# Abstract

**Purpose:** Eggbeater kick is an important skill in artistic swimming to lift the body above water level. Previous attempts to model eggbeater kick performance include complex biomechanical parameters that cannot be easily used to guide strength and conditioning training. In this study, we modelled the relationship between hip function and eggbeater kick performance in 92 elite artistic swimmers with a machine learning algorithm.

**Methods:** We assessed hip function with six isometric tests that can be easily performed on a weekly basis, without the need for a physiotherapist or a scientist.

**Results:** Our model may accurately predict future performances as the predictive error is similar to the resolution of the scale used by judges during competitions. We then provide a set of interpretation and simulations methods that showcase some of the important predictors of the eggbeater kick performance and highlight personalized strategy to reach a target performance.

**Conclusions:** By using a model that is both accurate and interpretable, practitioners could access objective decision-making support, allowing them to effectively select athletes and design personalized conditioning programs. We believe machine learning and data-driven decision-making provide sports scientists and coaches with new opportunities to enhance research and performance.

Keywords: artistic swimming; eggbeater kick; hip joint function; machine learning

# Introduction

In artistic swimming, about 40% of routine duration is spent with the head out of the water and most of that time is associated with the use of eggbeater kicks [1,2](https://paperpile.com/c/PPEkUn/wNlO+F0rd). The eggbeater kick involves a complex combination of hip, knee and foot motions [3,4](https://paperpile.com/c/PPEkUn/r7YR+kFap) to keep the body above the water surface while performing artistic arm movements. This water-treading technique is therefore a representative and important skill in artistic swimming, but also in water polo and lifesaving [3](https://paperpile.com/c/PPEkUn/r7YR). Eggbeater kicks can be performed with both legs simultaneously (termed as body boost) or alternatively (termed as sustained). Height relative to the water surface is one of the key factors of performance and is evaluated using the FINA guiding scale for height [5](https://paperpile.com/c/PPEkUn/4BnS). During team events, eggbeater kick is also needed during acrobatic moves by lower swimmers to create vertical propulsive force.

From studies on water polo, the pattern of muscle activation during sustained eggbeater kick is complex with high levels of activation, especially in the vastus medialis and biceps femoris [3,6,7](https://paperpile.com/c/PPEkUn/bGD9+r7YR+4xKy). Sanders [8](https://paperpile.com/c/PPEkUn/PScr/?noauthor=1) and Oliveira et al. [6](https://paperpile.com/c/PPEkUn/bGD9/?noauthor=1) conclude that effective technique involves high hip muscle strength, in addition to fast knee flexion and foot extension. Oliveira et al. [9](https://paperpile.com/c/PPEkUn/ebYG/?noauthor=1) found that a large abduction and flexion of the hips, as well as a fast extension and flexion of the knees are essential kinematic skills to perform an effective eggbeater kick. Homma et al. [4](https://paperpile.com/c/PPEkUn/kFap/?noauthor=1) added that knees should be maintained higher than hip joints to ensure a horizontal trajectory of the feet. Zinner et al. [10](https://paperpile.com/c/PPEkUn/hOMc/?noauthor=1) also showed that water polo players with higher isometric hip abductor muscle strength performed better in eggbeater kick. Hip joint strength appears as essential to maintain knee height and enable horizontal trajectories against water resistance using a combination of hip internal-external rotation and knee flexion-extension.

To understand the relationship between technique and performance, qualitative models of sustained and dynamic heights have been proposed by Sanders et al. [3,8](https://paperpile.com/c/PPEkUn/r7YR+PScr/?noauthor=1,0). They performed multiple linear regressions to predict height based on foot speed, range of knee extension and initial angle of the upper body. Their predictive models of sustained and dynamic height explain 90% and 79% of the variance. Oliveira [9](https://paperpile.com/c/PPEkUn/ebYG/?noauthor=1) and Homma [4](https://paperpile.com/c/PPEkUn/kFap/?noauthor=1) also reported modest to strong correlations between height and kinematics variables such as orientation and speed of the limbs. Unfortunately, these predictive models are based on small samples size () and were not cross-validated. Furthermore, they cannot easily be used as they rely on biomechanical parameters that require 3D kinematic analyses. Since the ability to generate high foot velocities is related to muscular function [6](https://paperpile.com/c/PPEkUn/bGD9), we hypothesized that hip joint strength could be used to predict eggbeater kick performance.

