COMPARISON OF TWO COMPUTER RANKING ALGORITHMS (ITA AND ELO) APPLIED TO COLLEGE SQUASH

April 2015

Introduction

This report describes two ranking algorithms (ITA[1] and ELO[2]) applied to the 2014-15 CSA season.

The first algorithm we describe and apply is currently used by the Intercollegiate Tennis Association (ITA) to rank college teams and individuals. At the start of each season a ranking committee determines a pre-season ranking list. After a "sufficient" number of matches have been played, the ITA switches to a computer system where points are accumulated for beating the "NBEST" highest ranked teams based on current rankings. The number of points gained for beating the top ranked team is 106; for beating the second ranked team 102; the third ranked 98, etc. In college tennis, the value of NBEST increases from 4 to 10 during the season as match results are accumulated. Teams are penalized for losses. The penalty for losing to higher ranked teams is less than the penalty for losses to lower ranked teams. The ITA has applied its algorithm for several years to college tennis. Details are described in a Ranking Manual available online[1] to coaches and players. The only clearly subjective input is the initial coaches ranking poll. By all accounts, coaches and players are satisfied with the ranking method. However, in applying the ITA method to CSA results we find some troubling characteristics: (a) initial ranking affects ranking position late into the season, and (b) details of match schedule affect ranking predictions. Nevertheless, widespread acceptance of the ITA ranking method by the ITA constituency should give us hope that if CSA chooses to adopt an objective computer ranking method, it will be accepted.

The second ranking we describe in detail (beginning on p 17) is a variant of the ELO chess rating system. ELO has been applied for many years to a variety of sports and is closely related to Bradley-Terry rating[3]. (**We thank Vir Seth for sharing a write-up of**

work he did as a senior at St. Lawrence University applying the Bradley-Terry method to the CSA 2013-14 season). ELO is one of the rating systems chosen by Jeff Sagarin[4], the well-known sports statistician, whose ratings help determine the participants in the NCAA Mens Division 1 Basketball Championship Tournament as well as the Bowl Championship Series of college football. ELO awards / penalizes rating points by an amount proportional to the difference between how teams are “expected” to perform (based on current ranking points) and how they actually perform. The probabilistic underpinnings of ELO will be explained, as will the ideas behind expectation of performance. The variant of ELO we use allows a self-consistent calculation of ranking that takes into account *all* matches played. This removes any dependence on match schedule. Pre-season ranking plays no role in ELO ranking (Sequential ELO or Self-Consistent ELO), and no adjustable parameters appear in the self-consistent version.

Both ITA and ELO ranking methods were applied to the 2014-15 men’s CSA team results through the end of the regular season (Feb 15). As a reality check we compared ITA and ELO ranking predictions with the CSA pre-tournament rankings. Whereas ELO provided sensible rankings through all five divisions of play, the ITA method produced unsatisfactory predictions outside of the top 25 teams. We also applied ELO to the women’s 2014-15 season, again finding sensible ranking predictions, confirming that Self-Consistent ELO is a good candidate for adoption by the CSA as a reliable, objective computer ranking system.

ITA Rankings

The first ranking method we study is one currently employed by the Intercollegiate Tennis Association (ITA). The following is extracted from the ITA Ranking Manual[1]

ITA Rankings GUIDELINES AND RULES – TEAM

1. ***The first six national top 75 team rankings of the spring will be decided by vote of the ITA National Ranking Committee. For the remainder of the spring dual match season, the rankings will be based on the ITA computer ranking system (beginning February 24). For each countable victory and all losses a team receives a prescribed number of points (see point chart) based upon the national ranking of the opponent defeated. Victories and losses in ITA-sanctioned college dual matches will count towards the team ranking.***

2. *A team is worth its current value/ranking. If a team drops or climbs during the season, it will always be worth its current ranking each given period. Each ranking period, the ranking average will be figured with the total of countable victories and all loses. If the team has fewer ranked victories than the countable victory total for the period, the rest of the counted victories will be its unranked victories. If the team has more ranked victories than the countable victory totals, the team's highest countable victories will be those counted. All losses will be considered as countable matches, but losses are also weighted according to opponent rank.*

3. ***The way the points are distributed – points for wins; percentages deducted for losses – they consider a team's won- loss record, strength of schedule and depth of wins and losses; and significant wins and losses (with bonus points for road wins). The formula works as follows: Sum of points from 'x' best wins for that rankings period divided by [the 'x' best countable wins for that particular ranking period + Points from all losses].***

4. *The ITA National Ranking Committee can review Nos. 51 through 75 in the first five ITA computer team rankings and has the authority to adjust the rankings in that area to ensure the most-deserving teams enter into the rankings.*

5. *Shortened or different formats approved by the ITA can also count towards rankings (if both coaches have agreed on this prior to the match).*

6. *Non-division I opponents count as unranked wins and/or losses.*

7. *The NCAA team champion automatically goes to No. 1 in final ranking. Bonus points are awarded for advancement in the NCAA Team Championships (see point chart).*

The ITA Rating formula (para 3 of ITA Rankings GUIDELINES AND RULES) has the form

$$R_i = \frac{\sum_{j=1}^{NBEST} \text{Winpoints}_j}{NBEST + \sum_{j=1}^{NWORST} \text{Losspoints}_j}, \text{ for all teams i} \qquad (1)$$

- Winpoints $_j$ are points won by team "i" for beating team "j". The number of points won depends on the rank of team j on the day the rankings are calculated - *not the rank of j on the day the match took place!*

- Similarly, Losspoints $_j$ are points which count against team i for losing to team j. The number of points that count against team i depends on the rank of team j. *Again, it is the rank of team j on the day of the ranking calculation that matters*.

- Once the rating points R_i have bean calculated for each team on the given ranking date, the team rank is calculated by sorting the rating points in decreasing order. The team with the largest number of rating points is the number one ranked team; the team with the second largest number is the second ranked team, etc.

- As the season progresses and team ranks change from one ranking date to the next, the value of the rating points for a given team may change, even if that team has played no matches during this period. This is also true in the present CSA ranking system.

- It is only the points for each of the NBEST "best" wins and NWORST "worst" losses that count. "Best" for team i means count matches from the NBEST highest ranked teams that team i has wins against. "Worst" for team i means count matches from the NWORST lowest ranked teams that team i has losses against.

- For tennis rankings, "NBEST" are the so-called "countable matches", and they increase from 4 to 10 as the season progresses.

- ITA tennis teams play many more matches than CSA ($\gtrsim$ 30 typically, 26 for Princeton this year)

- There is a strong argument for using less countable matches for CSA and limiting number of countable losses (eg NBEST = 5, NWORST = 5), as will be explained later.

The number of points won and lost are shown in Table 1 below[1].

ITA RATING POINT ASSIGNMENT CHART (for TEAMS)

ITA Rankings

Point chart and guidelines for NCAA Division I Men's and Women's Tennis – 2014-15

POINT CHART – TEAM

Win over No.	Pts.	Win over No.	Pts.	Win over No.	Pts
#1	106	#39	43	#76-85	16
#2	102	#40	42	#86-95	14
#3	98	#41	41	#96-105	10
#4	94	#42	40	#106-115	8
#5	91	#43	39	#116-125	6
#6	88	#44	38	Not ranked	4
#7	85	#45	37		
#8	82	#46	36		
#9	79	#47	35		
#10	77	#48	34		
#11	75	#49	33		
#12	73	#50	32		
#13	71	#51	32		
#14	69	#52	31		
#15	67	#53	31		
#16	66	#54	30		
#17	65	#55	30		
#18	64	#56	29		
#19	63	#57	29		
#20	62	#58	28		
#21	61	#59	28		
#22	60	#60	27		
#23	59	#61	27		
#24	58	#62	26		
#25	57	#63	26		
#26	56	#64	25		
#27	55	#65	25		
#28	54	#66	24		
#29	53	#67	24		
#30	52	#68	23		
#31	51	#69	23		
#32	50	#70	22		
#33	49	#71	22		
#34	48	#72	21		
#35	47	#73	21		
#36	46	#74	20		
#37	45	#75	20		
#38	44				

Note: The scale for below #75 is not in use until February 22.

Points deducted for losses:

Loss to team ranked #1-5	0.1 match played
Loss to team ranked #6-10	0.2 match played
Loss to team ranked #11-15	0.3 match played
Loss to team ranked #16-20	0.4 match played
Loss to team ranked #21-25	0.5 match played
Loss to team ranked #26-30	0.6 match played
Loss to team ranked #31-40	0.7 match played
Loss to team ranked #41-50	0.8 match played
Loss to team ranked #51-64	0.9 match played
Loss to team ranked #65-75	1 match played
Loss to team ranked #76-100	1.1 match played
Loss to team ranked #101-125	1.2 match played
Loss to unranked team	1.3 match played

Countable matches played (report dates):

Feb. 22	4 matches
March 1-8-15	5 matches
March 22	6 matches
March 29	7 matches
April 5	8 matches
April 12-19	9 matches
April 26 *	9 matches
May 11 **	10 matches
May 20 Final*	10 matches

Note: 125 teams will be ranked starting February 22, However Nos. 76-125 will not be published.

* April 26 & May 20 ranking are run twice on the computer, second run is published.

** May 11 ranking is unpublished

Bonus points for road victories:
All road wins 10%
Note: Points not awarded for wins at neutral site

Not used in CSA application

Bonus points for NCAA Team Championships

Advance to second Round	1 point added to overall average
Round of 16	2 points
Quarterfinals	3 points
Semifinals	4 points
Final	6 points
Champion	Automatically ranked No. 1

Table 1

- A similar chart exists for singles ratings.

The ITA college tennis ranking dates are shown in Table 2[1].

ITA RANKING DATES FOR COLLEGE TENNIS 2014-15

Division I - 2014-15 Ranking Dates (Men & Women)

(tentative dates)

Report Date	Release Date	Type of Ranking	Method	Report Date	Release Date
Sunday, August 24, 2014	Monday, September 08, 2014	Singles/Doubles	ballot	8/24/2014	9/8/2014
Tuesday, October 21, 2014	unpublished	National Singles/Doubles	computer	10/21/2014	unpublished
Friday, November 14, 2014	Tuesday, January 06, 2015	National Singles/Doubles	computer	11/14/2014	1/6/2015
Friday, November 14, 2014	Tuesday, January 06, 2015	National Team	team ballot	11/14/2014	1/6/2015
Sunday, January 18, 2015	Tuesday, January 20, 2015	National Team	ballot	1/18/2015	1/20/2015
Monday, January 26, 2015	Wednesday, January 28, 2015	National Team	ballot	1/26/2015	1/28/2015
Sunday, February 01, 2015	Tuesday, February 03, 2015	National Team (men only)	ballot	2/1/2015	2/3/2015
Monday, February 09, 2015	Tuesday, February 10, 2015	Team (women only)/Singles/Doubles	team ballot/ Singles&Doubles computer	2/9/2015	2/10/2015
Monday, February 16, 2015	Tuesday, February 17, 2015	National Team	ballot	2/16/2015	2/17/2015
Sunday, February 22, 2015	Tuesday, February 24, 2015	Team/Singles/Doubles(Computer Rankin	computer	2/22/2015	2/24/2015
Sunday, March 01, 2015	Tuesday, March 03, 2015	National Team	computer	3/1/2015	3/3/2015
Sunday, March 08, 2015	Tuesday, March 10, 2015	Team/Singles/Doubles	computer	3/8/2015	3/10/2015
Sunday, March 15, 2015	Tuesday, March 17, 2015	National Team	computer	3/15/2015	3/17/2015
Sunday, March 22, 2015	Tuesday, March 24, 2015	Team/Singles/Doubles	computer	3/22/2015	3/24/2015
Sunday, March 29, 2015	Tuesday, March 31, 2015	National Team	computer	3/29/2015	3/31/2015
Sunday, April 05, 2015	Tuesday, April 07, 2015	Team/Singles/Doubles	computer	4/5/2015	4/7/2015
Sunday, April 12, 2015	Tuesday, April 14, 2015	Team/Singles/Doubles	computer	4/12/2015	4/14/2015
Sunday, April 19, 2015	Tuesday, April 21, 2015	Team/Singles/Doubles	computer	4/19/2015	4/21/2015
4/26/2015*	Friday, May 01, 2015	Team/Singles/Doubles	computer (double-run for NCAA selections)	4/26/2015*	5/1/2015
Monday, May 11, 2015	unpublished	Team/Singles/Doubles (unpublished)	computer: results after 1st & 2nd rd. NCAA team championships	5/11/2015	unpublished
Wednesday, May 20, 2015	Tuesday, May 26, 2015	Team (Final)	computer	5/20/2015	5/26/2015
Wednesday, May 27, 2015	Wednesday, June 03, 2015	Singles/Doubles (Final)	computer	5/27/2015	6/3/2015
Wednesday, May 27, 2015	Wednesday, June 03, 2015	Regional (Final) Team/Singles/Doubles	ballot	5/27/2015	6/3/2015

Notes:
Report Date: Date that all match results must be entered into the Results Reporting system by each program to ensure that these results are included for weekly rankings purposes. Deadline for entry is Midnight ET; except for April 26, where deadline for entry is 10 pm ET.

