Research Project

Rabnawaz Jansher-rabsh696

January 3, 2020

Contents

1 Background

Performance of players in ice hockey is often measured using traditional metrics such as goals, assists, points and plus-minus statistics, Corsi and Fenwick.A commonly used method to evaluate player performance is to attribute values to the different actions that players perform and sum up these values every time a player performs these actions.Models have been proposed for dealing with weaknesses of plus-minus measure(Macdonald 2011) and (Macdonald 2011).

In the field of sports analytics, researches have been conducted introducing more advanced metrics which takes context of players' actions into account(Sans Fuentes, Carlsson, and Lambrix 2019).Another work illustrates the performance of player when on ice(Schuckers and Curro 2013).More work has been done by analyzing the pair wise player performance(Ljung, Carlsson, and Lambrix 2018).Other researches have used markov model for evaluating players' performance(Kaplan, Mongeon, and Ryan 2014) and (Routley and Schulte 2015). In his research (Routley and Schulte 2015) has quantified the impact of actions on goals using Q-value function. In his research, he has emphasized on the context of actions within a game and have used state space model to learn the Q-value function. Q-value function uses reward matrix.Different researchers have defined reward function differently. In general definition, reward is given when one goal is scored. In different approach, reward function also considers the context of goal.

2 Problem

The most important metric for performance evaluation of a player is the connection of player with the goal scored. But if equal reward is assigned to every goal, the final evaluation might be biased. To make the best predictions, reward function should be changed based on the context of goal scored that is instead of giving equal reward to all goals, reward should be assigned based on the time and scenario in which the goal was scored.That is if a team is leading by 5 goals, it should not get the equal reward as when the game is tied. Based on above assumption, we are working on extending the work previously done. In this research, we have assigned different rewards to goals based on when it was scored.

3 Method

In this research, we needed to take context of goals into account e.g goal difference at the time of goal scored, manpower difference at the time of goal scored and when the goal was scored.

On the basis "when" the goal was scored, we have broken down the time into intervals. We have taken two approaches.

1. Giving weightage if goal is scored in the beginning of interval.i.e Method-1
2. Giving weightage if goal is scored at the end of interval.i.e Method-2

3.1 Formula

Sum of count for = CF, Interval = I, time = t, ManPower Difference = MD, Goal Difference = GD, Win = W, Tie-lose = TL, Tie-win = TW

$$P(W|I_t, MD, GD) = \frac{CF\ (TW)\ at\ I_t\ , when\ (\ MD,\ GD)\ +\ CF\ (W)\ at\ I_t\ , when\ (\ MD,\ GD)}{CF(MD, GD)\ at\ I_t}$$

$$P(tie|I_t, MD, GD) = \frac{CF\ (TL)\ at\ I_t\ , when\ (\ MD,\ GD)}{CF(MD, GD)\ at\ I_t}$$

So, the Reward Function will be,

$$\begin{aligned} R(I_t\ ,\ MD\ , GD) = \ & 2\ [P(W|I_t, MD, GD) - P(W|I_t, MD, GD-1)] \\ & +\ 1[P(tie|I_t, MD, GD)\ -\ P(tie|I_t, MD, GD-1)] \end{aligned}$$

3.2 Idea

The main idea is to form a reward matrix which contains the goal difference, manpower difference, goal scored, time interval when the goal was scored and the weightage based on when was the goal scored.

We have used different methods for forming time intervals.

1. Fixed time intervals (by fixed difference).
2. Dynamic time interval (based on provided vector).

3.3 Data

In this paper we use play-by-play data for NHL regular seasons and playoff seasons during the year 2013-2014.Data contains 36 tables tables in total but we are using only 3 tables(goalie,play_by_play_events,game).

