Neural Network-Based Prediction of Disc Golf Throw Speed from Motion Capture, EMG and Force Plate Data

Utilizing Machine Learning for Enhanced Disc Golf Performance Analysis

Andreas G. Björtoft
Department of Computer Science,
Electrical and Space Engineering
Luleå University of Technology
Luleå, PA 97187
`Aneboj-9@student.ltu.se`

Felix S. Woxblom
Department of Computer Science,
Electrical and Space Engineering
Luleå University of Technology
Luleå, PA 97187
`Sanfel-1@student.ltu.se`

Oscar A. Östensson
Department of Computer Science,
Electrical and Space Engineering
Luleå University of Technology
Luleå, PA 97187
`Setosa-0@student.ltu.se`

April 23, 2025

Abstract

The biomechanics of the disc golf throw involve complex, high-speed movements that require coordinated muscle activation and force applications. The purpose of this study is to identify key biomechanical factors influencing throw velocity and develop predictive models for disc speed based on biomechanical data. A multi-modal data collection approach was employed, integrating 3D motion capture, electromyography (EMG), force plate analysis, and velocity measurements of the disc to capture kinematic and kinetic variables. Challenges such as marker mislabeling, synchronization issues across different sampling rates, and missing data points due to high-velocity movements and occlusions were addressed through data preprocessing. Techniques such as manual labeling and normalization were applied to improve data accuracy and consistency.

Two predictive modeling approaches were explored to predict the disc velocity: an Artificial Neural Network (ANN) for classifying disc speeds and a regression model using Slice Sampling. The results showed that the Slice Sampling model achieved a root mean square error (RMSE) of 6.58, while the ANN model had an RMSE of 7.60. For comparison, random guessing within the dataset's value range would result in an RMSE close to 20. These findings demonstrate the feasibility of using biomechanical and kinematic data to predict disc golf throw speed and highlight the potential of machine learning models in sports performance analysis.

Keywords: Artificial Neural Networks, Biomechanics, Disc Golf, Electromyography, Machine Learning, Markov Chain Monte Carlo, Motion Capture, Slice Sampling

[0]Examinor: Kerstin Ramser
Supervisors: Ulrik Röjezon, Joel Wahl, Johan Jirlén

Contents

Keywords Disc Golf · Biomechanics · Motion Capture · Electromyography · Machine Learning · Artificial Neural Networks · Markov Chain Monte Carlo · Slice Sampling

1 Introduction

Disc golf is a flying disc sport in which players aim to complete each hole in the fewest possible throws, similar to traditional golf. Instead of hitting a ball into a hole, players throw a disc into a metal basket, typically navigating a course made up of 9 or 18 holes across varied terrain. The sport combines precision, power, and strategy, and has grown in popularity worldwide due to its accessibility and low equipment costs.

The use of 3D human modeling in sports science has been increasingly recognized for its ability to enhance motion analysis and performance evaluation (Imamoglu, 2019) [1]. Such methodologies have been applied in various sports to assess movement patterns and optimize athlete training, further emphasizing the relevance of 3D motion capture in biomechanical research. In disc golf, previous studies have used EMG to analyze muscle activation during throws, underlining the role of specific muscle groups in generating velocity (Holcomb, 2023) [2]. Beyond performance, research has also explored injury patterns in disc golf, highlighting the need for proper technique to minimize risk (Nelson et al., 2015) [3]. From a mechanical perspective, biomechanical analyses of throwing techniques (Greenway, 2007) [4] and aerodynamic modeling of disc flight (Giljarhus et al., 2022) [5] further demonstrate the complexity of factors influencing performance. Despite these contributions, the application of machine learning to synthesize such data and provide actionable feedback for disc golf athletes remains underexplored, a gap this project aims to address.

Understanding the biomechanics of a disc golf throw is crucial for optimizing performance, injury prevention, and improving training methodologies. In this study, a comprehensive motion analysis was conducted to investigate the relationship between various biomechanical factors and disc speed. The primary objective was to collect and analyze kinematic, kinetic, and muscle activation data to develop predictive models capable of estimating throw velocity based on the collected data.

This study was conducted with a single participant, which limits the generalizability of the findings but allows for a focused investigation of the biomechanics involved in the throw. To achieve this, a multi-modal data collection approach was employed, integrating 3D motion capture, electromyography (EMG), force plate analysis, and velocity measurement tools. These methods allowed for the detailed examination of key performance indicators, such as muscle activation, ground reaction forces, and joint kinematics.

The second part of the study focused on modeling throw velocity using machine learning techniques. Two primary approaches were explored: an Artificial Neural Network (ANN) and a Markov Chain Monte Carlo (MCMC)-based regression model incorporating both EMG, force plate and 3D data. Given the nonlinear and multi-dimensional nature of biomechanical interactions, machine learning techniques offer a powerful tool for uncovering patterns that traditional statistical models may overlook. Despite their simplicity, MCMC regression models remain valuable for their interpretability and ability to assess output reliability, complementing more complex machine learning methods. These models were evaluated based on their ability to accurately predict throw speed.

This report outlines the experimental methodology, data processing techniques, and model development process, followed by an analysis of the results and a discussion on potential improvements. By investigating the biomechanics of the disc golf throw through motion analysis and machine learning models, this study contributes to the growing body of research on performance optimization in sports.

2 Theory

The theory is divided into two parts, the theory behind the models used and the experimental theory.

2.1 Experimental theory

2.1.1 3D data

The 3D data is recorded with motion capture, which record multiple reflective markers placed on the participants body in time and place. Obtaining skeleton data from a 3D recording provides valuable joint information for studying technique and movement patterns as previously demonstrated by (Imamoglu, 2019) [1]. A system called AIM (Automatic Identification of Markers) (Qualisys) [6] were utilized, which automatically detects and labels reflective markers based on predefined templates and spatial relationships. This is meant to significantly reduces the need for manual marker labeling and minimizes errors in marker identification, ensuring more efficient and reliable motion capture workflows.

The 3D-data captures full-body movement trajectories that are crucial for analyzing the biomechanics of the disc golf throw. Due to the dynamics of the throw, every part of the body could play a crucial role in how the disc will fly.

2.1.2 Planting foot

The planting/bracing leg with the combination of the planting foot is were the forward momentum of the lower body gets translated into force of the upper body. The planting foot can be visualized in Figure 1. See 3.3.2 for technical details. Travis Greenway [7] highlights the importance of the bracing leg in disc golf biomechanics. Due to this, the force plate was chosen as a data collection method to collect the ground reaction force and direction of applied force. A force plate measures the forces exerted on it when a person stands, steps, or moves on it. It uses sensors to detect vertical and horizontal ground reaction forces, providing data on balance, movement, and force output.

