

Oxford Croquet

Technical

Which Ranking System?

by Louis Nel, 1st January 2009

The splendid croquet ranking service rendered by Chris Williams to the croquet public via his 'Butedock Server', (<http://butedock.demon.co.uk/cgs/rank.php>), includes maintenance of no fewer than five different systems. So when Butedock visitors get there, which system do they look at? Mostly CGS (Croquet Grading System), I imagine. It has been there for over twenty years and long standing habits endure. Furthermore, it is the official system sanctioned by the World Croquet Federation (WCF). Which system(s) should they turn to in case they have a choice? This article aims to enable informed decisions. It reports on a systematic numerical comparison of the systems as regards their Performance Gauging Ability (PGA) and Volatility. PGA refers to how closely the system grade reflects the current performance level of the player. Volatility refers to how rapidly the rank positions change from one monthly ranking list to the next.

So which system gets the nod? Actually, none of the five. Contemplation of PGA inspired me to design a new system, or more precisely a variant of one of the above, with specific focus on Performance Gauging Ability. The new variant turned out to outperform all others in all test samples as far as PGA is concerned. As regards Volatility the situation is intriguing. The results suggest that some systems may have Volatility that is occasionally too low and others for which it may be generally too high, while that of the new variant seems to be about right.

When people need to take prescription drugs they generally don't try to understand the underlying biochemistry and physiology – they trust the experts about those technical things and consider only the effects and side-effects. I believe most players have, and indeed should have, the same attitude to ranking systems. Therefore, my advice to readers in general is to forget about the underlying algorithms. Leave that to the experts. Just focus on the effects and side-effects. For individual players the plots given on Butedock are helpful in this regard. This article supplements this information by showing effects on populations in nontechnical terms – everybody understands what percentages mean.

Readers may occasionally wish to calculate a post-game update. On **OxfordCroquet** there is an interactive calculator (see www.oxfordcroquet.com/rank/rcalc.asp) which enables readers to calculate approximate CGS updates. In principle, such interactive calculators could be created for most of the other systems too, as the need may arise. In view of this, *computational complexity* of ranking systems is in principle not a problem – not in our on-line age.

I look forward to the day when such interactive on-line services are expanded so that tournament managers could upload a list of game results together with entry data of the players and get updated ranking data for all participants right away. Thus block-based ranking or (flexible) swiss-based ranking could conveniently become replaced by updated world ranking for the purpose of seeding the final knock-out stage of a tournament – commendable in the interest of greater justice. (Of course, those that enjoy a little controversy could let a committee do the seeding instead).

There is a small but important group of readers who are interested in technical matters and for their sake algorithmic details are provided below. Those who do not belong to this exceptional group, should ignore the technical stuff.

Contents

[1. The Systems to be Compared](#)

- [1.1 The CGS Grade \(CG\)](#)
- [1.2 Elo Grade \(EG\)](#)
- [1.3 Bayesian Grade \(BG\)](#)
- [1.4 Index-only Based Grade, with Stepsize = 30 \(IG30\)](#)
- [1.5 Average Index Grade \(AvIG\)](#)
- [1.6 Adaptive Bayesian Grade \(ABG\)](#)

[2. Performance Gauging Ability Comparisons](#)

- [2.1 Numerical Measurement of PGA](#)
- [2.2 General Low Disparity Comparison](#)
- [2.3 Top Player Low Disparity Comparison](#)
- [2.4 General Top Player Comparison](#)
- [2.5 General Top Player Class 1 Games Comparison](#)

[3. Volatility Comparison](#)

- [3.1 Comparison Method](#)
- [3.2 General Population Comparison](#)
- [3.3 Top Player Population Comparison](#)
- [3.4 PGA vs Low Volatility](#)

[4. Side-effects](#)

[5. Notes](#)

- [5.1 CG\(k\) as Weighted Average of Indexes](#)
- [5.2 When CG Lags Behind Idx.](#)

1. The Systems to be Compared

The systems will be identified by their 'grade' i.e. ranking statistic.

