JOURNAL OF SPORTS SCIENCES, 2018
https://doi.org/10.1080/02640414.2018.1455261

Predicting biathlon shooting performance using machine learning

Thomas Maier[a], Daniel Meister[b], Severin Trösch[a] and Jon Peter Wehrlin[a]

[a]Swiss Federal Institute of Sport, Section for Elite Sport, Magglingen, Switzerland; [b]Datahouse AG, Zurich, Switzerland

ABSTRACT

Shooting in biathlon competitions substantially influences final rankings, but the predictability of hits and misses is unknown. The aims of the current study were A) to explore factors influencing biathlon shooting performance and B) to predict future hits and misses. We explored data from 118,300 shots from 4 seasons and trained various machine learning models before predicting 34,340 future shots (in the subsequent season). A) Lower hit rates were discovered in the sprint and pursuit disciplines compared to individual and mass start ($P < 0.01$, $h = 0.14$), in standing compared to prone shooting ($P < 0.01$, $h = 0.15$) and in the 1st prone and 5th standing shot ($P < 0.01$, $h = 0.08$ and $P < 0.05$, $h = 0.05$). B) A tree-based boosting model predicted future shots with an area under the ROC curve of 0.62, 95% CI [0.60, 0.63], slightly outperforming a simple logistic regression model and an artificial neural network ($P < 0.01$). The dominant predictor was an athlete's preceding mode-specific hit rate, but a high degree of randomness persisted, which complex models could not substantially reduce. Athletes should focus on overall mode-specific hit rates which epitomise shooting skill, while other influences seem minor.

ARTICLE HISTORY
Accepted 14 March 2018

KEYWORDS
Sport; competition; modelling

Introduction

In the Olympic sport of biathlon, athletes complete multiple laps on a cross-country ski track with intermittent rifle shootings. Disciplines vary in the total distance to be covered, the number of shootings and the start modality. The total distance in non-team disciplines (individual, sprint, pursuit, mass start) varies from 7.5 km (sprint woman) to 20 km (individual men) corresponding to race durations of around 20 – 50 min. The sprint discipline consists of 3 skiing laps with 2 shootings in between, whereas the others consist of 5 skiing laps with 4 shootings. Half of the shootings per race require athletes to be in the prone position and the other half require them to be standing. Each shooting consists of 5 shots that count either as hits or misses. The latter result in a time penalty (individual) or penalty lap the athlete must cover. Each miss results in a time loss corresponding to 1 – 2% of total race duration, substantially influencing final rankings (Nitzsche, 1998).

Different academic fields have investigated the biathlon shooting task. From a biomechanical viewpoint, low vertical rifle motion (for prone shooting) and low body sway (for standing shooting) have been advocated (Sattlecker, Buchecker, Gressenbauer, Müller, & Lindinger, 2017; Sattlecker, Buchecker, Müller, & Lindinger, 2014). Physiologists have reported on the impact of elevated heart rate and fatigue on shooting performance (Groslambert, Candau, Hoffman, Bardy, & Rouillon, 1999; Hoffman, Gilson, Westenburg, & Spencer, 1992; Hoffman & Street, 1992). Analyses of in-competition shooting performance have shown higher hit rates in top-10 athletes compared to those ranked lower, lower hit rates when athletes were standing compared to when they were shooting prone and lower hit rates in the sprint compared to the individual discipline (Luchsinger, Kocbach, Ettema, & Sandbakk, 2017; Nitzsche, 1998; Skattebo & Losnegard, 2017). Additionally, Luchsinger et al. (2017) have identified differences between the 5 shots per series. However, Skattebo and Losnegard (2017) reported a high within-athlete standard deviation of 12% for the hit rate in 10 shots.

In situations with inherent randomness like biathlon shooting, the question of predictability of the outcome arises. Machine learning summarises numerical techniques to extract patterns from data and includes models to generate predictions on new data. Such models are used in diverse fields e.g., to predict the weather, consumer behaviour, or the presence of diseases. In sports sciences, various sub-fields have applied machine learning techniques. For instance, Begg and Kamruzzaman (2005) used machine learning to classify movement patterns from accelerometer sensor data and Novatchkov and Baca (2013) demonstrated the potential for automated feedback during weight training. Only recently were such models applied in training or competition settings in elite sports. Studies in track cycling and triathlon (Ofoghi, Zeleznikow, Dwyer, & Macmahon, 2013; Ofoghi, Zeleznikow, Macmahon, Rehula, & Dwyer, 2016) or swimming (Mężyk & Unold, 2011) investigated patterns in competition outcomes or daily training respectively. Baca and Kornfeind (2012) already applied machine learning in biathlon by analysing the stability of the aiming process with an artificial neural network trained on video data.

