Recalibrating IPF GL points

Bo Arends
Joery Wesseling
Sam Rappange

January 2025

Contents

1 Introduction

In 2020, the International Powerlifting Federation (IPF) introduced a new scoring formula that can be used to compare lifters from different weightclasses.

The input of the formula is the bodyweight (in kilograms) and the sex of the lifter (male/female). The output is a coefficient, that needs to be multiplied by the total of the lifter to get the IPF GL points score for that performance.

The calculated coefficient is based on lifting data from the years 2011 to 2020. Since it has been four years since the IPF GL point formula was introduced, it would interesting to see what would happen to the coefficients of the IPF GL points formula if data from 2020 to 2024 is included. Powerlifting as a sport has grown tremendously, and we have seen a steady increase in not only the total worlds records, but also in the competitive depth of the weightclasses. This means that a recalibrated IPF GL points formula could lead to very different results in IPF GL points now. In this report we explain the method that we have used to recalibrate the IPF GL points formula using the most recent data (20th of December 2024).

For now, we chose to only recalibrate the formula for classic full power competitions, so not for equipped or bench press only competitions.

2 Method

The method section consists of two parts. First we go over the method used by the IPF, after which we go over the recalibration method of the IPF GL points formula.

2.1 IPF GL points method made by the IPF

In this section we explain what data the IPF used for their formula. The whole method and thought process can be found on the website of the IPF (1) (2).

IPF GL coefficients are derived statistically. Regression analysis is used with the best results of elite athletes, called the Golden Standard Samples. The Golden Standard Samples consist of the best result of athletes not achieved not less than 16% of the current world records within an IPF or EPF competition, from 2011 until 2020.

To recalibrate the IPF Goodlift points formula, we used data from 2011 until December 2024 to repeat the method of the IPF, but with up to date data.

2.2 Recalibrating the formula

The full Python code for this method can be found in the Jupyter notebook on our GitHub account. To get a better understanding of how different scores like Wilks and IPF GL points are calculated, we recommend to read the article written by Tommy Odland (3).

We started off by using the same formula of the IPF GL points formula, together with the original coefficients of that formula. The formula and coefficients (raw, full power competitions) are given below.

$$\text{IPF GL Coefficient} = \frac{100}{\text{A - B} \cdot e^{-\text{C} \cdot \text{Bwt}}}$$

Coefficient	Men	Women
A	1199.72839	610.32796
B	1025.18162	1045.59282
C	0.00921	0.03048

A, B and C are coefficients that are determined by regression analysis on the data, e is the base of the natural logarithm, and 100 is the normalization factor. This factor 100 is chosen so that the performance golden standard (average elite athlete's level) is 100 points for their respective bodyweight.

To start of our regression analysis, we plot the up to date data (from 2011 to December 2024) in a graph, where bodyweight of the lifter is on the x-axis, and the performed total is on the y-axis.

We start of the regression with the original coefficients as in the table above. We then use a method of a Python package called SciPy, called curve_fit to optimize the line that is used to fit the data. The function is used to fit a function to a set of data points by optimizing the parameters of the function, using non-linear least squares optimization to minimize the difference between the observed data and the predicted values of the function (the fitted line).

The full code of our method can be found on the GitHub repository:
https://github.com/JoeryWesseling/RecalibratingIPFPoints/tree/main.

3 Results

3.1 New Coefficients

Below you can see a table with the new coefficients of the calibrated data.

Coefficient	Men	Women
A	1038.271102	688.743183
B	1207.593924	1135.134454
C	0.017290	0.028594

3.2 Golden Standard Samples

In the graphs below we can see the amount of Golden Standard Samples per weightclass. This is the amount of data points that are available to fit our model on.

