Accelerated Pairing

by

Otto Milvang

2016-03-26, Rev 1

2016-07-10, Rev 2

2016-08-19, Rev 3

2016-11-05, Rev 4

2017-09-22, Rev 5

2017-10-09, Rev 6

Introduction

In Swiss tournaments with a big number of players and with a large span of playing strength, the first round(s) may be quite uninteresting. On each board in the first game it's a huge gap between the players, only a few percent of the games have another result than win to the stronger part. This may be repeated in round two.

Round 1 on 2015/07/25 at 14:00											
Bo.	**No.**		**Name**	**Rtg**	**Pts.**	**Result**	**Pts.**		**Name**	**Rtg**	**No.**
1	1	GM	Fressinet Laurent	2707	0	1 - 0	0		Hjornevik Knut J	1958	217
2	218		Andreassen Mikkel	1954	0	0 - 1	0	GM	Ragger Markus	2688	2
3	3	GM	Hammer Jon Ludvig	2677	0	1 - 0	0		Rasmussen Lars	1953	219
4	220		Mott Jens	1953	0	0 - 1	0	GM	Smirin Ilia	2663	4
5	5	GM	Shankland Samuel L	2656	0	1 - 0	0		Jensen Ib	1948	221
6	222		Quinn David	1947	0	0 - 1	0	GM	Nisipeanu Liviu-Dieter	2654	6
7	7	GM	Jones Gawain C B	2647	0	1 - 0	0		Hogeman Joel	1946	223
8	224		Stam Bart	1946	0	0 - 1	0	GM	Rakhmanov Aleksandr	2626	8
9	9	GM	Naroditsky Daniel	2622	0	1 - 0	0		Skjoldager Flemming	1944	225
10	226		Philipsen William Horup	1943	0	0 - 1	0	GM	Grigoriants Sergey	2594	10
11	11	GM	Marin Mihail	2579	0	1 - 0	0		Nielsen Morten	1942	227
12	228		Subramanian Aditya	1942	0	½ - ½	0	GM	Kunin Vitaly	2576	12
13	13	GM	Maze Sebastien	2575	0	1 - 0	0		Matras Ove	1940	229
14	230		Perrett David C	1937	0	0 - 1	0	GM	Hansen Sune Berg	2571	14
15	15	GM	Timman Jan H	2566	0	1 - 0	0		Brobakken Geir	1936	231
16	232		Alfven Jorgen	1932	0	0 - 1	0	GM	Hillarp Persson Tiger	2563	16
17	17	GM	Brunello Sabino	2539	0	1 - 0	0		Rasmussen Morten	1932	233
18	234		Roksvold Jan	1927	0	0 - 1	0	GM	Glud Jakob Vang	2531	18
19	19	GM	Schandorff Lars	2520	0	1 - 0	0		Molvig Julius	1926	235
20	236		Weishaeutel Moritz	1926	0	0 - 1	0	GM	Rasmussen Allan Stig	2507	20

An example is Politiken cup 2015, with mean rating difference in round 1 = 530. In the first round there was 193 win, 15 draw, and 3 lost games to the highest rated player.

One way to overcome this weakness of the Swiss system is to split score brackets into smaller sub-brackets, and maybe merge sub-brackets of different points based on rating. This process is called acceleration. A very neat way to do this is to add virtual points to strong players.

Total points = Real points + virtual points.

Then you may run your standard Swiss algorithm with score brackets based on Total points. Almost all variants of acceleration can be solved with virtual points.

Take for instance the first round in the Britain acceleration (assume that the number of players are a multiple of 4 to keep it simple). "Divide the players in rating/grading order into four quarters A,B,C,D respectively. In the first round quarter A is paired against B, and C against D. Alternate the colors". This is the exact the same as "Add 1 virtual point to the upper half of the players and draw the round".

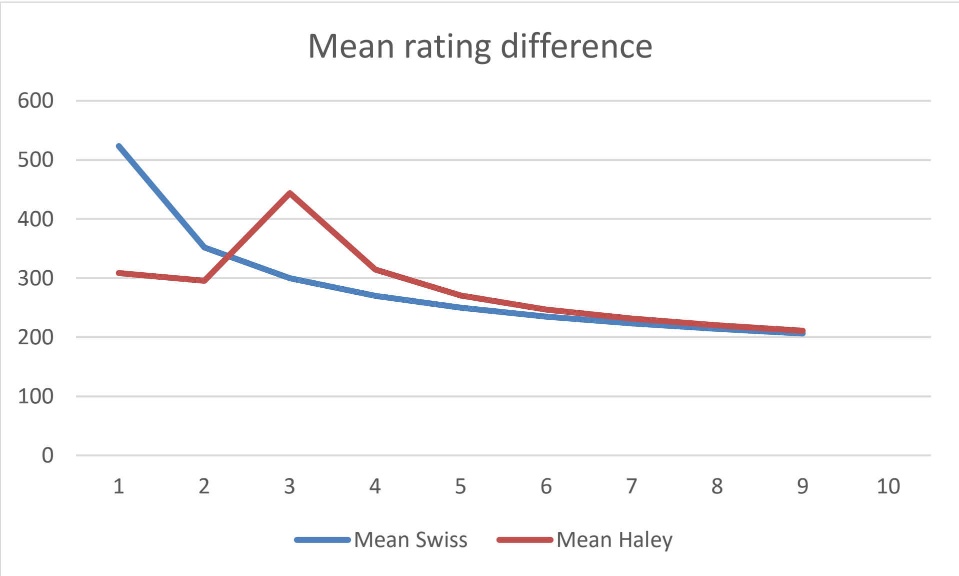
Accelerated paring systems are used worldwide. Nobody knows if this system really works or not. FIDE SPP Commission decided on the 86th FIDE Congress 2015 in Abu Dhabi to explore the behavior of different systems.

Different systems

The Haley system

In the Haley system the players are divided in equally spaced subgroups A, B, … based on rating. Each subgroup is assigned different amount of virtual points. The lowest subgroup is not assigned virtual points at all. Thus only players within the same subgroup will meet in the first round. The idea is that in round two weak players from A will meet strong players in B, weak players in B will meet strong players in C and so on. The virtual point will successively be removed in the following round. As the first glance it looks great just to divide into subgroups, but thing turns out to be a little more complicated. Let's assume that we have only one round with acceleration, and as described in the previous paragraph. If most of players in A and C win their games, players from A and C will meet in round two, and B and D will meet. This is actually the same as round 1 unaccelerated, and our round 1 is like round 2 unaccelerated, so we have just swapped round 1 and 2. Luckily we turn out to find better accelerated systems than this.

In the original Halley description all points are removed after round two, but this make no sense since both theory and practice shows that this is more or less to switch round 1 and 3.



The figure shows the mean rating difference between players in round 1-9 for Swiss and original Haley system for the same tournament.

The idea of the Haley system is good, but we need more rounds with acceleration. Four different schemas were tested. They are named Haley 1 to Haley 4 since they are all give acceleration point to some players for some rounds.

Haley 1 (Baku acceleration!)

Split the players in two groups A, and B. If the number of players are odd, put the extra player in B.

Add 1 virtual point to all players in A, and keep this in the first **three** rounds

Remove ½ virtual point from A, such than A players have ½ virtual point, and keep this for **two** rounds.

Haley 2

Split the players in two groups A, and B. If the number of players are odd, put the extra player in B.

Add 1 virtual point to all players in A, and keep this in the first **two** rounds

Remove ½ virtual point from A, such than A players have ½ virtual point, and keep this for **one** rounds.

Haley 3

Split the players in three groups A, B, and C. The number of players in A and B shall be even.

Add 2 virtual points to all players in A, 1 virtual point to all players in B, and keep this in the first **three** rounds

Remove 1 virtual point from A, and 1/2 virtual point from B, such than A players have 1 virtual point, players in B have ½ virtual point and keep this for **two** rounds.

Haley 4

Split the players in three groups A, B, and C. The number of players in A and B shall be even.

Add 2 virtual points to all players in A, 1 virtual point to all players in B, and keep this in the first **two** rounds

Remove 1 virtual point from A, and 1/2 virtual point from B, such than A players have 1 virtual point, players in B have ½ virtual point and keep this for **one** rounds.

Summary Haley

Method	Subgroups	Round1		Round2		Round3		Round4		Round5	
		A	B	A	B	A	B	A	B	A	B
Haley 1	A/B	1.0	0.0	1.0	0.0	1.0	0.0	0.5	0.0	0.5	0.0
Haley 2	A/B	1.0	0.0	1.0	0.0	0.5	0.0	-	-	-	-
Haley 3	A/B/C	2.0	1.0	2.0	1.0	2.0	1.0	1.0	0.5	1.0	0.5
Haley 4	A/B/C	2.0	1.0	2.0	1.0	1.0	0.5	-	-	-	-

Progressive acceleration

Players are divided in 3 subgroups A, B and C according to rating. Each group must include at least 25% and a maximum of 50% of the players. At the start of the tournament group A receive 2 virtual points, group B 1 virtual point and group C 0 virtual points. When a player in group B or C gains at least 1.5 real points, their virtual score is increased by another ½ virtual point. When a player gains their third real point, additional ½ virtual point is added, and for C group players, this will repeat on 4.5 real points. When a player achieves N/2 real points (where N is the number of rounds), their virtual points is set to 2. Before the penultimate round, all virtual points are cancelled and the system become a usual Swiss system.

Summary table:

Top number: real points. Number within brackets: virtual points. The rightmost cells of the table consider cases with more than 11 rounds (they don't include groups A or B).

Niveau (SG)	0	0,5	1	1,5	2	2,5	3	3,5	4	4,5	5	5,5	6	6,5	7
Gr. A					0	0,5	1	1,5	2	2,5	3	3,5	4	4,5	5
					(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)	(2)
Gr. B			0	0,5	1		1,5	2	2,5		3	3,5	4	4,5	5
			(1)	(1)	(1)		(1,5)	(1,5)	(1,5)		(2)	(2)	(2)	(2)	(2)
Gr. C	0	0,5	1		1,5	2	2,5		3	3,5	4			4,5	5
n=9	(0)	(0)	(0)		(0,5)	(0,5)	(0,5)		(1)	(1)	(1)			(2)	(2)
Gr. C	*0*	*0,5*	*1*		*1,5*	*2*	*2,5*		*3*	*3,5*	*4*		*4,5*	*5*	*5,5*
n>11	*(0)*	*(0)*	*(0)*		*(0,5)*	*(0,5)*	*(0,5)*		*(1)*	*(1)*	*(1)*		*(1,5)*	*(1,5)*	*(1,5)*

Britain acceleration

1 Divide the players in rating/grading order into four quarters, A, B, C, D respectively. The 'top half' players are in quarters A and B. If the number of players is not a multiple of four act as follows. Divide the number of players by 4. If the remainder is one or two enlarge the top half by two. If three enlarge the top half by two and the bottom half by 1.

2 In the first round, quarter A is paired against quarter B and quarter C is paired against quarter D. As normal the colour on board 1 is decided by lot (unless tournament regulations specify otherwise) and then alternates throughout.

3 For round 2 pairings, the score groups are, in order, (a) top half 1's (see 6), (b) bottom half 1's paired in grading order against an equal number of the highest-rated/graded top half players who did not win in round 1 taking account of colour requirements, (c) bottom half '1/2's and any remaining top half 1/2's , (d) bottom half 0's and any remaining top half 0's.

4 Acceleration continues with score groups (a) top half players with 100% (see 6), (b) bottom half players with 100% scores paired against the highest rated/graded top half players due the appropriate colour in the next two lower score groups, (c) remaining players in normal separate score groups, until either there are no bottom half players with 100% scores or there are only two rounds to play, whichever comes first.

5 In all subsequent rounds, each score group contains only players with the same score, with the possible addition of floaters.

6 A floater from the top score group will be paired against the highest rated/graded top half player who is a half point behind and is due the appropriate colour as normal. However, this pairing will be broken if the number of bottom half players with 100% exceeds the number of top half players remaining in pairing group (b) above. In this case the top half player with 100% is paired against the highest rated bottom half

player of the appropriate colour and the pairings then continue as detailed in 3(b) or 4(b) as appropriate.

Advanced Acceleration

In this method bottom half players within a half point of the lead are also paired against top half players. This can be used for longer tournaments. This method requires the top 'half' to be larger than the bottom so that there are sufficient players available to be paired against the bottom half within half point of lead. The top half should contain 56-60% of the entrants.

1 Divide the players in grading order into four 'quarters', A, B, C, D. The 'top half' players are in quarters A and B. The top half should be an even number between 56% and 60% of the total entry.

2 In the first round, quarter A is paired against quarter B and quarter C is paired against quarter D. As normal the colour on board 1 is decided by lot (unless specified otherwise) and then alternates throughout.

3 For round 2 pairings, the score groups are, in order, (a) top half l's, (b) bottom half l's paired in grading order against an equal number of the highest-rated/graded top half players who did not win in round 1 taking account of colour requirements, (c) remaining top half 1/2s (see 7) (d) bottom half 1/2s and top half 0's (see 6), (d) bottom half 0's and any remaining top half 0's.

4 Acceleration continues with score groups (a) top half players with 100% (see 6), (b) bottom half players with 100% scores paired against the highest rated/graded top half players due the appropriate colour in the next two lower score groups, (c) remaining players in normal separate score groups, until the designated number of rounds have been played or either there are no bottom half players with appropriate scores or there are only two rounds to play, whichever comes first.

5 In all subsequent rounds, each score group contains only players with the same score, with the possible addition of floaters.

6 A floater from the top score group will be paired against the highest rated/graded top half player who is a half point behind and is due the appropriate colour as normal. Similarly a floater from the next 'top half' score group will be paired against the highest rated/graded top half player who is a half point behind and is due the appropriate colour as normal. However this pairing will be broken if the number of bottom half players with 100% exceeds the number of top half players remaining in pairing group (b) above. In this case the top half player with 100% is paired against the highest rated/graded bottom half player of the appropriate colour and the pairings then continue as detailed in 3(b) or 4(b) as appropriate.

7 At this point it is worth counting the number of 'bottom half'A's and 'top half O's (or other appropriate scores in later rounds). If the former is greater then return to the previous step (b) and, working up from the bottom of the pairings made, substitute a 'top half player of the appropriate colour who has scored 0 with the highest rated unpaired appropriate 'top half who has scored Y. Repeat this process until you have sufficient 'top half O's to play the 'bottom half 'As. If even this fails to produce sufficient players it may be necessary to promote the highest rated 'bottom halfs' into the 'top half' group.

My comments

This method is more or less similar to the Haleys since it describes how subgroups shall be created and pared without using the term virtual points. There are some differences, and I have tried to implement it using virtual points. I have only implemented this for round 1 and 2, and then switch to Haley 2 in round 3 to see if there are any real differences between these methods.

Faded systems

These systems are quite similar to Haleys system, but instead of withdraw virtual points, virtual points are added, such that after 5 rounds (or less) all players have added maximum virtual points.

In this system players are divided in 6 subgroups, best 30%, 30-40%, 40-50%, 50-60%,60-70%, and 70% and up.

Faded 1:

Round	0-30	30-40	40-50	50-60	60-70	70+
1	2.0	1.0	1.0	0.0	0.0	0.0
2	2.0	2.0	1.0	1.0	1.0	0.0
3	2.0	2.0	1.0	1.0	1.0	1.0
4	2.0	2.0	2.0	2.0	1.0	1.0
5	2.0	2.0	2.0	2.0	2.0	1.0

Faded 2:

Round	0-30	30-40	40-50	50-60	60-70	70+
1	2.0	1.0	1.0	1.0	1.0	0.0
2	2.0	1.5	1.5	1.0	1.0	0.0
3	2.0	2.0	1.5	1.5	1.0	0.5
4	2.0	2.0	2.0	1.5	1.5	1.0
5	2.0	2.0	2.0	2.0	1.5	1.0
6	2.0	2.0	2.0	2.0	2.0	1.5

Faded 3:

Round	0-30	30-40	40-50	50-60	60-70	70+
1	1.5	1.0	0.5	0.5	0.5	0.0
2	1.5	1.0	1.0	0.5	0.5	0.0
3	1.5	1.5	1.0	1.0	0.5	0.5
4	1.5	1.5	1.5	1.0	1.0	0.5
5	1.5	1.5	1.5	1.5	1.0	1.0
6	1.5	1.5	1.5	1.5	1.5	1.0

It's of course an unlimited number of possibilities. Perhaps we should have worked more on this.

In a real implementation we always subtract the lowest virtual points from all players, since there is no sense of giving all players virtual points.

Rating

We will divide into several subgroups based on rating. The subgroups are: rating > 2200, 1900-2200, 1600-1900, 1300-1600, and below 1300.

Round	>2200	>1900	>1600	>1300
1	3.5	2.5	1.5	0.5
2	3.0	2.5	1.5	0.5
3	2.5	2.0	1.0	0.5
4	2.0	1.5	1.0	0.5
5	1.5	1.0	0.5	0.0
6	1.0	0.5	0.0	0.0
7	0.5	0.0	0.0	0.0

Split

The idea behind split is to split all score bracket into mini brackets based on rating. Each bracket is divided into > rating 2200, 1900-2200, 1600-1900, 1300-1600, and below 1300. Thus we will have no games with huge rating difference between players (max 300).

Tournament simulations

To test out the different methods, we will do tournament simulations. With the same set of players, we will run a big number of tournaments, and analyze the result. In the tournaments the results are drawn from a probability model described in «Probability for the outcome of a chess game based on rating». The input rating to this model may be the FIDE rating, a tournament rating performance, or other models.

The simulation is easy.

For each round make the pairing with JaVaFo Swiss pairing engine. Foe engine has options for adding virtual points.
Apply results according to http://www.nordstrandsjakk.no/documents/spp/Probability.pdf
Repeat this for 2000 to 10000 tournaments.

Model tournaments

The first attempt was to look at big tournaments already played and used the same set of players in the simulations. The test tournaments are:

1) Reykjavik Open 2015, 273 playes
2) Villa Be Benasque 2015, 412 players.
3) Politiken cup 2015, 432 players

Some test was run with these tournaments, however we found that the distribution of rating was rough, and it could be difficult to say something about the algorithms with these tournaments.

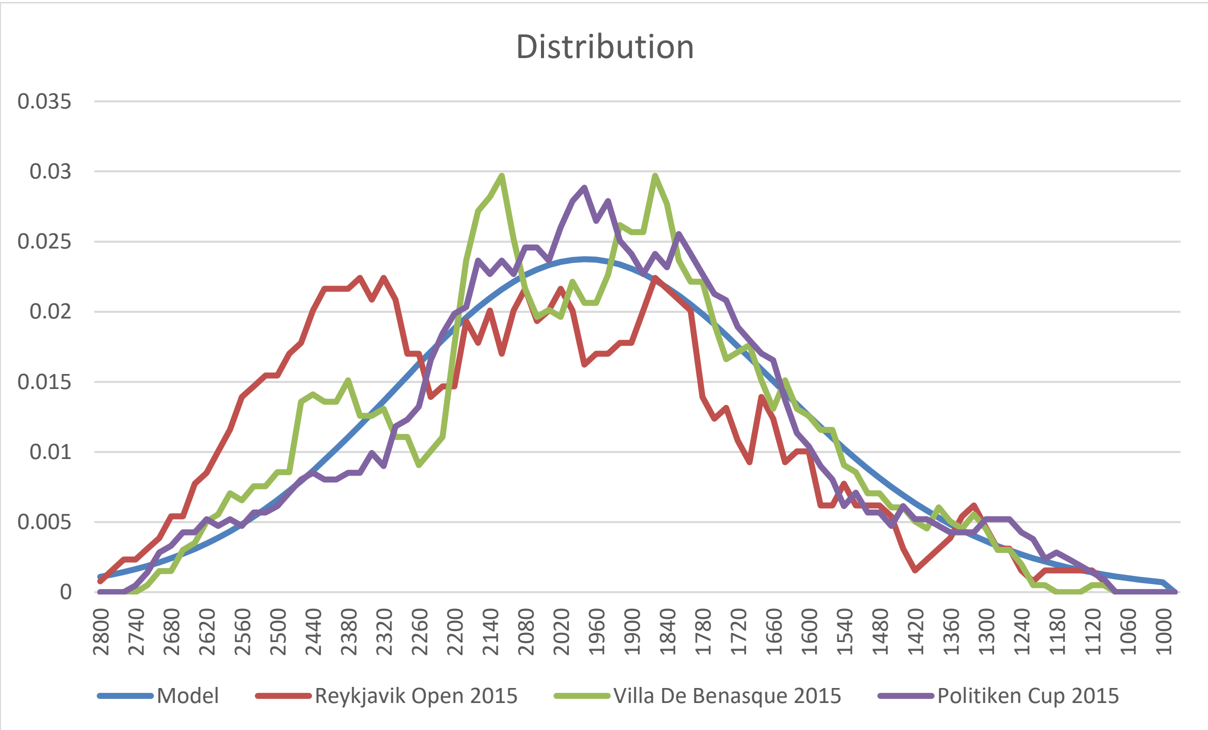


Figure 1. The density distribution of rating for the players in the model tournaments.