Figure 1: The planting foot, inside the red circle, during the end of a discgolf throw serves as the pivot point through which force is transferred from the lower body to the upper body

2.1.3 Muscles in arm

As with the planting foot, the muscles in the arm and shoulder were considered to be a very important part of the throw. The authur James Holcomb [2] highlights the importance of specific muscle groups in generating throw velocity. consequently, four muscles were chosen for the EMG recording in this study. Electromyography (EMG) measures the electrical activity produced by muscles during contraction. Electrodes placed on the skin detect these signals, which reflect muscle activation levels, see more in section 3.1.1. This is important since the timing of the muscle activation is important to ensure both good technique and a long distance throw.

2.1.4 Disc velocity

The velocity of the disc were measured using two different types of methods. The main recording method was a Gameproofer Disc [8]. This disc uses built-in accelerometers and gyroscopes to measure motion and rotation, allowing it to calculate the disc's speed immediately after release. The data is then visable on an app on the phone.

The other method was a speed sensor from Biltema [9]. It emits a microwave signal, and when the disc moves toward it, the frequency of the reflected signal changes slightly. ~~Then~~ it displays the speed.

2.2 Model theory

2.2.1 Slice Sampling and Regression Theory

Slice sampling is a Markov Chain Monte Carlo (MCMC) method used to sample from complex, high-dimensional distributions, meaning that the distribution does not have a nice, simple shape like a normal distribution and that there are many variables. An example would be if the dataset contained traits of a group of people it would not just contain their height, but it would contain height, weight, age, gender, nationality and maybe 15 other variables. It is particularly useful for drawing samples from the posterior distribution of the coefficients β in a linear regression model. Markov Chain Monte Carlo (MCMC) methods are a class of algorithms used to approximate distributions by sampling from them. They are particularly useful when the distribution is complex and cannot be directly computed.

The key idea behind slice sampling is to efficiently sample from a distribution without requiring knowledge of its normalization constant. This is achieved by iteratively defining a slice at a random level of the posterior distribution and updating each parameter β to stay within that slice. This means that first a random height is chosen on the curve. Then the regression coefficient β is picked randomly from the part of the distribution that is above the height that was previously chosen. In figure 2 it is visualized. The x-axis is the value of the β coefficients and the y-axis is the probability of the β-coefficient having those values. A slice height y is chosen randomly and the value x is chosen randomly so that the curve f(x) is above the slice height y.

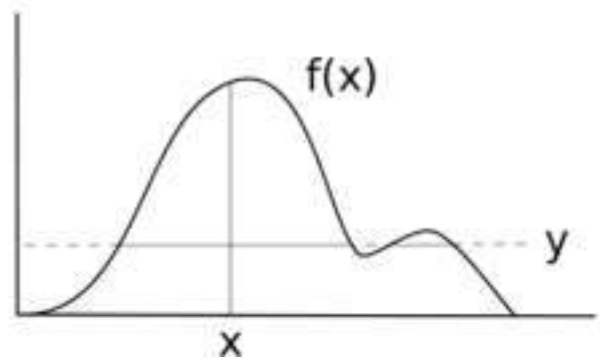

Figure 2: Slice sampling visualized. The x-axis is the value of the β coefficients and the y-axis is the probability of the β-coefficient having those values. A slice height y is chosen randomly and the value x is chosen randomly so that the curve f(x) is above the slice height y.

If the slice sampling process is successful, the β values will converge with small fluctuations, ensuring a stable estimation of the regression coefficients. This allows for the inference of relationships between the the input variables and the outcome in a probabilistic manner. For more information on slice sampling, see Jesper Møller [10].

The prior distribution for β is assumed to be normally distributed with a gaussian likelihood since this is a convenient and computationally efficient assumption. The posterior distribution is the updated beliefs after seeing the data and it is a combination of the prior distribution and the likelihood using bayes theorem. It is implemented in the algorithm as

$$log_{posterior} = log_{likelihood}(\beta) + log_{prior}(\beta)$$

Regression is a statistical method for finding a relationship between a dependent variable and one or more independent variables. A regression function typically looks like this:

$$Y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

For more details on regression, see Investopedia [11].

2.2.2 Neural networks

A Neural Network (NN) is a mathematical model inspired by the human brain. It consists of layers of interconnected nodes (neurons), where each neuron processes input data to produce an output. By adjusting the weights through feedback from a learning process, the network can learn patterns in the data over time. For further details, refer to Ian Goodfellow [12].

3 Experimental methods

The participant was age 24, male and had five year of experience with disc golf. He was picked due to his experience in disc golf and since he was readily available to participate in the data collection.

To gather the data, a baseline for the experiment was created. This involves external and internal conditions for the participant/participants performing the experiments to ensure that the gathering of data is consistent. Data were collected using six different methods. For collecting model data, the methods were muscle activation for four muscles, motion capture for the whole body and force distribution for the bracing foot in the throw. For the target label data collection, a tech disc and a velocity radar were used. During the experiment, a control person needs to manually note down the recorded speed of the velocity radar.

The data were recorded in larger section consisting of multiple throws. How the throws were extracted from these continuous sections is discussed more in section 4.1. In total the dataset consisted of about 60 throws, ranging from 70km/h to 115km/h.

3.1 Data Collection Methods

3.1.1 Biomechanical Data Collection

For recording of the muscle activation we used a surface EMG system from Noraxon [13]. The system looks at muscle activation at the placement of the sensor, therefore correct electrode placement is critical to ensure accurate data collection. The EMG system was recording at 1500Hz to capture high frequency changes in the muscle activation. To make the data easier to understand, a baseline max muscle activation test was performed by the participant. This is done to have something to normalize the output with after doing the real experiment. Instead of getting values in volt, percentages of maximum muscle activation is the output. Baseline max tests were performed based on SENIAM [14] (Surface ElectroMyoGraphy for the Non-Invasive Assessment of Muscles) guidelines and other methods was developed through experimentation. The max test duration was the following: three secounds of max activaiton followed by five secounds of rest, repeated three times for each muscle.

The recorded muscles included the following:

- Deltoideus posterior (backside of the shoulder)
- Triceps brachii (long head of the triceps)
- Extensor carpi radialis longus (a muscle responsible for wrist extension and radial deviation)
- Flexor digitorum (a muscle that flexes the fingers and wrist, aiding in gripping and holding)

SENIAM had guidelines for the tricep and shoulder muscle. For the flexor muscle a dynamometer was used, and for the wrist extention muscle, resistance during extention was used as the max test. See Appendix A for the description of the developed methods.

For capturing 3D data of the human body, QTM [15] (qualisys track manager) was used. The camera system was recording at 300 Hz to ensure synchronization with the EMG recording speed of 1500 Hz. The body markers are placed based on the setup BAHRT project (M. Pauelsen [16]). Additional markers were placed on the tech disc 3.1.2, enabling an accurate determination of the release time of the disc for data processing.

