

Scouting Automated Ratings Analyzing Habits (SARAH): A Statistical Methodology for Scouting and Player Development

Noah Chaikof¹, Katia Clément-Heydra², Carleen Markey³, Mikael Nahabedian²,
Adam Pilotte⁵, and Mairead Shaw²

¹ Ottawa 67's, Ottawa, Ontario, K1S 3W7, Canada

² McGill Marlets Hockey, Montreal, Quebec, H2W 1S4, Canada

³ Carnegie Mellon University, Pittsburgh, PA, 15213, USA

⁴ Saint Laurent Patriotes, Saint-Laurent, Quebec, H4L 3X7, Canada

⁵ McGill Redbirds Hockey, Montreal, Quebec, H2W 1S4, Canada

Abstract. The project serves a two-fold purpose: to reduce the time that scouts and coaches spend trying to identify what players have foundational on-ice habits, and to streamline the process of evaluating the developmental progress of a players' habits. Essentially what we did was first look at the various national women's hockey teams and identify the set of "habits" a player regularly executes (i.e., edgework, catching the puck in the hip pocket, pass placement, etc). Combining the dataset of players' habits with a set of players' microstats (entries via pass/stickhandling, exits via stickhandling/pass, accurate/inaccurate passes, etc.), we developed a random forest classification model to accurately predict if a player possesses a certain habit based on their set of microstats. We also used random forest regression on our data to see how habits impacted each specific microstat. Combining this with an estimate of how frequently players used each habit, we created a Player Development Matrix for a player's habits based entirely on their microstats. To help coaches, scouts, and anyone else access & use these tools, we've also created an interactive visualization for these models using our training dataset of national women's hockey teams in the last Worlds and Olympics.

Keywords: Player Development, Analytics, Coaching, Scouting.

1 Introduction

This paper provides a comprehensive overview of a newly designed player-evaluation framework for women skaters at the 2021 IIHF tournament and 2022 Olympic games using a 'habit-tracking' system. Building on the work of Bryce Chevallier[1], Jack Han[2], and Darryl Belfry[3], the goal of this study is to explore the validity of using micro habit-tracking as a supportive scouting technique (player-ranking system) and utilize habit-tracking as a foundation to uncover the highest priority areas for player development staff to hone in on meaningful skill improvements in their players or clients.

Our study demonstrates a statistically significant ability to accurately link a player's "habit-score" to statistical events for scouting purposes (micro stats such as zone exits, zone entries, type of pass,...), and uncovers 'habits-of-focus' for player development staff based on a player's advanced stats. Lastly, the study explores a habit-improvement framework using a Player Development Matrix[4] to analyze the habits of highest importance for development staff relative to the rest of that player's skill set.

The core technique of this study is the novel development of a complete list of habits and categorization of those habits into 7 different skill set areas. A comprehensive tracking model was used to obtain a baseline habit-score of all players, this data (combined with enriched data from InStat[5]) was used as the basis for the two models and the development matrix outlined below.

1.1 Motivation

The aim of this study is to offer a quantitative tool to both player evaluators (coaches and scouts) and player development staff as they are challenged with examining/improving the skill sets of large groups of players.

Scouting. The motivation for the study is to attempt to add a complimentary quantitative approach to traditional scouting and player evaluation analysis. Under the current model of scouting across hockey leagues, scouts are faced with a tremendous challenge of ranking players across broad skill categories, as evidenced by the sheer number of NHL draft rankings alone [6,7,8]. It is a significant challenge to rank a player's skill set (i.e. passing) on a scale from 1-10 and subsequently justify why a player's rating in that category will vary so significantly across scouts watching the same player.

The goal of the study with the creation of a binary habit-tracking system (said habit positively impacts a player's game or not) will enable certain player evaluators to bring a more quantitative approach to their rankings and give teams an edge in their scouting process.

Player Development. Similarly, player development staff are facing a tremendous challenge in trying to prioritize their limited time with each player and design a personalized skill development plan to drive improvement in their game[9]. The binary tracking system will allow player development staff to hone in on more exact skill gaps and work directly on improving those habits. Additionally, as a larger dataset of player habit-tracking is built over time, player development coaches can uncover which groups of habits are most critical to player success at different points in their careers, and how player habits may evolve over time.