For Report Date: April 26, 2015 - all match results must be entered by 10 pm ET as mandated by the NCAA to be considered for NCAA Championship selections.

Table 2

The first few team rankings (14 Nov - 16 Feb) are done by committee ballot!

- For computer rankings, ITA formula (Eq 1) requires an initial ranking list.
- All teams must have "enough" match data for computed ranking list to be stable (will see this when display evolution of CSA ranking results using ITA algorithm!). Eg., Princeton Men's tennis team has played 6 matches before first released computer ranking.

A sample page showing a typical summary sheet of data provided to each school after rankings have been updated is shown in Table 3 below.

Sample Sheet Summarizing Ranking Data

Type	League	Number of wins to count	Based on Past Rankings	Use point rules from ranking period	Ranking Date
Teams	Division I Men	6	3/15/2015 March 17 National Team Rankings [5 wins, 1000 run]	1/3/2000 12:00:00 AM Team5	03/22/2015
AVG.	University of Oklahoma	Total Win Points	Number of top wins counted	Total Loss Points	
85.77419		531.8	6	0.2	
AVG. Formula		Win Points Formula	Loss Points Formula		
85.77419355		531.8	0.2		

Rank	ALL LOSSES	Points	Match Location	Match Date
7	Texas A&M University	0.2	Away	2/20/2015
	ALL WINS			
3	Baylor University	98	Neutral	3/22/2015
3	Baylor University	98	Neutral	2/15/2015
4	University of Southern California	94	Neutral	2/16/2015
11	**Ohio State University**	**82.5**	**Away**	**3/6/2015**
12	**Wake Forest University**	**80.3**	**Away**	**1/30/2015**
9	University of Virginia	79	Home	3/10/2015
10	North Carolina	77	Home	2/22/2015
10	North Carolina	77	Neutral	2/14/2015
13	University of Mississippi	71	Neutral	2/13/2015
14	Northwestern University	69	Home	1/25/2015
21	University of South Florida	61	Home	2/6/2015
29	**Columbia University**	**58.3**	**Away**	**3/21/2015**
27	Florida State University	55	Home	2/28/2015
44	Wichita State University	38	Home	1/20/2015
48	University of New Mexico	34	Home	1/24/2015
60	University of Alabama	27	Home	2/8/2015
152	Purdue University	4	Neutral	3/20/2015

Table 3

- Results which “count” and rationale for present rank are clear for coaches and players to see!

We now turn to the application of the ITA method to the CSA 2014-15 squash season.

ITA Ranking Method applied to CSA 2014-15 Season (Men)

Table 4 summarizes the ranking dates we used for both ITA and ELO rankings. The first ranking date is not until the end of November to give Ivy schools time to complete a couple of matches. Subsequently, rankings were performed approximately every week. A total of 447 matches were played by 59 teams. Match results were extracted from the CSA website[5].

Ranking Dates for ITA and ELO Calculations

RANKING DATE	TOTAL NUMBER OF MATCHES PLAYED FROM START OF SEASON TO RANKING DATE
nov 30 2014	142 (142)
dec 08 2014	164 (22)
jan 10 2015	186 (22)
jan 18 2015	248 (62)
jan 25 2015	315 (67)
feb 04 2015	393 (78)
feb 08 2015	433 (40)
feb15 2015	447 (14)

Table 4

Figs ITA-1a/b, ITA-2a/b and ITA-3a/b show ranking results using the ITA formula Eq 1. An additional multiplicative scale factor, SCALEL, has been included in the denominator (compared with Eq.1) for ease of investigating the effect of increasing/decreasing the penalty for losses. The nominal value used for ITA tennis ranking is SCALEL = 1.0. They also use NBEST = 10 (ramped from 4 during the season), and NBEST = “all”.

$$R(i) = \frac{\sum_{j=1}^{NBEST} \mathrm{Winpoints(j)}}{NBEST + SCALEL \sum_{j=1}^{NWORST} \mathrm{Losspoints(j)}} \qquad (1')$$

ITA Ranking Method applied to CSA 2014-15 Season (Men)

Fig. ITA-1a shows the dependence of final ITA rank on parameters NBEST, NWORST and SCALEL for the top 30 teams according to the pre-season poll. The dependence on these parameters for teams ranked 31 and lower are shown in Fig. ITA-1b.

DEPENDENCE OF ITA RANKINGS ON MODEL PARAMETERS AND COMPARISON WITH CSA

COLUMNS IN RED SHOW DIFFERENCES BETWEEN ITA AND CSA PRE-TOURNAMENT RANKING PREDICTIONS. YELLOW SHADING IF DIFFERENCE > 3 SPOTS

NBEST	99	5	99	5	5
NWORST	99	99	5	5	5
SCALEL	1.0	1.0	1.0	1.0	0.5

(A)	(B) PRE-SEASON RANK	(C) CSA PRE-TOURN RANK	(D) ITA PRE-TOURNAMENT RANK	(E) D - C	(F)	(G) F - C	(H)	(I) H- C	(J)	(K) J- C	(L)	(M) L - C
Harvard	1	3	2	-1	4	1	2	-1	4	1	4	1
St. Lawrence	2	2	6	4	2	0	5	3	2	0	2	0
Trinity	3	1	1	0	1	0	1	0	1	0	1	0
Yale	4	5	4	-1	6	1	4	-1	6	1	6	1
Rochester	5	6	5	-1	5	-1	6	0	5	-1	5	-1
Columbia	6	4	3	-1	3	-1	3	-1	3	-1	3	-1
F&M	7	7	9	2	7	0	7	0	7	0	7	0
Cornell	8	10	11	1	11	1	11	1	11	1	11	1
U of Penn	9	8	7	-1	8	0	8	0	9	1	9	1
Princeton	10	9	10	1	9	0	9	0	8	-1	8	-1
Dartmouth	11	11	8	-3	10	-1	10	-1	10	-1	10	-1
Drexel	12	12	13	1	14	2	13	1	13	1	12	0
Bates	13	16	16	0	18	2	15	-1	19	3	15	-1
Williams	14	13	12	-1	15	2	12	-1	14	1	14	1
Western Ontario	15	19	23	4	23	4	21	2	22	3	22	3
Naval Academy	16	14	15	1	13	-1	17	3	12	-2	13	-1
GWU	17	15	18	3	17	2	18	3	18	3	18	3
Brown	18	18	17	-1	19	1	16	-2	15	-3	17	-1
Middlebury	19	17	14	-3	16	-1	14	-3	17	0	16	-1
Wesleyan	20	20	21	1	21	1	19	-1	20	0	19	-1
Amherst	21	24	28	4	29	5	23	-1	25	1	23	-1
Bowdoin	22	23	35	12	41	18	26	3	26	3	25	2
Colby	23	21	20	-1	20	-1	20	-1	21	0	20	-1
Hamilton	24	25	42	17	43	18	29	4	36	11	30	5
Conn College	25	26	32	6	37	11	24	-2	23	-3	26	0
Hobart	26	22	24	2	24	2	25	3	24	2	24	2
Bucknell	27	31	37	6	36	5	38	7	37	6	36	5
Stanford	28	27	40	13	38	11	35	8	34	7	35	8
Georgetown	29	33	22	-11	22	-11	27	-6	27	-6	28	-5
Johns Hopkins	30	34	34	0	35	1	39	5	39	5	38	4

Fig. ITA-1a

ITA Ranking Method applied to CSA 2014-15 Season (Men)

- Column A lists team names in order of CSA pre-season rank, numerated in column B.
- Pre-tournament rank determined by the CSA is shown in column C.
- Columns D and every second column thereafter lists the ITA prediction for pre-tournament rank using NBEST, NWORST and SCALEL values indicated by the green arrows.
- NBEST = 99 indicates that *all* wins were taken into account in the summation over Winpoints in Eqs 1′. Similarly, NWORST = 99 indicates that *all* losses were taken into account in the summation over Losspoints.
- Column E and every second column thereafter show differences in rank between ITA and CSA rank.
- As a visual aid in detecting anomalous results, yellow shading indicates where differences between ITA and CSA rankings differ by greater than 3 spots.

MIT	31	29	26	-3	28	-1	28	-1	28	-1	27	-2
Northeastern	32	38	30	-8	30	-8	33	-5	33	-5	33	-5
Tufts	33	28	36	8	40	12	30	2	29	1	29	1
Denison	34	35	53	18	53	18	53	18	53	18	53	18
Boston	35	42	44	2	44	2	45	3	45	3	46	4
Haverford	36	37	27	-10	26	-11	31	-6	31	-6	31	-6
Colgate	37	43	51	8	51	8	51	8	52	9	50	7
Lehigh	38	40	41	1	39	-1	42	2	42	2	41	1
Boston U	39	45	48	3	48	3	43	-2	43	-2	42	-3
Washington(St Louis)	40	46	39	-7	34	-12	40	-6	40	-6	43	-3
Chicago	41	32	19	-13	12	-20	22	-10	16	-16	21	-11
NYU	42	36	25	-11	27	-9	32	-4	30	-6	32	-4
Charleston	43	50	46	-4	45	-5	47	-3	47	-3	47	-3
Penn State	44	??	47		47		48		48		48	
Northwestern	45	41	29	-12	25	-16	34	-7	32	-9	34	-7
Richmond	46	47	38	-9	33	-14	41	-6	41	-6	40	-7
Bryant	47	44	45	1	46	2	46	2	46	2	45	1
Swarthmore	48	52	52	0	52	0	52	0	51	-1	52	0
Cal Berkeley	49	54	49	-5	49	-5	49	-5	49	-5	49	-5
Fordham	50	39	31	-8	31	-8	37	-2	38	-1	37	-2
Ithaca	51	61	56	-5	56	-5	56	-5	56	-5	56	-5
USC	52	??	57		57		57		57		57	
Washington, U of	53	51	33	-18	32	-19	36	-15	35	-16	39	-12
Vassar	54	57	58	1	58	1	58	1	58	1	58	1
Siena	55	56	55	-1	55	-1	55	-1	55	-1	55	-1
Notre Dame	56	60	54	-6	54	-6	54	-6	54	-6	54	-6
Vanderbilt	57	59	50	-9	50	-9	50	-9	50	-9	51	-8
Minnesota	58	53	43	-10	42	-11	44	-9	44	-9	44	-9
Sewanee	59	??	59		59		59		59			

Fig. ITA-1b

ITA Ranking Method applied to CSA 2014-15 Season (Men)

Observations:

- There is a lot of yellow! - significantly more than when we apply ELO - see later.

- The amount of yellow increases as we go down in ranking (compare, especially, top 30 according to pre-season poll - Fig. ITA-1a - with lower 30 - Fig. ITA-1b. Since agreement between ELO and CSA ranking is decent even for lower-ranked teams, we cannot blame the discrepancy between predicted ITA and CSA rank as due to a lack of validity of CSA rankings!

