NC Metric

Ryan Poché
Florida State University Women's Soccer

ABSTRACT

This document provides the rationale and use of a different wholistic metric to measure match performance in football. It will discuss reasons for creation, methodology, examples of how it can be used to identify players' contributions to their teams, and areas for future expansion.

Keywords: Women's Football, Women's Soccer, Football Analytics, Soccer Analytics

I. INTRODUCTION AND RATIONALE

While researching metrics for other professional sports, I always found myself searching for the all-in-one statistics to help quantify a player's value. Fortunately, almost all other sports have this, such as basketball's BPM/VORP, baseball's WAR, and American football's AV. Unfortunately, despite its popularity, football has never quite produced a sibling for these measures. New efforts such as OBV and xTA have pushed closer to the goal in this space but still lack some components in my opinion such as quantifying value added through additional types of actions.

This new attempt, Net Contribution (NC) seeks to create a way to measure player value added in the pursuit of scoring, as well as not conceding, and do so in a way that does not require 360 tracking data to be effective. It will create an interpretable product for those looking to quantify the value of their players and other players in recruitment, and not exclude those who do not have financial access to 360 tracking data, similar to the way BPM is accessible for basketball scouting.

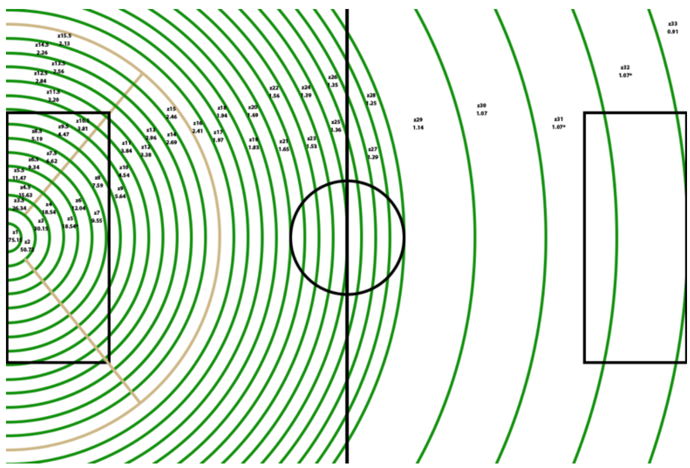
II. COMPONENTS

A. **Framework and Key Differences**

The first step in creating a metric quantifying scoring probability added is to assign values to the places on the field that the actions occur. For this, a model was built using over 12,000,000 open source StatsBomb data events to determine what percentage of possessions that have an action in that zone result in a goal, similar to the framework built in xT. The key differences between SPA and other calculations are:

1. NC's zones are curved and 2.5 meters long, corresponding to the distance that location is from the opposing goal, while xT's zones are rectangular. The values of these zones and the visual representation of this will follow this paragraph (Figures 1 and 2). Each attacking zone is split into a central and wide zone to gain further accuracy of the value of each section.
2. NC builds its outcome variable based on xG, not actual goals. This allows for increased accuracy, as the finishing ability of the forwards in the sample is not important when seeking to evaluate the value of the position. The quality of the chances created from the zone will lead to better outcomes than the chances that happened to be converted by the teams in the sample.
3. NC does not have an action limit to disqualify the goal, whereas other metrics such as xT only count the goal if it is under 5 further actions from the initial action. Practically, a team does not care whether a big chance comes 5 passes after a tackle rather than 7, they care about the chance itself, so the model will not make this distinction either.

