

Porque não se fala de EXPECTED GOALS no Brasil; e porque realmente deveríamos

No futebol de alta performance, já é unanimidade a diferença entre performance e resultados. Embora intimamente relacionados, a beleza do futebol é que nem sempre o resultado pode ser obviamente deduzido através da performance apresentada pelas equipes

Porém, a importância de interpretar, e assim quantificar, a performance de uma equipe, para fins de tomadas de decisão nas mais variadas esferas da gestão de equipe e clube, ainda esbarra na dificuldade de mensurar tal performance.

Atualmente, o futebol brasileiro tem sido inundado com uma enxurrada de dados: gols, finalizações, desarmes, passes, etc; que ajudam a entender o resultado de uma partida, porém, de pouca utilidade na interpretação da performance das equipes envolvidas.

É necessário entender que, ao contrário de outros esportes, mais suscetíveis à criação de métricas de avaliação, o futebol possui ações muito dinâmicas para serem isoladas e placares muito baixos para evitarem a presença da sorte/acaso nos resultados.

E foi então que esbarrei em um KPI (key performance indicator), me surpreendendo em como já vem sendo usado no futebol mundial (especialmente Europeu) para auxiliar na quantificação e qualificação de performances de equipes e jogadores e individuais. E me surpreendi ainda mais, com a inexplicável imunidade do futebol brasileiros a esse tipo de modelagem e uso de dados.

O KPI em questão é ***Expected Goals***.

Em termos simples: cada finalização (contra ou favor) terá um Expected Goal Value (Xg), ou seja, uma probabilidade de resultar em gol (assim, $0 < Xg < 1$). Tal probabilidade advém de algoritmos alimentados por milhares de *data points*, que reproduzem o contexto de milhares de finalizações feitas no passado, e seu resultado final (“gol” ou “não gol”).

(Mais abaixo o conceito ficará mais claro, mas recomendo muito a leitura de outras fontes, principalmente estrangeiras, com o modelo mais em detalhe).

Como economista, acho interessante modelos de “expected values” porque estes me permitem não apenas interpretar a realidade (actual values) como também fazer previsões (alterar inputs para alcançar expected values no futuro). Assim, Expected Goals torna-se útil na criação de modelos, sendo estes uma importante ferramenta na tentativa de livrar a interpretação de resultados de estatísticas defasadas (e altamente influenciadas pela sorte/acaso inerente ao jogo)

Como scout, Expected Goals me permite avaliar se uma equipe ou atleta está over-performing (performance acima do esperado) ou underperforming (abaixo do esperado), o que apresentaria inúmeras limitações se avaliados com as métricas usuais de futebol (finalizações, passes, gols, assistências, etc).

Porque deveríamos estar falando de Expected Goals

Para extrair valor de um “expected value”, é necessário usá-lo em termos de comparação com “actual values”, ou seja, comparar os valores esperados com os valores realmente observados. Tal comparação mostrará se uma equipe/atleta foi melhor, pior ou igual ao esperado. Assim, podemos identificar quem são os melhores/piores em termos de eficiência

Xg na análise coletiva:

Considerando que cada finalização apresenta uma probabilidade de resultar em gol (definição de Xg), se somadas as probabilidades de todas as finalizações realizadas por uma equipe em uma partida, teremos quantos gols era esperado que tal equipe marcasse naquele jogo.

Podemos assim, comparar com a quantidade de gols que REALMENTE foram marcados.

Por exemplo, uma equipe que teve um Expected Goals de $Xg=2.3$ em uma partida, e que acabou marcando 3 gols na realidade, foi eficiente ($3 > 2.3$). Alternativamente, se a mesma equipe tivesse marcado apenas 1 gol, teria sido ineficiente ($1 < 2.3$).

Campeonatos longos provém dados ainda mais interessantes: se somarmos os Xg de todas as partidas ao longo de um campeonato, podemos comparar com o número de “Gols Pró” de um time. Também é possível somar o Xg de finalizações concedidas, e assim teríamos Expected Goals Against (XgA) (gols esperados contra), podendo assim comparar com o número de “Gols Contra”.

Enquanto Expected Goals fornece informações sobre a eficiência do setor ofensivo de uma equipe, Expected Goals Against fornece informações valiosas da atuação do setor defensivo, uma vez que mostra qualitativa e quantitativamente as chances que permitiram ao adversário criar.