- Focusing on Fig. ITA-1a, the ITA method does somewhat better when we restrict the number of wins and losses using NBEST = 5 and NWORST = 5. The motivation for decreasing NBEST from the value 10 used in ITA tennis ranking is that college tennis teams play many more matches during the tennis season ($\gtrsim$ 30 typically, 26 for Princeton this year) than squash teams play during the CSA season. Assuming the ITA point allotments and number of countable matches are tuned to a typical ITA team schedule, scaling NBEST from 10 to 5 makes sense for the CSA based on the roughly 2:1 ratio between number of tennis and squash team matches. Preserving the *relative* importance between losses and wins in Eq 1′ demands that we also scale NWORST.

- Another strong argument exists for limiting NBEST and NWORST: Eq 1 can be re-written in terms of averages in the *exactly equivalent* form

$$R_i = \frac{\langle \text{Winpoints} \rangle}{1 + \frac{NWORST}{NBEST}\langle \text{Losspoints} \rangle}, \quad \text{where} \qquad (1'')$$

$$\langle \text{Winpoints} \rangle = \frac{1}{NBEST}\sum_{j=1}^{NBEST} \text{Winpoints}(j)$$

and

$$\langle \text{Losspoints} \rangle = \frac{1}{NWORST}\sum_{j=1}^{NWORST} Loss\text{points(j)}$$

The numerator acts as a credit for matches won, the denominator acts as a penalty for matches lost. Column 4 of Figs. ITA-2a/b shows the number of matches played by each team during the 2014-15 CSA season. There is a wide disparity in this number. A team such as Trinity fulfills its conference commitments *and* plays a healthy schedule of matches against potentially stronger opposition. Trinity, therefore, plays significantly more matches (18) than competitively strong teams such as St. Lawrence (who play 13), or Harvard (11). From the ITA points table we see that the number of points awarded for winning decreases as the rank of opposition decreases. A necessary consequence and flaw in the ITA system if NBEST is not appropriately limited, is that the *more* wins a team achieves, the *smaller* becomes the average Winpoints (numerator in Eq 1′′), and the *smaller* becomes the accumulated rating points upon which rank is determined.

ITA Ranking Method applied to CSA 2014-15 Season (Men)

FINAL ITA RANKINGS (using NBEST = 5, NWORST = 5, SCALEL = 0.5)

WITH WIN/LOSS DATA

Wins (+) are wins against opponents who finished higher in rank
Losses (-) are losses against opponents who finished lower in rank

ITA RANK		Ranking Points	Matches played	Won	Lost	Wins(+)	Losses(-)
1	Trinity	92.48	18	17	1	0	1
2	St. Lawrence	87.92	13	12	1	0	0
3	Columbia	83.33	14	12	2	0	1
4	Harvard	83.11	11	8	3	1	1
5	Rochester	82.64	13	9	4	2	2
6	Yale	79.62	14	10	4	1	0
7	F&M	71.59	18	12	6	1	1
8	Princeton	68.81	13	5	8	1	2
9	U of Penn	67.04	14	8	6	0	0
10	Dartmouth	65.93	13	6	7	0	0
11	Cornell	64.40	15	7	8	1	0
12	Drexel	58.57	15	7	8	0	1
13	Naval Academy	58.42	23	16	7	1	1
14	Williams	57.52	21	12	9	0	0
15	Bates	55.79	20	12	8	1	2
16	Middlebury	55.75	18	12	6	2	1
17	Brown	54.64	13	5	8	0	0
18	GWU	53.04	15	7	8	2	0
19	Wesleyan	50.86	18	11	7	0	1
20	Colby	49.91	18	11	7	0	0
21	Chicago	47.25	9	9	0	0	0
22	Western Ontario	46.02	10	4	6	0	0
23	Amherst	45.15	16	6	10	1	1
24	Hobart	44.00	17	9	8	0	0
25	Bowdoin	42.55	17	4	13	1	0
26	Conn College	42.27	21	7	14	0	3
27	MIT	37.78	17	9	8	0	1
28	Georgetown	36.53	11	7	4	0	0
29	Tufts	34.69	16	4	12	1	2
30	Hamilton	34.60	17	7	10	5	0

Fig. ITA-2a

ITA Ranking Method applied to CSA 2014-15 Season (Men)

31	Haverford	34.09	16	9	7	0	3
32	NYU	32.77	7	4	3	1	0
33	Northeastern	31.96	11	7	4	0	0
34	Northwestern	31.02	10	7	3	1	0
35	Stanford	29.59	13	5	8	0	0
36	Bucknell	29.04	11	4	7	1	0
37	Fordham	28.26	14	9	5	0	1
38	Johns Hopkins	28.25	8	3	5	1	0
39	Washington, U of	27.83	7	5	2	0	1
40	Richmond	26.67	7	3	4	1	1
41	Lehigh	26.60	9	5	4	0	0
42	Boston U	25.56	13	3	10	1	2
43	Washington(St Louis)	25.28	7	3	4	0	0
44	Minnesota	23.24	8	4	4	0	0
45	Bryant	23.17	12	4	8	1	2
46	Boston College	23.13	9	4	5	1	0
47	Charleston	20.00	5	2	3	0	0
48	Penn State	19.34	5	2	3	0	0
49	Cal Berkeley	19.03	5	2	3	0	0
50	Colgate	17.87	7	2	5	1	0
51	Vanderbilt	17.06	6	2	4	0	0
52	Swarthmore	16.86	9	2	7	0	0
53	Denison	15.83	8	1	7	1	0
54	Notre Dame	9.67	9	1	8	0	0
55	Siena	9.51	11	1	10	0	0
56	Ithaca	0.00	6	0	6	0	0
57	USC	0.00	8	0	8	0	0
58	Vassar	0.00	8	0	8	0	0
59	Sewanee	0.00	1	0	1	0	0

Fig. ITA-2b

Figs. ITA-3a/b show the evolution of rankings through the season.

ITA Ranking Method applied to CSA 2014-15 Season (Men)

ITA RANK HISTORY (8 RANKING DATES USED)

	PRE-SEASON RANK	Thru Nov	Thru Dec	10-Jan	18-Jan	25-Jan	4-Feb	8-Feb	15-Feb
Harvard	1	46	43	6	5	8	5	3	4
St. Lawrence	2	8	9	8	3	2	2	2	2
Trinity	3	4	4	1	1	1	1	1	1
Yale	4	47	1	5	4	4	4	6	6
Rochester	5	18	12	18	14	5	6	5	5
Columbia	6	1	3	3	2	3	3	4	3
F&M	7	11	14	9	13	6	7	7	7
Cornell	8	3	6	7	10	7	11	10	11
U of Penn	9	5	2	2	6	9	10	9	9
Princeton	10	2	19	25	18	15	8	8	8
Dartmouth	11	48	11	17	25	14	9	11	10
Drexel	12	49	50	50	20	20	13	13	12
Bates	13	10	10	4	12	17	17	16	15
Williams	14	7	18	16	7	10	12	14	14
Western Ontario	15	50	51	51	27	24	25	25	22
Naval Academy	16	9	5	15	8	11	14	12	13
GWU	17	13	20	14	9	16	16	18	18
Brown	18	6	7	13	15	13	18	17	17
Middlebury	19	12	13	10	16	12	15	15	16
Wesleyan	20	30	30	36	21	18	21	19	19
Amherst	21	14	8	11	19	21	22	24	23
Bowdoin	22	16	16	26	43	45	32	23	25
Colby	23	15	15	19	11	19	20	20	20
Hamilton	24	21	17	21	22	26	29	33	30
Conn College	25	19	22	22	33	32	26	26	26
Hobart	26	26	23	32	31	23	24	22	24
Bucknell	27	31	41	44	46	37	36	36	36
Stanford	28	20	32	39	45	44	37	37	35
Georgetown	29	17	21	20	24	22	23	27	28
Johns Hopkins	30	32	28	24	26	28	34	35	38

Fig. ITA-3a

- Although the rankings have settled down / converged to sensible values by early Feb, even as late as Jan 18 there are some ITA calculated ranks that are problematic and would cause consternation if published. (Will also be true of ELO, later!).

ITA Ranking Method applied to CSA 2014-15 Season (Men)

MIT	31	29	27	28	23	30	27	28	27
Northeastern	32	24	24	23	28	25	31	32	33
Tufts	33	34	42	40	34	33	30	29	29
Denison	34	51	52	52	53	53	52	52	53
Boston College	35	35	36	38	40	39	41	45	46
Haverford	36	33	26	29	29	35	28	31	31
Colgate	37	23	29	31	37	46	47	53	50
Lehigh	38	28	25	30	32	36	43	40	41
Boston U	39	41	39	42	42	43	39	42	42
Washington(St Louis)	40	42	45	46	49	42	44	43	43
Chicago	41	22	31	33	35	27	19	21	21
NYU	42	27	35	27	30	31	38	30	32
Charleston	43	52	53	53	17	48	48	47	47
Penn State	44	38	44	45	47	47	49	48	48
Northwestern	45	25	33	34	36	38	35	34	34
Richmond	46	40	38	41	41	29	40	41	40
Bryant	47	37	40	43	44	41	45	44	45
Swarthmore	48	43	46	47	50	50	50	51	52
Cal Berkeley	49	53	54	54	54	54	51	49	49
Fordham	50	39	37	37	39	34	33	38	37
Ithaca	51	54	55	55	55	55	56	56	56
USC	52	55	56	56	56	56	57	57	57
Washington, U of	53	36	34	35	38	40	42	39	39
Vassar	54	56	57	57	57	57	58	58	58
Siena	55	45	49	49	52	52	54	54	55
Notre Dame	56	57	58	58	58	58	55	55	54
Vanderbilt	57	58	48	12	48	49	53	50	51
Minnesota	58	44	47	48	51	51	46	46	44
Sewanee	59	59	59	59	59	59	59	59	59

Fig. ITA-3b

Since the ITA ranking method is a sequential algorithm where updated ranking points are determined by the most recent ranking positions (through the points assignment chart) there can be a strong dependence of final rank on match schedule. To illustrate this, Fig. ITA-4 shows the effect on final pre-tournament rank of assuming that matches which actually took place between Feb 08 and Feb 15 had, instead, taken place between Feb 04 and Feb 08, and vice-versa. We see a troubling number of changes to final rank position - troubling because of potential impact on tournament division selection.

ITA Ranking Method applied to CSA 2014-15 Season

INSTABILITY OF ITA RANKINGS w.r.t. MATCH SCHEDULING

ORIGINAL (DATES IN CORRECT ORDER)			MODIFIED (MATCHES DURING LAST 2 RANKING DATES ARE INTERCHANGED)		
1	Trinity	92.48	1	Trinity	91.88
2	St. Lawrence	87.92	2	St. Lawrence	87.13
3	Columbia	83.33	3	Columbia	83.92
4	Harvard	83.11	4	Harvard	83.88
5	Rochester	82.64	5	Rochester	81.70
6	Yale	79.62	6	Yale	80.19
7	F&M	71.59	7	F&M	71.78
8	Princeton	68.81	8	Princeton	67.82
9	U of Penn	67.04	9	U of Penn	66.67
10	Dartmouth	65.93	10	Dartmouth	64.82
11	Cornell	64.40	11	Cornell	64.26
12	Drexel	58.57	12	Naval Academy	58.94
13	Naval Academy	58.42	13	Drexel	58.57
14	Williams	57.52	14	Williams	58.05
15	Bates	55.79	15	Bates	56.81
16	Middlebury	55.75	16	Middlebury	54.47
17	Brown	54.64	17	GWU	53.62
18	GWU	53.04	18	Brown	53.57
19	Wesleyan	50.86	19	Colby	49.22
20	Colby	49.91	20	Chicago	49.00
21	Chicago	47.25	21	Wesleyan	48.74
22	Western Ontario	46.02	22	Western Ontario	44.68
23	Amherst	45.15	23	Amherst	43.88
24	Hobart	44.00	24	Bowdoin	43.69
25	Bowdoin	42.55	25	Hobart	43.47
26	Conn College	42.27	26	Conn College	40.83
27	MIT	37.78	27	Georgetown	38.38
28	Georgetown	36.53	28	MIT	37.03
29	Tufts	34.69	29	Tufts	35.67
30	Hamilton	34.60	30	NYU	33.62

YELLOW SHADING WHERE RANKING POSITION CHANGED

Fig. ITA-4

We now turn our discussion to ELO ranking and a version of ELO that avoids any dependence on match schedule, as well as having other attractive features

ELO BASICS

Sequential ELO[6]:

At the start of each season each team is assigned the same number of ranking points (the actual number has no effect on final rank, and here we assume the number is 1000). After each match, say between teams labeled “i” and “j”, ranking points for team i are updated according to

$$R_i' = R_i + K\left[S_i - E_i\right] \qquad (2)$$

where

$$E_i = \frac{1}{1+e^{-\Delta R/R_s}}, \quad \text{where } \Delta R = R_i - R_j \qquad (3)$$

Here

- R'_i is the new ranking points for team i
- R_i is the old.
- $K\,(S_i - E_i)$ is the points adjustment.
- S_i is a numerical expression of the match result from i’s perspective: 1 = win, 0 = loss (and 0.5 for a tie).
- E_i is the expectation that team i beat team j given their ranking point differential ΔR prior to the match. Later we will explain where this expression comes from.
- R_s is a scale factor tuned to set a reasonable probability that a team can pull an upset and beat a team with a chosen point differential.
- K is an exchange factor that governs the magnitude of rating changes (how rapidly the rating points can adjust from one ranking period to the next).

