Sushi and Sliders: NPB Pitching Scouting Report Using Multiple Linear Regression

By: Leonardo C. Gentile



Figure 1: Tatsuya Imai pitching for the Saitama Seibu Lions

Abstract and Executive Summary

Nippon Professional Baseball (NPB) is the highest level of professional baseball in Japan. This paper investigates the performance of NPB pitchers using multiple linear regression to predict Wins Above Replacement (WAR). By analyzing a data set of NPB pitcher statistics, we evaluate which pitching metrics are most predictive of overall value and assess their implications for projecting Major League Baseball (MLB) success. MLB is America's highest level of professional baseball. After narrowing to key predictors such as innings pitched, strikeouts, walks, and home runs, our model explains over 84% of the variance in WAR, with significant interaction effects enhancing predictive accuracy. Although certain violations of regression assumptions were observed, the results provide a data-driven framework for identifying MLB-caliber talent in NPB.

Based on my multiple linear regression model, stats from DeltaGraphs, WBC and Premier12 data, graphics and NPB updates from Yakyu Cosmopolitan and Gaijin Baseball, in-person scouting in Japan, and my own statistical analysis, I believe the top five MLB-ready pitchers currently in NPB are Tatsuya Imai, Jon Duplantier, Hiromi Itoh, Foster Griffin, and Hiroto Saiki. This study contributes to the growing field of international sabermetrics by bridging statistical insights across leagues and offering practical applications for scouting and player evaluation.

Introduction

The increasing globalization of baseball has created new opportunities for evaluating talent across leagues, with Nippon Professional Baseball (NPB) serving as one of the most competitive environments outside of Major League Baseball (MLB). Japanese pitchers have historically made significant impacts in MLB, from Hideo Nomo and Yu Darvish to Shohei Ohtani, demonstrating the importance of reliable scouting and projection methods. However, subjective scouting alone is often insufficient for accurately predicting success across leagues, where differences in competition level, ballpark dimensions, and run environments complicate evaluation.

Sabermetrics and statistical modeling provide an objective framework for bridging these gaps. Regression analysis in particular has been widely applied within MLB to measure player contributions and forecast future performance. Building on my prior work with Korean Baseball Organization (KBO) pitchers, this study applies multiple linear regression to NPB pitching data to identify which performance metrics most strongly predict Wins Above Replacement (WAR).

The goals of this case study are threefold: (1) to construct a multiple linear regression model that explains variance in NPB pitcher WAR, (2) to evaluate the reliability and performance of this model in light of statistical assumptions, and (3) to generate actionable insights for projecting NPB pitchers into MLB contexts. By combining statistical rigor with practical scouting considerations, this work contributes both to academic sabermetric research and to the applied practice of international player evaluation.

Exploratory Data Analysis (EDA)

Our training data set has 1,082 observations (pitchers) of 84 variables while our testing data set has 239 observations of 84 variables. Each pitcher in the data set has thrown a minimum of 10 innings to reduce the noise that may be introduced by pitchers with a small sample size. The 84 variables listed for each observation are season (2021, 2022, 2023, 2024, or 2025), Team, Wins (W), Loses (L), Earned Run Average (ERA), Games (G), Games Started (GS), Complete Games (CG), Shutouts (ShO), Saves (SV), Holds (HLD), Innings Pitched (IP), Total Batters Faced (TBF), Hits (H), Runs (R), Earned Runs (ER), Home Runs (HR), Walks (BB), Intentional Walks (IBB), Hit By Pitch (HBP), Wild Pitch (WP), Balk (BK), Strikeouts (SO), Stolen Bases (SB), Caught Stealing (CS), Strikeout Rate (K%), Walk Rate (BB%), Strikeout Minus Walk Rate (K-BB%), Home Run Rate (HR%), Batting Average Against (AVG), Walks and Hits per Inning Pitched (WHIP), Left On Base Percentage (LOB%), Runs Allowed Indexed (RA-), Fielding Independent Pitching Indexed (FIP-), Expected FIP Indexed (xFIP-), Runs Allowed (RA), Fielding Independent Pitching (FIP), Runs Allowed minus FIP (R-F), Expected Fielding Independent Pitching (xFIP), True Runs Allowed (tRA), Defensive Efficiency Ratio (DER), Ground Ball to Fly Ball Rate (GB/FB), Line Drive Percentage (LD%), Ground Ball Percentage (GB%), Fly Ball Percentage (FB%), Infield Fly Ball Percentage (IFFB%), Home Run to Fly Ball Percentage (HR/FB), Runs Scored (RS), Runs Scored Per Nine Innings (RS/9), Strikes, Balls, Total Pitches, Pull Percentage (Pull%), Center Percentage (Cent%), Opposite Field Percentage (Oppo%), Soft Contact Percentage (Soft%), Medium Contact Percentage (Mid%), Hard Contact Percentage (Hard%), Win Probability Added (WPA), Run Expectancy 24 Base-Out States (RE24), Run Expectancy Wins (REW), Leverage Index (pLI), In-Play Leverage Index (inLI), Game Leverage Index (gmLI), Expected Leverage Index (exLI), Win Probability Added per Leverage Index (WPA/LI), Clutch, Outside Swing Percentage (O-Swing%), Zone Swing Percentage (Z-Swing%), Overall Swing Percentage (Swing%), Outside Contact Percentage (O-Contact%), Zone Contact Percentage (Z-Contact%), Overall Contact Percentage (Contact%), Zone Rate (Zone%), First-Pitch Strike Rate (F-Strike%), Whiff Percentage (Whiff%), Put Away Percentage (Put Away%), Swinging Strike Percentage (SwStr%), Called Strike Percentage (CStr%), Called Strike Plus Whiff Rate (CSW%), Starting Pitcher Innings Pitched (SP IP), Reliever Innings Pitched (RP IP), and our response variable Wins Above Replacement (WAR). 6 pitchers had missing exLI (5 in train, 1 in test) and were removed from the data set. I also created a League variable based on whichever league the pitcher played in (either Central or Pacific). The amount of pitchers on each team in the training set is fairly even, with the Orix Buffaloes having the most with 97. The Pacific League also has the most among the two leagues with 551 pitchers. Some highlights include Los Angeles Dodgers ace Yoshinobu Yamamoto's legendary 2021 and 2022 campaigns with the Orix Buffaloes, where he dominated the data set with 18 wins (2021), back-to-back 193.2 IP seasons (2021–22), and over 200 strikeouts in both years, the only pitcher to surpass that mark. He also posted a peak WAR of 6.5 (2021), the highest WPA at 5.92 (2022), and an exceptional RE24 of 49.7 (2021). Other notable outliers include Hideaki Wakui's 13 losses (2023), Yaku Cho's extreme ERA of 13.06 (2021), and Noboru Shimizu's 72 appearances and 50 holds (2021). In the bullpen, Raidel Martinez notched 43 saves (2024), while Yoshihisa Hirano recorded the highest leverage index (pLI) of 3.23 (2024), reflecting heavy usage in critical situations. Finally, Tatsuya Imai posted a remarkable clutch score of 5.02 (2021), signaling rare performance in high-pressure moments. The distributions for each variable can be found in Tables 3-6 in the appendix as well as some interesting graphs in Figures 2-4.

Correlations

A correlation analysis shows that pitchers who log more innings, generate higher strikeout totals, and contribute positively to win-based and run-prevention metrics tend to post higher WAR values. Conversely, allowing more true runs, walks, or elevated FIP scores is associated with lower WAR. This reinforces the idea that starters, who accumulate more innings, are generally more valuable than relievers. The strongest positive relationships were between WAR and Starting Pitcher Innings Pitched ($r = 0.73$) and WAR and Strikeouts ($r = 0.63$), highlighting workload and strikeout ability as the most important drivers of value. The strongest negative relationships were between WAR and trueRunsAllowed ($r = -0.425$) and WAR and FIP ($r = -0.413$).

