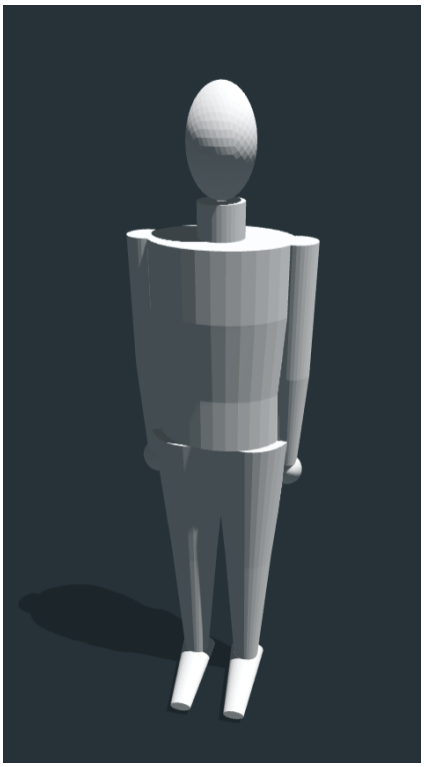
CHALMERS



Predicting Body Segment Properties

Using Height, Mass and Gender to Estimate the Mass, Center of Mass, and Moment of Inertia of Body Segments

Bachelor's thesis in Mechanics and Maritime Sciences

Augusta Fhager, Parsa Khodabakhsh,
Sanna Lingsten, Theodor Roegner Kinnmark, Elinor Schwartz

DEPARTMENT OF MECHANICS AND MARITIME SCIENCES

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2025
www.chalmers.se

Bachelor's thesis 2025

Predicting Body Segment Properties

Using Height, Mass and Gender to Estimate the Mass, Center of Mass, and Moment of Inertia of Body Segments

Augusta Fhager
Parsa Khodabakhsh
Sanna Lingsten
Theodor Roegner Kinnmark
Elinor Schwartz

CHALMERS

Department of Mechanics and Maritime Sciences
Divisions of Dynamics and Vehicle Safety
Chalmers University of Technology
Gothenburg, Sweden 2025

Predicting Body Segment Properties
Using Height, Mass and Gender to Estimate the Mass, Center of Mass, and Moment of Inertia of Body Segments
AUGUSTA FHAGER
PARSA KHODABAKHSH
SANNA LINGSTEN
THEODOR ROEGNER KINNMARK
ELINOR SCHWARTZ

Supervisor: Jobin John, Mechanics and Maritime Sciences
Supervisor: Shivesh Kumar, Mechanics and Maritime Sciences
Examiner: Håkan Johansson, Mechanics and Maritime Sciences

Bachelor's Thesis 2025
Department of Department of Mechanics and Maritime Sciences
Divisions of Dynamics and Vehicle Safety
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Visualization of a human body using predicted anthropometric measurements and estimated geometries. Constructed using URDF Standard.

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2025

Predicting Body Segment Properties
Using Height, Mass and Gender to Estimate the Mass, Center of Mass, and Moment of Inertia of Body Segments
Augusta Fhager
Parsa Khodabakhsh
Sanna Lingsten
Theodor Roegner Kinnmark
Elinor Schwartz
Department of Mechanics and Maritime Sciences
Chalmers University of Technology

Abstract

Most of the current models used to investigate the biomechanics of human body do not account for variation in the population. Creating a more inclusive model is necessary to have better biomechanical analysis tools. By using anthropometric data to predict measurements using regression models, it is possible to estimate body segments masses, body segments volumes, centers of mass, and moments of inertia. These properties can be used to make models of various bodies.

Regression models were created to estimate anthropometric measurements from height, mass, and gender. Two predictive models were developed, in parallel both with the same purpose. One model is based on the Bayesian modeling library Bambi and the other comes from the XGBoost library. Principal Component Analysis was used to investigate gender differences and to identify the most influential measurements. For the mass and inertial estimation part, the Modified Hanavan Model were used to estimate body segment mass, volume and moment of inertia. The body segment volumes were calculated using the 16 geometric shapes the Modified Hanavan geometric model consists of. By assuming uniform density the volume could be estimated and was validated by using the total body mass. Then the center of mass and the moment of inertia could be calculated using the volumes, masses and with formulas of the geometrical shapes. Together, anthropometric prediction and mass and inertial properties, gives the possibility to represent human variability in digital human modeling.

Keywords: Anthropometry, ANSUR, Bayesian Regression, Body segment parameters, PCA, Supervised learning

Sammandrag

De flesta av dagens biomekaniska modeller för att undersöka människokroppens egenskaper tar inte hänsyn till individuella variationer. Att skapa en mer inkluderande model är nödvändigt för att kunna göra bättre biomekaniska analyser. Med antropometrisk data kunde estimeringar av mått göras genom att använda regressionsmodeller. Måttestimeringarna användes sedan för att estimera kroppssegments massa, volym, tyngdpunkt samt tröghetsmoment. Dessa egenskaper kan användas för att modellera olika kroppar.

Regressionsmodeller skapades för att uppskatta antropometriska mått från längd, vikt och kön. Två olika modeller utvecklades parallellt. En modell är baserad på Bayesiansk modellering och kommer från biblioteket Bambi och en annan från biblioteket XGBoost. För att undersöka skillnader mellan könen samt identifiera viktiga mått användes en analysmetod som kallas Principal Component Analysis. För att uppskatta massa, volym och tröghetsmoment för olika kroppsdelar användes den modifierade Hanavan modellen. Vid volymberäkningar för de olika kroppsdelarna användes de 16 geometriska formerna den modifierade Hanavan modellen består av. Genom att anta en homogen densitet kan volymerna användas för att uppskatta den totala kroppsvikten. Med hjälp av volymerna, vikten och formler för de geometriska formerna beräknades tyngdpunkt och tröghetsmomentet. Slutligen kombinerades måttestimeringarna och mass- och tröghetsegenskaperna för att möjliggöra skapandet av en digital mänsklig modell.

Acknowledgements

This thesis was conducted at the department of Mechanics and Maritime Sciences, at Chalmers University of Technology, during the spring 2025.

We would like to express our sincere gratitude to our supervisors Jobin John and Shivesh Kumar for inspiring us and consistently approaching us with positivity. We are grateful for all their patience, invaluable guidance and their constant support. A special thanks to Jobin for answering our questions late into the night, and for both Jobin's and Shivesh's reactions to all Monday meeting memes.

We would also like to thank all authors of the report for good collaboration, great discussions and for all the hard work and hours poured into this thesis.

Augusta Fhager,
Parsa Khodabakhsh,
Sanna Lingsten,
Theodor Roegner Kinnmark,
Elinor Schwartz,
Gothenburg, May 2025

List of Acronyms and Terms

Below is the list of acronyms and terms used throughout this thesis, listed in alphabetical order.

Acronyms

ANSUR	Anthropometric Survey of the US Army
Bambi	Bayesian Model-Building Interface
BSM	Body Segment Mass
BSP	Body Segment Parameters
BSV	Body Segment Volume
COM	Center of Mass
MCMC	Markov Chain Monte Carlo
MOI	Moment of inertia
PC	Principal Component
PCA	Principal Component Analysis
RMSE	Root Mean Square Error
SVD	Singular Value Decomposition
URDF	Unified Robotics Description Format

Terms

Component	A categorical variable found in the data, ANSURII. This variable consists of three different groups, *Regular Army*, *Army National Guard* and *Army Reserve*.
Distal	Used to refer to lowest part of a limb or trunk, meaning the hand, foot or lower trunk.
Elbow	A distinct change of the trendline in a scree plot. The trendline changes from a steep slope to a shallow slope, where the point of change resembles an *elbow*.
Loading	The coefficients of a variable in a principal component.
Mesh	A digital geometry used to represent complex surfaces and solids.
Inertial properties	Mass, center of mass and moment of inertia of a specific body segment.
Lateral	Used to describe a position or direction away from the midline of the body. For example, the lateral side of the thigh refers to the outer side, farther from the center of the body..
Medial	Used to describe a position or direction towards the midline of the body. For example, the medial side of the thigh refers to the inside side, closest to the center of the body..
Proximal	Used to refer to the top part of a limb or trunk, meaning the shoulder, hip or upper trunk.
Shank	Refers to the lower part of the leg, typically defined from knee joint to ankle joint.

Contents

List of Figures

List of Tables

1

Introduction

How many males actually look like an average male? Can a single average body represent us all? Currently there is no widely available model for predicting the mass properties of human bodies with varying shapes and sizes. This presents a limitation in areas where accurate human body modeling is required, such as safety engineering, injury assessments and movement analysis. To contribute to research within the field of biomechanics, this project has developed a mass estimator that captures a variety of human bodies, including gender differences.

By using a publicly available data base, two regression models have been developed to predict anthropometric measurements of the human body. The measurements are then used in a geometric model to replicate the human body's shape and to calculate the mass properties.

1.1 Objective

This project aims to create a mathematical model for the human body used for predicting mass- and inertial properties of various body segments. Building the regression models on a database of anthropometric measurements enables the model to estimate these properties for a body with a given height, mass and gender.

This project will be used in a larger project where the purpose is to represent the variety of human bodies in our society (females, obese, elderly) in order to model humans of different shapes and sizes in biomechanics studies. Variety of human bodies will be captured from photographs to build models that replicate the individual in the photograph. The output from the current project will then be used to build or validate biomechanical models of the individuals represented in the photograph. The project therefore also aims to structure and sort the result in an applicable way, to make it accessible for continued work.

1.2 Background

Until recently, research and development in certain sectors of safety technology have been based on models of a standard male size [1]. Mechanical properties for a standard male has been estimated in earlier studies. A method for segmenting the body into simple geometrical shapes and determining their inertial properties was developed by Hanavan, and later modified by Hanavan and Clauser in 1975 using regression models developed by Clauser [2]. This project will use a similar version of this model as the basis for calculating the inertial properties of human body segments. In order to calculate inertial properties, regression models that provide the anthropometry measurements to the Clauser regression are needed. The regression models need to be trained on a large subset of anthropometric data. The data will be gathered from a report called 2012 Anthropometric Survey of U.S. Army personnel (ANSURII) published by the U.S. Army Natick Soldier Research, Development and Engineering Center [3]. The report is a survey that was completed in 2012 and includes approximately 4,000 men and 2,000 women. The report is one of the first publicly available of its kind and its data is a requisite for the creation of this project.

1.3 Limitations

In the project the model will be developed based on the input parameters height, mass and gender. This limitation is made due to the size of the project. In a future or continued project additional parameters, such as age and ethnicity, could be considered.

ANSURII is a set of data based on measurements of soldiers from the U.S army [3]. Therefore the data set is limited to a certain group of people, and will therefore be seen as a limitation of this project. Reasons for choosing ANSURII is the availability of this public data. The data is also fairly current and it is a large database with approximately 6000 people consisting of both male and female. This kind of data is often processed and presented in percent or means, but ANSURII has also published the raw data.

This project's geometric model and mass estimation model will be built using the Modified Hanavan odel. To this model the data from ANSURII will then be applied. Some measurements needed in the Hanavan model do not exist in ANSURII, therefore some measurements need to be approximated using the existing ones in ANSURII which can lead to a less accurate answer.

The Hanavan model is a simplified model of the human body, which is used in this project. The model is divided into 16 segments consisting of simple geometrical shapes representing; head, upper torso, middle torso, lower torso, upper arms, forearms, hands, thighs, shanks and feet. This is done in order to limit the amount of complexity when calculating mass- and inertial properties. The model will be limited to give estimations of mass, center of mass and moment of inertia.

The regression equations used to determine the mass of body segments will be based on previous research with a limited amount of data. The research was done by measuring the mass of parts from human cadavers [4]. This research is limited by the amount of test subjects available and will have an impact on the final accuracy of this project's model.

However at this time, these are the best available models for determining mass of specific body segments.

Body segments are assumed to be homogeneous, conservative systems, with a constant density and mass. Furthermore no regard is taken to the asymmetrical nature of the body's inner organs. This is done to simplify the estimation of mass distribution in each body segment.

In this project final validation such as motion capture will not be made, since it is not in our scope. The results of the mass of the segments can be validated by adding up the mass of all segments and compared it to the total body mass. To validate the center of the mass and the moment of inertia, comparison with literature could be made in order to ensure that the result is in a realistic range.

1.4 Overview

The structure of the thesis will in this section be explained to give the reader information on how this thesis is intended to be read. This project consist of several sub-studies which are presented in chapters. Chapter 2 called Anthropometry Estimation, lays out the groundwork for analyzing anthropometric data and estimating anthropometric measurements using PCA and anthropometric regression models. The third chapter, called Mass and Inertia Estimation, focuses on estimating body segment masses and moment of inertia, by using anthropometric measurements from Chapter 2. Be aware that the term *regression model(s)* is used in both 2 and 3 but refers to different models. Chapter 4, Human Body Model, explains how a human model is made by bringing together Chapter 2 and 3. Chapter 5 called Software workflow, explains how all sub-studies are connected, creating the final model, through software. The last chapter, Chapter 6, is the conclusion and summary of the findings. To facilitate the reader's understanding, Chapter 2, 3 and 4, which contain the sub-studies, each have their own sections on theory, method, result, and discussion.

Consequently, the structure of this thesis reflects the development of a biomechanical model of a human, from raw data to a functional code. In Figure 1.1 parts of the project are visualized in a workflow, where the blue box represent Chapter 2 and the orange box includes both Chapter 3 and 4.

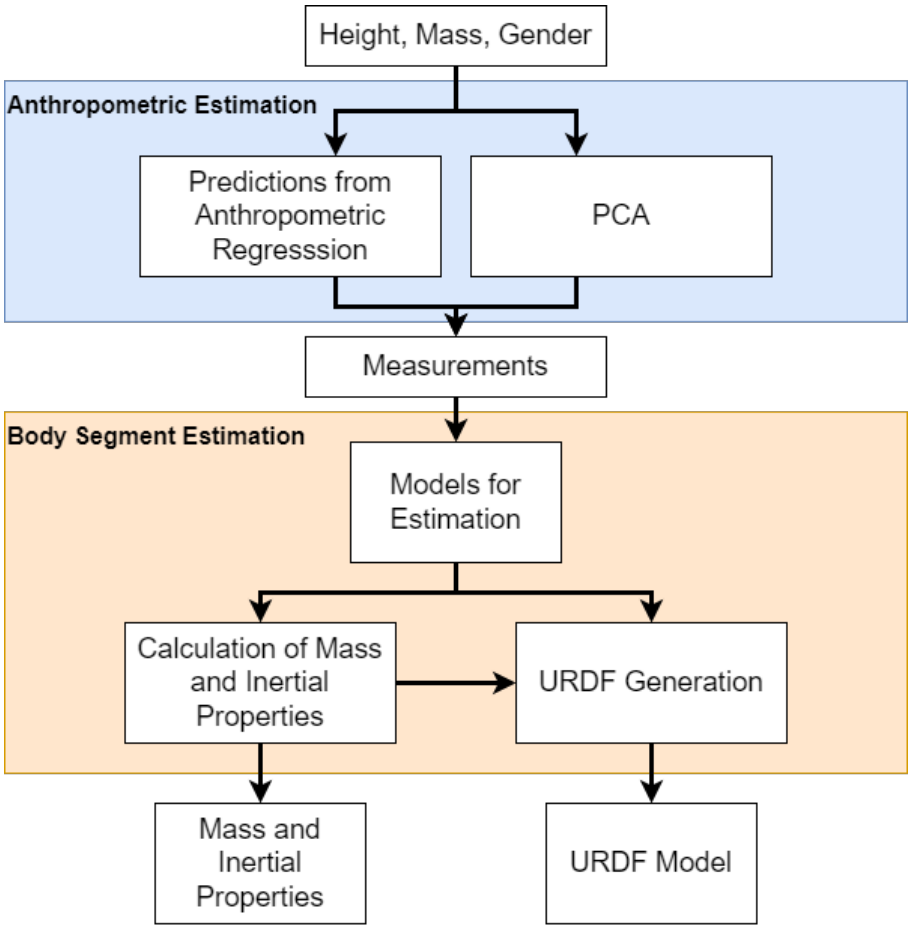


Figure 1.1: Structure of the thesis presented in sub-tasks

2

Anthropometry Estimation

In this chapter the theory, method, result and the discussion for two techniques, which are used for estimating anthropometry measurements, will be presented. The two techniques used is Principal Component Analysis (PCA) and anthropometry regression with two different models. PCA will identify the measurements with largest effect on the estimation made by the anthropometry regression models.

2.1 Theory

The purpose of the following section is to provide a theoretical background for PCA and anthropometric regression. The concept of PCA will be introduced, along with its potential applications. A Bayesian approach to linear regression will be presented as well as techniques to implement the approach. Finally, strategies on how to validate and evaluate the techniques will be described.

2.1.1 Principal Component Analysis: Capturing Structure in High-Dimensional Data

Principal component analysis (PCA) is a technique used to reduce data with many variables. The technique makes it possible to reduce multidimensional data into fewer dimensions, while still describing the variation, information, in the data without losing important information [5]. This statistical method takes the variation in the variables from the data and combines them into fewer variables, which still describe most of the variation in the data. These new variables are called principal components (PCs) [6]. The first PC, PC1, is the variable that explains most of the variation, the second, PC2, explains the second most variation, and so on.

2.1.1.1 Coefficients of the Principal Components

Each PC is defined by a linear combination of original variables and a coefficient that describes how much the variable contributes to the PC [6]. These coefficients are some times called *loadings* and can be both positive and negative depending on how they influence the PC. After finding the first PC, the condition to find PC2 is that it should be uncorrelated, orthogonal, to PC1. This condition makes the PCs measure different features in the data and include as much variance as possible.

The eigenvalues, $\boldsymbol{\lambda}$, of the PCA (matrix of PCs) describe the variance of each PC and the sum of the eigenvalues describes the total variance they describe [6]. The elements of the eigenvectors corresponding to these eigenvalues give the coefficients to the PCs. The largest eigenvalue, λ_i, indicates the largest variance and gives the information to the first PC.

An efficient way to obtain all the information needed for PCs is by deriving a singular value decomposition (SVD) [6]. The SVD consists of three parts, which are shown in figure 2.1, a matrix with left singular vectors, $\mathbf{U}$, a vector containing singular values, $\boldsymbol{\alpha}$, and another matrix with the right singular vectors, $\mathbf{V}$. The left singular vectors contain information about the the data and give, together with the singular values, the coordinates to the PCs. The right singular vectors, on the other hand, give the eigenvectors. The left and right singular matrices are orthonormal [6].

