PROBABILITY OF WIN

Part I: The local model for handicap games

Geoff Kaniuk *geoff@kaniuk.co.uk* July 2020

Contents

List of Tables

List of Figures

References

[1] www.europeangodatabase.eu/EGD/winning_stats.php

[2] www.kaniuk.co.uk/articles/pwin/prob-win.pdf

[3] en.wikipedia.org/wiki/Sigmoid_function

[4] en.wikipedia.org/wiki/Error_function

1 INTRODUCTION

$X_{\texttt{yellow}}$

black

white

In McMahon tournaments, players above shodan rarely play handicap games, but players below 10 kyu will more often need to give or take handicap stones. We are thus seeking to obtain information on the probability that a player i with grade g_i will win a game against a player of grade g_j, taking a handicap of h stones. Note that in a McMahon tournament it is possible for $g_i > g_j$ and yet i is playing black with a handicap.

A model for winning probability is a key components in any attempt to improve pairing programs or rating systems.

In even games there is no handicap, so we are interested in a simple representation for the winning probability: $P_{\texttt{win}}(g_i, g_j)$ where $g_i \leq g_j$. The probability of win between players in even games can readily be examined through the tables provided by the European Go Database [1]. Clearly, for handicap games the winning probability will depend on the handicap h, so we need to add an extra parameter to the representation, giving $P_{\texttt{win}}$in the form $P_{\texttt{win}}(g_i, g_j, h)$.

Values for $P_{\texttt{win}}$can be obtained from game data stored in EGD, and I am very grateful to Aldo Podavini (the previous EGD maintainer) for providing me with a file of such data containing results from 1996 to 2018. An initial attempt to model the full distribution $P_{\texttt{win}}(g_i, g_j, h)$, was not very succesful as the data is quite sparse and the resulting distributions were unacceptably noisy. The usual technique for managing such noise is to increase the time window for collection of game and result counts. But if the window is made too wide we lose any information on the time behaviour of the distribution - a feature important for assessing evidence of systematic changes in player ratings.

The handicapping system is designed to give the weaker player a handicap $h = g_{\texttt{white}} - g_{\texttt{black}}$. This should result in a fair game where the black player (grade $g_{\texttt{black}}$) should win about 50% of games[1]. However in McMahon tournaments it is normal to allocate handicaps (where needed) on the basis of the difference in McMahon *scores*. This means that the handicap actually allocated could be less than ideal and the win probability of the black player is consequently reduced.

So the main driver for a player's win probability in tournament handicap games is not the actual handicap h, but rather the *handicap deficit* defined by:

$$d = g_{\texttt{w}} - g_{\texttt{b}} - h \tag{1}$$

This allows us to collect more data for a given grade $g_{\texttt{b}}$ as *all* the games with different handicaps but fixed deficit d will contribute to the probability of win $P_{\texttt{win}}(g_{\texttt{b}}, d)$ for the deficit d.

The plan of this report is to firstly spell out the data needed for measuring both $P_{\texttt{win}}(g_{\texttt{b}}, g_{\texttt{w}}, h)$ and $P_{\texttt{win}}(g_{\texttt{b}}, d)$. We then present local models for $P_{\texttt{win}}(g_{\texttt{b}}, d)$ obtained for each of a limited range of values for d. It has become clear in this analysis that the model varies over time. Finally we illustrate what a global model fitted over both $g_{\texttt{b}}$ and d might look like.

2 $P_{\texttt{win}}$ FOR SPECIFIC HANDICAP

The dataset needed for the measurement of $P_{\texttt{win}}(g_{\texttt{b}}, g_{\texttt{w}}, h)$ is presented in Appendix B. It is useful to present $P_{\texttt{win}}$ separately for black and white:

$$P^{\texttt{B}}_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{w}}, h) = N^{\texttt{B}}_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{w}}, h)/\texttt{N}_{\texttt{game}}(g_{\texttt{b}}, g_{\texttt{w}}, h) \tag{2}$$

$$P^{\texttt{W}}_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{w}}, h) = N^{\texttt{W}}_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{w}}, h)/\texttt{N}_{\texttt{game}}(g_{\texttt{b}}, g_{\texttt{w}}, h) = 1 - P^{\texttt{B}}_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{w}}, h) \tag{3}$$

Note that in these expressions the grades $g_{\texttt{b}}$, $g_{\texttt{w}}$ for Black and White respectively can take any value in principle, but the handicap $h >= 0$.

[1]It is often claimed that this handicap should be increased by half a stone.

2.1 Even games

When $h = 0$ the game is treated as even no matter what the actual player grades are, and White will usually receive komi. In theory if the correct komi is given, then for equally graded players[2] the winning probability is $\frac{1}{2}$ for either player.

