
ANÁLISE DO DESEMPENHO DO CAMPEONATO BRASILEIRO DE PRIMEIRA DIVISÃO ENTRE 2003 E 2014

ANALYSIS OF THE BRAZILIAN CHAMPIONSHIP FIRST DIVISION PERFORMANCE BETWEEN 2003 AND 2014

Domingos Rodrigues Pandeló Junior¹, Roberto Melchior² e Paulo Henrique Silva Marques de Azevedo¹

¹Universidade Federal de São Paulo, Santos-SP, Brasil.

²Faculdade de Educação Física de Santos, Santos-SP, Brasil.

RESUMO

O objetivo do presente artigo é avaliar o desempenho dos times da primeira divisão do futebol brasileiro, com base em dados do campeonato brasileiro, no período entre 2003 e 2014. A técnica multivariada utilizada para a análise foi a construção de uma medida de dissimilaridade e de similaridade, bem como a análise de alguns indicadores. Como principais resultados observou-se uma grande concentração de desempenhos acima da média em clubes das regiões sul e sudeste. Além disso, observou-se que o São Paulo Futebol Clube foi o time com melhor desempenho no período de pontos corridos.

Palavras-chave: Desempenho. Análise. Similaridade. Dissimilaridade. Futebol. Campeonato brasileiro.

ABSTRACT

The aim of this paper is to evaluate the performance of teams in the first division of Brazilian soccer, based on data of the Brazilian championship, between 2003 and 2014. The multivariate technique used for the analysis was the construction of a measure of dissimilarity and similarity as well as the analysis of some indicators. The main results show that there was a large concentration of high performances in clubs of south and southeast regions. Furthermore, it was observed that the São Paulo Futebol Clube was the team with better performance in the period.

Keywords: Performance. Analysis. Similarity. Dissimilarity. Soccer. Brazilian championship.

Introdução

O futebol tem uma magia especial para a maior parte dos povos. No Brasil, a magia da bola atinge uma dimensão difícil de se explicar¹. O campeonato brasileiro é um dos mais equilibrados do mundo, com muitas equipes em condições teóricas (técnicas) de lutar pelo título, conforme pode ser visto nos resultados dos anos pesquisados. Nos últimos anos, de forma um pouco mais intensa, houve uma maior aproximação do futebol com a academia, com a ciência. Dessa forma, cada vez mais, um maior número de pesquisadores se dedica a estudar o futebol, nas suas múltiplas facetas².

Vários estudos já foram realizados na tentativa de avaliar ou tentar prever o desempenho de times em competições. Dentre eles pode-se citar os trabalhos de Añon et al³, Marcelino, Sampaio & Mesquita⁴, De Araújo et al⁵, sendo estes bem centrados na utilização de métodos estatísticos para a previsão. O trabalho de Almeida, Oliveira & Silva⁶, fez uma interessante análise, sobre times com mando de jogo em casa nas séries A (primeira divisão) e B (segunda divisão), tendo encontrado uma clara vantagem em se jogar em casa, com maior impacto na série B do campeonato. Da mesma forma, os trabalhos de Hass⁷ e Gomes e Mendo⁸ também procuraram avaliar o impacto do mando de jogo e apoio efetivo da torcida, no resultado. Trata-se de uma abordagem relevante, pois em certas condições, o mando de campo pode ser um fator complicador, especialmente se o time não estiver com uma boa relação com a torcida. Fatores mais específicos, tais como posse de bola, número de finalizações, por exemplo, podem explicar o resultado de uma equipe⁹⁻¹¹. Já o trabalho de

Carvalho, Scaglia e Costa¹² faz uma interessante análise sobre o desempenho tático e o resultado final (empate, vitória e derrota), encontrando diferenças significativas e possíveis explicações entre desempenho tático e resultado final.

A previsão de desempenho poderia auxiliar patrocinadores na tomada de decisão, por exemplo em qual equipe investir seu dinheiro, bem como, influir na decisão de jogadores a respeito da escolha da equipe com potencial para alcançar as primeiras colocações. O objetivo do presente artigo é avaliar o desempenho dos times da primeira divisão do futebol brasileiro, com base em resultados obtidos no campeonato brasileiro de primeira divisão, no período entre 2003 e 2014, através de indicadores que possam refletir o desempenho das equipes.

Métodos

Os dados utilizados no trabalho foram obtidos através do site Futdados¹³ e foram em seguida transferidos para o software SPSS 21 Pro. Trabalhou-se com dados de 2003 até 2014, pois o critério de pontos corridos, no campeonato brasileiro, começou em 2003. Foram efetuados os cálculos apenas para o campeonato da primeira divisão. A técnica multivariada utilizada para a análise foi a construção de uma medida de dissimilaridade e de similaridade, bem como a análise de alguns indicadores.