Typical parameter values used in Chess Federation rankings are $K = 32$ and $R_s = 175$.

The average number of rating points among all teams is conserved throughout the season in ELO rankings!

Familiarizing Example 1: Imagine teams i and j play each other and i beats j. Coming into the match assume both i and j are tied with the same number of ranking points ($R_i = R_j$). We would expect that each team is equally likely to win. Sure enough, E_i evaluates to 0.5 when $\Delta R = 0$. Since i won the match, $S_i = 1$, therefore team i’s points are adjusted to $R'_i = R_i + 0.5{*}K$ (a change proportional to the exchange factor K). From team j’s perspective, $S_j = 0$ and $E_j = 0.5$, therefore $R'_j = R_j - 0.5{*}K$. In the updated rankings team i will appear above team j because i has gained points through the win. Team j however has lost points to slip below i. This penalty for losing may even cause j to slip behind other nearby teams.

Familiarizing Example 2: Team i plays team j and i beats j. However, coming into the match assume there is some point differential $\Delta R = R_i - R_j \neq 0$ between the teams. In Eq 2 we must evaluate E_i, the expectation that i beats j given this ΔR. So we had better understand this function, and the scale factor R_s that appears in it.

Figure ELO-1 shows a plot of E_i, Eq 3, as a function of points differential for three assumed values of the scale factor R_s. We see that R_s controls how rapidly the expectation curve rises as a function of ΔR. If, instead of $R_s = 1.0$, 2.0 and 3.0 we choose 10 or 100 times these values, the plot *shape* does not change; we simply multiply the scale on the horizontal axis by 10 or 100. This shows that the appropriate choice of the scale factor is simply cosmetic. It controls the scale of the ranking points distribution.

$$E_i = \frac{1}{1+e^{-\Delta R/R_s}}$$

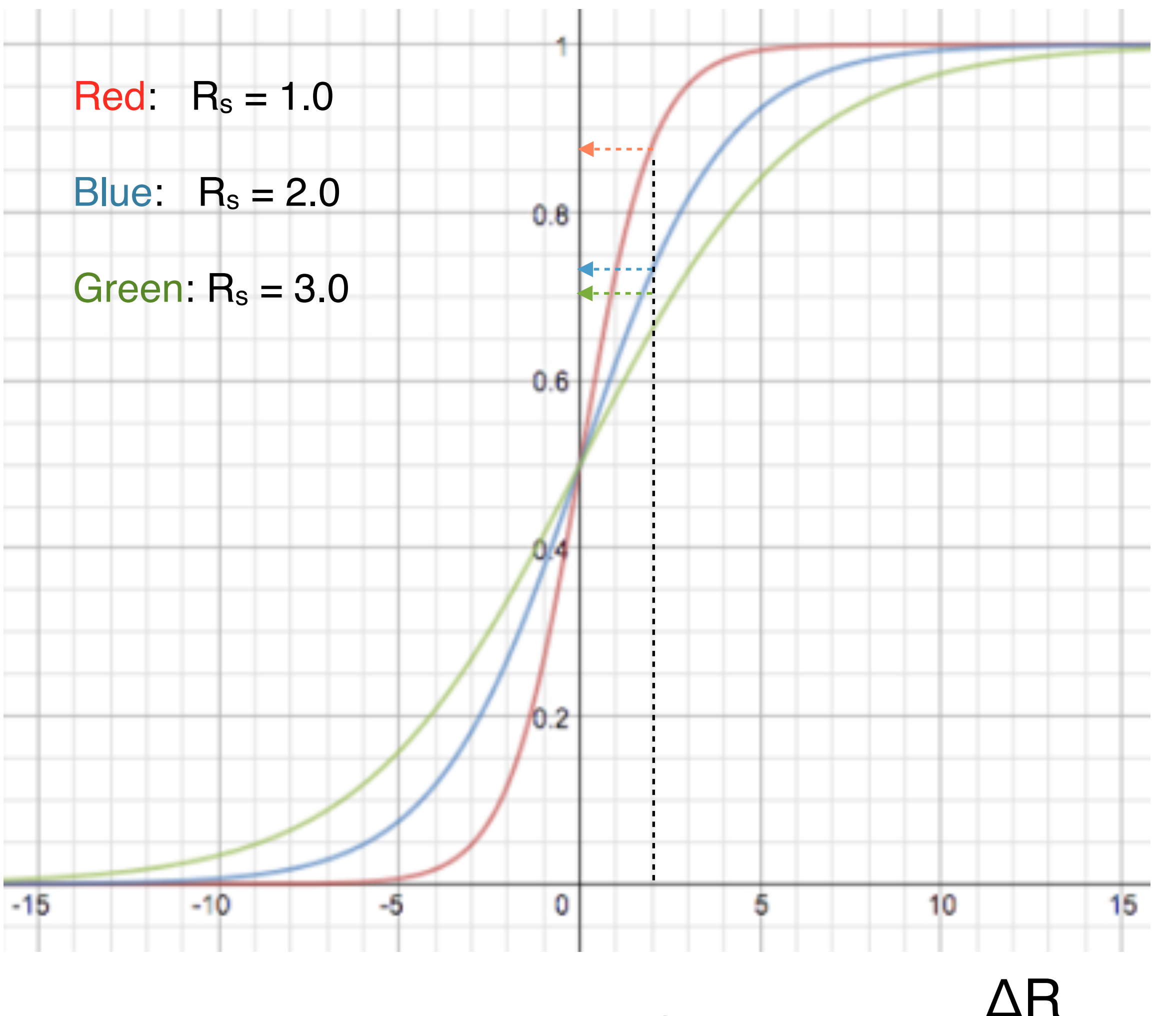

Fig. ELO-1

ELO BASICS

For a point differential of ΔR = 2.0, Fig. ELO-1 shows:

E_i = 0.66 when R_s =3.0;

E_i = 0.73 when R_s =2.0; There is a 73% chance that team i will beat team j if the point differential between i and j is 2.0 and, equivalently, a 27% chance that j will upset the point spread and beat i. For Chess Federation rankings with R_s = 175, these are the likely percentages for winning/losing matches with a point differential of 350.

E_i = 0.88 when R_s =1.0.

Before showing results, we dig deeper into ELO to understand what "expectation of winning" means and how probability arguments make sense in a ranking system.

Theoretical Underpinnings of ELO:

Most would agree that the outcome of a match between two teams (or competitors) depends on the (current) abilities of the two teams. The ElO method assumes a probability for competitor "i" beating competitor "j" as a ratio that can be written schematically as

$$P(\text{i beats j}) \sim \text{strength(i)} / [\, \text{strength(i)} + \text{strength(j)} \,]. \qquad (4)$$

But what does "strength" mean here?

The ELO rating system assigns to every team a numerical rating based on performance in matches. The rating is a number in some range (explained later) that changes over time depending strictly on the outcome of matches. When two teams compete, the rating system predicts that the team with the higher rating should win more often than the team with the lower rating. The larger the difference in ratings, the greater the likelihood that the higher rated team will win. **Once the ratings are calculated, they can be sorted in order of decreasing value to determine team rank**

There are many factors that determine how well players on a given team will perform on a given day (niggling injuries, fatigue from a recent challenge match, how well match preparation went, nerves, …). We can expect that the distribution of performance strength takes the shape of a curve such as shown in Fig. ELO-2 below. **ELO calculates the average rating of each team - the location of the peaks**. **This is quantity R_i in Eq 2.** (We can conjecture that the width of the "strength distribution" will be narrower for elite teams, where players have considerable competition experience, than it will be for lower ranked teams. But this is not an assumption we use!).

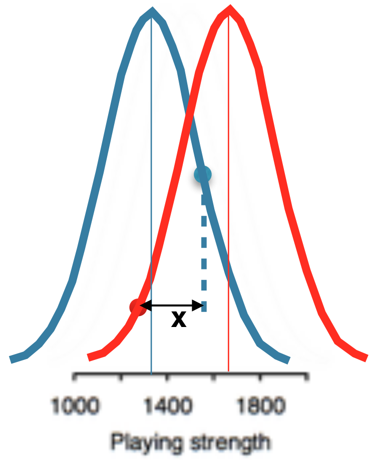

Fig ELO-2

Imagine two teams, named Blue and Red, that are scheduled to play each other. Assume Team Blue is ranked behind Team Red which means that the average rating of Blue is less than the average rating of Red. (The blue peak is to the left of the red peak in Fig. ELO-2). Each team has the *potential* to perform at a level corresponding to any point along its performance strength distribution curve. To simulate a match between Blue and Red, we ask a computer to select a pair of points *at random*, one from each strength distribution. The blue and red dots in Fig. ELO-2 illustrate one such simulated match. Although Red is ranked ahead of Blue, the simulation has chosen a scenario where Team Red significantly under-performs relative to its mean, and Team Blue over-performs relative to its mean. In fact, the combined relative performances have resulted in a simulated playing strength for Team Red that is *less* than the simulated playing strength of Team Blue. In this computer match, Team Red would lose to Team Blue in spite of the fact that Team Red is actually ranked ahead of Team Blue. The Navy-Princeton or F&M-Rochester "upsets" are good examples of a realization of Fig. ELO-2.

Let variable **x** denote the difference between the sampled performance strengths of any two teams (**x** is shown in Fig. ELO-2). We sample the strength distributions many times (as if simulating many matches between Blue and Red), each time sampling the two distributions, always taking the difference in the same order (eg Red minus Blue), and building a frequency distribution of results. By appropriately normalizing the frequencies we build a "probability distribution function (pdf)" p(**x**) for the difference between team performance strengths. A powerful theorem of mathematics, called the Central Limit Theorem, *guarantees* that if we sample enough times, and plot the distribution of the sample means of **x**, the resulting distribution is a bell curve - a Gaussian distribution with **x** in the range $-\infty < \mathbf{x} < \infty$. In fact, this result does not depend on the actual form of the strength distributions that were sampled and is an important reason that statistics, applied correctly, can be successfully applied in many real life situations!

The probability distribution function p(**x**) tells us how common is the occurrence that sampled differences between playing strengths of two teams takes on the value **x**. From p(**x**) we can infer the expected result of a match between two teams that differ in ranking strength by a particular value, ΔR. We need P(**x** < ΔR), the probability that a sampled **x** is smaller than the actual difference in average ranking strength of the two teams. Mathematically, we can write this as

$$P(\mathbf{x} < \Delta R \equiv R_i - R_j) = \int_{-\infty}^{\Delta R} p(\mathbf{x})\,dx \equiv \int_{-\infty}^{\Delta R} \frac{dE}{dx}dx = E(\Delta R) \qquad (5)$$

where E is known as the “cumulative distribution function (cdf)” and is defined such that its derivative is the probability distribution function p(**x**).