Exploratory Data Analysis (EDA) Plots

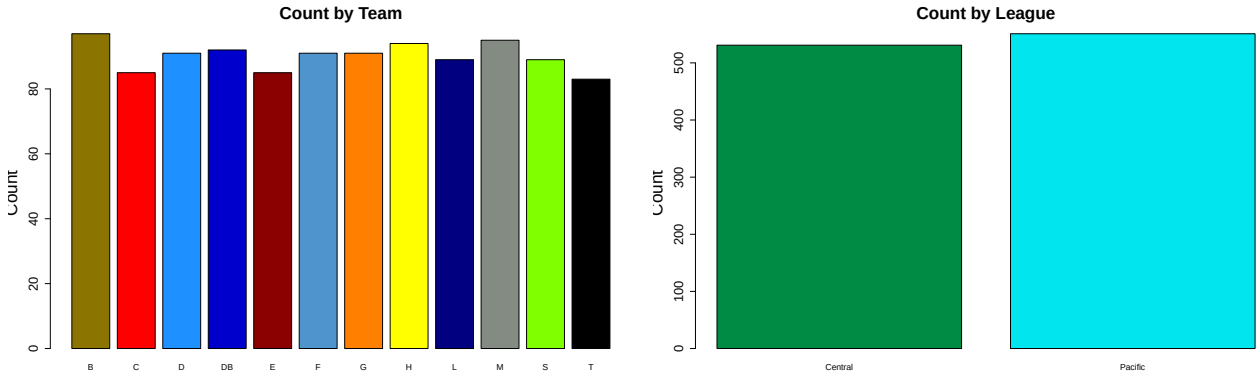


Figure 2: Barplots of Categorical Variables

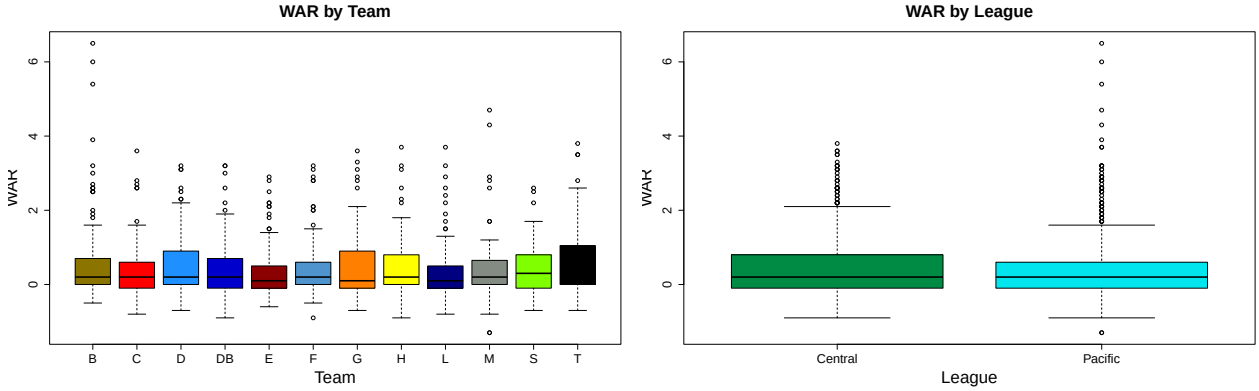


Figure 3: Boxplots of Categorical Variables vs. WAR

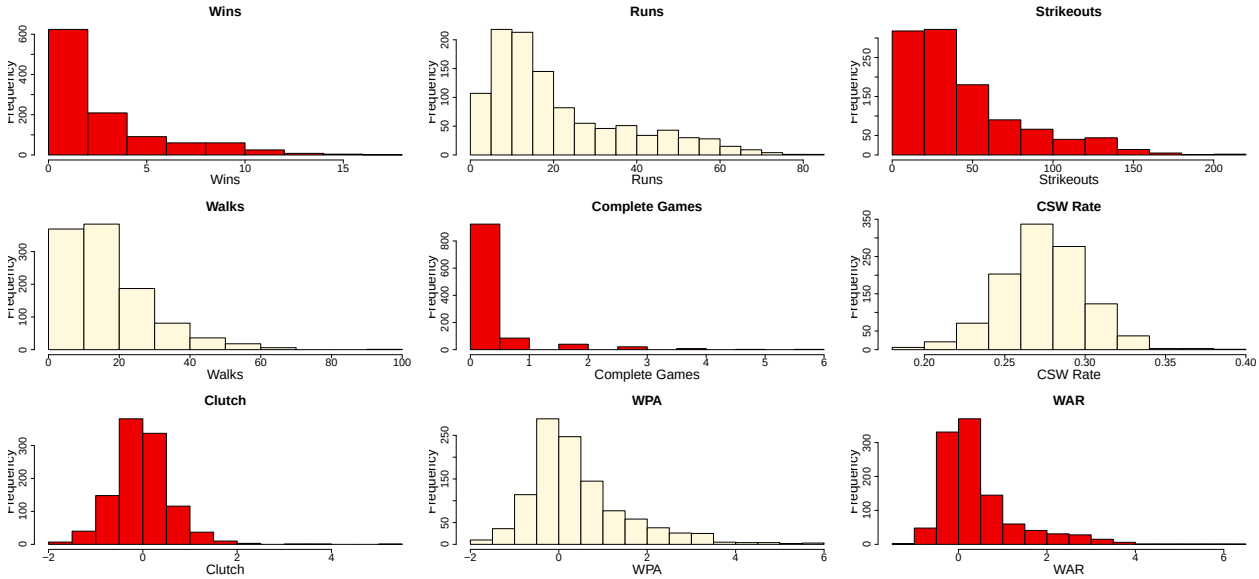


Figure 4: Histograms of Continuous Variables

Initial Modeling

Firstly, I removed the variables Balls, Strikes, and Pitches before fitting the initial model, as including them produced unusual error outputs in the R code. When running the initial model, I also excluded the Team and League factors, since neither was statistically significant and both added unnecessary degrees of freedom. The resulting model performed well, with a residual standard error of 0.3466 on 990 degrees of freedom, an adjusted R-squared of 0.8543, and an F-statistic of 70.67 on 91 and 990 DF ($p < 2.2e\text{-}16$). This initial model, however was far too complex for interpreting or generalizing to other data as there were over 100 coefficients. To narrow down which predictors were most significant, I ran a forward, backward, and stepwise selection algorithm. This gave me what I am calling a "core model." My core model consisted of 23 variables: StartingPitcherIP, trueRunsAllowed, RE24, Homeruns, RelieverInningsPitched, StrikeOut, Walks, REW, Season, Games, Hits, IntentionalWalks, GamesStarted, FIPIndexed, BattingAverageAgainst, WPA, StolenBases, Runs, WildPitch, RunsAllowedIndexed, RunsAllowedFIPDifference, Loses, pLI. I then checked the VIF values of these variables and saw major multicollinarity between many of these variables. Firstly, RE24 and REW are both run expectancy measurements; however, REW is scaled by WAR and is thus redundant with our response variable so it was removed in favor of RE24. I also removed Runs, RunsAllowedIndexed, and RunsAllowedFIPDifference as WAR already accounts for runs allowed, so these variables all increase multicollinearity and are redundant. I also removed Games, Season, and Wild Pitches as Games was redundant, the test set only uses data from 2025, and Wild Pitches was neither statistically significant nor a variable I found useful. This left me with a cleaner model of 16 variables.

I also explored the following six interactions: StartingPitcherIP:StrikeOut, pLI:WPA, IntentionalWalks:WPA, trueRunsAllowed:FIPIndexed, Walks:Homeruns, and RelieverInningsPitched:pLI. With StartingPitcherIP:StrikeOut, I wanted to see if pitchers with higher strikeout totals and more innings pitcher contributed more to WAR. With pLI:WPA, I wanted to see if pitchers in high leverage (pLI) games with high WPA gained an additional WAR boost. This interaction could capture value for relievers pitching in big moments. I also wanted to see if intentional walks, when issued strategically in high-leverage situations, could actually correlate with higher WPA and therefore more WAR. Using trueRunsAllowed:FIPIndexed, I wanted to test whether pitchers who both allow a high number of runs and post poor advanced metrics (high FIP) show consistent underperformance. I also wanted an interaction that demonstrated when WAR is hurt disproportionately so I chose Walks:Homeruns. Lastly, I wanted to see if relievers who throw more innings in high-leverage spots receive additional WAR credit compared to those who pitch the same workload in low-leverage situations. Of these interactions, only StartingPitcherIP:StrikeOut and Walks:Homeruns were statistically significant. These interactions are plotted below in Figure 5.

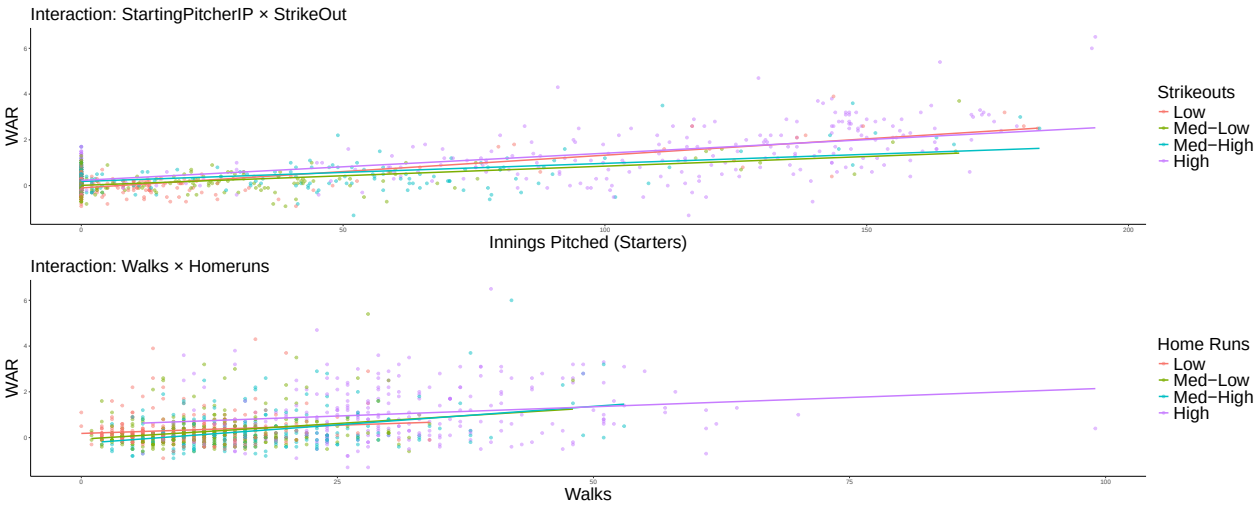


Figure 5: Interaction Plots of StartingPitcherIP:StrikeOut and Walks:Homeruns

Diagnostics and Model Selection

Next, I ran model diagnostics on my clean model. The diagnostic plots revealed that the residuals are fairly evenly scattered around 0 for lower fitted values but fan out towards higher fitted values, indicating non-constant variance and a violation of the homoscedasticity assumption. Additionally, the Q-Q plot displays significant right-tail and left-tail skewness, suggesting a violation of the normality assumption. These patterns motivate a transformation of the response variable. The residuals vs predictor plots showed no evidence of curvature so we can assume no assumptions are violated. Given the evidence of heteroscedasticity and non-normality, I will apply a Yeo-Johnson Transformation to the response variable (WAR). I used a Yeo-Johnson transformation instead of Box-Cox as Box-Cox does not work when the response variable has negative values. A player having negative WAR is meaningful data so I opted to use Yeo-Johnson. The optimal λ I got from the Yeo-Johnson transformation was ~0.42 but for ease of interpretability I will use $\lambda = 0.5$ which is equivalent to a square root transformation, transforming the data to $\sqrt{WAR}$. Because the transformed model ($\lambda = 0.5$ and $\lambda = 0.42$) provided little improvement and still exhibits evidence of heteroscedasticity and non-normality, I will proceed using the clean model, not the transformed model, for ease of interpretation. All of these results are laid out below in Figure 6.

