Forecasting League of Legends Player Performance through In-Game Metrics and Key Statistics*

A Hybrid Model Predict the KDA for Bilibili Gaming (BLG) team players

Colin Sihan Yang

November 26, 2024

This paper develops a hybrid model to forecast the 2024 performance of Bilibili Gaming (BLG) team players by analyzing team statistics from the League of Legends Pro League (LPL). The model focuses on predicting the Kill-Death-Assist ratio (KDA) for individual players across different games, leveraging both traditional in-game metrics such as gold difference, kills, assists, and vision score, and advanced variables like total creep score (total CS) to provide a comprehensive view of player performance. By integrating machine learning techniques with esports-specific insights, this model forecasting player KDAs, identifying performance trends, but also offers a clear assessment of individual contributions to the team. It highlights top-performing players while identifying those whose performance may negatively impact the team's success. The findings offer actionable insights, enabling the BLG organization to make informed decisions on player contracts.

1 Introduction

In recent years, esports has evolved into a global phenomenon, with professional gaming leagues such as the League of Legends Pro League (LPL) attracting millions of viewers and significant organizational investments. As the competitive landscape intensifies, teams and organizations increasingly rely on data-driven approaches to gain a strategic edge. Among these, predictive analytics has emerged as a powerful tool to evaluate player performance, optimize team strategies, and make informed management decisions.

*Code and data are available at: https://github.com/hoodiexxx/lol_win_rate_prediction.

Bilibili Gaming (BLG), a prominent team in the LPL, faces the challenge of maintaining a competitive roster in the ever-changing meta of League of Legends. Accurate performance forecasting is critical for assessing individual player contributions, identifying areas for improvement, and making decisions about player contracts and roster changes.

This paper developing a hybrid model that leverages both traditional in-game metrics and advanced variables, such as total creep score (CS), to forecast player performance for BLG in the 2024 season. The model specifically focuses on predicting the Kill-Death-Assist ratio (KDA), a widely accepted indicator of individual performance, across various matches. Finally offers actionable insights into player roles and impacts, highlighting those who can consistently carry games and identifying players whose performance may not justify their perceived value.

The remainder of this paper is structured as follows: Section 2 provides an overview of the data. Section 3 provides the modeling approach, including simple and multiple linear regression models and a Bayesian hierarchical model. We then present our results in Section 4 and discuss the implications, limitations, then we predict the election in **?@sec-prediction**, and future research directions in Section 5.

The data simulation is done in python 3.9(Python Software Foundation 2023) with the following packages: Numpy (NumPy Developers 2023) and pandas(pandas Development Team 2023).

the data gathering and analysis is done in R (R Core Team 2023) with the following packages: knitr (Xie 2014), tidyverse (Wickham et al. 2019), ggplot2 (Wickham 2016), dplyr (Wickham et al. 2023), arrow (Richardson et al. 2024), here (Müller 2020), and Rstan (Stan Development Team 2024)

1.1 Estimand

the estimand for this research paper is the predicted Kill-Death-Assist ratio (KDA) for the BLG team players. The prediction is based on quantifying various in-game factors, including total cs, gold earned, vision score, damage to champion, which are used as predictors.

2 Data

2.1 Measurement

The dataset we obtain from oracleselixir (2024) is accumlated from multiple league of legend esport league like LCS, LEC, LCK, LPL, and the rest of global pro LoL.

The dataset utilized in this study was compiled from multiple professional League of Legends (LoL) esports leagues, including the LCS (North America), LEC (Europe), LCK (South Korea), LPL (China), and other global professional leagues. The data aggregation spans from 2023-2024 (23-24 season) and encompasses a wide variety of matches played at different competitive levels. This diversity ensures a comprehensive representation of professional gameplay styles, strategies, and performances across regions.

Each data point corresponds to a single game of one player, identified by a unique gameid, teamid and playerid and includes contextual information such as the league, split (Spring, Summer, or playoffs), split (Winter, Sping, etc), and year. The dataset captures a rich array of in-game metrics and performance indicators for teams and individual players. Team-level data includes variables such as match result (win/loss), game duration (in minutes and seconds), and pivotal game objectives (e.g., First Blood, First Tower, Baron Nashor, and Dragon counts). These metrics are critical for understanding team performance and macro-level strategies.

Player-level statistics include detailed in-game actions such as kills, deaths, assists, gold earned, damage dealt, and wards placed or cleared. These variables allow for a granular examination of individual contributions to a team's success. Additionally, positional data (e.g., top lane, jungle, mid lane, ADC, and support) is included, enabling role-specific analyses. The dataset also tracks champion selections and bans, providing insights into meta trends and drafting strategies over time.

To ensure consistency, all metrics are measured using standardized definitions as provided by the data source, oracleselixir (2024). This source is widely regarded as reliable within the esports analytics community. The dataset is structured to facilitate longitudinal analyses, enabling researchers to inspect trends and patterns across regions, seasons, and competitive tiers. Furthermore, derived metrics, such as the champion xp difference at 25 minutes (xpdiffat25), Gold Difference at 25 minutes (golddiffat25), and Damage Per Minute (dpm), are calculated to provide advanced insights into player and team efficiency.

These variables combined allow researchers to reliably analyze and predict the 23-24 season LPL BLG team's player performance.

However there are limitations in these measurements.The dataset combines data from multiple leagues (LCS, LEC, LCK, LPL, and others), each of which has distinct playstyles, metas, and competitive environments. These regional differences might introduce biases, as strategies or performance metrics effective in one league may not translate directly to another. For instance, the LCK may prioritize slower, macro-focused playstyles, whereas the LPL might exhibit faster, skirmish-heavy games. This variation could confound cross-regional comparisons unless properly controlled for (Tan 2020).

2.2 Raw Data

The raw data from oracleselixir@oracleselixir_oracles_nodate contains 52 columns, all the column headers are displayed below:

gameid	datacompleteness	url
league	year	split
playoffs	date	game
patch	participantid	side
position	playername	playerid
teamname	teamid	champion
ban1	ban2	ban3
ban4	ban5	pick1
pick2	pick3	pick4
pick5	gamelength	result
kills	deaths	assists
teamkills	teamdeaths	doublekills
triplekills	quadrakills	pentakills
firstblood	firstbloodkill	firstbloodassist
firstbloodvictim	team kpm	ckpm
firstdragon	dragons	opp_dragons
elementaldrakes	opp_elementaldrakes	infernals
mountains	clouds	oceans
chemtechs	hextechs	dragons (type unknown)
elders	opp_elders	firstherald
heralds	opp_heralds	void_grubs
opp_void_grubs	firstbaron	barons
opp_barons	firsttower	towers
opp_towers	firstmidtower	firsttothreetowers
turretplates	opp_turretplates	inhibitors
opp_inhibitors	damagetochampions	dpm
damageshare	damagetakenperminute	damagemitigatedperminute
wardsplaced	wpm	wardskilled
wcpm	controlwardsbought	visionscore
vspm	totalgold	earnedgold
earned gpm	earnedgoldshare	goldspent
gspd	gpr	total cs
minionkills	monsterkills	monsterkillsownjungle
monsterkillsenemyjungle	cspm	goldat10
xpat10	csat10	opp_goldat10
opp_xpat10	opp_csat10	golddiffat10
xpdiffat10	csdiffat10	killsat10
assistsat10	deathsat10	opp_killsat10
opp_assistsat10	opp_deathsat10	goldat15
xpat15	csat15	opp_goldat15
opp_xpat15	opp_csat15	golddiffat15
xpdiffat15	csdiffat15	killsat15
assistsat15	deathsat15	opp_killsat15
opp_assistsat15	opp_deathsat15	goldat20
xpat20	csat20	opp_goldat20
opp_xpat20	opp_csat20	golddiffat20
xpdiffat20	csdiffat20	killsat20
assistsat20	deathsat20	opp_killsat20
opp_assistsat20	opp_deathsat20	goldat25

These columns can be categorized into three types, elements of response variable, numeric predictors, and categorical predictors.

the elements of response variable are: - kills: The number of enemy champions defeated by the player. - deaths: The number of times the player was defeated by opponents. - assists: The number of assists a player contributes to a teammate's kills.

