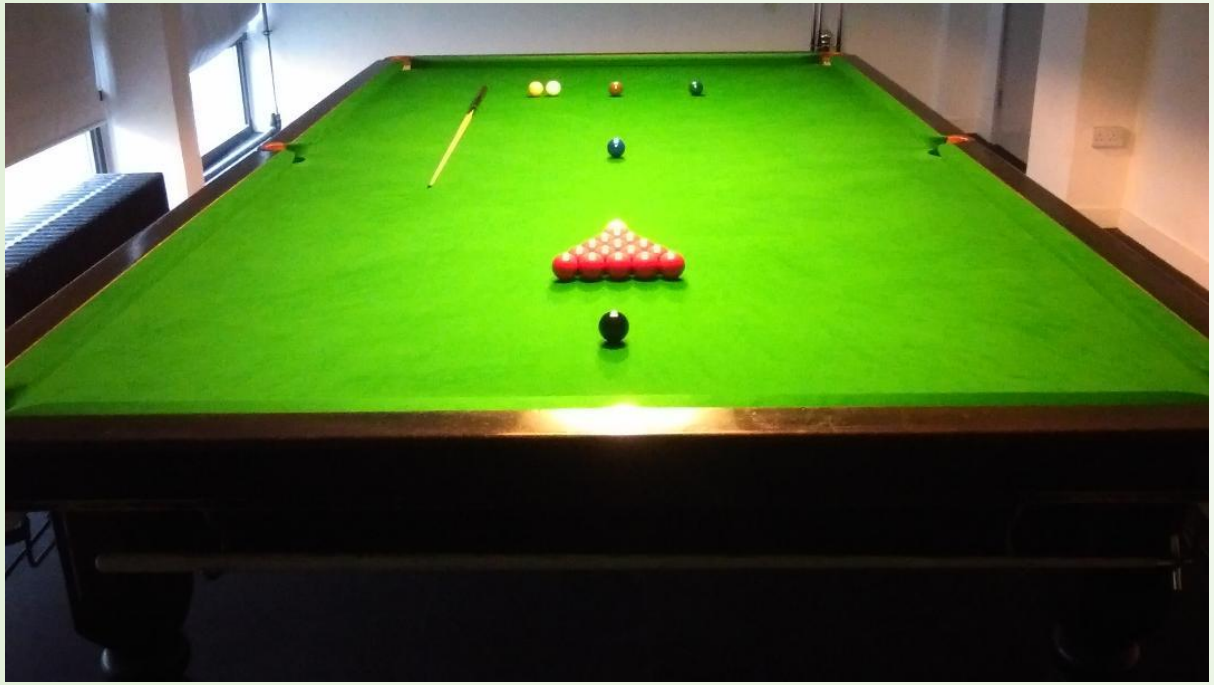
Modelling the outcomes of professional Snooker matches



James Collingwood
Mike Wright
Roger Brooks

Lancaster University

MathSport International 2019
1st July 2019

Research Aims

1. Develop a rating system which provides a reliable indication of the **relative ability of different players** and enables us to estimate the underlying probability that one player will beat another.
   - 4 types of model tested
   - Discussion of results: main limitations and differences between models

2. Understand what effect **current form / momentum** has on the outcome of the match / outcome of the next frame.

3. Review the use of **performance statistics** in snooker and their potential for explaining the outcome of a match and / or highlighting differences in ability between players.