In biathlon, numerous factors possibly influencing shooting performance and the predictability of individual shots in competitions have not yet been investigated. Machine learning models could support athletes and coaches in identifying

CONTACT Thomas Maier thomas.maier@baspo.admin.ch Swiss Federal Institute of Sport, Section for Elite Sport, Hauptstrasse 247, 2532 Magglingen, Switzerland

The complete analysis code and the raw data for this article can be accessed at Figshare (http://tandf.figshare.com/).

shots with an increased risk of missing the target, which could help to develop appropriate training routines.

The aims of this study were therefore A) to explore factors influencing biathlon shooting performance and B) to predict future hits and misses using machine learning models.

Method

Design

Detailed biathlon shooting data was collected from 5 consecutive competition seasons (2012/13 to 2016/17). Using feature engineering techniques, the data set was modified to maximally expose relevant information. The data set was split chronologically into training data (2012/13 to 2015/16) and test data (2016/17). With the training data, A) we investigated factors influencing shooting performance with exploratory data analysis and B) trained different machine learning models for predicting future hits and misses. These models were then used to predict hits and misses in the test data to evaluate their performance. The study was accepted by the institutional review board of the Swiss Federal Institute of Sport Magglingen.

Data source and processing

A major supplier of biathlon shooting target systems (HoRa Systemtechnik GmbH, Bad Endorf, Germany), used in about half of all world cup competitions, openly publishes detailed shooting results (http://www.hora2000.de/de/downloads/). Data from non-team world cups, world championships and Olympic Games from the 2012/13 to the 2016/17 season was downloaded, corresponding to the highest international level of biathlon competitions. Table 1 lists included competition locations and disciplines.

First, data was converted from PDF into Excel format with Adobe Acrobat X Pro, then imported into the statistical software package R (R Core Team, 2016) and finally processed to a tidy data structure (Wickham, 2014). The resulting data set described 152,640 individual shots (training data: 118,300 shots, test data: 34,340 shots).

Feature engineering

The original data set already contained presumable relevant information to predict individual hits and misses. Differences between locations, disciplines or athletes are plausible. Variables describing the shooting like the shooting mode, the shooting lane (is it easier to shoot at lanes in the front?) or the shot number (is the last shot easier than the first one?) likely contain further information. Still missing was information about preceding sequences or trends in hit rate, aiming times or run times of individual athletes. Therefore, we modified the data set as follows.

First, the data set was sorted chronologically by adding competition dates and start times. Then, we generated additional variables, which held data relating to preceding shots by the same athlete (run time of preceding laps, preceding shooting times, preceding targets and preceding hits or misses). Finally, further variables were calculated with moving averages of hit rates of preceding shots and with cumulative sums of preceding shots (per season, location or discipline) for each athlete. Care was taken to include only information that was available before the athlete fired the respective shot.

In summary, 29 variables described each shot with information about the competition, the athlete, the shooting, preceding run and shooting times, preceding hits, preceding numbers of shots and finally, whether the shot resulted in a hit or miss (Table 2).

Table 1. Included competitions in this study (men and woman).

Location (country)	Season 2012/13	2013/14	2014/15	2015/16	2016/17
Annecy (FRA)	-	SP	-	-	-
Antholz (ITA)	SP	SP	SP	SP	IM
Hochfilzen (AUT)	SP	SP	SP	SP	SPIM
Nove Mesto (CZE)	-	-	SP	-	SPM
Oberhof (DEU)	SP	SPM	SM	-	SPM
Pokljuka (SVN)	SPM	SPM	SPM	SPM	SP
Presque Isle (USA)	-	-	-	SP	-
Ruhpolding (DEU)	SM	IP	SM	SPMI	SP
Sochi (RUS)	IS	SPIM	-	-	-

Notes. Disciplines: S = sprint, P = pursuit, M = mass start, I = individual.