1.1 The CGS Grade (CG)

This is the well known Croquet Grading System, the system sanctioned by the World Croquet Federation. We denote its grade by CG and its index by Idx. Occasionally we use CG(k) and Idx(k) to denote the values of CG and Idx for a player after k games in the system.

Postgame updates for Idx are done as follows:

for Winner: $Idx(k+1) = Idx(k) + Stepsize \cdot cwp(Loser_Idx(k), Winner_Idx(k))$
for Loser: $Idx(k+1) = Idx(k) - Stepsize \cdot cwp(Loser_Idx(k), Winner_Idx(k))$

where Stepsize = 60, 50 or 40 for Class 1, 2, 3 games respectively, the symbol \cdot denotes multiplication and cwp is the well known Classical Win Probability function given by

$$cwp(x,y) = 1 / (1 + 10^{(y-x) / 500}).$$

Postgame updates for CG are done as follows:

$$CG(k+1) = s(k) \cdot CG(k) + (1 - s(k)) \cdot Idx(k+1)$$

where $s(k) = 0.80 + (CG(k) - 1000)/10000$. Thus the Smoothing Parameter s has a value that varies from 0.9 (when CG(k) = 2000 or less) to 0.97 (when CG(k) = 2600 or more).

1.2 Elo Grade (EG)

In postgame updates the player's EG is incremented by the amount $40 \cdot (OW - EW)$, where OW is the observed number of wins of the player in the event and EW is the expected number of wins i.e. the sum of the player's Classical Win Probabilities, calculated on the EG values with which the players entered the event. Thus EG becomes increased when $OW - EW > 0$ and decreased when $OW - EW < 0$.

1.3 Bayesian Grade (BG)

Ranking data consists of two numbers, BG and SD, respectively the Mean and Standard Deviation of the Bell Curve that represents the performance level of the player. For postgame updating purposes the Bell Curve of a player X can be replaced by a pregame histogram consisting of 8 (say) pairs of numbers (x_r, ξ_r) ($r = 1, 2, \dots, 8$) where x_r is a performance level on the croquet scale and ξ_r (Greek letter xi) is the probability that the player will perform at that level x_r . These numbers can be derived from the ranking data BG, SD by putting

$$x_r = BG + \sqrt{2} \cdot SD \cdot g_r \text{ and } \xi_r = h_r / \sqrt{\pi}, (r = 1, 2, \dots, 8)$$

where g_r and h_r are Gauss-Hermite nodes and weights used in numerical integration. [Google 'Gauss-Hermite quadrature'](#) for details.

When X beats Y the postgame histogram (x_r, ν_r) of X is given by the formula

$$\nu_k = \xi_k \cdot (\eta_1 \cdot \text{cwp}(x_k, y_1) + \dots + (\eta_8 \cdot \text{cwp}(x_k, y_8))) / \text{BWP}(X, Y), (k=1, 2, \dots, 8),$$

where (y_r, η_r) ($r = 1, 2, \dots, 8$) is the pregame histogram of Y, cwp is the Classical Win Probability function (used also in CG and EG), BWP(X, Y) is the Bayesian Win Probability of X over Y, i.e. the sum of all terms of the form $\xi_i \cdot \eta_j \cdot \text{cwp}(x_i, y_j)$ ($i, j=1, 2, \dots, 8$). By interchanging the roles of (x_r, ξ_r) and (y_r, η_r) we get the corresponding postgame histogram for Y.

From a postgame histogram (x_r, ν_k) ($k=1, 2, \dots, 8$) we obtain the updated ranking data as follows:

$$\text{newBG} = x_1 \cdot \nu_1 + \dots + x_8 \cdot \nu_8$$

$$\text{newSD} = \sqrt{((x_1 - \text{newBG})^2 \cdot \nu_1 + \dots + (x_8 - \text{newBG})^2 \cdot \nu_8)}$$

subject to the restriction $\text{newSD} \geq 55$.