These three attributes will later be combined to calculate the Kill-Death-Assist ratio (KDA), which serves as the response variable for the model in the cleaned data section Section 2.3.

example of Numeric Predictors are: - totalgold: Total gold earned by the player during the game. - total.cs: Total number of minions or monsters killed by the player. - visionscore: A composite measure of vision provided and vision denied. - damagetochampions: Total damage inflicted by the player on opponents. - golddiffat25: Gold advantage or deficit compared to the opposing player/team at the 25-minute mark.

example of categorical predictors are: - position: The player's role in the game (e.g., top, jungle, mid, ADC, support). - league: The league where the match was played (e.g., LPL, LCK). - side: team or the blue side or the red side

2.3 Cleaned Data

The cleaned dataset used in this study focuses on matches from the League of Legends Pro League (LPL) involving Bilibili Gaming (BLG) during the 2024 season. To ensure the data aligns with the research objectives, the dataset was filtered to include only games played by BLG, with key variables retained for performance analysis. The cleaning process involved removing rows with missing values and selecting relevant columns, including contextual information such as game ID, league, year, split, and side, as well as in-game performance metrics like kills, deaths, assists, total gold, and vision score. A critical addition to the dataset was the calculation of the Kill-Death-Assist ratio (KDA), a widely used indicator of individual performance. The KDA was derived using the formula:

$$KDA = \frac{Kills + Assists}{Max(1, Deaths)} \tag{1}$$

ensuring robustness even when deaths were zero.

and we mutate the gamelength unit from second into minute.

The cleaned dataset was saved in Parquet formats to support flexibility in analysis and compatibility with various tools. This focused and structured dataset ensures reliability and completeness, providing a robust foundation for forecasting player performance, analyzing trends, and identifying key contributors to the analysis of BLG's players performance.

The first 6 rows of the dataset are displayed in Table 1.

Table 1: Sample of cleaned league of legend esport data

side	position	playername	champion	kills	deaths	assists	visionscore	totalgold	total.cs
Blue	top	Bin	K'Sante	1	0	6	21	10913	251
Blue	jng	Xun	Brand	8	0	8	34	12311	200
Blue	mid	knight	LeBlanc	4	0	6	19	13156	247
Blue	bot	Elk	Varus	1	2	8	34	11495	244
Blue	sup	ON	Ashe	3	1	10	140	9104	43
Red	top	Bin	Aatrox	0	2	2	35	11821	319

2.4 DATA Insight

2.4.1 Champion Pool for players

Table 2: the table that shows the champion pool for the BLG players

Table 2: Top 5 Most Played Champions by Player with Win Rate

Player Name	Champion	Position	Count	Win Rate
Bin	Renekton	top	18	0.78
Bin	Kennen	top	12	0.92
Bin	Gnar	top	11	0.73
Bin	K'Sante	top	11	0.55
Bin	Aatrox	top	10	0.70
Elk	Kai'Sa	bot	10	0.80
Elk	Kalista	bot	10	0.90
Elk	Lucian	bot	10	0.90
Elk	Ezreal	bot	9	0.78
Elk	Senna	bot	9	0.78
ON	Rell	sup	15	0.80
ON	Rakan	sup	13	0.69
ON	Nautilus	sup	11	0.73
ON	Alistar	sup	9	1.00
ON	Renata Glasc	sup	9	0.78
Wei	Sejuani	jng	6	0.83
Wei	Maokai	jng	5	1.00
Wei	Lillia	jng	4	1.00
Wei	Zyra	jng	4	0.75
Wei	Brand	jng	2	1.00

Player Name	Champion	Position	Count	Win Rate
Xun	Kindred	jng	12	0.83
Xun	Vi	jng	11	0.55
Xun	Xin Zhao	jng	8	0.88
Xun	Brand	jng	6	0.50
Xun	Lee Sin	jng	6	0.83
knight	Corki	mid	14	0.93
knight	Yone	mid	10	0.70
knight	Ahri	mid	8	0.88
knight	Karma	mid	7	0.86
knight	Neeko	mid	7	1.00
knight	Taliyah	mid	7	0.86
knight	Tristana	mid	7	0.57

The table Table 2 summarizes the top 5 most played champions for each Bilibili Gaming (BLG) player, including their positions, usage counts, and win rates. It highlights the players' preferences and effectiveness with specific champions, offering insights into individual strengths and team strategies.

- Bin (top lane) favors durable, aggressive champions like Renekton (18 games, 77.8% win rate) and Kennen (12 games, 91.7%). While most of his champions are highly effective, K'Sante has a lower win rate of 54.5%, suggesting Bin may not good at Tank champion.
- Elk (bot lane) excels on meta ADCs like Kai'Sa, Kalista, and Lucian, all boasting win rates of 80-90%, with consistent performance on utility champions like Senna and Ezreal (77.8%).
- ON (support) shines with engage-heavy champions such as Rell (80%) and Rakan (69.2%), but his Alistar stands out with a perfect 100% win rate, indicating strong synergy in specific scenarios. Wei (substitute jungle) shows exceptional results on niche picks like Maokai, Lillia, and Brand, each with 100% win rates, while his most-played Sejuani achieves 83.3%.
- Xun (main jungle) demonstrates a balanced pool with high win rates on Kindred (83.3%) and Xin Zhao (87.5%), though his performance on Vi (54.5%) is less consistent. Knight (mid lane) showcases versatility, excelling on Corki (92.9%) and Neeko (100%), but his Tristana struggles with a 57.1% win rate.

Wei's (Secondary jungle) limited champion pool suggests fewer games played, with Sejuani (6 games) and Maokai (5 games) leading the count. Wei appears to specialize in tanky jungle champions and niche picks like Zyra and Brand. it may shows BLG want wei play tank champion to protect other teammates.

Overall, most players exhibit strong performance with their preferred champions, though some picks, like K'Sante and Vi, show room for improvement. These insights can guide BLG's drafting strategies and optimize player-champion pairings for better outcomes in competitive play.