Type	Number of games
Regular Season	1229
Playoffs	93

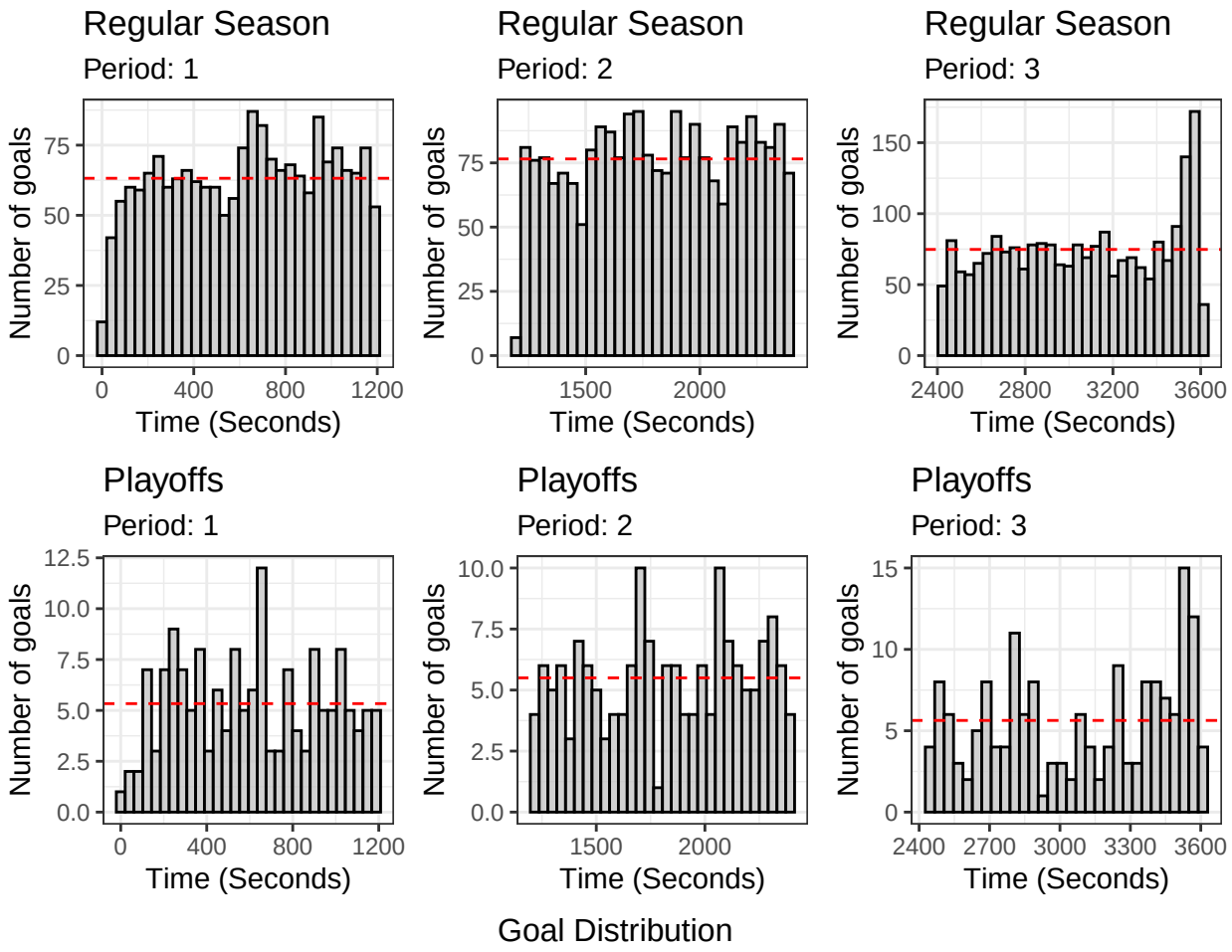
The attributes we have taken into account :

```
##  [1] "GameId"         "Season"         "SeasonType"     "AwayTeamId"
##  [5] "HomeTeamId"     "PeriodNumber"   "EventTime"      "ScoringTeamId"
##  [9] "manPowerAway"   "manPowerHome"   "seconds"        "GF"
## [13] "GA"             "fAwayTeamGoals" "fHomeTeamGoals" "outcome"
```

where GA = Gaol Away, GF = Goal For, fAwayTeamGoals = Final Away Team goals of game, fHomeTeamGoals = Final Away Team goals of game, outcome = result(win/lose/tie-win/tie-lose) and SeasonType(Regular Season, Playoffs)