A recent survey pointed out the benefits of machine learning algorithms over multiple linear regressions—as used in Sanders [3](https://paperpile.com/c/PPEkUn/r7YR/?noauthor=1)—to predict physical capability as they account for complex nonlinear relationships between variables [11](https://paperpile.com/c/PPEkUn/L65F). Machine learning is gaining popularity to predict performance in sports such as triathlon [12](https://paperpile.com/c/PPEkUn/ODSy), biathlon [13](https://paperpile.com/c/PPEkUn/KoKn) or track and field [14](https://paperpile.com/c/PPEkUn/FVZ1). A review of computational intelligence applications in sports underlined the need to develop data-driven support systems that would help to design training plans, to measure indicators of an athlete’s readiness and to analyze data during training sessions [15](https://paperpile.com/c/PPEkUn/whKM). To implement such systems, accurate performance prediction is crucial. Gradient boosting algorithms are one possible approach to achieve such accuracy. These are highly customizable and powerful tools for learning and analyzing problems with heterogeneous parameters and noisy data with complex interactions [16](https://paperpile.com/c/PPEkUn/wlY9). The current study aims to investigate the relationship between hip function in elite artistic swimmers and performance in eggbeater kicks using a gradient boosting algorithm. The predictive model will be considered valid if its accuracy is similar to the FINA guiding scale for height resolution (0.5). A supplementary objective is to illustrate how the model can be used to develop a data-driven conditioning support system.

# Methods

## Experimental procedures

Artistic swimmers with provincial, national and international levels participated in the study (; age: years old; height:  cm, weight:  kg; training load:  days/week). They were free from hip pain or injury at the time of testing and had no history of hip surgery. Prior to the experimental procedure, the participants—and their guardians for swimmers under 17 years old—signed an informed consent form approved by the ethics committee (17-163-CERES-D). Swimmers performed nine tests: six tests to evaluate hip joint function (Table 1) that were later used to predict three technical skills specific to artistic swimming (Table 2). Hip function was assessed using three maximal voluntary isometric contractions (MVIC) lasting three seconds for six positions (Table 1) after a standard dry-land warm-up. Trials were recorded using the Groin Bar (Vald Performance, Queensland, Australia; 0.5 N resolution) at 50 Hz. Participants were strongly encouraged and were given at least 5-s rest between contractions [17](https://paperpile.com/c/PPEkUn/LZkq).

[insert table 1]

Participants then performed three sport-specific tests. First, the height during a double-arm body boost (BB-H) and a 15-seconds sustained eggbeater kick (EB-H) was measured. Height was estimated with a 120 frames-per-second video and evaluated using the FINA guiding scale for height [5](https://paperpile.com/c/PPEkUn/4BnS). Then, the upward force produced during a 5-s sustained eggbeater (EB-F) was measured using a hand-held dynamometer at 40 Hz (Lafayette Model 01165, Indiana, USA). Each test was repeated twice.

[insert table 2]

## Modeling

From the tests previously described, MVIC force signals were low-pass filtered using a zero-lag fourth-order 10 Hz Butterworth filter and averaged using a moving root-mean-squared average on a 200-ms window. Timeframe envelopes of each MVIC were then reduced by taking the mean of the highest consecutive values during 0.2 second. From these values, forces from left and right legs were transformed such as . This transformation favors lefts and rights that are similar and therefore penalizes asymmetries. Left-right imbalance was also computed using the relative difference. The input variables consisted of these 12 variables (left-right F-values and imbalances for 6 tests) as well as anthropometric measurements (height and mass, described in appendix 1). We used the three sport-specific tests (maximum BB-H, mean EB-H and mean EB-F) as output variables. This dataset, composed of the input and output variables mentioned above, was randomly split into training (80%, ) and test (20%, ) sets.