As variáveis disponíveis para análise foram o número de pontos, de jogos, de vitórias, de empates, de derrotas, número de gols à favor, o número de gols contra, o saldo de gol e o número de edições do campeonato em que o time participou (na primeira divisão) no período considerado.

Para facilitar as análises foram criados quatro indicadores de desempenho: a relação entre o número de vitórias e o total de jogos (D1), a relação entre o número de empates e o total de jogos (D2), a relação entre o número de derrotas e o total de jogos (D3) e o índice de aproveitamento (D4), definido como sendo o complemento de total de derrotas, dividido pela soma de empates e vitórias.

A ideia de criação dos índices foi para poder normalizar e padronizar os dados, pois assim, o número de edições de cada time, o número de jogos, dentre outros aspectos, não seriam influenciados por uma maior ou menor participação em termos de número de campeonatos disputados na primeira divisão do campeonato brasileiro.

Assim, com base na tabela 1 criou-se os 4 indicadores que foram usados para o cálculo da medida de similaridade e dissimilaridade. No total, 41 clubes foram incluídos na amostra.

Medida de Dissimilaridade

A distância euclidiana foi usada como medida de dissimilaridade. A distância euclidiana para os indivíduos i e j , com a utilização de p parâmetros, é dada pela fórmula¹⁴:

$$D_{(i,j)} = \left(\sum_{j=1}^p (X_{i,j} - X_{(j)})^2 \right)^{1/2}$$

No presente artigo, i e j são os clubes considerados. Note que neste caso, o modelo atinge uma fórmula diferenciada, pois trata-se de múltiplas análises (41 times foram considerados, no total). No caso dos parâmetros utilizados, eles foram quatro: D1, D2, D3 e D4, já definidos anteriormente.

Medida de Similaridade

A similaridade, por sua vez, pode ser mensurada pelo coeficiente de correlação de Pearson, que pode ser calculado pela fórmula¹⁵:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \cdot \text{var}(Y)}}$$

onde x_1, x_2, \dots, x_n e y_1, y_2, \dots, y_n são os valores medidos de ambas as variáveis, e

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

$$\bar{y} = \frac{1}{n} \cdot \sum_{i=1}^n y_i$$

são as médias aritméticas de ambas as variáveis

No caso do presente estudo, trata-se da medida de correlação entre as variáveis consideradas, para os diversos parâmetros considerados (D1, D2, D3 e D4).

A análise da Tabela 1 dá uma ideia da concentração dos times, com melhor desempenho, nas regiões sul (26,83% do total) e sudeste (43,90%). As duas regiões perfazem 71,73% dos times que disputaram a primeira divisão do campeonato brasileiro, no período considerado. Já as regiões centro oeste (7,32%), nordeste (19,51%) e norte (2,44%) apresentaram uma participação bem menor.

Tabela 1. Dados brutos do campeonato brasileiro

Equipe	Pontos	Jogos	Vitórias	Empates	Derrotas	Gols Pró	Gols Contra	Saldo Gol	Edições
São Paulo	851	499	242	125	132	779	546	233	13
Cruzeiro	806	499	236	98	165	800	638	162	13
Internacional	777	499	217	124	158	686	579	107	13
Santos	761	499	210	131	158	776	632	144	13
Fluminense	735	499	200	133	166	711	652	59	13
Corinthians	728	461	198	132	131	616	513	103	12
Flamengo	706	499	187	149	163	655	635	20	13
Grêmio	693	457	193	114	150	632	527	105	12
Atlético Paranaense	659	461	183	110	168	651	624	27	12
Atlético Mineiro	657	461	179	120	162	663	614	49	12
Palmeiras	587	415	159	110	146	569	540	29	11
Goiás	583	423	161	100	162	608	578	30	11
Botafogo	581	430	152	125	153	588	567	21	11
Vasco	550	423	143	121	159	570	623	-53	11
Coritiba	508	385	135	103	147	484	488	-4	10
Figueirense	496	385	130	106	149	486	556	-70	10
Vitória	343	282	91	70	121	355	407	-52	7
Ponte Preta	317	271	82	72	117	310	416	-106	7
Paraná	281	210	79	44	87	294	312	-18	5
Juventude	266	210	71	50	89	268	327	-59	5
Sport	260	213	66	62	85	257	297	-40	6
Bahia	224	198	55	59	84	207	270	-63	5
São Caetano	215	172	65	41	66	209	199	10	4
Náutico	200	190	54	38	98	224	318	-94	5

Continuação da Tabela 1...

