УДК 796.015
DOI: 10.36979/1694-500X-2023-23-10-136-149

ИСПОЛЬЗОВАНИЕ УНИВЕРСАЛЬНОЙ СИСТЕМЫ ОЦЕНКИ СКВОША И ЦЕПЕЙ МАРКОВА ДЛЯ ТОЧНОГО ПРОГНОЗИРОВАНИЯ РЕЗУЛЬТАТОВ МАТЧЕЙ ПО СКВОШУ

Чейз Греппин

Аннотация. За последние несколько десятилетий спортсменов всё чаще оценивают с помощью объективных данных об их работоспособности и состоянии. Ярким примером такого перехода является сквош. Большинство игроков подключены к системе универсального рейтинга сквоша (USR), в которой каждому игроку присваивается однозначный рейтинг, отражающий его уровень игры. Рейтинги игроков могут расти или падать в зависимости от того, как они выступают против игроков с более высоким или более низким рейтингом, чем у них. В течение многих лет эта система широко использовалась как объективная мера уровня игрока. Но насколько это объективно и есть ли лучший способ прогнозировать результаты матчей, чем просто использовать рейтинги? Для этого исследования десятки тысяч предыдущих матчей были проанализированы на наличие тенденций в рейтинговой системе, и результаты более сотни матчей были предсказаны до того, как они произошли, с помощью оригинального алгоритма, который учитывает рейтинги игроков и их предыдущие результаты по сравнению с другими игроками. Этот алгоритм правильно предсказал на один матч меньше из 121 матча, чем Универсальная рейтинговая система в реальном турнире.

Ключевые слова: сквош; рейтинг; универсальная система рейтинга сквоша; воздействия; семя; импульс; расстройство.

СКВОШ БОЮНЧА МАТЧТЫН ЖЫЙЫНТЫГЫН ТАК БОЛЖОЛДОО ҮЧҮН СКВОШ УНИВЕРСАЛДУУ БААЛОО СИСТЕМАСЫН ЖАНА МАРКОВ ЧЫНЖЫРЛАРЫН КОЛДОНУУ

Чейз Греппин

Аннотация. Акыркы бир нече он жылдыкта спортчулар алардын көрсөткүчтөрү жана абалы жөнүндө объективдүү маалыматтарды колдонуу менен бааланат. Мындай өткөөлдүн ачык мисалы болуп сквош эсептелет. Көпчүлүк оюнчулар сквош универсалдуу рейтинг системасына (USR) туташкан, ал ар бир оюнчуга алардын оюн деңгээлин чагылдырган бир рейтингди ыйгарат. Оюнчулардын рейтинги алардан жогору же төмөн рейтингге ээ болгон оюнчуларга кандайча каршы чыккандыгына жараша жогорулашы же төмөндөшү мүмкүн. Көп жылдар бою бул система оюнчунун чеберчилигинин объективдүү ченеми катары кеңири колдонулуп келет. Бирок бул канчалык объективдүү жана матчтын жыйынтыгын алдын ала айтуунун жөн эле рейтингдерди колдонуудан башка жакшы жолу барбы? Бул изилдөө үчүн мурунку он миңдеген матчтар рейтинг системасындагы тенденциялар боюнча талдоого алынган жана жүздөн ашык матчтын натыйжалары оюнчулардын рейтингин жана башка оюнчуларга салыштырмалуу мурунку көрсөткүчтөрүн эске алган оригиналдуу алгоритмдин жардамы менен алар боло электе эле болжолдонгон. Бул алгоритм чыныгы турнирдеги Универсалдуу рейтинг системасына караганда 121 матчтан 120сын туура болжолдогон.

Түйүндүү сөздөр: сквош; рейтинг; универсалдуу сквош рейтинг системасы; таасири; үрөн; импульс; бузулуу.

UTILIZING THE UNIVERSAL SQUASH RATING SYSTEM AND MARKOV CHAINS TO ACCURATELY PREDICT OUTCOMES OF SQUASH MATCHES

Chase Greppin

Abstract. Over the past few decades, athletes have increasingly been evaluated with objective data on their performance and condition. A prime example of this evaluation has been in the sport of squash. Most players are connected within the Universal Squash Rating system (USR) in which each player is given a single number rating to reflect their level of play. Players' ratings can rise or fall depending on how they perform against players with a higher or lower rating than them. For years this system has been widely accepted as an objective measure of a player's level. But just how objective is it and is there a better way to predict results of matches than just using ratings? For this study, tens of thousands of previous matches were analyzed for trends in the rating system and over a hundred match results were predicted before they happened with an original algorithm that considers players' ratings and their previous performance against other players. This algorithm predicted one fewer match correctly out of 121 matches than the Universal Rating System in a real tournament.

Keywords: squash; rating; Universal Squash Rating System; exposures; seed; momentum; upset.

Introduction. Squash is a competitive racquet sport that is similar to racquet ball. Two players are enclosed in a four-wall court that is approximately 10m by 6.5 m and alternate hitting a small black ball against the front wall. The area of play (where the ball can be hit) is bounded by a line near the bottom of the court and near the top of the court on the front wall. The line near the top of the front wall then slopes down across the side walls where it generally meets a glass back wall for spectators to watch. Squash consistently ranks as one of the healthiest sports in the world because of how it helps players burn calories and improve overall health.