For recording forces from the bracing foot, more about technique in section 3.3.2, a 3D force plate by Kistler [17] was used. It was recording at a speed of 300 Hz matching the motion capture data. The force plate outputs the ground reaction forces, torques and center of pressure that occurs.

3.1.2 Disc Speed Measurement

A tech disc from Gameproofer [8] was used. This disc can capture arm speed, flight speed, spin, distance and much more. For this paper, only flight speed was used. See Appendix B for images of the tech disc.

To get more redundancy a velocity radar from Biltema [9] was used. This redundancy ensures that errors or missing data from the tech disc does not compromise the overall dataset.

3.2 Environment

The physical environment 3 has a camera system setup for capturing 3D data, force plate, a net to catch the disc. The eight cameras were placed in an $5x8.5$ meter rectangle allowing a good capturing area. The velocity radar is set up behind the net at an appropriate distance to capture the speed and not get hit. The net was placed approximately 2.25 meters from the center of the force plate. The velocity radar was placed approximately 2 meters behind the net. No reflective objects other than the markers are allowed since they can disturb the camera system.

Figure 3: The environmental setup, with the 3D cameras, net, force plate and a participant. Reflective markers (grey dots) are placed on the participant's body to facilitate motion capture

3.3 Participant Instructions and Technique

3.3.1 External and Internal

To ensure accurate motion capture data and muscle activation measurements the participant was unclothed to just underwear and socks. The tech disc mentioned in 3.1.2 is used to ensure external consistency. The participant did not have any soreness, injuries or diseases that might affect results. The participant did their usual warm up beforehand to feel comfortable throwing the disc. The participant performed test throws to acclimate to the experimental setup, as it differed from their usual environment. Proper foot placement, which is crucial for accurate data collection (see Section 3.1.1), was a key focus during this familiarization.

3.3.2 Technique

The participant threw the disc as usual when throwing long distances with a backhand throw. The type of throw analyzed in this study was the backhand throw. It is the typical throw used for control and distance by the majority of disc golf practitioners. The participant was asked to throw using the normal X-step, which is a foot-work method where you approach the throwing position by crossing your legs. The last step and the bracing leg must be placed on the force plate to ensure good data.

3.4 Ethical considerations

The participant could be injured during the experiment. This can be in the form of muscle and joint strains. The risk can be minimized by ensuring that the participant is properly warmed up and hydrated. If injury occurs, the experiment will be interrupted so that the participant may receive appropriate care. Missing the net with the disc could potentially damage equipment or other people. The net is 2x2m large, and the participant was asked to throw the disc about 1,5m away from the net, ensuring that missing the net is extremely unlikely.

Since participants was not fully clothed, it is important to ensure that no sensitive data is leaked and to respect integrity and personal space. The dataset contains no personal information, and all data collected (such as muscle activity, movement, and force data) is anonymized. Data is securely stored in a shared Google Drive folder with access limited to authorized personnel only. A blanket was provided to cover the participant between measurements, both to enhance privacy and to reduce the risk of getting cold. The research team will also ensure that the participant's personal space is respected throughout the process.

4 Data preprocessing

The data was intended to be used for a MCMC model and an ANN to take the collected data and predict what speeds the disc throws had. For this, some data preprocessing had to be done. The data was however first exported from QTM into a MAT-file and later used in both matlab and python.

4.1 Timesteps and Labeling

Using QTM one can determine the timesteps of when a data sample should start and stop by going through the recorded frames. That is, what is considered a throw for the models. In this case, the start of a sample was determined to be when the disc was at its lowest point vertically in the beginning of the throw just after the bracing leg had landed. The end of the sample was when the disc left the hand. This part of the throw could be seen to be the most influential to how fast the disc will be thrown, since this is where most of the forward acceleration occurs. Limiting the data to only this part made it more effective to work with later on, limiting unnecessary noise.

Based on all these timesteps a dataset could be generated in combination with the labels gathered in the experimental phase. As mentioned in section 3.1 , different equipment had different recording frequencies, so the frames had to be offset for the EMG compared to the 3D data the timesteps were gathered from.

4.2 Biomechanical Data Preprocessing

In the section 2.1.1, the AIM model was briefly described. The skeleton data from the AIM model did not always give a skeleton. Due to this, manual labeling was made in QTM. Due to time constraints and tedious work, only the marker on the right hand, between the index- and middle finger, was labeled fully. The 3D data was thus represented by a single marker point. Each point consists of x, y and z position and a confidence factor. From this data, the maximum speed was derived and extracted then used for the model. The raw data can be seen in the Appendix C.

The force plate produced raw data on with these outputs: 'Fx', 'Fy', 'Fz', 'Mx', 'My', 'Mz', 'COPx', 'COPy', 'COPz'. F = Force, M = Torque, COP = Center Of Pressure.

As mentioned in section 3 the data were recorded in sections with multiple throws. To extract individual throws from the force plate data, the same start and end timesteps described in Section 4.1 were used, with an added buffer of -80 and $+600$ frames before and after the throw to ensure the full motion was captured.

From the nine output features, three parameters were extracted. The 'Fx', 'Fy', 'Fz' from a throw were used to form vectors. The average magnitude of the vectors is the values used later on to train the models. The raw data can be seen in Appendix C.

As mentioned in section 3.1.1, normalization of the EMG data was performed. Based on the maximum activity of the muscle in the max activation test. It's worth noting that this is meant to give a value between 0 and 1 where 1 would be max activation of that muscle. Because muscles function differently in static (isometric) and dynamic (moving) conditions, it's possible that muscle activation during movement could exceed the maximum measured in a static test. If this happened, the calculated output would be greater than 1. However, in this experiment, this did not occur, so no additional adjustments were needed afterward. From this data the maximum activation over each sample was used for the model. The ground force and hand speed data recorded was normalized by max value from each separate sample. The raw data can be seen in Appendix C.

The training parameters are then shown in Table 1 and all the measured parameters can be found in Appendix D, Table 2.

Experimental Variables and Descriptions

Variable	Description
x_1	Rear Delt
x_2	Wrist Extensor
x_3	Triceps
x_4	Finger Flexor
x_5	Ground Force
x_6	Hand Speed
y	Disc Speed

Table 1:

The data collected is splitted into a 70/30 split, were 70% is for training and 30% is for validaiton. This is done randomly, see the Github [18] for the code. Since the dataset is small, no test dataset was utilized.

5 Models

5.1 Application of MCMC in Disc Throw Prediction

To model the speed of a disc throw, a linear regression approach was used with MCMC, as described in section 2.2.1,

$$y = \beta_1 x_1 n_1 + \beta_2 x_2 n_2 + \cdots + \beta_N x_N n_N$$

where y represents the throw speed, β the learned parameters, and x are variables extracted from the recorded data. N are the normalization constants for each variable.