For even games between players of different grades, the winning probability depends only on the individual grades and not the player colours. All of these statements can be tested by examination of the data, but for the moment we can treat these as axiomatic:

$$P^{\mathtt{B}}_{\mathtt{hwin}}(g, g, 0) = P^{\mathtt{W}}_{\mathtt{hwin}}(g, g, 0) = \frac{1}{2} \tag{4}$$

$$P^{\mathtt{B}}_{\mathtt{hwin}}(r, s, 0) = P^{\mathtt{W}}_{\mathtt{hwin}}(s, r, 0) \tag{5}$$

Note that in (5) the superscript picks the position of the winner in the parameter list. So both sides give us the probability that the player with grade r wins. From Equations (3) and (5), we obtain the expected result (dropping the term $h = 0$):

$$P^{\mathtt{B}}_{\mathtt{hwin}}(r, s) + P^{\mathtt{B}}_{\mathtt{hwin}}(s, r) = 1 \tag{6}$$

$$P^{\mathtt{W}}_{\mathtt{hwin}}(r, s) + P^{\mathtt{W}}_{\mathtt{hwin}}(s, r) = 1 \tag{7}$$

For even games we have found [2] that a sigmoid model for $P_{\mathtt{win}}$ fits the game data quite well - at least for players above 14 kyu until about 2011. There is no standard definition for a sigmoid function $S(x)$, but many useful examples [3] have these properties:

$$S(x) \to \mp 1 \text{ as } x \to \mp\infty \tag{8}$$

$$S'(x) > 0 \;\forall\, x \tag{9}$$

$$S(x) + S(-x) = 0 \;\forall\, x \tag{10}$$

For even games, we expect that the winning probability will decrease to zero as the grade of the player's opponent gets much larger than the player's own grade. This, together with the Sigmoid properties above, and the results in Equation (6) immediately suggests the form of the model we use in Go, namely:

$$P_{\mathtt{win}}(r, s) = \frac{1}{2}[1 - S(\Lambda(r, s))] \tag{11}$$

In particular, the anti-symmetry in Equation (10) implies that Λ is antisymmetric in r, s). Based on these properties good models have beed found using the error function $\mathrm{erf}(x)$ as the basic sigmoid. The other sigmoid much used in rating systems is of course logistic function:

$$P_{\mathtt{win}}(r, s) = \frac{1}{2}[1 - \tanh(\Lambda(r, s)] = \frac{1}{1 + \exp(\Lambda(r, s))} \tag{12}$$

No one sigmoid is sacrosanct as the 1-1 property in Equation (9) allows us to transform Λ as appropriate.

2.2 Handicap games

$P_{\mathtt{win}}$ is now more complicated as it depends on the full 3 variables $(g_{\mathtt{b}}, g_{\mathtt{w}}, h)$, all of which can take any value. The data available was too sparse and noisy to provide reliable measurements of $P_{\mathtt{win}}$ at each parameter point. This distribution will be revisited using more extensive data from EGD.

The handicap-deficit probabilities are estimated from:

$$P^{\mathtt{B}}_{\mathtt{dwin}}(g_{\mathtt{b}}, d) = N^{\mathtt{B}}_{\mathtt{hwin}}(g_{\mathtt{b}}, d)/N^{\mathtt{B}}_{\mathtt{game}}(g_{\mathtt{b}}, d) \tag{13}$$

$$P^{\mathtt{W}}_{\mathtt{dwin}}(g_{\mathtt{b}}, g_{\mathtt{w}}, h) = N^{\mathtt{W}}_{\mathtt{hwin}}(g_{\mathtt{w}}, d)/N^{\mathtt{W}}_{\mathtt{game}}(g_{\mathtt{w}}, d) \tag{14}$$

The relationship between the expected handicap-deficit probability and the full handicap probability function is examined in Appendix B. This leads to separate convolution expressions for these probabilities in terms of $P^{\mathtt{B}}_{\mathtt{hwin}}$ and $P^{\mathtt{W}}_{\mathtt{hwin}}$.

[2] The EGD tables [1] do not provide this data.

3 LOCAL MODELS

Local models are obtained by fitting the erf sigmoid function to the data for each specific handicap deficit d.

$$S(x) = \frac{1}{2}[1 - \mathrm{erf}(\Lambda(g, d))] \tag{15}$$

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{t=-\infty}^{x} \exp(t^2)\, dt \tag{16}$$

This gives us model parameters which depend (in some initially unkown way) on the parameter d. The measured value of Λ associated with each measured win probability $P^{\mathtt{B}}_{\mathtt{dwin}} = P_{\mathtt{raw}}$ is obtained by inversion [4] of Equation (15) at the point $P_{\mathtt{raw}}$:

$$\lambda^{\mathtt{raw}} = \mathrm{erf}^{-1}(1 - 2P_{\mathtt{raw}}) \tag{17}$$

Plots for $P_{\mathtt{raw}}$ and $\lambda^{\mathtt{raw}}$ are shown in Figure 1. The fixed value of handicap-deficit on each curve is indicated by the subscript $d = 0 \cdots 3$ on the graph labels. These are very noisy, but do provide some insight into the behaviour of winning probability for handicap games. The functional form used for fitting Λ to its raw values is a simple quadratic in grade for each fixed d. The details are presented in Appendix C.