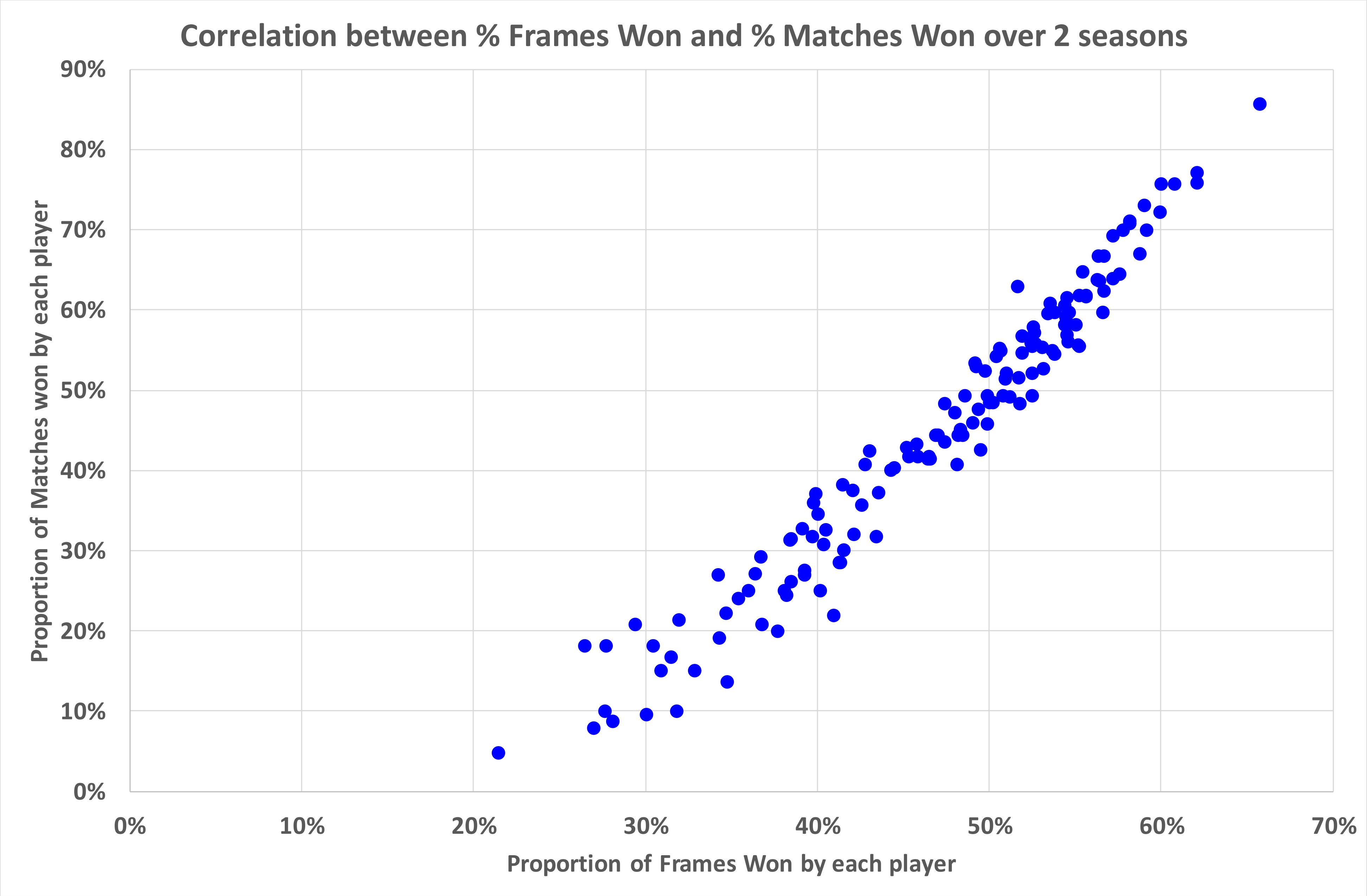
Methods of rating and ranking players

Official World Rankings

- 128 professional players
- 20 ranking events per season (all knockout competitions)
- Rankings based on prize money won over the last 2 seasons
- Officially updated around 10 times per season

Win Percentage

- Proportion of frames won by each player
- Allows for modelling of different lengths of match
- Very strongly correlated with proportion of matches won


Correlation between % Frames Won and % Matches Won over 2 seasons
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
0%
10%
20%
30%
40%
50%
60%
70%
Proportion of Matches won by each player
Proportion of Frames Won by each player

Modelling the outcome of matches

- Models used to estimate the probability of winning a frame P(F) against another player
- Probability of winning a match then derived as a series of Bernoulli trials

Prize Money model: $\boldsymbol{P(F)_1 = 0.5 +} \{\ln(\boldsymbol{PM_1}) - \ln(\boldsymbol{PM_2})\} * \boldsymbol{0.04707}$

Win Percentage model: $\boldsymbol{P(F)_1 = 0.5 +} \{\boldsymbol{WP_1} - \boldsymbol{WP_2}\}$ (capped at 0 and 1)

Bradley-Terry model: MLE based on wins and losses against each individual

Elo model: Reflects current ability rather than performance over a given period

(logistic distribution: standard deviation = 500, weight = 10)

Prediction Accuracy

Proportion of matches played in ranking events during 2017/18 and 2018/19 (4,430) won by the player with the higher rating

Model	Correct Predictions	% Correct
Win%_2 year	3,049	68.8
BT_2 year	3,048	68.8
Elo	3,047	68.8
Prize Money	3,024	68.3
BT_1 year	3,017	68.1
Win%_1 year	3,004	67.8

