

SCORING DYNAMICS OF AUSTRALIAN FOOTBALL LEAGUE TEAMS: ANALYSIS OF THE 2008-2016 SEASONS



Jacqueline Tran^{1,2} & Ryan Letter¹

¹ Centre for Sport Research, School of Exercise & Nutrition Sciences, Deakin University (Geelong, Australia),

² Football Department, Geelong Cats Football Club (Geelong, Australia).

BACKGROUND

Scoring patterns in basketball, American football, and ice hockey leagues tend to demonstrate 'end-spurts' where scoring events are more likely to occur as quarter-, half-, or full-time approaches (Merritt & Clauset, 2014; Peel & Clauset, 2015). Further investigations are warranted in the Australian Football League (AFL) due to unique aspects of the sport that could influence scoring patterns. In the AFL, the game clock fluctuates from game-to-game due to the 'time on' rule, and phases of play can be more free-flowing than in other team sports where play is regularly stopped for player substitutions and time-outs.

The aim of this study was to examine the scoring dynamics of AFL teams, using time series analyses to investigate the probability of scoring events throughout the course of games played in the 2008-2016 seasons.

METHODS

Data was scraped from www.afltables.com to obtain scoring timelines for all games played in the 2008-2016 seasons, inclusive (total of 86,066 scoring events (goals and behinds) from 1,739 games). Total game time was scaled to 100% for every match, due to variations in absolute game durations between matches. The observed probability of a scoring event was calculated for every 0.01% interval of game time. A simple moving average was applied to the raw probability time series, and a local regression was calculated using the moving average values to reduce noise and isolate major features. The moving average and local regression time series were then decomposed to their seasonal, trend, and random components.

All analyses were conducted in R (version 3.4.0). Several packages were used: 'rvest' (Wickham, 2016), 'purrr' (Henry & Wickham, 2017), 'stringr' (Wickham, 2017), 'dplyr' (Wickham & Francois, 2016), 'lubridate' (Golemund & Wickham, 2011), reshape2 (Wickham, 2007), TTR (Ulrich, 2016), and ggplot2 (Wickham, 2009). Code stored here: www.github.com/jacquietran/AFL-scoring-dynamics.

RESULTS

On initial inspection of the decomposed time series (Figure 1), a seasonal pattern appears to exist: within any given quarter, scoring probability shows a 'pulsing', three-wave pattern. Each wave is characterised by an increase in scoring probability to a peak, then a decrease in scoring probability. Probability is lower in the third wave (towards the end of a quarter) than in the first wave (shortly after the start of a quarter). The trend component shows that scoring probability is stable through much of the first half. At 40% of total game time, scoring probability increases markedly in a linear fashion until ~70% of total game time, at which point scoring probability stabilises and remains elevated until the end of the game. However, it is important to note that the magnitude of variation in the observed time series is very small, and is primarily made up of the trend component. For context, Figure 2 shows the average score margins across all games played by each team in each season (2008-2016 inclusive).

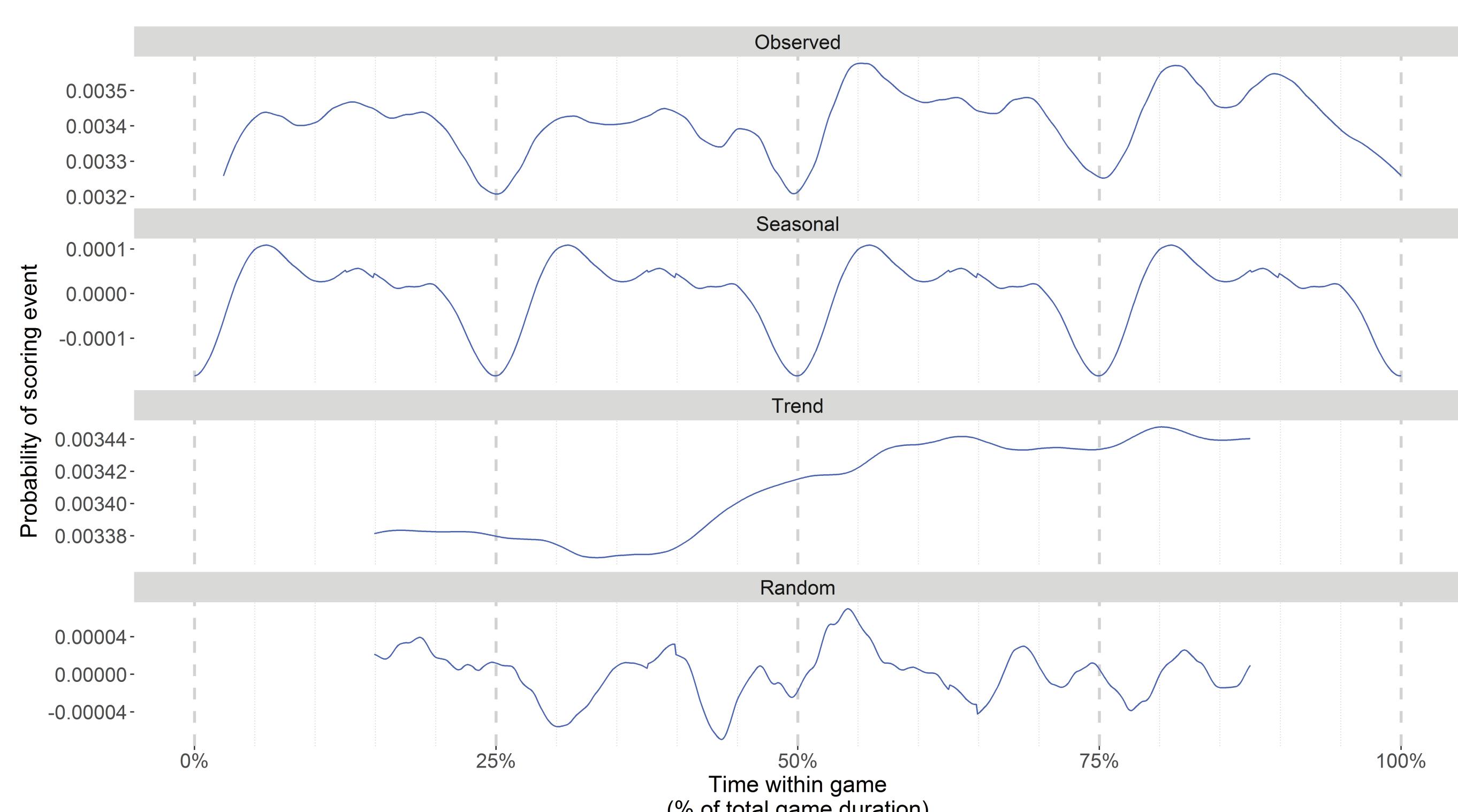


Figure 1. Decomposed time series of scoring probability versus game time.

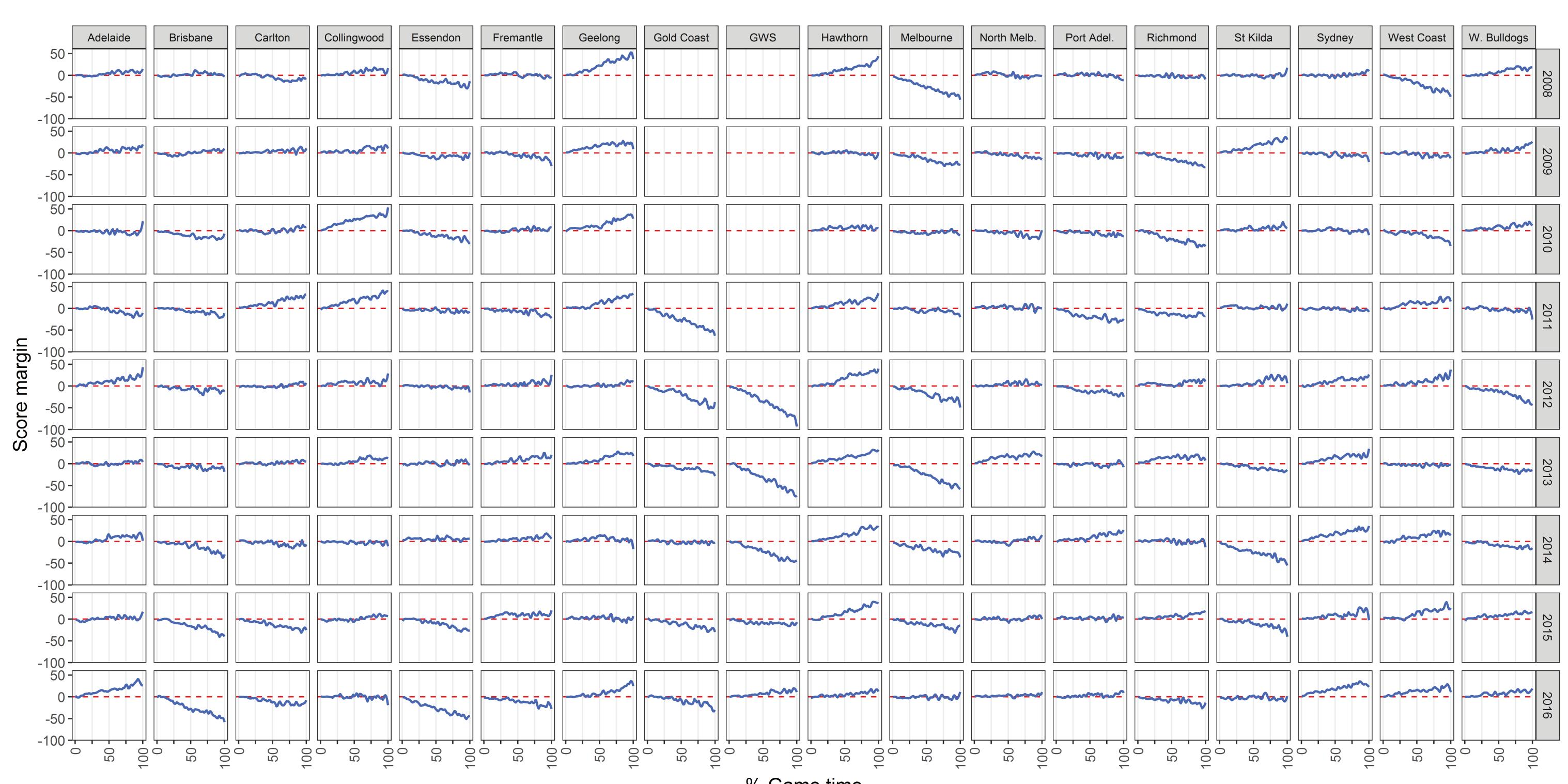


Figure 2. Average score margins for each AFL team and in each season, from 2008-2016 inclusive.

DISCUSSION

These findings indicate that the probability of scoring by AFL teams is largely independent from one instance to the next. The trend shows a higher probability of scoring events in the second half compared to the first half, which is consistent with the findings of Kiley et al. (2016). In contrast to previous research by Merritt & Clauset (2014), end-spurts are not apparent in this data. This likely reflects the differences in the strictly timed periods of sports such as basketball and ice hockey, compared to the somewhat flexible timing of quarters in Australian football and the absence of a game clock that is visible for AFL players to check throughout the course of a game. With limited published research on this topic, we suggest that future investigations focus on systematically testing popular beliefs about scoring patterns in Australian football, such as what constitutes a 'safe lead', and the relationships between the frequency / timing of lead changes and eventual match outcomes (Clauzet et al., 2015).

REFERENCES

- AFL Tables. (2017). Retrieved from www.afltables.com
Clauzet, A., Kogan, M., & Redner, S. (2015). Safe leads and lead changes in competitive team sports. *Physical Review E*, 91(6), 062815.
Golemund, G. & Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software*, 40(3), 1-25. Retrieved from <http://www.jstatsoft.org/v40/i03/>
Henry, L. & Wickham, H. (2017). purrr: Functional programming tools - R package version 0.2.2.2 [Computer software]. Retrieved from <https://CRAN.R-project.org/package=purrr>
Kiley, D.P., Reagan, A.J., Mitchell, L., Danforth, C.M., & Dodds, P.S. (2016). Game story space of professional sports: Australian rules football. *Physical Review E*, 93(5), 052314.
Merritt, S. & Clauzet, A. (2014). Scoring dynamics across professional team sports: Tempo, balance and predictability. *EPL Data Science*, 3(4).
Peel, L. & Clauzet, A. (2015). Predicting sports scoring dynamics with restoration and anti-persistence. Paper presented at the 2015 IEEE International Conference on Data Mining. doi: 10.1109/ICDM.2015.26
Ulrich, J. (2016). TTR: Technical Trading Rules - R package version 0.23-1 [Computer software]. Retrieved from <https://CRAN.R-project.org/package=TTR>
Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12), 1-20. Retrieved from <http://www.stats.umn.edu/~j211/12/>
Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
Wickham, H. (2016). rvest: Easily harvest (scrape) web pages - R package version 0.3.2 [Computer software]. Retrieved from <https://CRAN.R-project.org/package=rvest>
Wickham, H. (2017). stringr: Simple, consistent wrappers for common string operations - R package version 1.2.0 [Computer software]. Retrieved from <https://CRAN.R-project.org/package=stringr>
Wickham, H. & Francois, R. (2016). dplyr: A Grammar of Data Manipulation. R package version 0.5.0. Retrieved from <https://CRAN.R-project.org/package=dplyr>