As can be seen in Table 1, the selected input variables are: $x_1 - x_4$: The maximum EMG signals from four different muscles during each throw. x_5: The average force vector from the force plate during each throw. x_6: The maximum speed of the hand during the throw. n1-n4 are derived so that the values are normalized based on the max muscle activations test of x1-x4. n5-n6 are derived to normalize x5 and x6 between 0 and 1. They are normalized to ensure numerical stability and improve convergence.

The coefficients β are estimated using slice sampling. This technique samples efficiently from the posterior distribution of the regression parameters, iteratively refining the estimates. By doing so, it provides a robust way to model the relationship between muscle activation, force application, and disc throw speed.

The Slice sampling algorithm is implemented in MatLab 2024b with the help of functions included in the Statistics and Machine Learning Toolbox. See the Github [18] for the code.

5.2 ANN

The ANN is a fully connected feedforward network which means that every neuron in one layer is connected with every neuron in the next layer and that it has no cycles or feedback loops. It has two hidden layers. The first of size 4 and the second of size 2. This means that the input data is first mapped to and computed by 4 neurons and the output of these neurons are mapped to and computed by 2 neurons. Both layers have a ReLu (Rectified Linear Unit) activation function. It is a mathematical function used in neural networks that keeps the values computed by the network from becoming too small. The output from the second layer is mapped to an output layer which gives a single continuous value which is the predicted disc speed. It uses mean squared error as a loss function which means that the error of each prediction is squared and then the mean of all squared errors is taken to produce a score to define performance. It uses Adam (Adaptive Moment Estimation) optimizer with a learning rate of 0.01. This is an optimization function and its purpose is to fine-tune the parameters of the network so that it becomes better and the learning rate dictates how aggressively these parameters are changed. A technique called L2-regularization is implemented which penalizes large weight on the neurons. This is to prevent something called overfitting which is when the model gets increasing performance on the training data while getting decreasing performance on testing data. It runs for 1000 epochs and early stopping is implemented with a patience of 50. This means that if the model does not improve for 50 epochs, the training will finish and it also helps prevent overfitting.

The ANN is implemented using Python 3 and imported libraries such as numpy which is a library for numerical computing, torch which is a library including functions for neural networks and matplotlib which is a library for making plots. See the Github [18] for the code.

6 Results

The models are evaluated using Root mean squared error (RMSE). It takes the true values and subtracts the predicted values. It squares the result so that overestimations are valued equally to underestimations. It takes the mean of all these squared errors in the validation set and takes the square root of it to get a result that accurately describes the average error on the validation set. This means that the model should aim to get as low RMSE as possible.

6.1 Using all input variables

6.1.1 Slice Sampling

Slice Sampling was performed with the variables x_1 = max deltoideus posterior activation, x_2 = max extensor carpi radialis longus activation, x_3 = max triceps brachii activation, x_4 = max flexor digitorum activation, x_5 = average absolute ground force, x_6 = max right hand speed. Using these variables, coefficients for β are found that have an average RMSE on the validation set of 6.9106.

The trace plots 4 show the convergence behaviour of the regression coefficients β_1 to β_6. The plots show convergence, no long periods of stagnation and random fluctuations after convergence which are all good signs of good mixing.

The posterior distribution 5 shows what values that are the most likely to represent the coefficients. A peak at 0 or very close to 0 shows that a variable might be unnecessary. Based on the conergence time of the trace plots in figure 4, the posterior distribution does not include the first 400 iterations.

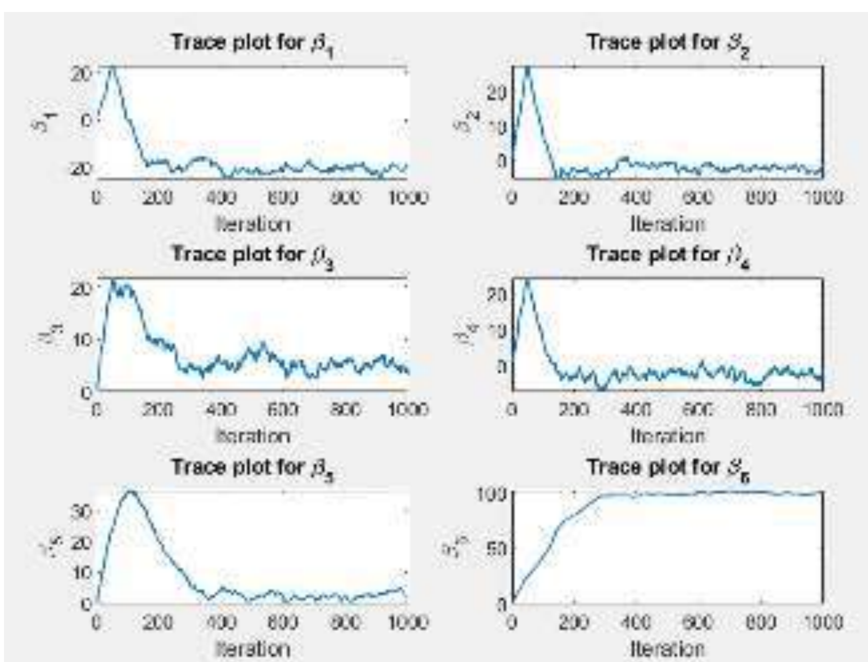

Figure 4: Trace plot showing convergence of β coefficients. The x-axis shows the number of iterations run by the algorithm and the y-axis shows the most likely value for the regression coefficient β. As the algorithm performs more iteration the plots converge meaning that there is a better idea of what values β can have.

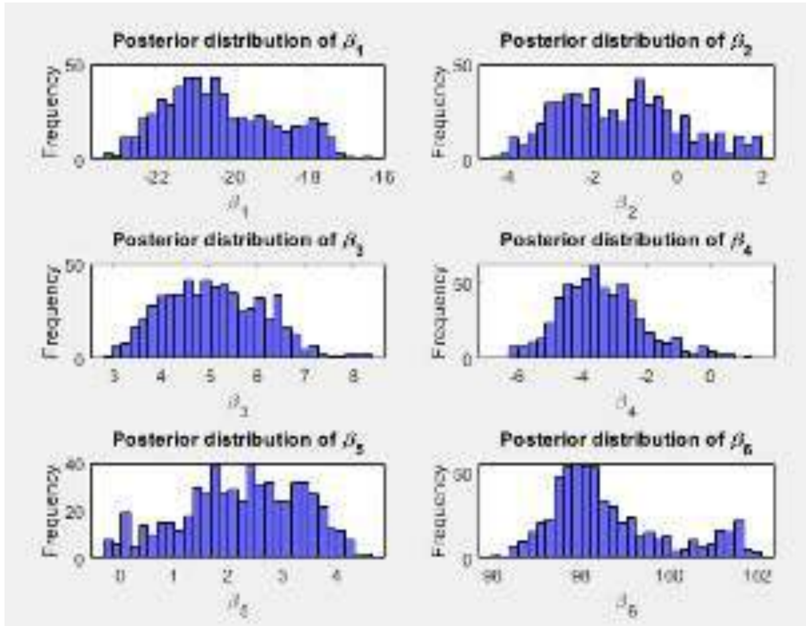

Figure 5: Posterior distirbution showing what values are most likely to represent the coefcients. Excluding the first 400 iterations it takes the values of β for each iteration from the trace plots to create a distribution plot where the x-axis shows the value of β and the y-axis shows the frequency of that value which directly relates to the probability of β having that value.