A) Exploratory data analysis

Confidence limits for mean hit rates depending on categorical variables (e.g., athlete's gender, shooting mode) were calculated. For continuous variables (e.g., athlete's start number, athlete's preceding hit rates), Pearson correlations between the respective variable and hit rate (overall and within athlete) or confidence intervals for the respective variable mean before subsequent hits and misses were calculated. Differences were tested with pairwise Chi-squared tests with Bonferroni-Holms corrections or non-parametric Mann-Whitney *U* tests ($\alpha = 0.05$). Effect sizes for differences in hit rates are presented with Cohen's *h*.

If not otherwise stated, values are presented as mean, 95% confidence interval (CI) [lower limit, upper limit]. Figures were created with the ggplot2 package (Wickham, 2016).

B) Predicting future hits and misses

Model training was performed with the caret package (Kuhn, 2017). Categorical variables were encoded with dummy variables (one-hot encoding, only the 10 most frequent nations were included, athlete names were excluded), resulting in 48 model

Table 2. Variables used to describe each shot after feature engineering.

Group	Variables	*N*	Type
Competition	Location, discipline	2	Cat
Athlete	Name, gender, nation, start number	4	Cat/Con
Shooting	Lap, mode, lane, shot number	4	Cat/Con
Preceding run times	Run time change *	1	Con
Preceding shots	Aiming times (3), target (1), results (3) **	7	Cat/Con
Preceding hit rates	Overall (10, 50, 200), mode-specific (10, 50, 200), mode and shot number specific (200) ***	7	Con
Cumulative shots	This season, this location, this discipline	3	Con
Target variable	Result of shot (hit/miss)	1	Cat

Notes. Cat = categorical, Con = continuous. * time of current lap divided by time of preceding lap. ** (*n*) = for 1 to *n* preceding shots. *** (*n*) = mean hit rate over *n* preceding shots.

input variables. Constants were imputed for unavailable values that arose by definition in the variables describing preceding shots or run times (we imputed 0.82 for unavailable preceding hit rates, 0 for unavailable preceding aiming times, 1 for run time changes in the first lap and random values for unavailable preceding targets).

We tuned model parameters and data transformations with cross-validation using a rolling forecasting origin (Hyndman & Athanasopoulos, 2013): models were trained on approximately one season (29,575 shots) and predictions for the subsequent half season (14,787 shots) were evaluated. The time windows were then repeatedly shifted a half season, resulting in 6 data folds to evaluate model performance. We used these chronological splits to prevent information leaks through the variables spanning multiple shots (e.g., preceding past hit rates).

The area under the receiver operating characteristic (ROC) curve (AUROC) was used as performance metric (Bradley, 1997). AUROC represents a summary metric for the predictive power of a two-class classification model, independent of class imbalances. For the final models, we selected the model parameters and data transformations that resulted in the highest mean AUROC during cross-validation.

Various machine learning models spanning different types (parametric, non-parametric) and different complexities were used: a logistic regression model (LogReg), a tree-based model with boosting (XGB) and an artificial neural network (NNet). LogReg and NNet have been widely used in research and industry and XGB has often excelled in recent machine learning competitions (www.kaggle.com). LogReg allowed estimation of the baseline performance of a simple parametric model. In contrast, XGB and NNet are both capable of modelling complex patterns but with different approaches: XGB uses a combination of sequentially built decision trees while NNet trains weights of an artificial neural network with error back-propagation.

LogReg was implemented using the glm function with only a single input variable and no specific data pre-processing.

XGB was implemented using the xgboost package (Chen & He, 2017). No specific data pre-processing was applied, because decision trees are robust against skewed data. Model complexity was initially tuned with a high learning rate before decreasing the learning rate and increasing the number of boosting rounds.

NNet was implemented using the nnet package (Venables & Ripley, 2016). A fully connected feed forward network with a single hidden layer and a sigmoid activation function was deployed. Different specific data pre-processing steps and the number of hidden neurons as well as the weight decay during training were tuned.

With the 3 final models we predicted the probabilities of hits and misses in the test data. No further changes were applied to the models in order to get a true representation of their predictive performance on previously unseen data. The following results are reported for the 3 models: their distributions of the predicted probabilities of a hit, their ROC curves and their AUROC. Confidence limits were calculated by bootstrap sampling.

