Hockey's Most Controversial Statistic: An Analysis of the Effectiveness of the Plus-Minus Statistic

Melanie Desroches
Department of Statistics
University of Connecticut

December 6, 2024

Abstract

The goal of this paper is to evaluate the plus minus statistic as a metric for determining the effectiveness of a player. This involves analysing the relationship between a players plus-minus rating and other offensive and defensive ratings, the effect of a players team performance on their plus-minus, and comparing plus-minus to other similar statistics. Based on the analysis performed, the plus minus does have some value as a predictor of offensive and defensive metrics. However, other advanced metrics like Corsi and Fenwick perform much better. The plus-minus is also heavily influenced by a player's team, making it less valuable as an indivual evaluator of player effectiveness.

1 Introduction

Within the realm of sports, many different statistics are used to determine how good a player is. This can include points, assists, time played, etc. But is there just one value that can be used to determine the effectiveness of a player? This is an important topic, as fans, coaches, players, and general managers often want to get a better understanding of how an individual is impacting their team. It is one thing to watch the players to determine how good a player is, often referred to as the "eye test", but with so many players, it is hard to quantify this. Enter the plus-minus statistic. Plus-minus is generally calculated by adding all the points scored by their team while they were playing and subtracting by points scored by the opposition while they were playing. The idea is that if a player has a generally positive impact on their team, they will have a highly positive plus-minus. In this project, plus-minus will be evaluated in the context of ice hockey in the National Hockey League (henceforth refered to as the NHL).

The use of the plus-minus originated in ice hockey by the NHL team Montréal Canadiens Hvattum (2019). The league started to keep track of plus-minus in 1956. The NHL defines plus-minus as "a team's goal differential while a particular player is on the ice, excluding power play goals for and against but including empty net situations. All the skaters on the ice receive a plus or minus when an even-strength goal or shorthanded goal is scored depending on which team scored" National Hockey League (2023). This means that when a goal is scored, as long as there are the same number of players from each team on the ice, the team that scored gets plus one and the other team gets a minus one.

While the plus-minus statistic is a great idea in theory, it does not come without its weaknesses. The biggest drawback to this statistic is its issues with independence. Many ice hockey players tend to play on "lines", meaning the same three forwards tend to play together and the same two defensemen usually play together. As a result, the performance of one player is highly dependent on the performance of their linemates. In "A Regression-based Adjusted Plus-Minus Statistic for NHL Players" Macdonald (2011), Brian MacDonald provides a perfect example of this with the Henrik and Daniel Sedin, Swedish twins who played together for the Vancouver Canucks. "Daniel spent 92% of his playing time with Henrik, the highest percentage of any other player combination where both players have played over 700 minutes. Because of this high collinearity between the twins, it is difficult to separate the individual effect that each player has on the net goals scored on the ice."Macdonald (2012) Many critics of the plus-minus argue that plus-minus is more of a team statistic, since it is heavily influenced by team dynamics rather than individual contributions. Furthermore, numerous confounding variables, such as the quality of the opponent and situational factors, make the calculation less reliable.

As a result of these criticisms, alternatives to the plus-minus statistic have been developed. Many of these alternatives are either different metrics or adjusted plus-minusHvattum

(2019). Two popular examples are Corsi or Fenwick, which are calculated using shot attempts rather than goals. Another popular metric is expected goals, which is based on the probability of a goal being scored depending on the shot. Brian MacDonald introduced an adjusted plus-minus statistic that is based on weighted least squares regressionMacdonald (2012). However, many of these alternatives are more complicated to calculate, which can be confusing to the average fan. Due to the fact that plus-minus is easy to interpret and calculate, it has remained in use despite its controversial background.

The goal of this project will be to shed light on this controversial statistic and determine if plus-minus is truly an effective predictor to individual contribution, specifically in the National Hockey League (henceforth refered to as the NHL). Is the plus-minus a good evaluator of offensive and defensive output? Is the plus-minus more reflective of how well a player is doing than how well a player is performing? Are there better alternatives to evaluating individual player contribution? Is the plus-minus an effective predictor to individual contribution? These are the questions that will be answered from this analysis.

The rest of the paper is organized as follows. The data will be presented in Section 2. The methods are described in Section 3. The results are reported in Section 4. A discussion concludes in Section 5.