2 Methodology

2.1 Statistical Methodology

The core statistical methodology/tracking technique used in this study is a novel binary-habit evaluation model developed below. In lay terms, the contributors of this study

developed a list of habits (edgework, neutral zone angling etc.) and categorized those habits into different skillset areas (skating, puck reception, stickhandling, physicality, play away from the puck, passing & shooting) in an attempt to break down a player's game into micro attributes.

The selection of these various habits cover a broad spectrum of skills that may be displayed over the course of a hockey game (both offensive and defensive) but are highly specific in nature. Each habit was selected only if it can be measured clearly in the tracking process and the presence of that habit in a player's game is associated with driving impactful results during their time on ice. The following table summarizes the different skill sets and habits identified as part of the project. Refer to Appendix for a brief description of each habit identified as part of this project.

Table 1. Skill Sets & Habits

Skating	Puck Reception	Stickhandling	Physicality	Play Away from the Puck	Passing	Shooting
Edgework Outside	Catching Puck in Hip Pocket	Loading Puck to Hip Pocket	Initiating Contact	Shoulder Checks	Slip Passes	Coordination
Backwards Skating	Dynamic Catch	Underhandling of Puck	Puck Protection with Body	NZ Angling	Leveraging & Creating Seams	Weighttransfer
Stride Recovery	Getting Off the Boards	Handedness Versatility	Fitness Level	Unassisted Stops	Pass Placement	Tip
Skating Mechanics		Deception with Puck		Jumping in Shot Lanes	Vision	
Crossovers				Awareness Without Puck		
Shouldering Speed				Net Front Presence		
Feet in Motion						

2.2 Tracking Technique

In order to build a sample with over 7500 observations to train the models on a period per period basis, the tracking technique used for the study relied on observing a minimum of three periods of a player's ice-time and assigning a binary score for each of the habits underscored above. The sample time-on-ice from the three periods were each tracked from three different games to adjust for strength of opponent and variances in a player's effort and effectiveness from game to game. In total, the data set included habits for 262 players from 12 different teams.

Based on whether a player demonstrated that habit more often than not when given the opportunity to do so during their observed ice-time, they were given a score of '1' (habit positively impacting a player's game) or '0' (habit not positively impacting a player's game). This resulted in a total unweighted score out of 30 for each roster player based on the number of habits they possessed during the sample period.

Table 2. Skill Sets & Habits

Rank	Name	Team	Position	Score (on 30)
1	Marie-Philip Poulin	Canada	F	30
2	Jenni Hiirikoski	Finland	D	29
T-3	Kendall Coyne Schofield	USA	F	28
T-3	Jocelyne Larocque	Canada	D	28
T-3	Ronja Savolainen	Finland	D	28
T-6	Mélodie Daoust	Canada	F	27
T-6	Brianna Decker	USA	F	27
T-6	Sarah Fillier	Canada	F	27
T-6	Rebecca Johnston	Canada	F	27
T-10	Michelle Karvinen	Finland	F	26
T-10	Claire Thompson	Canada	D	26

2.3 Modelling

SARAH 1 - Identifying events or advanced metrics expected based on player habits. The first model used in this project (random forest regression model[10]) was created to identify the different events or advanced statistics that one would expect to see a player possess based on whether they have a given habit. The random forest used in SARAH 1 and 2 consists of generating a number of decision trees, each of which are only given a random part of the dataset. Each decision tree then decides how each independent variable affects the dependent variable based on the random subset of the data it sees and makes predictions for each player in the entire dataset based on their independent variable data. The predictions from all the trees are then averaged to create one prediction for each player.

This model utilizes the event specific data from InStat (i.e. controlled entries and inner slot shots etc.)[5] for each player, with the intended goal of **finding which habits yield results in specific advance statistics or event categories.** *Subconsciously, scouts complete this same exercise when evaluating a player's effectiveness and instincts. For example, one would expect a player who exhibits linear crossovers and keeps their feet in motion following a puck catch, to complete successful controlled entries at a higher rate than a player without these habits.* In this model, the independent variables are the habits (variables X), with event data being treated as the dependent variable (variable y).