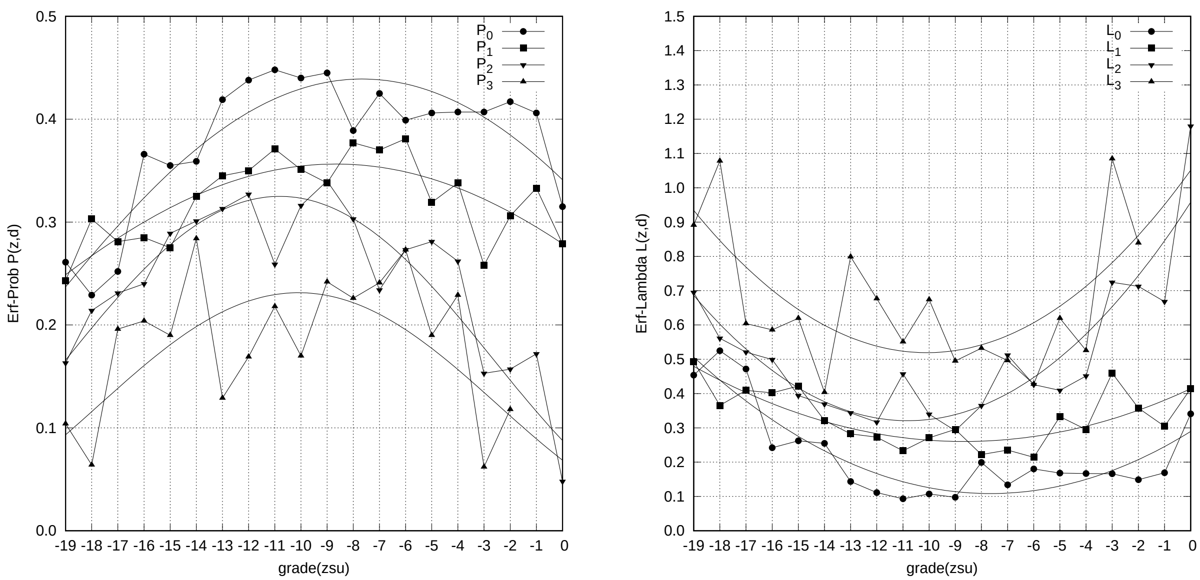

Figure 1: $P_{\mathtt{raw}}$ and $\lambda^{\mathtt{raw}}$ for time window 2014-2018

Notice that the grade scale starts at 19 kyu. Currently, the 20 kyu group contains a large mix of players of different actual grades below 20 kyu, and these unknown grades would significantly distort the fit. The grade scale runs up to shodan - beyond that there are very few games. The handicap-deficit runs from 0 to 3; handicap data beyond $d = 3$ becomes very sparse. The case $d = 0$ shows that players in the range 3 kyu to 12 kyu achieve near to the theoretical performance (0.45) for games with full handicap, but others fall well below.

The local model for Λ can be expressed in the form:

$$\Lambda_d(g) = K2_d\, g^2 + K1_d\, g + K0_d \tag{18}$$

The values for the coefficients are shown in Appendix C. The fit for Λ is exact in the sense that it the solution to a linear system of equations. We can anticipate that the fit to $P_{\mathtt{raw}}$ is less perfect, as a non-linear transform from the Λ model to $P_{\mathtt{win}}$is required. In fact, for the erf model the fit remains good with an average residual 0.034 across all d values. The logistic model gave the same good fit.

4 THE Λ_d COEFFICIENT SIGNAL

The process described in Section 3 evaluates $\Lambda_d(g_{\mathtt{b}})$ by fitting raw data over a time window of 5 years. This process was carried out for each year starting at 1999 through to 2014. The resulting coefficients for each d vary with time as shown in the following plots for each of the coefficients $K2_d$, $K1_d$, $K0_d$:

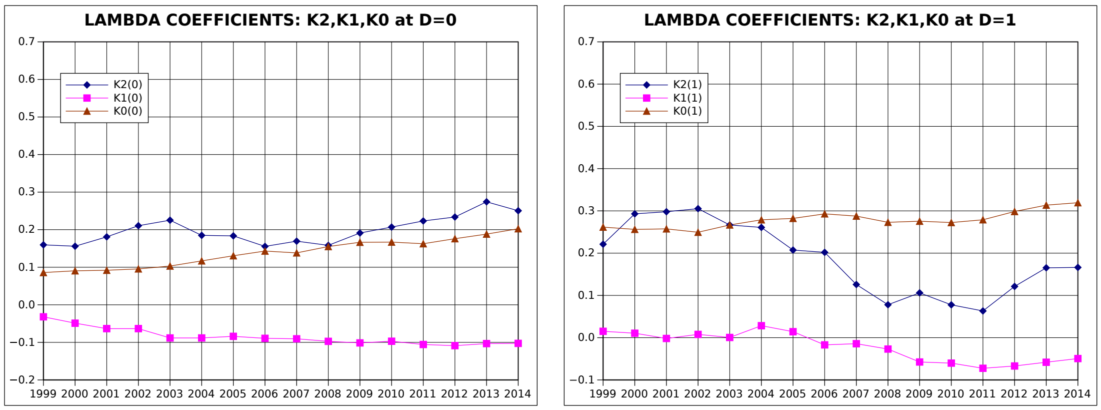

Figure 2: Λ_d coefficient signals at $d = 0$ and $d = 1$

For $d = 0$ or $d = 1$, the $K1(d), K0(d)$ signals are plausible. The signals $K0(d)$ are responsible for the vertical separation of the individual d curves. $K1(d)$ - the smallest of the three - controls the shift in the position of the peak seen in the plot for $P_{\mathtt{raw}}$ in Figure 1.

The $K2(0)$ signal for games with full handicap ($d = 0$), is plausible. But the signal $K2(1)$ for games with handicap reduced by 1, is perplexing to say the least!