A gradient-boosting algorithm was fitted for each output variable (BB-H, EB-H and EB-F) with the training set using the Python Catboost library [18](https://paperpile.com/c/PPEkUn/ZQOM). This particular algorithm was chosen as it provides the best cross-validation error on our dataset. Once trained, we evaluated the generalization error on unseen data from the test set and reported the difference in mean absolute error (MAE) and mean absolute percentage error (MAPE). MAPE differences between real and predicted performance were investigated using Bayesian estimation and the procedure described in Kruschke [19](https://paperpile.com/c/PPEkUn/notd/?noauthor=1), which provides distributions of credible values for the effect size (), the group means and their differences (). We reported the mean of the posterior distribution and the 95% highest posterior density (HPD) interval, which contains 95% of the posterior distribution. We define a statistically significant difference when the HPD of the difference between predicted and real values does not contain zero. To get an overview of which variables are most important and describe the impacts each variable has on the model output, we used the shapley additive explanations (SHAP) implemented in the Shap Python library [20](https://paperpile.com/c/PPEkUn/NceT). The strength of the linear relationship between each pair of variables was assessed using Pearson’s correlation coefficient.

# Results

## Variables distribution

Hip adduction-abduction generated the highest forces among the MVIC (ADD:  kg, ABD:  kg), followed by hip flexion-extension (FLEX:  kg, EXT:  kg) and hip internal-external rotations (IR:  kg, ER:  kg) (Figure 1, left panel). With %, EXT reached the highest left-right imbalance (Figure 1, right panel) while all other tests did not exceed 10% (FLEX: %, IR: %, ER: %, ABD: % and ADD: %).

[insert figure 1]

Based on the FINA guiding scale for height [5](https://paperpile.com/c/PPEkUn/4BnS), athletes from our sample scored () and () in BB-H and EB-H, respectively (Figure 2). The average force measured in EB-F was  kg ().

[insert figure 2]

## Model evaluation

Our predictive model averaged a MAPE of % on the test set. The largest errors (Figure 3) were found in EB-F (MAE:  kg, MAPE: %) although the predictions remained similar to the measured performances ( kg, , posterior distribution in appendix 3). The MAE error in BB-H (MAE: , MAPE: %) and EB-H (MAE: , MAPE: %) predictions was smaller than the resolution of the FINA guiding scale for height (0.5 point). The predicted performance was similar to the measured performance for both BB-H () and EB-H ().

[insert figure 3]

## Model interpretation

A different set of feature importance was reported for each sport-specific test (Figure 4) which suggests that these tests would require different physical capacities. Indeed, the correlation coefficients indicate a weak relationship between the three tests: ranging from 0.01 (EB-H and EB-F), to 0.16 (BB-H and EB-F) and 0.30 (BB-H and EB-F). While BB-H requires to be tall and have strong internal rotation according to the model (Figure 4, left panel), the most important variables in EB-H are a strong external and internal rotations with a low imbalance in abduction and internal rotation (Figure 4, middle panel). Heavier athletes with strong internal rotation, extension and abduction seem to perform better in EB-F (Figure 4, right panel).

[insert figure 4]

# Discussion

Research at the intersection of sports sciences and machine learning offers great promise to advance training decision-making and human movement research. In this study, we used a machine learning algorithm to model the relationship between a series of six hip MVIC and the performance of key skills in artistic swimming. In accordance with our hypothesis, our results show that hip joint isometric function can be used to predict eggbeater kick performance in elite artistic swimmers. The purpose of such a model is twofold. First, it can be used to predict future performance and is therefore useful in a selection setting. Second, the interpretation of this model can help to build personalized and potentially efficient conditioning programs.

