Distribution-free multiple imputation based on robust singular value decomposition: an empirical comparison

Author:
Daniel Andrés Becerra Sierra

Tutor:
Marisol García Peña
Sergio Arciniegas Alarcón

Data Science

Degree of thesis

Pontificia Universidad Javeriana

Faculty of Sciences

Bogotá D.C
2025



In loving memory of Luz Mariela Becerra, Lucy Arrieta and Alfonso Charum

Acknowledgments

First and foremost, I want to thank my advisors, **Dr. Marisol García Peña** and **Dr. Sergio Arciniegas Alarcón**, for their patience, guidance, and unwavering support throughout this journey. I felt truly accompanied at every step. There was not a single question left unanswered, and there was never a moment when I could not turn to you for your invaluable help. For that, I am eternally grateful. I also extend my gratitude to **ZINE, the High-Performance Computing Center at Pontificia Universidad Javeriana**, for their assistance and for allowing me to run my work on their clusters, which made the heavy computational load far more manageable. I would also like to express my sincere gratitude to **Dr. Lina Acosta**, with whom I had classes for three years during my studies and with whom I took nearly all of my statistics courses. She taught me far more than academic knowledge she inspired me to go beyond the basics, to think critically, and to always strive for more

To my mother, **Luz Stella**: thank you for every word of encouragement and each moment you listened, not only through this thesis but throughout my entire path as a student. You were the rock I leaned on whenever I could not stand on my own, a true star life sent my way. To my father, **John**: thank you for always reminding me that I could do it. You listened when I doubted myself and, with a few wise words, told me exactly what I needed to keep going. You were, are, and will always be my North Star. To **Kevin**, for your support, kindness, love, and patience through every vent, and for the many late-night study sessions we shared since high school, thank you. To my brothers, **Tomás** and **Juange**, for the constant love you have always given me.To my grandparents **Cecilia**, **Gabriel** and **Pedro** who understood me throughout this process every time I couldn't be with them.

To my university friends, **"Los Impedidos"**: **Samuel, Tomás, Sebas, Juanjo, and Mauro**, with whom I have walked this path since the very first semester. Thank you for the laughs, the companionship, and for still making me smile today. To **Anamaría**, my partner in countless projects and battles this 4 years. Thank you for teaching me more than I expected to learn, not only professionally but also personally. To **Sofía**, who followed me into many wild ideas. I still remember what I told you after the first exam in Introduction to Data Science: *"Keep trying to be as good as me."* The truth is that every day I wake up hoping to be only a fraction of the incredible professional and person that you are. To **Gali**, who has been my grounding force since high school and throughout university. Thank you for every time you listened.

To **Lau**, who accompanied me in every part of this thesis. She is truly a wonderful woman in every sense of the word. Her strength, kindness, and unwavering presence carried me through the hardest nights, and I do not know if I would have finished without her.

And finally, with an especially warm place in my heart, to my cousin and friend **Juanda, "el ingeniero."** We have lived a lifetime side by side. We graduated from the same high school class, and now we graduate from the same program at the same university. Your

capacity to create and your originality amaze me every day, like sending a probabilistic model without a seed. Thank you for everything, soulmate. This is one more milestone among the many that lie ahead. With you, I am ready for any battle.
This work is dedicated to each person named here, and to many others I surely missed, because it is thanks to you that I am who I am and that I am here today.

Abstract

This study develops, validates and compare a multiple imputation (MI) approach based on robust variants of the Singular Value Decomposition (SVD), specifically rSVD84, rSVD01, and a hybrid rSVD84 01, to address missing data under different mechanisms and contamination levels. The objective is to achieve reliable imputations under MCAR, MAR, and MNAR scenarios while maintaining robustness against outliers. The proposed models are benchmarked against Approximated Bayesian Bootstrap (ABB) and Multivariate Imputation by Chained Equations (MICE) through controlled simulations and real world datasets. Simulation results indicate that under moderate contamination, SVD based models perform competitively: they consistently outperform ABB and remain distribution free, although they generally fall short of MICE in efficiency. Under extreme outliers, the SVD models demonstrate a clear advantage, maintaining stable accuracy as missingness increases, with rSVD84 exhibiting the greatest robustness. The MNAR scenario remains the most challenging for all methods, yielding lower predictive accuracy and higher between imputation variability. In multivariate real datasets, SVD based methods produce mixed results and often underperform relative to MICE and ABB, likely because the bootstrap style randomization inherent to these methods is better suited to genotype by environment (GxE) structures than to generic multivariate data. In very small samples, these models are unable to fully exploit their randomization mechanism, leading to unstable imputations. Overall, robust SVD methods preserve predictive accuracy in the presence of extreme outliers but at the cost of increased between imputation variance, indicating greater imputation uncertainty. While robust SVD based imputation can approximate the performance of established baselines, it rarely achieves the lowest prediction error, suggesting that the strengths observed in simulations do not always generalize to diverse real world applications.

Table of Contents

Contents

1 Introduction **10**

2 Objectives **13**
2.1 General Objective 13
2.2 Specific Objectives 13

3 Theoretical Framework **14**
3.1 Multiple Imputation 14
3.1.1 Rubin's Rules 14
3.1.2 Classical MI Methods (MICE, Model-Based, Nonparametric) 15
3.1.3 Evolution in recent literature 16
3.2 Missing Data Mechanisms 17
3.2.1 Missing Data Patterns 17
3.2.2 Diagnosing Missing Data Patterns 18
3.2.3 Simulating Missing Data Mechanisms 20
3.3 Outliers 20
3.3.1 Definitions and Types of Outliers 20
3.3.2 Outlier Detection Methods 21
3.3.3 Generating Outlier Simulations 22
3.4 Singular Value Decomposition (SVD) 22
3.4.1 SVD Theoretical Foundations and Applications 22
3.4.2 SVD as imputation method 23
3.4.3 Sensitivity of SVD to Outliers and Missing Data 23
3.5 Robust Alternatives to SVD (rSVD84 and rSVD01) 24
3.5.1 rSVD84 24
3.5.2 rSVD01 25

4 Methodology **27**
4.1 Simulation study 27
4.2 Real Data Study 28
4.3 Evaluation method 29
4.4 Proposed MI method 29

5 Results **30**
5.1 Results for simulated data with 5 standard deviation environment 30
5.2 Results for simulated data 100 standard deviation environment 39
5.3 Results for real data with 5 standard deviation environment 46
5.4 Results for real data with 100 standard deviation environment 54

6 Conclusions **60**

7 Future Work **61**

8 References **62**

9 Appendices **67**

List of Figures

Figures

1 Boxplot of P_e under MAR at $p = 10\%$. 30
2 Boxplot of P_e under MAR at $p = 20\%$. 31
3 Boxplot of P_e under MAR at $p = 30\%$. 31
4 Boxplot of P_e under MCAR at $p = 10\%$. 32
5 Boxplot of P_e under MCAR at $p = 20\%$. 32
6 Boxplot of P_e under MCAR at $p = 30\%$. 32
7 Boxplot of P_e under MNAR at $p = 10\%$. 33
8 Boxplot of P_e under MNAR at $p = 20\%$. 33
9 Boxplot of P_e under MNAR at $p = 30\%$. 34
10 Boxplot of $\mathbf{V}$ under MAR at $p = 10\%$. 34
11 Boxplot of $\mathbf{V}$ under MAR at $p = 20\%$. 35
12 Boxplot of $\mathbf{V}$ under MAR at $p = 30\%$. 35
13 Boxplot of $\mathbf{V}$ under MCAR at $p = 10\%$..... 36
14 Boxplot of $\mathbf{V}$ under MCAR at $p = 20\%$..... 36
15 Boxplot of $\mathbf{V}$ under MCAR at $p = 30\%$..... 36
16 Boxplot of $\mathbf{V}$ under MNAR at $p = 10\%$..... 37
17 Boxplot of $\mathbf{V}$ under MNAR at $p = 20\%$..... 37
18 Boxplot of $\mathbf{V}$ under MNAR at $p = 30\%$..... 38
19 Boxplot of P_e under MAR at $p = 10\%$. 39
20 Boxplot of P_e under MAR at $p = 20\%$. 39
21 Boxplot of P_e under MAR at $p = 30\%$. 40
22 Boxplot of P_e under MCAR at $p = 10\%$. 40
23 Boxplot of P_e under MCAR at $p = 20\%$. 41
24 Boxplot of P_e under MCAR at $p = 30\%$. 41
25 Boxplot of P_e under MNAR at $p = 10\%$. 42
26 Boxplot of P_e under MNAR at $p = 20\%$. 42
27 Boxplot of P_e under MNAR at $p = 30\%$. 42
28 Boxplot of $\mathbf{V}$ under MAR at $p = 10\%$. 43
29 Boxplot of $\mathbf{V}$ under MAR at $p = 20\%$. 43
30 Boxplot of $\mathbf{V}$ under MAR at $p = 30\%$. 44
31 Boxplot of $\mathbf{V}$ under MCAR at $p = 10\%$..... 44
32 Boxplot of $\mathbf{V}$ under MCAR at $p = 20\%$..... 45
33 Boxplot of $\mathbf{V}$ under MCAR at $p = 30\%$..... 45
34 Boxplot of $\mathbf{V}$ under MAR at $p = 10\%$. 45
35 Boxplot of $\mathbf{V}$ under MAR at $p = 20\%$. 46
36 Boxplot of $\mathbf{V}$ under MAR at $p = 30\%$. 46

List of Tables

1 Multivariate datasets chosen to perform the real-data study 28
2 Summary of G×E datasets 28
3 P_e by percentage and method — Mechanism **MAR** (REP 1–5) 48
4 P_e by percentage and method — Mechanism **MCAR** (REP 1–5) 49
5 P_e by percentage and method — Mechanism **MNAR** (REP 1–5) 49
6 P_e by percentage and method — Mechanism **MAR** (REP 6–10) 50
7 P_e by percentage and method — Mechanism **MCAR** (REP 6–10) 50
8 P_e by percentage and method — Mechanism **MNAR** (REP 6–10) 51
9 P_e by percentage and method — Mechanism **MAR** (REP 11–15) 51
10 P_e by percentage and method — Mechanism **MCAR** (REP 11–15) 52
11 P_e by percentage and method — Mechanism **MNAR** (REP 11–15) 52
12 P_e by percentage and method — Mechanism **MAR** (REP 16–17) 52
13 P_e by percentage and method — Mechanism **MAR** (REP 16–17) 53
14 P_e by percentage and method — Mechanism **MCAR** (REP 16–17) 53
15 P_e by percentage and method — Mechanism **MNAR** (REP 16–17) 53
16 P_e by percentage and method — Mechanism **MAR** (REP 1–5) 54
17 P_e by percentage and method — Mechanism **MCAR** (REP 1–5) 55
18 P_e by percentage and method — Mechanism **MNAR** (REP 1–5) 55
19 P_e by percentage and method — Mechanism **MAR** (REP 6–10) 56
20 P_e by percentage and method — Mechanism **MCAR** (REP 6–10) 56
21 P_e by percentage and method — Mechanism **MNAR** (REP 6–10) 57
22 P_e by percentage and method — Mechanism **MAR** (REP 11–15) 57
23 P_e by percentage and method — Mechanism **MCAR** (REP 11–15) 58
24 P_e by percentage and method — Mechanism **MNAR** (REP 11–15) 58
25 P_e by percentage and method — Mechanism **MAR** (REP 16–17) 58
26 P_e by percentage and method — Mechanism **MCAR** (REP 16–17) 59
27 P_e (top) by percentage and method — Mechanism **MNAR** (REP 16–17) .. 59
28 Prediction error (P_e) under MAR (grouped by missingness percentage) 67
29 Prediction error (P_e) under MCAR (grouped by missingness percentage) 68
30 Prediction error (P_e) under MNAR (grouped by missingness percentage) 69
31 Variance among imputations (**V**) under MAR (grouped by missingness percentage) 70
32 Variance among imputations (**V**) under MCAR (grouped by missingness percentage) 71
33 Variance among imputations (**V**) under MNAR (grouped by missingness percentage) 72
34 Prediction error (P_e) under MAR (grouped by missingness percentage) 73
35 Prediction error (P_e) under MCAR (grouped by missingness percentage) 74
36 Prediction error (P_e) under MNAR (grouped by missingness percentage) 75
37 Variance among imputations (**V**) under MAR (grouped by missingness percentage) 76
38 Variance among imputations (**V**) under MCAR (grouped by missingness percentage) 77
39 Variance among imputations (**V**) under MNAR (grouped by missingness percentage) 78
40 V by percentage and method — Mechanism **MAR** (REP 1–5) 79

41 V by percentage and method — Mechanism **MCAR** (REP 1–5) 80
42 V by percentage and method — Mechanism **MNAR** (REP 1–5) 81
43 V by percentage and method — Mechanism **MAR** (REP 6–10) 82
44 V by percentage and method — Mechanism **MCAR** (REP 6–10) 83
45 V by percentage and method — Mechanism **MNAR** (REP 6–10) 84
46 V by percentage and method — Mechanism **MAR** (REP 11–15) 85
47 V by percentage and method — Mechanism **MCAR** (REP 11–15) 86
48 V by percentage and method — Mechanism **MNAR** (REP 11–15) 87
49 V by percentage and method — Mechanism **MAR** (REP 16–17) 88
50 V by percentage and method — Mechanism **MCAR** (REP 16–17) 88
51 V by percentage and method — Mechanism **MNAR** (REP 16–17) 89
52 V by percentage and method — Mechanism **MAR** (REP 1–5) 90
53 V by percentage and method — Mechanism **MCAR** (REP 1–5) 91
54 V by percentage and method — Mechanism **MNAR** (REP 1–5) 92
55 V by percentage and method — Mechanism **MAR** (REP 6–10) 93
56 V by percentage and method — Mechanism **MCAR** (REP 6–10) 94
57 V by percentage and method — Mechanism **MNAR** (REP 6–10) 95
58 V by percentage and method — Mechanism **MAR** (REP 11–15) 96
59 V by percentage and method — Mechanism **MCAR** (REP 11–15) 97
60 V by percentage and method — Mechanism **MNAR** (REP 11–15) 98
61 V by percentage and method — Mechanism **MAR** (REP 16–17) 99
62 V by percentage and method — Mechanism **MCAR** (REP 16–17) 99
63 V by percentage and method — Mechanism **MNAR** (REP 16–17) 100

1 Introduction

From the beginning of the career, it was learned that when a data scientist starts working with a dataset, it is essential to follow a systematic process to extract the maximum value possible. One of the most renowned methodologies is Cross-Industry Standard Process for Data Mining (CRISP-DM), which comprises several phases one of them being data preparation. Data preparation, or preprocessing, is a critical phase for ensuring the proper performance of any model, analysis, or subsequent value extraction. If your data is flawed from the outset, your conclusions, predictions, analyses, and results will inevitably suffer. It is akin to manufacturing a t-shirt from low-quality fabric eventually, it will deteriorate and not look good, among other issues.

One of the most notable aspects addressed during data preprocessing is missing data. This phenomenon occurs more frequently than expected. As mentioned in the book of Templ (2023) missing data can arise for various reasons depending on the study's context. In surveys and official statistics, missing data may result from overly extensive or complicated questionnaires, questions that are too personal, unrealistic answers (according to editing rules), missing responses, a lack of knowledge, or reluctance to answer sensitive questions. In measurement experiments, it can be caused by discontinuations in follow-up studies, longitudinal designs with high dropout rates, sensor failures or defective units, and implausible measurements. Additionally, censored values may be due to detection limits in chemical measurements, while the loss of monotonicity is observed, for example, when a patient withdraws from a clinical study or when related variables cease to provide consistent data.

Now, although ideally when a missing data point occurs one would try to collect it (for example, in surveys where it is possible to contact the respondent after all, there is no better estimate than reality), there are many cases in which, simply due to the nature of the data collection process or due to errors within it, this is impossible. In such situations, it is incumbent upon the data scientist to employ various techniques that "fill in" these missing gaps, always ensuring that the inserted value introduces as little bias as possible and represents the best estimate of the true value of that observation, so that it does not negatively affect the results of the study and/or subsequent model.

Now, why should we care about missing data? Well, missing data is a significant concern because its absence can render many analyses impractical. For example, Gabriel (2002) [9] notes that techniques based on singular value decomposition (SVD) become unfeasible when data is incomplete, limiting methods such as principal components analysis (PCA) and biplots. Similarly, Brahim et al. (2012) shows that directly applying Cox models to incomplete datasets may lead to erroneous conclusions. In fact, it is not just that functions in statistical software generate errors the improper handling of missing data can also reduce a study's statistical power and lead to biased estimates, which in turn may result in invalid conclusions. Furthermore, as mentioned by Schafer and Olsen (1998), methods that eliminate rows can lead to a high rate of data loss, potentially resulting in insufficient information to represent all possible samples of the data and thus introducing

bias.

For this, two methods are proposed: simple imputation and multiple imputation (MI). To highlight the main difference between the two, simple imputation involves replacing missing values with a single estimated value. This value can be, for example, the mean, median, mode, or even a prediction from a model. On the other hand, multiple imputations generate several imputed datasets. For each missing value, rather than assigning a single value, several plausible imputations are produced based on the probability distribution of the observed data.

Regarding which methodology to use, the literature clearly outlines the limitations of one approach over the other. For instance, as mentioned in Schafer and Olsen (1998) simple imputation techniques such as using the mean, this imputation replaces missing values with the average of the available values for a given variable, but it comes with certain drawbacks. It reduces variability, especially when many values are missing, and can create an artificial sense of similarity among individuals with imputed values (Krzanowski, 2007). To address these issues, a common solution is to introduce random variation by adding a value drawn from a normal distribution with a mean of zero and a variance equal to that of the original variable. This adjustment helps restore variability and minimizes the risk of misleading patterns in the data. Additionally, Arciniegas-Alarcón et al. (2014) highlights that simple imputation does not account for the uncertainty inherent in the imputation process. When parameters are estimated from imputed data, the standard errors tend to be underestimated, which can invalidate confidence intervals and hypothesis tests. These issues are more effectively addressed through multiple imputation methods. By generating multiple complete datasets and combining the results into a single inference, multiple imputation mitigates these problems. Moreover, as demonstrated by Madley et al. (2019), multiple imputation reduces bias and enhances efficiency regardless of the proportion of missing data, assuming that the data are missing at random (MAR).

Various multiple imputation techniques have been extensively studied, resulting in models and proposals for a range of methods from Bayesian methods based on Markov Chain Monte Carlo (MCMC) to techniques involving generalized additive models (Enders, 2023). However, for this particular work, the goal is to study more deeply into SVD-based models, some of those discussed in García-Peña et al. (2021). To provide some context, these models are variations of SVD designed to be more robust in the presence of outliers. Among the SVD variants presented, rSVD84 and rSVD01 (these are robust SVD methods, the numbers 84 and 01 refer to the years in which they were proposed), stood out as they performed exceptionally well in matrices with contamination levels above 2%. They also excelled in simulations with moderate contamination levels such as 10% missing data and, moreover, produced the best and most stable results in real datasets (García-Peña et al., 2021). Thus, the interest lies not only in evaluating the quality of imputations in contaminated matrices using both methods but also in developing a multiple imputation technique based on them. This new method should exhibit high performance in matrices with these characteristics and demonstrate outstanding results in both simulation studies and real datasets.

The idea of this work is to continue the research carried out in García-Peña et al. (2021), exploring in depth the techniques that showed the best performance in the tests con-

ducted. It is proposed to follow the research line that points to the strong performance of rSVD84 and rSVD01. The rSVD84 method starts by computing trimmed means for each row and column to form initial left and right vectors, then iteratively refines these estimates using medians of ratios (between data entries and current estimates) until convergence; the robust rank-1 approximation is then extracted, and further components are obtained by deflating the matrix. In contrast, the rSVD01 method employs L1 regression which minimizes the sum of absolute deviations to initially derive a left vector from the medians of each row, then estimates column coefficients through L1 regression to form a normalized right vector; it alternates between updating these vectors using L1-based regressions until convergence, and finally computes a scaling factor (eigenvalue) via another L1 regression, with further components extracted using deflation. For missing data of the missing completely at random (MCAR) and missing not at random (MNAR) types, while also assessing their effectiveness under the MAR mechanism, which was not addressed in the original article.

Furthermore, the results of the article cited recommend, for cases with more than 2% contamination in the data, using one of the robust variants of SVD. Therefore, there is an interest in exploring how, from these two methods, a multiple imputation approach can be developed to overcome the limitations of simple imputation having the principal benefit of the SVD that is no structural or distributional assumptions. We know that multiple imputation offers significant advantages by more appropriately handling uncertainty, reducing biases in parameter estimation, and improving the validity of confidence intervals under MAR (Rubin 1987). However, combining this strategy with robust methods for outliers and missing data remains an open area of research.

On the other hand, although the original article analyzes certain levels and types of contamination, it is of great interest to test these methods with other classes of outliers, which through simulation studies and application to real datasets.

2 Objectives

2.1 General Objective

Develop and validate a multiple imputation method based on the robust variants of the singular value decomposition (rSVD84 and rSVD01), enabling effective handling of missing data under different missing-data mechanisms (MCAR, MNAR, and MAR) and in the presence of various levels of outliers.

2.2 Specific Objectives

- Determine the effectiveness of rSVD84 and rSVD01 in imputing missing data under the MAR mechanism.
- Propose a multiple imputation approach based on rSVD84 and rSVD01, incorporating strategies that address the presence of outliers.
- Establish the robustness of the proposal to different contamination and *missing-data* scenarios.

3 Theoretical Framework

3.1 Multiple Imputation

As mentioned before, multiple imputation (Rubin, 1987) is a framework for handling missing data by creating several complete datasets, analyzing each, and combining results. In MI, it is a two-stage process (Rubin, 1987): an imputer fills in missing values with random draws from an appropriate model (creating m completed datasets), then an analysis is performed on each completed dataset and the estimates are pooled using Rubin's rules (Little and Rubin, 2002).

3.1.1 Rubin's Rules

These rules are a group of formulas that let us pool the different datasets considering the different sources of variability:

- **Uncertainty within each imputation:** the variability of the different parameters estimated by each dataset.
- **Variability between imputations:** the estimates generated by the different datasets can vary slightly.

If we take Q_i as the estimation of the parameter of interest in the i-th imputation (e.g., the mean), m as the number of completed datasets, and U_i as the variance of the estimate in the i-th imputation (squared standard error), then:

Mean estimation of the parameter:

$$\bar{Q} = \frac{1}{m}\sum_{i=1}^{m} Q_i$$

Within-imputation variance:

$$\bar{U} = \frac{1}{m}\sum_{i=1}^{m} U_i$$

Between-imputation variance:

$$B = \frac{1}{m-1}\sum_{i=1}^{m}(Q_i - \bar{Q})^2$$

Total variance (used in confidence intervals):

$$T = \bar{U} + \left(1 + \frac{1}{m}\right) B$$

It is important to clarify that multiple imputation was initially derived under the Bayesian framework (Murray, 2018). We are interested in the posterior distribution of the parameter Q given the observed data. When there are missing data $(Y_{\text{mis}}, Y_{\text{obs}})$, the full posterior is expressed by integrating over the possible realizations of the missing values:

$$P(Q \mid Y_{\text{obs}}) = \int P(Q \mid Y_{\text{mis}}, Y_{\text{obs}}) P(Y_{\text{mis}} \mid Y_{\text{obs}})\, dY_{\text{mis}}$$

This leads to expressions equivalent to the frequentist formulation:

$$\bar{Q} \approx \mathbb{E}\left[\mathbb{E}(Q \mid Y_{\text{mis}}, Y_{\text{obs}}) \mid Y_{\text{obs}}\right] = \mathbb{E}(Q \mid Y_{\text{obs}})$$

$$\bar{U} \approx \mathbb{E}\left[\text{Var}(Q \mid Y_{\text{mis}}, Y_{\text{obs}}) \mid Y_{\text{obs}}\right]$$

$$\left(1 + \frac{1}{m}\right) B \approx \text{Var}\left[\mathbb{E}(Q \mid Y_{\text{mis}}, Y_{\text{obs}}) \mid Y_{\text{obs}}\right]$$

$$T \approx \text{Var}(Q \mid Y_{\text{obs}}), \quad \hat{SE} = \sqrt{T}$$

The pooled standard error can be used to construct confidence intervals (Little and Rubin, 2002), noting that:

$$(\bar{Q} - Q) \sim t_\nu(0, T), \quad \nu = (m-1)\left(1 + \frac{1}{r}\right)^2, \quad r = \left(1 + \frac{1}{m}\right)\frac{B}{\bar{U}}$$

for large samples (Murray, 2018).

3.1.2 Classical MI Methods (MICE, Model-Based, Nonparametric)

Early MI methods were often based on joint modeling of the data for instance, assuming a multivariate normal distribution for all variables and drawing imputations from the posterior predictive distribution as in Schafer's Bayesian MVN imputation model (Schafer and Olsen, 1998). Another common approach is Fully Conditional Specification (FCS), also known as Multivariate Imputation by Chained Equations (MICE). MICE defines an imputation model for each variable with missing data and iteratively imputes each variable in turn, conditioning on the latest imputations of the others (van Buuren, 2012). This effectively samples from an implied multivariate distribution without requiring an explicit joint density formula. For example, one can specify a regression for a partially observed variable Y_o given other variables, cycle through all o, and repeat multiple iterations to generate imputations. Van Buuren's implementation of MICE in R (2011) popularized this approach (Murray, 2018).

