# Personalising Crutch Geometries through Bayesian Optimisation

**Riccardo Conci**
Department of Computer Science
University of Cambridge
rc667@cantab.ac.uk

## Abstract

Crutches are optimised for stable motion, but this safety comes at the cost of comfort and speed. In this paper, I employ Gaussian Processes (GPs) and Bayesian Optimisation (BO) as hypothesis generators to discover better crutch configurations, which are validated on a physical prototype. A novel loss function is defined indicating the quality of a crutch design which combines subjective metrics (joint pain, instability and effort) and the corresponding objective ones. Finally, I (1) use this methodology to build a more stable, less effortful and less painful personalised crutch design and (2) use the knowledge built by the GP through these experiments to enhance the understanding of the physical dynamics of crutching.

Code: https://github.com/rickconci/Personalising-Crutches

## 1 Introduction and Motivation

Crutches are used worldwide to regain mobility following lower limb injuries [7], with an estimated 400,000 crutches give out across UK hospitals (source unpublished). Historically, there have been two main crutch designs: axillary and forearm (Lofstrand [5]), both of which have well-documented complications due to their design. These include brachial plexus neuropathies [9] and severe wrist pain [8]. Despite the need for improved designs across a diverse population, there has been little innovation over the past century in effectively mapping crutch geometries to personalised quantitative metrics [11].

For the substantial subpopulation of crutch users which is forced into a three-point crutching gait (or 'place and swing'), I show a novel way to rapidly discover an optimal geometry for each user. Doing so from first principles is intractable. Furthermore, the scarcity of large crutching datasets limits the application of conventional machine learning methods. However, Gaussian Processes (GPs) [10] inherently incorporate uncertainty in their predictions, making them suited to situations with limited data. Additionally, they allow for the integration of prior knowledge through kernels, and can be queried to create a better understanding of the data they model.

In this paper, I successfully answer the following questions:

1. Can BO be used to rapidly discover an optimal crutch geometry tailored towards an individual, enabling us to move away from traditional one-size-fits-all approaches?
2. Can the GP in the BO loop be interrogated to give previously unknown insights into crutching dynamics?

## 2 Methods

To address the first question, an experimental loop is set up, where the GP Bayesian Optimiser acts as the hypothesis generator, as shown in figure 1.

1. An initial crutch geometry (section 2.2) is chosen and used as an experiment.
2. Qualitative and quantitative metrics are extracted and aggregated into a loss metric (sections 2.3, 2.4, and 2.5.1).
3. These are fed to the Bayesian optimiser with the crutch parameters. The BO suggests the next set of crutch parameters to experiment (section 2.5.3).
4. The new parameters are fed into a simulator that assesses the feasibility of the suggested crutch geometry (section 2.6).
5. If the suggested geometry is feasible, the experiment is run, and the loop is repeated.
6. After the BO loop is completed, the data is queried through Sensitivity Analysis (section 3.2.2)

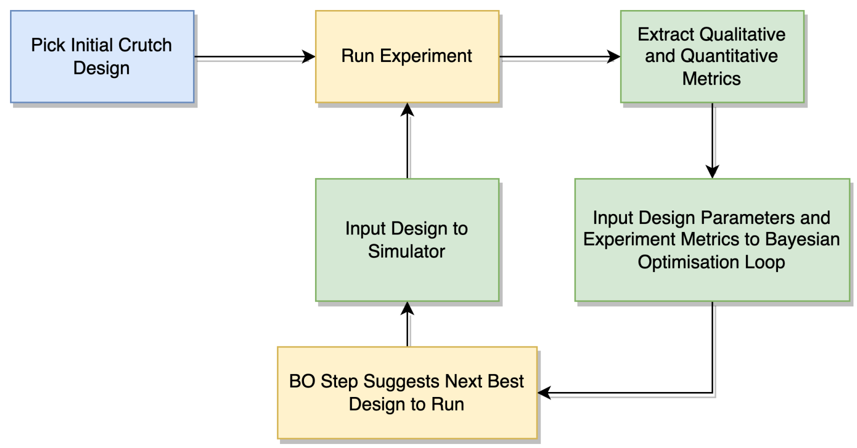



Figure 1: Experimental workflow.

## 2.1 Experimental design

Experiments were performed indoors. The user crutches 15 steps, then turns around and repeats this six times, for a total of 90 steps. Each set of 15 steps is a 'bout'. A three-point crutching technique was used with the same leg completely non-weight bearing, as is shown in Figure 3. No specific rate or step length was set. All the experiments were conducted by the author RC, to ensure consistency in the metrics gathered, especially the subjective ones (Section 2.4).

## 2.2 Crutch designs

To rapidly vary crutch parameters in our experiments, I collaborated with Dr. Christof Schwiening at the Physiology department at the University of Cambridge to design and manufacture a flexible crutch. For comparison purposes, experiments were also run on the standard NHS (Lofstrand) elbow crutch, the Smart Crutch [1], which is currently on the market as a replacement to the standard NHS design, and the OPMO crutch, the best design created by the Cambridge Engineering Department prior to this project. The flexible crutches are visible in Appendix A. A simple way to understand the crutch parameters is through figure 2. The design is characterised by 2 angles: $\alpha, \beta$ and a distance $\gamma$. $\gamma$ is the horizontal distance from the forearm contact to the vertical strut, $\beta$ is the angle between the forearm contact point and the hand contact point (with 0º starting at the horizontal), and $\alpha$ is the angle from a vertical line to the handle rod.

## 2.3 Data acquisition

The Polar H10 heart rate monitor uses a direct current (DC) tri-axial accelerometer to detect acceleration in the anterior-posterior ($Z$), vertical ($X$) and mediolateral ($Y$) directions. DC accelerometers are sensitive to static acceleration like gravity, such that the axis-aligned vertical ($X$) acceleration will measure -1$g$. Due to the lack of an internal gyroscope, the accelerometer data is also affected by the device's tilt. Appendix B shows a clear diagram of this.

