Addendum:

Review of

Method Proposals to Calculate Best Lifter Scores (Relative Scores) in IPF Powerlifting Competitions

Submitted to the IPF to replace the Wilks coefficients

Reviewers: Dr. C. Maiwald[1] | Dr. T. Mayer[1,2]

[1] Chemnitz University of Technology, Department of Research Methodology and Data Analysis in Biomechanics
[2] TecStat Analytics, Werdau

Chemnitz, Werdau 29.10.2018

Contents

1 Purpose of this Addendum 3
1.1 Proposal description: Author 5 3
1.2 Comparison of Author 5 to the previously evaluated proposals 3
2 Comparison of the scoring methodologies 4
2.1 Men's classic bench press (MEN.CL.BP) 4
2.2 Men's classic powerlifting (MEN.CL.PL) 7
2.3 Women's classic bench press (WMN.CL.BP) 10
2.4 Women's classic powerlifting (WMN.CL.PL) 13
3 Conclusion 16
4 R Code used for calculating relative scores 17
4.1 Author 2 17
4.1.1 PL: 17
4.1.2 BP: 17
4.2 Marksteiner 17
4.2.1 PL: 17
4.2.2 BP: 17
4.3 Author 4 18
4.3.1 MEN PL: 18
4.3.2 MEN BP: 18
4.3.3 WMN PL: 18
4.3.4 WMN BP: 18
4.4 Author 5 19
4.4.1 MEN CL PL: 19
4.4.2 MEN CL BP: 19
4.4.3 WMN CL PL: 19
4.4.4 WMN CL BP: 19

1 Purpose of this Addendum

An additional proposal was submitted by Author 5 after the initial evaluation. It shall hereafter be referred to as **Author 5** hereafter. The question arose as to how this approach would perform compared to the previously evaluated methods.

1.1 Proposal description: Author 5

In contrast to the other proposals, Author 5 does not contain a presentation of a theoretical or scientific basis. It is merely a statement about a new formula and a proposal on certain coefficients that should be used for men and women.

This proposal lacks the scientific foundation or elaborated theory on the statistical approach that is chosen. It also lacks a description of how the coefficient was determined that is given in the formula:

Score = Total / atan (bodyweight / coefficient)

Although the simplicity of the formula is compelling, we see no theoretical justification for using a trigonometric approach. Due to the lack of information given in the proposal, we are unable to comment on the quality of the scientific foundation of this approach.

1.2 Comparison of Author 5 to the previously evaluated proposals

Although little theoretical information is provided by Author 5, we exposed the proposed relative scores to the same criteria as those used previously for evaluating the other proposals.

1. Relative Scoring Distribution: Using a χ^2 summary statistic (cumulative sum of χ^2 values across percentile groups), **distributional fairness** was evaluated as described in the original report.
2. LOESS fit deviations were used to evaluate **scoring validity** in the sense that top performances within each scoring systems should result in similar score magnitudes across weight classes.

Due to the lack of a formula for the fitted model, we cannot plot the predicted average score for this method. Consequently, model fit plots have been eliminated from this addendum. Furthermore, Author 5 only provides a formula for calculating classic scores. We limited the comparisons to the datasets from classic (CL) results, since the results of equipped competitions (EQ) are not addressed by Author 5.

For the two above described evaluation criteria (distributional fairness & scoring validity) applies: **The lower the score, the better the method performs.**

2 Comparison of the scoring methodologies

2.1 Men's classic bench press (MEN.CL.BP)

Table 1: Summary statistics for men's classic bench press

	Wilks	Author 2	Marksteiner	Author 4	Author 5
Distribution χ^2 sum	829.96483	1014.02677	280.71897	772.38166	886.85953
LOESS residual sum	2.06503	1.91810	1.70323	1.83663	1.96343

Marksteiner's scoring represents the shape of the underlying population best for the distribution of relative scores (χ^2 sum Table 1 & Figure 2). **Author 5 takes fourth place for this evaluation criterion.** LOESS residual sums indicate that Marksteiner scores are most level across all performance bands (Table 1). **For LOESS scores, Author 5 is again fourth.** Figure 1 on the next page depicts LOESS-plots with relative scoring across weight classes and performance bands in MEN CL BP.

