

EXTENDING STATISTICAL DATA OF EEG BIOFEEDBACK QUALITY IMPROVEMENT WITH SOFT COMPUTING

Łukasz Czuba

Department of Naval Command and Operations, Polish Naval Academy of the Heroes of Westerplatte, Inżyniera Jana Śmidowicza 69, 81-127 Gdynia, Poland

E-mail address: lukasczuba@interia.pl

ORCID: <https://orcid.org/0009-0009-8601-975X>

Monika Łukowska

Department of Methodology, Statistics and Informatics, Faculty of Theory and Practice of Sport, The Jerzy Kukuczka Academy of Physical Education, ul. Mikołowska 72a 40-065 Katowice, Poland

E-mail address: the.monkas@gmail.com

ORCID: <https://orcid.org/0009-0000-2759-1654>

ABSTRACT

Aim. Automatic processing of the data in order to determine the status of work and identification of the activity and brain-wave frequencies becomes necessary for the modern systems in the in the diagnosis of biofeedback among athletes.

Concept. The study aimed to explore the effects of physical exertion on alterations in the manifestation of brain wave frequencies (pre/post exercises) in a group of 15 endurance athletes.

Results and conclusion. Statistic methods allowed an identification of data anomalies, such as extreme, outliers and missing values. Combining information with soft computing tool can distinguish the level of electrical activity of the analysed muscles. Used Big Data and Data Mining tools solution with a statistical approach while maintaining high measurement accuracy indicates the effectiveness of this method in medical diagnosis.

Keywords: soft computing, eeg, computer science, statistic methods

INTRODUCTION

Monitoring the electroencephalography frequency of brain waves is very important, both at the individual and social level of mechanisms governing the course of life activities of organisms. From the historical level, these traditional assumptions of clinical diagnostics are complemented with modern information

technologies and methodologies, which resulted in a significant improvement in quality. Collection and analysis of good quality data is effective in detection of anomalies in the neuromuscular system. Such advances in medicine cause faster diagnostics, and optimum adaptation of methods and therapeutic agents. A significant amount of the scientific achievements in the field of frequency of brain waves is based only on the results, which lack precise information about the patient and his other medical issues. In addition, reports on the effectiveness of certain medical interventions in population studies are often conducted without appropriate identifying the patients in the study. That's, why there are extreme values, that interfere with the test results. The challenge of generating and analysing large data sets is an adaptation to new methods of data collection and numerical conceptualisation.

The aim of the study was an attempt to create and determine the impact of physical effort on changes in the manifestation of brain wave frequencies among 15 athletes of endurance disciplines. It was decided to check whether the frequency of brain waves would change significantly in the measurement after exercise in relation to the measurement before exercise. At the same time, the research objects carried out exercise work with the same basic goal. All stages of the research will be grouped in detail, categorised and analysed using the Big Data method and verified using statistical tools. At the end of this stage, as well as during the previous one, regression models will be built and verified.

EEG BIOFEEDBACK EXAMINATION PROCESS

Electroencephalography is a tool to measure and present information about brain waves as a result of changes in physiological processes. In turn, the biofeedback method aims to teach the subject to consciously modify functions (brain waves) that are not consciously controlled on the basis of feedback signals about changes in the physiological state. This, in turn, allows the determination of these frequencies and the optimal control of their values in order to positively influence the development process, and in the case of these studies, objects on the training process. The whole process of electroencephalography is learning to control brain activity. On the basis of the obtained feedback (biofeedback), it is possible to describe the impact of changes in external conditions on the activities of realising the sports potential in preparation motivating the achievement of the highest sports results. There are different ranges of electromagnetic waves generated by the human brain. The production of brainwave components is always present. However, systematic training and a strong will of the subject are required to gain the advantage of certain desirable brainwaves. The essence of EEG biofeedback training is therefore to generate certain desired waves while reducing other undesirable ones.

Electroencephalography (EEG) measurement was performed using B-Alert X10 Retail Pricing Sheet and X-Series Basic Software.

MATERIALS AND DATA COLLECTION

There is a great need for a scientific approach in the field of information technology used in automatic data processing. Referring to the four paradigms of science, it can be argued that all of them allow the processing of large EEG data streams. The first three paradigms include an empirical description of phenomena, theoretical generalisations and computational simulations of complex phenomena. The fourth element includes data exploration which was discovered in the last years, suits Big Data analysis, also called Data-Intensive research.

