

Linear and Nonlinear Prediction Models Show Comparable Precision for Maximal Mean Speed in a 4x1000 m Field Test

Jäger, J. M., Kurz, J., Müller, H.

Institute of sport science, Justus-Liebig-University Gießen

Abstract

Maximal oxygen uptake ($\text{VO}_{2\text{max}}$) is one of the most distinguished parameters in endurance sports and plays an important role, for instance, in predicting endurance performance. Different models have been used to estimate $\text{VO}_{2\text{max}}$ or performance based on $\text{VO}_{2\text{max}}$. These models can use linear or nonlinear approaches for modeling endurance performance. The aim of this study was to estimate $\text{VO}_{2\text{max}}$ in healthy adults based on the Queens College Step Test (QCST) as well as the Shuttle Run Test (SRT) and to use these values for linear and nonlinear models in order to predict the performance in a maximal 1000 m run (i.e. the speed in an incremental 4x1000 m Field Test (FT)). 53 female subjects participated in these three tests (QCST, SRT, FT). Maximal oxygen uptake values from QCST and SRT were used as (a) predictor variables in a multiple linear regression (MLR) model and as (b) input variables in a multilayer perceptron (MLP) after scaling in preprocessing. Model output was speed [$\text{km}\cdot\text{h}^{-1}$] in a maximal 1000 m run. Maximal oxygen uptake values estimated from QCST ($40.8 \pm 3.5 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) and SRT ($46.7 \pm 4.5 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) were significantly correlated ($r = 0.38$, $p < 0.01$) and maximal mean speed in the FT was $12.8 \pm 1.6 \text{ km}\cdot\text{h}^{-1}$. Root mean squared error (RMSE) of the cross validated MLR model was $0.89 \text{ km}\cdot\text{h}^{-1}$ while it was $0.95 \text{ km}\cdot\text{h}^{-1}$ for MLP. Results showed that the accuracy of the applied MLP was comparable to the MLR, but did not outperform the linear approach.

KEYWORDS: MULTIPLE LINEAR REGRESSION, MULTILAYER PERCEPTRON, PERFORMANCE PREDICTION, ARTIFICIAL NEURAL NETWORK

Introduction

The maximal oxygen uptake (VO_2max) measures the maximum amount of oxygen that an individual can use per unit of time during strenuous physical exertion. It is one of the most distinguished parameters in endurance sports and serves as an index of cardiorespiratory function, general health, and aerobic fitness (Deuster & Heled, 2008). Besides its physiological meaning, VO_2max is also relevant in exercise prescription. The intensity of cardiorespiratory exercise is commonly quantified as a percentage of VO_2max (ACSM, 2011). Consequently, assessing the maximal oxygen uptake plays an important role in endurance sports. It provides a basis for measuring aerobic power or designing programs to improve cardiorespiratory fitness.

A broad variety of studies focused on determining VO_2max . Besides those studies which measure VO_2max directly in a maximal incremental exercise test on a treadmill or cycle ergometer, there are also some approaches that try to predict VO_2max based on non-exercise and submaximal exercise test data (e.g., Black, Vehrs, Fellingham, George & Hager, 2016) or even based on statistical models using solely non-exercise data (see Abut & Akay, 2015, for a review). In case of non-exercise models, variables such as sex, age or body mass index were used to predict VO_2max . Especially these non-exercise models may be of interest, since this does not require collection of expired respiratory gases and thus it does not depend on expensive indirect calorimetry equipment (Marshall, Coe & Pivarnik, 2014). The Shuttle Run Test is one of the most frequently applied test to assess VO_2max without directly measuring respiratory gases. Mayorga-Vega, Aguilar-Soto and Viciana (2015) found in a meta-analysis that the 20-m Shuttle Run Test (SRT) had a moderate-to-high criterion-related validity for estimating VO_2max . So, it seems to be a useful alternative for estimating cardiorespiratory fitness, if attaining the maximum oxygen uptake during a laboratory-based test is unfeasible.