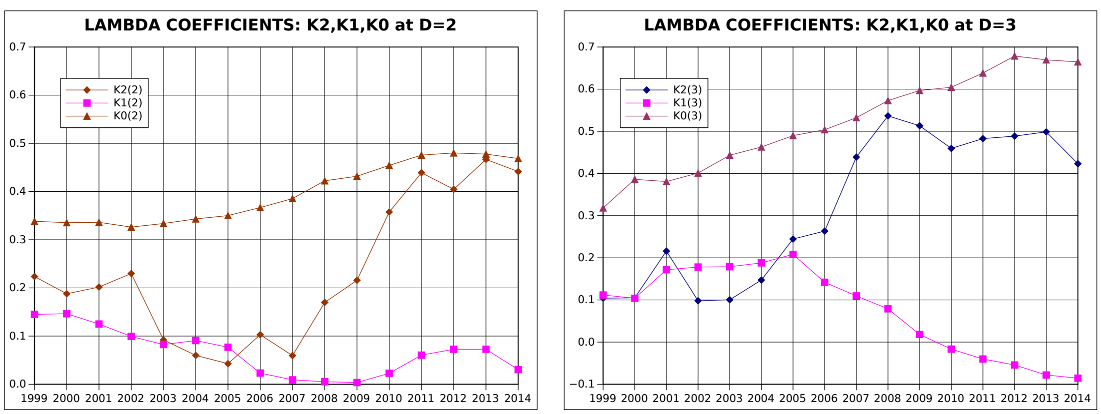

Figure 3: Λ_d coefficient signals at $d = 2$ and $d = 3$

For higher handicap deficit as shown in Figure 3, the $K0(d)$ and $K1(d)$ signals are plausible. The $K2(d)$ signal continues to show erratic behaviour. There is naturally a steady decline in the number of games as the deficit increases, but nothing to suggest a link with the erratic behaviour of $K2(d)$

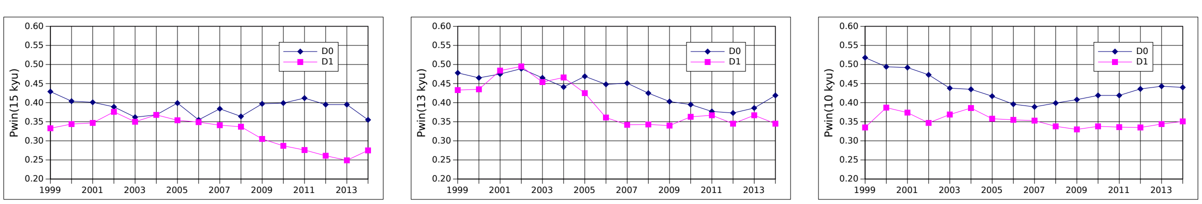

Figure 4: $P_{\mathtt{raw}}$ signals at for 15k,13k, 10k

The time dependent coefficients model the time-dependent raw probabilities shown in Figure 4 for 3 selected grades. The raw probabilities have moved in some case by as much as 20% from maximum to minimum over a period of 20 years.

5 GLOBAL PREVIEW

In this final analysis section, we examine what a global model for $P_{\mathtt{win}}$ might look like. We can achieve this

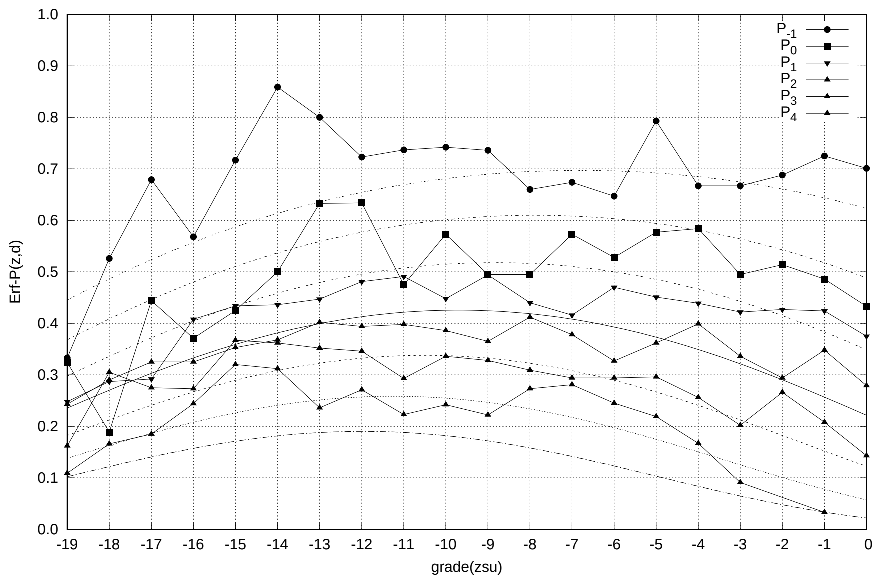

Figure 5: Raw $P_{\mathtt{win}}$ data and rough fit for time window 2006-2018

by fitting a model to an extended time window from 2006 to 2018. This more than doubles our previous time window. We might have expected more smoothing, but we do now get enough games for higher handicap deficits to see win probabilities for a range of d from -1 to 6. In this calculation games deemed to be unreliable (reset or new player with no wins) were excluded from the statistics. For the time series analysis with a 5 year window reported in Section 4, excluding such games made little difference.