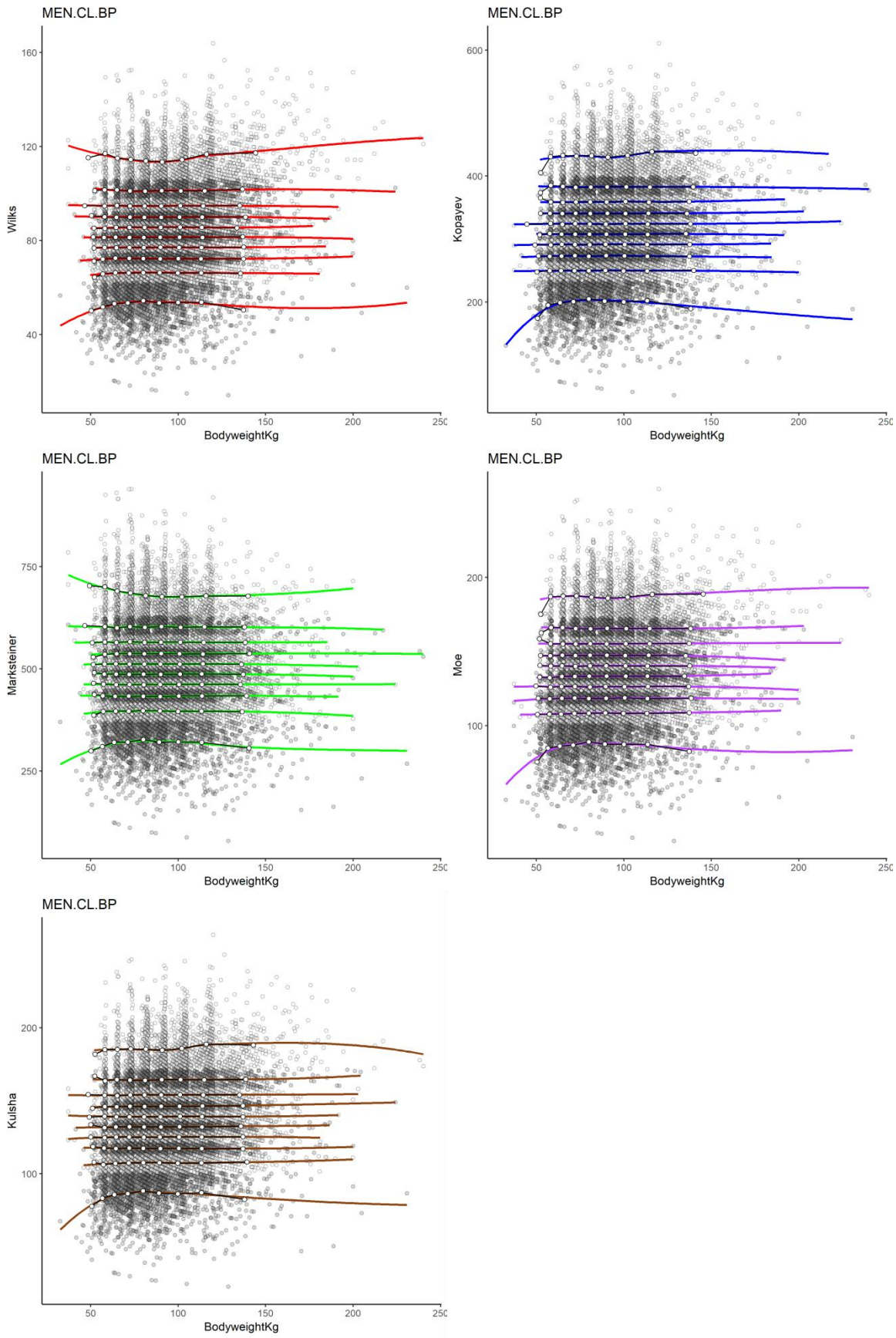

Figure 1: LOESS-plot for men's classic bench press

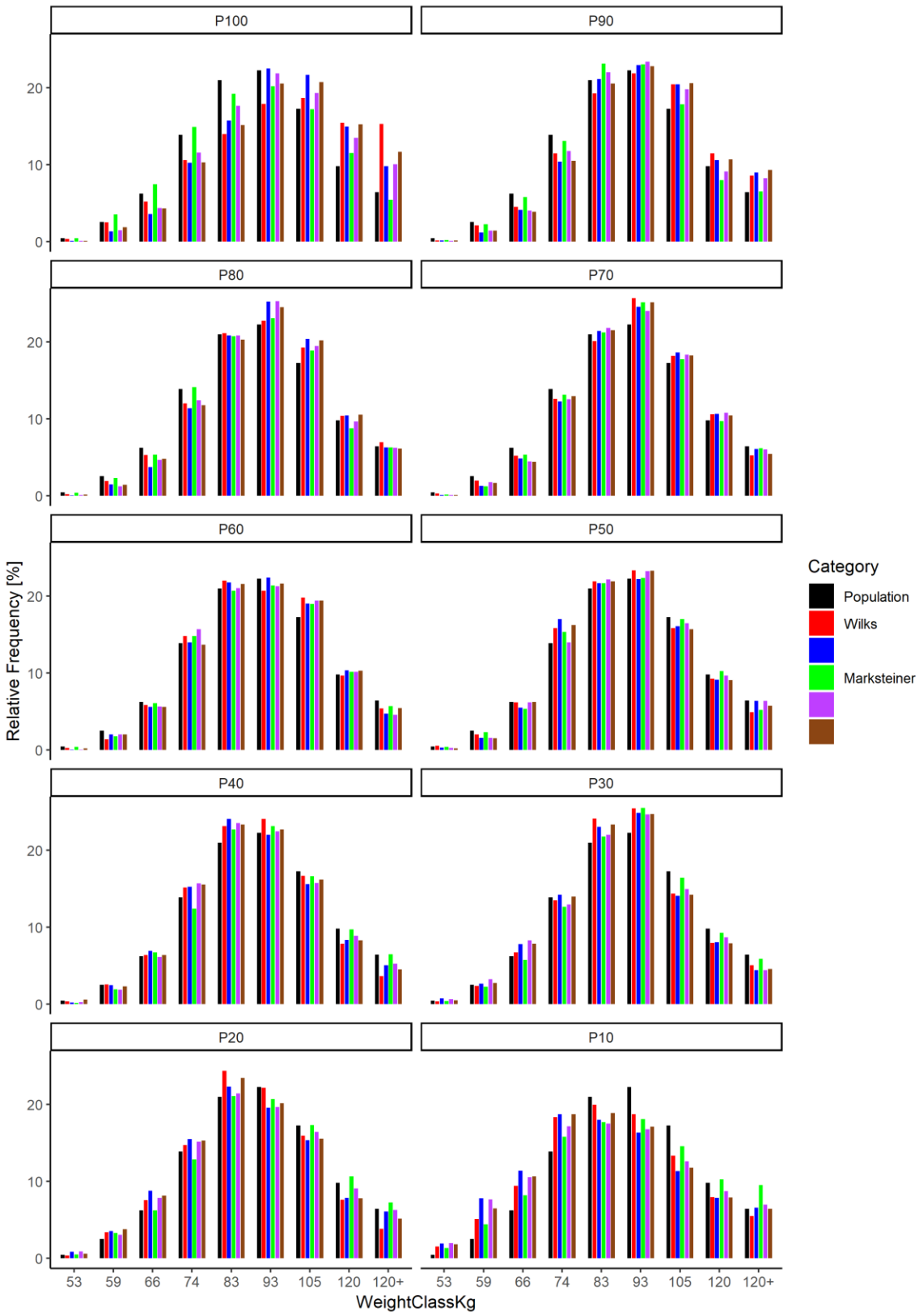

Figure 2: Distribution of relative scores across performance bands for men's classic bench press