## Variables distribution

The model is fed with body mass, height and six hip MVIC recorded with a GroinBar, an easy-to-use and reliable hip strength assessment system [17](https://paperpile.com/c/PPEkUn/LZkq). Swimmers generated the highest forces during hip adduction-abduction, followed by hip flexion-extension and hip internal-external rotations. These forces levels are consistent with those reported by Cichanowski et al. [21](https://paperpile.com/c/PPEkUn/8pk4/?noauthor=1) in female collegiate athletes (once normalized by body mass, see appendix 2). The extension test reached the highest force variability, as well as the highest left-right imbalance while all other tests are more balanced (<10%). This inter-participant and inter-leg variability suggest that hip extension is a difficult test to perform and replicate, as reported in Scott et al. [22](https://paperpile.com/c/PPEkUn/Ps0c/?noauthor=1). As the maximum height was calculated from the head in Sanders [3](https://paperpile.com/c/PPEkUn/r7YR/?noauthor=1), from the hand in Stirn et al. [23](https://paperpile.com/c/PPEkUn/h5Vf/?noauthor=1) and using the FINA guiding scale for height [5](https://paperpile.com/c/PPEkUn/4BnS) in the present study, it is difficult to compare performance with previous studies. Similarly, the upward force generated during eggbeater cannot be easily compared to the literature since Oliveira et al. [24,25](https://paperpile.com/c/PPEkUn/zRzi+cSGQ/?noauthor=1,1) and Sanders [3](https://paperpile.com/c/PPEkUn/r7YR/?noauthor=1) estimated the propulsive force using a biomechanical inverse dynamics model combined with hydrodynamic equations, while we measured a resultant force. Nevertheless, the swimmers were, on average, categorized as “very good” and “good” in the eggbeater kick performances, with some considered as “near perfect” according to the FINA guiding scale for height [5](https://paperpile.com/c/PPEkUn/4BnS).

## Model evaluation

A common study design in machine learning is to split the sample into a training set to train the model and an independent test set to evaluate its performance on unseen data. None of the previous models of sustained and dynamic height [3,4,9](https://paperpile.com/c/PPEkUn/r7YR+ebYG+kFap) used such design, which remains unusual in sports sciences due to small sample size [26](https://paperpile.com/c/PPEkUn/zp4m). In this study, we used a train-test split to make sure that our evaluation is representative of the generalization error of the model. The predictions are similar to the real performances in all three tests, with an average relative error of 4%, 5% and 9% in the predicted BB-H, EB-H and EB-F, respectively. As previously described, EB-F is the test with the largest variability which could explain the higher error on this test. The absolute error in EB-H and BB-H prediction is smaller than the resolution of the FINA guiding scale for height (0.5). Once validated, the model may be used to predict future performances and to better understand the interactions between variables and predictions.

## Model interpretation

A secondary objective was to illustrate how the model can be used to build personalized and potentially efficient conditioning programs. The rich set of interpretation methods implemented in the shapley additive explanations [20](https://paperpile.com/c/PPEkUn/NceT) improves our understanding of the model and give practical insights. Different sets of feature importance were reported for each of the three sport-specific tests, which suggest that these tests would require different physical capacities. The BB-H model predicts high performance for tall athletes with high internal rotation strength, while the EB-H model favors strong external rotation, internal rotation and low imbalance in abduction and internal rotation. Heavier athletes with strong internal rotation, extension and abduction seem to perform better in the EB-F test. Anthropometry contributes only in BB-H (taller athletes getting better performances) and EB-F (heavier athletes getting better performances) models. Being taller can be beneficial during a BB-H test because longer segments allow the contact surface to be increased [27,28](https://paperpile.com/c/PPEkUn/9sjb+Wf4W) while being heavier is often associated with larger muscle mass that can generate more power [29](https://paperpile.com/c/PPEkUn/TCGU). Despite these differences, three similarities arise among the three sport-specific tests models. First, having a high internal rotation strength appears to be important for all tests, which is reasonable as it is an inherent technical component of the eggbeater kick [25](https://paperpile.com/c/PPEkUn/cSGQ). Second, an increase in left-right imbalance-based variables is often associated with decreased performance, which is in line with the literature [9](https://paperpile.com/c/PPEkUn/ebYG). Third, an increase in abduction hip strength is not necessarily linked to better performance. From our sample, 77% of the participants have a higher maximal strength in adduction than abduction, and additional abduction strength could increase body mass. In fact, the coefficient of correlation between abduction and mass is about 0.70, which is the highest among all input variables. Rather than maximum strength, endurance in abduction could be a required quality to sustain high eggbeater kick performance and we could therefore include such tests in the hip function assessment [25](https://paperpile.com/c/PPEkUn/cSGQ).