Equipe	Pontos	Jogos	Vitórias	Empates	Derrotas	Gols Pró	Gols Contra	Saldo Gol	Edições
Criciúma	188	168	50	38	80	195	266	-71	4
Avaí	154	137	39	37	61	178	223	-45	4
Guarani	147	130	36	39	55	140	180	-40	3
Paysandu	146	134	41	31	62	193	245	-52	3
Fortaleza	142	126	36	34	56	155	200	-45	3
Portuguesa	127	114	31	38	45	137	157	-20	3
Atlético-GO	120	114	30	30	54	138	169	-31	3
Ceará	86	76	20	26	30	82	108	-26	2
Barueri	77	76	19	23	34	98	116	-18	2
Chapecoense	72	61	19	15	27	58	68	-10	2
Santo André	41	38	11	8	19	46	61	-15	1
Brasiliense	41	42	10	11	21	47	67	-20	1
América Mineiro	37	38	8	13	17	51	69	-18	1
Ipatinga	35	38	9	8	21	37	67	-30	1
Santa Cruz	28	38	7	7	24	41	76	-35	1
Joinville	21	23	5	6	12	17	25	-8	1
América-RN	17	38	4	5	29	24	80	-56	1

A Tabela 2, por sua vez, mostra os quatro indicadores criados para as análises. O grande objetivo da tabela 2 foi o de se criar indicadores relativos, que não levassem em consideração o número de participações dos times na competição, mas a sua efetividade, ou inefetividade, em termos de vitórias, empates, derrotas e aproveitamento.

Tabela 2. Indicadores de Efetividade e Inefetividade

Equipe	D1 (vitória/jogos)	D2 (empates/jogos)	D3 (derrotas/jogos)	D4 (aproveitamento)
São Paulo	48,50%	25,05%	26,45%	64,03%
Cruzeiro	47,29%	19,64%	33,07%	50,60%
Internacional	43,49%	24,85%	31,66%	53,67%
Santos	42,08%	26,25%	31,66%	53,67%
Fluminense	40,08%	26,65%	33,27%	50,15%
Corinthians	42,95%	28,63%	28,42%	60,30%
Flamengo	37,47%	29,86%	32,67%	51,49%
Grêmio	42,23%	24,95%	32,82%	51,14%
Atlético Paranaense	39,70%	23,86%	36,44%	42,66%
Atlético Mineiro	38,83%	26,03%	35,14%	45,82%
Palmeiras	38,31%	26,51%	35,18%	45,72%
Goiás	38,06%	23,64%	38,30%	37,93%
Botafogo	35,35%	29,07%	35,58%	44,77%
Vasco	33,81%	28,61%	37,59%	39,77%
Coritiba	35,06%	26,75%	38,18%	38,24%
Figueirense	33,77%	27,53%	38,70%	36,86%
Vitória	32,27%	24,82%	42,91%	24,84%

Continuação da Tabela 2...

Equipe	D1 (vitória/jogos)	D2 (empates/jogos)	D3 (derrotas/jogos)	D4 (aproveitamento)
Ponte Preta	30,26%	26,57%	43,17%	24,03%
Paraná	37,62%	20,95%	41,43%	29,27%
Juventude	33,81%	23,81%	42,38%	26,45%
Sport	30,99%	29,11%	39,91%	33,59%
Bahia	27,78%	29,80%	42,42%	26,32%
São Caetano	37,79%	23,84%	38,37%	37,74%
Náutico	28,42%	20,00%	51,58%	-6,52%
Criciúma	29,76%	22,62%	47,62%	9,09%
Avaí	28,47%	27,01%	44,53%	19,74%
Guarani	27,69%	30,00%	42,31%	26,67%
Paysandu	30,60%	23,13%	46,27%	13,89%
Fortaleza	28,57%	26,98%	44,44%	20,00%
Portuguesa	27,19%	33,33%	39,47%	34,78%
Atlético-GO	26,32%	26,32%	47,37%	10,00%
Ceará	26,32%	34,21%	39,47%	34,78%
Barueri	25,00%	30,26%	44,74%	19,05%
Chapecoense	31,15%	24,59%	44,26%	20,59%
Santo André	28,95%	21,05%	50,00%	0,00%
Brasiliense	23,81%	26,19%	50,00%	0,00%
América Mineiro	21,05%	34,21%	44,74%	19,05%
Ipatinga	23,68%	21,05%	55,26%	-23,53%
Santa Cruz	18,42%	18,42%	63,16%	-71,43%
Joinville	21,74%	26,09%	52,17%	-9,09%
América-RN	10,53%	13,16%	76,32%	-222,22%

