

titulo

Andres Chacoma*

*Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales,
Departamento de Física. Buenos Aires, Argentina and
CONICET - Universidad de Buenos Aires,
Instituto de Física Interdisciplinaria y Aplicada. Buenos Aires, Argentina.*

Juan I. Perotti

*Universidad Nacional de Córdoba, Facultad de Matemática,
Astronomía, Física y Computación. Córdoba, Argentina and
CONICET- Universidad Nacional de Córdoba,
Instituto de Física Enrique Gaviola (IFEG). Córdoba, Argentina.*

Orlando V. Billoni

*Universidad Nacional de Córdoba, Facultad de Matemática,
Astronomía, Física y Computación. Córdoba, Argentina and
CONICET- Universidad Nacional de Córdoba,
Instituto de Física Enrique Gaviola. Córdoba, Argentina.*

Abstract

We ...

I. INTRODUCTION

II. DATA

A. Recopilación de métricas

En este trabajo utilizamos la base de datos de evento provista por L. Pappalardo et al en [1]. En ese artículo, los autores visualizan todos los partidos de la temporada 2017-2018 de las principales ligas de fútbol europeas: La Liga (España), Premier league (Inglaterra), Serie A (Italia), Bundesliga (Alemania), Ligue 1 (Francia). Por cada partido detectan, clasifican, y localizan en tiempo y espacio todos los eventos: goles, tiros al arco, pases, saques de esquina, faltas, etc. En el sistema de referencia que utilizan para ubicar los eventos en tiempo y espacio, t expresa el tiempo transcurrido desde el inicio del partido, la coordenada x expresa la distancia respecto del arco del equipo creador del evento, y la coordenada y la distancia respecto a la banda lateral derecha. Las unidades de las coordenadas espaciales están dadas en porcentajes del campo juego, siendo por ejemplo $x = 0$, $x = 50$ y $x = 100$ las posiciones de la línea de meta propia, la línea de centro del campo y la línea de meta del rival, respectivamente.

En este marco, definimos intervalo de posesión de pelota (BPI) como al conjunto dado por una secuencia continua de eventos generados por un equipo. Note que cada BPI contiene información de un solo equipo. Recopilamos todos los BPI de todos los equipos de cada liga, y sobre esos datos extrajimos métricas que nos permiten detectar algunos de los recursos tácticos que están utilizando los equipos en esa ventana temporal del partido. Las métricas recopiladas para nuestro análisis están basadas en las propuestas por J. Fernandez-Navarro en [2]. El estudio de estas métricas fue utilizado también en un trabajo anterior para determinar estilos futbolísticos característicos [3]. A continuación describimos en detalle a cada una de estas,

* achacoma@df.uba.ar

1. *Direct play.* Cada vez que hay un pase o un tiro libre en un BPI, medimos la velocidad media en la dirección de ataque, dada por el cociente entre la distancia recorrida por la pelota en el eje x y el tiempo transcurrido. De cada BPI tomamos el valor máximo. Esto nos permite detectar que tan directo hacia la portería rival es el movimiento de la pelota en el equipo.
2. *Counterattack.* Dado dos eventos consecutivos en un BPI, si el primero se observa en $x_1 < 40$ y el segundo en $x_2 > 60$ a una diferencia temporal Δt , se informa la velocidad como $v = \frac{x_2 - x_1}{\Delta t}$. En otro caso se informa 0. Esto es una medida de que tan rápido un equipo pasa de una posición defensiva en su campo a una ofensiva en campo rival.
3. *Maintenance.* Dado un BPI, se calcula el promedio de las posiciones en el eje x , donde se generaron todos los eventos. Si se cumple $\bar{x} < 40$, es decir si los eventos se desarrollaron mayoritariamente en la zona defensiva del equipo, se informa el tiempo total de la posesión. En otro caso se informa 0. Esto nos permite detectar que tanto un equipo decide mantener y construir su juego desde su propio campo.
4. *Build up.* Si en un BPI se verifica que $\bar{x} > 60$, es decir la posesión se desarrolla mayoritariamente en campo rival, se informa el tiempo total de la posesión. En otro caso se informa 0. Esta métrica informa el tiempo de posesión en situaciones donde el equipo invade fuertemente el campo rival.
5. *Midfield play.* Si en un BPI se observa que $\bar{x} \leq 60$ y $\bar{x} \geq 40$, es decir la posesión se desarrolla mayoritariamente por el centro del campo de juego, se informa el tiempo total de la posesión, en otro caso se informa 0. La idea de esta variable es medir el tiempo que el equipo pasa en el sector medio del campo de juego.
6. *Flow rate.* En cada BPI donde $\bar{x} \geq 50$ se toma la diferencia temporal entre todos los eventos, y se calcula el valor medio, \bar{dt} . Luego se define la métrica como $1/\bar{dt}$. De esta manera se tiene una medida de que tan rápido el equipo mueve la pelota en el campo rival.
7. *Crossing.* Si en un BPI se observa un evento centro, se informa 1, en otro caso se informa 0. Esta métrica sirve para contabilizar los intentos de llegada por vía aérea.

8. *Pressure point*. De cada BPI se toma el primer evento, y se extrae la posición en la variable x , es decir donde el equipo comienza su posición. Esto nos permite medir si el equipo esta recuperando la pelota en su campo, en la zona media o en el campo rival.
9. *Pressure loss*. Si un BPI comienza en un evento donde $x > 40$, se informa el tiempo total de la posesión del adversario del BPI inmediatamente anterior. Este métrica es útil para observar si el equipo esta relajando o aumentando el nivel de presión que ejerce sobre el juego del rival en las zonas medias y altas del campo de juego.
10. *Shots*. Si en un BPI se registra un evento “Shot”, se informa 1, en otro caso se informa 0. Esta métrica permite contabilizar los tiros al arco de cada equipo.

