Thailand Statistician
July 2015; 13(2): 223-242
http://statassoc.or.th
Contributed paper

Longitudinal Linear Mixed Effect Model: An Application in Analyzing Age Effects in Twenty20 Cricket

Hemanta Saikia [a], Dibyojyoti Bhattacharjee [b], Atanu Bhattacharjee* [c] and Hermanus H Lemmer [d]

[a] School of Business, The Assam Kaziranga University, Jorhat (Assam), India.

[b] Department of Statistics, Assam University, Silchar (Assam), India.

[c] Division of Clinical Research and Biostatistics, Malabar Cancer Centre, Thalassery (Kerela), India.

[d] Department of Statistics, University of Johannesburg, Auckland Park, South Africa.

*Corresponding author; e-mail: atanustat@gmail.com

Received: 30 September 2014
Accepted: 27 April 2015

Abstract

Though most of the cricketing fraternity opines that Twenty20 cricket is a game of young cricketers, yet the performance of senior cricketers in the different seasons of IPL seems to be comparable to that of the youngsters. Thus, this study tries to examine the effect of age on on-field performance of the cricketers in Indian Premier League. To quantify the performance of cricketers, a measure is developed utilizing the four prime skills of the game *viz.*, batting, bowling, fielding and wicket keeping. Various cricketing factors are considered related to the performance of cricketers under the above-mentioned skills. All these factors are normalized and accordingly adjusted by using appropriate weights on the basis of their relative importance. The performance measures are obtained for each cricketer separately in all the four seasons of IPL played so far, and the regression model with random regression coefficient has been applied to determine the effect of age on cricketers' performance. The results obtained from the regression model confirm that the on-field performances of the cricketers are positively associated with the age of the players.

Keywords: age, performance measure, regression, statistics in sports, twenty20 cricket.

1. Introduction

Twenty20 cricket is the shortest version of cricket which was first proposed by Stuart Robertson, the former Marketing Manager of England and Wales Cricket board (ECB). In this format, each side would bat for twenty overs; and in effect, the game transformed into one that could be finished in three hours, thus not requiring the day-long commitment by spectators, unlike One-Day International (ODI) limited over game of cricket [1]. In 2001, ECB introduced a domestic Twenty20 tournament which turned to be a great success at England in terms of viewership. Gradually, it became a benchmark for the other major cricket playing countries like Australia, West Indies, South Africa, etc. Recognizing the popularity of the Twenty20 format, International Cricket Council (ICC) organized the first Twenty20 World Cup in 2007 that was played in South Africa. India won the inaugural Twenty20 World Cup and this proved to be an incredible boost to this type of cricket in India [1]. The fact that Twenty20 cricket has a huge plethora of excitement for the cricketing fans across the globe after the inaugural Twenty20 World Cup.

In April 2008, the game of cricket got a new dimension when the Board of Control for Cricket in India (BCCI) initiated the Twenty20 format of cricket called Indian Premier League (IPL). What makes the IPL different from the other Twenty20 cricket organizing nations is that the players are owned by franchisees [2]. The franchisees formed their teams by competitive bidding from a collection of Indian and international players and the best of Indian upcoming talent [3]. Each player has a base price fixed by the IPL authorities; however, there is no upper limit for their bid price. The valuation of players obtained through auction and availability of players' performances have allowed researchers to infer on different aspects of this format of the game of cricket [4].

However, it is believed that the recent format of Twenty20 cricket is a game for young cricketers only [5]. This belief got credibility when Younis Khan, the former captain of the Pakistan cricket team, had announced his retirement from Twenty20 cricket at the age of 32, after the team Pakistan won the Twenty20 world-cup in England in 2009. While announcing his retirement Khan commented that, he doesn't consider himself fit enough for this format of cricket [1].The former Australian spin bowler Shane Warne also stated that though retired players and emerging talented players both get a platform to perform and learn from each other in IPL, the Twenty20 national tournament (IPL) is a

great platform only for up-and-coming players rather than veteran cricketers [1]. The IPL has helped some young cricketers to make a mark [6]. Twenty20 cricket is not particularly taxing to a professional sportsperson, or even someone who has crossed 40 (*e.g.* S Warne) and is in the second life of his career. Indeed, it is true that a few young players like P Valthaty, R Ashwin, S Marsh, S Randiv, S Hasan, etc. have managed to draw the attention of selectors and spectators toward themselves through their remarkable performances in IPL [7]. But in all the seasons of IPL played so far, the performances of senior cricketers are noteworthy. The senior players make a huge impact in the IPL largely due to their skills which have made them everyday cricketers unlike the ones who fire only for a while [6]. This has been proven by players like S Tendulakar, R Dravid, S Warne, A Kumble, D Vettori, etc. Though the players like S Tendulkar, R Dravid, etc. have not played much of the Twenty20 format of cricket but their skill sets are so complete that they can crack the code in no time [8]. The performances of skilled players in Twenty20 cricket were a lesson for youngsters that there is no substitute for skill, if anyone is to succeed in any format of the game. The skills required for batsmen are that they must keep their heads, choose the right shot and resist pace, swing and wrist-spin and for bowlers that they must move the ball around, possess a wide range of deliveries and retain control under intense pressure [9].

