Statistical model predicts number of goals for each footballer

May 29 2013

Credit: Christopher Johnson

After analysing football league players over nine seasons, from 2000 to 2009, Spanish researchers have come up with a mathematical method for estimating the goal-scoring performance of each player. According to their ranking, the most able strikers are Messi, Ronaldo Nazário, Makkay, Villa and Etoo.

Over nine seasons, researchers from the faculties of Economic and Business Sciences at the Universities of Granada and Jaén analysed the performance of football players in the Spanish league from 2000/2001 to 2008/2009, with the aim of creating a mathematical model to evaluate their goal-scoring potential.

Their paper, published in the *European Journal of Sport Science*, presents a model based on Bayesian statistics, which they say enables them to predict the number of goals each player will score based on their own individual qualities.

As the researchers explain, whether a player scores or not depends on easily quantifiable extrinsic factors, such as the number of minutes or matches played, the position on the pitch - defender, midfielder or striker - and the team quality (measured by their position in the league table). But this is not everything.

"Other factors that are harder to measure can also influence performance, such as the individual characteristics of the player that make him different from all the rest," explains to SINC José María Pérez Sánchez, professor at the institution in Granada and co-author of the study.

Extrinsic factors aside, the model can quantify an additional individual factor that influences the qualities of each footballer as a goal-scorer. "With this, we can rank Spanish league players according to this factor of individual ability," the authors assert.

The best players

The model enables a player's goal-scoring performance over the nine seasons analysed to be evaluated, by comparing the number of goals scored by the player with the number that would be expected of a

footballer playing at the same time, in the same position and on the same team.

The results reveal that the five best defenders in the seasons studied were, in descending order, Ezequiel Garay, Roberto Carlos, Campano, Cristian Álvarez and Larrazábal.

According to the classification resulting from the model, the best midfielders were Rivaldo, Robert, Luis Cembranos, Mark González and Mostovoi; and the most able strikers were Messi, Ronaldo Nazário, Makkay, Villa and Etoo.

In terms of the overall assessment of all the players in the football league over the nine seasons studied, the best was Rivaldo, followed by Robert, Ezequiel Garay, Luis Cembranos and Roberto Carlos.

"It is remarkable that the strikers do not rank very highly in terms of overall performance," Pérez Sánchez continues. "In fact, apart from Messi, who ranks sixth, the rest all rank below 15th place."

The authors suggest that the reason might lie in the fact that the model takes into account that they are strikers and assesses them more harshly in relation to the expected number of goals. Furthermore, the ranking particularly values goals scored by defenders and midfielders, since this is not their primary objective on the pitch.

The Messi paradox

Lastly, the study analyses in depth the careers of several emblematic football players to assess their goal-scoring performance over this period, drawing attention to cases such as that of Leo Messi.

"Messi's development is interesting," the researchers highlight.

"Although his performance is high in early seasons, in 2007/2008 (the penultimate season of the study) his productivity decreases significantly."

The experts point out that this must be due to his change of position on the pitch, because he had previously been considered a midfielder but became a striker, demanding higher goal-scoring ability.

"If the latest seasons, in which the player's goal-scoring success has increased considerably, were included in the study, this effect would surely be offset, and Messi would clearly rank many places higher in the overall assessment," concludes Pérez Sánchez.

Provided by Plataforma SINC