Next, I examined the observations with large leverage, standardized residuals, studentized residuals, proportional influence, Cook's distance, and DFFITS. After running these diagnostics, it became clear that observations 49, 1004, and 988 were outliers. These observations correspond to some of the top NPB pitchers who have transitioned to MLB: Tomoyuki Sugano, Roki Sasaki, and Tatsuya Imai. Sugano, now with the Baltimore Orioles, is one of the most decorated pitchers in NPB history, having won three Central League MVP awards (2014, 2020, 2024), two Eiji Sawamura Awards (2017, 2018), three Central League wins titles (2017, 2018, 2020), four ERA titles (2014, 2016–2018), two strikeout titles (2016, 2018), eight NPB All-Star selections, three Golden Glove Awards (2016–2018), and pitched a no-hitter in 2018. Sasaki, who now pitches for the Los Angeles Dodgers, is a record-breaking phenom, having thrown a perfect game on April 10, 2022, achieved 19 strikeouts in a single game (an NPB record), 13 consecutive strikeouts in a single game (a world record), and earned two NPB All-Star selections. Imai, perhaps the most MLB-ready pitcher currently in NPB, is a two-time NPB All-Star (2021, 2024). The presence of these top-tier, MLB-bound pitchers as influential points is expected, and their influence is substantively meaningful; therefore, they will be retained in the model.

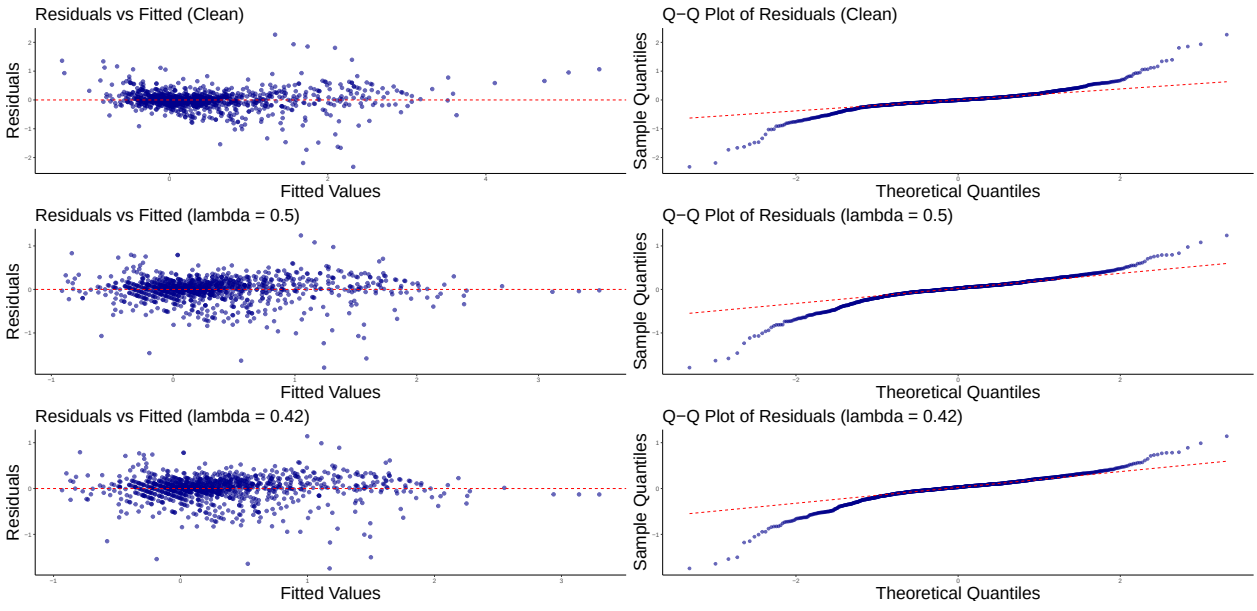


Figure 6: Clean, Y-J (lambda = 0.5), and Y-J (lambda = 0.42) Model Diagnostics

Final Model Summary Table and Diagnostics

	Estimate	Std_Error	CI_Lower	CI_Upper	T_Value	P_Value
(Intercept)	-0.121	0.105	-0.328	0.086	-1.151	2.499e-01
StartingPitcherIP	0.014	0.001	0.013	0.015	25.562	9.953e-113
trueRunsAllowed	-0.202	0.030	-0.260	-0.143	-6.766	2.185e-11
RE24	0.010	0.003	0.004	0.017	3.201	1.412e-03
Homeruns	-0.059	0.010	-0.078	-0.039	-5.956	3.503e-09
RelieverInningsPitched	0.016	0.001	0.014	0.018	14.527	9.467e-44
StrikeOut	0.014	0.001	0.012	0.017	11.659	1.204e-29
Walks	-0.017	0.003	-0.023	-0.012	-5.939	3.881e-09
Hits	-0.010	0.001	-0.013	-0.008	-7.807	1.398e-14
IntentionalWalks	0.020	0.011	-0.002	0.041	1.802	7.175e-02
GamesStarted	0.041	0.005	0.031	0.052	7.586	7.183e-14
FIPIndexed	0.004	0.001	0.002	0.006	3.595	3.396e-04
BattingAverageAgainst	1.198	0.376	0.459	1.936	3.181	1.510e-03
WPA	-0.101	0.023	-0.146	-0.057	-4.443	9.817e-06
StolenBases	-0.009	0.004	-0.018	-0.000	-2.069	3.881e-02
Loses	-0.021	0.009	-0.038	-0.004	-2.468	1.373e-02
pLI	0.074	0.035	0.005	0.144	2.101	3.592e-02
StartingPitcherIP:StrikeOut	0.000	0.000	0.000	0.000	7.636	4.971e-14
Homeruns:Walks	-0.001	0.000	-0.001	-0.001	-4.458	9.140e-06

