Descriptive Conversion of Performance Indicators in Rugby Union

Abstract

Objectives: The primary aim of this study was to examine whether accuracy of rugby union match prediction outcomes differed dependent on the method of data analysis (i.e., isolated vs. descriptively converted or relative data). A secondary aim was to then use the most appropriate method to investigate the performance indicators (PI's) most relevant to match outcome.

Methods: Data was 16 PI's from 127 matches across the 2016-17 English Premiership rugby season. Given the binary outcome (win/lose), a random forest classification model was built using these data sets. Predictive ability of the models was further assessed by predicting outcomes from data sets of 72 matches across the 2017-18 season.

Results: The relative data model attained a balanced prediction rate of 80% (95% CI – 75-85%) for 2016-17 data, whereas the isolated data model only achieved 64% (95% CI – 58-70%). In addition, the relative data model correctly predicted 76% (95% CI – 68-84%) of the 2017-18 data, compared with 70% (95% CI – 63-77%) for the isolated data model. From the relative data model, 10 PI's had significant relationships with game outcome; kicks from hand, clean breaks, average carry distance, penalties conceded when the opposition have the ball, turnovers conceded, total metres carried, defenders beaten, ratio of tackles missed to tackles made, total missed tackles, and turnovers won.

Conclusions: Outcomes of Premiership rugby matches are better predicted when relative data sets are utilised. Basic open-field abilities based around an effective kicking game, ball carrying abilities, and not conceding penalties when the opposition are in possession are the most relevant predictors of success.

Keywords: Team sport, random forest, performance indicators, partial dependence plots

Introduction

Success in sport can be assessed and quantified with performance indicators (PIs)[1]. Understanding PI's that relate to success in sport is important for coaches to improve future technical, tactical and physiological performance[2]. Whilst the most meaningful PI's should differentiate between successful and unsuccessful outcomes[1], no consensus can currently be drawn in rugby union regarding PI's associated with success [3–9].

Based on the available literature, the frequency of ball kicking differentiates success in both domestic and international rugby union matches[4,7,8]. Winning teams kick the ball more and kick away greater proportions of possession. Match winners also have lower error[4,9] and turnover[8,9] rates compared to losers. In addition, winners have an effective defensive game, with a superior success rate at the tackle[8] and make more tackles overall[4]. Attacking actions, such as higher distance of average carry[8] and making more clean breaks in the opposition's defensive line[3,7,8], are also associated with successful performances. Together with open field actions, set piece performance is important, with winners securing more opposition lineouts[9] and a greater effectiveness at the scrum[7]. However, some research has failed to uncover significant differences in PI's between successful and less successful teams. For example, at the 2011 World Cup competition, multiple indicators were examined and no differences were established that explained tournament ranking[5].

It is unlikely that the complex, dynamic and interactive games such as rugby union can be represented by simple analysis or frequency data[5]. The conflict in current literature with respect to PI's and match outcome is best represented by Vaz et al[4]. They reported significant predictors of match outcome in the Super Rugby competition, but the same PI's did not differentiate between winners and losers in an International competition. The authors suggested international level differences between winners and losers do not exist or are masked by variations in playing styles that underpin match outcome.

A significant limitation of the above research is the failure to acknowledge that, in rugby union, outcome depends on ability and performance of both teams. Therefore, when considering associations between PI's and competition results equal emphasis should be placed on data from each team[2]. Failure to do so will likely distort any relationships present[1]. Processing sports data to consider PI's as a differential between opponents is suggested as a better descriptor of a sport's nature[10] and a contest's outcome. In analysing sports data, this type of data processing method has been termed "descriptive conversion" but has not been applied in the literature concerning rugby union. Only isolated data has been considered, 'isolated' referring to the PI's of each participating team considered discretely and not relative to the opposition.

The primary aim of this study was to examine whether accuracy of match prediction outcomes differed dependent on the method of data analysis (i.e., isolated vs descriptively converted data). A secondary aim was to use the most appropriate method to identify the most relevant PI's for successful outcomes in rugby union and specify how this information can have practical relevance to sports practitioners.