Para nuestro análisis, se descartaron todas los BPI con menos de 3 eventos y con tiempo total menor a 2 segundos. La idea de esto es descartar pequeñas recuperaciones pasajeras y quedarnos con posesiones consolidadas. Del proceso de recopilación se obtuvieron 215681 BPI. Luego de calcular los valores de las métricas en cada uno de estos, estudiamos la distribución de los datos. Observamos que las métricas parecen seguir una distribución tipo log-normal, por lo tanto decidimos transformar los datos como $x \rightarrow \log(1 + x)$ para trabajar con distribuciones aproximadamente normales. Posteriormente, agrupamos la información por partido y por equipo, y sumamos los valores obtenidos en cada métrica. De esta manera por ejemplo el feature *Shots* cuantifica la cantidad de tiros al arco ejecutados por el equipo en ese partido. Asimismo el feature *Build-up* cuantifica la cantidad de tiempo neto en la cual un equipo sostuvo una posición de ataque frente al rival en ese partido. Note que en la temporada 2017/2018, en las ligas Española, Inglesa, Francesa e Italiana los equipos jugaron 38 partidos. Por lo tanto, al tomar los datos de los primeros 4 equipos cada liga aporta un total de $38 \times 4 = 152$ muestras al archivo de datos. Asimismo, en la liga alemana, al ver menos equipos, se jugaron 34 partidos, por lo tanto esta liga aporta 136 muestras. En consecuencia, la matriz de datos consta de 744 filas y 10 columnas. Por ultimo, en un dataset aparte recopilamos meta-data asociada a cada muestra, útil luego para realizar el análisis: a que equipo pertenece esa muestra, cual es la liga de pertenencia, y el resultado final en la tabla de posiciones.

B. Representación en redes complejas

En lo que sigue, presentamos nuestra propuesta para representar las métricas de rendimiento en términos de redes complejas. Definimos $M(i, j, g)$ como la métrica de rendimiento correspondiente al equipo i cuando enfrenta al equipo j en el partido g . Por ejemplo, puede representar la cantidad de tiros al arco realizados por el FC Barcelona al jugar contra el Real Madrid en el primer encuentro del torneo español *La Liga*. En nuestro conjunto de datos, todos los equipos participaron en un formato de liga todos contra todos, enfrentándose dos veces: un partido de ida (g_1) y otro de vuelta (g_2). Utilizando la información de ambos encuentros, definimos una métrica agregada que resume el desempeño observado entre esos dos equipos a lo largo del torneo:

$$M(i, j) = \sum_{g_1, g_2} M(i, j, g).$$

En el ejemplo anterior, $M(i, j)$ representa la cantidad total de tiros al arco realizados por el FC Barcelona contra el Real Madrid en ambos partidos del torneo. Calculando $M(i, j)$ para cada par de equipos en una liga L , es posible representar estas relaciones de desempeño mediante un grafo dirigido y ponderado $G(L, M)$, cuyos pesos se definen como

$$f_{ij} = M(j, i) - M(i, j).$$

Nótese que, en esta representación, $f_{ij} < 0$ indica que el equipo i superó al equipo j en la métrica considerada. En este marco, construimos 50 grafos en total, diez grafos por liga. Cada uno asociado a una métrica de rendimiento distinta.

III. THEORY

IV. RESULTADOS

A. Estadística del rating real

La idea de esta sección es definir un modelo estadístico para la distribución de probabilidad del rating real, R_T , definido como la cantidad de puntos totales obtenidos por un equipo durante la liga. R_T es una variable estocástica que depende de la cantidad de par-

tidos ganados, empatados y perdidos, por los equipos de la liga. En las ligas de futbol, un equipo obtiene 3 puntos cuando gana, 1 cuando empata y 0 cuando pierde. En este marco, R_T puede expresarse en terminos de las variables estocasticas W y D , que representan el numero de partidos ganados y empatados, respectivamente,

$$R_T = 3W + D.$$

Sea n el número total de partidos en la temporada. Modelamos W con una distribución binomial,

$$W \sim \text{Bin}(n|p_w)$$

donde p_w es la probabilidad de ganar un partido. Condicional al número de victorias $W = w$, los empates siguen,

$$D|W = w \sim \text{Bin}(n - w|\frac{p_d}{1 - p_w}),$$

con p_d probabilidad de empatar. Dado que los equipos tienen distinto nivel competitivo, cada equipo posee probabilidades intrínsecas p_w y p_d de ganar y empatar, respectivamente. Modelamos estas probabilidades mediante un enfoque jerárquico: cada equipo tiene un par de habilidades latentes $\boldsymbol{\eta} = (\eta_w, \eta_d)^T$, de ganar o empatar, que siguen una distribución normal bivariada,

$$\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

$$\boldsymbol{\mu} = (\mu_w, \mu_d)^T, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_w^2 & \rho\sigma_w\sigma_d \\ \rho\sigma_w\sigma_d & \sigma_d^2 \end{pmatrix},$$

donde $\boldsymbol{\mu}$ es el centro de la distribucion, $\boldsymbol{\Sigma}$ la matriz de covarianza. En este marco, ρ representa la correlación entre las habilidades para victoria y empate. Las probabilidades se obtienen mediante la transformación softmax,

$$p_w = \frac{e^{\eta_w}}{1 + e^{\eta_w} + e^{\eta_d}}, \quad p_d = \frac{e^{\eta_d}}{1 + e^{\eta_w} + e^{\eta_d}}.$$