Note that these interpretations just explain how the model works. Since the model is trained from observational data, it is not necessarily a causal model, and just because changing a factor makes the model’s prediction of performance go up, it does not always mean it will raise the actual performances—even if the generalization error is acceptable. In addition to the interpretation previously showcased, we provide two case studies to illustrate how our model can be used to help design conditioning programs in appendix 4. We hope that our results provide sports scientists and coaches with new opportunities upon which to build modern training programs that enhance the athlete’s performance.

## Methodological considerations

Our study has some limitations. First, we only included high-level artistic swimmers. While a large variability was found in propulsive eggbeater kick force, the variability in sustained and dynamic eggbeater kick was only 0.6 points. Since the FINA guiding scale for height resolution is 0.5 points, the low variability of these target variables may lead to a model with poor predictive capability for sports applications. Inclusion of low-level swimmers may, however, reduce the predictive accuracy since this population is more heterogeneous in terms of technique and flexibility, which are essential in such complex skills [3,4](https://paperpile.com/c/PPEkUn/r7YR+kFap). Inclusion of data from water polo players or lifesavers could also extend the reach of the model by increasing the variability in both predictor and target variables. It may, however, adversely affects its accuracy since these populations may differ in other physical qualities not considered in the model, such as flexibility. Additional tests on complementary skills—such as technique and flexibility—could help to include swimmers with various levels and coming from various sports. Second, we only assessed the isometric force while eggbeater kick is a dynamic skill. Even if isokinetic force assessment, as in Yamamura et al. [30](https://paperpile.com/c/PPEkUn/F5dm/?noauthor=1), would provide a more accurate prediction, isometric tests are more suited to a training setting with regard to cost, ease of use and implementation. Third, the performance in sustained eggbeater kick was defined as an average value. The standard deviation in force and height might also be a key performance indicator associated with hip function. Oliveira et al. [6](https://paperpile.com/c/PPEkUn/bGD9/?noauthor=1) showed force variation of about 40% throughout the eggbeater cycle. This variation may be related to bilateral asymmetries in force and technique between dominant and non-dominant limbs in line with the dynamic dominance theory [9](https://paperpile.com/c/PPEkUn/ebYG). Finally, sport performance must be considered as a complex multi-causal phenomenon. The results presented in this study only consider hip joint function and anthropometry, omitting the many other dimensions that interact in the production of sport performance. Moving forward, we would like to investigate the relevance of a predictive model to support decision-making to improve sports performance.

# Practical applications

Our results showed that hip joint isometric strength can be used to predict eggbeater kick performance in elite artistic swimmers within the resolution of the sport notation scale. We also highlighted some of the important predictors of key technical skills in artistic swimming. Those new findings support the use of predictive modelling to select athletes and to design personalized conditioning goals. It also provides coaches, athletes and other researchers in sports physiology and sports performance with hip strength normative data in elite artistic swimmers.

# Conclusion

Our model may accurately predict future performances as the generalization error is similar to the resolution of the FINA guiding scale for height. In addition, we used a set of interpretation and simulation methods to show that our model could provide practical guidelines to build effective and personalized conditioning programs. The model can be easily implemented in elite training structures as the required tests can easily be performed on a weekly basis, without the need for a physiotherapist or a scientist.

# Acknowledgments

This study was funded by Institut National du Sport du Québec. We wish to thank Artistic Swimming Canada for the data collection during their training camp, and the coaches of the clubs of Dollard Synchro, Montréal Synchro and Gatineau Synchro for providing us access to the athletes. We also thank Patrick Marion and Nicolas Bourret who helped with the data collection. The results of the current study do not constitute endorsement of the product by the authors or the journal.