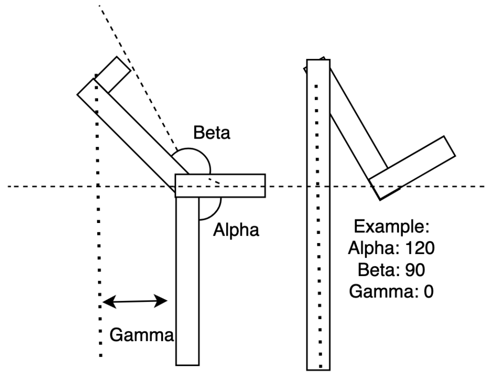



Figure 2: Alpha, Beta and Gamma across two designs.

The open-source *FingerPulseLatency* software ([12]) was utilised to pair with the H10 polar monitor and extract and save the accelerometer and heart rate data during the experiments. An interactive step detection algorithm developed by Schwiening was used to identify each crutch step through the $X + Z$ accelerometer traces. The steps detected match three-point crutching as shown in Figure 3.

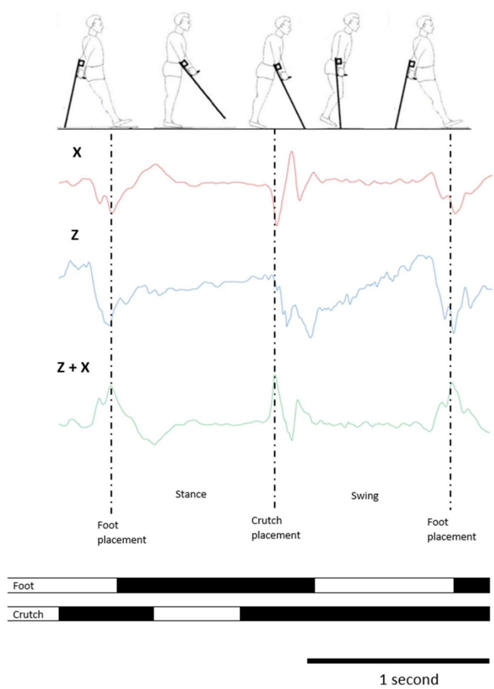



Figure 3: A crutching step where the cycle is divided into two phases as depicted by the peaks in $Z + X$. The bottom segments represent the periods when feet or crutches contact the ground (black strip). Illustration adapted from [4].

### 2.4 Metrics

Metrics for each experiment run include both subjective and objective indicators. The objective metrics are derived from the accelerometer data, as in Appendix B, and from ECG signals. The accelerometer-based metrics each describe aspects of instability during crutch use. Firstly, these include the average step frequency per bout $i$ denoted $\mathbb{E}[\text{stepfreq}_i]$ and the sum of the standard deviation of the step frequency for each bout $i$, which are scaled by a factor of 20 in the loss, giving $20\sum_i[\sigma_{\text{stepfreq}_i}]$.

The $X$ , $Y$ and $Z$ acceleration metrics are then used to extract metrics $\Delta(X+Z)$, $\Delta Y$ and $|Y|$. The metric $\Delta(X+Z)$ is calculated by summing the $X+Z$ acceleration in the last two bouts, then subtracting the summed $X+Z$ acceleration values in the first two bouts. This $\Delta$ shows if the $X+Z$ acceleration stays constant or increases (suggesting greater user confidence), or decreases (suggesting gradual instability increase or fatigue). The $X+Z$ are divided by their absolute mean to normalise.

$\Delta(Y)$ is calculated the same way as $\Delta(X+Z)$. $Y$ is specifically the sideways motion, so a change should be more sensitive to instability than $\Delta(X+Z)$. Finally, $|Y|$ is extracted to mitigate that $\Delta(Y)$ only looks at relative values, not absolute. $\Delta(Y)$ is the normalised sum of the $Y$ acceleration across all bouts. Finally, as shown in table 1, to align the magnitude of the metrics, $\Delta Y, \Delta(X+Z), |Y|$ and $\Delta$HR are scaled by $100, 100, 3000$ and $10$ respectively, so that they carry similar contributions to the loss function. The motivation behind this processing is further described in Section 2.5.1. Finally, Heart Rate (HR) change $\Delta$HR between the start and end of the crutching experiment is also considered.

Table 1: Metrics

| **Interpretation** | **Subjective metrics** | **Objective metrics** | **Weighting (for objective)** |
|---|---|---|---|
| Instability | User instability score | Total step frequency variation | 20 |
| | | Average Step Frequency | -1 |
| | | Change in Y Acceleration | 1/300 |
| | | Change in X & Z Acceleration | 1/100 |
| | | Total Y Acceleration | 1/100 |
| Effort | User effort score | Heart rate change | 1/5 |
| Pain | User pain score | | |

### 2.5 Modelling

A Gaussian Process (Section 2.5.2) is used to model an inaccessible loss function (Section 2.5.1) that informs the quality of a given crutch configuration. Because the loss function is inaccessible analytically, the experimental loop described in Section 2.1 is run to empirically find the loss associated with the suggested configuration. All the implementations are done with GPy [3] and GPyOpt [2].

