Including Preference in Anthropometry-Driven Models for Design

Christopher J. Garneau
Pennsylvania State University,
University Park, PA 16802
e-mail: cjgarneau@psu.edu

Matthew B. Parkinson
Department of Mechanical Engineering,
Engineering Design Program,
Pennsylvania State University,
University Park, PA 16802
e-mail: parkinson@psu.edu

1 Introduction

An understanding of the body dimensions and capabilities of the population of potential users can assist engineers in creating artifacts, tasks, and environments that meet goals of fit, safety, and other performance metrics and system-level design targets. Efforts since the 1950s have produced tools that assist in basic assessments of accommodation, the degree to which a design meets the needs of the user population. This work has culminated in recommended tools and practices that are in common use today [1–3]. These indicate to the engineer how to design for the variability in body dimensions (or anthropometry), capability, and age in the target user population. The application of design automation tools to this problem facilitates the simultaneous consideration of these and other aspects of designs [4–7].

There are a number of preference modeling efforts in engineering, business, marketing, and other fields. MacDonald et al. [8,9], for example, developed methods for constructing preference models for individuals and incorporating them into design decisions, particularly as they relate to “green” products. At a higher level, individual preferences are combined into aggregate models representing the preference of a target design population [10]. At the highest level, the preferences of market systems and corporations, and their impact on design and policy are studied [11,12]. The present work focuses on a population model of individual preference, specifically investigating scenarios in which the characteristics such as the body size, shape, and capability of the target user population are strongly correlated with—but do not explain all of—the variability in user requirements. In this and related works, preference is explicitly broken down into two classifications: the portion related to body dimensions correlated with the output measure of interest and the residual variance from these predictions.

The goal of dimensionally optimizing a product with respect to its users is to accommodate a certain percentage of those users [3], often through adjustability, the creation of separate sizes, or both. First, the target user population is identified, then the spatial requirements for accommodating the variability exhibited by this population are determined. There are three general approaches to achieving user accommodation: manikins, population models, and hybrids of the two. All three guide the determination of the spatial requirements of the designed artifact. These three are discussed briefly along with a hybrid approach that explicitly incorporates a model of variability not correlated with body dimensions. An experiment was conducted to obtain data for the population and hybrid approaches and to obtain the necessary preference data for the new approach. A univariate design problem is used to quantify the importance of including both components of preference in anthropometry-driven models for even simple design problems.

1.1 Manikin-Based Approach. “Manikins” are typically two- or three-dimensional representations of the human form with external contours intended to represent human body size and shape for design. They exist as 2D templates [13] and as 3D computerized manikins [14,15]. A boundary manikin refers to a body geometry that lies at the limit of acceptability. In general,
accommodation at the boundary is thought to ensure accommodation at interior points, as long as the adjustability of the product dimension is continuous [1]. Consider a design problem in which only one body dimension (e.g., stature) is relevant and both “small” and “large” people must be considered. In the typical boundary manikin approach, only these two cases are considered necessary to describe the upper and lower limits of acceptability. For example, to accommodate 95% of the population, one might use the 2.5th percentile and 97.5th values of the measure of interest as boundary cases. Note that since the distribution of any particular body measure across a population is generally continuous, the specific level of accommodation (95%) could be achieved by targeting any appropriate span (e.g., 0th to 95th percentile, and 2nd to 97th percentile). Generally the range is selected to minimize the amount of adjustability or material (and therefore cost) required. For a single variable this is often the lower (0th to 95th percentile) or central (2.5th to 97.5th percentile) portion of the distribution.

Stature and body mass index (BMI), a measure of weight-for-stature, are often used to describe the overall size of an individual. They are easily obtained experimentally and are the most readily available anthropometric data for various populations. Once a target user population is identified, the distribution of stature and BMI for a similar population is obtained from an anthropometric database (e.g., NHANES [16]) and the sizes of small and large virtual users are determined. In the event that the true measure of interest is something besides stature, a number of methods including proportionality constants (Fig. 1) or boundary constants can be utilized [17,18]. These represent the average length of a particular body segment as a proportion of stature. Both approaches have a number of limitations, however [18].