Com isso, podemos subtrair XgA de Xg, e teremos Net Expected Goals (Xg “Líquido”) , sendo que o valor para comparação seria o “saldo de gols” de uma equipe em uma competição. Sendo assim, Net Expected Goals é o melhor KPI para avaliar a eficiência de uma equipe como um todo (setores defensivo e ofensivo) dentro de uma competição.

Nº	Team	M	G	GA	PTS	xG	xGA	NPxGD	xPTS
1	Juventus	38	86	24	95	59.23 ^{+8.77}	28.58 ^{+4.58}	+27.61	73.51 ^{-21.49}
2	Napoli	38	77	29	91	70.45 ^{-8.55}	25.27 ^{-3.73}	+40.44	82.23 ^{-8.77}
3	Roma	38	61	28	77	65.63 ^{+4.63}	38.46 ^{+10.46}	+26.54	69.24 ^{-7.76}
4	Lazio	38	89	49	72	66.51 ^{-22.49}	42.57 ^{-5.43}	+23.18	66.05 ^{-5.95}
5	Inter	38	66	30	72	60.83 ^{-5.17}	37.33 ^{+7.33}	+19.70	65.25 ^{-8.75}
6	AC Milan	38	56	42	64	50.88 ^{-5.12}	38.78 ^{-5.22}	+11.81	62.35 ^{-1.85}
7	Atalanta	38	57	39	60	64.29 ^{+7.29}	36.50 ^{-2.50}	+22.46	69.88 ^{+9.88}
8	Florentina	38	54	46	57	54.12 ^{+0.12}	42.15 ^{-3.85}	+8.16	60.41 ^{+3.41}
9	Torino	38	54	46	54	46.59 ^{-7.41}	49.11 ^{+3.11}	+2.03	51.84 ^{-2.16}
10	Sampdoria	38	56	60	54	53.14 ^{-2.86}	62.26 ^{+2.26}	-7.44	46.87 ^{-7.13}
11	Sassuolo	38	29	59	43	47.15 ^{+8.15}	50.63 ^{-3.37}	-8.81	51.70 ^{+8.70}
12	Genoa	38	33	43	41	42.21 ^{+9.21}	42.87 ^{-0.13}	-1.42	50.38 ^{+9.38}
13	Udinese	38	48	63	40	45.20 ^{-2.80}	50.71 ^{+2.29}	-8.11	47.34 ^{+7.34}
14	Chievo	38	36	59	40	34.42 ^{+1.58}	55.28 ^{-3.72}	-17.03	38.66 ^{-1.34}
15	Bologna	38	40	52	39	36.11 ^{-3.89}	48.80 ^{-3.20}	-11.92	43.01 ^{+4.01}
16	Cagliari	38	33	61	39	38.27 ^{+5.27}	57.91 ^{-3.09}	-15.83	40.41 ^{+1.41}
17	SPAL 2013	38	39	59	38	34.61 ^{+4.39}	60.18 ^{+1.18}	-21.13	35.41 ^{-2.59}
18	Crotone	38	40	66	35	33.24 ^{-8.76}	59.43 ^{-8.57}	-22.55	35.44 ^{+0.44}
19	Verona	38	30	78	25	33.76 ^{+3.76}	72.71 ^{-5.23}	-38.19	31.08 ^{+6.08}
20	Benevento	38	33	84	21	38.21 ^{+5.21}	77.31 ^{-8.63}	-31.49	30.09 ^{+9.09}

A imagem acima (fonte: understat.com) nos permite tirar conclusões interessantes: Juventus marcou 144% dos gols esperados e assim, conquistou 129% dos pontos esperados (xPTS é calculado através da Net Expected Goals de cada partida somados), sendo assim, a Juve conquistou o título porque foi mais eficiente do que o Napoli, ainda que este criou mais chances de gol (Xg maior) e concedeu menos chances (XgA menor)

Xg na análise individual:

Expected Goals também pode ser usado como KPI de atletas individuais. Para jogadores de ataque, a avaliação é mais intuitiva: um atleta eficiente terá um valor observado maior do que o valor esperado (Xg).