#### 2.5.1 Loss function

This novel loss function $\mathcal{L}$ formalises what a 'good' crutch is in a quantifiable and precise way. It takes a crutch configuration consisting of 3 angles $\alpha, \beta, \gamma \in \mathbb{R}$ and outputs the weighted sum of the metrics described in Section 2.4 corresponding to that particular design.

$$\begin{aligned}\mathcal{L} = \text{JointPain} + \text{Instability} + \text{Effort} + \frac{1}{100}\Delta(X+Z) + \frac{1}{100}\Delta Y + \frac{1}{3000}|Y| + \frac{1}{10}\Delta\text{HR} \\ - \mathbb{E}[\text{stepfreq}_i] + 20\sum[\sigma_{\text{stepfreq}_i}]\end{aligned}$$

Combining subjective and objective measures in the function is necessary for a well-rounded representation of the problem and a common practice for complex modelling (e.g. [13, 6]). The subjective metrics make the intrinsic variability of these across users explicit. Combining objective and subjective at with an equal weighting is important, as there is not obvious higher fidelity. Whilst objective metrics are more transferable, the person's subjective impression is just as important.

#### 2.5.2 Gaussian Process Kernel

The Gaussian Process kernel allows us to input prior knowledge about the system to influence the posterior in small data settings. As such, the chosen kernel $K$ is the product of a Matern $\frac{5}{2}$ kernel, denoted $M$, and a white noise kernel $W$ that acts only on the $\beta$ angle with variance $\sigma_W = 0.2$, so that $K := MW$.

The first kernel, $M$, describes a prior assumption that similar crutch geometries will share a similar loss and that this loss will not be smooth. The kernel's *variance* is set to 1 to allow for variability inherent in the data and the measurements. As a prior, the kernel's *lengthscale* $l$ is also set to 3. When two angle configurations $x$ and $y$ are $l = 3$ apart, i.e. $||x - y||^2 = 3$, their degree of correlation will be $\sim 0.52$, giving it a intuitively correct correlation. However, as part of the GPy package, hyperparameters can be adjusted to best fit the data. Through this, we let the lengthscale and variance be learned during the BO loop.

$M$ is multiplied by $W$ acting only on $\beta$ to take into account the fact that the physical prototype does not allow $\beta$ to be fixed with certainty. This variability comes from the fact that $\beta$ is the angle from the hand location on the rod, which often varies during crutching. Hence, the variance of $\beta$ is scaled by $\sigma_W = 1.1$ to increase its inherent prior noise.

#### 2.5.3 Bayesian Optimisation loop

The BO is employed as a hypothesis generator for better crutch geometries. Our exploration domain for sampling has the following bounds: $\alpha$ from 70º to 120º, $\beta$ from 90º to 145º and $\gamma$ from 0cm to 30cm. A further constraint that $\alpha + \beta \geq 190$ is added, as shorter combinations of these angles is physically unusable. The acquisition function is set to *expected improvement* (EI), as it is a more exploitative strategy. This helps find an optimal design rapidly by refining promising designs found in the geometry space, rather than extensively exploring new and potentially less viable regions of the hypothesis space.

To initially guide our BO loop, 3 selected crutch designs are used as starting points. These 3 configurations correspond to the NHS crutches, the SmartCrutch, and the Opmo crutches. In scenarios where the BO-proposed geometries are not physically possible with the flexible crutch design, they are adjusted to the closest viable geometry and used run as an experiment.

### 2.6 Crutch Simulator

**SolveSpace**: A simulated model of the flexible crutch design was built using the parametric 3D modelling simulator, SolveSpace. For each suggested geometry by the BO loop, the parameters were checked on SolveSpace for physical feasibility. The simulated crutch is visible in Appendix C.

## 3 Results & Discussion

### 3.1 Bayesian Optimisation to discover better designs

Table 2: Bayesian Optimisation loop for crutch designs .

| Crutch | Alpha (º) | Beta (º) | Gamma (cm) | Combined Subjective Loss | Combined Objective Loss | Total combined loss |
|---|---|---|---|---|---|---|
| NHS | 100 | 95 | 3 | 10 | 36.8 | 46.8 |
| Smart | 140 | 140 | 0 | 11 | 43.7 | 54.7 |
| Opmo | 90 | 140 | 30 | 6 | 37.1 | 43.1 |
| BO1 | 110 | 100 | 12 | 11 | 45.6 | 56.6 |
| BO2 | 75 | 125 | 3 | 3 | 29.2 | 32.2 |
| BO3 | 80 | 120 | 25 | 9 | 37.4 | 46.4 |

Results from the BO loop are shown in Table 2 and in Figure 4. The top three rows show the crutch experiments that were given to the model as initial points. Remarkably, by the second BO-suggested geometry, the loss is 30% better than the standard elbow crutch. Then to explore the space it asks for a design that has a similar $\alpha$ and $\beta$, but a different $\gamma$.

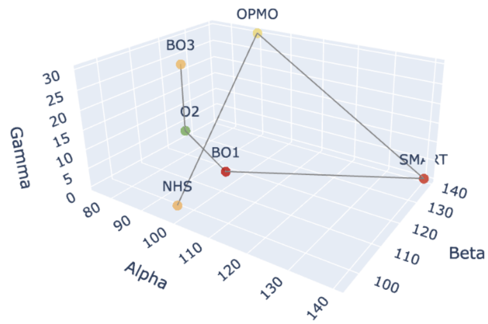



Figure 4: BO loop over the 3 dimensional space. Red is a higher loss, Green is a lower loss.

### 3.2 Knowledge extraction from data and models

Over the experimental period, a total of 12 different geometries were tested. These included 2 runs from the NHS crutches to give the GP a sense of the intrinsic variance of the loss for each geometry. To best extract the knowledge from this data, all experiments were aggregated and used for the following three sections, reviewing the correlation across metrics (3.2.1), sensitivity analysis (3.2.2) and finally through GP predictions across various hyperparameters (Section 3.2.3).

#### 3.2.1 Correlations

A pairwise correlation across the recorded metrics was computed, with the results shown on Figure 6. These correlations show some notable findings.