2.4.2 Players VisionScore

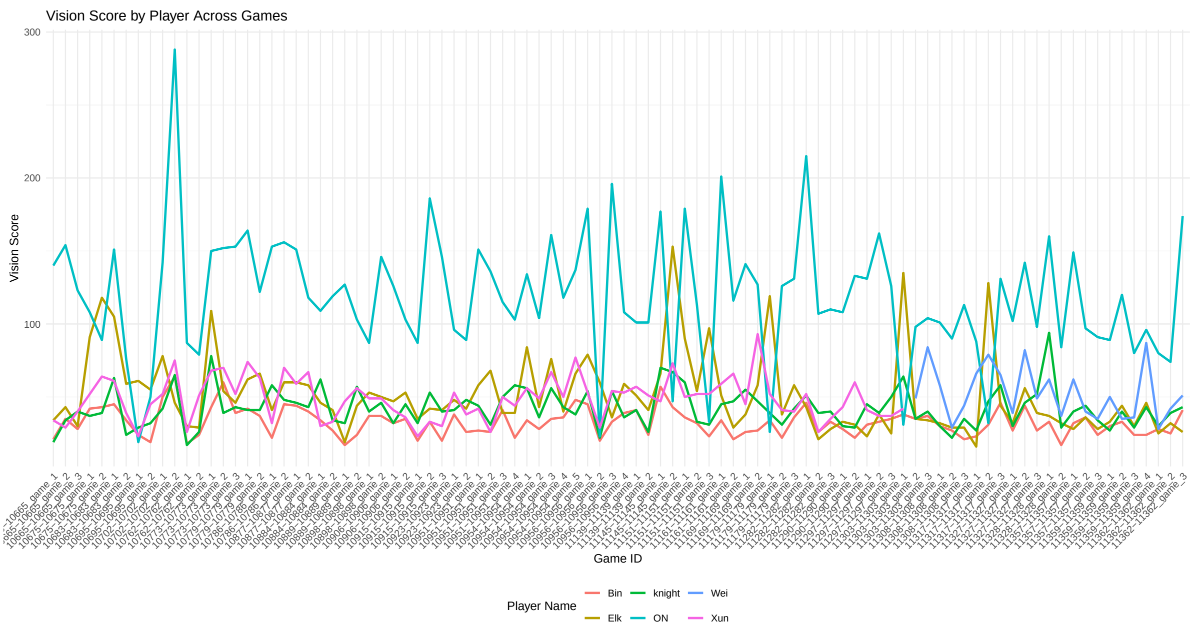


Figure 1: BLG players visionscore over game

The line chart Figure 1 visualizes the vision scores of individual players on the Bilibili Gaming (BLG) team across multiple games, offering insights into their vision control performance and consistency. Each line represents a player's cumulative vision score over the course of several matches, helping to compare and analyze their contributions to map awareness and team strategy.

Key observations highlight that ON, the team's support player, consistently records the highest vision scores across games. This is expected, as supports typically focus on warding and vision control, critical for map dominance. ELK, the ADC, also shows considerable vision score variability. Other roles, such as Bin (top lane), and Knight (mid lane), have lower and steadier vision scores, aligning with their roles' primary focus on laning and damage output rather than vision control.

2.4.3 Game duration

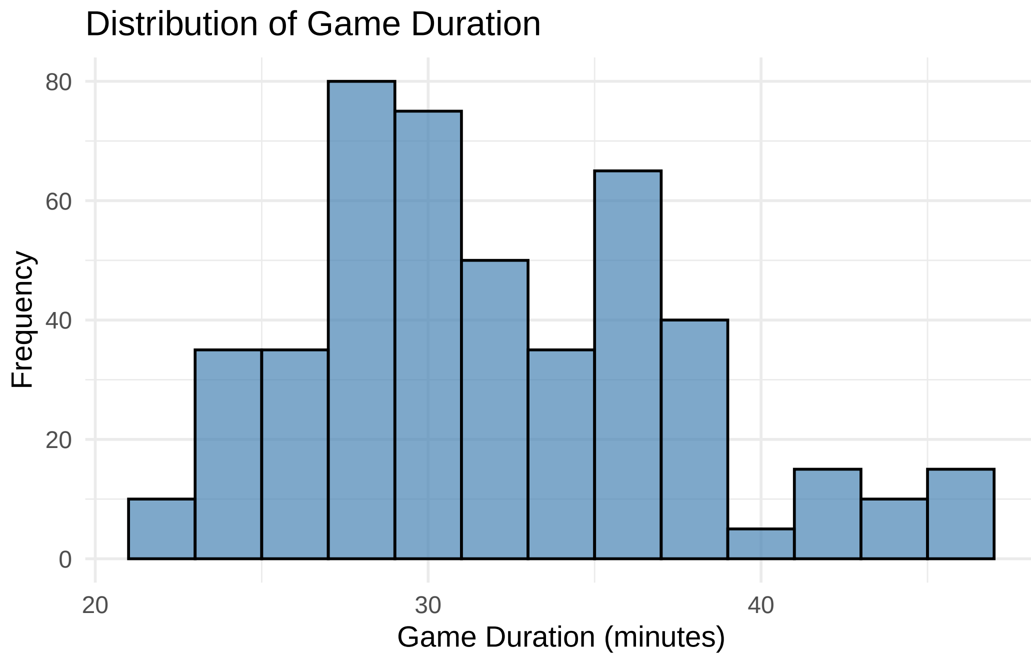


Figure 2: BLG team generally end the game around 30 minutes

The histogram Figure 2 depicts the distribution of game durations for the Bilibili Gaming (BLG) team, measured in minutes. The graph reveals that most games tend to last between 28 to 36 minutes, with a peak frequency around the 30-minute mark. This suggests that BLG typically ends their games within the expected duration for competitive matches.

2.4.4 players gold earned and players damage

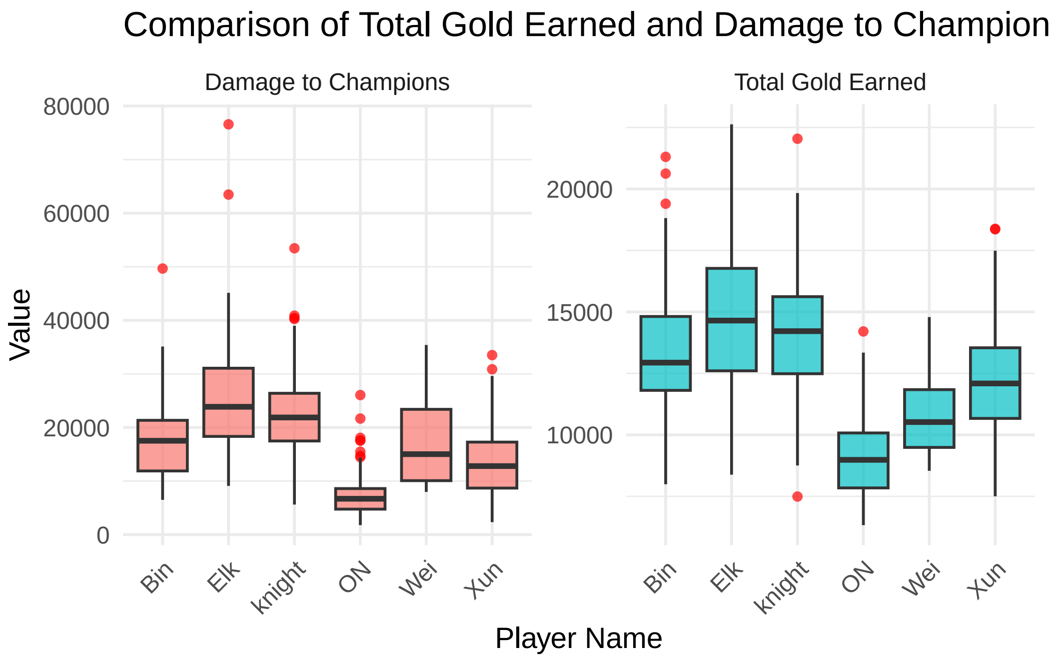


Figure 3: In BLG team, ELK is the highest damage and gold earned player.

The boxplots Figure 3 compare the total gold earned and damage to champions by each player on the Bilibili Gaming (BLG) team, offering insights into their in-game contributions and roles.

Key observations highlight that Elk, the ADC, consistently earns the most gold and deals the highest damage to champions. This reflects the ADC role's reliance on gold to build damage-focused items and their critical role in team fights. However, Wei stands out by earning a comparatively high amount of gold while contributing less damage, suggesting a potential mismatch between resource allocation and damage output. This could indicate a focus on tank or utility-oriented champions or inefficiency in converting gold into impactful performance.

3 Model

The goal of our modelling strategy is twofold. Firstly, we aim to accurately predict the player KDA performance based on relevant game data and key influencing factors. Secondly, we seek to evaluate the efficacy of different modeling approaches—from simple linear regression (SLR) to multiple linear regression (MLR) and Bayesian hierarchical models—to understand their predictive capabilities and assess the underlying relationships between variables. By comparing these models, we can determine which approach provides the most robust and reliable predictions, while considering the variability and potential uncertainty in the data.

