

## A BRIEF REVIEW OF SOCCER PERFORMANCE ANALYSIS MODELS

<sup>1</sup>Junshi Liu, <sup>1</sup>Andrew A. Pustina, <sup>1</sup>Kyle Uptmore, <sup>1</sup>Hugo Santana, <sup>1</sup>Evan Harvey, <sup>1</sup>Matt L. Sams

<sup>1</sup> Center of Excellence in Sport Science and Coach Education, East Tennessee State University, Department of Exercise and Sport Sciences, Johnson City, TN.

**INTRODUCTION:** Soccer performance is generally divided into physical abilities (Di Salvo, Pigozzi, González-Haro, Laughlin, & De Witt, 2013), technical skills (Rampinini, Impellizzeri, Castagna, Coutts, & Wisløff, 2009) and tactical behavior (Gonçalves, Figueira, Maçãs, & Sampaio, 2014; Sarmiento et al., 2013). Soccer play requires fundamental physical abilities which refer to flexibility, strength, balance, agility, power, and endurance. Soccer performance related to physical ability varies with positional role, and is quantified in the performance analysis literature as total distance covered, distance covered at high intensity, and distance covered at low intensity (Di Salvo et al., 2013). Technical skills include passing, dribbling, tackling, shooting, and the performance involved with individual ball control. Like physical abilities, different positions require different technical skills (Rampinini et al., 2009).

The tactical behavior includes the soccer-specific parameters such as “low intermediate positioning”, “counter-attack” (Sarmiento et al., 2013), and the activities and situations that involve more than one player. According to the formation or situation, players need to be able to make quick tactical decisions during gameplay (Gonçalves et al., 2014). Soccer performance can be expressed with these three performance indicators with a mathematical model. However, several models exist to explain soccer performance under different situations. The following will review common models of performance analysis used in soccer to determine which models are best for analyzing certain soccer performance indicators.

**METHODS:** Electronic searches used the SPORTDiscus (to September 2014) database with the following definition and strategy (Table 1):

Table 1. List of keywords and search strategy

keyword	soccer football athletic performance sport performance
search strategy	(soccer or football) and (athletic performance or sport performance)

**PERFORMANCE ANALYSIS MODELS:** Team soccer performance can be assessed by grouping players into defense, midfield, and attacking positions. Variables in the study were the mean value of position related centroid. The algorithm to calculate Approximate entropy (ApEn) that was used to evaluate the position centroid (Gonçalves et al., 2014) is listed below.

$$ApEn = \Phi^m(r) - \Phi^{m+1}(r)$$

$$\Phi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \log(C_i^m(r))$$

In the algorithm, m represented length, which was 1 because the comparison was between the defender, midfielder and forward. The value r was the filtering level, which was set as 0.5 in the study.

The N is the raw data related to time series. ApEn ranged from 0 to 2, with 2 representing more chaotic player movements relative to the centroid.

After a centroid position for defense, midfield and attackers is calculated, the Hilbert transform can be used to further analyze the relationship between the three centroids. The main function of the Hilbert transform is to show movement coordination of the defensive, midfield, and attacking groups. The Hilbert transform categorizes inter-position movement patterns into in-phase or anti-phase. In-phase movement patterns are when the groups are moving in similar directions, while anti-phase movement patterns are when groups are moving in opposite directions. Below is the Hilbert transform.

$$\xi(t) = s(t) + iH(t) = A(t)e^{i\phi(t)}$$

$H(t)$  is Hilbert transform to  $s(t)$ .  $A$  is amplitude and  $\phi$  is phase.

