the mathematical model should begin with formulation of aim. The first construction step is definition of quality index to be minimized (maximized) in the run of training process. The index generally contains time, after which the aim should be obtained - we call it T. The parameters (characteristics) of the person subjected to training, serving the definition of quality index and others logically connected, even when not present in the index in implicit way, will be called state variables. Their right choice will materially influence the final result of the project. We will denote them with X and assume their number is *n*. Next, it should be determined which actions (training elements) may significantly influence the state variables defined earlier. These actions should be defined quantitatively (percent intensity of exercises, total time of certain actions), as numerical functions, time dependent. It is important for each function to normalize each function values into the same numerical interval (preferably [0, 1]). It is easy to accomplish by dividing by maximum possible value. This function will be called control and will be denoted with U. Let's assume that there are *m* such functions.

Experiment planning

For determination of the optimal control a mathematical model is needed in the shape of the set of differential equations:

$$X_i' = f_i(t, X, \ldots, X_n, U_1, \ldots, U_m) \quad \text{i =1,...,n} \tag{1}$$

with conditions $X_i (0) = x_{0;}$ i = 1,...,n

These equations should be interpreted in the following way:

- the rate of change of state variables in any moment is dependent of time, all state variables and each control; at the beginning of the training state variables have values $x_{10},\ldots,x_{n0}$. In view of the methods of determination of optimal control it is preferable to deal with as simple form of equations - as close as possible to linear form.

To determine the parameters of functions in mathematical model, the experiment (training) should be planned and then implemented in proper way in the period [0, T], using certain exercising group, in which different persons will undergo training with different methods. Elements to be further discussed are: group, control (training), time of the experiment and results.

Control group

For the experiment a group of about 20 - 50 persons should be chosen with possible variegated state variables. In the following the number of this group will be denoted with N. After measuring all parameters in this group is advisable to make randomness test for each state variable. Single persons with results significantly (in sense of statistical value) differ from the rest of the group should be excluded from experiment. This means they should continue training but their data should be ignored in calculations.

Control (training)

Next, with the help of experts (experienced coaches), the logical time run of controls should be determined, that is a practice different for different people should be well planned. It would be preferable to have different (but only slightly) run of control for each person in the group. It is not always possible; the next best thing is to divide the group into subgroups of over ten persons each, subjected to the same course of practice. This way has advantage in form of averaging the influence of each kind of practice. The number of subgroups should be at least three. In the ideal solution with each of over ten courses of practice would be tested on a group of over ten persons. However this would necessitate using a group of 300 - 400 persons. The division into subgroups should be done in the random way to avoid systematic error. Arbitrary division may lead to grouping persons with like parameters in the same subgroup. Subjecting them to the same kind of exercise will decrease the value of information obtained from experiment. It is not absolutely necessary to strictly observe planned controls. The illnesses, contusions, equipment malfunctions may happen. Necessary is recording actual control of each person exercising.

Experiment duration

Total duration of the experiment is closely connected with optimized quality index. As continuous measurement of state variables is not possible, time interval for measurements must be determined. It may be a week, month, quarter, etc. Because of aiming at greatest possible accuracy of model, it is preferable to choose possibly short time between measurements. Excess should be avoided. The important thing is to be certain that the change of state variable is caused by exercise and not by any accidental factor, such as loss of weight caused by illness. In this work we will assume that the number of measuring intervals (time between two consecutive measurements) is M+1. A factor simplifying calculation is providing constant interval between measurements. This is not necessary, but when the intervals vary, the increments must be calculated in different way.

Experiment result

Recording for each person after each time interval the values of his or her state variables and actually effected control (practice), we obtain a set of numbers

$$x_i^{(1)}(k_j)\ \ u_i^{(1)}(k_j) \tag{2}$$

i = 1,...,n; 1=1,...,N; j = 1,..., m; k = 1,..., M+1

which is analyzed and then used for determination of the model. It is recommended to graph all state variables as well as time averages for whole group and for each subgroup.

Processing of the results

The derivative of each state variable in the model may be approximated by the increment of observed value corresponding to this variable of the characteristic in single measuring intervals. Therefore we first calculate

$$\Delta x_i^{(1)}(k) = x_i^{(1)}(k+1) - x_i^{(k)}(k) \tag{3}$$

i = 1,...,n; 1=1,...,N; k = 1,..., M

It is advisable to analyze the course of mean increment for whole group

$$\Delta x_i^{(1)}(k) = \frac{1}{N}\sum_{1=1}^{N} \Delta x_i^{(1)}(k) \tag{4}$$

i = 1,...,n; k = 1,..., M

and mean increment in subgroups

$$\Delta^r x_i(k) = \frac{1}{N_r}\sum_{1=1}^{N_r} \Delta x_i^{(1)}(k) \tag{5}$$

i = 1,...,n; k = 1,..., M

where r is a number of subgroup, and N_r its size.