Instead of using one of the three tournaments directly we made a model tournament as a sum of Gauss models. A Gauss distribution with mean = μ, and standard deviation δ is denoted N(μ, δ).

For Villa Be Benasque 2015 and Politiken Cup 2015, a good approximation is

Model1 = 0.4 * N(1900,400) + 0.6 *N(2000,300)

And for Reykjavik open 2015 an approximation is

Model2 = 0.25 * N(1900,380) + 0.75 *N(2200,280)

We will also test the algorithms against a linear distribution.

In both this models we use the number of players (432), the titles and names from Politiken Cup 2015. The reason is that we want to estimate the probability for GM norms and IM norms for different systems. For full list of tournaments are attached in the end of the document.

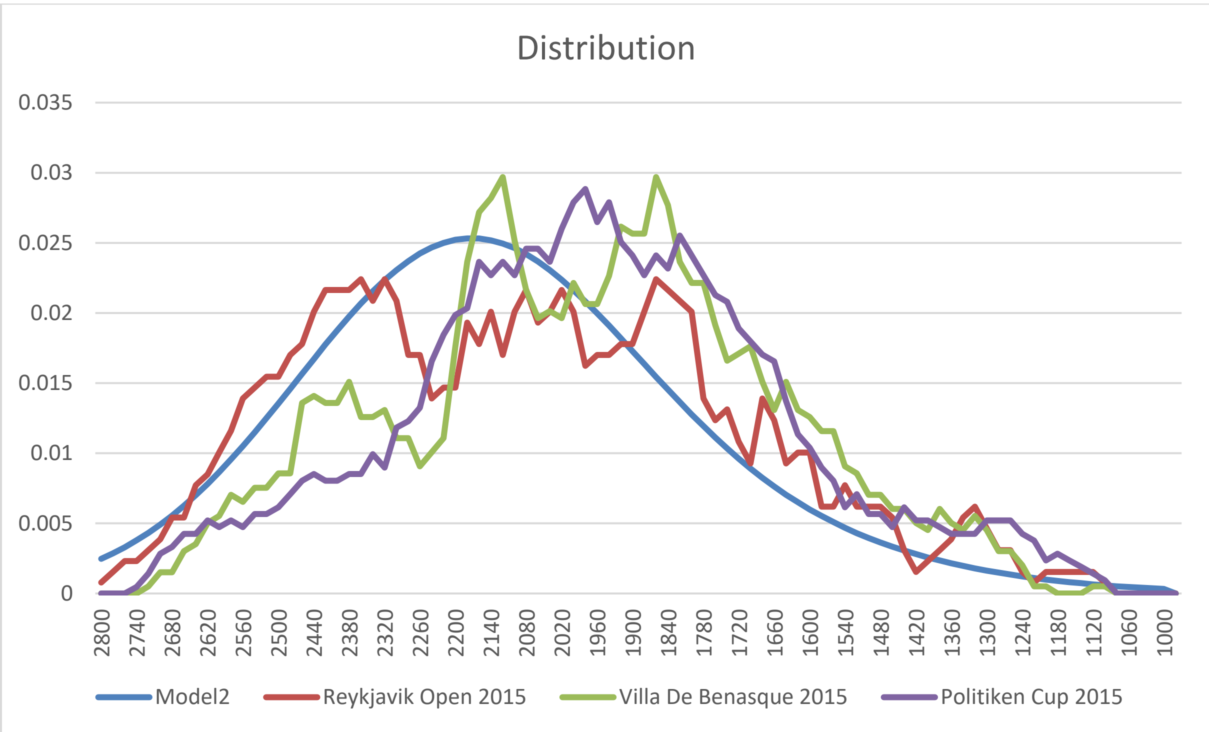


Figure 2. The density distribution of rating for the players in the second model tournaments.

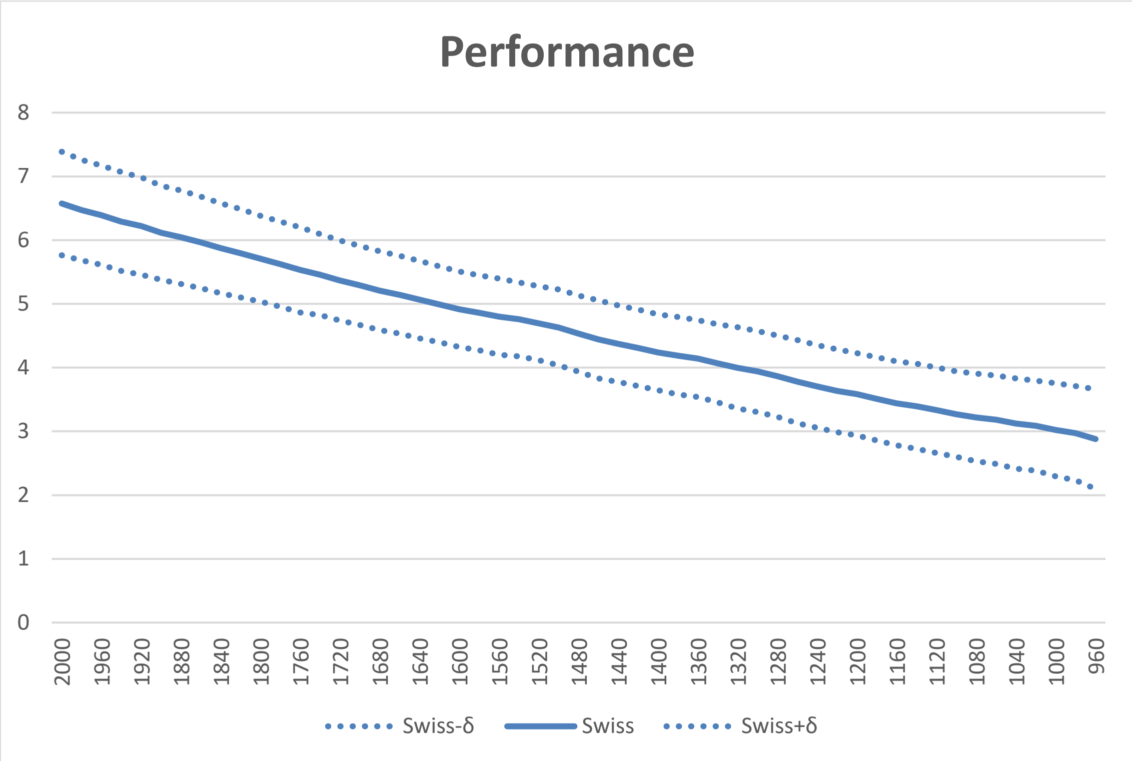
Evaluation

Introduction

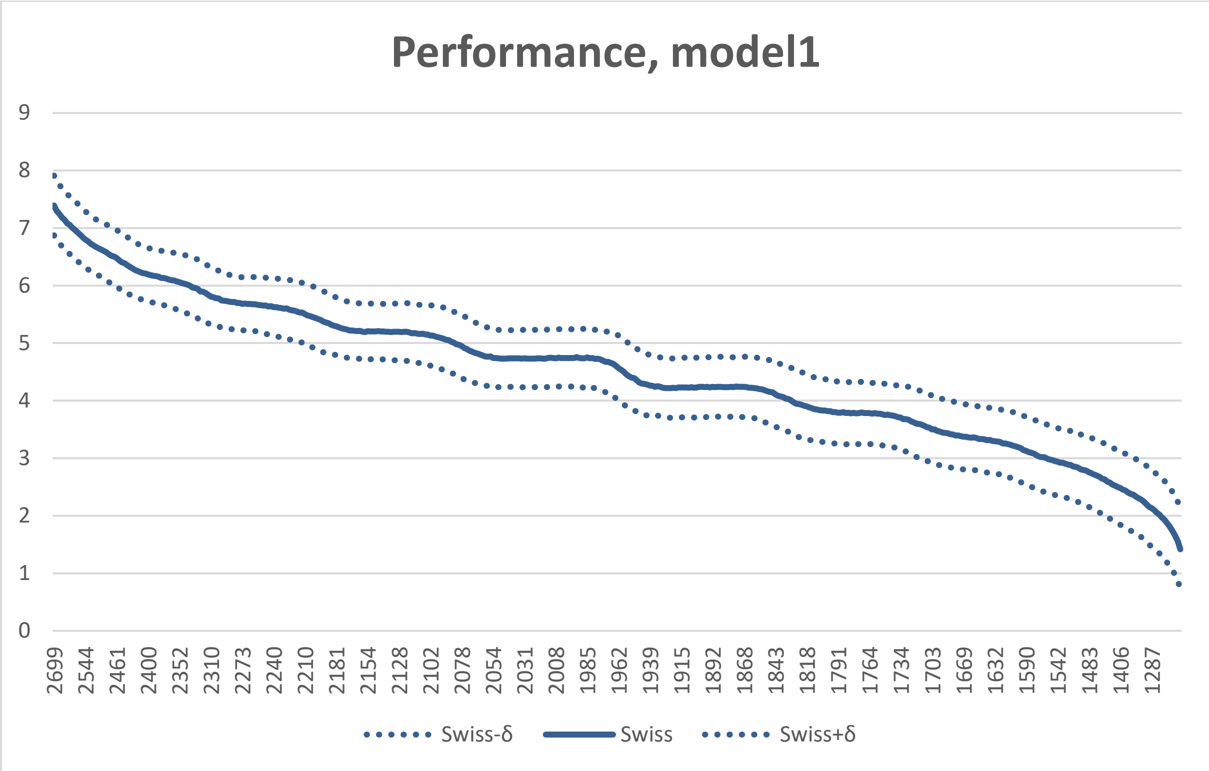
There are no official parameters to describe how good tournament systems performs. The parameters described below are a summary of what players think are good parameters.

Performance

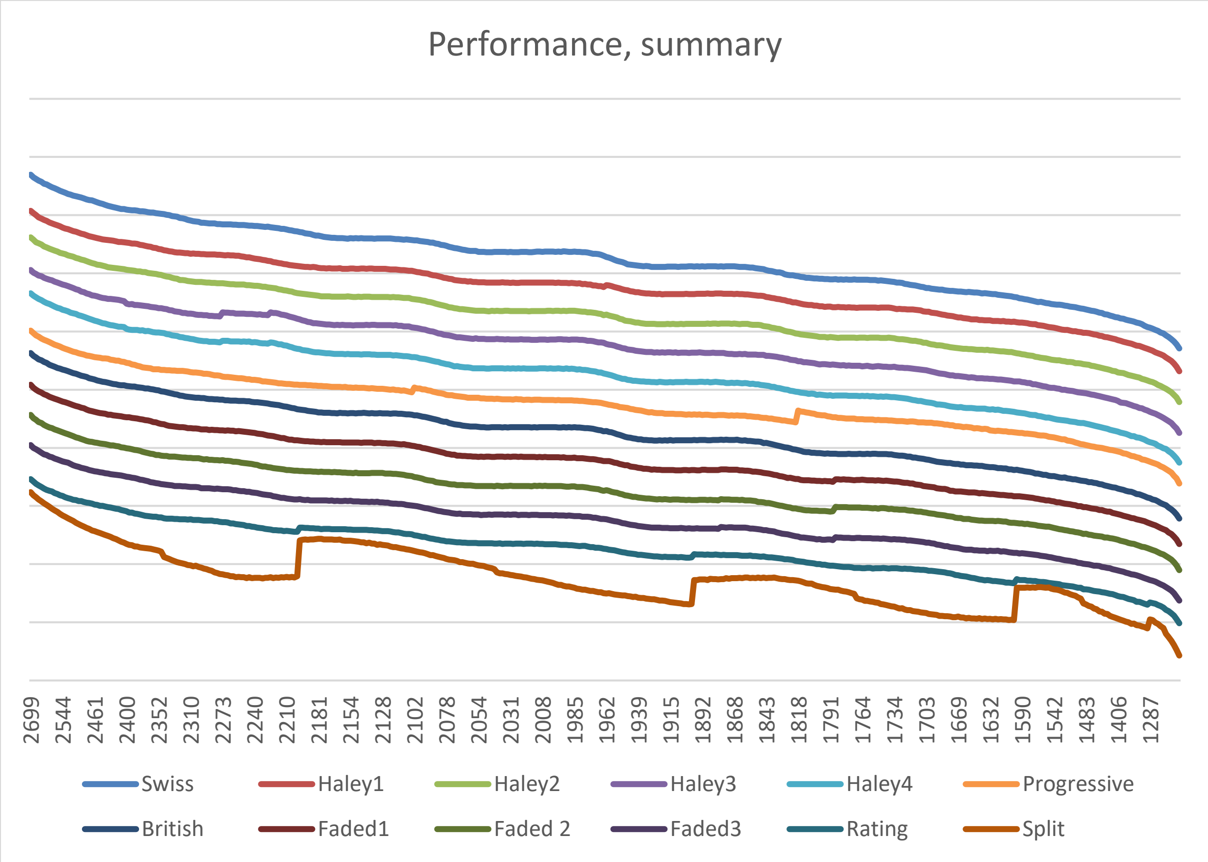
In the long run you want a score that is related to rating performance.



In a simulation of a tournament with 53 players with rating equally spaced from 2000 down to 960 the solid graph shows the expected scored of the player. The dotted graphs show the expected score ± one δ. The graph is based on 10000 simulations. For a low number of players, the Swiss system performs very well.



The same plot with the model1 tournament with 432 players, rating from 2699 down to 1006. The graph is based on 10000 simulations. The Swiss system does not perform very well.



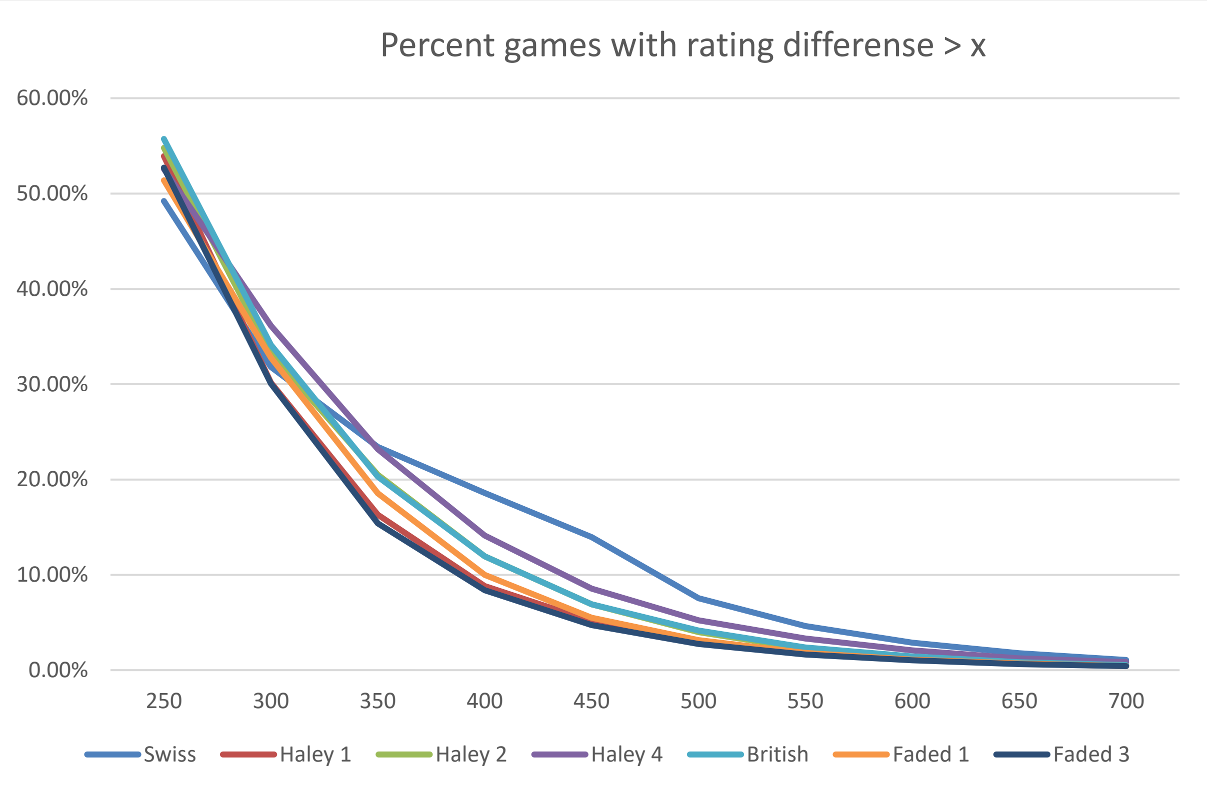
The plot shows the shape of the performance of different methods. Each method has an offset so it's easier to inspect the shapes. A strong requirement to a method must be that if you compare two players the highest rated player is expected to have the highest score. From visual inspection we must conclude that "Split" performs worst in this test, and a method that behaves like this cannot be used in tournaments.

I have extracted two parameters, the sum of positive jump between two adjacent players, and the maximum positive jumps between two players. The last parameter is probably the most important, and I have set all methods with Max jump > 0.1 to bad, and have not done further tests on these methods.

Method	Sum pos jump	Max pos jump	Good
Swiss	0.33	0.02	Yes
Haley1	0.41	0.07	Yes
Haley2	0.37	0.02	Yes
Haley3	0.59	0.14	No
Haley4	0.54	0.05	Yes
Progressive	0.84	0.41	No
British	0.39	0.02	Yes
Faded 1	0.45	0.09	Yes
Faded 2	0.55	0.17	No
Faded 3	0.52	0.10	Yes
Rating	0.70	0.15	No
Split	4.35	1.33	No

Uninteresting games

The problem with Swiss tournaments is the large number of uninteresting games in the first round(s). In the model tournament we have 432 players with a huge span in rating, and we must expect to have at least 50% of the games played with rating difference > 250 regardless what methods we are using.