The P_{-1} plot in Figure 5 shows $P_{\mathtt{win}}$ for a handicap deficit of -1. In this case the handicap given to Black is 1 more than would be required for full handicap. Note that the peak $P_{\mathtt{win}}$ value produced by the model gives a value $P_{\mathtt{win}}(d=-1)=0.7$ and a value $P_{\mathtt{win}}(d=0)=0.6$ for a 7 kyu player. By comparison, the local model shown in Figure 1 gives a value $P_{\mathtt{win}}(d=0)=0.45$ - closer to the theoretically expected result.

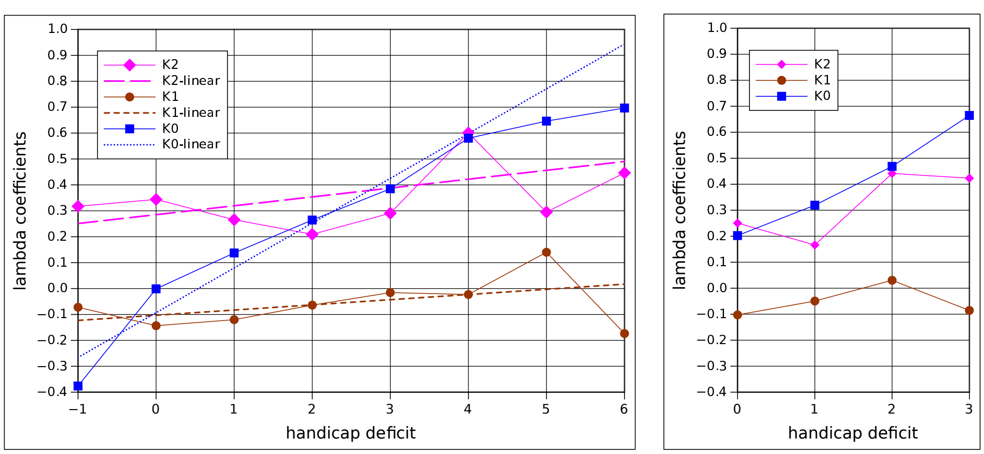

Figure 6: The left graph shows coefficents for 2006-2018, the right one for 2014-2018

The coefficients for the global model were obtained by a linear fit to the coefficients produced by the local models for each value of d from -1 to 4 as depicted by the left hand plot in Figure 6. This is a very quick way to get a first impression of the kind of global model that may be appropriate. For comparison, the right hand plot shows the local coefficients for the time window 2014-2018. In both plots we see that the $K0(d)$ and $K1(d)$ coefficients follow a roughly linear trend with d. But the $K2(d)$ coefficient shows chaotic behaviour in both. The graphs labelled Kn-linear in the left hand plot are linear models for $Kn(d)$ established for d in the range 0 to 4. The game counts are in the hundreds for this range, but drop off dramatically for higher values of d.

The above diagrams are useful for showing how well (or otherwise) our model fits the $P_{\mathtt{win}}$ data, but this picture does not relate very well with the sigmoid nature of $P_{\mathtt{win}}$. The sigmoid model says that the win probability for a player of known grade decreases as the opponent grade (or handicap deficit) increases. Now that we have d values from -1 to 6 we have a chance to see the sigmoid behaviour in our raw data.