In order to obtain a single regression coefficient for each variable, the mean of each posterior distribution is taken to compute a single value for each β. The complete regression function then becomes:

$$\begin{aligned} Y = & -16.2472 \cdot x_1 \cdot \frac{1}{\max x_1} + (-0.4449) \cdot x_2 \cdot \frac{1}{\max x_2} + 7.2976 \cdot x_3 \cdot \frac{1}{\max x_3} \\ & + (-1.1798) \cdot x_4 \cdot \frac{1}{\max x_4} + 8.5185 \cdot x_5 \cdot \frac{1}{\max x_5} + 86.5659 \cdot x_6 \cdot \frac{1}{\max x_6} \end{aligned}$$

Here, $\max x_1$ to $\max x_4$ are the maximum values from the muscle activation tests (in μV), $\max x_5$ is $1000\,\mathrm{N}$, and $\max x_6$ is the highest value of that variable in the dataset, which is $20.3756\,\mathrm{m\,s^{-1}}$. Plugging in these values, the final regression function becomes:

$$\begin{aligned} Y = & -16.2472 \cdot x_1 \cdot \frac{1}{3.6763 \times 10^3\ \mu\text{V}} + (-0.4449) \cdot x_2 \cdot \frac{1}{999.7864\ \mu\text{V}} + 7.2976 \cdot x_3 \cdot \frac{1}{1.9147 \times 10^3\ \mu\text{V}} \\ & + (-1.1798) \cdot x_4 \cdot \frac{1}{1.1356 \times 10^3\ \mu\text{V}} + 8.5185 \cdot x_5 \cdot \frac{1}{1000\text{N}} + 86.5659 \cdot x_6 \cdot \frac{1}{20.3756\ \text{m/s}} \end{aligned}$$

6.1.2 ANN

The ANN used the same input as the Slice Sampling mentioned above, that is: x_1 = max deltoideus posterior activation, x_2 = max extensor carpi radialis longus activation, x_3 = max triceps brachii activation, x_4 = max flexor digitorum activation, x_5 = average absolute ground force, x_6 = max right hand speed. The ANN gets an average RMSE of 8.5369 on the validation set.

As seen in Figure 6, the model learns rapidly during the first 19 epochs. After this initial phase, the loss fluctuates between 35 and 80 for the remaining epochs until early stopping occurs.

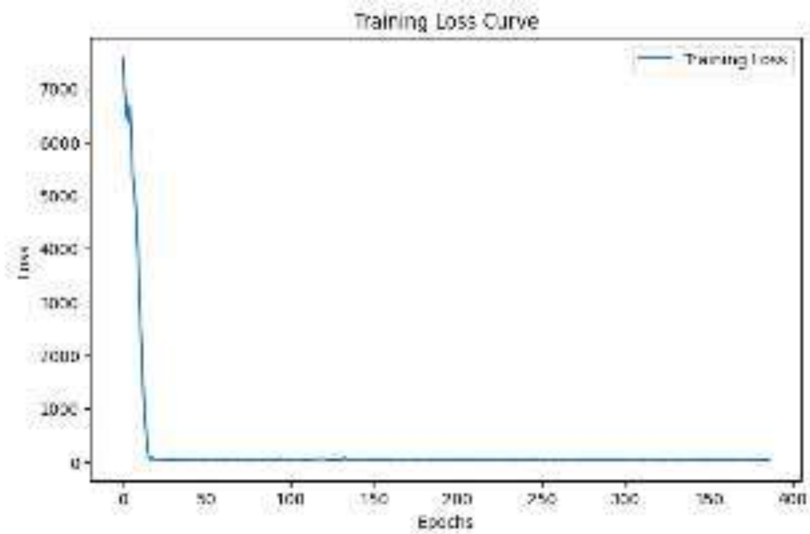

Figure 6: Training curve for ANN. The plot shows the x-axis number of epochs (training iterations through the dataset) and on the y-axis it shows the loss which is measured in MSE (RMSE^2). It can be seen that the loss starts at around 7500 and quickly improves during the first 20 epochs wherafter it fluctuates with really small changes around $35 - 80$.

6.2 Best variable selection

Based on the the posterior distribution 5, testing was done by removing input variables deemed unnecessary, based on experiments the best result was achieved when x_5 was removed from the input.

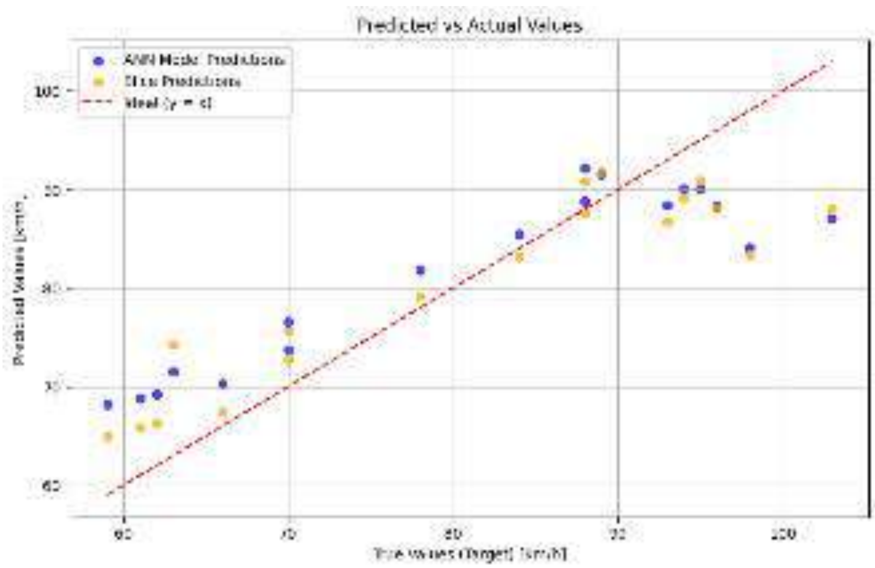

Figure 7: Scatter plot comparing predicted and actual disc speeds using two modeling approaches: the Artificial Neural Network (ANN) and the MCMC-based Slice Sampling method. The x-axis represents the true (target) disc speeds measured in km/h, while the y-axis shows the predicted values from each model. The red dashed line represents the ideal prediction line ($y = x$), indicating perfect prediction accuracy. The proximity of the predicted points to this line reflects the accuracy of each model.