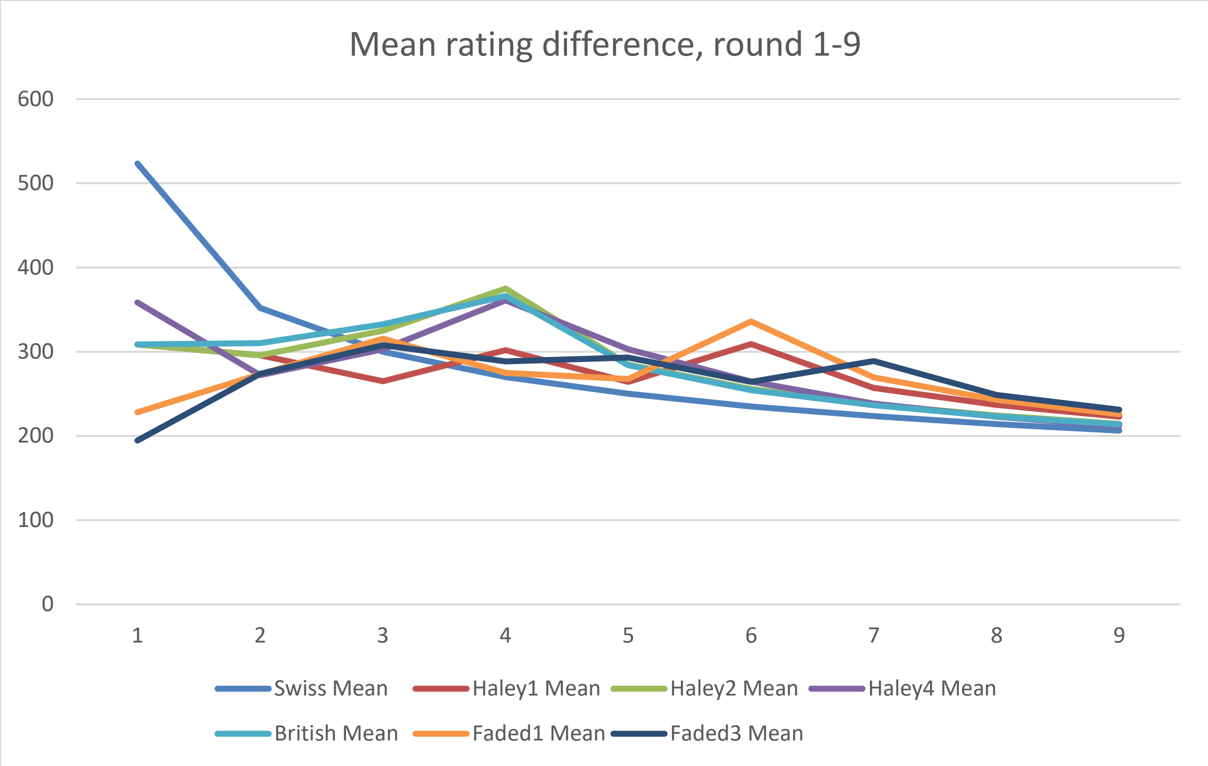


	Swiss	Haley 1	Haley 2	Haley 4	British	Faded 1	Faded 3
200	73.88 %	81.46 %	79.79 %	74.75 %	79.67 %	76.44 %	76.15 %
250	49.21 %	53.92 %	54.79 %	52.57 %	55.72 %	51.38 %	52.72 %
300	31.80 %	30.17 %	33.21 %	36.13 %	34.12 %	32.78 %	30.07 %
350	23.42 %	16.31 %	20.51 %	23.19 %	20.30 %	18.56 %	15.41 %
400	18.58 %	8.81 %	11.95 %	14.11 %	11.95 %	9.98 %	8.39 %
450	13.94 %	5.13 %	6.91 %	8.57 %	6.92 %	5.49 %	4.74 %
500	7.53 %	3.07 %	4.00 %	5.22 %	4.14 %	3.13 %	2.74 %
550	4.61 %	1.82 %	2.25 %	3.32 %	2.39 %	1.82 %	1.63 %
600	2.89 %	1.16 %	1.36 %	2.07 %	1.43 %	1.11 %	1.02 %
650	1.78 %	0.71 %	0.82 %	1.25 %	0.84 %	0.71 %	0.64 %
700	1.07 %	0.44 %	0.53 %	0.85 %	0.53 %	0.43 %	0.42 %
750	0.60 %	0.24 %	0.32 %	0.51 %	0.32 %	0.28 %	0.27 %
800	0.37 %	0.15 %	0.22 %	0.34 %	0.21 %	0.17 %	0.16 %

“Faded 3”, “Haley 1” and “Faded1” scores best in this test

Rating difference

This parameter in almost the same as “uninteresting games”. The difference is that this parameter finds the overall rating mean difference.



Method	Mean opponent	Max opponent
Swiss	286	523
Haley1	273	309
Haley2	280	375
Haley4	282	361
British	281	366
Faded 1	270	336
Faded 3	266	308

Mean opponent is overall mean of all games. Max opponent is the maximum of mean opponent per round. “Faded 3” and “Haley 1” performs best in this test. “Faded 1” is also good.

Expectation factor

If you win a game, you will expect to meet a stronger player in the next round. If you lose a game you will expect to meet a weaker player. The expectation factor measures the percent score for the expectation.

Method	Expectation
Swiss	90.4 %
Haley1	86.2 %
Haley2	85.5 %
Haley4	85.2 %
British	85.0 %
Faded 1	84.0 %
Faded 3	84.8 %

All accelerated methods are in the same level ~ 5% below Swiss pairing

GM / IM Norms

I have calculated the number of norms that we may expect for the different methods. This is done for the Model1 tournament, but with the norm distribution as in Politiken Cup 2015.

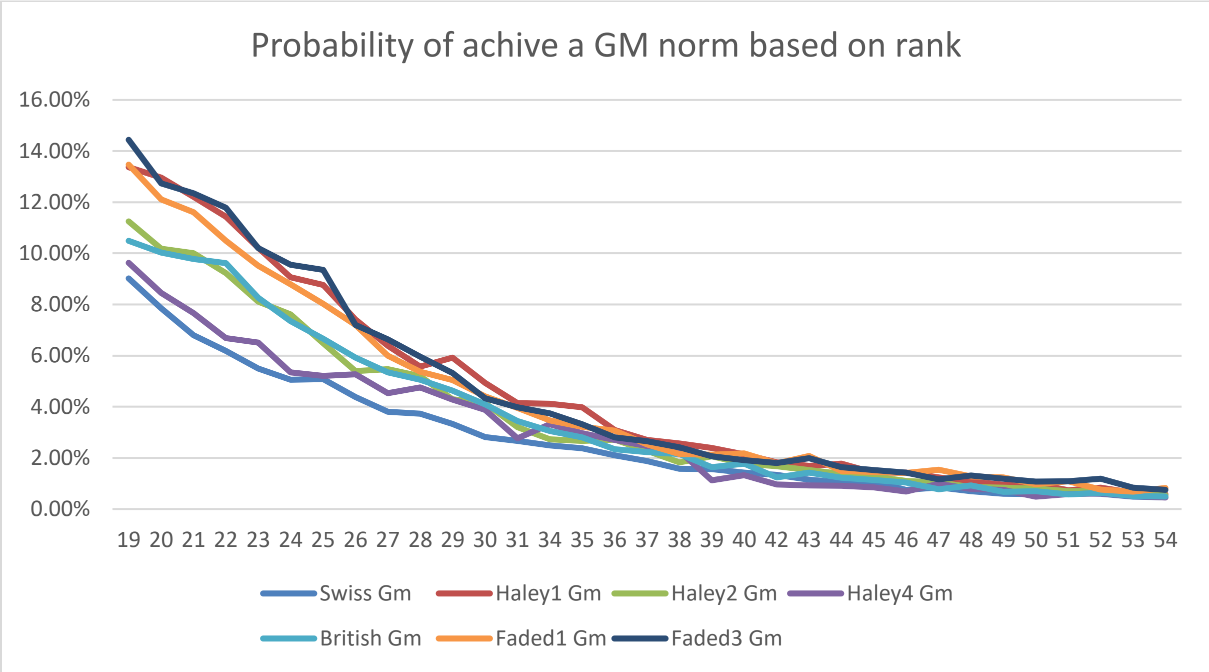
Method	GM	IM
Swiss	0.46	2.65
Haley1	0.72	3.17
Haley2	0.58	2.87
Haley4	0.53	2.62
Britain	0.58	2.93
Faded 1	0.72	3.03
Faded 3	0.74	3.17

“Faded 3” and “Haley 1” performs best in this test. “Faded 1” is also good.

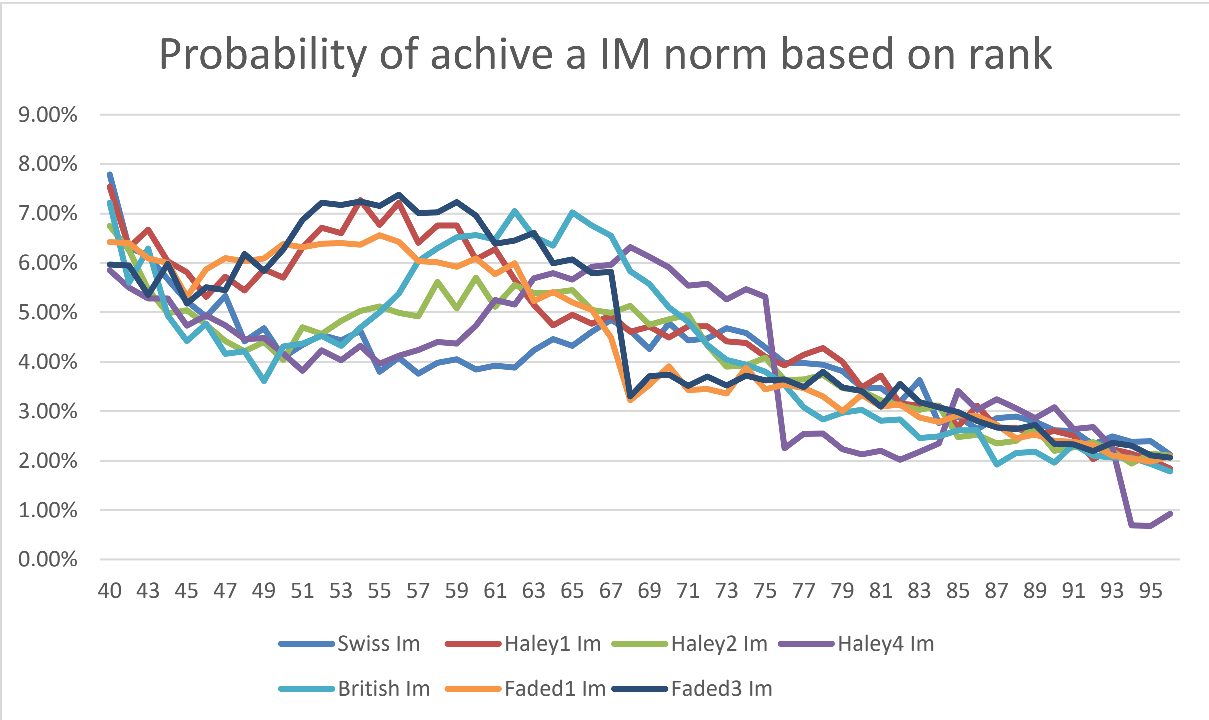
We must also expect that the highest rated players without title are the players with the highest probability to achieve a norm. The analyses for this is not very easy since the title regulations are very unmathematical. In Politiken Cup 2015 the player ranked as number 312 is a FM. Thus player number 100 will meet him in round 1 in a Swiss tournament.

To do this a little bit more mathematical I defined top 1/16 of the players as GM, the next 1/16 as IM and the third/16 FM and the reran the calculations.

For my test tournament player 1-27 is GM, 28-54 is IM and 55-81 is FM. For all players I tested the probability of getting norms if you were without titles.



Faded3, Haley1 and Faded1 is good for all players.



In the area around the weakest IM and strongest FM (player 54-55) we scan see a huge advantage of Faded3, Haley1 and Faded1. From player 67 to 68 Faded3 drops 2.5%, and from player 75 to 76 Faded3 drops 3%.

Discussion

Downfloaters

If a player with virtual points are downfloated, I think the most correct way to treat this downfloater is to temporarily remove 0.5 or more virtual points from the player such that the score bracket is treated as a homogenious score bracket.

Summary and conclusion

A lot of methods has been tested. Some simple, and some quite complicated. The British has behaved very close to Haley 2, and I guess they have the same properties.

I think the Faded systems may perform very well, and I am sure it is possible to improve the parameters. Faded3 had some abnormalities in the norm test. The Haley1 methods worked so well and are so simple that I will recommend Haley 1 as the method to use.

If this method is adopted as a FIDE system, my suggestion is to name it "**Baku acceleration**"

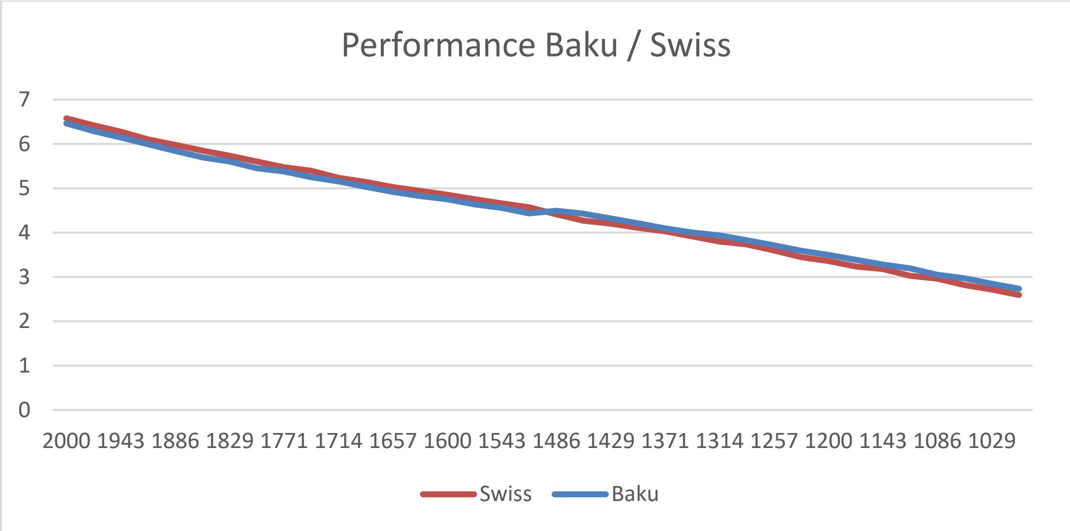
The Baku acceleration:

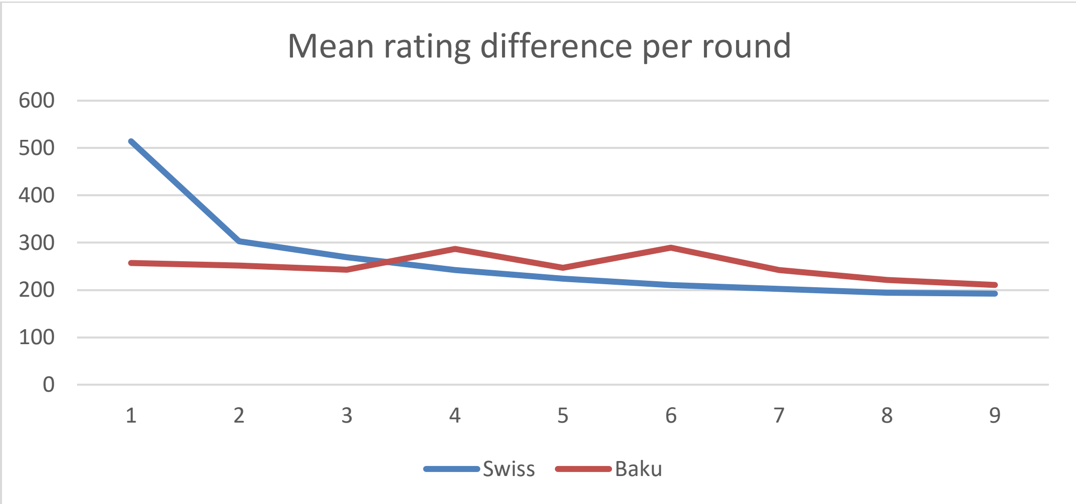
Divide the players in two equally sized subgroups A and B. If the number of players is odd, the number of players in A is rounded down.

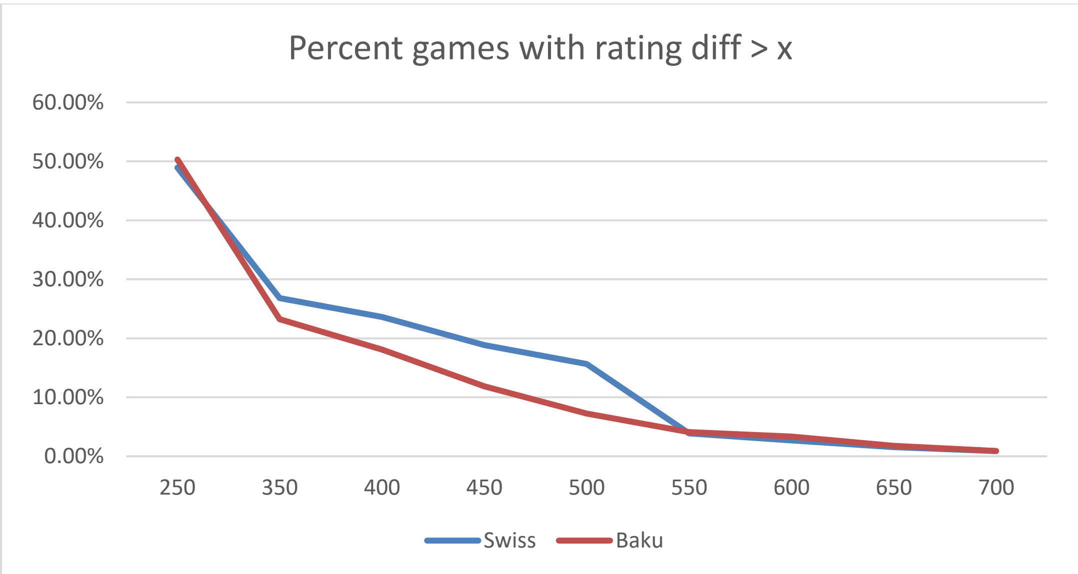
Add 1 virtual point to all players in A in the first 3 round, and ½ virtual point to all players in A in the next 2 rounds.

Test of Baku (Haley-1) acceleration for other tournaments

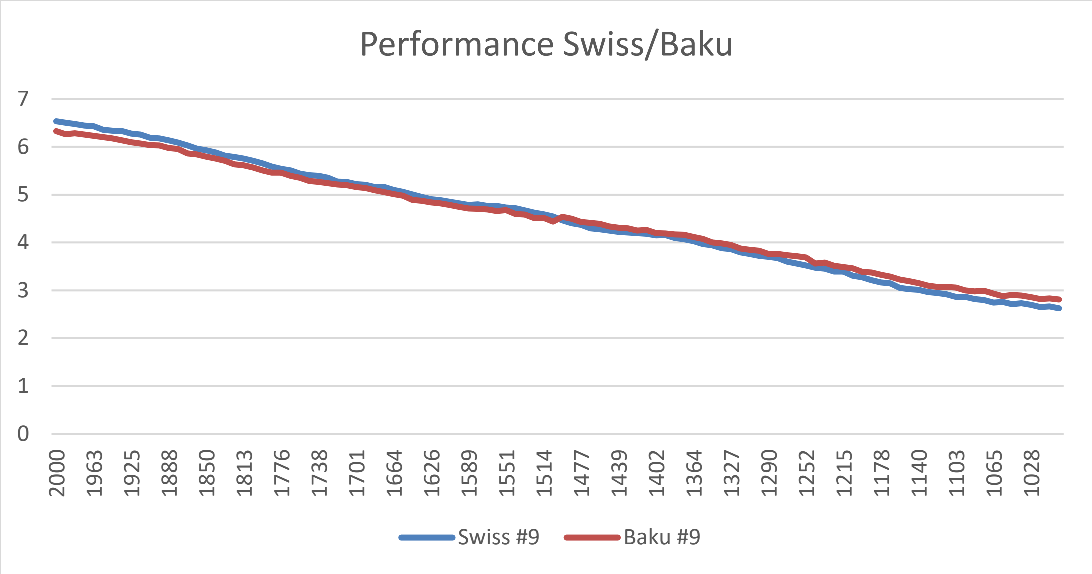
36 players range from 1000 to 2000, linear distribution