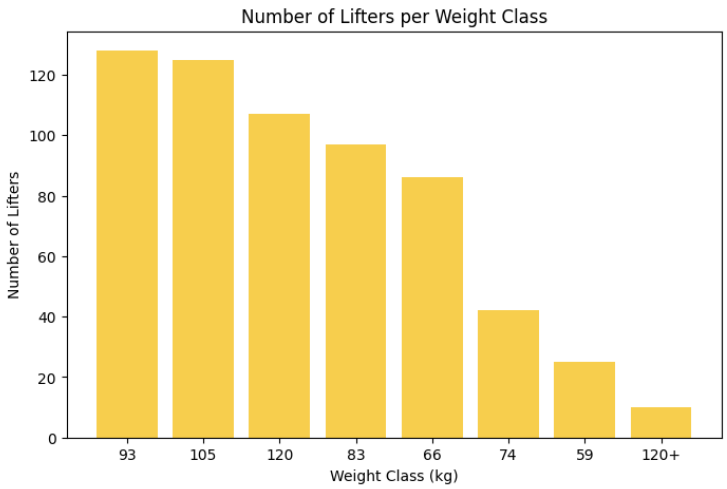

Figure 1: Old and refitted model for men

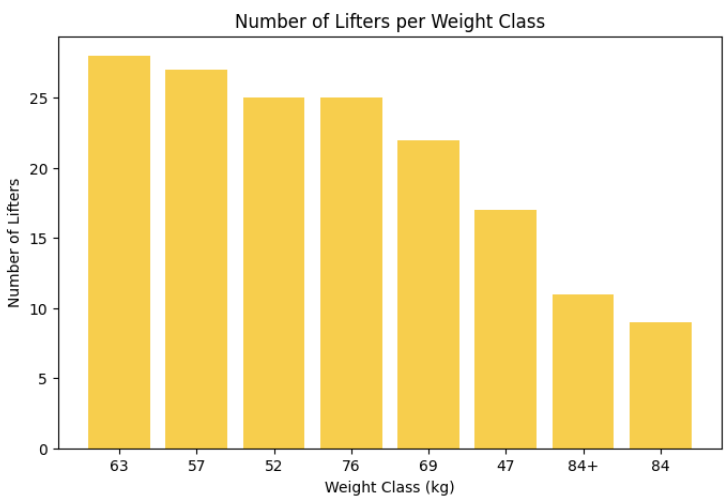

Figure 2: Old and refitted model for women

3.3 Old and refitted model

The blue dots are the data points, so these consist of all the Golden Standard Samples from 2011 until 2024. The red line is the old model, and the green line is the new refitted model on the up to date Golden Standard Sample data.