Methods

PI's for the 2016-17 English Premiership Rugby Union regular season and the first 12 rounds of the 2017-18 season were downloaded from the OPTA website (optaprorugby.com). The 2016-17 season data consisted of 22 rounds of 6 matches (132 matches total, 12 teams). As the study assessed the impact of PI's on a binomial outcome (win/loss), matches that finished with a draw (n = 5) were excluded from analysis. The full set of team PI's for each match were utilised in the analysis. These PI's were "carries made", "clean breaks", "offloads", "total number of defenders beaten", "total number of metres ball was carried", "tackles made", "tackles missed", "ratio tackles missed to tackles made", "turnovers a team won", "turnovers a team conceded", "lineout throws won on own ball", "lineout throws lost on own ball", "the number of kicks from hand", "penalties conceded offence" (with the ball), "penalties conceded defence" (without the ball) and "the average distance for each ball carry".

The PI's of a single team, from one match, were considered isolated data. For example, if team A carried 450 m in total during the game and team B 300 m, the assigned isolated data values would be 450 m for team A and 300 m for team B. From each game, descriptive conversion was also undertaken by calculating the differences between teams and this data set was termed the relative data set. From the previous example the relative data values would be +150 m for team A and -150 m for team B.

Collinearity between predictors, in both data sets, was investigated using the rfUtlities package[11]. No collinearity was noted between predictors in the isolated data set. Collinearity was noted between defenders beaten and tackles missed in the relative data set. A separate analysis was run for the relative data set, with these predictors eliminated. The results indicated that the collinearity had no effect on the predictive ability or the casual inferences from the random. forest. With this in mind the decision was made to run the analysis with the original "intact' data set.

PI's from each data set (relative and isolated) were used as predictors for match outcomes (win/lose). To interpret relationships between PI's and match outcome a random forest classification model was developed, using 2016-17 season data, with the randomForest[12] package in R[13]. A classification model predicts categorical outcome from a set of predictor variables[14]. The randomForest package uses ensembles of decision making trees to categorise data[15]. A decision tree repeatedly repartitions data, with binary splits, to maximise subset homogeneity, and estimates the class or distribution of a response[16]. The aggregate tree approach of a random forest algorithm, has improved performance when compared to a single tree[15]. Random forests also utilise bootstrapped data samples and random subsampling of predictors in each tree to improve prediction accuracy and prevent overfitting[15]. The mean decrease of accuracy (MDA)[15] and mean of the distribution of minimal depth[17] of each PI were utilised to assess the importance of each predictor towards classification of game outcome and Pearson's correlation coefficients used to assess agreement between PI MDA and mean of distribution of minimal depth in each model[18]. A negative MDA value represents a decrease in importance and not the presence of an inverse relationship[19]. The significance level ($p < 0.05$) of the MDA of each PI was calculated, using the rfPermute package[20], the rfPermute package permutes the response variable and produces a null distribution for each predictor MDA and a p value of observed.

Partial dependency plots were produced for each significant predictor in the model defined as most appropriate by the primary statistical analysis. Partial dependency plots are useful to summarise the relationships between predictor and outcome relationships[21] and are based on permeated data sets that calculate the relationship between outcome and particular predictor changes, accounting for averaged associations of all other predictors on outcome[16].

Data from the first 12 rounds of the 2017-18 (i.e. the subsequent season) English Premiership competition was then used to test the predictive relevance (i.e. overall accuracy of prediction and balance) of both the isolated and relative models. Balance ensured models were equally adept at picking winning or losing data sets and not having bias of success to either[22]. Statistical significance of predictive accuracy for each model was recorded, as were z-scores for McNemar's test[23], which was performed against the comparison of predictive ability of each model. McNemar's test produces a z-score which when above 1.64 is indicative of a confidence level of 95% that one model has better performance than another.

Results

The randomForest model based on the isolated data set from the 2016-17 season classified 85 from 127 losses (67%) and 78 from 127 wins (61%), giving an overall accuracy of 64% (95% CI 58-70%, $p<0.05$). The randomForest model based on the relative data set predicted 102 of 127 losses (80%) and 101 of 127 wins (80%), with an overall accuracy of 80% (95% CI 75-85%, $p<0.05$). The McNemar's value of 57.7 ($p<0.05$) confirmed that the relative model outperformed the isolated model.

When assessing the predictive ability of the isolated data model against the first 12-rounds of the 2017-18 season, 58 from 72 (81%) losses and 43 from 72 (60%) wins were correctly classified, giving an overall accuracy of 70% (95% CI 63-77%, $p<0.05$). Assessment of the model based on relative data resulted in correct predictions for 54 of 72 wins (75%) and 55 of 72 losses (76%). This equated to an overall accuracy of 76% (95% CI 68-84%, $p<0.05$). McNemar's z score (31.1, $p<0.05$) again confirmed the superior performance of the relative data model.

Data with respect to each individual predictor variable's MDA is summarised in Table 1 and Table 2 for the models based on the isolated and relative data sets, respectively. The isolated data set model contained eight predictors whose distribution varied significantly from the null. The relative data set model included ten predictors whose distribution varied significantly from the null. The magnitude of significant MDA values ranged from 13.8 to -1.8 in the isolated data model and 51.6 to -4.6 in the relative data model. Mean values for minimum depth value for predictors in the isolated set varied from 2.53 for the strongest predictor to 4.4 for the weakest. In the relative set these values were between 1.81 and 4.44. A strong, negative correlation existed between MDA values of predictor importance and mean minimum depth distribution within both models, the coefficient for the relative data model being significantly higher[18] (r^2=-0.63 isolated data predictors ($p<0.05$), r^2=-0.91 relative data predictors ($p<0.05$).

******Table 1 ******

******Table 2 *******

Partial dependence plots for the top four predictors (based on MDA) were produced for the relative data model (Figure 1a-d). Plots demonstrate positive associations between match outcome and numbers of relative kicks from hand, relative clean breaks and relative average carry. A negative relationship is present with penalties conceded in defence (when the opposition are in possession). Plots also reveals upper limits are present for each PI, beyond which no increase in the probability of a positive match outcome was noted.

******Figure 1 ********

Discussion

The primary aim of this study was to investigate for the first time whether a relative (a data set that has undergone descriptive conversion) or an isolated data set best predicted outcomes of rugby union matches. Results indicated relative data was more effective at predicting match outcome compared to isolated data. The model based on the relative data set outperformed the isolated data model in terms of overall accuracy and, as per previous research[24,25], the balance of prediction was poorer from the isolated model. Isolated data sets are a less accurate reflection of the association between PI's and match outcome[1,10]. If data used to produce classification models is not an entirely accurate reflection of competition results, a bias will be present in the predictive outcomes. The reduced accuracy and balance of the isolated model in this study may help explain the conflict in previous research that have used isolated data sets[4–7].

Stability of the ranking of predictors produced from random forests is key to their interpretation[26]. The stochastic nature of a random forest is a result of the bagging, randomisation and permutation of the data set that is intrinsic to the methodology used in the process[27]. Variable importance measures with small magnitudes of difference are more likely to have their rankings influenced by the processes that are central to the methodology. The MDA values of the models are presented in Tables 1 and 2. The PI's ranked first and fourth (for example) from the relative data model have larger magnitudes of differences between them than the first and forth ranked PI's from the isolated data model. This denotes greater stability to deviations in ranking from the inherent modelling process and likely perturbations in future data. The larger magnitude of the MDA vales for the model based on the relative data set also signify greater overall importance and relevance of the data's ability to predict match outcomes[28], bringing into question the use of isolated PI's in rugby union. This conclusion is supported by the mean minimum depth distribution for a variable (Table 1), confirming the primacy of the relative data model. Pearson's correlation coefficients between mean minimum depths of each predictor and its MDA value confirmed a greater level of agreement within the relative model.

A secondary aim was to specify how our information can have practical application to sports practitioners. Partial dependence plots are a novel method to examine a multitude of relationships[29] but have not been utilised extensively in a sports performance setting to interpret statistical information for practical use. They provide a useful summary of the relationships between predictor variables and the predicted probability of match outcome[21]. The partial dependence plots indicate there are upper limits for predictor levels, beyond which no advantages are inferred towards game outcome (but not necessarily points difference). These upper limits (and their associated lower limits) offer objective outcome measures for teams to base game plans on and assess where training time is spent to win more matches.

The top four predictors from the relative data model were represented in the partial dependence plots (Figure 1a-d) and show that increases in average carry, clean breaks made and kicks made are related to improved likelihood of positive match outcomes. Conversely increased penalties, whilst the opposition have the ball, make a negative outcome more likely. Of note, penalties conceded when the opposition have the ball had a significant relationship with match outcome but penalties conceded when in possession of the ball did not. Possibly, this relationship is not solely a reflection of the penalties given away but a vestige of possession levels of teams; a high number of penalties conceded when the opposition have the ball may merely be a function of increased quantity of possession of the opposition. A further investigation needs to be undertaken that directly examines the relationships between penalties conceded when the opposition possess the ball, team possession, and game outcome. Whilst it is problematic to make presumptions without these objective data, the relationship between match outcome and penalties is such that teams need to focus on defensive strategies that are within the laws of the game. Similar conclusions can be inferred between the relationship of game outcome and number of kicks from hand, with relative kicks being an expression of relative possession levels. Data was not available for the original 2016-17 season model to investigate further but, for the 2017-18 season, the number of possessions a team attained in a match was positively related with the number of kicks from hand (r^2=-0.42 ($p<0.05$)). Possession statistics therefore explain only 42% of the variance between kicks made in matches, the remainder provided by team attributes including match tactics and strategy. It can therefore be conjectured that kicking has an impact on game outcome outside of revealing a team's possession levels. In rugby union, kicking away possession might be advantageous when teams have exhausted other options and are under pressure of turning the ball over or being penalised in an unfavourable position. Equally, kicking the ball away before a team is under pressure may be advantageous, and the relationship between kicking and success could simply reflect the advantages inferred through good tactical kicking strategy. Previous research suggests a positive relationship between possession kicked and success in both international[7] and domestic[4] rugby. Ortega[7] discusses how successful teams kick more frequently, but not the proportion of possession kicked. Vaz[4] however suggests that successful teams kick a greater amount of their possession away allowing teams to gain territory more effectively than a carrying game. This suggestion being equally applicable to the relationship between penalties in defence and match outcome.

The MDAs for clean breaks made and average carry verify the positive impact of teams having a strong ball carrying game. Indeed, research indicates clean breaks differentiated between successful and unsuccessful teams in both domestic[3] and international[7] competitions. This research demonstrates that average carry appears a more important predictor than the total metres carried. Successful teams should have strategies and players who carry greater average distance, compared to the opposition. Also, teams who prevent the opposition from carrying ball past the gainline will have a positive impact on their relative average carry. This confers the importance of robust defence as well as attacking ability and is supported by MDA values for missed tackles and ratio of tackles missed to tackles made being significant predictors of match outcome. Indeed, tackle completion has previously been shown to be an important PI in determining success[7,8]. Within the current study, tackle completion only reached significance as a predictor of match outcome in the relative model. In rugby league, regression of tackle technique is associated with fatigue, the greatest reductions in technique occurring in the players with lowest aerobic fitness levels[30]. The same relationship may exist in rugby union, indicating aerobic fitness offers an advantage toward success. No work has demonstrated a link between aerobic fitness and match outcome in rugby union.

It seems feasible that successful and unsuccessful teams differ in ability to identify tactical processes. Average distance per carry is a more accurate predictor of outcome than overall metres carried. This, combined with the observation that successful teams kick away more ball compared to losing teams may indicate the ability of successful teams to identify when effective carries can be made or otherwise to kick ball tactically. Tactically superior teams may also use the kicking game to open up attacking options as well as a pressure relieving method. A successful kicking game means opposition teams invest greater resource in covering the backfield, resulting in a weakened defensive line and opportunities for effective ball carries. Similar can be said around the tackle area, the ability to select when there is a good chance of a turnover will mean the defensive line stays intact and gives the opposition less opportunity to find space. It also has the added advantage of decreasing the number of defensive penalties conceded in these situations.

This work offers insight into rugby union not reported in the literature to date. It advances evidence that relative data surpasses isolated data in explaining game outcome, therefore being more relevant to analysts and coaches trying to influence behaviours of players and teams[2]. For instance, in previous studies success at the lineout has been demonstrated to be a predictor of success[7,9]. In this study lineouts won and lost were significant indicators in the isolated data set, but not when considered as a relative data set. This is an appropriate example of predictor and outcome relationships distortion when isolated data sets are used[1]. It is plausible the equivocality of current literature respective to predictors of performance in rugby union is in part due to the exclusive use of isolated measures of PI's. Future research should investigate physical and technical strategies to improve ball carrying quality, whilst an in-depth exploration of kicking and its impact on game outcome would also provide valuable, practical information.

Conclusions

This study demonstrates the effectiveness of utilising data that has undergone descriptive conversion in predicting match outcomes. It also demonstrates game outcomes are more closely related to open field abilities and basic skills such as ball carrying, kicking and tackling ability than they are to set pieces and, despite the apparent complexity of the game, success can be explained by a small number of basic components.

Practical applications

- The use of relative data sets rather than isolated data sets, when evaluating match performance
- Devising game strategies to maximise average carry and tackles at or over the gainline.
- Having a focus on defensive strategies that minimise the likelihood of conceding penalties. This would include areas of the game where high numbers of penalties are conceded in matches, for example when defending driving line-outs.
- Using partial dependency plots to set objective team performance markers.

Acknowledgements

No funding was provided.

References

1. Hughes MD, Bartlett RM. The use of performance indicators in performance analysis. *J Sports Sci*. 2002;20(10):739-754.
2. McGarry T. Applied and theoretical perspectives of performance analysis in sport: Scientific issues and challenges. *Int J Perform Anal Sport*. 2009;9(1):128-140.
3. den Hollander S, Brown J, Lambert M, Treu P, Hendricks S. Skills associated with line breaks in elite rugby union. *J Sports Sci Med*. 2016;15(3):501.
4. Vaz L, Van Rooyen M, Sampaio J. Rugby game-related statistics that discriminate between winning and losing teams in IRB and Super twelve close games. *J Sports Sci Med*. 2010;9(1):51.

5. Hughes MT, Hughes MD, Williams J, James N, Vuckovic G, Locke D. Performance indicators in rugby union. *J Hum Sport Exerc*. 2012;7(2).
6. Prim S, van Rooyen M, Lambert M. A comparison of performance indicators between the four South African teams and the winners of the 2005 Super 12 Rugby competition. What separates top from bottom? *Int J Perform Anal Sport*. 2006;6(2):126-133. doi:10.1080/24748668.2006.11868378.
7. Ortega E, Villarejo D, Palao JM. Differences in game statistics between winning and losing rugby teams in the Six Nations Tournament. *J Sports Sci Med*. 2009;8(4):523.
8. Watson N, Durbach I, Hendricks S, Stewart T. On the validity of team performance indicators in rugby union. *Int J Perform Anal Sport*. 2017;17(4):609-621.
9. Jones NMP, Mellalieu SD, James N. Team performance indicators as a function of winning and losing in rugby union. *Int J Perform Anal Sport*. 2004;4(1):61-71.
10. Ofoghi B, Zeleznikow J, MacMahon C, Raab M. Data mining in elite sports: a review and a framework. *Meas Phys Educ Exerc Sci*. 2013;17(3):171-186.
11. Evans JS, Murphy MA. rfutilities. R package version 1.0-0. 2014.
12. Breiman L, Cutler A, Liaw A, Wiener M. Package "randomForest." *Softw available URL http//stat-www berkeley edu/users/breiman/RandomForests*. 2011.
13. Team RC, others. R foundation for statistical computing. *Vienna, Austria*. 2013;3(0).
14. Kabacoff RI. *R in Action*. manning; 2010.
15. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32.
16. Jones Z, Linder F. Exploratory data analysis using random forests. In: *Prepared for the 73rd Annual MPSA Conference*. ; 2015.
17. Ishwaran H, Kogalur UB, Gorodeski EZ, Minn AJ, Lauer MS. High-dimensional variable selection for survival data. *J Am Stat Assoc*. 2010;105(489):205-217.
18. Diedenhofen B, Musch J. cocor: A comprehensive solution for the statistical

comparison of correlations. *PLoS One*. 2015;10(4):e0121945.

19. Ng VW, Breiman L. Bivariate variable selection for classification problem. In: *Technical Report, Department of Statistics, University of California-Berkeley*. ; 2005.

20. Archer E. rfPermute: estimate permutation p-values for Random Forest importance metrics. *R Packag (Zenodo), Version*. 2016;2(1).

21. Friedman J, Hastie T, Tibshirani R. *The Elements of Statistical Learning*. Vol 1. Springer series in statistics New York; 2001.

22. Phillips ND, Neth H, Woike JK, Gaissmaier W. FFTrees: A toolbox to create, visualize, and evaluate fast-and-frugal decision trees. *Judgm Decis Mak*. 2017;12(4):344.

23. Bostanci B, Bostanci E. An evaluation of classification algorithms using Mc Nemar's test. In: *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012)*. ; 2013:15-26.

24. Woods CT, Sinclair W, Robertson S. Explaining match outcome and ladder position in the National Rugby League using team performance indicators. *J Sci Med Sport*. 2017.

25. Leicht AS, Gomez MA, Woods CT. Team Performance Indicators Explain Outcome during Women's Basketball Matches at the Olympic Games. *Sports*. 2017;5(4):96.

26. He Z, Yu W. Stable feature selection for biomarker discovery. *Comput Biol Chem*. 2010;34(4):215-225.

27. Huazhen W, Fan Y, Zhiyuan L. An Experimental Study of the Intrinsic Stability of Random Forest Variable Importance Measures. *BMC Bioinformatics*. 2016;17(60).

28. Pearson R. Assessing Variable Importance for Predictive Models of Arbitrary Type. Cran.r-project Vignettes. https://cran.r-project.org/web/packages/datarobot/vignettes/VariableImportance.html. Published 2018.

29. Cutler DR, Edwards TC, Beard KH, et al. Random forests for classification in ecology. *Ecology*. 2007;88(11):2783-2792.

30. Gabbett TJ. Influence of fatigue on tackling technique in rugby league players. *J Strength Cond Res*. 2008;22(2):625-632.

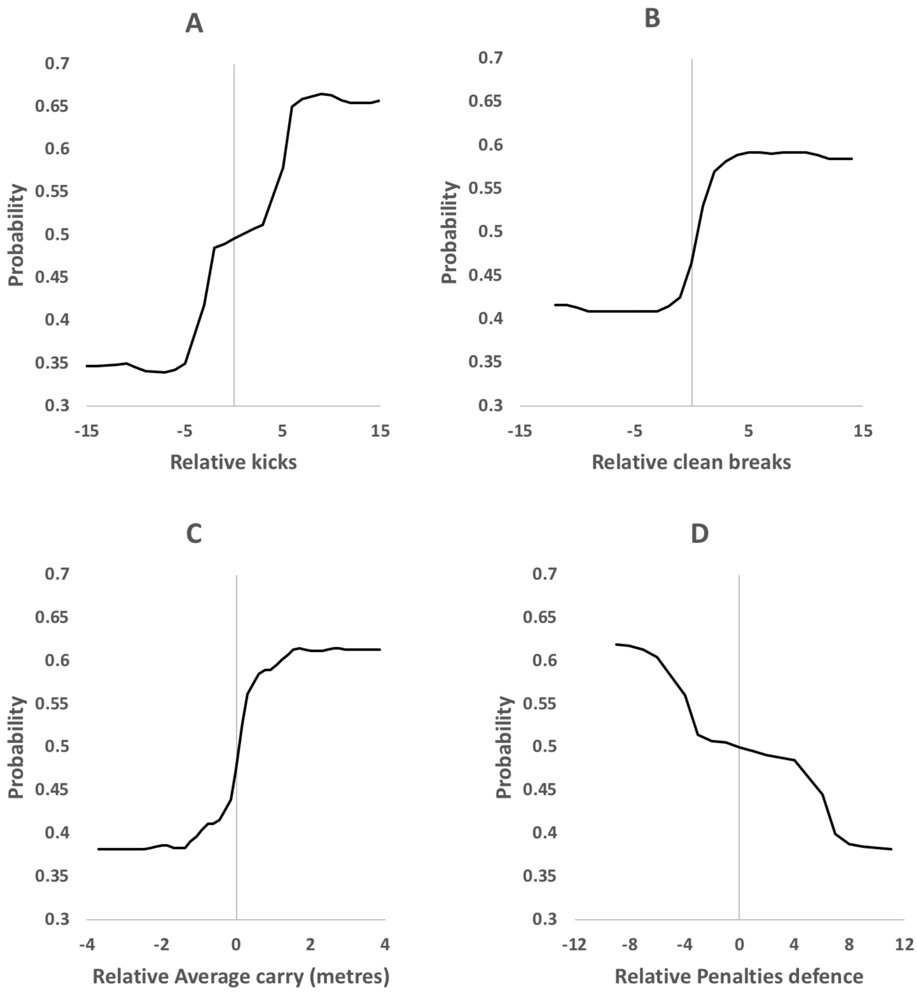

Figure 1. Partial dependence plots for Random forest model based on the relative data set. The plots show the effect of relative kicks (Panel A), relative clean breaks (Panel B), relative average carry (Panel C) and relative penalties in defence (Panel D) on the classification of match outcome.

Table 1. Mean decrease in accuracy, associated p values and mean value of minimum depth distribution for the Random Forest model, based on the isolated set.

Performance indicator	MDA	p value	Mean min depth
Average carry	13.8	0.0198	2.53
Turnovers conceded	13.4	0.0099	2.98
Clean breaks	11.0	0.0198	3.19
Total metres carried	10.7	0.0297	2.9
Missed tackles	9.8	0.0297	3.29
Tackles made/missed	8.7	0.0594	2.65
Kicks from hand	8.7	0.0495	3.10
Own LO won	8.5	0.0396	3.90
Own LO lost	6.7	0.0495	3.85
Defenders beaten	6.6	0.0693	3.46
Carries	4.1	0.1386	3.87
Penalties defence	2.4	0.2178	3.52
Tackles made	0.6	0.3663	3.62
Penalties offence	-0.3	0.4275	4.4
Turnovers won	-0.6	0.5050	3.9
Offloads	-1.8	0.6535	3.95

Table 2. Mean decrease in accuracy, associated p values and mean value of minimum depth distribution for the Random Forest model, based on the relative set.

Performance indicator	MDA	p value	Mean min depth
Kicks from hand	51.6	0.0099	1.81
Clean breaks	34.3	0.0099	2.31
Average carry	34.2	0.0099	2.17
Penalties defence	23.9	0.0099	2.62
Turnovers conceded	20.9	0.0099	2.79
Total metres carried	16.9	0.0099	2.88
Defenders beaten	12.3	0.0099	3.54
Tackle made: missed	12.2	0.0099	3.19
Missed tackles	12.0	0.0099	3.67
Turnovers won	6.2	0.0495	3.31
Carries	5.4	0.1800	3.89
Own LO won	3.5	0.2574	3.58
Offloads	1.8	0.2574	3.68
Tackles made	1.4	0.2673	3.93
Own LO lost	-0.1	0.4653	3.94
Penalties defence	-4.6	0.9505	4.44