The previous collection of large EEG data streams is preceded by the extension of the statistical results achieved during the test. An approach based on known techniques of computer science and statistics allowed increased sensitivity of monitoring by improving detection performance in order to obtain data of better quality. Statistics will allow the identification of data anomalies in recorded muscles electrical activity, such as extreme values, outliers and missing values. Outliers are data points overlapping the distribution of the remaining data. They reflect occurring anomalies of the electrical potential of muscles or neural circuits that interfere with process of modelling. Even one coming off observation may distort the significant coefficients of the decomposition of batch data, that's, why such observations should be taken into account in statistical modelling. Typically, these findings represent a random error and artificially increase or decrease the value of statistical coefficients. Extreme values are points far away from the range of distribution of batch data and are found beyond outliers. Extreme values are positioned above or below the limit defined by three times the length of non-standoff values (Min, 1Q, Median, 3Q, Max), whereas outliers are values that are above or below the limit of one and a half and not more than three times the length of non-standoff values. Effective explorative technique of verified distribution data is scatterplot 3W, which shows obscured patterns of data collections in real angle.

Somewhat different, but equally important stage of statistical analyses are calculations of expected marginal means, which are the best linear estimators with minimal oppress for marginal means of the system (Milliken & Johnson, 1992).

BIG DATA ANALYSIS

To prepare a multi-structural analysis of musculoskeletal and nervous systems with large amount of data we have to create a multithreaded architecture for parallel processing (Changqing, Yu, Wenming, Awada, & Keqiu, 2012). The data included 18 variables from EEG biofeedback measurements (Alpha, Theta, Delta, SMR, Beta: Beta 1, HighBeta, Beta 2, Beta 3, Beta 4, Beta 5). As a result, 46,208 observations were obtained from one record for each parameter (Shvachko, Kuang, Radia, & Chansler, 2010).

The Big Data Architecture creation is the combination of the measurement devices, special processing equipment and software to work together in an integrated way with each other during the processing of the data. An important issue in this type of concept is to divide the information in terms of their type and volume. Ensuring the appropriate data partitioning can be achieved by indexing

the records in the form of the fractal indexing (fractal tree indexing). The next step, during modelling is to optimise the computational units and aggregates information.

Using the appropriate computer interface, electromyographic data are processed in the form of model solutions to a given problem. Most frequently, such models are designed to process the nerve-muscle activities by mapping and reduction. A common system for storing files is Hadoop distributed file system, and the combination of data from different disks provides a programming model called MapReduce. Processing of the Big Data takes place in two steps: Map phase and Reduce phase. Programming is a calculation of the function of Map and Reduce, where the information is encoded in the form of pairs of keys and values as their inputs and outputs (Liu, Wang, Matwin, & Japkowicz, 2015). The first phase of map function includes a division of key-value pairs into subsets and their distribution to the various nodes as the cluster. The second phase of reduce function acts as the aggregation key-value pairs. The purpose of reduce function is to prepare the final value, assuming that the pairs with the same keys will go to the same nodes.

REGRESSION ANALYSIS

The procedure for assessing the relationship between EEG biofeedback data before and after exercise among 15 high-performance athletes consisted of regression modeling, followed by verification with a statistical tool. A simple regression model was used in which the form corresponds to the n -element sample (Changqing, Yu, Wenming, Awada, & Keqiu, 2012):

$$y_i = x_i\beta_1 + \beta_0 + \varepsilon_i, \quad 1 \leq i \leq n,$$

where: y_i – the value of y for case i ,
 x_i – the value of x for observation i ,
 ε_i – random disturbance about the distribution $\varepsilon_i \sim N(0, \sigma^2)$, independently, which means, $Cov_{i \neq j}(\varepsilon_i, \varepsilon_j) = 0$,
 β_1, β_0 are the coefficients of the model.

The vector record of the model is described by the formula:

$$y = x\beta_1 + \beta_0 + \varepsilon,$$

where: y and x are column vectors and the random disturbance $\varepsilon \sim N(0, \sigma^2 I_{n \times m})$. The feature y has a normal distribution with a variance σ^2 and average $x_i\beta_1 + \beta_0$. Parameter evaluations β_1 and β_0 of the model are determined after simplifying the model with formulas (Shvachko, Kuang, Radia, & Chansler, 2010):

$$\hat{\beta}_1 = \frac{cov(x, y)}{var(x)},$$

$$\hat{\beta}_0 = \bar{y} - \bar{x}\hat{\beta}_1,$$

where: $cov(x, y)$ – sample covariance for x and y vectors,
 $var(x)$ – vector sample variance x ,
 \bar{x} – the average of the vector x .