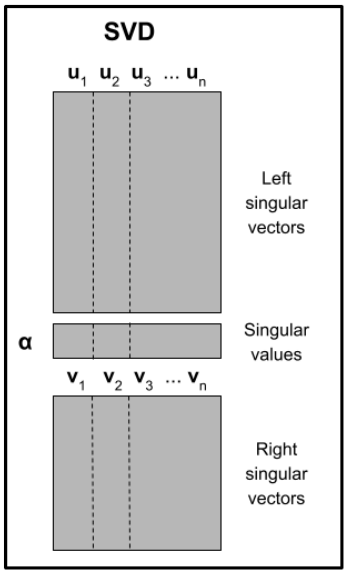


Figure 2.1: Figure showing a SVD matrix and its components. The matrix contains left singular vectors, $\mathbf{U}$, with columns $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_n$, singular values, $\boldsymbol{\alpha}$, and right singular vectors, $\mathbf{V}$, with columns $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_n$ (Modified from [6]).

2.1.1.2 Scaling Variables with Different Ranges and Units

When using PCA, the goal is to analyze the variation of different variables to find correlation. In a dataset the variables are not necessarily in the same range of magnitude or have the same unit. In this case, it can be hard to compare the variables and find the variation and obtain the PCs [6]. To make it easier to compare the different variables and their variance, the data can be scaled.

Scaling is an important step in the PCA method in case the variables have a large difference in the range of magnitude [6]. In another case where the variables are measured in the same range (for example, only positive values) with the same unit, the data does not necessarily have to be scaled and can be used as it is.

2.1.1.3 Determining the Optimal Number of PCs for Analysis

A visualization tool that is often used in combination with PCA is the scree plot [6]. This is a bar graph showing the percentage of variance explained by each PC. By analyzing the scree plot, it is possible to identify which PCs capture the majority of the data's variance and are therefore of interest for further analysis.

A common guideline on how many PCs is necessary include in further analysis is to find the *elbow* in the graph [6]. This method suggests that a distinct trend can be identified in the graph as the bar heights decrease with the number of PCs. By letting a line follow the tops of the bars, it is possible to see where the line makes a distinct bend, going from having a steep slope to a shallow slope. This bend is referred to as an *elbow*. Where the *elbow* is located gives information about where the information carried by each PCs decreases. The PCs before the *elbow* of the graph typically contain enough information on the data to describe the entire dataset [6].

2.1.1.4 Visualization and Analysis of PCs

A useful way to visualize the results of the PCA is by showing the information from the PCs in a biplot. In this kind of plot every point is an observation projected from two of the PCs, to get at 2D-plot, while the variables are shown as arrows [6].

The data points in a biplot for the two first PCs, PC1 and PC2, are represented by a combination of the two first left singular vectors (u_1, u_2) from the SVD [6]. Furthermore, the arrows are defined as pairs of coefficients from the two first right singular vectors (v_1, v_2). An example of PC1 and PC2 is shown in equations 2.1 and 2.2. In this example, A, B, ..., N represent different variables and a, b,...,n represent the loading of the variable.

$$PC1 = a_1 \cdot A + b_1 \cdot B + \ldots + n_1 \cdot N \tag{2.1}$$

$$PC2 = a_2 \cdot A + b_2 \cdot B + \ldots + n_2 \cdot N \tag{2.2}$$

In this case, the arrow representing variable A is shown as the combination of the coefficients [a_1, a_2] in the biplot. Each variable represents a biplot axis from which all points can be orthogonally projected. Due to centering of the data, all the arrows have their mean, and are therefore centered, in the origin.

The different arrows point in various directions, and the influence of the variable represented by the arrow increases in the direction of the arrow. Two arrows pointing in the same direction indicate that these variables have a strong correlation [6]. In contrast, a negative correlation is indicated by two arrows pointing in opposite directions. The length of the arrow increases with its influence on PC1 and PC2. The length is calculated as $\sqrt{a_1^2 + a_2^2}$. When a_1 and a_2 increase, the variable has a significant influence on PC1 and PC2 respectively, and the length of the arrow increases in each direction respectively.

2.1.2 Regression Models: Bayesian Approach, Techniques and Evaluation

In this section the theory behind the linear regression will be explained. Linear regression is a model that can be used to predict a dependent variable based on independent variables, such as predicting body measurements based on the individuals physical properties [7]. It is assumed that the body measurement depends on physical properties and that there therefore is a relationship. When using regression, the best fitted mathematical relationship can be found by estimating the model's intercept and slope. The model's intercept and slope are considered as the model's parameters and they are the values that define the relationship. This section will explain how the models work and what validation and evaluation methods there are for these models.

2.1.2.1 Bayesian Statistical Model: Update the Belief of Parameter Values

A Bayesian statistical model is a model based on Bayes theorem which describes how beliefs are updated after seeing new evidence. The formula for Bayes theorem can be found below in Equation 2.3 [8].

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{2.3}$$

This formula consists of four variables which are used to determine conditional probability. P(A|B) is the posterior, P(B|A) is the likelihood, P(A) is the prior and P(B) is the evidence, which is the observed data. The formula reveals that the priors together with the likelihood produce the posterior distribution. In practice, this updating process can be implemented through a fitting procedure using regression techniques.

The priors are the prior knowledge of a models parameter values, which is the intercept and slope of the models regression, before seeing the data [8]. Next, the likelihood explains how well a specific model with specific parameter values explain the observed data. The data is used, via likelihood, to update the prior knowledge and reach a new understanding of the parameter values which is called the posterior. The posterior is the updated belief of the parameter values and it can be used to make predictions. Both the prior and the posteriors are expressed in distributions which allows the model to intercept the full uncertainty estimates of each parameter.

2.1.2.2 Bambi as a Framework for Bayesian Regression Modeling

A Bayesian linear regression model can be constructed by using the **BA**yesian **M**odel-**B**uilding **I**nterface (Bambi). The model is built by passing a dataset and a model formula, which describes what model is to be fitted [9]. A general formula is shown below where x is the predictors and y is the outcome variable, the variable that is to be predicted. Depending on how the formula is written it can be decided whether a global intercept should be included or not. This can be done by writing a 0 or 1 behind the symbol tilde. See the example below:

$$y \sim 1 + x_1 + x_2 + ... + x_n \qquad\qquad y \sim 0 + x_1 + x_2 + ... + x_n \tag{2.4}$$

Adding a 1 as the first term in the formula means that a global intercept will be included. Having a global intercept means that the outcome variable y will have a non-zero value when the predictors x are zero. This is the default setting, meaning that if the term, 1 or 0, isn't added a global intercept will be included. Adding a 0 to the formula means that a global intercept should not be included. This forces the model to go through the origin, meaning that when the predictors x are zero the outcome variable y also is zero. However the term, 1 or 0, will have a different meaning if a categorical variable with n-levels is added after this term [10]. Adding a categorical variable after a 1 in the formula will result in the effect of one level being used as the global intercept. This level is called the reference level. The remaining levels n-1 will be used as dummy variables in the model. Dummy variables are binary variables which are used to represent the levels in a category, showing wether the level is present or absent. If a level is present the dummy variable will compare its level to the reference level. If however a categorical variable is added in the formula after a 0, all level of the categorical variable will be used as dummy variables. The global intercept is removed and instead a intercept for each level to the category is generated.

The formula is used to define the regression model. Below is an example of a multiple linear regression model, based on formula 2.4, shown.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ...\beta_n X_n + \varepsilon \tag{2.5}$$

Comparing formula 2.4 with the regression 2.5 it can be seen that an error term is added and that regression coefficients, β, have been added. The regression coefficients are the models parameter values, where β_0 is the intercept and the other β are the slopes [9]. These parameter values are the values that hold priors and are later updated when the model has seen the data.

When creating a bayesian linear regression model using Bambi, a prior distribution for the models parameters can be specified [9]. Since they are expressed in distributions the priors can be specified by applying a mean and a standard deviation for the input variables, the predictors. If priors are not set, Bambi will by default, loosely scale the predictors to the observed data that is used to build the model to set reasonable priors.

When creating the model the type of distribution of the outcome variable, called a family, can also be specified [9]. If no family is specified Bambi will use the gaussian family by default which is used for continuous decimal data with a normal distribution.

2.1.2.3 Tree-Based Regression Using XGBoost

Another type of model that is used for prediction in this project is called XGBoost, which stands for Extreme Gradient Boosting. It is an effective model, which makes accurate predictions with a relatively low computational time [11]. The model is based on decision trees, which can be illustrated as a tree where each branch represent a possible outcome.

XGBoost can be applied for tasks like ranking, classification, and, as in this report, regression.

2.1.2.4 Validation and Evaluation Strategies

Model validation and evaluation is a necessity when building models. The validation process serves as a tool to confirm that the model is correctly built while the evaluation process has a focus on the actual result of the model and therefore confirms the results accuracy.

The Bayesian linear regression model built using the Bambi library, has the ability to show what prior distributions are used, how the model is built and fitted, and what the posterior distribution looks like [9]. From both the priors and posterior distribution predictions can be made and shown as a distribution. These are called prior predictive distribution and posterior predictive distribution. The prior and posterior predictive distributions can be compared to understand how the model predicts before and after fitting and to understand the extent of the prediction's uncertainties. When plotting the predictive posteriors and the observed data, the distribution can be compared to ensure that the predictions resemble the observed data. This is a common method for validating the models fitting process [9].

The fitting process can be validated by making a trace plot. This process explores all possible parameter values to see which values fits the data best [9]. This process is done several times, where a round is called a chain, a Markov Chain Monte Carlo (MCMC) chain. These chains can be compared in the trace plot. In the trace plot two graphs can be found, trace line and density plot. The trace line shows which parameter values are explored, the sampling path, while the density plot shows the distributions of the parameter values, with one distribution for each round. To verify the fitting process the trace line should resemble white noise with a stable mean and the distributions in the density plot should be aligned [12]. That is to confirm that all rounds draws a similar conclusion and that the model has had a successful fitting process. In addition to this a numerical value of the chain's convergence can be obtained, so called r-hat. The r-hat value explains the variation in the chain and how well the chains are mixed. If the r-hat value is 1 the chains have converge successfully [13].

The above mentioned validation processes are important to use, however to find the most accurate model an evaluation process needs to take place. To find the optimal model the models that are to be compared need to follow the same evaluation process. One evaluation process that can be used for regression models is root mean square error (RMSE).

RMSE is a way to express the difference between the predicted data point and observed data point. The mean square error is calculated by summarizing the squared distance between the points, which is then divided by the number of data points to obtain a mean [14]. Taking the root of this value results in the RMSE, which is the average error for the model. If the distance is not squared there is a risk that some values will annihilate each other since the distance can both be positive and negative. Squaring the distance ensures that all distances are positive. The RMSE value can be used to evaluate the model itself but also to compare different models to see which model gives the least error.

2.2 Methods

In this chapter the method of the regression estimation is covered. The method will be presented in three sections. The first section includes which data is used and how it is processed. The second section explain how PCA is used and what was needed in order to use this analysis. The third section describes how the anthropometric regression models are built in order to predict and estimate anthropometric measurements from height, mass and gender. In this section the validation and evaluation of the models can also be found.

2.2.1 Data Processing

The project is built on data from the US Army anthropometric survey of 2010-2012, ANSURII [3]. The data is public and there are two files that can be found on the website of the Defense Health Agency [15]. One file contains data for approximately 2000 females and the other for approximately 4000 males. In these files 108 parameters are presented, where 93 are numerical anthropometric measurements. The rest is additional information about the subject, for instance, place of birth, age, writing preference, etc.

To be able to work effectively with the whole data set, the two files, female and male, are combined. Their gender is represented with 0 for females and 1 for males. Since all of the measurements in the data were approximately within the same range of magnitude and shared the same unit, no standardization action was needed. The units of the measurements were all millimeters, except for the mass, which was measured in kilograms.

Since the data is used for both the PCA and the regression models, the data needed to be processed in different ways for different purposes. Therefore, two data files were created, both based on the combined dataset.

To use the data for the analysis of PCA a scaling process was needed. For this part of the project, all measurements were divided by the mass and height of each person in the data. By doing this, the difference in mass and height between males and females was removed and only the differences in measurements between the genders could be observed. The scaled data were then used when applying the PCA technique.

To use the data for anthropometric regression, the data needed to be processed in a different way. To have the ability to evaluate and test the models, the data need to be split into a train and a test set. The train set consists of 80% of the data, and the test set consists of the remaining 20%. Individuals included in the train and test set are randomly assigned.

2.2.2 Investigating differences in the data

In order to develop a good anthropometric regression model, the data on which it is built on has to be investigated and analyzed. To investigate the data used in the construction of the model, PCA has been applied. This was done to examine whether there is a difference between genders, male and female, as well as to determine if there is variation in the contributions to information from different measurements. If such differences exist, this information could be valuable to consider in the development of the regression models.

2.2.2.1 Preparing the data for PCA

The first step in the analysis is to scale the data in a way that allows all variables to be compared with each other. How this was done is described in the section on data processing 2.2.1.

In order to get a better view of the measurements, twelve different body sections were created inspired by the Modified Hanavan Model [16]. Each section contains measurements relevant to the respective body region. The twelve sections were the following:

- Head
- Upper trunk
- Middle trunk
- Lower trunk
- Trunk
- Thigh
- Shank
- Foot
- Hand
- Upper arm
- Forearm
- Neck

Lastly, the measurements from all body sections were analyzed together, in a section called *All*.

In order to obtain the body sections that were desired, some needed measurements were added to the dataset. These were subtracted from other measurements in the data set.

2.2.2.2 PCA and plotting

PCA was then performed on each body section separately and the results were visualized in three different plots: a scree plot, a biplot and a plot of the top contributing loadings.

The scree plots were observed to get information about the most important PCs by finding the *elbow* in the plot. The biplots on the other hand, were used to observe a potential difference in the measurements between the genders. Finally, the plot of the loadings gives information on how much each measurement, which in this case are the variables, influences the PCs.

2.2.3 Modeling Anthropometric Relationships

In this section the method of predicting anthropometric measurements will be explained. Starting with model building, which also includes exploration of the data, followed by evaluation of the model.

2.2.3.1 Constructing the models using Bambi and XGBoost

The first step in model building is to explore the data by visualizing the correlations to understand which properties may affect the predicted measurement. The initial thought was to have height, mass, and gender as input variables. Therefore, the correlation between these parameters and the predicted measurement was explored by using a pairplot to ensure that the attributes are valid input variables, see one example in figure 2.2. Other attributes were also investigated and one became of interest, the attribute *Component*. *Component* is a categorical attribute which contain three categories, *Regular Army*, *Army National Guard* and *Army Reserve*. This attribute will however only be included in the Bayesian linear regression model.

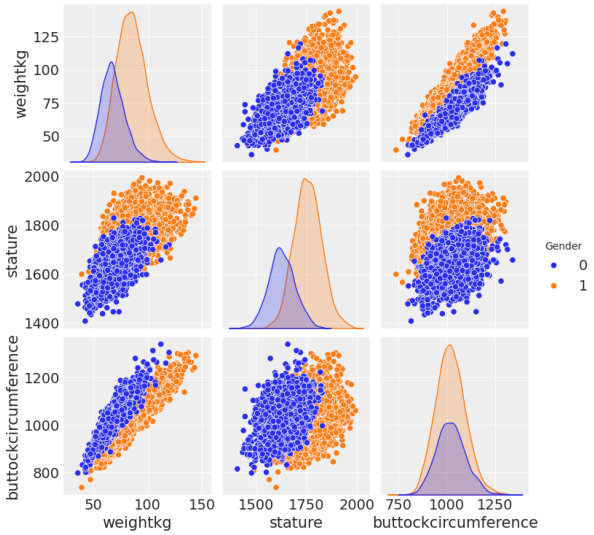


Figure 2.2: Pairplot used to investigate the correlation between a predicted measurement, *buttock circumference*, and the three input variables, height, mass and gender.

After understanding how the attributes relate to each other and which attributes are valid input variables the models could be created. Two regression models were built, one linear regression model from the XGBoost library and one Bayesian linear regression model using the Bambi library. Both of these models were built and fitted on the same train data. To build the models, information and instructions were collected from the respective library documentation.

From the theory section 2.1.2.2 it can be understood that a dataset was needed as well as a formula to be able to build the Bayesian model. The dataset that was used is the train data mentioned above and the formula can be seen below in equation 2.6, where m is the predicted measurement.

$$m \sim 0 + gender + height + mass + component \tag{2.6}$$

This formula shows that four input variables were used to predict an anthropometric measurement. However when using this model the only variables actually needed are the first three, height, mass and gender since the model is designed to use the component *Army Reserve* as default. The models primary purpose is to predict civilians. It is

assumed that civilians most likely resemble the component *Army Reserve* in terms of anthropometric properties, therefore this category was chosen as default.

In the formula a 0 can be found in front of the first input variable, gender. As mentioned in the theory section 2.1.2.2 a 0 in front of a categorical variable means that an intercept is set for each category. Since the 0 is in front of the category gender, two intercepts for each measurement is created one for males and one for females. When building the model the default family and priors were used, however it is possible to set a family and priors upon request. After having built the Bayesian model it needed to be fitted. The fitting process was conducted using 4 chains.

The XGBoost model predicted measurements by using height, mass and gender as input variables and therefore uses one global intercept. Due to time constraints in the workflow and coordination with the other subgroups of the project, *Component* was not addressed, but can however be included.

2.2.3.2 Validating and Evaluating Models

A validation process is needed to ensure that the models are constructed well, but since XGBoost does not have an explicit validation method, such as prior and posterior checks, this validation will only be able to be done for the Bayesian model. However, both models will be evaluated using RMSE.

As a first step in the validation process the posterior predictive distribution was produced. Figure 2.3 presents the distribution, which follows the observed data. The prior predictive distribution was also constructed to compare with the posterior predictive distribution, see the prior predictive plot in Figure 2.4.

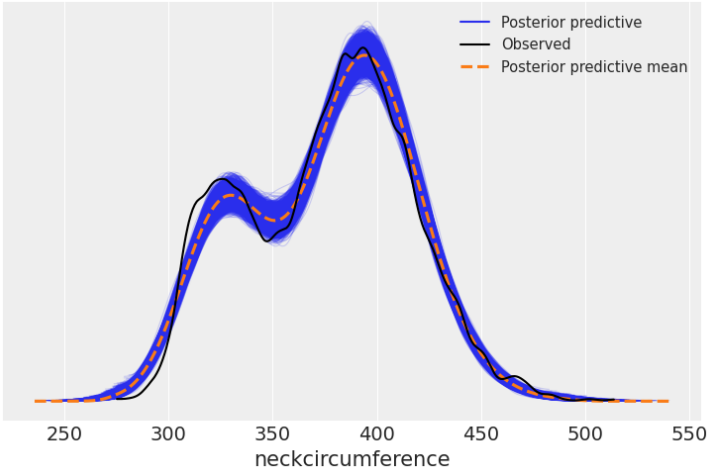


Figure 2.3: Posterior predictive plot where the blue lines are predictions of the posteriors and the orange line is a mean of all the blue lines. The black line is the observed data.