4 Results:

4.1 Goal Distribution Regular Season Vs Playsoff Season

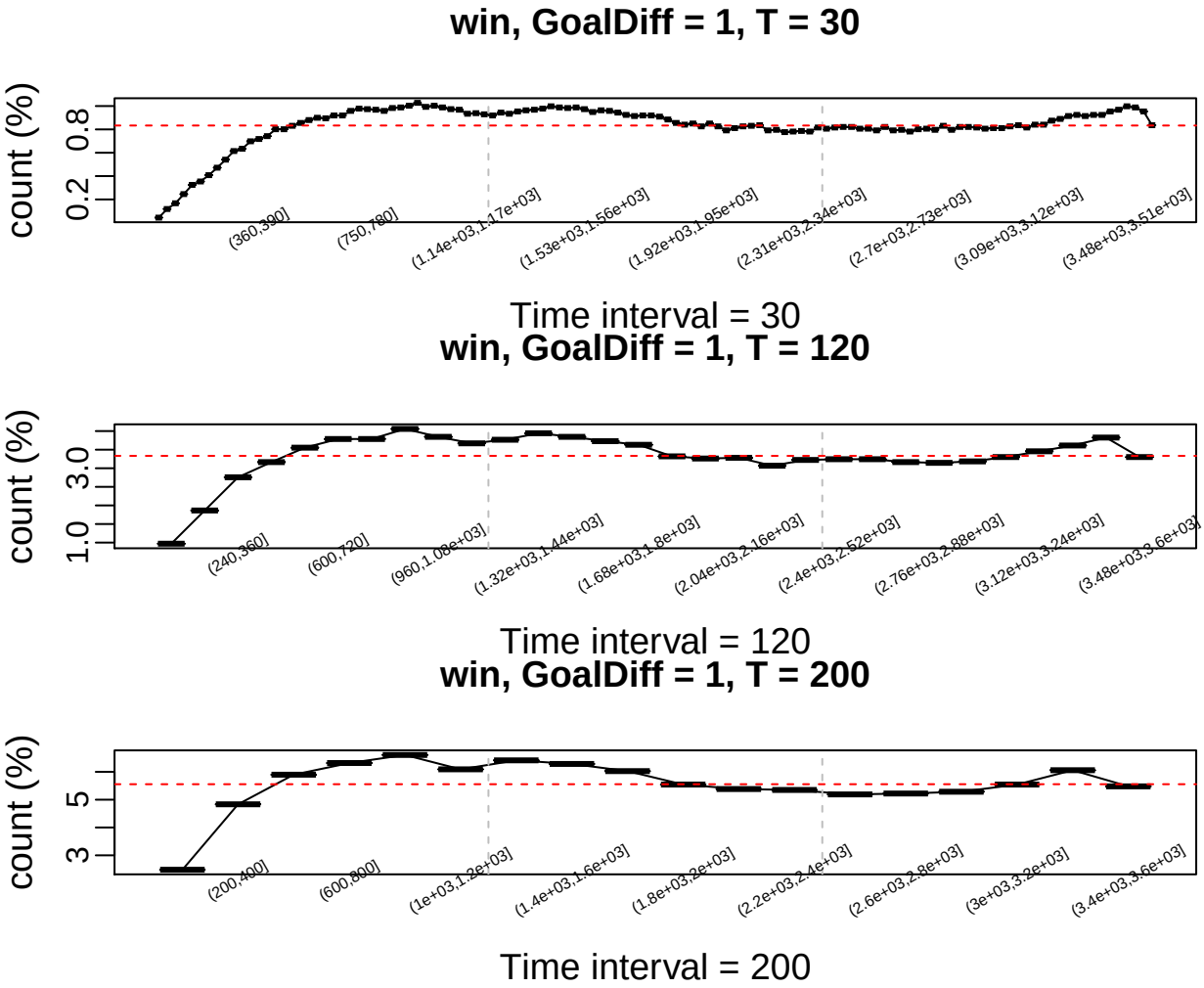


The distribution of goals over different periods is almost same with highest number of goals being scored at the end of period 3 in both seasons.