The parameter $\alpha$, which controls the angle of the handle from vertical and ranges from 70 to 120 degrees, has a correlation of 0.59 with subjective effort. A down-sloping handle means immediately that $\beta$ has to be greater than 90 for the crutch to be physically possible. The wrist is in a stronger position, so even with increased pressure, it is less painful. However, pain is really explained by the combination of alpha with beta, and in this case, the higher the $\beta$, the less the subjective pain (-0.56 correlation). Subjective effort is also very strongly correlated (0.82) with the overall loss, which combines the full subjective and objective metrics. Effort, indeed, is the most encompassing of the metrics. Low effort is when a crutch is not painful, stable, and easy to use, the aim.

Correlations with $\gamma$ also show interesting results. $\gamma$ is the vertical distance in cm from the forearm contact point to the vertical strut. It has positive correlations with step frequency (0.4), its variation (0.52) and $\Delta(X + Z)$ (0.45). This can be explained by subplot b in Figure 5. When $\gamma$ is 0, the forward moment in mid-swing is as extreme as it can get. The effect is that regardless of how one lands on the crutch the user will be rapidly pushed forward, decreasing the step variability. As $\gamma$ increases, the user has less initial moment, so more control but also increased variability and the possibility to crutch faster and more smoothly, rather than 'plant the crutch and fall forward'.

$\Delta Y$ is picked to assess the variability of instability during crutching. $|Y|$ allows us to compare absolute values across crutches. Despite their similar intent, and their high correlation with subjective instability, they have a correlation of 0. The differences arise across step frequencies, variations and $\Delta(X + Z)$. With a lower $\Delta Y$, and a more stable crutch, the user can crutch quickly. Instead, the results for $|Y|$ are more challenging. $|Y|$ is strongly correlated with subjective instability (0.6). Yet it has a strong negative correlation with step frequency variation and $\Delta X + Z$, both intuitive indicators of instability. A possible explanation for this is that when a certain level of instability ($|Y|$) is reached, the user changes the crutching style to be more prudent. However, this hypothesis would need further testing.

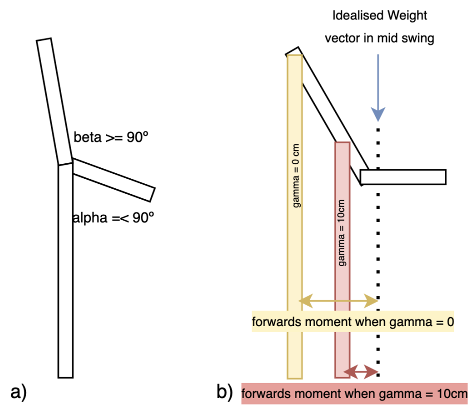



Figure 5: a) $\alpha + \beta \geq 190$ b) Decreasing $\gamma$ increases the forward moment

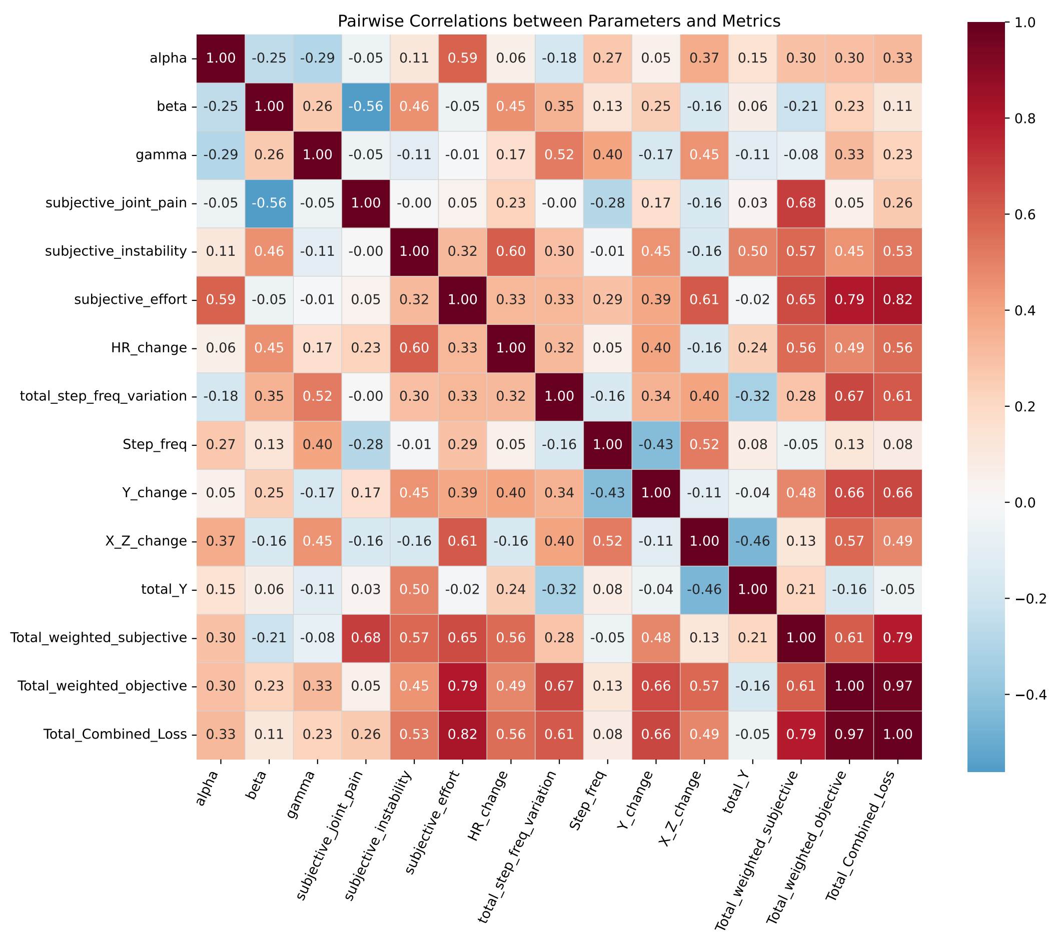



Figure 6: A pairwise correlation heatmap for each aspect and performance metric of the crutch