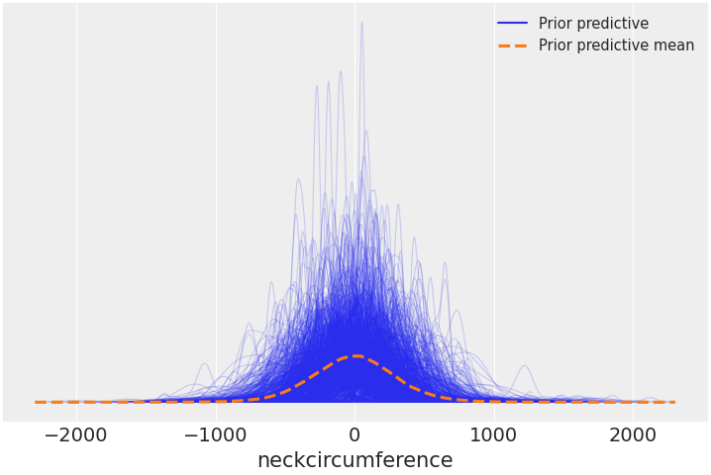


Figure 2.4: Prior predictive plot where the blue lines are predictions of the posteriors and the orange line is a mean of all the blue lines.

A trace plot was generated to follow the fitting process of the model, and can be found in Figure 2.5. The fitting was made using 4 chains, resulting in four trace lines and four corresponding distributions in the density plot. The trace lines showed random behavior similar to white noise, indicating a good mixing, and the four distributions were aligned. The r-hat value was calculated to be 1, suggesting that the chains has converged and that the fitting was successful.

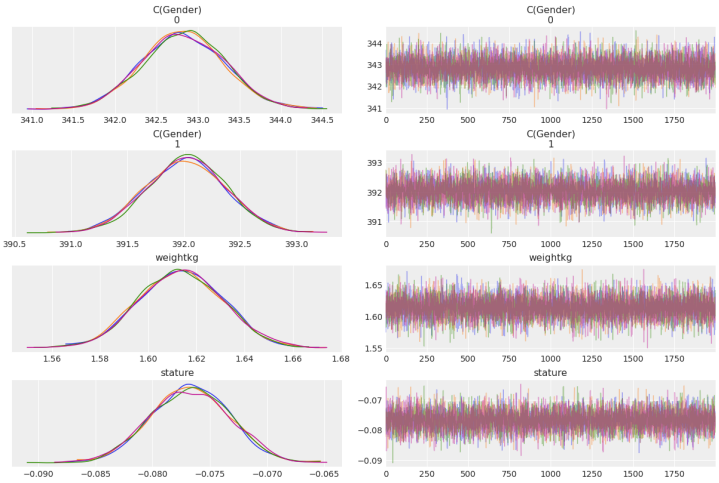


Figure 2.5: Trace plot including density plot, to the left, and trace line, to the right. Different colors correspond to different chains.

To ensure that the model's predictions are accurate, an evaluation process was needed for both regression models, the Bayesian model and XGBoost model. Since the models are built on a set of train data, this data cannot be used to evaluate the models accuracy. Therefore predictions were made from the test data based on height, mass and gender. The predictions were then compared with the actual value of the test data, by calculating the RMSE which provides an average error of the measurement. Since the same value of error can have a different impact depending on the total value of the measurement, a

percentage of the error was made. The error was therefore divided by the mean of the measurement from the data set. Due to long computational time for predictions, RMSE and the percentage were calculated for a subset of 100 individuals from the test set.

2.3 Results

The result of the anthropometric estimation are presented in this section, beginning with the PC analysis for the body sections. Followed by the regression analysis, where the accuracy of the models is presented.

2.3.1 PCA Results for Body Sections

In this part of the report, the results of the investigation of the data with PCA will be presented. The results from the *head* section and the *trunk* section as well as section *All* will be shown.

2.3.1.1 Results from the Head Section

The scree plot in Figure 2.6 shows that applying PCA on the data describing the *head* section, the first principal component, PC1, alone describes most of the data. This information is given by the *elbow* of the blue trendline that can be found at the bar representing PC2. PC1 is explaining 99% of the variation in the chosen data. For the *head* section PC1 is the PC of interest for further investigation.

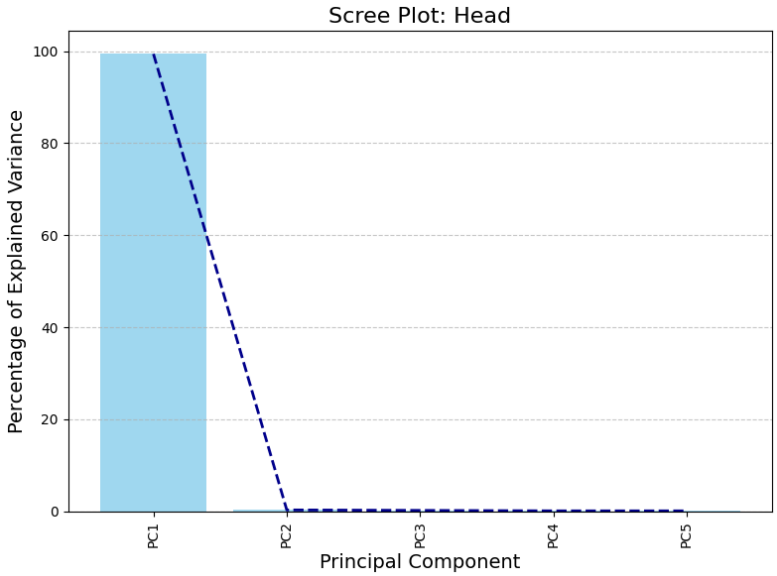


Figure 2.6: Scree plot of PCA on the head section. The *elbow* can be found at the bar representing PC2.

A biplot from PCA on the *head* section is shown in Figure 2.7. The biplot contains information on both the difference between genders and the importance of the measurements. Looking at the data points in the plot, it is possible to see a tendency for grouping between the two genders. The blue points representing the males tend to be skewed to the right in the figure, and the orange, representing the females, to the left. Additionally, the arrows in the plot show that the *head circumference* is the measurement influencing PC1 most. The loading of the *head circumference* is 0.89, which can be found in Table 2.1.

Figure 2.7: Biplot of PCA on the head section. The colors of the data points represent the genders, blue represents male and orange female.

Table 2.1: Loadings-values of PC1 for the head section.

Measurement	Loading of PC1
Head circumference	0.89
Head length	0.29
Head breadth	0.23
Bizygomatic breadth	0.20
Menton sellion length	0.17

2.3.1.2 Results from the Trunk Section

When working with PCA on the *trunk* section, PC1 again describes the majority of variation in the data, over 90%. This time in comparison to the *head* section PC2 describes about 3%. In Figure 2.8 the *elbow* can again be found at the bar representing PC2. This means that PC1 still is the PC of interest for further investigation, deciding on which measurements have the most influence on the data.

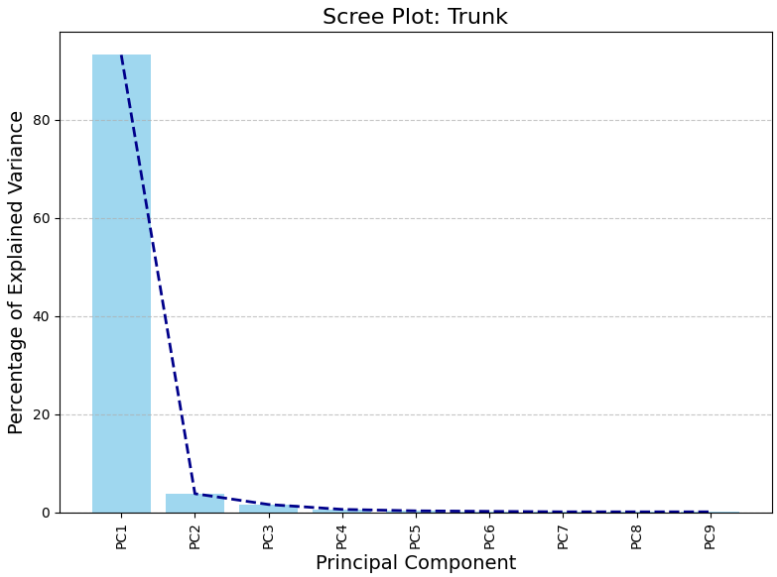


Figure 2.8: Scree plot of PCA on the trunk section. The *elbow* can be found at the bar representing PC2.

Figure 2.9 shows a biplot for PCA on the *trunk* section. The data points in the plot, representing the genders, tend to move in opposite directions in the figure. Although they are not completely separated it is possible to see a clearer grouping compared to the biplot of the *head* section in Figure 2.7.

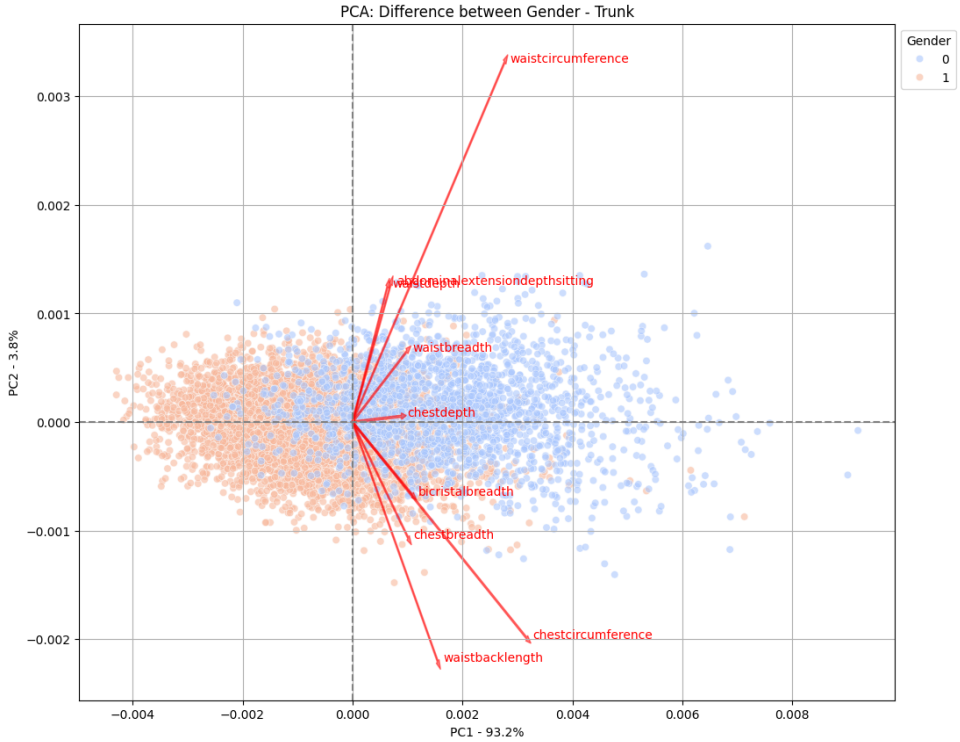


Figure 2.9: Biplot of PCA on the trunk section. The colors of the data points represent the genders, blue represents male and orange female.

The arrows in this biplot show that *chest circumference*, *waist circumference*, and *waist back length* are the measurements of greatest significance for PC1. The loadings for these measurements are 0.63 and 0.53, see Table 2.2.

Table 2.2: Loading-values of PC1 for different the trunk section.

Measurement	Loading of PC1
Chest circumference	0.63
Waist circumference	0.55
Waist back length	0.31
Bicristal breadth	0.22
Chest breadth	0.20
Waist breadth	0.20
Abdominal extension depth, sitting	0.14
Waist depth	0.13

2.3.1.3 Results from All Sections Combined

The results from the PCA of all the body sections together follow the same trend as those presented for the *head* and *trunk* sections. In Figure 2.10 the scree plot shows that PC1 still describes the majority of the variation in the data. The percentage decreases slightly, but is still above 90 %. The *elbow* can be found at the bar that represents PC2, which means that PC1 is the PC of interest for further discussion.

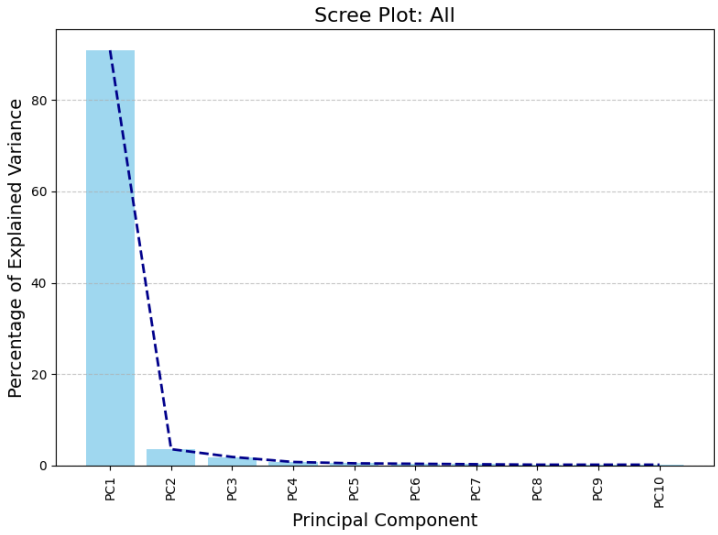


Figure 2.10: Scree plot of PCA on the measurements in all sections. The *elbow* can be found at the bar representing PC2.

Looking at the biplot showing the results of PCA of all body sections, Figure 2.11, a grouping between the colors can be seen. The two groups representing the genders are almost separated but slightly overlapping in the middle of the cluster of data points. Compared with the grouping in the biplots of the *head* and *trunk* sections, Figures 2.7 and 2.9, the separation between the blue and orange points tend to be skewed diagonally and not sideways.

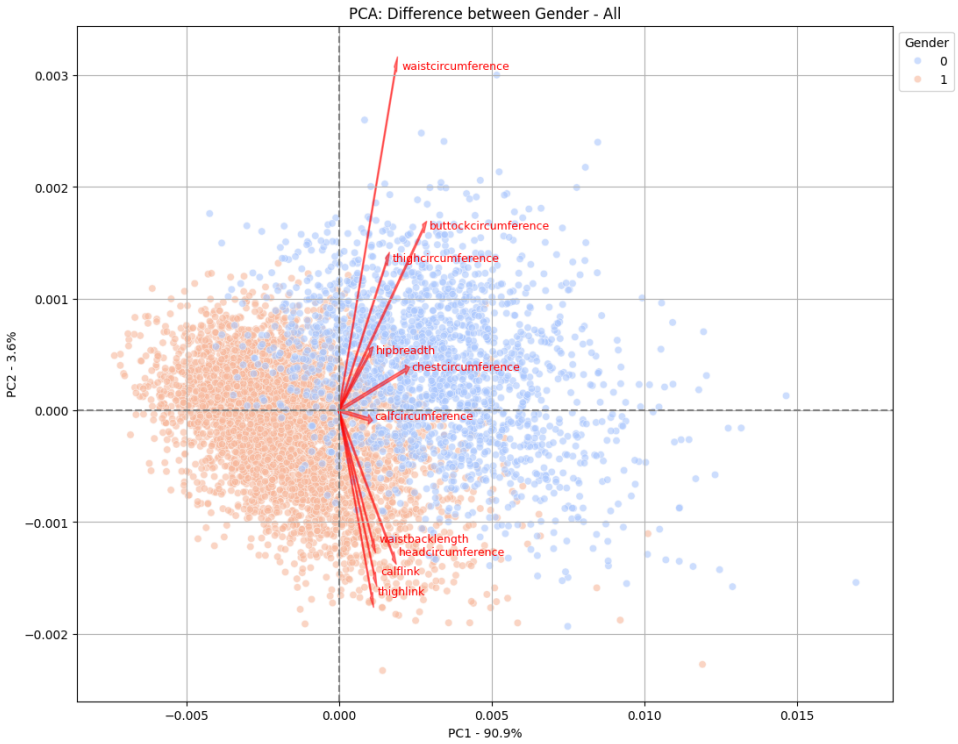


Figure 2.11: Biplot of PCA on the measurements in all sections. The colors of the data points represent the genders, blue represents male and orange female.

In the biplot, Figure 2.11, the measurements with the top 10 highest loadings for PC1 are shown as arrows. The measurements corresponding to the arrows that extend furthest to the right in the figure are the *buttock circumference*, the *chest circumference*, and the *waist circumference*. PC1 loading values for these measurements are shown in table 2.3.

Table 2.3: The top 10 largest loading-values of PC1 form PCA on the measurements from all body sections.

Measurement	**Loading of PC1**
Buttock circumference	0.46
Chest circumference	0.36
Waist circumference	0.31
Head circumference	0.30
Thigh circumference	0.26
Calf link	0.20
Waist back length	0.18
Thigh link	0.18
Hip breadth	0.17
Calf circumference	0.16

2.3.2 Anthropometric Regression

In this section the results from the model building process and regression will be presented. The result of the regression are the models that have been made and the accuracy of the models will be presented below. Each measurement has its own model which means a total of 93 models per kind of model were developed, therefore only distinctive models will be presented. After building models for each measurement, an input of height, mass and gender will result in a table of all predicted variables.

Comparable errors from the two models, the Bayesian model and XGBoost model, will be presented below. The error terms are calculated from a test set of 100 individuals.

2.3.2.1 Error terms for key measurements

The top five most influential measurements identified from the PCA study have been compiled. The Tables 2.4 and 2.5 presents the accuracy of predicting these measurements using the Bayesian model and the XGBoost model, respectively. The RMSE is an average of absolute error shown in the first row, and expressed in millimeters. The relative error is presented as a percentage in the second row.

Table 2.4: RMSE and relative error for most influential measurements when predicting with the Bayesian model.

	Buttock Circumference	Chest Circumference	Waist Circumference	Head Circumference	Thigh Circumference
RMSE (mm)	31.05	36.91	43.10	15.74	21.01
Percent (%)	3.04	3.61	4.71	2.76	3.38

Table 2.5: RMSE and relative error for most influential measurements when predicting with the XGBoost model.

	Buttock Circumference	Chest Circumference	Waist Circumference	Head Circumference	Thigh Circumference
RMSE (mm)	27.88	38.45	47.23	15.13	20.31
Percent (%)	2.73	3.76	5.16	2.65	3.26

2.3.2.2 Average error across all measurements

The mean prediction error for all 93 measurements has been calculated and is presented in Table 2.6. Both models, Bayesian and XGBoost, are presented to be able to make comparisons.

Table 2.6: Average percent error for Bayesian and XGBoost models.

	Bayesian	XGBoost
Average Percent (%)	4.16	4.19

2.3.2.3 Identifying the measurements with largest relative error

To identify the measurements with the poorest prediction performance, a limit of 7 % error was set. The measurements exceeding this, the corresponding RMSE and percentage is presented in Table 2.7 for the Bayesian model, and Table 2.8 for the XGBoost model.

Table 2.7: RMSE and relative error for measurements with highest relatively error when predicting with the Bayesian model.

	Crotch Length (Posterior Omphalion)	Ear Breadth	Ear Protrusion	Elbow Rest Height	Lateral Malleolus Height	Shoulder Length
RMSE (mm)	27.81	3.07	2.85	30.14	5.13	10.22
Percent (%)	7.92	8.74	12.80	12.51	7.37	7.04

Table 2.8: RMSE and relative error for each body measurement with highest relatively error when predicting with the XGBoost model.