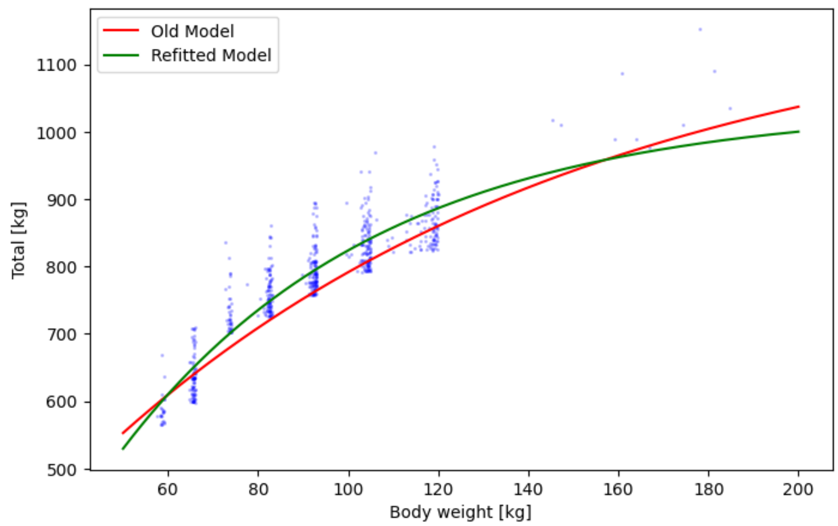

Figure 3: Old and refitted model for men

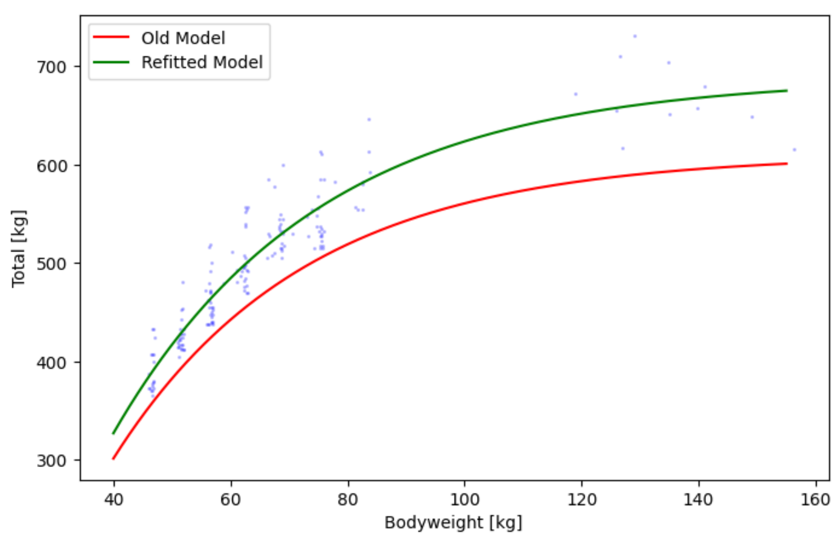

Figure 4: Old and refitted model for women

3.4 Old and new Goodlift coefficients

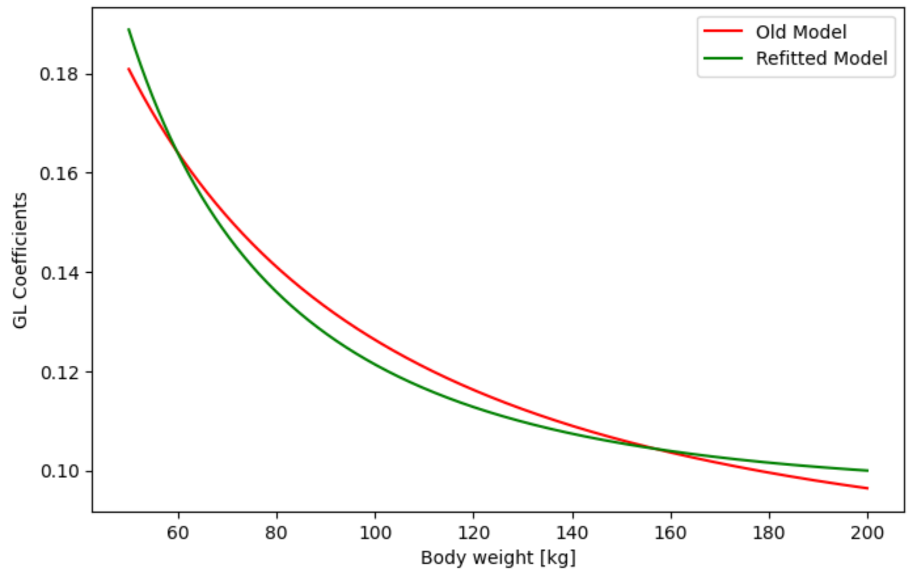

Figure 5: Old and new Goodlift coefficient for men

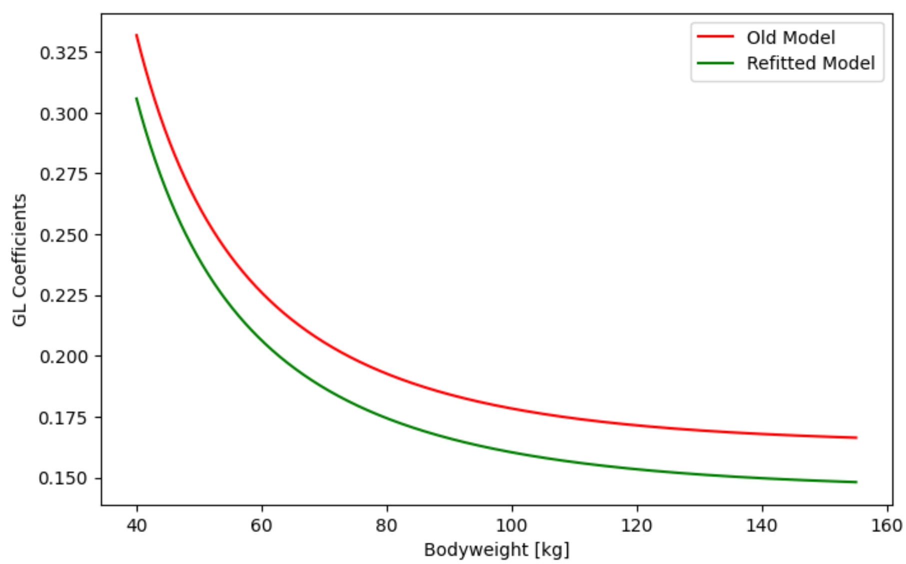

Figure 6: Old and new Goodlift coefficien for women

3.5 Relative changes

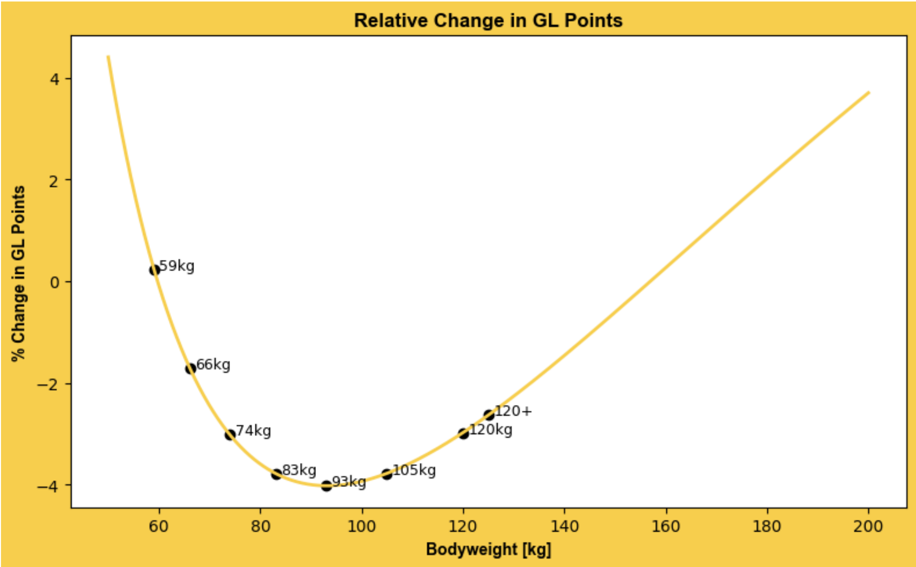

Figure 7: Relative changes for men

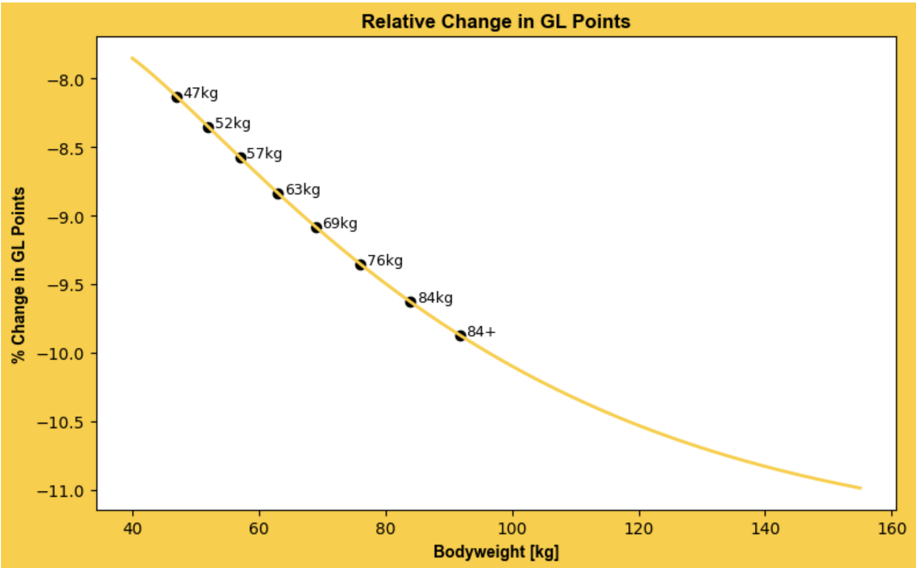

Figure 8: Relative changes for women

4 Conclusion

4.1 Golden Standard Samples

We can see that there are significantly more Golden Standard Samples available for men than for women, which may result in a better fit for the curve for men, since there is more data to fit the curve on.
However, we also see that there is a relatively low amount of super heavyweight Golden Standard Samples to fit the data on.

4.2 Conclusion old and refitted model

First, we look at the comparison of the old and refitted model, seen as in figure 3. We can see that to fit our new data better, the curve increases between bodyweights of 60 kg up to approximately 160 kg. However, from 160 kg onward the curve decreases compared to the old model. We think the new refitted model fits way better on data from the 59 kg weightclass to the 120 kg weightclass, but does not fit well to the 120+ kg class data.