Figure#1



Figure#2

Zone	Distance From Goal	Scoring Probability	Zone	Distance From Goal	Scoring Probability
1	0-2.5	75.15%	2	2.5-5	50.72%
3	5-7.5	30.15%	3.5	5-7.5	26.34%
4	7.5-10	18.54%	4.5	7.5-10	15.63%
5	10-12.5	18.54%*	5.5	10-12.5	11.47%
6	12.5-15	12.04%	6.5	12.5-15	9.34%
7	15-17.5	9.55%	7.5	15-17.5	6.62%
8	17.5-20	7.59%	8.5	17.5-20	5.19%
9	20-22.5	5.64%	9.5	20-22.5	4.47%
10	22.5-25	4.54%	10.5	22.5-25	3.81%
11	25-27.5	3.84%	11.5	25-27.5	3.20%
12	27.5-30	3.38%	12.5	27.5-30	2.84%
13	30-32.5	2.96%	13.5	30-32.5	2.56%
14	32.5-35	2.69%	14.5	32.5-35	2.26%
15	35-37.5	2.46%	15.5	35-37.5	2.13%
16	37.5-40	2.41%	17	40-42.5	1.97%
18	42.5-45	1.94%	19	45-47.5	1.83%
20	47.5-50	1.69%	21	50-52.5	1.65%
22	52.5-55	1.56%	23	55-57.5	1.53%
24	57.5-60	1.39%	25	60-62.5	1.36%
26	62.5-65	1.35%	27	65-67.5	1.29%
28	67.5-70	1.25%	29	70-82.5	1.14%
30	82.5-95	1.07%	31	95-107.5	1.07%
32	107.5-120	1.07%*	33	120+	0.91%

***on two occasions a zone's value was marginally higher than the one in front of it, and the value of that zone was capped at the value of the front zone.**

B. Selecting Ingredients

Now that the values of possessing the ball in our zones has been established, the actions that will be considered and how they will be valued in said zones must now be decided. Intuitively, when creating an all-encompassing metric, it follows that as many types of events as possible should be included to capture the most data. The 20 types of events that are included in the model and their values are laid out in Figure 3, and the rationale for their inclusion and calculation will follow. The calculations will be given using the decimal SP (or Scoring Probability) of the zone it takes place in as described earlier.

Figure#3

Action	Calculation	Action	Calculation
Complete Pass	SPend-SPstart+0.00375	Aerial Duel Lost	-0.5*SP-0.5*oppSP
Incomplete Pass	((1/4 * SPend)-0.5*SPstart)-oppSPend	Missed Tackle	-(0.5*SP)-oppSP
Progressive Carry	SPend-SPstart	Foul Won	SP
Tackle	oppSP+2*SP	Foul Against	-oppSP
Interception	oppSP+2*SP	Penalty Won	0.78
Aerial Duel Won	0.5*oppSP+0.5*SP	Penalty Against	-0.78
Shot	0.5*xG+(xGOT-xG)	Score Penalty	0.22
Successful Take On	oppSP+SP	Miss Penalty	-0.78
Turnover	-SP-oppSP	Goal Line Clearance	1
Recovery	oppSP+SP	Last Man Tackle	0.65

C. Inclusion and Calculation Rationale

1. Complete Pass
    a. The SPA of a completed pass is how much more valuable the ending zone is than the starting zone. An additional boost is given relative to the chance of not completing the pass in an average zone to value ball retention.
2. Incomplete Pass
    a. Similar to completed pass but negative, but weighted in a fashion that passes into incredibly valuable areas that don't come off are still seen as a net neutral or positive. Also penalized for the value of the opponent's possession at the end of the pass.
3. Progressive Carry
    a. How much more valuable the zone that the ball was carried to is than the zone the carry started from.
4. Tackle
    a. Credited for the opponent's zone that they no longer occupy and your own. Extra weight due to the nature of defender positioning after a takeaway as opposed to in set defence.
5. Interception
    a. Credited for the opponent's zone that they no longer occupy and your own. Extra weight due to the nature of defender positioning after a takeaway as opposed to in set defence.
6. Aerial Duel Won
    a. Half of their and their opponent's zone, different from tackles and interceptions because they both imply possession retention whereas aerial duels won do not.
7. Shot
    a. Credits the player for both the finishing quality of their shot and getting to the position to create the chance in the first place. This excludes penalties which will be addressed in this section.
8. Successful Take On
    a. Credited for retaining possession in the zone they are in and for the zone that the opponent would have possessed in had the ball carrier been tackled.
9. Turnover
    a. Penalized for both the zone they lost the ball in and the zone the opponent now possesses in.
10. Aerial Duel Lost
    a. Penalized half of their and their opponent's zone, opposite of aerial duel won.
11. Missed Tackle
    a. Penalized for both the zone the opponent still possesses the ball in and partially for the zone that they would have possessed in had they made the tackle.
12. Foul Won
    a. Credited for the zone the foul occurred in.
13. Foul Against
    a. Penalized for the zone the foul occurred in.
14. Penalty Won
    a. Credited with the usual following xG of a penalty kick.
15. Penalty Against
    a. Penalized for the usual following xG of a penalty kick.
16. Score Penalty