Here we briefly describe the Bayesian analysis model used to investigate… Background details and diagnostics are included in Appendix 7.

3.1 Precaution for Building Model

According to the KDA formula in cleaned data section Section 2.3, the response variable KDA is directly depend on three attributes, kills, deaths and assists, so we cannot use kills, deaths and assists as predictors in our model formula Since the response variable is mathematically derived from these three attributes, including kills, deaths, and assists as predictors in the model would create perfect multicollinearity. This means the predictors would have a deterministic relationship with the response variable, violating the assumption of independence between the predictors and the response variable in regression models. For a Bayesian model, this also violates the assumption that the prior information (or likelihood) is not over-specified. In Bayesian inference, incorporating variables that deterministically define the response variable introduces redundancy and inflates the certainty of the posterior distribution, making the model overly confident and prone to poor generalization.

3.2 Basic Simple Linear Regresion Model

3.2.1 Basic Simple Linear Regression Model Set-Up

$$Y = \beta_0 + \beta_1 \cdot X + \epsilon$$

Where:

- Y: The dependent variable (response variable) you are trying to predict.
- X: The independent variable (predictor variable) used to predict Y, in this case, X represent the predictor variable - total CS.
- β_0: The intercept of the regression line, representing the value of Y when $X = 0$.

- β_1: The slope of the regression line, representing the change in Y for a one-unit increase in X.

- ϵ: The error term, accounting for the variability in Y that X does not explain.

3.2.2 SLR Model

In this SLR model, the response variable is KDA and the only one predictor is the total CS. Figure 4a visualizes the relationship between the actual and predicted values of KDA for BLG players, based on a simple linear regression model with total cs as the sole predictor. eaach points represent individual comparisons between actual and predicted values, with the color of the dots representing different players. The red dashed line represents the line of perfect prediction, where actual values would equal predicted values.

Figure 4b examines the relationship between total CS and KDA for BLG players, with each point representing a game and each player distinguished by color. Separate regression lines for each player highlight role-specific trends.

The primary concern lies in the evident dispersion of data points, which are widely spread and do not cluster closely around the line of perfect prediction (the dashed red line). This suggestslimited dependence of KDA on total CS, total CS does not adequately explain the variability in KDA. The observed inconsistencies between actual and predicted values indicate that the relationship between KDA and total CS is not sufficiently captured by a linear model with just one predictor.

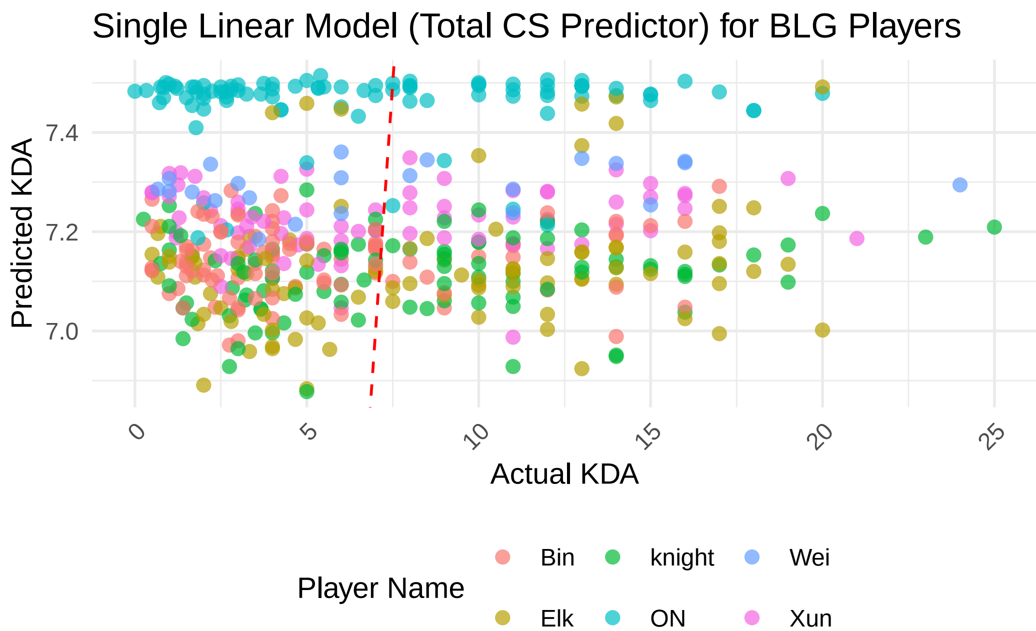
3.2.3 Compute R Squared for SLR Model

$$SS_{\text{total}} = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

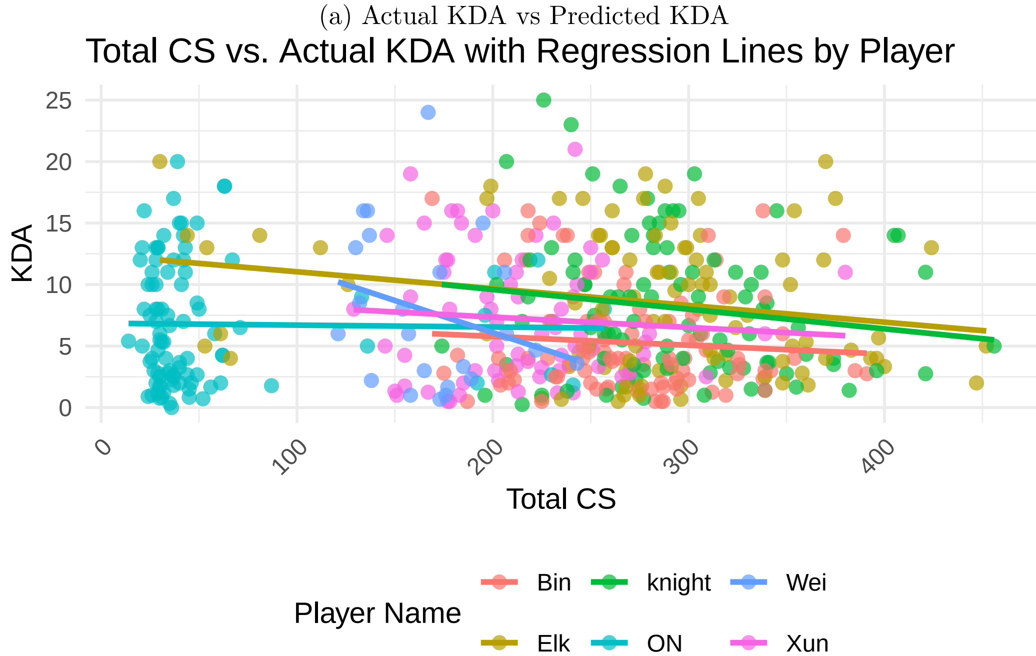
$$SS_{\text{residual}} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$R^2 = 1 - \frac{SS_{\text{residual}}}{SS_{\text{total}}}$$

$$\text{R-squared} = 0.0008659$$



(a) Actual KDA vs Predicted KDA



(b) predictor total CS vs Actual KDA with linear regression line

Figure 4: total cs has limited predictive accuracy for league of legend esport players KDA performance.