	Crotch Length (Posterior Omphalion)	Ear Breadth	Ear Protrusion	Elbow Rest Height	Lateral Malleolus Height	Shoulder Length
RMSE (mm)	27.73	3.15	2.95	30.85	5.00	10.32
Percent (%)	7.89	8.97	13.27	12.81	7.19	7.11

2.4 Discussion

The following section presents an interpretation and discussion of the results together with the theory of anthropometry estimation. This section will firstly analyze the results of PCA, assessing whether gender differences can be found in the measurements and identifying which measurements carry the most information about the data from ANSURII. This will be followed by a discussion of the predictions of the measurements obtained by anthropometric regression. Finally, the section addresses limitations of the study and suggestions for future research.

2.4.1 Potential Difference between the Genders

The investigation of whether there is a difference in measurements between males and females has been carried out in several steps. Initially, the data was scaled by dividing all measurements by both height and mass for each person. Subsequently, the scaled data was analyzed with the PCA method. The analysis provides insight into which measurements differ between the genders when the effect of difference in mass and height has been removed.

The biplots showing the results of the PCA for the different sections showed a grouping between the genders. This was mainly shown in the biplot for all sections, Figure 2.11. In this figure, it is possible to see how the data points representing the different genders are parted into two groups. An interesting observation is that the groups are not parted vertically, but diagonally. Why this occurs is not fully known, but it could be because it follows the direction of the arrows of the variables that most significantly contribute to the difference between the groups.

Looking at Figure 2.11, the groups are separating in each direction of a diagonal line. This diagonal goes in a direction between the arrows for *buttock circumference* and *waist circumference*, which are the two measurements that have the greatest influence on the grouping. This could be the reason why the division appears as such.

2.4.2 The Meaning of High Loadings

A consistent trend in the results is that PC1 accounts for the largest proportion of explained information and is the principal component of main focus regardless of which body section is investigated. From the theoretical chapter, it can be understood that this means that with only PC1, almost all the information in the data for a specific body section can be described. The data which is multidimensional, can be described using only a one-dimensional combination of variables and components, loadings, which is described in PC1.

Therefore, it is the loadings with high values for PC1 that most influences the data. The results of the PCA on the different body sections all give circumferences as the measurements with the hightest loadings. From the PCA of the data for all body sections, among the top 10 measurements that most significantly influence PC1, the 5 largest are all circumferences; see Table 2.3.

The measurement with the highest loading of all is the *buttock circumference*. This suggests that this is the measurement, among those examined, that differs the most between genders. It does not seem entirely unreasonable that these measurements differ between males and females. It is well known that for biological reasons, such as reproductive processes, females have wider hips, which contributes to a larger circumference in the pelvic region. Similarly, the fact that the chest circumference differs between males and females could be accounted for.

Given the differences in anthropometry between males and females as seen in the result, gender differences need to be considered when developing anthropometric regression models. The models need to take the gender into account when predicting the measurements.

2.4.3 Interpreting the Models Accuracy and Finding Improvements

Looking at the table 2.4 and 2.5 from the result section, indicate that the two models have similar error percentage with varying performance between the models depending on the measurement. Therefore conclusions on which model gives the least error cannot be drawn from these tables. Table 2.6 that compiled the mean percentage of all measurements gives a comparable percentage of error. The models total percentage average were similar but in total the Bayesian model was performing 0.3 percentage points better that XGBoost. Although the Bayesian model was better in total, there are some measurements which XGBoost was more successful in predicting.

Even though the models are built differently they perform with a similar error percentage. The building differences come with different advantages meaning that the two models can be used for different purposes and still give similar results. The ability to follow the model-building process in the Bayesian model makes it easier to understand the building process and allows for adjustments and customization of the model. However the computational time is longer than the computational time for XGBoost. The XGBoost model is fast and computational efficient but it lacks the ability to specify prior distributions and does not offer the visualization of the models validation process.

Both regression models were continuously updated and improved during the building process, in this process one adjustment was made for the Bayesian model. The categorical variable *Component*, was added as an input variable and allowed the model to predict each civilian as an *Army Reserve* individual. For future studies the same change could be applied to the XGBoost model.

Including *Component*, aimed to make the result more representative of the general population. The data that was available was the ANSURII, which consists of measurements for the American army. As a result, initially the model is developed for the people reflecting army characteristics. To make the predictions closer to average population, predictions were based on the *Army Reserve*, who are not full time working in the military. Although *Army Reserve* may be in good physical shape, they may reflect the average people more in terms of lifestyle. Due to limited data, the models have not been tested on other than military data. In future studies the project could be repeated based on another data set of measurements that better represents the general population.

The models that have been developed will, given a fixed height, mass, and gender, always predict the same measurements. However, it is intuitively understood that individuals can look different despite having the same height, mass, and gender, due to variations in body mass distribution. This is something the current models do not account for.

Improvements to the models could be made by adding more predictors, similar to the inclusion of the *component* variable. For example, having age and ethnicity as input variables may influence the predictions. However, investigating age based on the ANSURII data was not successful due to all individuals being in the same age span. Measurements with significant impact discovered in the PCA could also be included to improve the predictions. For example, *buttock circumference* was the measurement with the highest loading for PC1, and including it as an input variable may improve the prediction of other measurements.

In the result another two tables for the models was presented, 2.8 and 2.7. The table shows which predictions were unsuccessful with the largest error percentage. To find these measurements the limit was set to 7%. Here the two models share common error percentages, with *ear protrusion* having the largest term of relative error and *elbow rest height* showing a slightly lower relative error. The error of percentage for these two measurements and for both models lies around 13%. These measurements are not relevant for this study since they are not used to estimate the mass and inertia. The information ANSURII provides is primarily used in designing and sizing clothing, protective equipment and work stations [3]. Therefore several of these measurements are irrelevant for this project. If irrelevant measurements were to be excluded, a change in the average relative error, shown in the result 2.3.2.2, could occur.

3

Mass- and Inertial Estimations

This chapter details the method used to determine mass and inertial properties of geometries representing human body segments. Two methods for Body Segment Mass (BSM) estimation is explored. A method for determining geometries for Body Segment Volumes (BSVs) is also developed. Center of mass (COM) and Moments of Inertia (MOI) is then derived from the BSMs and BSVs.

3.1 Theory

The purpose of the following section is to present a relevant theoretical background for understanding methods used to determine Body Segment Parameters (BSPs), as well as a review of previous research conducted in this field.

3.1.1 Foundational modeling reference

A reference for this work that was foundational is a master thesis by Shivesh Kumar[17]. Which focuses on a robotic based model of human gait over a stride. With the associated conference paper[18] building on this, extends the analysis through experimental validation and highlights symmetry dynamics in walking patterns under different device placements. These two together, were used as starting point for modeling a humanoid model.

3.1.2 Body Segment Mass Estimation

To predict BSMs, various regression equations can be applied. These equations estimate the correlation between mass, or other properties, and various anthropometric measurements for specific body segments. The segmentation of the body looks different depending on what landmarks are intended to be used, but typically includes, the head, trunk, upper arms, lower arms, hands, thighs, shanks and feet. The anthropometric measurements that are needed to derive BSMs also depend on how the regression equations are defined but generally consist of height, mass, or specific measurements from the body segments. There are numerous regression models available from literature, each derived using different methods. Be aware that the term *regression models(s)* refers to these mass estimation equations when mentioned in Chapter 3 and 4 and is not to be confused with the *regression model(s)* mention in Chapter 2.

3.1.3 Zatsiorsky and Seluyanov's Regression Equations

This method originates from a study conducted by Zatsiorsky and Seluyanov in the 1970s. The work was originally published in Russian and was later summarized and translated into English in BIOMECHANICS VII-B published in 1983 [19]. The purpose of the study was to see if inertial characteristics and mass could be determined by measuring gamma ray absorption. This method compared the intensity of gamma-ray radiation before and after passing through body segments. Since the amount of gamma-ray absorption is directly proportional to the mass of tissue, BSMs could be determined. Measurements taken of segment lengths was then later used to construct regression equations. The study was conducted on 100 male and 15 female subjects and resulted in gender-specific regression equations for estimating BSMs based on total body height and mass, among other parameters.

Using regression models derived from this method offers advantages, as volume estimation relies on precise measurements from a large number of subjects. These equations are particularly suitable when anthropometric measurements are limited for a given subject, often requiring only total body mass and height. In contrast, a weakness of using this method is that the derived mass equations are not based on actual measured segment mass, only the estimated derived from volume and density.

3.1.4 De Leva Adjustments to Zatsiorsky Model

In 1996 De Leva [20] revised and expanded on the original inertial parameters of human body segments provided by Zatsiorsky and Seluyanov, adapting them to be more consistent with joint center based reference points commonly used in biomechanical modeling. These adjustments addressed limitations in the original dataset, which had used bony landmarks that could vary in location due to joint flexion. The modified dataset provided by De Leva includes updated values for segment masses, center of mass positions, and radii of gyration for both males and females. The adjustments are particularly useful for simulations and models requiring more anatomically and kinematically accurate segment definitions. Moving on the Zatsiorsky and Seluyanov's regression equations will be refereed to as "Zatsiorsky's method".

3.1.5 Clauser et al.'s Regression Equations

Regression equations developed by Clauser et al comes from a study done in 1969 [4]. The study developed regression equations for mass estimation. Six cadavers were measured in order to derive regression equation to predict mass, the location of center of mass and principal moments of inertia. The cadavers were segmented into 14 parts and measured for certain number of anthropometric measurements. The subsequent data resulted in regression equations used to predict segment masses based on these anthropometric measurements. This was then used to calculate the center of mass and the principal moments of inertia.

An advantage of this method is that the regression equations are directly derived from actual measured segment masses. Furthermore, such equations are advantageous because they correlate segment masses with anthropometric measurements, theoretically better capturing the relationship between variations in anthropometric measurements and mass.

However, a drawback of using these equations is that they rely on a limited amount of older data.

3.1.6 The Modified Hanavan Model

The Modified Hanavan Model is a model of the human body used to estimate BSMs, BSVs, COMs, MOIs. The model was developed by Y.-H. Kwon in 1993 [16] and builds on the original work by Hanavan [21]. By refining the geometric representations of body segments and updating the mass estimation equations based on the empirical data of Clauser et al, accuracy for BSMs, BSVs and inertial properties is increased.

The model defines 16 body segments using 10 unique geometric shapes to represent the head, upper trunk, middle trunk, lower trunk, upper arm, lower arm, hand, thigh, shank and foot. See Figure 3.1.

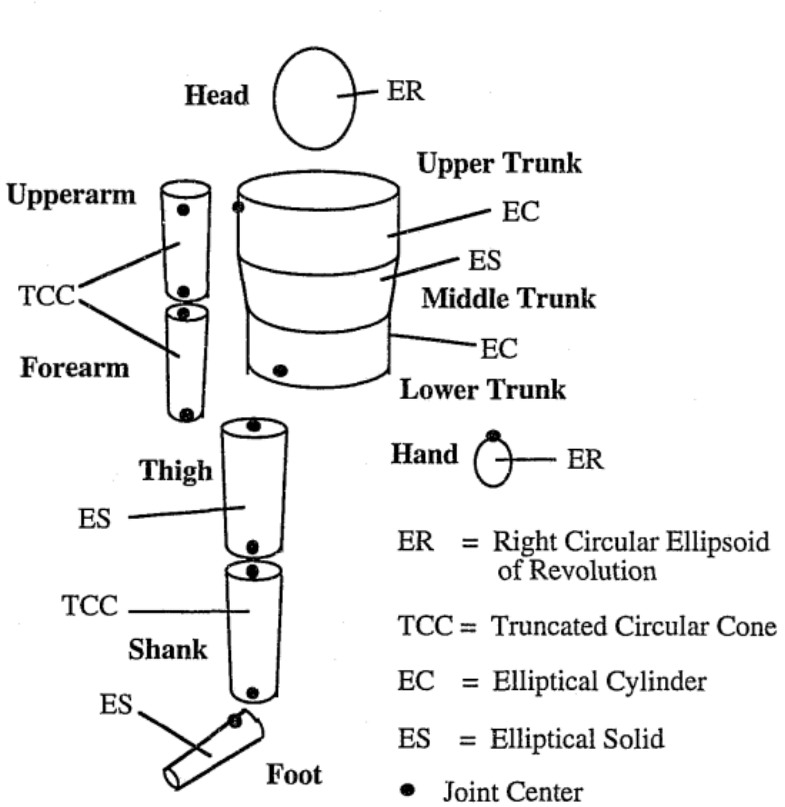


Figure 3.1: Image representing the body segmentation used in the Modified Hanavan Model created by Y.-H. Kwon [16].

r

Four primary types of geometric solids are used to model body segment geometries:

- **Right Circular Ellipsoid of Revolution** – used for the head and hands
- **Truncated Circular Cone** – used for the upper arm, lower arm, and shank
- **Elliptical Cylinder** – used for the upper and lower trunk
- **Elliptical Solid** – also called, *Frustum of Elliptical Solid*, is used for the middle trunk, thigh, and foot

A key difference between the original and modified versions of the model is the use of more anatomically realistic shapes, such as elliptical cylinders and solids in place of overly basic shapes like spheres and basic cones. These refined shapes enable more accurate estimation of segment volume and moment of inertia. Each segment is assumed to have a uniform density.

The Modified Hanavan Model incorporates segment-specific anthropometric measurements, making it suitable for estimating inertial properties for individual subjects. Mass estimation of segments is not derived directly from volume and density. Instead it uses multiple regression equations based on empirical relationships between body measurements and segment mass, as developed by Clauser et al.

The integration of both geometric volume modeling and regression-based mass estimation provides a flexible framework for biomechanical simulation that balances accuracy with practical data requirements. Moving on the Kwons method for estimating body segment masses and geometries for body segments using Hanavans modified model will be refereed to as "the Modified Hanavan Model".

3.1.7 Volume Calculation

The volume for the four different shapes used in the Modified Hanavan Model can be calculated using the well known formulas presented below. The formulas were taken from Wolfram MathWorld. [22] [23] [24]

Elliptical Cylinder:

$$V = \pi \cdot a \cdot b \cdot h \tag{3.1}$$

Where a and b are the semi-axes of the elliptical base, and h is the height.

Elliptical Frustum (truncated elliptical cone):

$$V = \frac{\pi h}{3}\left(a_{\text{bot}}b_{\text{bot}} + a_{\text{top}}b_{\text{top}} + \sqrt{a_{\text{bot}}b_{\text{bot}} \cdot a_{\text{top}}b_{\text{top}}}\right) \tag{3.2}$$

Where $a_{\text{bot}}, b_{\text{bot}}$ are the semi-axes of the bottom ellipse, and $a_{\text{top}}, b_{\text{top}}$ are those of the top ellipse.

Conical Frustum (truncated circular cone):

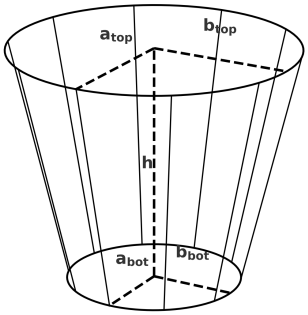
$$V = \frac{\pi h}{3}\left(r_{\text{bot}}^2 + r_{\text{top}}^2 + r_{\text{bot}} \cdot r_{\text{top}}\right) \tag{3.3}$$

Where r_{bot} and r_{top} are the bottom and top radii, respectively.

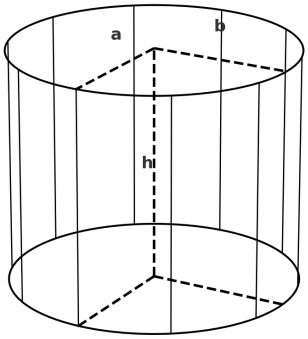
Ellipsoid:

$$V = \frac{4}{3}\pi a^2 b \tag{3.4}$$

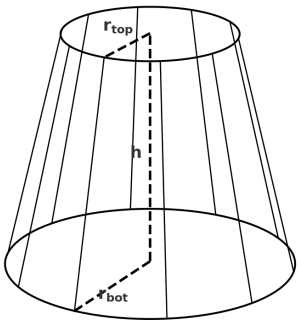
Where a is the horizontal semi-axis and b is the vertical semi-axis (e.g., for the head). Below illustrations of the four different shapes and what dimensions they have are provided.



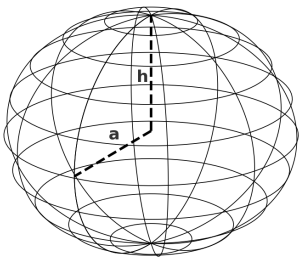
(a) Frustum of Elliptical Solid



(b) Elliptical Cylinder



(c) Truncated Circular Cone



(d) Right Circular Ellipsoid of Revolution

Figure 3.2: Image representations of the shapes and dimension used for the body segment models. (Created with Chat GPT from the prompt "from previous instruction generate illustrations of these", 2025)

3.1.8 Uniform Density Estimation

In a study done by Tan and Przekwas in 2011 [25], a computational model for articulated human body dynamics is made. In this model the human body is assumed to have a constant density of $1000 kg/m^3$.

3.1.9 Reference Values for Inertial Properties

A study was done by Gergana Stefanova Nikolova and Yuli Emilov Toshev[26] comparing different methods of estimating body segment parameters. The study included the use

of a 16-segmental mathematical model, similar to the Hanavan model[21], in order to estimate MOIs and COM positions.

3.1.10 Center of Mass Calculation

The simplified geometries used in the Modified Hanavan Model allow for the center of mass to be calculated. Each body segment was modeled as a solid of revolution, exhibiting geometric symmetry about its longitudinal (z) axis. As a result, the center of mass for each segment lies along this axis when oriented in a standard Cartesian coordinate system. For symmetrical shapes, such as *Right Circular Ellipsoid of Revolution* and *Elliptical Cylinder* the COMs will lie at the geometric center of the segments.

For the shapes such as *Truncated Circular Cone* and *Frustum of Elliptical Solid* the COM can be calculated using the following equations, which are consistent with those found in Wolfram MathWorld [27].

$$COM_z = h - \frac{h}{4} \cdot \frac{R_{\text{top}}^2 + 2R_{\text{top}}R_{\text{bot}} + 3R_{\text{bot}}^2}{R_{\text{top}}^2 + R_{\text{top}}R_{\text{bot}} + R_{\text{bot}}^2} \tag{3.5}$$

where R_{top} and R_{bot} are the effective radii of the top and bottom elliptical faces. In the special case of *Frustum of Elliptical Solid* the radii is defined as:

$$R_{\text{top}} = \sqrt{\frac{a_1^2 + b_1^2}{2}} \tag{3.6}$$

$$R_{\text{bot}} = \sqrt{\frac{a_2^2 + b_2^2}{2}} \tag{3.7}$$

with a_1, a_2 and b_1, b_2 being the semi-major and semi-minor axes of the top and bottom elliptical face, respectively.

3.1.11 Inertial Calculation

In order to calculate the MOI in the COM for certain geometries, a set of basic equations can be used. The set of equations is inertial formulas for *Right Circular Ellipsoid of Revolution*, *Elliptical Cylinder* and *Truncated Circular Cone*. The only equation used from the Modified Hanavan is the equation for *Frustum of Elliptical Solid*.