### 3.2.2 Sensitivity Analysis

The Sobel indices measure the effect of each parameter individually (S1), the effect of pairs of parameters (S2), and the total effect (ST) of each parameter on each individual loss (which includes interactions with all other parameters). The score can be interpreted as how much of the variance

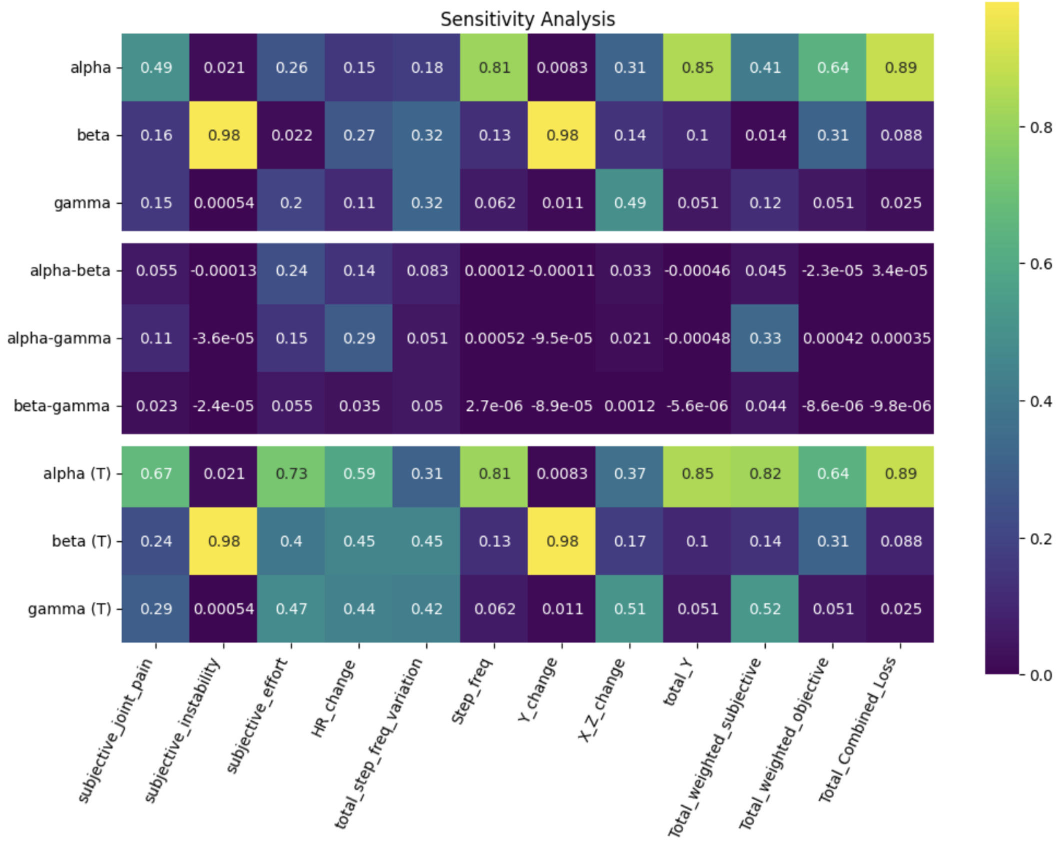



Figure 7: **Sensitivity Analysis** with Sobol Indices

of the output can be explained by the variance in the parameter, or potentially, what is the impact averaged across all the possible other parameters. As such, any negative values are due to numerical instabilities, and comparison with the correlation matrix directionality is useful to assign directionality.

Looking at the first order indices, the variance in $\alpha$ alone explains much of the variance in the step frequency, total Y and total combined loss. These all have a slight positive correlation, suggesting that $\alpha$ variance not only explains these, but increasing $\alpha$ also increases these metrics. Interestingly, alpha's strongest correlation, subjective effort, has a minimal first Sobel index, but is high instead on the total index for alpha, beta, gamma together. Changing $\alpha$ is the largest single change to the geometry: decreasing it allows the wrist to lengthen and $\beta$ to increase, and increasing $\alpha$ forces $\beta$ to also increase even more due to the geometry constraints, making $\gamma$ important as well.

Interestingly, $\beta$ has very first order high Sobel values for Y change and subjective instability. As a prior these are the two most important subjective and objective metrics for instability, and as positively correlated to each other and to $\beta$. However, the positive correlation is much smaller than the Sobel index, suggesting that even with higher $\beta$, there will be more stable crutches, and vice versa: the variance is explained more than the correlation. These results are also possibly positively skewed by the white kernel applied specifically on $\beta$.

Finally, $\gamma$ has minimal single or double effects, but explains more variance when taking into account both $\alpha$ and $\beta$ in the total effects. This matches intuition that $\gamma$'s role is strongly mediated by the values of $\alpha$ and $\beta$, rather than being a driving force to the geometry.