a. Credited with the value added by scoring a penalty using the xG of the penalty (100%-78%).

17. Miss Penalty
    a. Penalized for the usual xG of a penalty kick.
18. Goal Line Clearance
    a. Credited with an entire positive goal, if the player had not put themselves in that position then the team would have conceded.
19. Recovery
    a. Credited with the end of the opponent's possession and the start of theirs.
20. Last Man Tackle
    a. Credited with the xG of a 1v1 shot on the keeper from the penalty spot.

III. TESTING EXAMPLE 1

To use this metric practically, I ran it throughout the season on data I collected from the Florida State Soccer team. We are in a particularly good spot to see data like this when compared to what professional teams are using, as after each season our seniors and a select group of underclassmen that get great offers go to professional teams. Looking at their decisions both gives us an insight into their rationale and can let us predict which of these signings will likely be worth this investment in the short-term.

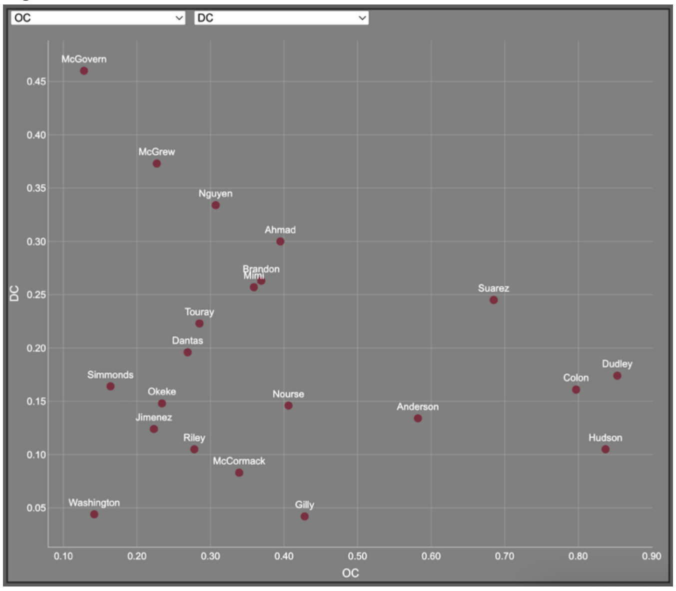
The four players that the NWSL picked from this team to give lucrative offers before they graduated were Dudley, Suarez, McCormack, and Mimi, a striker, two midfielders, and a fullback respectively.

Looking at the chart that divides Net Contribution into Offensive Contribution on the X axis and Defensive Contribution on the Y axis, Dudley and Suarez clearly stand out as players with the data of a professional prospect. Dudley is our most valuable attacker and is contributing on defense at a relatively average rate compared to her team. Suarez is the most 'balanced' contributor, as the only player to be in the top 7 on both sides of the ball.