Results

A) Exploratory data analysis

Individual hit rates varied between athletes (with > 200 shots) from 0.91, 95% CI [0.89, 0.93] to 0.69, [0.63, 0.75] (best and worst male, h = 0.57) and from 0.92, [0.88, 0.96] to 0.66, [0.61, 0.71] (best and worst female, h = 0.69). Hit rates between nations (with > 1000 shots) varied from 0.86, [0.85, 0.87] to 0.75, [0.72, 0.77] (best and worst nation, h = 0.30).

Hit rate did not differ between male and female athletes (P = 0.09), but differed between disciplines (Figure 1, h = 0.14).

Hit rate was 0.06, 95% CI [0.05, 0.07] lower in the standing mode compared to the prone mode (P < 0.01, h = 0.15). In the standing mode, the 5th shot had a lower hit rate than the other shots (P < 0.05, h = 0.05), whereas in the prone mode the 1st shot had the lowest hit rate (P < 0.01, h = 0.08, Figure 2). These effects did not depend on lap number.

Hit rate did not differ between laps when controlled for mode and discipline (P > 0.40). Between athletes, hit rate was negatively correlated with start number (r = − 0.74, P < 0.01), but not within athletes (P > 0.13). Between athletes, hit rate was negatively correlated with shooting lane (r = − 0.63, P < 0.01), but within athletes the correlation was slightly positive (r = 0.10, P < 0.01, only analysed in pursuit, due to confounding competition rules in the other disciplines).

Table 3 shows further differences in selected variables preceding hits or misses.

B) Predicting future hits and misses

Best cross-validation performances in the training data resulted with the final model configurations shown in Table 4.

On the test data, model output probabilities for a hit were mainly between 0.7 and 0.9. The complex models predicted a wider range than LogReg (Figure 3). Figure 4 illustrates the ROC curves for the 3 models. XGB achieved

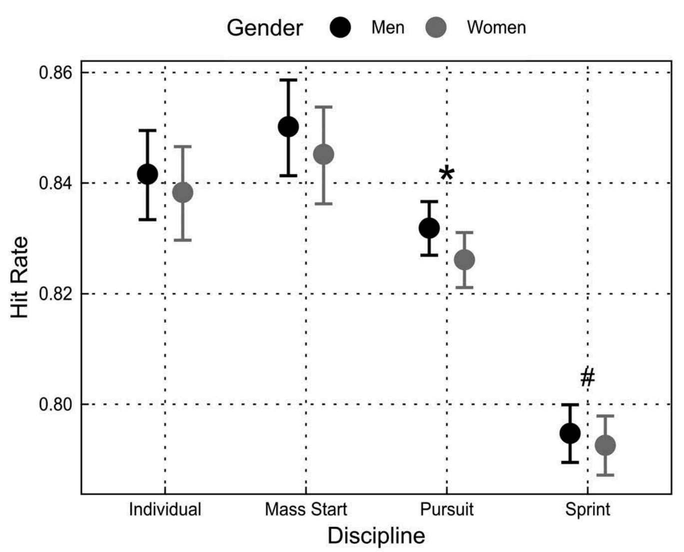

Figure 1. Hit rate differed between disciplines. * lower than individual and mass start (P < 0.01), # lower than others (P < 0.01). Data is illustrated as mean with 95% confidence interval.

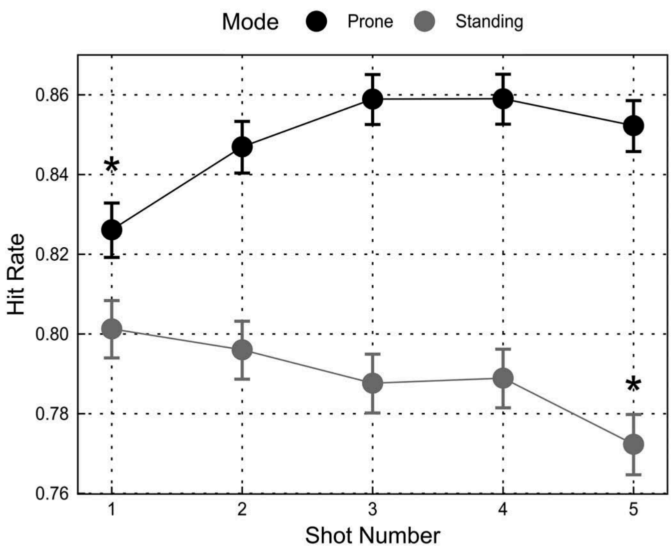

Figure 2. Hit rate differed between shooting mode and shot number. * lower hit rate than other shots in the same mode. Data is illustrated as mean with 95% confidence interval.

an AUROC of 0.62, 95% CI [0.60, 0.63], which was higher ($P < 0.01$) than NNet 0.61, [0.60, 0.62] and LogReg 0.60, [0.59, 0.61].