As can be seen in the figure 7 of the ANN and Slice Sampling, one can see a slight difference, in advantage for Slice Sampling. As stated earlier, Slice Sampling with RMSE = 6.5832 and ANN with an RMSE = 7.6011 on the validation set. Slice Sampling typically predicts slightly more accurate compared to the ANN.

7 Discussion

7.1 Model Result Discussion

In this section, we discuss the results obtained from the two predictive models used to estimate disc throw velocity: Slice Sampling and an Artificial Neural Network (ANN).

Looking at Figure 4, we can see that β_6 converges to a large value, which is most likely due to the importance of x_6 (right hand max speed) in the model's result. This suggests that the right hand speed plays a key role in predicting disc velocity. Additionally, Figure 5 shows that β_2 and β_4, which represent the EMG data for the wrist extensor and finger flexor, respectively, have posterior distributions close to 0. Similarly, β_5, related to the average absolute ground force, shows minimal contribution. These results indicate that these variables may not be particularly valuable for improving model performance.

When compared to Holcomb [2], we used a different muscle for the wrist, namely the extensor carpi radialis longus, while they used the extensor carpi ulnaris. This discrepancy could suggest that the muscle choice in our model might not be optimal, potentially explaining why this muscle data did not provide significant insight into the model. Additionally, we speculate that our max activation test for the finger flexor may not have been ideal. This will be discussed further in 7.4. This poor muscle activation testing could contribute to the lack of contribution from the finger flexor muscle in the model.

Regarding ground force measurements, since the participant performed the throws without shoes, the force direction and production from the bracing leg may have been affected. A shoe typically provides a lot of friction, which could lead to more stable and directed ground forces. The absence of shoes might have reduced the contribution of the ground force data to the model, as the lack of friction could lead to unpredictable or less consistent results.

Turning to the ANN results, as shown in 6, we can see that its progress stalls after the initial epochs. The loss fluctuates between 35 and 80 for the remaining epochs until early stopping occurs. Given the small dataset (60 throws), we might have benefited from implementing cross-validation or bootstrapping. These techniques could have helped the model generalize better and extract more value from the limited data available.

While the Slice Sampling model achieved an RMSE of 6.5832 and the ANN model resulted in an RMSE of 7.6011, both of these values indicate modest accuracy. However, these results are still a significant improvement over random guessing, which would result in an RMSE close to 20. This is to be expected, given that disc speed is influenced by numerous factors. For example, we only measured 4 muscles, but other muscles, such as the biceps (which act as antagonists to the triceps), could also contribute to disc speed. Moreover, fatigue plays a role in the results, as the participant likely became more fatigued as the throws progressed, necessitating greater muscle activation to generate the same force. This variability could make accurate predictions more difficult. It is also worth noting that ANNs generally require large datasets to perform well, which likely explains the ANN's relatively poorer performance compared to Slice Sampling in this context.

In Figure 7, we observe that both models predict similar values for all the test points. However, the Slice Sampling model's predictions are consistently closer to the ideal line. This is reflected in the RMSE values, which confirm that the Slice Sampling model is slightly more accurate.

There are additional data sources that could be utilized in future iterations. For instance, torque and center of pressure data from the force plate might offer further insight. However, given that the force data did not significantly contribute to the current models, we did not prioritize these factors. Other promising variables for future studies include the maximum force recorded from the force plate and the timing of key movements and muscle activations. A faster throw typically generates greater forces in the body, and incorporating max force data could improve the model's predictive accuracy. Moreover, as discussed in Holcomb [2], the precise timing of muscle activations is crucial for achieving higher velocities, which could be a valuable feature in future models. Since disc golf is a full-body movement, future research could explore the integration of 3D motion capture data to improve the model further. This would allow for a more comprehensive analysis of body mechanics during the throw.

In conclusion, these results demonstrate that machine learning models trained on data from motion capture, force plate, and EMG sensors can provide useful predictions of disc velocity. While the accuracy of the models could be improved with more data and further feature exploration, this study marks an important step towards enhancing the analysis of throwing techniques in disc golf.

7.2 Using continuous data for an SNN

The data was recorded in a continous manner over frames, due to this our initial plan was using a Spiking Neural Network (SNN). Since it is not covered in this report, we wont go in to much detail. Simply put, SNNs are good at handling temporal dynamics, that is, changes in a signal or pattern over time in the input. For further details on the underlying biology and encoding methods, see Wulfram Gerstner [19].

The biggest problem however, is that spiking neural network models require a lot of data, data that we did not have. This could however be explored in the future, especially since we now have a better understanding of what type of data collection methods exist and how they work.

7.3 Difference in recording Hz

For our data collection methods, we had different sampling frequencies (Hz) to choose from. For the camera and force plate setup, we could select between 100–500 Hz, while for the EMG system, we had options of 1500 or 3000 Hz.

To synchronize the data effectively, we needed to ensure that the camera's sampling frequency was a clean divisor of the EMG sampling frequency (i.e., without a remainder). This constraint limited our choices.

Initially, we aimed to use 500 Hz for the camera system to maximize data resolution. However, after discussions with our supervisors, we considered the potential drawbacks of increasing the camera system's sampling rate. One concern was that higher Hz could lead to a loss of accuracy due to system limitations.

As a result, we opted for a lower sampling rate, which reduced the amount of collected data but ensured higher accuracy and synchronization across all measurement systems.

7.4 EMG max activation test

After performing our maximum EMG tests and normalizing the EMG data, we identified a potential issue with the finger flexor max test. We theorize that a more sport-specific approach could have provided more accurate results for this particular muscle.

A potentially better method would have been to perform the max test in the same grip position used during the experiment, specifically by having the participant grip the disc as hard as possible in their natural throwing grip. This could lead to a more effective and relevant measurement of maximum muscle activation for this specific movement.

7.5 Qualisys labeling

A preliminary test run was conducted to ensure the data collection process was functioning as expected. The initial results appeared satisfactory, leading to the decision to proceed with full data collection. However, upon reviewing the complete dataset, several issues were identified within the QMT dataset, primarily related to missing or incorrectly assigned labels.