In recent years, junior squash (players up to 18 years old), especially in America, has become increasingly competitive and popular. Tournaments consistently waitlist players because demand for tournaments is greater than supply. Every junior squash player is given both a rating and a ranking. These measures are independent, and players are not ranked according to their rating. The Universal Rating System provides all players a single number rating to reflect their level of play. Rankings, however, are calculated through a point average system, and these rankings determine if players can get into tournaments. There are different levels of tournaments and depending on how players place in a certain level tournament, they get a certain number of points. For instance, if a player wins a gold level tournament, they get 2,000 points. Depending on how many exposures they have (the number of tournaments they have played over an 11-month period), their highest point values are averaged and then ranked against other players. This perpetual game to get one's ranking up is what draws players into the sport. However, there is a delicate balance in choosing which tournaments to go to. For instance, if one plays in 12 or fewer tournaments over an 11-month period, their ranking is calculated by averaging their four highest point values from tournaments. But if one plays in 13 tournaments, their ranking is the average of their five highest point values. So, players only want to go to tournaments where they know that they can get their point averages higher. This is why a match predictor is important.

There is no existing squash match predictor besides that of the Universal Rating System. The ranking system seems like an objective way to predict the outcome of a match. However, the rating system is the true testament to a player's level even though players are not ranked according to it. Even if Player A is ranked higher than Player B (player A has a higher point average), Player B would still be favored to win the match if he has a higher rating. So, in the development of the predictor algorithm, players' rankings were not considered. Rather, it focused on players' ratings in relation to their opponents, and their performance in their three prior matches. With access to a predictor algorithm that can predict match results with more or the same accuracy as rating, ranking, or seed (how a player ranks in relation to others in a tournament), squash players can make educated decisions on which tournaments to go to, rather than basing their choice on subjective factors [1].

Literature Review. In any sport, a system to judge the level of a team or player is important. Squash utilizes a unique system. Unlike most sports which base a player or team's level on their record, squash assesses a player's level with the use of the Universal Rating System. This system "connects all players through a global network of match results [and] is a long-term measure of their level of play" [1]. The ratings span from 1.5 for a beginner to 7 for a professional. Players' ratings go up or down depending on how they perform against other players of different ratings. Especially for competitive juniors who play many matches, the rating system is considered to be an objective measure of a player's level. The predictor algorithm discussed in this paper heavily relied on this rating system and its objectivity at certain intervals.

Other systems for ranking squash players against each other exist outside of the Universal Rating System as well. A more recent system called Squash Levels also provides squash ratings, albeit on a different scale, along with providing a social network for the sport. This system ranges from 0 for beginners to 80,000 for professionals. This system produces "levels [that] are accurate enough that you can make result predictions or calculate handicaps" [2]. Unlike the Universal Rating System, Squash Levels employs a "confidence level" when it predicts match outcomes. Confidence levels can increase when players play many matches or decrease if they are consistently losing to players that have lower ratings than them. In addition, levels not only can predict match results, but point breakdowns as well. For instance, if a player has a rating of 2000, they are predicted to win two out of every three points if they play another player with a level of 1000. This system was not considered in this paper, as it is still relatively new, and the majority of junior squash players rely on the Universal Squash Rating System to get their ratings.

Similar research has been done to determine the objectivity of the Universal Rating System. Assessing nearly 80,000 matches, Varun Fuloria, Rutwik Kharkar, and Ryan Rayfield found that a 0.1 increase in rating gap resulted in a 15 % increase in the likelihood of the higher rated player winning the match [3]. Denis L. Bourke and Robert H. Eather performed a similar analysis on the Universal Squash Rating System in which they determined how much more likely a player was to win a game or match when their PwP, the probability they had to win any given point based on their rating compared to their opponents, increased, or decreased [4]. They found that closely rated players would have an unexpectedly higher chance of winning a game or match when their PwP was slightly increased. Additionally, their empirical probability of winning was closely correlated to the theoretical probability of winning. This is also closely correlated with the data collected for this article.

While little prior research has been done on predicting squash match results, similar work has been done for tennis match results. A large portion of these tennis predictors are based on Markov Chains, the notion that the probability of one thing happening is dependent on prior things that have happened [5]. The prediction model developed for this article heavily relied on these chains. The likelihood of a player winning a match was dependent on how they performed against players of various ratings in the past. Other tennis prediction models such as that produced by William J. Knottenbelt, Demetris Spanias, and Agnieszka M. Madurska are based on how both players performed against a common opponent in the past [6]. The constant drive to ameliorate tennis prediction models is driven by the world of sports betting. One model created by Michal Sipko utilized machine learning methods when considering vast arrays of past data such as how tired a player was when he played a match and was able to generate a return of investment of 4,4 % when betting on ATP matches from 2013–2014 [7]. A contrasting study in 2022 found that considering other factors other than just ranking did not significantly improve the prediction accuracy [8]. While none of these studies directly impacted the model discussed in this paper due to having different foci and dealing with another sport, they were considered. For example, the initial model was going to have a factor that reflected how tired a player was going into their match, but the 2022 study proved that there was no need for such a factor.