2 Data

The data used to perform this analysis was collected from Natural Stat Trick and the NHL website. The observations in the data is from all players that played in the National Hockey League from the 2021-2022, 2022-2023, and 2023-2024 seasons. Only players that have played in over 25 games were used in the analysis. This is to ensure that the players being evaluated were all of similar calliber and were regular NHL players. The columns of the dataset will be combined from the two data sources. The following is a description of each of the columns in the dataset obtained from the website Natural Stat Trick (2023):

Player - Player name.

Team - Team or teams that the player has played for. Not displayed when filtering for specific teams.

Position - Position or positions that the player has been listed as playing by the NHL.

GP - Games played.

TOI - Total amount of time played.

Corsi - Any shot attempt (goals, shots on net, misses and blocks) outside of the shootout. Referred to as SAT by the NHL.

CF - Count of Corsi for that player's team while that player is on the ice.

CA - Count of Corsi against that player's team while that player is on the ice.

CF% - Percentage of total Corsi while that player is on the ice that are for that player's team. CF*100/(CF+CA)

Fenwick - any unblocked shot attempt (goals, shots on net and misses) outside of the shootout. Referred to as USAT by the NHL.

FF - Count of Fenwick for that player's team while that player is on the ice.

FA - Count of Fenwick against that player's team while that player is on the ice.

FF% - Percentage of total Fenwick while that player is on the ice that are for that player's team. FF*100/(FF+FA)

Shots - any shot attempt on net (goals and shots on net) outside of the shootout.

SF - Count of Shots for that player's team while that player is on the ice.

SA - Count of Shots against that player's team while that player is on the ice.

SF% - Percentage of total Shots while that player is on the ice that are for that player's team. SF*100/(SF+SA)

Goals - any goal, outside of the shootout.

GF - Count of Goals for that player's team while that player is on the ice.

GA - Count of Goals against that player's team while that player is on the ice.

GF% - Percentage of total Goals while that player is on the ice that are for that player's team. GF*100/(GF+GA)

Scoring Chances - a scoring chance, as originally defined by War-on-Ice

SCF - Count of Scoring Chances for that player's team while that player is on the ice.

SCA - Count of Scoring Chances against that player's team while that player is on the ice.

SCF% - Percentage of total Scoring Chances while that player is on the ice that are for that player's team. SCF*100/(SCF+SCA)

High Danger Scoring Chances - a scoring chance with a score of 3 or higher.

HDCF - Count of High Danger Scoring Chances for that player's team while that player is on the ice.

HDCA - Count of High Danger Scoring Chances against that player's team while that player is on the ice.

HDCF% - Percentage of total High Danger Scoring Chances while that player is on the ice that are for that player's team. HDCF*100/(HDCF+HDCA)

High Danger Goals - goals generated from High Danger Scoring Chances

HDGF - Count of Goals off of High Danger Scoring Chances for that player's team while that player is on the ice.

HDGA - Count of Goals off of High Danger Scoring Chances against that player's team while that player is on the ice.

HDGF% - Percentage of High Danger Goals while that player is on the ice that are for that player's team. HDGF*100/(HDGF+HDGA)

SH% - Percentage of Shots for that player's team while that player is on the ice that were Goals. GF*100/SF

SV% - Percentage of Shots against that player's team while that player is on the ice that were not Goals. GA*100/SA

PDO - Shooting percentage plus save percentage. (GF/SF)+(GA/SA)

SH% - Percentage of Shots for that player's team while that player is on the ice that were

Goals. GF*100/SF

SV% - Percentage of Shots against that player's team while that player is on the ice that were not Goals. GA*100/SA

Goals - Goals scored by the player, outside of the shootout.

Assists - Any assist by the player.

First Assists - Primary assists by the player.

Second Assists - Secondary assists by the player.

Total Points - Goals scored and assists by the player, outside of the shootout.

IPP - Individual Point Percentage, the percentage of goals for that player's team while that player is on the ice that the player earned a point on.

Shots - Any shot attempt on net (goals and shots on net) by the player, outside of the shootout.