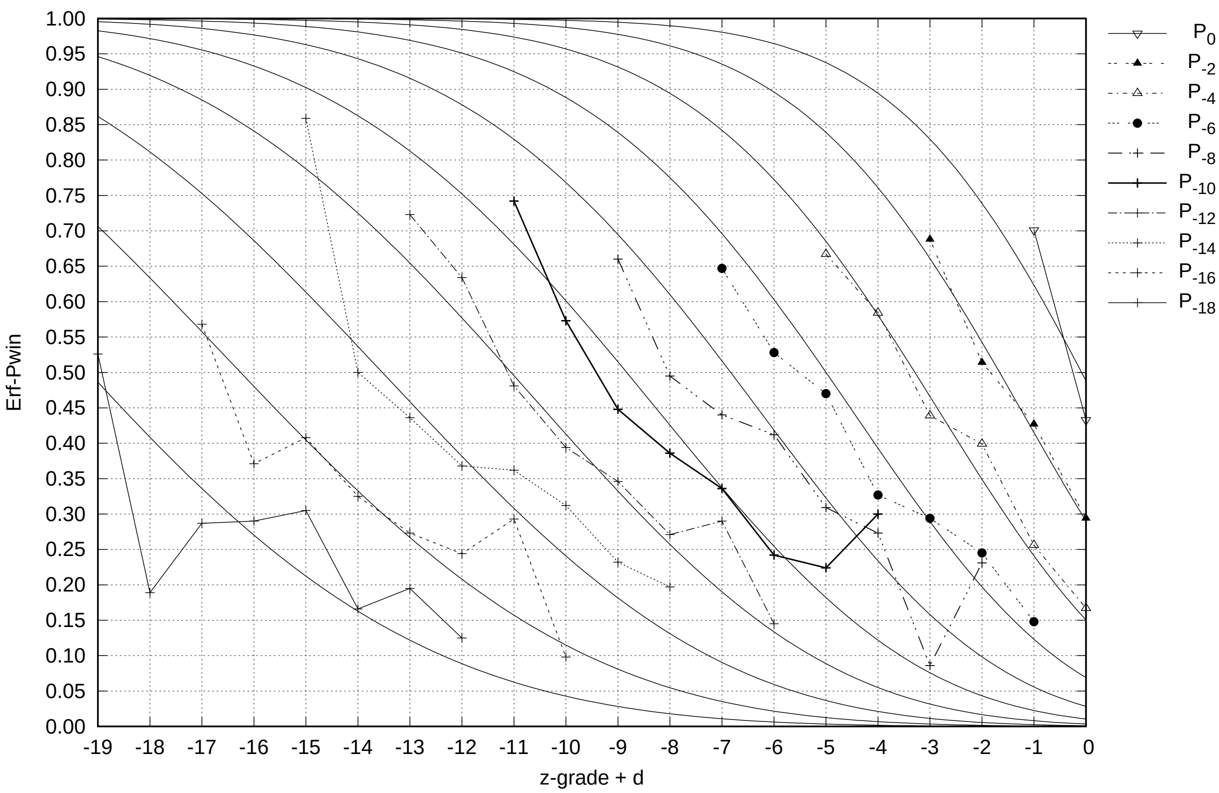

Figure 7: Raw $P_{\mathtt{win}}$ data and sigmoid model for time window 2006-2018

The functional form of a general model which applies to both handicap and even games is:

$$P_{\mathtt{win}}(g_{\mathtt{b}}, d) = \frac{1}{2}[1 - \mathrm{erf}(\Lambda(g_{\mathtt{b}}, d))] \tag{19}$$

We are looking for the behaviour of $P_{\mathtt{win}}$ for fixed values of $g_{\mathtt{b}}$. In this case graphs of $P_{\mathtt{win}}$ produced by Equation (19) for different $g_{\mathtt{b}}$ would overlay each other, so we introduce a transform to separate them out. Define

$$\tilde{\Lambda}(g_{\mathtt{b}},\ g_{\mathtt{b}} + d) = \Lambda(g_{\mathtt{b}}, d) \tag{20}$$

The graphs given in Figure 7 are plots of $\tilde{\Lambda}(g_{\mathtt{b}}, g_{\mathtt{b}} + d)$, for every *even* grade from 18 kyu to 1 dan. It is encouraging that in spite of all the chaos encountered, the model bears a plausible relation to the raw data for these grades.

6 FURTHER WORK

The model for the probability of win has applications in both the rating system and in tournament simulation with the aim of improving pairing algorithms. It should be possible for EGD to provide a single dataset which can serve both purposes. To this end, and to aid the development of the global $P_{\mathtt{win}}$ model further work should anticipate issues that may arise in the processing of the data. Some of these are:

- Unjustified resets.
- Unjustified first time entry grades.
- Players with just one tournament in a year.
- Stability of grade distributions.

An examination of these will be undertaken with a view to specifying an efficient dataset useful for EGF researchers in these fields.

The EGD tournament submission format is fairly loose. It is not necessary to specify the exact handicap in a game, instead it is deduced from $g_{\mathtt{white}} - g_{\mathtt{black}} - H_{\mathtt{reduced}}$. However, in most McMahon tournaments the handicap is usually based on MMS difference. This of course can be more than 1 stone different from the grade difference.

A DATASET FOR $P_{\texttt{win}}(g_{\texttt{b}}, g_{\texttt{w}}, h)$

A.1 Player Game Data

The low level data consists of a sequence of *games*. The essential data required to represent a game (denoted by γ) can be expressed as a five-tuple:

$$\gamma = (t, i, j, h, s) \tag{21}$$

In this formulation, t is the time[3] of the game, i identifies the black player, j is the white player, and s is Black's score. The handicap given to Black is $h \geq 0$, and if $h = 0$, the game is *even*.

We define the handicap given to White as $h_{\texttt{W}}$ and White's score as $s_{\texttt{W}}$. Ignoring Jigo, these satisfy the relations:

$$h_{\texttt{B}} + h_{\texttt{W}} = 0 \tag{22}$$

$$s_{\texttt{B}} + s_{\texttt{W}} = 1 \tag{23}$$

It is convenient to introduce projection operators T, $\mathsf{G_B}$, $\mathsf{G_W}$, $\mathsf{H_B}$, $\mathsf{H_W}$, $\mathsf{S_B}$, $\mathsf{S_W}$ to extract details about any given game γ as represented in Equation (21). They are defined as follows:

Operator	Description
$\mathsf{T}\gamma = t$	The time of the game.
$\mathsf{G_B}\gamma = \mathsf{G}(i)$	Grade of the player taking black.
$\mathsf{G_W}\gamma = \mathsf{G}(j)$	Grade of the player taking white.
$\mathsf{H_B}\gamma = h$	Black's handicap.
$\mathsf{H_W}\gamma = -h$	White's handicap.
$\mathsf{S_B}\gamma = s$	Black's score.
$\mathsf{S_W}\gamma = 1 - s$	White's score.