Table 1: Final Model Summary with 95 Percent Confidence Intervals

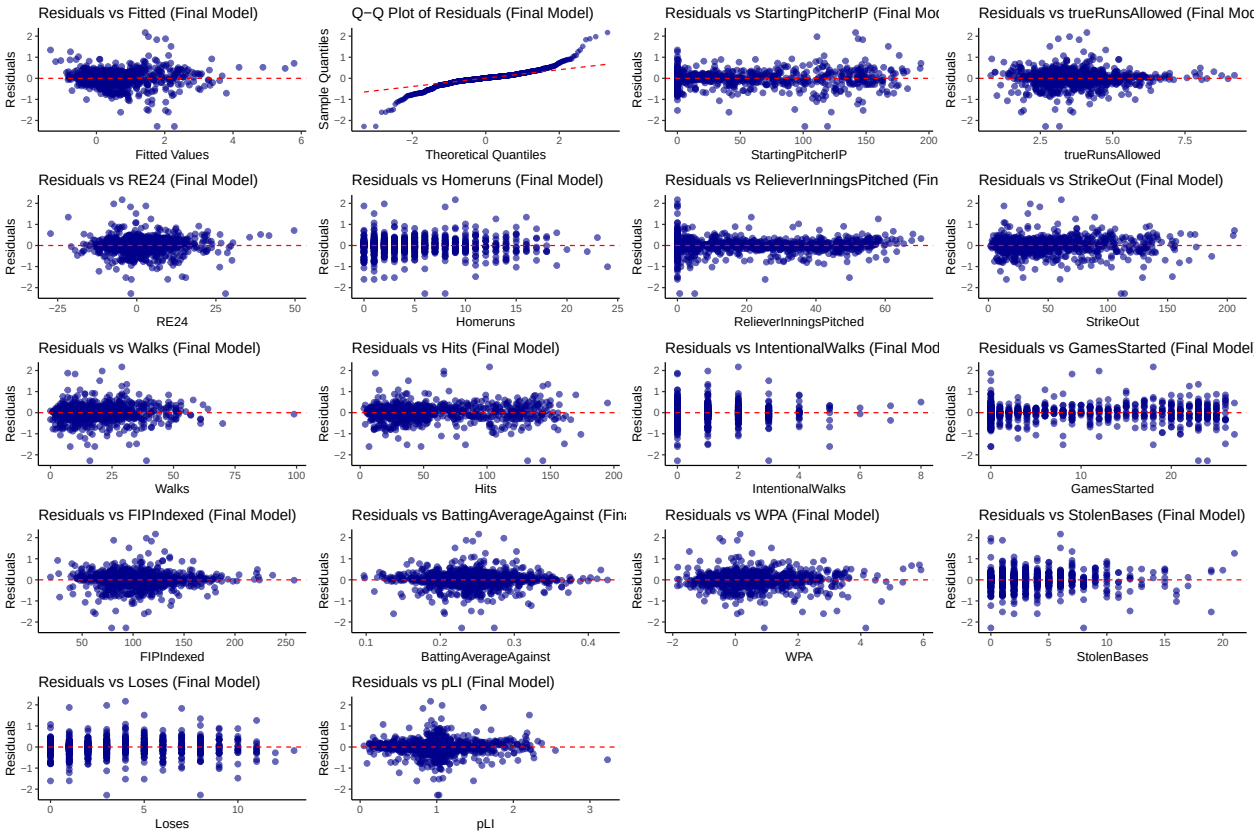


Figure 7: Final Model Diagnostics

Final Model Inference and Results

As shown in Figure 7 on the previous page, our final model has violated assumptions of constant variance, normality, and homoscedasticity so regardless of the findings that come from our final model, the results may be somewhat unreliable or hard to extrapolate to other data. We can see however, no curvature or patterns in the residual plots of our predictors.

To test the effectiveness of our final model compared to the intercept-only model, I performed an F-test. The null hypothesis (H_0) was that all the regression coefficients, except the intercept, are equal to zero, implying that the model with predictors does not explain significantly more variation than the intercept-only model. The alternative hypothesis (H_A) was that at least one predictor has a nonzero coefficient. From the summary output of our final model we have an F-statistic of 320.5 on 18 and 1063 degrees of freedom and a p-value of less than 2.2e-16, which is well below the significance level of $\alpha = 0.05$. Therefore, I reject the null hypothesis and conclude that the final regression was effective compared to the intercept-only model.

Our final model produced an adjusted R-squared of 0.8418, indicating that it explains approximately 84.18% of the variance in the data. While this is a strong result, it may also reflect some degree of over-fitting.

Based on the summary of our final model in Table 1, for a pitcher with all predictors set to zero, we would expect a WAR of approximately –0.12. The following predictors were statistically significant at least at the 0.1 level and should be interpreted while holding all other variables constant. Each additional inning pitched as a starter is associated with a 0.014 increase in WAR on average, while each additional true run allowed is associated with a 0.20 decrease in WAR on average. A one-unit increase in RE24 is associated with a 0.010 increase in WAR on average. Each additional home run allowed decreases WAR by about 0.059 on average. For relievers, each inning pitched is associated with a 0.016 increase in WAR on average. Each strikeout increases WAR by about 0.014 on average, while each walk decreases WAR by about 0.017. Similarly, each additional hit allowed decreases WAR by about 0.010 on average. Intentional walks were marginally significant ($p = 0.072$), suggesting a possible positive effect on WAR, though this should be interpreted cautiously. Each additional game started is associated with a 0.041 increase in WAR on average. A one-unit increase in FIP Indexed increases WAR by 0.004 on average. Batting average against was also significant, with a one-unit increase associated with an estimated 1.20 increase in WAR, though this is large and may capture interaction with other performance measures. A one-unit increase in WPA decreases WAR by 0.10 on average. Each stolen base allowed decreases WAR by about 0.009, while each loss decreases WAR by about 0.021. Pitch leverage (pLI) was positively associated with WAR, with a 0.074 increase for every unit increase in leverage. Two interaction effects were also significant. The interaction between starting pitcher innings pitched and strikeouts was positive, indicating that strikeouts contributed even more strongly to WAR for pitchers who threw more innings. Specifically, each additional strikeout by a starter added about 0.000048 WAR above and beyond the individual effects of innings and strikeouts. Conversely, the interaction between walks and home runs was negative, suggesting that pitchers who allowed both walks and home runs were penalized more heavily in terms of WAR, with each additional walk–home run combination decreasing WAR by about 0.001 on average. Returning to our research hypotheses, we observe a clear positive relationship between pitching workload and effectiveness and higher WAR. Metrics such as innings pitched, strikeouts, and game starts consistently predict higher WAR, while negative outcomes such as true runs allowed, walks, home runs, and hits reduce WAR. The interaction terms highlight the compounding effects of dominance (innings $\times$ strikeouts) as well as inefficiency (walks $\times$ home runs).

Based on the scatterplot in Figure 8 on the next page and the correlation coefficient of 0.955, our model performs extremely well in predicting new observations. Overall, our model captures both volume and quality of pitching performance, aligning with WAR's role as a comprehensive measure of pitcher value.

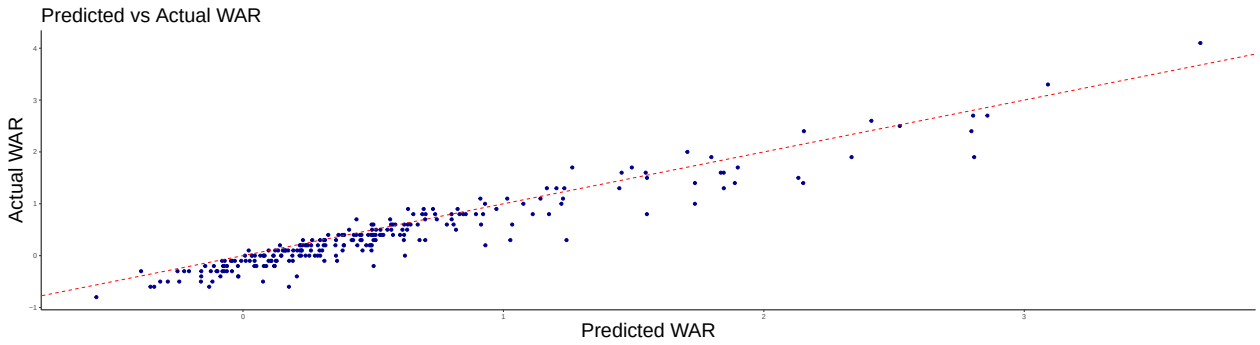


Figure 8: Predicted vs Actual WAR

To further evaluate the effectiveness of the model, I randomly selected one pitcher from the test set. The chosen observation (170) corresponded to Kazuki Sugiyama, a 27-year-old right-hander for the Fukuoka SoftBank Hawks, who had posted a 0.9 WAR as of August 17. The model's 95% confidence interval for the mean prediction of WAR ranged from 0.738 to 0.910, while the 95% prediction interval for a single observation ranged from 0.110 to 1.538. This indicates that the model's estimate closely aligned with the actual value, while also capturing the observed outcome within the broader prediction interval, consistent with expected model behavior.



Figure 9: Kazuki Sugiyama pitching for the Fukuoka SoftBank Hawks

The final model regression equation is:

$$
\begin{aligned}
\widehat{\text{WAR}} = \; & -0.12 + 0.014 \cdot \text{StartingPitcherIP} - 0.202 \cdot \text{trueRunsAllowed} + 0.010 \cdot \text{RE24} \\
& - 0.059 \cdot \text{Homeruns} + 0.016 \cdot \text{RelieverInningsPitched} + 0.014 \cdot \text{StrikeOut} \\
& - 0.017 \cdot \text{Walks} - 0.010 \cdot \text{Hits} + 0.020 \cdot \text{IntentionalWalks} \\
& + 0.041 \cdot \text{GamesStarted} + 0.004 \cdot \text{FIPIndexed} + 1.198 \cdot \text{BattingAverageAgainst} \\
& - 0.101 \cdot \text{WPA} - 0.009 \cdot \text{StolenBases} - 0.021 \cdot \text{Loses} \\
& + 0.074 \cdot \text{pLI} + 0.000048 \cdot \text{StartingPitcherIP:StrikeOut} \\
& - 0.001 \cdot \text{Homeruns:Walks}
\end{aligned}
$$

Top 20 NPB Pitchers by Predicted WAR

Player	predicted_WAR_2025	WAR_2025
Moinelo, Livan	3.68	4.10
Imai, Tatsuya	3.09	3.30
Duplantier, Jon	2.86	2.70
Murakami, Shoki	2.81	1.90
Miyagi, Hiroya	2.80	2.70
Itoh, Hiromi	2.80	2.40
Yamasaki, Iori	2.52	2.50
Sumida, Chihiro	2.41	2.60
Kitayama, Koki	2.34	1.90
Griffin, Foster	2.15	2.40
Takahashi, Hiroto	2.15	1.40
Saiki, Hiroto	2.13	1.50
Ohzeki, Tomohisa	1.90	1.70
Morishita, Masato	1.89	1.40
Akahoshi, Yuji	1.85	1.30
Ihara, Takato	1.85	1.60
Kay, Anthony	1.83	1.60
Kuri, Allen	1.80	1.90
Jackson, Andre	1.74	1.40
Azuma, Katsuki	1.74	1.00

Table 2: Top 20 NPB Pitchers by Predicted WAR

When evaluating these pitchers, I will focus on strikeouts, K-BB%, SwStr%, K%, K/9, and xFIP. I believe these stats are the best suited for predicting MLB success because they isolate a pitcher's raw "stuff" and their direct impact on the game. Right now, NPB is in the middle of a "dead ball era" (balls simply aren't carrying) which has produced some eye-popping numbers: 7 of 23 qualified pitchers have an ERA under 2.00, and 20 of 23 are under 3.00. That's unheard of, and it means many pitchers may be riding inflated stat lines. By focusing on what pitchers can actually control, such as missing bats, limiting walks, generating swing-and-miss, and defense-independent metrics like expected FIP, we get a clearer sense of who has real MLB-caliber ability versus who might just be a "dead ball merchant." The stats I've chosen "travel well" because they capture both the ability to overpower hitters and to avoid self-inflicted damage. A high K-BB% and strong strikeout totals are among the most predictive indicators of MLB success, and in today's NPB environment, any ERA above 3.00 should already raise serious doubts about a pitcher's viability as an MLB prospect. With these criteria in mind, I now turn to individual evaluations of the 2o top NPB pitchers laid out in Table 2 above (as well as two other interesting pitchers) and will discuss their statistical profiles and MLB potential.