Rather than working with a Gaussian distribution, ELO ratings work with a very similar distribution called a “logistic distribution[7]” which has the advantage of having an E which can be written in terms of simpler functions than would appear if one worked with a Gaussian. Specifically, the logistic cdf takes the form

$$E(x) = \frac{1}{\left[1 + e^{-x/x_s}\right]}$$

which was plotted earlier in this document. The scale parameter x_s controls the slope of the E(**x**) at **x** = 0

Using the logistic function, Eq (5) becomes

$$P(\mathbf{x} < \Delta R) = E(\Delta R) = \frac{1}{[1 + e^{-\Delta R/R_s}]} = \frac{e^{R_i/R_s}}{[e^{R_i/R_s} + e^{R_j/R_s}]} \qquad (6)$$

expressing the probability that Team i will beat Team j when their ratings differ by an amount ΔR. Comparing Eq 6 with Eq 4 we see a similarity to the intuitive ratio form for the probability of winning.

The ELO rating formula Eq 2 is seen to award / penalize rating points by an amount proportional to the difference between how the two teams were predicted to perform in their match and how they actually performed.

SELF-CONSISTENT ELO

The shape of the curve $E(\Delta R)$ shown in Fig. ELO-1 determines how much credit / penalty a team gets for a win / loss. The credit / penalty is given by the “points adjustment” factor $K\ (S_i - E_i)$ in Eq 2. For a WIN ($S_i = 1$) against a team with $\Delta R > 0$ (i.e., team i was favored to win over team j), the amount of credit for the win DECREASES with increasing points spread ΔR, and INCREASES with increasing points spread if $\Delta R < 0$ (in which case team i has scored an upset). Conversely, for a LOSS ($S_i = 0$) against a team with $\Delta R > 0$, the penalty for losing DECREASES with increasing ΔR, but INCREASES with increasing points spread if $\Delta R < 0$. In short: good wins are highly credited; bad losses are greatly penalized. I.e., strength of schedule is taken into account by ELO!

The ELO system most appropriate to college squash is **Self-Consistent ELO**, rather than **Sequential ELO** described so far. **In the self-consistent approach, each time the rankings are evaluated we take into account *all* matches that have taken place through that ranking date, going all the way back to the start of season**. Sequential ELO would simply *update* the rankings based on what came out of the previous ranking calculation. Should rankings reflect most recent form, or the body of work (wins and losses) over the entire season? If the same team was played multiple times a strong argument could be made for rankings to reflect most recent form. However, that is not the case in college squash. Most teams play each other only once during the regular season. Moreover, scheduling constraints may force a given team to play a rival early in the season. Why should that not count as much as another team playing the same rival later in the season when coaches have limited control over schedule ?!

Self Consistent ELO[8]:

First, we generalize Eq 2 using notation borrowed from Richard Brent[9].

N teams play a number of matches throughout any interval within a season. Each match can end with a win, loss or draw, with a win scoring 1 point, a draw 0.5 points, and a loss 0 points. The results are stored in a score matrix $\mathbf{S}$ where $S_{i\,j}$ is the number of points that team i scores against team j. The diagonal elements $S_{i\,i}$ are arbitrary, but conveniently set to 0. The sum $S_{i\,j} + S_{j\,i}$ is the total number of games played between teams i and j. Each team has a points rating R_i, updated according to

$$R_i' = R_i + K\left[\sum_{j=1}^{N} S_{i\,j} - \sum_{j=1}^{N} S_{i\,j}\,E_{i\,j}\right]; \quad i = 1, \cdots, N$$

where (7)

$$E_{ij} = \frac{1}{1 + e^{-(R_i - R_j)/R_s}}$$

is the probability of i beating j given their ranking points differential. These equations are entirely equivalent to Eqs 2 and 3. The first term in the square bracket is the actual number of wins of team i against *all* opponents; the second term in the bracket is the expected number of wins.

At the start of the season, all teams are assigned 1000 ranking points. **There is no subjective assignment of pre-season rank - every team has the same rank!!** A number of ranking dates are chosen throughout the season - days when rankings will be evaluated (such as in Table 4). ***All*** match results from the start of season through each ranking date are entered into the score matrix **S**, and **Eq (8) is iterated until a set of ranking points $\{R_i\}$ is found such that the expected wins for each team matches the actual number of wins**:

$$\sum_{j=1}^{N} S_{ij} - \sum_{j=1}^{N} S_{ij} E_{ij} = 0 \quad \text{for all teams i.} \tag{8}$$

From Eq 7 we see that when this condition is satisfied then $R_i' = R_i$ for all teams, implying consistency between ranking points and match results!

There are no adjustable parameters in Self-Consistent ELO:

(a) The factor K does not appear in Eq 8.

(b) From Eqs 8 and 7 we see that the value of R_s has no impact on the ratings since it can be eliminated by a change of variable. It turns out that, depending on the method of iteration, R_s can impact the number if iterations it takes for the ratings to converge, and that convergence may only be achieved in a finite range of R_s values.

Results from Running Self-Consistent ELO rankings code on 2014-15 season

The Figs ELO-3 to ELO-6 summarize the results of applying the Self-Consistent ELO ranking method to the CSA 2014-15 season.

Examining the theory behind ELO ratings, whether Sequential or Self-Consistent, presents no clear argument that teams playing more countable matches than the average should have their ranking affected (unlike our finding with ITA rankings). However, we felt this should be tested, and results are shown in Figs. ELO-3a/b. Here, columns D and E (indicated by green arrows) show final pre-tournament rankings predicted by Self-Consistent ELO when all matches played by each team are counted (col D) and when this number is limited to 13 (col F). We see little difference, as predicted.

Columns E and G of Fig. ELO-3a show differences between the ELO predicted pre-tournament (Feb 15 2015) rankings and the rankings assumed by the CSA. As in Figs. ITA-1a/b, yellow shading is used to indicate significant differences between ELO predictions and CSA rank, where a “significant difference” is defined as greater than 3 positions.

SELF-CONSISTENT ELO APPLIED TO THE CSA 2014-15 SEASON - Men

DEPENDENCE OF ELO ITERATED RANKINGS ON NMAX AND COMPARISON WITH CSA

NMAX represents # best matches that count

COLUMNS IN RED SHOW DIFFERENCES BETWEEN ELO AND CSA PRE-TOURNAMENT RANKING PREDICTIONS. YELLOW SHADING IF DIFFERENCE > 3 SPOTS

			NMAX = 99		NMAX = 13	
(A)	(B)	(C)	(D)	(E)	(F)	(G)
	PRE-SEASON RANK	CSA PRE-TOURN RANK	ELO PRE-TOURNAMENT RANK			
				D - C		F - C
1	Harvard	3	4	1	4	1
2	St. Lawrence	2	2	0	2	0
3	Trinity	1	1	0	1	0
4	Yale	5	5	0	5	0
5	Rochester	6	6	0	6	0
6	Columbia	4	3	-1	3	-1
7	F&M	7	7	0	7	0
8	Cornell	10	11	1	11	1
9	U of Penn	8	8	0	8	0
10	Princeton	9	9	0	9	0
11	Dartmouth	11	10	-1	10	-1
12	Drexel	12	13	1	13	1
13	Bates	16	17	1	17	1
14	Williams	13	14	1	14	1
15	Western Ontario	19	19	0	19	0
16	Naval Academy	14	12	-2	12	-2
17	GWU	15	18	3	18	3
18	Brown	18	15	-3	15	-3
19	Middlebury	17	16	-1	16	-1
20	Wesleyan	20	20	0	20	0
21	Amherst	24	22	-2	23	-1
22	Bowdoin	23	25	2	25	2
23	Colby	21	21	0	21	0
24	Hamilton	25	26	1	27	2
25	Conn College	26	28	2	31	5
26	Hobart	22	23	1	24	2
27	Bucknell	31	32	1	32	1
28	Stanford	27	30	3	29	2
29	Georgetown	33	27	-6	28	-5
30	Johns Hopkins	34	33	-1	33	-1

Fig. ELO-3a

SELF-CONSISTENT ELO APPLIED TO THE CSA 2014-15 SEASON - Men

continuation:

31	MIT	29	31	2	30	1
32	Northeastern	38	37	-1	37	-1
33	Tufts	28	29	1	26	-2
34	Denison	35	38	3	39	4
35	Boston	42	42	0	42	0
36	Haverford	37	35	-2	35	-2
37	Colgate	43	43	0	43	0
38	Lehigh	40	40	0	40	0
39	Boston U	45	45	0	45	0
40	Washington(St Louis)	46	41	-5	41	-5
41	Chicago	32	24	-8	22	-10
42	NYU	36	34	-2	34	-2
43	Charleston	50	50	0	50	0
44	Penn State	??	47		47	
45	Northwestern	41	36	-5	36	-5
46	Richmond	47	48	1	48	1
47	Bryant	44	44	0	44	0
48	Swarthmore	52	52	0	52	0
49	Cal Berkeley	54	54	0	54	0
50	Fordham	39	39	0	38	-1
51	Ithaca	61	58	-3	58	-3
52	USC	??	59		59	
53	Washington, U of	51	49	-2	49	-2
54	Vassar	57	46	-11	46	-11
55	Siena	56	55	-1	55	-1
56	Notre Dame	60	57	-3	57	-3
57	Vanderbilt	59	53	-6	53	-6
58	Minnesota	53	51	-2	51	-2
59	Sewanee	??	56		56	

Fig. ELO-3b

For teams in the bottom 30, ELO predictions are *much* closer to CSA ranking than was found using the ITA method. With no restriction on countable matches (in future all of our ELO result discussions will apply to this unrestricted case), only Georgetown in the top 30 had an ELO ranking significantly different than CSA ranking. Interestingly, even if were to modify our definition of significant difference to “greater than 2 positions”, only Brown, GWU and Stanford would be additionally flagged and we are aware that a provisional pre-tournament CSA ranking list had Brown and GWU in positions that were more consistent with the ELO predictions but those provisional rankings were subsequently adjusted for to penalize Brown for lacking a sufficient “strength-of-schedule”.

SELF-CONSISTENT ELO APPLIED TO THE CSA 2014-15 SEASON - Men

Figs. ELO-4a/b show final ELO rank, including data on matches played, matches won and lost, and quantities we denote by Wins(+) and Losses(-). The first of these, Wins(+) is the number of wins a team has against opponents who finished higher in the ELO rank; Losses(-) is the number of losses a team has against opponents who finish lower in rank.