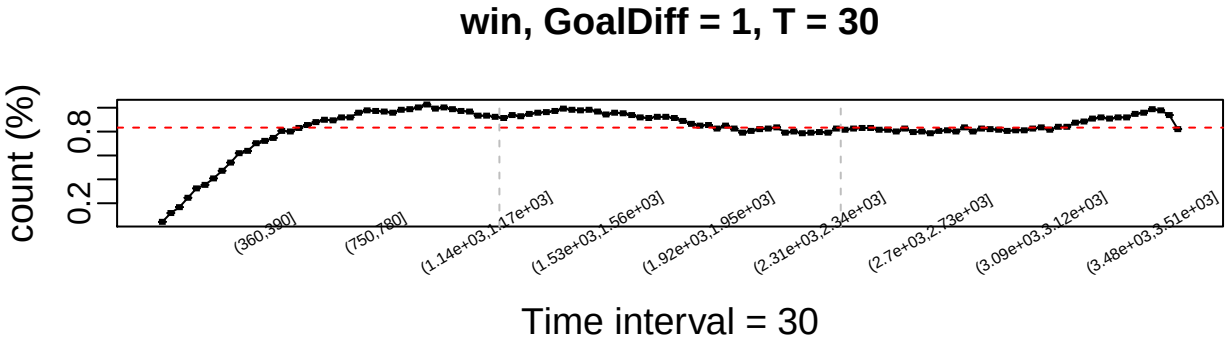
We ran our function on different time intervals and the following results were produced:

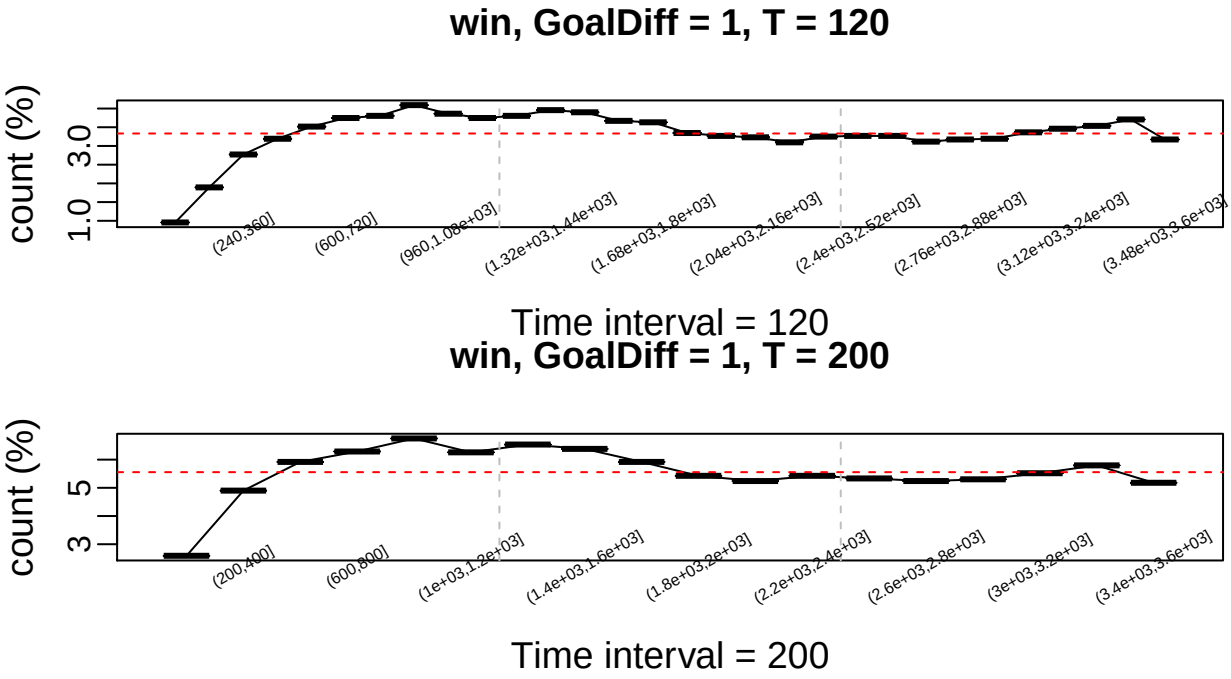
4.2 Comparison on fixed Time intervals

Compare the result on three different time intervals 30 seconds,120 seconds and 200 seconds by using **Method-1**.