The following equations 3.8 - 3.10 is for the calculation of MOI of a *Right Circular Ellipsoid of Revolution*. Where *a, b* and *c* stand for the semi-axis in x,y,z-direction depending on the moment of inertia that is required, and M stands for the mass of the frustum. I_{xx} takes the y and z dimension (b and c) into account since the rotation in the x-axis depends on how mass is spread in y,z-plane. Which is consistent with the ones

found in Wolfram MathWorld[28]

$$I_{xx} = \frac{M}{5}(b^2 + c^2) \tag{3.8}$$
$$I_{yy} = \frac{M}{5}(a^2 + c^2) \tag{3.9}$$
$$I_{zz} = \frac{M}{5}(a^2 + b^2) \tag{3.10}$$

The equations written bellow, 3.11 and 3.12, were used for calculating MOI for a *Truncated Circular Cone*. Where M stands for the mass of the frustum, R_1 stand for radius of the smaller base and R_2 for the radius of the larger base, h stands for the hight of the frustum. Which are derived from the ones given in the MIT handbook[29]

$$I_{zz} = \frac{M}{10}(R_1^2 + R_1R_2 + R_2^2) \tag{3.11}$$
$$I_{xx} = I_{yy} = \frac{M}{20}(3R_1^2 + 3R_2^2 + 4h^2) \tag{3.12}$$

Looking at Equation 3.11 the term R_1R_2 is a blend term that takes the frustum's tapered shape into account, since the radius increases from the R_1 to R_2. The reason why the equation are dived by 10 is because it comes from the volume integral of the mass elements, $dm = \rho dV$ times r^2, where ρ is the density and r is the distance from the axis. Part $3R_1^2 + 3R_2^2$ in Equation 3.12, reflect how the mass is spread horizontally with both bases taken in consideration. The reason why the equation is divided by 20 is because of integrating $y^2 + z^2$ over the frustum's volume, using polar coordinates and linearly varying radius.

When calculating the MOI of an elliptical cylinder the following Equations 3.13 - 3.15 were used.

$$I_{zz} = 0.5 \cdot M(a^2 + b^2) \tag{3.13}$$
$$I_{xx} = 0.25 \cdot M(a^2 + \frac{h^2}{3}) \tag{3.14}$$
$$I_{yy} = 0.25 \cdot M(b^2 + \frac{h^2}{3}) \tag{3.15}$$

Where M stands for the mass of the segment, a the semi-major axis, b is semi-minor axis, and height h. These equations are derived from integrating the mass distribution of an *Right Circular Ellipsoid of Revolution* [28].

3.2 Methods

In this chapter, the method used for mass-inertial estimation is presented. The approach is divided into three sections. The first section outlines how body segment masses were determined. The second section explains how body segment volumes were computed. The third section describes how BSMs and BSVs were used to calculate the center of mass and moments of inertia for individual body segments.

3.2.1 Body Mass Estimation

Two regression-based equations were implemented to estimate body segment masses. The first set of equations used are from the Modified Hanavan Model [16]. These were chosen for their ability to estimate BSMs using subject-specific anthropometric inputs available from the ANSURII dataset. The second set of equations where based on the Zatsiorsky method [19]. This method was chosen as a compliment to the Modified Hanavan Model, as it estimated BSMs without the use of any anthropometric measurements. This makes it particularly useful as a reference model in cases where measurement interpretation or data quality may introduce uncertainty in more input-dependent methods. After both models had been implemented, their accuracy was validated by comparison to each known subjects body mass.

3.2.1.1 Implementation of the Modified Hanavan Model for BSM Estimation

The first set of regression equations implemented was based on the Modified Hanavan Model, as described by Kwon [2]. The equations that were used are shown in Table 3.1. In these equations, m denotes the estimated segment mass, M is the total body mass, and $M(x)$ refers to a specific anthropometric measurement relevant to the segment.

Table 3.1: Modified Hanavan Model regression equations for predicting segment masses (Kwon, 1993).

Segment	**Prediction Equation**
Hand	$m = 0.038 \cdot M(15) + 0.080 \cdot M(32) - 0.660$
Forearm	$m = 0.081 \cdot M + 0.052 \cdot M(16) - 1.650$
Upperarm	$m = 0.007 \cdot M + 0.092 \cdot M(18) + 0.050 \cdot M(5) - 3.101$
Foot	$m = 0.003 \cdot M + 0.048 \cdot M(22) + 0.027 \cdot M(6) - 0.869$
Shank	$m = 0.135 \cdot M(23) - 1.318$
Thigh	$m = 0.074 \cdot M + 0.138 \cdot M(25) - 4.641$
Head	$m = 0.104 \cdot M(26) + 0.015 \cdot M - 2.189$
Trunk	$m = 0.349 \cdot M + 0.423 \cdot M(41) + 0.229 \cdot M(27) - 35.460$

The regression equations require a total of 13 anthropometric measurements. Each measurement and the correlating description are presented in Table 3.2. Note: A more extensive description is provided on pages 257-260 in Kwon of landmarks where each measurement is intended to be taken.

Some of the anthropometric measurements required by the Modified Hanavan Model were not directly available in the ANSURII dataset or were defined differently. This

Table 3.2: Anthropometric measurements used in the Modified Hanavan Model for segment mass prediction.

Code	Measurement Description
M	Total body mass
$M(5)$	Elbow circumference
$M(6)$	Toe circumference
$M(15)$	Wrist circumference
$M(16)$	Forearm circumference
$M(18)$	Axillary arm circumference
$M(22)$	Ankle circumference
$M(23)$	Shank circumference
$M(25)$	Upper thigh circumference
$M(26)$	Head circumference
$M(27)$	Chest circumference
$M(32)$	Wrist width
$M(41)$	Hip-to-chin/neck length (derived)

necessitated certain approximations and interpretations based on related available measurements. Specifically, measurements *Axillary arm circumference*, *Head circumference*, *Wrist width*, *chest circumference*, and *Hip to chin* had to be replaced with the best approximation in places where needed measurements were missing from the available ANSURII dataset. These were used with a varying degree of interpretation, depending on how closely related measurements aligned with the original definitions. Additionally, there was some uncertainty regarding measurement *Forearm circumference*, as ANSURII provides data only for the arm in a flexed position, whereas the Modified Hanavan Model implicitly requires the measurement in a relaxed, extended position. To counteract the extra size of the forearm being flexed a estimated multiplying coefficient of 0.97 was used. For a full list of how each measurement was derived, see Table 3.3 bellow.

Table 3.3: Mapping between original variables $M(xx)$ and their approximated ANSUR II definitions, including derived expressions.

Original Variable	Measurement Description (Derived from ANSURII)
M	Weight
$M(5)$	Acromion radiale length
$M(6)$	(Foot breadth horizontal $\div$ 5) $\times$ π
$M(15)$	Wrist circumference
$M(16)$	Forearm circumferenceflexed $\times$ 0.97
$M(18)$	Bicepscircumferenceflexed $\times$ 0.91
$M(22)$	Anklecircumference
$M(23)$	Calfcircumference
$M(25)$	Thighcircumference
$M(26)$	Head circumference
$M(27)$	Chestcircumference
$M(32)$	Wristcircumference $\div$ π
$M(41)$	Eye height sitting + Stature - Sitting Height - Mentonsellion length - Trochanterion height

3.2.1.2 Implementation of the Zatsiorsky's Method for BSM Estimation

The regression equation developed by Zatsiorsky only required body mass and total height as anthropometric measurements to make the BSM predictions. The regression equation was structured as:

$$\text{Segment Mass} = B_0 + B_1 \cdot M + B_2 \cdot H$$

Where:

- B_0, B_1, and B_2 are the regression coefficients (see Table 3.4),
- M is the subject's total body mass (in kg),
- H is the subject's height (in cm).
- R and SD are multiple correlation coefficient and standard deviation of the estimate respectively.

As shown in the table below 3.4 the coefficients B_0, B_1, and B_2 used for each body segment is listed.

Table 3.4: Coefficients of Multiple Regression Equations for Estimating Segment Masses (Zatsiorsky & Seluyanov).

Segment	B_0	B_1	B_2	R	**SD**
Foot	-0.829	0.0077	0.0073	0.702	0.101
Shank	-1.592	0.0362	0.0121	0.872	0.219
Hip	-2.649	0.1463	0.0137	0.891	0.727
Hand	-0.1165	0.0036	0.00175	0.516	0.063
Forearm	0.3185	0.01445	-0.00114	0.787	0.178
Upper Arm	0.250	0.03012	-0.0027	0.837	0.171
Head	1.296	0.0171	0.0143	0.800	0.206
Upper Torso	8.2144	0.1862	-0.0584	0.798	1.242
Middle Torso	7.181	0.2234	-0.0663	0.742	1.020
Lower Torso	-7.498	0.0976	0.04896	0.743	1.020

Unfortunately the septate regression equations intended for females were not available. Instead the regression equations intended for males were applied for females as well. After regression equations were implemented, their accuracy was evaluated by comparing the estimated total body mass to the known body mass of the corresponding ANSURII subjects. For each subject, the sum of all estimated body segment masses was calculated and plotted against the actual body mass, separately for males and females. Additionally, the average difference between the estimated and actual total body mass was computed for each group.

3.2.2 Implementation of the Modified Hanavan Model for Geometric Estimation

A geometric model was developed to estimate the COMs and MOIs of the body segments. The purpose of the geometric model was to approximate the shape and volume of 16 body segments, which were then used to calculate each segment's COM and MOI. The method is based on the Modified Hanavan Model and utilizes a set of 16 geometric shapes (comprising 10 unique figure types) to represent the body segments. The model's accuracy was validated by estimating total body mass using an assumed uniform density for the human body and comparing the result to the known body mass of each subject in the ANSURII data.

The implementation of body BSV estimation was based on the Modified Hanavan Models geometries. As with the implementation of the segment mass regression equations, the first step involved adapting the available ANSURII measurements. The measurements that were believed to best represent the required shape dimensions were chosen. Some anatomical landmarks used to define the endpoints of measurements were adapted from the Modified Hanavan Model to better align with the available ANSURII measurements. For the thigh segment, anthropometric measurements needed to model them as *Elliptical Solids* were not available, it was instead simplified and modeled as an *Truncated Circular Cone*. A total of 27 unique anthropometric measurements were used to define the dimensions. Some of the dimensions had to be approximated, especially when information about the radii were not available. In these cases the radii was approximated by solving for the radii of a circles circumference ($\div 2\pi$). The dimensions and their corresponding description of what ANSURII measurement they were adapted from, are shown in Table 3.5.

Table 3.5: Segment dimensions used to model shapes used in the Modified Hanavan Model. All measurements are derived from ANSURII and converted to meters.

Segment	Dimension	Derived From
Upper Trunk	a_1	(Chest breadth + Interscyeii) $\div$ 4
	b_1	Chest depth $\div$ 2
	h_1	Acromial height $-$ Chest height
Middle Trunk	a_2	(Chest breadth + Interscyeii) $\div$ 4
	b_2	Chest depth $\div$ 2
	a_3	Waist breadth $\div$ 2
	b_3	Waist depth $\div$ 2
	h_2	Chest height $-$ Waist height (omphalion)
Lower Trunk	a_4	Waist breadth $\div$ 2
	b_4	Waist depth $\div$ 2
	h_3	Waist height (omphalion) $-$ Trochanterion height
Thigh	c_1	Thigh circumference
	c_2	Lower thigh circumference
	$r_{1,\text{thigh}}$	$c_1 \div 2\pi$
	$r_{2,\text{thigh}}$	$c_2 \div 2\pi$
	h_4	Trochanterion height $-$ Knee height
Shank	c_3	Lower thigh circumference
	c_4	Ankle circumference
	$r_{1,\text{shank}}$	$c_3 \div 2\pi$
	$r_{2,\text{shank}}$	$c_4 \div 2\pi$
	h_5	Knee height $-$ Lateral malleolus height
Foot	a_5	Heel-ankle circumference $\div 2\pi$
	b_5	Lateral malleolus height $\div$ 2
	a_6	Foot breadth horizontal $\div$ 2
	b_6	Lateral malleolus height $\div$ 4
	h_6	Foot length
Upper Arm	r_1	Biceps circumference (flexed) $\div 2\pi$
	r_2	Forearm circumference (flexed) $\div 2\pi$
	h_7	Shoulder to elbow length
Lower Arm	r_3	Forearm circumference (flexed) $\div 2\pi$
	r_4	Wrist circumference $\div 2\pi$
	h_8	Forearm-hand length $-$ Hand length
Hand	a_7	Hand breadth $\div$ 2
	h_9	Hand length $\div$ 4
Head	a_8	(Stature - Acromaial height) $\div$ 2
	b_{10}	Head breadth $\div$ 2

With the dimensions for each geometric shape determined, volumes could be calculated as described in theory chapter 3.1.7. Dimensions labeled, a_1, a_2, a_3,... etc. correspond to dimension labels illustrated in Figure 3.2. The largest numbered dimension (in individual bodysegments) was used to refer to the most distal dimensions in each shape.

At a later stage of the project a neck was implemented to increase accuracy of visual representation of the heads position and COM placement. A neck segment was included to increase the accuracy of where the head is estimated to be located. This is similar to what De Leva [20] did to increase positional accuracy of certain inertial properties. When the geometric model was used to calculate COM and MOI, the neck implementation was used to increase the accuracy of positions. The neck geometry was estimated as a cylinder. The head dimensions were redefined to better represent the dimensions of a more realistic smaller head. The new estimation of the neck and new dimensions for the head are represented in Table 3.6 bellow.

Table 3.6: Segment dimensions used to model the neck. Measurements are derived from ANSURII and converted to meters.

Segment	**Dimension**	**Derived From**
(New) Head	a_8	Menton-sellion length
	b_{10}	Head breadth $\div$ 2
Neck	h_{10}	Eye height sitting + Stature $-$ Sitting height $-$ Menton-sellion length $-$ Acromial height
	a_9	Neck circumference $\div 2\pi$

These additions was however ignored when performing the mass estimation calculations. To keep the BSM estimations as intended by the Modified Hanavan Model, the larger head defined from shoulder to the top of the head was still estimated. The estimated mass for the larger head was then dissipated according to the amount of volume the head and neck represented separately. A visual representation of the different options used are shown in Figures 3.3 and 3.4 bellow. The figures were created using an adaption of the tool later created in chapter 4.2.1.

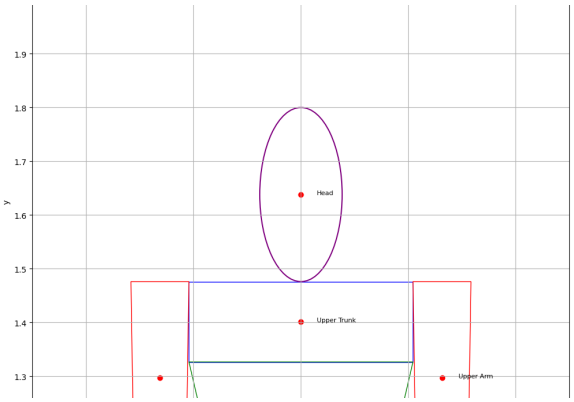


Figure 3.3: "Larger head" defined as the Modified Hanavan Model intended, used for BSM estimation.

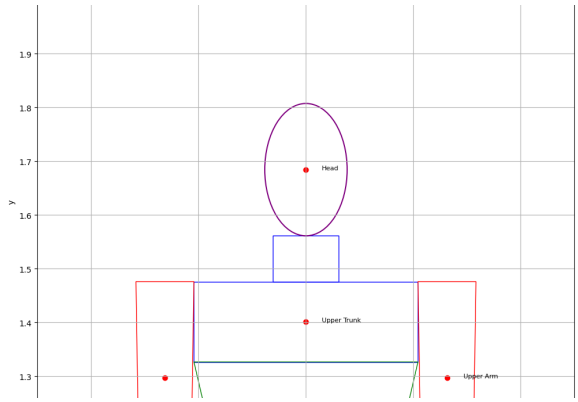


Figure 3.4: "Smaller head" used for BSV, total height, COM and MOI estimations.

To be able to validate the accuracy of the estimated volume and the accuracy of the dimensions representing vertical height, two methods were used. First, the estimated volume was multiplied with a uniform density of $1000 kg/m^3$. The specific density was chosen because it is the same as the density used by Tam and Prezekwas, described here 3.1.8, and is known to somewhat accurately reflect an assumed uniform density for the entire human body. Secondly a method for determining the accuracy total height estimation was employed. All dimensions representing the segments vertical heights were added to effectively get the total estimated height of each subject. The total height was derived by the following equation.

$$\textit{Total estimated height} = 2b_5 + h_5 + h_4 + h_3 + h_2 + h_1 + h_{10} + 2a_7 \tag{3.16}$$

Based on the final measurement definitions for the BSVs, a generalized program was developed to compute segment volumes from ANSURII data.

3.2.3 Center of Mass and Moment of Inertia Calculation

After calculating the masses and geometries for each body segment COMs and MOIs could be determined. Using the method described Theory section 3.1.10 COM positions for each body segment was calculated. The inertial calculations were made by using inertial formulas for each geometric shape. Masses used for this calculation were the ones estimated when using the Modified Hanavan Model for BSMs. For the individual trunk masses, they were derived by dividing the segment volumes with the total volume and then multiplying it with the total trunk mass. When calculating some of the more complex formulas, such as for a *Frustum of an Elliptical Cone*, it was given in Modified Hanavan Model which simplified the calculation. Once the method for determining the COM and MOI had been established, a generalized program was developed to calculate these values using the previously calculated body segment volumes and masses.

3.3 Results

In this chapter the result from mass, volume and moment of inertia estimations is presented. A comparison of the estimated segment masses to the actual body mass data from ANSURII[3] is presented, this is done for both mass estimation methods. The geometric method is evaluated in a similar way by assuming a uniform density.

3.3.1 Result BSM using the Modified Hanavan Model's Method

Below, the average body segment masses for males and females in the ANSURII database are presented. Each segment mass corresponds to the geometries described in Section 3.2.2. The trunk segment method consolidates the lower, middle and upper trunk into a single segment.

Table 3.7: Average estimated mass (kg) of individual body segments for male and female subjects using the Modified Hanavan Model's Method.

Segment	**Male (kg)**	**Female (kg)**
Hand	0.46	0.32
Forearm	1.34	0.94
Upper Arm	2.17	1.49
Foot	0.66	0.53
Shank	3.98	3.72
Thigh	10.31	8.88
Head	5.07	4.66
Trunk	44.81	33.60

Below the comparison between estimated mass and the actual mass of each ANSURII subject using the method from the Modified Hanavan [16] is presented. The sum of all BSMs, making up the entire body mass, is plotted against the actual mass of each subject in the ANSURII dataset. The plots in figures 3.5 and 3.6 presents for females and males separately.