The franchisee based Twenty20 cricket (now-a-days common to different cricket playing nations) has turned to a heaven sent opportunity for past masters. With the advent of franchisee cricket and the money it brings in, some of these senior cricketers are now travelling from country to country, playing their trade [10]. The senior cricketers who are busy '*playing their trade*' are mostly the superlative players of their generation. In most cases, with their performance in franchisee cricket like the IPL, such players are giving the averagely talented youngsters a run for their money. The above discussions set the background of this paper.

2. Review of literature and objectives of the study

On the basis of available literature, it has been revealed that the performance of players varies on aging. Relative age effect in sport is a worldwide phenomenon and it exists in many, but not at all competitive sports [11]. Yet the empirical evaluation of the aging effect in performance is difficult in sports. It is believed that the performances of the players reach to a peak at a given age and then shows a waning, but determination of the peak performing years of the players is always challenging. Also in terms of team games, the level of performance of the players can vary depending on the strength of the

opponents and different playing conditions. The effect of age has been studied in different kinds of sports. Most of these researches showed linear, curvilinear or exponential trends when modeling effects of aging [12]. Linear trend was found when evaluating the effects of aging on freely chosen walking speed [13]. Curvilinear trend was found by Seiler et al. [14] when investigating competitions in indoor rowing events and it was found that in freestyle swimming performances decreased exponentially [15]. In baseball, the effects of age using nonlinear fixed effect regression and found that aging effects are larger for pitchers than for batters [16]. The peak age of performance for professional baseball pitchers and batters are 26 and 28 years respectively. In many sports the classification system (*i.e.* senior, young, peak performance age, etc. through cut off values) based on biological age are difficult to organize [11]. The effect of aging for professional football players using performance-aging curves was evaluated by Young and Weckman [12]. Recently, the relative age effect (RAE) in the National Hockey League (NHL) examined [17] .

In cricket, Wood [18] examined consistency in the performance of batsmen by applying the geometric distribution to model cricket batting scores based on results from test cricket. A non-parametric approach to fit the geometric distribution based on runs scored for assessing batting performance of cricketers and found that it did not fit for scores of 0 as Wood [18] had found previously [19]. Two comprehensive measures, called the combined bowling rate (CBR) and the dynamic bowling rate (DBR) were developed by Lemmer [20, 21] to measure the current bowling performance of the bowlers in one-day and test cricket respectively. The use of stochastic dominance rules demonstrated by Damodaran [22] can be used to analyze the batting performance of Indian cricketers in ODI cricket.

The literature on performance measurement of players in Twenty20 cricket is not much enriched owing to the fact that it is a new format of cricket. Lemmer [23] discussed the performance of batsmen and bowlers in the first Twenty20 World Cup. Suleman and Saeed [24] attempted to test the hypothesis whether the performances of cricketers can be treated as a real option. A performance index called SS Index is developed to measure the performance of batsmen, bowler, all-rounder and wicket keeper separately [25]. In terms of the all-round performance of the players, Van Staden [26] developed a performance measure for cricketers in Twenty20 cricket using data from the first season of IPL. Based on the strike rate and economy rate, Saikia and Bhattacharjee [3] classified the performances of all-rounder who participated in the first three seasons of IPL. Then the factors responsible for classification of all-rounder were

identified and a naïve Bayesian classification model was used to predict the expected class of all-rounder who had played only in the fourth season of IPL. A suitable performance measure was developed by Saikia et al. [4] to measure the performance of batsmen, bowlers, all-rounder and wicket keepers but the fielding skill of cricketers were not included.

Though several measures of performance are reported in the literature, surrounding cricket, yet the impact of aging on players' feats did not enrich our search [14]. In cricket, it is difficult to generalize a relation between aging and performance as cricketers with different expertise peak at different ages. Batsmen tend to reach their peak in their late twenties after their technique have matured through experience. Conversely, fast bowlers often reach their peak in between early to mid-twenties when they are at the height of their physical capacity. Other bowlers, mostly spinners, even fast bowlers who can "swing" the ball, are most effective in the later part of their career [3].Therefore, this study tries to examine the effect of age on performance of the cricketers in terms of Twenty20 cricket. In view of that the following two objectives are proposed.

- Develop a tool to measure the performance of cricketers.
- To examine the effect of age on performance of the cricketers in Twenty20 cricket.

3. Data and model formulation

The data pertinent to the performances of the players in the first four seasons of IPL are collected from the website *www.espncricinfo.com*. It is necessary that players' statistics for large numbers of games should be considered to measure the performance of players. The actual quality of a player may not be properly judged from one or two games. The effects of outstanding or poor, single performances are smoothed over the larger number of games [27] and therefore individual performances across a series of matches are required to provide a suitable frame of reference [28]. Thus, some selection criterion needs to be set up while considering the players for performance measurement. Therefore, the players are selected based on the following criteria.

- The player was in the playing eleven in at least 5 matches in each of the IPLs (from the first season to the fourth season).
- Either the player has bowled at least 10 over's in each of the IPLs or has faced at least 100 balls in each of the IPLs as batsman or both.

There are 40 players who satisfied both the criteria. All these players are considered for the study and information about their ages was collected from the said website to determine the apposite ages of the selected players.