You can see that looking at the figure, the curve of the new refitted model is below a lot of the actual data points of the 120+ kg class. This is because there are a lot more data points for the 59 kg class up to the 120 kg class, than there are data points for the 120+ kg class lifters. There are only 10 Golden Standard Sample data points for the 120+ kg class. This means the fit of our curve mostly get influenced by those data points in the weight classes below 120+ kg. We can conclude that the new refitted model seems to do really well for all male weightclasses except the 120+ kg. Right now, the 120+ kg class will be favored with this new model. This means they will get a higher IPF Goodlift coefficient and therefore a higher IPF Goodlift score than they should have. Suggestions to solve this problem can be found in 'Discussion and Future Work'.

The new refitted model for women as seen in figure 4, seems to fit really well to the data. The curved line seems to fit well for all weightclasses, also for the 84+ kg class.
What we do see is a very sharp increase from the refitted model compared to the old model, especially when comparing this to the refitted model for men. We think that since powerlifting has become increasingly more popular in the past four years, the sport has seen a sharp increase in elite women lifters. As a result, the height of the curve has increased significantly for all weight classes, especially compared to the curve for the men.

4.3 Conclusion Old and New Goodlift coefficients

Here, we can see the resulting graphs for the refitted IPF Goodlift coefficients, which correspond to figure 1 and figure 2 with the refitted models. Hence, the same conclusions can be drawn as in section 4.2.

4.4 Conclusion Relative Changes

In this section, we visualized the percentage change in IPF Goodlift points compared to the old model.
For the men, we can see that the 93 kg class will drop the most, their IPF Goodlift points will drop with approximately 4%. The only class that increases a little bit is the 59 kg class, and the rest of the classes all see a decrease in IPF Goodlift Points.

For the women, the IPF Goodlift Points will drop even harder. It starts with around 8.1% for the 47 kg class, and the percentage drop only increases from there. The Goodlift Points of the 84+ kg class even drops with 10% or more. This is because there has been an increase in elite women powerlifters, but even a bigger increase in the heavier weightclasses.

It is interesting to see the percentage drop in IPF Goodlift points between the men and women. The biggest percentage drop in the men classes is 93 kg class with a drop of 4%, while the smallest percentage drop for women is already 8%. We can again conclude from this that since 2020, there has been a large increase in elite women powerlifters, especially in the heavier weightclasses.
In conclusion, we think that these calibrated IPF Goodlift Points are a fair and up to date change to the current IPF Goodlift Points, taking into account the growth of the sport in the past four years.