Model-based MI refers to methods that explicitly assume a parametric distribution for the data for example, a Bayesian parametric imputation where one draws parameter values from a posterior and then draws missing entries from the assumed model (Rubin's original approach).

Following Arciniegas-Alarcón et al. (2020), the Approximated Bayesian Bootstrap (ABB) is a nonparametric, resampling-based multiple imputation procedure that avoids specifying parametric distributions while still delivering well-calibrated (near-unbiased) variance estimates across imputations. The method operates with two successive sampling-with-replacement stages. Consider a vector $\mathbf{x} = (x_1, \ldots, x_n)^\top$ with the first r entries observed and the remaining $n - r$ missing. In the first stage, ABB forms a donor version of the observed data by bootstrapping the r observed values, drawing each donor element from $\{x_1, \ldots, x_r\}$ with equal probability, $\Pr(x_i^* = x_j) = 1/r$. In the second stage, each missing entry is imputed by sampling from this donor pool with replacement, again under equal probabilities, yielding a completed vector $\tilde{\mathbf{x}}$ that reflects uncertainty from both stages of resampling.

The two-layer resampling construction can be viewed as generating an approximate draw from the posterior predictive distribution of the missing values given the observed data, that is,

$$\tilde{\mathbf{X}}_{\text{miss}} \;\approx\; p(\mathbf{X}_{\text{miss}} \mid \mathbf{X}_{\text{obs}})\,.$$

For a two-way $G \times E$ data matrix $\mathbf{X}$ with observed and missing cells, ABB first bootstraps the observed entries to create a donor matrix $\mathbf{X}^*_{\text{obs}}$ and then imputes each missing cell by sampling with replacement from that donor set. Repeating the entire mechanism M times produces multiple completed datasets, and standard multiple-imputation combining rules can then be applied to integrate within- and between-imputation variability.

Nonparametric and machine learning MI methods have also emerged. For instance, the random forest imputation method *missForest* (Stekhoven and Bühlmann, 2012) uses an ensemble of regression trees to impute missing values. By averaging over many unpruned trees, random forests inherently produce an aggregated prediction that can be viewed as a form of "multiple" imputation, since each tree can be seen as a sample. Unlike parametric models, such algorithms can capture nonlinear relationships.However missForest excels an imputation prediction by outputting deterministic point estimates, whereas Rubin's Multiple Imputation (MI) requires stochastic draws to represent uncertainty across datasets; without this randomness, missForest can't capture between imputation variability or yield valid standard errors, and rerunning it with different seeds doesn't fix the problem it introduces bias and overly narrow confidence intervals (Shah et al., 2014, p. 768). Moreover, key Rubin-style techniques like estimating variance among imputations aren't possible with classic missForest; the paper suggests using mice with Random Forest as a proper MI alternative that generates multiple datasets via random draws and supports Rubin's combining rules. Other nonparametric approaches include nearest-neighbor hot-deck imputation (which defines distances from "similar" observed cases), and predictive mean matching, which ensures imputed values are actual observed values from k nearby donor cases (Murray, 2018).

3.1.3 Evolution in recent literature

Over the last 10–15 years, MI strategies have greatly diversified. Researchers have developed more flexible imputation models to relax the assumptions of normality or linearity. For example, tree-based and forest-based imputations have been incorporated into the MI paradigm (e.g., using CART or random forests within MICE). Semi-parametric and fully

nonparametric Bayesian approaches, like Bayesian Dirichlet process mixtures, have been proposed to handle arbitrary variable distributions (Murray, 2018).

Another line of recent work involves high-dimensional and large-scale multiple imputation. Traditional MICE can struggle with datasets that contain hundreds of variables or millions of observations due to computational cost and convergence issues. In response, strategies like sequential imputation (grouping variables), parallel computing, and dimension-reduction techniques (e.g., PCA-based imputation) have been introduced. Ensemble approaches that integrate imputation with predictive modeling such as lasso, random forests, or deep learning have also shown promise.

Multiple imputation has extended to new data structures: for longitudinal data, mixed-effects models can be used; for hierarchical data (e.g., students within schools), hierarchical imputation models are applied; and for genetic or image data, models that incorporate spatial or network structure are needed.

FCS has also seen refinements for example, using logistic regression for binary data, ordinal regression for categorical data, and predictive mean matching or random forests for continuous variables, all within the chained equations framework. This flexibility has made MI highly practical and popular in research and applied fields.

Looking ahead, recent work explores imputation with deep generative models such as variational autoencoders (VAEs) and generative adversarial networks (GANs) to produce multiple imputations. There is also growing interest in handling MNAR data by incorporating the missingness mechanism into the imputation model (Ma and Zhang, 2021).

In genotype-by-environment (G×E) agricultural research, for instance, MI can fill missing yield observations for certain genotype-environment combinations, allowing full data matrices for analyses like stability analysis or AMMI modeling. MI generally assumes data are missing at random (MAR). If data are missing not at random (MNAR), and the imputation model ignores the missingness process, biased estimates may result (Baraldi et al., 2010). Some limitations of MI include the difficulty of diagnosing imputation quality (e.g., verifying the distribution of imputed values) and the computational cost of storing multiple completed datasets and running analyses on each. However, software optimizations and parallelization can help mitigate these issues.

3.2 Missing Data Mechanisms

3.2.1 Missing Data Patterns

A critical part of any missing data analysis is understanding why the data are missing. Rubin (1976) introduced three broad classes of missing data mechanisms: Missing Completely at Random (MCAR), Missing at Random (MAR), and Missing Not at Random (MNAR). These have precise probabilistic definitions: let R be an indicator (mask) for missingness (0 or 1) and $Y_{\text{mis}}, Y_{\text{obs}}$ as defined in the section before. MCAR means the probability of missingness is independent of all data values, formally Carpenter and Kenward

(2005):

$$P(R \mid Y_{\text{mis}}, Y_{\text{obs}}) = P(R)$$

In simple terms, nothing in the data influences what is missing; the missing values are a random subset of all values, an example given by Templ (2023) is when respondents may decide to omit a question due to loss of interest.
In MAR, missingness may depend on observed data but not on other variables. Formally Carpenter and Kenward (2005):

$$P(R \mid Y_{\text{mis}}, Y_{\text{obs}}) = P(R \mid Y_{\text{obs}})$$

Intuitively, given what we do observe about a case, we can explain the missingness; there is no additional dependence on the unseen data. For instance, in a medical study participants with higher income (observed) might be less likely to report their age (making age MAR missingness, if income is recorded and age's missingness only depends on income) (Templ, 2023).
MNAR is when the probability of missingness depends on the values that are missing (even after considering observed data):

$$P(R \mid Y_{\text{mis}}, Y_{\text{obs}}) \neq P(R \mid Y_{\text{obs}})$$

A example is when the weight of heavy mammals often cannot be measured (Templ, 2023).

3.2.2 Diagnosing Missing Data Patterns

Before choosing a handling method, one should explore the patterns and potential causes of missingness. One of the simplest approaches is the t-test, it is applied when one variable $\mathbf{X}$ is completely observed, and another y has missing values. Two groups are created for the variable $\mathbf{X}$ based on whether the corresponding value in y is observed or missing. A t-test is performed to compare the means of $\mathbf{X}$ between these two groups. The null hypothesis is that the means are equal. This test helps assess whether missing data follows an MCAR or in case that rejects the null hypothesis it could suggest an MAR mechanism (because the means of the two groups are significantly different). Alternatively, a Wilcoxon test can be used, which is a non-parametric method that is less sensitive to outliers and does not require strict distributional assumptions, though it evaluates only one covariate at a time (Templ, 2023).
A widely recognized test for MCAR is Little's MCAR test, examines whether the missingness in multivariate data is truly random by comparing observed sample means across different missing data patterns. In this test, each observation y_i and covariance Σ, although certain variables may be missing. The data are divided into J groups according to

which parts of each observation were recorded. For each group j, calculate the average of the observed parts (denoted by $\bar{y}_{oj}$). If the missing data is completely random (null hypothesis, H_0), these averages should be equal to the expected mean μ_{oj} across all groups; otherwise, if they are different (denoted by v_{oj}), the missing data might not be random. Little proposes a test statistic:

$$d_o^2 = \sum_{j=1}^{J} n_j (\bar{y}_{oj} - \hat{\mu}_{oj})^T \left(\frac{n\hat{\Sigma}_{oj}^{-1}}{n-1} \right) (\bar{y}_{oj} - \hat{\mu}_{oj})$$

If H_0 is true, this statistic follows a chi-square distribution with degrees of freedom given by the sum of the group sizes p_j minus the total number of variables p.

$$H_0 : \quad y_{0i} \mid r_i \sim N(\mu_{oj}, \Sigma_{oj}) \quad \text{if } i \in I_j,\ 1 \leq j \leq J$$
$$H_1 : \quad y_{0i} \mid r_i \sim N(v_{oj}, \Sigma_{oj}) \quad \text{if } i \in I_j,\ 1 \leq j \leq J$$

The hypothesis of complete randomness is rejected if the value of the statistic exceeds a certain critical threshold:

$$d^2 > \chi^2_{\sum_{j=1}^{J}(p_j - p)}(1-\alpha) \quad \text{(Li, 2013)}$$

Other specialized tests, such as those in the R package `MissMech`, use methods like the Anderson-Darling k-samples test. If this is rejected, it is concluded that the covariances are not homogeneous (there is heteroscedasticity), which in turn indicates that the data are not MCAR.

For a more comprehensive assessment, generalized linear models (such as logistic regression) can be used, where the missingness indicator is regressed on fully observed or imputed covariates. For example, if the indicator r_i is defined, where $r_i = 1$ if the covariate vector is complete and $r_i = 0$ if any component is missing, it is suggested that, under the MAR (Missing at Random) assumption, modeling the probability $\pi_i = P(r_i = 1 \mid \ldots)$ using logistic regression (a type of GLM) is advisable. This approach leverages observed (or even imputed) covariates to predict missing data patterns (Ibrahim et al. 2011).

Detecting whether data are MNAR is inherently challenging and, in practice, often impossible based solely on the observed data. This is because the defining characteristic of MNAR is that the probability of missingness depends on the missing values themselves, which are, by definition, unobserved. As a result, distinguishing MNAR from other missing data mechanisms is frequently unfeasible. For instance, if the missingness correlates with observed variables, a MNAR scenario might be misclassified as MAR; if there is no correlation, it may be mistaken for MCAR. These challenges are further compounded in multivariate datasets with missing values across multiple variables. Unlike MNAR, MAR mechanisms can be investigated using visual tools (e.g., parallel coordinate plots, scatterplots, boxplots, matrix plots, and mosaic plots) that assess the relationship between missingness in one variable and observed values in others. Additionally, statistical tests such as Little's MCAR test can assess the null hypothesis of MCAR, though rejecting it does not confirm MNAR. Ultimately, reliably identifying MNAR typically requires domain expertise, additional data collection, or explicitly modeling the missingness mechanism tasks that are often complex and resource intensive (Templ, 2023).

3.2.3 Simulating Missing Data Mechanisms

To evaluate or demonstrate methods, it is often useful to simulate data with controlled missingness mechanisms. For MCAR, simulation is straightforward: after generating a complete dataset, one can "delete" values completely at random. For example, to impose 20% MCAR missingness in a variable, randomly choose 20% of the rows and set that variable to missing.
To simulate MAR, one needs to make missingness depend on other observed data. A common approach is to use a logistic model for the missingness indicator. For instance, generate a full dataset, then for each value of Y decide if it will be missing by:

$$P(\text{missing } Y \mid X) = \text{logit}^{-1}(\alpha + \beta X)$$

where X is some other variable. By adjusting β, Y's missingness can depend on X (which remains observed). If $\beta > 0$, maybe high values of X cause Y to be more likely missing. This yields a MAR mechanism (since X is observed).
For MNAR, the missingness must depend on the would-be value of the variable itself (or on other missing values). In simulation, this can be done by a two-step process: generate a complete dataset including the "true" values of Y, then remove values of Y with a probability depending on Y itself:

$$P(Y_{\text{mis}}) = \text{logit}^{-1}(\gamma + \delta Y)$$

Here, if $\delta > 0$, perhaps larger Y values are more likely to be missing. This is MNAR because whether a given Y is observed depends on its unseen true value. Another simple MNAR mechanism is threshold-based: e.g., "censor" data above a threshold values above T are recorded as missing. This mimics detection limit problems (common in chemistry or assay data) where readings too high or too low are not recorded inherently MNAR relative to the true value (Zhang, 2021).In practice, the R package `mice` provides an `ampute` function that can generate missing data under MCAR, MAR, or MNAR by specifying a logistic weighting scheme.

3.3 Outliers

3.3.1 Definitions and Types of Outliers

An outlier is generally understood as an observation that "lies far from the rest" of the data, an anomaly that doesn't fit the pattern of the majority. We can categorize outliers in several ways.
Outliers can be classified based on the number of instances involved. Point outliers are individual instances that significantly deviate from the rest of the data. Within this category, there are two subtypes. Global outliers are detected by comparing each instance with the entire dataset, identifying those that stand out on a broad scale. On the other

hand, local outliers are identified based on discrepancies observed only within their immediate neighborhood, meaning they may appear normal when compared to the overall dataset but are unusual within their local context (Boukerche and Zheng, 2020).
Another category is collective outliers, which refer to a group of instances that, while appearing normal individually, form an anomalous pattern when considered together. These types of anomalies are particularly relevant in cases where relationships between data points play a crucial role in identifying unusual behavior (Boukerche and Zheng, 2020).
Outliers can also be classified based on the type of input data. Vectorial outliers are found in multidimensional (or vector-based) datasets, where each instance has multiple attributes. These outliers are typically detected using distance-based measures, such as Euclidean or cosine distance, which allow for the comparison of vectors to determine deviations from expected patterns (Boukerche and Zheng, 2020).
In structured data represented as graphs, graph outliers emerge. These anomalies can manifest as nodes or edges that differ significantly from the rest of the graph, or even as groups of nodes (subgraphs) that collectively exhibit abnormal behavior. The detection of graph outliers is particularly relevant in network analysis, social media, and fraud detection, where structural relationships provide crucial insights (Boukerche and Zheng, 2020).

3.3.2 Outlier Detection Methods

A wide range of techniques exists to detect outliers, from simple statistical rules to complex algorithms. Some common approaches include:
Statistical rule-based methods (univariate): These are often the first step. For a single variable, one classic rule is the Z-score method: flag any data point whose value is more than say 3 standard deviations away from the mean (assuming a roughly normal distribution). Another is Tukey's IQR rule: compute the first and third quartiles (Q1, Q3) and the interquartile range

$$\text{IQR} = Q_3 - Q_1$$

Any observation below

$$Q_1 - 1.5 \cdot \text{IQR} \quad \text{or above} \quad Q_3 + 1.5 \cdot \text{IQR}$$

is considered an outlier. These rules are easy to apply but has the normal distribution asumption. They primarily target univariate outliers.
Distance-based methods (multivariate): For a p-dimensional observation vector X_i, compute

$$D_i = \sqrt{(X_i - \bar{X})^T \Sigma^{-1} (X_i - \bar{X})}$$

where $\bar{X}$ is the mean vector and Σ the covariance matrix of the data. D_i measures how far X_i is from the center of the data in multivariate units. If the data are roughly normal, D_i^2 follows a chi-square distribution with p degrees of freedom, so one can flag points with D_i exceeding a high quantile of χ_p^2 as outliers (Li et al., 2019).

Another, more recent approach casts PCA as a decomposition problem: one seeks to write the data matrix X as

$$X = L + S$$

where L is low-rank (captures the main structure) and S is a sparse matrix of outliers. This formulation, known as Robust PCA via Principal Component Pursuit, can be solved by convex optimization:

$$\min \left(\|L\|_* + \lambda \|S\|_1 \quad \text{subject to} \quad X = L + S \right)$$

where $\|L\|_*$ is the nuclear norm promoting low rank and $\|S\|_1$ promotes sparsity (Candès et al., 2011).

Remarkably, under certain conditions, one can perfectly recover the true low-rank structure despite a fraction of arbitrary outliers in S. Such robust methods (Candès et al., 2011) aim to automatically handle outliers without explicitly identifying and removing them; instead, they model them.

In the context of genotype-by-environment data, robust methods have been applied to the AMMI model (which uses SVD/PCA on interaction residuals). The AMMI (Additive Main effects and Multiplicative Interaction) model first fits additive genotype and environment effects (the "main effects") and then applies SVD to the residual matrix of genotype×environment interaction effects (Rodrigues et al., 2016). The resulting singular vectors provide interaction principal components, which can be used to visualize which genotypes and environments have particularly strong positive or negative interactions. A robust AMMI replaces the least-squares SVD with a robust SVD so that a few bad data points do not distort the interaction decomposition (Rodrigues et al., 2016).

3.3.3 Generating Outlier Simulations

To evaluate how methods perform in the presence of outliers, we can simulate data with controlled outlier contamination. A simple model is the contaminated normal model: generate the majority of data from a baseline distribution (say a normal $N(0, 1)$), and a small fraction from a different distribution that produces outliers, such as a normal with a much larger variance (García-Peña et al., 2021). Indeed, the robust AMMI study referenced earlier used Monte Carlo simulations with various "contamination schemes" injecting outliers of different sizes and positions to test the robust vs classical AMMI models (Rodrigues et al., 2016).

3.4 Singular Value Decomposition (SVD)

3.4.1 SVD Theoretical Foundations and Applications

The singular value decomposition (SVD) is a fundamental matrix factorization in linear algebra with wide applications in data analysis. For any $n \times p$ data matrix X, the SVD factors it as

$$X = U\Sigma V^T$$

where U is a $n \times n$ orthogonal matrix, V is a $p \times p$ orthogonal matrix, and Σ is a $n \times p$ diagonal matrix with nonnegative entries (the singular values) on the diagonal. Intuitively, SVD rotates the data into a new orthogonal coordinate system (given by columns of V for the columns of X, and columns of U for the rows), such that the axes are ordered by the amount of variability in X they explain. The singular values σ_i are the square roots of the eigenvalues of XX^T, and they indicate the magnitude of the data along the corresponding singular vectors.

3.4.2 SVD as imputation method

According to the paper of Krzanowski (1988) introduces a novel method for imputing missing values in multivariate data based on the Singular Value Decomposition (SVD) of a matrix. This technique is proposed as a "perfectly general" solution, designed to be free of the distributional or structural assumptions, such as multivariate normality, that underpin many existing imputation methods like regression-based approaches or the E-M algorithm. The core of the method involves estimating a missing value by using the SVD components derived from the remaining data. This method was "highly promising" and broadly applicable to any numerical data matrix, particularly in situations where principal component analysis is appropriate, due to its key advantage of not depending on any assumptions about the data.

3.4.3 Sensitivity of SVD to Outliers and Missing Data

While SVD/PCA is a powerful tool, it relies on the data as given and uses the least squares criterion implicitly (because it finds axes that minimize the sum of squared perpendicular distances to the data(García-Peña et al., 2025;Arciniegas-Alarcón et al., 2022)). This makes it sensitive to outliers. An extreme outlier can dominate the variance and thus skew the first principal component direction towards itself. For example, if in a G×E yield matrix one entry is erroneously high by a large amount, a standard PCA might devote one component largely to capturing that single high value (instead of reflecting the broader structure of interactions), as noted: even a single extreme outlier can draw the leading principal component toward itself (Rodrigues et al., 2016).
Similarly, SVD is not robust to violations of modeling assumptions (e.g., heavy-tailed noise). Because it minimizes squared errors, a point that is far off will contribute a disproportionately large amount to the sum of squared deviations, and PCA will rotate to accommodate that point.
Regarding missing data, standard SVD requires a complete matrix. The mathematics of SVD does not directly allow skipping missing entries. A naive approach to apply PCA with missing data might be to omit any case with missing values (which wastes data) or to fill missing values with some estimate (which can bias the components if done poorly, as mentioned before). Because of this, specialized algorithms for PCA with missing data have been developed.

One approach is an EM algorithm for PCA: start by guessing the missing values (fill with mean or zeros), do an SVD to get principal components, then project the data onto the principal components to estimate the missing values, and iterate. This is essentially what the SVD-impute method does for gene expression missing data (Kim et al., 2004 ; Krzanowski, 1988).

3.5 Robust Alternatives to SVD (rSVD84 and rSVD01)

3.5.1 rSVD84

rSVD84 is a robust variation of SVD designed to handle missing data and outliers effectively. It is based on the work of Gabriel and Odoroff (1984) but incorporates trimmed means and outlier detection techniques to improve robustness. Given a matrix $Y \in \mathbb{R}^{n \times p}$ with possible missing values, the method follows these steps (García-Peña et al., 2021):

Compute Initial Trimmed Means Compute the trimmed means (at 10% or 20%) for both rows and columns of Y, defining:

$$b_1 = \text{trimmed mean of columns of } Y \quad (1 \times p)$$

$$a_1 = \text{trimmed mean of rows of } Y \quad (n \times 1)$$

Detect and Replace Outliers in a_1 Detect outliers in a_1 using a univariate method and replace them with the trimmed mean of the elements of a_1.

Iterative Updating of a_1 **and** b_1 Update elements using the following median-based formulas:

$$b_c = \text{median}\left(\left|\frac{y_{r,c}}{a_r}\right|, \quad r = 1, \ldots, n\right)$$

$$a_r = \text{median}\left(\left|\frac{y_{r,c}}{b_c}\right|, \quad c = 1, \ldots, p\right)$$

Repeat Steps 2 and 3 iteratively until the values a_1 and b_1 converge (stabilize).

Compute Robust Low-Rank Approximation

Once a_1 and b_1 are stable, the rank-one approximation of Y is given by:

$$\tilde{Y} = a_1 b_1^T$$

where the signs of the elements match those of Y.

Compute First Singular Value and Vectors

Apply SVD to $\tilde{Y}$:

$$\tilde{Y} = U\Sigma V^T$$

Extract:

1. First singular value: σ_1
2. First left singular vector: u_1
3. First right singular vector: v_1

Compute Further Singular Values and Vectors

Deflate Y by subtracting the first rank-one approximation:

$$Y' = Y - a_1 b_1^T$$

Repeat the entire process (Steps 1–5) using Y' instead of Y to obtain the second singular value and singular vectors, and continue iteratively until reaching the desired number of singular components.

3.5.2 rSVD01

rSVD01 is a robust variation of SVD designed to handle missing values and outliers. Unlike traditional SVD, which is based on minimizing squared errors (L2 norm), this method uses L1 regression, minimizing absolute deviations, to obtain a rank-one approximation of the data matrix $Y \in \mathbb{R}^{n \times p}$. The method follows these steps:

Compute Initial Left Eigenvector a_1

We begin by estimating the first left eigenvector a_1. The initial estimate of a_1 is computed as the median of absolute values in each row:

$$a_{1,i} = \text{median}\left\{|y_{i,j}| : \text{for observed } y_{i,j}\right\}, \quad i = 1, \ldots, n$$

This results in a vector $a_1 \in \mathbb{R}^{n \times 1}$.

Compute Right Eigenvector b_1 via L1 Regression

With a_1 in hand, we estimate the first right eigenvector b_1. For each column $j = 1, \ldots, p$, we perform an L1 regression to find a scalar coefficient c_j that minimizes:

$$\min_{c_j} \sum_{i=1}^{n} |y_{i,j} - c_j \cdot a_{1,i}|$$

The resulting set of coefficients forms a vector $c \in \mathbb{R}^{p \times 1}$. Normalize c using the Euclidean norm to obtain b_1:

$$b_1 = \frac{c}{\|c\|}, \quad \text{where } \|c\| = \sqrt{\sum_{j=1}^{p} c_j^2}$$

This ensures b_1 is a unit vector.

Update Left Eigenvector a_1 via L1 Regression

Now, using b_1, we refine a_1 by performing another L1 regression.

For each row $i = 1, \ldots, n$, find a coefficient d_i that minimizes:

$$\min_{d_i} \sum_{j=1}^{p} |y_{i,j} - d_i b_{1,j}|$$

Collect these coefficients into a vector $d \in \mathbb{R}^{n \times 1}$. Normalize d to get an updated a_1:

$$a_1 = \frac{d}{\|d\|}, \quad \text{where } \|d\| = \sqrt{\sum_{i=1}^{n} d_i^2}$$

Iterate Until Convergence

Steps 2 and 3 are repeated iteratively, refining a_1 and b_1, until the difference between successive iterations is below a given threshold. Once convergence is reached, we obtain a robust rank-one approximation of Y:

$$Y_1 \approx \lambda_1 a_1 b_1^T$$

Compute First Singular Value λ_1

Once a_1 and b_1 are stabilized, the first singular value λ_1 is estimated by solving:

$$\min_{\lambda_1} \sum_{i=1}^{n} \sum_{j=1}^{p} |y_{i,j} - \lambda_1 a_{1,i} b_{1,j}|$$

Compute Higher-Order Singular Components

To compute additional singular values and vectors, we deflate the matrix:

$$Y' = Y - \lambda_1 a_1 b_1^T$$

Repeat the entire process (Steps 1–5) using Y' to extract the next component λ_2, a_2, b_2, and continue iteratively until the desired number of components is reached.