2.2 Men's classic powerlifting (MEN.CL.PL)

Table 2: Summary statistics for men's classic powerlifting

	Wilks	Author 2	Marksteiner	Author 4	Author 5
Distribution χ^2 sum	300.97666	426.17741	333.56628	485.76790	409.48174
LOESS residual sum	1.74409	1.51711	1.45324	1.86049	1.72047

Wilks scoring represents the shape of the underlying population best for the distribution of relative scores in MEN.CL.PL (Table 2 & Figure 4). **Author 5 follows Marksteiner and takes third place.** LOESS residual sums indicate that Marksteiner scores are most level across all performance bands (Table 2 & Figure 3). **For LOESS residual sum, Author 5 again takes third place.**

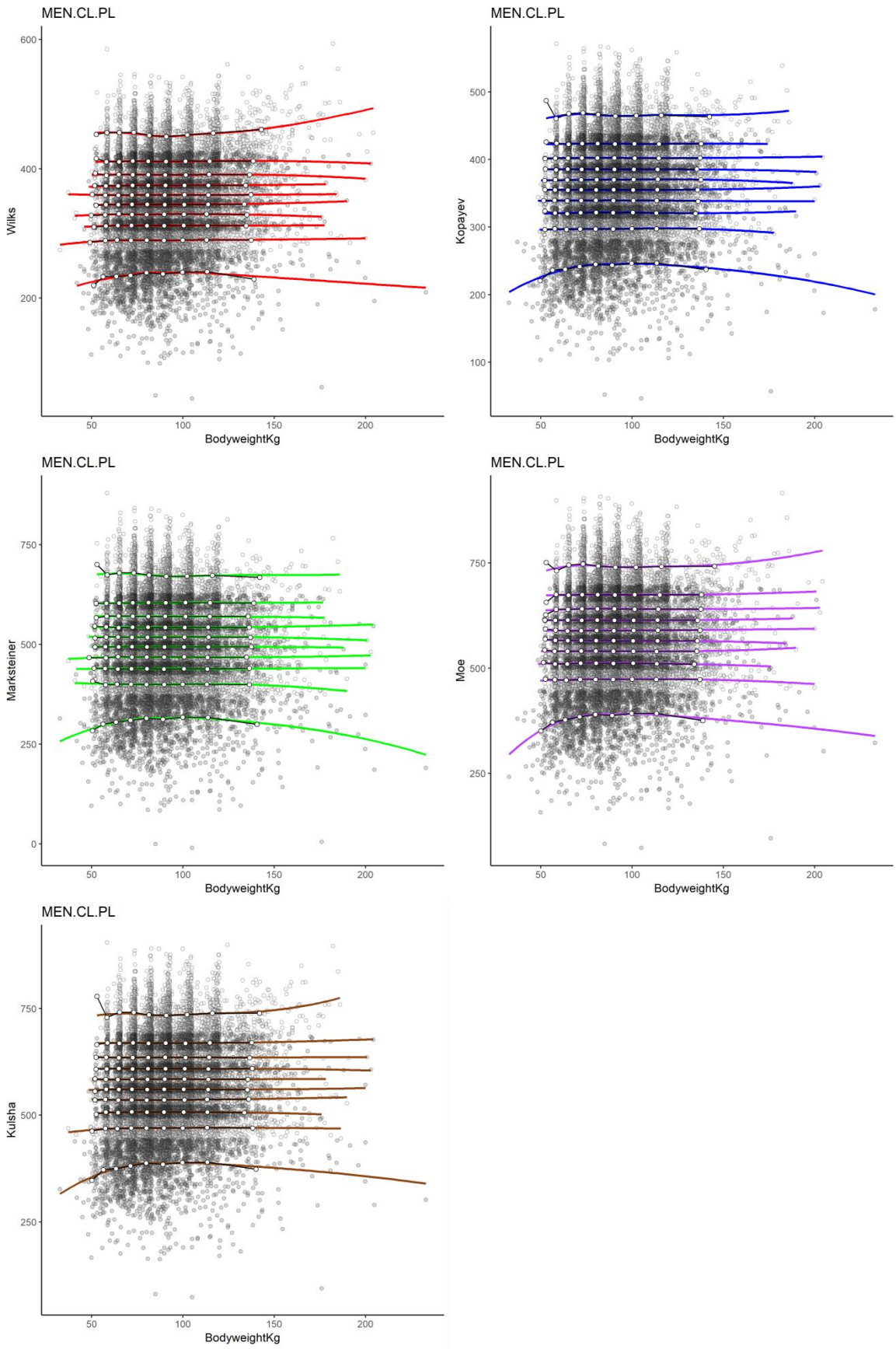