Methodology

Data Analysis Methodology. For this paper, over 35,000 previous matches were assessed to find trends within the Universal Rating System.

Additionally, the matches that were assessed were held within three parameters: the matches were male vs. male, the ratings had a value of 3 or greater, and the ages of the players were not more than 18 years old (the oldest age in junior squash). Junior squash was the only age group assessed because it has the highest frequency of tournaments. Males were the only gender assessed because there were more male matches than female matches and the rating trends between the two genders could have differed. The reason that no match with a rating below 3 was assessed was because those players are generally beginners whose ratings are much more volatile than better, more developed players.

The data was received from Club Locker, an affiliate of US Squash that runs the rating system and logs data on matches such as players' genders and the scores of games in the matches. The data set contained information on 305,000 squash matches from 2018 to 2022. These matches were then filtered to meet the three parameters discussed above. Ultimately, 35,238 matches met the parameters. The only data that was considered was the rating of the winner and the loser, as well as how many games the matches comprised (squash matches are scored in a best of 5 format). The principal goal of this analysis was to determine the likelihood a player had at winning a match, given a specific rating difference. This data would then be used when making the predictor algorithm.

The 35,238 matches were further filtered into specific rating groups to determine how accurate the rating system was at certain intervals of rating difference. The matches were classified as either an expected win or an upset (where the lower rated player beat the higher rated player) before finding the trends. The rating gaps assessed were 0.01–0.04, 0.05–0.09, 0.10–0.14, 0.15–0.19, 0.20–0.24, 0.25–0.29, 0.30–0.34, 0.35–0.39, 0.40–0.44, 0.45–0.49, and anything greater than 0.49. For an expected win, the winning player had a rating that was that much higher than the player they beat. For an upset, the lower rated player was rated an equal amount lower than the higher rated player. The number of matches in each interval were counted with a true/false analysis in Excel. If a match met a certain criterion, such as having a rating gap between 0.1 and 0.14, the match was classified as true. The number of expected wins and upsets were all counted (Table 1, Figure 1).

Table 1 – Likelihoods of expected wins and upsets at various rating intervals. The group percentage of expected results represents an objective measure of the Universal Rating System's accuracy

Rating gap	Percentage of matches in which the higher rated player won	Percentage of matches in which the lower rated player won
Group (every match)	85 %	15 %
0.01–0.04	56 %	44 %
0.05–0.09	62 %	38 %
0.10–0.14	70 %	30 %
0.15–0.19	75 %	25 %
0.20–0.24	81 %	19 %
0.25–0.29	86 %	14 %
0.30–0.34	89 %	11 %
0.35–0.39	92 %	8 %
0.40–0.44	94 %	6 %
0.45–0.49	95 %	5 %
>0.49	99 %	1 %

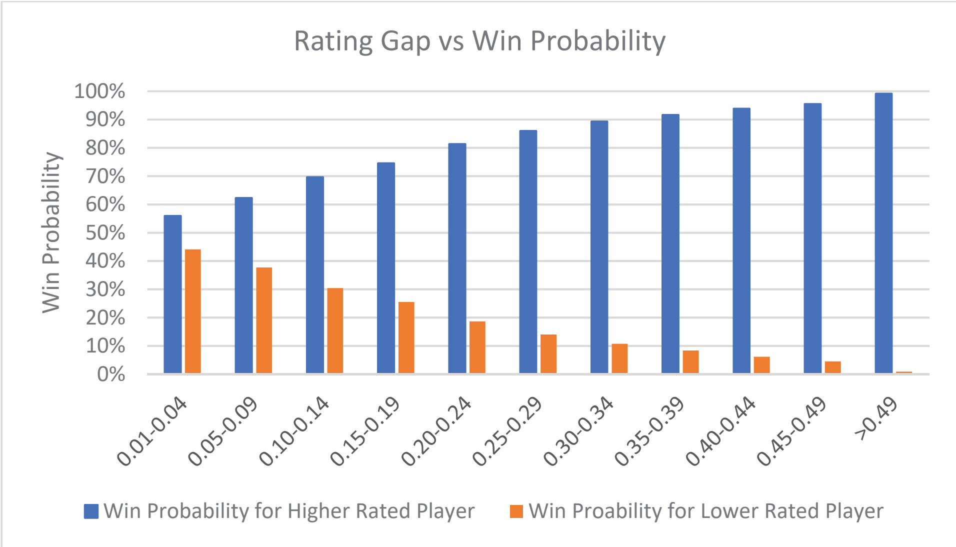

Figure 1 – Percentage of expected results vs upsets at various intervals. The graph reveals expected trends such as a consistent increase in win percentage as rating intervals increased

The findings showed expected trends within the Universal Rating System and gave a clear measure to the objectivity of the system. The greater the rating gap, the greater the chance there was of an expected win (where the higher rated player won).