Esta transformación garantiza que $p_w + p_d \leq 1$ y que el modelo sea identificable al fijar implícitamente $\eta_l = 0$ para las derrotas. Con estos elementos, podemos escribir la distribución conjunta de victorias y empates marginalizando sobre los efectos aleatorios $\boldsymbol{\eta}$,

$$P(W = w, D = d) = \int_{\mathbb{R}^2} \text{Bin}(n|p_w(\boldsymbol{\eta})) \text{Bin}(n - w|\frac{p_d(\boldsymbol{\eta})}{1 - p_w(\boldsymbol{\eta})}) \phi(\boldsymbol{\eta}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) d\boldsymbol{\eta}, \quad (1)$$

Notar que la integral en 1 no tiene forma cerrada, para resolverla utilizamos el metodo de monte carlo. Con estos elementos, podemos finalmente escribir la distribucion de probabilidad teorica para el true rating como,

$$P(R_T = r_T) = \sum_X P(W = w, D = d), \quad (2)$$

$$X = \{w, d > 0, 3w + d = r_T, w + d \leq n\}.$$

Los grados de libertad que otorgan los parametros μ_w , μ_d , σ_w , σ_d y ρ , permiten ajustar la curva teorica 2 a la curva empirica. El ajuste fue realizado sobre el conjunto de datos asociado a las ligas Inglesa, Francesa, Italiana y Española. Para este analisis, decidimos no utilizar los datos de la liga Alemana, ya que su torneo se disputa entre un numero menor de equipos (18), y por lo tanto se genera una distribucion de valores de true rating menor respecto de las otras ligas, en las cuales participa un total de 20 equipos por liga. Para hacer el ajuste, utilizamos el algoritmo Nelder-Mead minimizando el RMSE entre las CDF empírica y teórica. Posteriormente, con los parámetros óptimos, realizamos bootstrap paramétrico (1000 réplicas) para estimar el valor medio y la incertidumbre de las probabilidades p_w y p_d . Para la probabilidad de victoria obtuvimos $\bar{p}_w = 0.366$ ($SD = 0.010$) con un intervalo de confianza del 95% de $[0.3527, 0.3798]$. Asimismo, para la probabilidad de empate obtuvimos $\bar{p}_d = 0.280$ ($SD = 0.020$) con un intervalo de confianza del 95% de $[0.2548, 0.3074]$. Para contrastar estos resultados con los datos empiricos, utilizando la informacion de los resultados de los partidos, calculamos para cada equipo i las probabilidades de victoria y empate empiricas, q_w y q_d ,

$$q_w^{(i)} = \frac{\#wins^{(i)}}{\#matchs^{(i)}}, \quad q_d^{(i)} = \frac{\#draws^{(i)}}{\#matchs^{(i)}}.$$

Calculando el valor medio y el desvio estandar sobre el conjunto de todos los equipos, obtuvimos para la probabilidad de victoria $\bar{q}_w = 0.379$ ($SD = 0.168$) con un intervalo de confianza del 95% de $[0.1572, 0.7638]$, y para la probabilidad de empate $\bar{q}_d = 0.243$ ($SD = 0.081$) con un intervalo de confianza del 95% de $[0.1046, 0.3954]$. Podemos observar, (1) Las medias son similares (0.366 vs 0.379 para victorias, 0.280 vs 0.243 para empates), lo que valida el centro de la distribución. (2) las desviaciones estándar empíricas (0.168, 0.081) son mucho mayores que las implícitas en el modelo (0.01, 0.02), indicando que el modelo subestima la heterogeneidad entre equipos.

Con respecto al parametro de correlacion, del bootstraping parametrio obtuvimos $\bar{\rho} = -0.130$ ($SD = 0.07$) con un intervalo de confianza del 95% de $[-0.3040, -0.0371]$. Esto indica la presencia de una correlacion estructural negativa debil entre las habilidades de victoria, η_w , y empate, η_d , de los equipos. Esto sugiere que equipos con mayor habilidad para ganar tienden ligeramente a tener menor habilidad para empatar.

En la Fig. 1 (a), mostramos la CDF del true rating junto con el ajuste. Podemos ver que el modelo propuesto captura muy bien el comportamiento de la curva en todo el soporte. En Fig. 1 (b) mostramos un grafico con la relacion entre los quantiles teoricos y de los datos. A modo de referencia, mostramos tambien la relacion entre los quantiles de los datos y los asociados a una distribucion gaussiana con media y desvio estandar igual a los valores muestrales obtenidos de los datos, $\bar{R}_T = 52.45$ ($SD = 18.56$). En primer lugar, podemos ver que el modelo propuesto captura bien el comportamiento en la zona central izquierda y de manera aceptable el comportamiento de la cola de la derecha. La comparaci3n con el modelo gaussiano nos permite ver que la distribucion real presenta colas mas livianas a la izquierda y mas pesadas a la derecha, respecto del comportamiento gaussiano, develando as3 la presencia de una asimetria en la distribucion. La concabidad hacia abajo en el centro del grafico, indica tambien que la mediana de los datos esta corrida hacia la izquierda.

B. Comparacion entre el rating real y el rating obtenido de las metricas

En esta secci3n, la idea es caracteriar las diferencias entre el true rating, dado por el sistema de puntos de las ligas, y el rating obtenido a partir del metodo de HodgeRank. En este marco, definimos como el metric rating como el valor de rating calculado, para un equipo de una liga, con metodo HodgeRank utilizando la red asociada a una metrica en particular. Dado que son ocho las metricas utilizadas (ver seccion II A), por cada equipo calcularemos ocho valores de metric rating distintos, que compararemos con el true rating. Es importante remarcar que el true rating y el metric rating estan expresados en escalas de valores distintas. El primero esta expresado en cantidad de puntos obtenidos y el segundo en un numero estandarizado. No obstante esto no nos impide compararlos, ya que lo importante de estos valores no es el valor absoluto que obtiene cada equipo, si no mas bien el valor relativo entre ellos. Por tal motivo, para comparar el comportamiento del true rating con el metric rating, decidimos estandarizar ambas variables y trabajar con el *z-score* de cada una

de estas. En lo siguiente discutimos algunos resultados interesantes obtenidos al comparar true rating con metric rating.

