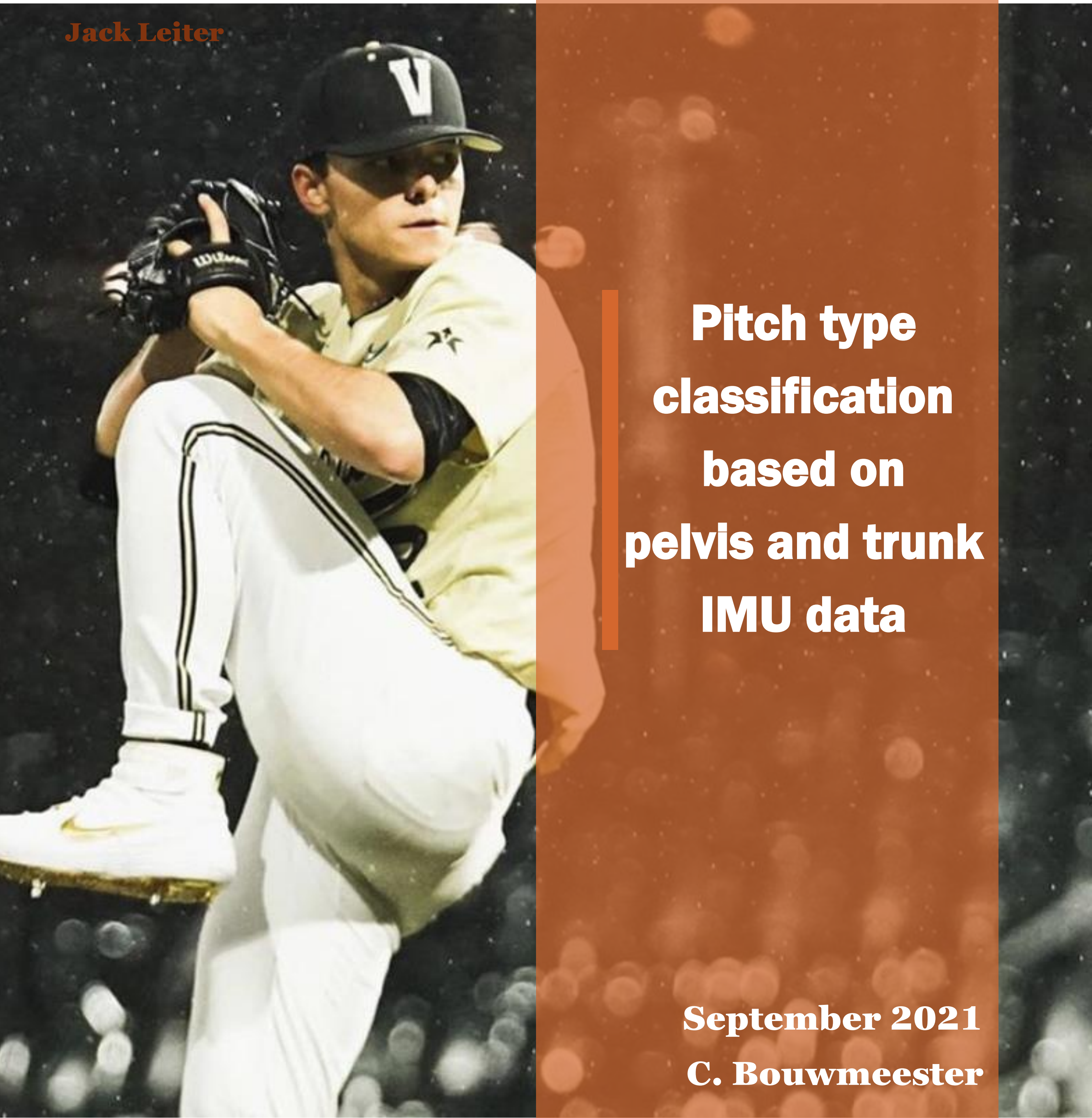


Pitch type classification based on pelvis and trunk IMU data

September 2021

C. Bouwmeester

PITCH PERFECT

TU Delft Delft University of Technology

Page left blank intentionally

Pitch type classification based on pelvis and trunk IMU data.

By

C. Bouwmeester, MSc

in partial fulfilment of the requirements for the degree of

Master of Science

in Biomedical Engineering

at the Delft University of Technology

Thesis committee:	**Prof. dr. H.E.J Veeger,**	TU Delft, Biomechanical Engineering
	L. Gomaz, PhD Candidate	TU Delft, Applied Mathematics
	Dr. E. van der Graaff	PITCHPERFECT
	B. van Trigt, PhD Candidate	TU Delft, Biomechanical Engineering
	Dr. ir. E. van der Kruk	TU Delft, Biomechatronics & Human-Machine control

Page left blank intentionally

Preface

I believe we can all agree that this was an exceptional year. Not only was working from home the new standard, there were also a lot of limitations regarding measurements. Because of Covid restrictions, the start of the data collection was later than planned. Consequently, the number of participants was lower than planned in the time available for this thesis. Although this period had all the signs of being a year where nothing could be done, I am extremely proud to present this thesis.

This thesis is the result of two passions of mine, namely data and sports. With a background in Human Movement Science and Sport Science, I think this thesis contains the best of both worlds. As part of my internship at PitchPerfect, a company that has developed sensors for biomechanical feedback in baseball, I had the opportunity to be on the field and have worked with various professional pitchers. My gratitude goes to the pitchers and coaches from Twins Oosterhout, where I have learned a lot about pitching and could test the sensors during their training sessions.

During my project, many people have helped me. First of all, I would like to thank PitchPerfect and in particular my external supervisor Dr. Erik van der Graaff. He taught me about (pitching) biomechanics, writing and what it takes to be an entrepreneur. I would also like to thank my other supervisors Larisa Gomaz, Bart van Trigt and Prof. Dr. DirkJan Veeger. Their academic experience have guided me through this project and have given me new insights. We have had some really interesting discussions. Lastly, I would like to thank my friends and family for being supportive and even assisting me with some measurements.