Discussion

A) Exploratory data analysis

The analysis in the current study confirms already identified patterns in biathlon shooting results and reveals additional insights.

The largest differences between hit rates were evident between individual athletes (0.66 – 0.92, $h > 0.57$), dominating all other influences. The observed differences between disciplines (0.79 – 0.85, $h = 0.14$) confirm the reported lower hit rate in sprint compared to individual competitions (Nitzsche, 1998). The same author as well as recent studies (Luchsinger et al., 2017; Skattebo & Losnegard, 2017) reported the difference between the prone and standing mode (0.79 and 0.85, $h = 0.15$). However, while the current study illustrates slightly lower hit rates for the 1st shot in the prone position ($h = 0.08$) and the 5th shot in the standing position ($h = 0.05$), Luchsinger et al. (2017) reported increasing hit rates in the standing position (but without indications of statistical confidence). Furthermore, the current study reports correlations that have not previously been investigated, which are between hit rates and start number/shooting lane and patterns between run times, shooting times or individual hit rates and subsequent hits or misses.

It is important to distinguish between patterns caused by the athletes and patterns caused by competition rules. The overall negative correlation between hit rate and start number/shooting lane likely result from the current competition rules, where the best athletes get the lowest start numbers in mass start and pursuit competitions and have to shoot at lanes in the front when arriving early. This could explain why these variables are not correlated within athletes. The higher hit rates in pursuit and mass start races compared to the sprint could be caused by the limitations on competing athletes in these disciplines (only the best 60 and the best 30 athletes can start in the pursuit and in the mass start respectively).

Other reported patterns could be more relevant for athletes and coaches: the large differences between individual athletes emphasise the importance of their overall hit rate and shots identified as having a decreased probability of a hit (1st in prone, 5th in standing) could be targeted by specific training routines. Decreasing postural stability could cause the decreasing hit rate in the standing position, as discussed in traditional rifle shooting (Ball, Best, & Wrigley, 2003; Era, Konttinen, Mehto, Saarela, & Lyytinen, 1996), whereas rifle stability seems to be more important in prone shooting (Sattlecker et al., 2017). Further studies should investigate the causal relation between preceding run times or shooting times and hit rates.

B) Predicting future hits and misses

Training XGB and NNet resulted in rather regularised and simple model configurations (XGB: low maximal tree depth, NNet: a single hidden neuron), indicating that the predictive patterns in the data are rather simple. Specifically, the resulting architecture of NNet corresponds to a simple perceptron, incapable of modelling interactions between variables.

Table 3. Differences in selected variables before a subsequent hit or miss.

	Subsequent shot		
Variable	Hit	Miss	*P*-value
Run time change	1.06, [1.05, 1.06]	1.07, [1.06, 1.08]	< 0.01
Aiming time of shot (s, without 1st shot)	2.97, [2.95, 2.98]	3.16, [3.13, 3.20]	< 0.01
Hit rate over 200 shots	0.83, [0.82, 0.83]	0.80, [0.80, 0.81]	< 0.01
Shots this season	75.5, [75.1, 75.9]	69.8, [69.0, 70.6]	< 0.01

Notes. Values are presented as mean, 95% confidence interval [lower limit, upper limit].

Table 4. Final model configurations chosen after cross-validation on the training data.

Model	Data pre-processing	Model parameters	AUROC
LogReg	No pre-processing	only 1 input variable (preceding mode-specific hit rate over 200 shots)	0.60, [0.59, 0.62]
XGB	No pre-processing	eta = 0.02, nrounds = 300, max_depth = 3, min_child_weight = 10, gamma = 1, colsample_bytree = 0.5, subsample = 0.8	0.62, [0.60, 0.63]
NNet	Range scaled to [0, 1]	layers = 1, number of hidden neurons = 1, weight decay = 0.1, activation function = sigmoid	0.61, [0.59, 0.64]

Notes. LogReg = logistic regression model, XGB = tree-based boosting model, NNet = artificial neural network, AUROC = area under the receiver operating characteristic curve. Values are presented as mean, 95% confidence interval [lower limit, upper limit].

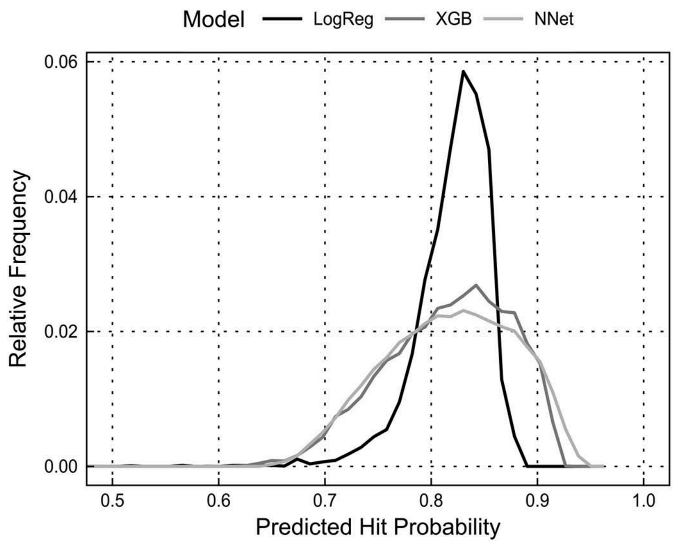

Figure 3. Distribution of model output probability of a hit. LogReg = logistic regression model, XGB = tree-based boosting model, NNet = artificial neural network.

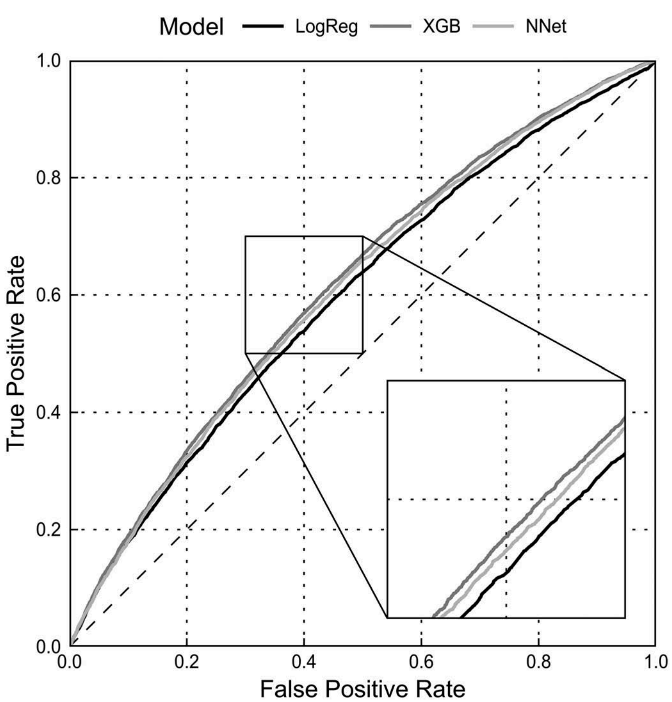

Figure 4. Receiver operating characteristic curves of the models. LogReg = logistic regression model, XGB = tree-based boosting model, NNet = artificial neural network.

Furthermore, LogReg achieved a comparable predictive performance during training using only a single input variable (individual hit rates).

Distributions of model output probabilities of a hit illustrated that almost every shot is more likely to be a hit than a miss. This seems reasonable, as the overall hit rate in biathlon is around 0.82.

The predictive performance of the models on the test data illustrates the difficulty of classifying hits and misses with the available information during competitions. A high degree of variance could not be explained by the models as evidenced by the relatively low AUROC of around 0.61, presumably caused by the high degree of randomness involved. Furthermore, as already seen during model training, LogReg using only the information of athletes' preceding mode-specific hit rates achieved most of the predictive performance of the two complex models. However, XGB's performance was slightly superior to that of LogReg and NNet. An AUROC of 0.62 implies that given a hit and a miss at random, XGB will rank the hit probabilities correctly 62% of the time. Other studies comparing machine learning models for classification tasks have shown that model performance is specific to the data set and its inherent complexity (Bauer & Kohavi, 1999; Bradley, 1997).

In summary, this suggests that most of the predictive information lies in individual mode-specific hit rates, superimposed by a high degree of randomness involved in every shot.

Relevance for biathlon competitions

Our results imply that each shot in biathlon can be seen as a realisation of a Bernoulli trial, where preceding hit rate indicates the "probability of success" and other influences are negligible. While the outcome of one shot involves a high degree of randomness, it is important to consider the total amount of shots an athlete has to fire in non-team competitions (10 or 20). This can be analysed with a binomial distribution (a sum of Bernoulli trials), what suggests a standard deviation of 1.2 hits (in 10 shots) or 1.7 hits (in 20 shots) per athlete and competition (standard deviation of a binomial distribution = $\sqrt{np(1-p)}$, mean hit rate $p = 0.82$, number of shots $n = 10$ or 20). This is perfectly in-line with the reported within-athlete variability of 12% in 10 shots (Skattebo & Losnegard, 2017). Thus, given 2 athletes with the same preceding hit rate ($p = 0.82$), the probability that one of them will score 2 hits more than the other in 20 shots equates to approximately 53%. The randomness involved in biathlon shooting is comparable to free throws in basketball, which are also solely predictable by a players typical hit rate (Gilovich, Vallone, & Tversky, 1985).

Therefore, athletes should focus on improving their overall mode-specific hit rates which epitomise shooting skill and their probabilities of good competition results while the influence of other factors (like the course of events in the competition) seems minor. Nevertheless, as every miss can substantially decrease an athlete's final ranking, predictions from machine learning models could be valuable in competition preparation to analyse the probabilities of a miss for each shot (simulating different scenarios). Furthermore, live predictions of the probability of individual hits during competitions could be attractive for broadcasting purposes.

Future studies could combine competition data with biomechanical measurements, which could increase predictive performance, as already applied in pistol shooting (Hawkins, 2011).

Limitations

In the current study, a vast data set of biathlon shots was analysed. Nevertheless, the data included only about half of the competitions in the analysed time window. Thus, possible effects of competition locations not included in the data are missing in the final models. The discovered patterns and models could also differ at biathlon competitions below the highest international level.

Different research groups have identified biomechanical predictors (Baca & Kornfeind, 2012; Ball et al., 2003; Hawkins, 2011; Sattlecker et al., 2017) or physiological predictors (Hoffman et al., 1992) of shooting performance. However, as the current study only used publicly available information during competitions, no such predictors were available for model training. The preceding individual hit rate presumably combines such factors in a single performance score. Nonetheless, the discussed inherent randomness in individual shots can at least partly be attributed to such explaining factors (e.g., body or rifle sway).

The predictive power of different machine learning models was investigated, but it has been shown that combining the predictions of differing models often results in improved predictions, an approach which was not used in the current study.

Conclusion

Hit rates in biathlon competitions differ between athletes, disciplines, shooting modes and shot numbers (from large to small effects respectively). To predict future shots, a simple machine learning model using only an athlete's preceding mode-specific hit rate showed some predictive power. However, a high degree of randomness involved in every shot persisted, which complex models could not substantially reduce.

Athletes and coaches should focus on improving overall mode-specific hit rates which epitomise shooting skill and the probability of good competition results. Supplementary, they could develop specific training routines targeting possible misses like the 1st prone and the 5th standing shot. Machine learning models could be used to analyse hit probabilities for individual athletes to aid in the preparation for important competitions. Furthermore, live predictions of the probability of individual hits could be attractive for broadcasting purposes.

Acknowledgments

This study was supported by the Swiss Federal Institute of Sport Magglingen (SFISM). We would like to thank Dr Joachim Steinwendner and Dr Beat Tödtli from the Swiss Distance UAS for the training in machine learning.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Baca, A., & Kornfeind, P. (2012). Stability analysis of motion patterns in biathlon shooting. *Human Movement Science*, *31*(2), 295–302.

Ball, K. A., Best, R. J., & Wrigley, T. V. (2003). Body sway, aim point fluctuation and performance in rifle shooters: Inter- and intra-individual analysis. *Journal of Sports Sciences*, *21*(7), 559–566.

Bauer, E., & Kohavi, R. (1999). An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine Learning*, *36*(1), 105–139.

Begg, R., & Kamruzzaman, J. (2005). A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data. *Journal of Biomechanics*, *38*(3), 401–408.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, *30* (7), 1145–1159.

Chen, T., & He, T. (2017). xgboost: Extreme gradient boosting (Version 0.6-4) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/xgboost

Era, P., Konttinen, N., Mehto, P., Saarela, P., & Lyytinen, H. (1996). Postural stability and skilled performance–A study on top-level and naive rifle shooters. *Journal of Biomechanics*, *29*(3), 301–306.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, *17*(3), 295–314.

Groslambert, A., Candau, R., Hoffman, M. D., Bardy, B., & Rouillon, J. D. (1999). Validation of simple tests of biathlon shooting ability. *International Journal of Sports Medicine*, *20*(3), 179–182.

Hawkins, R. (2011). Identifying mechanic measures that best predict air-pistol shooting performance. *International Journal of Performance Analysis in Sport*, *11*(3), 499–509.

Hoffman, M. D., Gilson, P. M., Westenburg, T. M., & Spencer, W. A. (1992). Biathlon shooting performance after exercise of different intensities. *International Journal of Sports Medicine*, *13*(3), 270–273.

Hoffman, M. D., & Street, G. M. (1992). Characterization of the heart rate response during biathlon. *International Journal of Sports Medicine*, *13*(5), 390–394.

Hyndman, R. J., & Athanasopoulos, G. (2013). *Forecasting: Principles and practice*. Melbourne: OTexts.

Kuhn, M. (2017). caret: Classification and regression training (Version 6.0-77) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/caret

Luchsinger, H., Kocbach, J., Ettema, G., & Sandbakk, Ø. (2017). Comparison of performance-levels and sex on sprint race performance in the biathlon world cup. *International Journal of Sports Physiology and Performance*. Advance online publication.

Mężyk, E., & Unold, O. (2011). Machine learning approach to model sport training. *Computers in Human Behavior*, *27*(5), 1499–1506.

Nitzsche, K. (1998). *Biathlon: Leistung - Training - Wettkampf: ein Lehrbuch für Trainer, Übungsleiter und Aktive* [Biathlon: performance - training - competition: a textbook for coaches, instructors and athletes]. Wiebelsheim: Limpert.

Novatchkov, H., & Baca, A. (2013). Artificial intelligence in sports on the example of weight training. *Journal of Sports Science and Medicine*, *12*(1), 27–37.

Ofoghi, B., Zeleznikow, J., Dwyer, D., & Macmahon, C. (2013). Modelling and analysing track cycling Omnium performances using statistical and machine learning techniques. *Journal of Sports Sciences*, *31*(9), 954–962.

Ofoghi, B., Zeleznikow, J., Macmahon, C., Rehula, J., & Dwyer, D. B. (2016). Performance analysis and prediction in triathlon. *Journal of Sports Sciences*, *34*(7), 607–612.

R Core Team. (2016). R: A language and environment for statistical computing [Computer software]. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Sattlecker, G., Buchecker, M., Gressenbauer, C., Müller, E., & Lindinger, S. J. (2017). Factors discriminating high from low score performance in biathlon shooting. *International Journal of Sports Physiology and Performance*, *12*(3), 377–384.

Sattlecker, G., Buchecker, M., Müller, E., & Lindinger, S. J. (2014). Postural balance and rifle stability during standing shooting on an indoor gun range without physical stress in different groups of biathletes. *International Journal of Sports Science & Coaching*, *9*(1), 171–184.

Skattebo, Ø., & Losnegard, T. (2017). Variability, predictability and race factors affecting performance in elite biathlon. *International Journal of Sports Physiology and Performance*. Advance online publication.

Venables, W. N., & Ripley, B. D. (2016). nnet: Feed-forward neural networks and multinomial log-linear models (Version 7.3-12) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/nnet

Wickham, H. (2014). Tidy data. *Journal of Statistical Software*, *59*(10), 1–23.

Wickham, H. (2016). ggplot2: Create elegant data visualisations using the grammar of graphics (Version 2.2.1) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/ggplot2