Mimi is another player that the model is clearly identifying as being valuable to the team, even if the raw number is not as high as the two forward players mentioned before her. I believe this is not necessarily a flaw in the model but the reality of the sport, that finishing is the most important predictor of goals and that defenders will be 'undervalued' as they do not get into these positions as frequently as forwards and midfielders.

McCormack is an interesting one, as she plays the same position as Suarez, Nourse, Ahmad, and Touray, and is garnering more national team interest than all four and apart from possibly Suarez also garnering the best professional offers. Of the five players she is having the 4th most production offensively, marginally above the one below her and less than half of Suarez's production. Defensively she has the least production of the five and is marginally above half the defensive contribution of the player in 4th. To add on to this, she is the oldest of the five as she is a junior and the other four are two sophomores and two freshmen. It will be very interesting over the next two years to see the four players that went pro and the five players in this sample, and be able to apply hindsight to if the modeling was seeing something or not.

Figure#4



IV. TESTING EXAMPLE 2

To test this metric's effectiveness in identifying players that are helping their team succeed, I ran the model on all the publicly available seasons of StatsBomb data. This includes a wide range of men's seasons from various leagues, and select WSL and NWSL seasons for women's. I separated the stats into categories to more easily identify what parts of a player's game the metric are identifying as elite, poor, or anywhere in between. The final number was then standardized per 100 possessions in this case, as the seasons had vastly different match counts. In a sample with consistent season sizes, there is an argument against doing this as a player's involvement in itself can be valuable and should be included in analyzing their contribution. The top players for both women's and men's seasons are shown at the bottom of this page in Figures 4 and 5.

Thankfully for the sake of the metric, the players pulled out of the sample of 1000s of footballers are either among the best players in the world or were among the best in the world at the time. Among women's players the top 6 seasons all came from top performers in the WSL, and come from a mix of strikers, wingers, and midfielders. While no defenders feature, only one player has over half their value creation from shots, showing that the data does not seem to be overly reliant on goals. This is especially relevant even in modern organizations and punditry where raw goalscoring output is seen as the end all be all for a player's value.

On the men's side, 4 Lionel Messi seasons are joined by Gareth Bale's best Real Madrid season and Kylian Mbappe's best season for PSG. The breakdown is generally useful in connecting to what we know the players to be good at, as Messi is providing a huge amount of value through carries, passes, and finishing, but he is not renowned for his defensive work rate or for competing in the air. Overall, I believe these results from the women's and men's games are positive in the ability of this metric to not only quantify how valuable a player is to their team's positive scoring output but also the ways in which they are providing this value.

*The modeling changed marginally between July and December 2025, and will inflate the current numbers in Example #1 relative to Example #2.

V. CONCLUSION

The aim of the NC metric was to create a wholistic stat that would be able to identify players that are contributing to their team scoring and not conceding goals, and on the whole I believe it was successful at this. There are certainly ways that it could be improved through future research and if given access to tracking data, but given the results of the testing sets, the wide ranging events that are included in the formula, and the rationale behind their inclusion and importance, it was successful in its aim.

VI. AREAS FOR EXPANSION

A. New Inclusions via Tracking Data

- Measures of space: Having measures of both how much space a player is receiving the ball in and how much space they are playing teammates into would be very beneficial. This would be especially useful offensively in my opinion as it would identify players that are not playing their teammates into pressure and credit them for this.
- Pressure: Having if a player is under pressure or not would inform their ability to still play-make under said pressure. This would be especially important for midfielders, as players that can maintain a high level under pressure from a defender are incredibly valuable.
- Defensive Positioning: This is certainly the hardest addition to quantify, and is largely subjective in many cases, but with a large enough data set, it would be possible to determine the general range of locations that position is taking up relative to the ball, and record if a player is in that range or if they are out of position. Again, due to subjectivity and tactical differences this would be incredibly difficult but could certainly be attempted with a large enough data set.