# References

1. [Alentejano T, Marshall D, Bell G. A time-motion analysis of elite solo synchronized swimming. *Int J Sports Physiol Perform*. 2008;3(1):31-40.](http://paperpile.com/b/PPEkUn/wNlO) <https://www.ncbi.nlm.nih.gov/pubmed/19193952>[.](http://paperpile.com/b/PPEkUn/wNlO)

2. [Homma M. A study of the components of the team free routines in synchronized swimming at the Atlanta Olympic Games. *Suiei Suichu Undo Kagaku*. 1997;13:9-20.](http://paperpile.com/b/PPEkUn/F0rd)

3. [Sanders RH. Analysis of the Eggbeater Kick Used to Maintain Height in Water Polo. *J Appl Biomech*. 1999;15(3):284-291. doi:](http://paperpile.com/b/PPEkUn/r7YR)[10.1123/jab.15.3.284](http://dx.doi.org/10.1123/jab.15.3.284)

4. [Homma M, Homma M. Coaching points for the technique of the eggbeater kick in synchronized swimming based on three-dimensional motion analysis. *Sports Biomech*. 2005;4(1):73-87. doi:](http://paperpile.com/b/PPEkUn/kFap)[10.1080/14763140508522853](http://dx.doi.org/10.1080/14763140508522853)

5. [FINA. FINA Artistic swimming manual for judges, coaches and referees. *Fédération Internationale de Natation*. 2018.](http://paperpile.com/b/PPEkUn/4BnS)

6. [Oliveira N, Fernandes RJ, Sarmento M, et al. Muscle Activity During the Typical Water Polo Eggbeater Kick. *International Journal of Aquatic Research and Education*. 2010;4(2):7. doi:](http://paperpile.com/b/PPEkUn/bGD9)[10.25035/ijare.04.02.07](http://dx.doi.org/10.25035/ijare.04.02.07)

7. [Viana E, Bentley DJ, Logan-Sprenger HM. A Physiological Overview of the Demands, Characteristics, and Adaptations of Highly Trained Artistic Swimmers: a Literature Review. *Sports Med Open*. 2019;5(1):16. doi:](http://paperpile.com/b/PPEkUn/4xKy)[10.1186/s40798-019-0190-3](http://dx.doi.org/10.1186/s40798-019-0190-3)

8. [Sanders RH. A Model of Kinematic Variables Determining Height Achieved in Water Polo Boosts. *J Appl Biomech*. 1999;15(3):270-283. doi:](http://paperpile.com/b/PPEkUn/PScr)[10.1123/jab.15.3.270](http://dx.doi.org/10.1123/jab.15.3.270)

9. [Oliveira N, Sanders RH. Kinematic and kinetic evidence for functional lateralization in a symmetrical motor task: the water polo eggbeater kick. *Exp Brain Res*. 2015;233(3):947-957. doi:](http://paperpile.com/b/PPEkUn/ebYG)[10.1007/s00221-014-4166-8](http://dx.doi.org/10.1007/s00221-014-4166-8)

10. [Zinner C, Sperlich B, Krueger M, Focke T, Reed J, Mester J. Strength, Endurance, Throwing Velocity and in-Water Jump Performance of Elite German Water Polo Players. *J Hum Kinet*. 2015;45:149-156. doi:](http://paperpile.com/b/PPEkUn/hOMc)[10.1515/hukin-2015-0015](http://dx.doi.org/10.1515/hukin-2015-0015)

11. [Abut F, Akay MF. Machine learning and statistical methods for the prediction of maximal oxygen uptake: recent advances. *Med Devices* . 2015;8:369-379. doi:](http://paperpile.com/b/PPEkUn/L65F)[10.2147/MDER.S57281](http://dx.doi.org/10.2147/MDER.S57281)

12. [Ofoghi B, Zeleznikow J, Dwyer D, Macmahon C. Modelling and analysing track cycling Omnium performances using statistical and machine learning techniques. *J Sports Sci*. 2013;31(9):954-962. doi:](http://paperpile.com/b/PPEkUn/ODSy)[10.1080/02640414.2012.757344](http://dx.doi.org/10.1080/02640414.2012.757344)

