

# QOP Cross-Validation Results

---

We performed a 2-fold cross-validation of QOP in the following way, using the 2015 regular + post season data. For each pitcher, we randomly divided their QOPVs into two groups and took the mean of each group giving *mean1* and *mean2*. This was repeated for each pitcher, giving a total of 734 pairs of (*mean1*, *mean2*). Pearson's correlation was calculated between the two groups. This was repeated 10,000 times. The mean (standard deviation in parentheses) of the 10,000 correlations is shown in Figure 1. The above process was repeated for all pitches combined, fastballs only, and off-speed pitches only. While the results were adequate for all pitches combined ( $R^2=52\%$ ) and fastballs ( $R^2=56\%$ ), the off-speed pitch correlation was too low for us to consider the data validated ( $R^2=29\%$ ). Since pitch quality can vary substantially by pitch, and some pitchers have a low sample size, we re-ran the simulation excluding pitchers with minimums below 30, 100, and 200. All correlations rose to very good cross-validation levels.

	Minimum Number of Pitches			
	1	30	100	200
<b>All Pitches</b>	0.722 (0.029)	0.822 (0.016)	0.883 (0.009)	0.901 (0.007)
<b>Fastballs</b>	0.747 (0.035)	0.844 (0.013)	0.900 (0.007)	0.925 (0.006)
<b>Off-Speed</b>	0.542 (0.038)	0.684 (0.023)	0.798 (0.014)	0.843 (0.012)

Figure 1: QOP Cross-Validation. The 2015 QOPVs for each of the 734 pitchers were randomly split into two groups. Cell entries are the mean (SD) Pearson's correlation of 10,000 random samples between the two groups.