During the regression analysis, the occurrence of a significant relationship between the variables x and y was examined. For this purpose, the description of the simple regression was translated into the language of a linear model in the form:

$$y = X\beta + \varepsilon,$$

where: $\beta = (\beta_0, \beta_1)$ - columnar vector,

matrix X - a matrix with the first column filled with ones, and the second column with the values of the variable X .

Answering the question of the existence of dependencies between variables x and y in the language of the linear model, a null hypothesis is obtained regarding the β_1 coefficient verifying the lack of dependence $H_0 : \beta_1 = 0$ and an alternative hypothesis concerning the non-zero dependence $H_A : \beta_1 \neq 0$. If the test result gives a p-value lower than the assumed significance level α , the null hypothesis is rejected, which means that the relationship between variable x and y is significant. In this paper, a significance level of $\alpha=0.05$ was chosen, but this value may vary depending on the issue being tested or other factors.

Based on the above formulas, regression models were built examining the relationships between two variables for each brain wave, respectively, from EEG biofeedback and post-exercise EMG measurements, which were then compared with each other using a statistical test. The relationships between Alpha, Theta, Delta, SMR and Beta waves were examined: Beta 1, HighBeta, Beta 2, Beta 3, Beta 4, Beta 5. Subsequently, 3-minute brainwave and EEG recordings were compared before and after exercise, and descriptive statistics. Finally, the above stages made it possible to estimate the frequency dependence of EEG waves.

RESULTS AND DISCUSSION

It was decided to create 34 simple regression models in order to show the dependence. The results of the statistical regression model with their residues between electrical waves (Alpha, Theta, Delta, SMR and Beta: Beta 1, HighBeta, Beta 2, Beta 3, Beta 4, Beta 5) are presented in table 1 and table 2. Knowing that the variables in the model are not treated symmetrically, the criterion for estimating the coefficients in the model, i.e., the mean square error and residuals (Min, 1Q, Median, 3Q, Max) will be different along with the results of the evaluation of the model coefficients. Therefore, the explained and explanatory variables were treated interchangeably in the models in order to check which relationship is more strongly related to the data.

Table 1
Frequency dependencies of Alpha, Theta, Delta, SMR brainwaves in EEG biofeedback measurements

Residuals	Min	1Q	Median	3Q	Max	F-statistic	Multiple R-squared	Estimate	Intercept Estimate	
Delta~Alpha	-29.41	6.61	1.79	6.48	42.70	2.61	5.654	-0.0106	-0.0009	67.49
Alpha~Delta	-15.01	-4.79	0.09	4.46	13.95			-5.321	6.658e-05	33.80
Delta~Theta	-29.59	-6.76	1.81	6.53	42.57	17.8 ***	0.0004	0.0234	-0.0009	67.47
Theta~Delta	-24.24	-4.69	-0.09	5.30	20.29			0.0164	0.0009	47.30
Delta~SMR	-29.50	-6.60	1.76	6.50	42.80	0.0007	1.62e-1	-0.0002	-0.0009	67.49
SMR~Delta	-15.25	-3.83	-0.14	3.78	15.03			-8.14e-05	-2.283e-04	27.62
Theta~Alpha	-22.26	-4.89	-0.07	4.67	20.50	1915 ***	0.0398	0.2360	0.0008	45.43
Alpha~Theta	-12.08	-4.78	0.17	4.66	14.04			1.69e-01	-7.117e-05	32.45
Theta~SMR	-24.39	-4.66	-0.09	5.36	20.90	0.178	3.85e-06	0.003	0.0008	47.31
SMR~Theta	-15.26	-3.84	-0.15	3.78	15.01			0.001	-0.0002	27.62
Alpha~SMR	-14.48	-4.68	0.07	4.44	14.21	5067 ***	0.09883	0.348	0.0002	30.46
SMR~Alpha	-15.36	-3.38	0.05	3.34	16.11			0.284	-0.0002	24.89