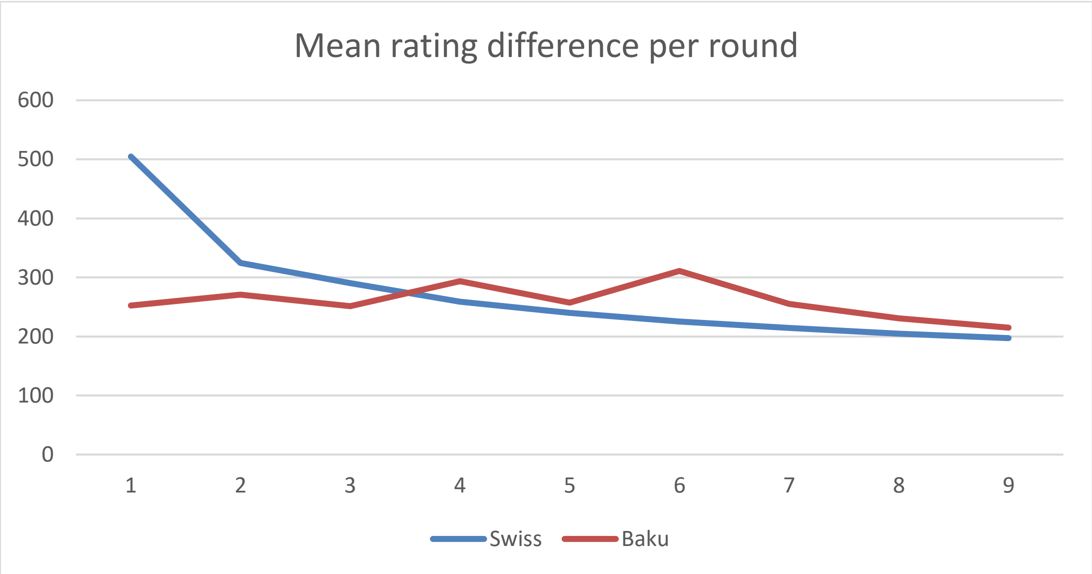


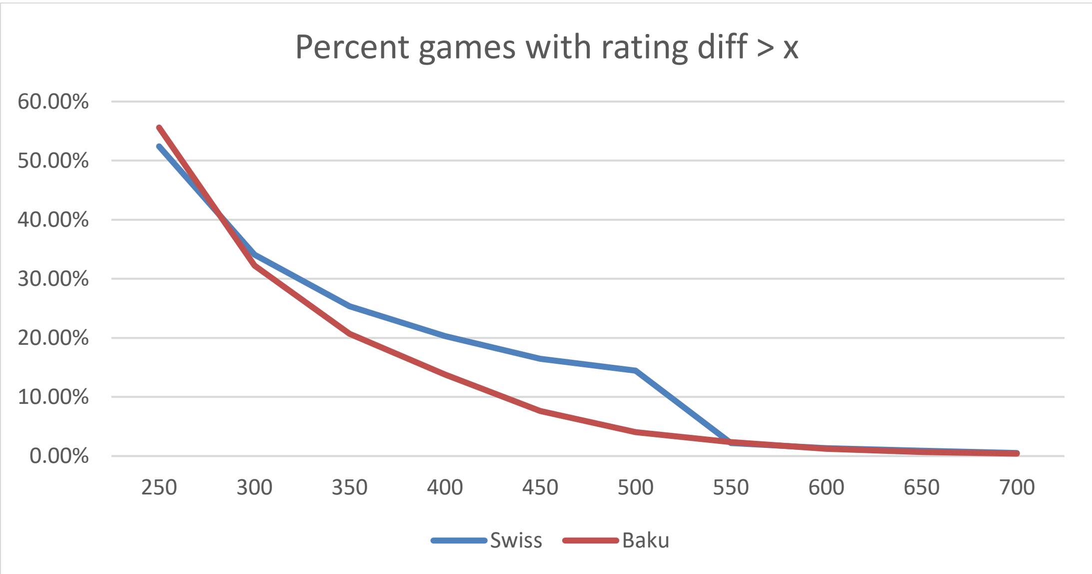




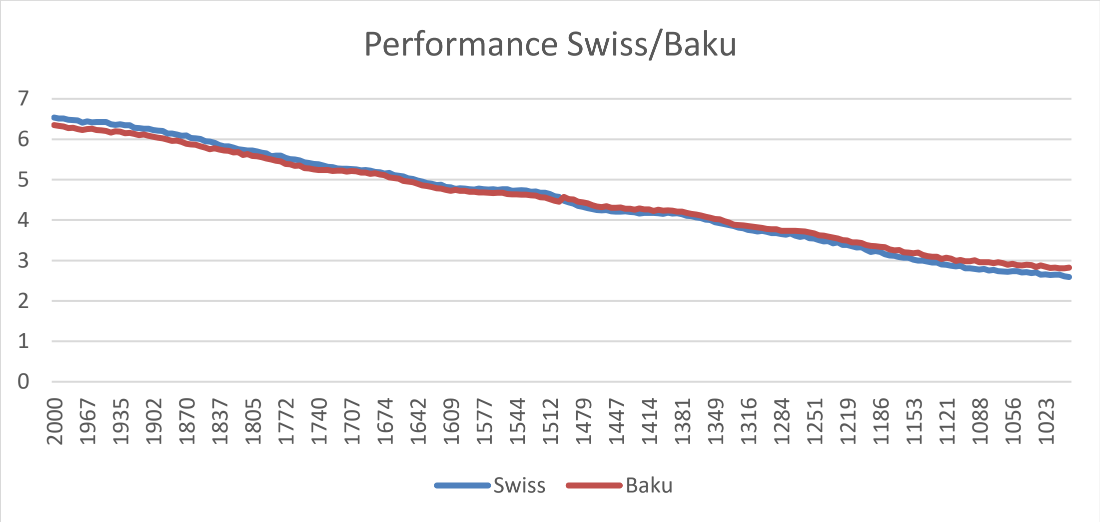
108 players range from 1000 to 2000, linear distribution

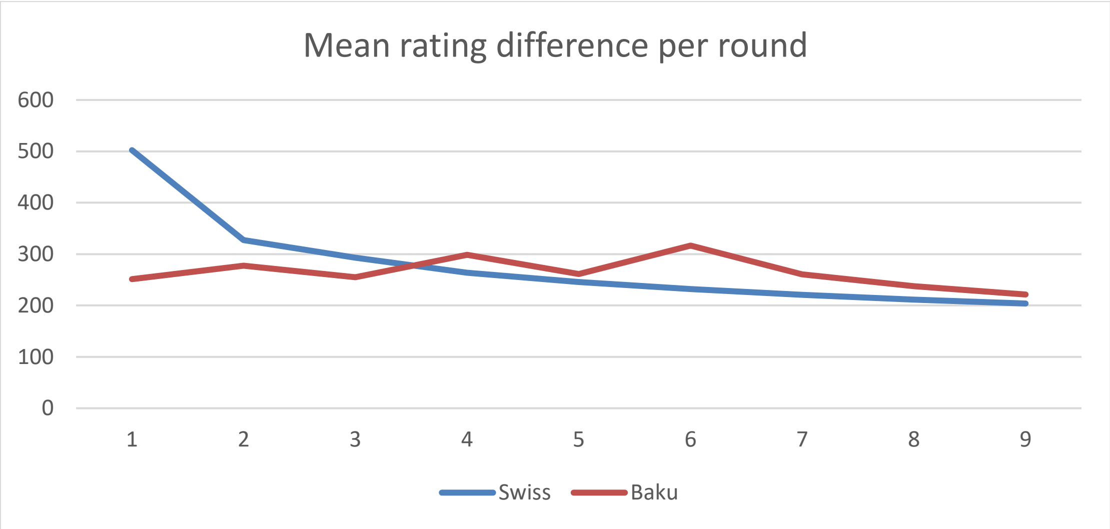


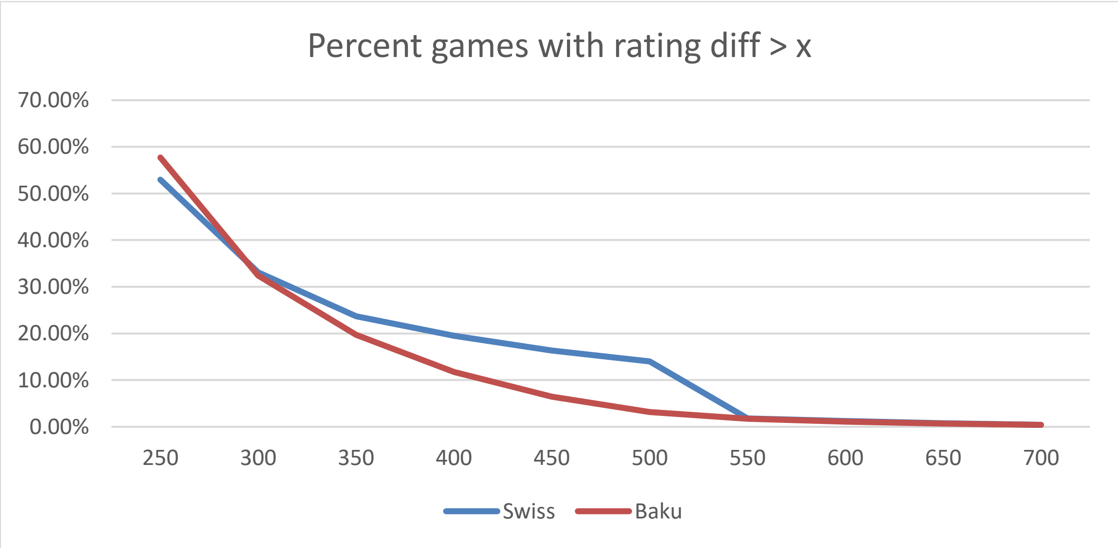




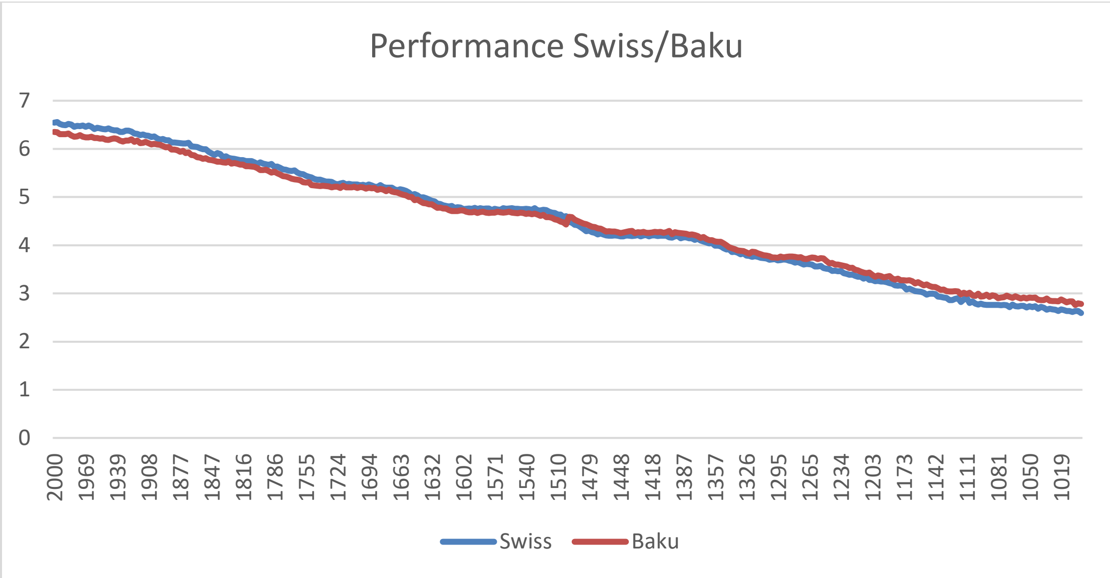
216 players range from 1000 to 2000, linear distribution

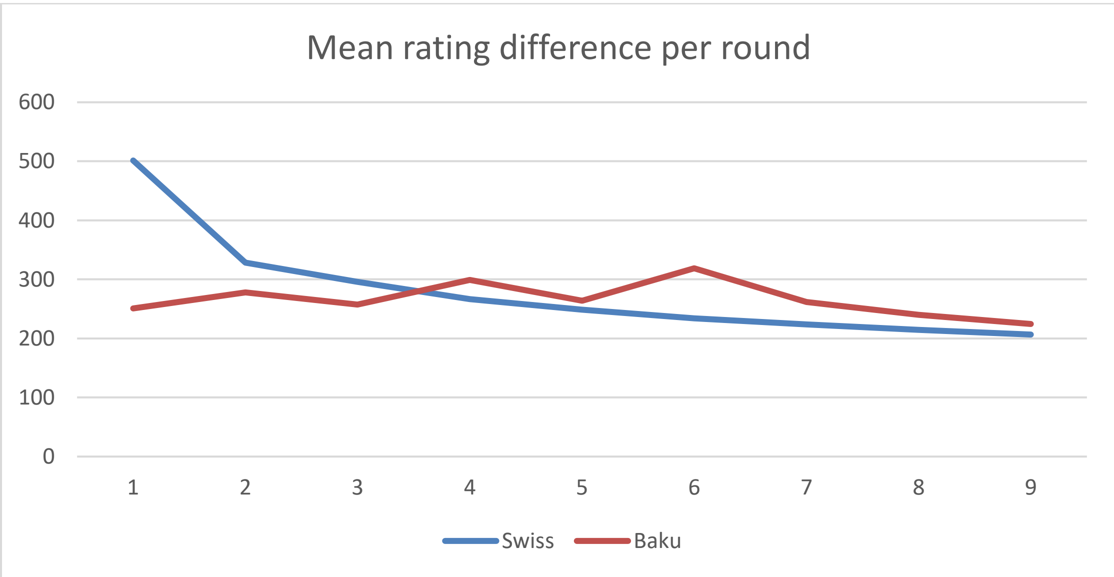


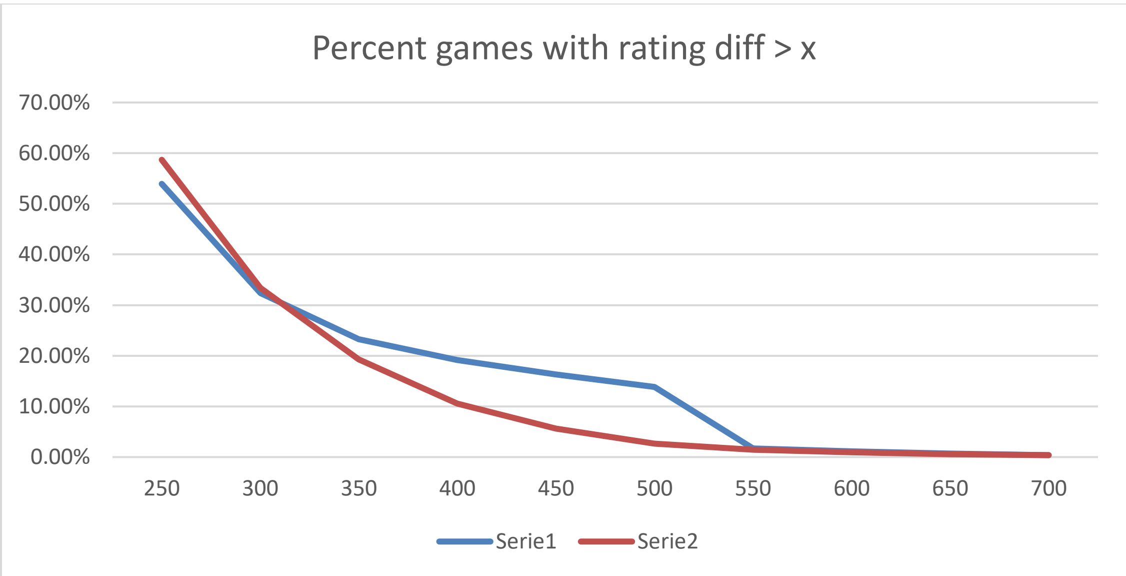




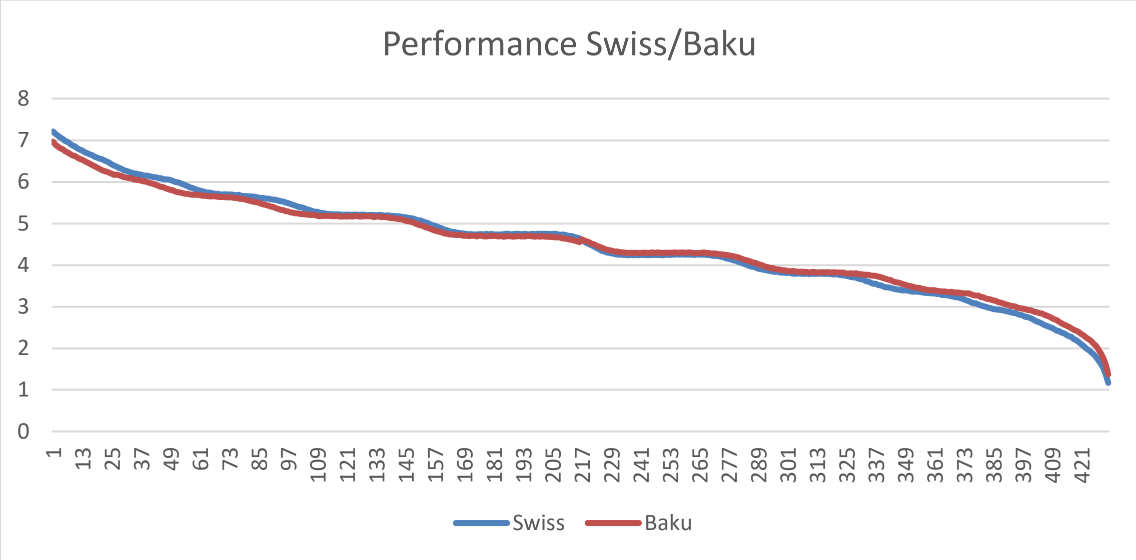
380 players range from 1000 to 2000, linear distribution