4 Methodology

To address each of the study's objectives, two strategies will be implemented: a simulation study, and testing it on real datasets, also comparing the performance against 3 gold standard models: ABB, Mice+pmm and Mice+rf. To evaluate the effectiveness of these methods, missing data will be inserted using each of the mechanisms. The idea is to test all three mechanisms on simulated data and real data. To do this, the procedure described by Arciniegas-Alarcón et al. (2020) will be followed.

4.1 Simulation study

One hundred complete matrices $X_{100\times 8}$ are simulated with observations drawn from a uniform distribution $[-0.5, 0.5]$, then do SVD of $X_{100\times 8}$ and obtain the matrices U, V, and D, containing, respectively, the left and right singular vectors and the singular values of $X_{100\times 8}$.
Simulate the main effects mean, row and column:

$$\mu \sim \mathcal{N}(15, 3), \quad \alpha \sim \mathcal{N}(5, 1), \quad \beta \sim \mathcal{N}(8, 2)$$

Following an AMMI2 structure, simulate a matrix:

$$Y = \mathbf{1}_n \mathbf{1}_p^T \mu + \alpha \mathbf{1}_p^T + \mathbf{1}_n \beta^T + 28 \cdot U_{[:,1]} D_{1,1} V_{[:,1]}^T + 15 \cdot U_{[:,2]} D_{2,2} V_{[:,2]}^T$$

The percentages of missing values will be 10%, 20% and 30% for all 3 missing data mechanisms.
For **MCAR** elements are randomly drawn from a uniform distribution (UNIF) in the interval $[0, 1]$, using the `runif` function in R. Then, the q-th quantile of the matrix elements is calculated, where $q = 10\%, 20\%$, or 30%. The positions in Y corresponding to elements in UNIF smaller than the q-th quantile are deleted and considered as missing.
This process involves simulating the main effects (μ, α, β) from normal distributions and combining them with SVD components from a randomly initialized matrix. In the case of **MAR** the columns of matrix Y are divided into pairs (A_x, A_y), and the percentage of missing data in each column is adjusted so that the overall proportion of missing observations remains constant. If one column in the pair has no missing values, the other must have $2q$ missing observations to achieve a mean of q across the entire dataset.
The $2q/100$ quantile of values in A_x is estimated, and based on a defined condition, the corresponding values in A_y are marked as missing. For instance, values in A_y are considered missing when the corresponding values in A_x are greater than or equal to this quantile.
Following this procedure, MAR values are simulated as follows:

- For matrix $X_{100\times 8}$ the $\frac{3q}{2}$-th quantile is calculated for the odd-numbered columns $(j = 1, 3, 5, 7)$, and these values are denoted as q_j^{odd}. If an element at position (i, j) is smaller than q_j^{odd}, the corresponding value at position $(i, j + 1)$ is set as missing.

- Similarly, the $\frac{q}{2}$-th quantile is calculated for the even-numbered columns ($j = 2, 4, 6, 8$), denoted as q_j^{even}. If an element at position (i, j) is smaller than q_j^{even}, the value at position $(i, j-1)$ is considered missing.

MNAR: For each column j, calculate the q_j-th quantile. Elements in each column that are smaller than the corresponding q_j value are marked as missing.

Outliers. Following García-Peña et al.(2021), we generate outliers columnwise (by environment). For each environment j, let $\mu_{j,\text{env}}$ and $\sigma^2_{j,\text{env}}$ denote the mean and variance, respectively, computed from the non-missing entries in column j. We then replace selected cells with draws from two Normal laws that shift the mean by a multiple of the column standard deviation (while preserving the column variance):

$$X \sim \mathcal{N}\big(\mu_{j,\text{env}} + 5\,\sigma_{j,\text{env}},\ \sigma^2_{j,\text{env}}\big) \quad \text{and} \quad X \sim \mathcal{N}\big(\mu_{j,\text{env}} + 100\,\sigma_{j,\text{env}},\ \sigma^2_{j,\text{env}}\big).$$

This design lets us, first, reproduce the moderate-contamination procedure in García-Peña et al. (2021) and, second, probe an extreme-contamination regime in parallel. Evaluating both settings side-by-side reveals how the models behave under usual contamination scenarios and under severe, stress-test conditions.The percentage of shortages is 2%.

4.2 Real Data Study

The selected datasets are those used in Arciniegas-Alarcón et al. (2024).

[htbp]

Reference	Dataset Type	No. of Columns	No. of Rows
Krzanowski (1988)	Oil	5	20
Bro and Smilde (2014)	Red Wine	14	44

Table 1. Multivariate datasets chosen to perform the real-data study

[htbp]

Reference	Species	No. of Genotypes	No. of Env.
Forkman and Piepho (2014)	Soybean	7	10
Forkman and Piepho (2014)	Pea	5	4
Cornelius et al. (1996)	Maize	9	20
Corsten and Denis (1990)	Corn	20	7
Yan et al. (2007)	Wheat	18	9
Lavoranti (2003)	Eucalyptus grandis	20	7
Calinski et al. (2009a)	Pea	18	9
Calinski et al. (2009b)	Winter rye	18	15
Farias (2005)	Cotton	15	27
Filho et al. (2008)	Cotton	17	23
Flores et al. (1998)	Faba beans	15	12
Mattos et al. (2013)	Sugarcane	22	5
Rad et al. (2013)	Wheat	36	6
Yang (2007)	Barley	6	18
Zobel et al. (1988)	Soybean	7	35

Table 2. Summary of G×E datasets

To introduce the missing data, we use the same distribution in the simulation part.

$$X \sim \mathcal{N}\big(\mu_{j,\text{env}} + 5\,\sigma_{j,\text{env}},\ \sigma^2_{j,\text{env}}\big) \quad \text{and} \quad X \sim \mathcal{N}\big(\mu_{j,\text{env}} + 100\,\sigma_{j,\text{env}},\ \sigma^2_{j,\text{env}}\big).$$

4.3 Evaluation method

To evaluate the performance of the MI method, we use the two statistics used in Arciniegas-Alarcón et al. (2020), the prediction error and the variance among the imputations:

$$P_e = \left[\frac{1}{I} \sum_{i=1}^{I} (y_i - \hat{y}_i)^2 \right]^{1/2}$$

$$V = \frac{1}{I} \sum_{i=1}^{I} \left[\frac{1}{M-1} \sum_{m=1}^{M} \left(\hat{y}_{ij(m)} - \bar{Y}_i \right)^2 \right]$$

Where y_i is the real value (the deleted one), $\hat{y}_i$ is the value imputed by the MI method,M is the number of imputations, $\hat{\mathbf{y}}_{\mathbf{ij(m)}}$ is the **m**-th imputation for a missing value, $\bar{Y}_i$ is the mean of the imputations for that observation and **I** is the number of missing entries.

4.4 Proposed MI method

Preparation and Initial Imputation: Start from a matrix $\mathcal{Y}$ with missing values. Replace missing values in each column by the median of observed values, obtaining a complete matrix Y_0.
Application of the Methods (rSVD84 and rSVD01): Apply both algorithms independently on Y_0.
Note: A threshold is also defined that indicates the moment at which matrix deflation stops, i.e., new components are no longer extracted when the improvement in the approximation falls below the threshold.
Each algorithm provides an imputed version of $\mathcal{Y}$:

$$\tilde{Y}_{84}, \quad \tilde{Y}_{01}$$

Combination of Imputations Using Weighted Average:

$$\text{Imputed value} = w \cdot \tilde{Y}_{84} + (1 - w) \cdot \tilde{Y}_{01}, \quad \text{with w} = 0.5,\ 0 \text{ and } 1$$

Generating Multiple Imputed Sets Using Bootstrap (Hu & Yang, 2013):

- Define the number of imputed datasets (e.g. $m = 5$, $m = 10$).
- For each m:
  - Draw row indices $I^{(m)} = (i_1, ..., i_n)$ with replacement.
  - Draw column indices $J^{(m)} = (j_1, ..., j_p)$ with replacement.
  - Create an $n \times p$ matrix $Y^{*(m)}$:

$$Y_{kl}^{*(m)} = Y_{i_k, j_l} \quad \text{for } k = 1, ..., n;\ l = 1, ..., p$$

  - Apply previous steps on $Y^{*(m)}$ to get m complete datasets.
  - This process won't be done with missing positions; it will only be done with positions with observed data.

Statistical Analysis and Combination Using Rubin's Rules:

- Conduct analysis/modeling on each imputed dataset.
- Combine results using Rubin's rules to obtain pooled estimates and standard errors.

5 Results

5.1 Results for simulated data with 5 standard deviation environment

We evaluated six imputation methods (ABB, MICE, missForest,rSVD01-84 MI, rSVD01 MI, and rSVD84 MI) on all simulated datasets. For each dataset, we induced missingness under three mechanisms (MCAR, MAR, and MNAR) and at three levels (p = 10%, 20%, and 30%). In addition, we contaminated 2% of the observations with outliers to assess robustness. Performance is summarized with the Prediction error (P_e); we report the mean, median, variance, and interquartile range across n = 100 replications per condition, with smaller values indicating better accuracy. It was used an M=500.

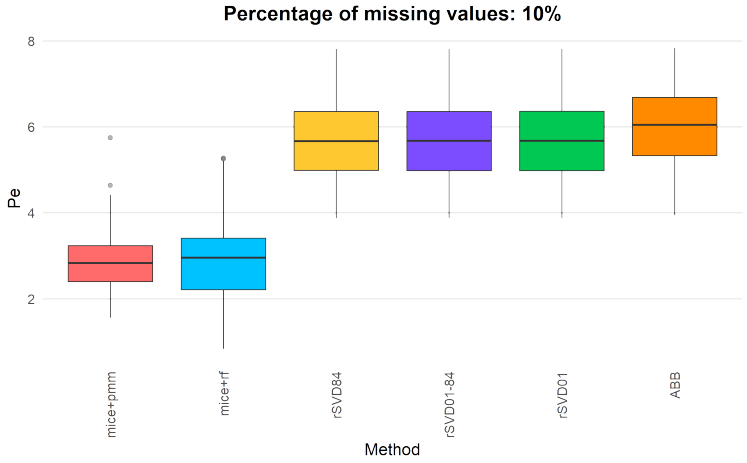


Figure 1. Boxplot of P_e under MAR at $p = 10\%$.

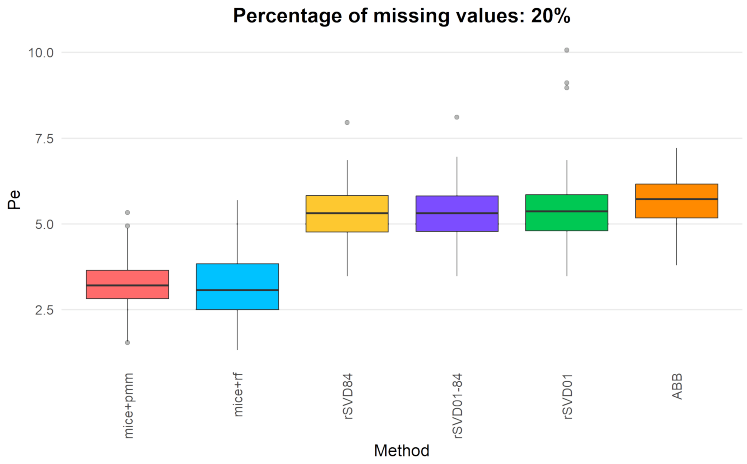


Figure 2. Boxplot of P_e under MAR at $p = 20\%$.

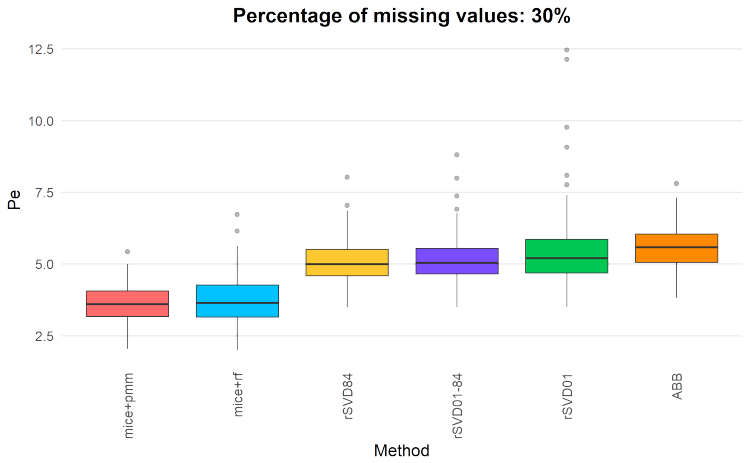


Figure 3. Boxplot of P_e under MAR at $p = 30\%$.

Under the MAR mechanism across missingness levels p, the non–SVD baselines **mice+pmm** and **mice+rf** serve as the gold standards and achieve the lowest prediction error (P_e) with relatively small dispersion; **mice+pmm** is the most stable (lowest variance/IQR) at each p. The robust SVD family performs worse, but shows a clear internal ranking: **rSVD84** consistently outperforms **rSVD01**, and remains competitive with **ABB** (often slightly better). As p increases from 10% to 30%, the gold standards degrade mildly yet preserve a decisive lead, while robust SVD methods narrow the gap only marginally.



Figure 4. Boxplot of P_e under MCAR at $p = 10\%$.

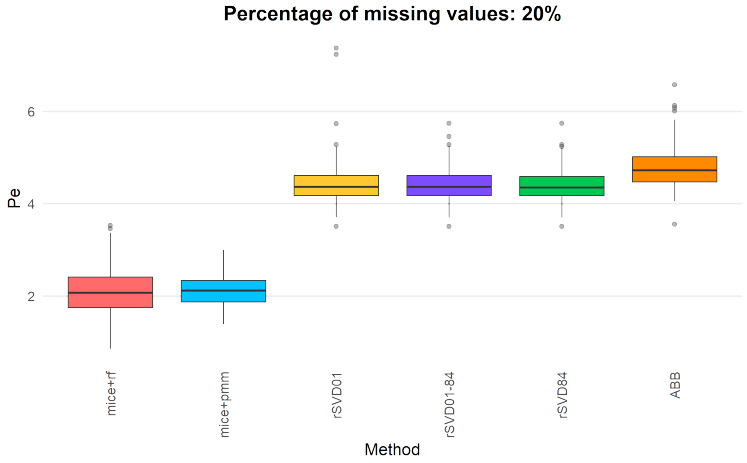


Figure 5. Boxplot of P_e under MCAR at $p = 20\%$.

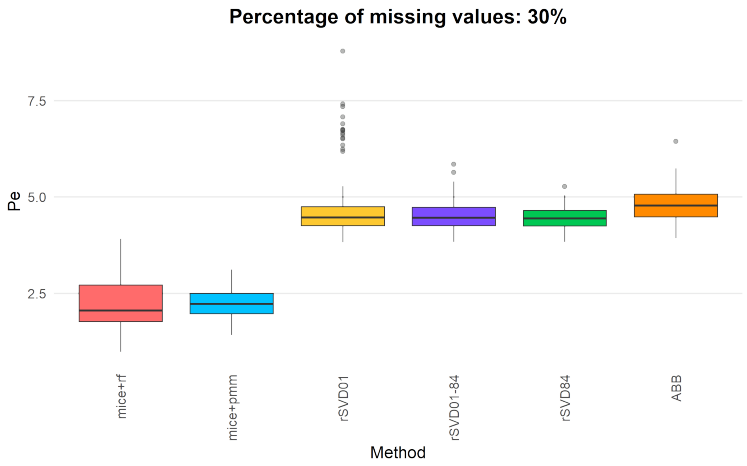


Figure 6. Boxplot of P_e under MCAR at $p = 30\%$.

Under the MCAR mechanism, all models perform significantly better compared to MAR and will see later on the document taht also to (MNAR. For example, in the scenario of missing data 10% (Figure 4), both **mice+pmm** and **mice+rf** achieve the lowest mean (P_e) around 1.9), with **mice+rf** slightly ahead in performance. The SVD-based methods (**rSVD01**, **rSVD01-84**, and **rSVD84**) show a noticeable gap in comparison, with **rSVD01** performing the worst at approximately 4.4. As the percentage of missing values increases to 20% and 30% (Figures 5 and 6), this trend remains consistent, with **MICE** methods outperforming the **SVD** alternatives.

Interestingly, **rSVD01** shows a steady increase in error as the missingness increases (from 4.4 to 4.8), while **MICE** methods maintain their low variance and consistency. The variance and IQR metrics reflect a more stable performance for the **MICE** variants, highlighting their robustness across different missingness percentages.

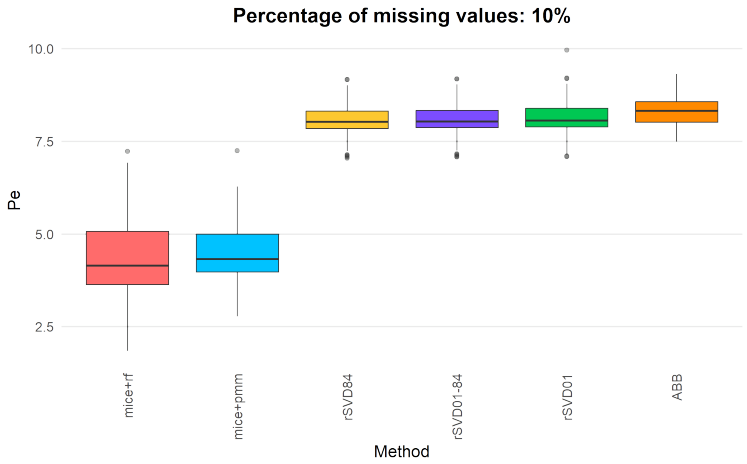


Figure 7. Boxplot of P_e under MNAR at $p = 10\%$.

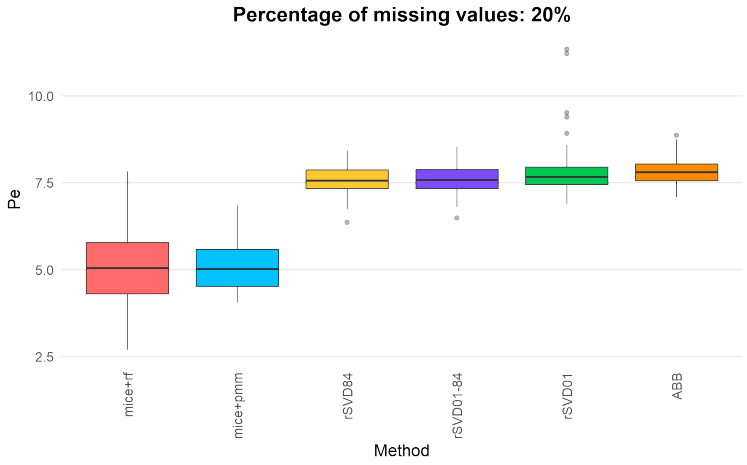


Figure 8. Boxplot of P_e under MNAR at $p = 20\%$.

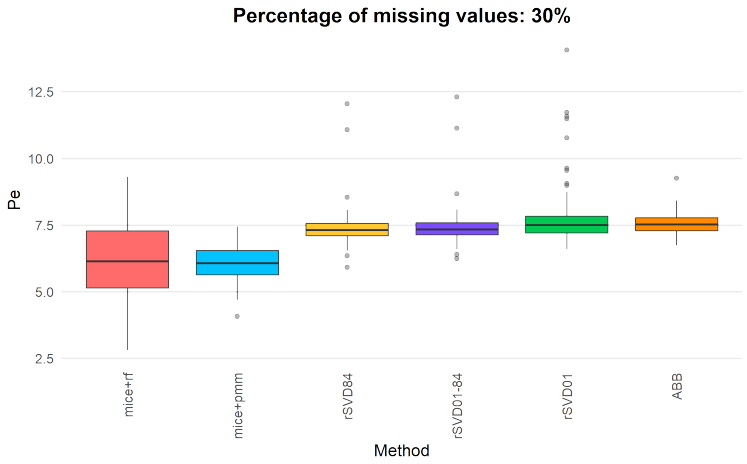


Figure 9. Boxplot of P_e under MNAR at $p = 30\%$.

Under the MNAR mechanism, all models show worse performance compared to MCAR, especially in terms of mean prediction error (P_e). In the 10% missing data scenario (Figure 7), the **MICE** methods (**mice+pmm** and **mice+rf**) achieve relatively low *pe* values (around 4.4–4.7), but the **SVD**-based methods (**rSVD01**, **rSVD01-84**, **rSVD84**) have much higher prediction errors, ranging from 8.0 to 8.3. As the percentage of missing values increases to 20% and 30% (Figures 8 and 9), this performance gap narrows. The **MICE** methods continue to outperform the **SVD**-based models, with the **MICE+rf** variant maintaining the lowest P_e) in the 20% and 30% missing data scenarios. Likewise, mice+rf appears to increase the variance of its performance with increasing missing percentage.

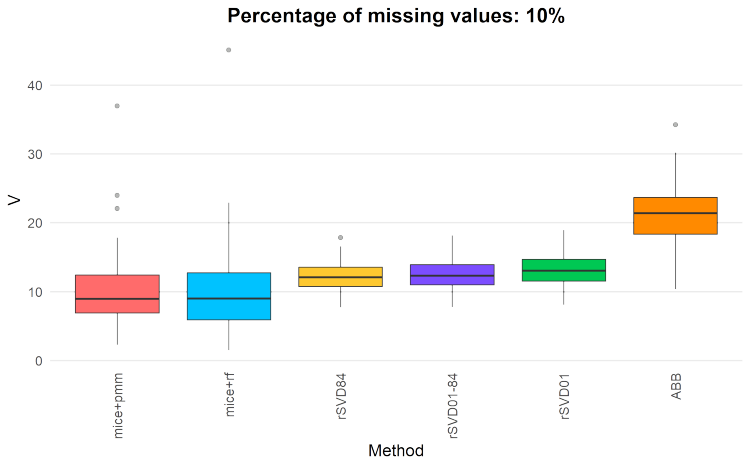


Figure 10. Boxplot of **V** under MAR at $p = 10\%$.

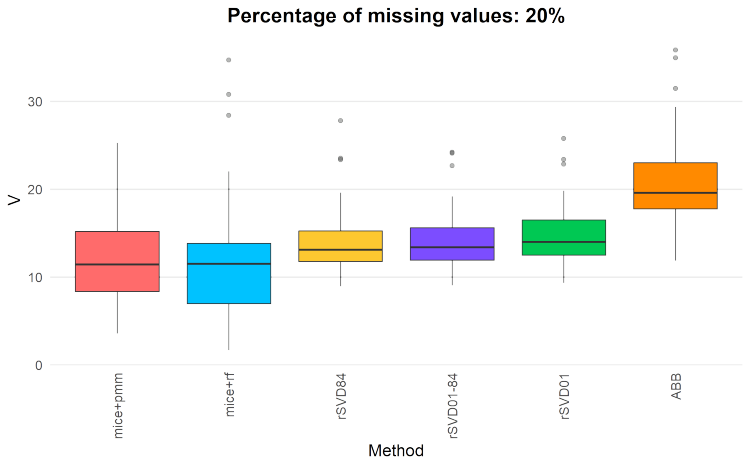


Figure 11. Boxplot of **V** under MAR at $p = 20\%$.

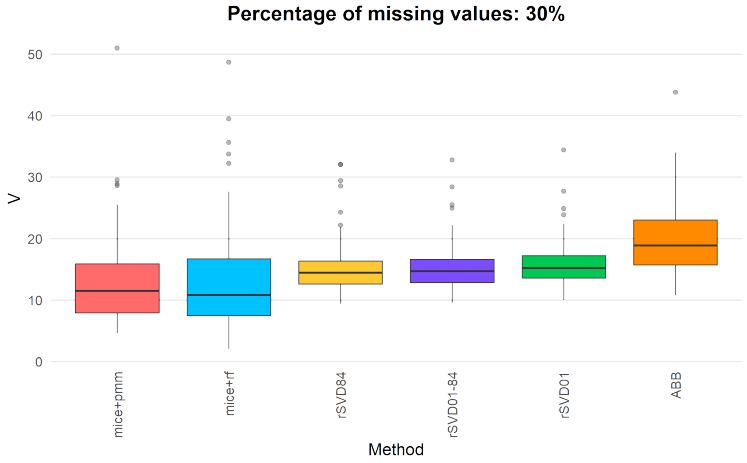


Figure 12. Boxplot of **V** under MAR at $p = 30\%$.