SARAH 1 included 17 separate sub-models, with each of the sub-models representing one of the 17 different event types adjusted per 60 minutes that were observed in the study. This is also referred to as “event-based advanced stats” later in the paper. The events included in the model are the following:

Table 3. Event Types (Microstats)

Accurate passes	Puck battles lost
Breakouts via pass	Puck battles won
Breakouts via stickhandling	Puck losses
Dump ins	Puck recoveries
Dump outs	Shots blocking
Entries via pass	CF
Entries via stickhandling	CA
Inaccurate passes	Shots
Passes to slot	

Significance Threshold for Linking Habit to Event and Selection Process. A critical component of this event-to-habit linking methodology is to identify habits that meaningfully impact the event/advanced statistical metrics. In this study - any habit with an importance above the 0.0325 threshold is considered having a strong influence on the likelihood of a player-habit meaningfully impacting that statistic or advanced stat category. Below is an example of the 10 main habits that meet the threshold for the event pertaining to “puck battles won”.

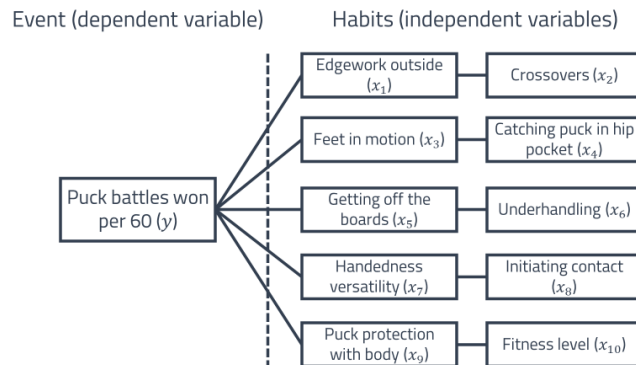


Fig. 1. An example of the 10 main habits that meet the threshold for the event pertaining to “puck battles won”

The considerations used in selecting the threshold of 0.0325 include;

1. Finding a threshold figure that aligns with the knowledge of the contributors to the report and removes results that do not align with hockey-logic (i.e. neutral zone angling should have no relation to shots statistic).
2. We wanted to select a threshold that ensured each habit would be meaningfully connected to a minimum of five events. If this was not the case - the habit was removed for lack of importance to the model

Weighted Average Consideration for Event Statistics. Lastly, a weighted-average accounting for both the number of events completed and time-on-ice in the period was relied on in the SARAH 1 analysis.

This was done to adjust for problematic tracking outcomes when a player may have a high volume of events on a low base of ice-time (i.e. 4 successful completed passes in 3 minutes of ice-time in a given period) that would result in non-representative per/60 minute data. Therefore, greater weight was assigned to events that occurred over a larger period of ice time than in smaller sample sizes.

SARAH 2 - Predicting the probability of a habit meaningfully impacting a player's game. After establishing the impactful event-habit relationships in the first set of models, SARAH 2 reverses the variables and attempts to make a prediction about the probability of a habit successfully being completed by a given player.

This second set of models serves a dual purpose. First, it provides scouts with a baseline to precisely quantify habit evaluation. In other words, if the event-based advanced stats are available, this model can be seen as an automated habit-evaluation tool. However, SARAH 2 can also be used in conjunction with video scouting, allowing player evaluators to compare the statistical results versus their personal assessment of habits for different skaters.

Secondly, by precisely evaluating the success probability of various habits for skaters through the steps described below, this set of models enables skills coaches to uncover development opportunities for players and measure their progress over time in a systematic way.

The starting point of SARAH 2 is the meaningful event-habit relationships identified as part of SARAH 1, based on the 0.0325 threshold discussed in the previous section. However, flipping the variables in the case of SARAH 2 allows us to statistically estimate the probability of successful habit completion based upon a set of event-based advanced statistics for a given player.