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

In Table 1 there is a significant linear relationship between the models for Alpha and SMR waves. For the first model, where Alpha was the explained variable and the SMR wave was the explanatory variable, the assessment of $\beta_0 = 0.0002$, $\beta_{SMR} = 0.35$, $\sigma^2 = 30.46$ was completely different and higher than the evaluation of the coefficients of the second model $\beta_0 = -0.0002$, $\beta_{Alpha} = 0.28$, $\sigma^2 = 24.89$. A significant linear relationship, but much weaker than the one described above, occurred between the Theta and Alpha and Delta and Theta waves. The first Theta-Alpha relationship model showed higher coefficient estimates $\beta_0 = 0.0008$, $\beta_{Alpha} = 0.24$, $\sigma^2 = 45.43$ than the second model $\beta_0 = -7.117e-05$, $\beta_{Theta} = 1.686e-01$, $\sigma^2 = 32.45$. The Delta~Theta brainwave dependence in the first model also had higher ratings of the coefficients $\beta_0 = -0.0009$, $\beta_{Theta} = 0.02$, $\sigma^2 = 67.47$ than the ratings of the coefficients of the second model $\beta_0 = 0.0009$, $\beta_{Delta} = 0.016$, $\sigma^2 = 47.30$. However, in the test for the β_0 values of the test statistic for the edge tests are different from zero in each described model. The tests carried out allowed rejection of the null hypothesis about the zero effect of the mean effect, which allowed the selection of models with the mean, i.e., models with higher ratings of the coefficients $\beta_0, \beta_1, \sigma^2$: Alpha~SMR, Theta~Alpha, Delta~Theta. Thus, the significant predictor for the Alpha wave was the SMR wave, for the Theta wave was the Alpha wave, and for the Delta wave was the Theta wave.

Table 2
The dependence of the frequency of Beta 1-5 and HighBeta brainwaves in EEG biofeedback measurements

Residuals	Min	1Q	Median	3Q	Max	F-statistic	Multiple R-squared	Estimate	Intercept Estimate	
B1~HB	-13.17	-5.02	-0.11	4.92	12.80	14.13 ***	0.0003	0.004	0.0003	40.51
HB~B1	-69.05	-15.39	-0.04	15.57	71.93			6.909e-02	-4.832e-05	632.79
B1~B2	-13.42	-4.94	0.00	4.81	13.83	396.9 ***	0.009	0.080*	0.0003	40.18
B2~B1	-14.65	-6.15	-0.02	6.27	14.78			1.07e-01	-2.283e-05	54.08
B1~B3	-13.14	-4.96	-0.10	4.95	12.82	0.20	4.37e-06	-0.001	0.0003	40.53
B3~B1	-22.62	-8.22	0.10	8.18	23.08			-0.003	-0.0003	98.60
B1~B4	-13.15	-4.96	-0.10	4.95	12.84	0.06	1.21e-06	-0.0004	0.0003	40.53
B4~B1	-51.92	-12.73	0.34	12.34	52.05			-0.003	-0.0003	346.05
B1~B5	-13.16	-4.97	-0.11	4.95	12.83	0.003	6.507e-08	3.801e-05	2.577e-04	40.53
B5~B1	-108.07	-28.75	-0.10	29.04	108.26			0.002	0.0001	1825.4
B2~B3	-14.59	-6.29	-0.07	6.40	14.52	515 ***	0.011	7.809e-02	2.923e-05	53.94
B3~B2	-23.83	-8.27	0.03	8.05	24.57			0.141	-0.003	97.51
B2~B4	-14.07	-6.31	-0.05	6.62	14.25	4.11*	8.886e-05	-3.742e-03	3.743e-06	54.54
B4~B2	-52.11	-12.56	0.25	12.42	52.20			-0.024	-0.0003	346.02
B2~B5	-13.99	-6.37	-0.04	6.57	14.23	0.15	3.122e-06	3.060e-04	4.723e-06	54.55
B5~B2	-107.95	-28.71	-0.10	29.00	108.19			0.0102	0.0001	1825.4
B3~B4	-24.80	-8.22	-0.06	8.29	25.02	560 ***	0.0119	0.0584	-0.0003	97.41
B4~B3	-47.42	-12.45	0.75	12.57	49.02			0.2051	-0.0002	341.90
B3~B5	-22.66	-8.18	0.09	8.12	23.02	2.417	5.231e-05	0.00168	-0.0003	98.60
B5~B3	-107.57	-28.80	-0.07	29.15	107.81			0.0311	0.0001	1825.3
B4~B5	-42.13	12.06	0.06	11.82	42.51	9215 ***	0.1663	0.1775	-0.0003	288.51
B5~B4	-100.95	-25.28	0.67	26.20	101.94			0.9365	0.0004	1521.9