Resultados e Discussões

A análise da Tabela 1 mostra que dos 41 clubes analisados, 6 deles - São Paulo, Cruzeiro, Internacional, Santos, Fluminense e Flamengo participaram das 13 edições do campeonato brasileiro, por pontos corridos, consideradas. Já Corinthians, Grêmio, Atlético Mineiro e Atlético Paranaense participaram de 12 edições. Outro ponto que chama a atenção é que 71% dos clubes considerados na análise pertencem às regiões sul e sudeste do Brasil. Caso sejam considerados os clubes que tiveram 5 ou mais participações, na primeira divisão, no período em análise, apenas Goiás, Vitória, Sport e Bahia, entrariam como representantes fora do eixo da região Sul/Sudeste. Dessa forma, pode-se inferir uma maior concentração, em termos de poderio, dos clubes das regiões sul e sudeste, no campeonato brasileiro de primeira divisão.

A Tabela 2 é interessante para todas as análises. Em primeiro lugar ela representa os quatro indicadores que foram utilizados para o cálculo da similaridade e dissimilaridade. Em segundo lugar ela torna as análises mais justas, pois como são índices, não levam em consideração o número de participações de cada clube, mas sim a "efetividade" em cada quesito analisado. O índice D1 mede a relação entre o número de vitórias e o total de jogos. O índice D2 mede a relação entre o número de empates e o total de jogos. O índice D3 mede a relação entre o total de derrotas e o total de jogos. Finalmente, o índice D4, é uma medida de aproveitamento, considerando o total de vitórias e empates em relação ao número de jogos.

Uma vez que se pontua nas vitórias e nos empates, o índice de aproveitamento, dá uma ideia da efetividade do clube na busca por pontos. É evidente que a vitória deve ser o grande objetivo, pois pontua mais do que um empate (3 pontos contra 1 ponto), mas em última instância, dependendo da situação, um empate pode ser considerado um resultado razoável. Do ponto de vista da efetividade em termos de resultados, o total de pontos (TP) pode ser obtido por meio de vitória (V), empate (E), ou derrota (D). Evidentemente, a derrota seria o resultado menos efetivo, em termos de pontuação para o campeonato, mas é um resultado que invariavelmente ocorre, mesmo em equipes que venham a se sagrar campeãs. Como a vitória vale 3 pontos, vamos assumir que a vitória seja o maior índice de efetividade, o empate seria o segundo melhor índice de efetividade e a derrota o pior. Dessa forma, em termos de efetividade teríamos um índice de 1 para a vitória 1/3 para o empate e 0 para a derrota.

A Tabela 3 mostra a efetividade em termos de vitória por número de jogo. Na coluna 1 tem-se a classificação original, com base no número de pontos do clube nas 13 edições do campeonato brasileiro consideradas. Na coluna 2 tem-se a classificação com base no percentual de vitórias em relação ao número de jogos. Nas 3 primeiras posições não há nenhuma mudança na ordem das colunas, mas a partir da 4 posição algumas mudanças ocorrem. Do ponto de vista do campeonato, como um todo, a segunda coluna é extremamente importante, pois mostra a efetividade das equipes em termos de vitórias, que como visto anteriormente, é onde ocorre a maior pontuação para o campeonato. Obviamente uma equipe que conseguir uma maior consistência em termos de vitórias e, ao mesmo tempo, conseguir um menor número de derrotas. Se olharmos para os 10 primeiros, no total de pontos corridos, pela classificação original (coluna 1), verifica-se que apenas o Flamengo apresentou queda expressiva quando se compara com a coluna 2. No caso do Flamengo foi uma queda relativa de 7 posições. Grêmio e Corinthians foram os que apresentaram as maiores variações positivas, subindo 3 e 2 posições, respectivamente.

Tabela 3. Percentual de vitória

Equipe (Class. Original)	Equipe (Class. Maior nº Vitória)	Vitória/Total
São Paulo	São Paulo	48,50%
Cruzeiro	Cruzeiro	47,29%
Internacional	Internacional	43,49%
Santos	Corinthians	42,95%
Fluminense	Grêmio	42,23%
Corinthians	Santos	42,08%
Flamengo	Fluminense	40,08%
Grêmio	Atlético Paranaense	39,70%
Atlético Paranaense	Atlético Mineiro	38,83%
Atlético Mineiro	Palmeiras	38,31%
Palmeiras	Goiás	38,06%
Goiás	São Caetano	37,79%
Botafogo	Paraná	37,62%
Vasco	Flamengo	37,47%
Coritiba	Botafogo	35,35%
Figueirense	Coritiba	35,06%
Vitória	Juventude	33,81%
Ponte Preta	Vasco	33,81%

Continuação da Tabela 3...