Figure 3: LOESS-plot for men's classic powerlifting

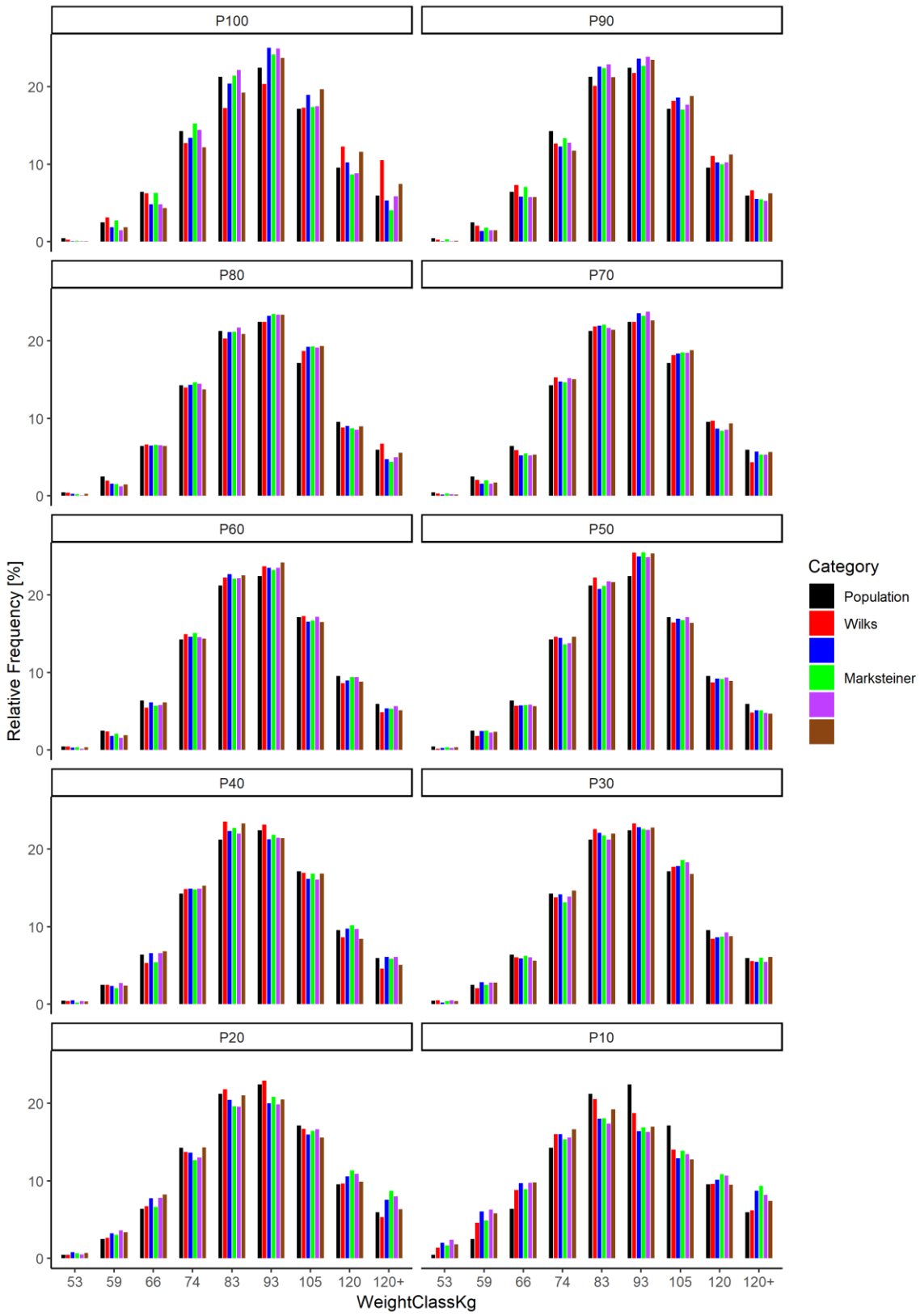

Figure 4: Distribution of relative scores across performance bands for men's classic powerlifting

2.3 Women's classic bench press (WMN.CL.BP)

Table 3: Summary statistics for women's classic bench press

	Wilks	Author 2	Marksteiner	Author 4	Author 5
Distribution χ^2 sum	551.38732	190.66368	171.41021	148.97693	210.97744
LOESS residual sum	2.02875	1.60337	1.73783	1.80993	1.59840

For the distribution of relative scores, Author 4's scoring represents the shape of the underlying population best. **Author 5 takes third place.** LOESS residual sums indicate that **Author 5 scores are most level across all performance bands** (Table 3, Figure 5 & Figure 5).