En la Fig. 2 (a) mostramos las CDF del true rating asociado a los datos de todos los equipos y todas las ligas, con la del metric rating asociado a los datos de todos los equipos de todas las ligas y de las ocho metricas estudiadas. A ambas curvas se les estrajeron los valores outliers considerados como aquellos casos que se encuentran por encima del quantile $Q_{99.7}$. Podemos observar una coincidencia notable en casi todo el rango, observando algunas diferencias en las colas: el metric rating muestra valores positivos mas alejados de la media, y el true rating muestra valores negativos levemente mas alejados de la media. A continuacion evaluamos la distribucion de valores de metric rating de cada una de las ligas. Para esto, por liga, agrupamos los valores de metric rating asociado a todas las metricas y estandarizamos los valores utilizando el z-score. En la Fig. 2 (b) se muestran los resultados. A modo de referencia se grafica tambien, en linea negra punteada, la CDF de una distribucion normal estandar. Si bien vemos que las curvas asociadas a las diferentes ligas difieren levemente, indicando un comportamiento dependiente de la liga, para todos los casos podemos ver una acumulacion de probabilidad en la zona media. Esto se compensa con colas livianas a izquierda y pesadas a la derecha, siempre utilizando como referencia la distribucion normal estandar. Por otro lado, en la Fig. 2 (c) graficamos las curva true rating y metric rating en función del true ranking, para el caso de la metrica *Pressure Point* en la liga Inglesa. La idea es evaluar cualitativamente las diferencias observadas equipo a equipo. En primer lugar podemos notar que el true rating decrece monotonamente a medida que crece el ranking, esto es trivial por defnicon: el primer equipo tiene un true rating mayor o igual que el segundo, el segundo tiene un true rating mayor o igual que el tercero, y asi hasta el ultimo puesto. En la curva de metric rating observamos una tendencia tambien decreciente, pero no monotona. En este caso, podemos decir que en terminos generales el metric rating sigue la tendencia del true rating, a menos de esas diferencias locales. En la Fig. 2 (d) mostramos una comparacion analoga a la anterior para el caso de la metrica *Direct play* en la liga española. En este caso vemos que la curva de metric rating no sigue una clara tendencia decreciente y se observa una fuerte descorrelacion con los valores de true rating. Vemos entonces que el true rating de una dada liga puede diferir poco o mucho del metric rating dependiendo de la metrica que utilizamos para inferir el rating.

Nuestro objetivo ahora es cuantificar la correlación entre la curva de true rating de una

liga y las curvas de metric rating asociadas a cada metrica. Nuestra idea es cuantificar con que metrica se obtiene el rating mas sigue la tendencia del real, y si eso depende de la liga. Para esto, estudiamos el coeficiente de correlacion de Pearson entre las curvas de true rating de cada liga y la de metric rating en cada metrica. En Fig. 3, mostramos los resultados. Cada panel contiene la información asociada a una metrica, y dentro de los panes, cada barra contiene la información asociada a una liga. Los paneles se muestran ordenados de mayor a menor por correlacion promedio y las barras, dentro de cada panel, odenadas de mayor a menor por correlacion. Las barras con hatch indican que para esos casos no se puede asegurar significancia estadistica. En primer lugar podemos ver que las ligas italiana e inglesa muestran la mayor correlacion en todas las metricas. En segundo lugar, vemos que las ligas frances y alemana se alteran el tercer y cuarto puesto entre los casos con mas correlacion en casi todas las metricas. Por ultimo, vemos que la liga española muestra los peores resultados, mostrando el peor valor de correlacion en 7 de las 8 metricas. Aun así, presenta valores que indican correlacion fuerte ($\rho > 0.4$) en los casos de las metricas *Flow rate* y *Counterattack*.

C. Comparacion entre el true ranking y el ranking obtenido de las metricas

La idea de esta sección es complementar el analisis de la seccion anterior haciendo una comparacion utilizando los rankings. La posicion final de un equipo en una liga se obtiene al ordenar los equipos de esa liga de manera descendente por cantidad de puntos obtenidos. Ese ordenamiento define lo que llamaremos el true ranking de esa liga. Analogamente, para los equipos de una liga, definimos el metric ranking ordenando los equipos de manera descendente por valor de metric ranking obtenido con el metodo HodgeRank. Note que por cada liga tenemos ocho ranking en total. De manera similar a lo realizado en la seccion anterior, compararemos el true ranking de los equipos con los metric rankings obtenidos a partir del metodo HodgeRank. Para esto utilizaremos el coeficiente de correlacion de rango Kendall, τ [ref].

D. Composite rating

En esta seccion, desarrollamos para cada liga un Composite Rating (CR), definido como una combinacion lineal de los metric rating obtenidos de las metricas de rendimiento. La idea es proporcionar una evaluacion mas general de la calidad de los equipos. Sea n la cantidad de equipos en la liga y m la cantidad de metricas, definimos $R^{n \times m}$ como la matriz donde las filas contienen los valores de metric rating obtenidos para cada equipo. Asimismo definimos el $T^{n \times 1}$ como el vector que contiene los true ratings de cada equipo de una liga. En este marco, nuestro objetivo es encontrar el vector de pesos $\alpha^{m \times 1} = (\alpha_1, \dots, \alpha_m)^T$ tal que,

$$T = \alpha R.$$

Para resolver este sistema, y con la de idea de garantizar la parsimonia y la capacidad explicativa del Composite Rating (CR), empleamos un algoritmo de Selección Exhaustiva de Características (Exhaustive Feature Selector o EFS). Luego de normalizar las metricas mediante un escalado estandar ($Z - score$), este procedimiento evalua todas las combinaciones posibles de las ocho métricas de rendimiento iniciales, desde modelos univariados hasta el modelo completo con ocho variables. Con el fin de evitar el sobreajuste, utilizamos el Criterio de Información Bayesiano (BIC) como función de selección. El BIC permite identificar el modelo que maximiza la verosimilitud de los datos introduciendo una penalización proporcional a la complejidad del modelo (número de variables), favoreciendo así la estructura más eficiente. En este marco, el modelo óptimo seleccionado fue aquel que minimiza el valor del BIC. De esta manera aseguramos que el rating final sea estadísticamente robusto y representativo del rendimiento real de los equipos.