FINAL ELO RANKINGS (using NMAX = 99) WITH WIN/LOSS DATA

Wins (+) are wins against opponents who finished higher in rank

Losses (-) are losses against opponents who finished lower in rank

ELO RANK		Ranking Points	Matches played	Won	Lost	Wins (+)	Losses(-)
1	Trinity	1177.19	18	17	1	0	1
2	St. Lawrence	1175.64	13	12	1	0	0
3	Columbia	1167.44	14	12	2	0	1
4	Harvard	1166.34	11	8	3	1	1
5	Yale	1161.66	14	10	4	0	0
6	Rochester	1161.36	13	9	4	2	1
7	F&M	1150.65	18	12	6	1	1
8	U of Penn	1142.83	14	8	6	0	1
9	Princeton	1142.45	13	5	8	2	2
10	Dartmouth	1139.06	13	6	7	0	0
11	Cornell	1138.46	15	7	8	1	0
12	Naval Academy	1125.23	23	16	7	1	2
13	Drexel	1124.70	15	7	8	1	1
14	Williams	1123.07	21	12	9	0	0
15	Brown	1117.40	13	5	8	0	0
16	Middlebury	1115.98	18	12	6	1	2
17	Bates	1115.83	20	12	8	2	1
18	GWU	1115.24	15	7	8	2	0
19	Western Ontario	1090.66	10	4	6	0	0
20	Wesleyan	1082.01	18	11	7	0	1
21	Colby	1080.69	18	11	7	0	0
22	Amherst	1071.82	16	6	10	1	1
23	Hobart	1070.06	17	9	8	0	0
24	Chicago	1067.64	9	9	0	0	0
25	Bowdoin	1065.38	17	4	13	1	0
26	Hamilton	1046.36	17	7	10	0	0
27	Georgetown	1025.96	11	7	4	0	0
28	Conn College	1024.73	21	7	14	0	0
29	Tufts	1003.35	16	4	12	0	0
30	Stanford	1003.01	13	5	8	0	0

Fig. ELO-4a

SELF-CONSISTENT ELO APPLIED TO THE CSA 2014-15 SEASON - Men

continuation:

31	MIT	987.10	17	9	8	0	0
32	Bucknell	981.73	11	4	7	0	0
33	Johns Hopkins	963.64	8	3	5	0	0
34	NYU	961.20	7	4	3	0	1
35	Haverford	960.65	16	9	7	1	1
36	Northwestern	960.26	10	7	3	1	1
37	Northeastern	960.25	11	7	4	1	0
38	Denison	942.89	8	1	7	0	0
39	Fordham	939.94	14	9	5	0	0
40	Lehigh	925.27	9	5	4	0	0
41	Washington(St Louis)	914.94	7	3	4	0	0
42	Boston College	907.19	9	4	5	0	1
43	Colgate	905.19	7	2	5	0	0
44	Bryant	903.24	12	4	8	1	1
45	Boston U	900.79	13	3	10	1	0
46	Vassar	900.36	8	0	8	0	0
47	Penn State	888.45	5	2	3	0	0
48	Richmond	885.47	7	3	4	0	0
49	Washington, U of	867.40	7	5	2	0	0
50	Charleston	860.84	5	2	3	0	0
51	Minnesota	845.59	8	4	4	0	0
52	Swarthmore	840.98	9	2	7	0	0
53	Vanderbilt	834.17	6	2	4	0	0
54	Cal Berkeley	820.74	5	2	3	0	0
55	Siena	811.27	11	1	10	0	0
56	Sewanee	802.37	1	0	1	0	0
57	Notre Dame	793.35	9	1	8	0	0
58	Ithaca	779.99	6	0	6	0	0
59	USC	762.58	8	0	8	0	0

Fig. ELO-4b

There is a class of ranking methods called Minimum Violations Ranking (MVR)[10] which algorithmically seek to minimize the number of so-called ranking violations, which occur when a lower ranked team beats a higher ranked team. Summing Wins(+) (= Losses(-)) over all 59 teams gives the total number of rank violations. Comparing the data shown in Figs. ELO-4a/b with those in Figs. ITA-2a/b we find 21 violations for ELO compared with 31 for ITA. If we adopt the number of violations as a metric for effectiveness of ranking scheme, ELO “wins” over ITA.

SELF-CONSISTENT ELO RANKING HISTORY CSA 2014-15 SEASON - Men

Figs. ELO-5a/b show the evolution of ELO rankings through the season. Here, teams are sorted according to their final ELO rank. For most teams, the rank has stabilized by the second ranking date in Jan.

ELO RANK HISTORY (8 RANKING DATES), MEN 2014-15

Final ELO rank		Thru Nov	Thru Dec	10-Jan	18-Jan	25-Jan	4-Feb	8-Feb	Feb 15 (pre-tour)
1	Trinity	3	4	1	1	1	1	1	1
2	St. Lawrence	7	1	3	3	2	2	2	2
3	Columbia	2	3	2	2	3	3	3	3
4	Harvard	30	5	8	7	6	4	4	4
5	Yale	32	6	7	4	5	5	5	5
6	Rochester	14	2	5	5	4	6	6	6
7	F&M	8	11	13	9	9	7	7	7
8	U of Penn	5	10	11	10	10	9	8	8
9	Princeton	6	12	16	11	11	8	9	9
10	Dartmouth	34	9	15	12	12	10	11	10
11	Cornell	4	8	4	6	7	11	10	11
12	Naval Academy	1	13	17	13	13	12	12	12
13	Drexel	41	38	40	18	18	13	13	13
14	Williams	12	16	12	14	14	14	14	14
15	Brown	9	7	6	8	8	16	15	15
16	Middlebury	11	14	10	17	16	15	16	16
17	Bates	10	17	14	16	17	18	17	17
18	GWU	13	15	9	15	15	17	18	18
19	Western Ontario	42	40	36	24	22	24	25	19
20	Wesleyan	21	25	25	21	19	20	19	20
21	Colby	17	22	21	19	20	19	20	21
22	Amherst	16	21	19	22	21	21	21	22
23	Hobart	24	23	24	23	23	23	22	23
24	Chicago	19	19	20	26	25	22	23	24
25	Bowdoin	15	20	22	28	27	26	24	25
26	Hamilton	20	26	23	27	26	25	26	26
27	Georgetown	23	24	26	25	24	27	27	27
28	Conn College	22	27	32	30	28	28	28	28
29	Tufts	28	37	38	32	29	30	29	29
30	Stanford	36	34	35	34	32	29	30	30

Fig. ELO-5a

At the start of each season, every team starts with the same number of ranking points (1000). This is part of the objective assumption. At any point in the season, when two teams play one another there is a transfer of ranking points between just those teams. The winner gains a certain number of points and the loser loses the same number of points. Exactly what that number is depends on what the ranking points differential is between the teams immediately prior to them playing. So, let's consider what happens after the first week of matches. Half of the teams that played (the winning teams) gain

SELF-CONSISTENT ELO RANKING HISTORY CSA 2014-15 SEASON - Men

ranking points, and the other half (the losing teams) lose ranking points. Teams that didn't play retain their previous ranking points. No matter how good one imagines the teams are that didn't play during the first week of play are, they will be ranked behind all of the teams that won during that week, and be ranked AHEAD of all the teams that lost during that week. A team that continues to win continues to gain ranking points; a team that loses continues to lose ranking points. Drexel scheduled many of its toughest matches early in the season and did not win until after the 10 jan ranking date. Therefore, on 10 jan its ranking points total will be its starting value (1000) minus a bunch of points whose magnitude depends on the quality of the teams it has lost to. This is why Drexel has a weak early ranking (lower than 30 - the "average" rank since there are approximately 60 teams). Once Drexel starts winning matches its ranking rapidly improves. Chicago had an unbeaten season so it must, by the ELO method, end with a number of ranking points equal to its starting value (1000) plus a bunch of points. Chicago's strength of schedule(SoS) was "weak" (the highest ranked team it played was Georgetown (#27)) and the CSA must be diligent in enforcing adequate team SoS.

continuation:

31	MIT	39	42	42	35	31	32	31	31
32	Bucknell	44	43	43	42	38	33	32	32
33	Johns Hopkins	38	35	37	31	30	31	33	33
34	NYU	27	29	28	38	34	34	34	34
35	Haverford	26	28	27	39	35	35	35	35
36	Northwestern	18	18	18	40	36	37	36	36
37	Northeastern	35	39	39	37	33	36	37	37
38	Denison	51	52	52	50	46	38	38	38
39	Fordham	47	44	44	44	39	39	39	39
40	Lehigh	43	46	46	41	37	40	40	40
41	Washington(St Louis)	37	33	33	47	45	43	41	41
42	Boston College	50	49	49	46	41	41	42	42
43	Colgate	25	41	41	33	42	42	43	43
44	Bryant	48	50	50	48	43	44	44	44
45	Boston U	53	48	48	45	44	45	45	45
46	Vassar	45	51	51	53	49	46	46	46
47	Penn State	52	53	53	51	48	48	47	47
48	Richmond	54	54	54	52	47	47	48	48
49	Washington, U of	55	55	55	55	51	49	49	49
50	Charleston	33	32	31	20	52	50	50	50
51	Minnesota	57	56	56	57	55	51	51	51
52	Swarthmore	56	57	57	56	54	52	52	52
53	Vanderbilt	31	31	30	29	53	53	53	53
54	Cal Berkeley	46	47	47	43	40	54	54	54
55	Siena	58	58	58	58	58	55	55	55
56	Sewanee	29	30	29	36	56	56	56	56
57	Notre Dame	40	36	34	49	57	57	57	57
58	Ithaca	59	59	59	59	59	58	58	58
59	USC	49	45	45	54	50	59	59	59

Fig. ELO-5b

SELF-CONSISTENT ELO RANKING HISTORY CSA 2014-15 SEASON - Women

ELO RANK HISTORY (7 RANKING DATES), WOMEN 2014-15

	Final ELO rank	Thru Nov	Thru Dec	10-Jan	18-Jan	25-Jan	4-Feb	8-Feb
1	U of Pennsylvania	2	3	3	2	2	1	1
2	Harvard U	23	19	1	4	3	2	2
3	Trinity C	1	2	5	1	1	3	3
4	Princeton U	4	4	4	6	5	4	4
5	Yale U	24	5	6	3	4	5	5
6	Cornell U	3	1	2	5	6	6	6
7	Dartmouth C	22	31	23	7	7	8	8
8	Columbia U	6	6	13	8	9	7	7
9	GWU	16	18	8	9	8	9	9
10	Stanford U	10	7	7	10	10	10	10
11	Brown U	5	8	12	18	15	11	11
12	Williams C	19	17	14	11	12	12	12
13	Middlebury C	7	9	9	15	11	13	13
14	F&M	15	15	17	12	13	15	14
15	Bates C	8	10	15	13	17	16	15
16	Drexel U	29	30	32	16	16	17	16
17	Amherst C	18	22	19	17	18	18	17
18	Hamilton C	30	25	25	20	21	20	18
19	U of Virginia	12	11	11	14	14	14	19
20	Bowdoin C	25	23	26	21	22	21	20
21	St. Lawrence U	11	14	18	19	20	23	21
22	Colby C	13	13	16	23	23	22	22
23	Georgetown U	26	26	22	22	19	19	23
24	Wesleyan U	14	16	20	24	24	25	24
25	Wellesley C	21	20	24	25	25	26	25
26	William Smith C	20	24	27	27	27	27	26
27	Johns Hopkins U	34	34	29	26	26	24	27
28	Tufts U	40	33	34	30	31	31	28
29	Mount Holyoke C	31	28	31	29	29	30	29
30	Connecticut C	33	27	30	28	28	28	30
31	Haverford C	32	39	39	34	32	32	31
32	Vassar C	27	36	38	32	33	33	32
33	Boston C	39	35	37	33	34	34	34
34	Northwestern U	9	12	10	36	35	35	33
35	Washington(St.Louis)	17	21	21	39	39	36	35
36	Dickinson C	38	42	42	38	37	37	36
37	Northeastern U	35	32	33	40	40	40	37
38	Bucknell U	41	38	36	31	30	29	38
39	Smith C	45	40	40	35	38	39	39
40	U of Rochester	36	41	41	37	36	38	40
41	Colgate U	44	45	45	43	44	44	41
42	Fordham U	42	43	43	41	42	41	43
43	New York U	43	44	44	42	41	43	44
44	U of Minnesota	28	29	28	44	43	42	42
45	U of Notre Dame	37	37	35	45	45	45	45

Fig. ELO-5c

Early ranking “anomalies” will always be resolved by ELO before the end of the regular season. Consideration can be given to “publishing” traditional CSA rankings until some agreed date (eg second ranking date in Jan) with a switch to ELO computer rankings for the remainder of the season.

SELF-CONSISTENT ELO - INTERPRETATION OF RANKING POINTS

Finally, we discuss how to interpret the ELO ranking points that appear in the third column of Figs. ELO-4a/b and are repeated in Fig. ELO-6 below for the top 30 ranked teams according to the Feb 15 rankings. In particular, how should we interpret magnitudes of point differentials between teams? If we simply take the difference in ranking points to form ΔR, substitute into the expression for E(ΔR) given, for example, in Eq 3, we obtain the expectation of winning and losing if the two teams were to play one another again. If the reader is uncomfortable with evaluating the expression for E, he/she can simply estimate the value by interpolating from the Table that appears on the right hand side of the Figure.

