Common metrics

Expected Goals (xG)
Probability that a shot be converted into a goal, mostly based on the distance to goal and the shooting angle. The xG models can be enriched by using the type of action (set piece or open play), the type of pass received, the number of defenders in the cone in front of the goal, the ball height and other contextual information.

Expected Assists (xA) - *Common definition*
Calculates the probability of a pass resulting in a goal, considering various factors such as pass type, pass distance, assist location, and the nature of the attacking move. (see an example of why it could be misleading)

Expected Assisted Goals (xAG):
Quantifies the probability of an assist resulting in an expected goal by considering all passes that lead to a scoring chance, regardless of whether the chance is ultimately converted into a goal.

Possession State Value Models:

• **Expected Threat (xT)**:
Divides the pitch into a 16x12 grid, assigning each cell a probability of an action initiated there to result in a goal in the next N actions. An action can be a shot or a ball move (i.e. a pass or a carry).
xT is calculated by summing up two terms:
1. The product of shot probabilty and goals/shots rate from each zone
2. The sum over each zone of the probability of moving the ball to another cell (using the transition matrix) times the xT of each of the zone the ball can be moved to.
The formula is iterative and thus it needs a starting state that is xT of all cells equal to 0.
Performing N=5 iterations should imply convergence, where N is the number of actions at which we look after the one being evaluated.
(see the math explained by Karun Singh)

• **Valuing Actions by Estimating Probabilities (VAEP)**:
Values all actions performed by players - not just passes and carries, but shots and defensive actions too. It also considers the impact an action has on a team's chances of conceding as a result of the action - not just the impact on their chances of scoring. Considering a pre and a post action state, it is calculated as the difference of two subtractions:
- The scoring probability before the action - the scoring probability after the action
- The probability of conceding a goal before the action - the same probability after the action.

Expected Pass (xPass)
Is the likelihood of a pass being completed. It factors in distance, angle, pressure, body part, pattern of play (open play or set piece) and possibly other contextual information.

Expected Goals on Target (xGoT)
Is a post-shot goal probability meaning that it takes into account where the ball finished in the goal mouth. The model has only three variables:
- xG of the shot on target that encodes the positional and contextual information
- x coordinate of the ball destination in the goal mouth
- y coordinate of the ball destination in the goal mouth.

Goals prevented and Shooting Goals Added (SGA)
Stemming from xGoT:
- Goals prevented measure the ability of a goalkeeper to save shots by calculating the difference between xGoT and Goals allowed
- SGA measure the quality of the shots of a player by calculating the difference between xGoT and Goals scored.

Expected Points (xPts) - *Common definition*
Quantifies match outcomes based on the total Expected Goals (xG) for each team. It simulates matches several times and extract the probabilities of winning, drawing and losing for both teams based on the fraction of victories, draws and defeats over the simulations.

Field tilt
The share of possession in the final third in terms of touches and passes.

Passes allowed Per Defensive Action (PPDA)
Measures the pressure that the defending team puts on the opposition players when they are in possession of the ball in the attacking third. It is calculated by dividing the number of opponents' passes by the number of defensive actions of the defending team in that zone of the pitch.

Proprietary metrics

Expected Assists (xA) - *Soccerment*
Applied to any pass, quantifies the likelihood of a pass leading to a goal. It takes into account various factors such as the pass location, the position of the receiving player, and the historical probability of similar passes resulting in goals. xA helps in evaluating the creative contribution of a player in creating goal-scoring opportunities.

Expected Offensive Value Added (xOVA) - *Soccerment*
Measures the offensive value that a player adds with respect to that received from their teammates.
The formula is: `xOVA = (non-p enalty xG + xA) - xA rece ived`

Proprietary metrics (cont)

Possession State Value Models:

• **Possession Value (PV)** - *Stats Perform*

It is trained on goals scored and uses a time-based approach, measuring the probability that a team in possession will score during the next 10 seconds of play.

PV opened the street to a metric built on top of it called Match Momentum.

• **On-Ball Value (OBV)** - *StatsBomb*

It is trained on StatsBomb xG and evaluates all actions containing Goals For and Goals Against to accurately measure the risk/reward of each action. It does not include possession history features, such as details of previous events in the possession to avoid team strength bias.

Expected Points (xPts) - *Soccerment*

The common definition suffers the issue of *probabilities of non-independent events* when multiple shots occur in the same action. This flaw is addressed by simulating matches not based on shots (and their xG) but on possessions. Indeed, the possession xG is calculated as `p(goal) = 1 - p(no goal)` and `p(no goal) = (1 - xG1) * (1 - xG2) * ...`

Buildup Disruption Percentage (BDP) - *Soccerment & Antonio Gagliardi*

Tells how successful the pressing is in disrupting the opponents' buildup phase.

The metric is calculated by computing the opponent team's pass completion rate for each match, comparing it with the team's average rate, and computing a percentage difference. Then switching the point of view and looking at the team whose BDP is measured, average these differences weighting them by the opponent's average pass accuracy and change the sign.

Gegenpressing Intensity (GPI) - *Soccerment & Antonio Gagliardi*

It is the fraction of times a team immediately attempts to regain the ball in its offensive half after losing possession in the attacking 40% of the pitch. The tally takes into account defensive actions performed in the attacking half in the six seconds following a change of possession happened in the last 40% of the rectangle as well as a wrong opponents' pass that starts in their half and is recovered in the other one.

One-twos - *Soccerment*

Open-play completed passes followed by another completed pass of the same team, received by the same player who made the first pass then filtered using a progression threshold and a temporal threshold. Consider only the exchanges where the progression between the start coordinates of the first pass and the end coordinates of the second pass bring the initiating player closer to either the center of the goal or the goal line by at least 25%. Plus, no more than four seconds pass between the first and the second pass.

Proprietary metrics (cont)

Finally, discard all events where the carry distance between the end of the first pass and the start of the second is longer than five meters.. **Aerial Elo Rating Optimization (AERO)** - *Soccerment*

Measures the aerial skills, based on the Elo ranking algorithm. It takes into account the skill level matchup of each individual duel and is divided in Offensive and Defensive AERO.

Starting with a common Elo of 1000, after each aerial duel, the score of both players involved is updated by `K*(W/L -P(W))` where `W/L` is a dummy equal 1 if player wins and 0 if they lose, `P(W)` is the probability of winning the duel and `K` is a scaling constant usually equal to 32.

The probability of winning a duel is calculated as:

```
P(W) = 1 / (1+10^(( ELOa-ELOb ) / 400))
```

Normalizations

Normalization P90 minutes

Normalizatio per 90 minutes to compare players with different playing time or teams with a different number of matches played.

Normalization P100 touches

Normalization per 100 ball touches intended as events on the ball and not actual touches.

Useful for metrics like loose balls or key passes.

Normalization per Possession

Normalization based on the team's possession percentage to account for the more time to potentially create attempts.

Team strength adjustment

Based on a team strength indicator (like Elo), adjusts the metrics of two or more teams to factor in their relative strength.

*A very difficult as well as interesting improvement would be an Elo system for leagues.

Standardization (Z-score)

Transforming data to have a mean of 0 and a standard deviation of 1 to compare players or teams in terms of standard deviations from the mean.

Time decay

Normalization technique that assigns less weight to the metrics regarding matches far in time.