As illustrated in Figure 8, a top view of the 3D world in QTM, the participant was positioned outside the cameras' visible range. This was repeated multiple instances since the beginning of the throw started around that point. Given that a disc golf throw involves rapid and dynamic movements, some markers may have shifted slightly during the motion. Additionally, the high velocities and occlusions caused by body positioning likely contributed to data loss or mislabeling in certain instances.

Furthermore, the calibration of the measurement volume was not done by our group. This process was done by previous researchers and due to poor communication between us, them and our lab assistant, no recalibration was done. In the end, this caused the area of the calibrated volume to be much smaller than anticipated when reviewing the data after the fact. This as the previous point, likely contributed to data loss.

One particularly important data point was the marker on the back of the right hand (RHM2). It was observed that this marker was occasionally mislabeled as the left hand (LHM2). As a result, manual verification and relabeling were required to ensure the accuracy of the dataset.

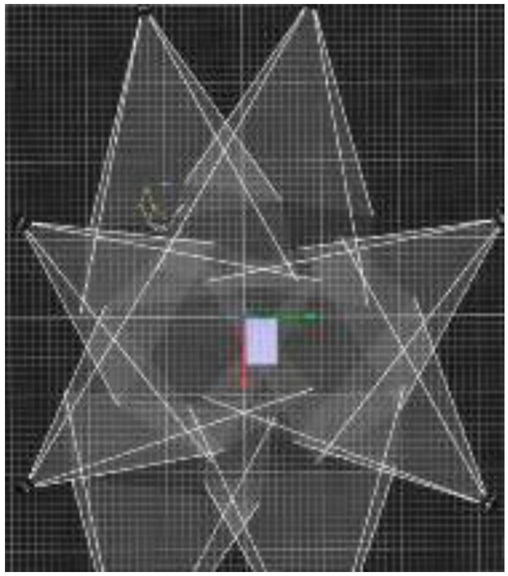

Figure 8: Top view of 3D world in QTM. Cameras field of view shows. Participant can be seen out of view of multiple cameras

7.6 Generalization of models

It is important to note that the Artificial Neural Network (ANN) and Markov Chain Monte Carlo (MCMC) models developed in this study were trained exclusively on data collected from a single participant. As a result, the performance and predictive accuracy of these models are highly specific to the unique biomechanical patterns and physiological characteristics of that individual. Applying these same models to data from other athletes would likely result in reduced accuracy, as the models have not been exposed to inter-individual variability in throwing mechanics, muscle activation patterns, or physical attributes.

To develop models that generalize well across a broader population of disc golf athletes, it is essential to collect data from a larger and more diverse group of participants. This would allow the models to learn a wider range of biomechanical strategies and adapt to different body types, throwing styles, and experience levels.

8 Conclusions

We have used motion capture, EMG, a force plate and a tech disc to collect data from a disc golf throw. The data was then preprocessed and analyzed using a slice sampling algorithm and an ANN. The models got mediocre results and the slice sampling algorithm was slightly better than the ANN. This result was not surprising since the data is very complex and that there was not a lot of data to work with. Therefore the conclusion is drawn that patterns can be found in the data to predict a disc speed but that the data we collected is not descriptive and plentiful enough to make accurate predictions. What we could have done differently is to do a better job when collecting the data. That could be to make adjustments to avoid missing data points, making adjustments to avoid or take into account fatigue, choosing more or different data points that could be more descriptive of the disc speed and collecting more samples. These results are an indication that machine learning models used on data from motion capture, fore plate and EMG sensors can predict disc velocity. This is a stepping stone towards better throwing technique analysis in disc golf.

References

[1] M. İmamoğlu. "Three-Dimensional Human Modeling Applications in Sport Sciences". In: *Journal of Sports Science Technology* 16 (2019), pp. 59–67. URL: https://files.eric.ed.gov/fulltext/ED598673.pdf.

[2] James Holcomb. "EMG Activation of the Upper Extremity During a Standing Backhand Disc Golf Throw". In: *East Tennessee State University Honors Theses* 6.1 (2023), p. 2014. URL: https://dc.etsu.edu/cgi/viewcontent.cgi?article=2014&context=honors.

[3] J.T. Nelson, R.E. Jones, and J. Hardy. "Disc Golf, a Growing Sport: Description and Epidemiology of Injuries". In: *Orthopaedic Journal of Sports Medicine* 3.6 (2015), pp. 1–6. DOI: 10.1177/2325967115590370.

[4] Travis Greenway. “A Biomechanical Analysis of the Backhand Disc Golf Drive for Distance”. Master’s thesis. Oklahoma State University, 2007.

[5] K.E.T. Giljarhus, M.T. Gooding, and J. Njærheim. “Disc Golf Trajectory Modelling Combining Computational Fluid Dynamics and Rigid Body Dynamics”. In: *Sports Engineering* 25.1 (2022), pp. 1–12. DOI: 10.1007/s12283-022-00388-9.

[6] Qualisys AB. *Qualisys Track Manager (QTM) 2024.1*. Includes AIM (Automatic Identification of Markers) functionality. 2024. URL: https://docs.qualisys.com/getting-started/content/getting_started/processing_your_data/using_aim_models/using_aim_models.htm.

[7] Travis Greenway. “A biomechanical analysis of the backhand disc golf drive for distance”. ProQuest ID: 1443026. PhD thesis. Oklahoma State University, 2007.

[8] GameProofer. *Game-changing smart tag and app for disc golf*. Online. Accessed: 2025-02-25. URL: https://www.gameproofer.com/?srsltid=AfmBOooZG_ySb3Az_Ln725Fr9ETIDOHpKbSdWj0uFy_jqgkScgqni4Rf.

[9] Biltema. *Skottmätare – Hastighetsmätare för sportaktiviteter*. Online. Accessed: 2025-03-13. URL: https://www.biltema.se/fritid/sport/tekniktraning/skottmatare-2000045487?utm_source=google&utm_medium=cpc&utm_campaign=p-shopping-LIA-low&gad_source=1&gclid=Cj0KCQjwkN--BhDkARIsAD_mnIoKMybBdT2GsGooe-rwjtV_gqNgec--0Zj6uMg6PTZUHjb9oB64K6MaAnBgEALw_wcB.

[10] Jesper Møller, Jana Novovicova, and Jan Rasmussen. “Slice sampling”. In: *The Annals of Statistics* 31.3 (2003), pp. 580–611. DOI: 10.1214/aos/1056562461. URL: https://projecteuclid.org/journals/annals-of-statistics/volume-31/issue-3/Slice-sampling/10.1214/aos/1056562461.full?tab=ArticleLink.

[11] Investopedia. *Regression*. Accessed: 2023-03-17. 2023. URL: https://www.investopedia.com/terms/r/regression.asp.

[12] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. URL: https://www.deeplearningbook.org.