On the basis of increments and measuring data, a model of linear multiple regression may be formulated for each state variable. In this way the linear form of mathematical model is determined [Findeisen et al. 1979, Leitmann 1972]. Such model has the form:

$$\Delta x_i = \sum_{p=1}^{n} a_{pi} x_p + \sum_{q=1}^{n} b_{qi} u_q + c_i \qquad (6)$$

i=1,..., 7

This model may be interpreted as a following problem: How the mean increment of state variable in the group during single measuring interval may be linearly explained with state at the beginning of the period and implemented control (exercise). Using numeric values (4) and (5) regression coefficients may be approximated in formula (6). When determining regression coefficients values a, b, c, their significance should be tested, i.e. verify the hypothesis that they are nonzero. The coefficients with small probability of zero value, are the most significant factors influencing increase (or generally change) of state variable. In this way it may be determined, which control and how influences the change of value of each characteristic. Next, the observed increments should be compared with calculated with the use of regression, that is the following values should be analyzed

$$\sum\sum \Delta\Delta_i^{(1)}(k) = \Delta_i^{(1)}(k) - \sum_{p=1}^{n} a_{pi} x_i^{(1)}(k) - \sum_{q=1}^{m} b_{qi} u_q(k) - c \qquad (7)$$

i = 1,...,n; 1=1,...,N; k = 1,..., M

Time functions of these differences for all individuals may differ significantly, therefore it is advisable to calculate means in the group.

$$\Delta\Delta_i(k) = \frac{1}{N} \sum_{1=1}^{N} \Delta\Delta x_i^{(1)}(k) \qquad (8)$$

i = 1,...,n

These values will help to determine *h* functions dependent solely on time, influencing increments of state variables. It should be expected that they will be functions oscillating around zero. Therefore they should have the form:

$$h_i(t) = \alpha_i \cos\left(\frac{2\pi t}{\lambda_i}\right) + \beta_i \sin\left(\frac{2\pi t}{\lambda_i}\right) \qquad (9)$$

i = 1,...,n

Coefficients $\alpha_1, \beta_1, \lambda_1$ should be determined with mean square approximation methods, that is they should be values minimizing function

$$H(\alpha_i, \beta_i, \lambda_i) = (\sum_{k=1}^{M} h_i(k) - \Delta\Delta_i(k)) \tag{10}$$

i = 1,...,n

Formulation of a model

Regression coefficients obtained from the model (6) and h_i functions determined with formula (9) lead to mathematical model in the form:

$$x_i^{'}(t) = \sum_{p=1}^{n} a_{pi} x_p(t) + \sum_{q=1}^{m} b_{qi} u_q(t) + c_i + h_i(t) \tag{11}$$

In vector-matrix notation the set of differential equation will have the form:

$$X_i^{'}(t) = AX(t) + BU(t) + C + h(t) \tag{12}$$

and in accordance with procedure described in [Ryguła and Wyderka 1990, Szczotka 1980] may be used to determine optimum control for different quality indexes.

The problem of construction of mathematical model will be analyzed on the example of swimming.

Construction of a model of 13-14-year old swimmers

In this chapter, on the example of a year of exercises of young swimmers, will be described a method of construction of the mathematical model described above. We have decided to use minimal score for 25 m free style as a quality index of the optimization model. Therefore the following state variables were assumed:

X_1 - score in 25 m free style [points]

X_2 - score in 800 m free style [points]

X_3 - score in static long jump [cm]

X_4 - evaluation of free style swimming technique [number of RRcycles/25 m]

X_5 - height [cm]

X_6 - body mass [kg]

X_7 - vital capacity [cm^3]

therefore $n = 7$

Elements of exercise considered as controls, were chosen in the following way:

U_1 - warm-ups and free exercises [min]

U_2 - runs, plays, amusement [min]

U_3 - swimming on distances up to 25 m [min]

U_4 - swimming on distances 25 - 50 m [min]

U_5 - swimming on distances 100 - 200 m [min]

U_6 - swimming on distances 400 - 800 m [min]

U_7 - warm-up swimming, compensation [min]

U_8 - exercising technique of styles [min]

U_9 - exercising jumps and turn-backs [min]

U_{10} - competitions [min]

Number of controls therefore equals 10. Because T - the total length of exercise period - is 8 months, as measuring interval one month has been chosen. Each control is expressed with the number of minutes in a month, during which given element was exercised. Next, to normalize all controls, they were divided by maximal time that could be used for given exercise in a month.