Para determinar la composición del Composite rating en cada liga, se ajustó un modelo de regresión por mínimos cuadrados ordinarios (OLS) utilizando las métricas seleccionadas por el algoritmo EFS. Dado que las variables fueron estandarizadas previamente, los coeficientes obtenidos representan la importancia relativa de cada dimensión en la predicción de los puntos obtenidos en la liga. En la tabla I, se muestran los resultados de la regresion. El análisis de los coeficientes revela áreas críticas para la mejora del rendimiento. Para la liga inglesa, podemos ver que las metricas mas relevantes, en orden de importancia relativa, son *Build up*, *Flow rate* y *Pressure point*, las cuales explican casi el 90% de la varianza de los datos. El coeficiente negativo en la métrica *Build up* (-24.56) sugiere que los equipos que reducen sus tiempos de elaboración en el ultimo cuarto del campo de juego, incrementan

significativamente su probabilidad de obtener puntos. Por el contrario, un aumento en esta métrica penaliza el Composite Rating, lo que indica una ineficiencia en estilos de juego que utilicen esta táctica en esa liga en particular. También podemos decir que el coeficiente positivo en *Flow rate* (22.69) sugiere que los equipos que incrementan la velocidad con la que mueven la pelota en el campo de juego, incrementa sus chances de obtener más puntos en esta liga. Asimismo, el coeficiente positivo en la métrica *Pressure point* sugiere que los equipos que presionan más alto en el campo de juego incrementa su probabilidad de obtener puntos. Con respecto a la liga francesa, haciendo un análisis similar basado en el análisis de los coeficientes de la regresión, podemos ver que los equipos que presionan alto en el campo de juego y disminuyen su juego directo, aumentan sus chances de obtener puntos. El caso de la liga alemana es probablemente el más curioso, note que con una sola variable, *Direct play*, es posible explicar el 46% de la varianza. Es decir, entender esta variable, nos permite intuir por qué un equipo gana o pierde puntos en esa liga. En este caso encontramos un modelo extremadamente parsimonioso. Lo obtenido para la liga italiana sugiere que el éxito estaría asociado a un incremento en las variables *Flow rate* y *Direct play* y un decremento en *Midfield play*. Es decir, los equipos más exitosos de esta liga parecen estar relacionados con un estilo de juego que fomenta el movimiento rápido de la pelota y relega el tiempo de control en la mitad del campo. Por último, en el caso de la liga española podemos ver que un estilo de juego asociado a la obtención de puntos parece estar relacionado con una disminución los niveles de presión alta, y un aumento de la velocidad con la que se mueve la pelota y de los contrataques.

El composite rating obtenido para cada liga puede ser comparado con el true rating como se hizo en las secciones anteriores con los ratings asociados a las métricas individuales. En Fig. 5 paneles (a), (b), (c), (e) y (e) mostramos la comparación liga a liga, equipo a equipo, en un gráfico rating contra true ranking. En el caso de la liga inglesa, en Fig. 5 (a), podemos ver que para los primeros dos equipos el true rating es mayor al composite rating. Estos equipos muestran entonces un supra-rendimiento estadístico, parecen estar rindiendo más de lo que deberían. Esto puede deberse a diversos factores externos atribuibles al azar. Contrariamente el equipo que quedó tercero en la tabla, muestra un true rating menor al composite rating, lo cual indica un comportamiento infra-estadístico, es decir, parece ser que hicieron lo estadísticamente correcto pero fallaron en la ejecución o simplemente tuvieron "mala suerte". Ejemplos similares se observan en todas las ligas, en todas las posiciones de