The model of  $k$  archetypes was used to establish an individual performance profile of soccer players. In the model, there are  $n$  players' performance scores,  $m$  performance indicators making up the matrix  $X$ . According to the performance provided, the  $k$  archetypes can make up a new matrix  $Z$ . The algorithm for choosing the  $k$  archetypes is that  $k$  value could produce the minimal residual sum of squares (RSS), which testified the tightness of fit of the  $k$  archetypes model to the data. Manuel J.A. (Eugster, 2012) used the model of  $k$  archetypes to establish the performance profile of soccer players from European leagues. Manuel assessed 25 different individual soccer skills. Next, the 25 soccer skills were rated on a 100 point scale. Data was then used to create 25 dimensions with  $i_{th}$  number. Next, the matrix of  $Z_k$  was established, which included  $Z_1, \dots, Z_p$  for the  $k$  archetype model. Using the 25 dimensions, Manuel derived 4 dimensions, which he used as archetype models because they provided the minimal RSS value. Therefore, a  $\{\alpha_{ik}\}$  with  $Z_k$  could produce the minimal RSS.

$$RSS = \sum_{i=1}^n \left\| x_i - \sum_{k=1}^p \alpha_{ik} Z_k \right\|_2$$

**CONCLUSION:** The model for soccer performance was divided in the aspects of group and individual analysis. Soccer performance analysis contained a large number of variables such as: physical abilities, technical skills, and tactical behaviors. The ApEn was used for the chaotic data like players movement patterns. ApEn is a model based on time series. Therefore, the time-related information like movement at various velocities could be assessed.

The Hilbert transform can be applied to analyze inter-position movement patterns. It is unknown whether relative synchronized or unsynchronized movement patterns are desired or not. To date there is no research that applies this transform to game situations. Thus, more research examining game situations and relative inter-positional movement patterns is granted. This research may be helpful in understanding and creating tactical behavior.

Eugster (2012) found large individual technical skill differences between athletes using the archetype principle. The archetype principle demarcates the performance level based on the performance indicator and could assess the player individually. The advantage of this method is that the soccer performance turns into an ordinal value. The ordinal sequence was matched to the coach's judgment. One disadvantage of using the ordinal rating scale data was on the basis of assumption that the difference is within the same dimension could be interpreted in the 100 point system. For an indicator like teamwork, using 100 point scale to assess the indicator was arbitrarily considered. This is because different players would cooperate with each other at varying levels from game to game.

Because of this, it would be difficult to compare data between games. The ordinal scale may show differences, but there may be a high degree of variability between ordinal scores. While the model indicates individual performance status, it neglects group performance. These two variables are interrelated. Thus, separating them can weaken archetype measures generated by this model. The situation of intra player relationships should still be applied and it should be assumed that all the indicators were consistent without changes in different game situations.

In summary, the models that had been used were separated into group and individual analysis. The group analysis model could be applied into the position-related movement pattern. The team tactical behavior in a 4-3-3 formation could be explained in the ApEn model. The individual performance model could offer a whole view of players' technical information related to the soccer performance. However, the validity and reliability of the performance indicator should be re-analyzed.

## REFERENCES:

- Di Salvo, V., Pigozzi, F., González-Haro, C., Laughlin, M., & De Witt, J. (2013). Match performance comparison in top English soccer leagues. *Int J Sports Med*, 34(06), 526-532.
- Eugster, M. J. A. (2012). Performance Profiles based on Archetypal Athletes. *Int J Perf Analysis Sport*, 12(1), 166-187.
- Gonçalves, B. V., Figueira, B. E., Maças, V., & Sampaio, J. (2014). Effect of player position on movement behaviour, physical and physiological performances during an 11-a-side football game. *J Sports Sci*, 32(2), 191-199.
- Rampinini, E., Impellizzeri, F. M., Castagna, C., Coutts, A. J., & Wisløff, U. (2009). Technical performance during soccer matches of the Italian Serie A league: Effect of fatigue and competitive level. *J Sci Med Sport*, 12(1), 227-233.
- Sarmiento, H., Pereira, A., Matos, N., Campaniço, J., Anguera, M. T., & Leitão, J. (2013). English Premier League, Spain's La Liga and Italy's Serie A—What's Different? *Int J Perf Analysis Sport*, 13(3), 773-789.