For instance, when attempting to predict the success probability of the “outside edgework” habit, the first step is to highlight that this habit is strongly impacting the following 8 event-based advanced stats in SARAH 1. After identifying these strong event-habit relationships, the idea of SARAH 2 is to use these events to predict the successful completion of the “outside edgework” habit, as exemplified below:

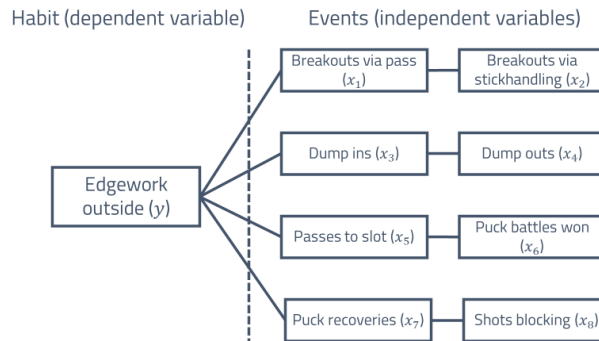


Fig. 2. An example of the 8 main events that meet the threshold for the “outside edgework” habit

In this example, as part of SARAH 1, we had identified that the “puck battles won-outside edgework” event-habit relationship was meaningful. For this reason, as part of SARAH 2, the “puck battles won” statistic is incorporated, among other events, as one of the predictors of the “outside edgework” habit.

While not visualized in the previous section, similarly for the 7 other events listed above (e.g., breakouts via pass, puck recoveries,...), it was established in SARAH 1 that the “outside edgework” habit is also meaningfully driving part the results for these other events. As such, in addition to “puck battles won”, these 7 other event-based advanced stats are also incorporated as predictors in SARAH 2 for this specific habit.

In short, SARAH 2 is built as a *random forest classification model* [11] in which the event-based advanced statistics are the independent variables (X variables) and the habits are the dependent variable (y variable).

SARAH 2 included 30 separate sub-models, with each of the sub-models representing one of the 30 different habits that were tracked in the study.

The outcome of SARAH 2 is that for each player, all of the habits measured will be assigned a value between 0 and 1 (considered a percentage probability) that a respective habit yields positive results while on ice.

For instance, in the case of Laura Stacey, a Canadian forward who initiates a high volume of controlled exits, dump entries and puck recoveries, the probability that she successfully completes the “outside edgework” habit is around 80%.

It is important to note that the outcome of this model is only identifying the success probability of a habit completion (i.e. a 0.8 score is not necessarily better than a 0.65), it is only significant in that it creates a probability based prediction on which habits are likely strengths and weaknesses for a given player.

For this random forest model, any habit with a score above 0.5 implies that when a player has the opportunity to exhibit this habit, they are more likely to complete this micro-ability well. As we had established that the event-habit relationships are meaningful, the successful completion of said habit is inherently related to driving impactful results on the ice.

SARAH 2 Testing - Hyper-Parameter Tuning. SARAH 2 went through hyper-parameter tuning in order to optimize the number of trees to use for probabilistic prediction of habits. The process described below yielded an accuracy score 82%.

For this hyper-parameter tuning, part of the data was used as the test set and was separated from the training data. The test set was utilized to compare predictions to tracked habits.

The resulting closeness of the predicted outcomes made by the training data set compared to the actual test-data enables us to be confident in the prediction made by our model.

3 Outcome - Player Matrices of Success Probability and Frequency

The outcome of this study is that each player will have their habits mapped out in a 2x2 matrix based on the amount of times that habit is exhibited (driven by event-data) and the success probability expected when that habit is completed (probabilistic figure uncovered in SARAH 2).

3.1 Frequency and Success Probability – Measurement Techniques

Frequency. This number is driven by the number of times a player exhibited that habit - which is uncovered through their time adjusted event data.

Example - A player with a significant volume controlled entry via pass or stickhandling (after establishing the connection between those events and the efficient use of crossovers as a habit) allows us to conclude that crossovers are frequently utilized by this player.

We can predict that a player will utilize crossovers habit more often because of this higher volume of event data.

Success Probability. The probabilistic figure between 0-1 discussed in SARAH 2 that provides a percentage probability that a player will complete that habit successfully when the opportunity presents itself, which is inherently related to driving impactful results on the ice.

3.2 Matrix Deep Dive - Quadrant Breakdown (Player Development Matrix)

As introduced in the public sphere by Jack Han in his newsletter[4], the matrix presented below has four quadrants, which is designed to enable player-development staff and scouts to identify the habits of strength and weakness for players. In its current form, skills on the the Player Development Matrix are estimated qualitatively and plotted on the chart. To instead quantitatively determine where skills should go on this matrix, we plot the calculated frequency against the success probability for each player. An example of this novel quantitative iteration of matrix is included below.