The second part of this analysis determined how many games it took the higher or lower rated player to win the match. A similar process was used to find the number of games it took to win to the previous analysis on win probability. The spreadsheet listed the score of the winner and the loser for each of the five games. If games four and five had 0's in them then the match lasted three games. If only game 5 had a 0 in it, then the match was four games. If every game had a number in it, then the match was five games. A match would be classified as "true" if it met certain criteria such as having a rating gap of 0.26 and being only a three-game match. Each match could then be classified as either an expected win or an upset along with the number of games that the match lasted.

The findings for expected wins demonstrated expected trends such as an increase in the number of 3 game wins as the rating differential increased and an increase in the number of four game and five game matches as the rating differential decreased. The only anomaly was that the percentage of 3 game wins with a rating differential greater than 0.49 was lower than that of the 0.45–0.49 differential. A possible explanation could be that when a player is rated that much higher, they might not try as hard even though their opponent is giving their full effort (Table 2, Figura 2).

The findings for upsets followed similar expected trends but were less defined. Overall, upsets were more likely to occur in three games if there was a lower rating differential. Likewise, the percentage of 4 game upsets, and especially 5 game upsets increased as the rating differential grew (Table 3, Figura 3).

Predictor Methodology. The prediction algorithm was based on two factors: a rating factor and a momentum factor. In the end, each player was given a score that was the sum of their rating and momentum score, and the higher of the two scores would be the player that was expected to win.

The rating factor was based on trends that I found by analyzing the data. The score that a player received for their rating depended on how much higher or lower their rating was than their opponents'. A player's score for their rating was the accuracy of the rating system at that point multiplied

Table 2 – Percentage of 3, 4, and 5 game expected wins at various intervals

Rating Gap	Percent Chance of a 3-game win	Percent Chance of a 4-game win	Percent Change of a 5-game win
0.01–0.04	40 %	35 %	25 %
0.05–0.09	42 %	36 %	22 %
0.10–0.14	46 %	33 %	20 %
0.15–0.19	48 %	35 %	17 %
0.20–0.24	54 %	31 %	15 %
0.25–0.29	60 %	28 %	12 %
0.30–0.34	63 %	26 %	11 %
0.35–0.39	70 %	22 %	8 %
0.40–0.44	73 %	21 %	7 %
0.45–0.49	77 %	17 %	6 %
>0.49	72 %	19 %	9 %

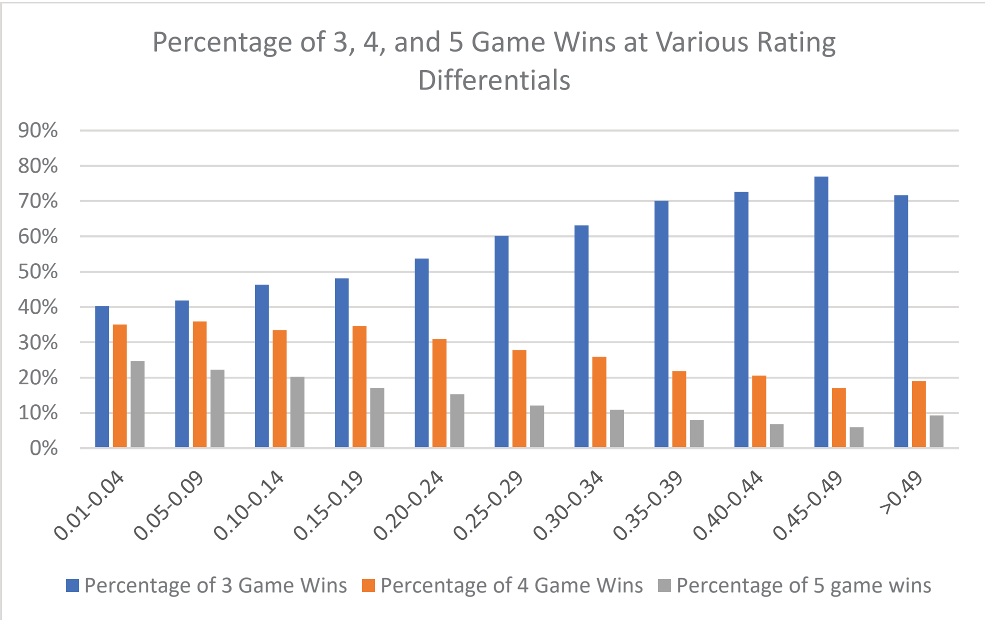

Figure 2 – Percentage of 3, 4, and 5 game expected wins at various rating intervals broken down visually

Table 3 – Percentage of 3, 4, and 5 game upsets at various rating intervals

Rating Gap	Percentage of 3 game upsets	Percentage of 4 game upsets	Percentage of 5 game upsets
0.01–0.04	34 %	39 %	27 %
0.05–0.09	34 %	36 %	30 %
0.10–0.14	33 %	35 %	31 %
0.15–0.19	29 %	35 %	36 %
0.20–0.24	28 %	39 %	32 %
0.25–0.29	30 %	36 %	34 %
0.30–0.34	27 %	36 %	37 %
0.35–0.39	24 %	32 %	44 %
0.40–0.44	29 %	32 %	39 %
0.45–0.49	17 %	52 %	31 %
>0.49	31 %	27 %	42 %