3.3 Multiple Linear Regression Model

3.3.1 Multiple Linear Regression Model Set-up

The formula is:

$$Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \beta_3 \cdot X_3 + \beta_4 \cdot X_4 + \beta_5 \cdot X_5 + \beta_6 \cdot X_6 + \epsilon$$

Where:

- Y: The dependent variable (response variable), representing the KDA.
- X_1: Side (categorical variable: Blue or Red), representing the team side during the game.
- X_2: Vision Score, measuring the player's vision contribution during the game.
- X_3: Total Gold, representing the total gold earned by the player.
- X_4: Total CS, representing the total creep score of the player.
- X_5: Damage to Champions, indicating the total damage dealt to enemy champions.
- X_6: Game Duration, representing the length of the game in seconds.
- β_0: The intercept, representing the expected value of Y when all predictors are 0.
- $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$: The coefficients, representing the change in Y for a one-unit increase in the corresponding predictor while holding other predictors constant.
- ϵ: The error term, accounting for the variability in Y not explained by the predictors.

This multiple linear regression model accounts for both categorical and continuous variables to predict the player's **KDA**, capturing the combined influence of gameplay factors such as team side, vision control, resource management, damage output, and game duration.

3.3.2 MLR

The plot Figure 5 visualizes the results of a Multiple Linear Regression (MLR) model predicting KDA using predictors such as side, vision score, total gold, total CS, damage to champions, and game duration. The x-axis represents the actual KDA values, while the y-axis represents the predicted KDA values. Each dot corresponds to an individual game, with colors indicating different players. The red dashed line represents the perfect prediction line (y=x), where predicted values would perfectly match the actual values.

The model demonstrates an improved ability to capture variance in KDA compared to a simple linear regression model, as evidenced by many points clustering closer to the red dashed line.

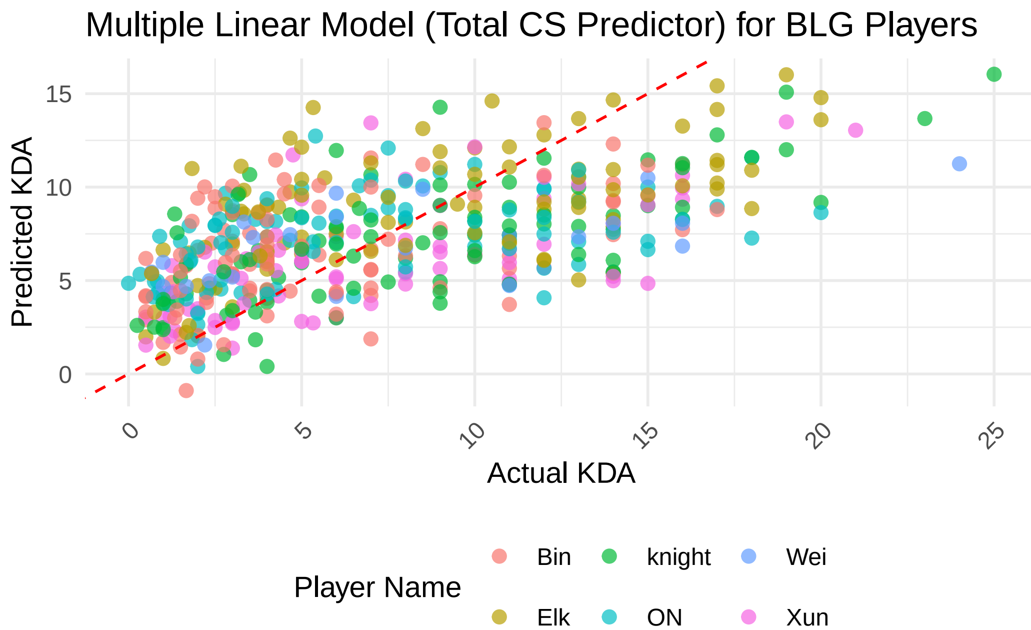


Figure 5: The MLR shows better performance on captured significant variance in players KDA performance.

3.3.3 Compute R Squared for SLR Model

$$SS_{\text{total}} = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$SS_{\text{residual}} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$R^2 = 1 - \frac{SS_{\text{residual}}}{SS_{\text{total}}}$$

$$\text{R-squared} = 0.337$$

We could see that there is a strong improvement on the R-squared values, which mean the MLR model have better performance on illustration for the response variable but still reflects limited predictive accuracy, as only 33.7% of the variance in the KDA is explained by the MLR model.

3.4 Bayesian Model

3.4.1 Bayesian Model Set-UP

$$
\begin{aligned}
\text{KDA}_i &\sim \mathcal{N}(\mu_i, \sigma) \\
\mu_i &= \alpha + \beta_1 \cdot \text{visionscore}_i + \beta_2 \cdot \text{totalgold}_i + \beta_3 \cdot \text{totalcs}_i \\
&\quad + \beta_4 \cdot \text{damagetochampions}_i \\
&\quad + u_{\text{side}[i]} + u_{\text{gamelength}[i]} + u_{\text{position}[i]} \\
\alpha &\sim \text{Normal}(0, 2.5) \\
\beta_j &\sim \text{Normal}(0, 2.5), \quad j = 1, 2, 3, 4 \\
u_{\text{side}} &\sim \text{Normal}(0, \sigma_{\text{side}}) \\
u_{\text{gamelength}} &\sim \text{Normal}(0, \sigma_{\text{gamelength}}) \\
u_{\text{position}} &\sim \text{Normal}(0, \sigma_{\text{position}})
\end{aligned}
$$

This hierarchical Bayesian model is designed to capture both the fixed effects of predictors and the random effects of grouping factors. The response variable is KDA, predictors are visionscore, totalgold, total.cs, damagetochampions, side,gamelength and position. A normal prior with mean 0 and standard deviation 2.5 is set for all coefficients and the intercept, scaled automatically.

3.4.2 Posterior Predictive Checks

Figure 6 The posterior predictive check (PPC) plot compares the observed KDA distribution y(dark blue line) with the posterior predictive distributions y_{rep}(lighter blue lines) generated by the Bayesian model. The alignment between the observed and predictive distributions suggests that the model captures the central tendencies and overall variability of KDA effectively. The peak and general shape of the observed distribution are closely mirrored by the simulated predictions, indicating that the included predictors—vision score, total gold, total CS, damage to champions, position and game length are relevant in explaining the main patterns in KDA.

However, the model shows slight underperformance in capturing extreme values, as evidenced by discrepancies in the tails of the distributions.

In conclusion, the Bayesian model performs well in approximating the central tendencies and variance of KDA, as demonstrated by the close alignment between observed and predictive distributions. However, some refinements may be necessary to improve the model's ability to capture extreme values and better represent the full range of variability in the data.