Top 12 NPB Pitcher with MLB Potential

- **Liván Moinelo** is a 29-year-old Cuban left-hander having arguably NPB's most dominant 2025 season for the Fukuoka SoftBank Hawks. He transitioned from coming out of the bullpen as one of NPB's best relievers to a starter in 2024. He currently ranks first in NPB among qualified pitchers with a 1.13 ERA, second in strikeouts with 144 and SwStr% with 14.4%, first in K-BB% with 22.4%, first in K% with 28.3%, third in K/9 with 9.58, first in WPA with 4.87, and lowest xFIP of 2.31 in 19 starts on the season. He mixes in a slider, curveball, and change-up along with his four-seam fastball, which averages around 93 mph and is statistically the most valuable pitch in NPB among qualified pitchers. His curveball is very effective with a 35% CSW% and 66% GB%. In his last start on August 10, he threw a complete-game shutout where he racked up an unbelievable 13 strikeouts and 28 whiffs. He's already had four games this year with 12+ strikeouts, including an 18-strikeout game against the Swallows. He has thrown two other complete games this year, one of which was also a shutout. Due to his Cuban citizenship, he likely will never go to MLB (although he would be very successful if he did) so I will not consider him in my final ranking.
- I believe **Tatsuya Imai** is the best Japanese pitcher in NPB. The 27-year-old right-handed ace of the Saitama Seibu Lions ranks among the league's very best overall and is expected to make the jump to MLB soon. He has stayed healthy throughout his career despite his small and skinny frame. He dominates right handed hitters with a fastball that tops out at 100 and regularly sits 95 (the fastest in NPB), as well as the fastest slider in NPB, a change-up, and a splitter. Against righties, he primarily uses his four-seam fastball and slider and against lefties he mixes in his change-up which has a remarkable 73% GB%. Among pitchers with at least 80 innings of work, he currently has the fourth highest K-BB% in NPB at 20.8%, third highest K% at 27.9%, and third most strikeouts with 137 (including a 17 strikeout game against the BayStars in June), and the second highest K/9 at 9.81. He leads NPB in Whiff% with 31.9% and SwStr% with 15.5% and is fourth in NPB with a CSW% of 31.7%. His walk rate has decreased steadily from 16.8% in 2020 to 7.1% in 2025 while increasing his K% from 14.2% in 2020 to 27.9% in 2025 as he continues to improve his command of his arsenal and overwhelm hitters with his velocity. He has been on an upward trajectory his whole career and will make a strong addition to any rotation in MLB in the near future. I expect Imai to be the most prized NPB prospect when it comes time for him to be posted (either this off-season or next), especially considering he is only 27 years old.
- The former major leaguer for the Arizona Diamondbacks and MLB Pipeline's 2017 Pipeline Pitcher of the Year, **Jon Duplantier** is a 31-year-old American right-handed pitcher for the Hanshin Tigers. While not yet qualified as he missed two starts for load management, he has currently pitched to a 1.39 ERA with 113 strikeouts while tallying 2 complete game shutouts in just 15 starts. Among pitchers with at least 80 innings of work, he leads all of NPB with a K% of 32.4%, K/9 of 11.22, K-BB% of 26.6%, HR% of 0.3%, CSW% of 33.2% and xFIP of 2.00. He is second in NPB with a Whiff% of 28.1%, and fourth with a SwStr% of 12.6%. Duplantier has displayed a wide arsenal of pitches from a curveball that piles on the whiffs and is faster than most, a four-seam fastball that sits 93, as well as one of the three fastest sliders in NPB. His slider is elite with a 40% CSW% and 68% GB%, although he doesn't use it against lefties. His change-up also has a 39% CSW% and 76% GB%. By the time he finishes this dominant season, I expect the 6'4" Jon Duplantier to be gearing up for a return to MLB despite his older age.
- **Shoki Murakami** is a 27-year-old right-hander for the Hanshin Tigers. He had a historic rookie season in 2023, when he was named an NPB All-Star, led the Central League in ERA, and became the first Central League rookie ever to win both Rookie of the Year and Central League MVP in the same season. Murakami is a workhorse, going at least six innings in all but three starts this season. He doesn't strike many batters out, however, tallying only 106 strikeouts so far this season. His K-BB% of 16.0%, K% of 20.2%, 6.98 K/9, Whiff% of 22.0%, and SwStr% of 10.7%, don't exactly jump off the page either. That being said, he doesn't walk many batters, sporting a top-five BB% of 4.2%, nor does he give up a lot of hard contact, with a 1.1% HR%. He has a wide arsenal, with six different pitches he can throw for strikes, including a 65 mph slow curve, a 75 mph slider, and a 90 mph fastball. His 1.91 ERA contrasts with his 2.89 xFIP. Despite clearly not being back to his former MVP form, Murakami has

thrown three complete-game shutouts and provides length every time he pitches, with above-average command and a unique pitch mix. He also leads NPB with a 31.7% Soft-Contact rate. He wouldn't be my fourth-ranked prospect as the model suggests, but I am confident an MLB team could find value in a relatively young pitcher with such a deep arsenal.

- **Hiroya Miyagi** is the 23-year-old left-handed ace of the Orix Buffaloes. He was the 2021 Pacific League Rookie of the Year and has steadily improved every year since joining NPB. Miyagi pitched 5 innings in relief, striking out 7, in Japan's pool play game against the Czech Republic in the 2023 World Baseball Classic. He currently has a 2.55 ERA, which reflects more of Orix's poor defense than Miyagi's ability to miss bats, as exemplified by his 2.30 trueRunsAveraged, 25.7% K%, and 31.7% CSW%—all top-three marks—as well as his xFIP, which is second in NPB at 2.36. He also ranks top four in K-BB% (20.6%), Whiff% (25.1%), K/9 (9.33), and SwStr% (12.0%), and is top four in NPB in strikeouts with 121 despite throwing only 116 innings thus far. He relies heavily on his four-seam fastball, slider, splitter, slow curveball, and change-up, all of which rack up whiffs despite his relatively low velocity across the board. His slider is especially valuable, ranking second in NPB behind Imai's. Lefties like Miyagi need deep enough pitch mixes to build confidence that they can get out right-handed batters on MLB. His BB% is up and his fastball velocity is down this season; however, he has still managed to pitch at least six innings in 15 of 17 starts, though he has yet to complete a full game. He also features a noteworthy pickoff move. His improvement since his rookie year, combined with his consistency and command, are all promising traits that should warrant an MLB nod once he gains a few more years of experience, despite his 5'6" build.
- Another member of the 2023 World Baseball Classic gold medal team as well as the 2020 Summer Olympic gold medal team, **Hiromi Itoh** is a 27-year-old right-handed pitcher for the Hokkaido Nippon-Ham Fighters. Despite standing just 5'8", he currently leads NPB with 146 strikeouts, the lowest BB% at 3.5%, and five complete games including one shutout—a testament to his excellent mechanics and body control. Itoh features a seven-pitch arsenal but relies primarily on his four-seam fastball (93 mph), slider (83 mph), and splitter (88 mph), all of which have strong movement. His wide repertoire keeps batters guessing. He uses his change-up mostly against lefties. He currently ranks second in NPB with a 21.1% K-BB%, fourth in K% at 24.7%, fifth in SwStr% at 11.7%, and posts a 9.02 K/9 which is also a top-5 mark. Despite pitching to a 2.66 ERA, he holds the third-lowest xFIP in NPB at 2.41, underscoring that he is not solely responsible for the Fighters' defensive shortcomings. He is improving every year, showing gains in both control and velocity, and is currently in the midst of the best season of his career. An MLB opportunity seems inevitable in the near future.
- **Iori Yamasaki** is a 26-year-old right-hander for the Yomiuri Giants and was an NPB All-Star in 2024. So far in 2025, he has pitched to a 1.73 ERA over 19 starts with 103 strikeouts. He ranks top four in NPB in tRA (2.44), ground-ball rate (53.3%), and soft-contact rate (28.2%). He also holds top-10 marks with a 20.9% K%, 5.5% BB%, 15.4% K-BB%, 2.64 xFIP, 22.4% Whiff%, and 35.9% O-Swing%. A high O-Swing% means hitters have a tendency to chance Yamasaki's pitches outside of the strike zone. Yamasaki features a six-pitch arsenal. He throws both his four-seam fastball (which sits around 92 mph) and his splitter roughly one-third of the time, while also mixing in a cutter, slider, two-seam fastball, and curveball. Earlier this year, he broke an NPB record by opening the season with 35 consecutive scoreless innings. Despite that achievement, he has already had a few rough outings this season, has yet to complete a game, and has only recorded double-digit strikeouts twice. I struggle to decide whether Yamasaki is benefiting from the current dead-ball era or if he is truly emerging as one of NPB's next top pitchers. While there may be more MLB-ready options in Japan right now, his steady improvement over the years suggests that his stock will continue to rise.
- **Chihiro Sumida** is a 25-year-old lefty for the Saitama Seibu Lions. In the 2024 WBSC Premier12 championship, Sumida also struck out 16 batters in just 8.0 innings while walking none. Across 18 starts (including two complete games, one of which was a shutout), Sumida has pitched to a 2.42 ERA with 114 strikeouts against just 22 walks. He relies on a four-seam fastball, cutter, curveball, and an impressive change-up–splitter combo. His fastball averages 92 mph, his newly developed splitter grades as the most valuable in NPB, and his change-up ranks third. His splitter has a 33% CSW% and a 72% GB%. Statistically, he currently boasts a 22.7% K% and a 30.5% CSW% (6th in NPB), along with an