Celine Bouwmeester, MSc

Gouda, September 2021

Contents

Pitch type classification based on pelvis and trunk IMU data

C. Bouwmeester, L. Gomaz MSc, B. van Trigt MSc, Dr. E. van der Graaff, Prof. dr. H.E.J. Veeger

Abstract—**Classification of pitch types outside the laboratory or game environment provide benefits in designing and outlining training routines to reduce injury and improve performance. Current classification approaches are based on vision-based optical data, however these optical data is often not available in training sessions. The aim of the current study is to use machine learning algorithms to classify pitch types, i.e., fastball, change-up and curveball, based on pelvis and trunk IMU data in training sessions. A total of 406 successful pitches thrown by 19 pitchers were used to classify pitch types. Three conditions were tested and evaluated in a binary (fastball versus others) and multi-class (fastball versus change-up versus curveball) classification approach. Condition A included the direct output of the PitchPerfect software, whereas condition B included features normalized by fastball characteristics. Condition C is a combination of both condition A and B. The random forest algorithm demonstrated the best predictions in both the binary and multi-class classification approach based on the highest accuracy and F1 scores, i.e., the harmonic mean between sensitivity and precision. Therefore, the random forest algorithm is the best fit for classifying pitch types based on pelvis and trunk IMU data in training sessions. In addition, the performance of the classification algorithm improved when using a binary classification approach. There were no relevant improvements when using additional features. The random forest algorithm can directly be implemented in the PitchPerfect application. Pitchers can use the pitch type data to track and tune their performances, whereas coaches can use the data to design match and training routines. Future research should focus on larger datasets, i.e., more pitchers and pitches, to pre-classify pitchers with similar pitching characteristics to improve the classification algorithms.**

Keywords: Supervised learning, Sports, Biomechanics, Off-speed pitches

NOMENCLATURE

μ_{FB}	Mean of the first five fastballs [deg/s or ms]
FN	False Negatives [-]
FP	False Positives [-]
$Pelvis$	Pelvis angular velocity [deg/s]
$Prop$	Ratio [-]
Sep	Separation time [ms]
TN	True Negatives [-]
TP	True Positives [-]
$Trunk$	Trunk angular velocity [deg/s]

I. INTRODUCTION

Technological advances have been commonly used in baseball. More pitchers and coaches use objective information to improve their training and performance over the last decade. Baseball pitchers in professional leagues have the opportunity to acquire in-game data with Pitchf/x (SportsVision Inc.,Chicago, IL, USA). This information is, however, not available in training environments.

One element of information is pitch types. The most frequent pitch types are the fastball, change-up and curveball [1], [2]. The fastball is the fastest and straightest pitch. The change-up is thrown with the same motion as a fastball but with less ball velocity. The curveball is designed to deceive the batter with ball spin and movement away from the pitcher's arm side. Major League Baseball provides an extensive explanation of each pitch type [3]. A complicating factor in acquiring pitch type data, i.e., knowing the pitch type that is thrown, is that training sessions are often not recorded on video. Pitch type data outside the laboratory or game environment provide benefits in designing and outlining training routines to reduce injury and improve performance. Another benefit is that counting errors can be prevented as pitch type data can be used to automatically count and register the thrown pitches in training sessions.

The use of Inertial Measurements Units (IMUs) in baseball is rising as it has the potential to be used outside the laboratory [4]–[11]. IMUs often consist of a three-axis accelerometer and gyroscope that measure linear accelerations and angular velocities, respectively. Technological innovations such as PitchPerfect (PitchPerfect, The Netherlands) use synchronized IMUs to measure pelvis and trunk angular velocities, and timing between these peaks in training sessions. These segments are measured because the trunk is the main force generator of the kinetic chain due to the large segment mass [12], [13]. Moreover, proper timing between rotation of the pelvis and trunk is seen as a critical component to maximize the contribution of each link of the kinetic chain [14], [15].

IMUs could be used not only to measure mechanics, but also to determine pitch types. Kinematic parameters, such as the magnitudes of the pelvis and trunk angular velocity, significantly differ among pitch types [16]–[20]. The magnitudes of the pelvis and trunk angular velocity are the highest when throwing a fastball [16]–[20] and the lowest when throwing a change-up [16], [17], [19]. However, there are some inconsistencies when comparing the change-up and curveball as some studies did not find a significant difference [17], [18], [20]. This suggests that the fastball is the most distinctive pitch type and that a binary classification approach (fastball versus other pitch types) will likely perform better than a multi-class classification approach (fastball versus change-up versus curveball).

The studies described above indicate that there are differences in pelvis and trunk angular velocities among pitch types. However, there is still a gap regarding how these kinematic parameters translate in practice when using IMUs in training sessions. Only one study has classified pitch types based on IMU data with a machine learning approach [11]. Although this study gives a first impression of classifying pitch types based on IMU data, the study used six IMUs located on the upper extremities and chest. Also, the study was based on a small dataset. Other studies have focused on classifying pitch types based on video-based optical data and ball behaviour [21], [22]. The accuracy of the model improved when classifying pitchers with similar ball characteristics of the fastball, i.e., relative speed, spin rate, vertical and horizontal break distance, prior to the classification of the pitch types [22]. This suggests that taking fastball characteristics into account might improve the accuracy of the used machine learning algorithms.

The aim of the current study is to use machine learning algorithms to classify pitch types, i.e., fastball, change-up and curveball, based on pelvis and trunk IMU data in training sessions. It was hypothesized that adding more features that take fastball characteristics into account and by using a binary classification approach results in a better performance of the machine learning algorithm.

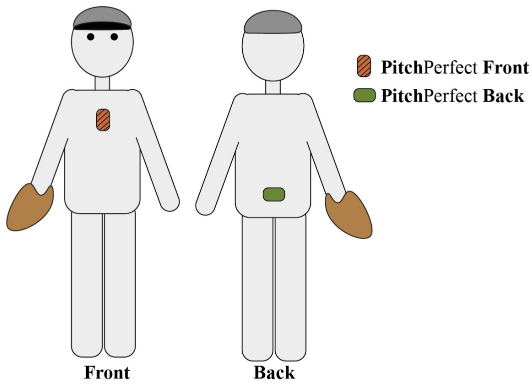


Fig. 1: Placement of the sensors.

II. Methods

A. Participants

Data were collected from 24 baseball pitchers. To be included in the current study, pitchers had to be pain- and injury-free at the moment of measurements. Pitchers throwing higher than the thresholds for pelvis and trunk angular velocities set by PitchPerfect were included. A total of 19 pitchers (mean $\pm$ SD; age: 19.1 ± 3.89 years, $experience$: 8.1 ± 3.95 years, $height$: 181.3 ± 11.35 cm, $weight$: 75.2 ± 17.8 kg, $throwing$ $velocity$: 98.9 ± 9.99 mph) fit these criteria and were included in the current study. The study protocol followed the guidelines stated in the Declaration of Helsinki and was approved by the Ethics Committee of the Technical University of Delft, Netherlands (C201109130). Informed consent was signed by the participants or leader of the team.

B. Procedures

The measurements were performed during the pitchers' regular training at their own training facility. After performing a regular warming-up, pitchers were instructed to throw pitch types that they are familiar to throw in a game. The pitchers followed their own training routine in accordance with their coach with a minimum of 20 pitches. Pitchers threw from a mound in the bullpen towards a catcher at the official distance of 18.45 and 16.45 meter depending on their age. Sensors were taped with Leukoplast FixoMull® stretch (BSN Medical GmbH, Hamburg, Germany) on the processus Xiphoideus and in the middle of the left and right posterior superior iliac spine before starting the bullpen (Figure 1). As a thank you, the pitchers received a summary of their mechanics during the bullpen session (Appendix I).