13. [Maier T, Meister D, Trösch S, Wehrlin JP. Predicting biathlon shooting performance using machine learning. *J Sports Sci*. 2018;36(20):2333-2339. doi:](http://paperpile.com/b/PPEkUn/KoKn)[10.1080/02640414.2018.1455261](http://dx.doi.org/10.1080/02640414.2018.1455261)

14. [Pantazopoulos A, Maragoudakis M. Sports & Nutrition Data Science Using Gradient Boosting Machines. In: *Proceedings of the 10th Hellenic Conference on Artificial Intelligence*. SETN ’18. New York, NY, USA: ACM; 2018:56:1-56:7. doi:](http://paperpile.com/b/PPEkUn/FVZ1)[10.1145/3200947.3201060](http://dx.doi.org/10.1145/3200947.3201060)

15. [Fister I, Ljubič K, Suganthan PN, Perc M, Fister I. Computational intelligence in sports: Challenges and opportunities within a new research domain. *Appl Math Comput*. 2015;262:178-186. doi:](http://paperpile.com/b/PPEkUn/whKM)[10.1016/j.amc.2015.04.004](http://dx.doi.org/10.1016/j.amc.2015.04.004)

16. [Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Front Neurorobot*. 2013;7:21. doi:](http://paperpile.com/b/PPEkUn/wlY9)[10.3389/fnbot.2013.00021](http://dx.doi.org/10.3389/fnbot.2013.00021)

17. [Desmyttere G, Gaudet S, Begon M. Test-retest reliability of a hip strength assessment system in varsity soccer players. *Phys Ther Sport*. 2019;37:138-143. doi:](http://paperpile.com/b/PPEkUn/LZkq)[10.1016/j.ptsp.2019.03.013](http://dx.doi.org/10.1016/j.ptsp.2019.03.013)

18. [Dorogush AV, Ershov V, Gulin A. CatBoost: gradient boosting with categorical features support. *arXiv [csLG]*. October 2018.](http://paperpile.com/b/PPEkUn/ZQOM) <http://arxiv.org/abs/1810.11363>[.](http://paperpile.com/b/PPEkUn/ZQOM)

19. [Kruschke JK. Bayesian estimation supersedes the t test. *J Exp Psychol Gen*. 2013;142(2):573-603. doi:](http://paperpile.com/b/PPEkUn/notd)[10.1037/a0029146](http://dx.doi.org/10.1037/a0029146)

20. [Lundberg SM, Lee S-I. A Unified Approach to Interpreting Model Predictions. In: Guyon I, Luxburg UV, Bengio S, et al., eds. *Advances in Neural Information Processing Systems 30*. Curran Associates, Inc.; 2017:4765-4774.](http://paperpile.com/b/PPEkUn/NceT) <http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf>[.](http://paperpile.com/b/PPEkUn/NceT)

21. [Cichanowski HR, Schmitt JS, Johnson RJ, Niemuth PE. Hip strength in collegiate female athletes with patellofemoral pain. *Med Sci Sports Exerc*. 2007;39(8):1227-1232. doi:](http://paperpile.com/b/PPEkUn/8pk4)[10.1249/mss.0b013e3180601109](http://dx.doi.org/10.1249/mss.0b013e3180601109)

22. [Scott DA, Bond EQ, Sisto SA, Nadler SF. The intra- and interrater reliability of hip muscle strength assessments using a handheld versus a portable dynamometer anchoring station. *Arch Phys Med Rehabil*. 2004;85(4):598-603.](http://paperpile.com/b/PPEkUn/Ps0c) <https://www.ncbi.nlm.nih.gov/pubmed/15083436>[.](http://paperpile.com/b/PPEkUn/Ps0c)

23. [Stirn I, Strmecki J, Strojnik V. The examination of different tests for the evaluation of the efficiency of the eggbeater kicks. *J Hum Kinet*. 2014;41:215-226. doi:](http://paperpile.com/b/PPEkUn/h5Vf)[10.2478/hukin-2014-0049](http://dx.doi.org/10.2478/hukin-2014-0049)