Equipe (Class. Original)	Equipe (Class. Maior nº Vitória)	Vitória/Total
Paraná	Figueirense	33,77%
Juventude	Vitória	32,27%
Sport	Chapecoense	31,15%
Bahia	Sport	30,99%
São Caetano	Paysandu	30,60%
Náutico	Ponte Preta	30,26%
Criciúma	Criciúma	29,76%
Avaí	Santo André	28,95%
Guarani	Fortaleza	28,57%
Paysandu	Avaí	28,47%
Fortaleza	Náutico	28,42%
Portuguesa	Bahia	27,78%
Atlético-GO	Guarani	27,69%
Ceará	Portuguesa	27,19%
Barueri	Atlético-GO	26,32%
Chapecoense	Ceará	26,32%
Santo André	Barueri	25,00%
Brasiliense	Brasiliense	23,81%
América Mineiro	Ipatinga	23,68%
Ipatinga	Joinville	21,74%
Santa Cruz	América Mineiro	21,05%
Joinville	Santa Cruz	18,42%
América-RN	América-RN	10,53%

Quando se analisa o percentual de empates, em relação ao total de jogos, na Tabela 4, verifica-se que, de um modo geral, os clubes menores jogam mais pelo empate, ou apresentam menor potencial para vitória. Se olharmos para a tabela, vemos que entre os 10 times que mais empataram, apenas Flamengo e Corinthians estão na classificação original, entre os 10 times que mais pontuaram em campeonatos brasileiros, no período considerado. Do ponto de vista da estratégia de jogo, talvez a opção pelo empate, como filosofia tática, implique em abrir mão da vitória, evitando o risco de derrota. Dependendo, todavia, do objetivo do clube, em função das suas especificidades, esta pode ser uma estratégia ótima, no sentido de otimização (melhor uso) dos recursos disponíveis.

Tabela 4. Percentual de empate

Equipe (Class. Original)	Equipe (Class. Maior nº Empates)	Empate/Total
São Paulo	Ceará	34,21%
Cruzeiro	América Mineiro	34,21%
Internacional	Portuguesa	33,33%
Santos	Barueri	30,26%
Fluminense	Guarani	30,00%
Corinthians	Flamengo	29,86%
Flamengo	Bahia	29,80%

Continuação da Tabela 4...

Equipe (Class. Original)	Equipe (Class. Maior nº Empates)	Empate/Total
Grêmio	Sport	29,11%
Atlético Paranaense	Botafogo	29,07%
Atlético Mineiro	Corinthians	28,63%
Palmeiras	Vasco	28,61%
Goiás	Figueirense	27,53%
Botafogo	Avaí	27,01%
Vasco	Fortaleza	26,98%
Coritiba	Coritiba	26,75%
Figueirense	Fluminense	26,65%
Vitória	Ponte Preta	26,57%
Ponte Preta	Palmeiras	26,51%
Paraná	Atlético-GO	26,32%
Juventude	Santos	26,25%
Sport	Brasiliense	26,19%
Bahia	Joinville	26,09%
São Caetano	Atlético Mineiro	26,03%
Náutico	São Paulo	25,05%
Criciúma	Grêmio	24,95%
Avaí	Internacional	24,85%
Guarani	Vitória	24,82%
Paysandu	Chapecoense	24,59%
Fortaleza	Atlético Paranaense	23,86%
Portuguesa	São Caetano	23,84%
Atlético-GO	Juventude	23,81%
Ceará	Goiás	23,64%
Barueri	Paysandu	23,13%
Chapecoense	Criciúma	22,62%
Santo André	Santo André	21,05%
Brasiliense	Ipatinga	21,05%
América Mineiro	Paraná	20,95%
Ipatinga	Náutico	20,00%
Santa Cruz	Cruzeiro	19,64%
Joinville	Santa Cruz	18,42%
América-RN	América-RN	13,16%

A análise da Tabela 5 torna evidente a importância de tentar se evitar a derrota. Se observarmos a relação entre a classificação original, com os times que mais pontuaram no campeonato brasileiro e a coluna com os times que apresentam o maior número de derrotas observa-se uma relação inversa em termos de posição. Neste quesito o São Paulo mostra-se o mais efetivo, seguido do Corinthians. Como a derrota não pontua, perder menos é um dos requisitos para se buscar uma posição confortável num campeonato por pontos corridos, onde a regularidade conta muito, especialmente quando se têm times com bastante equilíbrio técnico.