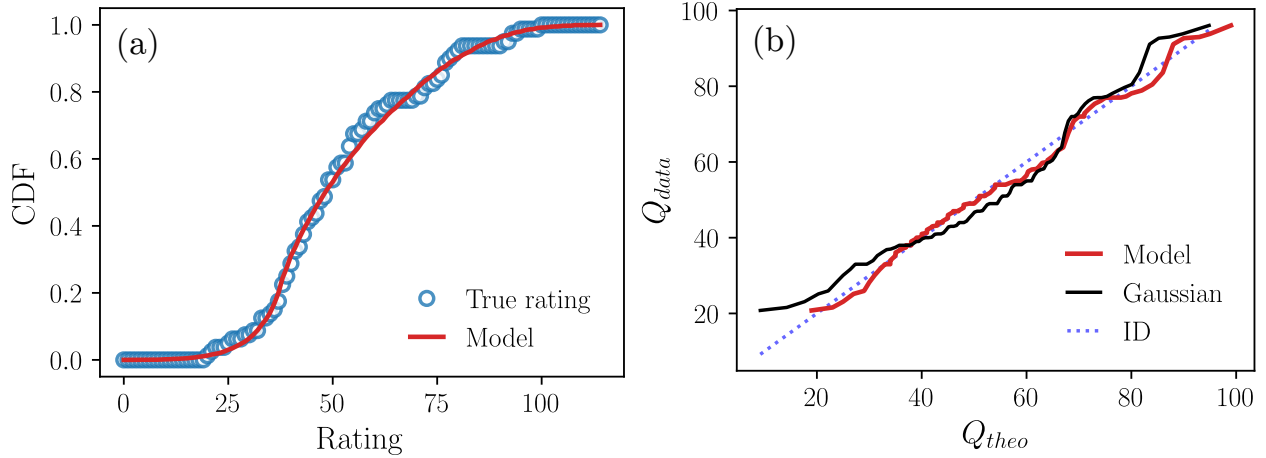


FIG. 1. Estadística del rating real. (a) Comparación de la Cumulative distribution function (CDF) de los valores de rating real con lo obtenido a partir del modelo. (b) Relación entre los Quantile asociado a los datos y los obtenidos a partir del modelo. A modo de referencia, se agrega también una comparación con una distribución gaussiana de media y desvío estándar igual al de los datos.

la tabla. En términos generales podemos decir que las curvas de composite rating siguen una tendencia similar a las curvas de true rating, destacándose a ojo desnudo los casos de la liga inglesa e italiana. En Fig. 5 (f) mostramos un gráfico de barras con los valores de los coeficientes de correlación de Pearson, ρ , y del coeficiente de Kendall, τ . En todos los casos, podemos ver que los valores obtenidos son superiores a los obtenidos a partir de las métricas individuales (ver Fig. 3 y Fig. 4). Destacamos principalmente los valores obtenidos en la liga Italiana y las mejoras significativas en las ligas española y alemana.

V. DISCUSION Y CONCLUSION

-
- [1] Luca Pappalardo, Paolo Cintia, Alessio Rossi, Emanuele Massucco, Paolo Ferragina, Dino Pedreschi, and Fosca Giannotti. A public data set of spatio-temporal match events in soccer competitions. *Scientific data*, 6(1):236, 2019.
 - [2] Javier Fernandez-Navarro, Luis Fradua, Asier Zubillaga, and Allistair P. McRobert. Influence of contextual variables on styles of play in soccer. *International Journal of Performance Analysis in Sport*, 18(3):423–436, 2018.

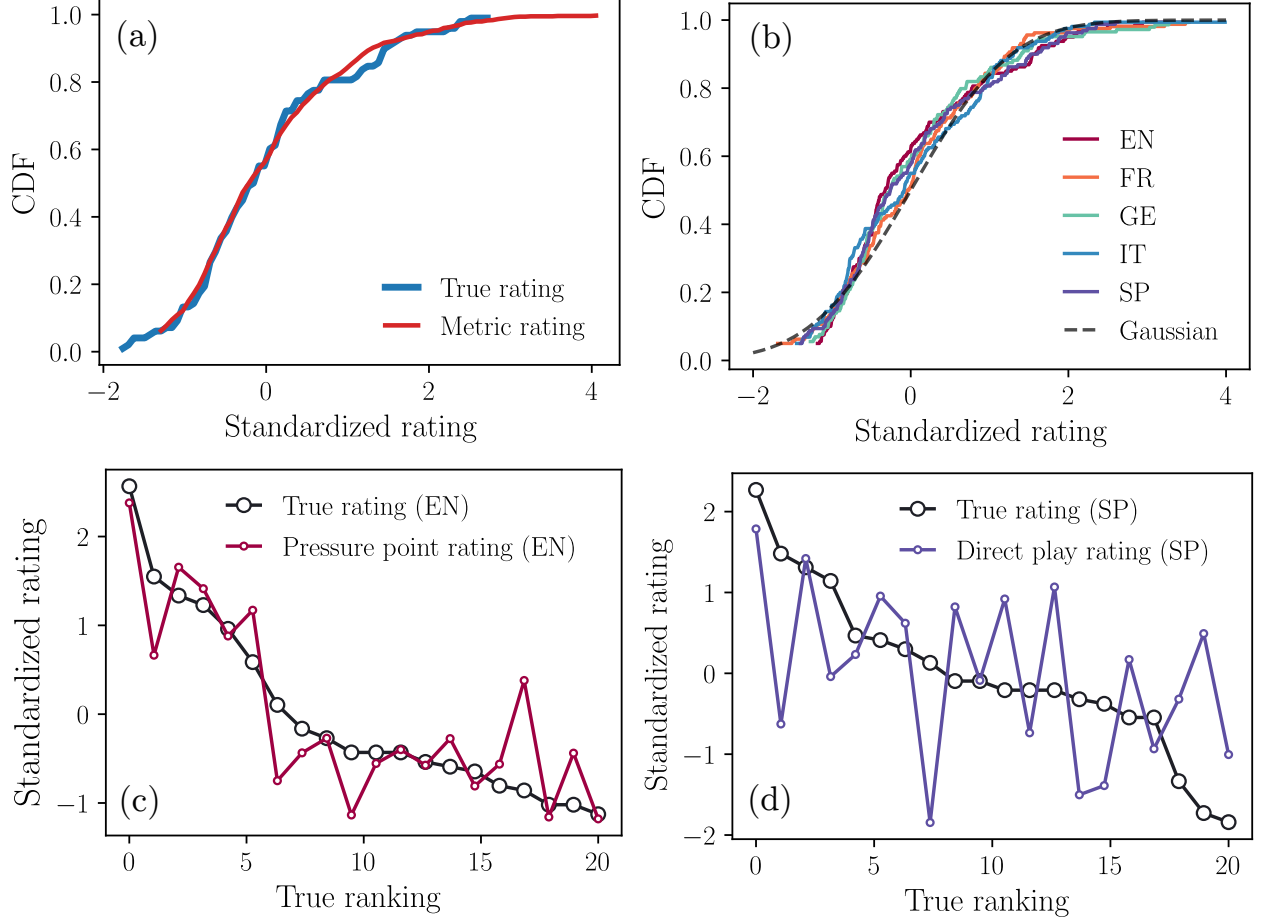


FIG. 2. Comparacion entre true rating y metric rating. Debido a que estas cantidades estan en unidades distintas, para comparar se usan los datos estandarizados. (a) Comparacion de las cumulative distribution function asociadas al true rating y al metric rating. (b) Cumulative distribution function del metric rating por liga. En lineas negras punteadas se grafica a modo de referencia la curva de una distribucion normal estandar. (c) Comparación en el caso Pressure point en la liga inglesa. (d) Comparación en el caso Direct play en la liga española.