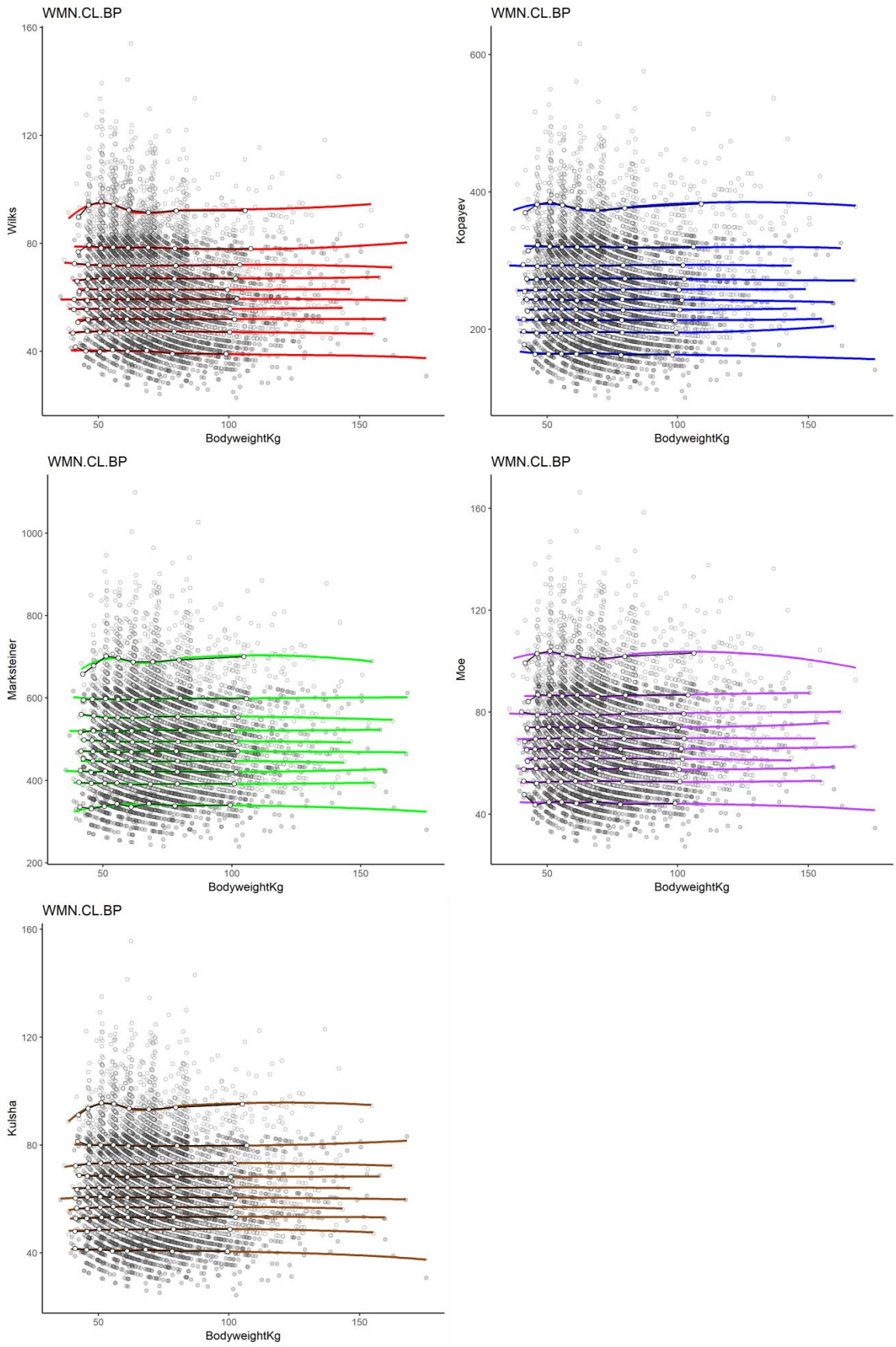

Figure 5: LOESS-plot for women's classic bench press

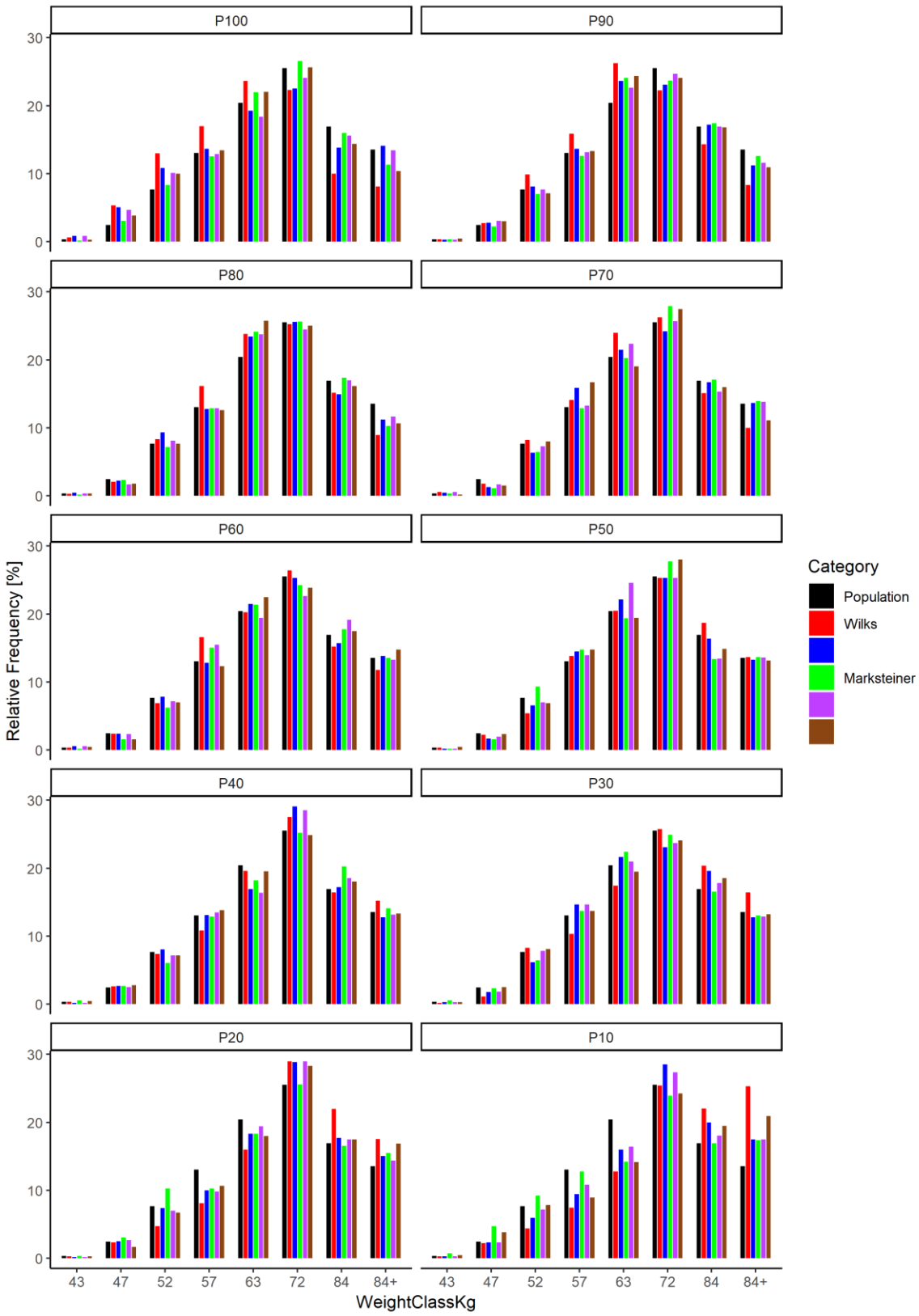

Figure 6: Distribution of relative scores across performance bands for women's classic bench press