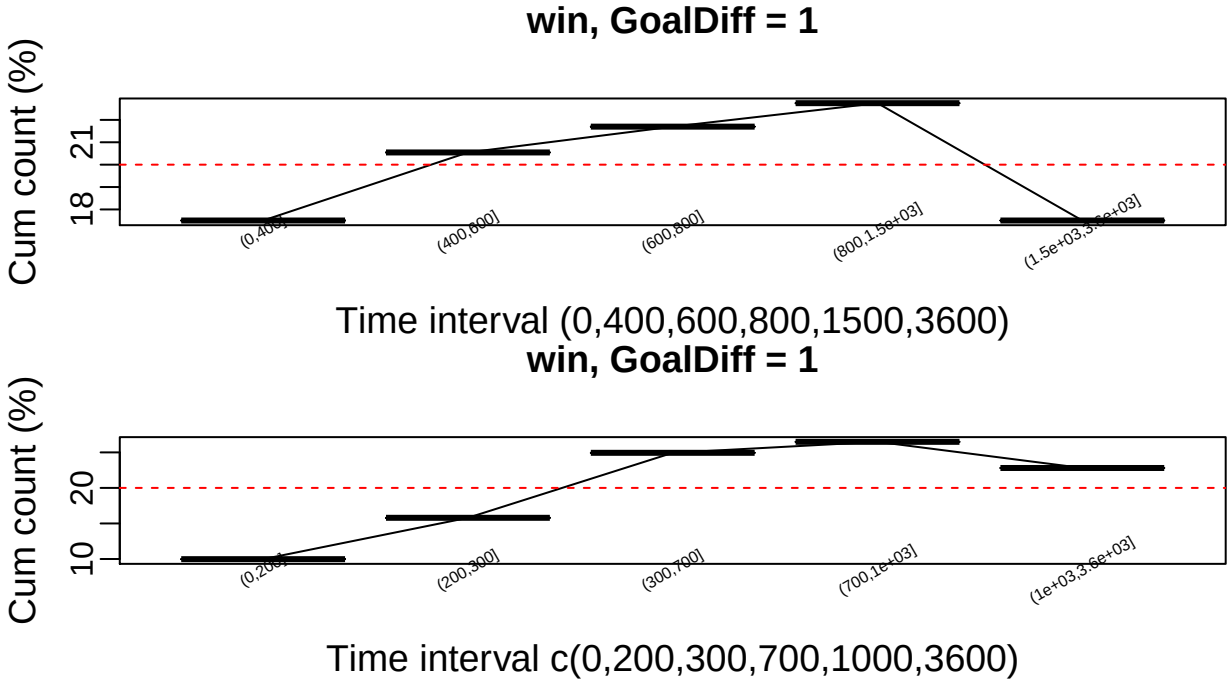
Compare the result on three different time intervals 30 seconds,120 seconds and 200 seconds by using **Method-2**.

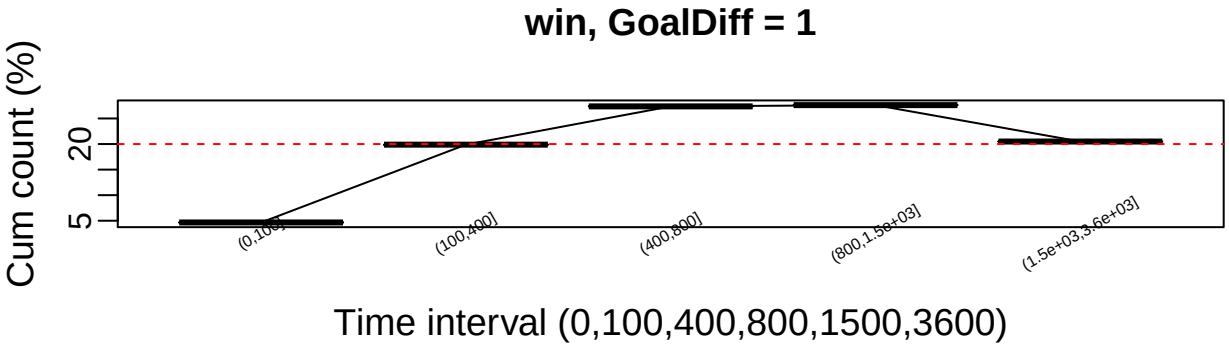




It can be observed that the sum of counts rate for hometeam winning games in 3 different time intervals using "Method-1" and "Method-2" is almost same.