There is a second updating procedure, called Temporal Update, done at the start of each tournament for all players.

It adjusts only the Standard Deviation by increasing it as follows:

$$\text{newSD} = \sqrt{(\text{SD}^2 + 4489 \cdot \text{Days} / 365)}$$

where Days = the number of days since the player last played. A larger SD has the effect of increasing the size of the postgame adjustment to BG.

Upon admission to the database, players start with an initial SD of 320 and an initial BG assigned by the ranking officer.

(BG was introduced in the article '[Bayesian Ranking for Croquet](#)' and the post-game updating simplified in '[Bayesian Updating Simplified](#)').

1.4 Index-only Based Grade, with Stepsize = 30 (IG30)

Postgame updates are extremely simple: the winner's IG30 is increased by $30 \cdot \text{cwp}(\text{Loser IG30}, \text{Winner IG30})$ and the loser's IG30 is decreased by the same amount.

1.5 Average Index Grade (AvIG)

After each game, AvIG is the straight average of all Idx values the player had in the preceding 12 month period.

1.6 Adaptive Bayesian Grade (ABG)

This is the new system, herewith introduced.

The performance level of a player may sometimes rapidly change for various reasons, including randomness. When that happens, the grade may deviate unusually far from the actual performance level, because it is then difficult for a ranking system to keep track. The system ABG is essentially BG with incorporation of an additional adaptive algorithm. When, and only when, it detects a rapid change in the form of a player it takes action to address the grade deviation.

To this end we define the Grade Deviation (GD) of a player by putting $GD = OW - EW$, where OW is the observed number of wins of the player and EW is the expected number of wins. This expression occurs also in Elo post-event updating, but it is implemented differently here. EW is now the sum of the Bayesian win probabilities of the player and it is applied over 5 games at a time. For example, if 5 games have been played since the previous review and the player had 3 wins while his win probabilities in the five games were 0.4, 0.6, 0.7, 0.3, 0.2 respectively, then that player has $EW = 0.4+0.6+0.7+0.3+0.2 = 2.2$. So $GD = 3 - 2.2 = 0.8 > 0$ and the player is performing slightly above expectation. If $GD < 0$, the performance is below expectation. For a reasonably accurately graded player, GD is close to zero. A large positive value for GD signifies that the player is seriously underrated; a large negative value makes him seriously overrated. The system ABG works in Game Cycles of 5 games.

For example, when GIS (the player's number of Games In the System) reaches 115, the GD that accumulated for games 111, 112, 113, 114, 115 is looked at. When $|GD| \leq 1.88$ no adaptive action is taken because the GD is deemed small enough to ignore (notation: $|GD|$ means GD or -GD according to whether $GD > 0$ or not). When the player's Standard Deviation is 104 or higher, no adaptive action is taken either, because an SD of 104 is deemed high enough for autocorrection to take place fast enough. (In Bayesian ranking, the SD regulates the size of the grade increment in postgame updating; an SD of 104 gives a grade increase of 23 for a win against an equally graded player.) When the review finds that $|GD| > 1.88$ and the standard deviation is below 104, twofold action takes place. Firstly, the SD is raised to the value of 104. In cases where the standard deviation was much below 104 this speeds up autocorrection considerably. Secondly, there is the consideration that the grade of the player should be given a little push towards correction right away. It seems reasonable that the grade should be increased to the extent that $|GD| > 1.88$. However, towards avoiding overcorrection, a slower than linear growth with respect to the variable $x = |GD| - 1.88$ seems preferable; similarly for $y = SD - 104$. In view of these considerations, an increment of the form $5 \cdot \sqrt{(|GD| - 1.88) \cdot (104 - SD)}$ was chosen. Experiments with a gradual increase of SD instead of the jump to 104 showed the jump to be clearly preferable. All told, the procedure can be summarized as follows.