1	Trinity	1177.19
2	St. Lawrence	1175.64
3	Columbia	1167.44
4	Harvard	1166.34
5	Yale	1161.66
6	Rochester	1161.36
7	F&M	1150.65
8	U of Penn	1142.83
9	Princeton	1142.45
10	Dartmouth	1139.06
11	Cornell	1138.46
12	Naval Academy	1125.23
13	Drexel	1124.70
14	Williams	1123.07
15	Brown	1117.40
16	Middlebury	1115.98
17	Bates	1115.83
18	GWU	1115.24
19	Western Ontario	1090.66
20	Wesleyan	1082.01
21	Colby	1080.69
22	Amherst	1071.82
23	Hobart	1070.06
24	Chicago	1067.64
25	Bowdoin	1065.38
26	Hamilton	1046.36
27	Georgetown	1025.96
28	Conn College	1024.73
29	Tufts	1003.35
30	Stanford	1003.01

Expectation of Winning and Losing given ranking points spread = ΔR (using the same R_s = 6.667 that produced the rankings)

ΔR	%E(ΔR)	%E(-ΔR)
0	50.0	50.0
0.5	51.9	48.1
1.0	53.7	46.3
2.0	57.4	42.6
4.0	64.6	35.4
8.0	76.9	23.1
16.0	91.7	8.3

$$E(\Delta R) = \frac{1}{[1 + e^{-\Delta R / R_s}]}$$

Fig. ELO-6

SELF-CONSISTENT ELO - INTERPRETATION OF RANKING POINTS

Example 1: The points gap between Trinity and St Lawrence (1177.19 - 1175.64 = 1.55) implies an expectation / probability of Trinity beating St Lawrence approximately 56% of the time. Equivalently, St Lawrence is predicted to beat Trinity 44% of the time.

Example 2: Princeton vs Navy points gap at end of season is 17.22. Navy beat Princeton and the magnitude of this upset is quantified by $E(\Delta R) = E(17.22) = 0.93$. The ELO-predicted expectation of Princeton winning, given the season results, is 93%; and of Navy winning is 7%. **ELO agrees that Navy pulled a big upset over Princeton!**

Example 3: Consider the following question:
Q: Is a win by the 35th ranked team over a team ranked 30 equivalent to the 6th ranked team beating the number 1 ranked team?
A: The way to look at the ELO rankings is that the number of points "gained" for winning a match is proportional to the quantity in square brackets on the RHS of Eq 2, where E is given by the expression on the RHS of Eq 3. For the S term in Eq 2 you use the value 1 if you win, and the value 0 is you lose. The crucial thing is that the points gained or lost depends only on the difference in rating points for the two teams. So it is not necessarily true that a win by the 35th ranked team over the 30th ranked team is the same as 6 beating 1 UNLESS the difference in rating points between the 35th and 30th teams is the same as the difference in points between the 6th and 1st. Specifically, from p27 Fig ELO-4a in the case of the Men's 2014-15 season, the 1st ranked team has 1177.19 rating points, the 6th team has 1161.36 for a difference of 15.83. The 30th ranked team has 1003.01 points and the 35th team has 960.65 points for a difference of 42.36. This is MUCH more than the difference between the 1st and 6th ranked teams. So there is a much greater difference in computed strength between the 35th and 30th teams than between the 6th and 1st, and a much lower probability of winning as a result (from plugging into the expression for E). Note also that the rating points gap (computed difference in level) between teams 31 and 30 was 15.91 ... more than twice he points gap (computed difference in level) between Rochester and Columbia who were ranked 6 and 3 respectively. This example is specific to the points distribution for the 2014-15 season!

SUMMARY COMPARISON BETWEEN CSA, ITA AND ELO - Men

COMPARISON OF FINAL (PRE-TOURNAMENT) RANKING

	CSA	ITA	ELO
1	Trinity	Trinity	Trinity
2	St. Lawrence	St. Lawrence	St. Lawrence
3	Harvard	Columbia	Columbia
4	Columbia	Harvard	Harvard
5	Yale	Rochester	Yale
6	Rochester	Yale	Rochester
7	F&M	F&M	F&M
8	U of Penn	Princeton	U of Penn
9	Princeton	U of Penn	Princeton
10	Cornell	Dartmouth	Dartmouth
11	Dartmouth	Cornell	Cornell
12	Drexel	Drexel	Naval Academy
13	Williams	Naval Academy	Drexel
14	Naval Academy	Williams	Williams
15	GWU	Bates	Brown
16	Bates	Middlebury	Middlebury
17	Middlebury	Brown	Bates
18	Brown	GWU	GWU
19	Western Ontario	Wesleyan	Western Ontario
20	Wesleyan	Colby	Wesleyan
21	Colby	Chicago	Colby
22	Hobart	Western Ontario	Amherst
23	Bowdoin	Amherst	Hobart
24	Amherst	Hobart	Chicago
25	Hamilton	Bowdoin	Bowdoin
26	Conn College	Conn College	Hamilton
27	Stanford	MIT	Georgetown
28	Tufts	Georgetown	Conn College
29	MIT	Tufts	Tufts
30	Virginia	Hamilton	Stanford

Fig. SUMMARY-1a

SUMMARY COMPARISON BETWEEN CSA, ITA AND ELO - Men

continuation:

31	Bucknell	Haverford	MIT
32	Chicago	NYU	Bucknell
33	Georgetown	Northeastern	Johns Hopkins
34	Johns Hopkins	Northwestern	NYU
35	Denison	Stanford	Haverford
36	NYU	Bucknell	Northwestern
37	Haverford	Fordham	Northeastern
38	Northeastern	Johns Hopkins	Denison
39	Fordham	Washington, U of	Fordham
40	Lehigh	Richmond	Lehigh
41	Northwestern	Lehigh	Washington(St Louis)
42	Boston College	Boston U	Boston College
43	Colgate	Washington(St Louis)	Colgate
44	Bryant	Minnesota	Bryant
45	Boston U	Bryant	Boston U
46	Washington(St Louis)	Boston College	Vassar
47	Richmond	Charleston	Penn State
48	Davidson	Penn State	Richmond
49	Miami	Cal Berkeley	Washington, U of
50	Charleston	Colgate	Charleston
51	Washington, U of	Vanderbilt	Minnesota
52	Swarthmore	Swarthmore	Swarthmore
53	Minnesota	Denison	Vanderbilt
54	Cal Berkeley	Notre Dame	Cal Berkeley
55	Dickinson	Siena	Siena
56	Siena	Ithaca	Sewanee
57	Vassar	USC	Notre Dame
58	Bard	Vassar	Ithaca
59	Vanderbilt	Sewanee	USC

Fig. SUMMARY-1b

DISCUSSION

When comparing ITA predictions with ELO predictions it is important to take a dispassionate view of the results. For example Princeton, Drexel, and Bates would surely prefer the ITA predictions shown in Fig. SUMMARY-1a over the ELO predictions, whereas Penn, Navy, and Brown would likely prefer ELO predictions over ITA predictions! However, it is best to review the findings discussed previously.

First, we must note that there is no such thing as a “correct” ranking system. At best, our job is to seek a robust system which gives sensible results and produces an acceptably small number of ranking anomalies. The alternative is to maintain the hands-on approach used by the CSA until now. However, one of the most contentious aspects within the association is rankings, whether individual or team. Bubble positions between the various divisions will always be a particular focus and to minimize contention the CSA should eliminate human influence and apply an objective ranking system.

In choosing ranking methods to test on college squash results we were initially attracted to the ITA method since it has been applied for a number of years to rank teams and individuals in college tennis. If the ITA system proved to be satisfactory for squash, a closely related racket sport, there would be advantages to advertising that the CSA was adopting the same approach used by the ITA. For all the criticism that the CSA receives for its team rankings we know that, for the most part, the CSA gets team rankings right! This is the reason for comparing predictions of candidate computer rankings with the CSA’s rankings. We should hope for good, but not identical, agreement. Although the ITA method (with parameters NBEST and NWORST tuned to squash) was found to produce sensible results for teams ranked in the top 25, the results were strikingly deficient for teams ranked below the top 25 (see Figs. ITA-2a/b). Additionally, the ITA method shows an unfortunate dependence of ranking results on match date schedule (see Fig. ITA-4 for a simple demonstration). This is especially troubling since detailed scheduling is beyond the control of team coaches.

The ELO ranking method has been applied by the US Chess Federation since 1960, and by the World Chess Federation since 1970. By November 2012, over 11,000 chess players worldwide had an active ELO rating! The ELO system has been applied to many team sports, including professional basketball, football and soccer. The particular brand of ELO that is usually discussed in the literature (and is the version used in chess) is called Sequential ELO in this report. However, for college athletics where there is a 100% turnover of players in each team over the course of four years, and where a single particularly strong recruiting year can completely change a team’s prospects for having a successful season, we believe that the most appropriate form of ELO to use is the iterated Self-Consistent ELO method. This method does not require a subjective pre-season rank - all teams have equal rank at the start of each season, its results are completely independent of match date schedule since it takes into account all matches that have been played to date in the season, and there are no adjustable (by human) parameters in the method. Figs. ELO-3a/b shows that Self-Consistent ELO produces sensible results for teams ranked in the top 25 and, for the most part, for teams ranked below the top 25 as well.

DISCUSSION

Based on the men's CSA 2014-15 season, it appears that Self-Consistent ELO is a promising candidate for adoption by the CSA as an objective computer ranking system, whereas ITA is not. To ensure that the ELO success is not specific to the men's 2014-15 dataset, we have also applied Self-Consistent ELO to the women's 2014-15 season and the men's 2013-14 and 2012-13 seasons:

FINAL ELO WOMEN'S RANKINGS (using NMAX = 99) WITH WIN/LOSS DATA

Wins (+) are wins against opponents who finished higher in rank

Losses (-) are losses against opponents who finished lower in rank

TEAM		CSA RANK	ELO RANK	Ranking Points	Matches played	Won	Lost	Wins(+)	Losses(-)	ELO - CSA
Amherst C	1	Harvard U	U of Pennsylvania	1265.72	13	12	1	0	1	-1
Bates C	2	U of Pennsylvania	Harvard U	1265.72	12	11	1	0	0	1
Boston C	3	Trinity C	Trinity C	1265.59	15	14	1	1	0	0
Bowdoin C	4	Princeton U	Princeton U	1237.42	11	9	2	0	0	0
Brown U	5	Yale U	Yale U	1215.45	14	10	4	0	0	0
Bucknell U	6	Cornell U	Cornell U	1194.81	14	9	5	0	0	0
Colby C	7	Columbia U	Dartmouth C	1168.58	13	6	7	0	1	-2
Colgate U	8	GWU	Columbia U	1168.57	13	7	6	1	1	1
Columbia U	9	Dartmouth C	GWU	1163.65	15	9	6	1	1	1
Connecticut C	10	Stanford U	Stanford U	1158.94	13	4	9	1	0	0
Cornell U	11	Brown U	Brown U	1139.20	12	5	7	0	0	0
Dartmouth C	12	Williams C	Williams C	1122.86	20	13	7	0	0	0
Dickinson C	13	Middlebury C	Middlebury C	1106.89	18	13	5	0	0	0
Drexel U	14	F&M	F&M	1090.57	16	9	7	0	0	0
Fordham U	15	Drexel U	Bates C	1071.37	19	10	9	0	0	-1
F&M	16	Bates C	Drexel U	1068.95	18	8	10	0	0	1
Georgetown U	17	Amherst C	Amherst C	1048.11	16	10	6	0	0	0
GWU	18	Hamilton C	Hamilton C	1028.30	20	13	7	0	0	0
Hamilton C	19	Bowdoin C	U of Virginia	1025.59	12	10	2	0	0	-4
Harvard U	20	St. Lawrence U	Bowdoin C	1013.12	20	9	11	0	0	1
Haverford C	21	Colby C	St. Lawrence U	990.55	10	4	6	0	0	1
Johns Hopkins U	22	Wesleyan U	Colby C	990.46	22	9	13	0	0	1
Middlebury C	23	U of Virginia	Georgetown U	980.44	10	6	4	0	0	-6
U of Minnesota	24	Wellesley C	Wesleyan U	968.76	17	6	11	0	0	3
Mount Holyoke C	25	William Smith C	Wellesley C	961.43	21	13	8	0	0	1
New York U	26	Tufts U	William Smith C	951.12	17	12	5	0	0	1
U of Notre Dame	27	Connecticut C	Johns Hopkins U	933.76	8	5	3	0	0	-3
Northeastern U	28	Mount Holyoke C	Tufts U	925.94	19	6	13	0	2	2
Northwestern U	29	Georgetown U	Mount Holyoke C	925.81	21	6	15	1	1	1
U of Pennsylvania	30	Johns Hopkins U	Connecticut C	925.71	22	6	16	2	0	3
Princeton U	31	Haverford C	Haverford C	902.82	15	10	5	0	1	0
U of Rochester	32	Boston C	Vassar C	902.13	12	6	6	1	0	-1
Smith C	33	Vassar C	Boston C	880.76	12	4	8	0	0	1
Stanford U	34	Dickinson C	Northwestern U	880.37	9	6	3	0	1	-3
St. Lawrence U	35	Bucknell U	Washington(St.Louis)	875.79	7	5	2	1	0	-3
Trinity C	36	Smith C	Dickinson C	871.36	11	3	8	0	0	2
Tufts U	37	Northwestern U	Northeastern U	855.64	7	1	6	0	0	-6
Vassar C	38	Washington(St.Louis)	Bucknell U	845.83	7	0	7	0	0	3
U of Virginia	39	Colgate U	Smith C	845.75	16	0	16	0	0	3
Washington(St.Louis)	40	U of Rochester	U of Rochester	835.32	9	3	6	0	1	0
Wellesley C	41	Fordham U	Colgate U	834.77	9	1	8	1	0	2
Wesleyan U	42	New York U	Fordham U	812.89	9	2	7	0	0	1
Williams C	43	Northeastern U	New York U	788.80	9	1	8	0	0	1
William Smith C	44	U of Minnesota	U of Minnesota	762.53	8	1	7	0	0	0
Yale U	45	U of Notre Dame **	U of Notre Dame	731.88	3	0	3	0	0	0