Figure 12 clearly illustrates that as the percentage of missingness increases to 30%, the variability in the variance among imputations (V) becomes more pronounced. This is particularly evident for the **mice+rf** method, which displays a wide distribution and numerous outliers, with some values reaching more than double its median, indicating significant instability. Another important observation is that the **ABB** method consistently yields the highest mean V. Notably, its average value is nearly double that of the **mice**-based models, suggesting that it systematically captures a substantially higher degree of uncertainty related to the imputation process.

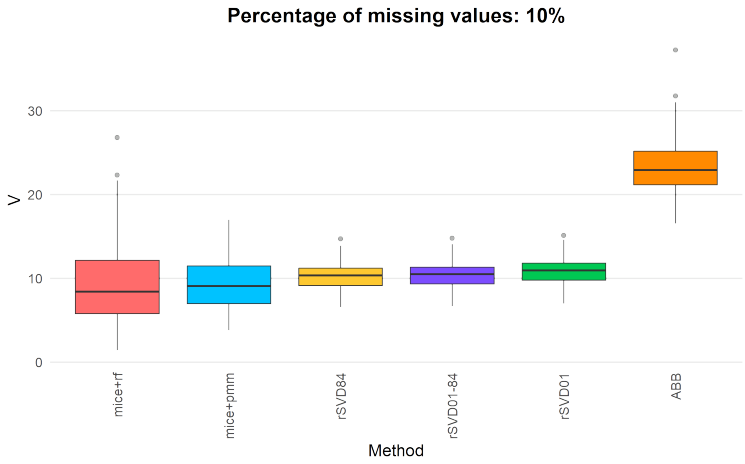


Figure 13. Boxplot of $\mathbf{V}$ under MCAR at $p = 10\%$.

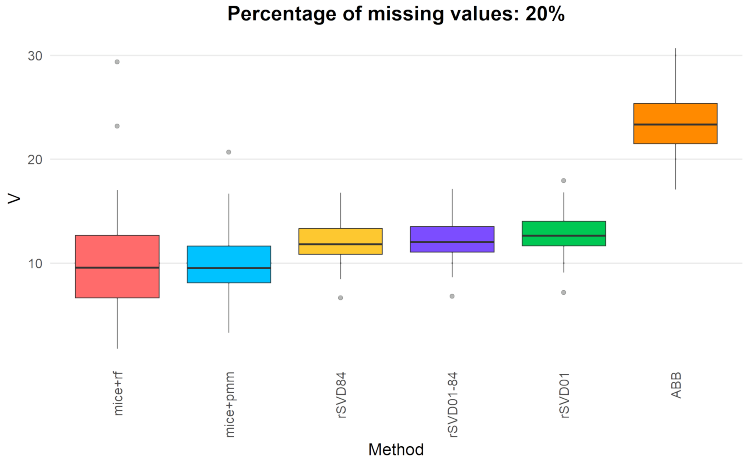


Figure 14. Boxplot of $\mathbf{V}$ under MCAR at $p = 20\%$.

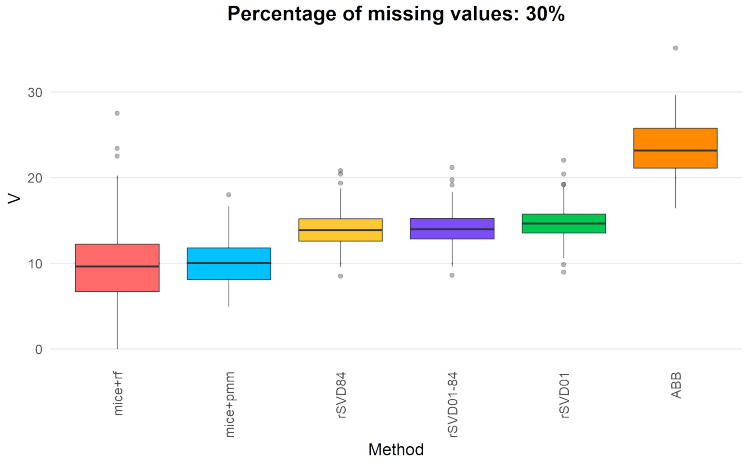


Figure 15. Boxplot of $\mathbf{V}$ under MCAR at $p = 30\%$.

In general, the performance trends hold, but the behavior of the **ABB** method is particularly noteworthy, as it continues to distinguish itself by producing a substantially higher variance among its imputations.

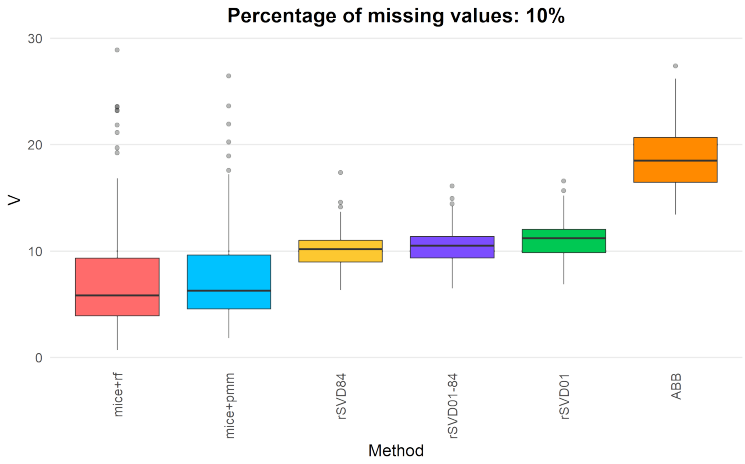


Figure 16. Boxplot of $\mathbf{V}$ under MNAR at $p = 10\%$.

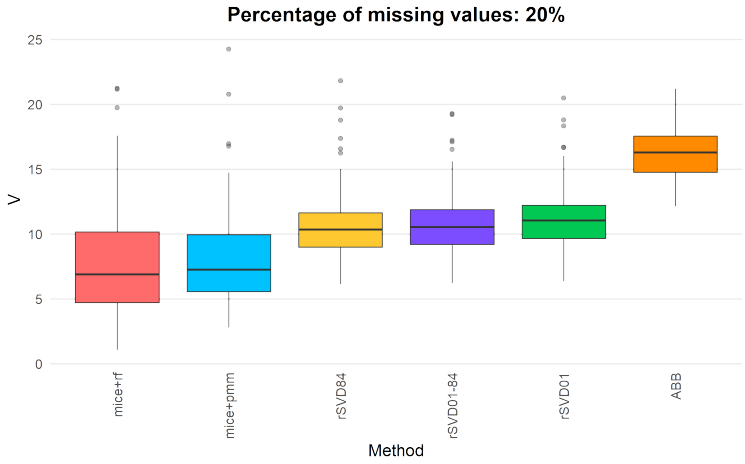


Figure 17. Boxplot of $\mathbf{V}$ under MNAR at $p = 20\%$.

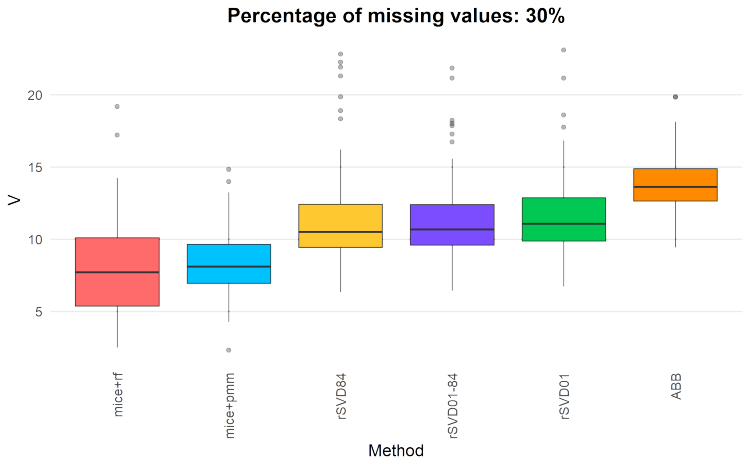


Figure 18. Boxplot of **V** under MNAR at $p = 30\%$.

ABB, in this case, decreases its variability among imputations as p increases, a behavior contrary to what was observed under other missing data mechanisms. Similarly, in MNAR, it appears to be the mechanism with the lowest variance.
During the initial phase of this study, the **mice** models consistently outperformed the others. They not only achieved the lowest P_e values across all mechanisms and loss percentages but also maintained a relatively stable variance between imputations. While the proposed SVD models performed similarly to **mice**, they were not as efficient under the given data scenarios when compared to the two gold standards. Nevertheless, the SVD models showed superior performance compared to **ABB**, without the limitation of relying on a specific data distribution. Additionally, a preliminary evaluation of the **rSVD84** and **rSVD01** models under the MAR mechanism showed promising results, although they still lagged behind the MICE models, one of the main objectives of this study. Further investigation of this behavior will be conducted in the continuation of the study.

5.2 Results for simulated data 100 standard deviation environment

The idea of this second part of the study is to see how the models behave under simulated data in the presence of quite extreme outliers, in this case 100 standard deviation environment.

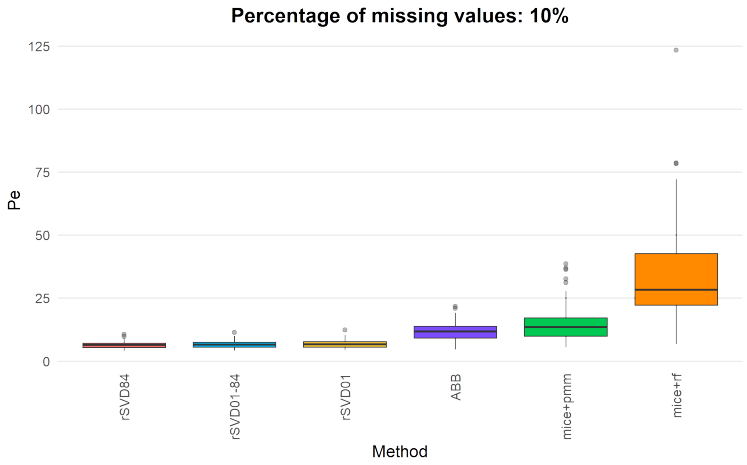


Figure 19. Boxplot of P_e under MAR at $p = 10\%$.

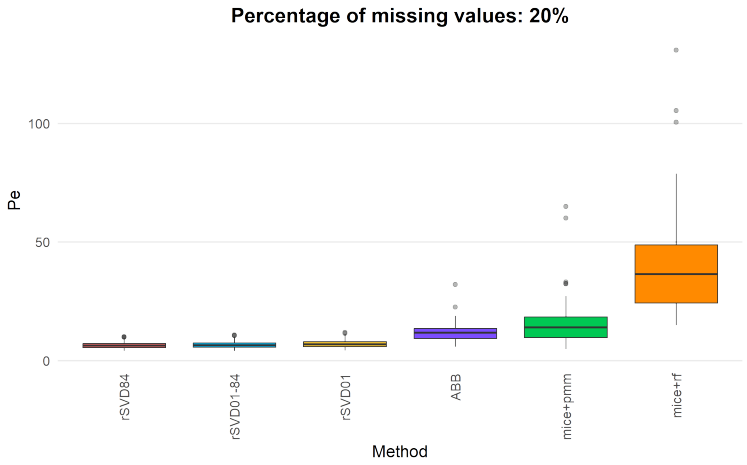


Figure 20. Boxplot of P_e under MAR at $p = 20\%$.

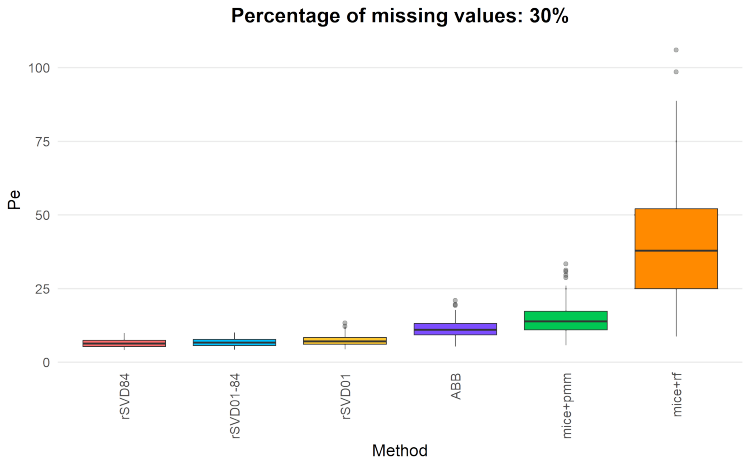


Figure 21. Boxplot of P_e under MAR at $p = 30\%$.

Under MAR, one of the first aspects that can be noticed is that, on the one hand, the SVD models are much less influenced by the magnitude of the outlier than the mice; in fact, here, regardless of the percentage of loss, both the mean and median of the P_e favor the robust models orders of magnitude more than, for example, the **mice+rf** model. Likewise, another notable aspect is the low variance of the P_e data from the SVD models. In fact, under this mechanism and conditions, the performance of the models has remained stable, and here, the one that has been best managed and maintained despite an increase in data loss is **rSVD84**.

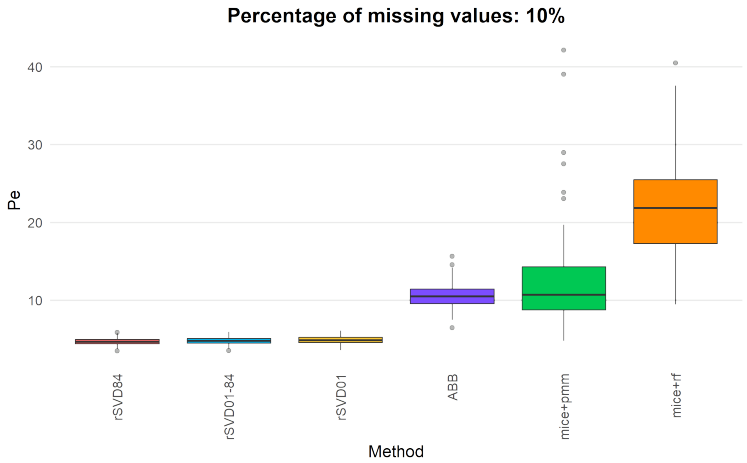


Figure 22. Boxplot of P_e under MCAR at $p = 10\%$.

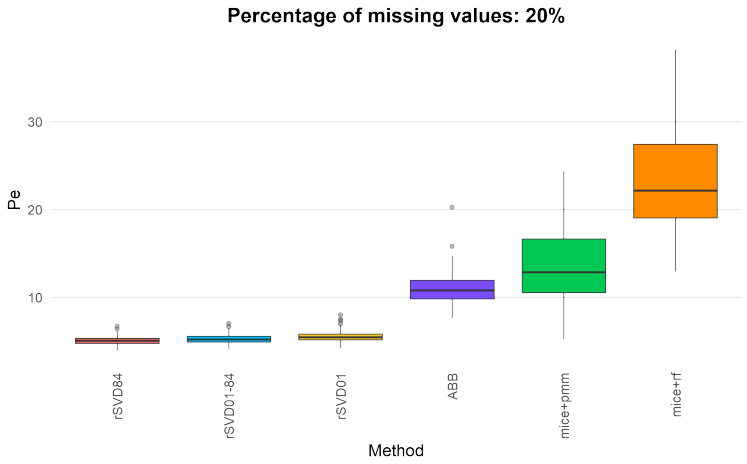


Figure 23. Boxplot of P_e under MCAR at $p = 20\%$.

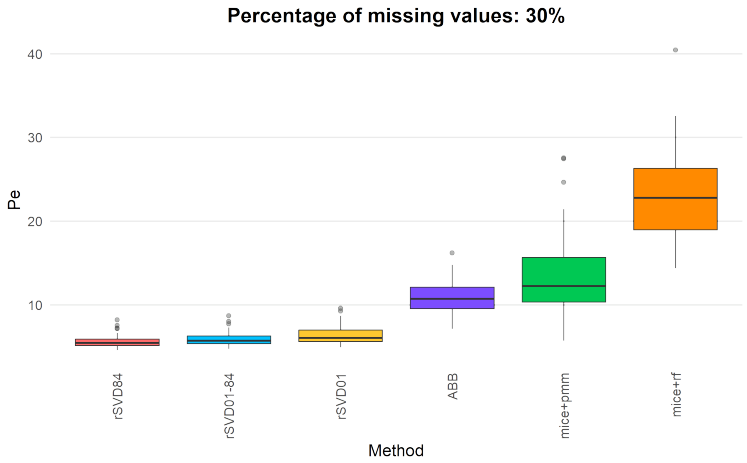


Figure 24. Boxplot of P_e under MCAR at $p = 30\%$.

Under MCAR, the behavior of the previous mechanism is maintained, with mice+rf being the one that is most maintained, even so, it is under this mechanism where the SVD models have had the best P_e score and especially the **rSVD84** model, here a slightly stronger change is noted compared to the increase in the percentage of loss, however the variance of these results remains quite low, which, reviewing previous literature such as García-Peña et al. (2021), seems to be a behavior of these models, in that paper the IQR also remained low in both models and so did the deviation.

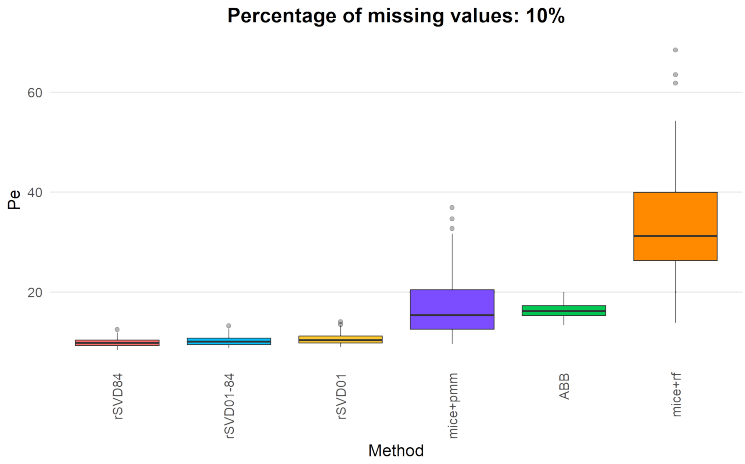


Figure 25. Boxplot of P_e under MNAR at $p = 10\%$.

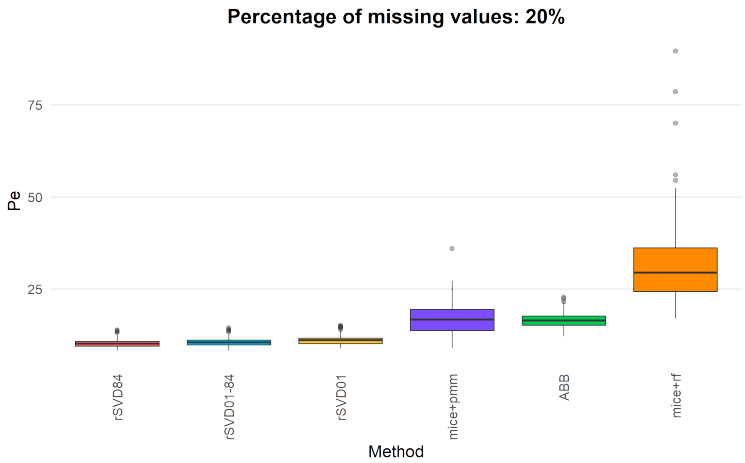


Figure 26. Boxplot of P_e under MNAR at $p = 20\%$.

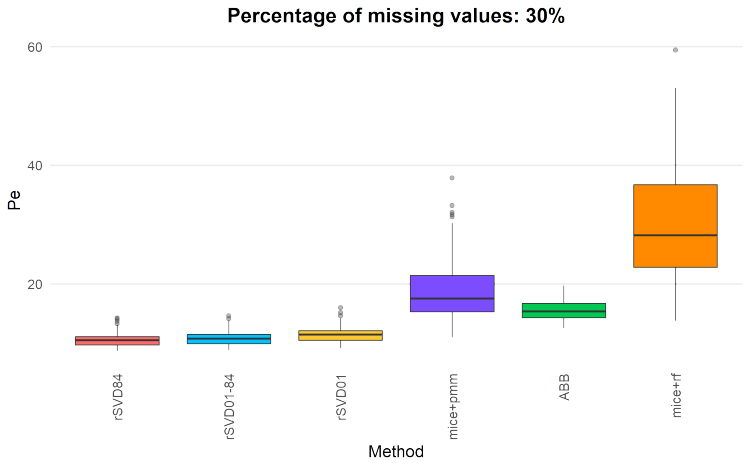


Figure 27. Boxplot of P_e under MNAR at $p = 30\%$.

One of the first aspects to note in this study is that for the proposed MI models, the MNAR mechanism seems to be the one that penalizes the models the most in terms of

precision. In this case, although the three models have a low average P_e compared to the gold standard, in reality, for what We have been presenting in the previous mechanisms, a substantial increase is noted that still remains stable, as shown in Figures 25, 26, and 27, where even though there is an increase in missing values, this does not affect the robustness of the models.

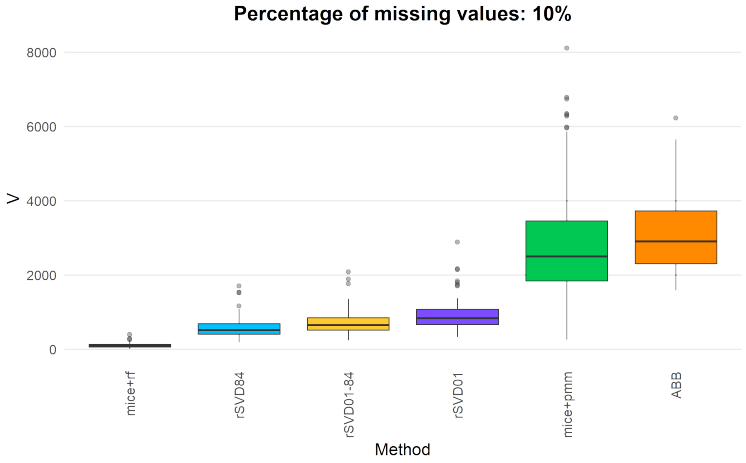


Figure 28. Boxplot of $\mathbf{V}$ under MAR at $p = 10\%$.

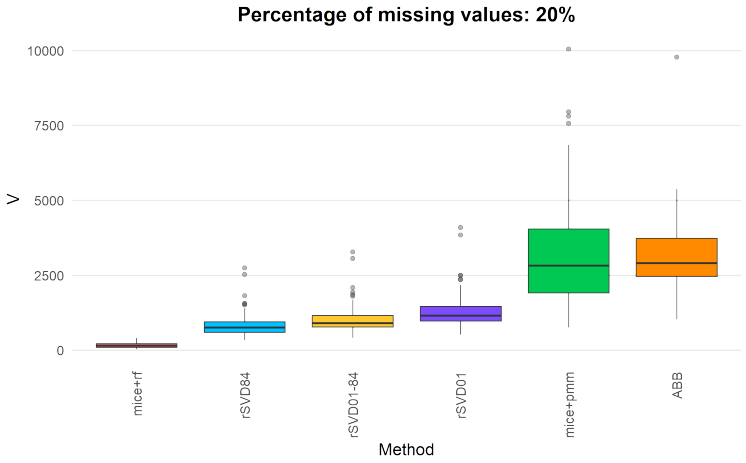


Figure 29. Boxplot of $\mathbf{V}$ under MAR at $p = 20\%$.

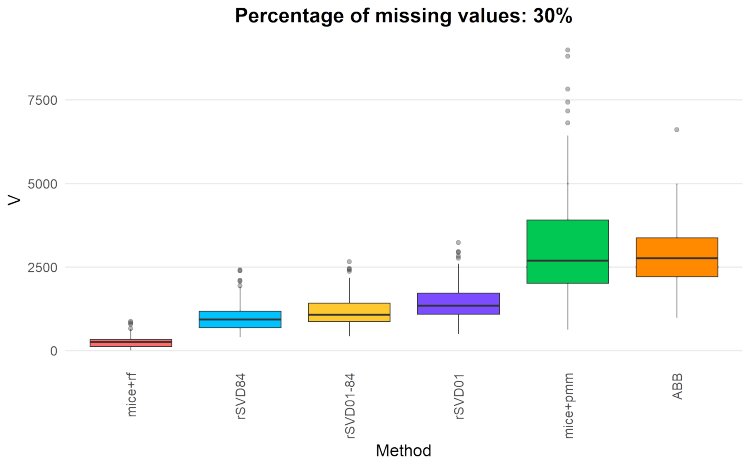


Figure 30. Boxplot of $\mathbf{V}$ under MAR at $p = 30\%$.

When such large magnitudes were assigned to the outliers as in our When such large magnitudes were assigned to the outliers, as specified in our methodology, the variance across imputations increased exponentially reaching values in the order of thousands in some cases and was therefore substantially inflated. This behavior can be partially explained by the fact that variance calculations involve squaring deviations, which amplifies the effect of any imputer that incorporates these extremely large values. As shown in Figures 29 and 30, the scale of the Y axis reaches several thousand units, and even at this elevated scale, some imputers such as **mice+pmm** and **ABB** exhibit particularly high variance across imputations. However, it is noteworthy that none of the imputers achieve variance values close to lower, more typical ranges.