2.4 Women's classic powerlifting (WMN.CL.PL)

Table 4: Summary statistics for women's classic powerlifting

	Wilks	Author 2	Marksteiner	Author 4	Author 5
Distribution χ^2 sum	654.80202	158.34773	131.90024	119.80617	183.25622
LOESS residual sum	2.35176	1.95914	1.97693	2.05277	2.17448

Author 4's scoring represents the shape of the underlying population best for the distribution of relative scores. **Author 5 takes fourth place for this evaluation criterion.** LOESS residual sums indicate that Author 2's scores are most level across all performance bands. **Author 5 again takes fourth place here** (Table 4, Figure 7 & Figure 8).

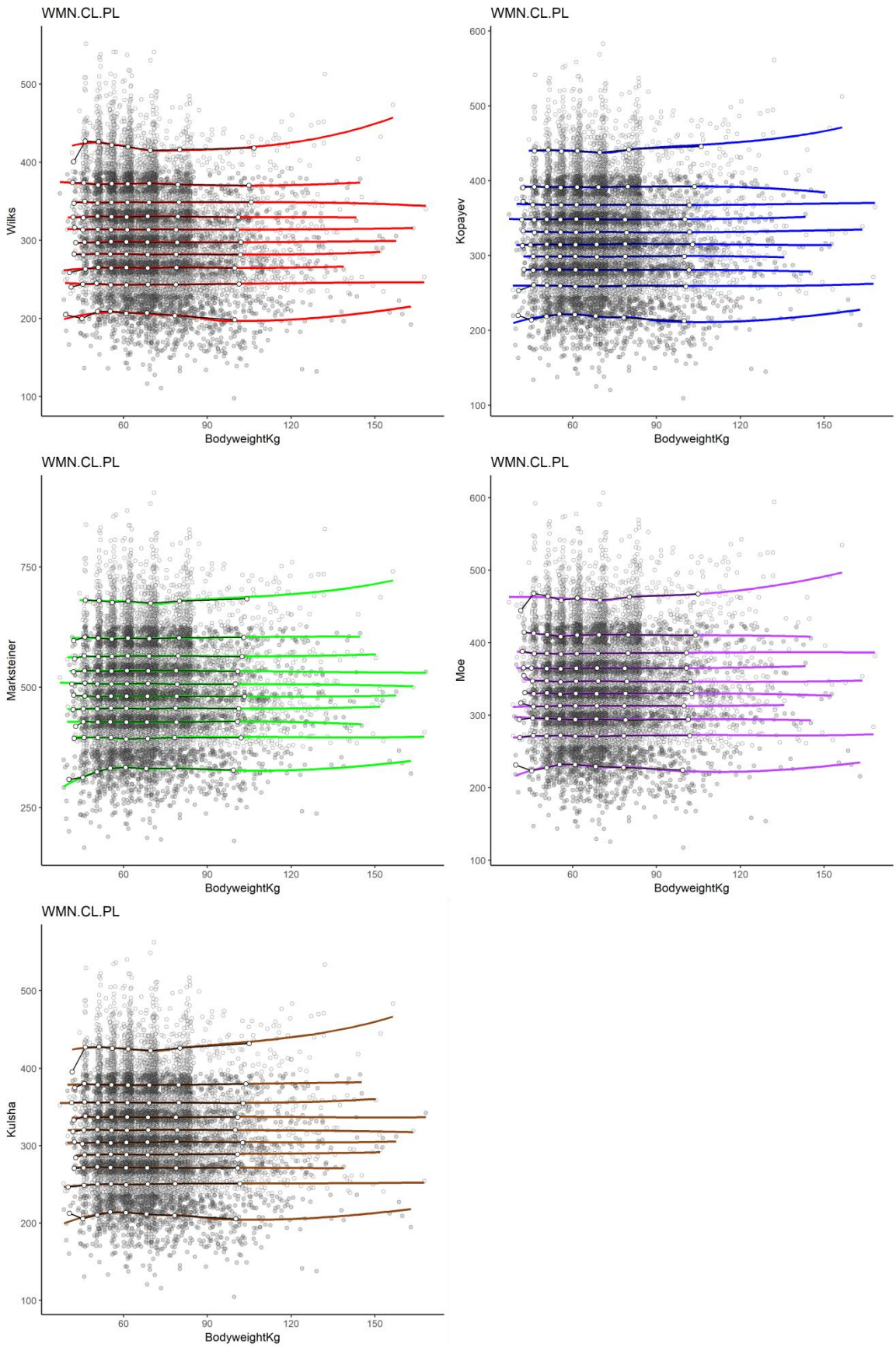

Figure 7: LOESS-plot for women's classic powerlifting

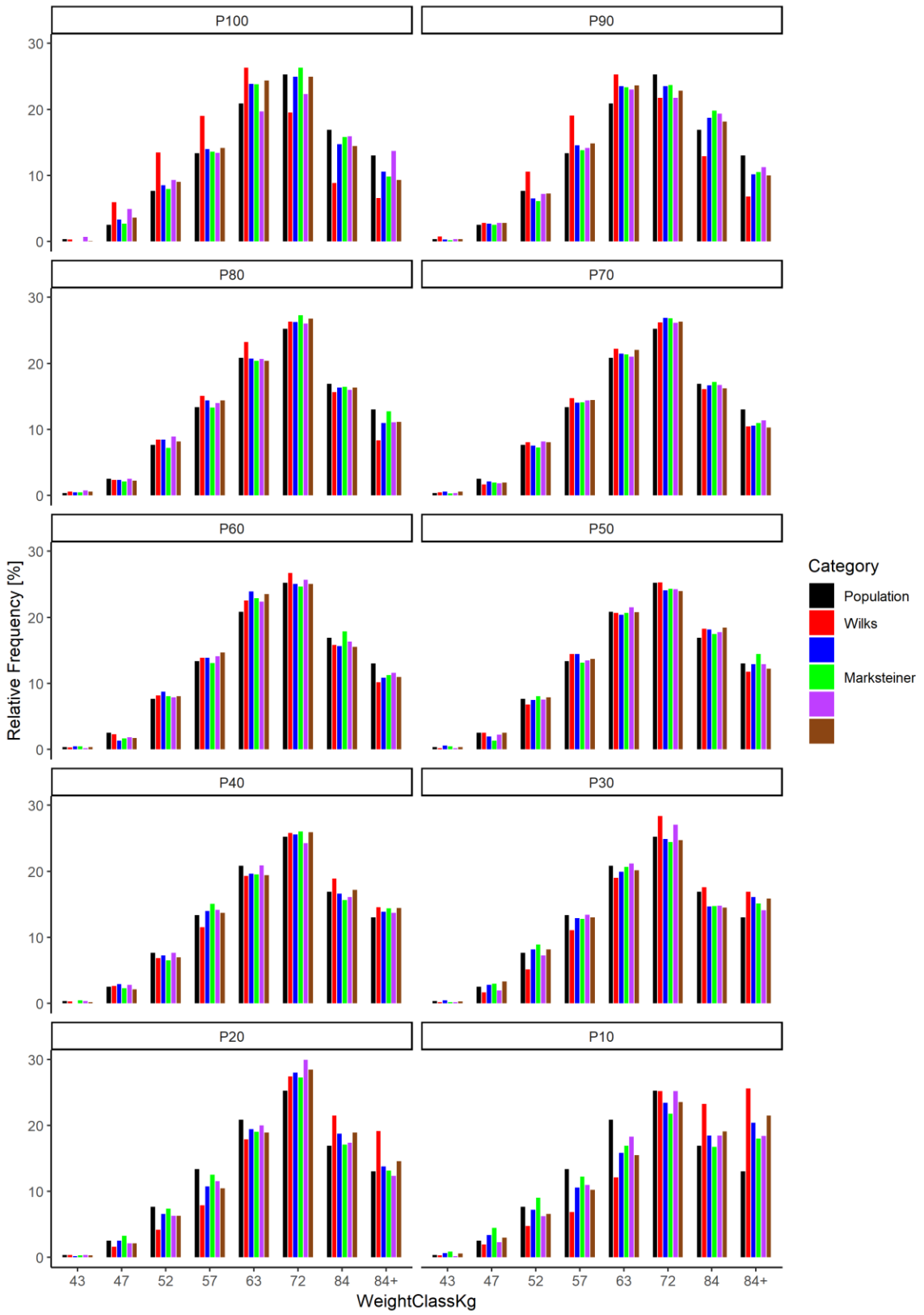

Figure 8: Distribution of relative scores across performance bands for women's classic powerlifting

3 Conclusion

Table 5: Summary of statistics across all subdisciplines

		Wilks	Author 2	Marksteiner	Author 4	Author 5
MEN.CL.BP	Distr	829.96483	1014.02677	280.71897	772.38166	886.85953
	LOESS	2.06503	1.91810	1.70323	1.83663	1.96343
MEN.CL.PL	Distr	300.97666	426.17741	333.56628	485.76790	409.48174
	LOESS	1.74409	1.51711	1.45324	1.86049	1.72047
WMN.CL.BP	Distr	551.38732	190.66368	171.41021	148.97693	210.97744
	LOESS	2.02875	1.60337	1.73783	1.80993	1.59840
WMN.CL.PL	Distr	654.80202	158.34773	131.90024	119.80617	183.25622
	LOESS	2.35176	1.95914	1.97693	2.05277	2.17448
			z-scores			