Fig. ELO-women

DISCUSSION

Women's results are shown in Fig. ELO-women. We see good agreement through the top 18 spots between ELO and CSA. The last column, in red, displays the difference between ELO and CSA ranking, with yellow highlighting where differences are greater than 3. Were it not for Virginia, Georgetown and Northeastern we would probably make a blanket statement here that the ELO ranks make sense throughout. However, it is striking that the ELO predicts these three teams should be ranked much higher than their CSA rank. After a cursory review of all match results for these teams we were unable to find compelling arguments for preferring the ELO ranking of these teams over CSA's (or vice versa!). We note, however, that the CSA ranking system makes use of pre-season rank. In the absence of registered "upsets" memory inherent to the method preserves rank (whether high or low). ELO, on the other hand, makes teams earn their rating points, positive or negative with respect to their starting mean of 1000 in this report.

Consider now CSA's ranking of Columbia, GWU and Dartmouth (7, 8 and 9 respectively) compared with ELO's 8, 9 and 7. CSA chose to invoke a triangle for these teams since Columbia beat Dartmouth, GWU beat Columbia, and Dartmouth beat GWU. However, if instead of invoking triangles the CSA had adopted a different decision mechanism - one where records are compared *against opposition **excluding** teams in the triad.* Then we would find Dartmouth has "best" wins against Brown (#11) and Williams (#12); Columbia has best wins against Brown (#11) and Middlebury (#13); and GWU has best wins over Middlebury (#13) and F&M (#14). This would decide rank in precisely the order that ELO has predicted. This shows that the ELO order is, in fact, a perfectly logical choice. It just happens not to be the one that the CSA has chosen!

Both Harvard and Penn are seen to be ranked at the top with an identical number of rating points. The fact that Penn is listed #1 in the ELO ranking is an arbitrary convention buried in the logic of writing our version of ELO! In the event of obtaining a tie in points such as found here, a tie-breaking convention must be adopted. For completely different reasons, CSA invoked yet another triangle for settling final order between the Harvard, Penn and Trinity women. In the ELO context this is not necessary since Trinity has fewer rating points than Harvard and Penn (albeit by a very small margin). Nevertheless, adhering to the ideal of avoiding subjective decisions, ELO would declare that Trinity is unambiguously #3 and only the question is how to split the tie between Harvard and Penn. The resolution is uncontroversial - The tie is broken by determining who won the regular season dual meet. Penn won this encounter, therefore Penn would be declared #1 based on ELO plus objective decision making.

Finally, we consider the Men's 2013-14 and 2012-13 seasons along side the previously shown Men's 2014-15 season.

CSA AND ELO PREDICTIONS FOR MEN'S SEASONS 2012-13, 2013-14, and 2014-15

	2014-15			2013-14			2012-13		
	ELO		ELO - CSA	ELO		ELO - CSA	ELO		ELO - CSA
1	1177.19	Trinity	0	1300.05	Harvard U	0	1219.93	Trinity C	0
2	1175.64	St. Lawrence	0	1273.47	Trinity C	0	1182.53	Princeton U *	0
3	1167.44	Columbia	-1	1251.16	Yale U	0	1182.53	Harvard U *	0
4	1166.34	Harvard	1	1231.20	St. Lawrence U	0	1175.32	Yale U **	0
5	1161.66	Yale	0	1212.81	F&M	0	1175.32	Cornell U **	0
6	1161.36	Rochester	0	1195.56	U of Rochester	0	1147.53	U of Rochester	0
7	1150.65	F&M	0	1179.83	Cornell U	0	1127.26	F&M	0
8	1142.83	U of Penn	0	1164.28	U of Pennsylvania	0	1101.53	Dartmouth C	-1
9	1142.45	Princeton	0	1148.66	Princeton U	0	1100.73	St. Lawrence U	1
10	1139.06	Dartmouth	-1	1146.85	Western Ontario	0	1092.99	Williams C	0
11	1138.46	Cornell	1	1133.57	Columbia U	0	1090.53	Columbia U	0
12	1125.23	Naval Academy	-2	1119.86	Dartmouth C	0	1089.94	Western Ontario	-1
13	1124.70	Drexel	1	1108.07	Naval Academy	0	1082.20	U of Pennsylvania	1
14	1123.07	Williams	1	1095.37	Drexel U	0	1081.28	Naval Academy	0
15	1117.40	Brown	-3	1082.38	Bates C	-1	1080.09	Bates C	0
16	1115.98	Middlebury	-1	1082.10	Williams C	1	1079.49	Brown U	0
17	1115.83	Bates	1	1063.07	Wesleyan U	-1	1076.16	Middlebury C	0
18	1115.24	GWU	3	1061.90	Middlebury C	1	1070.72	GWU	0
19	1090.66	Western Ontario	0	1040.29	Bowdoin C	-2	1064.76	Bowdoin C	-2
20	1082.01	Wesleyan	0	1037.15	Brown U	1	1056.62	Wesleyan U	-2
21	1080.69	Colby	0	1036.93	GWU	1	1053.85	Amherst C	1
22	1071.82	Amherst	-2	1034.28	Amherst C	0	1049.45	Drexel U	3
23	1070.06	Hobart	1	1015.79	Colby C	0	1043.36	Hamilton C	0
24	1067.64	Chicago	-8	1000.27	Hamilton C	0	1037.65	Connecticut C	0
25	1065.38	Bowdoin	2	994.96	Bucknell U	-3	1034.10	Bucknell U	-2
26	1046.36	Hamilton	1	981.89	Connecticut C	1	1032.44	Colby C	1
27	1025.96	Georgetown	-6	970.00	MIT	-5	1014.77	Haverford C	-1
28	1024.73	Conn College	2	952.58	Northeastern U	-5	1011.23	Stanford U	2
29	1003.35	Tufts	1	950.04	Johns Hopkins U	-1	994.19	Georgetown U	0
30	1003.01	Stanford	3	948.75	Hobart C	4	993.91	Johns Hopkins U	0

Fig. ELO-CSA COMPARISONS

* Princeton, Harvard equal ranking points - rank decided by dual match result.
** Yale, Cornell rank decided similarly.

Notes:

- The agreement between ELO predictions and CSA is excellent for teams through the top two divisions.
- Anomalous differences are most often associated with emerging and club teams such as Chicago, Georgetown and Stanford (2014-15 results), Bucknell, Northeastern (2013-14 season). All emerging and club teams were included in the ELO ranking calculations and treated on an equal basis with the varsity teams!

CSA AND ELO PREDICTIONS FOR MEN'S SEASONS 2012-13, 2013-14, and 2014-15

Does the ELO choice of ranking Columbia ahead of Harvard in 2014-15 have a rational basis?:

- Columbia (#3) had 0 upset wins and 1 minimal upset loss to Harvard (#4).
- Harvard (#4) had 1 minimal upset win over Columbia (#3) and 1 upset loss to Rochester (#6) ranked two places behind
- If rank positions were to be reversed, Columbia would have 0 upset wins and 0 upset losses (ie in a relatively better situation). However, Harvard would no longer have any upset wins and would have an even worse loss to Rochester (who would be 3 spots lower) ⇒ the ELO rank has a logical basis!

In conclusion, Self-Consistent ELO does, indeed, hold promise. It's application to ranking individuals for the CSA Individual Tournaments and All-American awards would be equally straightforward.

References

[1] *ITA Division I Rankings Manual 2014 -15:* http://www.itatennis.com/Assets/ita_assets/pdf/Rankings/2014-15/2014-15+ITA+Division+I+Rankings+Manual.pdf

[2] *A Comprehensive Guide to Chess Ratings:* http://www.glicko.net/research/acjpaper.pdf

[3] *The Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons:* R A Bradley and M E Terry, Biometrika **39**, 324-45 (1952)

[4] Eg., http://sagarin.com

[5] http://collegesquashassociation.com/scheduleresults/mens-results/

[6] Eg., http://en.wikipedia.org/wiki/Elo_rating_system

[7] Eg., http://en.wikipedia.org/wiki/Logistic_distribution

[8] Eg., http://www.pro-football-reference.com/blog/?p=839

[9] *Note on Computing Ratings From Eigenvectors:* Richard Brent (2010) http://maths-people.anu.edu.au/~brent/pd/rpb237.pdf

[10] Eg., *Minimizing Game Score Violations in College Football Rankings:* B J Coleman http://digitalcommons.unf.edu/cgi/viewcontent.cgi?article=1000&context=bmgt_facpub

RPI Comparison with ELO and CSA

COMPARISON OF CSA, ELO AND RPI (M Bello) PRE-TOUR RANK

	CSA	ELO	ELO - CSA	RPI*	RPI - CSA
1	Trinity	Trinity	0	Trinity	0
2	St. Lawrence	St. Lawrence	0	Columbia	-2
3	Harvard	Columbia	-1	Harvard	0
4	Columbia	Harvard	1	Yale	-1
5	Yale	Yale	0	St. Lawrence	3
6	Rochester	Rochester	0	Rochester	0
7	F&M	F&M	0	F&M	0
8	U of Penn	U of Penn	0	U of Penn	0
9	Princeton	Princeton	0	Chicago	-23
10	Cornell	Dartmouth	-1	Williams	-3
11	Dartmouth	Cornell	1	Princeton	2
12	Drexel	Naval Academy	-2	Dartmouth	1
13	Williams	Drexel	1	Middlebury	-4
14	Naval Academy	Williams	1	Cornell	4
15	GWU	Brown	-3	Bates	-1
16	Bates	Middlebury	-1	Naval Academy	2
17	Middlebury	Bates	1	GWU	2
18	Brown	GWU	3	Drexel	6
19	Western Ontario	Western Ontario	0	Northwestern	-22
20	Wesleyan	Wesleyan	0	NYU	-16
21	Colby	Colby	0	Georgetown	-12
22	Hobart	Amherst	-2	Hobart	0
23	Bowdoin	Hobart	1	Wesleyan	3
24	Amherst	Chicago	-8	Virginia	-6
25	Hamilton	Bowdoin	2	Haverford	-12
26	Conn College	Hamilton	1	MIT	-3
27	Stanford	Georgetown	-6	Brown	9
28	Tufts	Conn College	2	Western Ontario	9
29	MIT	Tufts	1	Fordham	-10
30	Virginia	Stanford	3	Colby	9

ELO - CSA is difference in rank between ELO and CSA
RPI - CSA is similar difference between RPI and CSA