Tabela 5. Percentual de derrotas

Equipe (Class. Original)	Equipe (Class. Maior nº Derrotas)	Derrota/Total
São Paulo	América-RN	76,32%
Cruzeiro	Santa Cruz	63,16%
Internacional	Ipatinga	55,26%
Santos	Joinville	52,17%
Fluminense	Náutico	51,58%
Corinthians	Santo André	50,00%
Flamengo	Brasiliense	50,00%
Grêmio	Criciúma	47,62%
Atlético Paranaense	Atlético-GO	47,37%
Atlético Mineiro	Paysandu	46,27%
Palmeiras	Barueri	44,74%
Goiás	América Mineiro	44,74%
Botafogo	Avaí	44,53%
Vasco	Fortaleza	44,44%
Coritiba	Chapecoense	44,26%
Figueirense	Ponte Preta	43,17%
Vitória	Vitória	42,91%
Ponte Preta	Bahia	42,42%
Paraná	Juventude	42,38%
Juventude	Guarani	42,31%
Sport	Paraná	41,43%
Bahia	Sport	39,91%
São Caetano	Portuguesa	39,47%
Náutico	Ceará	39,47%
Criciúma	Figueirense	38,70%
Avaí	São Caetano	38,37%
Guarani	Goiás	38,30%
Paysandu	Coritiba	38,18%
Fortaleza	Vasco	37,59%
Portuguesa	Atlético Paranaense	36,44%
Atlético-GO	Botafogo	35,58%
Ceará	Palmeiras	35,18%
Barueri	Atlético Mineiro	35,14%
Chapecoense	Fluminense	33,27%
Santo André	Cruzeiro	33,07%
Brasiliense	Grêmio	32,82%
América Mineiro	Flamengo	32,67%
Ipatinga	Internacional	31,66%
Santa Cruz	Santos	31,66%
Joinville	Corinthians	28,42%
América-RN	São Paulo	26,45%

A Tabela 6 mede o índice de aproveitamento de cada equipe. Como definido anteriormente, o índice mede o complemento da divisão entre o número de derrotas (D) e a soma de empates (E) e vitórias (V). Assim, o índice de aproveitamento é dado por $1 - [D/(E+V)]$.

Como esperado, o índice de aproveitamento é um dos mais efetivos para mensurar o desempenho de um clube num campeonato por pontos corridos. A análise da tabela 6 chama a atenção para a posição do Corinthians que salta da sexta posição da classificação original, pelo total de pontos corridos, para a segunda posição, quando se considera o índice de aproveitamento. O time do Cruzeiro, por sua vez, apresenta um comportamento inverso, apresentando uma queda acentuada, saindo da segunda colocação, na classificação original, para a sétima na classificação com base no índice de aproveitamento.

Tabela 6. Percentual de aproveitamento

Equipe (Class. Original)	Equipe (Class. Original)	Aproveitamento
São Paulo	São Paulo	64,03%
Cruzeiro	Corinthians	60,30%
Internacional	Internacional	53,67%
Santos	Santos	53,67%
Fluminense	Flamengo	51,49%
Corinthians	Grêmio	51,14%
Flamengo	Cruzeiro	50,60%
Grêmio	Fluminense	50,15%
Atlético Paranaense	Atlético Mineiro	45,82%
Atlético Mineiro	Palmeiras	45,72%
Palmeiras	Botafogo	44,77%
Goiás	Atlético Paranaense	42,66%
Botafogo	Vasco	39,77%
Vasco	Coritiba	38,24%
Coritiba	Goiás	37,93%
Figueirense	São Caetano	37,74%
Vitória	Figueirense	36,86%
Ponte Preta	Portuguesa	34,78%
Paraná	Ceará	34,78%
Juventude	Sport	33,59%
Sport	Paraná	29,27%
Bahia	Guarani	26,67%
São Caetano	Juventude	26,45%
Náutico	Bahia	26,32%
Criciúma	Vitória	24,84%
Avaí	Ponte Preta	24,03%
Guarani	Chapecoense	20,59%
Paysandu	Fortaleza	20,00%
Fortaleza	Avaí	19,74%
Portuguesa	Barueri	19,05%

Continuação da Tabela 6...