B. Women's Football Specific Additions

- Retraining the zones on only women's matches: While the differences would likely not be staggering, for the best accuracy it would always be best to use only use women's data when creating metrics for the women's game. StatsBomb does not have enough open-source data to do this right now, but hopefully they have more seasons tagged in the future to have the ability to retrain the zones.
- Testing the formula on more women's seasons: Currently, only select WSL and NWSL seasons are tagged, giving a rather small sample size in the example. Ideally, more (and more recent) seasons of these leagues and some Liga F, Women's Bundesliga and Première Ligue seasons would be tagged to see our player identification in use over a larger dataset.

VII. REFERENCES

"About Box plus/Minus (BPM)." Basketball, Feb. 2020, www.basketball-reference.com/about/bpm2.html.

Singh, Karun. "Introducing Expected Threat (XT)." Karun Singh, 24 Dec. 2018, karun.in/blog/expected-threat.html.

"Using Hudl Statsbomb Free Data in R." Hudl, Hudl, 25 June 2025, www.hudl.com/blog/using-hudl-statsbomb-free-data-in-r.

VIII. OVERSIZED FIGURES

Figure #4

	player.name	season	totalspa	SPApasses	SPACarries	SPAShots	SPATakeons	SPATurnovers	SPATacklesand	SPAAerialduels	SPAFouls	numaction	spaper100actions
1	Bethany Mead	WSL 2018/19	17.771232	8.507375	4.2758	3.96403183	1.9086	−2.341	0.646	−0.20865	1.0666	2075	0.8564449
2	Rachel Daly	WSL 2020/21	9.013303	4.465725	0.4771	1.581474147	0.738	−0.3346	0.7457	0.54385	0.9652	1082	0.8330224
3	Claire Emslie	WSL 2018/19	8.529893	5.2601	0.8867	2.142543498	0.7149	−0.8198	0.3199	−0.14785	0.0791	1087	0.7847188
4	Vivianne Miedema	WSL 2019/20	9.467038	1.0664	1.3457	6.547399989	0.929	−1.2779	0.47825	−0.0691	0.0134	1291	0.7333104
5	Chloe Kelly	WSL 2019/20	10.005178	2.593225	1.7535	3.912627942	1.6133	−1.3977	1.287	−0.13265	−0.3978	1373	0.7287092
6	Jordan Nobbs	WSL 2018/19	7.867029	3.14355	0.3412	2.753357733	0.9089	−0.2799	0.4785	0.09385	0.1205	1138	0.6913031

Figure #5

	player.name	season	totalspa	SPApasses	SPACarries	SPAShots	SPATakeons	SPATurnovers	SPATacklesand	SPAAerialduels	SPAFouls	numaction	spaper100actions
1	Gareth Frank Bale	LaLiga 2015/16	18.235584	3.4408	2.6277	10.01167238	1.2393	−1.0898	0.3174	0.38825	1.2479	2167	0.8415129
2	Lionel Andrés Messi Cuccittini	LaLiga 2012/13	41.036938	3.348925	7.2961	24.00861971	8.1215	−3.4738	0.28205	−0.84265	2.1712	4961	0.8271908
3	Lionel Andrés Messi Cuccittini	LaLiga 2018/19	39.128207	12.119775	7.6929	15.12298032	5.5998	−3.5906	0.1903	0.1012	1.6479	5086	0.7693316
4	Lionel Andrés Messi Cuccittini	LaLiga 2009/10	38.470156	5.940125	8.5914	13.05540878	10.4643	−4.9882	0.7682	0.3056	3.0377	5195	0.7405227
5	Lionel Andrés Messi Cuccittini	LaLiga 2016/17	35.604069	9.290825	6.7388	14.07025214	7.4635	−4.8648	0.5688	−0.07325	2.0948	4911	0.7249861
6	Kylian Mbappé Lottin	Ligue 1 2021/22	18.730581	0.939675	5.6228	6.779656097	3.4496	−2.1324	0.3915	0.134	4.1229	2629	0.7124603