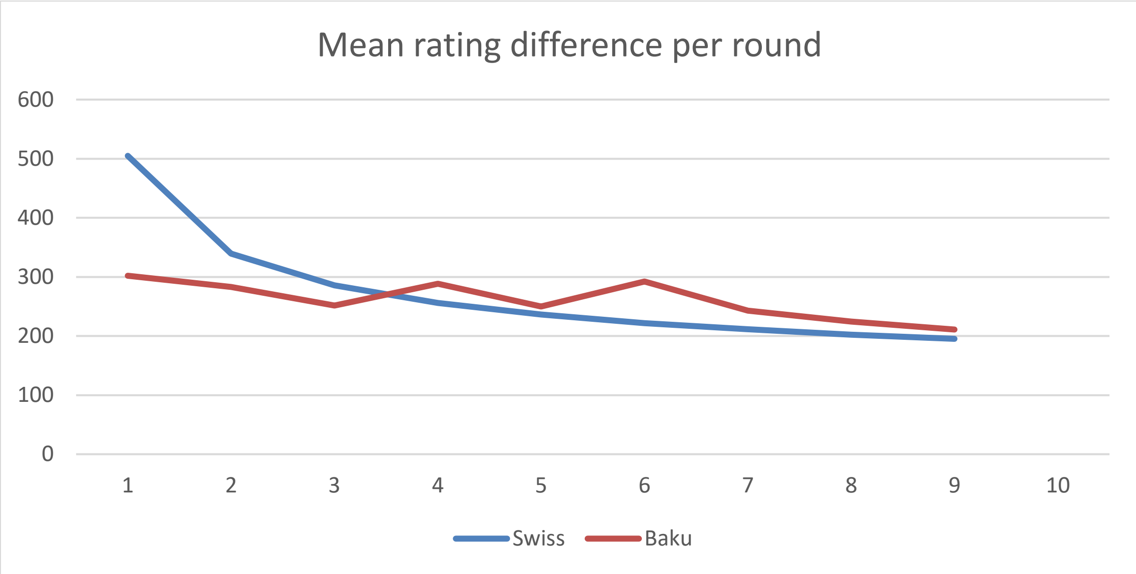


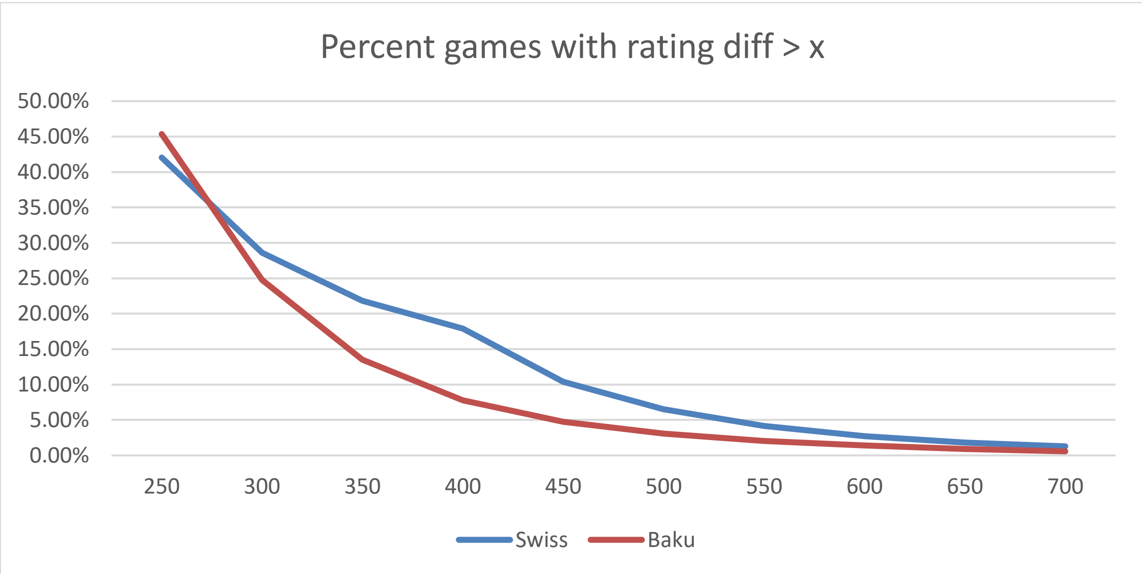




432 players range from 1000 to 2770, Gaussian model2

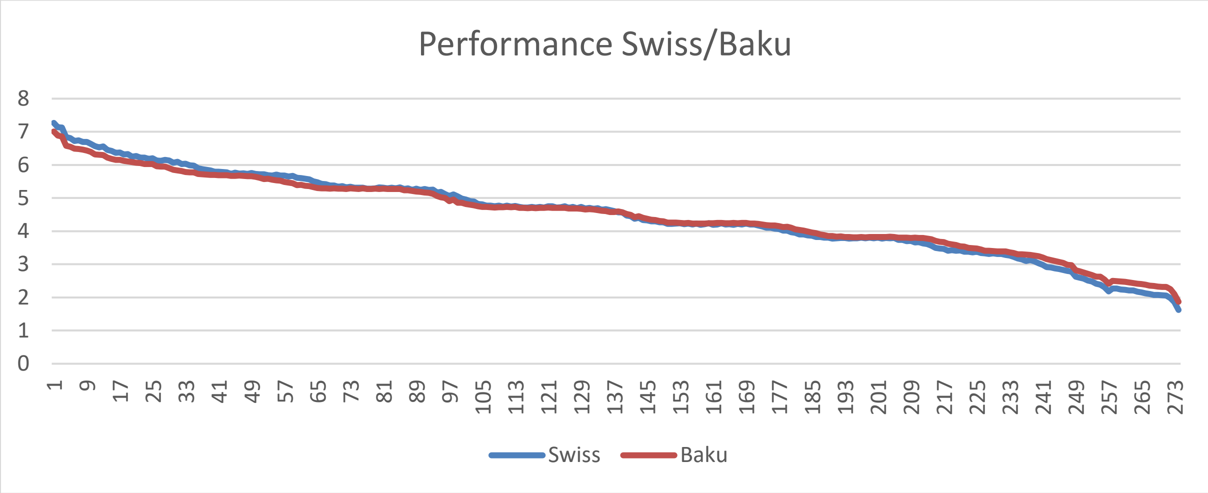


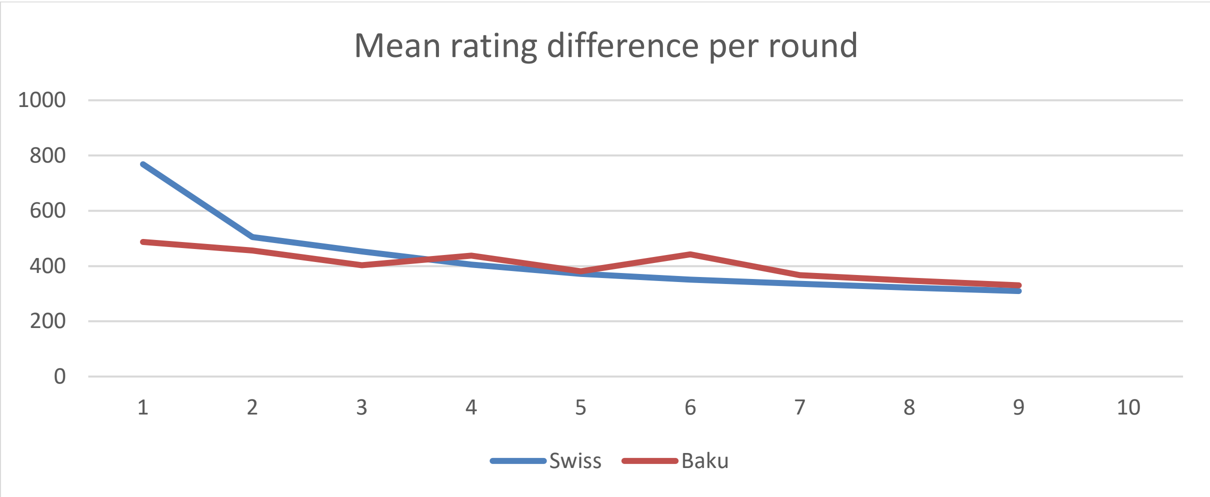


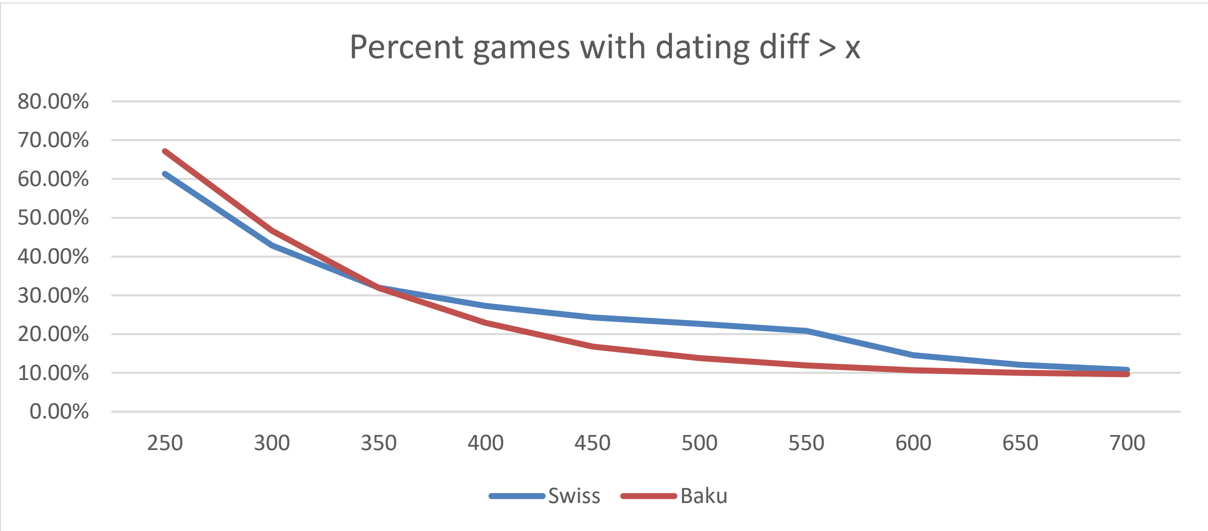


Test of Baku (Haley-1) acceleration for real tournaments

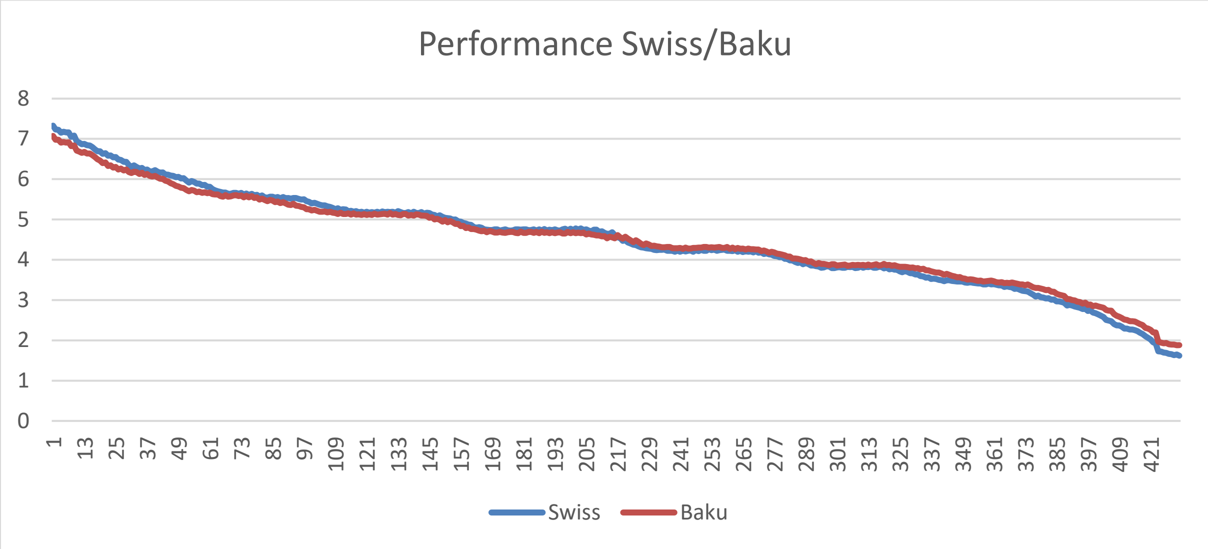
Reykjavik Open 2015, 274 players, rating from 2756 to unrated

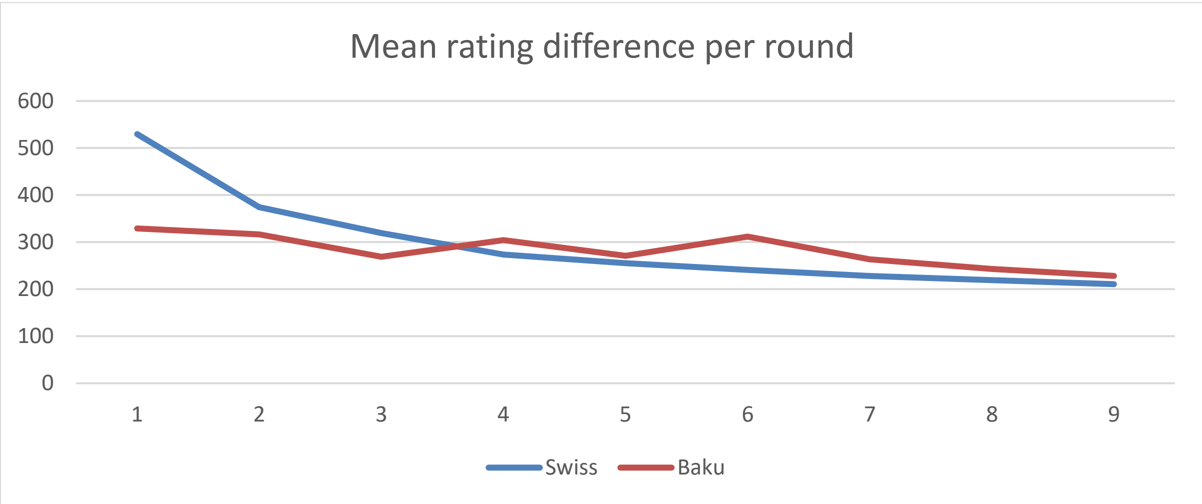


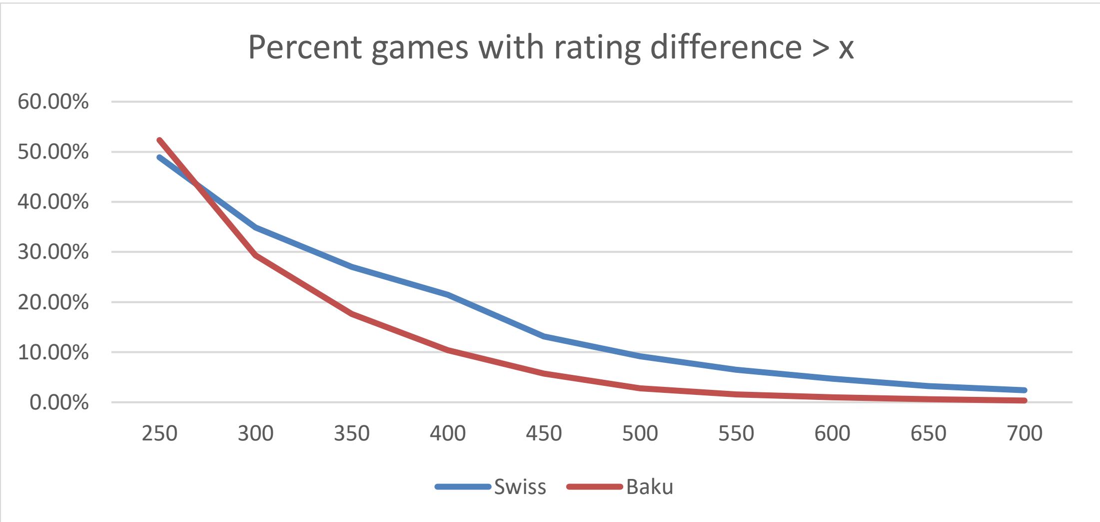




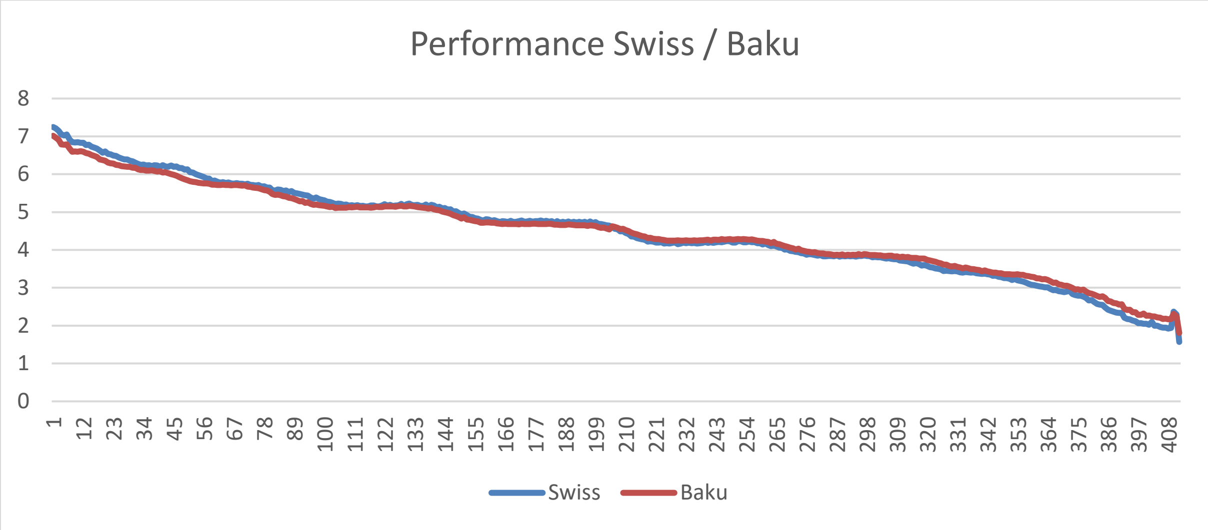
Politiken cup 2015, 432 players, rating from 2707 to unrated

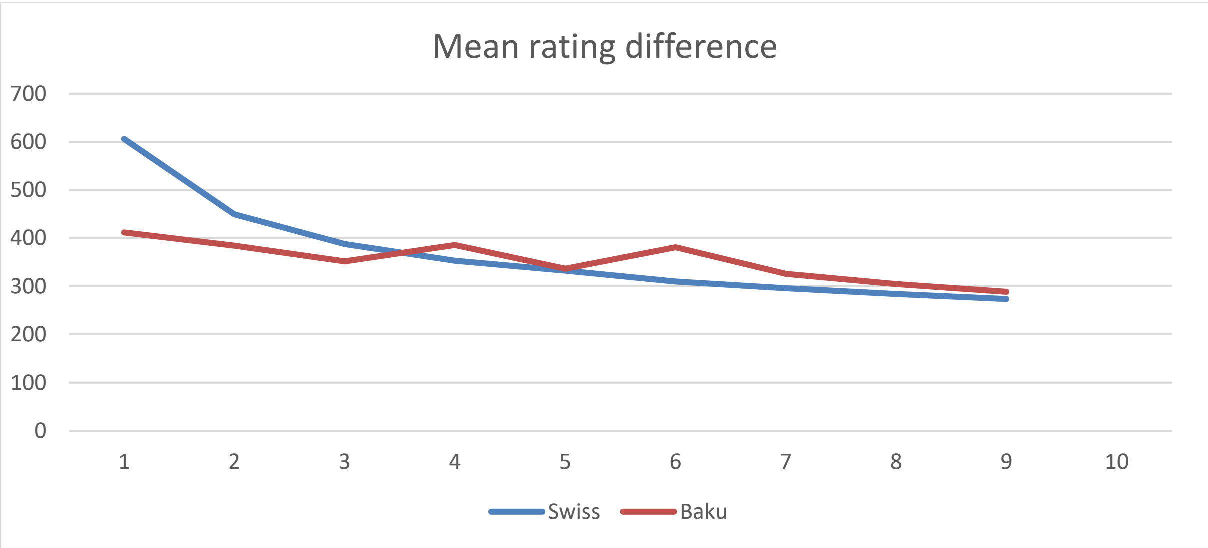


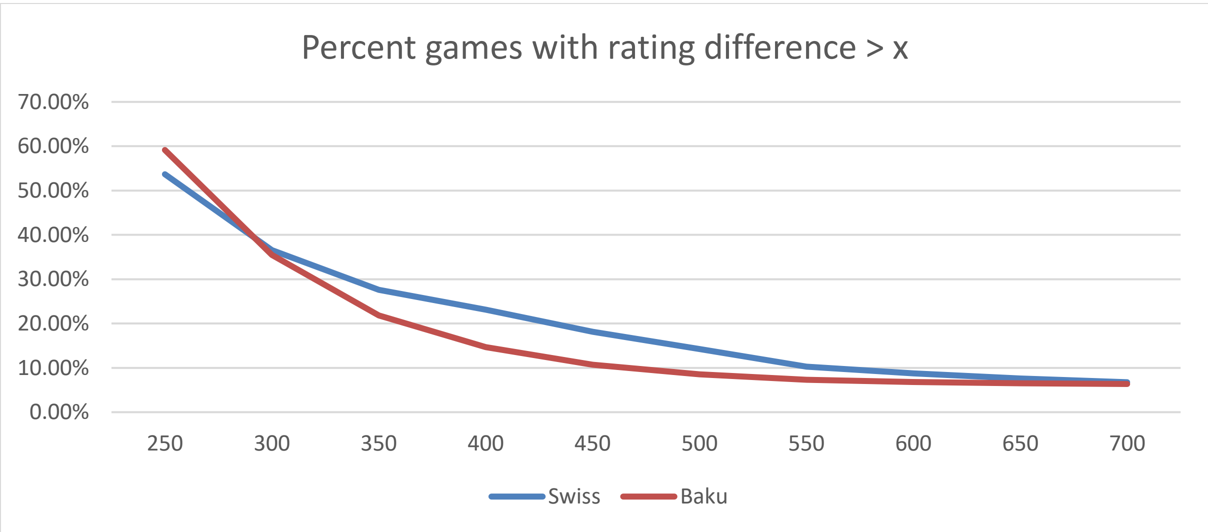




Villa Be Benasque 2015, 412 players, rating from 2678 to unrated







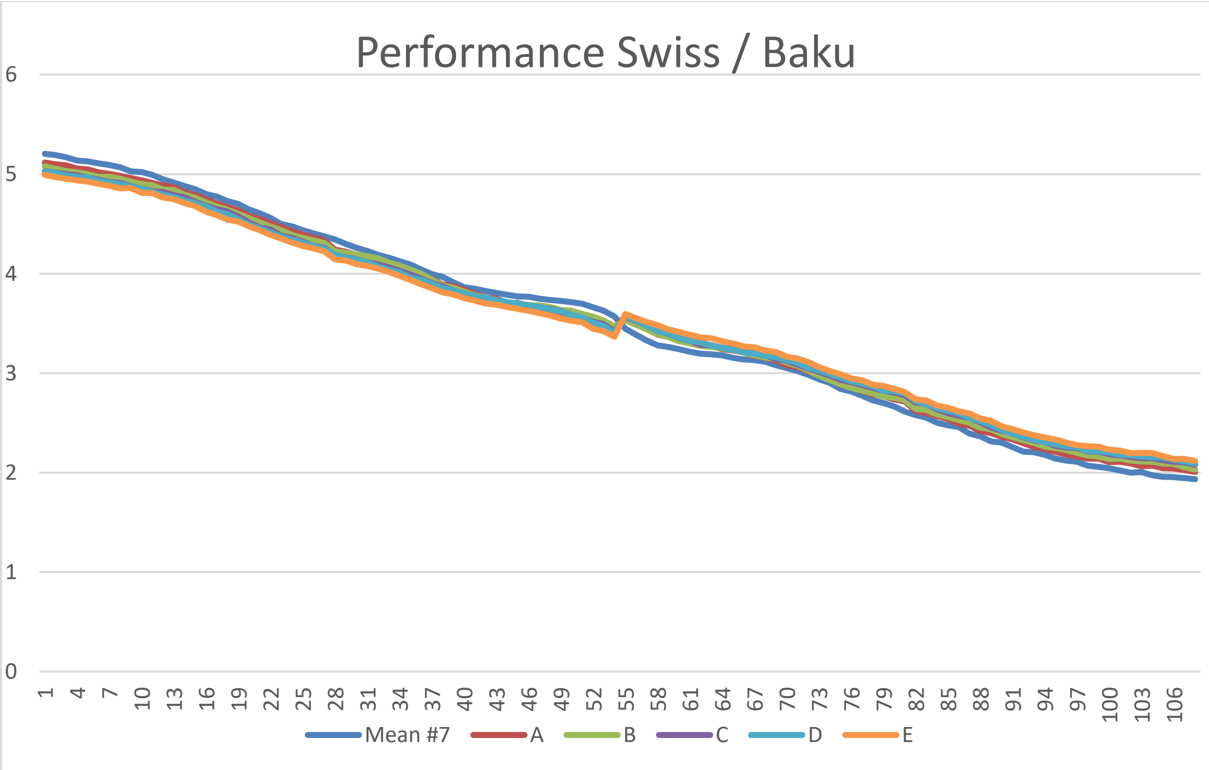
Baku acceleration, 7 and 5 rounds

From the very good results with 9 round, it's time to test algorithm's for 7 round.