C. Data collection

Peak pelvis and trunk angular velocity [in deg/s], and separation time between both peaks [in ms] were registered with the PitchPerfect sensors (PitchPerfect, The Netherlands). The PitchPerfect sensors consist of a digital 3-axis gyroscope measuring up to 2000 deg/s. The PitchPerfect software automatically calculates peak pelvis angular velocity ($Pelvis$), peak trunk angular velocity ($Trunk$) and separation time (Sep) from the two IMUs. These parameters, also known as features, were extracted from the database of PitchPerfect. Pitch types were collected based on hand signal and agreement of the pitch type between pitcher and catcher prior to the throw. The ball velocity [in mph] was measured behind the pitcher with a Pocket radar Ball coach, Model PR1000-BC (Pocket Radar Inc., Santa Rosa, CA, USA) to check if ball velocity significantly differs among pitch types, i.e., fastball is thrown with highest ball velocity whereas curveball is thrown with lowest ball velocity [7], [9], [17], [19], [23].

D. Data processing

Data were pre-processed and analysed with Rstudio (version 1.3.1093, Rstudio, PBC). As PitchPerfect has an integrated, no additional filters were used to pre-process the data. Pitches were included based on three inclusion criteria. First, the pitch type should be a fastball, change-up or curveball. Second, the ball should be able to be caught by the catcher. This means that the ball could be a outside the striking zone, but the catcher needs to be able to catch the ball. Third, Sep should not be higher than 500 ms or lower than -500 ms. These limits were based on the average rotation velocity profile where most rotation velocity of the pelvis and thorax occurred between 300 and 800 ms [15]. The package $Caret$ was used to create the predictive models after pre-processing the data [24]. The data were then used to train the supervised machine learning algorithms (See II.F Classification approaches) and classify the pitch types. The machine learning algorithm with the best performance was used for comparison (See II.G Classification methods).

E. Feature extraction

Besides the three features directly extracted from the PitchPerfect database, i.e., $Pelvis$, $Trunk$ and Sep, four other features were calculated in order to investigate whether the model improved when normalizing the features by fastball characteristics. First, the ratio of pelvis and trunk angular velocity, $Prop_{Pelvis-Trunk}$, was calculated to investigate whether there was a relationship between trunk and pelvis angular velocities and pitch types (Eq. 1). Second, FB_{Pelvis}, FB_{Trunk} and FB_{Sep} were calculated (Eq. 2-4, respectively). These features were normalized by the mean of first five fastballs (μ_{FB}).

$$Prop_{TrunkPelvis} = \frac{Trunk}{Pelvis} \tag{1}$$

$$FB_{Pelvis} = \frac{Pelvis}{\mu_{FBPelvis}} * 100\% \tag{2}$$

$$FB_{Trunk} = \frac{Trunk}{\mu_{FBTrunk}} * 100\% \tag{3}$$

$$FB_{Sep} = \frac{Sep}{\mu_{FBSep}} * 100\% \tag{4}$$

All continuous features were scaled and centred [24]. An overview of the used features is shown in Table I.

TABLE I: Included features for pitch type classification.

Features	Definitions and measurements	Condition
$Pelvis$ (deg/s)	Peak angular velocity of the pelvis. Directly given in PitchPerfect App	A, C
$Trunk$ (deg/s)	Peak angular velocity of the trunk. Directly given in PitchPerfect App.	A, C
Sep (ms)	Separation time between both peaks. Directly given in PitchPerfect App.	A, C
$Prop_{Pelvis-Trunk}$	Ratio between trunk and pelvis. Calculated with $Trunk$ and $Pelvis$	B, C
FB_{Pelvis}	$Pelvis$ normalized with the mean of the first five fastballs.	B, C
FB_{Trunk}	$Trunk$ normalized with the mean of the first five fastballs.	B, C
FB_{Sep}	Sep normalized with the mean of the first five fastballs.	B, C

F. Classification approaches

A binary and multi-class classification approach was used to establish if an algorithm can classify pitch types based on pelvis and trunk IMU data. The binary classification approach classified the pitches in either a fastball or others, whereas the multi-class classification classified the pitches in a fastball, change-up or curveball.

The binary and multi-class classification approach both evaluated three conditions. The first condition, condition A, investigated whether features directly extracted from the PitchPerfect database were sufficient to classify pitch types. Condition A included $Pelvis$, $Trunk$ and Sep. Condition B investigated whether the normalized features were sufficient to classify pitch types and included $Prop_{Pelvis-Trunk}$, FB_{Pelvis}, FB_{Trunk} and FB_{Sep}. The last condition, condition C, investigated whether the models improved when taking all the features into account. An overview of the classification approaches is shown in Figure 2.

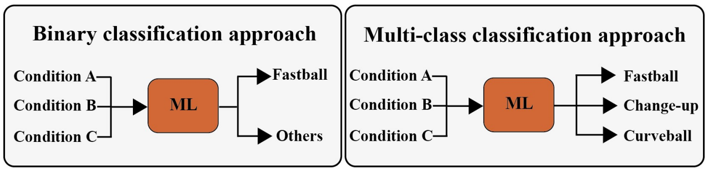


Fig. 2: The evaluated approaches. Left. Binary classification approach. Right. Multi-class classification approach. ML = Machine learning algorithms, Condition A includes $Pelvis$, $Trunk$ and Sep. Condition B includes $Prop_{Pelvis-Trunk}$, FB_{Pelvis}, FB_{Trunk} and FB_{Sep}. Condition C includes both condition A and B.

G. Classification methods

The current study utilized several widely used classifiers to classify pitch types. The most widely used classifiers include, but not limited to, Discriminant Analysis (DA), K-Nearest Neighbours (K-NN), Decision Tree (DT), Random Forest (RF) and Support Vector Machine (SVM) [25]. DA divides the space into the number of classes and assigns the variable to a class if the variable is in the region of the class [26]. K-NN assumes that similar things are close together. It calculates the distances between K points and the variable and assigns the variable to the class with the most frequent label [27]. DT creates a training model that can be used to predict the class of the variable by learning simple decision rules from training data [28]. RF builds decision trees on random samples and assigns the variable to the most frequent label [29]. SVM tries to find an optimal hyperplane that separates the classes with a minimum margin [30]. Included algorithms are explained in more detail in Appendix II. The models were generated using repeated K-fold cross-validation (repeated CV) with three times repetition and 10 folds to train the classifiers. All algorithms were used in both the binary as well as the multi-class classification approach.

H. Evaluation of Classification

The confusion matrix provides a summary of the prediction results of a classification algorithm. In this matrix, the number of correct and incorrect predictions are summarized with count values and broken down by each class. An example of a confusion matrix is shown in Table II.

TABLE II: Example Confusion Matrix.

Prediction	***Reference***	
	Positive	Negative
Positive	TP	FP
Negative	FN	TN

TP = True positives, FP = False positives, FN = False negatives, TP = True positives.

The output True Positive (TP) represents the number of positives classified accurately, where True Negative (TN) represents the number of negatives classified correctly. False

Positive (FP) shows the number of negatives that are predicted as positives, whereas False Negatives (FN) shows the number of positives that are predicted as negatives.