18.3% K-BB% and 2.54 xFIP (5th), 23.9% Whiff% (7th), 11.5% SwStr% (8th), and an 8.10 K/9. Since 2022, he has significantly improved his profile by lowering his BB% from 8.4% to 4.4% while raising his GB% from 43.9% to 49%. Durable and consistent, Sumida has gone 7+ innings in 13 of his 18 starts. I view Sumida as one of NPB's premier left-handed pitchers, and at just 25 years old, he remains a highly valuable asset with significant upside ahead. His young age might make the Lions hesitate to post him for a few seasons, however.

- **Koki Kitayama** is a 26-year-old right-handed pitcher for the Hokkaido Nippon-Ham Fighters. In the 2024 WBSC Premier12 championship, Kitayama struck out 10 batters in just 7.0 innings. He debuted in 2022 as a key bullpen piece before transitioning into the starting rotation, though he made only 14 starts in both 2023 and 2024. Now, through 16 starts in 2025 (possibly a result of load management, as he often gets 10 days between outings), he ranks second in NPB among qualified pitchers with a 1.54 ERA and third with four complete games. He currently ranks fifth in NPB with a 23.9% K%, sixth with a 17.7% K-BB% and 8.41 K/9, eighth with 104 strikeouts and a 2.83 xFIP, and ninth with a 28.4% CSW%. He also has the highest line-drive rate (LD%) among qualified pitchers at 11.6%. Since last year, his GB% has dropped from 44.8% to 39.5%, his K% from 27.2% to 23.9%, his SwStr% from 13.2% to 10.8%, and his Whiff% from 27.9% to 22.0%. Kitayama features an excellent four-seam fastball (rated sixth-best in NPB) which he throws about 50% of the time at 93.5 mph, along with a two-seam fastball, cutter, curveball, change-up (only to lefties), and splitter. Interestingly, he throws his curveball harder than his change-up. His splitter also ranks fourth in NPB. He really broke out in June when he struck out a career-high 14 batters against the Swallows, followed by taking a no-hit bid into the ninth inning in his next start against the Giants. I would consider Kitayama more of a work in progress in NPB with significant upside, rather than an MLB-ready finished product.
- **Foster Griffin** is a 6'3", 30-year-old American left-hander for the Yomiuri Giants. He appeared in seven games for the Kansas City Royals and Toronto Blue Jays between 2020 and 2022. Griffin began this season injured and missed the first month, but has since logged 76 innings across 12 starts. Among pitchers with 70 or more innings pitched, however, he has been extremely effective. He owns a 1.18 ERA, second in NPB only to Liván Moinelo, along with 77 strikeouts. He ranks fourth in NPB with a 26.1% K% and a 2.37 xFIP, fifth with a 20.7% K-BB%, and holds the second-lowest tRA at 1.53, trailing only Jon Duplantier. Impressively, he has yet to allow a home run this season. Griffin also carries a 9.12 K/9, 26.1% Whiff%, 31.0% CSW%, and 12% SwStr%. Last season, his 33.1% CSW led all of NPB. Griffin features a seven-pitch arsenal but relies less on his fastball than most pitchers, leaning instead on his slider (mostly against lefties), cutter, and change-up (against righties), along with a two-seam fastball and curveball. His velocity is relatively modest, with his fastball averaging just under 90 mph. While he has not pitched past the eighth inning nor reached double-digit strikeouts in a game this season, his consistency and effectiveness suggest he could receive another opportunity in MLB at the conclusion of this season, after three strong years in NPB.
- **Hiroto Takahashi** is a 23-year-old right-hander for the Chunichi Dragons. He was an NPB All-Star and the Central League ERA leader in 2024. He pitched three relief innings over three games for Samurai Japan in the 2023 WBC, striking out five batters. He also had 16 strikeouts in just 8.0 innings in the 2024 WBSC Premier12 championship. In 2025, he has started 20 games, including three complete games, two of which were shutouts. He has pitched to a 2.70 ERA and amassed 104 strikeouts in these starts. Among qualified pitchers, Takahashi's K% and K-BB% both rank thirteenth (19.5% and 12.4%, respectively). He is the first pitcher in this report with an xFIP over 3.00 (3.01). Takahashi ranks second in NPB in ground-ball rate (55.1%). He has the third-fastest four-seam fastball among qualified pitchers, averaging around 94 mph. He also throws a two-seamer, a cutter (mostly to righties), a curveball, and a splitter (primarily to lefties). His O-Swing% ranks eighth at 34.1%, third in Whiff% at 26.1%, and his SwStr% is 12.6%. He currently sports a 7.02 K/9. Statistically speaking, Takahashi is having a down year, with his K%, BB%, K-BB%, HR%, GB%, and CSW% all having declined. However, his soft-contact rate, as well as his Whiff% and SwStr%, have gone up since last year, showing his ability to miss bats and induce weak contact. Takahashi might be benefiting from the dead-ball era, but his velocity, pitch mix, and workhorse ability definitely provide a lot of upside for Takahashi. He likely will not be posted for a few years, however, due to his young age.

- **Hiroto Saiki** is a 26-year-old right-hander for the Hanshin Tigers. In the 2024 WBSC Premier12 championship, Saiki threw to a 1.69 ERA across 10.2 innings, racking up 10 punch-outs along the way for Japan. Across 19 starts in 2025, Saiki has pitched to a 1.57 ERA, including three complete games and two shutouts. He has accumulated 100 strikeouts along with 40 walks, the most walks of any pitcher in this report. Despite this sub-2 ERA, the 2024 NPB All-Star has shown some regression this year. Among qualified pitchers, his K% ranks thirteenth (19.5%), while his Whiff%, SwStr%, CSW%, BB%, K-BB%, and GB% all rank fifteenth or lower. He also owns a 3.19 xFIP and a 7.12 K/9. Despite this, he has posted the fourth-lowest soft-contact rate at 27.1% and the second-highest WPA at 3.84. Saiki throws a four-seam fastball, slider, slow curve (which is not very effective), and splitter. His fastball is notable for its elite spin rate, averaging around 2,600 rpm or higher (roughly 300 rpm above league average) and sits at 94 mph, an uptick from 2024. It is currently rated as the second-best fastball in NPB. Despite his walk issues, Saiki has not hit a single batter all season. It is also worth noting that he underwent Tommy John surgery as well as elbow arthroscopic surgery in November 2020, but he has since proven capable of pitching deep into games even without gaudy strikeout totals (his season-high is nine). His most notable performance came on March 15, 2024, in a preseason tune-up against the Los Angeles Dodgers, when Saiki was nearly untouchable, striking out seven hitters (including Shohei Ohtani) across five dominant innings while allowing just one hit. Dodgers manager Dave Roberts even called Saiki "MLB-caliber." If he can rein in his control and perhaps add another pitch to complement his fastball, Saiki could have the tools necessary to return to his 2024 form and make a jump to MLB this off-season.