		Wilks	**Author 2**	**Marksteiner**	**Author 4**	**Author 5**
MEN.CL.BP	Distr	0.26062	0.91618	-1.69558	0.05553	0.46326
	LOESS	1.23095	0.15272	-1.42398	-0.44508	0.48539
MEN.CL.PL	Distr	-1.21752	0.47212	-0.77771	1.27632	0.24680
	LOESS	0.50338	-0.84065	-1.21882	1.19259	0.36350
WMN.CL.BP	Distr	1.77196	-0.38233	-0.49732	-0.63129	-0.26102
	LOESS	1.54015	-0.85885	-0.10052	0.30609	-0.88687
WMN.CL.PL	Distr	1.77847	-0.40064	-0.51672	-0.56981	-0.29130
	LOESS	1.52728	-0.88340	-0.77417	-0.30850	0.43879
Sum z-scores		**7.39528**	**-1.82485**	**-7.00483**	**0.87585**	**0.55855**

To rate the methods against each other, we will first standardize these statistics, and then calculate cumulative sums for each method. The reviewers are not aware of any weighting preference of the IPF, so we will operate under the assumption that all subdisciplines are of equal importance. Furthermore, we will not introduce any performance band weighting within the methods. Table 5 contains the unweighted raw statistics already given in the previous sections, with their z-scores and cumulative sums in the lower area of the table. Standardization is done as follows: First, the five scores of the different methods within one domain (e. g. distribution scores for Wilks, Author 2, Marksteiner, Author 4 and Author 5 for MEN.CL.BP) are scaled to a mean of zero and standard deviation of 1. These z-scores are then summed across all subdisciplines for each method. The resulting z-score sum can be found at the bottom of Table 5.