Adaptive Algorithm: If GIS is a multiple of 5 and $SD < 104$ and $|GD| > 1.88$, then

- the standard deviation SD is assigned the value 104,
- the grade ABG is increased by $5 \cdot \sqrt{(|GD| - 1.88) \cdot (104 - SD)}$ in case $GD > 1.88$ or decreased by this quantity in case $GD < -1.88$.

After these adjustments GD is reset to 0.

The constant (parameter) values 1.88, 104, 5 that occur in this algorithm were all experimentally determined so as to give the best performance collectively.

Players seldom qualify for application of the algorithm. For example, for the year 2007, the qualification rate for all players on the ranking list was 1.15 per 100 games played; for the top 100 players it was 0.97 per 100 games. However, a rising star like Rutger Beijderwellen had 4 applications over the 125 games he played. The rarity of application is largely due to the requirement $SD < 104$. Without it, the application rate would be much higher and would not improve overall performance. Since players start with an $SD > 300$, they need to play a number of games before there can be any chance of qualifying for adaptive adjustment.

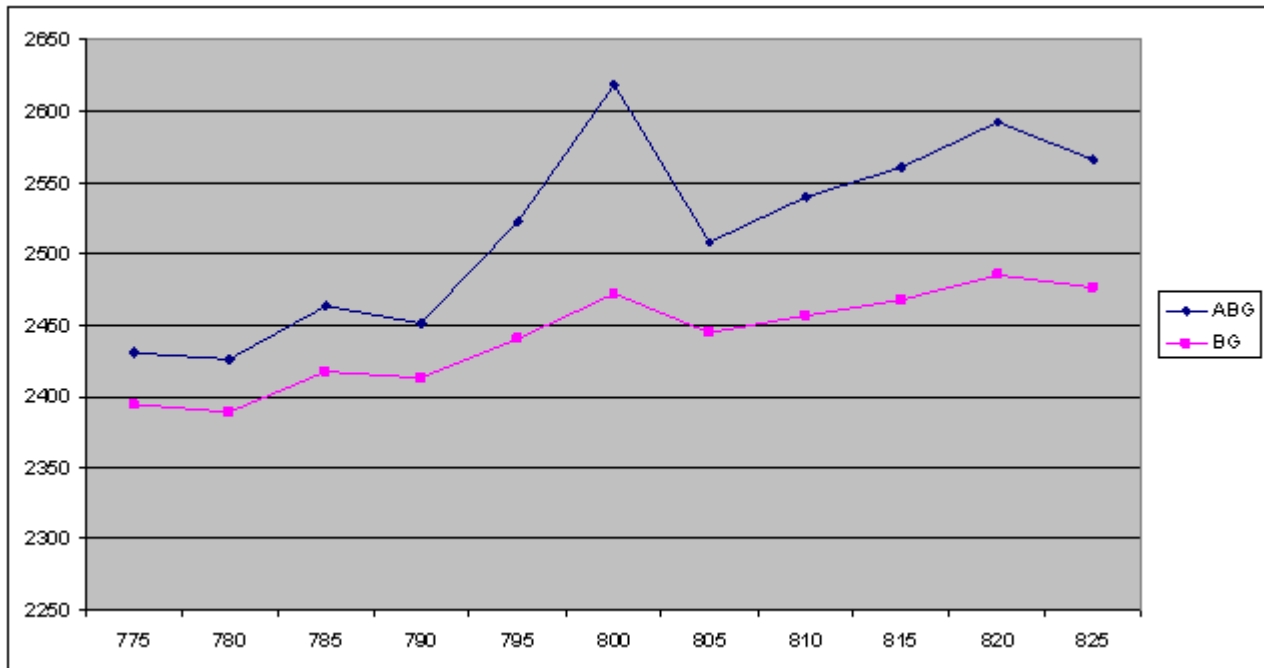
The Temporal Update of BG (see [1.3](#)) is changed in ABG as follows:

$$\text{newSD} = \sqrt{SD^2 + 3364 \cdot \text{Days} / 365}$$

and the constraint $\text{newSD} \geq 55$ is removed, being redundant in ABG.