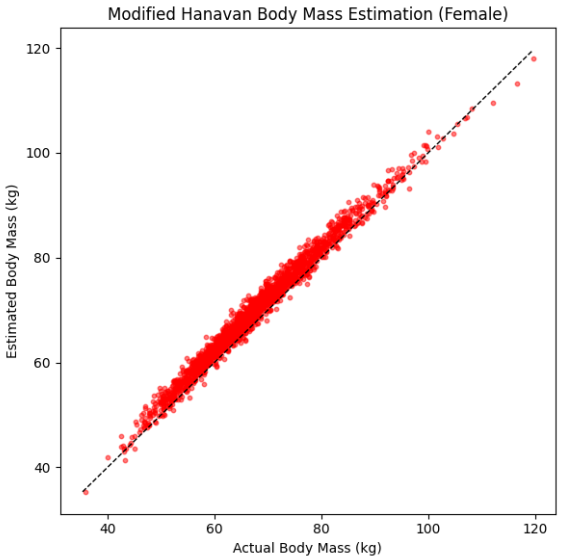


Figure 3.5: Estimation results for female subjects.

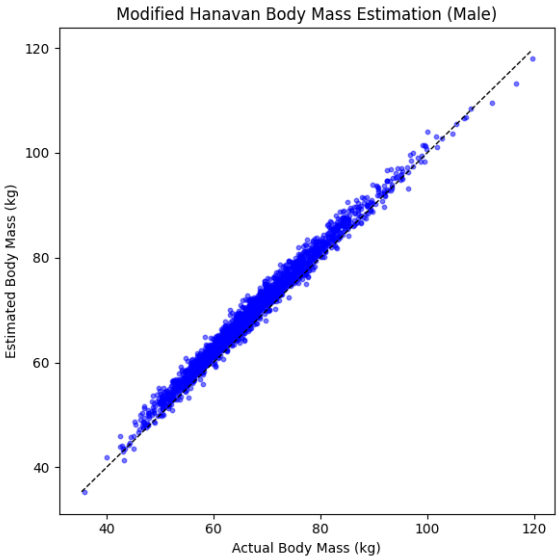


Figure 3.6: Estimation results for male subjects.

Table 3.8 presents the average calculated differences, in mass and percentage between estimated total mass and actual mass for every ANSURII subject and respective gender. The average estimated mass of each specific body segment is presented in table 3.7.

Table 3.8: Comparison of actual and estimated body mass for male and female subjects.

Gender	Actual Mass	Estimated Mass	Difference	Difference
Female	67.76 kg	70.01 kg	2.25 kg	3.42%
Male	85.52 kg	87.72 kg	2.20 kg	2.83%

Results from the comparison shows a slight trend of overestimating total body mass for both females and males. The comparison shows that total body mass is on average overestimated more for females.

3.3.2 Result BSM Estimation using Zatsiorsky's Method

Below, the average body segment masses for males and females in the ANSURII database is presented. In Table 3.9 below, the average BSMs in terms of percent of total body mass for males and females in the ANSURII database is presented.

Table 3.9: Average estimated mass (kg) of individual body segments for male and female subjects using Zatsiorsky's regression equations with added thigh mass. Values in percent represent the proportion of total body mass (85.67 kg for males, 67.88 kg for females).

Segment	Male (kg)	Female (kg)	Male (%)	Female (%)
Foot	1.11	0.88	1.30	1.30
Shank	3.63	2.83	4.24	4.17
Hip	12.27	9.50	14.32	14.00
Hand	0.50	0.41	0.58	0.60
Forearm	1.35	1.11	1.58	1.64
Upper Arm	2.35	1.85	2.74	2.73
Head	5.27	4.78	6.15	7.04
Upper Torso	13.88	11.32	16.21	16.68
Middle Torso	14.64	11.52	17.09	16.97
Lower Torso	9.45	7.09	11.03	10.45

In the following table 3.10, the average body segment masses (BSMs) as a percentage of total body mass (see Table 3.9) are compared with the corresponding average BSM percentages reported by De Leva (1996).

Table 3.10: Comparison of estimated body segment mass percentages between Zatsiorsky's regression model and De Leva (1996, Table 4) for males and females. Thigh (De Leva) corresponds to Hip (Zatsiorsky).

Segment	Est. Male	De Leva Male	Est. Female	De Leva Female
Foot	1.30%	1.37%	1.30%	1.29%
Shank	4.24%	4.33%	4.17%	4.81%
Hip (Thigh)	14.32%	14.16%	14.00%	15.45%
Hand	0.58%	0.61%	0.60%	0.56%
Forearm	1.58%	1.11%	1.64%	1.35%
Upper Arm	2.74%	1.85%	2.73%	2.17%
Head	6.15%	6.94%	7.04%	6.68%
Upper Torso	16.21%	11.32%	16.68	13.88%
Middle Torso	17.09%	11.52%	16.97%	14.64%
Lower Torso	11.03%	7.09%	10.45%	9.45%

The comparison shows that five of the ten calculated BSMs in % are closer to De Levals estimated BSM distribution for males. Likewise five out of ten calculated BSMs is closer for females. The largest difference seen for males is seen in the *Middle Trunk* segment with a difference of 5.57% points. In contrast, for females, the largest observed difference is for the *Upper Trunk* segment with 2.80% points.

The following figures, 3.7 and 3.8 shows plots comparing the total estimated mass and the actual mass of each ANSURII subject using Zatsiorsky's[19] method. The sum of all BSMs, making up the entire body mass, is plotted against the actual mass of each subject in the ANSURII dataset. The plot is presented for females and males separately.

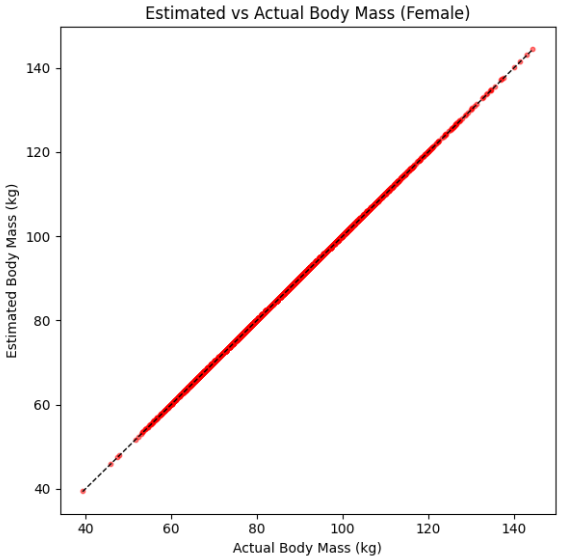


Figure 3.7: Estimation results for female subjects.

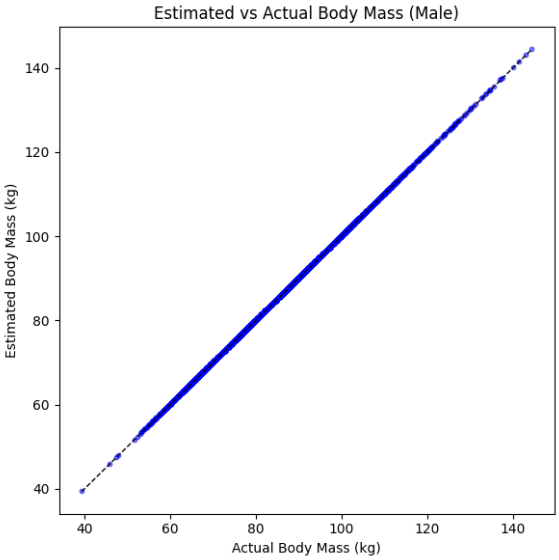


Figure 3.8: Estimation results for male subjects.

Table 3.11 presents the average calculated differences, in mass and percentage between estimated total mass and actual mass for every ANSURII subject and respective gender.

Table 3.11: Comparison of actual and total estimated body mass for male and female subjects using Zatsiorsky's method with thigh mass included.

Gender	Actual Mass	Estimated Mass	Difference	Difference
Female	67.76 kg	67.88 kg	0.12 kg	0.18%
Male	85.52 kg	85.67 kg	0.15 kg	0.17%

Results from the comparison shows an almost perfect total body mass estimation for both females and males. The comparison shows that total body mass is overestimated by less then 0.19% for both females and males.

3.3.3 Result Geometric Estimation

The figures 3.9 and 3.10 show the total estimated mass by adding all the computed BSV for each subject multiplied by a uniform density of 1000 kg/m^3. Each body segments volume is defined as in the Modified Hanavan Model. The resulting volume in m^3 is then multiplied with 1000 kg/m^3 and plotted against the known body mass for each subject.

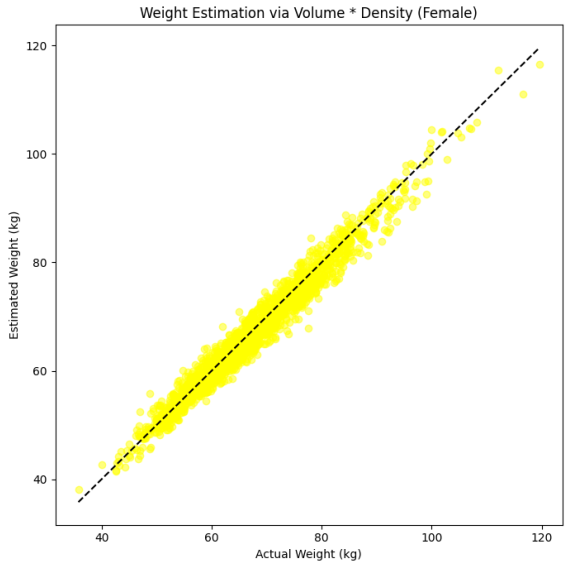


Figure 3.9: Volume estimation results for female subjects using a uniform density.

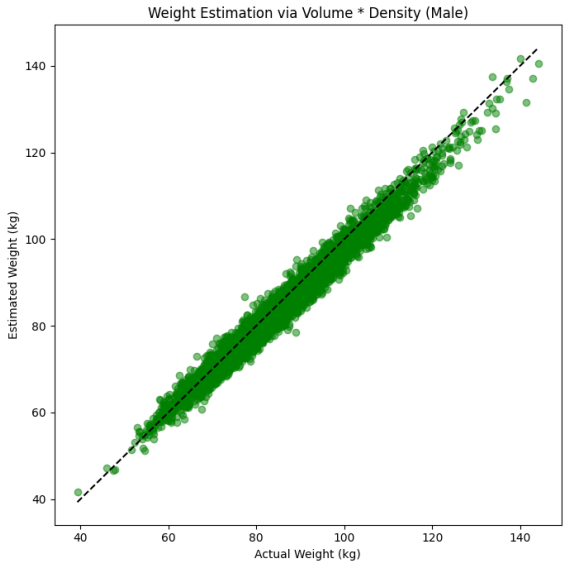


Figure 3.10: Volume estimation results for male subjects using a uniform density.

Table 3.12 presents the average calculated differences, in mass and percentage between estimated total mass and actual mass for every subject and respective gender when using the volume estimation model.

Table 3.12: Comparison of actual and volume-based estimated body mass for female and male subjects.

Gender	Actual Mass	Estimated Mass	Difference	Difference
Female	67.76 kg	67.26 kg	-0.50 kg	-0.64%
Male	85.52 kg	84.15 kg	-1.37 kg	-1.48%

Results from mass estimation using the volume to determine mass shows a small difference when using a uniform density of 1000 kg/m^3 for both females and males. The comparison shows that total body mass is underestimated with an average of 0.64% for females and 1.48% for males.

Figure 3.11 and 3.12 presented below represent the total estimated height plotted against the known height for each ANSURII subject. Each subjects estimated total height was estimated using equation 3.16 and plotted against the known subject height for each ANSURII subject.

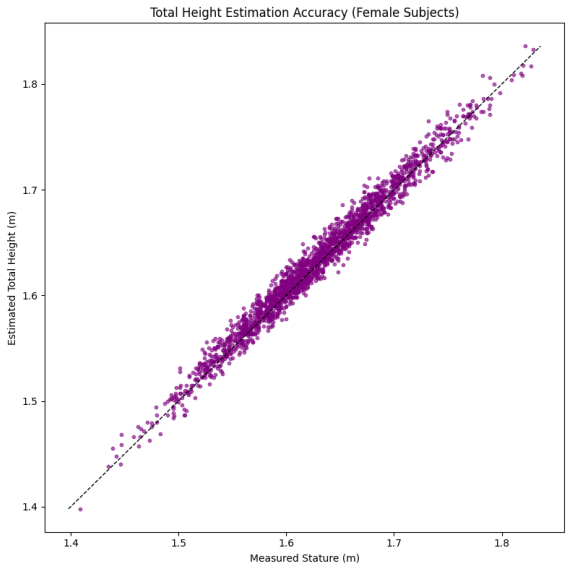


Figure 3.11: Height estimation results for female subjects.

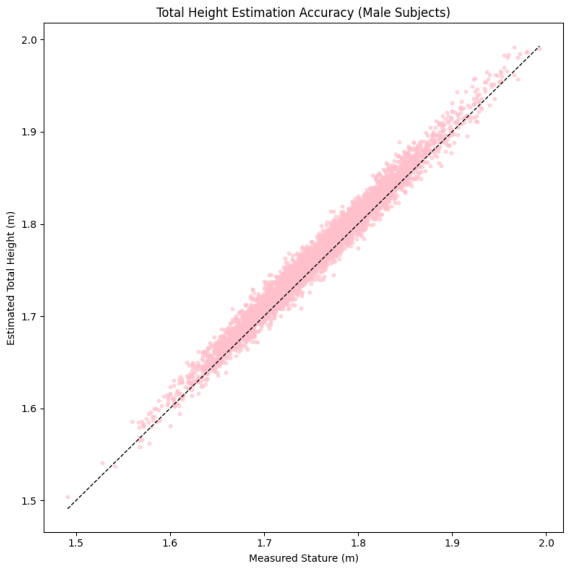


Figure 3.12: Height estimation results for male subjects.

In Table 3.13 the average estimated total height for both genders is presented numerically. It is compared to the actual known height for each subject and the offset is presented in meters and percent.

Table 3.13: Comparison of actual and estimated average height for male and female subjects.

Gender	Actual Height	Estimated Height	Offset	Offset
Male	1.7562 m	1.7651 m	0.0089 m	0.50%
Female	1.6285 m	1.6326 m	0.0042 m	0.26%

Results from height estimation shows very a small difference in offset. For both genders the estimated total height is less than one centimeter off the average actual height for all the subjects.

3.3.3.1 Result Center of Mass and Moment of Inertia

Table 3.14 presents the result of the calculated COM placements. They are presented as percentage of segment length measured from the distal end, meaning the distance from the lowest part for each body segment. The positions are visualized later on and can be seen represented in Figure 4.1 and Figure 4.2.

Table 3.14: Average center of mass location as a percentage of segment length, measured from the distal end, for males and females.

Segment	**Male (%)**	**Female (%)**
Head	50.00	50.00
Upper Trunk	50.00	50.00
Middle Trunk	53.22	53.12
Lower Trunk	50.00	50.00
Upper Arm	52.36	52.37
Lower Arm	58.95	58.48
Hand	50.00	50.00
Thigh	56.83	56.94
Shank	59.11	59.64
Foot	50.00	50.00

Below in table 3.15, the average moment of inertia is presented in $kg \cdot m^2$ for females and males. These values represent the rational resistance of each body segment around its principal axis.

Table 3.15: Average calculated moments of inertia of each body segment for females and males ($\mathrm{kg} \cdot \mathrm{m}^2$)

Segment	**Female**			**Male**		
	I_{XX}	I_{YY}	I_{ZZ}	I_{XX}	I_{YY}	I_{ZZ}
Head	0.0507	0.0507	0.0303	0.0702	0.0702	0.0398
Upper torso	1.0352	0.6099	2.3797	1.2936	0.6169	3.0301
Middle torso	1.1780	1.9472	0.0050	2.4246	2.4246	0.1435
Lower torso	0.4133	0.2551	0.9820	0.6961	0.4504	1.6403
Upper arm	0.1338	0.1338	0.0037	0.2085	0.2085	0.0066
Lower arm	0.0130	0.0130	0.0003	0.0229	0.0229	0.0006
Hand	0.0024	0.0024	0.0002	0.0043	0.0044	0.0004
Thigh	1.6865	1.6865	0.1113	2.0960	4.0539	0.0049
Shank	0.1249	0.1249	0.0028	0.1436	0.1436	0.0032
Foot	0.0216	0.0439	0.0423	0.0286	0.0667	0.0630

3.4 Discussion

From the result it can be seen that the Zatsiorsky's method[19] shows a lower average estimation error and is recommended for the final implementation of this project. In contrast the Modified Hanavan Model's method[16] for BSM estimation would be better suited in cases where further development of the project allows for the incorporation of a broader set of anthropometric measurements as initial input values. The accuracy of the geometric estimation model appear to be high. Validation using a uniform density shows that overall volume is estimated to an accurate degree, likewise dimensions making up total height also appear to be accurate. COM and MOI results are not directly validated but are believed to follow the general trends of what is seen in the literature.

3.4.1 Interpretation of BSM Estimation Results

In general, both the Modified Hanavan Model and Zatsiorsky's method provide total body segment mass estimates within an acceptable range of accuracy. While Zatsiorsky's regression based approach yields more accurate results on average, the Modified Hanavan Model[16] may be better suited for capturing individual variation due to its reliance on detailed, subject-specific anthropometric inputs. As presented in theory section 3.1.2 the BSMs using Zatsiorsky's method does not correspond to the BSMs described in Section 3.2.1.2 as they use different landmarks to define body segmentation. Specific segment masses should therefore not be compared between methods.

3.4.1.1 Reason for Modified Hanavan Result

The slightly larger inaccuracy observed in the BSM estimates using the Modified Hanavan Model 3.3.1 may be attributed to two factors. The first is because of limitations in the original dataset used to develop the regression equations. The second might be caused from uncertainties introduced during the approximations of required anthropometric measurements. While a certain degree of inaccuracy is to be expected, the larger estimation inaccuracy compared to Zatsiorsky's method might be attributed to the data used to develop Clauser et als[4] regression equation. As described in the theory 3.1.5 Clauser et al.s's regression equation is based on data from 6 cadavers with a narrow demographic. This is a limited amount of data to be applied on a generalized set of regression equations and might prove to be unsuitable (to a degree) for the larger more diverse population in ANSURII[3]. Unfortunately not a lot of information about the 6 cadavers is available so the amount of impact this had on the accuracy of the regression equations can only be speculated.

Another contributor to the difference in mass estimation is a cause of the approximations made to the anthropometric measurements. We know that the approximations made when adapting the ANSURII measurements to those required by the Modified Hanavan Model will have an impact on the final BSM estimation 3.2.1.2. For example, if an approximation made in the adaptation of the measurements proved to be an over approximation the resulting mass of that body segment will be overestimated. By comparison the data plotted in figures 3.5 and 3.6 appear more spread out than those in figures 3.7 and 3.8. This might be explained by the fact that the estimation using the Modified Hanavan Model is more individualized. Since this method accounts for individual differences in

the anthropometric measurements, it allows for a more individualized estimation of body segment masses. It also explains why subjects with same body mass can receive differently estimated body masses. Essentially this means that, if two individuals have the same height and mass but different body segment shapes, the Modified Hanavan Model will produce different segment masses and Zatsiorsky's will not.