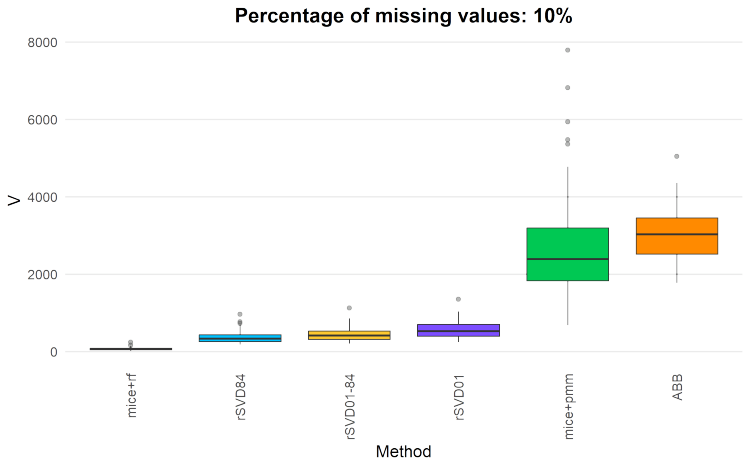


Figure 31. Boxplot of $\mathbf{V}$ under MCAR at $p = 10\%$.



Figure 32. Boxplot of $\mathbf{V}$ under MCAR at $p = 20\%$.

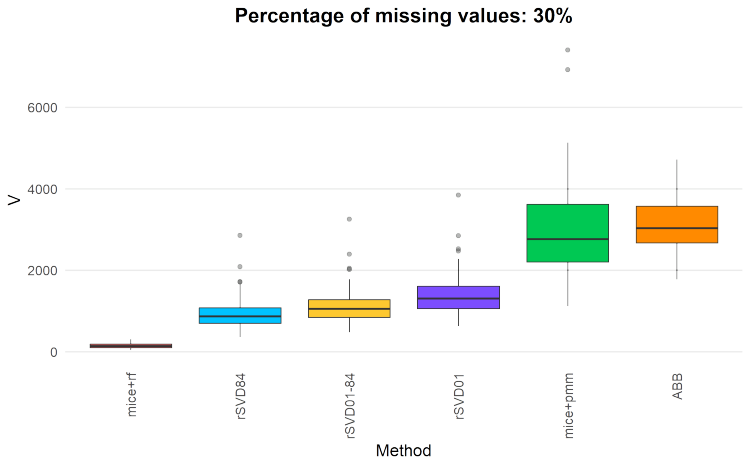


Figure 33. Boxplot of $\mathbf{V}$ under MCAR at $p = 30\%$.

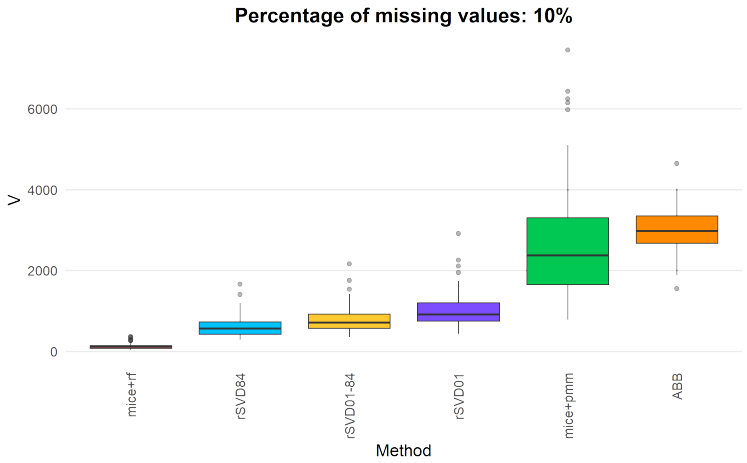


Figure 34. Boxplot of $\mathbf{V}$ under MAR at $p = 10\%$.

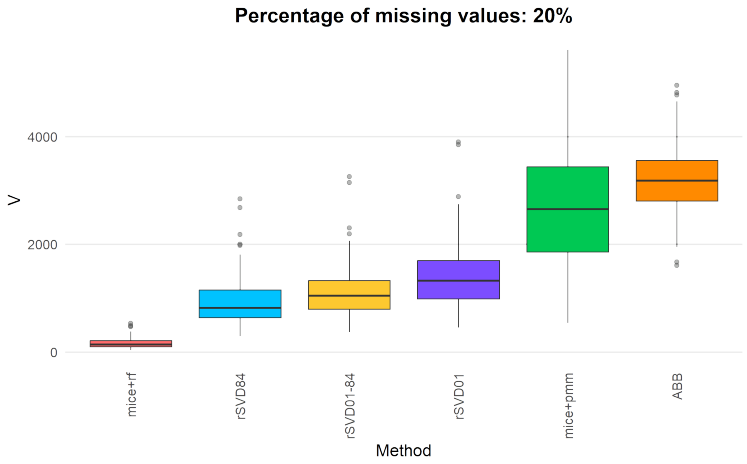


Figure 35. Boxplot of **V** under MAR at $p = 20\%$.

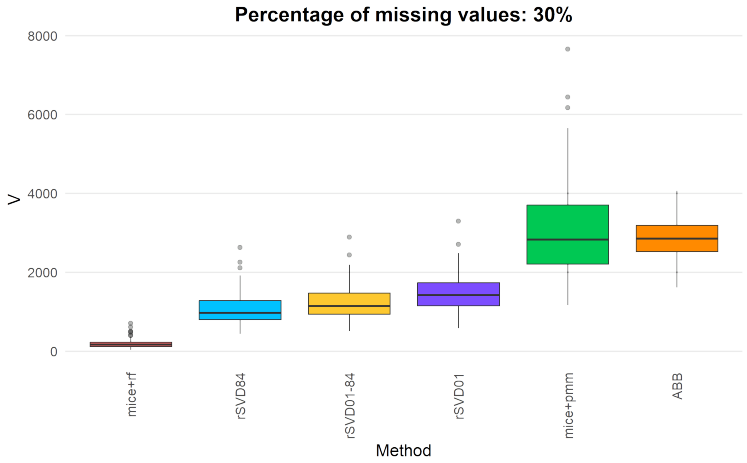


Figure 36. Boxplot of **V** under MAR at $p = 30\%$.

The proposed MI models based on SVD perform well under extreme scenarios where preserving accuracy is critical. However, as observed in the variance analysis, the reliability of the imputations decreases under these conditions due to the exponential growth in variability across imputations. In the normal scenario, these models achieve competitive performance, although they still do not outperform the **mice** models. Among the SVD-based approaches, **rSVD84** showed the best performance overall. An additional noteworthy finding is that, for the models evaluated in this study, the **MNAR** missingness mechanism proved to be the most challenging, yielding lower predictive accuracy and less stable variance across imputations.

5.3 Results for real data with 5 standard deviation environment

For the real datasets mentioned above, 100 data-loss simulations were generated for each percentage–mechanism combination and for each model. Imputation was then performed using the six models, and the same evaluation metrics were applied. This procedure was conducted under both outlier magnitudes (5 and 100).

Given the large volume of results, the analysis must proceed from a macro-level perspective to a more detailed micro-level examination. In order to facilitate the interpretation of the results, the datasets can be classified into two main categories: (1) those that produced consistent and reasonable results, and (2) those that, due to the inherent characteristics of the data, yielded unstable or unrealistic outcomes.

The results are presented in tables 3–15. These tables are organized by missingness mechanism, and each table contains between two and five rows, with each row corresponding to one dataset. The three columns represent the three percentages of missingness. For each percentage, the evaluation metrics either P_e or V are summarized using their mean, median, variance, and interquartile range (IQR).

The datasets Krzanowski (1988) and Bro and Smilde (2014) produced realistic and consistent results. Notably, these are the only multivariate datasets in the study. As shown in table 3, the gold-standard imputers achieve stable and reasonable P_e values across the three missingness percentages. However, the SVD-based models show considerably poorer performance in this context, with Pevalues far from those obtained by the other imputers. A similar pattern is observed for the Bro and Smilde dataset. For example, **rSVD84** achieves a mean P_e of 10.6 for p10, which may appear acceptable in isolation, but remains substantially worse when compared to the performance of **mice** and **ABB**, which achieve mean Pe values of 3.25, 3.06, and 4.26 respectively (see table 3). This behavior persists across all missingness mechanisms. Even under the MNAR mechanism, the performance in terms of P_e remains notably poor (see table 5). A possible explanation for this phenomenon is that the randomized factor proposed for the model involved a different bootstrap approach, which is more suitable for GXE type data. Some of this effect is also reflected in the variance results (tables 41, 42, and 43 on appendices), where the variance of the SVD-based models consistently exceeds 100.

If we focus on the datasets that pose challenges for all models, one notable example is the Forkman and Piepho (2014) dataset, which has dimensions 5×4. This is a very small dataset, particularly considering that the proposed SVD models rely on iterative operations that repeatedly move and swap data positions to introduce randomness. With such limited data, the bootstrap-based procedure used to generate multiple imputed datasets cannot fully exploit its randomization mechanism, leading to unstable imputations and, in some cases, effectively deterministic results, as observed in table 5.

In addition to this dataset, other datasets such as the second Forkman and Piepho (2014) dataset, Cornelius et al. (1996), Farias (2005), Flores et al. (1998), and Zobel et al. (1988)also produce poor results across all imputers, including the gold-standard methods. This is primarily due to the inherent characteristics of the data. For instance, in the Zobel et al. (1988) dataset, most imputers yield a P_e greater than 1000. A descriptive analysis reveals that these datasets exhibit extremely high variance across environments and genotypes, so even imputations based on the median, mean, or other methods result in very high predictive error.

When analyzing the other datasets, it is instructive to compare the performance of the SVD-based imputers with the gold-standard methods. The results on real data confirm the behavior observed in the simulations: the SVD imputers often approach the performance

of the gold-standard methods but, at least in terms of P_e, they never achieve the best score.

For example, in the Yan et al. (2007) dataset (Tables 6, 7, and 8), all three SVD models perform exceptionally well, obtaining P_e values below 1 across all three missingness percentages under both MAR and MCAR mechanisms. As expected from previous simulations, the models perform worse under MNAR conditions. Interestingly, in this dataset and in contrast with the Lavoranti (2003) dataset the SVD models achieve very good P_e scores while maintaining low variance among imputations. For instance, in Table 8, **rSVD01** achieves a mean P_e of 1.30 for p10, while the corresponding variance among imputations under MNAR is only 0.66. In this case, the best P_e values coincide with the lowest variance.

However, when analyzing other datasets, a different pattern emerges. Among the SVD models, the one achieving the best P_e often exhibits the highest variance. Moreover, in real datasets, the **ABB** imputer frequently outperforms the SVD models and, in some cases, even surpasses the **mice** models. This suggests that the behavior observed in simulated data does not always fully translate to real world datasets.

Table 3. P_e by percentage and method — Mechanism **MAR** (REP 1–5)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	8.05	8.21	0.68	0.65	8.29	8.28	0.17	0.56	8.78	8.71	0.16	0.43
	MICE+PMM	4.74	4.33	4.17	2.65	5.75	5.18	3.99	3.41	7.62	7.26	1.14	1.96
	MICE+RF	5.62	5.43	1.60	1.62	4.26	4.32	0.63	1.17	7.22	7.09	2.09	1.76
	rSVD01	33.81	33.88	0.13	0.49	29.14	28.80	0.82	1.35	28.06	28.00	0.24	0.65
	rSVD84	33.42	33.48	0.04	0.23	28.54	28.53	0.03	0.23	28.62	28.64	0.03	0.20
	rSVD84-01	33.93	33.98	0.50	0.83	29.11	29.16	0.32	0.75	28.13	28.15	0.30	0.76
Bro and Smilde (2014) Red Wine 14×44	ABB	3.25	3.20	0.05	0.28	2.65	2.61	0.03	0.25	2.52	2.50	0.03	0.26
	MICE+PMM	3.06	2.93	0.34	0.66	2.62	2.58	0.18	0.57	2.66	2.60	0.15	0.42
	MICE+RF	4.26	4.24	0.72	1.31	2.11	2.10	0.14	0.48	2.58	2.57	0.29	0.64
	rSVD01	20.91	19.40	36.37	7.23	28.27	23.81	130.73	10.10	33.48	26.33	1.29e+03	7.12
	rSVD84	10.60	10.60	0.00	0.02	10.74	10.74	0.00	0.02	10.80	10.79	0.01	0.08
	rSVD84-01	14.33	13.47	10.16	3.51	17.22	15.34	88.46	5.17	25.22	19.76	575.34	6.37
Forkman and Piepho (2014) Soybean 7×10	ABB	800.18	784.97	2.23e+03	54.34	789.87	776.57	2.48e+03	70.45	745.43	732.22	2.94e+03	69.84
	MICE+PMM	736.20	733.05	3.06e+03	73.37	739.63	721.27	2.77e+03	63.80	734.11	715.05	4.83e+03	72.12
	MICE+RF	710.90	705.68	2.16e+04	218.70	890.82	884.06	3.14e+04	224.68	776.18	769.94	2.27e+04	206.42
	rSVD01	1.03e+03	1.03e+03	554.82	35.41	1.06e+03	1.06e+03	899.62	46.64	1.03e+03	1.03e+03	289.36	22.91
	rSVD84	1.10e+03	1.09e+03	146.01	14.12	1.05e+03	1.05e+03	256.47	16.77	1.06e+03	1.06e+03	129.86	14.41
	rSVD84-01	1.06e+03	1.06e+03	1.56e+03	54.69	1.01e+03	1.01e+03	1.74e+03	56.62	1.02e+03	1.02e+03	884.37	41.27
Forkman and Piepho (2014) Pea 5×4	ABB	1.13e+03	1.15e+03	2.00e+04	210.40	1.14e+03	1.14e+03	2.24e+04	208.28	1.07e+03	1.04e+03	3.32e+04	315.25
	MICE+PMM	1.24e+03	1.25e+03	4.27e+04	300.16	1.24e+03	1.24e+03	4.10e+04	278.30	1.28e+03	1.16e+03	6.61e+04	320.36
	MICE+RF	1.35e+03	1.32e+03	8.60e+04	414.74	1.24e+03	1.23e+03	6.46e+04	314.97	1.61e+03	1.58e+03	1.19e+05	488.21
	rSVD01	3.16e+03	3.16e+03	0.00	0.00	3.16e+03	3.16e+03	0.00	0.00	2.58e+03	2.58e+03	4.85	1.52
	rSVD84	513.84	471.65	1.80e+04	146.76	504.30	489.56	7.19e+03	110.72	622.71	591.04	1.36e+04	127.81
	rSVD84-01	1.31e+03	1.31e+03	1.20e+04	150.42	1.33e+03	1.34e+03	1.25e+04	142.94	1.10e+03	1.10e+03	9.87e+03	156.07
Cornelius et al. (1996) Maize 9×20	ABB	2.49e+03	2.49e+03	7.42e+03	98.41	2.25e+03	2.24e+03	7.55e+03	81.39	2.06e+03	2.03e+03	7.86e+03	107.32
	MICE+PMM	1.95e+03	1.92e+03	8.40e+04	506.02	1.97e+03	1.97e+03	3.08e+04	238.30	1.88e+03	1.88e+03	2.73e+04	197.74
	MICE+RF	1.97e+03	1.98e+03	1.21e+05	449.89	2.42e+03	2.41e+03	1.70e+05	591.03	1.74e+03	1.69e+03	1.18e+05	436.89
	rSVD01	2.20e+03	2.19e+03	6.43e+03	97.94	1.84e+03	1.83e+03	845.75	34.75	1.71e+03	1.72e+03	3.53e+03	77.22
	rSVD84	2.12e+03	2.10e+03	8.20e+03	62.17	1.71e+03	1.70e+03	2.88e+03	59.31	1.46e+03	1.44e+03	4.99e+03	70.38
	rSVD84-01	2.17e+03	2.16e+03	1.57e+04	169.39	1.24e+03	1.24e+03	4.84e+03	99.73	1.26e+03	1.26e+03	7.09e+03	115.28

Table 4. P_e by percentage and method — Mechanism **MCAR** (REP 1–5)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	7.07	7.03	4.29	2.64	7.10	7.13	1.46	1.59	7.08	6.99	1.50	1.21
	MICE+PMM	4.71	4.56	3.41	2.16	5.13	5.06	2.34	2.25	5.72	5.62	2.39	1.99
	MICE+RF	4.20	4.23	0.79	1.35	7.18	7.32	2.21	2.21	6.50	6.44	1.62	1.80
	rSVD01	22.77	23.70	54.69	10.58	23.44	23.57	17.46	5.24	23.24	23.28	9.53	4.26
	rSVD84	22.61	23.26	51.71	9.86	23.26	23.22	17.76	5.87	22.88	22.56	8.29	4.33
	rSVD84-01	22.73	23.47	53.94	10.61	23.27	23.29	16.54	5.77	23.05	23.11	9.27	3.93
Bro and Smilde (2014) Red Wine 14×44	ABB	37.37	39.07	184.46	16.47	37.88	38.56	67.93	10.77	37.84	38.50	50.33	9.85
	MICE+PMM	38.38	37.85	241.67	21.22	37.95	37.07	74.00	12.47	38.19	38.40	57.75	9.97
	MICE+RF	29.64	30.29	47.06	10.46	20.84	20.84	18.02	5.97	23.30	23.36	30.01	6.81
	rSVD01	87.99	89.79	700.61	32.36	90.66	92.17	335.69	25.19	90.39	90.08	191.65	17.10
	rSVD84	87.79	89.95	754.73	34.43	88.83	90.38	343.67	25.85	89.16	88.98	209.02	15.38
	rSVD84-01	87.72	90.37	690.10	33.71	88.73	89.49	325.32	26.68	89.26	90.75	194.97	14.91
Forkman and Piepho (2014) Soybean 7×10	ABB	454.89	450.60	2.68e+04	253.51	526.51	500.61	3.20e+04	123.69	509.28	512.87	1.55e+04	117.24
	MICE+PMM	410.41	377.37	2.59e+04	225.03	444.97	438.62	2.11e+04	187.67	459.52	436.58	1.95e+04	178.16
	MICE+RF	753.67	764.62	2.53e+04	257.09	512.74	512.27	1.34e+04	167.07	779.51	773.38	2.00e+04	201.14
	rSVD01	867.94	879.36	6.12e+04	367.02	934.34	917.62	4.36e+04	205.29	915.87	908.22	2.61e+04	194.44
	rSVD84	863.78	855.68	6.23e+04	356.57	904.58	903.13	2.24e+04	201.33	910.42	898.38	1.71e+04	185.78
	rSVD84-01	861.33	873.96	5.95e+04	349.18	914.68	918.18	2.14e+04	181.40	907.89	901.14	1.85e+04	190.67
Forkman and Piepho (2014) Pea 5×4	ABB	729.06	663.28	1.49e+05	322.95	817.95	692.40	2.54e+05	276.63	1.18e+03	876.73	4.41e+05	1.18e+03
	MICE+PMM	884.25	763.60	2.16e+05	338.61	693.21	637.66	1.57e+05	446.83	817.97	664.98	3.21e+05	446.36
	MICE+RF	610.33	522.14	1.76e+05	419.87	672.52	631.97	8.37e+04	273.58	58.69	58.33	135.36	13.94
	rSVD01	476.05	427.87	6.41e+04	406.52	671.88	500.18	4.97e+05	259.16	1.41e+03	629.12	1.43e+06	2.32e+03
	rSVD84	470.42	454.16	5.28e+04	345.73	502.40	496.58	3.10e+04	243.16	553.48	508.54	1.04e+05	221.80
	rSVD84-01	478.86	471.38	6.03e+04	377.97	579.24	517.26	1.02e+05	301.18	900.55	635.13	3.41e+05	857.29
Cornelius et al. (1996) Maize 9×20	ABB	1.42e+03	1.42e+03	4.50e+04	251.85	1.44e+03	1.46e+03	2.92e+04	208.41	1.45e+03	1.45e+03	1.15e+04	119.34
	MICE+PMM	1.03e+03	1.01e+03	5.54e+04	316.78	1.07e+03	1.06e+03	2.66e+04	244.06	1.14e+03	1.14e+03	2.35e+04	162.63
	MICE+RF	773.28	762.37	1.86e+04	167.22	1.35e+03	1.36e+03	8.20e+04	355.63	1.15e+03	1.16e+03	5.59e+04	275.82
	rSVD01	1.31e+03	1.30e+03	2.75e+04	224.01	1.35e+03	1.36e+03	1.72e+04	158.55	1.37e+03	1.35e+03	2.64e+04	119.50
	rSVD84	1.34e+03	1.34e+03	2.86e+04	204.21	1.43e+03	1.45e+03	2.08e+04	165.40	1.36e+03	1.36e+03	7.44e+03	115.54
	rSVD84-01	1.38e+03	1.38e+03	3.33e+04	242.92	1.37e+03	1.38e+03	1.88e+04	157.44	1.39e+03	1.39e+03	1.35e+04	149.66

Table 5. P_e by percentage and method — Mechanism **MNAR** (REP 1–5)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	12.18	12.00	0.58	0.95	12.20	12.12	0.46	0.89	12.48	12.42	0.42	0.99
	MICE+PMM	8.12	7.47	4.14	2.43	10.19	9.78	3.35	2.61	12.67	12.63	0.80	1.20
	MICE+RF	6.89	7.01	2.28	2.23	10.10	10.03	3.55	2.31	12.66	12.48	5.46	3.35
	rSVD01	18.14	18.22	0.41	0.94	18.11	18.04	0.60	1.22	20.21	20.21	0.27	0.55
	rSVD84	17.20	17.20	0.03	0.21	17.81	17.78	0.03	0.20	20.21	20.19	0.05	0.24
	rSVD84-01	17.37	17.28	0.92	1.25	18.34	18.36	0.44	0.81	20.16	20.12	0.52	0.95
Bro and Smilde (2014) Red Wine 14×44	ABB	68.31	67.26	33.60	6.98	65.02	63.51	20.12	5.56	64.91	63.02	25.29	6.20
	MICE+PMM	62.12	61.07	122.44	13.61	62.28	60.31	86.40	8.72	65.10	62.93	70.23	9.82
	MICE+RF	64.27	65.11	211.02	17.62	78.81	77.92	262.94	19.59	65.69	64.02	131.81	16.87
	rSVD01	48.49	41.13	221.42	21.59	45.44	44.07	59.43	2.75	58.98	52.10	1.11e+03	6.36
	rSVD84	34.78	34.79	0.00	0.02	43.05	43.05	0.00	0.02	47.84	47.83	0.00	0.02
	rSVD84-01	35.67	34.85	12.76	1.46	44.05	42.98	16.89	2.23	52.96	48.29	609.09	2.83
Forkman and Piepho (2014) Soybean 7×10	ABB	832.09	827.37	1.94e+03	47.75	801.58	791.86	2.08e+03	66.19	791.10	779.17	1.91e+03	51.98
	MICE+PMM	784.59	776.09	4.10e+03	74.80	756.32	747.44	1.68e+03	55.60	752.72	740.26	1.82e+03	42.40
	MICE+RF	772.33	773.38	2.49e+04	189.15	697.07	686.90	1.99e+04	222.31	744.83	754.91	2.09e+04	197.58
	rSVD01	1.11e+03	1.11e+03	481.85	29.44	1.04e+03	1.04e+03	318.21	23.76	1.03e+03	1.03e+03	514.04	32.89
	rSVD84	1.11e+03	1.10e+03	173.80	14.05	1.03e+03	1.02e+03	286.54	19.87	984.89	979.24	287.39	18.38
	rSVD84-01	1.12e+03	1.11e+03	1.38e+03	53.99	1.03e+03	1.03e+03	882.71	39.62	1.03e+03	1.03e+03	732.50	34.91
Forkman and Piepho (2014) Pea 5×4	ABB	1.18e+03	1.17e+03	8.20e+03	140.39	1.16e+03	1.14e+03	8.61e+03	117.86	1.18e+03	1.16e+03	9.07e+03	108.22
	MICE+PMM	1.25e+03	1.23e+03	2.00e+04	190.38	1.23e+03	1.21e+03	1.46e+04	141.39	1.24e+03	1.21e+03	1.56e+04	140.18
	MICE+RF	1.13e+03	1.14e+03	5.60e+04	324.59	1.14e+03	1.15e+03	5.65e+04	319.42	1.41e+03	1.41e+03	5.83e+04	310.86
	rSVD01	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00
	rSVD84	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00
	rSVD84-01	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00
Cornelius et al. (1996) Maize 9×20	ABB	2.65e+03	2.64e+03	4.86e+03	98.41	2.61e+03	2.62e+03	3.20e+03	81.10	2.55e+03	2.55e+03	2.81e+03	62.89
	MICE+PMM	2.30e+03	2.28e+03	4.38e+04	194.32	2.44e+03	2.42e+03	7.60e+03	102.41	2.45e+03	2.42e+03	1.37e+04	165.33
	MICE+RF	2.24e+03	2.24e+03	2.39e+05	687.05	2.39e+03	2.33e+03	2.04e+05	559.63	2.26e+03	2.27e+03	2.22e+05	643.86
	rSVD01	2.46e+03	2.46e+03	2.45e+03	69.51	2.71e+03	2.71e+03	302.10	18.65	2.70e+03	2.69e+03	1.08e+03	17.29
	rSVD84	2.39e+03	2.38e+03	5.37e+03	78.58	2.52e+03	2.52e+03	1.01e+03	36.80	2.43e+03	2.42e+03	1.15e+03	22.15
	rSVD84-01	1.72e+03	1.72e+03	6.85e+03	97.86	2.29e+03	2.29e+03	960.13	44.12	2.36e+03	2.36e+03	728.72	28.48