3.1 The performance measure

Most sports have measures of performance that purport to indicate some level of achievement [29]. In every cricket match, an award baptized “man of the match” is given to a particular player for his all-around performance in that match [27]. However, to assess the all-around performance of a player, organizers as well as decision makers need quantitative information about the on-field performance of that player. As each game of cricket generates a huge amount of performance related statistics, the on-field performance of a player can easily be perceived through the scorecard of a match. But a match scorecard only provides information about the traditional performance measures. The batting average and strike rate are usually used to understand the performance of batsmen. On the other hand, bowling average, economy rate and bowling strike rate are used to measure the performance of bowlers. However, such statistics have severe limitations in assessing the true abilities of a players’ performance [27]. Moreover, the different traditional performance measures are in different units of measurement; so to assess a player’s all-round performance, it is very difficult to combine them. All these shortcomings to measure the performances of the cricketers were well discussed by Lewis [27].

Attempts have been made by Barr and Kantor [30] and Lemmer [23] to develop performance measures for batsmen by combining the traditional measures of performances. But both these measures cannot be used to measure the performance of all-rounders. Lewis [27] proposed an alternative measure of player performance to overcome the above-discussed limitations on the basis of the well-established Duckworth and Lewis [31] method. It is based on a simple model involving a two-factor relationship giving the number of runs which can be scored on average in the reminder of an innings as a function of the number of overs remaining and the number of wickets that have fallen. Using the Duckworth and Lewis [31] method, the proposed alternative measure of player performance by Lewis [27] can be used not only to measure the performance of batsmen and bowlers, but it can also be used to measure the performance of all-rounder. However, one has to note ball-by-ball information of a match or series of matches to apply the performance measure developed by Lewis [27]. To collect the ball-by-ball information of a match or series of matches is a tedious job, as one has to watch each ball of a match and then all the matches of the entire series.

Nowadays, some of the websites like *espncricinfo.com*, *cricwaves.com*, etc. provide ball-by-ball information of the match in descriptive forms. However, one has to read the information and note them down in quantitative terms for the necessary computation, a task which is by no means simple. Moreover, this ball-by-ball information of a match provided in *html* sites like *espncricinfo.com* is not sufficient to quantify the performance of a player. Thus, it would be easier if the scorecard of a match can be utilized for measuring the performance of players. This has led to the development of a performance measurement model in this paper. Except the 'number of catches' and 'run out' the fielding information of a player is not available in the scorecard of a match. Hence, the performance measure developed in this paper considers only these two factors to quantify the fielding performance of cricketers. This performance measure can be used to measure the batting, fielding, bowling or wicket keeping performances of cricketers and combines them into a single index based on data from the scorecard of the match.

The performance measure of the i^{th} player is defined by,

$$S_i = S_{i1} + S_{i2} + \delta_i \tag{1}$$

where

$$\delta_i = \begin{cases} S_{i3}^{a_i} + S_{i4}^{1-a_i} - 1, & \text{if } i^{th} \text{ player is either a bowler or wicket keeper} \\ 0, & \text{if } i^{th} \text{ player is neither a bowler nor wicket keeper} \end{cases}$$

where a_i is an indicator variable with,

$$a_i = \begin{cases} 1, & \text{if } i^{th} \text{ player is a bowler} \\ 0, & \text{if } i^{th} \text{ player is a wicket keeper} \end{cases}$$

with S_{i1} = Performance score for batting, S_{i2} = Performance score for fielding, S_{i3} = Performance score for bowling and S_{i4} = Performance score for wicket keeping. The computation of the performance indicators for each of the abilities *viz.* batting, fielding, bowling and wicket keeping (*i.e.* S_{i1}, S_{i2}, S_{i3} and S_{i4}) and also the factors considered under each of the ability is explained in the subsequent sections.

3.2 Factors considered for batting performance measure

To measure the batting performance of cricketers, a number of factors are considered. These factors are batting average of the batsman, strike rate of the batsman and the average percentage contribution of a batsman to the team total. The values of

these factors for each of the players are normalized and weights of these factors are calculated based on their relative importance. All the normalized scores for the different factors are multiplied by their corresponding weights and then added together to get S_{i1}.

3.3 Factors considered for fielding performance measure

The different factors that are considered to quantify the fielding performance of a cricketer are number of catches taken by the player as a fielder and the number of run outs caused by the player in a match or series of matches. Information of both these aspects of fielding is available from the scorecard of the match. The values of these two factors for each of the players are normalized and then weights of these two factors are calculated based on their relative importance. All the normalized scores for these two factors are multiplied by their corresponding weights and then added together to get S_{i2}.

3.4 Factors considered for bowling performance measure

The different factors that are considered while computing the performance score of bowlers are bowling average, economy rate and bowling strike rate of the bowler. The values of these factors for each of the players are normalized and then weights of these factors are calculated based on their relative importance. As earlier, all the normalized scores for the considered factors are multiplied by their corresponding weights and then added together to get S_{i3}.

3.5 Factors considered for performance measure of wicket keeping

To measure the performance of a wicket keeper the different factors are the number of catches taken per match, the number of stumping per match and the number of bye runs conceded per match. Here the phrase ‘per match’ means the number of matches when the player kept the wickets for his team. This is done because some of the teams have more than one player in their playing eleven who are capable of wicket keeping. Similarly, the values of these factors for each of the players are normalized and then weights of these factors are calculated based on their relative importance. Finally, all the normalized scores for the considered factors are multiplied by their corresponding weights and then added together to get S_{i4}.