Accuracy, Sensitively and Precision were calculated from the confusion matrix (Eq. 5-7). These metrics were calculated to compare the performance of each model. Accuracy describes the fraction the model predicts right. Sensitivity defines the proportion e.g., fastballs, that are correctly identified as fastballs. Precision describes the proportion of predicted fastballs that actually are fastballs.

$$Accuracy = \frac{TP + TN}{Totalsample} \tag{5}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

The current study used both a binary and multi-class classification approach. As the number of pitches within the multi-class classification is imbalanced (ratio of fastball to change-up and curveball is 2:1:1), the F1 score was calculated to evaluate and compare the models. The F1 score is calculated based on the precision and sensitivity score and is able to handle imbalance data. The F1 score is the harmonic mean of the sensitivity and precision (Eq. 8).

$$F1 = \frac{2 * Precision * Sensitivity}{Precision + Sensitivity} \tag{8}$$

In the multi-class classification, the macro-specificity and macro-precision were calculated. The macro-specificity and macro-precision were calculated as the mean of the specificity and precision of every class separately. The machine learning algorithm with the highest accuracy and F1-score was used for comparison between the conditions and approaches.

III. Results

Twenty-four pitchers threw 763 pitches in total. The pitchers and pitches that did not match the inclusion criteria stated in the method section were excluded. A total of 406 successful pitches thrown by 19 pitchers were included in the current study[1]. Descriptive statistics are presented in Table III.

TABLE III: Descriptive statistics of the included pitches.

Feature	Fastball (n = 213)		Change-up (n = 94)		Curveball (n = 97)		Total (n = 404)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Trunk (°/s)	748	236	832	276	809	244	782	250
Pelvis (°/s)	725	132	708	128	679	109	710	127
Separation (ms)	24.1	138	61.8	148	59.0	104	41.2	134

Figure 3 shows the ball velocity. One-way ANOVA showed significant differences in ball velocity among pitch types (F(2) = 83.07, p <.001). A post hoc Tukey test showed that every pitch type significantly differed from the other two groups. The pelvis and trunk angular velocity, and separation time for every pitch are shown in Appendix III.

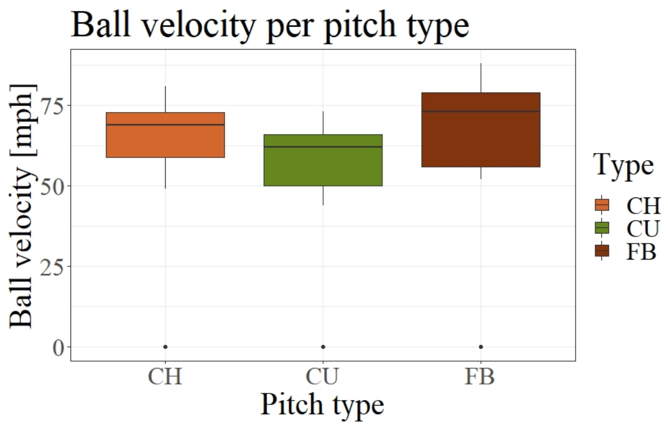


Fig. 3: Ball velocity per pitch type. CH = Change-up, CU = Curveball, FB = Fastball.

A. Binary classification approach

The four evaluation scores from the three best performing algorithms of every condition are shown in Figure 4.

In all conditions, the random forest (RF) algorithm demonstrated the highest evaluation scores (Condition A: Accuracy = .70, Precision = .72, Sensitivity = .71, F1 score = .71; Condition B: Accuracy = .69, Precision = .71, Sensitivity = .73, F1 score = .71; Condition C: Accuracy = .72, Precision = .73, Sensitivity = .75, F1 score = .74). Accuracy scores for RF were 10.9% and 12.8% in condition A, 4.60% and 9.11% in condition B and 12.3% and 15.8% in condition C higher compared to KNN and SvmRad, respectively. F1 scores for RF were 10.8% and 10.1% in condition A, 5.93% and 7.64% in condition B and 12.5% and 18.2% in condition C higher compared to KNN and SvmRad, respectively. Based on the highest evaluation scores, RF was selected as the best fit for the binary classification approach. The corresponding confusion matrices are shown in Table IV.

TABLE IV: Confusion matrices of RF in condition A, B and C in the binary classification approach.

Condition A			**Condition B**			**Condition C**		
Prediction	***Reference***		***Prediction***	***Reference***		***Prediction***	***Reference***	
	FB	Other		FB	Other		FB	Other
FB	453	181	FB	466	201	FB	480	178
Other	186	392	Other	173	372	Other	159	395

FB = Fastball, Other = Change-up and Curveball. Condition A includes $Pelvis$, $Trunk$ and Sep. Condition B includes $Prop_{Pelvis-Trunk}$, FB_{Pelvis}, FB_{Trunk} and FB_{Sep}. Condition C includes both condition A and B.

B. Multi-class classification approach

The four evaluation scores from the two best performing algorithms of every condition are shown in Figure 4.

[1]This number was lower than planned and maybe too low for this study. Due to Covid restrictions, it was however not possible to obtain more data within the time available for this thesis