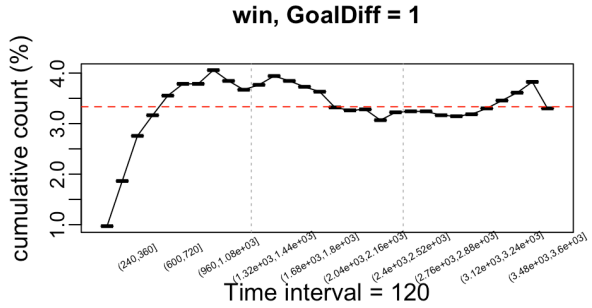
4.3 Compare on User define Time Interval

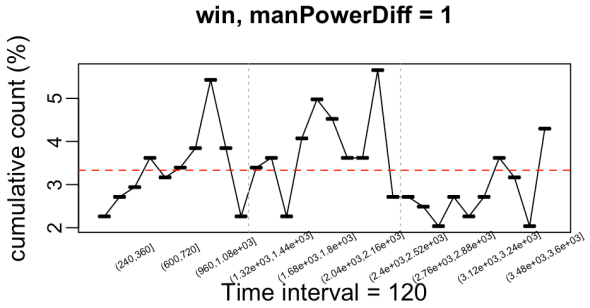


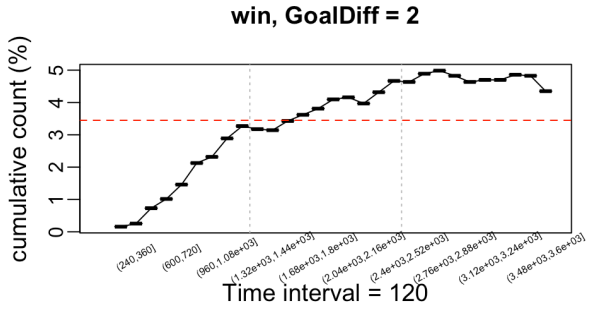


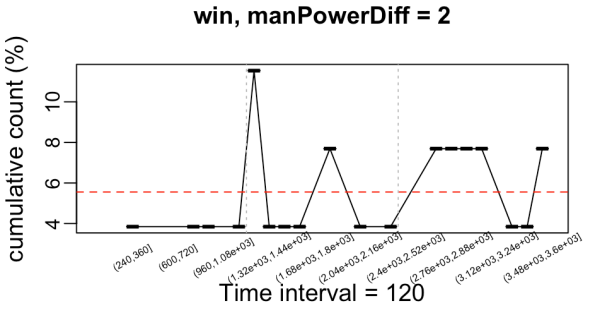
We have now determined the cummulative count of hometeam winning games, with goal difference =1 over dynamic time intervals, that is time intervals of different sizes. The winning ratio is more clear as it can be seen that the count increases in time interval 800 onwards.

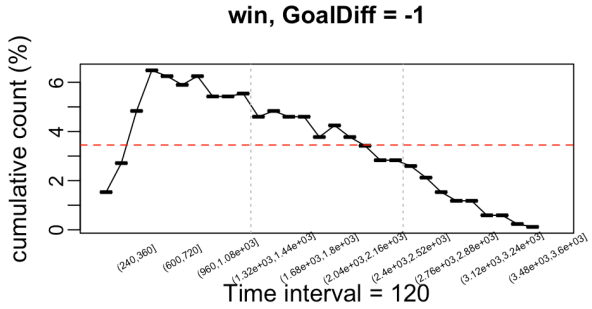
4.4 Goal Difference & ManPower Difference