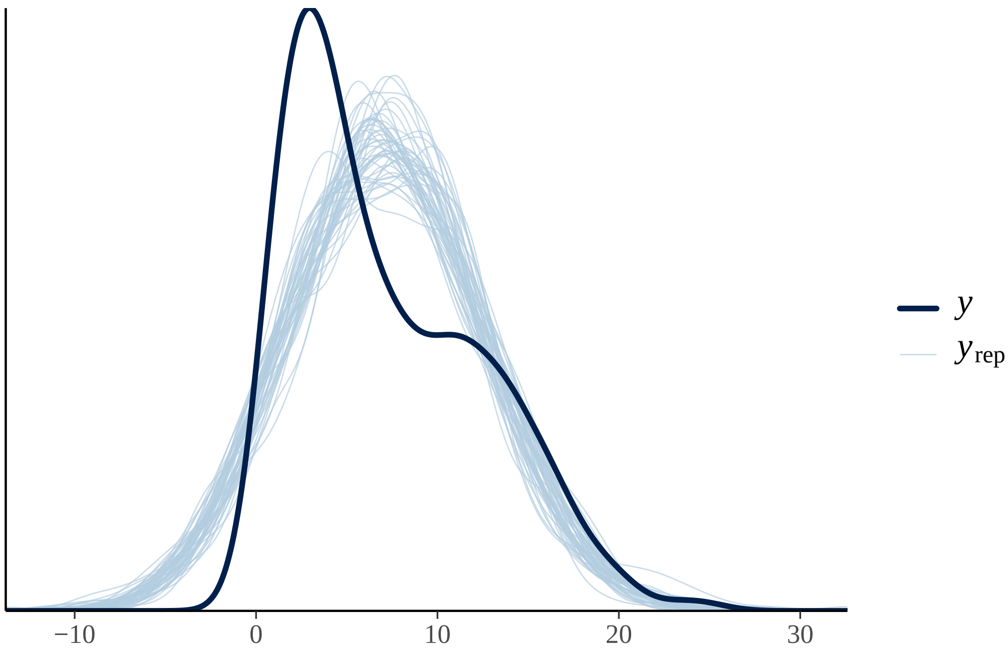


Figure 6: Posterior Predictive Check for the Bayesian Model Predicting pct

3.4.3 Train Test Validation

we implemented a train-test split validation approach to evaluate the performance of a Bayesian hierarchical model predicting KDA for BLG players. The dataset was divided into training (70%) and test (30%) sets to fit and validate the model, respectively. We used the stan_glmer function to fit the model on the training data, with predictors including vision score, total gold, total CS, damage to champions, position and game length. After training, we test the model by generated predictions on the test set and calculated an R squared value to quantify the model's predictive accuracy.

```
print(r_squared)
```

```
[1] 0.6274858
```

$$SS_{\text{total}} = \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{2}$$

$$SS_{\text{residual}} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3}$$

$$R^2 = 1 - \frac{SS_{\text{residual}}}{SS_{\text{total}}} \tag{4}$$

$$\text{R-squared} = 0.6281344 \tag{5}$$

Where:

- y_i represents the actual values,
- $\bar{y}$ is the mean of the actual values,
- $\hat{y}_i$ are the predicted values,
- SS_{total} is the total sum of squares,
- SS_{residual} is the sum of squared residuals.

An (R^2) value of 0.6281 indicates that the Bayesian model explains approximately 62.8% of the variability in the test dataset's KDA values. This suggests that the predictors used in the model—such as vision score, total gold, total CS, damage to champions, and the random effects—capture a substantial portion of the variation in KDA. The (R^2) value show the bayesian model has strong improvement compare to previous SLR and MLR model.

By visual checking actual pct and predicted pct plot in Figure 7, the points are plotted against a 45-degree line, which represents the ideal scenario where predicted values match the actual values perfectly. As there are a single indentifiable pattern and the dots are random scatter around the line. It provide a strong evidence that the model is valid.

In addition, the residual plot Figure 8 is fairly centered around zero with no major trend, suggesting that the model is not heavily biased in its predictions.

In conclusion, from the visual checks from the predicted vs. actual plot and the residual plot, there is no robust evidence that the Bayesian model is overfitting. It appears to generalize well to the data it was trained on without showing signs of capturing noise or irrelevant patterns.

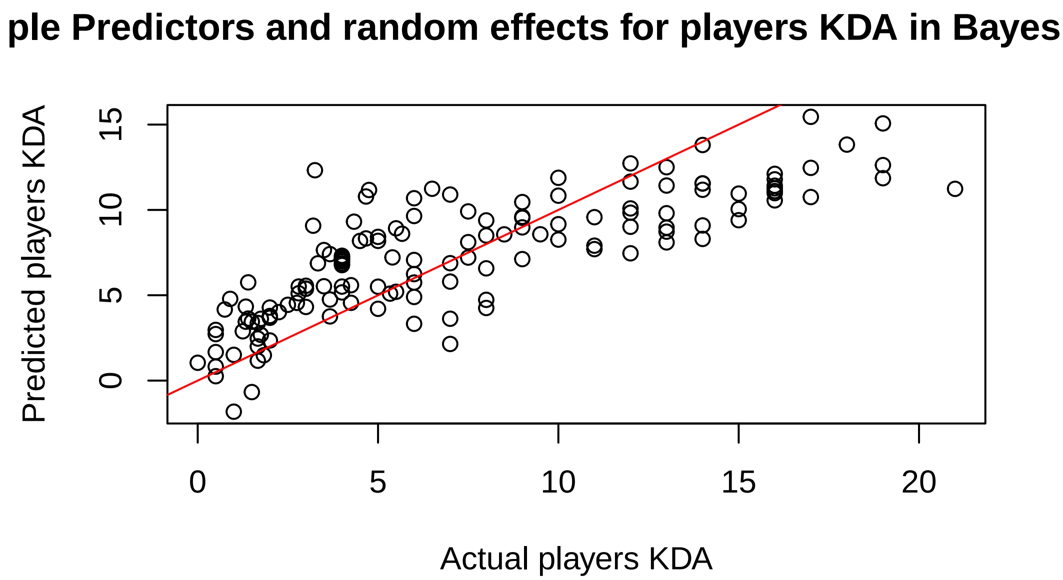


Figure 7: The Bayesian Model Accurately Predicts the Testing Dataset

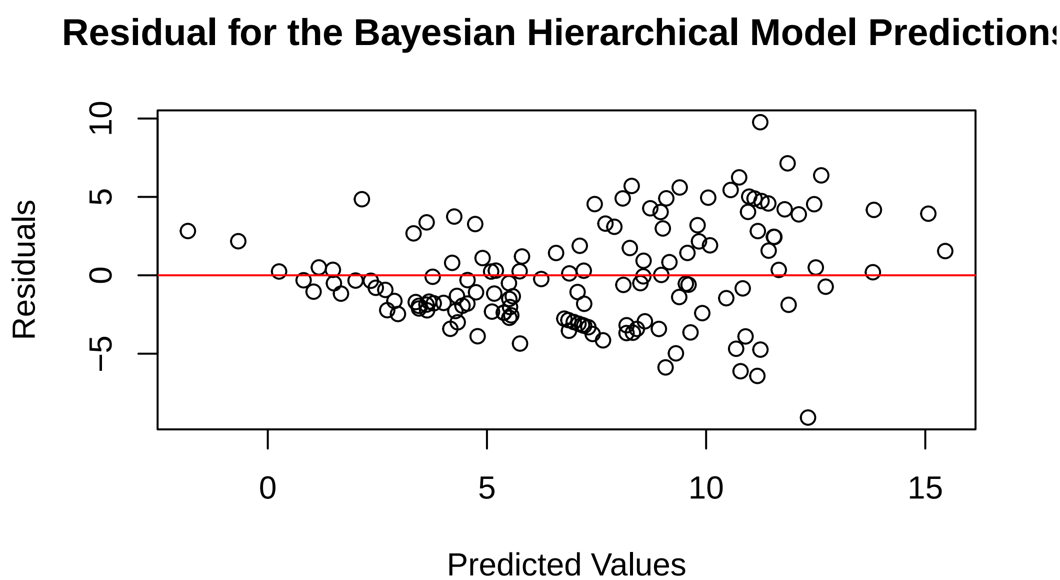


Figure 8: Bayesian Model Residuals Appear to Show No Trends

3.5 Model justification

We chose vision score, total gold, total CS, damage to champions, position and game length as key in-game metrics and contextual variability to predict KDA. The visualizations Figure 1 highlight significant relationships between predictors and performance. For instance, vision score, which varies widely among players, particularly supports like ON, reflects role-specific contributions to team performance, justifying its inclusion as a fixed effect. Similarly, metrics such as total gold and damage to champions, shown to vary by role (e.g., Elk consistently excelling as an ADC) in Figure 3, emphasize the importance of resource generation and damage output in predicting KDA. Including these predictors ensures that the model captures individual player contributions effectively.

Random effects for team side, position, and game length further enhance the model by accounting for contextual variability. The histogram Figure 2 of game durations shows that while most games last 30-35 minutes, variability exists, necessitating its inclusion to adjust for its impact on player performance. The hierarchical structure also captures differences in roles and sides, ensuring the model adapts to diverse scenarios. Overall, the chosen model balances fixed effects for key performance metrics with random effects for contextual factors, ensuring robust and accurate KDA predictions.