SH% - Percentage of Shots by the player that were Goals. Goals*100/Shots

iCF - Any shot attempt (goals, shots on net, misses and blocks) by the player, outside of the shootout.

iFF - Any unblocked shot attempt (goals, shots on net and misses) by the player, outside of the shootout.

iSCF - Any scoring chance by the player, outside of the shootout.

iHDCF - Any high danger scoring chance by the player, outside of the shootout.

Rush Attempts - Any rush shot attempt (goals, shots on net, misses and blocks) by the player, outside of the shootout.

Rebounds Created - Any shot attempt (shots on net, misses and blocks) that results in a rebound shot attempt.

Penalties Drawn - Number of penalties committed against the player.

Giveaways - Number of unforced turnovers made by the player.

Takeaways - Number of times the player takes the puck away from the opposition.

Hits - Number of hits made by the player.

Hits Taken - Number of hits taken by the player.

Shots Blocked - Number of opposition shot attempts blocked by the player.

The goal of this dataset was to provide a broader scope of variables that evaluate a players performance. By introducing more advanced and unconventional statistics, there are more factors to consider in evaluating a player's effectiveness on the ice.

3 Methods

3.1 Correlation Analysis

The main issue with the data is multicollinearity, where independent variables are highly correlated. A correlation matrix was created with some of the more basic nhl statistics to highlight this issue, shown in Table 1.

Table 1: Correlation Matrix of Variables

	GP	G	A	P	+/-	P/GP	EVG	EVP	PPG	PPP	SHG	SHP	OTG	GWG	S	S%	TOI
GP	1.00	0.66	0.74	0.74	0.21	0.54	0.69	0.81	0.48	0.53	0.41	0.51	0.44	0.60	0.83	0.25	0.56
G	0.66	1.00	0.82	0.93	0.23	0.88	0.98	0.93	0.90	0.84	0.44	0.42	0.72	0.92	0.91	0.59	0.42
A	0.74	0.82	1.00	0.97	0.34	0.90	0.80	0.95	0.75	0.91	0.35	0.43	0.71	0.79	0.89	0.34	0.68
P	0.74	0.93	0.97	1.00	0.31	0.93	0.91	0.98	0.85	0.93	0.40	0.45	0.74	0.88	0.94	0.46	0.60
+/-	0.21	0.23	0.34	0.31	1.00	0.28	0.24	0.33	0.19	0.23	0.18	0.28	0.16	0.30	0.25	0.05	0.24
P/GP	0.54	0.88	0.90	0.93	0.28	1.00	0.85	0.89	0.83	0.91	0.33	0.34	0.73	0.84	0.84	0.56	0.57
EVG	0.69	0.98	0.80	0.91	0.24	0.85	1.00	0.93	0.80	0.78	0.42	0.40	0.69	0.90	0.90	0.60	0.39
EVP	0.81	0.93	0.95	0.98	0.33	0.89	0.93	1.00	0.78	0.84	0.41	0.45	0.71	0.87	0.95	0.46	0.59
PPG	0.48	0.90	0.75	0.85	0.19	0.83	0.80	0.78	1.00	0.88	0.30	0.29	0.68	0.84	0.78	0.50	0.40
PPP	0.53	0.84	0.91	0.93	0.23	0.91	0.78	0.84	0.88	1.00	0.28	0.30	0.74	0.81	0.82	0.40	0.55
SHG	0.41	0.44	0.35	0.40	0.18	0.33	0.42	0.41	0.30	0.28	1.00	0.88	0.25	0.39	0.42	0.30	0.17
SHP	0.51	0.42	0.43	0.45	0.28	0.34	0.40	0.45	0.29	0.30	0.88	1.00	0.24	0.38	0.45	0.23	0.29
OTG	0.44	0.72	0.71	0.74	0.16	0.73	0.69	0.71	0.68	0.74	0.25	0.24	1.00	0.76	0.68	0.33	0.48
GWG	0.60	0.92	0.79	0.88	0.30	0.84	0.90	0.87	0.84	0.81	0.39	0.38	0.76	1.00	0.84	0.53	0.42
S	0.83	0.91	0.89	0.94	0.25	0.84	0.90	0.95	0.78	0.82	0.42	0.45	0.68	0.84	1.00	0.37	0.60
S%	0.25	0.59	0.34	0.46	0.05	0.56	0.60	0.46	0.50	0.40	0.30	0.23	0.33	0.53	0.37	1.00	-0.03
TOI	0.56	0.42	0.68	0.60	0.24	0.57	0.39	0.59	0.40	0.55	0.17	0.29	0.48	0.42	0.60	-0.03	1.00