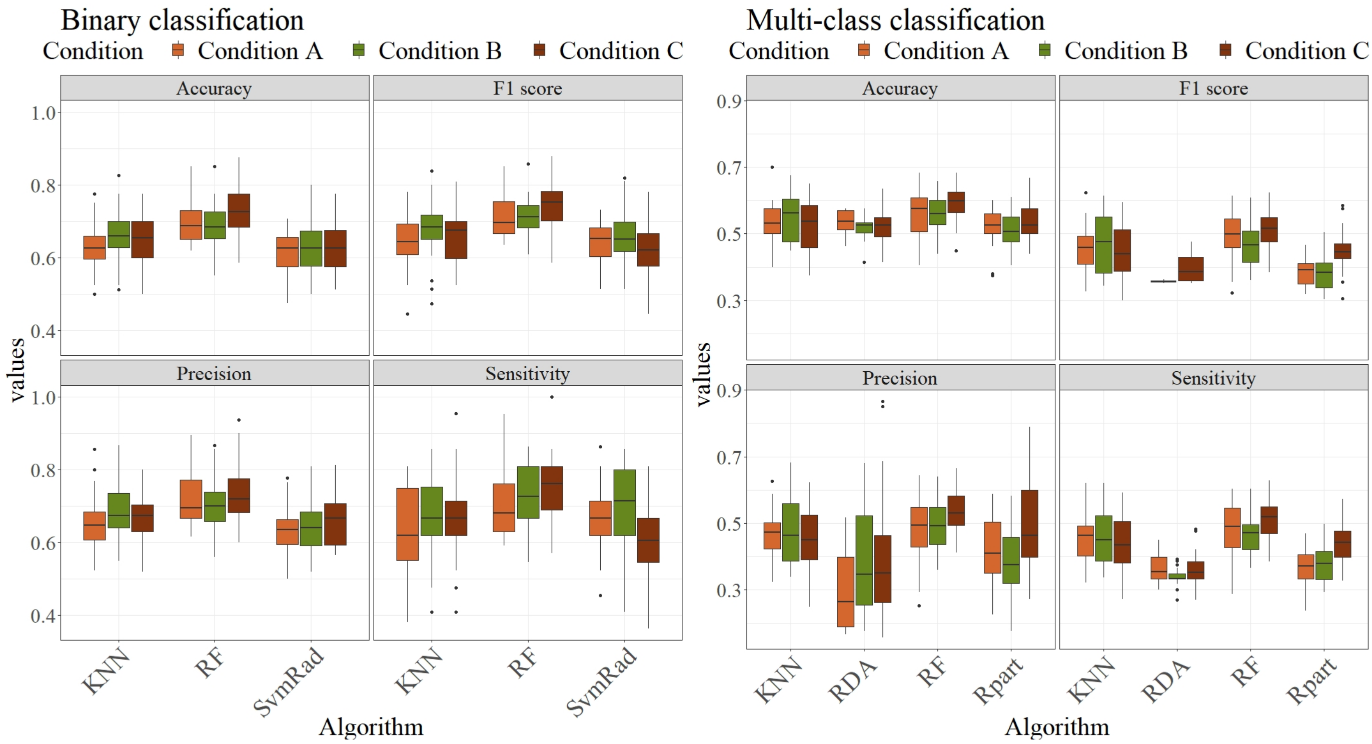


Fig. 4: Evaluation scores. Left. Binary classification. Right. Multi-class classification. KNN = K-nearest Neighbours, RF = Random Forest, SvmRad = Support Vector Machine, RDA = Discriminant Analysis, RPart = Decision trees. Condition A includes $Pelvis$, $Trunk$ and Sep. Condition B includes $Prop_{Pelvis-Trunk}$, FB_{Pelvis}, FB_{Trunk} and FB_{Sep}. Condition C includes both condition A and B.

RF demonstrated the highest evaluation scores in condition A (Accuracy = .56, Precision = .48, Sensitivity = .48, F1 = .50), condition B (Accuracy = .56, Precision = 0.49, Sensitivity = .47, F1 = .47) and condition C (Accuracy = .60, Precision = .54, Sensitivity = 0.51, F1 = .52). Accuracy scores for RF were 5.72%, 5.32% and 8.30% in condition A, 3.52%, 8.82% and 10.7% in condition B, and 14.0%, 14.1% and 10.4% in condition C higher compared to KNN, RDA and RPart, respectively. F1 scores for RF were 9.18%, NaN, 29.5% in condition A, 0.13%, 31.1% and 22.1% in condition B, and 14.1%, 29.3% and 14.8% in condition C higher compared to KNN, RDA and RPart, respectively. Based on the highest accuracy and F1 score, RF was selected as the best fit for the multi-class classification approach. Table V shows the corresponding confusion matrices.

TABLE V: Confusion matrices of RF in condition A, B and C in the multi-class classification approach.

Condition A				**Condition B**				**Condition C**			
Prediction	***Reference***			***Prediction***	***Reference***			***Prediction***	***Reference***		
	CH	CU	FB		CH	CU	FB		CH	CU	FB
CH	74	69	74	CH	82	63	68	CH	99	56	70
CU	71	112	76	CU	62	96	65	CU	59	117	63
FB	137	110	489	FB	138	132	506	FB	124	118	506

FB = Fastball, Other = Change-up and Curveball. Condition A includes $Pelvis$, $Trunk$ and Sep. Condition B includes $Prop_{Pelvis-Trunk}$, FB_{Pelvis}, FB_{Trunk} and FB_{Sep}. Condition C includes both condition A and B.

IV. Discussion

The aim of the current study was to use a machine learning algorithm to classify pitch types based on pelvis and trunk IMU data in training sessions. We demonstrated that the random forest (RF) is the best classification algorithm with a maximum accuracy of .72 and .60 and maximum F1 score of .74 and .52 in the binary and multi-class classification, respectively. Evaluation scores were between .12 and .24 higher when using a binary classification approach compared to a multi-class classification approach. Minor differences were seen when comparing condition A, B and C in both the binary as well as the multi-class classification approach, suggesting adding more features had almost no effect on the performance of the machine learning algorithms.

RF demonstrated to be the best fit for this specific classification problem. RF consists of various decision tree classifiers and uses randomly picked data to establish a different decision tree. RF then averages these decision trees to create a prediction. The improved performance of RF compared to the other algorithms might be explained by the nature of the current data. As can be seen in Appendix III, there is between-subject variability in pelvis and trunk angular velocity, and separation time. Therefore, there might not be a uniform solution to classify the pitch types. RF is based on ensemble learning, i.e., use of multiple learning algorithms, whereas the other algorithms are based on only one learning algorithm.

Descriptive statistics showed, surprisingly, that throwing a fastball resulted in the lowest trunk angular velocity compared to throwing a change-up or curveball. The pitchers had the highest trunk angular velocity when throwing the change-up. This contrasts with other research that have demonstrated that trunk angular velocity was the highest when throwing a fastball and the lowest when throwing a curveball [16]–[20]. Another interesting observation were the low values for all trunk angular velocities. The differences in trunk angular velocity were between 347-472 deg/s, 148-288 deg/s and 222-351 deg/s for the fastball, change-up and curveball, respectively [16]–[20]. Younger pitchers (12.5 $\pm$ 1.7 years) still had a mean trunk angular velocity between 979 and 1097 deg/s for every pitch type [16]. The differences between previous studies and the current study cannot be explained by age or level. However, the previous studies have measured the angular velocities from marker-based optical systems and in a laboratory setting. As PitchPerfect is designed to cope with practical issues in real-life situations, such as hitting the sensors, the filter is adjusted accordingly. Therefore, the differences between the obtained angular velocities of the current study and the previous studies might be device-specific.

Previous research has demonstrated that pitchers had the highest trunk and pelvis angular velocities when throwing the fastball compared to the other pitch types [16]–[20]. However, those studies were not consistent when comparing the angular velocities when throwing a change-up compared to a curveball. This is supported by the current results as there is an improvement in evaluation scores when using the binary classification approach compared to multi-class classification approach. Based on the improved performance when using the binary classification approach, it can be suggested that the algorithm has the most problems to classify curveballs and change-ups. The confusion matrices of the multi-class classification showed contradicting results. The number of predicted fastballs when the actual class was curveball or change-up was higher compared to the number of curveballs or change-ups when the actual class was a change-up or curveball, respectively. This suggests that the currently used algorithm has the most problems classifying fastballs.

Results of the current study demonstrated that there were some minor differences among the three conditions in both approaches. It is often assumed that using more features also means higher accuracy scores as there are more factors that could discriminate between the classes. Condition C included more features, but only a minor improvement was seen compared to the other conditions. Therefore, condition A is the most preferable as it includes only a few features and the features are directly extracted from the PitchPerfect software. Consequently, condition A is computationally less taxing compared to condition C and, therefore, easier to use in practice.