The z-score sums indicate that for the classic disciplines the lowest scoring and therefore the best method is Marksteiner. Marksteiner and with reservations Author 2 score substantially lower (better) than Author 4, Author 5 and Wilks. **Overall Author 5 is third, just ahead of Author 4.**

Even though Author 5's method clearly performs better than the Wilks method, his method does not reach the quality of Marksteiner and Author 2. **Against the background of two substantially better-performing methods Author 5's method should not be considered as a replacement for the Wilks coefficients.**

4 R Code used for calculating relative scores

For Marksteiner and Author 2, vector `C` contains the specific column of coefficients taken from data frame `COEF`, matching the variable `ID` (which is the filename) for each dataset with the column names. Lifter performance for PL is included in the variable `TotalKg`, for BenchPress in the variable `BestBenchKg`.

4.1 Author 2

```
COEF.AUT2 <- data.frame(MEN.EQ.PL = c(1256.96, 1440.39, 0.01566),
                        MEN.CL.PL = c(1225.49, 1108.60, 0.00940),
                        MEN.EQ.BP = c(399.49, 626.78, 0.01943),
                        MEN.CL.BP = c(344.45, 288.97, 0.00796),
                        WMN.EQ.PL = c(746.81, 1001.1, 0.02616),
                        WMN.CL.PL = c(632.84, 616.93, 0.01886),
                        WMN.EQ.BP = c(241.88, 266.63, 0.01857),
                        WMN.CL.BP = c(142.57, 321.15, 0.03967))
C <- COEF.AUT2[,which(names(COEF.AUT2)==ID)]
```

(Note that coefficient `c` for `COEF.AUT2$WMN.CL.BP` is taken from the original R output on page 39, not from table 3.4, which seems to contain an erroneous last digit.)

4.1.1 PL:

```
Author_2Points <- data$TotalKg * 500 / (C[1] - C[2] * exp(1)^(-C[3] *
data$BodyweightKg))
```

4.1.2 BP:

```
Author_2Points <- data$BestBenchKg * 500 / (C[1] - C[2] * exp(1)^(-C[3] *
data$BodyweightKg))
```

4.2 Marksteiner

```
COEF.MKS <- data.frame(MEN.CL.PL = c(310.67, 857.7850, 53.2160, 147.0835),
                       MEN.CL.BP = c(86.4745, 259.155, 17.5785, 53.122),
                       MEN.EQ.PL = c(387.265, 1121.28, 80.6324, 222.4896),
                       MEN.EQ.BP = c(133.94, 441.465, 35.3938, 113.0057),
                       WMN.CL.PL = c(125.1435, 228.03, 34.5246, 86.8301),
                       WMN.CL.BP = c(25.0485, 43.848, 6.7172, 13.952),
                       WMN.EQ.PL = c(176.58, 373.315, 48.4534, 110.0103),
                       WMN.EQ.BP = c(49.106, 124.2090, 23.199, 67.4926))
C <- COEF.MKS[,which(names(COEF.MKS)==ID)]
```

4.2.1 PL:

```
MarksteinerPoints <- 500 + (data$TotalKg - (C[1] * log(data$BodyweightKg) -
C[2]))/(C[3] * log(data$BodyweightKg) - C[4]) * 100
```

4.2.2 BP:

```
MarksteinerPoints <- 500 + (data$BestBenchKg - (C[1] * log(data$BodyweightKg) -
C[2]))/(C[3] * log(data$BodyweightKg) - C[4]) * 100
```

4.3 Author 4

4.3.1 MEN PL:

```
Author_4Points <- (842.04/(567.17331903331046-
            3.8274022516399691*data$BodyweightKg+
            0.085540790528515487*data$BodyweightKg^2+
            0.00038089371816506300*data$BodyweightKg^3-
            9.6534028292103307e-06*data$BodyweightKg^4+
            4.5466325500142581e-08*data$BodyweightKg^5-
            6.8016890038424778e-11*data$BodyweightKg^6))*data$TotalKg
```

4.3.2 MEN BP:

```
Author_4Points <- (842.04/(567.17331903331046-
            3.8274022516399691*data$BodyweightKg+
            0.085540790528515487*data$BodyweightKg^2+
            0.00038089371816506300*data$BodyweightKg^3-
            9.6534028292103307e-06*data$BodyweightKg^4+
            4.5466325500142581e-08*data$BodyweightKg^5-
            6.8016890038424778e-11*data$BodyweightKg^6))*data$BestBenchKg
```

4.3.3 WMN PL:

```
Author_4Points <- (540.34/(-817.96025420079411+
            52.980402587880548*data$BodyweightKg-
            0.85759063714003292*data$BodyweightKg^2+
            6.9435316930538046e-03*data$BodyweightKg^3-
            2.7144538222479236e-05*data$BodyweightKg^4+
            4.0708823967668707e-08*data$BodyweightKg^5))*data$TotalKg
```

4.3.4 WMN BP:

```
Author_4Points <- (540.34/(-817.96025420079411+
            52.980402587880548*data$BodyweightKg-
            0.85759063714003292*data$BodyweightKg^2+
            6.9435316930538046e-03*data$BodyweightKg^3-
            2.7144538222479236e-05*data$BodyweightKg^4+
            4.0708823967668707e-08*data$BodyweightKg^5))*data$BestBenchKg
```

4.4 Author 5

4.4.1 MEN CL PL:

```
Author_5Points <- data$TotalKg/atan(data$BodyweightKg/64)
```

4.4.2 MEN CL BP:

```
Author_5Points <- data$BestBenchKg/atan(data$BodyweightKg/64)
```

4.4.3 WMN CL PL:

```
Author_5Points <- data$TotalKg/atan(data$BodyweightKg/48)
```

4.4.4 WMN CL BP:

```
Author_5Points <- data$BestBenchKg/atan(data$BodyweightKg/48)
```