5 Discussion and Future Work

We recalibrated IPF Goodlift points using up to date data (December 2024). We did this by using the same method as the IPF did, to keep our results consistent and comparable. There are a few things that are not completely clear to us.
In the IPF documents, it is not stated why a percentage of no less than 16% of the world record is used for the Golden Standard Samples. A percentage of no less than 16% for super heavy weight classes will look different than for other weight classes, since the 120+ kg class has a long tailed distribution.
A result of this is that some weightclasses have way more data points to fit the curve on than other weightclasses, which can skew the curve, which we saw with the 120+ kg class lifters.
Another problem with the 16% is possible outliers. If one elite lifters is an extreme outlier that puts the world record in his or her weightclass at an extremely high total, the Golden Standard Samples will contain a lot less lifters.

We also do not know why only IPF and EPF results are used for the Golden Standard Samples. Of course, we want the Golden Standard Samples to be results with a high standard. International competitions have international referees and a jury, setting a high standard for competitions. However, not all international competitions are included, since AfricanPF, AsianPF, FESUPO, NAPF and ORPF are not included in the Golden Standard Samples.

Of course, these results could be extended to include equipped powerlifting, and bench press only competitions.

5.1 Changing the IPF GL formula for the 120+ kg class

Looking at our curve fit, we don't think our curve is a good fit for the 120+ kg class, and now the 120+ kg class gets favored heavily with our new model. One reason is because of the lack of data points for this weightclass. One way to tackle this problem would be to increase the amount of data points in the 120+ kg class.

First, you look at the average amount of data points in the other weight classes. Then you compare this with the amount of data points in the 120+ class. The difference in these two values is the amount of data points you generate, looking at the average performances of the Golden Standard Samples. You add these data points to the 'actual' data points of the Golden Standard Samples, to make sure that the curve fits better through the 120+ kg class data points.

We calculated an average amount of Golden Standard Samples per weightclass of 87 lifters. The 120+ kg class only has 10 Golden Standard Samples. So we wanted to add 77 data points to the 120+ kg class. We then looked at the average performance of those 10 elite lifters of the Golden Standard Samples in the 120+ kg class. This was 1036.15 kg total at 166.23 kg bodyweight.

After doing this, the new curve fitted on the data points is seen in the graph below. The green line is the calibrated formula discussed in this report. The blue line is the augmented line where we add data points for the 120+ kg class. Below that we also added the new values of the coefficients for men.

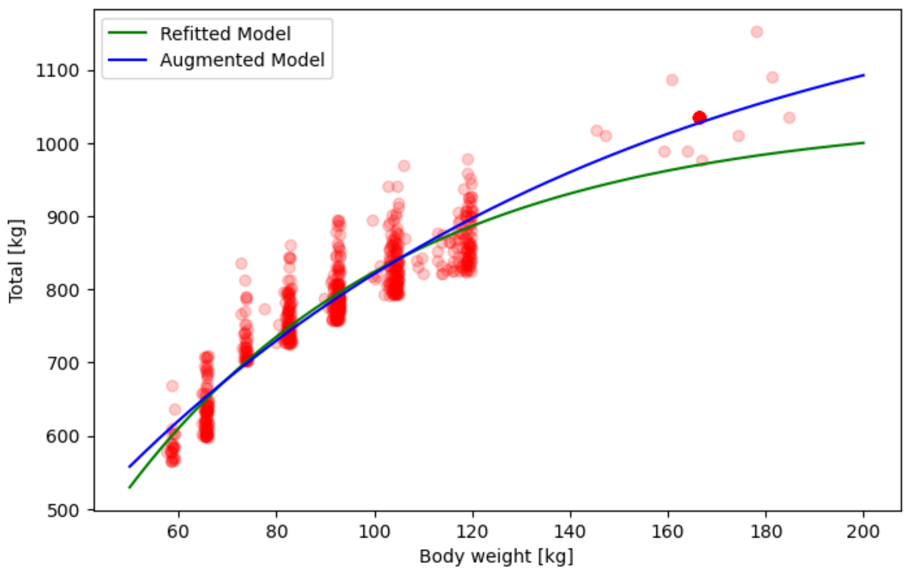

Figure 9: Our calibrated model and our augmented model

Coefficient	Men
A	1272.58362
B	1131.82948
C	0.00919515505

We can see that the curve fit on the data stays almost exactly the same for all the weight classes up until the 120- kg class, but fits way better through the data points of the 120+ kg class, making it more fair for everyone when comparing lifters of different weight classes.
An argument can still be made that 10 Golden Standard Samples in the 120+ kg class is still not enough, and should not be taken averages of to augment data. However, this seemed like the only feasible method to make the curve more fair for the 120+ kg class.