There is a high correlation between different shot related statistics and point related statistics. For example, there is a 0.826371 correlation between Goals and Shots. This makes sense because in order to score a goal, the player needs to shoot first. Correlation analysis can be used to identify the strength and direction of the relationship between various advanced metrics and the plus/minus statistic. Metrics with high positive correlations (e.g., GF% and PDO) suggest a strong alignment with the plus/minus, indicating that they may reflect similar aspects of offensive or defensive performance. This was done by calculating the correlation coefficient between +/- and each metric (like CF%, GF%, SCF%, etc.).The correlation values, ranging from -1 to 1, tell us how closely each metric aligns with the plus/minus. Higher positive correlations (e.g., with GF% and PDO) indicate metrics that vary similarly to plus/minus, suggesting that they may capture overlapping aspects of offensive or defensive performance.

3.2 Ridge Regression

Ridge regression is used in this project to evaluate the relationship between various performance metrics and plus-minus, while addressing multicollinearity among predictor variables. This regularization technique is particularly valuable in this context because many hockey performance metrics, such as Scoring Chances For percentage (SCF%) and Goals For percentage (GF%), are often highly correlated. Due to the presence of correlation among variables, ridge regression was performed in order to identify how plus/minus can be employed to assess offensive and defensive contribution. It was also used to determine if other advanced metrics Corsi or Fenwick have the same, better, or worse predictive abilities compared to plus-minus. Ridge regression, a form of regularized linear regression, is beneficial in handling datasets where predictor variables are highly interrelated, as is the case with advanced hockey metrics. Ridge regression estimates the contribution of metrics like Corsi, Fenwick, and scoring chances while controlling for their interdependencies. This applies to hockey metrics like Corsi, Fenwick, scoring chances, etc. since these statistics often interact or overlap in measuring aspects of performance. Ridge regression is appropriate because the data satisfies the model assumptions of independence, homoscedasticity, and linearity. It is ideal for this project because it effectively manages multicollinearity, retains all variables for interpretability, and improves predictive accuracy. Other regularization techniques like Lasso or Elastic Net are better suited for sparse models or when variable selection is a priority, which is not the project's focus.

3.3 Cross-Validation

Coupled with ridge regression, cross validation was utilized to assess the predictive power of different groups of variables (offensive, defensive, possession-based) on the plus/minus statistic. Cross-validation splits the data into training and testing sets multiple times, computing a model's predictive accuracy each time. In this case, the data was split into five subsets. In each iteration, one subset was kept as the test set and the other four were used to train the ridge regression model. The R-squared value was reported for each iteration to measure how well offensive, defensive, and possession-based metrics (like Corsi and Fenwick) predict plus/minus. This process ensures that the model's performance is consistent across different data partitions.

3.4 Mixed Model Effects

In order to separate individual contributions to plus/minus from team-level effects, a mixed-effects model was used. The mixed-effects model combines:

- **Fixed Effects**, which capture the influence of player-specific variables that directly relate to individual performance, such as Corsi For Percentage (CF%), Goals For Percentage (GF%), and other advanced metrics.
- **Random Effects**, which account for variability at the team level, recognizing that a player's plus/minus statistic can be influenced by the overall performance and style of their team. Including team-level random effects helps control for unobserved team factors that may affect each player similarly.

To apply this model, individual metrics that describe on-ice performance (e.g., CF%, GF%) were used as fixed effects to estimate each player's contribution to plus/minus. At the same time, team-level averages (e.g., TeamMeanCF%, TeamMeanGF%) were included as random effects. This approach distinguishes how much of a player's plus/minus statistic is attributable to their own performance versus the performance of their team. By using this mixed-effects approach, we can assess to what extent the plus/minus statistic reflects individual skill as opposed to team strength, thus helping clarify if plus/minus can serve as a reliable individual performance measure. The model was tested to ensure that the assumptions of linearity, independence, normality, homoscedasticity, and random sampling were not violated.

4 Results

The results of the correlation analysis and statistical modeling of plus-minus reveal several insights into how this statistic reflects both individual and team-level contributions.