Five methods were tested. Acceleration points per round in the tested methods:

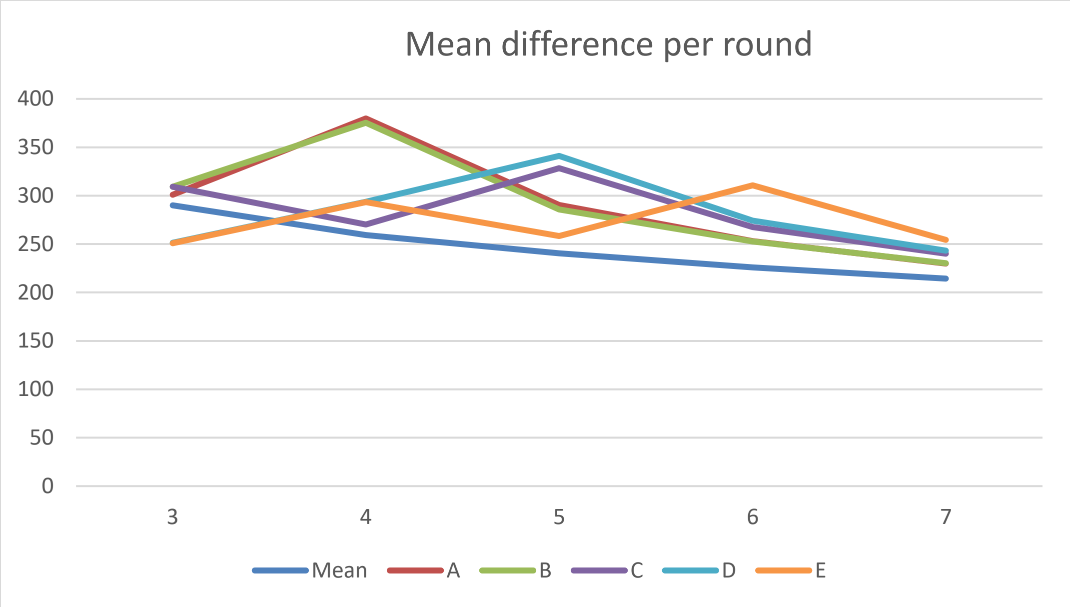
	R1	R2	R3	R4	R5	R6	R7
A	1	0.5	0.5	0	0	0	0
B	1	1	0.5	0	0	0	0
C	1	1	0.5	0.5	0	0	0
D	1	1	1	0.5	0	0	0
E	1	1	1	0.5	0.5	0	0

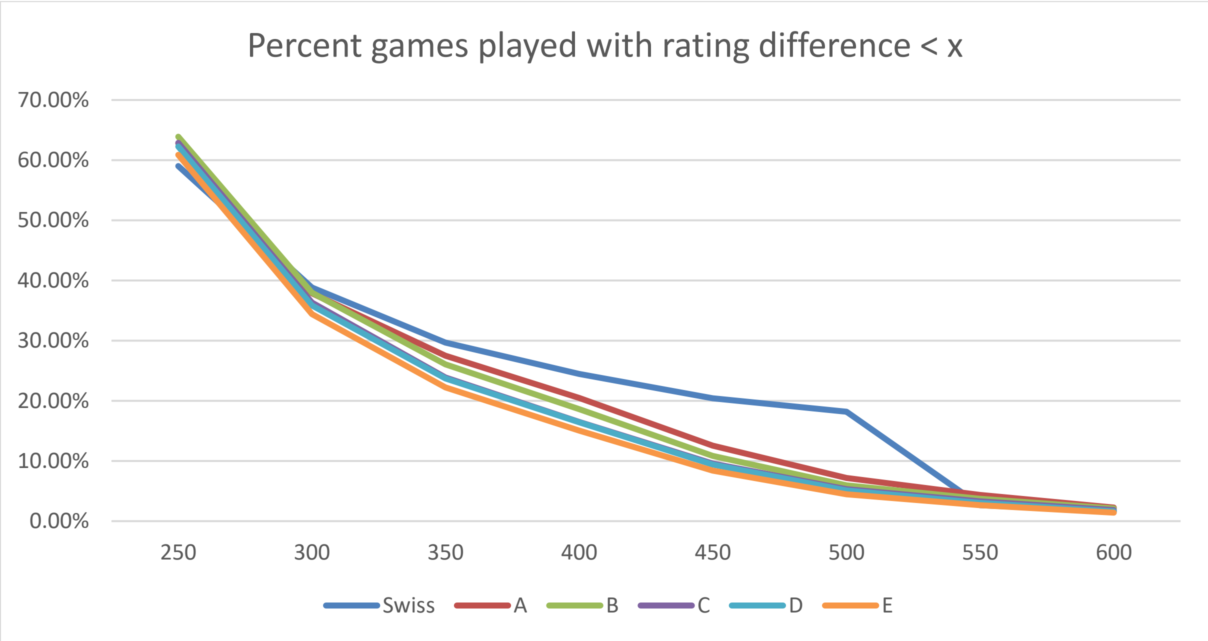
108 players range from 1000 to 2000, linear distribution



All methods have a small jump between the accelerated and the unaccelerated players:

A	0.12
B	0.06
C	0.15
D	0.18
E	0.23





Summary 7 rounds

The main difference is the number of rounds acceleration. E performs best on the last two tests, but has also the biggest step in test 1. C is the best of C and D (in the last test there are no visible difference) while C is best in test 1 and 2. A and B is OK, but do not perform well compared to the others.

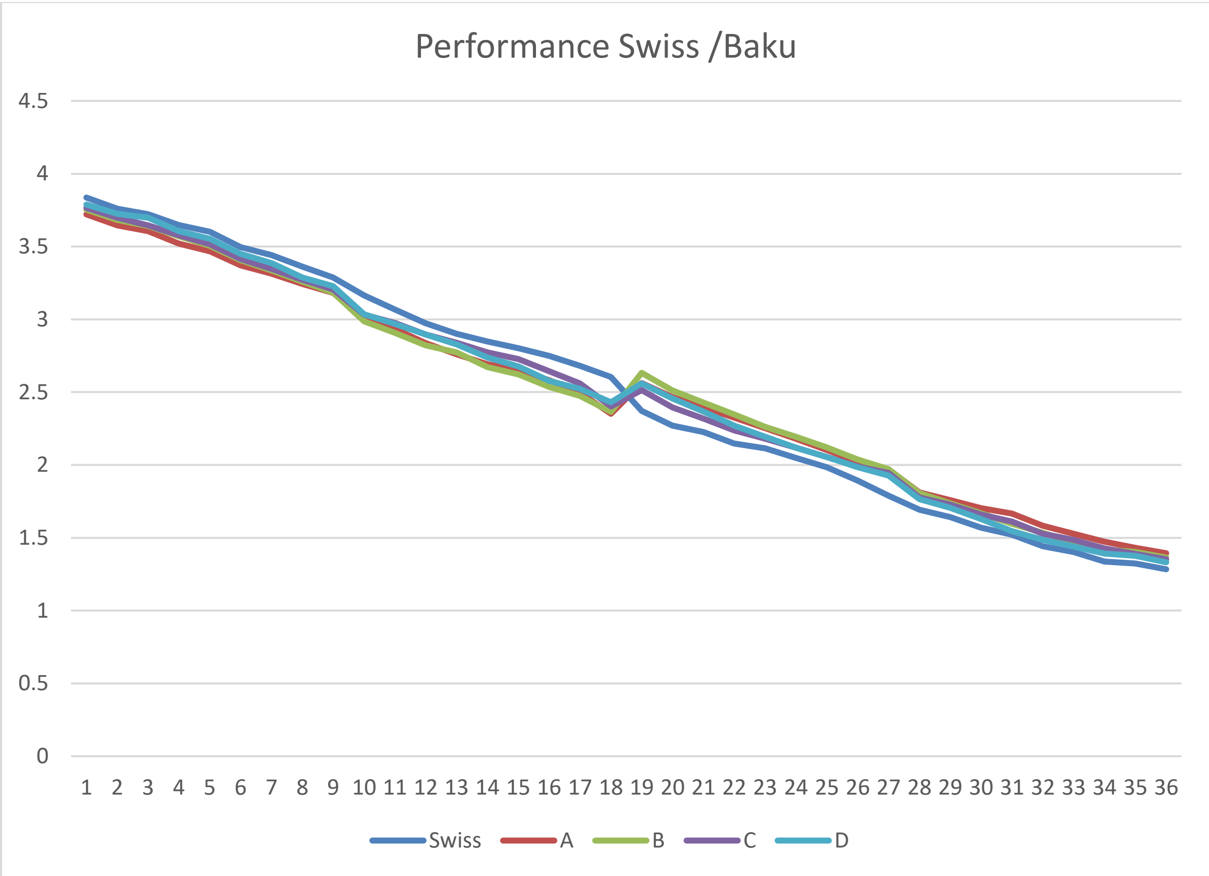
As a tradeoff between stability and performance I will consider C as the best method, 2 rounds acceleration with 1 point and 2 rounds with ½ point.

5 rounds

With 5 rounds four methods were tested. Acceleration points per round in the tested methods:

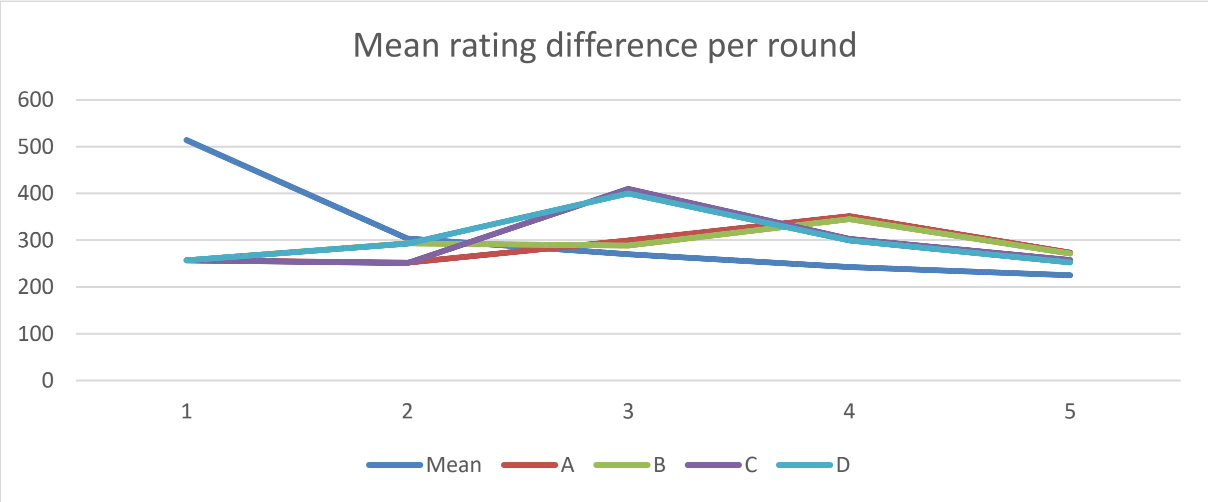
	R1	R2	R3	R4	R5
A	1	1	0.5	0	0
B	1	0.5	0.5	0	0
C	1	1	0	0	0
D	1	0.5	0	0	0

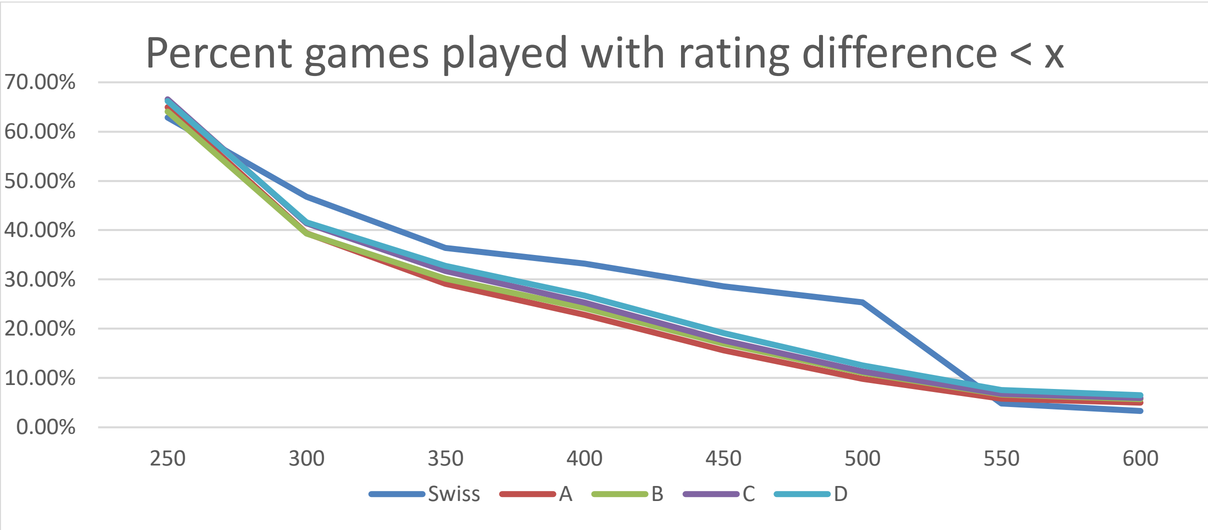
36 players range from 1000 to 2000, linear distribution



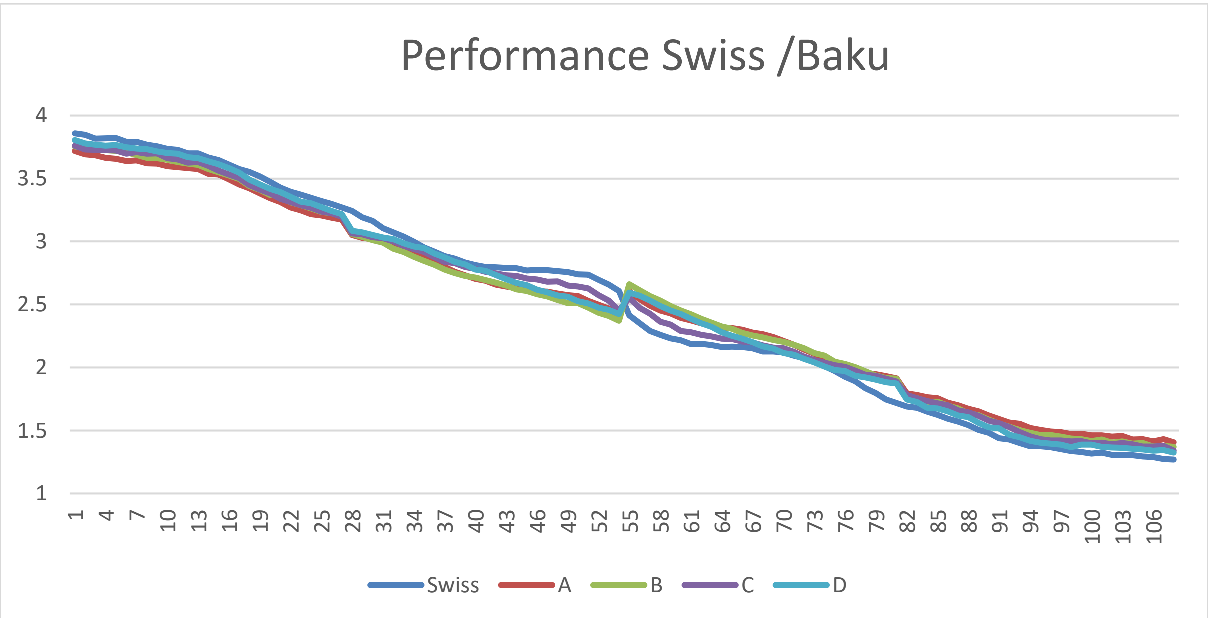
All methods have a small jump between the accelerated and the unaccelerated players:

A	0.21
B	0.27
B	0.11
D	0.13





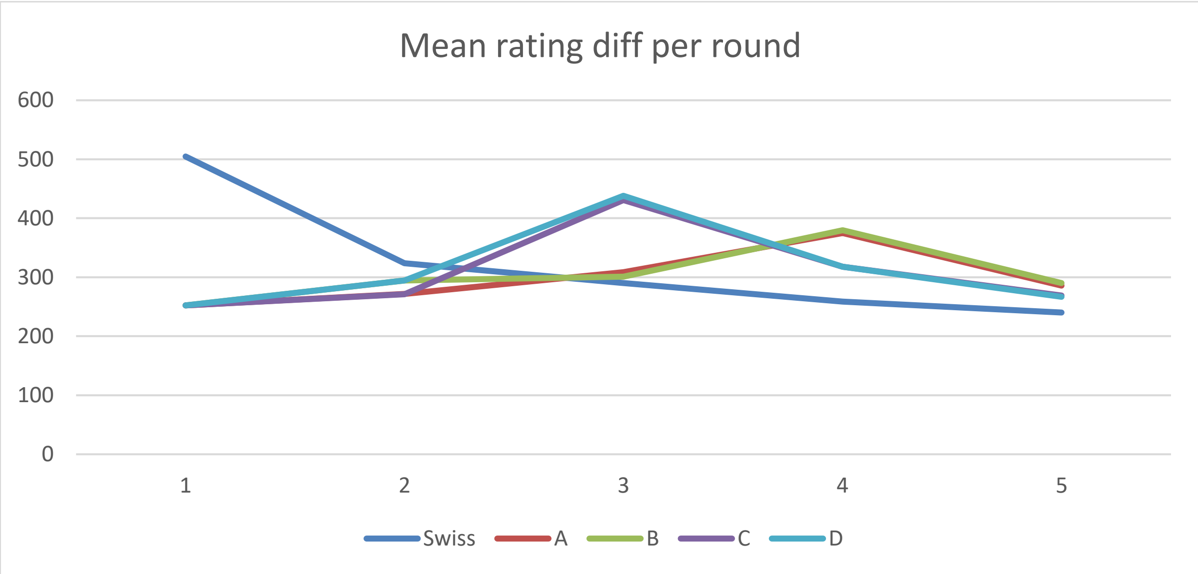
108 players range from 1000 to 2000, linear distribution

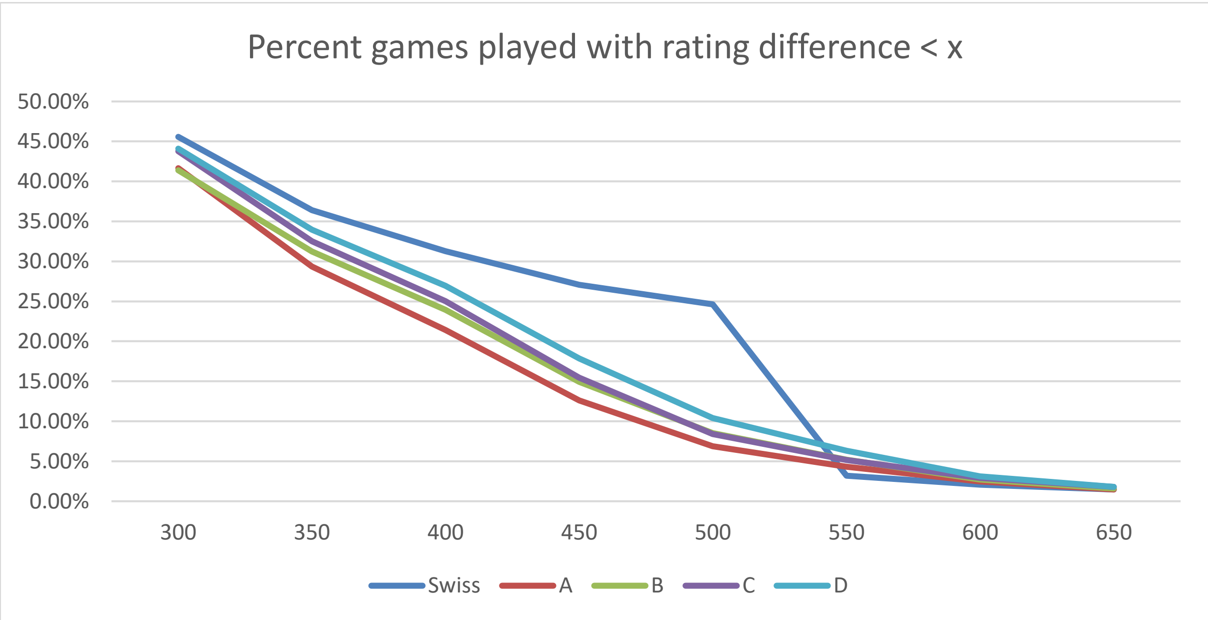


All methods have a small jump between the accelerated and the unaccelerated players:

A	0.21
B	0.29
C	0.10
D	0.17

Also standard Swiss pairing suffers in this area and drops 0.31 points from player 54 to 57.