3.6 Normalization

The performance measure developed in this paper is a linear combination of traditional performance measures under the abilities of batting, fielding, bowling and wicket keeping. Lewis [27] mentioned that the traditional measures of performances do not allow combining of the abilities of batting and bowling as they are based on incompatible scales. To overcome this limitation, data of cricketers corresponding to each of the factors discussed above are normalized. The process of normalization helps

to eliminate the unit of measurement and variability effect of all the traditional performance measures. Based on normalization, the traditional performance measures under the abilities of batting, fielding, bowling and wicket keeping come within a similar range from 0 to 1. Since normalization makes the measures unit free, so they can be combined through addition.

Let X_{ijk} be the value of the j^{th} factor (*i.e.* batting average, strike rate, etc.) for the i^{th} player in the k^{th} ability (*i.e.* batting, fielding, bowling and wicket keeping). Out of the different factors considered for performance measurement, some have a positive dimension like batting average, batting strike rate, etc. as they are directly related to the ability of the player. Other factors like economy rate, number of bye runs conceded, etc. have a negative dimension as they are negatively related to the abilities of the player. Now, if the factor represents a positive dimension associated with the ability of the player then it is normalized as,

$$Y_{ijk} = \frac{X_{ijk} - \min(X_{ijk})}{\max(X_{ijk}) - \min(X_{ijk})} \tag{2}$$

and if the factor represents a negative dimension associated with the ability of the player then it is normalized as

$$Y_{ijk} = \frac{\max(X_{ijk}) - X_{ijk}}{\max(X_{ijk}) - \min(X_{ijk})}. \tag{3}$$

3.7 Determination of weights

The next step would be to determine the weights for different factors in all the four seasons of IPL played so far. While simple averages provide an equal importance to all the variables, a composite measure is weighted and the relative importance of the variables can be considered. The weights vary inversely proportional to the variation in the respective variables [32]. This method is applied in this study to determine the weights of the different factors that are associated with the various abilities of the cricketers.

Let Y_{ijk} be the normalized value of the i^{th} player for the j^{th} factor of the k^{th} ability where i (= 1, 2, …, n) represents players; j (= 1, 2, 3) for the three different factors considered under each of the k^{th} ability (for example under batting ability different factors are the strike rate, batting average, average percentage of contribution to the team total);

k(= 1, 2, 3 4) represents the different abilities *viz.* batting (1), fielding (2), bowling(3) and wicket keeping(4). If w_{jk} represents the weight of the j^{th} factor under the k^{th} ability, it is calculated as

$$w_{jk} = \frac{C_k}{\sqrt{\underset{i}{Var}(Y_{ijk})}}, \qquad j = 1, 2, 3 \text{ and } k = 1, 3, 4 \tag{4}$$

where

$$\sum_{j=1}^{3} w_{jk} = 1 \quad for \; k = 1, 3 \; and \; 4$$

C_k is a normalizing constant that follows

$$C_k = \left[\sum_{j=1}^{3} \frac{1}{\sqrt{\underset{i}{Var}\,(Y_{ijk})}} \right]^{-1}.$$

Since this study has considered only two factors under fielding ability to quantify the fielding performance of a cricketer *viz.* number of catches taken and run outs, therefore if w_{jk} represents the weight of the j^{th} factor under the fielding ability then it is calculated as

$$w_{jk} = \frac{C_k}{\sqrt{\underset{i}{Var}(Y_{ijk})}}, \quad j = 1, 2 \text{ and } k = 2 \tag{5}$$

where

$$\sum_{j=1}^{2} w_{jk} = 1 \quad for \; k = 2$$

C_k is a normalizing constant that follows

$$C_k = \left[\sum_{j=1}^{2} \frac{1}{\sqrt{\underset{i}{Var}\,(Y_{ijk})}} \right]^{-1}.$$

The choice of the weights in this manner would ensure that the large variation in any one of the factors would not unduly dominate the contribution of the rest of the factors [32]. One can see the different factors along with their respective weights obtained from using equations (4) and (5) in Table 1.

Table 1. Different factors with its respective weights in IPLs.

	Factors	IPL1 Weights	IPL2 Weights	IPL3 Weights	IPL4 Weights
Batting	Batting average	0.298137	0.33224	0.325174	0.345992
	Batting strike rate	0.387143	0.420151	0.361271	0.396733
	Average percentage of contribution	0.314721	0.247608	0.313555	0.25727
Field	Number of catches taken	0.536173	0.471893	0.510203	0.472517
	Number of run outs	0.463827	0.528107	0.489797	0.527483
Bowling	Bowling average	0.327927	0.300937	0.34182	0.348438
	Economy rate	0.373001	0.379924	0.348853	0.33092
	Bowler's strike rate	0.299072	0.31914	0.309328	0.320641
Wicket keeper	Number of catches taken as wicket keeper	0.3753	0.313673	0.279731	0.328676
	Number of stumping	0.265377	0.328087	0.380585	0.314683
	Number of bye runs conceded	0.359323	0.35824	0.339684	0.356641

3.8 Computation of performance score

The performance scores for batting (k=1), fielding (k = 2), bowling (k = 3) and wicket keeping (k = 4) of the i^{th} player is calculated by

$$S_{ik} = \sum_{j=1}^{n_k} w_{jk} Y_{ijk} \tag{6}$$

where, n_k = 3 for k = 1, 3 and 4 and n_2 = 2.