Table 6. P_e by percentage and method — Mechanism **MAR** (REP 6–10)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	9.67	9.57	0.25	0.57	7.72	7.72	0.07	0.31	6.94	6.88	0.08	0.30
	MICE+PMM	9.25	9.01	0.86	1.08	7.70	7.64	0.23	0.51	6.85	6.73	0.18	0.49
	MICE+RF	8.43	8.40	2.52	2.14	8.70	8.72	2.77	2.27	6.20	6.17	1.63	1.56
	rSVD01	8.94	8.97	0.23	0.87	9.41	9.38	0.03	0.28	25.51	25.50	0.00	0.11
	rSVD84	8.59	8.57	0.02	0.15	9.51	9.50	0.01	0.12	8.88	8.86	0.02	0.10
	rSVD84-01	9.00	8.96	0.17	0.50	9.31	9.31	0.06	0.35	13.14	13.15	0.11	0.41
Yan et al. (2007) Wheat 18×9	ABB	0.89	0.89	0.00	0.05	0.80	0.80	0.00	0.03	0.70	0.70	0.00	0.02
	MICE+PMM	0.79	0.78	0.01	0.09	0.71	0.71	0.00	0.09	0.62	0.62	0.00	0.06
	MICE+RF	0.65	0.64	0.02	0.17	0.68	0.67	0.02	0.17	0.69	0.69	0.02	0.18
	rSVD01	0.92	0.91	0.00	0.09	0.96	0.96	0.00	0.04	0.96	0.96	0.00	0.03
	rSVD84	0.91	0.91	0.00	0.02	0.94	0.94	0.00	0.02	0.94	0.94	0.00	0.01
	rSVD84-01	0.95	0.94	0.00	0.06	0.95	0.95	0.00	0.05	0.94	0.94	0.00	0.04
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	2.21	2.22	0.00	0.07	1.97	1.97	0.00	0.06	1.89	1.88	0.00	0.07
	MICE+PMM	1.87	1.87	0.02	0.19	1.64	1.64	0.01	0.11	1.65	1.63	0.01	0.15
	MICE+RF	1.95	1.98	0.17	0.52	1.72	1.72	0.11	0.44	1.57	1.59	0.09	0.41
	rSVD01	3.11	3.11	0.00	0.05	5.34	5.33	0.00	0.06	6.06	6.06	0.00	0.02
	rSVD84	3.08	3.08	0.00	0.02	3.47	3.47	0.00	0.02	3.80	3.80	0.00	0.03
	rSVD84-01	3.13	3.13	0.00	0.07	3.87	3.86	0.00	0.06	4.38	4.38	0.00	0.07
Calinski et al. (2009a) Pea 18×9	ABB	4.57	4.56	0.04	0.29	4.05	4.01	0.04	0.24	3.93	3.90	0.02	0.25
	MICE+PMM	4.45	4.39	0.19	0.64	3.97	3.90	0.17	0.48	3.91	3.85	0.09	0.35
	MICE+RF	4.05	4.09	0.67	1.15	4.72	4.72	0.88	1.11	4.41	4.42	0.76	1.01
	rSVD01	9.03	9.06	0.04	0.24	8.08	8.05	0.08	0.53	8.18	8.22	0.02	0.23
	rSVD84	9.20	9.20	0.01	0.11	8.22	8.23	0.01	0.10	8.61	8.62	0.00	0.07
	rSVD84-01	9.01	9.02	0.06	0.29	8.19	8.20	0.05	0.29	8.40	8.40	0.06	0.28
Calinski et al. (2009b) Winter rye 18×15	ABB	10.13	10.09	0.15	0.50	10.54	10.48	0.12	0.50	10.42	10.44	0.08	0.40
	MICE+PMM	8.86	8.85	0.94	1.04	9.48	9.49	1.07	1.52	9.56	9.44	0.56	0.92
	MICE+RF	6.98	6.97	1.57	1.78	7.98	8.02	2.01	1.88	11.54	11.71	5.59	3.08
	rSVD01	16.23	16.22	0.03	0.26	15.05	15.05	0.03	0.25	14.35	14.35	0.02	0.19
	rSVD84	16.00	16.00	0.00	0.09	15.49	15.47	0.01	0.14	14.66	14.65	0.01	0.14
	rSVD84-01	16.29	16.31	0.08	0.35	15.25	15.24	0.08	0.36	14.35	14.38	0.07	0.37

Table 7. P_e by percentage and method — Mechanism **MCAR** (REP 6–10)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	5.04	4.86	1.72	1.50	5.46	5.49	1.03	1.58	5.34	5.25	0.58	1.08
	MICE+PMM	5.10	4.90	1.77	1.88	5.53	5.35	1.17	1.67	5.45	5.42	0.56	1.16
	MICE+RF	4.09	4.03	0.74	1.08	3.77	3.85	0.74	1.14	4.57	4.67	0.69	1.06
	rSVD01	8.28	8.20	1.60	1.79	8.47	8.46	0.64	1.00	9.38	8.46	17.08	0.97
	rSVD84	8.32	8.28	1.68	1.76	8.48	8.50	0.66	1.09	8.49	8.48	0.45	0.89
	rSVD84-01	8.30	8.31	1.69	1.87	8.47	8.49	0.58	0.97	8.81	8.45	2.81	0.93
Yan et al. (2007) Wheat 18×9	ABB	0.55	0.56	0.01	0.15	0.56	0.55	0.01	0.10	0.56	0.55	0.00	0.09
	MICE+PMM	0.51	0.50	0.01	0.14	0.52	0.51	0.01	0.10	0.51	0.51	0.01	0.10
	MICE+RF	0.63	0.63	0.01	0.14	0.65	0.63	0.02	0.18	0.60	0.60	0.01	0.16
	rSVD01	0.99	0.99	0.02	0.19	0.97	0.98	0.01	0.12	1.01	0.99	0.02	0.10
	rSVD84	0.98	0.97	0.02	0.20	0.98	0.98	0.01	0.14	0.99	0.99	0.01	0.10
	rSVD84-01	0.98	1.00	0.02	0.18	0.98	1.00	0.01	0.13	1.01	1.01	0.01	0.10
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	1.26	1.25	0.02	0.18	1.28	1.28	0.01	0.14	1.26	1.25	0.01	0.10
	MICE+PMM	1.01	1.01	0.02	0.18	1.03	1.02	0.01	0.13	1.04	1.03	0.01	0.12
	MICE+RF	1.19	1.19	0.05	0.29	0.86	0.86	0.03	0.22	1.18	1.15	0.05	0.32
	rSVD01	3.36	3.37	0.07	0.34	3.47	3.45	0.18	0.28	3.66	3.45	0.36	0.22
	rSVD84	3.36	3.37	0.08	0.38	3.41	3.43	0.04	0.28	3.42	3.42	0.04	0.16
	rSVD84-01	3.37	3.36	0.08	0.41	3.43	3.41	0.06	0.27	3.49	3.44	0.07	0.20
Calinski et al. (2009a) Pea 18×9	ABB	3.42	3.36	0.33	0.89	3.36	3.37	0.14	0.54	3.36	3.37	0.10	0.36
	MICE+PMM	3.51	3.45	0.41	0.83	3.49	3.51	0.15	0.53	3.46	3.47	0.11	0.45
	MICE+RF	4.65	4.61	0.79	1.18	4.83	4.79	0.90	1.08	3.43	3.48	0.41	0.90
	rSVD01	10.28	10.29	3.37	2.39	10.32	10.28	1.54	1.67	10.29	10.20	0.73	1.18
	rSVD84	10.24	10.37	3.49	2.62	10.29	10.14	1.56	1.74	10.25	10.28	0.73	1.11
	rSVD84-01	10.24	10.28	3.29	2.66	10.32	10.37	1.54	1.59	10.27	10.28	0.78	1.28
Calinski et al. (2009b) Winter rye 18×15	ABB	7.49	7.54	0.76	1.21	7.48	7.51	0.36	0.83	7.56	7.51	0.26	0.63
	MICE+PMM	6.31	6.16	1.18	1.43	6.13	6.17	0.41	0.89	6.33	6.26	0.30	0.75
	MICE+RF	7.17	7.04	1.84	1.66	6.77	6.77	2.05	2.00	5.84	5.82	1.49	1.82
	rSVD01	12.83	12.90	4.20	2.50	12.85	12.76	1.69	2.05	12.83	12.87	1.53	1.73
	rSVD84	12.96	12.94	4.20	2.60	12.98	12.90	1.85	1.86	13.08	13.14	1.56	1.80
	rSVD84-01	12.84	12.73	4.34	2.48	12.84	12.80	1.84	2.08	12.93	13.02	1.55	1.62

Table 8. P_e by percentage and method — Mechanism **MNAR** (REP 6–10)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	12.05	12.02	0.10	0.38	10.29	10.21	0.07	0.37	9.59	9.57	0.06	0.39
	MICE+PMM	11.71	11.64	0.52	0.86	9.95	9.86	0.20	0.62	9.56	9.53	0.10	0.41
	MICE+RF	12.12	12.34	6.42	3.95	9.56	9.46	3.63	2.24	9.25	9.44	4.49	3.21
	rSVD01	11.73	11.67	0.32	0.88	28.62	28.61	0.00	0.08	24.02	24.03	0.02	0.21
	rSVD84	11.77	11.73	0.05	0.27	9.38	9.35	0.03	0.21	9.58	9.57	0.02	0.18
	rSVD84-01	11.68	11.64	0.21	0.54	14.00	13.96	0.13	0.45	12.44	12.41	0.12	0.49
Yan et al. (2007) Wheat 18×9	ABB	1.17	1.17	0.00	0.03	1.09	1.09	0.00	0.03	0.99	0.98	0.00	0.03
	MICE+PMM	1.09	1.08	0.01	0.08	1.09	1.09	0.00	0.05	1.00	0.99	0.00	0.04
	MICE+RF	1.07	1.08	0.05	0.25	1.15	1.17	0.06	0.37	1.05	1.07	0.04	0.27
	rSVD01	1.30	1.27	0.01	0.14	2.34	2.34	0.00	0.01	2.49	2.49	0.00	0.01
	rSVD84	1.31	1.31	0.00	0.02	1.12	1.12	0.00	0.02	1.66	1.66	0.00	0.01
	rSVD84-01	1.33	1.34	0.00	0.07	1.17	1.17	0.00	0.03	1.76	1.76	0.00	0.02
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	2.86	2.86	0.00	0.04	2.55	2.55	0.00	0.04	2.35	2.35	0.00	0.03
	MICE+PMM	2.50	2.50	0.01	0.14	2.38	2.38	0.00	0.08	2.25	2.25	0.00	0.05
	MICE+RF	2.43	2.40	0.22	0.64	2.34	2.36	0.22	0.62	2.29	2.28	0.23	0.56
	rSVD01	6.41	6.42	0.00	0.04	8.15	8.15	0.00	0.02	10.43	10.43	0.00	0.02
	rSVD84	3.69	3.69	0.00	0.03	3.44	3.44	0.00	0.03	4.65	4.65	0.00	0.01
	rSVD84-01	4.27	4.27	0.01	0.12	4.63	4.63	0.00	0.09	6.29	6.30	0.00	0.06
Calinski et al. (2009a) Pea 18×9	ABB	6.63	6.60	0.02	0.20	6.13	6.12	0.02	0.20	5.82	5.82	0.02	0.17
	MICE+PMM	6.79	6.75	0.15	0.48	6.24	6.21	0.10	0.49	5.71	5.71	0.03	0.26
	MICE+RF	6.41	6.31	1.61	1.90	5.96	6.07	1.36	1.79	6.03	5.98	1.32	1.40
	rSVD01	12.63	12.61	0.02	0.17	23.20	23.20	0.00	0.06	20.14	20.14	0.00	0.08
	rSVD84	12.65	12.63	0.01	0.10	11.04	11.03	0.01	0.12	11.30	11.30	0.00	0.06
	rSVD84-01	12.46	12.48	0.04	0.29	14.52	14.50	0.09	0.39	13.46	13.44	0.06	0.29
Calinski et al. (2009b) Winter rye 18×15	ABB	12.47	12.46	0.10	0.41	12.16	12.14	0.06	0.29	12.52	12.46	0.07	0.32
	MICE+PMM	10.80	10.64	0.93	1.14	11.78	11.85	0.29	0.75	12.46	12.41	0.27	0.58
	MICE+RF	11.33	11.26	4.52	2.68	11.85	12.06	5.69	2.97	11.95	11.95	6.64	3.48
	rSVD01	15.03	15.00	0.04	0.24	15.80	15.81	0.04	0.24	30.93	30.90	0.02	0.19
	rSVD84	15.13	15.13	0.01	0.12	16.16	16.14	0.03	0.13	14.73	14.62	0.07	0.32
	rSVD84-01	14.99	14.99	0.08	0.40	15.94	15.95	0.09	0.41	18.48	18.47	0.19	0.46

Table 9. P_e by percentage and method — Mechanism **MAR** (REP 11–15)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	294.73	295.23	95.62	13.44	299.17	297.80	116.38	14.16	325.23	323.04	146.50	15.62
	MICE+PMM	313.16	309.78	450.49	22.98	316.44	313.56	268.18	22.19	330.32	328.45	216.08	16.47
	MICE+RF	306.65	312.69	3.15e+03	64.58	314.24	318.32	3.15e+03	75.94	329.25	332.41	4.87e+03	78.21
	rSVD01	1.41e+03	1.41e+03	468.78	31.70	1.42e+03	1.42e+03	239.75	19.32	1.39e+03	1.39e+03	250.85	21.90
	rSVD84	1.39e+03	1.39e+03	19.19	6.76	1.42e+03	1.42e+03	18.64	5.10	1.40e+03	1.40e+03	24.34	5.53
	rSVD84-01	1.42e+03	1.43e+03	399.33	25.03	1.41e+03	1.42e+03	315.84	25.58	1.39e+03	1.39e+03	275.42	21.31
Filho et al. (2008) Cotton 17×23	ABB	1.07	1.05	0.00	0.05	0.92	0.91	0.00	0.04	0.80	0.79	0.00	0.03
	MICE+PMM	1.04	1.02	0.01	0.11	0.93	0.91	0.00	0.05	0.81	0.80	0.00	0.04
	MICE+RF	1.08	1.07	0.05	0.28	0.90	0.90	0.03	0.22	0.82	0.82	0.03	0.25
	rSVD01	1.35	1.35	0.00	0.02	1.29	1.29	0.00	0.01	1.28	1.28	0.00	0.02
	rSVD84	1.33	1.33	0.00	0.01	1.28	1.28	0.00	0.01	1.29	1.29	0.00	0.01
	rSVD84-01	1.35	1.35	0.00	0.03	1.29	1.29	0.00	0.02	1.27	1.27	0.00	0.02
Flores et al. (1998) Faba beans 15×12	ABB	793.44	791.12	810.19	34.87	764.16	757.04	1.46e+03	47.35	712.17	709.85	364.61	22.20
	MICE+PMM	793.82	787.09	2.69e+03	61.56	768.15	758.67	2.47e+03	55.97	693.81	694.66	469.97	25.03
	MICE+RF	811.11	815.46	2.45e+04	213.59	789.71	796.92	2.39e+04	234.35	670.92	669.85	2.02e+04	169.92
	rSVD01	1.29e+03	1.29e+03	618.05	31.61	1.12e+03	1.12e+03	732.44	42.60	1.08e+03	1.09e+03	1.07e+03	55.01
	rSVD84	1.27e+03	1.27e+03	115.72	12.72	1.13e+03	1.13e+03	96.74	11.66	1.07e+03	1.07e+03	97.49	7.88
	rSVD84-01	1.28e+03	1.28e+03	1.19e+03	46.70	1.12e+03	1.12e+03	580.96	29.78	1.07e+03	1.07e+03	560.03	28.82
Mattos et al. (2013) Sugarcane 22×5	ABB	1.80	1.77	0.01	0.12	1.68	1.65	0.02	0.15	1.58	1.58	0.01	0.10
	MICE+PMM	1.77	1.75	0.03	0.20	1.69	1.65	0.05	0.15	1.61	1.61	0.02	0.15
	MICE+RF	1.72	1.72	0.11	0.42	1.65	1.67	0.13	0.45	1.60	1.56	0.12	0.43
	rSVD01	2.81	2.83	0.01	0.16	2.51	2.51	0.01	0.13	2.67	2.68	0.01	0.15
	rSVD84	3.12	3.11	0.00	0.09	2.65	2.65	0.00	0.09	2.60	2.60	0.00	0.05
	rSVD84-01	2.76	2.76	0.02	0.20	2.50	2.51	0.02	0.18	2.61	2.60	0.01	0.12
Rad et al. (2013) Wheat 36×6	ABB	3.80	3.79	0.01	0.11	3.59	3.58	0.00	0.08	3.22	3.21	0.00	0.08
	MICE+PMM	2.98	2.96	0.17	0.60	3.35	3.33	0.03	0.22	3.02	3.02	0.03	0.20
	MICE+RF	3.52	3.50	0.56	0.96	3.39	3.38	0.55	1.09	2.91	2.91	0.37	0.81
	rSVD01	4.14	4.14	0.00	0.09	4.62	4.62	0.00	0.04	4.49	4.49	0.00	0.02
	rSVD84	3.57	3.55	0.01	0.09	4.27	4.27	0.00	0.03	4.24	4.23	0.00	0.03
	rSVD84-01	2.62	2.61	0.01	0.12	4.17	4.17	0.00	0.03	4.13	4.13	0.00	0.03

Table 10. P_e by percentage and method — Mechanism **MCAR** (REP 11–15)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	324.05	315.80	2.17e+03	69.60	327.95	327.66	1.06e+03	44.14	325.27	323.75	725.23	37.38
	MICE+PMM	347.12	341.21	2.79e+03	65.34	341.55	338.51	1.21e+03	56.64	332.99	329.00	879.02	35.61
	MICE+RF	289.45	294.91	3.68e+03	79.84	301.23	306.36	4.20e+03	99.67	327.45	332.11	4.87e+03	104.05
	rSVD01	1.35e+03	1.35e+03	1.29e+04	150.25	1.34e+03	1.35e+03	8.80e+03	133.27	1.34e+03	1.34e+03	4.29e+03	75.81
	rSVD84	1.35e+03	1.35e+03	1.13e+04	166.09	1.35e+03	1.34e+03	9.12e+03	135.22	1.34e+03	1.33e+03	4.30e+03	87.43
	rSVD84-01	1.35e+03	1.34e+03	1.33e+04	183.47	1.34e+03	1.34e+03	8.85e+03	129.34	1.34e+03	1.34e+03	4.56e+03	84.21
Filho et al. (2008) Cotton 17×23	ABB	0.54	0.53	0.01	0.16	0.57	0.57	0.01	0.10	0.57	0.56	0.00	0.06
	MICE+PMM	0.56	0.55	0.01	0.14	0.56	0.56	0.00	0.09	0.56	0.55	0.00	0.07
	MICE+RF	0.57	0.58	0.01	0.15	0.65	0.65	0.01	0.17	0.67	0.68	0.02	0.19
	rSVD01	1.23	1.23	0.01	0.13	1.22	1.22	0.00	0.07	1.22	1.22	0.00	0.09
	rSVD84	1.23	1.23	0.01	0.14	1.21	1.21	0.00	0.07	1.22	1.22	0.00	0.07
	rSVD84-01	1.23	1.23	0.01	0.13	1.21	1.21	0.00	0.08	1.22	1.21	0.00	0.08
Flores et al. (1998) Faba beans 15×12	ABB	464.49	467.14	7.18e+03	113.53	491.87	492.36	4.23e+03	71.91	496.71	496.52	2.11e+03	47.20
	MICE+PMM	486.07	494.81	7.81e+03	112.57	484.01	477.26	4.42e+03	81.06	489.96	481.79	4.48e+03	69.91
	MICE+RF	528.54	530.53	9.81e+03	135.54	608.88	605.88	1.37e+04	147.73	449.12	454.40	9.35e+03	130.24
	rSVD01	960.48	973.05	3.13e+04	257.92	994.52	987.43	9.68e+03	129.28	982.61	991.59	8.27e+03	126.02
	rSVD84	956.30	979.81	2.92e+04	243.01	987.93	985.35	9.74e+03	132.87	992.80	995.15	7.96e+03	127.80
	rSVD84-01	954.05	959.61	2.96e+04	237.65	992.32	989.68	9.89e+03	144.18	991.60	996.48	8.15e+03	129.95
Mattos et al. (2013) Sugarcane 22×5	ABB	1.76	1.76	0.18	0.61	1.80	1.76	0.08	0.37	1.82	1.84	0.05	0.32
	MICE+PMM	1.80	1.75	0.21	0.68	1.81	1.77	0.07	0.33	1.83	1.83	0.05	0.29
	MICE+RF	1.77	1.74	0.12	0.44	1.39	1.41	0.07	0.28	1.68	1.67	0.12	0.47
	rSVD01	3.04	3.07	0.29	0.64	3.56	3.16	1.32	0.51	4.30	3.39	2.00	2.41
	rSVD84	3.02	2.96	0.28	0.72	3.14	3.14	0.30	0.52	3.29	3.13	0.44	0.42
	rSVD84-01	3.02	3.02	0.27	0.71	3.27	3.16	0.39	0.59	3.60	3.37	0.56	0.73
Rad et al. (2013) Wheat 36×6	ABB	1.88	1.90	0.11	0.45	1.93	1.92	0.04	0.29	1.96	1.95	0.02	0.21
	MICE+PMM	1.08	1.04	0.10	0.43	1.11	1.09	0.07	0.41	1.16	1.16	0.04	0.25
	MICE+RF	1.13	1.10	0.07	0.39	0.85	0.85	0.03	0.21	0.95	0.98	0.04	0.25
	rSVD01	2.84	2.84	0.10	0.40	2.95	2.88	0.16	0.32	3.19	3.00	0.26	0.77
	rSVD84	2.78	2.80	0.11	0.41	2.88	2.88	0.04	0.24	2.87	2.88	0.02	0.24
	rSVD84-01	2.83	2.86	0.10	0.42	2.90	2.88	0.06	0.25	2.96	2.95	0.05	0.27