All models – predicted same winner for 82% matches, with 72.2% success rate

Top 3 models – predicted same winner for 91% matches, with 70.5% success rate

Calibration measure

Ratio of expected over actual wins for higher-rated player in each match (ideal = 1.00)

Model	Frames Won	Matches Won
Win%_2 year	1.01	1.03
Elo	1.02	1.04
Win%_1 year	1.02	1.05
BT_2 year	1.03	1.05
Prize Money	1.04	1.06
BT_1 year	1.04	1.06

All models have a bias towards the higher-rated players (i.e. predict that they will win more matches than they actually do)

Modelling of ‘new’ players

Official Rankings:

- New players start with £0, all Amateur players are unranked (i.e. £0)

[£0 modelled as £250 to enable logs to be taken]

Win Percentage & Bradley-Terry models:

- Players given an individual rating after 10 matches
- Combined rating for players contesting <10 matches based on aggregated results

Elo model:

- Players allocated a start rating which is subsequently updated
- Differentiation between professionals and amateur players

Calibration scores based on experience of players (1)

Ratio of expected and actual wins where player had contested **X** matches over the past 2 years

Matches played in last 2 years	#	% won	PM	Win%_2	BT_2	Win%_1	BT_1	Elo
< 10	524	22%	0.56	0.67	0.60	0.85	0.78	0.94
10 - 20	329	27%	0.77	0.95	0.95	0.86	0.85	0.87
20 - 50	1,537	33%	0.97	0.93	0.95	0.91	0.94	0.94
50 - 100	5,390	55%	1.03	1.03	1.02	1.03	1.01	1.01
100+	1,080	70%	1.00	0.99	1.04	0.99	1.05	1.03

All models tend to under-estimate chances of less-experienced players.

Calibration scores based on experience of players (2)

Matches played in last 2 years	PM	Win%_2	BT_2	Win%_1	BT_1	Elo
< 10	**0.56**	**0.67**	**0.60**	**0.85**	**0.78**	**0.94**
10 - 20	**0.77**	**0.95**	**0.95**	**0.86**	**0.85**	**0.87**
20 - 50	0.97	0.93	0.95	0.91	0.94	0.94

1. Prize Money (PM) model heavily under-estimates players who have contested fewer than 20 matches in past 2 years

2. Elo model is the least biased for players who have contested fewer than 10 matches, although not as strong for players contesting 10 – 20 matches

3. For unrated players, models based on 1 year of results are less biased than models based on 2 years of results – although the 2-year models are less biased for those contesting 10-20 matches.