Now, on obtaining the values of S_{i1}, S_{i2}, S_{i3} and S_{i4} the performance score S_i of the i^{th} player is computed using equation (1). The performance score of all the players are computed and then converted into corresponding performance index (P_i). The performance index of the i^{th} player is denoted by P_i and is given by,

$$P_i = \frac{S_i}{\max_i (S_i)}. \tag{7}$$

The performance index value for a player is a number lying between 0 and 1 (*i.e.* $0 < P_i \leq 1$). Higher the value of the performance index better is the performance of

the player. The calculated performance scores of the cricketers obtained using equation (7) for the first four seasons of IPL can be seen in Table 2.

Table 2. Performance scores of the players in IPL generated from equation (7).

Players Number	Player Name	Age in IPL1	Performance score in _____			
			IPL4	IPL3	IPL2	IPL1
P1	A Mishra	25	0.681243	0.63577	0.641301	0.535339
P2	A Agarkar	30	0.556378	0.446764	0.543827	0.565319
P3	AB de Villiars	24	0.690542	0.453639	0.67476	0.367985
P4	A Gilchrist	32	0.578283	0.725423	1	0.625813
P5	D Bravo	24	0.676809	0.389733	0.848975	0.73816
P6	M Jayawardene	30	0.452346	0.601269	0.494187	0.423777
P7	G Gambhir	26	0.428163	0.438193	0.310502	0.45168
P8	I Sharma	19	0.548408	0.277208	0.561771	0.428701
P9	IkPathan	23	0.716967	0.722317	0.754609	0.728782
P10	JA Morkel	26	0.703198	0.670476	0.694835	0.72688
P11	JH Kallis	32	1	0.881844	0.563801	0.469916
P12	K Sangakara	30	0.480839	0.511915	0.528522	0.512606
P13	KD Kartik	22	0.429615	0.965849	0.829328	0.815161
P14	L Balaji	26	0.618041	0.393178	0.504498	0.351726
P15	M Muralidharan	36	0.314425	0.441	0.644643	0.393377
P16	M Dhoni	21	0.871602	0.657396	0.691283	0.62549
P17	P Kumar	21	0.503723	0.271864	0.636943	0.585205
P18	P Chawla	19	0.619663	0.589717	0.630257	0.8034
P19	P Ojha	21	0.604896	0.575239	0.467041	0.45364
P20	R Bhatia	28	0.700541	0.416477	0.58101	0.608063
P21	V Kumar	24	0.733077	0.430528	0.591218	0.508987
P22	RG Sharma	21	0.521624	0.694785	0.866129	0.894491
P23	RP Singh	22	0.660123	0.535246	0.609452	0.605867
P24	R Dravid	35	0.38551	0.446137	0.323311	0.407859

Table 2. (Continued).

Players Number	Player Name	Age in IPL1	Performance score in _____			
			IPL4	IPL3	IPL2	IPL1
P25	R Uthappa	22	0.673668	0.71802	0.559508	0.529213
P26	S Badrinath	27	0.364504	0.385303	0.311831	0.484803
P27	S Dhawan	22	0.942267	0.333841	0.189836	0.548704
P28	S Sreesanth	25	0.434683	0.17293	0.428089	0.384216
P29	SK Raina	21	0.700804	1	0.908265	1
P30	S Trivedi	25	0.471347	0.55091	0.470043	0.483473
P31	S Warne	34	0.541455	0.334797	0.688831	0.704004
P32	S Tendulkar	35	0.495892	0.509732	0.601663	0.442312
P33	T Dilshan	31	0.364419	0.519035	0.672127	0.900681
P34	V Kohli	19	0.751153	0.457017	0.61254	0.561261
P35	V Sehwag	29	0.597311	0.891077	0.421776	0.935588
P36	W Saha	23	0.389807	0.300878	0.44883	0.764305
P37	V Rao	26	0.321063	0.12045	0.713643	0.36983
P38	YK Pathan	25	0.908481	0.775398	0.729007	0.838675
P39	Y Singh	26	0.915371	0.602192	0.856075	0.905738
P40	Z Khan	29	0.588516	0.677007	0.572843	0.512818

4. Relation between age and performance

In order to examine the effect of age on performance of the cricketers, the information about the ages of selected cricketers (*i.e.* date of birth) in completed years are collected from the website *www.espncricinfo.com*. Since the performance scores of the cricketers and their ages in the different seasons of IPL are obtained, so the effect of the later on the former is studied through longitudinal data analysis. Menard [33] defined longitudinal research as one in which for the same subjects or cases, data are collected for each variable or items from two or more distinct time periods. Adopting this broad definition of longitudinal research, we can conceptualize the effect of age on cricketers' performance in IPL. However, modeling longitudinal data raises certain statistical complexities, mainly because observations are correlated and thus violating the key assumption of independence [34].