Table 11. P_e by percentage and method — Mechanism **MNAR** (REP 11–15)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	631.11	628.93	148.87	14.00	590.87	590.73	89.76	12.80	551.66	550.27	120.89	15.84
	MICE+PMM	642.89	641.05	554.94	30.16	604.86	605.69	211.46	18.76	560.49	559.21	232.97	20.72
	MICE+RF	618.10	639.62	1.53e+04	181.18	604.15	603.08	1.47e+04	172.13	586.07	574.55	1.81e+04	174.13
	rSVD01	1.29e+03	1.30e+03	111.28	16.42	1.30e+03	1.30e+03	128.39	16.72	1.32e+03	1.33e+03	460.65	35.82
	rSVD84	1.29e+03	1.29e+03	23.47	5.94	1.28e+03	1.28e+03	16.13	4.80	1.32e+03	1.32e+03	21.69	4.91
	rSVD84-01	1.30e+03	1.30e+03	299.42	25.82	1.30e+03	1.30e+03	223.04	21.88	1.31e+03	1.31e+03	244.33	19.61
Filho et al. (2008) Cotton 17×23	ABB	1.18	1.17	0.00	0.04	1.06	1.05	0.00	0.04	1.01	1.01	0.00	0.03
	MICE+PMM	1.18	1.17	0.00	0.07	1.06	1.05	0.00	0.05	1.03	1.03	0.00	0.04
	MICE+RF	1.29	1.28	0.07	0.39	1.07	1.06	0.04	0.27	1.04	1.05	0.04	0.24
	rSVD01	1.35	1.35	0.00	0.03	1.25	1.25	0.00	0.02	1.25	1.25	0.00	0.02
	rSVD84	1.35	1.35	0.00	0.01	1.27	1.27	0.00	0.01	1.27	1.27	0.00	0.01
	rSVD84-01	1.33	1.33	0.00	0.03	1.26	1.26	0.00	0.02	1.26	1.25	0.00	0.03
Flores et al. (1998) Faba beans 15×12	ABB	985.33	983.48	511.05	26.56	956.45	955.79	404.74	30.15	837.85	834.76	479.02	30.07
	MICE+PMM	919.49	907.22	2.69e+03	62.22	950.04	945.64	1.92e+03	58.48	837.28	834.14	881.59	43.72
	MICE+RF	895.34	867.34	2.94e+04	220.28	950.71	985.77	5.04e+04	285.36	849.67	821.79	3.65e+04	218.48
	rSVD01	1.26e+03	1.26e+03	448.53	32.78	1.25e+03	1.25e+03	322.08	20.55	1.12e+03	1.12e+03	304.03	19.33
	rSVD84	1.27e+03	1.27e+03	161.16	14.10	1.25e+03	1.25e+03	134.30	13.52	1.11e+03	1.11e+03	104.52	11.18
	rSVD84-01	1.28e+03	1.28e+03	982.37	42.18	1.26e+03	1.26e+03	622.39	35.80	1.11e+03	1.11e+03	685.70	35.47
Mattos et al. (2013) Sugarcane 22×5	ABB	3.28	3.26	0.02	0.16	3.28	3.26	0.01	0.17	3.32	3.28	0.01	0.16
	MICE+PMM	3.12	3.10	0.04	0.26	3.23	3.19	0.02	0.16	3.25	3.24	0.01	0.16
	MICE+RF	3.15	3.23	0.46	0.79	3.15	3.16	0.26	0.83	3.27	3.41	0.38	0.80
	rSVD01	3.72	3.70	0.02	0.16	6.51	6.49	0.00	0.07	7.58	7.58	0.00	0.02
	rSVD84	3.64	3.62	0.01	0.09	3.26	3.25	0.00	0.10	3.04	3.03	0.00	0.08
	rSVD84-01	3.61	3.60	0.03	0.22	3.47	3.48	0.01	0.16	3.71	3.71	0.01	0.13
Rad et al. (2013) Wheat 36×6	ABB	4.37	4.36	0.01	0.09	4.10	4.09	0.00	0.07	3.71	3.71	0.00	0.07
	MICE+PMM	4.15	4.14	0.03	0.21	3.92	3.92	0.01	0.15	3.60	3.59	0.02	0.16
	MICE+RF	4.28	4.32	0.76	1.26	3.86	3.88	0.48	0.79	3.55	3.53	0.51	0.88
	rSVD01	5.12	5.12	0.00	0.06	4.89	4.89	0.00	0.07	4.79	4.78	0.00	0.04
	rSVD84	4.67	4.66	0.00	0.09	4.65	4.65	0.00	0.03	4.72	4.72	0.00	0.01
	rSVD84-01	3.77	3.79	0.01	0.11	4.47	4.47	0.00	0.07	4.76	4.75	0.00	0.05

Table 12. P_e by percentage and method — Mechanism **MAR** (REP 16–17)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	3.23	3.22	0.03	0.24	3.13	3.11	0.02	0.20	3.26	3.24	0.03	0.29
	MICE+PMM	2.33	2.29	0.20	0.53	2.49	2.44	0.16	0.54	2.63	2.56	0.12	0.47
	MICE+RF	1.94	1.97	0.14	0.55	2.17	2.20	0.20	0.58	2.26	2.26	0.21	0.54
	rSVD01	2.32	2.32	0.03	0.27	2.32	2.31	0.00	0.08	2.25	2.24	0.00	0.05
	rSVD84	2.55	2.53	0.01	0.15	2.63	2.62	0.01	0.09	2.25	2.24	0.00	0.05
	rSVD84-01	1.78	1.77	0.03	0.24	2.06	2.05	0.01	0.12	2.25	2.25	0.00	0.06
Zobel et al. (1988) Soybean 7×35	ABB	1.36e+03	1.36e+03	1.47e+03	45.43	1.29e+03	1.30e+03	1.41e+03	50.46	1.33e+03	1.33e+03	1.75e+03	55.76
	MICE+PMM	917.42	896.32	1.18e+04	135.16	935.31	921.32	1.00e+04	117.58	1.08e+03	1.06e+03	1.31e+04	159.83
	MICE+RF	1.14e+03	1.15e+03	5.02e+04	338.60	899.29	913.18	3.93e+04	266.58	1.33e+03	1.38e+03	9.00e+04	394.69
	rSVD01	1.18e+03	1.18e+03	1.07e+03	46.32	1.01e+03	1.01e+03	1.20e+03	52.38	1.00e+03	997.09	446.74	26.38
	rSVD84	1.16e+03	1.15e+03	470.76	27.11	1.02e+03	1.02e+03	1.09e+03	27.81	891.27	887.05	598.14	27.44
	rSVD84-01	1.17e+03	1.16e+03	1.38e+03	53.91	886.75	888.07	1.51e+03	51.75	751.55	751.61	895.85	42.38

Table 13. P_e by percentage and method — Mechanism **MAR** (REP 16–17)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	3.23	3.22	0.03	0.24	3.13	3.11	0.02	0.20	3.26	3.24	0.03	0.29
	MICE+PMM	2.33	2.29	0.20	0.53	2.49	2.44	0.16	0.54	2.63	2.56	0.12	0.47
	MICE+RF	1.94	1.97	0.14	0.55	2.17	2.20	0.20	0.58	2.26	2.26	0.21	0.54
	rSVD01	2.32	2.32	0.03	0.27	2.32	2.31	0.00	0.08	2.25	2.24	0.00	0.05
	rSVD84	2.55	2.53	0.01	0.15	2.63	2.62	0.01	0.09	2.25	2.24	0.00	0.05
	rSVD84-01	1.78	1.77	0.03	0.24	2.06	2.05	0.01	0.12	2.25	2.25	0.00	0.06
Zobel et al. (1988) Soybean 7×35	ABB	1.36e+03	1.36e+03	1.47e+03	45.43	1.29e+03	1.30e+03	1.41e+03	50.46	1.33e+03	1.33e+03	1.75e+03	55.76
	MICE+PMM	917.42	896.32	1.18e+04	135.16	935.31	921.32	1.00e+04	117.58	1.08e+03	1.06e+03	1.31e+04	159.83
	MICE+RF	1.14e+03	1.15e+03	5.02e+04	338.60	899.29	913.18	3.93e+04	266.58	1.33e+03	1.38e+03	9.00e+04	394.69
	rSVD01	1.18e+03	1.18e+03	1.07e+03	46.32	1.01e+03	1.01e+03	1.20e+03	52.38	1.00e+03	997.09	446.74	26.38
	rSVD84	1.16e+03	1.15e+03	470.76	27.11	1.02e+03	1.02e+03	1.09e+03	27.81	891.27	887.05	598.14	27.44
	rSVD84-01	1.17e+03	1.16e+03	1.38e+03	53.91	886.75	888.07	1.51e+03	51.75	751.55	751.61	895.85	42.38

Table 14. P_e by percentage and method — Mechanism **MCAR** (REP 16–17)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	2.01	2.00	0.18	0.66	2.06	2.04	0.07	0.35	2.15	2.13	0.05	0.32
	MICE+PMM	1.16	1.15	0.22	0.61	1.25	1.16	0.13	0.47	1.39	1.37	0.07	0.39
	MICE+RF	1.46	1.51	0.09	0.44	0.91	0.93	0.04	0.23	1.71	1.69	0.09	0.35
	rSVD01	1.93	1.92	0.18	0.69	1.94	1.88	0.22	0.39	2.18	2.04	0.30	0.31
	rSVD84	1.90	1.84	0.16	0.60	1.98	1.95	0.08	0.40	2.04	2.02	0.08	0.34
	rSVD84-01	1.90	1.85	0.20	0.72	1.93	1.84	0.12	0.39	2.08	2.00	0.13	0.26
Zobel et al. (1988) Soybean 7×35	ABB	789.51	795.01	7.10e+03	108.09	777.50	778.47	3.76e+03	77.21	773.49	772.63	2.06e+03	64.20
	MICE+PMM	491.73	468.44	1.80e+04	154.64	492.23	475.26	1.13e+04	151.07	472.93	472.60	4.07e+03	80.91
	MICE+RF	477.63	487.07	8.42e+03	130.30	551.46	534.49	1.24e+04	150.87	416.96	420.64	6.68e+03	100.31
	rSVD01	776.40	796.92	7.84e+03	135.05	762.79	756.59	4.77e+03	93.45	797.79	777.30	1.08e+04	71.32
	rSVD84	762.38	767.25	7.25e+03	117.88	776.38	775.77	3.80e+03	81.85	787.77	776.46	5.92e+03	61.14
	rSVD84-01	777.14	786.01	9.42e+03	137.09	769.64	764.63	3.97e+03	85.69	781.76	772.88	6.12e+03	70.63

Table 15. P_e by percentage and method — Mechanism **MNAR** (REP 16–17)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	3.20	3.21	0.01	0.15	3.18	3.17	0.01	0.11	3.16	3.15	0.01	0.10
	MICE+PMM	2.47	2.42	0.12	0.44	2.81	2.81	0.02	0.20	2.71	2.69	0.03	0.24
	MICE+RF	2.25	2.32	0.22	0.59	2.71	2.69	0.25	0.71	2.85	2.80	0.26	0.59
	rSVD01	2.42	2.41	0.01	0.14	2.72	2.71	0.00	0.04	2.75	2.74	0.00	0.04
	rSVD84	2.68	2.65	0.02	0.12	2.66	2.66	0.00	0.02	2.57	2.56	0.00	0.06
	rSVD84-01	2.00	1.99	0.03	0.22	2.71	2.70	0.00	0.04	2.38	2.37	0.00	0.07
Zobel et al. (1988) Soybean 7×35	ABB	1.42e+03	1.43e+03	736.27	35.98	1.35e+03	1.35e+03	379.04	23.30	1.31e+03	1.30e+03	304.10	24.82
	MICE+PMM	1.06e+03	1.03e+03	8.32e+03	128.95	1.11e+03	1.10e+03	5.30e+03	95.21	1.15e+03	1.14e+03	1.91e+03	64.16
	MICE+RF	970.88	961.38	3.83e+04	268.50	1.33e+03	1.33e+03	8.46e+04	392.68	1.20e+03	1.23e+03	6.26e+04	333.73
	rSVD01	1.35e+03	1.34e+03	754.86	38.95	1.28e+03	1.27e+03	1.16e+03	36.64	1.40e+03	1.40e+03	316.30	24.94
	rSVD84	1.36e+03	1.35e+03	307.48	20.37	1.27e+03	1.26e+03	615.71	27.99	1.35e+03	1.34e+03	254.10	16.83
	rSVD84-01	1.34e+03	1.34e+03	1.49e+03	42.67	1.26e+03	1.26e+03	1.19e+03	41.59	1.31e+03	1.30e+03	330.06	19.99

5.4 Results for real data with 100 standard deviation environment

In this case, the behavior observed in the simulations is confirmed: when extreme outliers are introduced, the SVD-based models improve their relative performance compared to the other imputers. As shown in Table 19, the mean P_e for the other models increases substantially, while the P_e for the rSVD models remains close to the original values. Another example can be seen in table 19 in compared with table 6. Here, in most data sets, P_e values did not increase dramatically in the rSVD models, This also occurred with the data sets Filho et al. (2008), Mattos et al. (2013) and Rad et al. (2013), as seen in Tables 22, 23, 24 compared to the results in Tables 9, 10 and 11.
At this point it is logical to think that the main advantage of SVD models is that, in the presence of extreme outliers, their performance does not deteriorate significantly, making them a reliable choice when precision is critical under such conditions.

However, one notable observation consistent with the simulation results is that the variance among imputations increases substantially for all models in the presence of extreme outliers. This indicates that, although SVD models maintain good predictive accuracy, the uncertainty associated with their imputations grows considerably, highlighting the inherent unreliability of imputing matrices containing extreme outliers.

Table 16. P_e by percentage and method — Mechanism **MAR** (REP 1–5)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	23.45	19.24	233.69	15.78	26.29	21.83	253.60	22.19	30.92	21.03	504.42	20.41
	MICE+PMM	32.31	16.75	1.89e+03	26.46	33.75	23.96	964.99	33.16	39.53	24.39	1.65e+03	38.51
	MICE+RF	185.27	183.91	1.81e+03	63.37	25.82	25.83	23.28	6.22	25.65	25.77	21.16	6.14
	rSVD01	35.88	33.65	287.86	2.13	43.93	31.22	995.06	11.31	42.39	31.26	1.11e+03	11.15
	rSVD84	33.22	33.20	3.78	1.05	30.51	29.21	19.64	2.35	31.52	29.51	22.40	4.01
	rSVD84-01	34.12	33.42	27.24	1.74	37.30	31.03	790.39	6.07	36.07	30.24	513.91	7.02
Bro and Smilde (2014) Red Wine 14×44	ABB	7.51	7.09	17.62	6.13	5.87	4.88	10.75	4.27	6.16	6.29	10.08	4.85
	MICE+PMM	8.82	6.00	75.27	4.67	8.73	5.53	66.20	7.51	8.91	7.30	35.00	8.66
	MICE+RF	30.59	30.43	33.12	7.18	3.05	3.02	0.31	0.89	6.55	6.51	1.80	1.85
	rSVD01	95.90	23.18	1.38e+05	24.93	89.38	48.13	2.29e+04	66.85	139.89	72.46	1.07e+05	120.06
	rSVD84	10.54	10.56	0.00	0.07	10.72	10.71	0.01	0.10	12.82	11.02	62.49	0.35
	rSVD84-01	31.93	15.01	2.27e+03	11.32	46.65	22.38	3.74e+03	28.99	156.29	36.82	2.87e+05	90.02
Forkman and Piepho (2014) Soybean 7×10	ABB	2.59e+03	2.40e+03	1.87e+06	2.01e+03	2.70e+03	2.64e+03	2.38e+06	2.41e+03	2.71e+03	2.11e+03	2.88e+06	2.93e+03
	MICE+PMM	1.85e+03	1.37e+03	1.62e+06	1.53e+03	2.06e+03	1.44e+03	1.72e+06	1.91e+03	1.82e+03	1.15e+03	2.32e+06	1.44e+03
	MICE+RF	985.60	992.79	3.67e+04	260.92	3.66e+03	3.65e+03	4.55e+05	705.12	739.89	734.12	2.70e+04	194.47
	rSVD01	1.11e+03	1.04e+03	3.96e+04	93.01	1.21e+03	1.08e+03	1.27e+05	131.49	1.24e+03	1.10e+03	1.40e+05	204.41
	rSVD84	1.16e+03	1.10e+03	2.04e+04	34.50	1.10e+03	1.06e+03	1.60e+04	30.10	1.20e+03	1.11e+03	4.34e+04	184.72
	rSVD84-01	1.15e+03	1.08e+03	4.62e+04	82.73	1.11e+03	1.04e+03	5.68e+04	116.75	1.20e+03	1.09e+03	1.27e+05	183.37
Forkman and Piepho (2014) Pea 5×4	ABB	7.08e+03	8.37e+03	1.53e+07	5.70e+03	7.45e+03	8.35e+03	1.71e+07	5.02e+03	7.90e+03	7.75e+03	2.89e+07	1.10e+04
	MICE+PMM	8.43e+03	9.57e+03	2.29e+07	6.67e+03	8.62e+03	1.01e+04	2.26e+07	5.75e+03	7.25e+03	1.18e+03	4.92e+07	1.21e+04
	MICE+RF	9.17e+03	9.19e+03	3.80e+06	2.73e+03	9.19e+03	9.25e+03	3.77e+06	2.98e+03	1.30e+04	1.31e+04	6.54e+06	3.35e+03
	rSVD01	3.16e+03	3.16e+03	0.00	0.00	3.16e+03	3.16e+03	0.00	0.00	4.31e+03	3.80e+03	3.29e+06	1.31e+03
	rSVD84	3.20e+03	2.92e+03	2.76e+06	2.94e+03	3.74e+03	2.88e+03	7.26e+06	3.79e+03	5.65e+03	4.40e+03	8.81e+06	3.98e+03
	rSVD84-01	958.25	748.59	6.62e+05	741.97	870.32	699.81	5.82e+05	615.31	3.23e+03	2.83e+03	3.87e+06	2.36e+03
Cornelius et al. (1996) Maize 9×20	ABB	6.61e+03	6.49e+03	3.69e+06	2.60e+03	6.96e+03	6.69e+03	5.56e+06	2.87e+03	7.20e+03	6.87e+03	7.56e+06	3.42e+03
	MICE+PMM	8.19e+03	7.07e+03	3.91e+07	8.42e+03	9.59e+03	9.09e+03	2.65e+07	6.11e+03	9.98e+03	8.62e+03	3.28e+07	6.70e+03
	MICE+RF	1.15e+04	1.14e+04	4.89e+06	2.51e+03	2.69e+04	2.75e+04	2.80e+07	7.25e+03	1.01e+04	1.02e+04	4.19e+06	2.59e+03
	rSVD01	4.58e+03	4.35e+03	4.08e+06	2.26e+03	2.52e+03	2.36e+03	4.38e+05	780.61	3.87e+03	3.48e+03	3.44e+06	1.66e+03
	rSVD84	3.66e+03	3.28e+03	2.55e+06	1.58e+03	3.20e+03	2.99e+03	1.57e+06	1.68e+03	3.59e+03	3.20e+03	2.40e+06	1.60e+03
	rSVD84-01	4.21e+03	3.44e+03	4.35e+06	2.21e+03	2.16e+03	1.97e+03	7.21e+05	1.05e+03	3.11e+03	2.87e+03	2.10e+06	1.42e+03

Table 17. P_e by percentage and method — Mechanism **MCAR** (REP 1–5)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	21.73	17.56	275.91	18.45	23.99	22.16	216.99	22.81	31.02	31.97	303.95	27.31
	MICE+PMM	23.12	13.29	639.30	22.75	28.28	18.92	625.14	31.39	32.60	27.11	584.46	35.10
	MICE+RF	4.19	4.16	0.62	1.04	52.88	52.19	124.22	15.02	26.88	27.01	26.71	7.04
	rSVD01	24.59	25.07	45.23	8.70	26.32	25.42	43.78	6.12	31.72	27.71	174.00	10.92
	rSVD84	24.00	24.67	41.45	7.67	24.38	23.99	18.84	5.26	26.15	25.49	20.61	4.59
	rSVD84-01	24.19	24.73	45.79	9.60	25.46	24.74	29.74	5.00	31.01	26.57	385.62	6.43
Bro and Smilde (2014) Red Wine 14×44	ABB	91.05	52.42	5.10e+03	109.78	105.51	80.82	6.65e+03	111.33	114.66	82.06	1.29e+04	120.90
	MICE+PMM	128.71	48.34	2.95e+04	134.71	126.72	54.51	1.56e+04	176.38	154.15	54.48	3.12e+04	194.10
	MICE+RF	40.93	41.17	56.45	9.80	42.19	41.45	44.39	8.03	269.32	271.25	2.55e+03	71.37
	rSVD01	125.12	99.75	3.88e+04	42.95	142.59	100.81	4.58e+04	27.39	318.66	101.68	2.21e+06	40.66
	rSVD84	87.72	89.73	754.47	34.53	88.87	90.32	345.28	26.86	89.22	90.22	210.99	15.35
	rSVD84-01	95.86	97.13	1.06e+03	37.43	101.76	97.38	1.19e+03	28.60	107.21	93.25	3.94e+03	24.96
Forkman and Piepho (2014) Soybean 7×10	ABB	1.54e+03	591.57	3.67e+06	1.58e+03	2.20e+03	1.83e+03	3.32e+06	2.27e+03	2.43e+03	2.16e+03	2.53e+06	1.99e+03
	MICE+PMM	1.08e+03	476.78	2.35e+06	712.21	1.78e+03	851.76	3.32e+06	2.12e+03	1.54e+03	908.93	1.83e+06	1.93e+03
	MICE+RF	1.59e+03	1.56e+03	9.12e+04	490.74	2.57e+03	2.55e+03	2.09e+05	669.15	3.00e+03	3.05e+03	4.44e+05	999.24
	rSVD01	895.52	887.65	7.76e+04	366.12	984.26	959.87	5.43e+04	261.81	1.04e+03	995.75	6.65e+04	236.48
	rSVD84	872.10	888.68	6.54e+04	360.69	927.04	930.15	2.67e+04	219.58	968.10	948.28	3.53e+04	212.29
	rSVD84-01	882.39	872.74	7.08e+04	363.39	947.74	950.37	2.32e+04	194.58	1.01e+03	980.94	4.96e+04	199.87
Forkman and Piepho (2014) Pea 5×4	ABB	4.50e+03	758.31	4.37e+07	9.04e+03	5.61e+03	3.89e+03	3.39e+07	8.51e+03	7.94e+03	6.98e+03	2.86e+07	5.97e+03
	MICE+PMM	5.88e+03	885.60	8.92e+07	7.68e+03	5.25e+03	784.79	5.75e+07	1.07e+04	5.61e+03	859.16	6.71e+07	1.02e+04
	MICE+RF	3.97e+03	607.52	5.79e+07	1.14e+03	5.95e+03	912.86	5.64e+07	9.28e+03	62.30	62.42	158.19	17.92
	rSVD01	3.67e+03	2.76e+03	1.48e+07	2.45e+03	5.09e+03	3.97e+03	1.65e+07	3.12e+03	5.18e+03	4.26e+03	7.81e+06	2.67e+03
	rSVD84	2.71e+03	2.02e+03	6.59e+06	1.91e+03	4.02e+03	3.57e+03	4.20e+06	2.43e+03	5.85e+03	4.89e+03	1.13e+07	3.65e+03
	rSVD84-01	2.77e+03	2.37e+03	3.57e+06	1.90e+03	4.31e+03	3.62e+03	5.86e+06	2.70e+03	4.35e+03	3.99e+03	5.11e+06	2.54e+03
Cornelius et al. (1996) Maize 9×20	ABB	5.30e+03	5.11e+03	4.03e+06	2.45e+03	6.15e+03	6.06e+03	3.13e+06	1.77e+03	6.63e+03	6.43e+03	2.89e+06	2.07e+03
	MICE+PMM	6.62e+03	6.03e+03	1.46e+07	5.38e+03	6.35e+03	6.25e+03	7.61e+06	2.85e+03	6.71e+03	6.55e+03	5.35e+06	3.15e+03
	MICE+RF	7.14e+03	7.01e+03	1.71e+06	1.62e+03	7.03e+03	7.00e+03	1.99e+06	1.90e+03	3.51e+03	3.47e+03	5.33e+05	983.47
	rSVD01	1.55e+03	1.46e+03	1.58e+05	397.07	1.91e+03	1.83e+03	2.69e+05	540.59	2.36e+03	2.25e+03	5.55e+05	1.05e+03
	rSVD84	1.46e+03	1.41e+03	7.81e+04	327.30	1.80e+03	1.70e+03	1.77e+05	523.63	2.04e+03	1.91e+03	3.11e+05	647.81
	rSVD84-01	1.54e+03	1.47e+03	1.02e+05	375.44	1.79e+03	1.67e+03	2.57e+05	416.14	2.18e+03	2.08e+03	3.48e+05	733.05