[3] Andres Chacoma and Orlando V Billoni. Data-driven approach to defining football styles in major leagues. *Chaos, Solitons & Fractals*, 200:116926, 2025.

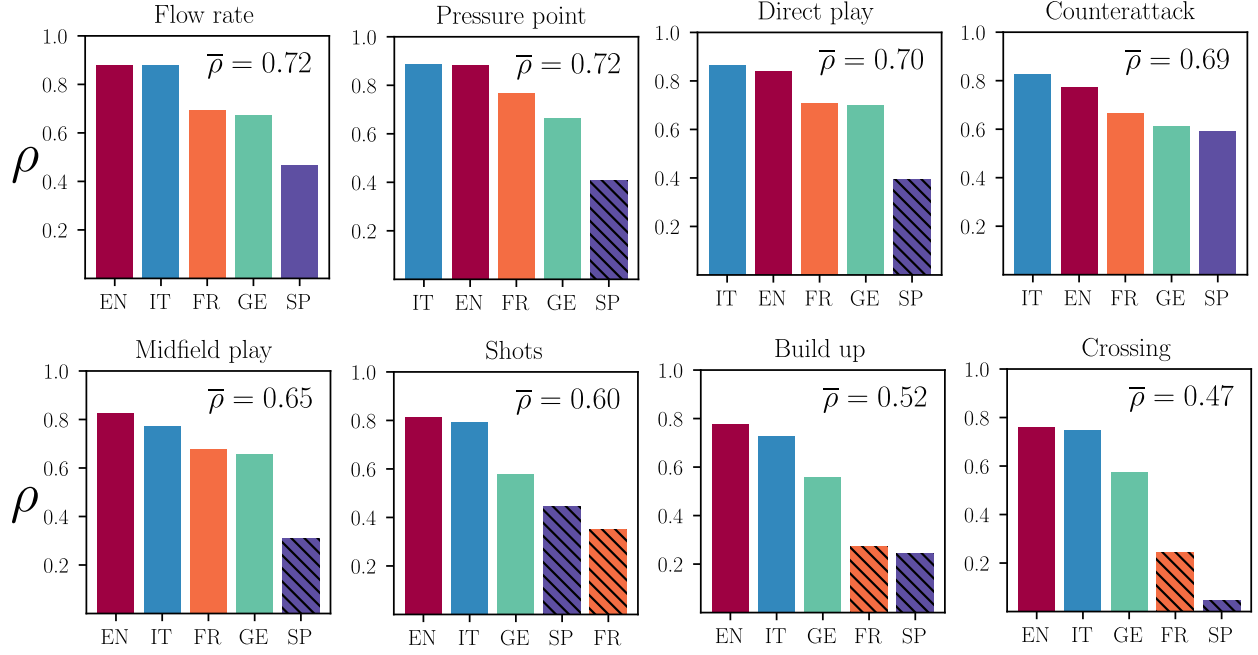


FIG. 3. Coeficiente de correlación de person, ρ , entre el rating verdadero y el rating dado por las metricas. Cada panel contiene la informacion de una metrica y las barras muestran el valor de ρ obtenido en cada liga. Las barras en cada panel estan ordenadas en orden creciente respecto del valor de ρ . Asimismo, los paneles estan ordenado en orden creciente respecto del valor promedio en cada metrica, $\bar{\rho}$. Las barras con hatch indica los casos donde el $p - value > 0.05$. En esos casos no se encontró una correlación estadísticamente significativa entre el true rating y el metric rating.