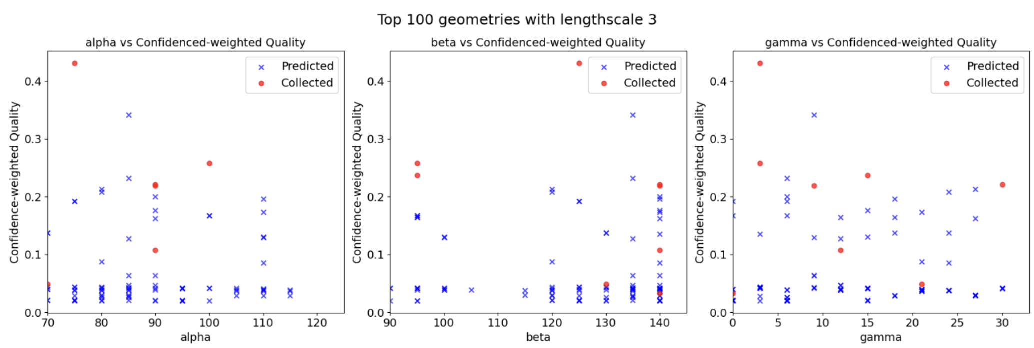
#### 3.2.3 GP predictions and anatomical interpretation

The great benefit of having a fully Bayesian model is that each prediction not only has a value but an associated posterior distribution. Given we have a search space of 1584 possible crutch geometries and only 13 known values, it is important to combine the predictions with their uncertainties to give confidence weighted predictions. We can further visualise these predictions to understand what combination of $'\alpha$ $\beta$ and $\gamma$ the model believes to be best.

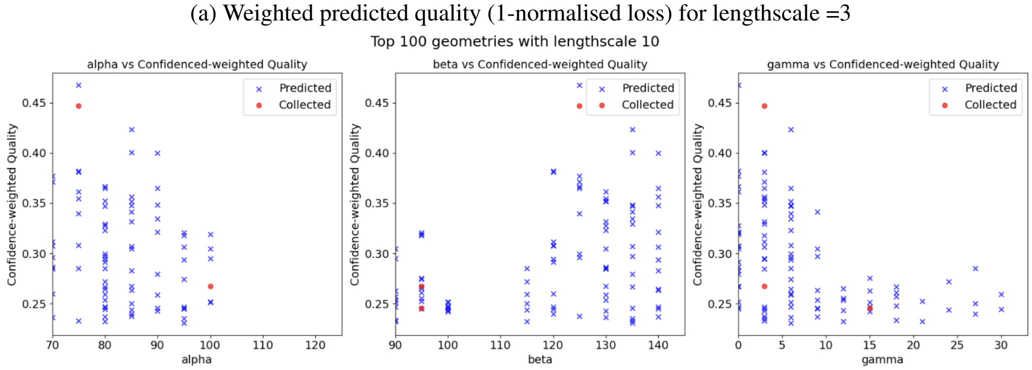
From Figure 8 we can see that as the lengthscale increases (meaning that experimental datapoints have an increasing influence on the neighbouring geometries), specific patterns from the top 100 confidence-weighted crutches arise. With a lengthscale of 10, the top geometries have an $\alpha < 100$, with a $\beta$ that peaks at 125 and a gamma that the lower the better. With a smaller lengthscale, the model is more open to alternatives, such as larger $\alpha$ values and the full range of $\gamma$.

While helpful, viewing the geometries this way makes it seem that the parameters are independent, which they are not. In Figure 9, we show the $\alpha + \beta$ angle across increasing lengthscales. With lower lengthscales, the top crutches have an $\alpha + \beta$ that peaks either at 200º or 230º. As the lengthscale increases, the peak shifts towards the middle of these, around 210º and 220º. In both cases, this shows a marked difference compared to the NHS crutch which has an $\alpha + \beta$ of 190º, likely too small. A further way to visualise these in a fully 3D way is showin in Appendix D and E.

**Anatomical interpretation:** these plots show the essence of what the GP has learned from the experiments: $\alpha$ should be less than 90º, forcing beta to be >110º. A lower $gamma$ is also beneficial. This combination encourages a lot of forward momentum, but in a position where the wrist can take the weight and the the latissimus dorsi are the main muscle group to suspend the user. This is in stark contrast to the NHS crutch, where the elbows are locked and the wrist takes all the load in an extremely weak position.




(a) Weighted predicted quality (1-normalised loss) for lengthscale =3




(b) Weighted predicted quality (1-normalised loss) for lengthscale =10

Figure 8: Quality (1-normalised loss) weighted by confidence (probability) across alpha, beta and gamma

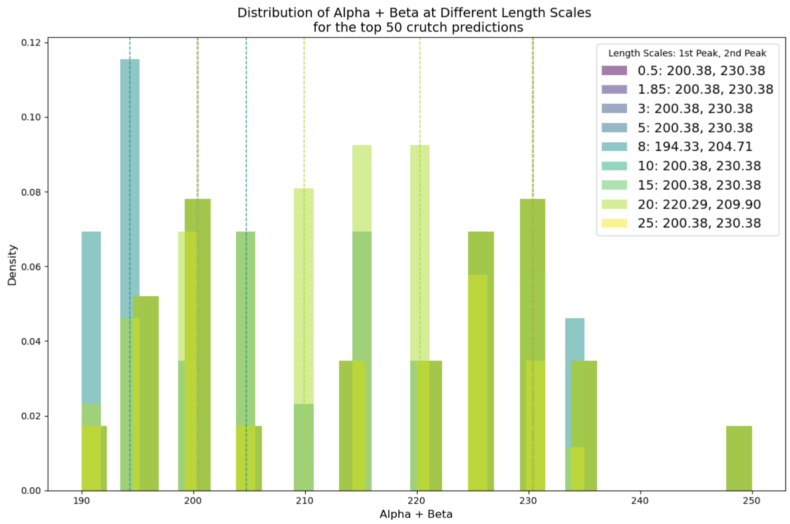



Figure 9: Histogram across lengthscales of $\alpha + \beta$ for the top 50 confidence adjusted predicted crutches.

# 4 Limitations

Despite promising results, there are certain aspects of the method to be improved. Firstly, there is uncertainty over which lengthscale and data noise should be best to use. The automatically optimised lengthscale of 1.85 by GPyOpt is helpful, however, if correct suggests that much more data is required to populate the space, as we cannot rely on a strong correlation between geometries that are relatively far away for a high quality prediction. Secondly, this paper highlights a method validated on a single user, making it difficult to interpret if these results would be valid for others.