This individualization would not however be represented in the final implementation of the project. Because the anthropometric measurements for a given height and mass will be predefined (see Section 2.4.3), the corresponding geometric shapes of the body segments will likewise be predetermined for that same height and mass.

3.4.1.2 Reason for Zatsiorsky's Methods Result

The Zatsiorsky methods ability to make better estimations on average might be explained by a more diverse and fitting dataset used when developing the regression equations along with a limited amount of needed anthropometric measurements.

As described in the theory chapter 3.1.3, Zatsiorsky's method used a significantly larger population to derive regression compared to Clauser et al. The population consisted of physical education students. It would be reasonable to assume that this type of population might posses a similar fitness level to those of military personnel. This type of population sample might therefore be better suited for a database such as ANSURII.

As there are no anthropometric measurements used in the regression equations there are also no added sources of estimation errors. In contrast to the Modified Hanavan Model's method of calculating BSMs, Zatsiorsky's method only require the height and mass for a specific subject. The absence of approximated anthropometric measurements ensures that no additional approximation error is introduced during the calculation of body segment masses.

3.4.1.3 Gender Differences in Mass Estimation

Both BSM estimation models were developed using male data, which may in theory reduce their accuracy when applied to female subjects. However, the extent of this impact on the final estimations is most likely low. Unfortunately the regression equations from the Modified Hanavan Model and Zatsiorsky's method is only intended to be used for male subjects. Despite this fact, total body mass estimation has been shown to be comparably accurate on average for both methods (see Section 3.3.1 and 3.3.2.), with Zatsiorsky's method exhibiting more consistent accuracy between male and female subjects.

Even though the equations used by Zatsiorsky's method were developed to provide estimations of what average male BSMs weigh, they appear to be equally applicable to both genders when estimating BSMs. The comparison done to the results presented by De Leva (1996), seen in section 3.10 show that there is no clear trend in better accuracy for BSM estimations for any specific gender. The reason for this absence of disparity between genders may lie in the population averaged nature of Zatsiorsky's regression equations, which apply uniform scaling factors across sexes, thereby reducing gender specific variation in the estimates.

No comparison to literature could be done for female segment mass distributions when using the Modified Hanavan Model but this distribution also likely captures gender differences. As discussed in section 3.4.1.1, the modified Hanavan Model's method represents differences between individuals, this would be equally true for differences between genders of individuals.

3.4.2 Interpretation of Geometric Estimation Results

The results from the volume estimation suggest that the BSV are well estimated. However, the estimated total mass is not inherently meaningful, rather, it serves as an indirect validation that the BSVs have been modeled in a way that accurately reflects the subject's total body volume but does not necessarily the shape of the BSVs. We see that when assuming the uniform density of 1000 kg/m^3, an accurate mass estimation can be obtained. Since kg/m^3 is only an assumed approximation of uniform human body density, the resulting mass estimate should likewise be considered an approximation, rather than a precise measurement. Because the total estimated mass falls in a reasonable range this must in term imply that the total estimated volume also falls in a reasonable range.

The results from the total estimated heights indicate that the geometric shapes used in the model are accurate with respect to the vertical dimensions. The small offset observed between the estimated and actual total height suggests that the segment dimensions contributing to overall height are well approximated. This outcome also shows that the modification made to the Modified Hanavan Model by implementing a neck and smaller head works well. The precise estimation to total height is expected, as only approximation for the height of the head and neck was made from the ANSURII measurements for the vertical components (see Section 3.2.2).

It is important to note that these results only validate the accuracy of the vertical dimension and volume, not the full geometric shape of each segment. BSVs and the vertical dimensions do not provide full information about the accuracy of all dimension used for the shapes.

3.4.3 Center of Mass and Inertial Calculations

When comparing results from table 3.14 and table 3.15 with the reference results presented by Nikolova and Toshev (See section 3.1.9), it is observed that the results are reasonable but present some differences.

Differences should be expected. Due to differences in geometric modeling assumptions and the use of distinct anthropometric datasets, the resulting body segment parameter estimates will exhibit some degree of variation. Even the different ways the models were simplified play a large role in the result. Therefore, comparisons with values from the literature should be interpreted primarily as a means to understand ballpark estimates, rather than as exact benchmarks.

Despite these differences, the results are still consistent enough to be considered appropriate. The results seem to follow the general trends and has a relative magnitude of COM location percentage and MOI. Therefore, even tough numerical variations exist, they do not reduce the validity of the result.

4

Human Body Model

The purpose of this chapter is to describe the final step in the project where the results of Chapter 2 and 3 are combined. Measurements estimated with anthropometry regression from Chapter 2 together with the inertial properties from Chapter 3 are used to generate a two-dimensional (2D) plot and a rigid human body dynamic model.

4.1 Theory

The following section presents the required theory and information to create a dynamic model of a human body. It explains what the URDF language is and what a URDF loader is.

4.1.1 URDF

Unified Robotics Description Format or URDF is a programming language used to describe mechanical systems. The format is widely used in different applications to provide physical descriptions of robots [30]. A URDF file can include information about positions of links, definitions about how joints are connected to links and their attributes. The joints can, for example, be defined to allow different ranges of rotation or resistances to movement. It can also provide information about inertial properties in the links [31].

4.1.2 URDF Loaders

The URDF files can be used in combination with simulators called *URDF loaders*. This can be used to simulate the information stored in the URDF files and provide visual feedback. There are several available simulators online with a variation of features. One example of a simple URDF simulator is *URDF viewer example* developed by Garrett Johnson [32].

4.1.3 Hip joint position

A study made by Harrington et al [33] focuses on how accurately three predictive equations estimate the hip joint center by using MRI data. Data was used from adults, healthy children and children with cerebral palsy.

4.2 Methods

The estimated measurements from the results of the anthropometric regression were implemented with strategies for estimating the inertial poperies. The measurements were use to calculate BSMs using both regression equations described in Chapter 3. Likewise, geometries for BSVs were calculated using the same estimated measurements. Subsequently, the COMs and MOIs were derived from these.

To achieve a better understanding of the resulting model, a 2D plot illustrating the COM was generated. Additionally, a three-dimensional (3D) model was designed to visualize the interaction of the inertial properties. Furthermore, both visualization models are intended to provide visual feedback on how anthropometric measurements and geometric assumptions affect the final estimation of inertial properties.

4.2.1 2D Visualization Tool

A generalized program for the illustration of a 2D projection of the BSVs was developed. Since the volumes exist in a 3D space, but a 2D visualization of the body segments is desired, a transformation of the coordinate system is necessary. To project the volume onto two dimensions, the y and z coordinates from the BSV were used. The height of the segment in 3D is represented in the z-axis, while it is represented in the y-axis in 2D.

This projection was made for all BSVs and then placed in a coordinate system. The resulting shapes from the projections included differently shaped polygons and ellipses. The previously determined COMs were also placed in the same global coordinate system. Projections and COM placements where implemented in a python program using the matplotlib library.

Joint center positions were estimated based on regression relationships reported in the literature. Joint positions in the proximal-distal direction were positioned using segmental offset percentages derived from the De Leva study [20], in which average joint locations were expressed as a percentage of segment lengths. These percentages were applied to the geometrically estimated segment lengths from our model to compute joint center locations. No joint center was estimated for the head due to the lack of a clearly defined rotational center for that segment. For the hip joint center in the medial-lateral direction, offsets were estimated using regression data reported by Harrington et al [33]. The joint positions for shoulders, elbows and wrists in medial-lateral direction were assumed to be at the position of *Acrominal breadth* according to the ANSURII definitions. Knee and ankle positions were assumed to be placed in the middle of each leg. Joint center positions were then added to the plot.

4.2.2 Rigid Human Body Dynamic Model

In order to visualize the BSVs together with the different inertial properties a rigid human body dynamic model was created. This was done by developing a program for generating rigid human body dynamic models based on the previously implemented geometric model and calculated MOIs. The program was designed to be compatible with the URDF standard. The first step meant developing a mesh generation program. The program was

designed to generate body segment meshes based on the output data from the final software, ensuring that the mesh dimensions correspond directly to the calculated geometrys described in Chapter 3.

Likewise, another program was developed for generating the main URDF script. The program generates a URDF compatible script that tells each mesh where it should be placed. It also generates information about where joints should be incorporated and how they should behave. Furthermore it also provides information about the calculated MOI for each body segment.

4.3 Results

The results of the implementation of the estimated measurement to the inertial properties calculations will in this section be presented for the two different visualization tools.

4.3.1 2D Results

In the two figures below, Figures 4.1 and 4.2, the result of the 2D plot is presented for each gender using the average heights and mass from ANSURII. A 2D projection of each geometry is presented in the global coordinate system at the their respective calculated positions. The resulting shapes are based on the final estimation of every geometry dimension needed for the 2D projection. The position of the COM for every segment is also communicated with a red marker accompanied by corresponding labels for each body segment. Positions for estimated joint centers are also presented with black markers.

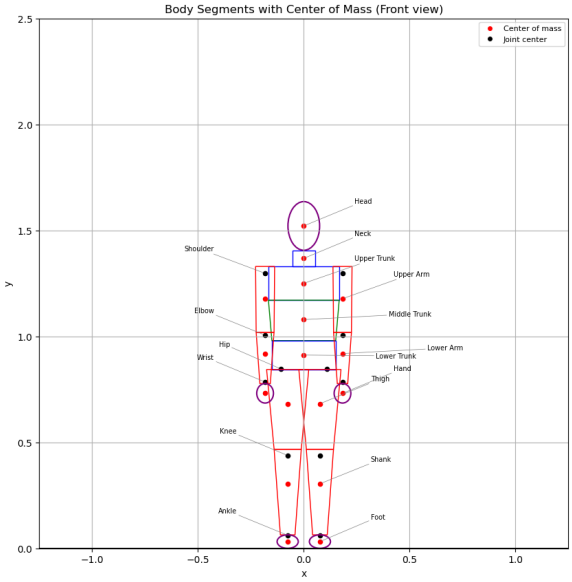


Figure 4.1: 2D plot of a 68 kg and 163 cm tall female.

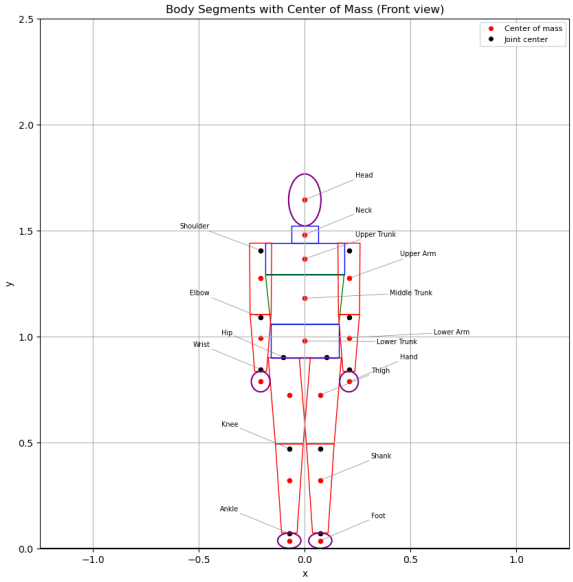


Figure 4.2: 2D plot of a 86 kg and 176 cm tall male.

Figures 4.3 and 4.4 below, the result of the 2D plot is presented for each gender using the same height and mass for each gender to illustrate the differences between the genders. The height 170 cm and mass 75 kg was used.

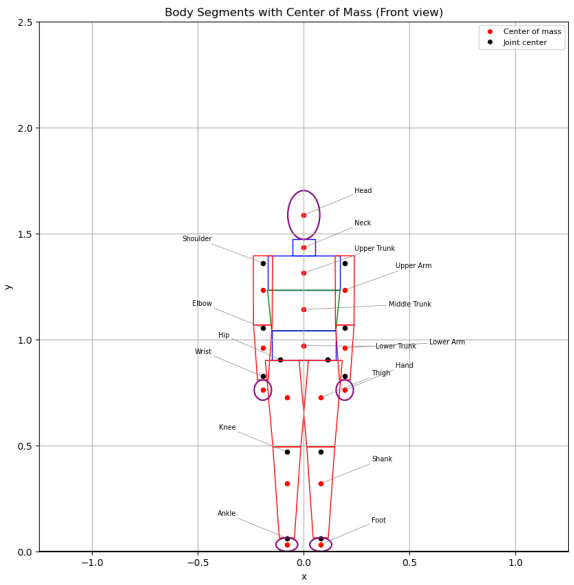


Figure 4.3: 2D plot of a 75 kg and 170 cm tall female.

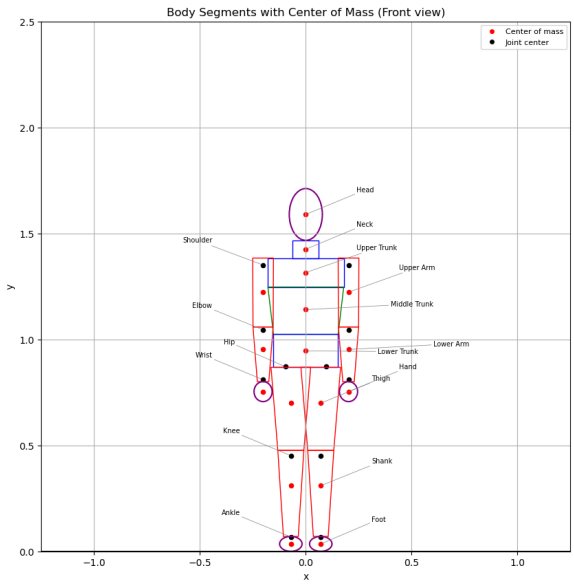


Figure 4.4: 2D plot of a 75 kg and 170 cm tall male.

Several small, but noticeable differences between genders are observable in the resulting figures. Namely, differences in the size of *upper arms*, *lower arms*, *thighs*, *shanks* and *feet* can be distinguished. The female plot appears to have smaller *upper-* and *lower arms* with larger *thighs* and a wider hip positions. Additionally, a difference in proportions in the *trunk* segments with the male figure appearing to have a larger *middle trunk*.

4.3.2 Results of the Rigid Human Body Dynamic Model

The resulting rigid human body model for each gender is presented in Figures 4.5 and 4.6. 3D meshes of each body segment are generated and placed in the resulting URDF model according to their calculated positions. The resulting meshes are based on the final estimation of every geometry dimension. The positions for every joint are also calculated within the URDF scrip. However, these are not visualized when imported into the URDF viewer.

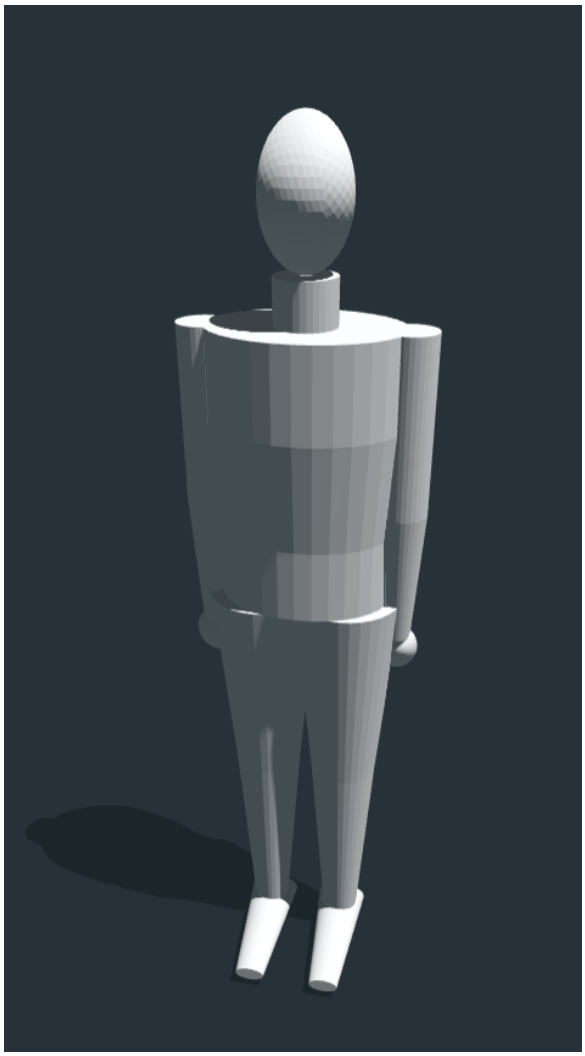


Figure 4.5: URDF model of a 68 kg and 163 cm tall female.

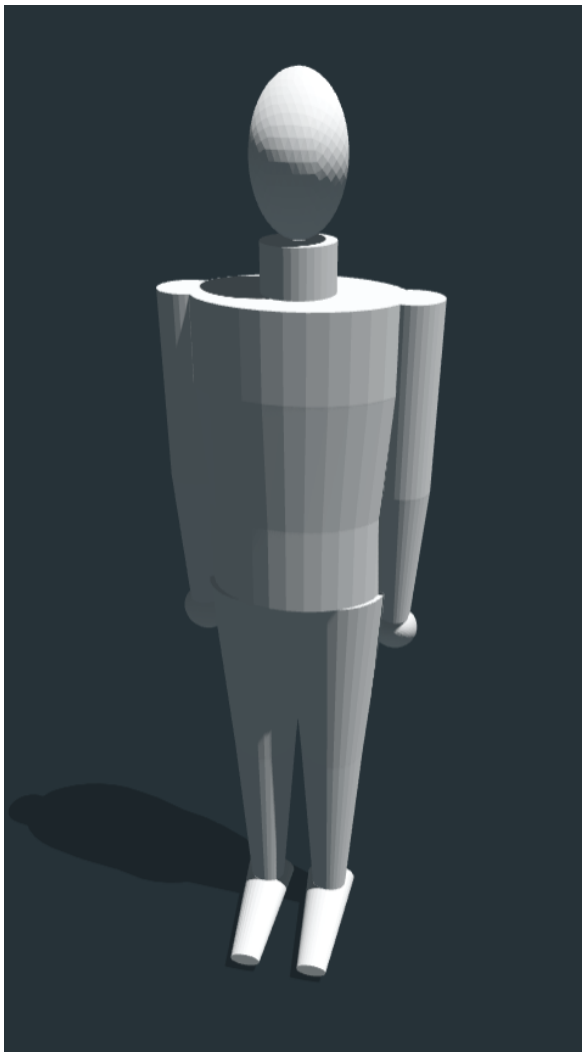


Figure 4.6: URDF model of a 86 kg and 176 cm tall male.

Images of the 3D models for each gender using same height and mass. The height 170 cm and mass 75 kg was used.