Since VO_2max is an important measure for cardiorespiratory fitness, it is also one of the predominant limiting factors in endurance exercises and training, and may also serve as a predictor for performance in endurance sports. Therefore, some attempts have been made to predict the performance in endurance sports based on VO_2max . For instance, McLaughlin, Howley, Bassett Jr., Thompson and Fitzhugh (2010) used VO_2max (as well as other variables that are linked to endurance performance) in order to predict running performance in a 16 km time trial. Also in shorter distances such as 5 and 10 km, VO_2max seems to be a strong predictor of running speed (Bird, Theakston, Owen & Nevill, 2003). High VO_2max values could be considered – among other factors – as one of the most influential determinants of performance in distance running (Takeshima & Tanaka, 1995). However, in homogeneous groups like elite runners, the prediction of running performance based on VO_2max might not be appropriate. In marathon runners, for instance, other variables like lactate threshold are better predictors since the speed associated with lactate threshold closely equates to the race pace of these athletes (Shave & Franco, 2006).

Different models can be used to predict performance in sports, including linear models such as linear or multiple linear regression or nonlinear models such as artificial neural networks and especially multilayer perceptrons. There are many applications of artificial neural networks in sports (Schöllhorn, Jäger & Janssen, 2008). As an example of performance prediction, Maszczyk, Rocznoik, Waskiewicz, Czuba, Mikolajec, Zajac and Stanula (2012) modeled competitive swimming performance in 50 m front crawl with multilayer perceptron neural models using anthropometric data or specific swimming skills as input parameters. When it comes to comparing linear and nonlinear models with regard to prediction accuracy, the quality of prediction of neural models seems to be similar to that of regression analysis and regression models, potentially even better. This seems to hold true in VO_2max prediction (Abut, & Akay, 2015; Akay, Zayid, Aktürk, & George, 2011) as well as in performance

prediction studies (Edelmann-Nusser, Hohmann, & Henneberg, 2002; Maszczyk, Zajac, & Rygula, 2011; Maszczyk, Rocznoik, Waskiewicz, Czuba, Mikolajec, Zajac, & Stanula, 2012). Hence, Jäger, Kurz and Müller (2016) assume that nonlinear neural network approaches possibly provide more sophisticated methods for predicting maximal mean speed in a 4x1000 m Field Test which may enhance the accuracy of their linear model.

This study aimed to: (a) estimate the VO_2max in healthy adults by two common endurance tests and (b) construct and evaluate both a linear and a nonlinear prediction model for the maximal mean speed in a 4x1000 m Field Test (FT) based on these estimated VO_2max values. Therefore, we expanded the approach of Jäger, Kurz and Müller (2016) in the following ways. Firstly, we used an increased number of subjects to build and validate the models for women. Secondly, an artificial neural network (multilayer perceptron) was built in order to compare the accuracy for predicting speed between linear and nonlinear models. Additionally, we used the improved regression formulas from Stickland, Petersen and Bouffard (2003) in this study for estimating the VO_2max based on the shuttle run test.

Methods

53 female subjects participated in this study (age 23.0 ± 2.9 years, mass 60.2 ± 6.1 kg, height 169.0 ± 6.4 cm). Informed written consent was obtained from each participant, and the study was in accordance with the ethical guidelines of the Helsinki Declaration. Each subject was tested in the following three exercise tests:

Endurance tests

The Shuttle Run Test (SRT):

The Shuttle Run Test (e.g., Léger, Mercier, Gadoury & Lambert, 1988; Ramsbottom, Brewer & Williams, 1988) is a well-established endurance test. During this test, subjects ran shuttles between two marked lines placed 20 m apart at increasing fast speeds. Running speed was increased each minute from one level to another. Accordingly, each level consisted of a different number of shuttles within that minute (from 7 runs in level 1 up to 16 runs in level 21). Subjects were verbally encouraged to give maximum effort in this test. VO_2max of subjects were determined based on their successfully completed levels and runs (X equals the last half-stage of the SRT completed; see Stickland, Petersen and Bouffard (2003) for details). The following regression equation was developed for women:

$$\text{VO}_2\text{max} [\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}] = 2.85 \cdot X + 25.1 \quad (1)$$

The SRT was conducted one week before the following two tests and within the same 90 min time slot on that day. Thus, a participant who was tested antemeridian was also tested antemeridian seven days later.