4 Result

This plot Figure 9 illustrates the posterior credible intervals for the predicted KDA of each player on the BLG team, based on the Bayesian mixed-effects model. The mean predicted KDA for each player is represented by the points, while the vertical lines show the 95% credible intervals. Each player is assigned a distinct color, making it easier to distinguish between their predictions and associated uncertainties.

The credible intervals vary in width across players, reflecting differences in the model's certainty about their performance. Players like Elk and Knight, who play key roles such as ADC and Mid Lane, have slightly higher predicted KDA values with credible intervals that extend across a moderate range, indicating variability in their potential performance. On the other hand, roles like Support (ON) and Jungle (Xun) also show notable ranges in their credible intervals, reflecting the dynamic nature of their contributions during games.

One thing worth mentioning is that Bin, the superstar Top Lane player of the BLG team, shows a disappointing KDA prediction within his 95% credible interval. Despite his reputation and the high expectations from media and fans, the model's prediction suggests that Bin's actual performance may not align with the hype. This could indicate that Bin is overrated by his supporters and might be more dependent on his teammates than previously thought. The data implies that Bin is often carried by his team, highlighting a significant gap between his perceived status and actual contributions. This reinforces the importance of using objective

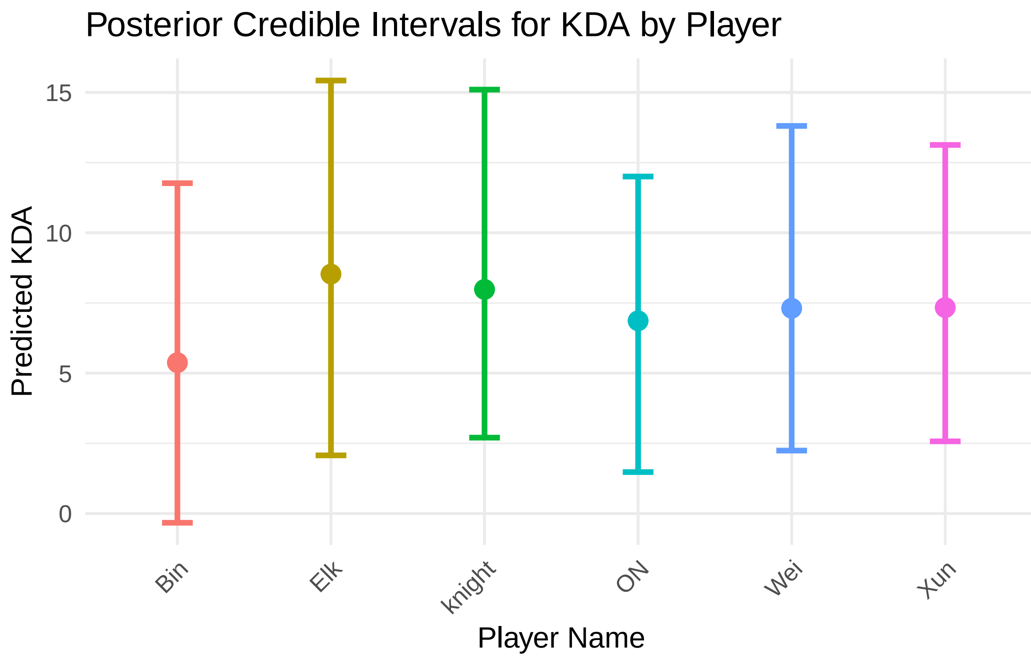


Figure 9: Bin is the superstar of BLG team, but show disappointed KDA prediction in 95% credible interval

metrics to assess individual player impact rather than relying solely on public perception or narrative.

Table 3: Bayesian Model Result Summary

Table 3: Summary of Actual vs Predicted KDA with 95% Credible Intervals

Player Name	Actual KDA	Predicted KDA	Lower Bound Credible Interval	Upper Bound Credible Interval
Bin	7.00	5.71	-2.12	13.74
Xun	16.00	11.39	3.50	19.36
knight	10.00	11.94	4.00	19.90
Elk	4.50	8.50	0.66	16.31
ON	13.00	9.34	1.53	17.28
Bin	1.00	-1.83	-9.68	5.81
Xun	1.00	3.23	-4.61	11.09
knight	1.67	3.40	-4.38	10.98
Elk	0.50	0.68	-7.17	8.36
ON	0.33	1.88	-5.87	9.96

Table 3 provides a summary of the actual KDA, predicted KDA, and the 95% credible intervals for the predictions of BLG players. The results reveal varying levels of alignment between the actual and predicted KDA.

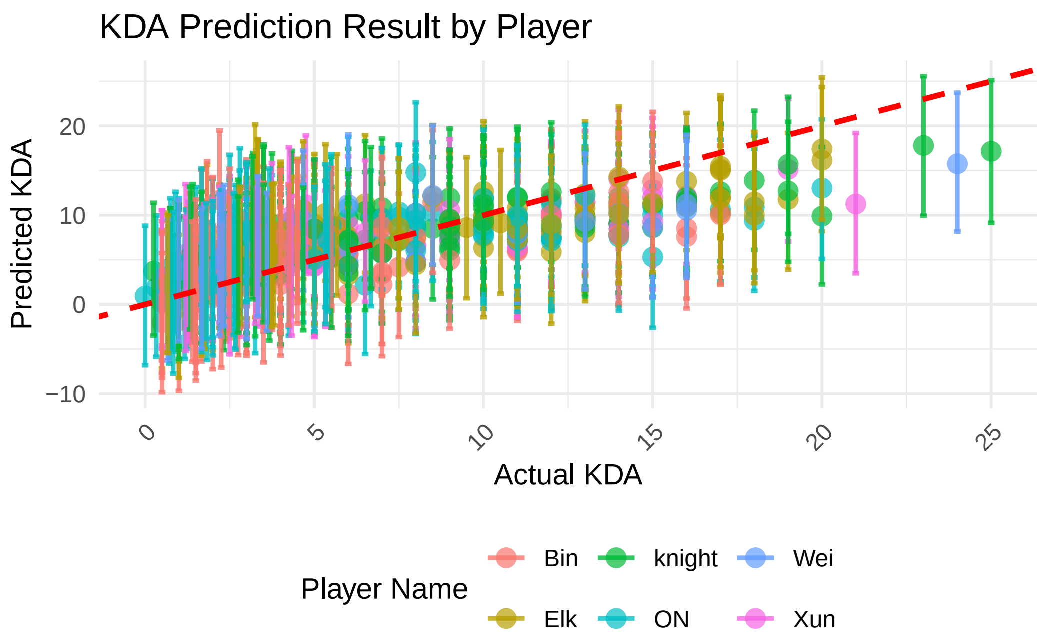


Figure 10: The Model Shows a Moderate-Well Alignment Between Actual and Predicted KDA

Figure 10 This plot visualizes the predicted KDA values and their associated 95% credible intervals for each BLG player, compared against their actual KDA. The points represent the mean predicted KDA, color-coded by player name, while the vertical error bars show the uncertainty in the predictions (credible intervals). The red dashed line represents perfect prediction, where the predicted KDA equals the actual KDA. Most predictions cluster around the red line, indicating that the Bayesian model aligns reasonably well with the actual KDA values. However, there are notable deviations, especially in cases of high actual KDA values. These deviations are likely due to the scarcity of data for games with extremely high KDA, leading to greater uncertainty in the model's predictions for these cases. This emphasizes the importance of accounting for data imbalance in future modeling efforts.