4.1 Correlation Analysis

The initial analysis involved calculating the correlation coefficients between plus-minus and various performance metrics. The correlation coefficients are displayed in Table 2.

Metrics such as GF%, FF%, SF%, PDO, xGF%, and HDGF% show moderate to strong positive correlations with plus-minus (0.705, 0.554, 0.569, 0.546, 0.586, and 0.614, respectively). This result aligns with expectations, as many of these metrics are related to goal-scoring chances and shot control, which contribute directly to team scoring and subsequently affect plus-minus. Higher correlations among these metrics indicate a similar variation pattern to plus-minus, suggesting that they capture overlapping aspects of a player's offensive and defensive performance.

Table 2: Correlation Coefficients with Plus-Minus

Variable	Correlation Coefficient
GP	0.232
G	0.237
A	0.343
P	0.315
+/-	1.000
P/GP	0.278
EVG	0.239
EVP	0.336
PPG	0.187
PPP	0.227
SHG	0.177
SHP	0.283
OTG	0.162
GWG	0.306
S	0.256
S%	0.048
TOI	0.292
CF	0.369
CA	0.211
CF%	0.534
FF	0.374
FA	0.213
FF%	0.554
SF	0.376
SA	0.212
SF%	0.569
GF	0.424
GA	0.083
GF%	0.705
xGF	0.387
xGA	0.196
xGF%	0.586
SCF	0.387
SCA	0.192
SCF%	0.585
HDCF	0.399
HDCA	0.196
HDCF%	0.551
HDGF	0.442
HDGA	0.093
HDGF%	0.614
On-Ice SH%	0.324
On-Ice SV%	0.419
PDO	0.546

Table 3: Cross Validation Scores

Both	Offensive	Defensive
0.5696297	0.64692431	0.73265605
0.50125171	0.62667921	0.70862251
0.47617049	0.62263538	0.69467537
0.48944826	0.58484799	0.65315835
0.57586267	0.69145036	0.76239753

4.2 Ridge Regression Analysis of Offensive and Defensive Contributions

Ridge regression was performed on three different sets of variables: combined metrics, offensive only, and defensive only. The combined metrics included GF%, SF%, PDO, Takeaways, and SCF%, metrics that combine both offensive and defensive factors, similar to plus-minus.The offensive only metrics were CF, FF, SF, GF, SCF, HDCF, HDGF, GF%, G, A, P/GP, CF%, SCF%, Rush Attempts, and On-Ice SH%. The defensive only metrics were SA, GA, SCF, HDCA, HDGA, On-Ice SV%, Hits, Shots Blocked, and Penalties Drawn. The results can be seen in the table below.

The results in Table 3 reveal:

- **Combined Metrics**: A moderate cross-validation score around 0.53 suggests that these metrics explain a moderate amount of variance in plus-minus.

- **Offensive Metrics**: With an average score of 0.63, offensive metrics are predictive of plus-minus to a certain extent, indicating the offensive contributions captured by plus-minus.

- **Defensive Metrics**: The highest average score (around 0.71) suggests that defensive metrics have a more substantial impact on plus-minus, aligning with the fact that plus-minus also reflects defensive contributions.

Mixed Effects Model:

The results of the mixed effects model can be summarized in the table.

The coefficient for individual Corsi For percentage is a small, positive effect on plus/minus, though it is not statistically significant (p = 0.478). This suggests that while Corsi percent-

Table 4: Mixed Effects Model Results for Plus-Minus Prediction

Predictor	Coefficient	Significance
GF%	1.947	0.000
CF%	0.446	0.478
SCF%	-0.109	0.848
Team Mean CF%	-3.111	0.000
Team Mean GF%	0.846	0.000
Team SCF%	3.127	0.001

age (CF%) has a slight positive association with plus/minus, it may not have a substantial or reliable impact on explaining variance in plus/minus at the individual level. Similarly, Scoring Chances For percentage (SCF%) also does not have a significant relationship with plus/minus ($p = 0.848$).The positive and highly significant coefficient for individual GF% (p ¡ 0.000) indicates a strong and reliable positive relationship between goals-for percentage and plus/minus. Players with a higher individual GF% are likely to have a higher plus/minus score, suggesting that goal-scoring and offensive contribution are important for explaining plus/minus.