Table 1: Extraction operators

A.2 Grade Group Data - Specific Handicap

We combine all the game data for individual players to produce winning probabilities for each grade group. The datasets are the number of games played and the number of games won for each combination of $(g_{\texttt{b}}, g_{\texttt{w}}, h)$. For convenience we express grades in *zero-shodan-units(zsu)*[4]. The minimum grade is $g_{\texttt{min}}$, the maximum grade is $g_{\texttt{max}}$[5]. We can also limit handicaps to 9 as there are very few games with a higher handicap. Furthermore, as the collection of all the games in the database spans a period of more than 2 decades, we will be measuring $P_{\texttt{win}}$on a rolling window for much shorter durations. For the moment however, we can drop the appearance of the time parameter in order to simplify the notation.

Let $\Gamma_{\texttt{game}}$ be the set of all games in the time period. Let $\Gamma_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{w}}, h)$ be the set of all games won by Black between players i and j with grades $g_{\texttt{b}}$ and $g_{\texttt{w}}$ and h the handicap given to black. Likewise $\Gamma_{\texttt{hlose}}$ is the set of games where Black loses. It follows from the definitions in Equation (21) and Equation (23) that:

$$\Gamma_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{w}}, h) = \{\gamma \in \Gamma_{\texttt{game}} : \mathsf{G_B}\gamma = g_{\texttt{b}},\ \mathsf{G_W}\gamma = g_{\texttt{w}},\ \mathsf{H_B}\gamma = h,\ \mathsf{S_B}(\gamma) = 1\} \tag{24}$$

$$\Gamma_{\texttt{hlose}}(g_{\texttt{b}}, g_{\texttt{w}}, h) = \{\gamma \in \Gamma_{\texttt{game}} : \mathsf{G_B}\gamma = g_{\texttt{b}},\ \mathsf{G_W}\gamma = g_{\texttt{w}},\ \mathsf{H_B}\gamma = h,\ \mathsf{S_B}(\gamma) = 0\} \tag{25}$$

$$\Gamma_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{w}}, h) \bigcup \Gamma_{\texttt{hlose}}(g_{\texttt{b}}, g_{\texttt{w}}, h) = \Gamma_{\texttt{game}}(g_{\texttt{b}}, g_{\texttt{w}}, h) \tag{26}$$

[3]The tournament code gives time and acccess to other features which might affect $P_{\texttt{win}}$.
[4]The zero point is Shodan. Kyu grades decrease by 1, Dan grades increase by 1 from zero.
[5]Typical values might be $g_{\texttt{min}} = -20$, $g_{\texttt{max}} = 7$.

We can now identify the raw components of the data for the $P_{\texttt{win}}$ evaluation. These are just the number N of members in each of the above sets, i.e. the set cardinality. The defintions are:

$$N^{\texttt{B}}_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{w}}, h) = |\Gamma_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{w}}, h)| \tag{27}$$

$$N^{\texttt{B}}_{\texttt{hlose}}(g_{\texttt{b}}, g_{\texttt{w}}, h) = |\Gamma_{\texttt{hlose}}(g_{\texttt{b}}, g_{\texttt{w}}, h)| \tag{28}$$

$$\texttt{N}_{\texttt{game}}(g_{\texttt{b}}, g_{\texttt{w}}, h) = |\Gamma_{\texttt{game}}(g_{\texttt{b}}, g_{\texttt{w}}, h)| = \texttt{N}_{\texttt{hwin}} + \texttt{N}_{\texttt{hlose}} \tag{29}$$

B HANDICAP-DEFICIT CONVOLUTION

Recalling the definition of the handicap-deficit in Equation (1), the data required is expressed in terms of functions of just two variables: $N^{\texttt{B}}_{\texttt{game}}(g_{\texttt{b}}, d)$, $N^{\texttt{W}}_{\texttt{game}}(g_{\texttt{w}}, d)$, $N^{\texttt{B}}_{\texttt{hwin}}(g_{\texttt{b}}, d)$, and $N^{\texttt{W}}_{\texttt{hwin}}(g_{\texttt{w}}, d)$, where d is the handicap deficit. These functions are obtained from the full handicap counts $N^{\texttt{B}}_{\texttt{hwin}}$, $N^{\texttt{W}}_{\texttt{hwin}}$, and $\texttt{N}_{\texttt{game}}$ via a summation over variable h with $g_{\texttt{w}} = g_{\texttt{b}} + h + d$ keeping $g_{\texttt{b}}$ and d fixed. This leads to a convolution expession for the handicap-deficit probability function.

Firstly, the number of games at fixed Black grade and handicap-deficit:

$$N^{\texttt{B}}_{\texttt{game}}(g_{\texttt{b}}, d) = \sum_{h=h_{\min}}^{h=h_{\max}} \texttt{N}_{\texttt{game}}(g_{\texttt{b}},\ g_{\texttt{b}} + h + d,\ h) \tag{30}$$

Since all grades must lie in the range $g_{\texttt{min}} \leq g \leq g_{\texttt{max}}$, some values of h may be excluded depending on the values of $g_{\texttt{b}}$ and d. If we set $d = -1$, we can capture games where Black is taking slightly more handicap than usual i.e. $g_{\texttt{w}} - g_{\texttt{b}} < h$. This will help to estimate the value of $P_{\texttt{win}}(g_{\texttt{b}}, -\frac{1}{2})$, which in theory should have the value $\frac{1}{2}$.