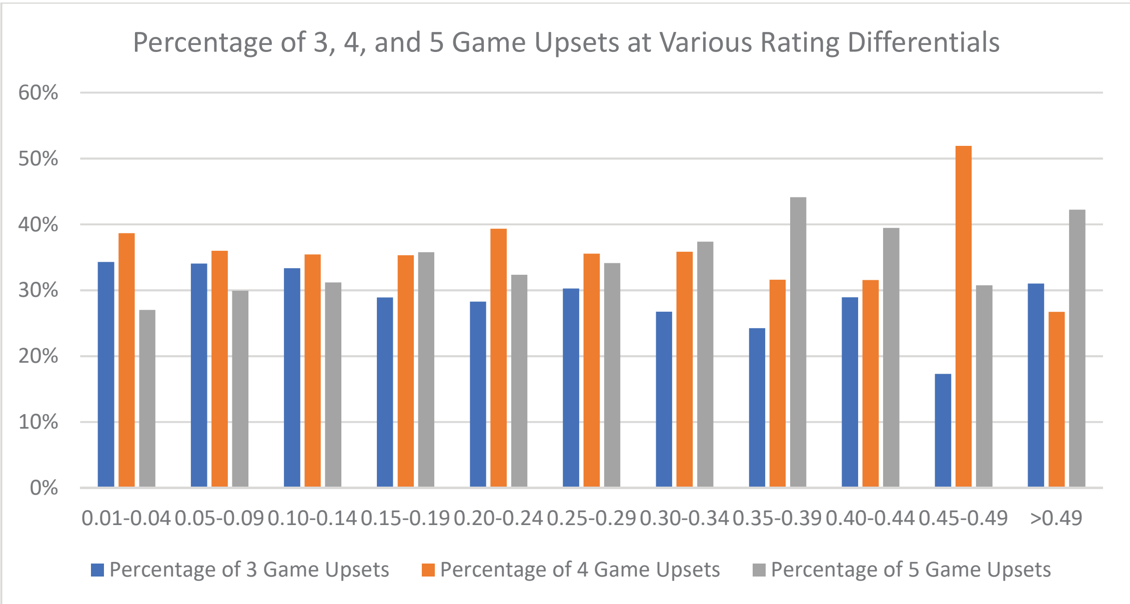

Figure 3 – Percentage of 3, 4, and 5 game upsets at various rating intervals visualized

by 100. For instance, if a player had a rating that was 0.52 higher than their opponents', they would have a very high rating score because at that rating gap the Universal Rating System was very accurate. However, if a player had a rating that was only 0.02 higher than their opponents, their score would be much lower because the rating system is less accurate when the interval is smaller. The player with the lower rating would get the difference of the higher rated players score and 100. Matches where players had equal ratings were 50 points for each player (Table 4).

The points reflect how likely a certain player is to win at a certain rating interval

The momentum factor was designed to address the Universal Rating System's inaccuracies. Its purpose was to determine whether players were proving that they deserved the rating that they had. For instance, if a player had a rating 0.30 higher than his opponent but was consistently losing to players with ratings 0.35 lower, then his momentum score would essentially cancel out his rating score because the player could not prove that he deserved the rating he had.

There were four possible outcomes of a match: 1) the higher rated player beat the lower rated player, 2) the higher rated player lost to the lower rated player, 3) the lower rated player beat the higher rated player, and 4) the lower rated player lost to the higher rated player. If the higher rated player beat the lower rated player, he would get a high score because it showed that he deserved the rating that he had. A player would get more points for winning in 3 games vs winning in 5 games and more points if the rating gap was larger. If the higher rated player lost to the lower rated player, he received a negative score because losing demonstrated that he did not deserve the rating he had. If the lower rated player lost to the higher rated player, he did not receive very many points because that showed that he could not beat someone of a higher rating. If the lower rated player beat a higher rated player, he received a very high score, especially for three games wins because this demonstrated that he could beat players that had higher ratings.

In each scenario, a played would receive a score based on the data analysis which would make up a portion of their final momentum score. The algorithm looked three matches into a player's past and assessed the scores of those matches and the rating differences. The prior match represented 50 % of the final momentum score, the second match before 30 %, and the third match before 20 %. Consistent poor prior performance would significantly lower a player's chances of winning, but a single bad match would not.

Below are the tables for the scores players would receive for their performance in a match with 1 of the four scenarios.

Scenario 1: Higher rated player beats lower rated player (Table 5).