Summary 5 rounds

The main difference is the number of rounds acceleration. This is again a tradeoff between stability and performance. Despite the gap on 0.21 points between accelerated and unaccelerated players I will consider A as the best method, 2 rounds acceleration with 1 point and 1 rounds with ½ point.

General acceleration scheme

1. Number of accelerated round shall be 50% of the rounds rounded up.
2. In 50% of the accelerated rounds (rounded up) the accelerated players shall receive 1 point, and in the rest of the accelerated rounds 0.5 point.

Rounds	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13
5	1	1	0.5	0	0								
6	1	1	0.5	0	0	0							
7	1	1	0.5	0.5	0	0	0						
8	1	1	0.5	0.5	0	0	0	0					
9	1	1	1	0.5	0.5	0	0	0	0				
10	1	1	1	0.5	0.5	0	0	0	0	0			
11	1	1	1	0.5	0.5	0.5	0	0	0	0	0		
12	1	1	1	0.5	0.5	0.5	0	0	0	0	0	0	
13	1	1	1	1	0.5	0.5	0.5	0	0	0	0	0	0

Conclusion

The Baku acceleration is tested with different distribution models for playing strength and different numbers of players. The method is stable and reliable. The method shows a significant reduction of games played with rating difference > 350, and it also nearly double the probability for achieving GM or IM norms.

As far as I have analyzed the data the Swiss Dutch + Baku acceleration will work better than normal Dutch Swiss for all types of tournaments with 9 rounds or more.

Appendix A, Experiences so far, sep 2017

German U25 Youth Chamionship

Hi SPPC friends,

> we are considering using the Baku Accelerated System for the Open
> German Youth Championship U25 in June.

I just wanted to let you know that we have indeed applied the Baku Accelerated System in the two groups of our German U25 Youth Championship. In fact, the A-group (Event code: 160838) had 160 players and was therefore a very good fit for a practical test. The B-group took place with 86 participants.

We used Swiss-Chess for the pairings generation, using its ability to handle special points, which we adjusted after R3 and R5 according to the Baku rules. The official tournament website gives you a good impression of the pairings:

A: http://www.deutsche-schachjugend.de/2017/odjm-a/

B: http://www.deutsche-schachjugend.de/2017/odjm-b/

Although the website is in German, I think you will easily find the pairings :)

One scenario discussed previously on this mailing list occurred in R3 of the A-tournament:

http://www.deutsche-schachjugend.de/2017/odjm-a/runde/3/

On board 48 there was the game of a player with zero points against another with 1.5 points. The first one had one virtual point and was upfloated, so the pairing is correct after all.

Although we are by no means a title tournament, our experiences with the Baku system were quite good. We will recapitulate this year's experiences in near future, but currently I am optimistic that we use the Baku System next year again.

All the best from Germany,

Falco

Nordstrand GP 2017 – 5 round tournament

Nordstrand GP is an annual tournament, and it's therefore convenient to compare the 2017 to previous tournaments. In the group for players > 1750 rating there were 34 players.

We had 3 round acceleration, 1p, 1p and 1/2p for the 18 highest rated players.

If we compare the results with the 2016 tournament that was held without acceleration, we can list some properties:

If we compare the final results with the results from a similar tournament from 2016 that was not accelerated, we can compare some properties:

Parameter:	NGP 2017	NGP 2016
Number of players	34	36
Average rating, all players	2008	2015
Average of opponents, 3 strongest players	2225	2125
Average of opponents, 3 weakest players	1863	1946
Opponents, best player	9,2,5,4,6	18,14,7,9,2
Opponents, weakest player	24,12,32,33,30	34,32,27,bye,28

The acceleration gives the strong players stronger opponents, and the weak players weaker opponent. We had no complaints from the players, and think this experiment was successful.

This article is a compressed version of a report written by Fridrik Karlson translated from Icelandic

Purpose of the analysis:

On Iceland, there has been some dissatisfaction about how Iceland's open chess tournaments is organized, and with focus on the fact that the tournaments have no lower limit for participation. The first question is about is a waste to open international tournaments for all players, the second question is about having an open tournament have a negative impact on GM/IM norms.

After Reykjavik Open 2017, Gunnar Björnsson, President of the Icelandic Chess Federation, asked Fridrik Karlson to review the pros and cons of accelerated acceleration pairings for next year, and he focused on examining a system with which he has personal experience as a player in Capo d'Orso, the "Progressive Acceleration".

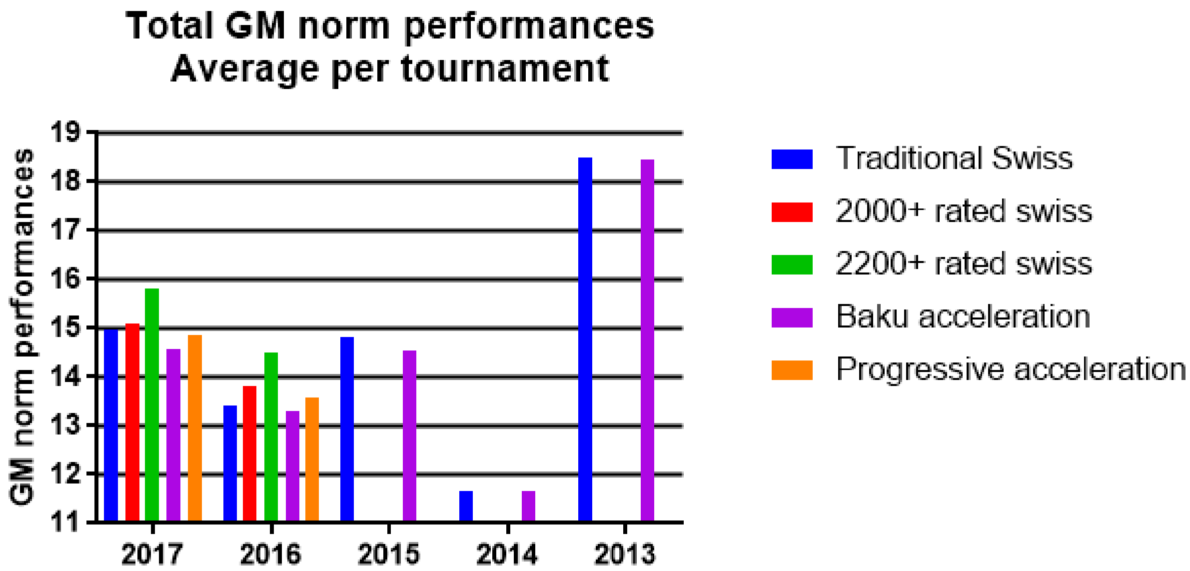
The author then began to look at three other possible formats for the tournament, the FIDE Baku acceleration, traditional Swiss system with a 2000 level minimum, and a 2200-point minimum. All these four systems were compared to the Swiss system that has been used in recent years.

Methods:

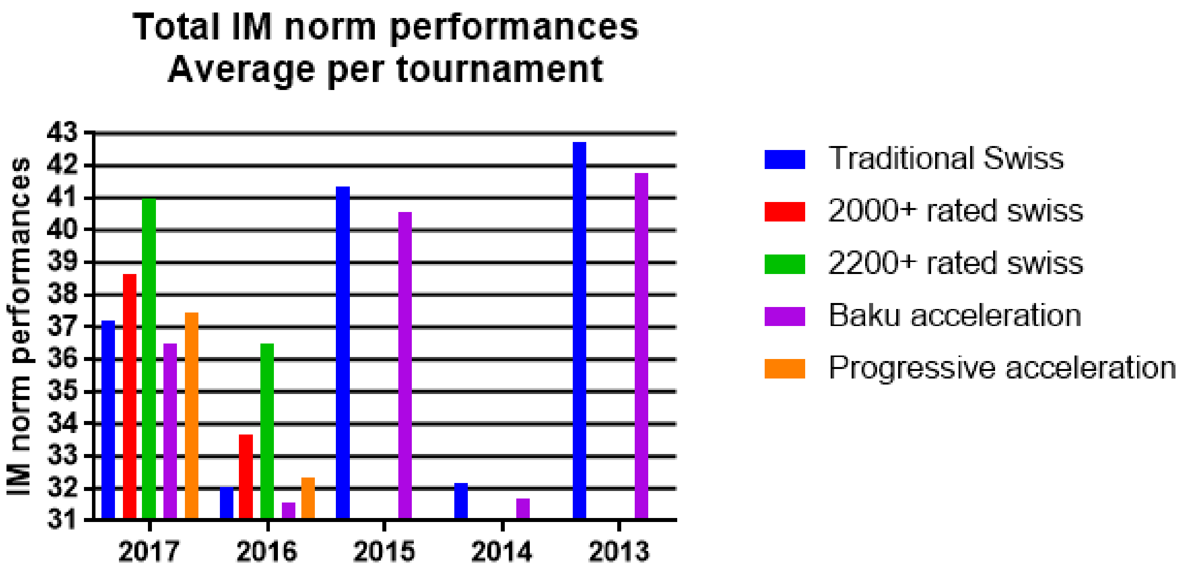
The mathematical model designed by Otto Milvang was used to simulate how these different systems worked. The results from Reykjavik Open for the past 5 years , and the following 16 cases were available with the simulations of 1000 tournaments:

Years					Pairing system
2017	2016	2015	2014	2013	Swiss
2017	2016				Swiss (2000+)
2017	2016				Swiss (2200+)
2017	2016	2015	2014	2013	Baku accelleration
2017	2016				Prorgressive accelleration

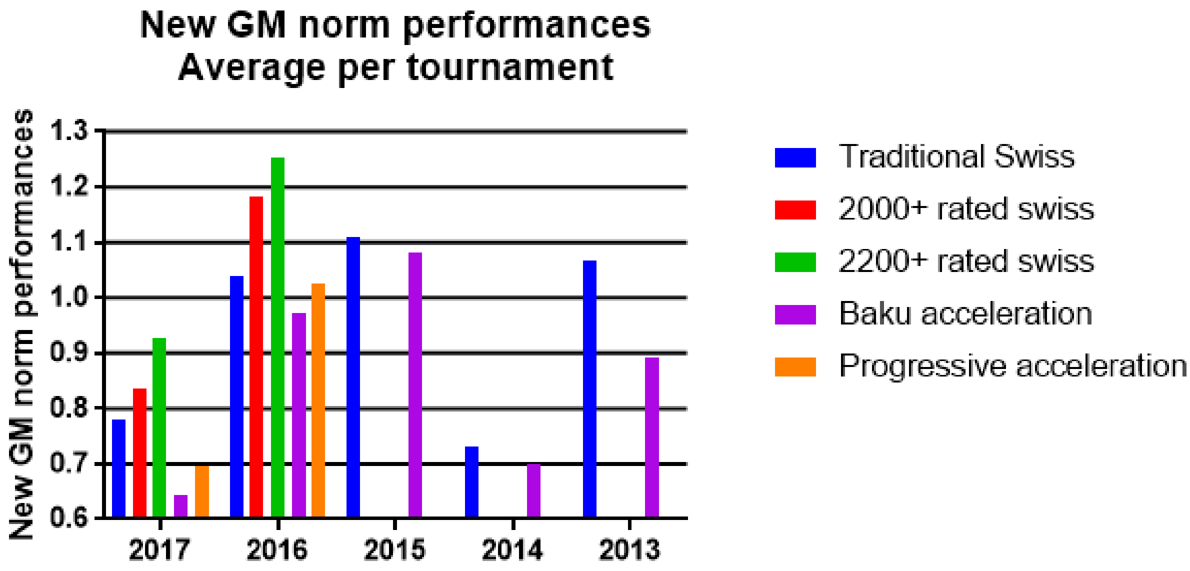
Results:



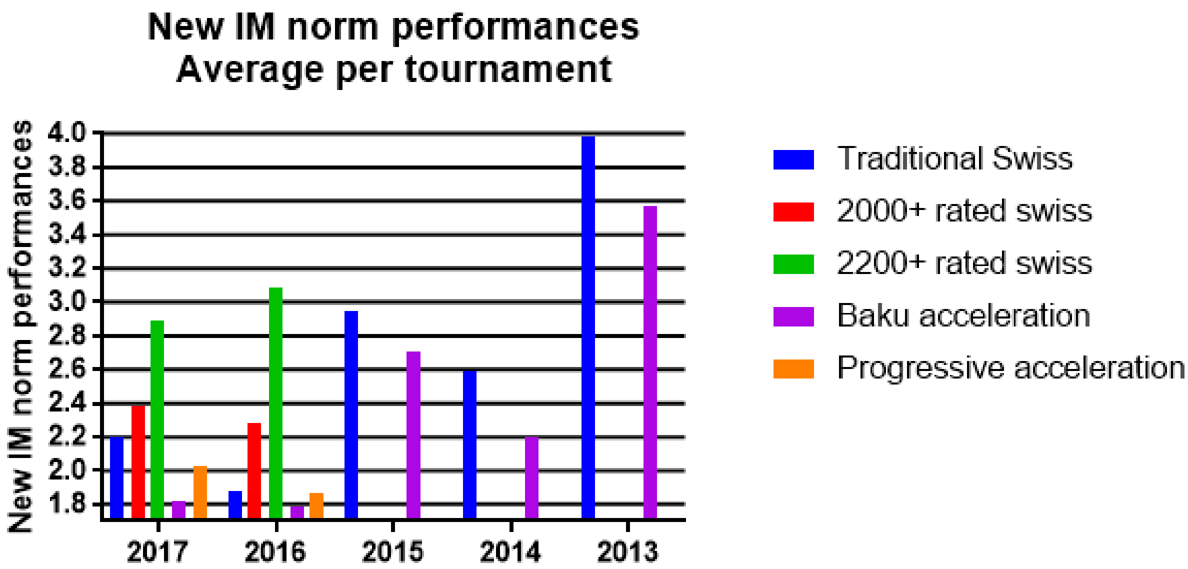
The figure shows how many chess players on average, performing GM norms. The Swiss system for 2200+ limit gives the best result, but the difference is not huge compared to other systems.



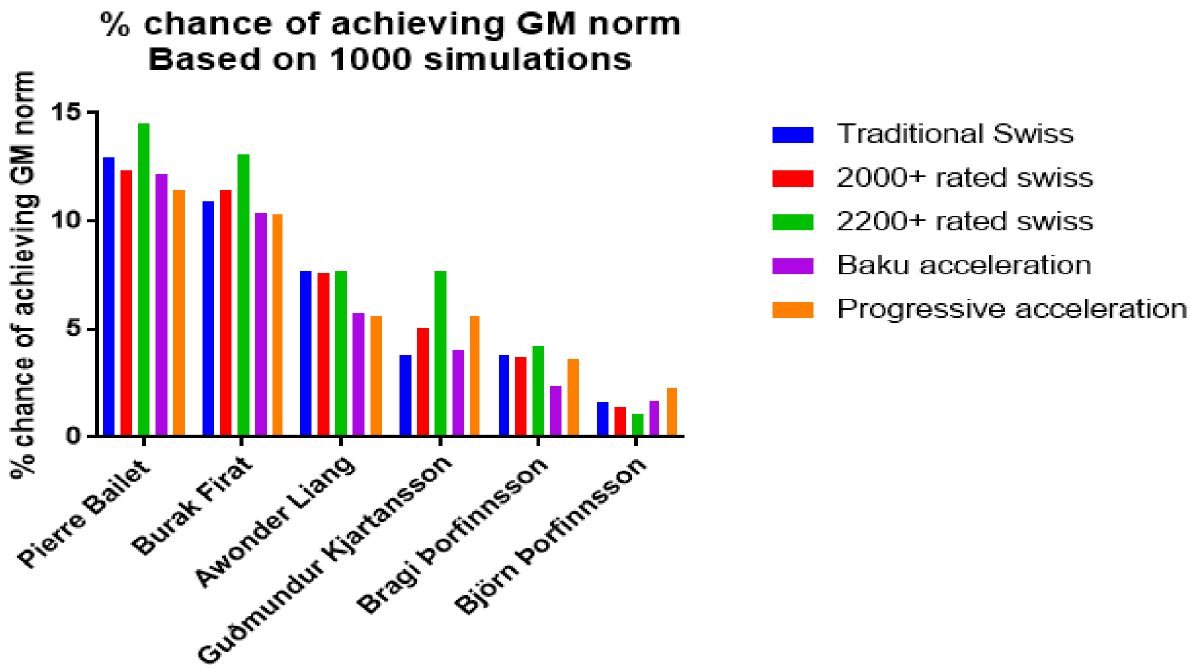
However, the difference is greater when considering how many players performing IM norm. In comparison with 2016 and 2017, a 2200+ limit group will be approx. four more norms than in a traditional Swiss system. Progressive acceleration is similar to conventional Swiss system.



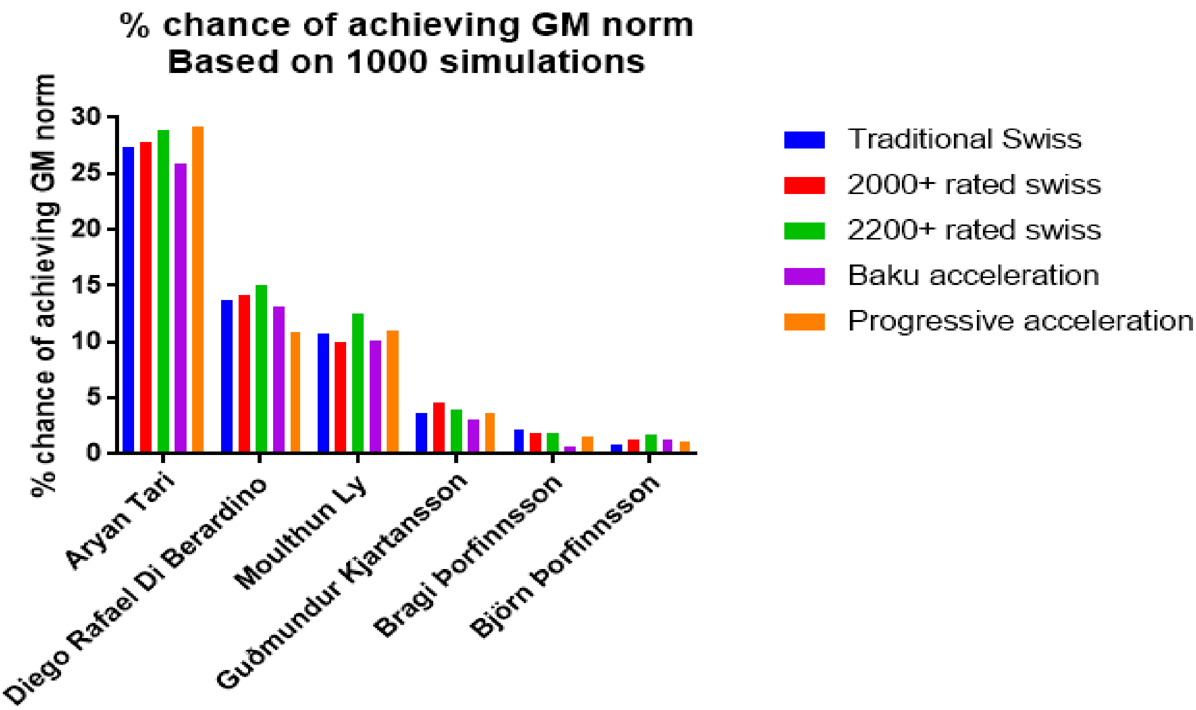
In this figure only those players who do not have a GM title are considered. The figure indicates that no method expects more than approx. one new GM norm in average.