13-20 NPB Pitchers with MLB Potential

- **Tomohisa Ohzeki** is a 27-year-old left-hander for the Fukuoka SoftBank Hawks who has expressed interest in going to MLB. While he has the third-best soft-contact rate in NPB (28.1%), his overall lack of strikeout ability raises questions about how successful he would be in MLB. He currently has a low K%, K-BB%, and K/9, with only 85 strikeouts in 124-plus innings of work. He also has one of the highest fly-ball rates in NPB (51.0%), and his four-seam fastball doesn't reach 90 mph despite being fourth-highest rated in NPB. His curveball is also especially ineffective. Despite being a workhorse, I don't see Ohzeki as an MLB-quality pitcher as he is likely helped by a combination of good defense and dead balls.
- **Masato Morishita** is a 27-year-old right-hander for the Hiroshima Toyo Carp. He was the 2020 Central League Rookie of the Year and a two-time NPB All-Star (2021, 2022). He also won gold with Japan in the 2020 Summer Olympics. Morishita has thrown the most innings in NPB this season, including three complete games, posting a 2.54 ERA with just 88 strikeouts for a 5.45 K/9. His K-BB% ranks 31st out of 40 among pitchers with at least 90 innings of work, although he does carry a 50.0% GB%. None of his 6 pitches really stand out, and his CSW% ranks 37th out of 40. His K% has dropped from 25.6% to 14.8% since his rookie year, a trend that is reflected across his other stats as well. He has regressed tremendously since his rookie season, and I no longer see him as worthy of a flyer for MLB even if he is inevitably posted.
- **Yuji Akahoshi** is a 26-year-old right-hander for the Yomiuri Giants. In 20 starts, he has posted a 2.48 ERA with 85 strikeouts, a 6.59 K/9, and just 16 walks. He currently has the lowest BB% in NPB at 3.5%, and his 14.9% K-BB% ranks tenth in the league. His xFIP ranks 7th at 2.71, while his GB% ranks 3rd at 53.6%. His CSW% is 28.7%, ranking 8th among qualified pitchers. Akahoshi throws six pitches, none of which are overly dominant. He is still relatively young to be posted and may be benefiting from the dead-ball era, given his limited ability to generate whiffs and strikeouts. However, he has shown flashes in previous seasons and his exceptional control is a notable strength.
- **Takato Ihara** is a 25-year-old left-hander for the Hanshin Tigers. He won't be posted as he is just a rookie, but he is currently in the running for Central League Rookie of the Year, having pitched to a 2.03 ERA (10th in the league) with 68 strikeouts across 22 games, 15 of which were starts. He has allowed only one home run in 93 innings, the fewest in NPB among pitchers with at least 90 innings. His 2.24 tRA ranks 5th, and his hard-hit rate is second to Jon Duplantier at 26.4%. Of his six pitches, his fastball sits at 88 mph and rates poorly, as does his slider, while his cutter may be his best pitch.

His Whiff%, SwStr%, and CSW% rank 30th, 30th, and 32nd among pitchers with at least 90 innings. He throws a four-seam fastball, cutter, curveball, and slider to lefties as well as an additional change-up and sinker to righties. While Ihara is having a solid rookie campaign, only time will tell if he can develop into an MLB-ready prospect.

- **Anthony Kay** is a 30-year-old American left-hander for the Yokohama DeNA BayStars who previously appeared in 44 games for the Toronto Blue Jays, New York Mets, and Chicago Cubs. In 18 starts, Kay has pitched to a 2.00 ERA with 90 strikeouts, a 2.87 xFIP, a 2.85 tRA, and an 8.1% LD%. He owns a league-leading 55.6% GB%, the second-lowest FB% at 36.3, and a bottom-five hard-hit rate at 29.4%. His four-seam fastball averages 94 mph, the second-fastest among qualified NPB pitchers, and his cutter currently grades as the most valuable in the league. Against lefties, Kay utilizes primarily a four-seam fastball, cutter, slider, and sinker against lefties. Against righties, however, he lowers his sinker and slider usage in favor of more cutter and change-up usage. Kay also ranks in the top ten in both Whiff% and SwStr%, while cutting his BB% from 9.0% last year to 6.8% in 2025. Kay's profile is promising, but I would like to see him raise his K% and truly overpower hitters with his fastball before projecting a successful return to MLB.
- **Allen Kuri** is a 33-year-old right-hander for the Orix Buffaloes and was an NPB All-Star in 2023. He will not be posted, having already signed a multi-year deal with Orix, but has pitched to a 2.64 ERA with 90 strikeouts across 19 starts. Among qualified pitchers, his K%, BB%, K-BB%, and xFIP all rank middle of the pack or lower, though his tRA sits tenth at 2.92. He owns the third-highest LD%, fifth-highest GB%, fourth-lowest FB%, and fourth-highest hard-hit rate in the league. His Whiff%, SwStr%, and CSW% also rank in the middle of the pack or worse. Kuri has likely benefited from an unusually strong ability to limit home runs this season, but beyond that he appears to be having a down year, and regression seems inevitable. He throws seven pitches, all with relatively low velocity, with his splitter (his most-used pitch at 42%) ranking as the third most valuable in NPB while his four-seam fastball is particularly ineffective.
- **Andre Jackson** is a 29-year-old American right-hander for the Yokohama DeNA BayStars who previously pitched for the Los Angeles Dodgers and Pittsburgh Pirates. In 19 starts, he has posted a 2.06 ERA with 88 strikeouts, though he has also issued 39 walks for a 7.9% BB%. His relatively low K% leaves him with just a 9.9% K-BB%. Among pitchers with at least 100 innings, his 3.42 xFIP ranks near the bottom, and his Whiff%, SwStr%, and CSW% are all mediocre at best. Jackson does benefit from a bottom-four LD% and bottom-three hard-hit rate, and he has shown the ability to pitch into the 6th and 7th innings. He throws a four-seam fastball that averages around 94 mph, along with a slider, cutter, curveball, and change-up that has a 39% CSW% and a 60% GB%. His fastball is his only above-average pitch. Overall, he has not shown the level of dominance expected from a former major leaguer seeking an MLB return.
- **Katsuki Azuma** is a 29-year-old left-hander for the Yokohama DeNA BayStars. He debuted in 2018, winning Central League Rookie of the Year, and has since been selected to three NPB All-Star teams (2018, 2023, 2024). In 19 starts this season, he has posted a 2.33 ERA with 97 strikeouts against just 18 walks. However, he has been hurt by the long ball, trailing only teammate Trevor Bauer for the most home runs allowed among qualified pitchers (14). His 2.7% HR% also leads NPB. Azuma's walk rate of 3.5% is the lowest among qualified pitchers, while his K% ranks 14th at 18.9%. His 3.58 tRA is the third-highest in the league, and his 36.1% hard-hit rate ranks second. His fastball, which averages around 88 mph, grades negatively but he does have a deep arsenal. He uses a slider against lefties and opts to use his change-up against righties. On the other hand, his O-Swing% leads NPB at 38.0%, and his CSW% ranks fourth at 31.3%. Remarkably, Azuma has pitched at least five innings in every start this year. In fact, the last time he failed to reach five innings in a start was May 15, 2022, an incredible streak of consistency and durability. His reliability, efficiency, and elite walk rate give him considerable upside on any given night. However, his extreme home run vulnerability is a major red flag, and leading NPB in HR% suggests limited MLB success.

Other Notable NPB Pitchers

- **Haruya Tanaka** is a 21-year-old rookie left-hander for the Chiba Lotte Marines. In 2025, he has posted a 2.48 ERA across 13 starts and 76 innings, recording 77 strikeouts against 27 walks. His 24.4% K% ranks 8th among pitchers with at least 70 innings, and his 2.56 tRA ranks 9th. Tanaka mixes a four-seam fastball (which has risen in velocity to 94 mph since last season) with a sinker, slider, cutter, and change-up. His sinker has a 65% GB%. His Whiff% of 27.8% ranks 4th in NPB, and his 12.7% SwStr% ranks 3rd, trailing only Duplantier, Imai, and Moinelo, the top three pitchers in the league. While his 8.5% BB% is relatively high, he is expected to refine his command as he gains experience. Although it is still early in his career and he will not be posted for quite some time, Tanaka's pitch mix, rising velocity, and bat-missing ability suggest significant upside. I fully expect Haruya Tanaka to pitch in MLB one day.
- **Kona Takahashi** is a 27-year-old right-hander for the Saitama Seibu Lions. An NPB All-Star in 2019, he has pitched to a 2.76 ERA across 17 starts, striking out 57 batters while issuing 31 walks. His K% (13.5%) has dropped from last year's 16.9%, while his BB% has risen from 6.6% to 7.3%. He currently sports a poor 5.08 K/9. Among pitchers with at least 100 innings, his K% and K-BB% are both the third lowest, while his xFIP (3.86) and tRA (3.79) are the third highest. He also owns the fifth-lowest Whiff% (16.9%), fourth-lowest soft-contact rate (19.0%), third-lowest SwStr% (7.6%), second-lowest CSW% (25.1%), and the lowest O-Swing% (26.9%) in NPB. Takahashi does have top-10 four-seam fastball velocity at 93 mph, along with a slider, cutter, curveball, and splitter in his arsenal. However, simply put, the 2025 version of Kona Takahashi is not an MLB-caliber pitcher, regardless of whether or not he is posted by the Lions.

Conclusion and Discussion

In conclusion, this case study demonstrates the utility of multiple linear regression for evaluating pitcher performance in Nippon Professional Baseball (NPB) and identifying players with Major League Baseball (MLB) potential. By narrowing the model to key predictors such as innings pitched, strikeouts, home runs, and walks, we explained over 84% of the variance in pitcher WAR. The inclusion of significant interaction effects, such as StartingPitcherIP $\times$ Strikeouts and Walks $\times$ Home runs, highlights how both volume and efficiency jointly determine player value. While the model produced strong results, it also exhibited violations of key regression assumptions, including non-constant variance and non-normality of residuals, suggesting caution when extrapolating beyond the data set. Furthermore, the current "dead ball era" in NPB emphasizes the importance of defense-independent metrics when projecting MLB success. By focusing on strikeout ability, command, and advanced measures like xFIP, we can better distinguish true MLB-caliber pitchers from those benefiting disproportionately from the run environment. Ultimately, this analysis identifies a shortlist of pitchers whose skill profiles should translate effectively to MLB competition. Future extensions could incorporate non-linear methods or machine learning approaches to enhance predictive accuracy, account for more complex interactions in the data, and integrate biomechanical metrics such as spin rate, horizontal extension, release point, and arm angle.