5 Discussion

5.1 Ethics, Sportsmanship, and the Broader Impact of League of Legends Esports

League of Legends (LoL) esports is more than a competitive gaming platform; it serves as a cultural and ethical touchpoint that influences players, fans, and the broader entertainment industry. The competitive environment of LoL is not only about winning games but also about fostering sportsmanship and collaboration. Professional players are expected to uphold ethical standards, demonstrating respect for opponents, referees, and fans, regardless of the game's outcome. This culture of respect and fair play is crucial, as it sets a positive example for the millions of young fans who view these players as role models. Encouraging good sportsmanship in esports reinforces values such as teamwork, discipline, and perseverance, which are important in both gaming and broader life contexts.

Moreover, League of Legends esports contributes significantly to the global entertainment industry, combining intense competition with narrative-driven content to create a highly engaging spectator experience. Events like the World Championship attract millions of viewers, showcasing esports as a legitimate entertainment medium on par with traditional sports. This global reach provides opportunities to inspire young generations by emphasizing the importance of never giving up, learning from failures, and working towards personal and collective goals. By integrating ethical discussions, promoting mental resilience, and fostering an inclusive community, the League of Legends esports scene not only entertains but also educates and empowers young people to strive for excellence in gaming and beyond.

5.2 Summary of Findings

This study develops a predictive model to assess player performance for Bilibili Gaming (BLG) in the League of Legends Pro League (LPL). The findings reveal that metrics such as vision score, total gold, total CS, damage to champions, and game side significantly contribute to predicting (KDA). The Bayesian mixed-effects model achieved an (R^2) of 0.63, demonstrating strong predictive accuracy and capturing substantial variability in player performance.

The analysis highlights that players like Elk and Knight consistently show high predicted (KDA) values with narrower credible intervals, indicating strong and stable contributions. Conversely, Bin, despite having a significantly higher contract than Elk, demonstrates lower predicted performance with wide credible intervals, suggesting inconsistencies and greater reliance on teammates. This disparity between investment and contribution indicates potential inefficiencies in resource allocation. It is recommended that BLG's management reassess contract structures and performance metrics to ensure alignment between player compensation and impact. Additionally, It is recommended that BLG's management lower Bin's in-game priority and focus team resources on higher-impact players like Elk and Knight. This approach could help mitigate Bin's negative impact on team performance and ensure the team's overall strategy remains efficient and competitive.

5.3 Sources of Bias in the Polls

Several sources of bias may affect the findings and predictions of this study Based on the figures in Section 2. First, the dataset relies heavily on historical in-game metrics from the League of Legends Pro League (LPL). These metrics may not fully capture all dimensions of player performance, such as intangible contributions like shot-calling, morale boosting, or adapting to unforeseen situations during matches. Consequently, the model may undervalue players who excel in these non-quantifiable aspects of gameplay.

Second, the inclusion of only certain predictors, such as vision score, total CS, and damage to champions, introduces potential bias by excluding other relevant variables, such as synergy with teammates, champion pool diversity, or adaptability to meta changes. This limited scope could lead to an overemphasis on specific roles or playstyles, particularly those that align well with the chosen predictors.

Third, the dataset may reflect team-specific dynamics or strategies that do not generalize well across all competitive scenarios. For example, players like Bin may appear to underperform relative to metrics like KDA but could be fulfilling strategic roles not reflected in the data. Additionally, external factors, such as the influence of the blue-side advantage or variability in opponent strength, may introduce bias if not adequately controlled for.

Finally, the predictive model assumes that past performance reliably predicts future outcomes, which may not always hold true in a rapidly evolving esports environment. Changes in game patches, meta shifts, or individual player development can significantly alter performance, potentially leading to inaccuracies in predictions. Acknowledging and addressing these biases is essential for refining the model and improving its applicability in real-world decision-making for esports teams.

5.4 Limitations of the Model

The Bayesian mixed-effects model, while effective, has several inherent limitations. First, it relies on a predefined set of in-game metrics, such as vision score, total CS, total gold, damage to champions, and game length. These metrics, while significant, do not fully account for intangible factors like leadership, communication, or decision-making, which are critical components of a player's performance. As a result, the model may undervalue players whose contributions extend beyond quantifiable metrics.

Second, the model assumes a linear relationship between predictors and the response variable (KDA), which may oversimplify the complex interactions in competitive gameplay. Non-linear dynamics, such as synergistic interactions between players or champion-specific performance nuances, are not explicitly captured, potentially limiting the model's ability to reflect real-world complexities.

Lastly, the model is trained on data from a single team, BLG, in the LPL, which could restrict its generalizability. Regional differences in playstyle, team dynamics, or meta-strategies may lead to variations in predictor-response relationships that the model is not designed to address. As a result, its applicability outside this specific context is limited.

5.5 Limitation of the Prediction

In@sec-result, the predictive accuracy of the model is subject to certain constraints that impact its reliability. First, predictions are heavily influenced by historical data and may fail to account for future changes, such as meta shifts, patch updates, or evolving team strategies. Players like Bin, whose performance appears inconsistent in the current data, might improve under different conditions, but such adjustments are beyond the scope of the model's predictions.

Second, predictions are affected by biases and variability in the data. Factors like opponent strength, game-side advantages, and external influences (e.g., player fatigue or morale) are not fully controlled for, which can introduce noise and uncertainty. This is reflected in the broad credible intervals for some players, particularly those with inconsistent performances, limiting the precision of the forecasts.

Finally, the predictions focus on quantifiable metrics like KDA, which may not fully encapsulate a player's overall contribution to team success. For example, players with lower KDA but critical strategic roles, such as shot-calling, may appear undervalued in the predictions. Future iterations could incorporate additional predictors, such as champion pool diversity or team synergy, to improve the model's predictive capacity and better capture the multifaceted nature of player performance.

5.6 Future Directions

Future research could expand on this study by incorporating additional predictors to better capture the multifaceted nature of player performance in esports. Metrics such as champion pool diversity, synergy with teammates, and adaptability to meta shifts could provide a more comprehensive understanding of individual contributions. Exploring non-linear relationships and interaction effects among in-game metrics may also enhance the model's ability to reflect the complexities of competitive gameplay. Additionally, validating the model across different teams, regions, and competitive levels could improve its generalizability and robustness. Integrating real-time data or patch-specific adjustments would further refine predictions, making them more adaptable to the dynamic nature of esports. Lastly, examining the broader implications of performance predictions, such as their influence on player development, team strategies, and fan engagement, could provide valuable insights for both academics and practitioners in the esports industry.

Appendix

6 Additional data details

7 Model details

7.1 Posterior predictive check

In **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows…

In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows…

7.2 Diagnostics

?@fig-stanareyouokay-1 is a trace plot. It shows… This suggests…

?@fig-stanareyouokay-2 is a Rhat plot. It shows… This suggests…

References

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

NumPy Developers. 2023. *NumPy: The Fundamental Package for Scientific Computing with Python*. NumPy Developers. https://numpy.org/.

oracleselixir. 2024. "Oracle's Elixir - LoL Esports Stats." https://oracleselixir.com/tools/downloads.

pandas Development Team. 2023. *pandas: Powerful Python Data Analysis Toolkit*. pandas Development Team. https://pandas.pydata.org/.

Python Software Foundation. 2023. *Python: A Dynamic, Open Source Programming Language*. Python Software Foundation. https://www.python.org/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

Stan Development Team. 2024. "RStan: The R Interface to Stan." https://mc-stan.org/.

Tan, Amanda "Tania Mae". 2020. "Aggression or Control: A Deep Dive into the Huge Difference Between the LPL and LCK | ONE Esports." https://www.oneesports.gg/league-of-legends/aggression-or-control-a-deep-dive-into-the-huge-difference-between-the-lpl-and-lck/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.