International Review of Management and Marketing

ISSN: 2146-4405

available at http: www.econjournals.com

International Review of Management and Marketing, 2016, 6(S8) 294-298.

Special Issue for "International Conference on Applied Science and Technology (ICAST) 2016, Malaysia"

E-compare of Soccer Tournament Structures

Muhammad Mat Yusof[1]*, Ruzelan Khalid[2], Tajularipin Sulaiman[3], Mohamad Shukri Abdul Hamid[4], Rosnalini Mansor[5], Nurzahidah Che Rosli[6]

[1]School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia, [2]School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia, [3]Sports Academy, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia, [4]School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia, [5]School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia, [6]School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia. *Email: mmy@uum.edu.my

ABSTRACT

This paper develops a system integrating the knowledge of statistical modeling and Monte Carlo simulation to evaluate various types of soccer tournament structures. A system called as "E-compare of Soccer Tournament Structures" aims to assist decision makers to choose the competitive soccer design. The system reports various tournaments metrics such as the expected number of goals scored and conceded, the expected number of wins, draws, and losses, and the expected final ranking at the end of the tournament. Based on a large number of simulations using teams participated in the Malaysian soccer super league, our analysis showed that different designs gave different impacts on the final ranking of the teams. Round robin is found to be the best structure in terms of identifying the strongest team to win the league compared to a knockout structure.

Keywords: Decision Support System, Soccer, Match Prediction Outcome, Simulation.
JEL Classifications: C63, C15

1. INTRODUCTION

Sports prediction is an obviously interesting topic and always attracts sport fans' interests. In match prediction outcomes in soccer for example, there are various match prediction models which have been developed; either to predict the number of goals scored and conceded (Koopman and Lit, 2015; Groll et al., 2015; McHale and Szcepanski, 2014; Baker and McHale, 2013; McHale and Scarf, 2011) or to directly predict the result outcomes such as the number of wins, draws and loses (Koning et al., 2003; Dobson and Goddard, 2003; Boulier and Stekler, 2003). However, in terms of prediction of effect of various types of tournament structures, there are few researches focusing on that issue. Scarf and Yusof (2011) and Scarf et al. (2009) have studied the effect in the context of a sport tournament design.

In this paper, we develop a system to compare various types of tournament structures using decision support system (DSS), called as "E-compare of Soccer Tournament Structures." The system is a computerized information system that helps the process of decision making for a sport designer and sport team manager. As a sport designer, this system can assist them to identify the competitive structure and to determine the fairest structure. As a team manager, the expected outcomes for individual matches are meaningful for team tactical strategies.

We adopt our system in the context of soccer tournament. In fact, soccer is a well-known sporting contest in the world. The arrangement of soccer tournament can be organized in three different ways; either in knockout (KO), round-robin (RR) or hybrid variations. KO is a structure where competitors compete head to head matches in k rounds. In each round, half of the paired teams are eliminated from the tournament. The winning team in the final round is declared as the champion of the tournament. RR or all-plays tournament is a structure where all teams play every other team an equal number of times. In each

match, points are awarded according to the match outcome. After all matches are played, the team with the highest total number of points is declared as the tournament winner. Finally, the hybrid tournament is created from the combination of the RR and KO structures. The most common designs are RR followed by the KO structure (RR-KO) and RR followed by the RR structure (RR-RR).

In order to describe the details of our system, we organized this paper into four main sections. Section 1, that we conclude here, elaborates the background, the aims and the significance of the study and also the general overview of tournament structures. Section 2 provides the general methodology and the framework of the study. Section 3 presents the results and analysis of the study. Finally, the conclusion and future work recommendation are stated in Section 4.

2. METHODOLOGY/FRAMEWORK

The system we developed uses the knowledge of statistical modeling and simulation technique to compare various types of tournament structures. In particular, the combination of that knowledge has developed various systems for assisting decision making processes. In a recent paper by Tani et al. (2015), a sports strategy decision system called SportsViz has been designed for the use of ball games. The statistical analysis results and player motions are visualized in an integrated way to facilitate easy decision making processes. Other than that, a decision support tool is also used in rugby to assist the performance evaluation processes. Calder and Durbach (2015) designed a tool to help the process by providing the framework for making decisions in a structured and distinct way.