Next consider the summation for $N^{\texttt{B}}_{\texttt{hwin}}$:

$$N^{\texttt{B}}_{\texttt{hwin}}(g_{\texttt{b}}, d) = \sum_{h=h_{\min}}^{h=h_{\max}} N^{\texttt{B}}_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{b}}+h + d, h) \tag{31}$$

Now from Equation (2) in Section 2 we see that

$$N^{\texttt{B}}_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{b}} + h + d,\ h) = P^{\texttt{B}}_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{b}} + h + d, h) \times N^{\texttt{B}}_{\texttt{game}}(g_{\texttt{b}}, g_{\texttt{b}}+h + d, h) \tag{32}$$

So from Equation (13) and Equation (31) we obtain:

$$P^{\texttt{B}}_{\texttt{dwin}}(g_{\texttt{b}}, d) = \sum_{h=h_{\min}}^{h=h_{\max}} \eta(g_{\texttt{b}}, g_{\texttt{b}} + h + d, h) \times P^{\texttt{B}}_{\texttt{hwin}}(g_{\texttt{b}}, g_{\texttt{b}} + h + d, h) \tag{33}$$

$$\text{where } \eta(g_{\texttt{b}}, g, h) = \texttt{N}_{\texttt{game}}(g_{\texttt{b}}, g, h) \big/ N^{\texttt{B}}_{\texttt{game}}(g_{\texttt{b}}, d) \tag{34}$$

The same procedure can be applied to obtain $P^{\texttt{W}}_{\texttt{hwin}}(g_{\texttt{w}}, d)$. In this case we are keeping $g_{\texttt{w}}$ and d fixed and summing over h with $g_{\texttt{b}} = g_{\texttt{w}} - d - h$. So the number of games for given $g_{\texttt{w}}$ and d is

$$N^{\texttt{W}}_{\texttt{hwin}}(g_{\texttt{w}}, d) = \sum_{h=h_{\min}}^{h=h_{\max}} N^{\texttt{W}}_{\texttt{hwin}}(g_{\texttt{w}} - d - h, g_{\texttt{w}}, h) \tag{35}$$

From Equations (3), (35), and (31), we find:

$$P^{\texttt{W}}_{\texttt{dwin}}(g_{\texttt{w}}, d) = \sum_{h=h_{\min}}^{h=h_{\max}} \eta(g_{\texttt{w}} - d - h, g_{\texttt{w}}, h) \times P^{\texttt{W}}_{\texttt{hwin}}(g_{\texttt{w}} - d - h, g_{\texttt{w}}, h) \tag{36}$$

$$\text{where } \eta(g, g_{\texttt{w}}, h) = \texttt{N}_{\texttt{game}}(g, g_{\texttt{w}}, h) \big/ N^{\texttt{W}}_{\texttt{game}}(g_{\texttt{w}}, d) \tag{37}$$

C LOCAL Λ MODEL

A standard least squares (LSQ) fitting algorithm is used to fit the model in Equation (18), Section 3. We minimise the square residual

$$R(k2, k1, k0) = \sum_{g=g_{\mathtt{min}}}^{g=g_{\mathtt{max}}} [\Lambda_d(g) - \lambda_{\mathtt{dg}}^{\mathtt{raw}})]^2 \quad (38)$$

The LSQ process produces a matrix equation whose elements are $\sum_g g^m\, g^n$. If g is transformed to lie in the range $-1 \leq g \leq 1$ then these elements can be approximated by integrals of orthogonal polynomials. This renders the matrix more diagonalised. This is especially useful if later we are required to develop non-linear fitting techniques to improve the fit.

Consequently the following transforms are introduced for the Λ model on each dataset for fixed d:

$$g_{\mathtt{mid}} = (g_{\mathtt{max}} + g_{\mathtt{min}})/2 \qquad g_{\mathtt{size}} = (g_{\mathtt{max}} - g_{\mathtt{min}} - 1)/2 \quad (39)$$

$$r(g, d) = (g - g_{\mathtt{mid}} + \frac{1}{2}d)/g_{\mathtt{size}} \quad (40)$$

$$\Lambda_d(g) = K2_d\,(r(g, d)^2 - \frac{1}{3}) + K1_d\, r(g, d) + K0_d \quad (41)$$

This transform ensures that $r(g, d)$ lies in the required range. The $\frac{1}{3}$ term is needed to make the polynomials $r^2 - \frac{1}{3}, r^1, r^0$ orthogonal.

The coefficients and residuals obtained for the three time periods reported in Figure 4 are:

coefficient	1999-2003	2005-2009	2014-2018
$K2_0$	0.159604	0.183485	0.250406
$K2_1$	0.221138	0.207311	0.166217
$K2_2$	0.223642	0.042741	0.441478
$K2_3$	0.104156	0.244479	0.423155
$K1_0$	-0.032017	-0.083937	-0.102667
$K1_1$	0.015072	0.014209	-0.049421
$K1_2$	0.145025	0.076889	0.030466
$K1_3$	0.111765	0.208150	-0.085558
$K0_0$	0.085855	0.130542	0.202454
$K0_1$	0.261589	0.282214	0.319359
$K0_2$	0.338246	0.350129	0.468540
$K0_3$	0.318277	0.489726	0.664631
residual	0.038	0.041	0.034

Table 2: Λ_d LSQ Coefficients

The residual quoted is the average of the individual residuals for all the d values in the time period.