In this study, the individual level of players' performance scores have been observed through the first four seasons of IPL. The regression model with random regression coefficient has been applied to determine the effect of age on cricketers' performance in IPL, where P_{ij} is the response variable (*i.e.* performance score of the i^{th} cricketer in the j^{th} season of IPL) and x_{ij} the explanatory variable (*i.e.* age of the i^{th} cricketer during the j^{th} season of IPL). As mentioned earlier the presence of correlation of the same individual's measurements is quite natural in longitudinal data analysis, therefore the presence of serial correlation of the player's serial performance has been measured through regression modeling. The model is fitted with the assumption that a common intercept and slope hold for all players. The within players correlation has been explained by serial correlation and residual errors. The first order autoregressive process has been assumed to deal with the covariance structure. The performance score of the i^{th} player in the j^{th} IPL season (*cf.* Table 2) has been modeled as,

$$P_{ij} = \mu + \alpha_i + \beta x_{ij} + e_{ij} \tag{8}$$

where, i = 1, 2, … , 40 (as there are 40 cricketers in the training sample) and j = 1, 2, 3, 4 (as four IPL seasons are considered) and $e_{ij} \sim N\left[\varphi e_{ij-1}, (1-\varphi^2)\sigma^2\right]$ with -1 ≤ ϕ ≤ 1. Here μ is the model constant and β is the associated regression coefficient of the covariates (x_{ij}) age. Let equation (8) it is called as model 1. In model 1, the term α_i represents the mean performance score of the i^{th} player in IPL, ϕ denotes the degree of auto regression between residual errors and individuals and the term σ^2 denotes the unconditional variance of e_{ij}. If ϕ = 0, then $e_{ij} \sim N\left(0, \sigma^2\right)$ and the model 1 can be reduced to a simple normal linear model.

Apart from that, the presence of correlation of the residuals can be captured through φ. A high level of positive correlation gives φ = 1 while the value of φ approaches -1 for high level of negative correlation amongst residual and e_{ij}. It should be noted that the measurements of IPL performance scores of all the players are in time order and the time interval is equal for all the players. However, the provision of time difference can be provided through AR1 (PlayersPerformance/Age). The model can be specified through

$$e_{ij} \sim N\left[\phi^{\left(\tau_{ij}-\tau_{ij-1}\right)} e_{ij-1}, \left(1-\phi^{2(\tau_{ij}-\tau_{ij-1})}\right)\sigma^2\right] \tag{9}$$

where (τ_{ij}-τ_{ij-1}) is the time gap between j^{th} and (j-1)th IPL performance score of the i^{th} player. This model be called model 2. Here φ is used to present the correlation between residuals one unit of time apart. If the time difference tends to 0, the correlation between residual errors tends to 1 and if the time difference tends to infinity, the correlation tends to zero for all φ <1.The serial correlation and measurement terms can be observed through (10), which is called model 3.

$$P_{ij} = \mu + \alpha_i + \beta x_{ij} + e_{ij} \quad (10)$$

The models have been run with the software WINBUGS 2.14.1. The Bayesian approach was used to perform the analysis. The mean performance score of players about four IPLs are computed and AB de Villiers performed least in comparison to others. Thus, AB de Villiers is considered as the reference player (group) for performance of the model in (10).The Bayesian approach comes with likelihood function, prior information and probability density function of the parameters. The posterior distribution observed through Monte-Carlo simulation is more powerful and flexible with sampling procedure than classical approach. The posterior mean with credible interval is considered as a tool for statistical inference. The detailed explanation of the model is provided in Table 3.

In Model 1, the posterior mean of Age (0.13) shows the positive influence of age on performance scores in IPL. However, the posterior mean obtained through model 2 and model 3 indicates insignificant influence of age on IPL performance scores. The figures in brackets corresponding to Age are the credible intervals of the posterior mean of Age. Table 3, provides summary statistics that can be used to compare the three models. The model constants for Model 1, Model 2 and Model 3 respectively are 0.79 (0.09), 0.63 (0.02) and 0.64 (0.02). The values in brackets indicate the standard deviation of the model constants. The terms provide the mean IPL score over the population of players at the mean time point. The model constant is nearly similar for all the comparable models. It should be noted that in each of the models, the parameter φ (AR) is estimated to be quite low with relatively low precision. The serial correlation terms has been addressed through the models.

Table 3. Longitudinal models for analyzing age effect on cricketers' performance in IPL**.**

Models	Model 1	Model 2	Model 3
Fixed Effect	Players		
Random Effect		Players	
Independent Covariate	Age	Age	
Random Coefficient			Players/Age
Run Time (Seconds)	6	10	8
Constant	0.79 (0.09)	0.63 (0.02)	0.64 (0.02)
Slope mean	1.13 (1.84)	0.78 (0.12)	0.80 (0.11)
Age	0.13 (0.02,0.25)	-0.005 (-0.01,0.004)	-0.006 (-.01,0.003)
SD(Residual)	0.16 (0.14,0.23)	0.17 (0.14,0.21)	0.19 (0.17,0.22)
ϕ(AR)	0.19 (0.18,0.49)	0.28 (0.01,0.52)	0.41 (0.26,0.56)
Deviance Information Criterion (DIC)	81.8	95.8	87.54

The presence of serial correlation of repeated observations shall be addressed in regression modeling between the performance score and the covariates of interest. The Deviance Information Criterion (DIC) is useful in model comparison. The models with smaller DIC are preferred to models with larger DIC. The details about DIC can be found in Spiegelhalter et al. [35]. The minimum DIC value of model 1 confirmed that Model 1 is the best in this scenario. It might be believed that age plays a positive role to determine the on-field performances of the cricketers. Thus, it can be concluded that the senior cricketers are performing better than the young cricketers in the IPL. Thus, there is no sufficient evidence to believe that Twenty20 cricket is a game of youngsters.