The current study used a small dataset because of the Covid restrictions. Small datasets are more prone to overfitting compared to larger datasets [31]. Cross-validation was used to overcome this problem, however, there might still be overfitting of the model. A larger dataset, i.e., more pitchers, would also provide a better representation of more combinations and variance of the features. Consequently, the model is able to learn better and classify future pitches with higher accuracy and/or F1 score. A larger dataset, i.e., more pitchers and more pitches, could also be used to take individual pitching characteristics into account. The current study tried to equalize pitchers by normalizing features by the mean of the fastballs. However, another way of equalizing pitchers and taking individual pitching characteristics into account would be to use pre-classification based on fastball characteristics [22]. As they tried to classify pitchers based on ball kinematics, the same can be done for trunk angular velocity. For example, pitchers with a relatively high trunk angular velocity could be classified into one group, whereas another group represents pitchers with a relatively low trunk angular velocity. Future research should therefore focus on pre-classifying pitchers based on pitching characteristics of the trunk angular velocity.

Another practical implication due to the Covid restrictions was that the current study made the assumption that age is not a confounder of the relationship between the angular velocities and pitch types. Previous research has demonstrated that pitching mechanics when throwing a fastball did not significantly differ with level [32]. In addition, the same pattern, i.e., angular velocity of the pelvis and trunk were the highest when throwing a fastball and the lowest when throwing a change-up, occurred at all levels of play. They also found that there were no significant interactions between pitch type and competition level [20]. Therefore, we have made the decision to include all pitchers without any age restrictions.

An alternative route to classify pitch types is to train classifiers for every pitcher individual. As stated previously, there seems to be between-subject variability in the collected data. Therefore, classifying pitch types for individual pitchers might seem like a promising alternative. However, pitchers show within-subject variability when investigating the variability of kinematic sequences [33]. They found that 10% of the pitchers performed only one kinematic sequence pattern, whereas 50% performed two types of kinematic sequence patterns. Therefore, the currently used route seems to be the best and computationally less taxing approach to classify pitch types.

We have made objective information available in training sessions and have demonstrated that RF is able to classify pitch types in PitchPerfect. Pitchers can use the classification algorithm in the PitchPerfect application to track their performance in training(-sessions) and use the information to tune their pitching mechanics. Coaches can use the classification to compare pitch count and pitch mechanics between pitchers and decide the pitcher order or maximum number of innings for every pitcher in matches. Pitchers and coaches can now design and outline training sessions and matches more efficient and pitchers are able to take control of their own mechanics and improvements. Standalone or together with other technological innovations focusing on the ball kinematics, such as Rapsodo pitching 2.0 (© Rapsodo LLC), pitchers are now able to understand their own pitching mechanics in relation to ball kinematics for every pitch type and act accordingly.

V. Conclusion

The random forest algorithm is the best machine learning algorithm to classify pitch types, i.e., fastballs, change-ups and curveballs, based on pelvis and trunk IMU data with a reasonable accuracy. Adding additional features besides the three outcome parameters of PitchPerfect, pelvis and trunk angular velocity, and separation time, had almost no effect on the performance of the algorithms. The evaluation scores were better when using a binary classification (fastball vs others) compared to a multi-class classification (fastball vs change-up vs curveball) approach. The random forest algorithm can be implemented in PitchPerfect application where pitchers and coaches could use the classification for designing training and match routines. Future research should focus on larger datasets, i.e., more pitchers and pitches, to pre-classify pitchers with similar pitching characteristics in order to improve the classification algorithms.

Acknowledgements

We would like to thank Jason Halman and Erik van der Graaff from PitchPerfect for arranging the measurements. We would also thank the pitchers and their coaches for participating in the current study and for being so hospitable. In particular, the staff and players of Twins Oosterhout as we could test the sensors at their training sessions.

References

[1] FanGraphs, "Fangraphs," *https://www.fangraphs.com*, 2021.

[2] J. Prodromo, N. Patel, N. Kumar, K. Denehy, L. P. Tabb, and J. Tom, "Pitch characteristics before ulnar collateral ligament reconstruction in major league pitchers compared with age-matched controls," *Orthopaedic Journal of Sports Medicine*, vol. 4, 6 2016.

[3] MLB, "What is the difference between pitches? and why are there so many? — quick question (mlb originals)," *https://www.youtube.com/watch?v=0DFYJkneoMo*, 2020.

[4] N. P. Brouwer, T. Yeung, M. F. Bobbert, and T. F. Besier, "3d trunk orientation measured using inertial measurement units during anatomical and dynamic sports motions," *Scandinavian Journal of Medicine Science in Sports*, vol. 31, 2 2021.

[5] M. Lapinski, C. B. Medeiros, D. M. Scarborough, E. Berkson, T. J. Gill, T. Kepple, and J. A. Paradiso, "A wide-range, wireless wearable inertial motion sensing system for capturing fast athletic biomechanics in overhead pitching," *Sensors*, vol. 19, 8 2019.

[6] K.-Y. Lian, W.-H. Hsu, D. Balram, and C.-Y. Lee, "A real-time wearable assist system for upper extremity throwing action based on accelerometers," *Sensors*, vol. 20, 2 2020.

[7] V. A. Lizzio, C. M. Gulledge, D. G. Smith, J. E. Meldau, P. A. Borowsky, V. Moutzouros, and E. C. Makhni, "Predictors of elbow torque among professional baseball pitchers," *Journal of Shoulder and Elbow Surgery*, vol. 29, pp. 316–320, 2 2020.

[8] E. C. Makhni, V. A. Lizzio, F. Meta, J. P. Stephens, K. R. Okoroha, and V. Moutzouros, "Assessment of elbow torque and other parameters during the pitching motion: Comparison of fastball, curveball, and change-up," *Arthroscopy - Journal of Arthroscopic and Related Surgery*, vol. 34, pp. 816–822, 3 2018.

[9] K. R. Okoroha, V. A. Lizzio, F. Meta, C. S. Ahmad, V. Moutzouros, and E. C. Makhni, "Predictors of elbow torque among youth and adolescent baseball pitchers," *American Journal of Sports Medicine*, vol. 46, pp. 2148–2153, 7 2018.

[10] M. J. Rose, K. A. McCollum, M. T. Freehill, and S. M. Cain, "Quantifying throw counts and intensities throughout a season in youth baseball players: A pilot study," *Journal of Biomechanical Engineering*, vol. 143, 3 2021.

[11] G. Dublon and J. Markowitz, "Baseball activity and player recognition from inertial sensor data."

[12] A. L. Aguinaldo, J. Buttermore, and H. Chambers, "Effects of upper trunk rotation on shoulder joint torque among baseball pitchers of various levels," *Journal of Applied Biomechanics*, vol. 23, 2 2007.

[13] S. S. Burkhart, C. D. Morgan, and W. Kibler, "The disabled throwing shoulder: Spectrum of pathology part i: Pathoanatomy and biomechanics," *Arthroscopy: The Journal of Arthroscopic Related Surgery*, vol. 19, 4 2003.

[14] D. Fortenbaugh, G. S. Fleisig, and J. R. Andrews, "Baseball pitching biomechanics in relation to injury risk and performance," *Sports Health: A Multidisciplinary Approach*, vol. 1, 7 2009.

[15] E. van der Graaff, M. M. Hoozemans, M. Nijhoff, M. Davidson, M. Hoezen, and D. H. Veeger, "Timing of peak pelvis and thorax rotation velocity in baseball pitching," *The Journal of Physical Fitness and Sports Medicine*, vol. 7, 9 2018.