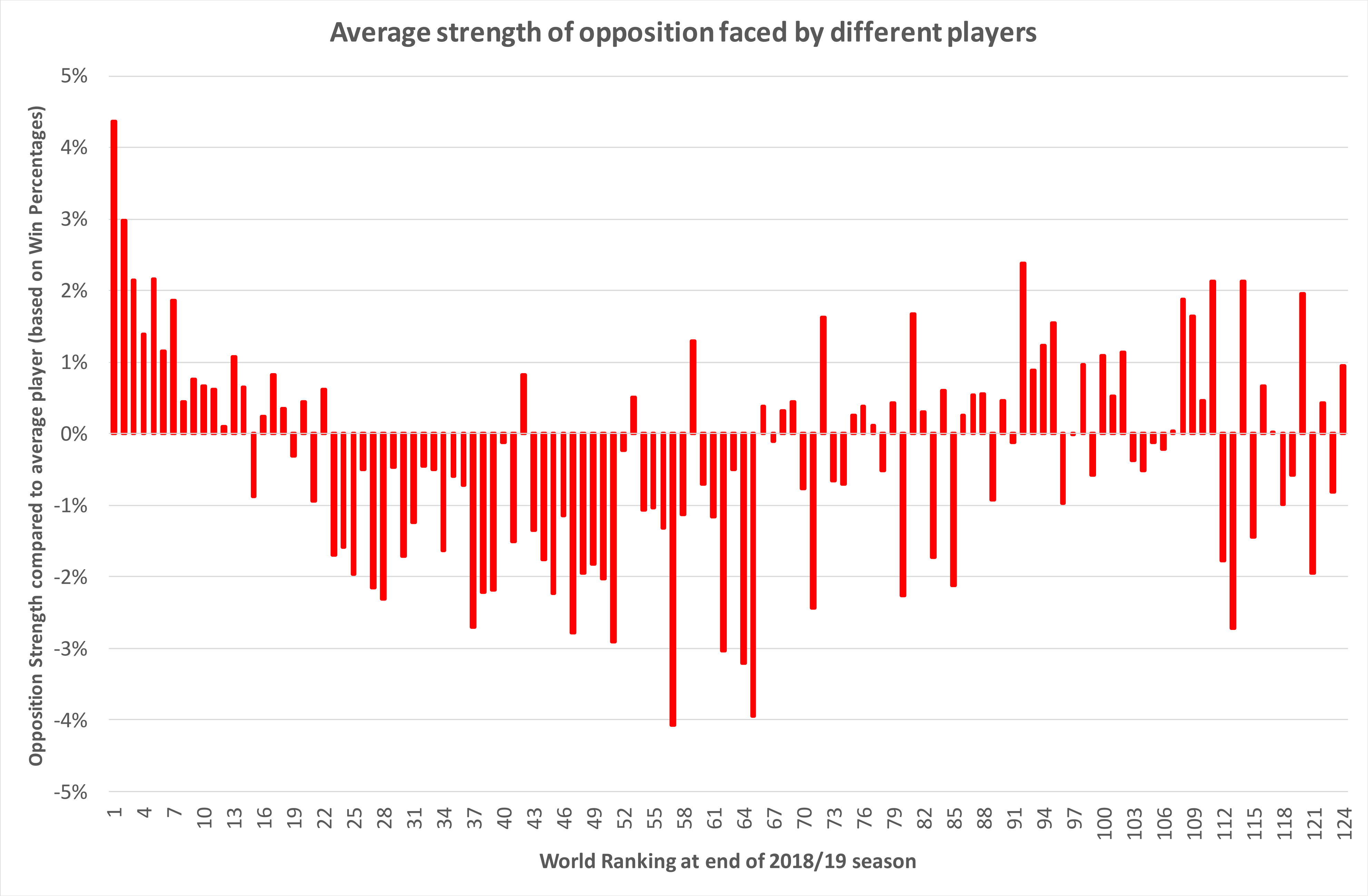
Strength of opposition

Anticipated limitation of Win Percentage model is that it doesn't take into account strength of opposition faced.

Bradley-Terry model produces a relative rating based on wins and losses against each player, which effectively takes this into account.

A measure of Strength of opposition is the weighted average of each opponent's Win Percentage – e.g. …

	Frames Played	Opponent Rating	FP * OR
Opponent 1	10	60%	6
Opponent 2	6	50%	3
Weighted average = ∑ (FP * OR) / FP =		**56.25%**	


Average strength of opposition faced by different players
Opposition Strength compared to average player (based on Win Percentages)
5%
4%
3%
2%
1%
0%
-1%
-2%
-3%
-4%
-5%
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61 64 67 70 73 76 79 82 85 88 91 94 97 100 103 106 109 112 115 118 121 124
World Ranking at end of 2018/19 season

Win Percentage v Bradley-Terry

For each match played, estimate the strength of opponent faced by either player over the last 2 seasons

- 543 cases where opponent strength > 52.5%
(player's Win Percentage is potentially an under-estimate of their performance)

- 446 cases where opponent strength < 47.5%
(player's Win Percentage is potentially an over-estimate of their performance)

	Calibration (Matches won)	
	Win%_ 2 yr model	**BT_ 2yr model**
Opponent strength > 52.5%	0.97	1.06
Opponent strength < 47.5%	1.03	0.90