## 4.1 Future work

The most exciting aspect of this work is the potential to transfer across people. However, as more data is gathered, transfers promise new individuals with similar characteristics to previous users to more rapidly discover their own personal crutch geometry. There are various challenges to this. Firstly, the crutching method and quality have to be such that the data platform can extract useful and relevant metrics. In particular, the quality and material of the crutch should be comparable across different users, as well as their skill in using crutches. Given the likely heterogeneity, the accelerometer extraction pipeline should be made more robust.

Secondly, subjective metrics likely do not transfer well and may need to be separated, perhaps by employing multiple GPs. Finally, while the flexible crutch we built is able to reach most of the full space of potential crutches, there are some obvious gaps (such as high $'alpha$ and low $\gamma$) that are not possible. Designing a flexible crutch that can reach this parameter space would also be helpful.

Finally, to scale this method requires prospective clinical validation. This can first be done on students, which will also help to increase the data, and finally on patients, likely from Addenbrooke's Hospital Fracture Clinic. From there, with enough high quality data, the optimal crutch could gradually be predicted from user metrics without them requiring any crutch geometry experimentation.

## References


[1] Smartcrutch forearm crutches. `https://smartcrutch.uk/`, 2024. Accessed: 2024-01-18.

[2] The GPyOpt authors. GPyOpt: A bayesian optimization framework in python. `http://github.com/SheffieldML/GPyOpt`, 2016.

[3] GPy. GPy: A gaussian process framework in python. `http://github.com/SheffieldML/GPy`, since 2012.

[4] B W Heller, M H Granat, and B J Andrews. Swing-through gait with free-knees produced by surface functional electrical stimulation. *Paraplegia*, 34(1):8–15, January 1996.

[5] J. A. R. Lofstrand. Crutch. U.S. Patent No. 2,453,632, November 1948. Patent issued November 16, 1948.

[6] Eshan D Mitra and William S Hlavacek. Bayesian inference using qualitative observations of underlying continuous variables. *Bioinformatics*, 36(10):3177–3184, May 2020.

[7] N. Papadosifos. The complications of crutches - a summary of a systematic review of the literature. *ARGnote*, 1(4), 2014.

[8] Rachele Quested, Danielle Wiltshire, Scott Sommerville, and Michael Lutz. The impact of non weight bearing: A prospective cohort study. *Injury*, 48(6):1129–1132, 2017.

[9] S Raikin and M I Froimson. Bilateral brachial plexus compressive neuropathy (crutch palsy). *J. Orthop. Trauma*, 11(2):136–138, 1997.

[10] Carl Edward Rasmussen and Christopher K I Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, November 2005.

[11] Fatemeh Rasouli and Kyle B Reed. Walking assistance using crutches: A state of the art review. *J. Biomech.*, 98:109489, January 2020.

[12] Christof Schwiening. FingerPulseLatency: A javascript based HTML program to calculate the latency between electrical activity of the heart (r wave from the polar H10 ECG) to the point of peak blood flow rate (PPG signal from the polar verity sense).

[13] Yichi Zhang, Daniel Apley, and Wei Chen. Bayesian optimization for materials design with mixed quantitative and qualitative variables, 2019.

## A Flexible Crutches

Figure 10: Flexible crutches used to run experiments

## B Accelerometer

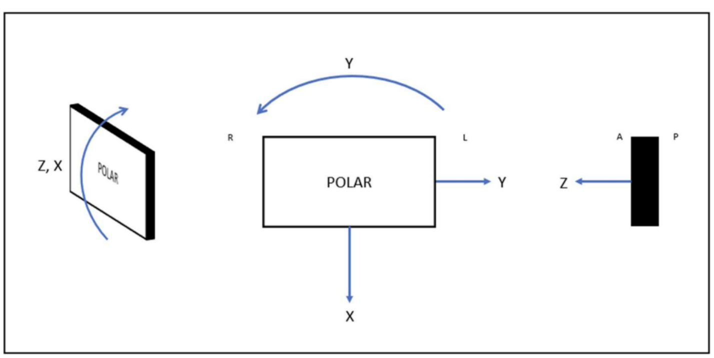



Figure 11: **Properties of Tri-axial DC accelerometer**. The X, Y and Z axes of the polar H10 respond positively to downward, right-left and forward acceleration respectively. These axes are also affected by lean as demonstrated by the curved arrows. *R=Right, L=Left, A=Anterior, P=posterior*. Image from Jonathan Bennett thesis: "Forearm crutch design: designing a platform for physiological testing" - supervised by R Conci in 2022.

## C SolveSpace simulator

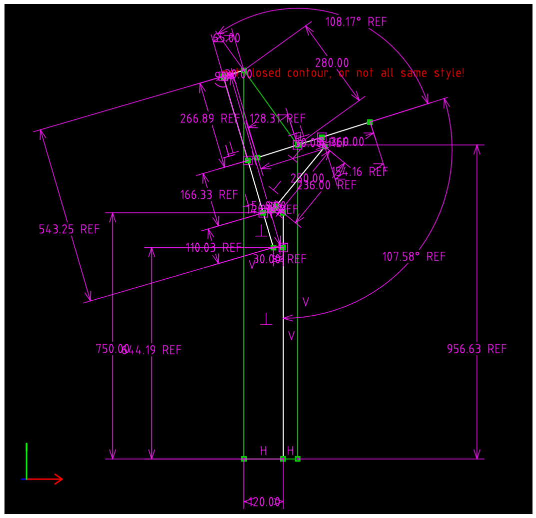


Figure 12: SolveSpace simulator to check feasibility of BO proposed design.