Table 4 – Points a player would receive for their rating factor score.
Reference Table 1

Rating difference	Points for Higher Rated Player	Points for Lower Rated Player
0.00	50.0	50.0
0.01–0.04	55.9	44.1
0.05–0.09	62.3	37.7
0.10–0.14	69.6	30.4
0.15–0.19	74.5	25.5
0.20–0.24	81.3	18.7
0.25–0.29	86.0	14.0
0.30–0.34	89.3	10.7
0.35–0.39	91.7	8.30
0.40–0.44	93.9	6.10
0.45–0.49	95.5	4.50
>0.49	99.1	0.90

Table 5 – Scores a higher rated player received for beating a lower rated player at various rating intervals and match scores

Rating difference	3 game Win	4 game Win	5 game Win
0.00	20.00	16.00	14.00
0.01–0.04	22.49	19.59	13.83
0.05–0.09	26.07	22.33	13.86
0.10–0.14	32.25	23.26	14.10
0.15–0.19	35.86	25.87	12.78
0.20–0.24	43.71	25.21	12.42
0.25–0.29	51.71	23.88	10.38
0.30–0.34	56.34	23.16	9.76
0.35–0.39	64.29	20.00	7.36
0.40–0.44	68.12	19.34	6.39
0.45–0.49	73.50	16.33	5.65
>0.49	71.04	18.87	9.20

Table 6 – Scores a lower rated player received for losing to a higher rated player at various rating intervals and match scores

Rating difference	3 Game Loss	4 Game Loss	5 Game Loss
0.00	9.00	11.00	13.00
0.01–0.04	13.83	18.16	22.49
0.05–0.09	13.86	19.96	26.07
0.10–0.14	14.10	23.17	32.25
0.15–0.19	12.78	24.32	35.86
0.20–0.24	12.42	28.06	43.71
0.25–0.29	10.38	31.04	51.71
0.30–0.34	9.76	33.05	56.34
0.35–0.39	7.36	35.83	64.29
0.40–0.44	6.39	37.26	68.12
0.45–0.49	5.65	39.57	73.50
>0.49	9.20	40.12	71.04

Table 7 – Scores a lower rated player received for beating a higher rated player at various rating intervals and match scores

Rating difference	3 Game Win	4 Game Win	5 Game Win
0.01–0.04	86.17	59.88	28.96
0.05–0.09	86.14	60.43	26.50
0.10–0.14	85.90	62.74	31.88
0.15–0.19	87.22	64.17	35.71
0.20–0.24	87.58	66.95	43.66
0.25–0.29	89.62	68.96	48.29
0.30–0.34	90.24	71.94	56.29
0.35–0.39	92.64	75.68	64.14
0.40–0.44	93.61	76.83	67.75
0.45–0.49	94.35	80.04	73.93
>0.49	90.80	81.84	77.51

In this scenario a player received the highest score by winning in three games with a large rating differential. No matter the rating gap, a three-game win always gave a higher score than a four game or five game wins. The larger the rating gap is, the less likely the higher rated player should be to lose one game or two. Therefore, the higher rated player received much lower scores, especially with high rating gaps, if he dropped one or two games. These scores were calculated by multiplying 100 by the percentage a player had at winning the match at a given interval by the percentage the player had at winning that match in three, four, our five games. For example, for the rating gap 0.45-0.49, a player had a 95% chance of winning the match along with a 77% chance of winning that match in three games resulting in a score of 73.50.

Scenario 2: Lower rated player loses to higher rated player (Table 6).

The general theme was that players did not receive a high score for losing if they were expected to lose because they demonstrated that they could not beat players that had higher ratings than them. However, players would get a larger score if they took one game or two games, especially against players with much higher ratings than theirs. With the exception of the scores players received in the fourth game, the scores were a mirror image of the scores for if the higher rated player beat the lower rated player. That is, players would receive the same number of points for winning a match in 5 games if they were rated higher than if they lost the match in three games with a lower rating. The reason that a loss in four games did not carry the same score as a win in four games was because a player should receive more points for winning in four games than losing in four games. For that reason, the score a player received for losing in four games if they were the lower rated player was the average of the score for three games and five games.

Scenario 3: Lower rated player beats higher rated player (Table 7).

This was the scenario that resulted in the greatest number of points for a player. In this scenario, players were heavily rewarded for beating a player with a higher rating, even if the rating differential was close because it demonstrated that they could beat players that were rated higher than them. The scores were based off the scores a player received for losing to a player with a higher rating. If a player beat another player with a higher rating in three games, the score they received was 100 minus the score they would have gotten for losing to that player in three games. If the upset was in four games, the score was 100 minus the mirror from top to bottom of the scores that player would have received for losing in four games. For instance, if a player beat another player who had a rating 0.35 higher than his in four games, his score would be 100 minus the score he would have received for losing to a player that was rated 0.15–0.19 higher than him. The reasoning for this top to bottom mirroring was to account for the fact that a player should receive more points the larger the rating differential got. If the scores were not mirrored, they would have received fewer points the larger the differential was. The score a player received for beating a higher rated player in 5 games was calculated the same way a four-game win was.

Scenario 4: higher rated player loses to lower rated player (Table 8).

If a higher rated player lost to a lower rated player, his momentum score would be significantly lowered because he proved that he could not maintain his rating in a match scenario. Players were especially hurt if they lost in three games with larger rating differentials. The scores were simply the negative value of the score a lower rated player would have received for beating a higher rated player.

Here is an example of how a match was predicted (Table 9).

Based on this analysis player 2 would be favored to win even though he has a lower rating than player 1.