Form

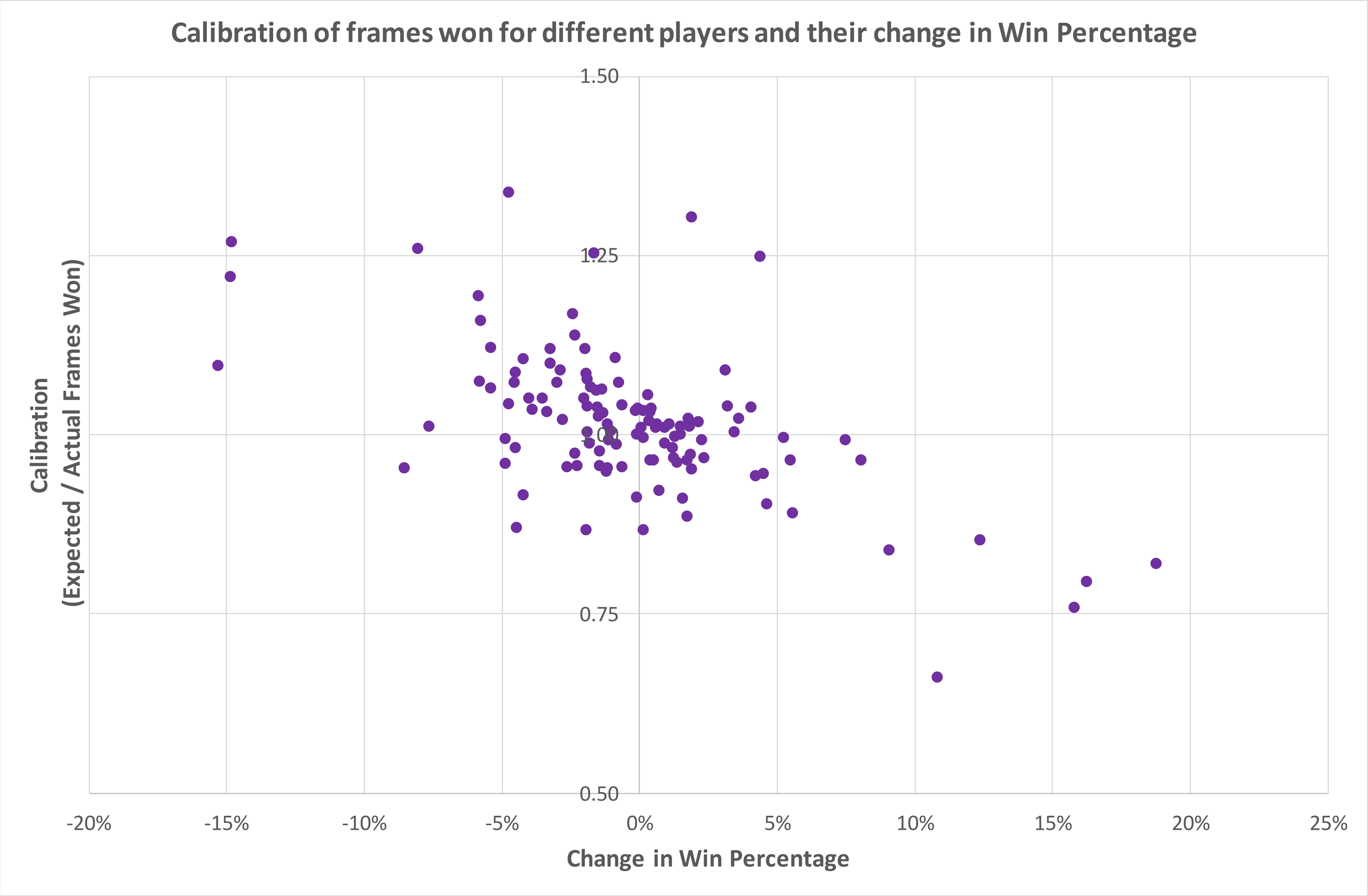
Analysis of individual players

X – axis:

Comparison of Win Percentage between start of 2017/18 and end of 2018/19

Y – axis:

Expected / Actual frames won (Win%_2 yr model)



2 years v 1 year

For each match played, compare each player’s 1-year Win % and 2-year Win %

- 428 cases where 1-year Win % is > 3% higher than 2-year Win %
(2-year model potentially under-estimates current level of performance)

- 734 cases where 1-year Win % is < -3% lower than 2-year Win %
(2-year model potentially over-estimates current level of performance)

	Calibration (Matches won)		
	Win%_ 2 yr model	**Win%_ 1yr model**	**Elo**
1-yr > 2-yr	0.93	1.10	1.03
2-yr > 1-yr	1.04	0.85	0.91

Conclusions

Modelling 'New' Players

Look for earliest viable time to base ratings on individual performance

Try and differentiate between unranked players

Strength of Opposition

Looks to be a significant factor but may need to be captured in another way

Current Form

A potentially key factor which is not fully taken into account by existing models.

Could be more significant for positive changes in performance

Performance measures

Alternative approach would be to derive a player rating from different components of their performance.

Data limited to 2 events per season, and 2 meaningful measures.

Based on 146 matches a plausible model would be:

$$\begin{aligned}\boldsymbol{P(F)_1} &= \mathbf{0.5}\\ &\quad + \mathbf{2.091} \times [\boldsymbol{Pot\ Success_1} - \boldsymbol{Pot\ Success_2}]\\ &\quad + \mathbf{0.717} \times [\boldsymbol{Safety\ Success_1} - \boldsymbol{Safety\ Success_2}]\end{aligned}$$