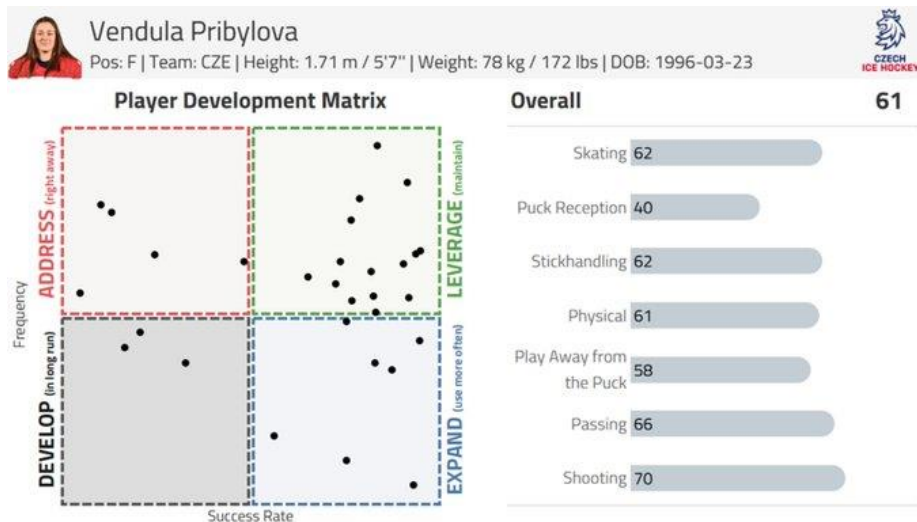


Fig. 3. The development matrix of Vendula Pribylova. Her data and development matrix has been included in this publication with her permission. The development matrix is used with permission from its creator, Jack Han.

A breakdown of the interpretation of the four quadrants is provided below:

Green Quadrant (LEVERAGE) - High success probability and high frequency; a player is expected to use this habit quite frequently and when completed it is done well (these are the skills that enable them to drive strong play).

Blue Quadrant (EXPAND) - High success probability and low frequency; these are habits completed well when attempted, but player development staff should encourage these habits to occur more often because they are being underutilized.

Red Quadrant (ADDRESS) - High frequency and low success rate; highest priority items to fix for player development given it occurs often but is done very poorly (high failure rates and likely holding the player back).

Black Quadrant (DEVELOP) - Low frequency and low success probability; staff should target long run improvement for these habits, the player does not have the opportunity to complete these habits often, but they are not executed well when the situation presents itself. This should be the lowest priority items for player development staff and may be unimportant to a player's archetype (i.e., grinder does not need to exhibit x skill).

Each matrix is relative to only that player's broader skill set. For example, Marie-Philip Poulin's red quadrant habits may still be elite in comparison to 95%+ of her opponents

but it is weak relative to the rest of her habit score. The reason this matrix was created on a relative basis was to allow player development staff to focus on a personalized plan for each player, rather than the most elite players having almost no areas of improvement.

3.3 Skill Set Scores Methodology

In order to estimate the score on different skill sets, a weighted average calculation was used to incorporate both the effects of success probability and frequency of habits. As initially outlined in the tracking methodology, 7 different skill sets were determined with the goal of linking statistical techniques to more traditional scouting techniques (video analysis) containing the following habits.

As such, weighting was applied to the frequency of different habits in each skill set to calculate the average success probability for the skill set.

4 Conclusion and Future Works

In short, this paper introduces a new approach to linking traditional scouting methods to advanced and micro stats in hockey through an automated scouting tool that can be used to improve the quantitative evaluation and player development processes of organizations. In terms of future work, three possible model expansions that could be explored are the following:

- Developing a multi classification model combined with a non-binary habit tracking system would allow the incorporation positive impact (or lack thereof) of a habit to different degrees. For instance, a player that is developing a habit, while not fully mastering it could receive a score of 0.5 for said habit instead of simply limiting the choices to binary options (0 or 1).
- The current model could also be extended to identify player archetypes at the micro-habit level in order to characterize the strengths and the weaknesses of different groups of players more precisely.
- Finally, the idea of skill stacking could be incorporated into the modeling process in the form of interactions between the different habits and multilevel targets in SARAH 1 and 2 respectively.