Data on experiment

For the experiment 40 young swimmers were chosen and they were divided into 4 subgroups. As said in the preceding paragraph, M = 7. The obtained results of measurements and assumed controls were collected with the EXCEL spreadsheet. Results were evaluated with the STATISTICA program.

Results of the investigation

In accordance with described methodology, the increments of state variables were calculated. On the basis of them, a mathematical model has been constructed in the form of the set of 7 differential equations with 10 control parameters. As an example, one of them is shown below:

$$\frac{dx_1}{dt} =$$

$$0.030*X_1-0.325*X_2+0.349*X_3-2.98*X_4+0.240*X_5+0.649*X_6-$$

$$0.011*X_7+0.326*X_1*U_1+0.280*X_1*U_2+1.46*X_1*U_3+$$

$0.311*X_1*U_4+0.249*X_1*U_5+0.420*X_1*U_6-0.150*X_1*U_7-$
$1.53*X_1*U_8+0.009*X_1*U_9+0.237*X_1*U_{10}+1.25*X_2*U_1-0.170*X_2*U_2-$
$1.86*X_2*U_3-0.680*X_2*U_4-0.197*X_2*U_5-0.611*X_2*U_6-$
$0.125*X_2*U_7+1.39*X_2*U_8+0.534*X_2*U_9-0.926*X_2*U_{10}-$
$2.35*X_3*U_1+0.551*X_3*U_2-0.198*X_3*U_3+$
$0.002*X_3*U_4-0.186*X_3*U_5-0.002*X_3*U_6+1.05*X_3*U_7-0.337*X_3*U_8-$
$0.153*X_3*U_9+2.69*X_3*U_{10}+ 4.78*X_4*U_1+$
$6.75*X_4*U_2+0.980*X_4*U_3-8.89*X_4*U_4+0.823*X_4*U_5-1.48*X_4*U_6-$
$5.09*X_4*U_7-11.9*X_4*U_8+9.19*X_4*U_9+4.54*X_4*U_{10}+$
$11.4*X_5*U_1-0.955*X_5*U_2-0.429*X_5*U_3-4.55*X_5*U_4-0.254*X_5*U_5-$
$3.75*X_5*U_6-8.55*X_5*U_7-6.86*X_5*U_8+1.16*X_5*U_9+$
$3.30*X_5*U_{10}-6.57*X_6*U_1+1.97*X_6*U_2-3.99*X_6*U_3+$
$0.003*X_6*U_4-0.028*X_6*U_5+ 0.289*X_6*U_6+8.33*X_6*U_7+$
$3.60*X_6*U_8+0.133*X_6*U_9-4.54*X_6*U_{10}-0.010*X_7*U_1+$
$0.079*X_7*U_2+0.098*X_7*U_3-$
$0.034*X_7*U_4+0.019*X_7*U_5 0.054*X_7*U_6+0.011*X_7*U_7-0.172*X_7*U_8-$
$0.014*X_7*U_9+$
$0.059*X_7*U_{10}-1132*U_1-326*U_2-63.1*U_3+894*U_4-$
$9.76*U_5+391*U_6+843*U_7+1571*U_8-304*U_9+54.4*U_{10}-0.005$

In possession of a mathematical model we may start to determine optimal controls of exercise loads for assumed initial conditions of given athlete. Optimal control for maximization of the 25-m distance score has been determined. Because in numerical solution of model equations a step has been assumed equal to half of observation period, optimal control was determined with the interval of 2 weeks. The values of separate control variables therefore denote total time of effecting given exercise during 2 weeks, which was presented in Table 1.

Table 1. Control leading to maximum point score on 25 m distance for chosen athlete.