Findings. The algorithm was then tested in a real scenario. Every boy's main draw match at the Lifetime City Center Junior Gold tournament in Houston Texas was predicted with the model. Draws varied in sizes ranging from 16 to 32 players and encompassed all the age divisions: Boys Under 11 (BU11), BU13, BU15, BU17, and BU19. In total, 121 match results were predicted. The accuracy of the prediction model was compared with that of the Universal Rating System and the seeding system in that tournament. A seed is assigned to every player before the tournament starts and is their

Table 8 – Scores a higher rated player received for losing to a lower rated player at various rating intervals and match scores

Rating differential	3 Game Loss	4 Game Loss	5 Game Loss
0.01–0.04	-86.17	-59.88	-28.96
0.05–0.09	-86.14	-60.43	-26.50
0.10–0.14	-85.90	-62.74	-31.88
0.15–0.19	-87.22	-64.17	-35.71
0.20–0.24	-87.58	-66.95	-43.66
0.25–0.29	-89.62	-68.96	-48.29
0.30–0.34	-90.24	-71.94	-56.29
0.35–0.39	-92.64	-75.68	-64.14
0.40–0.44	-93.61	-76.83	-67.75
0.45–0.49	-94.35	-80.04	-73.93
>0.49	-90.80	-81.84	-77.51

Table 9 – An example of what a prediction model looked like taking into account the rating and momentum factors

	Player 1	Player 2
Rating	3.38	3.36
Rating Gap	0.02	-0.02
Rating Factor Score	55.9	44.1
3rd match before result	1–3	3–1
3rd match before rating gap	0.04	-0.1
3rd match before score	-59.88	62.74
2nd match before result	2–3	3–2
2nd match before rating gap	-0.05	-0.07
2nd match before score	26.07	26.50
1st match before result	3–0	3–0
1st match before rating gap	0.53	0.4
1st match before score	73.50	56.34
Momentum Score	32.6	48.7
Final Score	88.5	92.8

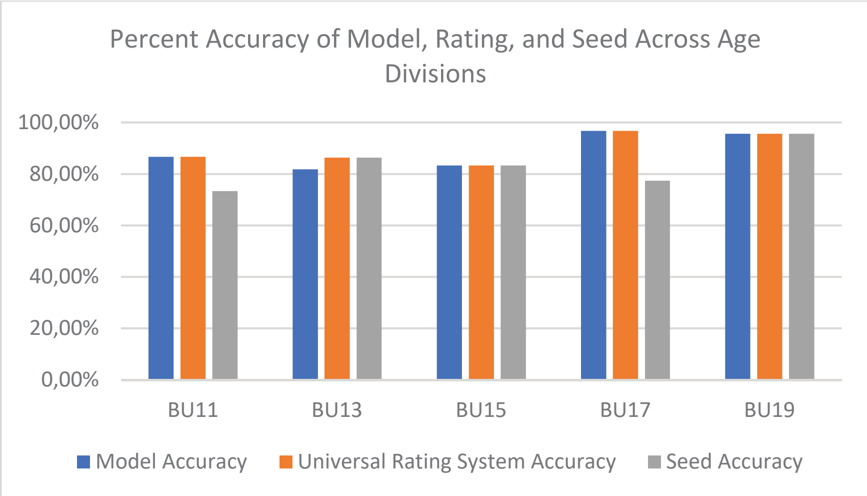

Figure 4 – Comparing the prediction accuracies of the model, the Universal Rating System, and the seeding system in various divisions

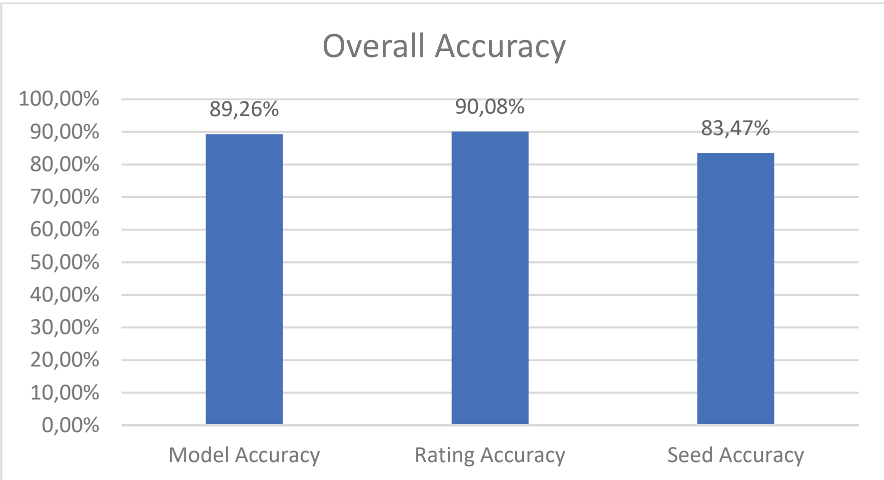

Figure 5 – Overall accuracies between the model, Universal Rating System, and seeding system visualized

expected finishing position according to their ranking in relation to others in the tournament.