The following chart provides a detailed illustration of effects produced by ABG. The Standard Deviation for ABG is written ABSD when it could be confused with the SD for BG.

Grades of Rutger Beijderwellen at 5 game intervals



The numbers below the horizontal line denotes the player's GIS (= games in system). Adaptations occurred at the games 785, 795, 800, 805. The first one occurred when ABSD was only 51, so it substantially amplified his grade adjustments for a while. The later adaptations had less dramatic effect, because the ABSD was still fairly high when they occurred. We can never be sure that the situation $|OW - EW| > 1.88$ happens through a change in player form rather than randomness. The peak at game 800 seems to be out of place in view of what came before and after it. It is gratifying to see that ABG quickly brought the grade back in line. Similar quick corrections of sudden sharp changes in grade was noticed also in other special cases. They suggest that no harm is done when ABG mistakes a random change for a real change in form. The steeper gradient after game 805 shows that the increased SD produces an effect that lasts for several games. The above history is that of a player with unusual upward mobility. For most players there is only a slight difference between the grades BG and ABG. It is also unusual for a player to get four adaptations in rapid succession.

Effects produced in individual cases are instructive but this needs to be supplemented by a study of effects on populations. Before turning to that, let us note that from 1985 to Aug 2008 a total of 82812 GD reviews at 5 game intervals were made by ABG. The reviews resulting in adaptive action numbered 4715 (5.6%). Of these, the upward adjustments numbered 2441 (51.8%). Among randomly induced adjustments one could expect the upward ones to be cancelled by downward ones. So among non-random instances of a rapid change the percentage of upward adjustments may be somewhat higher than the mentioned 51.8%. This is not surprising if one considers that players may acquire a better technique more readily than a worse one.

2. Performance Gauging Ability Comparisons

2.1 Numerical Measurement of PGA

A down-to-earth method for comparing the PGA of systems is to look at the Percentage of Correct Predictions (PCP) implied by their grades over a chosen set of test games. This method is applicable to all ranking systems - even systems not based on win probabilities. After all the mathematical modelling that goes into the design of a ranking system, PCP provides an objective reality check - to see how the system's rankings relate to observable real world results. Such reality check is an essential feature of all mathematical modelling of real world phenomena. Without it, the best theory could lead to unsatisfactory results. For example, an essential parameter could be missing or poorly chosen.

We will use, as [general sample](#), the 148250 games in the CGS database that were played in the period 1 January 1996 until 3 August 2008 in which both players had GIS ≥ 10 . However, PCP comparison on the

entire general sample is unsatisfactory. The average CG difference between opponents in this sample is 234. This means that in the vast majority of games all systems will predict the same winner, thus all will be simultaneously correct or simultaneously wrong. Any difference between systems that emerges is bound to look very small after division by 148250 games, most of which did not do any real testing. So we use subsamples of low disparity games to get better insight. Another consideration arises. The system CG is really a different system for grades above 2000 than for grades below that level. It is unique among the systems under consideration in its focus on top players. Its smoothing parameter gets progressively smoother until it reaches grades of 2600. Its class factors are introduced with elite players and prestigious events in mind. It is therefore inappropriate to judge it on a general sample only, which is bound to obscure whatever special attributes it may bring to these situations. Our tests therefore also include samples that involve top players specifically, in order to see how the various ranking systems perform at high altitudes.