Therefore, in order to compare various types of tournament structures, we firstly need to simulate a tournament. In order to simulate a tournament, we need to simulate the individual matches. To simulate the individual matches, we need a statistical match prediction model to predict the outcomes of the individual matches. Thus, in this paper, we predict the match outcomes based on the number of goals scored and conceded for every individual match. By letting X_{ij} and Y_{ij} be the number of goals scored in the match between teams indexed i and j, then a model for match scores is given by:

$(X_{ij}, Y_{ij}) \sim$ Poisson $(\alpha_i \beta_j \gamma, \alpha_j \beta_i)$ for team i playing at home
$\sim$ Poisson $(\alpha_i \beta_j \gamma, \alpha_j \beta_i)$ for team i playing at home
$\sim$ Poisson $(\alpha_i \beta_j, \alpha_j \beta_i \gamma)$ for team j playing at home

Where, X_{ij} and Y_{ij} are independent variables, and $\alpha_i, \beta_j > 0$ □ i. The α_i measures the attacking strength of the teams, the β_j measures the defensive weakness, and γ and κ are parameters that allow for home and neutral effects respectively. All the parameter strengths and weaknesses are calculated based on the maximum likelihood estimation using the R-packages.

Using the information of team strengths and weaknesses, we then simulate various types of tournament structures using the Visual Basic language. In this paper, we only consider four types of structures as follows:

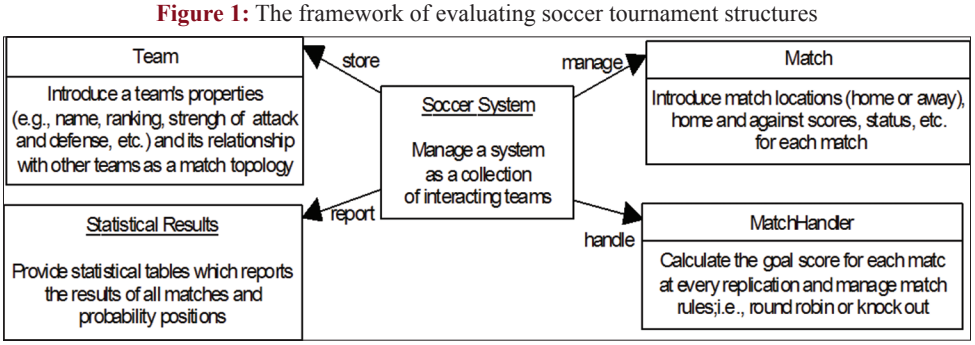
1. RR 1 leg: Each team plays each other team once
2. RR 2 legs: Each team plays each other team twice, once at home and once at away
3. KO (1 leg) - Random seeding: The pairing of team is chosen at random
4. KO (1 leg) – Standard seeding: The pairing team is based on their standing orders in the previous league season. The strongest team plays with the weakest team, the second stronger team plays with the second weaker team and so on.

In order to develop the DSS for evaluating sports tournament design, we design its framework which is illustrated in Figure 1. The framework stores the collection of teams and manages their relationships in terms of match information and scores. For fulfilling the purposes, three main classes have been structured; i.e., Team, Match and MatchHandler.

The Team class is to store a particular team's properties. The important properties include its name, ranking position and values of attacking strength and defensive weakness. The class also stores the list of matches with other teams and the scores in forms of array variables which will later be used to rank the team. The Match class meanwhile manages the match locations, the score and the status (win or lose) of each match before the information is transferred to a relevant team's properties. The MatchHandler stores the engine calculating the goal scored for each match at each replication, the number of replications to be performed and the type of tournament (e.g., RR and random) to simulate and analyze the match outcomes.

The MatchHandler implements the double Poisson and strength ratio methods to predict the match outcomes based on the values of attacking strength and defense weakness of each team. The statistical results in the framework meanwhile is a module which provides overall information on analyzing and evaluating the

Figure 1: The framework of evaluating soccer tournament structures



tournament design and statistical tables reporting the results of all matches and probability positions of participated teams.

3. DATA ANALYSIS AND RESULTS

Our DSS was employed to analyze the effect of various types of tournament structures using teams participated in the Malaysian Super League (MSL). To do this, we firstly estimated the teams' strengths and weaknesses based on the maximum likelihood estimation using the R-packages. Figure 2 shows the system screenshot of their strength and weaknesses. For example, the attacking strength for Lion XII is 1.0493 and its defense weakness is 0.7042. Meanwhile, the attacking strength for ATM_FC is 0.886 and its defense weakness is 1.4266.

Figure 3 shows the simulation setup for the DSS. This setup includes the four tournament structures (the RR one and two legs and KO random and standard ranking), the algorithm used to calculate the score for each match, the tournament metrics and the number of simulations or replications.

In Figure 4, we show the system screenshot of summary match results for the first replication out of 30 simulations using the RR 1 leg structure. The LionXII team is expected to be the champion with 8 wins, 4 draws and 3 losses with 28 points, followed by Perak_FA with 7 wins, 4 draws and 4 losses with 25 points.

Based on the position probability summary of 30 simulations using a RR 1-leg structure (Figure 5), LionXII has the highest probability to become the champion with 8 times appeared in the first league ranking, followed by Pahang_FA with 7 times.

In Figures 6 and 7, we show the simulation setup and the outcome of tournament metrics using a ranking KO structure (standard seeding) based on the structure, LionXII also has the highest probability to become the champion with 9 times appeared in the first league ranking, followed by Kedah_FA with 5 times.

4. CONCLUSION AND FUTURE WORK

This paper develops a DSS to evaluate various tournament structures. The effect of the tournament structures on match results were explored in the context of the MSL. The advantage of the system is to provide information to decision makers on tournament metric such as the probability of wins for each team, the number of matches involved, the effects of design changes based on various seeding technique.

Our analysis shows that each design gave different impacts on the participating teams. The RR structure is a better option to determine the strongest team to win the tournament, which is then followed by the KO standard seeding and KO random seeding structures. In the RR structure, each team plays each other teams and it includes more matches compared to the KO structure.

This study still needs further improvement. For example, the teams' attacking strengths and defense weaknesses need to flexibly be adjusted through the tournament. Their strengths and weaknesses need to be re-estimated accordingly since the current practices in an intermediate season, team can change their players.

This study only considers four types of tournament structures. Other tournament structures such as the combination of RR and KO in later stages should also be considered in our system. The

Figure 2: The attacking strength and defensive weaknesses

Team Information			
Team	Ranking	Attack	Defense
Sarawak_FA	1	0.8795	0.3845
LionXII	2	1.0493	0.7042
Kedah_FA	3	0.6571	0.4507
PKNS_FA	4	1.4698	1.1798
Selangor_FA	5	0.9862	0.8169
Perak_FA	6	1.0264	0.8589
Pahang_FA	7	1.3699	1.2896
Johor_Darul_Takzim	8	1.089	1.1387
Kelantan_FA	9	0.9654	1.0096
Terengganu_FA	10	0.9787	1.1662
T_Team_FC	11	0.8968	1.0805
Felda_United_FC	12	0.8159	0.9961
Sime_Darby_FC	13	0.6916	0.9742
ATM_FC	14	0.886	1.4266
Johor_FA	15	0.5069	0.9142
Negeri_Sembilan_FA	16	0.5164	0.933

Figure 3: The soccer simulation setup

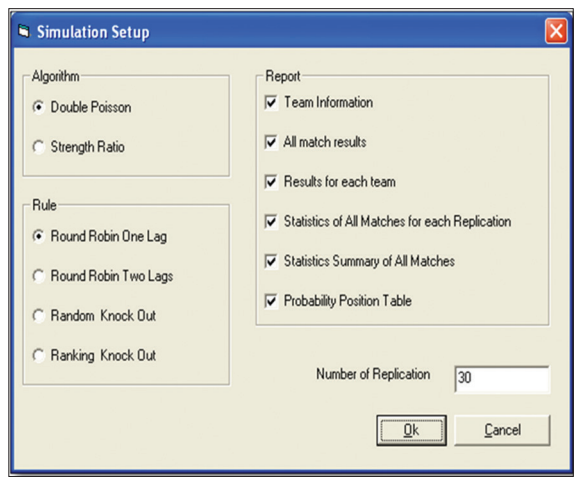


Figure 4: Summary match results

Summary Match Results (Replication 1 of 30)									
POS	TEAM	P	W	D	L	GF	GA	GD	PTS
1	LionXII	15	8	4	3	21	10	11	28
2	Perak_FA	15	7	4	4	20	13	7	25
3	PKNS_FA	15	6	4	5	26	25	1	22
4	Selangor_FA	15	6	5	4	18	11	7	23
5	Sarawak_FA	15	8	5	2	10	2	8	29
6	Terengganu_FA	15	8	3	4	15	10	5	27
7	Pahang_FA	15	4	8	3	19	13	6	20
8	T_Team_FC	15	4	5	6	18	16	2	17
9	Felda_United_FC	15	5	6	4	14	17	-3	21
10	Kelantan_FA	15	5	4	6	12	12	0	19
11	Kedah_FA	15	5	6	4	8	8	0	21
12	Johor_Darul_Takzim	15	4	6	5	12	16	-4	18
13	Negeri_Sembilan_FA	15	3	7	5	9	15	-6	16
14	Sime_Darby_FC	15	1	8	6	5	11	-6	11
15	ATM_FC	15	3	3	9	8	23	-15	12
16	Johor_FA	15	2	4	9	7	20	-13	10

Figure 5: The probabilities of team positions using a round-robin 1 leg structure

Position Number Probability																
TEAM	Post 1	Post 2	Post 3	Post 4	Post 5	Post 6	Post 7	Post 8	Post 9	Post 10	Post 11	Post 12	Post 13	Post 14	Post 15	Post 16
Sarawak_FA	1	10	6	8	1	2	0	1	0	1	0	0	0	0	0	0
LionXII	8	4	3	2	2	4	3	2	1	1	0	0	0	0	0	0
Kedah_FA	1	0	0	2	3	3	4	1	4	1	4	3	2	2	0	0
PKNS_FA	5	4	4	3	3	2	3	4	2	0	0	0	0	0	0	0
Selangor_FA	2	1	3	2	3	3	2	6	2	0	3	2	0	0	0	1
Perak_FA	3	2	2	4	3	1	3	1	4	2	4	0	1	0	0	0
Pahang_FA	7	2	2	1	2	4	4	4	1	2	0	0	0	1	0	0
Johor_Darul_Takzim	0	2	3	3	4	1	3	1	3	1	3	3	2	1	0	0
Kelantan_FA	1	2	1	0	3	2	2	6	0	6	3	1	3	0	0	0
Terengganu_FA	2	2	1	1	2	3	1	1	2	4	3	1	1	0	5	1
T_Team_FC	0	1	2	3	1	2	0	2	3	1	0	5	2	5	3	0
Felda_United_FC	0	0	1	1	1	2	3	1	3	2	3	4	3	4	2	0
Sime_Darby_FC	0	0	2	0	1	0	2	0	3	2	1	2	4	6	4	3
ATM_FC	0	0	0	0	1	1	0	0	2	2	4	1	2	5	5	7
Johor_FA	0	0	0	0	0	0	0	0	0	2	2	4	7	3	7	5
Negeri_Sembilan_FA	0	0	0	0	0	0	0	0	0	3	0	4	3	3	4	13

Figure 6: Simulation setup using knockout (standard seeding)

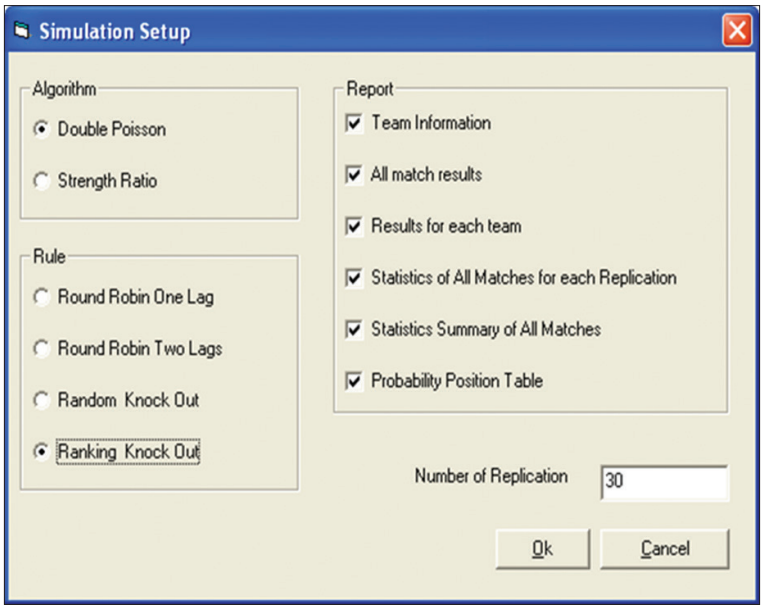


Figure 7: The probabilities of team positions using a knockout (standard seeding) structure

Position Number Probability																
TEAM	Post 1	Post 2	Post 3	Post 4	Post 5	Post 6	Post 7	Post 8	Post 9	Post 10	Post 11	Post 12	Post 13	Post 14	Post 15	Post 16
Sarawak_FA	4	6	3	4	0	2	1	3	2	1	2	1	0	0	1	0
LionXII	9	4	2	3	2	3	1	0	0	1	0	3	0	2	0	0
Kedah_FA	5	3	5	1	3	1	1	3	0	0	1	2	1	1	2	1
PKNS_FA	4	3	3	5	1	2	1	1	3	1	2	1	0	0	1	2
Selangor_FA	0	3	2	1	5	5	3	1	2	0	3	3	2	0	0	0
Perak_FA	3	2	2	1	5	2	5	2	1	1	2	0	0	1	1	2
Pahang_FA	1	2	1	4	1	2	3	3	4	2	1	0	2	2	1	1
Johor_Darul_Takzim	0	1	2	2	0	7	4	3	2	0	2	4	1	1	0	1
Kelantan_FA	1	1	4	2	2	1	3	1	2	4	2	3	3	0	1	0
Terengganu_FA	0	0	3	1	5	1	2	2	1	4	3	2	2	2	1	1
T_Team_FC	2	3	2	0	0	0	0	1	1	0	5	4	2	5	2	3
Felda_United_FC	0	0	1	1	2	1	0	5	2	4	2	4	4	1	1	2
Sime_Darby_FC	1	2	0	1	2	1	3	1	2	0	1	1	5	3	1	6
ATM_FC	0	0	0	2	1	1	1	2	2	2	0	2	6	4	6	1
Johor_FA	0	0	0	1	1	1	1	0	3	5	1	0	2	2	9	4
Negeri_Sembilan_FA	0	0	0	1	0	0	1	2	3	5	3	0	0	6	3	6

variety of tournament structures can give a wider view on the impact of changing structures to tournament decision makers and team management.

5. ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Education Malaysia's under the Research Acculturation Collaborative Effort 2013 grant (S/O Code: 12968).

REFERENCES

Baker, R.D., McHale, I.G. (2013), Forecasting exact scores in national football league games. International Journal of Forecasting, 29(1), 122-130.

Boulier, B.L., Stekler, H.O. (2003), Predicting the outcomes of national football league games. International Journal of Forecasting, 19(2), 257-270.

Calder, J.M., Durbach, I.N. (2015), Decision support for evaluating player performance in rugby union. International Journal of Sports Science and Coaching, 10(1), 21-38.

Dobson, S., Goddard, J. (2003), Persistence in sequences of football match results: A Monte Carlo analysis. European Journal of Operational Research, 148(2), 247-256.

Groll, A., Schauberger, G., Tutz, G. (2015), Prediction of major international soccer tournaments based on team-specific regularized poisson regression: An application to the FIFA World Cup 2014. Journal of Quantitative Analysis in Sports, 11(2), 97-115.

Koning, R., Koolhaus, M., Renes, G., Ridder, G. (2003), A simulation model for football championship. European Journal of Operational Research, 148(2), 268-276.

Koopman, S.J., Lit, R. (2015), A dynamic bivariate poisson model for analyzing and forecasting match results in the English Premier League. Journal of the Royal Statistical Society: Series A, 178, 167-186.

McHale, I.G., Scarf, P.A. (2011), Modeling the dependence goals scored by opposing teams in international soccer matches. Statistical Modeling, 11, 219-236.

McHale, I.G., Szcepanski, L. (2014), A mixed effects model for identifying goals scoring ability of footballs. Journal of the Royal Statistical Society: Series A, 177(4), 397-417.

Scarf, P.A., Bilbao, M., Yusof, M.M. (2009), A numerical study of designs for sporting contests. European Journal of Operational Research, 198(1), 190-198.

Scarf, P.A., Yusof, M.M. (2011), A numerical study of tournament structures and seeding policy for the soccer World Cup finals. Statistica Neerlandica, 65(1), 43-57.

Tani, T., Huang, H., Kawagoe, K. (2015), Sports play visualization system for American football. Lecture Note in Engineering and Computer Science, 2215(1), 161-166.