Tydzień	Exercise type										25 m score
	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	
1-2	25	135	13	25	0	33	0	0	0	1	133.9
3-4	10	98	5	10	0	15	0	0	0	1	142.2
5-6	25	45	11	0	0	25	3	25	0	2	149.9
7-8	25	45	11	0	0	25	3	25	0	2	167.8
9-10	25	45	11	0	0	25	3	25	0	2	186.0
11-12	25	45	11	0	0	25	3	25	0	2	191.8
13-14	60	135	0	0	140	0	30	30	30	4	212.9
15-16	60	135	0	0	140	0	30	30	30	4	235.8
17-18	60	135	0	0	140	0	30	30	30	4	258.5
19-20	60	135	0	0	140	0	30	30	30	4	281.3
21-22	60	135	0	0	140	0	30	30	30	4	304.6
23-24	60	135	0	0	140	0	30	30	30	4	328.5
25-26	60	135	0	0	140	0	30	30	30	4	352.9
27-28	60	135	0	0	140	0	30	30	30	4	368.4

Discussion of the results

The basis for defining detailed model was pedagogic experiment made on the group of 13-14-year boys. Before starting the experiment it was defined, which variables will be observed. Such elements of exercise were chosen as can be expressed with numbers and changed during its duration. During the experiment methods, forms and means of exercise were changed for separate athletes, so the observation material can contain as much information as possible. Due to this, the model is more close to reality. Collected date were given a thorough mathematical treatment.

The form of model suggested in the paper is nonlinear, most simple of all. At the same time, a simple form of conjugated set. The paper shows simple, but very effective method of model construction on the basis of possessed data.

The uses of constructed model may be manifold. The paper shows only evaluation of the influence of different variables on the development i formation of speed abilities in swimming (competition score on 25 m distance).

Analysis of optimal course of control functions has indicated that they are time variable. It should be pointed out, that each training period is characterized with variable controls. Variability range is very wide and proportions between used controls greatly differ between them depending on the training period [Counsilman 1977].

The results of this investigation and the experience of other workers indicate that optimization implies individualization of children and teenagers [Filin 1986, Komor 1984, Morawski 1980]. For each athlete there is defined, individual volume and intensity of exercise, absolutely necessary for keeping the exercise level of physiological characteristics on the current level. This applies to each stage of exercise, where the workload should be adequate to given development period and assumed sporting target of given athlete [Svoboda 1987, Schlich and Jannsen1990].

Conclusions

1. The basic mathematical tool that may be used for planning exercise workloads, is theory of optimal control, on the basis of which the mathematical model has been constructed and determined.

2. The mathematical model should constitute a set of differential equations. Each equation should describe dynamics of development of different characteristic, dependent on all analyzed characteristics, used controls and products of state variables and controls.

REFERENCES

Counsilman J.1977. *Swimming Power*. Swimming World, 10.

Filin V.P.1986. *Vospitanie fiziceskich kacestv u junych sportsmenov*. Moskwa, F i S.

Findeisen W., Szymanowski J., Wierzbicki A. 1979.*Theory and optimal control calculation methods*. Warszawa, PWN.

Komor J.A. et al. 1984. *Modelling and optimization of individual movement techniques.* Raport Inst. Sportu, Warszawa (In Polish).

Leitmann G.1972. *Introduction to optimal control theory*. Warszawa, PWN (In Polish).

Maas S., Mester J. 1996. *Diagnosis of individual physiological response in elite sport by means of time-series-analyses*. (In:) Marconnet, P. et al. (Eds.): *Frontiers in Sport Science.* The European Perspective. University of Nice. Nice, 98-99.

Mader, A. 1988. *A Transcription-Translation activation feedback circuit as a function of protein degradation, with the quality of protein mass adaptation related to the average functional load*. Jour. Theo. Biol. 134, 2 – 17.

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, 163.

Morawski J.M. 1980. *Methodological and physical bases of experimental model in sports*. nr 137, AWF Poznań (In Polish).

Nowikow A.A. et al. 1976. *O razrabotkie modielnych charaktieristik sportsmienow.* Tieorija i Praktika Fiziczeskoj Kultury nr 6.

Perl J.1977. *Aspekte unkonventioneller Modellbildung – exemplarische Anwendungsmöglichkeiten in der Sportwissenschaft. Zur Veröffentlichung angenommen bei*: Sportwissenschaft 7.

Ryguła I., Wyderka Z.: *Construction of optimization model of mechanical power development of 13-15-year old boys*. Wychowanie Fizyczne i Sport nr 2 (In Polish).

Ryguła I., Wyderka Z.1993. *Elements of control and optimization in sport training*. AWF Katowice (In Polish).

Schlich W., Janssen, J.P.1990. *Der Einzelfall in der empirischen Forschung der Sportwissenschaft: Begründung und Demonstration zeitreihen-analytischer Methoden*. Sportwissenschaft 3, 263–280.

Svoboda G.1987. *Indywidualizácia sportovej pripravy mládeže v plávani*. Tréner nr 5.

Szczotka F.A.1980. *Matemathical model of sport training. Materials of I scientific session on"Mothematical modelling in Physical Culture Sciences".* AWF Poznań (In Polish).