24. [Oliveira N, Chiu C-Y, Sanders RH. Kinematic Patterns Associated with the Vertical Force Produced during the Eggbeater Kick. *J Sports Sci*. 2015;33(16):1675-1681. doi:](http://paperpile.com/b/PPEkUn/zRzi)[10.1080/02640414.2014.1003590](http://dx.doi.org/10.1080/02640414.2014.1003590)

25. [Oliveira N, Saunders DH, Sanders RH. The Effect of Fatigue-Induced Changes in Eggbeater-Kick Kinematics on Performance and Risk of Injury. *Int J Sports Physiol Perform*. 2016;11(1):141-145. doi:](http://paperpile.com/b/PPEkUn/cSGQ)[10.1123/ijspp.2015-0057](http://dx.doi.org/10.1123/ijspp.2015-0057)

26. [Bernards JR, Sato K, Haff GG, Bazyler CD. Current Research and Statistical Practices in Sport Science and a Need for Change. *Sports (Basel)*. 2017;5(4). doi:](http://paperpile.com/b/PPEkUn/zp4m)[10.3390/sports5040087](http://dx.doi.org/10.3390/sports5040087)

27. [Carter JEL, Ackland TR. Kinanthropometry in Aquatic Sport: A Study of World Class Athletes. 1994.](http://paperpile.com/b/PPEkUn/9sjb)

28. [Grimston SK, Hay JG. Relationships among anthropometric and stroking characteristics of college swimmers. *Med Sci Sports Exerc*. 1986;18(1):60-68.](http://paperpile.com/b/PPEkUn/Wf4W) <https://www.ncbi.nlm.nih.gov/pubmed/3959866>[.](http://paperpile.com/b/PPEkUn/Wf4W)

29. [Harbo T, Brincks J, Andersen H. Maximal isokinetic and isometric muscle strength of major muscle groups related to age, body mass, height, and sex in 178 healthy subjects. *Eur J Appl Physiol*. 2012;112(1):267-275. doi:](http://paperpile.com/b/PPEkUn/TCGU)[10.1007/s00421-011-1975-3](http://dx.doi.org/10.1007/s00421-011-1975-3)

30. [Yamamura C, Zushi S, Takata K, Ishiko T, Matsui N, Kitagawa K. Physiological characteristics of well-trained synchronized swimmers in relation to performance scores. *Int J Sports Med*. 1999;20(4):246-251. doi:](http://paperpile.com/b/PPEkUn/F5dm)[10.1055/s-2007-971125](http://dx.doi.org/10.1055/s-2007-971125)

# Figure captions

Figure 1. Tukey box plot showing force (left panel) and imbalance (right panel) evaluated on the MVIC with median (vertical lines), first-third interquartile range (bars), minimum-maximum range (horizontal lines) and outliers (values that are > 3 standard deviations) (circles).

Figure 2. Tukey box plot showing the three sport-specific tests performances with median (vertical lines), first-third interquartile range (bars), minimum-maximum range (horizontal lines) and outliers (values that are > 3 standard deviations) (circles).

Figure 3. Empirical cumulative distribution function (ECDF) of the MAE (left panel) and MAPE (right panel) measured on the test set (n=19) for BB-H (blue), EB-H (orange) and EB-F (red). The ECDF evaluated at *x* is defined as the fraction of data points that are *≤ x*. Mean value are also displayed (vertical lines).

Figure 4. SHAP summary plot of the three gradient boosting models (left: BB-H, middle: EB-H, right: EB-F). The higher the SHAP value (x-axis) of a feature (y-axis), the higher the log of the target output. Only the five most important variables are displayed and ranked from most important (top) to least (bottom). Every participant is run through each model and a dot is created for each feature attribution value. Dots are colored by the feature value (red when the variable is high, blue when it is low) and pile up vertically to show density. For example, BB-H predicted performance increases if IR increases.