Alternatively, data from anthropometric surveys can be used to obtain actual distributions of segment lengths. For example, the ANSUR database [19,20] consists of 240 measurements taken from thousands of male and female U.S. Army personnel in the late 1980s. From these data, statistical measures such as the nth percentile value of some segment length can be calculated and used for design purposes. However, as with any anthropometric database, caution must be exercised when using a data set from one population to represent a different population (e.g., the body dimensions of a military population can vary widely from those observed in a civilian one).

For multidimensional problems, a similar approach is sometimes used, in which small and large (by stature) manikins are created and adjusted to fit relevant product dimensions. For example, a designer seeking an overall accommodation of 95% in a vehicle interior might assume that if 95% of the population is accommodated on headroom, and 95% on legroom, and 95% on seat adjustability, the overall goal has been achieved. However, considering each dimension separately and then combining the results will usually yield a significantly smaller than desired accommodation [21] since it is unlikely that the same 5% are accommodated on each measure. Consequently, methods incorporating principal components analysis have been developed to simultaneously consider the variability exhibited by a population across multiple anthropometric measures. These types of analyses yield many (e.g., 17 in Ref. [22]) “boundary” manikins, which are used to evaluate the design, rather than just two or three. If all are accommodated, the engineer expects to achieve the level of accommodation equivalent to the amount of anthropometric variability encompassed by the selected manikins. Unfortunately, this approach only works in extremely constrained cases, can produce misleading results, and provides little information when the design is unable to accommodate one or more of the manikins [23,24]. This paper only considers a univariate case, so these limitations are not fully investigated.

These boundary manikin approaches all intend to quantify the anthropometric variability expected within the target user population. However, realistic posturing is required for any use of manikins in design [4]. In the purely manikin-based approaches, this posturing may be performed manually by the designer, but to improve accuracy and repeatability it is increasingly done algorithmically. This algorithmic posturing may be accomplished by gathering data on a number of people performing similar tasks to the one in question (e.g., vehicle ingress/egress, sitting in a seat, interacting with an artifact, etc.) and subsequently creating statistical models of these data to describe the mean anticipated behavior as a function of anthropometric and task conditions [25,26]. Algorithmic posturing may also be achieved by using optimization to determine postures and interaction by minimizing objective functions such as torque at a joint and energy required [27]. However, this approach can produce inaccurate postures and lead to erroneous analyses.

1.2 Population Model Approach. An alternative to the boundary manikin approach is the task-oriented percentile model, described in Ref. [3]. These models are created from experimental data taken from a sample population performing a task related to the dimension under consideration. Spatial or adjustability requirements are then defined by the selections or capabilities of the desired proportion of users. Both a sufficiently large representative sample population and a workable prototype are required. This approach forms the basis for the SAE International recommended practices [2], which are used for vehicle design.

These models are an improvement on manikin-based approaches in some ways, since they specifically model the outcome measure of interest, e.g., reach, eye location, and driver-selected seat position, rather than trying to predict the population distributions of those outcomes from boundary cases defined by anthropometry. Consequently they do contain a quantification of preference in that the range of preferred conditions is contained within the data. However, they (1) require extensive human-subject data from a similar task scenario, parametrized using the design variables of interest and representing a large amount of variability on the population descriptors, e.g., body dimensions; (2) have not historically been parametrized for population attributes; and (3) they are essentially univariate, dealing with only a single outcome measure (e.g., preferred seat height) at one time.

1.3 Hybrid Approaches. In practice, some have sought to expand on the manikin- and population-based approaches, by combining them into a hybrid method that has some of the advantages of each. Regressions of experimental data predict the outcome measure of interest as a function of related anthropometry
(such as stature). This allows the population model to be extrapolated to populations different from the one from which the data were gathered. Unfortunately this procedure results in a practice where artifacts are designed to meet the mean behavior associated with a particular body size (i.e., two people with the same predictor value, such as stature, will have the same predicted performance/preference). This ignores the residual variance in the experimental data, which describes how the preference of an individual of a given size differs from the mean.

Recent research expands on this hybrid method by including a second component that models preference unrelated to body dimensions in the outcome measure of interest. Reed and Flannagan [23] investigated the effect of variability unrelated to body dimensions on a driver seat position and eye location in an automobile. Reed and co-workers [4,24] demonstrated a methodology for integrating design automation tools such as optimization and robust design methodologies with models of anthropometric and behavioral variability that include preference. The present work expands on this research by generally considering how variability unrelated to body dimensions relates to traditional design approaches that consider only anthrometry. A case study is presented involving the adjustment of the seat height of a particular exercise cycle. This simple problem was selected because the preferred cycle seat height for an individual should be predicted very well by a single anthropometric measure such as stature or leg length.

2 Methodology

This paper considers several common ways in which manikins and population models are applied in product design. The product is an upright exercise cycle (PRO-FORM XP70). The metric of interest is the minimum seat height and its range of adjustability. In each scenario, the tools or methods are applied such that the engineer or designer would expect to accommodate the central 95% of the target user population. The first two methods utilize traditional boundary manikin approaches considering only anthrometry to estimate the preferred seat height. As discussed previously, one advantage to these methods is that they do not require any experimental data; either hip height or stature is used as the measure of interest. Following these, the use of a population model is illustrated. Then a hybrid manikin/population model that predicts the mean behavior for a given body size is illustrated. Finally, these are all compared against a new methodology in which the residual variance in the regression model is retained to account for variability in preference not correlated with body dimensions.

This case study is used only for the illustration of various methods of dimensionally optimizing a product. The actual results are highly dependent on the product and task to be performed, so general conclusions regarding exercise cycle seat height may not be drawn from this study. The target users are taken to be the male ANSUR population [19], selected because of the extensive anthropometric measures available for the sample.

2.1 Boundary Manikins: Proportionality Constants and ANSUR. In the first two basic approaches toward determining the range of seat adjustability, the way in which users will sit on the cycle, and therefore their desired seat height, is entirely unknown. It is proposed that users will sit such that their heel just touches the ground while seated (with legs unbent). Such an assumption may be based on personal experience or company best practices. This means that the seat height setting for each user is assumed to roughly be equivalent to hip height. Boundary manikins with dimensions belonging to the 2.5th and 97.5th percentile persons are taken to represent 95% of the population.

Following common practice, the first approach uses two boundary manikins with hip height derived from the 2.5th and 97.5th percentile statures (for males) using the proportionality constants in Fig. 1. This approach is termed manikin-k (boundary manikins using proportionality constant k). The second approach uses two boundary manikins with hip height taken directly from the actual 2.5th and 97.5th percentile hip heights from ANSUR. This approach is termed manikin-ANSUR.

2.2 Population Model. Experimental data are required for the creation of the population model. Stature, sitting height (which is used to derive leg length), and preferred exercise bicycle seat height were measured for 42 male engineering students aged 20–26 in a study conducted at Pennsylvania State University. The study received approval from an appropriately constituted internal review board. Since females are not adequately represented in the sample, the restriction that only adult males are to be accommodated is imposed. This leads to our choice of using only ANSUR males as a representative database but does not impact the validity of the results or the methodology.

Each participant was asked to get on the exercise cycle, pedal a few revolutions, and then adjust seat height setting. Participants could repeat the process as many times as they felt necessary to achieve a desired position. After completing this task, the height of the seat top from the ground was recorded for each participant. A stature adjustment of 25 mm was included to account for the thickness of shoes (stature measures are generally taken without them). The bicycle seat post provided 243 mm of adjustability evenly distributed across 11 discrete adjustment locations. Despite limits on the range of adjustability, the data are not thought to be censored since no member of the sample expressed a preference for a higher or lower setting than those available.

To create the population model, statistical analysis is performed directly on the experimental data. The data were assumed to be normally distributed and the 2.5th and 97.5th percentile values are determined using the mean and standard deviation. These results are termed population model.

2.3 Hybrid: Mean Behavior. In a hybrid of the manikin and population model approaches, linear regression analysis is performed using the selected seat height and anthropometry from the participant population. Stature ($R^2=0.41$), leg length ($R^2=0.34$), and both measures ($R^2=0.42$) were considered as predictors, with stature being the ultimate choice for the seat height preference model. The resulting regression line is used to predict the selected seat height of two boundary manikins, characterized by 2.5th and 97.5th percentile statures. This approach is termed hybrid-mean.

2.4 Hybrid: With Residual Variance. Using a methodology described in Refs. [4,28], the fifth approach uses the regression parameters of the seat height preference model to generate a virtual population of 1000 users randomly sampled from the ANSUR database. The preferred seat height is calculated from two components. The first is calculated as in the hybrid-mean approach, using anthropometric predictors and a regression model. The second component explains how an individual’s preference deviates from the mean. This is done by randomly sampling from a normal distribution about the mean seat height for that stature. This distribution is characterized by a mean equal to the seat height predicted by the regression equation and a standard deviation equal to the root-mean-square error of the regression (i.e., $\sigma=\text{RMSE}$) [29]. For this reason, this method will be termed hybrid-ResVar.

3 Results

The adjustment ranges calculated using each of the methods are reported in Table 1. The engineer anticipates achieving 95% accommodation with each of the methods, although they all produce different results. The results are all compared with the hybrid-ResVar method, which has been shown in other studies to accurately model and predict this kind of experimental data [23,30]. This comparison is performed by imposing the limits of adjustment prescribed by each of the methods on the virtual population created by the hybrid-ResVar method. The numbers of virtual
people that lie within the limits are considered accommodated, and percentages reflecting the proportion of accommodated users appear in Table 1.

3.1 Boundary Manikins: Proportionality Constants and ANSUR. The 2.5th and 97.5th percentile values of male stature from ANSUR are 1625 mm and 1887 mm, respectively. For the manikin-\(k\) method, hip height is found as a proportion of stature \((H=0.530)\), and so the lower and upper values are 861 mm and 1000 mm. Following the assumption that hip height equals selected seat height, and adding the 25 mm shoe thickness, the lower and upper limits of seat height adjustment become 886 mm and 1025 mm. For the manikin-ANSUR method, hip height is found directly from ANSUR; the 2.5th and 97.5th percentile hip height measures are 835 mm and 1022 mm. Therefore, the lower and upper limits of seat height adjustment are 860 mm and 1047 mm, with shoe thickness added.

3.2 Population Model. The mean of the seat height selections of the sample population is 943.2 mm and the standard deviation is 46.3 mm. The data are assumed to be normally distributed and a \(k\)-value of 1.96 is selected. The resulting lower and upper limits for the central 95% of the distribution are 852 mm and 1034 mm, respectively.

3.3 Hybrid: Mean Behavior. Figure 2 shows the plot of selected seat height versus stature (with shoes) for the sample population. The equation for the resulting regression line is given by

\[
P = 0.515S + 22.1
\]  

where \(P\) is the selected seat height position and \(S\) is the stature. \(R^2\) for the regression is 0.41 and the RMSE is 35.2 mm. Table 1 denotes the results of entering the 2.5th and 97.5th percentile statures (with shoe thickness added) into Eq. (1), which subsequently defines the limits of required seat height adjustment for the hybrid-mean method.

Table 1  Summary of the results for all methods, showing the required low and high seat settings, adjustable range, and percentage accommodated (compared with hybrid-ResVar method). Dimensions in millimeters.

<table>
<thead>
<tr>
<th>Method</th>
<th>Low (mm)</th>
<th>High (mm)</th>
<th>Range (mm)</th>
<th>% accommodated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manikin-(k)</td>
<td>886</td>
<td>1025</td>
<td>139</td>
<td>82.5</td>
</tr>
<tr>
<td>Manikin-ANSUR</td>
<td>860</td>
<td>1047</td>
<td>187</td>
<td>93.2</td>
</tr>
<tr>
<td>Population</td>
<td>852</td>
<td>1034</td>
<td>181</td>
<td>93.6</td>
</tr>
<tr>
<td>Hybrid-mean</td>
<td>872</td>
<td>1007</td>
<td>135</td>
<td>82.5</td>
</tr>
<tr>
<td>Hybrid-ResVar</td>
<td>844</td>
<td>1036</td>
<td>192</td>
<td>95.0</td>
</tr>
</tbody>
</table>

3.4 Hybrid: With Residual Variance. As Fig. 2 shows, there is, as expected, a strong correlation between stature and the outcome measure of interest, preferred seat height. There is also, however, a large amount of variance in the data. This could be due to a number of factors including cycling skill level, the relative position of the handlebars (which was fixed for this study), and comfort. Rather than attempting to quantify these factors individually (which may certainly be appropriate in some cases), they are lumped into a single stochastic term as described below.

Figure 3 shows the distribution of the stature of the 1000 virtual users selected from the ANSUR database for use in the hybrid-ResVar method. These statures are selected at random from the entire pool of male ANSUR data and are assumed to be approximately normally distributed. The mean of the distribution is 1756 mm, identical to the mean of all males in ANSUR. Figure 4 shows the result of plotting these statures using the regression line of Eq. (1). Introducing the second term describing preference not correlated with body dimensions gives

\[
P = 0.515S + 22.1 + N(0, 35.2)
\]  

where \(N\) is a normal distribution with a mean of 0 and a standard deviation of 35.2 (RMSE of the regression). Figure 5 shows the 1000 seat positions predicted for the 1000 statures using Eq. (2). 95% accommodation is achieved by selecting the central 950 users with their seat height selections placed in order. Note that since this is a stochastic simulation, the individual statures and
predicted seat heights will vary from one simulation to the next, but the outcome (predicted seat height percentile cutoffs) does not vary substantially across simulations.

Figure 5 also compares the results of the manikin-\(k\), manikin-ANSUR, hybrid-mean, and population model methods with the result of the hybrid-ResVar method. The band of accommodated users given by the defined lower and upper limits in Table 1 is shown for each method by the bars at left, and the results for the hybrid-ResVar method are extended across the figure by the solid lines. The points above or below these limits represent disaccommodated users. Table 1 also denotes the number of accommodated users, expressed as a percentage.

4 Discussion

The results show the wide range of recommended designs that these various methods—all of which are in common use—provide to this extremely simple univariate design problem. They also indicate the importance of including both sources of preference, that which correlates with body dimensions and that which does not, in spatial analyses. For this simple example, the range of recommended designs is large and the accommodation estimates based on the experimental data show that the configurations exhibit far too little adjustability (manikin-\(k\) and hybrid-mean) or adjustability that might be better allocated (manikin-ANSUR and population). The bias errors primarily affect accommodation in the tails of the distribution, which lessens the number of potential users impacted. Nevertheless this is a simple, univariate example and these types of errors are usually additive and would be compounded for multivariate problems.

The effect of including the residual variance is seen in the hybrid-ResVar method. Since the regression model relating stature to seat height has an \(R^2\) value of 0.41, much of the variance in seat height cannot be explained exclusively by stature. The effects are clearly visible when comparing Fig. 4, which includes no preference, with Fig. 5, which does. It shows that when the stochastic component is included, for any particular stature, a variety of seat positions might be chosen about a mean value. The benefit of this approach may be seen in the 12.5% increase in accommodation from the hybrid-mean method to the hybrid-ResVar method, corresponding to 125 additional people from the virtual 1000-member sample.

Another way to think of the hybrid-ResVar method is the separation of user behavior into two quantifiable components: one defined by anthropometry (measured by the slope of the preference model regression) and a second not explained by anthropometry (measured by RMSE). The interrelation of anthropometry and preference may be easily examined by the shape and scatter of the preference model. In cases where anthropometry dominates the response, the model will exhibit a high \(R^2\) value and there will be little residual variance. In these situations, traditional manikin analyses may be sufficient. In contrast, design scenarios in which preference unrelated to body dimensions dominates will produce a model with a low \(R^2\) and a large amount of residual variance. When this occurs, population models or hybrid approaches are more likely to produce accurate results. For this example, the \(R^2\) is 0.41, so anthropometry has a bearing on the problem and guides the designer, but differences that cannot be explained by body size are a very important component.

Stature was used as the predictor for the hybrid models because it correlated as well with the outcome of interest (seat height) as any other anthropometry. Although additional measures could have been incorporated into the model, this often does not appreciably improve the predictive power when involving measures of length (e.g., leg length, arm length, and sitting height) since they correlate well with stature. In this particular case, stature produces even better results than leg length. This is likely because in gripping the handlebars of the bicycle, measures beyond leg length such as torso and arm length were a factor in selecting preferred seat heights. It is important, however, not to ignore the possibility that additional anthropometric measures could significantly improve the regression in more complicated or multivariate problems. Not all design problems will be neatly described by a single anthropometric measure such as stature. Some problems may require, at a minimum, an additional predictor for measures of breadth (e.g., hip breadth and shoulder breadth). Ultimately, however, there will be a portion of the variability in the data that cannot be explained by any number of additional anthropometric measures, and it is important not to ignore this component.

Ignoring variability that is not correlated with body dimensions can result in designs with too much adjustability (thereby increasing cost unnecessarily) or, more commonly, too little adjustability (causing users to interact with the artifact, task, or environment in unexpected ways). Designers are often surprised when actual use indicates lower than expected accommodation and are unable to explain how this occurred. To prevent this under-approximation of accommodation, designers might overcompensate by allocating an extremely large amount of adjustability to the problem, such that an accommodation level approaching 100% is expected. This would not constitute an optimal design, however, and in multivariate problems can be extremely impractical. Although the differences in results achieved using some of these methods (Table 1) may seem small or trivial for this application, many other applications that have limited physical space for adjustability or high costs for added adjustability benefit greatly from optimized solutions.

Although the hybrid-ResVar method illustrated here still requires experimental tests, one advantage of the method is that data from a relatively small sample of users evaluating physical prototypes can be used to make models that can be used for quantitative analysis. In practice the sample population would deliberately contain a diverse group of people, with care taken to oversample the tails of the distributions of relevant parameters (e.g., lots of short and tall people in this example).

This study has limitations. Although the results of the different analyses exhibit many of the issues prevalent in industry, they are not necessarily typical of every analysis of that type. For example, proportionality constants tend to under-approximate the amount of adjustability required, but not by a fixed percentage. Similarly, the results obtained using the population model were relatively good since the design population and the population from which the data were gathered were relatively similar (i.e., the data were collected from males aged 20–26 and the design population was males in the U.S. military). Were these populations more distinct, the population model approach would be less useful since it is not reconfigurable to new populations.

Other limitations include the size and composition of the data sample. A larger, more diverse sample would be required to validate the model across an entire population. Further study will seek...
to enlarge the sample size, particularly in the tails of the anthropometric distribution. To include females in the pool of prospective customers, data would need to be collected from a large sample of female users as well. The inclusion of different types of bicycles would allow the model to be extrapolated more broadly. The mode of adjustability should also be improved, reconfiguring the cycle in such a way so as to have continuous, rather than discrete, adjustment. The preferred seat height was selected using a “quick sit” methodology. This might correlate well with acceptability in a store or for a short ride but is likely to differ from what might be selected for longer duration rides. Additionally, the expertise of the rider was not considered; it is likely that more experienced riders will select relatively higher seat heights.

One could imagine that a design with more than one adjustable dimension experiences a compounding of the unexplained variance seen in this problem. Therefore, it is even more important in such problems to include effects owing to preference. Future work will further examine the effects of preference on higher-dimensional problems. The notion of a “just noticeable difference,” that is, the degree to which a user could be accommodated without experiencing negative effects, and its impact on preference and accommodation will also be investigated.

Typical approaches to designing for human variability focus on the spatial variability in body dimensions. Other sources of variability are often important and should be considered in assessment. This work shows that incorporating them into traditional statistical models improves accommodation and the confidence the engineer can have in their design.

Acknowledgment

We gratefully acknowledge the support of the Department of Mechanical Engineering, the Schreyer Honors College, and the Engineering Design Program at the Pennsylvania State University. This work was partially funded by the National Science Foundation under Award No. 0729386.

References