The predictions will be on the basis of pregame grades i.e. the grades as they are immediately before each game. Predictions on the basis of pre-event grades (i.e. grades on arrival at the event) are of interest because those are the grades seen by event organizers. However, such grades have various ages, which depend on how long ago the player was active. PCP on the basis of grades several months before the event are also of interest, because that is what selection committees are faced with. Such grades are even more variable in age. Grades in the month of March are typically about six months old in the northern hemisphere and September grades are also typically several months old in the southern hemisphere. Earlier studies have shown that PCP based on older grades of these kinds generally echo the behaviour of the pregame PCP; the predictions of all systems just get weaker as the grades get older. So pregame grades form a better basis on which to judge the inherent PGA of the systems. Since the systems under study are all based on continuous maintenance of win probabilities, PCP also provides, albeit indirectly, a test of how well each system performs this maintenance. Pregame grades do this better than older grades could do.

When looking at the PCP over more than one sample, it should always be born in mind that it is the relative positions of the systems that is of interest, not the absolute value of the PCPs, because the latter reflects the average disparity present in the sample. This average disparity varies from one sample to the next.

2.2 General Low Disparity Comparison

Sample: Games in the general sample with absolute CG difference between opponents below 70.
Sample size: 31627 games.

Table 2.2

Predictor	PCP
ABG	55.74
BG	55.29
IG30	54.81
AvIG	53.72
CG	53.32
EG	52.40

There is a simple principle at work, namely **older grades give weaker prediction**. The three systems at the bottom have in common that older ranking data has a direct and prominent role in the value of the updated grade. Indeed, CG is a weighted average of earlier values of Idx and **as grades increase from 2000 to 2600, the relative contribution of indexes more than 10 games ago to the total value of CG increases roughly from 0.31 to 0.67** (see note [5.1](#) below for elaboration). AvIG is a straight average of earlier index values. If the number of games in the preceding 12 months is 30, then the indexes more than 10 games ago will contribute 20/30 of the Grade i.e. 67%; for a player with 50 games it would be 80%. In the case of EG, postgame updates are done in terms of entry grades. They are often several weeks old.

The top three systems all have postgame updates expressed entirely in terms of the most recent grade only. Thus CG, AvIG and EG are born hampered compared to the top three systems. In view of this it is not surprising to find the mentioned systems at the bottom.

2.3 Top Player Low Disparity Comparison

Sample: Games of Comparison [2.2](#) in which both players have CG above 2500.
Sample size: 856 games.

Table 2.3

Predictor	PCP
AB	57.01
BG	55.02
IG30	53.27
CG	51.64
EG	48.95
AvIG	48.01

CG is unique among systems on Butedock in the way it is top-player oriented, expressed through the continuous increase of the smoothing factor s from 0.9 to 0.97 as CG increases from 2000 to 2600. Whatever be the reason for its introduction, this feature did not do its PGA any good. The numbers of Tables [2.3](#) and [2.2](#) suggest that [the PGA gap between CG and the top systems is wider for top players than for the general population](#). In view of the *older grades give weaker prediction* principle, this is not surprising. Indeed, as shown in [2.2](#), for grades near 2600 the older data component is significantly higher than for grades near 2000.

The same PGA gap for AvIG is also wider. That is not surprising either, because top players are generally more active than the general population and that also means (as shown in [2.2](#)) that the older data component of the grade is higher. In the case of EG one could speculate that the wider gap is due to an algorithmic weakness that becomes more conspicuous with low disparity.

2.4 General Top Player Comparison

Sample: Games in the general sample with both CG above 2500.

Sample size: 2151 games.

Table 2.4

Predictor	PCP
AB	62.02
BG	61.04
IG30	60.07
CG	59.69
AvIG	57.74
EG	57.60

2.5 General Top Player Class 1 Games Comparison

Sample: Class 1 games in the general sample with both CG above 2500.

Sample size: 1080 games.

Table 2.5

Predictor	PCP
AB	63.06
BG	61.94
IG30	61.94
CG	60.74
EG	59.44
AvIG	59.07

The latter two samples were not constrained to have low disparity, so the smaller gap between the systems is not surprising (see [2.1](#)). Since Class 1 games are not played exclusively by anybody, comparison [2.4](#) does not tell us much if anything about the effectiveness or not of the Class 1 factor. If it does improve the PGA of CG, the improvement is not conspicuous. This sample is included mainly because of the prominent

role of Class 1 games in the algorithms of CG and AvIG. They have no special role in any of the other systems.

3. Volatility Comparison

3.1 Comparison Method

Since the world ranking list typically appears at monthly intervals, we quantify the Volatility of ranking systems in terms of the rank variation per player per month. It works as follows. If a player is ranked at position 19 on a given list and at position 21 a month later, then the rank positions give a difference of $19 - 21 = -2$. The absolute value of this difference is therefore 2. We call this the **variation of that player for that month**. Such a variation can be computed for every player which appears on both of the lists. After dividing by the number of players on the list, we arrive **the rank variation per player for that month**. This value varies from month to month, partly influenced by the dormant seasons in the two hemispheres. By taking its average value over a number of consecutive months, we arrive at the statistic **Rvar**. It expresses **rank variation per player per month for the chosen period**. One can compute this statistic also for subpopulations. By **RvarTop** we will denote the statistic obtained by considering the top 100 players on the ranking list in question. In all these calculations, if a player under consideration did not appear on the previous monthly ranking list, then that player is ignored in the calculation for that month. Since the same players appear on the ranking list for every system, the same players are ignored for all systems. In the comparison to follow we used the 36 consecutive monthly ranking lists from January 2005 through December 2007.

3.2 General Population Comparison

Table 3.2

System	Rvar
BG	9.54
ABG	10.05
AvIG	16.26
CG	17.54
IG30	19.19
EG	21.14

3.3 Top Player Population Comparison

Table 3.3

System	RvarTop
AvIG	2.94
BG	3.53
CG	4.12
ABG	4.27
IG30	5.20
EG	6.59

3.4 PGA vs Low Volatility

Good PGA goes hand in hand with reasonable Volatility. Indeed, Volatility that is either excessively high or excessively low would preclude good PGA. It therefore seems prudent to design a ranking system in the first place so as to give good PGA. If its Volatility seems higher than desirable, two alternatives may reasonably be explored. One is to change parameters so as to reduce Volatility. The other is to look for a better system whose natural Volatility is acceptable. The above ranking systems illustrate pursuit of both alternatives, not with equal success.

CG with a smoothing parameter fixed at $s = 0.9$ gives reasonable PGA but with unpleasantly high Volatility (the $Rvar = 17.54$ indicates that). The increased value of s for top players dramatically reduces Volatility but does so at the expense of reduced PGA, even when it is already mediocre to start with. It also aggravates the lag effect described in notes [5.1](#) and [5.2](#) below.

There is reason to believe that the Volatility of BG for top players is actually too low. Top players are generally more active than the general population and so they end up having a low SD, because the only reason for SD to increase is through inactive periods. An SD which is too low could impair PGA. That is why the Minimum SD parameter was introduced in BG. The introduction of the new system ABG is a more potent measure to address this. It provides a natural additional way for the SD to increase when necessary and thus to improve PGA. The resulting Volatility has a credible claim to be appropriate.

The system IG30 gives further support to the view that PGA goes accompanied with appropriate Volatility rather than low Volatility per se. One can reduce IG30 Volatility to very low levels by reducing its Stepsize, but at the cost of much reduced PGA. High values of the Stepsize also reduces PGA while increasing Volatility to unpleasantly high levels. The simplicity of this system leaves little room for maneuvering.

AvIG sheds further light on the matter. Since top players are generally more active, they call on a greater proportion of their AvIG to come from older Idx values (as explained in 2.2). While this spectacularly reduces Volatility (as seen in 3.3), it does so at the expense of worse PGA (see its relative positions in 2.2 and 2.3).

4. Side-effects

The Class Factors used in CGS may produce side-effects but this is not well understood yet. One could speculate that players who do well in Class 1 games will tend to become overrated while those who don't do well will tend to be underrated. Since these two effects will cancel as far as the population as a whole is concerned, the effect will not show up in population studies.

The lag effect of CG (detailed in notes 5.1 and 5.2 below) is long known. It is particularly troublesome at critical moments. For example, in the 2008 World Championship the top seed turned out not to be in his best form. After block play, he was still top-ranked by CG and seeding of the Knock-out stage was done accordingly, as prescribed by WCF regulation. None of the systems with better PGA had him top-ranked at that point. This kind of thing will happen repeatedly in future for as long as a system with such lag effect is prescribed for seeding purposes. It is like an accident waiting all the time for a place to happen.

5. Notes

5.1 CG(k) as Weighted Average of Indexes

Recall (see 1.1) that

$$(a) \text{CG}(k) = s(k-1) \cdot \text{CG}(k-1) + (1 - s(k-1)) \cdot \text{Idx}(k).$$

Let us consider the case of a player that has at least 100 games in the system and that $\text{CG}(k) > 2600$ for the last 100 of those games. In such a case $s(k) = 0.97 = S$ (say) for all $k=1,2,\dots,100$.

Then similarly $\text{CG}(k-1) = S \cdot \text{CG}(k-2) + (1 - S) \cdot \text{Idx}(k-1)$. After substituting this expression of $\text{CG}(k-1)$ into formula (a) we arrive at an expression of $\text{CG}(k)$ in terms of $\text{CG}(k)$, $\text{CG}(k-1)$, $\text{Idx}(k-2)$. We can repeatedly do such substitution and ultimately express $\text{CG}(k)$ in terms of $\text{Idx}(k-100)$, $\text{Idx}(k-99)$, $\text{Idx}(k-98)$, ... $\text{Idx}(k)$. After algebraic manipulation we obtain

$$\begin{aligned} \text{CG}(k) &= (1 - S) \cdot \text{Idx}(k) + S \cdot (1 - S) \cdot \text{Idx}(k-1) + S^2 \cdot (1 - S) \cdot \text{Idx}(k-2) + S^3 \cdot (1 - S) \cdot \text{Idx}(k-3) + \dots \\ &= w(k) \cdot \text{Idx}(k) + w(k-1) \cdot \text{Idx}(k-1) + w(k-2) \cdot \text{Idx}(k-2) + \dots \end{aligned}$$

where $w(k-j) = S^j \cdot (1 - S)$. This shows $\text{CG}(k)$ to be a weighted average of earlier values of Idx . Although the weight $w(k-j)$ gets reduced by a factor S for each next j , it does not instantly become small. By plugging in $k = 100$, $s = 0.97$ and adding, we find that

$$w(89) + w(88) + w(87) + \dots + w(1) = 0.67 \text{ approximately.}$$

In other words [the terms involving indexes more than 10 games ago will have a combined weight that exceeds 0.669](#). A similar analysis for the situation where all 100 Grades are 2000 and below, so that $S = 0.9$ gives $w(89) + w(88) + w(87) + \dots + w(1) = 0.31$ approximately. So we see that [as grades increase from 2000 to 2600, the relative contribution of indexes more than 10 games ago to the total value of CG increases roughly from 0.313 to 0.669](#).

5.2 When CG Lags Behind Idx.

When the CG of a player becomes much smaller or much larger than the Idx, the older data effect discussed in 5.1 causes anomalous behavior, as follows. For several games in a row the value of CG will increase even after a game is lost, in case $CG < Idx$ and the value of CG will decrease even after a game is won, in case $CG > Idx$.

Author: Louis Nel
All rights reserved © 2009

Updated 5.i.09
[About](#), [Feedback](#)

www.oxfordcroquet.com/tech/nel-rank/index.asp
on www.oxfordcroquet.com