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 2 shows that the strongest significant linear relationship occurred between beta 5 and beta 4 waves. The second model, in which beta 5 was the predicted variable and beta 4 was the explanatory variable, is the evaluation of the coefficients $\beta_0 = 0.0004$, $\beta_{B4} = 0.94$, $\sigma^2 = 1521.9$ was significantly higher than the estimate of the coefficients of the first model $\beta_0 = -0.0003$, $\beta_{B5} = 0.18$, $\sigma^2 = 288.51$. Linear relationships between beta waves occurred in order of decreasing strength for: beta 3 and beta 4, beta 3 and beta 2, beta 1 and beta 2, beta 1 and highbeta, beta 2 and beta 4. Beta 3 and beta 4 brainwave relationship, in the first model, the coefficients $\beta_0 = -0.0003$, $\beta_{B4} = 0.06$, $\sigma^2 = 97.41$ were lower than the coefficients of the second model $\beta_0 = -0.0002$, $\beta_{B3} = 0.21$, $\sigma^2 = 341.90$. The second model of the relationship between beta 3 and beta 2 showed higher coefficients $\beta_0 = -0.0003$, $\beta_{B2} = 0.14$, $\sigma^2 = 97.51$ than the first model $\beta_0 = 2.923e-05$, $\beta_{B3} = 7.809e-02$, $\sigma^2 = 53.94$. The relationship between beta 1 and beta 2 waves achieved higher coefficients in the first model $\beta_0 = 0.0003$, $\beta_{B2} = 0.08$, $\sigma^2 = 40.18$ than in the second model $\beta_0 = -2.283e-05$, $\beta_{B2} = 1.07e-01$, $\sigma^2 = 54.08$. The Highbeta and beta 1 dependency model achieved

higher coefficient ratings in the first model $\beta_0 = 0.0003$, $\beta_{B1} = 0.004$, $\sigma^2 = 40.51$ than in the second model $\beta_0 = -4.832e-05$, $\beta_{HB} = 6.909e-02$, $\sigma^2 = 632.79$. A significant relationship, but the weakest of those described above, occurred between the beta 2 and beta 4 waves and vice versa, where the assessment of the coefficients of the second model $\beta_0 = -0.0003$, $\beta_{B2} = -0.02$, $\sigma^2 = 346.02$ was higher than that of the first model $\beta_0 = 3.743e-06$, $\beta_{B4} = -3.742e-03$, $\sigma^2 = 40.53$.

In the next step, it was checked whether there were significant differences in the measurements after and before the exercise. For this purpose, the difference was calculated by subtracting the average value of measurements before exercise from the average value of measurements after exercise. The percentage ratio was calculated from the numerical difference from the pre-exercise measurement and presented as an absolute value.

Table 3

Descriptive statistics of the frequency of EEG recordings before and after exercise and the ratio of the number and percentage of their differences.

EEG recording	Exercise endurance	Mean	Standard Dev	Skewness	Kurtosis	Average difference	
						Numeric	%
recording	after	-98,313	50,490	0,075	0,08967	0.099	0.001%
	before	-98,412	50,440	0,089	0,09584		

After a three-minute measurement, the EEG record differed only by 0.001% from the recording before exercise, which means that the average frequency reached slightly higher values after training, i.e., increased by 0.099.

Table 4

Descriptive statistics of Delta, Theta, Alpha and SMR wave frequencies before and after exercise, and the ratio of the number and percentage of their differences

Brain waves	Exercise endurance	Mean	Standard Dev	Skewness	Kurtosis	Average difference	
						Numeric	%
Delta	after	-0,0009	8,215	-0,410	-0,605	0.0045	5%
	before	-0,0054	8,830	-0,393	-0,365		
Theta (4-8 Hz)	after	0,0008	6,879	-0,009	-0,642	-0.0012	1.5%
	before	0,002	6,937	-0,011	-0,632		
Alpha (8-12 Hz)	after	0,0001	5,814	-0,002	-0,990	-0.0005	5%
	before	0,0006	5,814	-0,002	-0,969		
SMR (12-15 Hz)	after	-0,0002	5,255	0,000	0,007	0.0001	0.5%
	before	-0,0003	5,229	0,000	0,043		

The post-training delta waveform recording differed by 5% from the pre-training recording, which means that the average frequency of the waveform increased by 0.0045. Also, the mean SMR frequency wave after training was increased by 0.0001, and the percentage difference after and before training was small 0.5%. A reduction in wave frequency after training was observed for theta wave -0.0012 (1.5%) and alpha wave -0.0005 (5%).

