Research Master's programme: Methodology and Statistics for the Behavioural, Biomedical and Social Sciences
Utrecht University

Research Report Anaïs Fopma (6199356)
Title: A blended distance to define "people-like-me"
Date: 09-05-2022

Supervisors: Prof. Dr. Stef van Buuren, Dr. Gerko Vink & Mingyang Cai

Preferred journal of publication: Statistics in Medicine
Word count: —-

Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

MASTER'S THESIS

A blended distance to define "people-like-me"

Anaïs M. Fopma*

[1]Methodology & Statistics, Utrecht University, the Netherlands

Correspondence
*A.M. Fopma, Padualaan 14, 3584 CH Utrecht,the Netherlands. Email: a.m.fopma@uu.nl

Present Address
Padualaan14,3584CH Utrecht, the Netherlands

Summary

Curve matching is a powerful tool to predict the development of a child (the target) with the data of other children (donors). The technique relies on predictive mean matching, which matches donors that are most similar to the target based on the predictive distance. Even though this approach ensures high prediction accuracy, there are two disadvantages. Firstly, it requires users of curve matching to select a particular future time point to base the matches on. In some cases, it may be difficult to choose this time point. Secondly, the predictive distance may make matches look unconvincing, as the profiles of the matched donors can substantially differ from the profile of the target, even if they are close on the predicted time point. To counterbalance these disadvantages, similarity between the curves of the donors and the target can be taken into account when selecting donors. The objective of the current study is to do so by combining the predictive distance with the Mahalanobis distance, thus creating a 'blended distance' measure.

KEYWORDS:
Curve matching, predictive mean matching, distance measures, metrics

1 | INTRODUCTION

The first three years of childhood form a crucial stage in determining children's subsequent development and health outcomes.[1] For this reason, growth monitoring is considered to be an integral part of paediatrics. It can aid in the identification of problems in development such as growth stunting, and ensure timely treatment or intervention to improve the child's health.[2] However, growth monitoring solely provides insights in the past and current developmental stages of the child. Growth curve modeling, on the other hand, can be used to predict future development. It could therefore provide more specific answers to questions health professionals, parents, and insurance companies may have, such as: 'Given what I know of the child, how will it develop in the future?' and 'Does this child get the most effective treatment available?'[3]

1.1 | Curve matching

An approach currently used for growth curve modeling is curve matching. Curve matching[3] is a nearest neighbour technique for individual prediction that constructs a prediction by aggregating the histories of "people-like-me". Its aim is to predict the growth of a target child by using the data of other children that are most similar to the target child.

In order to select these donors, some form of similarity needs to be defined to match the donors to the target child. Therefore, the key question is: How are good matches obtained? The current approach uses predictive mean matching (PMM). PMM makes use of an existing donor database, containing the growth data of children who are older than the target child, and of which information at a later age is available. The first step is to fit a linear regression model on the donor database. Then, this model is

used to predict the values for all donors and for the target at a certain point in the future, for example at 14 months. Finally, the distance between the predicted value of each of the donors and the predicted value of the target is calculated, which is referred to as the predictive distance. A number of donors – usually five - with the smallest predictive distance are selected as the best matches. Their growth curves are then plotted and point estimates can be calculated by averaging the measurements. The growth patterns of the matched children thus suggest how the target child might develop in the future.

An advantage of this technique is its high prediction accuracy.[3] Moreover, the applications of curve matching can be extended to settings other than the prediction of child development, such as patient recovery after an operation, prediction of longevity, and decision-making when multiple treatments are available.[3]

1.2 | Alternative approach

Even though PMM has proven to be promising in growth curve matching,[3] there are two reasons to move beyond the predictive distance used in PMM and investigate an alternative metric. Firstly, PMM requires users of curve matching to select a particular future time point to base the matches on (e.g. 14 months of age). In some cases, it may be difficult to choose this time point, especially when the 'future' is more vaguely defined as a time interval.[4] Secondly, the predictive distance may make the matches look unconvincing. The trajectories of the selected donors may all be close to the prediction for the target child at 14 months, but this does not imply that the histories are identical. After all, different profiles may lead to the same predicted value. Consequently, the curves of some of the matches may be quite far from the curve of the target child. Some users of curve matching feel that such discrepancies are undesirable, as these matches do not appear to be *people-like-me*.[4]

For these reasons, the practical implementation and use of curve matching can be improved by combining the predictive distance with the Mahalanobis distance, thus creating a "blended distance" measure. This blended metric would take into account historical similarity, by giving more weight to similarities between units in the full predictor space. The objective of this study is to investigate what the properties of such a blended distance measure are.

2 | METHODS

Two simulations studies will be conducted. The following sections describe each study in accordance with the ADEMP-structure for reporting simulation research.[5] The different versions of the blended metric (i.e. the methods), the aims of the study, the data-generating mechanisms, and the estimand and performance measures are discussed. In addition, an application of the blended metric to an empirical data set is outlined.

2.1 | Simulation study I

2.1.1 | Blended metric

The blended distance measure to be evaluated is a weighted version of the predictive distance and the Mahalanobis distance. As described above, the predictive distance is the distance between the predicted value of a donor and the predicted value of the target at a particular future time point. The Mahalanobis distance is defined as the distance between two N dimensional points scaled by the variation in each component of the point. For example, if $\vec{x}$ and $\vec{y}$ are two points from the same distribution which has covariance matrix $\mathbf{C}$, then the Mahalanobis distance is given by

$$((\vec{x} - \vec{y})' \mathbf{C}^{-1} (\vec{x} - \vec{y}))^{\frac{1}{2}}. \tag{1}$$

Two potential versions of the blended metric will be compared: one that uses ranking and one that uses scaling. The ranked blended metric is created as follows. First, the predictive distance and the Mahalanobis distance are calculated for each donor. Then, the k donors with a low value for both the predictive distance and the Mahalanobis distance are selected. In order to do so, the rank is calculated for the predictive distance PD and the Mahalanobis distance MD, where ties are randomly broken. The ranked blended distance RBD is then given by:

$$RBD = p \cdot \text{rank}_{PD} + (1 - p) \cdot \text{rank}_{MD}, \tag{2}$$

where rank_{PD} is the rank for the predictive distance PD, rank_{MD} is the rank for the Mahalanobis distance MD, p is the blending factor (or weight) assigned to rank_{PD}, and $0 \leq p \leq 1$. The k donors with the lowest values on RBD are selected as the best matches.

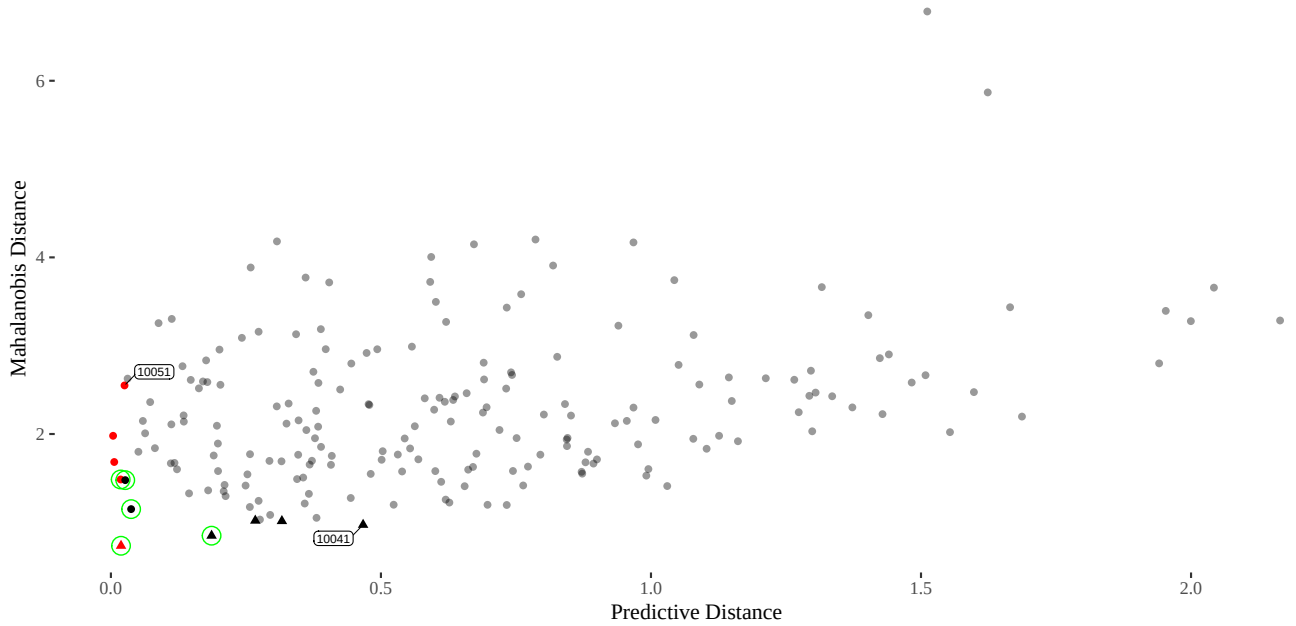

FIGURE 1 Mahalanobis distance plotted against predictive distance for each of the 199 donors. The donors in red are the five matches with the smallest predictive distance, the triangular donors those with the smallest Mahalanobis distance, and the donors circled in green those with the smallest blended distance.

The scaled blended metric is created similarly, but scales the predictive distance *PD* and the Mahalanobis distance *MD* before combining them. The scaled blended distance *SBD* is then given by:

$$SBD = p \cdot \frac{PD - \bar{x}_{PD}}{\sigma_{PD}} + (1 - p) \cdot \frac{MD - \bar{x}_{MD}}{\sigma_{MD}}, \tag{3}$$

where $\bar{x}_{PD}$ is the mean of the predictive distances, σ_{PD} their standard deviation, $\bar{x}_{MD}$ is the mean of the Mahalanobis distances, and σ_{MD} their standard deviation.

In theory, these two versions of the blended distance should yield identical results. However, the scaled version would be computationally more efficient. Both versions are included in this study to confirm that they produce the same results, and if this is indeed the case, the scaled version could be implemented in the `mice`[6] package.

As an example, the blended distance is illustrated in Figure 1. Here, the data of 200 children from the *Sociaal Medisch Onderzoek Consultatiebureau Kinderen* (SMOCK) study are used.[7] The first subject is taken as the target, the 199 other subjects as the donors. For all donors, the Mahalanobis distance for the measurements during the first six months of growth is calculated. In addition, the predictive distance between each donor and the target is calculated. In the figure, the Mahalanobis distance and predictive distance are plotted against each other. The red donors are the five matches with the smallest predictive distance, where especially subject 10051 has a large Mahalanobis distance. The triangular donors are the five matches with the smallest Mahalanobis distance, where especially subject 10041 has a large predictive distance. A weighted blended distance measure would balance the distance measures, such that the donors with a low value for both distance measures are chosen. These are circled in green.

In the current study, blending factors of respectively 1, 0.5, and 0 will be evaluated for the blended metric. A blending factor of 1 implies that the blended distance is equal to the predictive distance, whereas a weight of 0 implies that it is equal to the Mahalanobis distance. Therefore, a blending factor of 0.5 gives equal weight to both distance measures. PMM will be used as a reference in order to evaluate whether we do indeed obtain the same results with a blending factor of 1. Using PMM, both the ranking and scaling methods with three different blending factors each, this results in seven different versions of the blended metric to be evaluated.

2.1.2 | Aims

The main objective of this study is to investigate what the properties of the blended metric are. More specifically, we want to answer the following questions:

1. Do a ranked and a scaled version of the blended distance measure yield identical results?
2. Does a blending factor of 1 yield results identical to those obtained by PMM, as intended?
3. How is the performance of the blended metric related to the missingness mechanism, the proportion of missingness in the data, the distribution of the data, and the correlation in the data?
4. Is there a penalty from blending in terms of reduced predictability?

It is expected that the ranked and scaled versions do yield identical results, and that blending with a factor of 1 does indeed give the same results as PMM does. As pointed out before, PMM has been shown to result in high prediction accuracy. Therefore, it is expected that the predictability of the blended distance will decrease as the blending factor favours the Mahalanobis distance. When the correlation in the data is low, the prediction model will fit poorly and the blended metric is expected to perform worse when more weight is given to the predictive distance. When the correlation in the data is high, the prediction model will fit better, and the prediction model will explain more variance in the outcome. In this case, the blended metric is expected to perform better when more weight is given to the predictive distance. Finally, it is expected that the blended metric will perform worse in skewed data, when more weight is given to the Mahalanobis distance.

2.1.3 | Data-generating mechanisms

I don't know if this section is still correct after redesigning the simulations? In order to answer the previous questions, the blended distance measure will be evaluated in simulated data that meet different conditions. All data are generated from one of 24 data-generating mechanisms, with equal means, but with varying missingness proportions, missingness mechanisms, distributions, and variance-covariance matrices.

Three continuous predictor variables $\mathbf{X_1}$, $\mathbf{X_2}$, and $\mathbf{X_3}$ are defined, corresponding to the standardised height measurements (Z-scores) at birth, at 1 month, and at 2 months. One continuous outcome $\mathbf{Y}$ is defined, corresponding to the predicted standardised height measurement at the age of 14 months.

The distribution of the data is varied over two conditions. The data generating mechanism of the predictor space is a multivariate normal distribution for the first condition, and a strongly skewed multivariate distribution for the second condition, $\mathbf{X} = \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, with mean vector $\boldsymbol{\mu} = [0, 0, ..., 0]$. In order to achieve this, the predictors are transformed,[8] where

$$X = X$$

for the first condition and

$$X = X^{12}/max\{X^{11}\}$$

for the second condition.

The correlation in the data is varied over three conditions. The covariance matrix Σ for the populations with two predictors is given by:

$$\Sigma = \begin{bmatrix} 1 & \sigma^2\rho & \sigma^2\rho \\ & 1 & \sigma^2\rho \\ & & 1 \end{bmatrix},$$

where the off-diagonal elements are 0 for the first condition, 0.1 for the second condition, and 0.7 for the third condition.

The proportion of missingess in the outcome variable is varied over two conditions. The first condition simulates a setting with 25% missingness, the second a setting with 50% missingness.

Finally, the missingness mechanism is varied over two conditions. The first concerns a missing completely at random (MCAR) mechanism, where missingness does not depend on the values of the data, missing or observed.[9] The second concerns a missing at random (MAR) right mechanism. This means that missingness does depend on the data, but only through observed components of the data.[9] Is there a reference I can use to explain that the latter is the more extreme mechanism?

We consider a full-factorial simulation study design, where each of the possible combinations of weighting and data-generating mechanisms are evaluated. As there are seven different methods to be evaluated and 24 different data-generating mechanisms, the simulation will yield 168 results. From each data-generating mechanism, a sample of size 500 is drawn. The number of simulations run for each setting is set to 1000.

Note to self: explain that the simulation is seed-dependent, and add the same simulation with a different seed to the research archive like Hanne said

2.2 | Estimand and performance measures

The estimands of interest in this study are the predicted or imputed values. To assess the performance of each metric under each combination of conditions, the parameter estimate (qbar), standard error (se), total variance about the parameter estimate (t), degrees of freedom (df), variance between imputations (b), upper and lower 95% confidence limits, true value, coverage, and bias are computed.

2.3 | Simulation study II

2.3.1 | Aims

The objective of the second simulation study is to evaluate the version of the blended metric that performed best in Simulation study I. In this second study, a setting will be simulated that reflects the practical implementation of the blended metric in the prediction of the height of a target child by imputing a single value. The blended metric will be evaluated with blending factors ranging from 0 to 1, with intervals of 0.1.

2.3.2 | Data-generating mechanisms

Data is simulated from a single data-generating mechanism. The conditions chosen are the ones under which the blended metric performed the worst in Simulation study I. Therefore, the same variables are defined as in the first simulation study, where the data is skewed and the off-diagonal elements in the covariance are set to 0.7. A sample of size 500 is drawn. To simulate the prediction of the height measurement at 14 months for a single target child, a random case in the data will be made incomplete for the outcome.

2.3.3 | Estimand and performance measures

The missing value is imputed 50 times and evaluated against its true value. The number of simulations is set to 10000 I ended up doing this because the results were a bit weird with only 1000 simulations. The performance measures of interest are the (absolute) bias, sum of squared deviations (SSD), standard error, lower and upper confidence limits, coverage, and root mean square error (RMSE).

2.4 | Study on empirical data

After the simulation study is conducted, the blended metric will be evaluated in an application to empirical data from the SMOCK study.[7] How exactly should I approach this? The weights used will be the same as to those used in the simulation study: 1, 0.5, and 0. The SMOCK database contains the anonymised growth data of 1,933 children aged 0-15 months. In addition, the database contains covariates that influence growth, such as the sex, gestational age, birth weight, and height of the father and mother.

The growth data in the SMOCK database consist of the height measurements of children at different observation times. It is important to note that the actual time points of data collection will sometimes differ substantially from the scheduled times. This may be due to a doctor's visit being planned during a holiday, the subject not showing up at the appointment, or the measurement device being out of order at the time of the scheduled observation.[4] As a consequence, the observation times will vary across subjects, and are said to be irregular. Irregular observation times present significant challenges for quantitative analysis, as it becomes more complex to predict the future from past data. Usually, a linear mixed model with time-varying outcomes is applied. However, an alternative is offered by the broken stick model,[4] which converts irregularly observed data into a set of repeated measures. As a result, each child's growth trajectory can be approximated by a series of connected straight lines. The breakpoints between these lines are set to be the pre-specified, scheduled observation times. The advantage is that repeated measures data

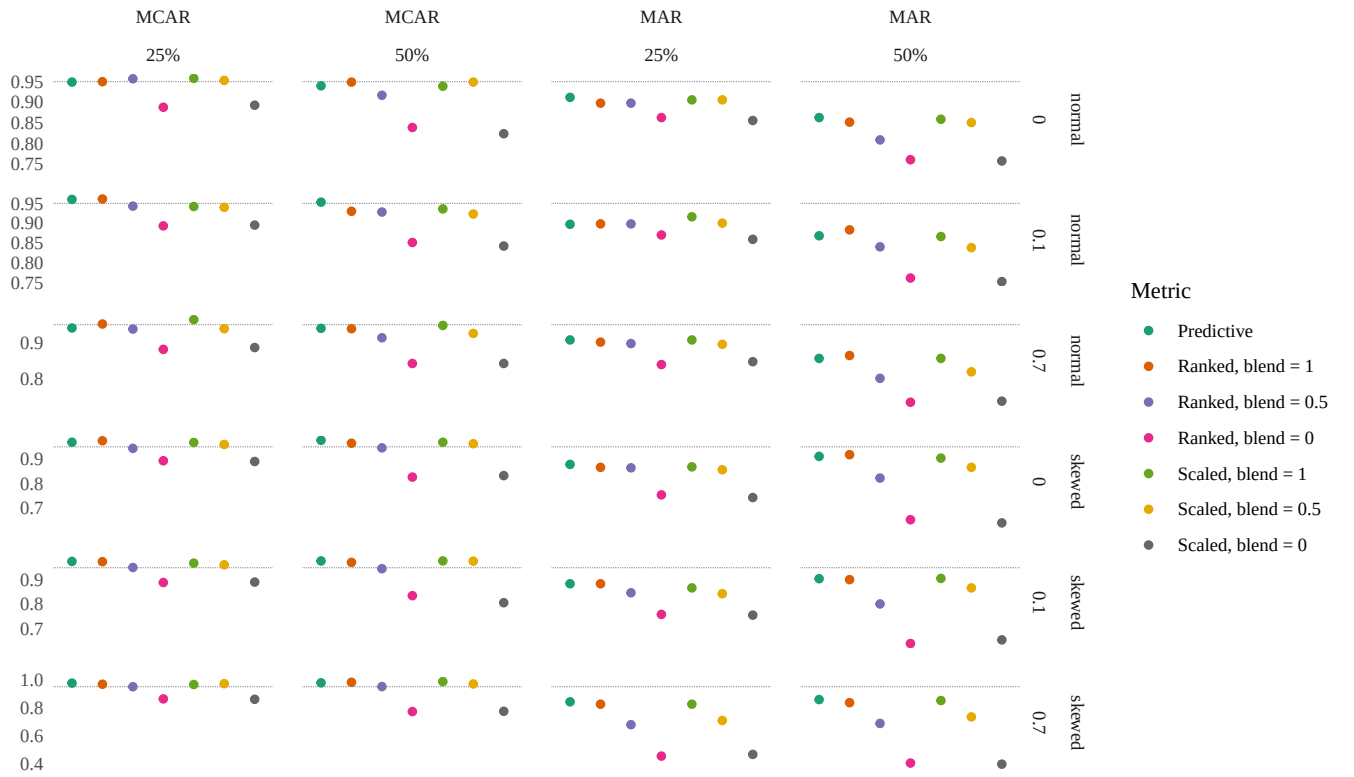

FIGURE 2 Coverage results per condition, where each individual plot shows the results for the seven methods. A reference line is given at coverage = 0.95. Above the plots, the condition combinations of missingness mechanism (MCAR, MAR right) and missingness proportion (25%, 50%) are given. On the right, the condition combinations of distribution (normal, skewed) and correlation (0, 0.1, 0.7) are given.

offer a lot more simplicity than the use of linear mixed models. Therefore, the empirical data will be analysed using the broken stick model.

2.5 | Software

R version 4.1.2 (2021-11-01)[10] will be used to simulate the data and perform the analyses. The `mice.impute.pmm` function in the `mice`[6] package will be used to perform PMM and an adaptation of this function will be used to calculate the blended distance. As the empirical data consist of irregular observation times, the `brokenstick` package[4] will be used for estimating the growth models. Instructions and scripts to reproduce the simulation results are available in the research archive of this project.

3 | RESULTS

3.1 | Simulation study I

The simulation results for each of the seven methods are displayed in Table A1 through Table A7 in Appendix A. In each table, the data-generating mechanisms are specified in the left columns, by indicating the missingness mechanism, missingness proportion, skewness of the distribution, and correlation in the data. The results for the coverage are visualised in Figure 2, those for the bias in Figure 3, and those for the explained variance in Figure 4. I have used free scales for the plots now, but I would rather not change them any further/create new figures. I understand they contain a lot of information, but I think they are useful to give an overview of the results, as the details are explained in the text already. The results are discussed below on the basis of the research questions.

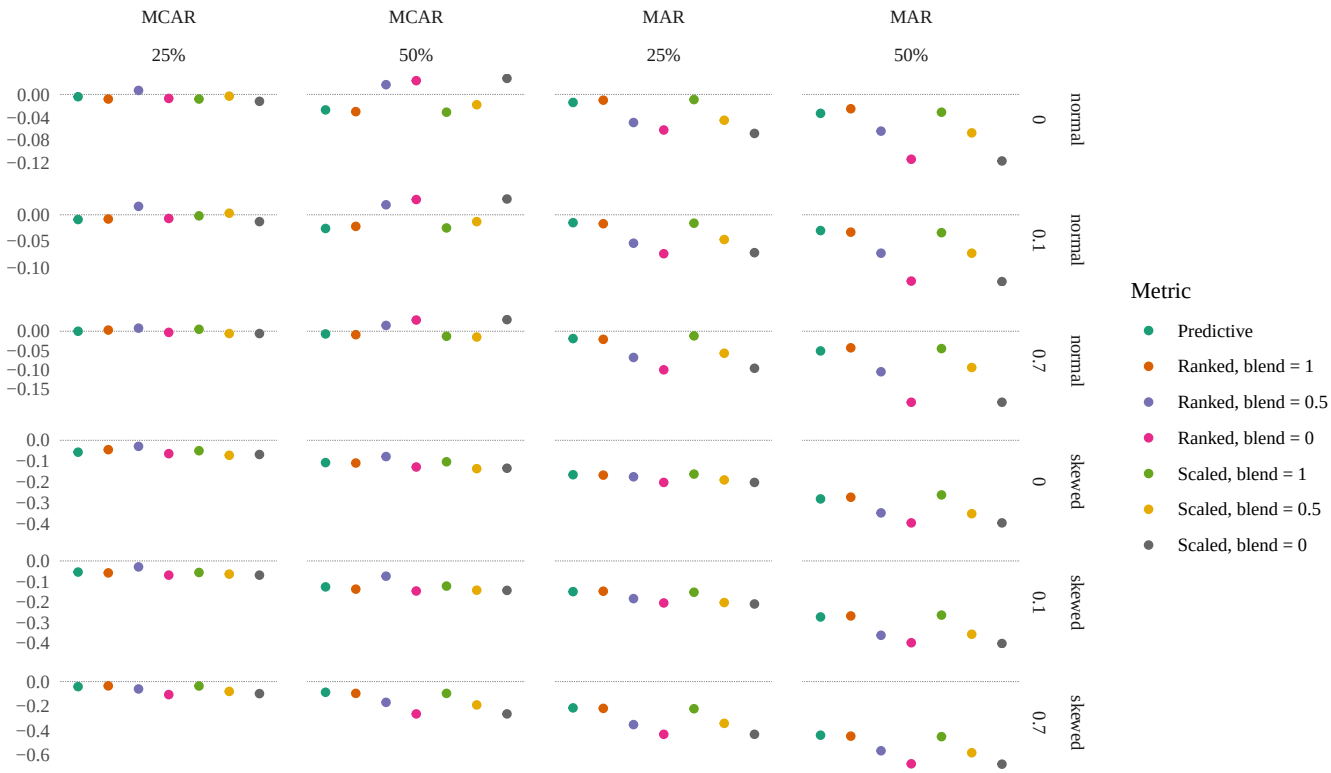

FIGURE 3 Bias results per condition, where each individual plot shows the results for the seven methods. A reference line is given at bias = 0. Above the plots, the condition combinations of missingness mechanism (MCAR, MAR right) and missingness proportion (25%, 50%) are given. On the right, the condition combinations of distribution (normal, skewed) and correlation (0, 0.1, 0.7) are given.

3.1.1 | Comparison of ranked and scaled blended metric

It was expected that the ranked and scaled versions of the metrics would yield identical results. The results show that this is true in some cases, but that not all are identical. When comparing the results for blending factor = 1 in Table A2 and Table A5, they show that the scaled method yields slightly higher coverages but larger biases overall. When comparing the results for blending factor = 0.5 in Table A3 and Table A6, they show that the scaled method outperforms the ranked method. The results for blending factor = 0 in Table A4 and Table A7 show that the ranked method outperforms the scaled method. Overall, the ranked method performs slightly better, except when the blending factor is set to 0.5. As most of the results are similar, however, the scaled version of the blended metric might be preferable, as it is computationally more efficient to use.

3.1.2 | Comparison of PMM and blending factor = 1

In both blended metrics, a blending factor of 1 indicates that full weight is given to the predictive distance. Therefore, using a blending factor of 1 should yield results identical to those obtained by PMM. Even though the results in Table A1, Table A2 and Table A5 are similar, they are not identical. This is likely due to the fact that a matching function written in C is used in the original The `mice.impute.pmm` function but not in the `mice.impute.blended` function. In some cases, particularly in the MCAR conditions, the both the ranked and scaled versions of the blended metric with a blending factor of 1 perform slightly better than the predictive metric.

3.1.3 | Effect of data generation conditions on performance

In the data generating models, the missingness mechanisms, proportions, skewness of the data, and correlation in the data were varied, resulting in 24 different simulation conditions. The impact of each of these conditions on the performance of the metrics

FIGURE 4 R squared results per condition, where each individual plot shows the results for the seven methods. A reference line is given at 0. Above the plots, the condition combinations of missingness mechanism (MCAR, MAR right) and missingness proportion (25%, 50%) are given. On the right, the condition combinations of distribution (normal, skewed) and correlation (0, 0.1, 0.7) are given.

in terms of coverage and bias is mostly as expected, and can be derived from the plots displayed in Figure 2, 3 and 4. The MCAR conditions show higher performance when compared to the MAR right conditions, and a higher proportion of missingness in the data leads to lower performance. The skewness of the data does not always impact the performance negatively. Under the MCAR conditions, a skewed distribution of the data results in higher coverage rates for some cases when compared to a normal distribution. Under the MAR right conditions, however, the opposite is true. The skewness does influence R squared, which is quite stable under the normal conditions but shows more variation under the skewed conditions. Finally, a higher correlation in the data under MCAR conditions does not lead to decreased performance, and in some cases to increased performance. It does lead to lower performance under the MAR right conditions. The proportion of explained variance increases with the correlation in the data, and is the highest under the MCAR condition with 25% missingness, skewed distribution and correlation of 0.7.

3.1.4 | Effect of blending on performance

The effect of the blending factor on performance of the metric can be evaluated for both the ranked version and the scaled version of the metric. For the ranked version, the coverage is almost always higher, the bias is often smaller, and the explained variance is almost always larger when the blended metric is weighted more towards the predictive distance. For the scaled version, the results are similar: in most cases, a higher blending factor leads to better performance.

There are a few exceptions where a blended metric with a blending factor of 0.5 performs better in terms of coverage. The ranked metric with blending factor = 0.5 performs better than both the predictive metric and the ranked metric with blending factor = 1 in the condition of MCAR with 25% missingness, a normal distribution and a correlation of 0.7. Finally, it performs better than the predictive metric but not the ranked metric with blending factor = 1 in the condition of MAR right with 25% missingness, a normal distribution and correlation of 0.7.

The scaled metric with a blending factor of 0.5 outperforms both the predictive metric and the scaled metric with blending factor = 1 in the coverage rate in some cases as well. It performs better than the scaled metric with blending factor = 1, but not the predictive metric, in the conditions of MCAR with 25% missingness and a normal distribution with both a correlation of 0 and a correlation of 0.7, and a skewed distribution with a correlation of 0.7. It performs better than the predictive metric, but not the scaled metric with blending factor = 1, in the conditions of MCAR with 25% missingness, a skewed distribution and correlation of 0, MCAR with 50% missingness, a normal distribution and a correlation of 0.7, and finally, MCAR with 50% missingness, a skewed distribution and a correlation of 0.7.

Additionally, there are exceptions where a blended metric with a blending factor of 0.5 or 0 performs better in terms of bias. The ranked metric with blending factor = 0.5 has a smaller bias than both the predictive metric and the ranked metric with blending factor = 1 in the MCAR conditions with 25% missingness, a skewed distribution and a correlation of 0 or 0.1, with 50% missingness, with both the normal and skewed distributions and a correlation of 0 or 0.1.

The scaled metric with a blending factor of 0.5 has a smaller bias than both the predictive metric and the scaled metric with blending factor = 1 in the MCAR conditions with 25% missingness, a skewed distribution and a correlation of 0 or 0.1, with 50% missingness, with a normal distribution and a correlation of 0 or 0.1.

The ranked metric with a blending factor of 0 has a smaller bias than both the predictive metric and the ranked metric with blending factor = 1 in the MCAR conditions with 25% missingness, a normal distribution and a correlation of 0.1, and with 50% missingness, a normal distribution and a correlation of 0. It performs better than the ranked metric with blending factor = 1, but not the predictive metric, in the condition MCAR with 25% missingness, a normal distribution and correlation of 0. Finally, it performs better than the predictive metric, but not the ranked metric with blending factor = 1, in the condition of MCAR with 50% missingness, a normal distribution, and correlation of 0.1.

The scaled metric with a blending factor of 0 has a smaller bias than both the predictive metric and the ranked metric with blending factor = 1 in the MCAR conditions with 25% missingness, a normal distribution and a correlation of 0 or 0.1, and with 50% missingness, a normal distribution, and a correlation of 0. It also performs better than the predictive distance, but not the scaled metric with blending factor = 1, in the condition of MCAR with 50% missingness, a normal distribution, and correlation of 0.1.

It is important to note, however, that these differences are small and likely due to chance. The overall trend shows that blending towards the Mahalanobis distance leads to worse performance, especially under the more extreme conditions such as MAR, a large proportion of missingness, and a skewed distribution.

3.2 | Simulation study II

As the ranked version of the blended metric performed best overall in Simulation study I, this version was evaluated in the second simulation study. The simulation results are displayed in Table B8 in Appendix B. The results show that as the blending factor decreases (which implies weighting in the direction of the Mahalanobis distance), the bias increases, but there is no clear trend in the absolute bias. In addition, a smaller blending factor leads to a smaller standard error and lower coverage. The RMSE also shows a decreasing trend, however, there is a small increase in the RMSE from a blending factor of 0.1 to 0. The performance measures show that a property of blending is the bias-variance tradeoff. That is, weighting towards the Mahalanobis distance results in more precise, but less accurate estimates. The predictive distance results in less precise, but more accurate estimates. This is further illustrated in Figure B1, where the distributions of the estimates obtained by the ranked metric with a blending factor of 0 and 1 are shown. What conclusions can be derived from this figure?

3.3 | Application to empirical data

What exactly should I do with the empirical data?

4 | DISCUSSION

This study investigated the properties of a blended metric through simulations and an application to empirical data. In simulation study I, seven metrics were compared: the predictive metric, a ranked version of the blended metric with blending factors of 1,

0.5, and 0, respectively, and a scaled version of the blended metric with blending factors of 1, 0.5, and 0, respectively. The data-generating mechanisms were varied in their missingness mechanism, missingness proportion, distribution, and correlation. A full-factorial design was used, where all possible combinations of metrics and data-generating mechanisms were simulated. The purpose of this study was to investigate whether a ranked and scaled version of the blended metric would yield identical results, a blending factor of 1 would perform the same as PMM, how performance is related to missingness mechanism, proportion, distribution, and correlation, and if blending reduces predictability. The results show that the ranked and scaled versions do not yield identical, but similar results. The scaled version might be favourable to implement because of its computational efficiency. A blending factor of 1 does not yield results identical to those obtained by PMM, but they are similar as well. Performance is higher when the missingness mechanism is MCAR as opposed to MAR right and when the missingness proportion is 25% as opposed to 50%, as would be expected. The skewness of the data does not always impact the performance negatively. A higher correlation in the data under MCAR conditions leads to a small increase in performance in some cases, but to lower performance under the MAR right conditions. Finally, a smaller blending factor, meaning that the metric gives more weight to the Mahalanobis distance, generally leads to lower performance. However, a smaller blending factor results in higher performance in some of the MCAR conditions. It is important to note that these differences are very small, and thus are not reasons to suggest that the blended metric should replace the predictive distance under these conditions. It does imply, however, that the blended metric could be implemented in situations where these conditions are true, without compromising in performance. This might be useful to do in cases where the 'future' is vaguely defined, or where a user is more interested in similarity between the histories of the donors and the target. On the other hand, the underperformance of the blended metric under conditions like MAR right is an indication that it should not be used under these circumstances. Furthermore, the results confirm that PMM is a stable and reliable technique for donor selection, even under extreme data-generating conditions.

In simulation study II, the ranked version of the blended metric was evaluated under the condition of a skewed distribution and correlation of 0.7. The results show that a property of blending is the bias-variance tradeoff: weighting towards the predictive distance results in less precise, but more accurate estimates.

The overall conclusion is that the blended metric can be implemented in situations where the missingness proportion is small and is MCAR. This is generally the case in the context of height prediction, as only a single value is to be imputed. It may be attractive for users of growth curve matching to use the blended metric when they have difficulty selecting a particular future time point to base matches on, or when they have more interest in the similarity between donors and the target in the predictor space. However, it is important to keep in mind the severe underperformance of the blended metric under more extreme conditions, and the reliable performance of the predictive distance under any conditions. PMM should therefore always be the preferred method, and the results of this study may minimize doubts about discrepancies between the trajectories of the matched donors. After all, our past does not always define our future, and selecting matches may not so much be a matter of finding "people-like-me" in terms of historic growth, as it is one of future similarity.

just some draft suggestions The current study investigated the influence of missingness proportion, missingness mechanisms, skewness of the data, and correlation in the data for the use of the blended metric. For further investigations of the properties of the blended metric, other factors could be varied, such as the sample size and the number of k matched donors. In addition, it would be interesting to evaluate alternative combinations of similarity measures and the predictive metric. Examples of such measures would be the Frechet distance,[11] and the locally supervised metric learning (LSML) measure.[12] Finally, the first simulation study solely used the blending factors of 1, 0.5 and 0. Further research could determine what the optimal blending factor is to predict outcomes.

ACKNOWLEDGMENTS

SUPPORTING INFORMATION

The scripts used for the simulation study are available in the research archive of this study. The study was approved by the Ethical Review Board of the Faculty of Social and Behavioural Sciences of Utrecht University. The approval is based on the documents sent by the researchers as requested in the form of the Ethics committee and filed under number 21-1906.

References

1. Straatmann VS, Pearce A, Hope S, et al. How well can poor child health and development be predicted by data collected in early childhood?. *J Epidemiol Community Health* 2018; 72(12): 1132–1140. Publisher: BMJ Publishing Group Ltd.

2. Cordeiro JR, Postolache O, Ferreira JC. Child's target height prediction evolution. *Applied Sciences* 2019; 9(24): 5447. Publisher: Multidisciplinary Digital Publishing Institute.

3. Van Buuren S. Curve matching: a data-driven technique to improve individual prediction of childhood growth. *Annals of Nutrition and Metabolism* 2014; 65(2-3): 227–233. Publisher: Karger Publishers.

4. Van Buuren S. Broken stick model for irregular longitudinal data. *Journal of Statistical Software* 2020; Submitted for publication: 1–47.

5. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Statistics in Medicine* 2019; 38(11): 2074–2102. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.8086doi: 10.1002/sim.8086

6. Van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* 2011; 45(3): 1–67.

7. Herngreen WP, Van Buuren S, Van Wieringen JC, Reerink JD, Verloove-Vanhorick SP, Ruys JH. Growth in length and weight from birth to 2 years of a representative sample of Netherlands children (born in 1988–89) related to socioeconomic status and other background characteristics. *Annals of human biology* 1994; 21(5): 449–463. Publisher: Taylor & Francis.

8. Vink G, Frank LE, Pannekoek J, Van Buuren S. Predictive mean matching imputation of semicontinuous variables. *Statistica Neerlandica* 2014; 68(1): 61–90. Publisher: Wiley Online Library.

9. Little RJ, Rubin DB. *Statistical analysis with missing data*. 793. John Wiley & Sons . 2019.

10. R Core Team . *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing . 2021.

11. Eiter T, Mannila H. Computing discrete Fréchet distance. tech. rep., Citeseer; 1994.

12. Ng K, Sun J, Hu J, Wang F. Personalized predictive modeling and risk factor identification using patient similarity. *AMIA Summits on Translational Science Proceedings* 2015; 2015: 132. Publisher: American Medical Informatics Association.

How to cite this article: Fopma A.M (2022), A blended distance to define "people-like-me", *Statistics in Medicine*, *x;x:x–x*.

APPENDIX

A RESULTS OF SIMULATION STUDY I

TABLE A1 Method PMM.

mech	mis	dist	cor	qbar	se	t	df	b	2.5%	97.5%	true	cov	bias	R2
MCAR	25%	normal	0	30.466	0.191	0.153	97.773	0.035	29.935	30.996	30.47	0.949	-0.004	0.079
			0.1	30.466	0.192	0.155	98.311	0.035	29.933	30.998	30.475	0.96	-0.009	0.09
			0.7	30.502	0.187	0.161	112.474	0.033	29.982	31.022	30.502	0.941	0	0.152
		skewed	0	3.199	0.248	0.187	65.411	0.06	2.511	3.888	3.256	0.969	-0.057	0.106
			0.1	3.238	0.251	0.189	64.18	0.061	2.543	3.934	3.292	0.976	-0.054	0.112
			0.7	4.068	0.259	0.213	67.895	0.064	3.349	4.788	4.109	0.976	-0.041	0.239
	50%	normal	0	30.443	0.339	0.243	30.409	0.109	29.502	31.383	30.47	0.94	-0.027	0.083
			0.1	30.45	0.342	0.248	31.72	0.113	29.499	31.4	30.475	0.953	-0.026	0.093
			0.7	30.495	0.327	0.243	35.429	0.102	29.587	31.403	30.502	0.94	-0.007	0.153
		skewed	0	3.149	0.462	0.373	21.577	0.214	1.867	4.431	3.256	0.977	-0.107	0.106
			0.1	3.167	0.49	0.402	17.865	0.238	1.807	4.526	3.292	0.978	-0.126	0.114
			0.7	4.022	0.516	0.444	18.404	0.257	2.59	5.455	4.109	0.978	-0.087	0.24
MAR	25%	normal	0	30.456	0.185	0.151	105.503	0.033	29.943	30.969	30.47	0.912	-0.014	0.082
			0.1	30.46	0.186	0.153	105.455	0.033	29.943	30.977	30.475	0.897	-0.015	0.093
			0.7	30.483	0.188	0.162	113.281	0.034	29.96	31.005	30.502	0.908	-0.019	0.153
		skewed	0	3.09	0.193	0.154	101.507	0.036	2.556	3.625	3.256	0.878	-0.165	0.077
			0.1	3.143	0.195	0.156	96.792	0.036	2.601	3.685	3.292	0.884	-0.149	0.088
			0.7	3.895	0.194	0.17	107.547	0.036	3.356	4.434	4.109	0.842	-0.214	0.176
	50%	normal	0	30.437	0.303	0.218	41.511	0.089	29.596	31.277	30.47	0.863	-0.033	0.086
			0.1	30.446	0.305	0.221	37.774	0.09	29.599	31.292	30.475	0.868	-0.03	0.097
			0.7	30.451	0.305	0.228	43.159	0.09	29.605	31.296	30.502	0.857	-0.051	0.153
		skewed	0	2.975	0.372	0.271	26.075	0.133	1.942	4.008	3.256	0.911	-0.281	0.081
			0.1	3.021	0.365	0.267	26.939	0.129	2.007	4.035	3.292	0.905	-0.272	0.089
			0.7	3.674	0.37	0.28	28.378	0.132	2.646	4.702	4.109	0.858	-0.436	0.141

TABLE A2 Method ranked, blend = 1.

mech	mis	dist	cor	qbar	se	t	df	b	2.5%	97.5%	true	cov	bias	R2
MCAR	25%	normal	0	30.462	0.192	0.154	97.885	0.035	29.929	30.995	30.47	0.95	-0.008	0.079
			0.1	30.467	0.195	0.156	95.776	0.036	29.927	31.008	30.475	0.961	-0.008	0.091
			0.7	30.504	0.185	0.16	111.671	0.033	29.989	31.019	30.502	0.952	0.003	0.152
		skewed	0	3.211	0.242	0.183	66.725	0.056	2.539	3.883	3.256	0.975	-0.045	0.107
			0.1	3.234	0.254	0.191	60.949	0.062	2.529	3.939	3.292	0.975	-0.058	0.112
			0.7	4.074	0.253	0.21	71.264	0.062	3.37	4.778	4.109	0.968	-0.035	0.24
	50%	normal	0	30.44	0.34	0.244	30.283	0.11	29.496	31.384	30.47	0.949	-0.03	0.083
			0.1	30.454	0.329	0.238	35.708	0.104	29.541	31.366	30.475	0.93	-0.022	0.094
			0.7	30.492	0.328	0.244	35.911	0.103	29.58	31.404	30.502	0.939	-0.009	0.153
		skewed	0	3.147	0.461	0.375	21.325	0.217	1.866	4.428	3.256	0.965	-0.109	0.106
			0.1	3.155	0.487	0.403	19.639	0.238	1.803	4.507	3.292	0.972	-0.137	0.115
			0.7	4.014	0.513	0.442	17.427	0.256	2.589	5.438	4.109	0.982	-0.096	0.24
MAR	25%	normal	0	30.46	0.185	0.151	105.089	0.032	29.946	30.973	30.47	0.898	-0.01	0.082
			0.1	30.458	0.188	0.153	103.006	0.033	29.937	30.979	30.475	0.898	-0.017	0.094
			0.7	30.481	0.188	0.162	110.033	0.034	29.958	31.004	30.502	0.902	-0.021	0.153
		skewed	0	3.088	0.191	0.153	100.427	0.035	2.558	3.619	3.256	0.866	-0.167	0.077
			0.1	3.145	0.193	0.155	98.57	0.035	2.611	3.68	3.292	0.884	-0.147	0.088
			0.7	3.892	0.192	0.169	107.928	0.035	3.358	4.426	4.109	0.825	-0.218	0.175
	50%	normal	0	30.444	0.301	0.218	41.213	0.088	29.607	31.281	30.47	0.852	-0.025	0.086
			0.1	30.442	0.304	0.22	41.027	0.089	29.598	31.286	30.475	0.883	-0.033	0.097
			0.7	30.459	0.301	0.226	45.189	0.087	29.623	31.295	30.502	0.865	-0.043	0.153
		skewed	0	2.983	0.381	0.279	24.606	0.14	1.923	4.042	3.256	0.918	-0.273	0.08
			0.1	3.025	0.372	0.273	27.808	0.134	1.994	4.057	3.292	0.901	-0.267	0.089
			0.7	3.665	0.367	0.276	27.827	0.129	2.646	4.684	4.109	0.836	-0.444	0.142

TABLE A3 Method ranked, blend = 0.5.

mech	mis	dist	cor	qbar	se	t	df	b	2.5%	97.5%	true	cov	bias	R2
MCAR	25%	normal	0	30.477	0.181	0.149	107.792	0.031	29.974	30.98	30.47	0.957	0.007	0.076
			0.1	30.492	0.183	0.151	106.919	0.031	29.985	30.999	30.475	0.943	0.016	0.088
			0.7	30.51	0.18	0.158	119.4	0.031	30.01	31.009	30.502	0.938	0.008	0.148
		skewed	0	3.227	0.204	0.161	90.943	0.04	2.662	3.792	3.256	0.944	-0.029	0.094
			0.1	3.264	0.214	0.167	84.099	0.044	2.67	3.857	3.292	0.951	-0.029	0.102
			0.7	4.05	0.224	0.191	85.852	0.048	3.426	4.673	4.109	0.95	-0.06	0.226
	50%	normal	0	30.486	0.303	0.216	38.173	0.087	29.646	31.326	30.47	0.917	0.017	0.074
			0.1	30.494	0.296	0.212	39.051	0.083	29.673	31.315	30.475	0.928	0.019	0.085
			0.7	30.516	0.293	0.219	47.093	0.082	29.704	31.328	30.502	0.914	0.015	0.142
		skewed	0	3.178	0.351	0.256	31.295	0.12	2.202	4.154	3.256	0.946	-0.078	0.087
			0.1	3.218	0.37	0.271	26.642	0.131	2.191	4.245	3.292	0.946	-0.074	0.094
			0.7	3.941	0.409	0.325	26.368	0.161	2.805	5.076	4.109	0.951	-0.169	0.206
MAR	25%	normal	0	30.421	0.178	0.148	111.804	0.03	29.927	30.915	30.47	0.898	-0.049	0.076
			0.1	30.421	0.177	0.149	116.968	0.03	29.93	30.912	30.475	0.898	-0.054	0.088
			0.7	30.434	0.18	0.158	117.612	0.031	29.934	30.934	30.502	0.898	-0.068	0.145
		skewed	0	3.081	0.177	0.146	114.292	0.03	2.589	3.573	3.256	0.864	-0.175	0.07
			0.1	3.109	0.175	0.146	118.274	0.029	2.624	3.594	3.292	0.847	-0.183	0.078
			0.7	3.76	0.17	0.155	133.762	0.028	3.287	4.233	4.109	0.678	-0.35	0.141
	50%	normal	0	30.406	0.262	0.191	52.803	0.066	29.678	31.133	30.47	0.809	-0.064	0.082
			0.1	30.402	0.273	0.199	48.497	0.071	29.643	31.161	30.475	0.84	-0.073	0.092
			0.7	30.396	0.264	0.202	58.445	0.068	29.662	31.13	30.502	0.802	-0.105	0.145
		skewed	0	2.907	0.301	0.213	37.629	0.087	2.071	3.744	3.256	0.822	-0.348	0.059
			0.1	2.932	0.293	0.207	41.762	0.081	2.12	3.744	3.292	0.801	-0.361	0.065
			0.7	3.547	0.29	0.214	46.431	0.081	2.741	4.353	4.109	0.687	-0.563	0.109

TABLE A4 Method ranked, blend = 0.

mech	mis	dist	cor	qbar	se	t	df	b	2.5%	97.5%	true	cov	bias	R2
MCAR	25%	normal	0	30.462	0.145	0.134	156.897	0.02	30.059	30.866	30.47	0.888	-0.007	0.074
			0.1	30.468	0.146	0.136	159.425	0.02	30.062	30.874	30.475	0.893	-0.007	0.086
			0.7	30.499	0.148	0.145	166.98	0.021	30.088	30.91	30.502	0.882	-0.003	0.144
		skewed	0	3.192	0.147	0.138	158.72	0.021	2.782	3.601	3.256	0.893	-0.064	0.097
			0.1	3.223	0.147	0.139	162.631	0.02	2.816	3.63	3.292	0.889	-0.069	0.103
			0.7	4.003	0.146	0.157	186.987	0.02	3.599	4.407	4.109	0.863	-0.106	0.223
	50%	normal	0	30.494	0.211	0.159	77.005	0.042	29.909	31.079	30.47	0.839	0.024	0.069
			0.1	30.504	0.211	0.161	77.508	0.042	29.918	31.091	30.475	0.851	0.029	0.079
			0.7	30.531	0.212	0.169	87.214	0.043	29.944	31.118	30.502	0.843	0.029	0.132
		skewed	0	3.128	0.209	0.16	79.65	0.041	2.548	3.707	3.256	0.826	-0.128	0.091
			0.1	3.146	0.207	0.161	82.816	0.04	2.571	3.721	3.292	0.835	-0.146	0.097
			0.7	3.846	0.213	0.181	92.397	0.043	3.255	4.438	4.109	0.772	-0.263	0.2
MAR	25%	normal	0	30.408	0.15	0.136	148.803	0.021	29.99	30.825	30.47	0.863	-0.062	0.078
			0.1	30.401	0.151	0.138	152.097	0.021	29.983	30.819	30.475	0.87	-0.074	0.088
			0.7	30.402	0.149	0.145	160.168	0.021	29.987	30.817	30.502	0.84	-0.1	0.144
		skewed	0	3.054	0.145	0.134	158.032	0.02	2.651	3.457	3.256	0.753	-0.202	0.07
			0.1	3.088	0.147	0.135	156.613	0.02	2.681	3.496	3.292	0.758	-0.204	0.074
			0.7	3.681	0.143	0.143	174.906	0.019	3.284	4.077	4.109	0.453	-0.429	0.13
	50%	normal	0	30.357	0.214	0.163	77.74	0.044	29.762	30.952	30.47	0.761	-0.113	0.08
			0.1	30.349	0.215	0.164	74.64	0.044	29.751	30.947	30.475	0.761	-0.126	0.089
			0.7	30.318	0.214	0.171	82.206	0.043	29.724	30.911	30.502	0.736	-0.184	0.141
		skewed	0	2.86	0.205	0.154	78.597	0.04	2.29	3.43	3.256	0.651	-0.396	0.059
			0.1	2.895	0.201	0.154	84.367	0.038	2.336	3.455	3.292	0.638	-0.397	0.065
			0.7	3.44	0.208	0.165	84.385	0.041	2.863	4.018	4.109	0.403	-0.669	0.1

TABLE A5 Method scaled, blend = 1.

mech	mis	dist	cor	qbar	se	t	df	b	2.5%	97.5%	true	cov	bias	R2
MCAR	25%	normal	0	30.461	0.197	0.156	93.518	0.037	29.914	31.009	30.47	0.958	-0.008	0.079
			0.1	30.473	0.189	0.154	103.355	0.034	29.947	30.999	30.475	0.942	-0.002	0.09
			0.7	30.507	0.194	0.164	100.127	0.035	29.968	31.046	30.502	0.964	0.005	0.152
		skewed	0	3.205	0.25	0.188	60.828	0.06	2.51	3.901	3.256	0.968	-0.05	0.107
			0.1	3.236	0.248	0.188	62.727	0.059	2.548	3.924	3.292	0.969	-0.056	0.112
			0.7	4.073	0.255	0.211	72.87	0.063	3.365	4.782	4.109	0.966	-0.036	0.24
	50%	normal	0	30.439	0.338	0.243	31.954	0.11	29.501	31.376	30.47	0.939	-0.031	0.084
			0.1	30.45	0.33	0.24	35.211	0.106	29.534	31.367	30.475	0.936	-0.025	0.093
			0.7	30.488	0.33	0.245	35.929	0.103	29.572	31.405	30.502	0.948	-0.013	0.153
		skewed	0	3.152	0.458	0.371	22.021	0.213	1.882	4.423	3.256	0.969	-0.103	0.107
			0.1	3.171	0.484	0.398	18.633	0.234	1.826	4.515	3.292	0.978	-0.122	0.113
			0.7	4.013	0.518	0.447	17.036	0.259	2.575	5.451	4.109	0.987	-0.096	0.241
MAR	25%	normal	0	30.461	0.185	0.151	104.176	0.032	29.947	30.975	30.47	0.906	-0.009	0.082
			0.1	30.46	0.189	0.154	102.232	0.034	29.935	30.984	30.475	0.916	-0.016	0.093
			0.7	30.489	0.188	0.162	109.703	0.033	29.968	31.011	30.502	0.908	-0.012	0.154
		skewed	0	3.094	0.191	0.153	104.554	0.036	2.563	3.626	3.256	0.868	-0.162	0.077
			0.1	3.14	0.193	0.155	96.462	0.036	2.603	3.677	3.292	0.867	-0.152	0.088
			0.7	3.889	0.193	0.17	111.041	0.035	3.354	4.424	4.109	0.825	-0.221	0.175
	50%	normal	0	30.439	0.299	0.217	44.292	0.088	29.608	31.27	30.47	0.859	-0.031	0.086
			0.1	30.441	0.301	0.218	39.103	0.088	29.604	31.278	30.475	0.866	-0.034	0.097
			0.7	30.457	0.296	0.222	46.435	0.085	29.635	31.278	30.502	0.857	-0.045	0.153
		skewed	0	2.994	0.378	0.279	27.141	0.14	1.944	4.044	3.256	0.904	-0.262	0.08
			0.1	3.029	0.374	0.275	25.454	0.136	1.99	4.068	3.292	0.906	-0.263	0.089
			0.7	3.661	0.371	0.278	27.739	0.131	2.632	4.69	4.109	0.851	-0.448	0.141

TABLE A6 Method scaled, blend = 0.5.

mech	mis	dist	cor	qbar	se	t	df	b	2.5%	97.5%	true	cov	bias	R2
MCAR	25%	normal	0	30.467	0.188	0.152	102.834	0.034	29.946	30.988	30.47	0.953	-0.003	0.077
			0.1	30.478	0.18	0.15	112.559	0.031	29.979	30.978	30.475	0.94	0.003	0.088
			0.7	30.496	0.178	0.158	125.64	0.03	30.002	30.99	30.502	0.939	-0.006	0.149
		skewed	0	3.184	0.246	0.186	68.218	0.059	2.502	3.866	3.256	0.96	-0.072	0.098
			0.1	3.228	0.251	0.189	61.883	0.061	2.531	3.925	3.292	0.962	-0.064	0.106
			0.7	4.03	0.256	0.21	69.202	0.063	3.319	4.74	4.109	0.972	-0.08	0.231
	50%	normal	0	30.452	0.31	0.221	35.27	0.09	29.591	31.312	30.47	0.949	-0.018	0.077
			0.1	30.462	0.307	0.221	38.968	0.09	29.61	31.314	30.475	0.923	-0.013	0.087
			0.7	30.487	0.301	0.224	42.769	0.086	29.65	31.323	30.502	0.926	-0.015	0.146
		skewed	0	3.119	0.454	0.365	20.317	0.208	1.858	4.381	3.256	0.963	-0.136	0.091
			0.1	3.151	0.475	0.388	19.111	0.227	1.833	4.469	3.292	0.977	-0.142	0.1
			0.7	3.919	0.509	0.435	17.998	0.252	2.505	5.333	4.109	0.97	-0.19	0.219
MAR	25%	normal	0	30.425	0.184	0.151	105.503	0.032	29.915	30.935	30.47	0.906	-0.045	0.078
			0.1	30.429	0.18	0.151	114.155	0.031	29.93	30.927	30.475	0.9	-0.047	0.089
			0.7	30.444	0.181	0.159	118.405	0.031	29.94	30.948	30.502	0.896	-0.057	0.148
		skewed	0	3.066	0.184	0.149	103.034	0.032	2.554	3.577	3.256	0.856	-0.19	0.072
			0.1	3.091	0.18	0.149	108.706	0.031	2.59	3.592	3.292	0.843	-0.202	0.078
			0.7	3.77	0.183	0.162	119.094	0.032	3.261	4.278	4.109	0.708	-0.34	0.149
	50%	normal	0	30.403	0.284	0.207	47.457	0.079	29.615	31.191	30.47	0.851	-0.067	0.083
			0.1	30.402	0.281	0.205	46.188	0.076	29.621	31.182	30.475	0.838	-0.073	0.093
			0.7	30.407	0.279	0.211	48.836	0.074	29.633	31.182	30.502	0.82	-0.094	0.148
		skewed	0	2.904	0.327	0.234	34.697	0.103	1.996	3.811	3.256	0.866	-0.352	0.064
			0.1	2.936	0.332	0.237	33.535	0.106	2.013	3.859	3.292	0.867	-0.356	0.07
			0.7	3.53	0.34	0.253	33.878	0.112	2.586	4.474	4.109	0.734	-0.579	0.114

TABLE A7 Method scaled, blend = 0.

mech	mis	dist	cor	qbar	se	t	df	b	2.5%	97.5%	true	cov	bias	R2
MCAR	25%	normal	0	30.457	0.143	0.134	163.429	0.019	30.059	30.855	30.47	0.893	-0.012	0.074
			0.1	30.462	0.146	0.136	157.712	0.02	30.058	30.867	30.475	0.895	-0.013	0.085
			0.7	30.496	0.146	0.144	167.413	0.02	30.091	30.901	30.502	0.887	-0.006	0.144
		skewed	0	3.188	0.145	0.137	165.544	0.02	2.785	3.59	3.256	0.89	-0.068	0.096
			0.1	3.223	0.147	0.139	159.723	0.021	2.814	3.633	3.292	0.891	-0.069	0.103
			0.7	4.011	0.147	0.157	183.983	0.021	3.603	4.419	4.109	0.86	-0.098	0.223
	50%	normal	0	30.497	0.202	0.156	86.643	0.039	29.936	31.059	30.47	0.824	0.028	0.069
			0.1	30.505	0.208	0.16	83.654	0.041	29.928	31.082	30.475	0.842	0.03	0.079
			0.7	30.532	0.209	0.167	85.635	0.041	29.953	31.111	30.502	0.843	0.03	0.132
		skewed	0	3.122	0.206	0.159	84.192	0.04	2.55	3.694	3.256	0.832	-0.134	0.09
			0.1	3.15	0.207	0.161	83.346	0.04	2.574	3.725	3.292	0.806	-0.143	0.097
			0.7	3.847	0.209	0.18	97.013	0.042	3.266	4.428	4.109	0.774	-0.263	0.2
MAR	25%	normal	0	30.402	0.15	0.136	153.443	0.021	29.987	30.817	30.47	0.856	-0.068	0.077
			0.1	30.403	0.15	0.138	151.516	0.021	29.985	30.82	30.475	0.859	-0.072	0.088
			0.7	30.406	0.154	0.146	153.178	0.022	29.979	30.833	30.502	0.848	-0.096	0.145
		skewed	0	3.054	0.143	0.133	162.082	0.019	2.657	3.451	3.256	0.742	-0.202	0.069
			0.1	3.084	0.145	0.135	158.785	0.02	2.681	3.486	3.292	0.755	-0.209	0.074
			0.7	3.681	0.144	0.144	175.877	0.02	3.28	4.081	4.109	0.465	-0.429	0.13
	50%	normal	0	30.354	0.214	0.163	77.214	0.044	29.759	30.949	30.47	0.758	-0.116	0.08
			0.1	30.348	0.211	0.162	82.081	0.042	29.763	30.933	30.475	0.752	-0.127	0.09
			0.7	30.318	0.214	0.171	85.563	0.043	29.725	30.911	30.502	0.739	-0.184	0.141
		skewed	0	2.86	0.201	0.154	89.028	0.039	2.301	3.419	3.256	0.638	-0.396	0.058
			0.1	2.891	0.211	0.158	76.828	0.042	2.306	3.477	3.292	0.653	-0.401	0.065
			0.7	3.438	0.204	0.163	86.894	0.039	2.871	4.005	4.109	0.395	-0.672	0.1

B RESULTS OF SIMULATION STUDY II

TABLE B8 Performance results for Simulation study II.

method	estimate	true	bias	absbias	ssd	se	lwr	upr	cov	rmse
PMM	4.10	4.29	-0.18	8.13	5133.00	7.12	-10.21	18.42	0.94	9.68
Blending factor = 1	4.10	4.29	-0.19	8.12	5126.81	7.12	-10.22	18.41	0.95	9.67
Blending factor = 0.9	3.98	4.29	-0.31	8.15	5182.68	7.15	-10.38	18.34	0.94	9.70
Blending factor = 0.8	3.95	4.29	-0.33	8.13	5153.59	7.06	-10.24	18.15	0.94	9.65
Blending factor = 0.7	3.99	4.29	-0.30	8.10	5116.94	6.97	-10.01	17.99	0.94	9.60
Blending factor = 0.6	4.00	4.29	-0.29	8.08	5107.43	6.90	-9.87	17.87	0.93	9.57
Blending factor = 0.5	3.99	4.29	-0.30	8.10	5118.92	6.85	-9.77	17.75	0.93	9.57
Blending factor = 0.4	3.96	4.29	-0.32	8.05	5082.51	6.77	-9.64	17.57	0.92	9.51
Blending factor = 0.3	3.94	4.29	-0.35	8.00	5025.93	6.66	-9.43	17.32	0.91	9.43
Blending factor = 0.2	3.86	4.29	-0.42	7.99	5014.08	6.52	-9.24	16.97	0.90	9.38
Blending factor = 0.1	3.74	4.29	-0.55	7.93	4948.25	6.29	-8.91	16.38	0.87	9.25
Blending factor = 0	3.74	4.29	-0.54	8.09	5121.67	6.04	-8.40	15.89	0.82	9.32

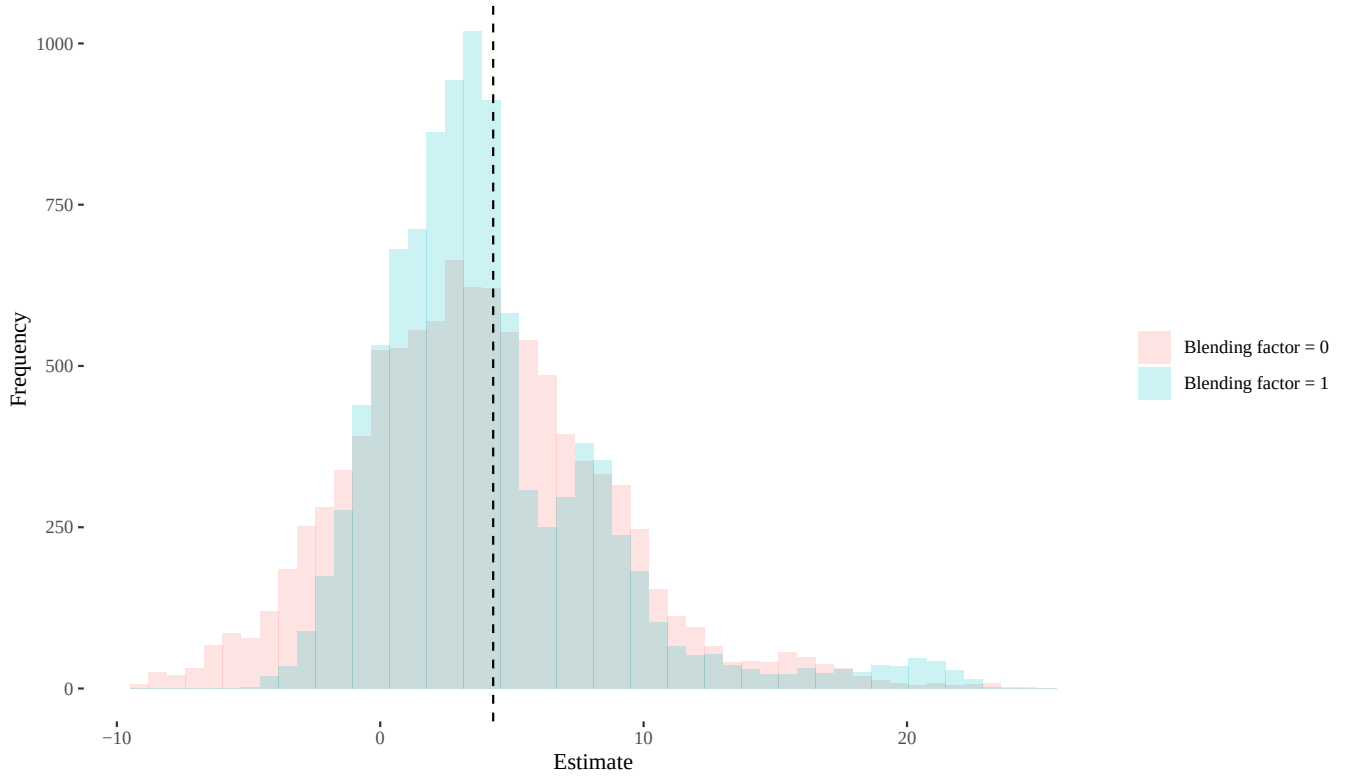

FIGURE B1 Distribution of the estimates obtained with the ranked version of the blended metric, with a blending factor of 0 (Mahalanobis distance) and 1 (predictive distance), respectively. The true value is marked by the dashed line.