## D 3D confidence-weighted predictions

Figures 14 and 13 depict the parameter space $\alpha$, $\beta$, and $\gamma$ through a three-dimensional plot. The interactive version of this plot can be found in the accompanying code. The "cropped" space in the bottom right corners corresponds to the area of the parameter space representing impossible crutch configurations, as detailed in Section 2.5.2.

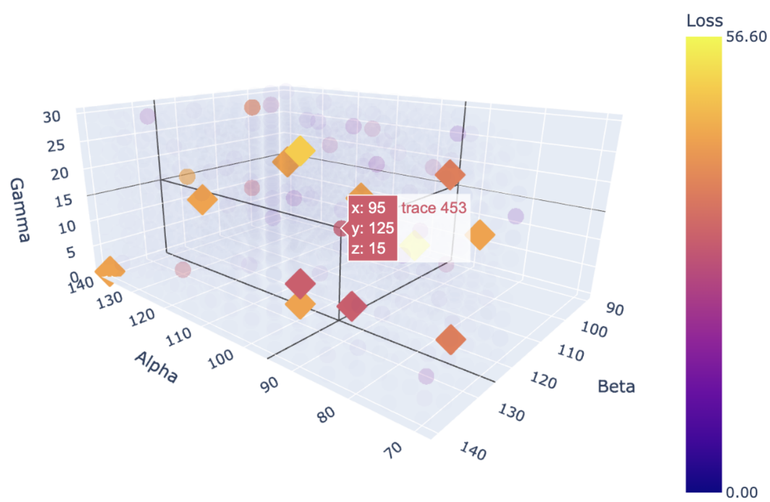



Figure 13: 3D Weighted predicted quality (1-normalised loss) for lengthscale = 3. Diamonds: experimental data. Circles: predicted geometries. The opacity is indicative of the confidence (more opaque = more confidence). The colour is indicative of the predicted loss.

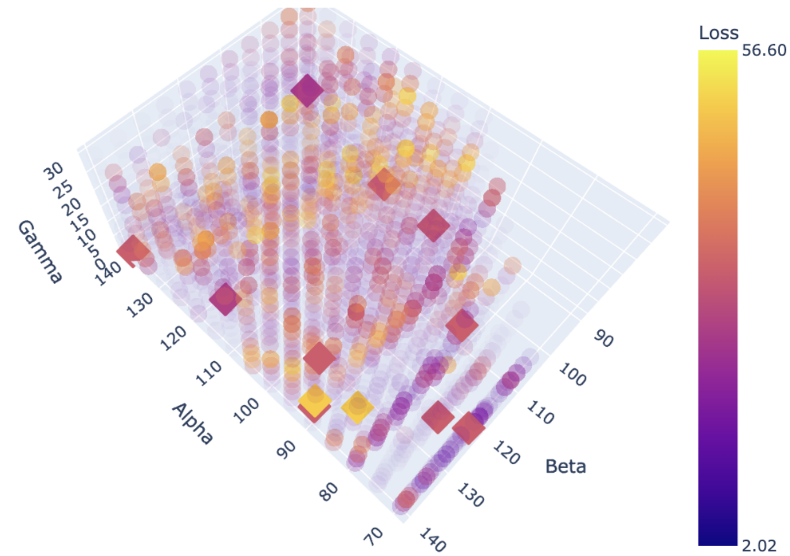
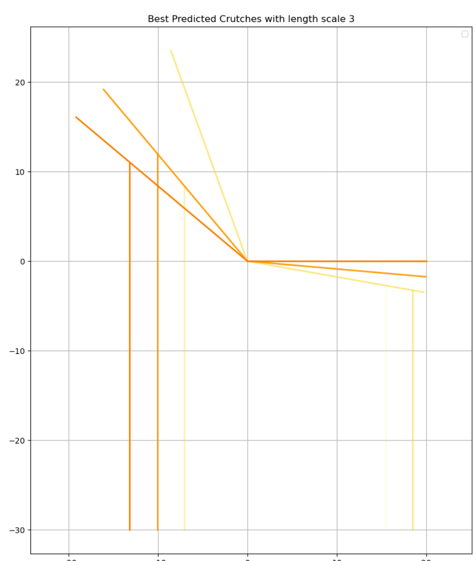
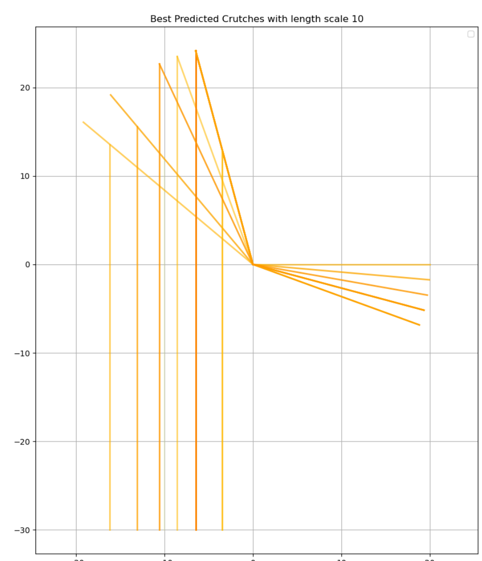



Figure 14: Weighted predicted quality (1-normalised loss) for lengthscale = 10. Diamonds: experimental data. Circles: predicted geometries. The opacity is indicative of the confidence (more opaque = more confidence). The colour is indicative of the predicted loss.

# E Crutch diagrams of top confidence-weighted predictions




(a) Visualisation of best crutches with lengthscale 3. Darker is better.




(b) Visualisation of best crutches with lengthscale 10. Darker is better.

Figure 15: Visualisation of the best crutches for lengthscales 3 and 10.