5.2 Future Work

An argument can be made that it would be more fair to use the previous data augmentation method to make sure all weightclasses are balanced, by adding data points for each class until all classes have the same amount of data points. This is definitively something that could be looked at in the future.
Another point that could be looked at is looking at how the Golden Standard Samples are derived, and if there is a more fair way to do this, especially if the IPF GL formula will be used a competitions for scoring.

References

(1) **IPF**. "The IPF GL coefficients for relative scoring". In: (2020).

(2) **A. Stetsenko O. Kopayev B. Onyshchenko**. "Evaluation of Wilks, Wilks-2, DOTS, IPF and Goodlift formulas for calculating relative scores in powerlifting competitions". In: (2020).

(3) **Tommy Odland**. "Relative strength - Wilks, IPF GL and allometry". In: (2023).

6 Appendix A

Changes per weightclass.

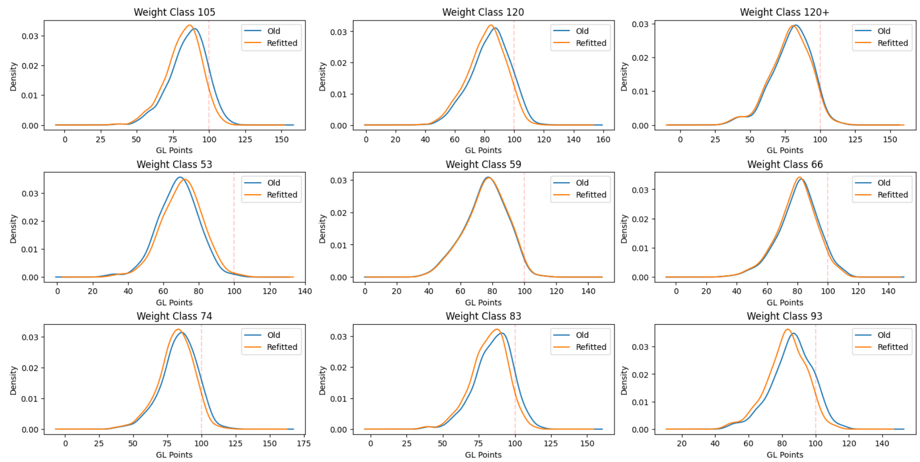

Figure 10: Changes per weightclass for men

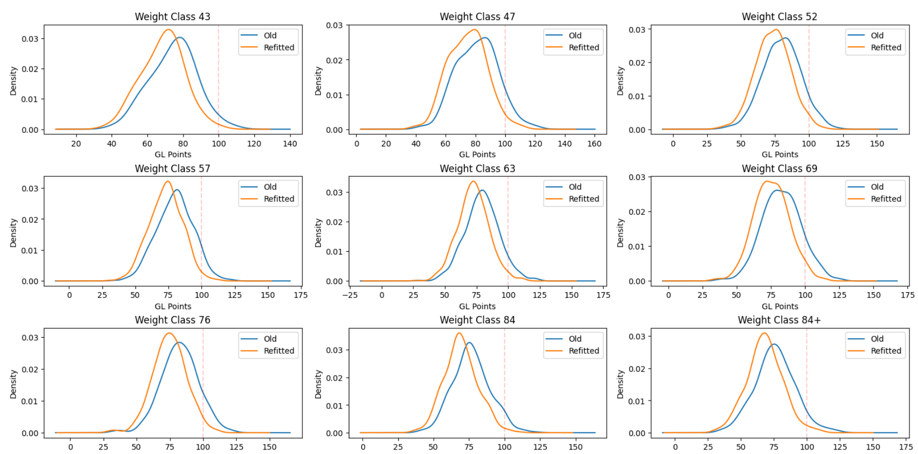

Figure 11: Changes per weightclass for women