[13] Noraxon. *Noraxon EMG System*. Online. Accessed: 2025-02-25. System used: Noraxon DTS 16 channels wireless EMG system (Noraxon Inc., USA). URL: https://www.noraxon.com/our-products/semg/.

[14] SENIAM. *Surface ElectroMyoGraphy for the Non-Invasive Assessment of Muscles*. Online. Accessed: 2025-02-25. URL: http://seniam.org.

[15] QTM. *Qualisys 3D Body Recorder – Track Manager*. Online. Accessed: 2025-02-25. System used: Qualisys Oqus 4 (Qualisys, Sweden) with 8 cameras. URL: https://www.qualisys.com/software/qualisys-track-manager/.

[16] M. Pauelsen. “Losing control and developing concerns: The complexities of ageing postural control and fall-related concerns”. PhD thesis. Luleå University of Technology, 2021.

[17] Kistler. *Kistler 3D Force Plate*. Online. Accessed: 2025-02-25. System used: Kistler force plate (Kistler, Winterthur, Switzerland). URL: https://www.kistler.com/INT/en/3d-force-plate/C00000090.

[18] Group 4 et al. *F7053T Project*. GitHub Repository. Accessed: March 18, 2025. 2025. URL: https://github.com/bigbjornn/F7053T_project.

[19] Wulfram Gerstner et al. *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge University Press, 2014. URL: https://neuronaldynamics.epfl.ch/index.html.

A Instructions for derived max tests

These max tests followed the same logic of time as described in 3.1.1

For the max test of the extensor carpi radialis longus, the participant tried as hard as possible to resist the counter force being a applied by another person while flexing the muscle.

Figure 9: Max test of wrist extension muscle. Counterforce being applied my an external person.

For the max test of the flexor digitorum, the participant tried to apply as much gripping strength as possible to the dynamometer. It’s worth noting that the dynamometer data was not looked at.

Figure 10: Max test of finger flexor muscle using a dynamometer.

B Tech Disc Images

Figure 11: The tech disc from the front.

Figure 12: The tech disc from behind, with the chip in the middle.

C Visualization of raw data

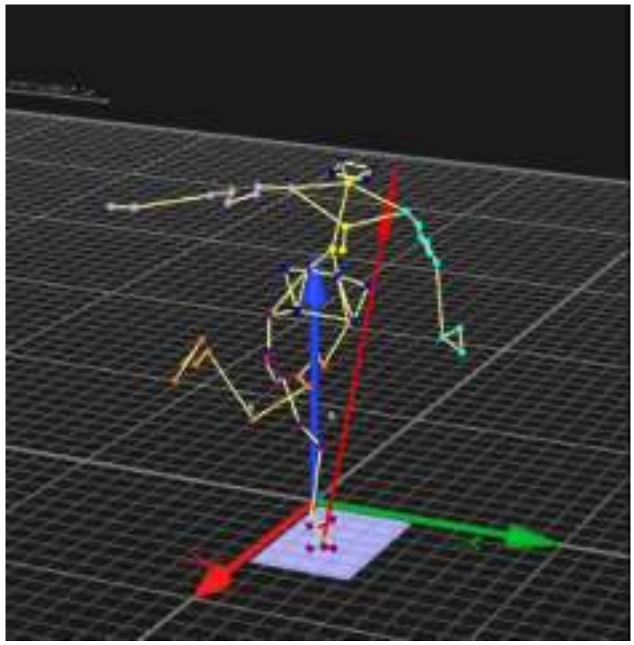

Figure 13: Picture showing data points in 3D space of participant mid throw.

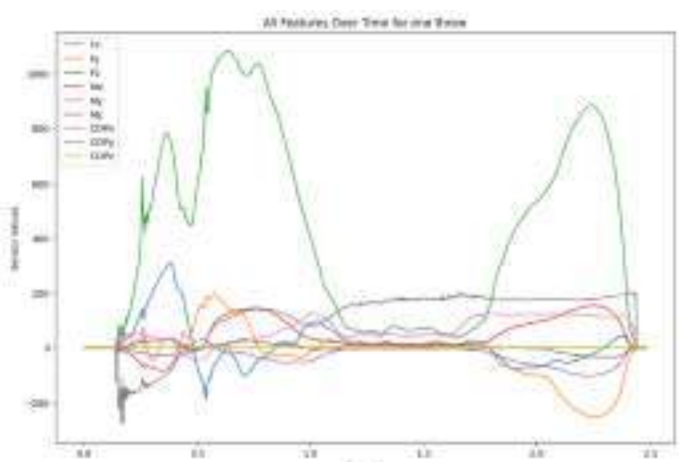

Figure 14: Picture showing raw data from the force plate.

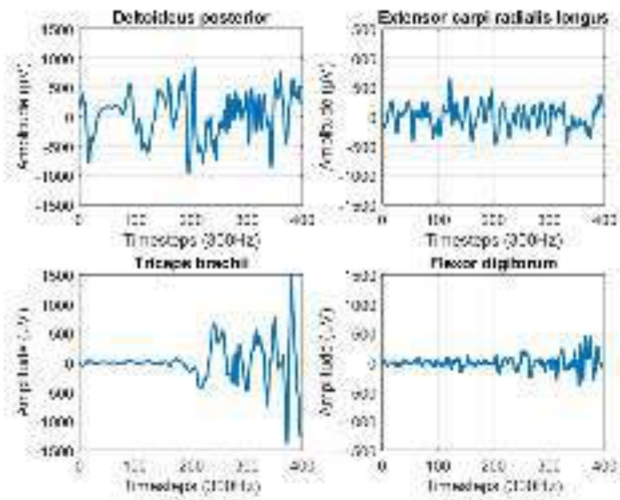

Figure 15: Picture showing raw data from the EMG measurements.

D Table of parameters

Variable	Description
EMG_1	EMG Rear Delt
EMG_2	EMG Wrist Extensor
EMG_3	EMG Triceps
EMG_4	EMG Finger Flexor
Fx	Ground Force X
Fy	Ground Force Y
Fz	Ground Force Z
Mx	Ground Torque X
My	Ground Torque Y
Mz	Ground Torque Z
$COPx$	Ground Center of pressure X
$COPy$	Ground Center of pressure Y
$COPz$	Ground Center of pressure Z
$MoCap_x$	Dot X (for 40 dots)
$MoCap_y$	Dot Y (for 40 dots)
$MoCap_z$	Dot Z (for 40 dots)
$MoCap_c$	Dot confidence (for 40 dots)
$Disc_v$	Disc velocity
$Disc_s$	Disc spin
$Disc_w$	Disc wobble
$Disc_n$	Disc nose angle
$Disc_h$	Disc hyzer angle
$Disc_l$	Disc launch angle

Table 2: Experimental Variables and Descriptions