[16] S. Dun, J. Loftice, G. S. Fleisig, D. Kingsley, and J. R. Andrews, "A biomechanical comparison of youth baseball pitches: Is the curveball potentially harmful?" *American Journal of Sports Medicine*, vol. 36, pp. 686–692, 4 2008.

[17] R. F. Escamilla, G. S. Fleisig, S. W. Barrentine, N. Zheng, and J. R. Andrews, "Kinematic comparisons of throwing different types of baseball pitches," *Journal of Applied Biomechanics*, vol. 14, pp. 1–23, 1998.

[18] R. F. Escamilla, G. S. Fleisig, D. Groeschner, and K. Akizuki, "Biomechanical comparisons among fastball, slider, curveball, and changeup pitch types and between balls and strikes in professional baseball pitchers," *American Journal of Sports Medicine*, vol. 45, pp. 3358–3367, 12 2017.

[19] G. S. Fleisig, D. S. Kingsley, J. W. Loftice, K. P. Dinnen, R. Ranganathan, S. Dun, R. F. Escamilla, and J. R. Andrews, "Kinetic comparison among the fastball, curveball, change-up, and slider in collegiate baseball pitchers," *American Journal of Sports Medicine*, vol. 34, pp. 423–430, 3 2006.

[20] G. S. Fleisig, W. A. Laughlin, K. T. Aune, E. L. Cain, J. R. Dugas, and J. R. Andrews, "Differences among fastball, curveball, and change-up pitching biomechanics across various levels of baseball," *Sports Biomechanics*, vol. 15, pp. 128–138, 4 2016.

[21] A. Attarian, G. Danis, J. Gronsbell, G. Iervolino, and H. Tran, "A comparison of classification methods with an application to classifying baseball pitches," *Conference: International MultiConference of Engineers and Computer Scientists*, vol. 1, 2013.

[22] K. Umemura, T. Yanai, and Y. Nagata, "Application of vbgmm for pitch type classification: analysis of trackman's pitch tracking data," *Japanese Journal of Statistics and Data Science*, 4 2020.

[23] S. W. Barrentine, T. Matsuo, R. F. Escamilla, G. S. Fleisig, and J. R. Andrews, "Kinematic analysis of the wrist and forearm during baseball pitching," *Journal of Applied Biomechanics*, vol. 14, 2 1998.

[24] M. Kuhn, "Building predictive models in r using the caret package," *Journal of Statistical Software*, vol. 28, 2008.

[25] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*. Springer New York, 2013, vol. 103.

[26] Y. Xiaozhou, "Linear discriminant analysis, explained," https://towardsdatascience.com/linear-discriminant-analysis-explained-f88be6c1e00b, 5 2020.

[27] O. Harrison, "Machine learning basics with the k-nearest neighbors algorithm," https://towardsdatascience.com/machine-learning-basics-with-the-k-nearest-neighbors-algorithm-6a6e71d01761, 8 2018.

[28] N. S. Chauhan, "Decision tree algorithm, explained," ://www.kdnuggets.com/2020/01/decision-tree-algorithm-explained.html, 1 2020.

[29] Sruthi94, "Understanding random forest," https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/, 6 2021.

[30] R. Gandhi, "Support vector machine â introduction to machine learning algorithms," https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47, 6 2018.

[31] R. Kohavi and D. Sommerfield, "Feature subset selection using the wrapper method: overfitting and dynamic search space topology." AAAI Press, 1995, pp. 192–197.

[32] G. S. Fleisig, S. W. Barrentine, N. Zheng, R. F. Escamilla, and J. R. Andrews, "Kinematic and kinetic comparison of baseball pitching among various levels of development," *Journal of Biomechanics*, vol. 32, 12 1999.

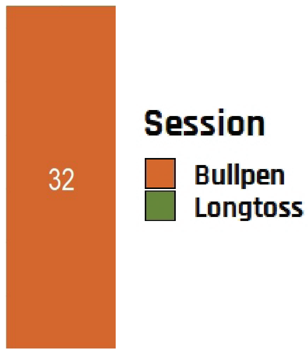
[33] D. M. Scarborough, A. J. Bassett, L. W. Mayer, and E. M. Berkson, "Kinematic sequence patterns in the overhead baseball pitch," *Sports Biomechanics*, vol. 19, 9 2020.

PITCHPERFECT

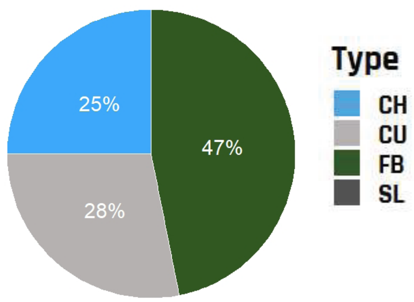
Name:
Date:
Team:

Pitch Count

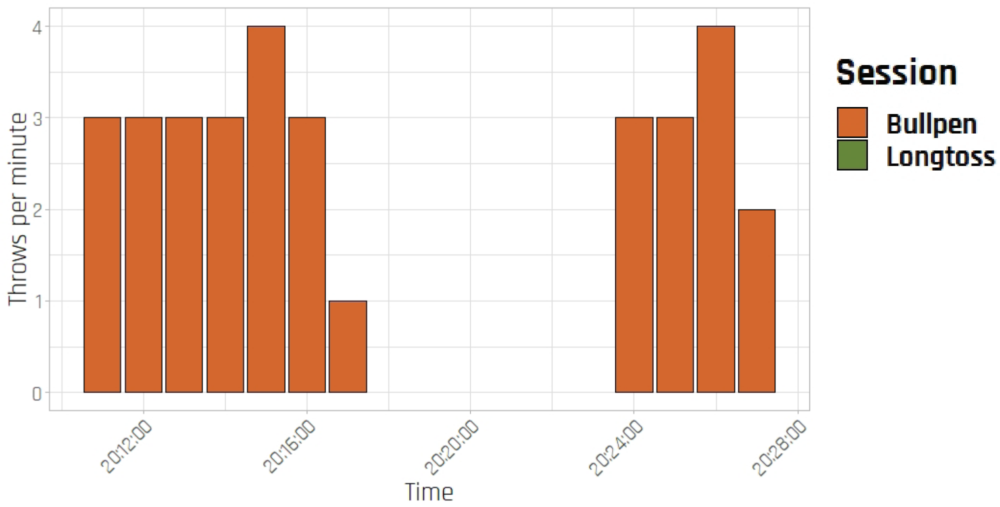
Total Pitch Count



Bullpen: Pitch Count per Pitch



Timeline



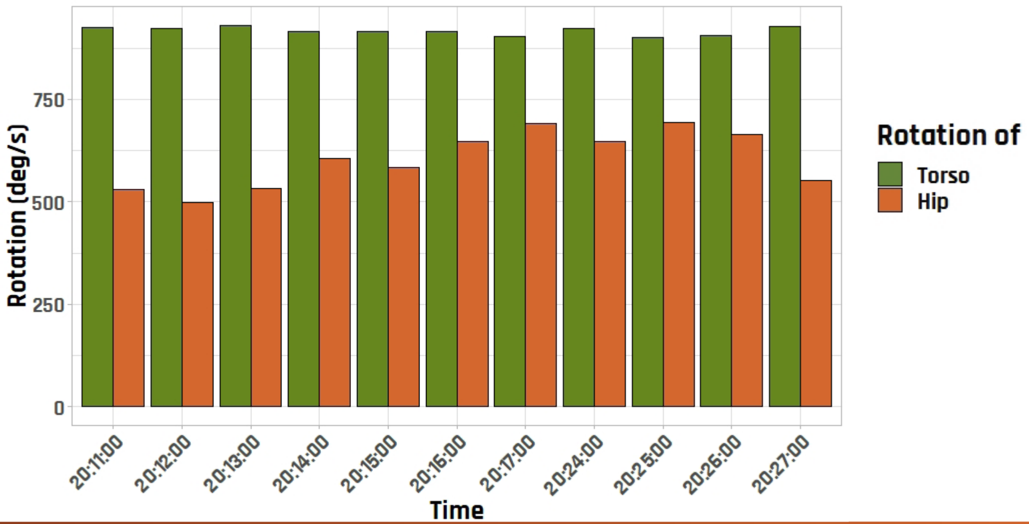
PITCHPERFECT

Name:
Date:
Team:

Consistency

Bullpen session

	Hip (deg/s)		Trunk (deg/s)		Speed (mph)	
	Average	Range	Average	Range	Average	Range
Total	599	444-753	917	890-943	75	63-87
CH	546	364-729	915	899-932	75	73-77
CU	655	522-788	904	887-922	66	64-69
FB	589	478-699	925	902-948	80	78-82



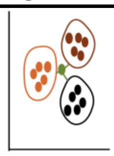
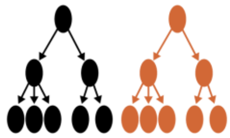
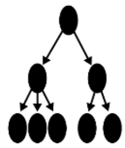
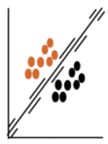
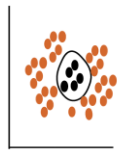
Efficiency

Seperation time: Good timing between upper and lower body is essential to transport energy through the body and results in reliable and consistent pitching. Good timing between the hips and trunk is between 5-50 ms.

info@pitchperfect-baseball.com
www.pitchperfect-baseball.com
@PitchPerfectBaseball

Appendix II. Machine learning algorithms

Table VI. Included machine learning algorithms in *Caret*

Name	Caret function	Tuning parameters*	Figure
K-Nearest Neighbours	knn	k	
Naïve Bayes	nb	fL, usekernel, adjust	
Random Forest	rf	mtry	
Classification and regression trees (CART)	rpart	cp	
discriminant analysis	rda	Gamma, lambda	
Support vector machines			
Linear	svmLinear	C	
Radial	svmRadial	Sigma, C	

*Tuning parameters were set at default.

Appendix III. Included pitches

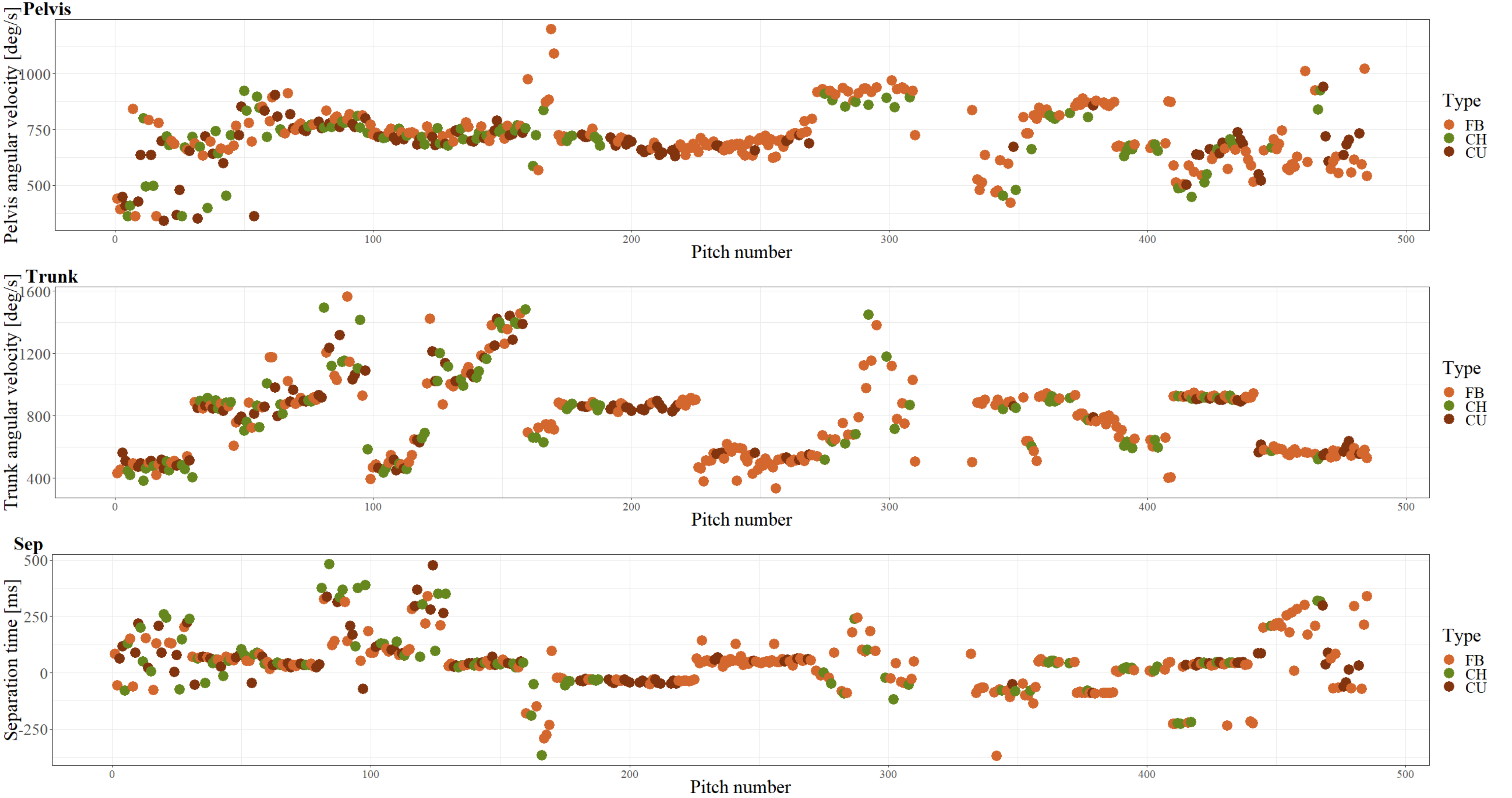


Figure V. Trunk and pelvis angular velocity, and separation time. FB = Fastball, CH = Change-up, CU = Curveball.