Equipe (Class. Original)	Equipe (Class. Original)	Aproveitamento
Atlético-GO	América Mineiro	19,05%
Ceará	Paysandu	13,89%
Barueri	Atlético-GO	10,00%
Chapecoense	Criciúma	9,09%
Santo André	Santo André	0,00%
Brasiliense	Brasiliense	0,00%
América Mineiro	Náutico	-6,52%
Ipatinga	Joinville	-9,09%
Santa Cruz	Ipatinga	-23,53%
Joinville	Santa Cruz	-71,43%
América-RN	América-RN	-222,22%

Medidas de Similaridade e de Dissimilaridade

O maior objetivo do presente estudo é analisar o desempenho das equipes da série A, do campeonato brasileiro, com base em medidas de análise multivariada. As medidas escolhidas foram as análises de similaridade e de dissimilaridade. São técnicas de classificação, de agrupamento. Como medida de dissimilaridade será usada a distância euclidiana e como medida de similaridade o coeficiente de correlação de Pearson. Assim, quanto maior a medida de dissimilaridade, mais afastado do melhor desempenho o time estará, ao passo que quanto maior a medida de similaridade, mais próximo do melhor desempenho a equipe estará.

Na análise da Tabela 7 pode-se verificar algumas diferenças nas classificações entre as medidas de similaridade e dissimilaridade. Tal fato pode ocorrer, pois são técnicas que utilizam diferentes metodologias para os cálculos, todavia, não é de se esperar mudanças muito grandes quando se avalia os dois resultados. Em todas as análises efetuadas no presente estudo, o São Paulo Futebol Clube (SPFC) ficou com a primeira posição, assim, tanto a medida de similaridade, quanto de dissimilaridade foram elaboradas com o SPFC como referência. Importante ressaltar que isso não foi escolha dos autores do trabalho, mas função da posição do time nas análises efetuadas.

Tabela 7. Medida de similaridade e dissimilaridade

Similaridade Pearson		Dissimilaridade distância euclidiana	
1: São Paulo	1	1: São Paulo	0
6: Corinthians	0,989	6: Corinthians	7,834
4: Santos	0,986	3: Internacional	12,64
3: Internacional	0,982	4: Santos	13,311
5: Fluminense	0,967	8: Grêmio	15,686
8: Grêmio	0,967	2: Cruzeiro	15,966
7: Flamengo	0,952	5: Fluminense	17,679
2: Cruzeiro	0,907	7: Flamengo	18,455
11: Palmeiras	0,905	10: Atlético Mineiro	22,397
10: Atlético Mineiro	0,904	11: Palmeiras	22,742
13: Botafogo	0,863	9: Atlético Paranaense	25,206
9: Atlético Paranaense	0,793	13: Botafogo	25,369

Continuação da Tabela 7...

Similaridade Pearson		Dissimilaridade distância euclidiana	
14:Vasco	0,601	12:Goiás	30,537
12:Goiás	0,551	14:Vasco	30,676
23:São Caetano	0,536	23:São Caetano	30,817
15:Coritiba	0,527	15:Coritiba	31,407
16:Figueirense	0,36	16:Figueirense	33,336
19:Paraná	0,036	21:Sport	37,824
21:Sport	-0,15	30:Portuguesa	39,339
20:Juventude	-0,288	19:Paraná	39,597
30:Portuguesa	-0,358	32:Ceará	40,013
32:Ceará	-0,4	20:Juventude	43,402
17:Vitória	-0,431	17:Vitória	45,496
34:Chapecoense	-0,565	27:Guarani	45,88
18:Ponte Preta	-0,57	22:Bahia	46,146
28:Paysandu	-0,642	18:Ponte Preta	47,065
27:Guarani	-0,672	34:Chapecoense	50,058
22:Bahia	-0,674	29:Fortaleza	51,608
25:Criciúma	-0,69	26:Avai	51,902
29:Fortaleza	-0,695	33:Barueri	54,196
26:Avai	-0,701	37:América Mineiro	56,525
35:Santo André	-0,729	28:Paysandu	56,843
24:Náutico	-0,747	25:Criciúma	61,835
31:Atlético-GO	-0,792	31:Atlético-GO	62,053
33:Barueri	-0,823	35:Santo André	71,083
38:Ipatinga	-0,826	36:Brasiliense	72,563
36:Brasiliense	-0,842	24:Náutico	77,703
39:Santa Cruz	-0,857	40:Joinville	82,01
41:América-RN	-0,861	38:Ipatinga	95,545
40:Joinville	-0,866	39:Santa Cruz	143,685
37:América Mineiro	-0,888	41:América-RN	293,277