5 Appendix

The code for this project can be found at <https://github.com/mnahabedian1/WHKY-Player-Development-Project>. The interactive player development matrix tool can be found at <https://public.tableau.com/app/profile/mikael.nahabedian1483/viz/PlayerDevelopmentProject-PublicVersion/Dashboard32>.

Below are the definitions for the habits included in Table 1.

5.1 Skating

Edgework Outside – Ability to access outside edges with ease (usually with a bow-legged basic posture).

Backwards Skating – Focus on pivot (without crossing feet) + stride mechanics yielding grip & smoothness.

Stride Recovery – Back leg just under full extension and recovers underneath the body to allow for recovery in the next stride.

Skating Mechanics – Knee flexion to generate power on each stride. Joints are stacked (shoulders, knees and toe caps form a line).

Crossovers – Use of crossovers when carrying the puck to change direction or build speed (every 4 to 5 strides).

Shouldering Speed – Movement patterns allowing smooth transition during changes of direction or to move from one play to the next.

Feet in motion – Following cutbacks or puck receptions, ability to create separation with the opponent.

5.2 Puck Reception

Catching puck in Hip Pocket – Ability to receive the puck on the side of the body (let it through body).

Dynamic Catch – Feet position (open) + catch in a weight shift or crossover.

Getting off the boards – Ability to catch the puck along the boards in a favourable posture to get away.

5.3 Stickhandling

Loading Puck to Hip Pocket – Ability to load the puck on the side of the body (good attack position).

Underhandling of Puck – Handling the puck efficiently without unnecessary stick motions.

Handedness Versatility – Being able to play the puck both on the forehand and backhand.

Deception w/ puck – Able to pull in players with the puck or give the illusion of making a specific play.

5.4 Physical

Initiating Contact – In board battles, willingness to initiate contact with the opponent to win the puck.

Puck Protection with Body – Ability to use body as a shield between puck and opponent.

Fitness Level – Overall ability to keep up with the pace of the game (& have reasonable shift lengths).

5.5 Play Away from Puck

Shoulder Checks – Making meaningful checks behind the play before retrieving the puck/in the DZ.

NZ Angling – Close space to ensure that threats are angled and neutralized in the NZ.

Unassisted Stops – Getting out of structure and swiftly killing plays early without opening seams in DZ.

Jumping in Shot Lanes – Purposefully & voluntarily jumping in front of shots in DZ.

Awareness without puck – Reading plays correctly yet understanding the purpose of playing inside structure.

Net Front Presence – Box out + goalie presence in DZ and OZ respectively.

5.6 Passing

Slip Passes – Ability to identify seams under or above the stick of opponents.

Leveraging & creating seams – Ability to create seams through movement and accurately leverage them.

Pass Placement – Ability to provide good pucks to teammates.

Vision – Ability to identify the best passing option.

5.7 Shooting

Coordination – Feet placement (front towards net) + application of downward force for accuracy/power.

Weight transfer – Transfer of weight to generate velocity on the shot.

Tip – Ability to tip shots/generate shots that are tip-able (usually low and through the defense).

References

1. Chevallier, B, Architecte Hockey, <https://architecte-hockey.com/methods/>
2. Han, J., The Hockey Tactics Newsletter, <https://jhanhky.substack.com/>
3. Belfry, D., Belfry Hockey, <https://belfryhockey.com/>
4. Han, J., <https://jhanhky.substack.com/p/how-to-ruin-a-player?s=r>
5. InStat Hockey, <https://instatsport.com/hockey>
6. NHL Central Scouting, <http://www.nhl.com/ice/draftprospectbrowse.htm>
7. Pronman, C. <https://theathletic.com/3297887/2022/05/10/nhl-draft-lottery-ranking/>
8. Mckeen's Hockey, <https://www.mckeenshockey.com/draft/>
9. Han, J. <https://jhanhky.substack.com/p/upgrading-the-player-development>
10. Scikit-learn, <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html>
11. Scikit-learn, <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>