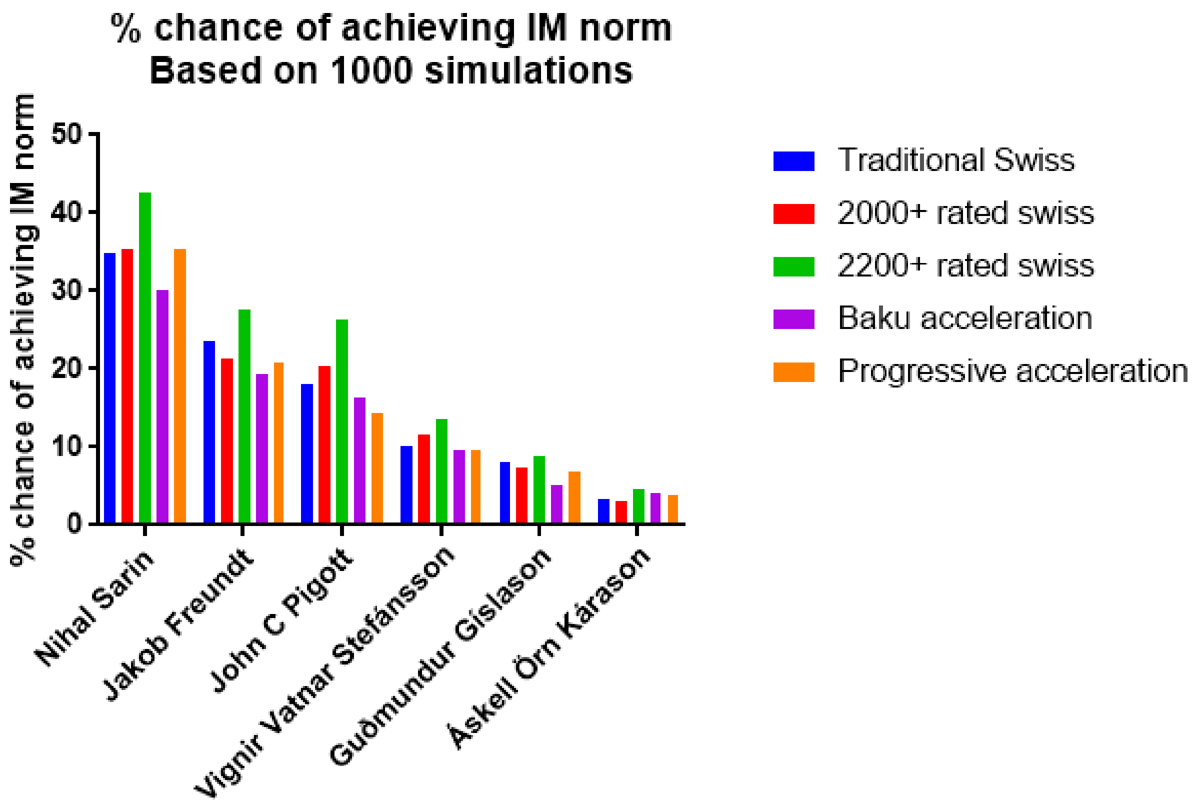
It seems appropriate for the IM norm most to have the tournament closed, and their benefits are considerably higher than the for GM norm. Note that the accelerated systems give no benefit to the standard Swiss system. I will come back with my hypothesis later.

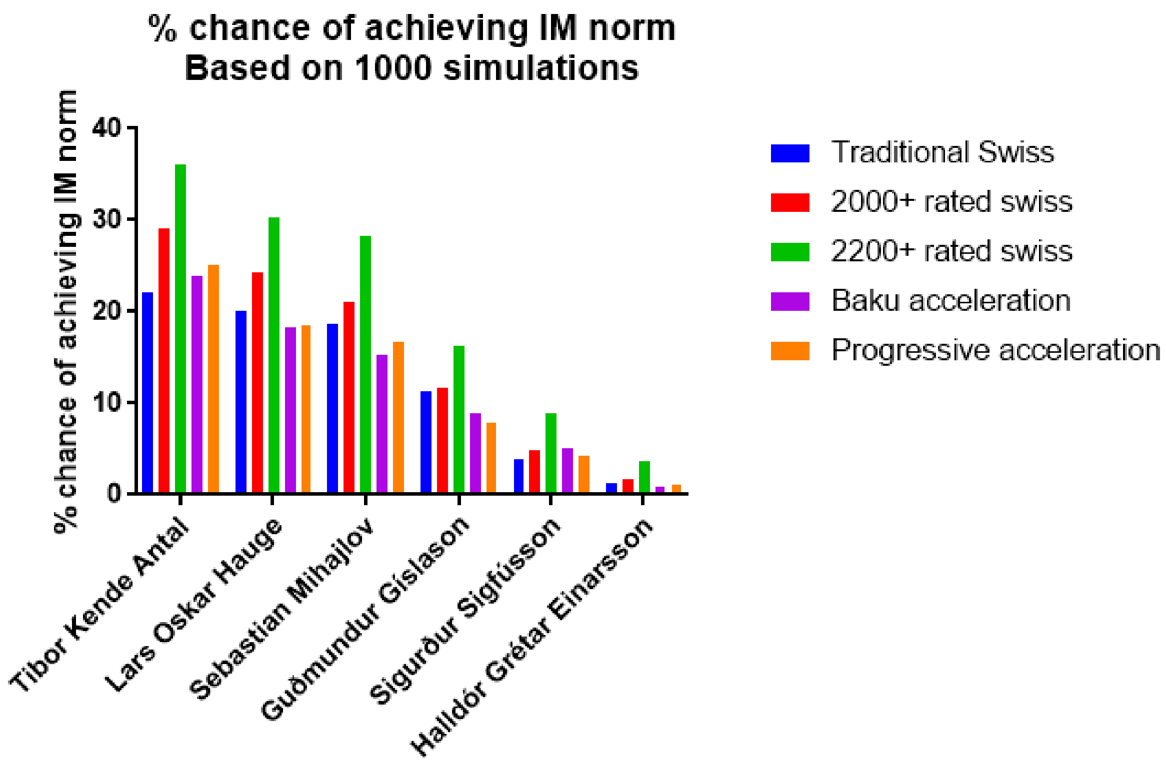


This is a close look at the three highest rated international players and the three highest rated Icelanders that could have a GM norm in Reykjavik Open 2017. We can see that is very difficult to achieve a GM norm. The highest rated player Pierre Bailet has over 2500 in FIDE rating, but nevertheless, he can only expect 10-15% chance of achieving a GM norm.

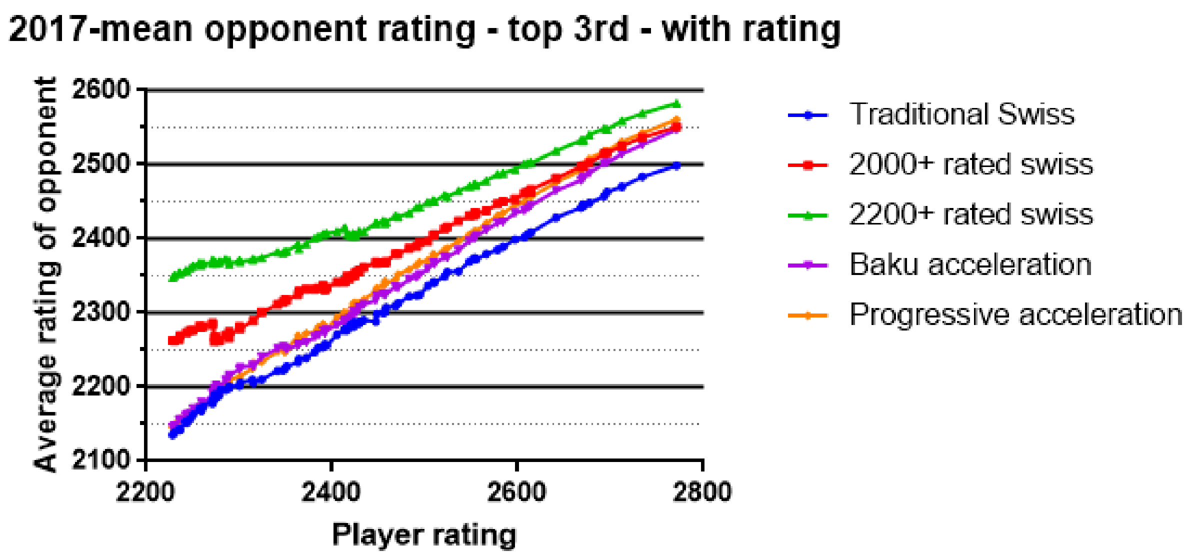


If we compare this to the 2016 tournament, it appears that Aryan Tari had more than 25% chance of achieving a GM norm, and it is worth remembering that he had 2553 chess level. This shows the most important factor for achieving a norm is obviously its own chess strength and not how the tournament is set up.

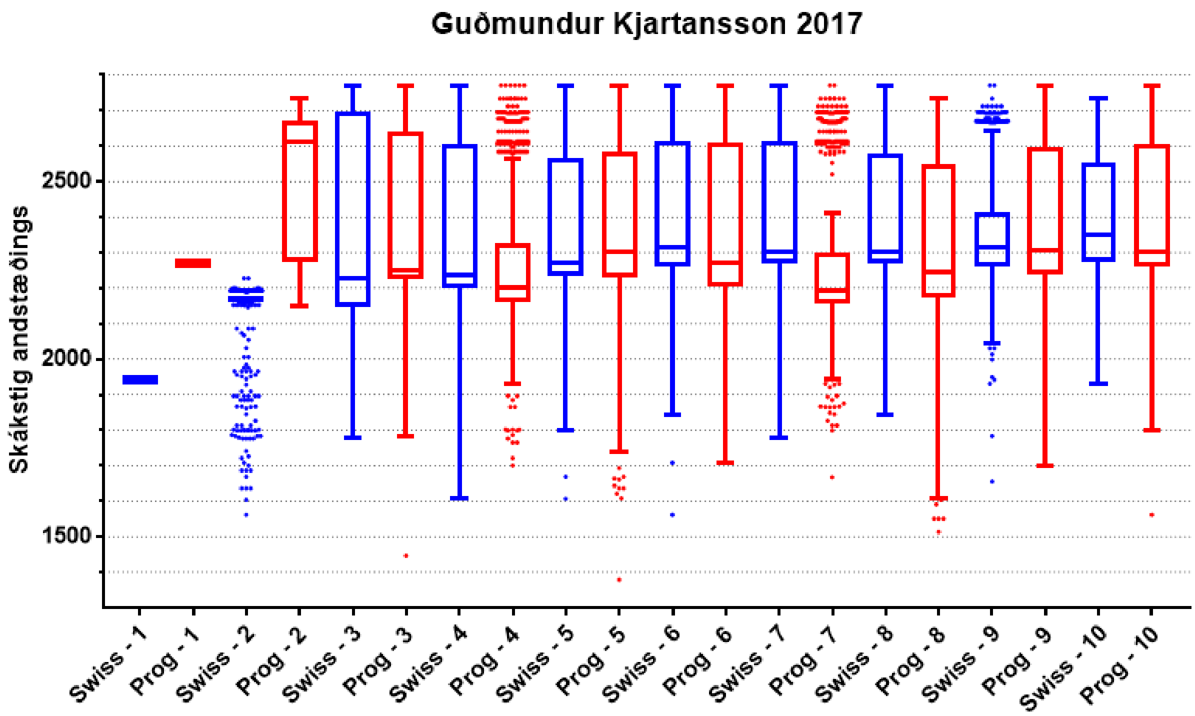




As previously mentioned, there is a greater variation among those seeking the IM norms, and especially in terms of closed tournaments verses open tournaments.



Looking at the opponents rating, it's clear that the accelerated systems seem to give 2500+ chess players a stronger opponent on average than the traditional Swiss system but the difference seems decrease as the chess strength decreases. As a result, acceleration do not seem to have a major impact on what refers to average counterparts for e.g. 2400 rated players.



For more information, I look at Gudmund Kjartansson as an example from 2017. In this “Tukey-Box-Andwhisker Graph”. The box shows where 25%, 50% and 75% of the distribution of Guðmundur's opponents. To explain the figure we can review round 2 in Progressive Acceleration (second red graph from the left), where shows that at 25% quartile in the model, Guðmundur is playing with opponents with over 2650 FIDE rating, in 50% quartile is just over 2600 points, and in 75% quartile just over 2300 points. The reason for that acceleration does not give more norms may be that after two relative easy rounds, there are 7-8 rounds with strong opponents, while in the accelerated tournament the player must fight in 8-9 round. I think this single extra round may lead to what is called "Regression to the mean" and that's why I think that the rankings are statistically not as big as in traditional Swiss system.

Conclusion:

If the Icelandic Chess Federation thinks it's better to have a completely open tournament as it has in recent years, there is not much in my mind that distinguishes between accelerated and the other traditional Swiss system. However, on the other side, I do not see anything against testing an accelerated system like Progressive Acceleration and it would be interesting to see how it stands in real life, not only in theoretical calculation. According to Gunnari Björnsson, it has been a great pleasure to use the system in Portu Mannu.

Finally, I would like to say that there is a large amount of data that came into play with this project and only a fraction of it is shown in this short report. If anyone wants to access the documents to make their own an analysis of this, please contact me.

Friðrik Jensen Karlsson
karlsson (dot) fridrik at gmail (dot) com
Brookline, MA, USA
10 July 2017

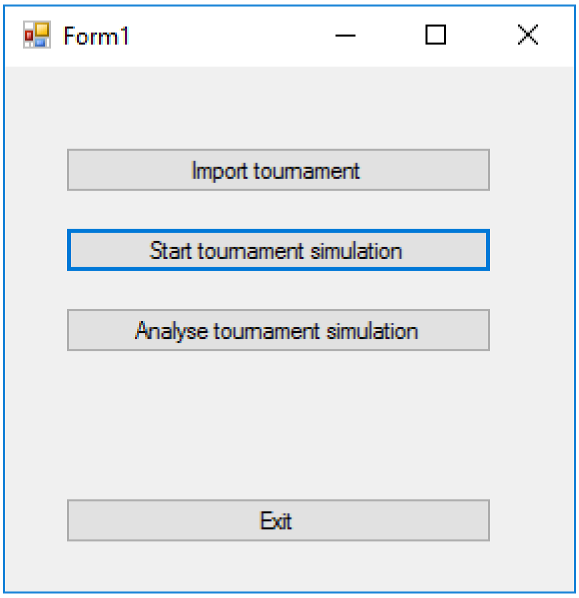
Appendix B. The program

The directory structure is:

```
C:\SPP\
...	accelerated\
...		csv\
...		failed\
...		JaVaFo\
...		results\
...		simulations\
...			mode1-Haley1-Fiderating\
...			mode1-Haley2-Fiderating\
...			mode1-Haley3-Fiderating\
```

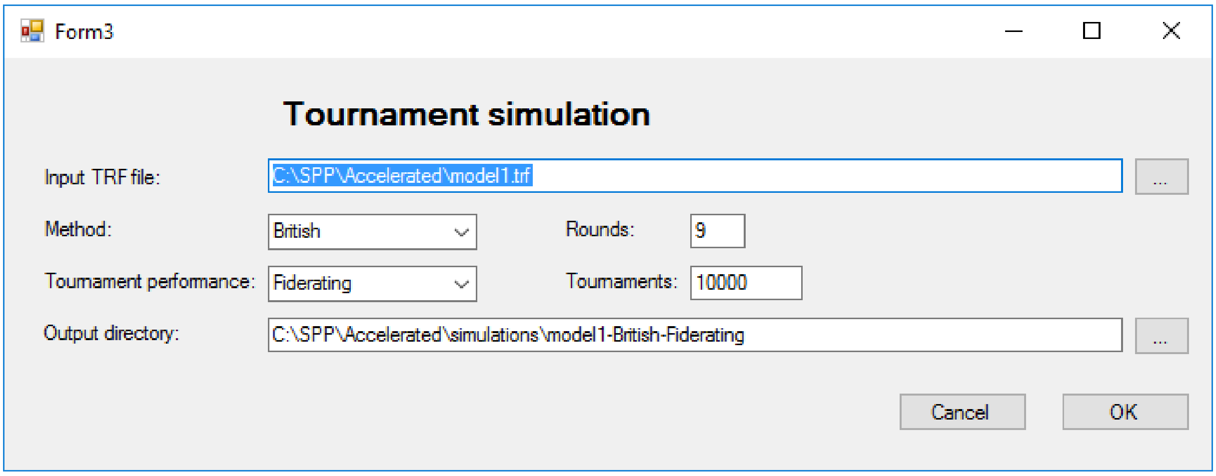
... mode1-Swiss-Fiderating\
... ...
... src\
... summary

The program written in Visual Studio 2013 is in the src catalog.



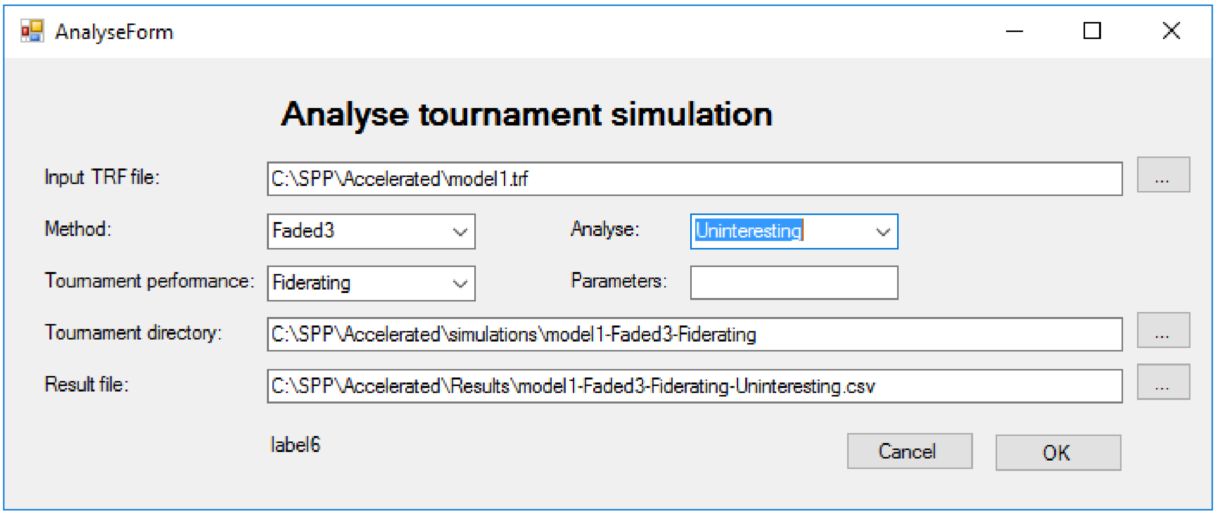
The program can import tournaments from Swiss tournament server. Download files in csv format and save to the csv catalog. Then run import tournament.

Start tournament simulations



Select an input file, a method, a rating, rounds, number of tournament and an output directory. This will run for a while, and it based on a correct directory structure.

Analyse tournament.



The Input file, method and rating performance will help to find correct input directory. Select what you want to analyze.

The result file is on csv format, comma separated file. I have used tab as delimiter. This makes it easy to import in programs like python and excel.

Example file

Rank	Num	Mean	Stdev
1	10000	6,5745	0,812188247883457
2	10000	6,4726	0,78790814185411
3	10000	6,3907	0,779489262273705
4	10000	6,29255	0,775041610173281
5	10000	6,2185	0,763745867419263
6	10000	6,11525	0,736405077046596
7	10000	6,0437	0,734397923472005
8	10000	5,9635	0,719143761705548
9	10000	5,87185	0,705905501820179
10	10000	5,794	0,69438029925971

Etc.

In Norway we have , as decimal sign!

Note the GNU General Public License

```
/*
 *  This file is part of SPP accelerated pairing software.
 *  Author: Otto Milvang, sjakk@milvang.no
 *
 *  SPP accelerated pairing software is free software: you can redistribute
 *  it and/or modifyit under the terms of the GNU General Public License as
 *  published by the Free Software Foundation, either version 3 of the License, or
 *  (at your option) any later version.
 *
 *  SPP accelerated pairing software  is distributed in the hope that it
 *  will be useful, but WITHOUT ANY WARRANTY; without even the implied
 *  warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.
 *  warranty of GNU General Public License for more details.
 *
 *  You should have received a copy of the GNU General Public License
 *  along with Foobar.  If not, see <http://www.gnu.org/licenses/>.
 */
```

Have fun!