Conclusões

O presente artigo fez uma análise do desempenho dos times de futebol, no campeonato brasileiro da primeira divisão. Para tanto, com base em informações disponíveis acerca do desempenho dos mesmos, foram criados alguns indicadores adicionais, com a finalidade de comparação relativa do desempenho dos mesmos. Os indicadores criados ajudaram na análise da efetividade, e inefetividade, dos times, de forma objetiva, ao longo do período considerado. Com base nos mesmos indicadores analisou-se a similaridade, medida pela correlação linear de Pearson, bem como da dissimilaridade, medida pela distância euclidiana. De uma maneira geral as análises mostraram que existe uma concentração, no eixo das regiões sul e sudeste, dos clubes com melhor desempenho. Além disso, pôde-se observar que o campeonato brasileiro, da primeira divisão, é bastante disputado, com mais de 10 times com desempenho muito próximos (medidos, por exemplo, pela similaridade). Uma aplicação prática interessante, do presente estudo seria a sua utilização para se detectar possíveis estratégias a serem adotadas visando ajustar o desempenho da equipe ao esperado. O trabalho pode,

também apresentar interessante *insight* no que se refere a possíveis estratégias no que se refere ao comportamento da equipe (jogar mais fechado buscando um empate, ou tentar ganhar arriscando não pontuar), conforme verificado com base no índice de aproveitamento. Finalmente, parece evidente que, no período considerado, que foi o período dos pontos corridos, o São Paulo Futebol Clube apresentou o melhor desempenho, por diversos critérios. Também chama a atenção, o desempenho do Sport Clube Corinthians Paulista, em termos de efetividade, o que pode ser verificado tanto pelo seu índice de aproveitamento, quanto pela medida de similaridade, ou mesmo pela medida de dissimilaridade.

Referências

1. Valentin BR, Coelho M. Sobre as escolinhas de futebol: processo civilizador e práticas pedagógicas. *Motriz Rev Edu Fis* 2005;11(3):185-197.
2. Reilly T, Gilbourne D. Science and football: a review of applied research in the football codes. *J Sports Sci* 2003;21(9):693-705.
3. Añon IC, Yamanaka GK, Machado JC, Scaglia A. Performance da equipe da Espanha e seus adversários nos jogos da Copa do Mundo FIFA 2010. *RBF* 2013;06(1):33-44.
4. Marcelino R, Sampaio J, Mesquita I. Investigação centrada na análise do jogo: da modelação estática à modelação dinâmica. *Rev Port Cien Desp* 2011;11(1):481-499.
5. De Araujo CTP, Tavares L, Alvares LG, Neto FL, Suzuki AK. Modelagem estatística para a previsão de jogos de futebol: Uma aplicação no campeonato brasileiro de futebol 2014. *Revista da Estatística UFOP* 2015;4(2):12-20.
6. Almeida LG, Oliveira ML, Silva CD. Uma análise da vantagem de jogar em casa nas duas principais divisões do futebol profissional brasileiro. *Rev bras educ fis esporte* 2011;25(1):49-54.
7. Haas DJ. Productive Efficiency of English Football Teams - A Data Envelopment Analysis Approach. *Manage Decis Econ* 2003;24(5):403-410.
8. Gómez R, Mendo H. Revisión de Indicadores de Rendimiento en Fútbol. *RICCAFD* 2012;1(1):1-14.
9. Lago-Ballesteros J, Lago-Peñas C. Performance in Team Sports: Identifying the Keys to Success in Soccer. *J Hum Kinet* 2010;25(1):85-91.
10. Lago-Peñas C, Lago-Ballesteros J, Rey E. Differences in performance indicators between winning and losing teams in the UEFA Champions League. *J Hum Kinet* 2011;27(1):135-146.
11. Vázquez AV, Gayo AA, Pita HB, Fernández CA. Diseño y aplicación de una batería multidimensional de indicadores de rendimiento para evaluar la prestación competitiva en el fútbol de alto nivel. *Int J Sport Sci* 2011;7(23):103-112.
12. Carvalho FM, Scaglia AJ, Costa IT. Influência do Desempenho Tático sobre o Resultado Final em Jogo Reduzido de Futebol. *Rev Educ Fis UEM* 2013;24(3):393-400.
13. Futdados.com [Internet]. Pontos corridos: Campanhas acumuladas desde 2003 - Todas as equipes. [acesso em: 17 jun 2015]. Disponível em: <http://futdados.com/pontos-corridos-campanhas-acumuladas/>.

14. Dattorro J. Convex Optimization & Euclidean Distance Geometry. Palo Alto, Califórnia: Meboo Publishing; 2008.
15. Hair Jr JF, Anderson RE, Tatham RL, Black WC. Multivariate Data Analysis. 5.ed. New Jersey: Prentice Hall; 1998.

Recebido em 04/10/15.

Revisado em 15/05/16.

Aceito em 02/06/16.

Autor para correspondência: Domingos Rodrigues Pandeló Júnior. Endereço: Rua Professor Reinaldo Porchat, 74 ap 12 - Vila Belmiro, Santos – SP. Email: pandelo@hotmail.com