The predicted posterior mean of performance scores and highest posterior density (HPD) of cricketers in IPL for individual level are shown in the box plot in Figure 1. p2, p3,…, p40 represent the consecutive names of the cricketers in the data set, the details of which are provided in Table 2. It shows that p15 (M Muralitharan), has the highest level of inconsistency in performance in comparison to others indicated by the longest box and its whiskers (*cf.* Figure 1). However, the best performance in terms of point estimates is observed with p15 (M Murlitharan) having the highest posterior mean. It may be noted that M Murlitharan is having the maximum age amongst the selected players.

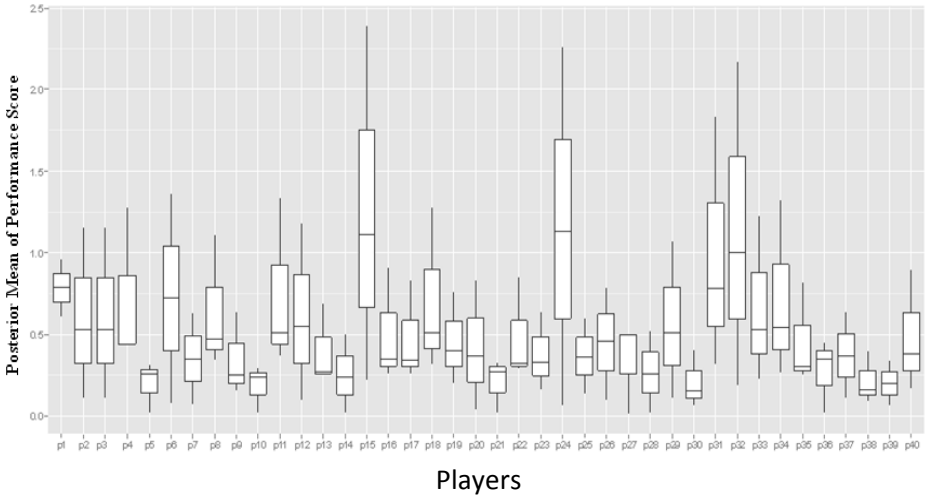

Figure 1. A box-plot representing the posterior mean and deviation of the performance score of the players.

5. Conclusion

The results obtained from the regression model confirm that the on-field performances of the cricketers are positively associated with the age of the players. It has been found that the experience (acquired over age) facilitates better performances in the IPL game. It is not true that the younger players perform better in Twenty20 cricket. Thus, the general belief that Twenty20 cricket is a game for young cricketers only has very little substance.

However, it is true that the fitness of the younger players can be better than that of the senior players and that young players are often more agile in the field. Due to the fact that the fielding measure only counts the number of catches and run outs, it may be true that this aspect is not reflected sufficiently in quantifying the performance. But fielding performance is relatively less important than batting, bowling and wicket keeping, which are the prime skills of the game. The senior cricketers with their experience can adapt quickly to the Twenty20 format though most of them have not experienced much of this format of cricket at the international level. It is interesting to note that only senior cricketers with superlative talent like Tendulkar, Dravid, Warne and Muralitharan make it to the IPL for consecutive seasons. So adjusting to the latest format and performing well may be attributed to the cricketing abilities with which the players are blessed. Many of the younger players who get the opportunity to play in the IPL do not have the same talent as the senior players as very few go through to the national or international level.

The proposed measure of performance can also be used to quantify the performances of professional football players, film stars and even politicians by changing the indicator variables. It can also be used to measure the performances of a whole cricket team in terms of batting, fielding, bowling and wicket keeping. Furthermore, this measure can also be used for comparing the performance of a cricket team in home, away and neutral grounds.

References

[1] Gupta, A., India and the IPL: Cricket's Globalized Empire, *The Round Table*, 2009; 98: 201-211.

[2] Mitra, S., The IPL: India's Foray into World Sports Business, *Sport in Society*, 2010; 13: 1314-1333.

[3] Saikia, H. and Bhattacharjee, D., On Classification of All-rounders of the Indian Premier League (IPL): A Bayesian Approach, *Vikalpa: A Journal of Decision Makers*, 2011; 36: 51-66.

[4] Saikia, H., Bhattacharjee, D. and Lemmer, H.H. Predicting the Performance of Bowlers in IPL: An Application of Artificial Neural Network*, International Journal of Performance Analysis in Sports*, 2012*a*; 12: 75-89.

[5] Raman, W.V., IPL is here to stay. [Online article] 2009*a*. Available from URL: www.tssonnet.com/tss3217/stories/20090523505202800.htm [accessed June 28, 2011].

[6] Raman, W.V., There's no substitute for skill. [Online article] 2009*b*. Available from URL: www.tssonnet.com/tss3217/stories/200904255010004000.htm [accessed June 28, 2011].

[7] Krishnaswamy, K., Of veterans and extended careers. [Online article] 2010. Available from URL: www.tssonnet.com/tss3313/stories/20100401502300400/htm [accessed June 28, 2011].

[8] Moonda, F., The IPL XI. [Online article] 2011. Available from URL: www.espncricinfo.com/indian-premier-league-2011/content/story.htm [accessed May 29, 2011].

[9] Roebuck, P., T20 will improve performance. [Online article] 2008. Available from URL: www.tssonnet.com/tss3121/stories/20080524503602600.htm [accessed June 28, 2011].

[10] Nayakan, S., T20 turns good bet for the retired stars. Business Line [serial online], 2013. Available from URL: http://www.thehindubusinessline.com/news/sports/t20-turns-good-bet-for-the-retired-stars/article4591551 [assessed April 28, 2013].

[11] Musch, J. and Grondin, S., Unequal Competition as an Impediment to Personal Development: A Review of the Relative Age Effect in Sports, *Developmental Review*, 2001; 21: 147-167.

[12] Young, W.A. and Weckman, G.R., Evaluating The Effects of Aging for Professional Football Players in Combine Events Using Performance-Aging Curves, *International Journal of Sports Science and Engineering*, 2008; 131-143.

[13] Himann, J., Cunningham, D.,Rechnitzer P. and Paterson, D., Age-Related Changes in Speed of Walking, *Medicine and Science in Sports and Exercise*, 1988; 121-127.

[14] Seiler, K., Spirduso, W. and Martin, J., Gender Differences in Rowing Performance and Power with Aging, *Medicine and Science in Sports and Exercise*, 1998; 121-127.

[15] Tanaka, H. and Seals, D., Age and Gender Interactions In Physiological Functional Capacity: Insight From Swimming Performance, *Journal of Applied Physiology*, 1997: 846-851.

[16] Fair, R. Estimated Age Effects in Baseball, *Journal of Quantitavive Analysis in Sports*, 2008; 1-39.

[17] Addona, V. and Yates, P.A., A Closer Look at the Relative Age Effect in the National Hockey League, *Journal of Quantitative Analysis in Sports*, 2010; 6: 1-19.

[18] Wood, G.H., Cricket Scores and Geometric Progression, *Journal of the Royal Statistical Society,* 1945; 12-22.

[19] Kimber, A.C. and Hansford, A.R., A Statistical Analysis of Batting in Cricket, *Journal of the Royal Statistical Society*, 1993; 156: 443-455.

[20] Lemmer, H., The Combined Bowling Rate as a Measure of Bowling Performance in Cricket, *South African Journal of Reasearch in Sport*, 2002; 37-44.

[21] Lemmer, H., A Method for the Comparison of the Bowling Performances of Bowlers in a Match or Series of Matches, *South African Journal for Research in Sport*, 2005; 91-103.

[22] Damodaran, U., Stochastic Dominance and Analysis of ODI Batting Performance: The Indian Cricket Team, 1989-2005, *Journal of Sports Science and Medicine*, 2006; 503-508.

[23] Lemmer, H., An Analysis of Player's Performances in the First Cricket Twenty20 World Cup Series, *South African Journal for Reasearch in Sport*, 2008; 71-77.

[24] Suleman, M.T., Saeed, M.A., Option on Human Performance: A Case Study of Indian Premier League. [Online article] 2008. Available from URL: www.papers.ssrn.com/sol3.../SSRN_ID1474093_code1208558.pdf [accessed December 12, 2009].

[25] Saikia, H., Bhattacharjee, D. and Bhattacharjee, A., Is IPL Responsible for Cricketers' Performance in Twenty20 World Cup?, *International Journal of Sports Science and Engineering*, 2012*b*; 6: 96-110.

[26] Van Staden, P.J., Comparison of Cricketers' Bowling and Batting Performances Using Graphical Displays, *Current Science*, 2009; 764-766.

[27] Lewis, A.J., Towards Fairer Measures of Player Performance in One Day Cricket, *Journal of the Operational Research Society*, 2005; 804-815.

[28] Bracewell, P.J. and Ruggiero, K., A Parametric Control Chart for Monitoring Individual Batting Performance in Cricket, *Journal of Quantitative Analysis in Sports*, 2009; 1-19.

[29] Clarke, S.R., Studying Variablity in Statistics Via Performance Measures in Sport. *[Invited talk on 56th session]*, Kolkata: Indian Statistical Institute, 2007.

[30] Barr, G.D.I. and Kantor, B.S., A Criterion for Comparing Selecting Batsmen in Limited Overs Cricket, *Journal of the Operational Research Society*, 2004; 55: 1266-1274.

[31] Duckworth, F.C. and Lewis, A.J., A Fair Method for Resetting the Target in Interrupted One-Day Cricket Matches, *Journal of the Operational Research Society*, 1998; 220-227.

[32] Iyenger, N.S. and Sudarshan, P., A Method of Classifying Regions from Multivariate Data, *Economic and Political Weekly*, 1982; 2048-2052.

[33] Menard, S., Longitudinal Research, Quantitative Applications in Social Sciences, London, *Sage*, 76: 1991.

[34] Tarling, R., *Statistical Modelling for Social Researchers: Principles and Practice*, Routledge: Taylor & Francis Group, New York, 2009.

[35] Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and Vander Linde, A., Bayesian Measures of Model Complexity and Fit (with discussion), *Journal of the Royal Statistical Society*, *Series B (Statistical Methodology)*, 2002; 64: 583–639.