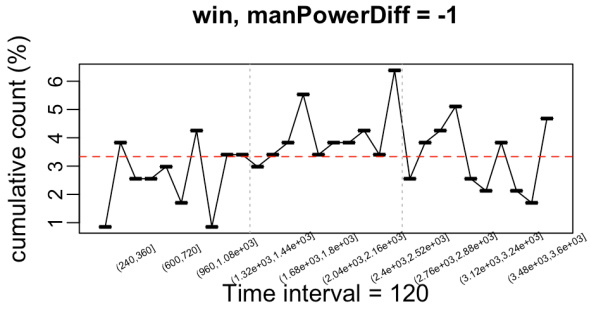


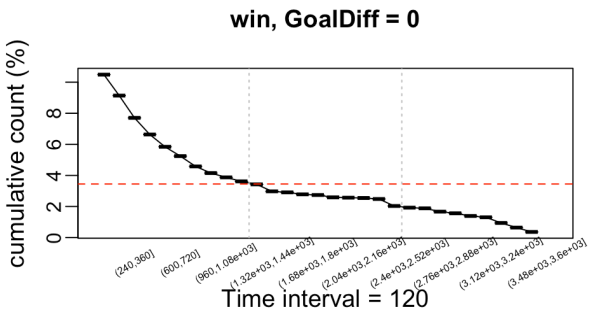


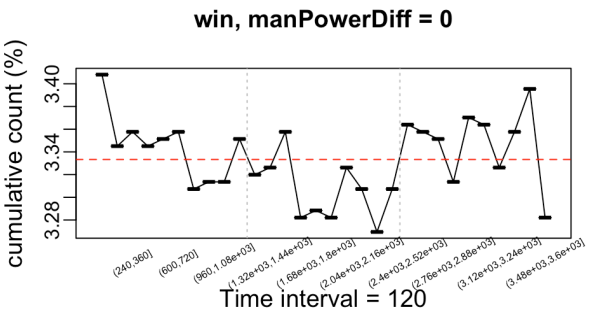












One hometeam leads by 1 and 2 goals the count ratio of win is high in almost all time intervals while the winning count ratio decreases if hometeam falls behind by one goal in later intervals, though if hometeam falls behind by 1 goal in the beginning of game, the winning ratio is still high.

The graphs of manpower difference also makes sense, as winning count ratio is fairly high when hometeam has advantage of manpower difference (that is by 1 or 2 players). But the count ratio decreases when there is negative manpower difference for the hometeam.

4.5 Outcome Comparison

	Number of Games (Outcome)			
Games	**Win**	**Loss**	**Tie-Win**	**Tie-Loss**
GD(After)	239	237	174	153
GD (before)	240	235	177	154
MD-Diff (After)	119	113	90	92
MD-Diff (before)	121	115	91	92

5 Conclusion

1. The trend is almost same for fixed intervals though trend is more visible for the user defined intervals as the intervals can be small or big in this case.
2. From the plots as well as from the table in section 4.5, it can be observed that there is not much difference in outcome using Method 1 or Method 2. The outcomes are almost same.

References

Kaplan, Edward, Kevin Mongeon, and John Ryan. 2014. "A Markov Model for Hockey: Manpower Differential and Win Probability Added." *INFOR: Information Systems and Operational Research* 52 (May): 39–50. https://doi.org/10.3138/infor.52.2.39.

Ljung, Dennis, Niklas Carlsson, and Patrick Lambrix. 2018. "Player Pairs Valuation in Ice Hockey." In.

Macdonald, Brian. 2011. "An Improved Adjusted Plus-Minus Statistic for NHL Players." http://www.sloansportsconference.com/content/an-improved-adjusted-plus-minus-statistic-for-nhl-players/.

Macdonald, Brian. 2011. "A Regression-Based Adjusted Plus-Minus Statistic for Nhl Players." *Journal of Quantitative Analysis in Sports* 7 (July). https://doi.org/10.2202/1559-0410.1284.

Routley, Kurt, and Oliver Schulte. 2015. "A Markov Game Model for Valuing Player Actions in Ice Hockey." In *Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence*, 782–91. UAI'15. Amsterdam, Netherlands: AUAI Press.

Sans Fuentes, Carles, Niklas Carlsson, and Patrick Lambrix. 2019. "Player Impact Measures for Scoring in Ice Hockey." In *Proceedings of Mathsport International 2019 Conference :* 307–17. Linköping University, Faculty of Science & Engineering; Athens University of Economics; Business.

Schuckers, Michael E., and James Curro. 2013. "Total Hockey Rating (Thor) : A Comprehensive Statistical Rating of National Hockey League Forwards and Defensemen Based Upon All on-Ice Events." In.