The Queens College Step Test (QCST):

According to Haff and Dumke (2012) the test started with a 3 min rest when the subjects sat on a bench step (height: 41 cm). For women, a metronome was set at 88 beats/min leading to 22 steps per minute. Subjects made contact with a foot on each beep of the metronome in an up-up-down-down manner. After exactly 3 min of stepping, the subjects stopped and palpated for the radial pulse at exactly 3:05 to 3:20 min. Recovery heart rate (HR) was derived for a full minute (bpm) and used to calculate VO_2max based on the following formula for female subjects from Haff and Dumke (2012):

$$\text{VO}_2\text{max} [\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}] = 65.81 - (0.1847 \cdot \text{HR}) \quad (2)$$

The 4x1000 m Field Test (FT):

This test from Held (2000) was done on the same day as the QCST. For this test, written instructions defined the different 1000 m running intensities according to (a) the usual durations of such runs, (b) a rating of perceived exertion after these runs, and (c) information about the breathing during the runs (e.g., “slow” refers to an intensity of a 1 h jog which is a bit tiring but not exhausting, where breathing and talking is easy). Participants were also instructed to run a preferably constant pace in each trial (Held, Steiner, Hübner, Tschopp, Peltola & Marti, 2000). Accordingly, the subjects were to choose their running speeds individually. The first three runs requested incremental running speeds corresponding to common training intensities (slow, medium, and fast). The last run had to be performed with maximal effort and thus corresponded to the maximal mean speed that the subjects were able to run for 1000 m. Between each run a two-minute rest was allowed. Due to the QCST, which is a submaximal test and was done a few minutes before the FT, and due to the first 1000 m run, which was “slow” and equaled a low intensity jog, no particular warm up was done before the FT. The test ended with an individual cool down (walk or jog).

One week before the first test, the test instructions were handed out to the participants and reminded them to avoid exercise for the previous 24 h, to fasten for at least 2 h prior to the test, and to avoid the use of foods and drugs that alter heart rate (e.g., coffee, soda, energy drinks, diet pills, beta-blockers) (Haff & Dumke, 2012).

The SRT and QCST were chosen, because they estimated the subjects' VO_2max from different points of view. SRT is a maximal effort test that aims at an athlete's ability to continue an incremental endurance activity and primarily to resist fatigue. In contrast, QCST is a submaximal effort test which measures the recovery heart rate of an athlete, because it returns to resting values more quickly in fitter people than it does in those who are less fit (Darr, Bassett, Morgan & Thomas, 1988; Dimpka, 2009).

Linear prediction model: Multiple linear regression

For the comparison of linear and nonlinear prediction models a multiple linear regression (MLR) was used to predict the maximal mean speed in a FT. In general, this approach helps to quantify the relationship between several independent (predictor) variables and the dependent (criterion) variable based on the assumption of linearity. The VO_2max estimates from SRT and QCST were used in this multiple regression model as predictor variables. Both predictor variables were expressed in units of $\text{ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$. Maximal mean speed in the last 1000 m run (in FT) served as the criterion variable (i.e., the outcome) and was measured in $\text{km}\cdot\text{h}^{-1}$. The predictors were tested for multicollinearity using variance inflation factors (VIF) and yielded VIF for both predictor variables below 1.17. Therefore, multicollinearity does not seem to be an issue in this regression analysis and it is acceptable for both variables to be included in the model, although the interpretation of VIF still has to be done carefully (O'Brian, 2007).

Nonlinear prediction model: Multilayer perceptron

A multilayer perceptron was used in a nonlinear approach to predict maximal mean speed. As a feedforward artificial neural network, the multilayer perceptron (MLP) is capable to classify even not linearly separable data and can be applied to learn nonlinear function mappings. The MLP architecture consisted of two input neurons (for the VO_2max values from the subjects' in QCST and in SRT) and three hidden neurons in the second layer with logistic activation functions. The output layer, i.e. the third layer, had one neuron representing the maximal mean speed [$\text{km}\cdot\text{h}^{-1}$] in the FT and used a linear activation function. Before training the network, data were preprocessed by scaling input to [0;1] so that they have a minimum of zero, a

maximum of one, and also a range of one. In the training phase, a resilient backpropagation algorithm with weight backtracking was used. This algorithm is based on the traditional backpropagation algorithm (e.g., Kattan, Abdullah & Geem, 2011). But, separate learning rates which can be changed during training process are used for each weight of the network (Günther & Fritsch, 2010). Therefore no over-all learning rate had to be predefined. A threshold of 0.005 was set as stopping criteria in the training process.

Model validation

For validating the models, first, the root mean squared prediction error (RMSE) was determined. This error was calculated for both the linear and the nonlinear model built on the whole data set and was used to measure the model fit. Second, a 10 fold cross validation was applied. Therefore, the whole data set was split randomly into 10 mutually disjoint subsets (i.e., the folds) leading to seven subsets with five cases and three subsets with six cases. Subsequently, one fold was excluded and the models (MLR and MLP) were built on the other folds. This procedure was repeated 10-times, while every single fold was excluded once in order to provide a test set (Schöllhorn, Jäger & Janssen, 2008). The cross validated root mean squared prediction error (cvRMSE) of the models' predictions on the test sets was not only calculated to compare the accuracy of the two models but also in order to account for overfitting problems in modeling.

All calculations were done with R (3.2.4) (R Core Team, 2016) – especially the neuralnet and DAAG packages.

Results

Estimated VO_2max values were $40.8 \pm 3.5 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ in QCST and $46.7 \pm 4.5 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ in SRT. Results from QCST and SRT were significantly correlated ($r = 0.38$, $p < 0.01$). Maximal mean speed in the FT was $12.8 \pm 1.6 \text{ km}\cdot\text{h}^{-1}$.

MLR analysis was used to test if the estimated VO_2max values from QCST and SRT significantly predicted participants' maximal mean speed. The results of the regression model indicated that 70% of the variance in maximal mean speed can be explained ($R^2_{\text{ad}} = 0.70$, $F(2,50) = 62.94$, $p < 0.01$). However, in this model only the VO_2max estimate from the SRT had a significant ($p < 0.01$) partial effect as a predictor. The accuracy of this model was obtained through the RMSE of prediction. An RMSE of $0.82 \text{ km}\cdot\text{h}^{-1}$ was found for the model fit (built on the whole data set). Furthermore, the cvRMSE of the predicted speed compared to the real speed was calculated and resulted in an error of $0.89 \text{ km}\cdot\text{h}^{-1}$.

The architecture of the net and the results of the neural network trained on the whole data set are shown in figure 1. RMSE was $0.74 \text{ km}\cdot\text{h}^{-1}$, which was the smallest error found in this study. Thus, the neural network model was able to fit the data best. However, cvRMSE of the net was derived by comparing the networks' predicted results, which were built on the training folds, to the measured maximal mean speed in the FT in the test sets. This led to a cvRMSE of $0.95 \text{ km}\cdot\text{h}^{-1}$, which is about $0.06 \text{ km}\cdot\text{h}^{-1}$ higher than the cross validated error of the linear model and $0.13 \text{ km}\cdot\text{h}^{-1}$ higher than its RMSE.

A detailed plot of the predictions from the linear model (MLR) and nonlinear neural network model (MLP) is shown in figure 2. Predictions are plotted against the measured speeds of subjects.

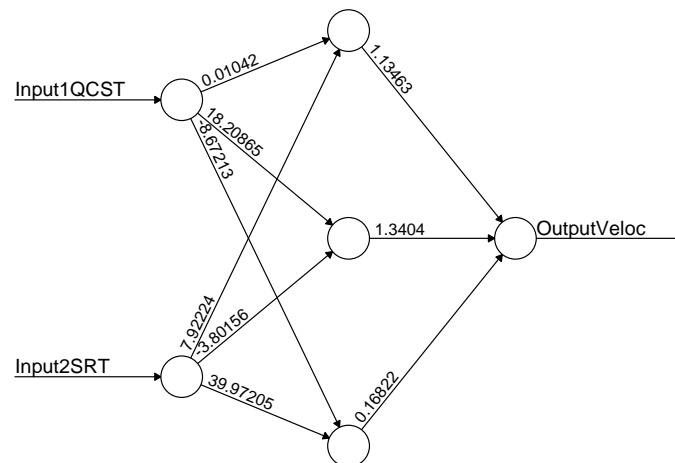


Figure 1. Model of the trained feed-forward artificial neural network (multilayer perceptron) for predicting maximal mean speed in a 1000 m run.

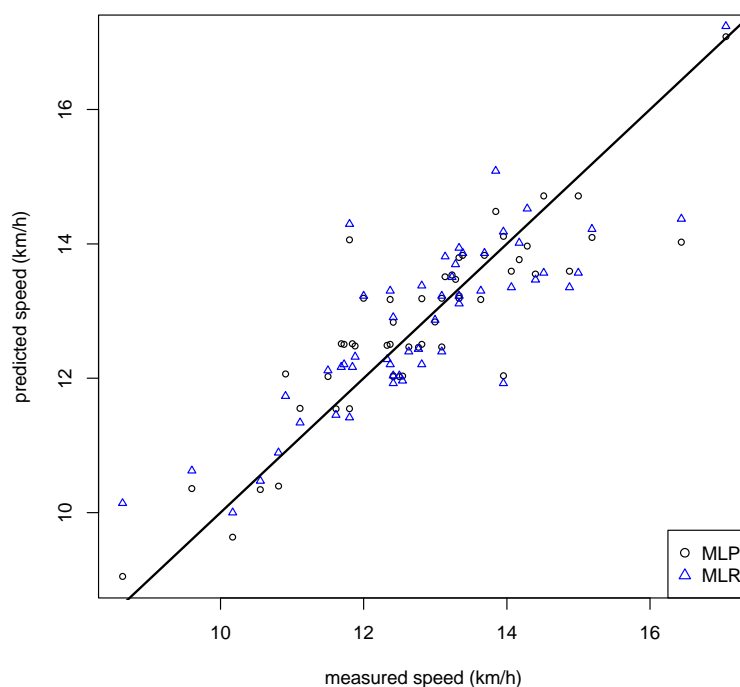


Figure 2: Predicted versus measured speed of both the linear model (MLR) and the nonlinear model (MLP) for all 53 subjects.

Discussion

The estimated VO_2max from both the SRT and the QCST are able to explain a remarkable amount (70%) of variance in maximal mean speed in a 1000 m run. This result of the MLR analysis is in line with the concept that VO_2max is a good predictor of endurance running performance in recreationally active individuals. However, it is mainly the VO_2max estimate from the SRT that has a major influence on the predicted speed in MLR analysis. Therefore, the SRT appears to be more appropriate than the QCST for assessing the maximal mean speed in the FT. That might be due to the faster and more complex movement of running in the SRT

compared to the slower, less complex movement of stepping up and down a bench. Since the participants in this study vary in their ability to run economically and because this ability is also inherent in the VO_2max estimates, the prediction of the maximal mean speed may rely stronger on the SRT's VO_2max estimate.

However, there is still 30% variance left that could not be explained by the linear model. Many aspects may have contributed to this unexplained variance and error in prediction. For instance, the motivation of the subjects in maximal effort tests (such as the SRT and the last trial in FT) and environmental factors could have influenced the results profoundly. With respect to environmental factors, the tests were performed on varying floor types and different days and therefore with varying weather conditions. Other aspects may be related to the different facets of endurance. The FT requires resistance to fatigue (in each run) as well as a quick recovery (between two runs). Thus, a specific combination of these two aspects is important for this test and affects the maximal mean speed in the last 1000 m run. The SRT and the QCST, however, focus mainly on one of these two aspects rather than the combination of both. Last but not least, the athletes' anaerobic energetics might have an impact, too. In a short duration 1000 m run, anaerobic metabolism contributes substantially to the total energy expenditure (Newsholm, Blomstrand & Ekblom, 1992; van Someren, 2006). Therefore, variations in the anaerobic energetics between athletes might also contribute to the observed variance in running speed. Taken together, these aspects may limit the prediction of speed exclusively based on estimated VO_2max values of subjects.

Furthermore, even though the estimated VO_2max values from SRT and QCST were significantly correlated, both tests did not measure oxygen uptake directly with a spirometer. Although the SRT was validated with respect to estimating VO_2max (Mayorga-Vega, Aguilar-Soto & Viciano, 2015), we still have to consider an estimation error. For instance, Léger and Lambert (1982) found a standard error of estimate (SEE) of $5.4 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ in an early version of the SRT and Stickland, Petersen and Bouffard (2003) a SEE of $4.07 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ for males and $3.64 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ for females. This applies analogously to the inaccuracy of the QCST results. While formulas from Haff and Dumke (2012) appear to be very accurate, McArdle, Katch, Pechar, Jacobson and Ruck (1972) found a standard error of prediction of $2.9 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ in their test population. Hence, the two predictor variables in our MLR and accordingly the input variables of the MLP may be inaccurate to a certain degree, which also affects the accuracy in predicting the speed as outcome variable. Finally, modeling in this study also neglects other potential predictors like age or further experiences with endurance sports, which can provide a basis to improve the quality of prediction. Altogether, these aspects certainly have contributed to the prediction errors.

Although a RMSE of prediction with approximately $0.82 \text{ km}\cdot\text{h}^{-1}$ in the applied linear model and $0.74 \text{ km}\cdot\text{h}^{-1}$ in the nonlinear model seems to be small, it may still be too big to allow detailed recommendations for training. Especially, when we take the cross validated errors with $0.89 \text{ km}\cdot\text{h}^{-1}$ (linear model) and $0.95 \text{ km}\cdot\text{h}^{-1}$ (nonlinear model) into account. It should be noted, though, that this difference in cross validated prediction accuracy of both the linear and nonlinear models is not substantial. The MLP with its chosen specifications (architecture, algorithm etc.) was not superior to the linear regression model in this study. Nevertheless, other studies had found nonlinear models to allow better predictions (Edelmann-Nusser, Hohmann, & Henneberg, 2002; Maszczyk, Zajac, & Ryguła, 2011; Maszczyk, Rocznoik, Waskiewicz, Czuba, Mikolajec, Zajac, & Stanula, 2012).

A next step should be to build and evaluate models for male athletes or to add further variables that might influence and explain performance in a 1000 m run. Additionally, other statistical approaches like random forests analyses (Breiman, 2001; Verikas, Gelzinis & Bacauskiene,

2011) may enhance the performance prediction. However, a major amount of variance in maximal mean speed can already be explained from the linear model approach chosen in this study.

Conclusion

The current study compared the accuracy of linear and nonlinear models for predicting performance in a 1000 m run. The models, i.e. a multiple linear regression model and a multilayer perceptron, were built using estimated VO_2max from Queens College Step Test as well as the Shuttle Run Test and predicted the maximal mean speed in an incremental 4x1000 m field test. Although there is some evidence that nonlinear models may outperform linear models in performance prediction, both models in this study showed a comparable precision. Therefore, it is incidental that possible advantages of linear or nonlinear models depend on the specific phenomenon they are applied to.

References

- Abut, F., & Akay, M. F. (2015). Machine learning and statistical methods for the prediction of maximal oxygen uptake: recent advances. *Medical Devices: Evidence and Research*, 8, 369–379. doi: 10.2147/MDER.S57281
- Akay, M. F., Zayid, E. I. M., Aktürk, E., & George, J. D. (2011). Artificial neural network-based model for predicting VO_2max from a submaximal exercise test. *Expert Systems with Applications*, 38, 2007–2010. doi: 10.1016/j.eswa.2010.07.135
- ACSM. (2011). Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Medicine & Science in Sports & Exercise*, 43(7), 1334–1359. doi: 10.1249/MSS.0b013e318213febf
- Black, N. E., Vehrs, P. R., Fellingham, G. W., George, J. D., & Hager, R. (2016) Prediction of VO_2max in Children and Adolescents Using Exercise Testing and Physical Activity Questionnaire Data. *Research Quarterly for Exercise and Sport*, 87(1), 89–100, doi: 10.1080/02701367.2015.1124969
- Bird, S. R., Theakston, S. C., Owen, A. & Nevill, A. M. (2003). Characteristics Associated With 10-km Running Performance Among a Group of Highly Trained Male Endurance Runners Age 21–63 Years. *Journal of Aging and Physical Activity*, 11, 333–350. doi: 10.1123/japa.11.3.333
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.
- Darr, K. C., Bassett, D. R., Morgan, B. J., Thomas, D. P. (1988). Effects of age and training status on heart rate recovery after peak exercise. *American Journal of Physiol*, 254(2), H340-343.
- Deuster, P., & Heled, Y. (2008). Testing for maximal aerobic power. In P. H. Seidenberg, & A. I. Beutler (Eds.), *The sports medicine resource manual* (pp. 520–528). Philadelphia: Elsevier. doi: 10.1016/B978-141603197-0.10069-2
- Dimpka, U. (2009). Post-exercise heart rate recovery: an index of cardiovascular fitness. *Journal of Exercise Physiology online*, 12(1), 19-22.
- Edelmann-Nusser, J., Hohmann, A., & Henneberg, B. (2002). Modeling and Prediction of Competitive Performance in Swimming Upon Neural Networks. *European Journal of Sport Science*, 2(2) [electronic resource].
- Günther, F., & Fritsch, S. (2010). neuralnet: Training of neural networks. *The R Journal*, 1(2), 30–38.

- Haff, G. G., & Dumke, C. (2012). *Laboratory manual for exercise physiology*. Champaign, IL: Human Kinetics.
- Held, T. (2000). Überprüfung der Ausdauerleistungsfähigkeit. Der 4x1000-m-Lauftest (mobilepraxis) [Assessing endurance performance. The 4x1000-m-Test]. *mobile - Die Fachzeitschrift für Sport*, 6, 5–9.
- Held, T., Steiner, R., Hübner, K., Tschopp, M., Peltola, K., & Marti, B. (2000). Selbst gewählte submaximale Laufgeschwindigkeiten als Prädiktoren des Dauerleistungsvermögens [Self-selected submaximal running velocities as predictors of endurance capacity]. *Schweizerische Zeitschrift für Sportmedizin und Sporttraumatologie*, 2(48), 64–69.
- Jäger, J. M., Kurz, J., & Müller, H. (2016). Predicting maximum speed in a 4x1000 m Field Test based on estimated VO₂max values from Shuttle Run Test and Queens College Step Test. In D. Saupe, T. Dahmen, C. Widmann, I. Baiker, M. Gratkowski, S. Wolf, & R. Bertschinger (Eds.), *Proceedings dvs-Workshop Modelling in Endurance Sports* (pp. 20–24). Konstanz: Universität Konstanz. URI <http://nbn-resolving.de/urn:nbn:de:bsz:352-0-371586>
- Kattan, A., Abdullah, R., & Geem, Z. W. (2011). *Artificial neural network training and software implementation techniques*. New York: Nova Science.
- Léger, L. A., & Lambert, J. (1982). A maximal multistage 20-m shuttle run test to predict VO₂max. *European Journal of Applied Physiology*, 49(1), 1–12.
- Léger, L. A., Mercier, D., Gadoury, C., & Lambert, J. (1988). The multistage 20 metre shuttle run test for aerobic fitness. *Journal of Sports Sciences*, 6(2), 93–101. doi: 10.1080/02640418808729800
- Marshall, M. R., Coe, D. P., & Pivarnik, J. M. (2014). Development of a Prediction Model to Predict VO₂peak in Adolescent Girls Using the Bruce Protocol to Exhaustion. *Research Quarterly for Exercise and Sport*, 85(2), 251–256. doi: 10.1080/02701367.2014.893053
- Mayorga-Vega, D., Aguilar-Soto, P., & Viciano, J. (2015). Criterion-Related Validity of the 20-M Shuttle Run Test for Estimating Cardiorespiratory Fitness: A Meta-Analysis. *Journal of Sports Science and Medicine*, 14(3), 536–547.
- Maszczyk, A., Rocznoik, R., Waskiewicz, Z., Czuba, M., Mikolajec, K., Zajac, A. & Stanula, A. (2012). Application of regression and neural models to predict competitive swimming performance. *Perceptual and Motor Skills*, 114(2), 610–626. doi: 10.2466/05.10.PMS.114.2.610-626
- Maszczyk, A., Zajac, A., & Ryguła, I. (2011). A neural network model approach to athlete selection. *Sports Engineering*, 13(2), 83–93. doi: 10.1007/s12283-010-0055-y
- McArdle, W. D., Katch, F. I., Pechar, G. S., Jacobson, L., & Ruck, S. (1972). Reliability and interrelationships between maximal oxygen intake, physical work capacity and step-test scores in college women. *Medicine & Science in Sports & Exercise*, 4(4), 182–186. doi: 10.1249/00005768-197200440-00019
- McLaughlin, J. E., Howley, E. T., Bassett Jr., D. R., Thompson, D. L., & Fitzhugh, E. C. (2010). Test of the Classic Model for Predicting Endurance Running Performance. *Medicine & Science in Sports & Exercise*, 42(5), 991–997. doi: 10.1249/MSS.0b013e3181c0669d
- Newsholme, E. A., Blomstrand, E., & Ekblom, B. (1992). Physical and mental fatigue: Metabolic mechanisms and importance of plasma amino acids. *British medical bulletin*, 48(3), 477–495.
- O'Brian, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41, 673–690. doi: 10.1007/s11135-006-9018-6

- Schöllhorn, W. I., Jäger, J. M., & Janssen, D. (2008). Artificial neural network models of sports motion. In Y. Hong, & R. Bartlett, *Handbook of Biomechanics and Human Movement Science* (pp. 50–64). London: Routledge.
- Shave, R., & Franco, A. (2006). The physiology of endurance training. In G. P. Whyte, *The physiology of training* (pp. 61-84). Philadelphia, PA: Elsevier.
- Stickland, M. K., Petersen, S. R., & Bouffard, M. (2003). Prediction of maximal aerobic power from the 20-m multi-stage shuttle run test. *Canadian Journal of Applied Physiology*, 2(28), 272–282.
- van Someren, K. A. (2006). The physiology of anaerobic endurance training. In G. P. Whyte, *The physiology of training* (pp. 85-115). Philadelphia, PA: Elsevier.
- Verikas, A., Gelzinis, A., & Bacauskiene, M. (2011). Mining data with random forests: A survey and results of new tests. *Pattern recognition*, 44, 330-349.
- R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ramsbottom, R., Brewer, J., & Williams, C. (1988). A progressive shuttle run test to estimate maximal oxygen uptake. *British Journal of Sports Medicine*, 22(4), 141–144.
- Takeshima, N., & Tanaka, K. (1995). Prediction of endurance running performance for middle-aged and older runners. *British Journal of Sports Medicine*, 29(1), 20–23. doi: 10.1136/bjism.29.1.20