The coefficent for Team Mean CF% is negative, statistically significant coefficient (p ¡ 0.000) indicates that the team's mean CF% negatively impacts individual plus/minus. In teams with higher Corsi percentages, individual players might have lower plus/minus scores, possibly due to the distribution of possession-based contributions across the team. The positive coefficient of Team Mean GF%, significant at p ¡ 0.01, confirms that team-level GF% positively influences individual plus/minus. This suggests that players benefit from being on teams that are generally good at scoring, supporting the notion that plus/minus partially reflects team-level offensive strength. Finally, the positive and significant coefficient for Team SCF ($p=0.001$) indicates that teams generating more scoring chances positively affect individual plus/minus. This highlights the role of team offensive strength in influencing this statistic.

These results confirm that plus/minus is heavily influenced by team-level performance metrics (e.g., Team Mean GF% and Team SCF) and less so by individual possession metrics

Table 5: Cross Validation Scores

Corsi	Fenwick
0.93898296	0.95099819
0.90054897	0.93012992
0.93915158	0.9528202
0.92841222	0.95201592
0.9325304	0.95921829

like CF% and SCF%. The significant relationship between GF% and plus/minus supports the notion that offensive contributions and team success are key drivers of this statistic, reinforcing that plus/minus is a team-oriented measure rather than solely reflective of individual performance.

4.3 Plus-Minus vs Corsi vs Fenwick

Both Corsi and Fenwick were used as predictors in ridge regression with the variables GF%, SCF%, HDCF%, PDO, On-Ice SH%, On-Ice SV%, G, A. The results of the cross validation are shown in the table below

These values are notably better that the cross validation scores for plus-minus, regardless of which set of predictors. This indicates that both Corsi and Fenwick are significantly better at predicting player effectiveness than plus-minus.

5 Discussion

Like the plus-minus statistic, this project does not come without its limitations. This analysis relies on specific datasets from Natural Stat Trick and includes only a selection of individual and team-based metrics. Other relevant factors like zone entries, exits, or additional situational metrics might provide further insights into player contributions but were not available in this dataset. The model does not capture all nuances of game context, such as player fatigue, line changes, and shifts against specific opponents, which can all influence a player's plus/minus. The mixed-effects model approximates some of these factors with team-level random effects, but a more complex model might be required for a comprehensive understanding.

In this paper, the effectiveness of the plus-minus statistic was evaluated to see if it is an effective predictor of player contribution in the National Hockey League. The plus minus does have some value as a predictor of offensive and defensive metrics. However, other advanced metrics like Corsi and Fenwick perform much better. The plus-minus is also heavily influenced by a player's team, making it less valuable as an individual evaluator of player effectiveness. Overall, the NHL should consider phasing out the use of the plus-minus statistic and consider more advanced metrics such as Corsi or Fenwick.

In order for fans, players, and sports management to have a better understanding of what make a player "effective", you have to summarize and quantify a large number of variables into one condensed number. That is, understandably, a tall order. While the plus-minus lays the ground-work for getting a better understanding of player effectiveness, there is more at play. General managers of sports teams often want to make the best team possible, getting players that will have a positive impact. Often times, people put a lot of emphasis on direct offensive output, such as points, to determine a players contribution. But there is so much more to consider, such as defensive skills, generating chances to score, etc, while still accounting for team effects. Many have tried to do this, but lack the interpretability and simplicity of the plus-minus. While plus-minus may not be the solution, the next step in the world of sports statistics is to find a way to quantify the concept of "player effectiveness" in a way that includes the simplicity of the plus-minus without its controversial nature.

References

L. M. Hvattum (2019). "A comprehensive review of plus-minus ratings for evaluating individual players in team sports." *International Journal of Computer Science in Sport*

18.

B. Macdonald (2011). "A regression-based adjusted plus-minus statistic for nhl players." *Journal of Quantitative Analysis in Sports* **7**.

B. Macdonald (2012). "Adjusted plus-minus for nhl players using ridge regression with goals, shots, fenwick, and corsi." *Journal of Quantitative Analysis in Sports* **8**.

National Hockey League (2023). "Nhl official website." Accessed: 2024-10-10.

Natural Stat Trick (2023). "Natural stat trick hockey statistics." Accessed: 2024-10-10.