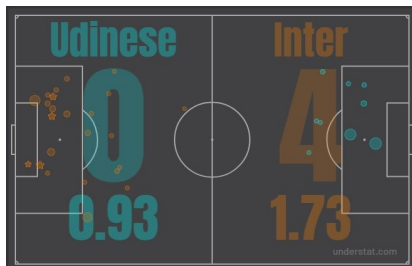
Na imagem abaixo vemos alguns números de atletas na Copa do Mundo da Fifa 2018 (Fonte: Wyscout):

Rank	Player	Expected goals	Goals	Team
1	H. Kane	4.53	6	England
2	Neymar	4.19	2	Brazil
3	M. Berg	3.81	0	Sweden
4	A. Griezmann	3.41	4	France
5	M. Mandžukić	2.95	3	Croatia
6	R. Lukaku	2.89	4	Belgium
7	E. Cavani	2.78	3	Uruguay
8	E. Hazard	2.73	3	Belgium
9	Cristiano Ronaldo	2.52	4	Portugal
10	L. Messi	2.51	1	Argentina
11	A. Mitrović	2.25	1	Serbia
12	Roberto Firmino	2.23	1	Brazil
13	O. Giroud	2.15	0	France
14	I. Perišić	2.1	3	Croatia
15	M. Jedinak	2.07	2	Australia
16	L. Modrić	2.04	2	Croatia
17	Diego Costa	2.01	3	Spain
18	Philippe Coutinho	2	2	Brazil
19	A. Rebić	1.98	1	Croatia
20	K. Mbappé	1.94	4	France
21	A. Granqvist	1.8	2	Sweden
22	M. Batshuayi	1.73	1	Belgium
23	B. Džemali	1.68	1	Switzerland
24	L. Suárez	1.66	2	Uruguay
25	Gabriel Jesus	1.65	0	Brazil
26	C. Vela	1.64	1	Mexico
27	A. Dzyuba	1.56	3	Russia
28	J. Lingard	1.52	1	England
29	V. Moses	1.48	1	Nigeria
30	Mohamed Salah	1.47	2	Egypt

Fica evidente a diferença da eficiência de Neymar (marcou 47,7% dos gols esperados) e Mbappé (206,2% dos gols esperados) na competição, diferença essa facilmente percebida quase que por unanimidade por todos que acompanharam os jogos. Porém, quem diria que o russo Artem Dzyuba teve eficiência similar (192,3% dos gols esperados) ao craque Mbappé? Se um clube, com recursos limitados, precisasse contratar alguém da lista acima, quem você indicaria?

O uso de Xg como KPI de jogadores de outras posições porém, é mais subjetiva. Para jogadores de defesa, Expected Goals Against fornece uma interpretação da quantidade e qualidade de chances criadas pelo adversário, ou seja, quanto menor o número de XgA (Expected Goals Against), melhor o sistema defensivo de uma equipe.

Representações visuais de XgA permitem identificar as áreas do campo com a maior quantidade de chances concedidas, como podemos ver na imagem abaixo (áreas dos círculos correspondem ao valor de Xg) (Fonte: unserstat.com):



Uma equipe com número de “Gols Sofridos” menor do que XgA pode também indicar a qualidade de peças individuais no sistema defensivo, especialmente goleiros. Já meio campistas, são melhores avaliados através de Expected Assists, outro KPI que segue a lógica de Xg mas que não será tratado a fundo nesse artigo;

Porque não se fala de Expected Goals no Brasil

Mas por que, apesar das vantagens mencionadas, ainda não se fala de Expected Goals no Brasil?

Acredito que os motivos para tal fato está mais relacionado à cultura e padrões vigentes ao futebol brasileiro do que simplesmente falta de interesse dos profissionais.

Apesar dos grandes avanços já alcançados pelos profissionais da área técnica e de gestão, o futebol brasileiro continua com grande aversão a riscos. Implementar novas filosofias ainda requer grande ousadia e muitas vezes tal comportamento chega a ser mal visto dentro dos clubes, sendo estes extremamente tradicionais.

Além disso, a inovação continua pouco fomentada no meio do futebol brasileiro. A produção de conhecimento, embora tenha se desenvolvido exponencialmente nos últimos anos, ainda não acompanha o ritmo da de outros países

Porém, o aspecto que mais pesa contra a adoção de métodos como Expected Goals e outros que requerem um certo nível de tecnologia, é o ambiente pouco atrativo que empresas detentoras desse conhecimento encontram no ecossistema do futebol brasileiro em geral. Um ambiente que evidentemente afeta o produto colocado pelos clubes dentro de campo, mas que começa a mostrar seus efeitos também fora dele, nas áreas técnicas e de gestão