Based on my multiple linear regression model, stats from DeltaGraphs, WBC and Premier12 data, graphics and NPB updates from Yakyu Cosmopolitan and Gaijin Baseball, in-person scouting in Japan, and my own statistical analysis, I believe the top five MLB-ready pitchers currently in NPB are Tatsuya Imai, Jon Duplantier, Hiromi Itoh, Foster Griffin, and Hiroto Saiki. I also see MLB potential in Hiroya Miyagi, Chihiro Sumida, Koki Kitayama, Hiroto Takahashi, Shoki Murakami, and Iori Yamasaki with their stock likely to rise over the next few seasons as they continue to develop. Additionally, I believe Haruya Tanaka will almost certainly pitch in MLB when the time comes, as he is only a rookie, and Katsuki Azuma could be an intriguing dark horse candidate to make the transition later in his career.

Acknowledgments

I would like to thank Yuri Karasawa, Tosh Nagata, and Daisuke Haba for taking the time to meet with me in Japan, answering all of my NPB-related questions both in person and over the phone, and sharing valuable resources that helped me compile this report. I am equally grateful for their genuine interest in my career aspirations. Additionally, I would like to thank my friends within MLB for helping me connect with these contacts.

Appendix

	Variable	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.	SD
1	Wins	0.000	1.000	2.000	2.990	4.000	18.000	3.221
2	Loses	0.000	1.000	2.000	2.933	5.000	13.000	2.835
3	ERA	0.000	2.393	3.190	3.450	4.178	13.060	1.640
4	Games	2.000	13.000	22.000	25.387	34.000	72.000	15.860
5	GamesStarted	0.000	0.000	1.000	6.244	12.000	27.000	8.497
6	CompleteGames	0.000	0.000	0.000	0.261	0.000	6.000	0.752
7	ShutOuts	0.000	0.000	0.000	0.139	0.000	4.000	0.458
8	Saves	0.000	0.000	0.000	1.523	0.000	43.000	5.837
9	Holds	0.000	0.000	1.000	5.105	7.000	50.000	8.154
10	InningsPitched	10.000	23.083	41.000	54.995	69.250	193.667	43.377
11	TotalBattersFaced	34.000	100.000	172.000	228.974	291.750	764.000	175.973
12	Hits	3.000	21.250	35.000	49.434	66.000	195.000	39.397
13	Runs	0.000	9.000	16.000	21.327	30.000	81.000	16.372
14	EarnedRuns	0.000	8.000	14.000	19.361	26.750	76.000	15.089
15	Homeruns	0.000	1.000	3.000	4.418	6.000	24.000	4.196
16	Walks	0.000	9.000	14.000	17.286	23.000	99.000	12.114
17	IntentionalWalks	0.000	0.000	1.000	0.879	1.000	8.000	1.145
18	HitByPitch	0.000	1.000	2.000	2.125	3.000	17.000	2.125
19	WildPitch	0.000	0.000	1.000	1.403	2.000	13.000	1.607
20	Balk	0.000	0.000	0.000	0.085	0.000	3.000	0.313
21	StrikeOut	2.000	17.000	33.000	44.571	59.000	206.000	36.470
22	StolenBases	0.000	1.000	2.000	2.970	4.000	21.000	3.195
23	CaughtStealing	0.000	0.000	1.000	1.459	2.000	11.000	1.761
24	StrikeoutRate	0.032	0.155	0.191	0.193	0.227	0.431	0.056
25	WalkRate	0.000	0.060	0.081	0.084	0.104	0.265	0.033
26	StrikeoutMinusWalkRate	-0.107	0.068	0.110	0.109	0.149	0.342	0.065
27	HRRate	0.000	0.011	0.018	0.020	0.027	0.100	0.013
28	BattingAverageAgainst	0.097	0.216	0.243	0.245	0.272	0.426	0.047
29	WHIP	0.480	1.090	1.240	1.281	1.420	2.510	0.277
30	LeftOnBaseRate	0.325	0.704	0.762	0.756	0.813	1.111	0.093
31	RunsAllowedIndexed	0.000	75.000	100.000	107.194	129.000	465.000	48.649
32	FIPIndexed	19.000	86.000	102.000	104.165	119.000	258.000	28.750
33	ExpectedFIPIndexed	36.000	91.000	102.000	103.647	116.000	193.000	20.881
34	RunsAllowed	0.000	2.700	3.540	3.812	4.578	15.680	1.729
35	FieldingIndependentPitching	0.670	3.040	3.630	3.705	4.240	9.050	1.040
36	RunsAllowedFIPDifference	-3.470	-0.708	-0.090	0.106	0.710	10.370	1.372
37	ExpectedFIP	1.260	3.180	3.650	3.688	4.130	6.740	0.768
38	trueRunsAllowed	0.790	3.010	3.590	3.697	4.260	9.210	1.108
39	DefensiveEfficiencyRatio	0.485	0.675	0.709	0.705	0.733	0.909	0.050

Table 3: Standard and Advanced Statistics Summary Table

	Variable	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.	SD
1	GBToFBRate	0.330	0.810	1.000	1.054	1.230	3.560	0.378
2	LDRate	0.000	0.073	0.092	0.094	0.111	0.245	0.031
3	GBRate	0.222	0.400	0.453	0.452	0.500	0.731	0.077
4	FBRate	0.205	0.404	0.451	0.454	0.500	0.667	0.075
5	IFFBRate	0.000	0.077	0.110	0.112	0.143	0.385	0.054
6	HRToFBRate	0.000	0.036	0.058	0.062	0.083	0.333	0.041
7	RunsScored	0.000	7.000	15.000	21.836	31.000	111.000	19.981
8	RSPerNine	0.000	2.450	3.460	3.475	4.350	10.950	1.618
9	Strikes	85.000	245.250	433.000	579.477	739.000	2008.000	450.839
10	Balls	37.000	142.000	241.000	314.304	404.750	1141.000	233.707
11	Pitches	135.000	390.250	675.000	893.781	1143.250	2936.000	682.109
12	PullRate	0.129	0.321	0.354	0.354	0.388	0.593	0.056
13	CentRate	0.185	0.372	0.400	0.401	0.429	0.694	0.053
14	OppoRate	0.061	0.217	0.244	0.245	0.273	0.521	0.049
15	SoftContactRate	0.055	0.192	0.225	0.231	0.267	0.419	0.059
16	MidRate	0.194	0.379	0.424	0.424	0.467	0.688	0.069
17	HardHitRate	0.147	0.311	0.344	0.345	0.377	0.559	0.055

Table 4: Batted Ball Statistics Summary Table

	Variable	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.	SD
1	OutsideSwingRate	0.122	0.270	0.299	0.298	0.326	0.474	0.043
2	ZoneSwingRate	0.505	0.626	0.653	0.655	0.683	0.819	0.046
3	SwingRate	0.283	0.438	0.459	0.460	0.483	0.593	0.036
4	OutsideContactRate	0.338	0.567	0.625	0.623	0.683	0.929	0.087
5	ZoneContactRate	0.678	0.845	0.875	0.870	0.899	0.987	0.042
6	ContactRate	0.556	0.750	0.786	0.782	0.816	0.933	0.051
7	ZoneRate	0.338	0.431	0.454	0.454	0.477	0.558	0.034
8	FirstPitchStrikeRate	0.231	0.444	0.484	0.481	0.521	0.702	0.057
9	WhiffRate	0.067	0.184	0.214	0.218	0.250	0.444	0.051
10	PutAwayRate	0.038	0.139	0.164	0.166	0.192	0.356	0.043
11	SwStrRate	0.030	0.083	0.098	0.101	0.117	0.239	0.027
12	CStrRate	0.099	0.158	0.174	0.174	0.190	0.267	0.024
13	CSWRate	0.185	0.258	0.276	0.275	0.292	0.390	0.027

Table 5: Plate Discipline Statistics Summary Table

	Variable	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.	SD
1	WPA	-1.850	-0.248	0.160	0.445	0.898	5.920	1.129
2	RE24	-27.250	-3.007	1.015	1.769	6.080	49.700	8.039
3	REW	-3.150	-0.340	0.110	0.198	0.688	5.400	0.905
4	pLI	0.050	0.820	0.990	1.000	1.130	3.230	0.395
5	inLI	0.050	0.690	0.880	0.857	0.970	2.230	0.350
6	gmLI	0.050	0.780	0.870	0.938	1.110	2.210	0.355
7	exLI	0.010	0.872	1.040	1.064	1.220	2.700	0.330
8	ContextNeutralWPA	-2.080	-0.310	0.200	0.343	0.827	6.050	1.048
9	Clutch	-1.950	-0.380	-0.040	-0.010	0.320	5.020	0.612
10	StartingPitcherIP	0.000	0.000	2.000	36.026	61.000	193.667	52.404
11	RelieverInningsPitched	0.000	0.000	14.333	18.968	33.000	70.333	18.663
12	WAR	-1.300	-0.100	0.200	0.478	0.700	6.500	0.908

Table 6: WPA and Value Statistics Summary Table