The model predicted matches with the same accuracy as the Universal Rating System in all but one of the divisions. Additionally, it either had a higher accuracy or the same accuracy as the seeding system in all but one of the divisions. Overall, the model had an accuracy of 89.26 % compared to the Universal Rating System's 90.08 % and the seeding system's 83.47 % accuracy. The Universal Rating System predicted one more match correctly than the model in this specific tournament. Of the more than 35,000 matches that were assessed, the Universal Rating System held an overall accuracy of 85 %. The model's accuracy of 89.26 % surpasses that of the Universal Rating System as a whole in junior boys' squash, although the Universal Rating System had a higher accuracy in this specific tournament. The discrepancy between the Universal Rating System and the prediction algorithm occurred when the algorithm incorrectly predicted the lower rated player to win (Figure 4, Figure 5).

Discussion. Like many of the tennis match prediction models, such as that created by Michal Sipko in 2015, this squash prediction model utilized Markov Chains to accurately predict match results [7]. Additionally, this model was based on tennis match prediction studies that found that considering data outside of just rankings did not significantly improve prediction accuracy [8]. While the squash model considered primarily rating and past match results rather than ranking, it did not consider factors such as how tired a player was when he played a match or a player's record against their opponent.

One limitation to this model is that it can only predict one match at a time. The model can predict who wins and who loses the match, but it cannot predict the score of that match. Because the momentum factor relies on the score of a player's previous match, the model can only predict one match into the future rather than several. Perhaps by assessing trends in the differences in players final scores based on the prediction models, trends could be determined to predict the number of games the match would last. For example, a difference of 30 in two players scores could mean that the match would be won in 4 games.

Tying into the previous point, fewer than 150 matches were predicted with the current model. If more matches were predicted, trends could be identified and assessed to improve the model's accuracy. In the previous case, the more matches that are predicted, the easier it would be to find relationships between score differences and final match scores.

Perhaps the most significant factor that led to the model being less accurate were unprecedented results. This meant that neither a player's rating score nor momentum score could account for the result of the match. In a particular instance a player with a 5.50 rating who was consistently beating players rated lower than him lost to another player with a 5.26 rating who in his most recent match failed to beat a player with a lower rating than him. Both in terms of momentum and rating, the player with the 5.50 rating was favored to win, and yet he lost 3–0. There is no statistical precent for this result so the player who had a 5.50 rating simply could have had a bad day on court.

The current model is very flexible to new research and methods. Future research could include a past performance factor, separate from the momentum factor. For example, in the previously discussed scenario with the player with a 5.50 rating, perhaps that player consistently performed poorly when he played opponents that were rated 0.24 lower than him but performed well against players that were rated 0.10 lower than him. In theory, it does not make sense that a player would perform better against closer opponents, but that could just be the nature of the individual player. A sensible addition could be a factor that looks farther into a player's past and assesses his overall performance against players of different ratings.

Conclusions. This study, assessing 35,238 matches, showed that using Markov Chains to assess a player's past performance combined with a player's rating produced a prediction model that came within 1 % of the accuracy of the Universal Rating System. The model focused on boys' matches in ages up to 18 years old. Data analysis was performed before the formation of the model so that trends within the Universal Rating System could be found and applied. Every score that a player received in the prediction model was based on assessed data. With a model that is more accurate than the Universal Rating System, junior players on the US Squash circuit of tournaments can make more educated decisions about which tournaments to attend based on their probability of winning various matches.

Поступила: 21.08.23; рецензирована: 05.09.23; принята: 08.09.23.

References

1. United States Squash (2023). "Universal Squash Rating". URL: https://www.ussquash.org/ratings/ (accessed: 23.06.2023).
2. SquashLevels (2023). About the SquashLevels System. URL: https://es.squashlevels.com/doc.php?doc=about_squash_levels.htm (accessed: 23.06.2023).
3. *Fuloria, Varun, Rutwik Kharkar, and Ryan Rayfield*. "The Ratings Game – an Analysis of US Squash Ratings". Squash Magazine (11.06.2021). URL: https://www.squashmagazine.com/2021/11/the-ratings-game-an-analysis-of-us-squash-ratings/
4. *Bourke Denis L., and Robert H. Eather.* "Community Perspectives: A Probability Analysis of the Game of Squash". Squash Magazine (26.06.2023). URL: https://www.squashmagazine.com/2023/06/community-perspectives-a-probability-analysis-of-the-game-of-squash/
5. *Madurska, Agnieszka M*. "A Set-by-Set Analysis Method for Predicting the Outcome of Professional Singles Tennis Matches". Imperial College London, Department of Computing: London, UK (2012).
6. *Knottenbelt, William J, Demetris Spanias, and Agnieszka M Madurska*. "A Common-Opponent Stochastic Model for Predicting the Outcome of Professional Tennis Matches". Computers & Mathematics with Applications 64, no. 12 (2012): 3820–27.
7. *Sipko, Michal, and William Knottenbelt.* "Machine Learning for the Prediction of Professional Tennis Matches".MEng computing-final year project, Imperial College London 2 (2015).
8. *Yue, Jack C, Elizabeth P Chou, Ming-Hui Hsieh, and Li-Chen Hsiao*. "A Study of Forecasting Tennis Matches Via the Glicko Model". Plos one 17, no. 4 (2022): e0266838.