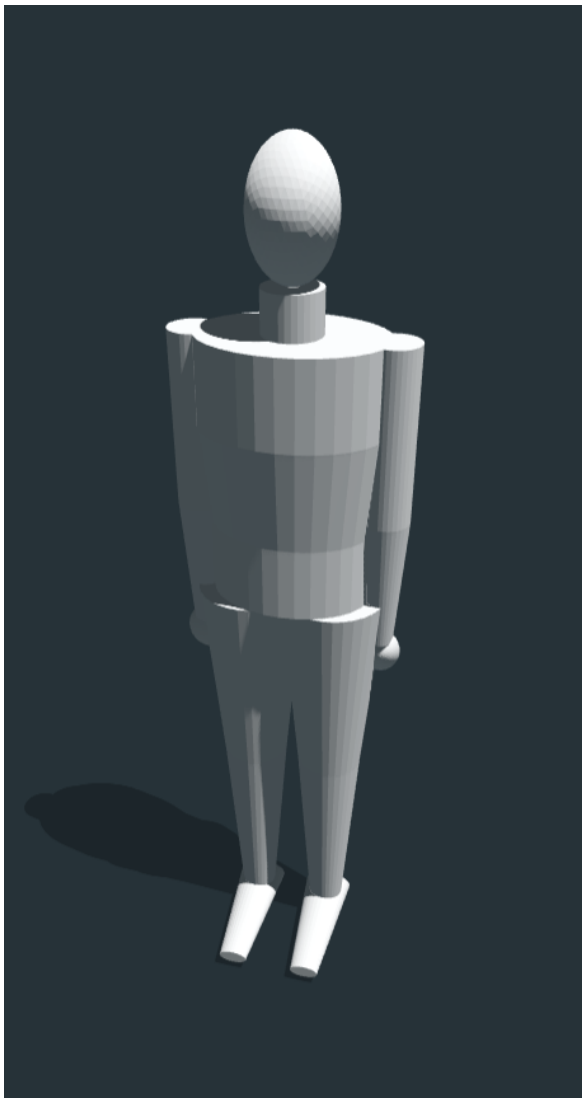


Figure 4.7: URDF model of a 75 kg and 170 cm tall female.

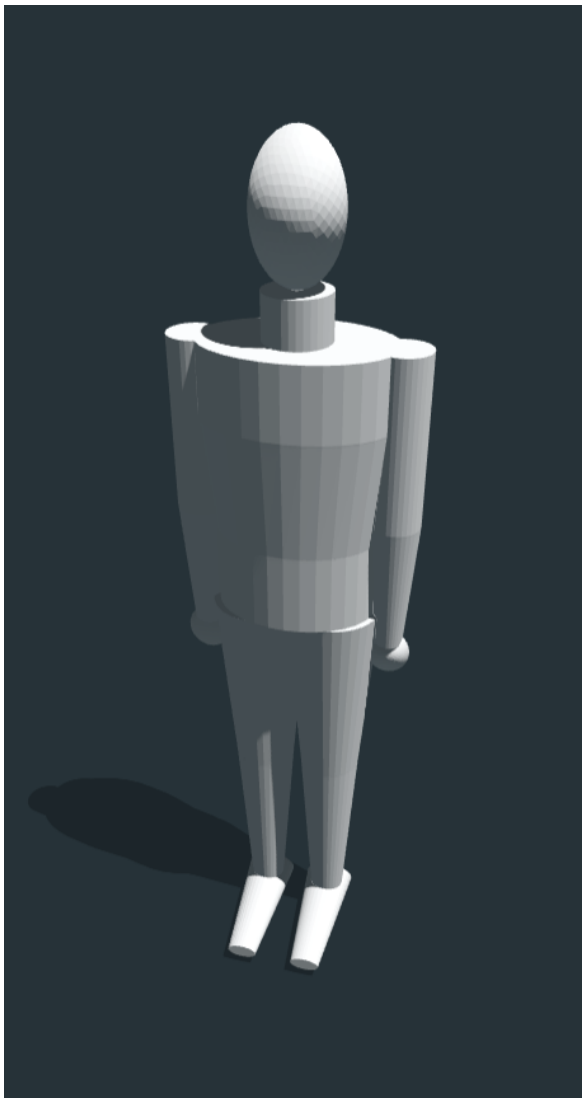


Figure 4.8: URDF model of a 75 kg and 170 cm tall male.

Differences observed between male and females in the 3D URDF models are identical to the differences discussed in the 2D model. Be aware that the images might be slightly misrepresented as the taken screenshots might incorporate slight deviations in viewing angle and zoom.

4.4 Discussion

The results from both the 2D and 3D models reveals differences between male and female. The visual differences depend on the underlying measurements. The measurements are derived from the anthropometric estimation.

In the visualization, the most notable differences appear in the hip and chest regions. These observations are consistent with the results presented in Chapter 2 where the PCA indicated that the *buttock circumference*, *chest circumference* and *waist circumference* were the measurements most influenced by gender.

The visualization also shows that the human tend to reflect typical army characteristics, such as broad shoulders and narrower waist. In contrast, a less physically trained body may have mass placed in other regions, such as narrower shoulders and wider waist. This is something the current model does not take into account. The reason for these characteristics is the data on which the model is built, as mentioned in Chapter 2. Including an extra measurement found in the PCA as an input parameter could improve the accuracy of the mass distribution.

Using another data set with non military people would make the predictions more accurate to the average population. Thus, it would still only be giving one distribution of the mass per gender. A male with a mass of 80 kg and height of 180 cm will be visualized regardless of the distribution of fat and muscles. To improve the mass distribution, an additional input measurement could be included, such as *waist circumference*. The 180 cm and 80 kg male with a larger waist would look different than a male with same height and mass but with a smaller waist.

When visualizing the body results from Chapter 2 and 3 are combined. That means that the errors from both the anthropometric estimation and the mass- and inertial estimations are combined. The estimation of measurements contributes to a source of error, and the way that the measurements are put together is another source of error. It is therefore important to be aware of the estimation of estimations.

5

Software Workflow

Since this project is part of a bigger initiative it is essential that the project is well organized and usable for extern people. Therefore, special emphasis was placed on creating a clear and functional structure that effectively integrates all parts of the project. A tool that was used for achieving this is called nbdev.

The code for this project was written in Python using multiple Jupyter Notebooks. This format allows combining executable code cells with Markdown cells to make explanations to the code. When writing functions that were needed to be reused across other notebooks, the function were exported to a python file. This export process could be automatically done by using the python library nbdev, which converts notebook code into python modules. The platform nbdev is based on the simplicity of having all code in one place, and then making it usability by automating the synchronization between notebooks and source code and by generating documentation automatically [34].

The structure of our project starts with a folder 'nbs', short for notebooks, where all code is written. The files are numerated, and the files for handling the data starts at 100. The regression files starts at 200, followed by the PCA at 300, and mass estimation at 600. Finally at 1000 all different parts are put together in a function where input is height, mass, and gender and the output is a visualization of the estimated human. The Python modules which hold all reusable functions are generated in another folder called antropmass. The data used for the project is located in a folder called data. This folder includes two sub folders, one for raw data and one for processed data. Lastly, there is an output folder where tables, figures and the models are saved. A map of the structure can be seen in Figure 5.1.

5.1 Map of Folder Structure

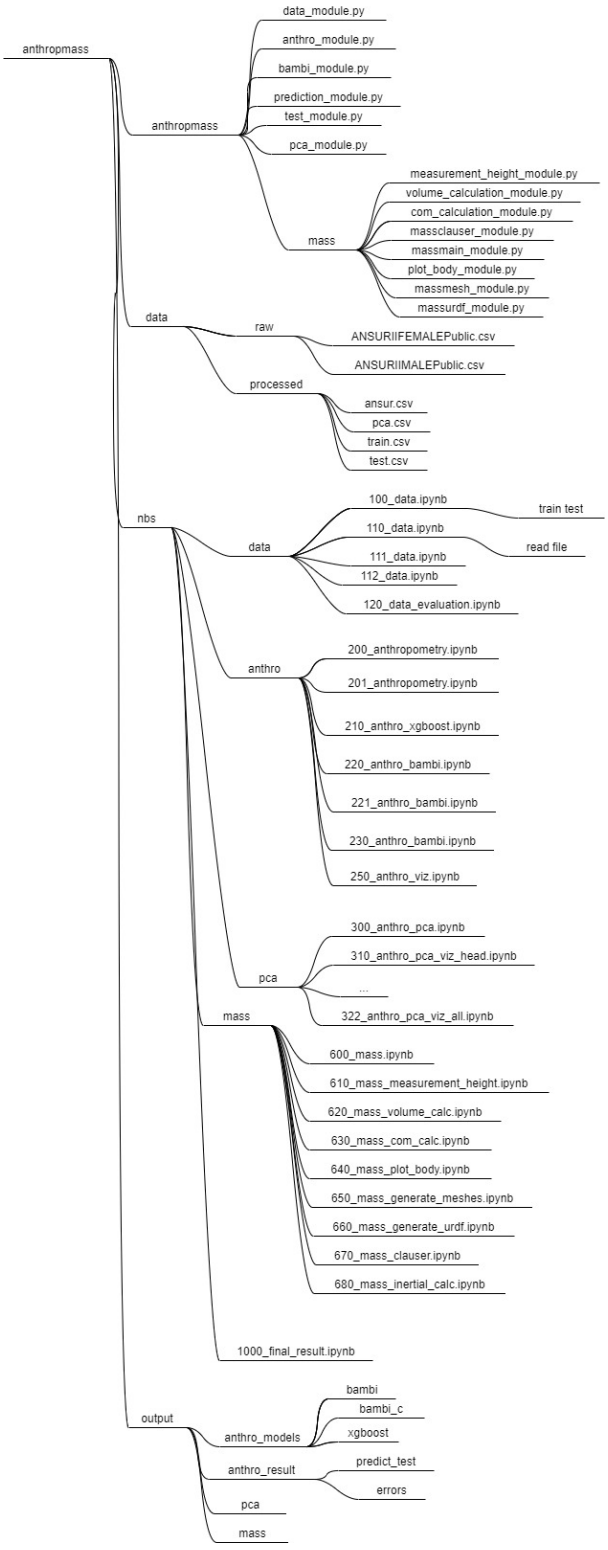


Figure 5.1: Structure of the files, where the main folders are antropmass, data, nbs, and output.

6

Conclusion

The aim of this project, to create a model which predicts body segments mass- and inertial properties, was fulfilled. The first step was to estimate plausible measurements for body segments using regression. In this project two different regression models were created. Both models were able to predict body measurements with an average error of 4 %, based on the three input variables, height, mass and gender. Measurements containing significant amount of information about other measurements were found by the Principal Component Analysis, such as *buttock circumference*. Incorporating these identified measurements as input variables could further improve the prediction accuracy.

The predicted body measurements were used in models for predicting volumes and inertial properties. To develop mass and volume estimation models, an adaptation of the Modified Hanavan Model and Zatsiorsky's model was made. Estimating body segment mass and volume appear to have an accurate overall estimation but the accuracy of individual body segments can not be directly validated. The estimation might lose accuracy when estimating for female subjects. However, this cannot be verified for individual body segments.

Further research could benefit from using a dataset that includes information from civilians. This project used military data, meaning the resulting models reflect properties from army personnel. Expanding the dataset or using a dataset that includes civilians would include people with different lifestyles, leading to a more inclusive and representative study. This project also included a Principle Component Analysis (PCA). Incorporating the results from the PCA in the anthropometric regression models could enhance the predictions' accuracy and thereby reduce the error term. Another improvement that could be included in future research is either finding or developing a regression model designed for females.

Bibliography

[1] E. Epker, "Fasten your seatbelts: A female car crash test dummy represents average women for the first time in 60+ years," 2023, available: urlhttps://www.forbes.com/sites/evaepker/2023/09/12/fasten-your-seatbelts-a-female-car-crash-test-dummy-represents-average-women-for-the-first-time-in-60-years/(Accessed: 2025-02-11).

[2] Y.-H. Kwon, "The effects of body segment parameter estimation on the experimental simulation of a complex airborne movement," Ph.D. thesis, The Pennsylvania State University, University Park, PA, 1993.

[3] C. Gordon *et al.*, "2012 anthropometric survey of u.s. army personnel: Methods and summary statistics," U.S. Army Natick Soldier Research, Development and Engineering Center, Natick, Massachusetts 01760-2642, Tech. Rep., 2012.

[4] C. E. Clauser, J. T. McConville, and J. W. Young, "Weight, volume, and center of mass of segments of the human body," Aerospace Medical Research Laboratories, Wright-Patterson Air Force Base, Ohio, Tech. Rep. AMRL-TR-69-70, 1969.

[5] E. Saccenti, "A gentle introduction to principal component analysis using tea-pots, dinosaurs, and pizza," *Teaching Statistics*, vol. 46, no. 1, pp. 38–52, 2024.

[6] M. Greenacre *et al.*, "Principal component analysis," *Nature Reviews Methods Primers*, vol. 2, no. 1, p. 100, 2022.

[7] Nationalencyklopedin, "Regressionsanalys," 2025, hämtad 2025-05-06. [Online]. Available: https://www.ne.se/uppslagsverk/encyklopedi/lÃěng/regressionsanalys

[8] K. Matematik, "Bayes och bayesiansk statistik," 2004, hämtad 2025-05-06. [Online]. Available: https://www.math.kth.se/matstat/gru/godis/bayes.pdf

[9] T. Capretto, C. Piho, R. Kumar, J. Westfall, T. Yarkoni, and O. A. Martin, "Bambi: A Simple Interface for Fitting Bayesian Linear Models in Python," *Journal of Statistical Software*, vol. 103, no. 15, pp. 1–35, 2022. [Online]. Available: https://www.jstatsoft.org/article/view/v103i15

[10] B. Team, "Getting started with bambi," 2025, accessed: 2025-05-06. [Online]. Available: https://bambinos.github.io/bambi/notebooks/getting_started.html

[11] XGBoost Developers, *XGBoost Documentation (Version 3.0.0)*, 2023, accessed: 2024-05-02. [Online]. Available: https://xgboost.readthedocs.io/en/release_3.0.0/

[12] R. McElreath, "Statistical rethinking 2024 – lecture 9: Introduction to glms," https://www.youtube.com/watch?v=rZk2FqX2XnY&list=PLDcUM9US4XdPz-KxHM4XHt7uUVGWWVSus&index=9, 2024, accessed: 2025-05-01.

[13] A. Vehtari, A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner, "Rank-normalization, folding, and localization: An improved r for assessing convergence of mcmc," *Bayesian Analysis*, vol. 16, no. 2, pp. 667–718, 2021. [Online]. Available: https://projecteuclid.org/journals/bayesian-analysis/volume-16/issue-2/Rank-Normalization-Folding-and-Localization--An-Improved-R%CB%86-for/10.1214/20-BA1221.full

[14] S. learn developers, "Mean squared error — scikit-learn documentation," https://scikit-learn.org/stable/modules/model_evaluation.html#mean-squared-error, 2024, accessed: 2025-05-04.

[15] U. D. of Defense, "Anthropometric database," https://ph.health.mil/topics/workplacehealth/ergo/Pages/Anthropometric-Database.aspx, 2025, accessed: 2025-04-29.

[16] Y.-H. Kwon, "Body segment parameter estimation: Hanavan model," http://www.kwon3d.com/theory/bspeq/hanavan.html, 2000, accessed: 2025-05-04.

[17] S. Kumar, "Contribution to modeling of human walking gait over a stride based on robotics for pedestrian navigation solution," Master's Thesis, Ecole Centrale de Nantes, Nantes, France, 2015, supervised by Prof. Yannick Aoustin, Prof. Eric Le-Carpentier, and Prof. Valerie Renaudin. [Online]. Available: https://www.researchgate.net/publication/304181267

[18] S. Kumar, V. Renaudin, Y. Aoustin, E. Le-Carpentier, and C. Combettes, "Model-based and experimental analysis of the symmetry in human walking in different device carrying modes," in *2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2016, pp. 1172–1179.

[19] V. Zatsiorsky and V. Seluyanov, "The mass and inertia characteristics of the main segments of the human body," in *Biomechanics VIII-B: Proceedings of the Eighth International Congress of Biomechanics*, ser. International Series on Biomechanics, H. Matsui and K. Kobayashi, Eds. Champaign, Illinois: Human Kinetics Publishers, 1983, vol. 4B, pp. 1152–1159.

[20] P. de Leva, "Adjustments to zatsiorsky–seluyanov's segment inertia parameters," *Journal of Biomechanics*, vol. 29, no. 9, pp. 1223–1230, 1996.

[21] E. P. Hanavan, "A mathematical model of the human body," Aerospace Medical Research Laboratories, Wright-Patterson Air Force Base, Ohio, Tech. Rep. AMRL-TR-64-102, AD-608-463, 1964.

[22] Wolfram MathWorld, "Elliptic cylinder," https://mathworld.wolfram.com/EllipticCylinder.html, n.d., accessed: 2025-05-12.

[23] ——, "Frustum," https://mathworld.wolfram.com/Frustum.html, n.d., accessed: 2025-05-12.

[24] ——, "Ellipsoid," https://mathworld.wolfram.com/Ellipsoid.html, n.d., accessed: 2025-05-12.

[25] X. G. Tan and A. J. Przekwas, "A computational model for articulated human body dynamics," *International Journal of Human Factors Modelling and Simulation*, vol. 2, no. 1/2, pp. 69–86, 2011.

[26] G. S. Nikolova and Y. E. Toshev, "Estimation of male and female body segment parameters of the bulgarian population using a 16-segmental mathematical model," *Journal of Biomechanics*, vol. 40, no. 16, pp. 3700–3707, 2007.

[27] Wolfram MathWorld, "Frustum," https://mathworld.wolfram.com/Frustum.html, n.d., accessed: 2025-03-11.

[28] W. F. Repository, "Moment of inertia of an ellipsoid," https://resources.wolframcloud.com/FormulaRepository/resources/Moment-of-Inertia-of-an-Ellipsoid, 2024, accessed: 2025-04-29.

[29] A. I. of Physics, "Aip handbook: Section 2c," https://web.mit.edu/8.13/8.13c/references-fall/aip/aip-handbook-section2c.pdf, 1990, accessed: 2025-04-29.

[30] Formant Inc., "Urdf (unified robot description format)," https://formant.io/resources/glossary/urdf/, n.d., accessed: 2025-05-12.

[31] S. Kumar, "Tutorial 02: Getting the right model of your robot," Online video, Chalmers Play, 2025, available only to Chalmers University students. [Online]. Available: https://play.chalmers.se/playlist/dedicated/746353/0_9us6ergl/0_haewb42v

[32] G. Johnson, "Urdf loader javascript demo," https://gkjohnson.github.io/urdf-loaders/javascript/example/bundle/, n.d., accessed: 2025-04-12.

[33] M. E. Harrington, A. B. Zavatsky, S. E. Lawson, Z. Yuan, and T. N. Theologis, "Prediction of the hip joint centre in adults, children, and patients with cerebral palsy based on magnetic resonance imaging," *Journal of Biomechanics*, vol. 40, no. 3, pp. 595–602, 2007.

[34] J. Howard and S. Gugger, "nbdev: A literate programming environment for python," https://nbdev1.fast.ai/, 2020, accessed: 2025-05-04.

MECHANICS AND MARITIME SCIENCE
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS