Journal of Science and Medicine in Sport 22 (2019) 467–471

Contents lists available at ScienceDirect

Journal of Science and Medicine in Sport

journal homepage: www.elsevier.com/locate/jsams

Original research

The relationship between match performance indicators and outcome in Australian Football

Christopher M. Young[a,*], Wei Luo[b], Paul Gastin[a], Jacqueline Tran[a,c], Dan B. Dwyer[a]

[a] *Centre for Sport Research, Deakin University, Geelong, Australia*
[b] *School of Information Technology, Deakin University, Geelong, Australia*
[c] *Geelong Cats Football Club, Geelong, Australia*

ARTICLE INFO

Article history:
Received 29 June 2018
Received in revised form 17 August 2018
Accepted 21 September 2018
Available online 10 October 2018

Keywords:
Decision support techniques
Data mining
Decision trees
Models
Performance analysis
Sports analytics

ABSTRACT

Objectives: To identify novel insights about performance in Australian Football (AF), by modelling the relationships between player actions and match outcomes. This study extends and improves on previous studies by utilising a wider range of performance indicators (PIs) and a longer time frame for the development of predictive models.
Design: Observational.
Methods: Ninety-one team PIs from the 2001 to 2016 Australian Football League seasons were used as independent variables. The categorical Win–Loss and continuous Score Margin match outcome measures were used as dependent variables. Decision tree and Generalised Linear Models were created to describe the relationships between the values of the PIs and match outcome.
Results: Decision tree models predicted Win–Loss and Score Margin with up to 88.9% and 70.3% accuracy, respectively. The Generalised Linear Models predicted Score Margin to within 6.8 points (RMSE) and Win–Loss with up to 95.1% accuracy. The PIs that are most predictive of match outcome include; Turnovers Forced score, Inside 50 s per shot, Metres Gained and Time in Possession, all in their relative (to opposition) form. The decision trees illustrate how combinations of the values of these PIs are associated with match outcome, and they indicate target values for these PIs.
Conclusions: This work used a wider range of PIs and more historical data than previous reports and consequently demonstrated higher prediction accuracies and additional insights about important indicators of performance. The methods used in this work can be implemented by other sport analysts to generate further insights that support the strategic decision-making processes of coaches.

1. Introduction

The analysis of team performance indicators (PIs) is useful from a strategic standpoint; providing the opportunity to improve one's understanding about the characteristics of performances that are associated with successful outcomes.[2] The data revolution has swept through professional sport and a transition from traditional; qualitative analysis methods to modern; data-driven analyses of sports performance has followed.[1,3] The use of mathematical models to represent the relationships between the values of PIs and measures of match outcome can reveal the relative importance of those PIs and desirable ranges of values for those PIs.[4] The performance of a model provides an indication of the certainty of the inferences that can be made about important PIs and their optimal values. Therefore; progressive improvements in the modelling of performance in sport has the potential to provide ever more valuable insights about performance enhancement. These improvements can arise from the use of a wider range of PIs that encapsulate more player actions and events; and additional seasons of data to train models. Improvements in the predictive accuracy may also arise from the use of novel modelling algorithms.

Australian (Rules) Football (AF) is a field-based, team invasion sport with two teams of 18 players on the field (see Gray and Jenkins,[5] for a thorough description of the sport). Previous investigations in AF have explored the relationship between performance within a match (PIs) and match outcomes represented as Win–Loss[6,7,9] and Score Margin.[6] This work has analysed performance in AF matches using a number of methods including; multiple regression,[6] decision trees,[7] logistic regression,[7] generalised estimating equations[9] and cumulative mixed models.[8] To

* Corresponding author.
E-mail address: yochr@deakin.edu.au (C.M. Young).

https://doi.org/10.1016/j.jsams.2018.09.235
1440-2440/

date, these reports have used a small to moderate number of PIs and a limited number of seasons.[6,7]

Stewart et al.[6] analysed 17 PIs from 5 AFL seasons and they identified Inside 50's and Kicks as the strongest predictors of final match Score Margin. More recently, Robertson et al.[7] analysed 14 publicly available PIs to predict match outcome defined as Win–Loss, achieving a prediction accuracy of 78.9% when the model was trained and tested on two seasons of data. Recently, Woods[8] investigated the relationship between 11 team PIs and end of season ladder position in AF. A significant negative association was found for Hit-outs, Clearances and Inside 50's, with final ladder position. Although the findings of Woods,[8] are novel, they were based on a single season only. Some limitations to the approaches used by these authors[6–8] are apparent, including the number (11–17) and scope of PIs and the number of seasons (1–5) which restricted the opportunity for model development and validation.

Robertson et al.[7] suggested that the use of data mining approaches may be advantageous in identifying non-linear relationships in AF and that the use of a more comprehensive set of PIs may be beneficial. Moreover, Score Margin should be utilised as a measure of match outcome (as well as Win–Loss) as it provides a more granular understanding of how PIs affect match outcomes. The use of interpretable modelling techniques in a sporting context is important, as it offers the ability to explain and understand the reasons behind model predictions.[10] Decision trees, linear and rule fit models are some of the most interpretable models,[10] while logistic regression may be considered partially interpretable.

Increasing the number and scope of PIs and the number of seasons of data that are used for modelling match outcome, may provide novel insights about performance in AF and improved certainty about those insights. Since the work of Stewart et al.[6] the number of PIs collected for an AFL game has increased. Consequently, this study is based on a wider range of PIs collected over a larger number of seasons than previous studies in AF. Furthermore, while aspects of the methodology employed here have been applied in other football codes,[1,3] its application in AF is novel this study aimed to: (1) develop interpretable match outcome prediction models, using a comprehensive set of PIs and seasons of data and (2) develop a model that provides an indication of the relative importance of PIs as predictors of Score Margin. The results of this study will inform coaches, analysts, and players about the characteristics of their performance that increase the likelihood of successful match outcomes.

2. Methods

Team PIs from the 2001 to 2016 AFL home-and-away seasons were obtained from Champion Data (Champion Data, Southbank, Australia). This consisted of a total of 54 primary PIs (i.e. raw team aggregate values) from 3145 AFL matches in total. Reliability and validity of the data collected by Champion Data has been assessed previously, with a high degree of accuracy reported.[9,11] A number of secondary PIs were also created from the existing dataset. They include PIs that represent the difference to the opposition teams' PI value or when one PI was divided by another (Inside 50's per shot) or summed together (Ball Gains/Losses). This meant that the complete database consisted of 103 PIs for analysis (54 primary and 49 secondary PIs), plus two outcome measures Win–Loss and Score Margin. The primary PIs consisted of those statistics gathered directly by Champion Data and secondary were those created from the original PIs (ratios, percentages, sum of two or more PIs). Drawn matches ($n = 25$) were removed from the analysis. Ethical approval for this study was granted by the University Human Research Ethics Committee before the study was undertaken.

All PIs were checked for collinearity using a correlation matrix (Pearson's r). If the correlation coefficient between any two PIs was equal to or greater than 0.95, the attribute with the weaker correlation with Score Margin was removed from further analysis.[12] A total of 12 PIs were removed prior to analysis, six PIs due to high collinearity (Uncontested Possessions relative, Marks uncontested, Marks uncontested relative, Clearances relative, Turnovers relative, Frees against relative) and six score-related variables were also removed as these are considered mathematically related to the outcome measures and are unlikely to provide insights on how a team needs to perform to have successful match outcomes (Goal Assists, Score Assists, Goal Conversion both raw and relative) therefore total of 97 PIs were available for analysis.

The 16-year database was arbitrarily partitioned into two equal-sized time frames. This resulted in three time frames for subsequent analyses: 2001–2008; 2009–2016; and combined, 2001–2016. Given the large size of the dataset, feature selection was applied. This process gave a weight to each PI using four methods; gini index, gain ratio, information gain and correlation. This process was iterated 100 times and appended to give an averaged weight to each PI. The feature selection process identified the top 45 most important PIs for each time frame and outcome measure, resulting in six unique datasets.

Data from each time frame was partitioned into model training and model testing sets using a 70:30 ratio, in accordance with published recommendations.[13] Data from the above time frames were split for model training: 2001–2005; 2009–2013; and 2001–2010 respectively. The remaining years within each time frame was used for model testing: 2006–2008; 2014–2016; and 2011–2016. Models for each time frame were trained using 10-fold cross-validation. Models were tested for each time frame using a bootstrapped with replacement sampling method, iterated 100 times.[14]

Decision tree models based on the C4.5 algorithm[15] were created for each time frame, using Win–Loss and Score Margin as the outcome measures. Prior to the execution of Score Margin-based models, a process of discretisation was implemented to split Score Margin into four bins with the same number of matches in each bin. The four bins represent matches with final Score Margins $-\infty$ to -30.5, -30.5 to 0, 0–30.5, and 30.5 to $+\infty$. The top 45 selected PIs identified previously were used in decision tree modelling for each time frame and outcome measure. The decision tree criterion Gini-index was used to assess model performance, as this provided the most consistent prediction accuracies. Furthermore, several parameters were adjusted to optimise the results; pruning was applied with confidence set to 0.95. The maximal depth of decision trees was also limited to five levels to reduce overfitting and ensure the models remained parsimonious.[7]

Decision tree models were also created with Metres Gained Relative used as the outcome measure. Metres Gained is defined as "the net metres gained with the ball by a player, by running, kicking, or handballing, combining measures towards attacking goal and away from defensive goal".[16] Metres Gained relative represents the difference between metres gained by opposing teams and is expressed as either a positive or negative value. Metres Gained relative was discretised into four bins according to the identified benchmark values in the decision tree model for 2009–2016 (Fig. 1), these bins included $-\infty$ to -360, -360 to -41.5, -41.5 to 171.5 and 171.5 to $+\infty$.

Generalised Linear Models (GLMs) were also created for each time frame for Score Margin in its continuous format. GLMs were chosen as they can be considered a partially interpretable model that provides a measure of the relative importance of each PI as a predictor of Score Margin. Rapidminer Studio (Rapidminer Studio. Version 7.6.001. Dortmund, North Rhine-Westphalia, Germany) was utilised for the feature selection process, decision tree models and GLMs.

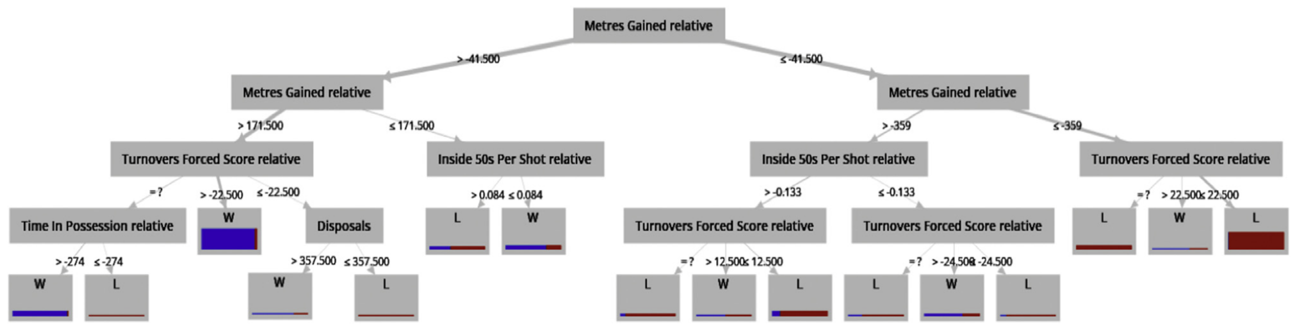

Fig. 1. A decision tree that predicts match Win–Loss for the 2009–2016 seasons. Nodes that are labelled "W" represent a win and "L" is for loss. The numerical values on the branches represent the benchmark values of the performance indicator above them. The height of the coloured bars in each terminal node represent the number of examples (i.e. matches) in each path.

Table 1
Decision tree model prediction accuracies for each time frame and outcome measure. Prediction of Score Margin was based on four score margin categories.

Time frame	Prediction accuracy	
	Win–Loss	Score margin
2001–2008	83.5%	64.4%
2009–2016	88.4%	70.3%
2001–2016	88.9%	69.7%

3. Results

The decision tree models achieved higher prediction accuracies for Win–Loss than for Score Margin (Table 1). Furthermore, the models for 2001–2016 and 2009–2016 were approximately 5% more accurate than for 2001–2008, for both outcome measures.

The decision tree that predicts whether a match outcome was Win–Loss during 2009–2016, branches first on the PI, Metres Gained relative (Fig. 1), which indicates the importance of this PI in this time frame. Consequently, a decision tree that predicts the value of Metres Gained relative was created (Fig. 2). The PIs that feature prominently in the decision tree that predicts match outcome also include; Turnovers Forced score relative, Inside 50s Per Shot relative and Time in Possession.

The decision tree model that predicts the value of Metres Gained relative, infers that Time in Possession, Inside 50s, Inside 50s Per Shot and Turnovers Forced Score, all in their relative form, are the most important factors in obtaining higher values of Metres Gained relative. This model achieved a prediction accuracy of 67.4%. Both decision trees illustrate a variety of combinations of PI values that were associated with winning and losing matches in AF. These combinations are represented via the different pathways through the decision tree/s. Importantly, the thickness of the lines that join any pair of nodes, represent the proportion of matches that exist on that branch of the tree.

GLMs were also created to predict Score Margin in each time frame and the prediction accuracies are similar for each time frame but are higher than the decision tree prediction accuracies. The root mean squared error of predicted Score Margin for 2001–2008 was 7.0 points, 2009–2016 was 6.8 points and 2001–2016 was 7.4 points. When these Score Margins were converted to simple Win–Loss outcomes, the prediction accuracies were 95.1, 94.5, and 93.1% respectively. The GLMs provided coefficients for the PIs that reflect their relative importance in the prediction of Score Margin. The most important PIs were relatively consistent between the time frames and these were; Inside 50s Per Shot, Inside 50s, and Rebound 50s, each in their relative form. (See online Supplementary materials for ranked PIs and their coefficients).

4. Discussion

Our findings indicate that using a wider range of PIs can improve model accuracy for predicting AFL match outcomes. The results also confirm the importance of specific PIs with respect to winning matches, and they also demonstrate the importance of 'new' PIs that have not previously been analysed. The GLMs also provide an indication of the relative importance of PIs with respect to Score Margin at the end of a match. This, in turn, highlights the most important PIs that may be worth monitoring for the purposes of tracking team performance during matches.

The performance of decision tree models for each time frame and outcome measure achieved prediction accuracies between 64.4–88.9%. Specifically, the Win–Loss prediction accuracies (83.5–88.9%) were as high as another type of decision tree used in previous relevant work.[7] Our modelling indicates that the most important PIs were the relative forms of Inside 50s Per Shot, Inside 50s, Rebound 50s and Metres Gained. In their raw (absolute) form, Intercepts, Disposals and Turnovers were also important predictors of match outcome. These findings are in agreement with previous work,[6,7] who have identified the same or similar PIs as being most closely related to match outcomes. There were also several additional PIs that we identified as being of importance that have not previously been analysed, including Metres Gained relative, Time in Possession relative and Turnovers Forced Score relative. These PIs appear to be important in our analysis, this may simply be because they were not available for analysis by previous authors.

The decision tree models indicated Meters Gained relative was an important predictor of match outcome. As this PI is not a player action per se, but it is the outcome of player actions. Coaches could implement game tactics that focus on maximising Metres Gained for their team, whilst minimising them for opposition teams. We felt that it was prudent to investigate the PIs that predict the value of Meters Gained relative, and these included, Time in Possession, Inside 50s relative, Turnovers Forced Score, Inside 50s Per Shot and Rebound 50s. When both models are combined, they provide the opportunity to identify the key PIs (KPIs) values that are associated with winning a match and gaining metres with ball possession. They also provide an indication of the values of these KPIs that could be set as targets for teams to achieve.

Previous authors[6,7,9] have suggested models that use Score Margin may provide additional insights beyond specifying simplistic outcome measures, such as the binary Win–Loss variable. The results from the Score Margin decision tree models were promising, with performance accuracies ranging from 64.4 to 70.3%, although these were less accurate than Win–Loss models. Models for 2009–2016 and 2001–2016 outperformed the 2001–2008 model. This could be attributable to the increased range of PIs measured as time progressed through each time frame.

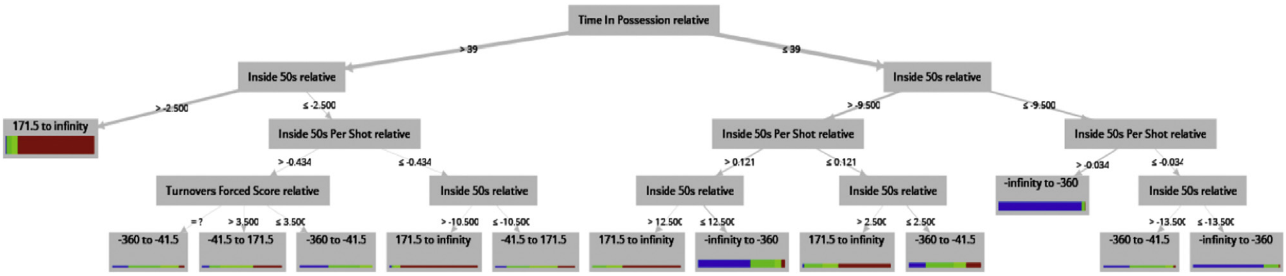

Fig. 2. A decision tree that predicts the value of Metres Gained relative for the 2009–2016 seasons. The numerical values on the branches represent the benchmark values of the performance indicator above them. The height of the coloured bars in each terminal node represent the number of examples (i.e. matches) in each path.

The application of GLMs to predict Score Margin proved to be worthwhile, as they achieved prediction accuracies that were higher than the decision tree models. Previously, generalised linear modelling has been applied in rugby[17] although these authors did not report the accuracy of Score Margin prediction. Unlike decision tree models, GLMs do not uncover the interrelatedness of PIs, nor do they identify benchmark values of PIs associated with different outcomes. The GLM-ranked PI list (see Supplementary materials) presents an alternative approach to modelling. We recommend interpreting these results in combination with the decision tree models to provide additional evidence of the KPIs important to match outcomes in AF. The relationships between PIs and Score Margin were very similar but not the same, in each time frame. It could be speculated that the relationship has changed between time frames due to a change in the typical PI values a team achieves when winning a match. Furthermore, this may provide additional evidence regarding the importance of modelling match outcomes in AF by time frames.

In general, the models in the present work achieved prediction accuracies that were as high or higher than achieved in previous reports.[7] In addition, these accuracies apply to longer time frames (3–6 seasons) than have been used in previous reports and therefore may be more reliable. The relatively high accuracies we achieved over longer time frames (seasons) may be due to several reasons. The increased number of PIs may have captured a wider range of playing characteristics, which improved the opportunity to model match outcome. The additional seasons of data may also have provided a better opportunity to train our models. Finally, the inclusion of PIs in their relative form was valuable as they tended to have a greater importance than in their raw form, which is consistent with previous work.[7]

This work presents the analysis of the largest database of PIs from AF that has been published to date. However, there are more PIs available that have not yet been analysed and these may provide further insights about performance in AF. This limitation of our study may be overcome in subsequent analyses that reveal further insights about performance that were not captured in this work. It is also acknowledged there are several approaches to identify and then reduce the number of PIs that are used for analysis. Previous work[18–20] has used factor analysis (Principal components) to analyse match-related statistics in soccer and future work in AF may benefit from this approach. The 16-year timespan of our database required us to consider the possibility that the style of play may have changed over time, as has been seen previously.[21] We felt that we could not use the time frames reported elsewhere[21] as they were too short for the modelling we have done and our database included additional PIs that may have led to the identification of different time frames. Furthermore, given that there may be changes in this sport in the future (e.g., rules changes), the PIs that we have identified as being important and the accuracy of our models may change in the future.

5. Conclusions

This work identified a set of PIs in AF that are most closely related to match outcome. Many of the PIs we identified are similar to previous reports, but we also identify new PIs such as Metres Gained. The decision tree models also indicate KPIs and the values of these KPIs that can be targeted by coaches and teams. The accuracy of our models is as high and higher than in previous reports, which underscores the value of conducting an analysis using a larger database with a wider range of PIs. There is evidence to suggest that the relationship between PIs and match outcome in AF has changed between time frames, indicating that longitudinal changes in the nature of AF have occurred. It was also found that relative form of PIs is more valuable than the absolute PIs. It may prove to be advantageous for the current results from the decision tree and GLMs to be interpreted in combination. Our methodology and findings can be used by analysts and coaches to inform decisions about playing strategy before and during a match. Furthermore, the methodologies we used could be applied to other similar sports to provide novel insights about performance.

Practical implications

- Decision tree models and Generalised Linear Models are capable of identifying how the characteristics of performance in a match relate to the outcome of a match.
- The analysis of a wider range of PIs over a longer period of time enables improvements in the accuracy of models of match outcome.
- Our models of match outcome in AF confirm the importance of Intercepts, Disposals and Turnovers, and identify the importance of PIs such as Metres Gained and Time in Possession.

Acknowledgements

We would like to acknowledge Champion Data for providing the data. There was no external financial support for this study.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jsams.2018.09.235.

References

1. Memmert D, Raabe D. *Data Analytics in Football: Positional Data Collection Modelling and Analysis*, Routledge, 2018.
2. Hughes MD, Bartlett RM. The use of performance indicators in performance analysis. *J Sports Sci* 2002; 20(10):739–754.
3. Memmert D, Rein R. Match analysis big data and tactics current trends in elite soccer. *Dtsch Z Sportmed* 2018; 69(3).
4. Moura FA, Martins LEB, Cunha SA. Analysis of football game-related statistics using multivariate techniques. *J Sports Sci* 2014; 32(20):1881–1887.

5. Gray AJ, Jenkins DG. Match analysis and the physiological demands of Australian football. *Sports Med* 2010; 40(4):347–360.
6. Stewart M, Mitchell H, Stavros C. Moneyball applied: econometrics and the identification and recruitment of elite Australian footballers. *Int J Sport Finance* 2007; 2(4):231–248.
7. Robertson S, Back N, Bartlett JD. Explaining match outcome in elite Australian rules football using team performance indicators. *J Sports Sci* 2016; 34(7):637–644.
8. Woods CT. The use of team performance indicator characteristics to explain ladder position at the conclusion of the Australian Football League home and away season. *Int J Perform Anal Sport* 2016; 16(3):837–847.
9. Robertson S, Gupta R, McIntosh S. A method to assess the influence of individual player performance distribution on match outcome in team sports. *J Sports Sci* 2016:1–8.
10. Ribeiro MT, Singh S, Guestrin C. Why should I trust you?: Explaining the predictions of any classifier, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, p. 1135–1144.
11. O'Shaughnessy DM. Possession versus position: strategic evaluation in AFL. *J Sports Sci Med* 2006; 5(4):533–540.
12. O'Brien RM. Dropping highly collinear variables from a model: why it typically is not a good idea. *Social Sci Q* 2017; 98(1):360–375.
13. Hall M, Witten I, Frank E. *Data Mining: Practical Machine Learning Tools and Techniques*, Kaufmann, Burlington, 2011.
14. Luo W, Phung D, Tran T et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. *J Med Internet Res* 2016; 18(12).
15. Quinlan JR. *C4. 5: Programs for Machine Learning*, Elsevier, 1993.
16. *AFL prospectus: the essential number-cuncher for season 2017*, 12th ed Champion Data Pty Ltd., 2017.
17. Higham DG, Hopkins WG, Pyne DB et al. Performance indicators related to points scoring and winning in international rugby sevens. *J Sports Sci Med* 2014; 13(2):358.
18. Fernandez-Navarro J, Fradua L, Zubillaga A et al. Attacking and defensive styles of play in soccer: analysis of Spanish and English elite teams. *J Sports Sci* 2016; 34(24):2195–2204.
19. Gómez MA, Gómez-Lopez M, Lago C et al. Effects of game location and final outcome on game-related statistics in each zone of the pitch in professional football. *Eur J Sport Sci* 2012; 12(5):393–398.
20. Lago-Peñas C, Gómez-Ruano M, Yang G. Styles of play in professional soccer: an approach of the Chinese Soccer Super League. *Int J Perform Anal Sport* 2017; 17(6):1073–1084.
21. Woods CT, Robertson S, Collier NF. Evolution of game-play in the Australian Football League from 2001 to 2015. *J Sports Sci* 2017; 35(19):1879–1887.