Table 5

Descriptive statistics of Beta 1-5 and HighBeta wave frequencies before and after exercise, and the ratio of the number and percentage of their differences.

Beta Waves	Exercise endurance	Mean	Standard Dev	Skewness	Kurtosis	Average difference	
						Numeric	%
Beta1 (15-18 Hz)	after	0,0003	6,366	-0,0006	-1,002	-0,0006	2%
	before	0,0009	6,346	-0,0008	-1,004		
HighBeta (18-30 Hz)	after	x	25,159	-0,0019	0,294	0	0%
	before	x	25,117	-0,0019	0,305		
Beta 2 (18-22 Hz)	after	0	7,386	0,0000	-1,159	-0,0003	-
	before	0,0003	7,371	-0,0001	-1,156		
Beta 3 (22-26 Hz)	after	-0,0003	9,930	0,0001	-0,790	0,0002	-0.67%
	before	-0,0005	9,916	0,0001	-0,781		
Beta 4 (26-30 Hz)	after	-0,0003	18,603	0,0000	-0,096	0,0009	3%
	before	-0,0012	18,584	0,0002	-0,093		
Beta 5 (30-38 Hz)	after	0,0001	42,725	0,0000	0,014	-0,0001	1%
	before	0,0002	42,599	0,0000	0,028		

The greatest differences in beta wave recordings were observed for Beta 4. The wave frequency after training differed by 3% from the recording before training, which means that the average wave frequency was increased by 0.0009. Significantly different changes were observed for the Beta 1 wave. After training, the frequency was reduced (-0.0006) and the percentage difference after and before training was 2%. A reduction in wave frequency after training was also observed for wave Beta 5 -0.0001 (1%) and Beta 3 -0.0003 (0.67%).

CONCLUSION

Based on the scientific reports it can be concluded that new technologies pose scientists broad analytical and methodical possibilities. In this paper we presented an overview of the broadening of usage of the statistical data and analysis of the implementation of Big Data for monitoring the peripheral nervous system and electroencephalography biofeedback. It is especially important for sports medical science, as provided information is useful to establish the correct diagnosis and plan appropriate and personalised treatment for disparities of human movement. Big Data brings also new perspectives in the sciences of physical culture (Welsch, Bird, & Mayhew, 2005). The most important capabilities include: monitoring frequency of brain waves before and after exercise, matching and optimisation of the process of preparation adaptive system for exercise (ex. for athletes) and detecting differences motivated psycho neuronal (Santana, Vera-Garcia, & McGill, 2007).

REFERENCES

- Milliken G. A., & Johnson D. E. (1992). *Analysis of messy data volume 1: Designed experiments*. CA: Belmont. Lifetime Learning Publications.
- Shvachko K., Kuang H., Radia S., & Chansler R. (2010). The Hadoop Distributed File System. *IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST)*. 1-10.
- Changqing J., Yu L., Wenming Q., Awada U., & Keqiu L. (2012). Big Data Processing in Cloud Computing Environments. *IEEE 12th International Symposium on Pervasive Systems, Algorithms and Networks (ISPAN)*. 17-23.
- Liu X., Wang X., Matwin S., & Japkowicz N. (2015). Meta-MapReduce for scalable data mining. *Journal of Big Data*, 2, 1-23.
- Welsch E.A., Bird M., Mayhew J.L. (2005). Electromyographic activity of the pectoralis major and anterior deltoid muscles during three upper-body lifts. *J. Strength Cond. Res.*, 19, 449-452.
- Santana J.C., Vera-Garcia F.J., & McGill S.M. (2007). A kinetic and electromyographic comparison of the standing cable press and bench press. *J. Strength Cond. Res.*, 21(4), 1271-1277.