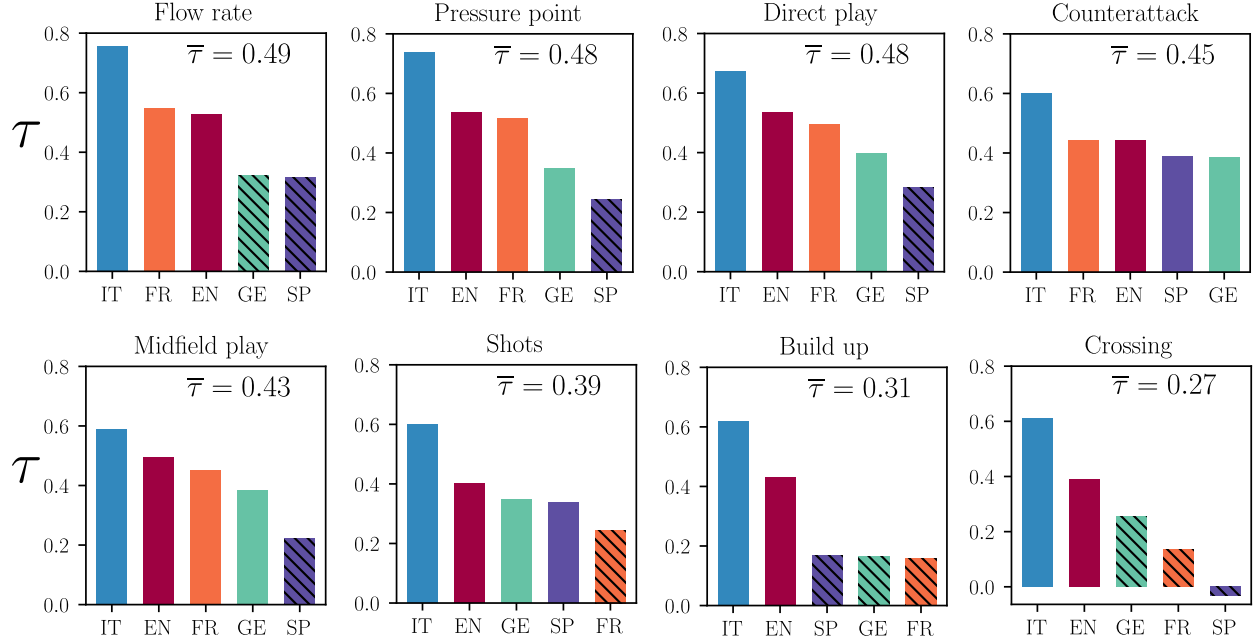


FIG. 4. Coeficiente de Kendall, τ , entre los rankings verdaderos y los rankings dados por las metricas. Cada panel contiene la informacion de una metrica y las barras muestran el valor de τ obtenido en cada liga. Las barras en cada panel estan ordenadas en orden decreciente respecto del valor de τ . Asimismo, los paneles estan ordenado en orden decreciente respecto del valor promedio en cada metrica, $\bar{\tau}$. Las barras con hatch indica los casos donde el $p - value > 0.05$. En esos casos, los resultados no proporcionan evidencia suficiente para afirmar la existencia de una relación lineal entre el ranking real y el inferido.

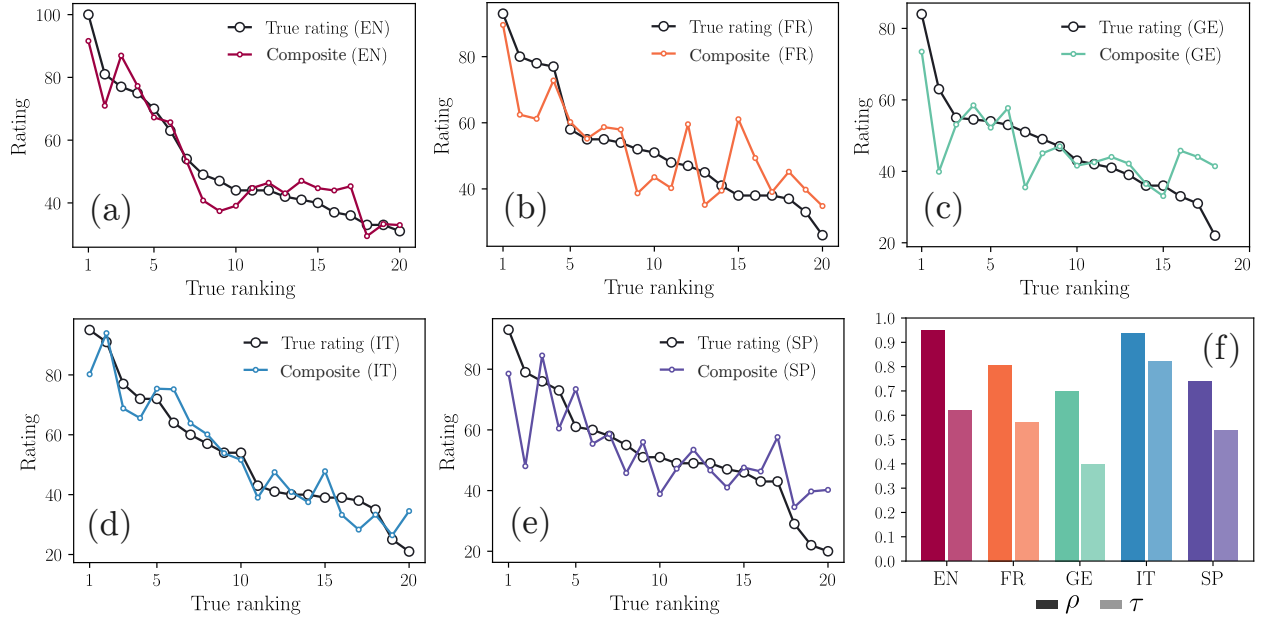


FIG. 5. Comparación entre el true rating y el composite rating. En los panes (a), (b), (c), (d) y (e) comparamos el true rating con el composite rating de cada uno de los equipos en las ligas inglesa, francesa, alemana, italiana y española respectivamente. En el panel (f) mostramos los valores del coeficiente de Pearson, ρ , y el coeficiente de Kendall, τ , obtenido para cada liga. Notar que, en cada caso, la barra de la izquierda se corresponde con el valor de ρ y la de la derecha con el valor de τ .

TABLE I. Componentes del Composite rating por Liga (basado en selección BIC)

Metric	Coef.	P-val	Imp. (%)	R^2
<i>England</i>				
Build up	-24.56	0.0006	37.65	0.88
Flow rate	22.69	0.0045	34.79	
Pressure point	17.98	0.0117	27.57	
<i>France</i>				
Pressure point	33.68	0.0145	61.71	0.61
Direct play	-20.90	0.1096	38.29	
<i>Germany</i>				
Direct play	9.48	0.0012	100	0.46
<i>Italy</i>				
Flow rate	17.84	0.0006	37.15	0.86
Direct play	15.46	0.0018	32.18	
Midfield play	-14.73	0.0119	30.67	
<i>Spain</i>				
Pressure point	-34.93	0.0185	43.08	0.46
Flow rate	26.38	0.0297	32.53	
Counterattack	19.78	0.0046	24.39	