Table 18. P_e by percentage and method — Mechanism **MNAR** (REP 1–5)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	30.19	28.94	182.16	21.07	33.07	31.82	241.78	21.08	36.63	37.25	254.86	24.95
	MICE+PMM	40.45	25.19	1.38e+03	49.63	35.43	31.23	526.61	26.26	45.20	38.91	664.75	41.85
	MICE+RF	16.37	16.43	9.07	3.41	31.02	30.58	40.33	7.36	58.52	59.13	124.55	15.21
	rSVD01	27.35	22.67	608.61	8.26	43.31	21.86	1.90e+04	8.11	33.50	26.60	279.13	14.30
	rSVD84	20.06	18.28	15.24	4.03	21.16	20.16	11.20	4.22	25.98	24.69	42.17	5.13
	rSVD84-01	30.28	20.48	6.37e+03	5.08	26.54	21.10	1.12e+03	5.57	29.34	25.04	164.80	10.46
Bro and Smilde (2014) Red Wine 14×44	ABB	177.51	174.45	1.26e+04	180.67	150.52	147.12	9.41e+03	153.43	150.11	64.83	1.26e+04	155.85
	MICE+PMM	212.68	116.91	4.57e+04	234.39	201.99	74.91	4.82e+04	216.37	198.62	67.47	3.83e+04	229.19
	MICE+RF	184.12	183.59	1.17e+03	47.31	523.20	512.46	1.06e+04	137.68	5.68	5.71	1.05	1.16
	rSVD01	92.86	52.28	1.27e+04	35.91	88.98	51.76	2.13e+04	27.80	137.18	61.08	3.35e+04	105.75
	rSVD84	34.71	34.72	0.01	0.06	43.31	42.96	2.45	0.06	49.29	47.74	55.90	0.07
	rSVD84-01	122.96	35.29	3.76e+05	17.05	60.72	43.87	2.73e+03	11.90	84.30	53.02	1.01e+04	22.13
Forkman and Piepho (2014) Soybean 7×10	ABB	2.53e+03	2.42e+03	1.71e+06	1.64e+03	3.03e+03	2.99e+03	2.13e+06	2.41e+03	2.73e+03	2.37e+03	1.82e+06	2.10e+03
	MICE+PMM	2.11e+03	1.46e+03	2.15e+06	1.69e+03	2.09e+03	1.75e+03	1.36e+06	1.56e+03	1.83e+03	1.41e+03	1.06e+06	1.45e+03
	MICE+RF	1.08e+03	1.08e+03	5.51e+04	329.62	767.36	768.95	3.06e+04	264.17	1.85e+03	1.82e+03	1.30e+05	467.12
	rSVD01	1.18e+03	1.12e+03	2.00e+04	63.28	1.20e+03	1.13e+03	3.79e+04	259.30	1.22e+03	1.10e+03	7.25e+04	261.77
	rSVD84	1.15e+03	1.12e+03	6.55e+03	45.15	1.15e+03	1.05e+03	3.48e+04	236.62	1.11e+03	1.04e+03	3.14e+04	179.32
	rSVD84-01	1.18e+03	1.13e+03	1.64e+04	100.07	1.18e+03	1.11e+03	4.69e+04	134.29	1.15e+03	1.07e+03	5.66e+04	131.72
Forkman and Piepho (2014) Pea 5×4	ABB	8.49e+03	8.44e+03	2.98e+06	2.42e+03	8.02e+03	7.86e+03	3.56e+06	2.25e+03	8.34e+03	7.93e+03	3.30e+06	2.00e+03
	MICE+PMM	9.79e+03	9.43e+03	5.42e+06	3.38e+03	9.48e+03	9.27e+03	5.43e+06	3.05e+03	9.85e+03	9.52e+03	4.79e+06	3.01e+03
	MICE+RF	7.11e+03	7.17e+03	2.02e+06	1.60e+03	8.05e+03	8.11e+03	1.88e+06	1.50e+03	7.76e+03	7.71e+03	2.75e+06	2.37e+03
	rSVD01	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00
	rSVD84	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00
	rSVD84-01	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00	3.17e+03	3.17e+03	0.00	0.00
Cornelius et al. (1996) Maize 9×20	ABB	6.87e+03	6.80e+03	1.55e+06	1.36e+03	7.33e+03	7.18e+03	1.63e+06	1.60e+03	8.17e+03	8.13e+03	1.58e+06	1.60e+03
	MICE+PMM	9.45e+03	8.89e+03	1.50e+07	5.30e+03	8.25e+03	8.04e+03	6.28e+06	3.07e+03	8.47e+03	8.04e+03	7.68e+06	3.19e+03
	MICE+RF	7.62e+03	7.63e+03	2.33e+06	2.34e+03	8.24e+03	8.22e+03	2.59e+06	2.21e+03	6.50e+03	6.47e+03	1.65e+06	1.86e+03
	rSVD01	2.82e+03	2.56e+03	2.53e+05	650.79	3.01e+03	2.90e+03	2.32e+05	358.92	3.57e+03	3.12e+03	1.75e+06	713.87
	rSVD84	3.89e+03	3.76e+03	2.05e+06	1.68e+03	3.55e+03	3.35e+03	1.10e+06	1.27e+03	3.56e+03	3.21e+03	1.35e+06	1.10e+03
	rSVD84-01	2.22e+03	1.95e+03	4.35e+05	799.66	2.69e+03	2.59e+03	1.87e+05	589.11	3.26e+03	3.02e+03	1.43e+06	799.43

Table 19. P_e by percentage and method — Mechanism **MAR** (REP 6–10)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	26.09	23.26	111.70	13.02	21.55	20.40	86.98	12.88	23.55	19.95	138.02	10.09
	MICE+PMM	35.94	28.85	706.40	28.62	26.07	19.54	279.79	22.42	26.31	23.28	255.87	18.64
	MICE+RF	21.47	21.54	20.84	6.95	23.25	23.13	22.89	6.09	7.32	7.44	2.24	1.98
	rSVD01	13.31	12.40	15.58	5.54	15.17	13.38	25.70	5.58	26.55	26.25	1.00	1.38
	rSVD84	11.99	10.84	11.02	4.53	13.25	12.66	7.00	3.49	12.15	11.73	7.75	3.38
	rSVD84-01	12.55	11.31	12.55	3.94	14.73	13.18	25.51	5.87	13.73	13.43	3.91	2.29
Yan et al. (2007) Wheat 18×9	ABB	2.50	2.41	0.42	0.89	2.65	2.63	0.47	0.82	2.27	2.26	0.38	0.80
	MICE+PMM	3.40	2.99	4.80	2.24	3.55	3.34	2.43	1.91	2.59	2.61	1.56	1.67
	MICE+RF	1.28	1.27	0.06	0.28	3.59	3.72	0.59	1.12	4.34	4.47	0.80	1.14
	rSVD01	1.92	1.67	0.69	1.01	2.17	1.89	1.10	0.93	1.58	1.49	0.19	0.44
	rSVD84	1.49	1.41	0.18	0.57	1.60	1.45	0.32	0.53	1.28	1.18	0.10	0.30
	rSVD84-01	1.74	1.69	0.46	0.89	1.91	1.70	0.60	0.97	1.45	1.31	0.20	0.46
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	5.47	5.36	1.61	1.81	5.25	5.06	1.47	1.49	5.37	5.05	2.54	1.93
	MICE+PMM	6.51	5.11	17.43	3.24	6.10	4.95	11.01	3.33	7.16	6.46	12.51	4.33
	MICE+RF	3.77	3.81	0.69	1.11	11.29	11.35	5.66	2.52	6.94	6.96	1.58	1.57
	rSVD01	3.87	3.69	0.33	0.73	5.54	5.49	0.04	0.19	6.24	6.23	0.04	0.24
	rSVD84	3.59	3.56	0.12	0.42	3.91	3.86	0.10	0.44	3.91	3.86	0.06	0.30
	rSVD84-01	3.86	3.79	0.29	0.64	3.98	3.96	0.06	0.30	4.30	4.30	0.04	0.25
Calinski et al. (2009a) Pea 18×9	ABB	12.78	12.04	22.55	8.14	15.55	15.76	28.43	7.10	13.53	13.03	43.25	9.25
	MICE+PMM	18.31	17.49	103.58	14.50	20.65	18.69	133.83	15.98	15.79	13.79	98.34	14.23
	MICE+RF	26.78	27.42	28.28	7.61	33.24	33.95	50.13	9.12	32.71	32.47	45.47	10.72
	rSVD01	9.28	9.17	0.42	0.37	8.77	8.55	0.94	1.19	9.36	9.06	1.41	1.50
	rSVD84	9.12	9.16	0.15	0.32	8.29	8.28	0.16	0.29	9.13	9.07	0.40	0.93
	rSVD84-01	9.08	9.01	0.30	0.44	8.59	8.47	0.47	0.75	9.20	8.99	0.90	0.79
Calinski et al. (2009b) Winter rye 18×15	ABB	34.72	33.83	76.10	11.78	33.62	33.40	101.82	13.74	39.03	40.16	82.96	13.45
	MICE+PMM	44.05	38.52	486.16	24.72	48.50	47.15	480.74	27.24	51.67	47.78	338.94	25.93
	MICE+RF	27.41	27.16	22.54	5.55	27.70	27.58	34.10	7.11	62.84	62.93	185.40	18.02
	rSVD01	16.49	16.28	0.52	0.30	16.55	16.31	3.90	2.26	17.31	16.24	14.09	3.59
	rSVD84	16.07	16.01	0.15	0.10	16.34	15.56	2.14	1.17	16.73	16.36	4.26	2.30
	rSVD84-01	16.40	16.35	0.27	0.39	16.35	15.84	2.48	1.96	16.52	15.85	5.72	3.15

Table 20. P_e by percentage and method — Mechanism **MCAR** (REP 6–10)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	21.67	19.34	100.14	12.11	19.88	18.48	79.73	11.17	22.15	21.20	36.56	8.96
	MICE+PMM	25.95	24.29	371.62	24.95	22.78	20.36	174.44	15.28	23.73	21.82	84.05	14.71
	MICE+RF	11.19	11.15	4.76	3.08	13.97	13.87	6.22	3.00	16.10	15.74	9.50	4.11
	rSVD01	11.48	10.61	12.97	4.13	11.81	11.02	13.23	3.76	13.99	12.18	25.91	5.05
	rSVD84	10.13	9.68	6.26	2.66	10.35	9.68	4.57	1.79	11.30	10.86	5.12	2.88
	rSVD84-01	10.75	10.13	8.57	3.35	10.94	10.00	11.81	2.35	12.42	11.93	10.05	3.92
Yan et al. (2007) Wheat 18×9	ABB	2.12	1.98	0.68	0.95	2.51	2.43	0.61	0.85	2.13	2.11	0.38	0.76
	MICE+PMM	2.65	2.59	2.61	2.77	2.90	2.93	1.46	1.33	2.27	2.20	0.69	1.08
	MICE+RF	1.73	1.75	0.11	0.44	2.37	2.42	0.28	0.70	1.22	1.24	0.05	0.28
	rSVD01	1.07	1.06	0.05	0.26	1.12	1.10	0.03	0.16	1.16	1.14	0.04	0.21
	rSVD84	1.06	1.03	0.06	0.27	1.09	1.06	0.02	0.17	1.09	1.07	0.01	0.15
	rSVD84-01	1.05	1.04	0.05	0.21	1.11	1.09	0.03	0.16	1.15	1.11	0.03	0.23
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	4.30	4.18	0.68	1.04	4.20	4.01	0.47	0.79	4.33	4.22	0.69	0.90
	MICE+PMM	4.95	4.72	3.97	2.56	5.08	4.49	3.53	2.35	4.96	4.62	2.05	2.05
	MICE+RF	4.31	4.42	0.74	1.07	6.87	7.07	2.05	1.94	6.35	6.31	1.67	1.48
	rSVD01	3.73	3.69	0.13	0.49	4.01	3.90	0.23	0.42	4.35	4.16	0.34	0.72
	rSVD84	3.58	3.59	0.11	0.43	3.78	3.78	0.07	0.36	3.93	3.84	0.12	0.30
	rSVD84-01	3.65	3.67	0.13	0.52	3.86	3.85	0.07	0.38	4.08	4.07	0.15	0.43
Calinski et al. (2009a) Pea 18×9	ABB	13.62	13.62	27.52	6.77	15.24	14.39	22.84	7.12	13.75	13.21	16.56	6.15
	MICE+PMM	19.41	15.45	201.91	15.75	18.87	18.06	66.56	12.53	15.24	14.69	37.95	7.83
	MICE+RF	77.87	78.71	293.63	21.89	33.86	33.89	49.95	10.08	18.68	18.37	14.61	5.01
	rSVD01	10.55	10.55	3.86	2.62	11.05	11.08	2.63	1.82	11.16	10.98	1.87	1.73
	rSVD84	10.38	10.41	3.88	2.63	10.77	10.68	2.26	1.86	10.83	10.77	1.45	1.58
	rSVD84-01	10.48	10.39	3.81	2.73	10.93	10.99	2.62	1.99	11.10	10.91	1.79	1.60
Calinski et al. (2009b) Winter rye 18×15	ABB	32.35	31.77	101.83	11.09	31.35	31.29	51.77	8.99	36.26	33.91	81.87	8.96
	MICE+PMM	42.89	42.64	396.36	22.78	38.52	38.38	192.44	19.42	41.68	41.10	99.34	11.59
	MICE+RF	103.44	103.45	345.16	25.19	30.82	30.62	40.47	8.48	40.00	39.17	73.87	11.31
	rSVD01	12.86	12.97	4.21	2.54	12.94	13.04	1.78	1.98	13.13	13.32	1.96	1.83
	rSVD84	12.99	12.96	4.35	2.61	13.06	12.86	2.06	2.05	13.29	13.35	1.66	1.97
	rSVD84-01	12.85	12.72	4.32	2.46	12.93	12.82	1.96	2.04	13.09	13.19	1.69	1.78

Table 21. P_e by percentage and method — Mechanism **MNAR** (REP 6–10)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	29.34	28.92	43.07	9.70	25.29	23.71	49.07	8.68	26.86	25.95	47.14	9.12
	MICE+PMM	36.20	31.06	490.93	27.57	25.38	22.99	95.64	13.03	28.80	26.29	143.09	13.64
	MICE+RF	57.35	59.96	127.82	16.47	17.56	17.56	15.74	5.64	21.83	22.06	19.24	5.76
	rSVD01	22.99	21.36	65.80	10.01	29.97	29.65	1.45	1.31	28.44	27.24	10.43	4.28
	rSVD84	18.81	17.90	25.41	6.32	14.44	13.10	24.16	4.77	15.84	15.60	11.84	4.67
	rSVD84-01	19.99	18.83	33.63	7.67	15.16	14.81	6.33	2.30	16.56	16.07	11.21	3.72
Yan et al. (2007) Wheat 18×9	ABB	2.95	2.84	0.37	0.71	3.21	3.14	0.33	0.72	2.81	2.76	0.23	0.67
	MICE+PMM	3.67	3.25	3.24	2.07	3.71	3.91	0.90	1.73	2.73	2.72	0.47	0.99
	MICE+RF	1.82	1.81	0.15	0.55	4.63	4.74	0.73	1.05	3.59	3.51	0.42	0.86
	rSVD01	2.12	2.01	0.45	0.74	2.43	2.39	0.02	0.14	2.57	2.54	0.01	0.12
	rSVD84	1.84	1.73	0.19	0.47	1.82	1.73	0.27	0.57	2.02	1.93	0.12	0.43
	rSVD84-01	2.14	2.04	0.32	0.77	1.20	1.15	0.05	0.22	1.81	1.77	0.02	0.11
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	6.10	5.96	0.47	0.84	5.69	5.63	0.28	0.73	5.68	5.53	0.36	0.72
	MICE+PMM	6.13	5.69	2.75	1.77	5.95	5.68	2.76	2.15	5.68	5.33	2.82	1.55
	MICE+RF	6.27	6.26	1.35	1.53	9.60	9.79	3.61	2.89	4.45	4.48	0.82	1.28
	rSVD01	7.14	7.12	0.15	0.63	8.74	8.67	0.09	0.36	10.66	10.61	0.02	0.18
	rSVD84	4.74	4.68	0.25	0.58	4.74	4.63	0.31	0.69	5.58	5.50	0.16	0.55
	rSVD84-01	4.92	4.94	0.18	0.50	5.16	5.05	0.15	0.44	6.27	6.25	0.04	0.24
Calinski et al. (2009a) Pea 18×9	ABB	17.10	16.94	11.31	4.55	19.20	18.80	18.05	5.93	16.31	15.96	11.24	4.30
	MICE+PMM	21.94	21.55	75.97	12.37	21.60	21.04	59.78	9.17	16.31	16.22	25.18	8.04
	MICE+RF	21.60	21.53	16.94	5.04	11.20	11.38	3.79	2.80	16.35	16.16	9.74	3.80
	rSVD01	13.29	13.15	0.50	0.85	23.61	23.45	0.21	0.39	20.88	20.73	0.35	0.76
	rSVD84	13.12	12.91	0.46	0.64	12.66	12.22	1.93	1.74	12.56	12.35	0.73	0.85
	rSVD84-01	13.03	12.93	0.42	0.72	14.40	14.45	1.05	1.52	13.67	13.62	0.50	0.89
Calinski et al. (2009b) Winter rye 18×15	ABB	37.39	36.92	50.40	8.55	35.64	34.64	29.00	6.68	40.90	39.51	41.82	7.13
	MICE+PMM	46.49	44.02	364.77	23.30	51.01	49.65	194.20	16.91	50.92	49.08	157.18	11.86
	MICE+RF	78.26	79.38	261.89	23.05	48.11	47.98	74.52	10.63	45.81	45.32	92.55	12.37
	rSVD01	15.41	15.03	1.57	0.33	18.23	17.64	7.68	3.42	38.81	36.07	49.95	8.95
	rSVD84	15.22	15.16	0.07	0.13	17.69	16.27	8.86	1.75	25.39	23.03	72.15	9.60
	rSVD84-01	15.19	15.04	0.52	0.44	18.01	17.25	6.70	2.86	26.17	23.92	71.19	8.33

Table 22. P_e by percentage and method — Mechanism **MAR** (REP 11–15)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	1.42e+03	1.35e+03	1.24e+05	460.29	1.48e+03	1.44e+03	1.76e+05	519.87	1.70e+03	1.70e+03	2.35e+05	545.91
	MICE+PMM	2.21e+03	2.22e+03	4.91e+05	662.63	1.92e+03	1.78e+03	3.83e+05	626.93	2.02e+03	1.94e+03	3.63e+05	585.92
	MICE+RF	2.66e+03	2.69e+03	3.15e+05	761.91	1.35e+03	1.36e+03	7.11e+04	379.97	2.07e+03	2.06e+03	1.74e+05	548.54
	rSVD01	1.41e+03	1.41e+03	514.59	35.03	1.42e+03	1.42e+03	252.39	21.60	1.39e+03	1.39e+03	262.43	24.98
	rSVD84	1.39e+03	1.39e+03	37.75	8.65	1.42e+03	1.42e+03	38.72	7.33	1.40e+03	1.40e+03	36.63	7.91
	rSVD84-01	1.42e+03	1.42e+03	407.37	26.13	1.42e+03	1.42e+03	344.82	28.88	1.39e+03	1.39e+03	291.72	21.04
Filho et al. (2008) Cotton 17×23	ABB	2.88	2.81	0.98	1.21	2.86	2.63	0.90	1.44	2.68	2.38	1.21	1.16
	MICE+PMM	3.80	3.70	3.28	2.26	3.42	2.97	2.40	1.83	3.04	2.91	2.01	1.69
	MICE+RF	5.65	5.65	1.43	1.48	2.58	2.64	0.31	0.72	4.47	4.43	0.71	1.03
	rSVD01	1.35	1.35	0.00	0.02	1.29	1.29	0.00	0.02	1.28	1.28	0.00	0.02
	rSVD84	1.33	1.33	0.00	0.01	1.28	1.28	0.00	0.01	1.30	1.29	0.00	0.01
	rSVD84-01	1.35	1.35	0.00	0.02	1.29	1.29	0.00	0.03	1.28	1.28	0.00	0.02
Flores et al. (1998) Faba beans 15×12	ABB	2.28e+03	2.22e+03	6.86e+05	997.80	2.75e+03	2.60e+03	6.71e+05	1.04e+03	2.28e+03	2.15e+03	4.93e+05	1.04e+03
	MICE+PMM	2.96e+03	2.62e+03	1.32e+06	1.63e+03	3.09e+03	3.02e+03	1.20e+06	1.19e+03	2.19e+03	2.10e+03	4.93e+05	789.40
	MICE+RF	2.88e+03	2.90e+03	3.54e+05	761.98	1.99e+03	1.93e+03	1.77e+05	537.92	2.58e+03	2.59e+03	2.11e+05	593.94
	rSVD01	1.36e+03	1.30e+03	2.38e+04	79.99	1.26e+03	1.24e+03	2.07e+04	170.58	1.39e+03	1.31e+03	1.15e+05	280.17
	rSVD84	1.35e+03	1.28e+03	2.19e+04	98.87	1.25e+03	1.19e+03	2.68e+04	155.02	1.28e+03	1.22e+03	6.79e+04	176.86
	rSVD84-01	1.35e+03	1.29e+03	2.41e+04	63.26	1.24e+03	1.20e+03	2.25e+04	180.46	1.31e+03	1.21e+03	8.05e+04	254.96
Mattos et al. (2013) Sugarcane 22×5	ABB	5.91	5.45	7.81	3.41	6.49	5.67	13.46	5.18	6.53	6.17	14.07	4.62
	MICE+PMM	8.37	6.38	53.12	7.95	9.84	7.47	104.60	11.58	8.35	5.93	41.81	7.08
	MICE+RF	3.33	3.42	0.65	0.93	14.02	13.72	7.61	3.68	1.64	1.65	0.09	0.42
	rSVD01	6.86	6.44	6.70	3.29	6.24	5.63	4.90	2.63	6.74	6.02	5.27	3.30
	rSVD84	5.67	5.47	3.15	2.82	5.61	5.43	4.01	2.15	5.43	5.14	2.17	2.14
	rSVD84-01	5.98	5.53	3.61	2.50	5.71	5.37	3.57	2.13	5.86	5.39	3.43	2.49
Rad et al. (2013) Wheat 36×6	ABB	9.21	9.00	4.83	2.72	8.88	8.53	4.20	2.66	8.94	8.86	4.48	2.65
	MICE+PMM	12.07	9.83	59.14	11.39	13.07	10.02	59.10	11.35	10.48	9.12	26.97	6.95
	MICE+RF	8.69	8.70	3.12	2.14	7.56	7.64	2.43	1.83	6.48	6.51	1.84	1.78
	rSVD01	5.94	5.74	1.45	1.27	5.20	5.15	0.13	0.47	5.40	5.29	0.31	0.78
	rSVD84	6.29	6.01	3.02	2.25	5.10	4.96	0.40	0.92	5.14	5.06	0.40	0.76
	rSVD84-01	4.56	4.36	1.45	1.60	4.61	4.50	0.15	0.42	4.91	4.83	0.33	0.57

Table 23. P_e by percentage and method — Mechanism **MCAR** (REP 11–15)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	1.37e+03	1.30e+03	2.27e+05	519.85	1.39e+03	1.33e+03	2.11e+05	377.72	1.62e+03	1.60e+03	2.19e+05	473.89
	MICE+PMM	1.93e+03	1.80e+03	5.25e+05	860.28	1.83e+03	1.66e+03	6.29e+05	707.76	1.87e+03	1.80e+03	2.53e+05	531.96
	MICE+RF	2.35e+03	2.40e+03	2.39e+05	690.96	792.63	781.77	2.46e+04	206.52	1.98e+03	2.02e+03	1.94e+05	587.98
	rSVD01	1.35e+03	1.34e+03	1.29e+04	141.33	1.34e+03	1.34e+03	8.43e+03	133.99	1.34e+03	1.34e+03	4.23e+03	77.23
	rSVD84	1.35e+03	1.35e+03	1.13e+04	163.08	1.34e+03	1.34e+03	8.88e+03	129.67	1.33e+03	1.33e+03	4.32e+03	82.91
	rSVD84-01	1.35e+03	1.34e+03	1.32e+04	193.93	1.34e+03	1.34e+03	8.61e+03	133.29	1.34e+03	1.33e+03	4.57e+03	83.80
Filho et al. (2008) Cotton 17×23	ABB	2.38	2.19	1.18	1.33	2.38	2.24	0.80	1.03	2.34	2.20	0.69	0.78
	MICE+PMM	3.79	3.41	4.82	2.07	3.31	3.22	1.69	1.40	2.71	2.52	0.88	1.17
	MICE+RF	3.00	3.09	0.31	0.73	4.36	4.43	0.80	1.09	3.98	4.10	0.76	1.20
	rSVD01	1.23	1.23	0.01	0.13	1.22	1.22	0.00	0.07	1.22	1.22	0.00	0.08
	rSVD84	1.23	1.23	0.01	0.14	1.21	1.21	0.00	0.07	1.22	1.22	0.00	0.07
	rSVD84-01	1.23	1.23	0.01	0.14	1.21	1.21	0.00	0.07	1.22	1.22	0.00	0.08
Flores et al. (1998) Faba beans 15×12	ABB	1.87e+03	1.79e+03	5.79e+05	933.64	2.15e+03	2.03e+03	5.42e+05	805.69	2.55e+03	2.40e+03	5.89e+05	800.14
	MICE+PMM	2.20e+03	1.82e+03	1.90e+06	1.96e+03	2.53e+03	2.61e+03	1.05e+06	1.32e+03	2.68e+03	2.64e+03	8.57e+05	1.04e+03
	MICE+RF	2.31e+03	2.34e+03	2.52e+05	718.59	1.27e+03	1.32e+03	8.96e+04	405.27	3.01e+03	3.00e+03	3.70e+05	925.59
	rSVD01	967.06	986.45	3.29e+04	254.95	1.02e+03	1.00e+03	1.46e+04	136.84	1.05e+03	1.04e+03	2.09e+04	161.74
	rSVD84	962.00	985.12	2.82e+04	232.31	1.01e+03	994.31	1.44e+04	149.47	1.05e+03	1.03e+03	1.36e+04	132.04
	rSVD84-01	961.27	965.49	3.02e+04	239.25	1.01e+03	1.00e+03	1.56e+04	160.15	1.04e+03	1.04e+03	1.50e+04	126.96
Mattos et al. (2013) Sugarcane 22×5	ABB	6.34	5.87	7.11	3.50	7.03	6.51	6.25	3.59	8.88	8.46	8.86	3.79
	MICE+PMM	6.86	4.92	27.19	5.82	8.18	7.54	16.70	5.16	9.98	8.84	28.92	5.85
	MICE+RF	9.74	9.84	3.17	2.39	7.56	7.61	2.14	2.04	3.94	3.93	0.72	1.21
	rSVD01	4.98	4.75	2.56	1.44	5.75	5.33	2.91	1.94	7.19	7.07	3.46	2.16
	rSVD84	4.09	3.94	1.23	1.28	4.71	4.48	1.18	1.07	5.73	5.23	2.44	2.15
	rSVD84-01	4.55	4.19	2.38	1.38	4.95	4.74	1.53	1.32	5.99	5.56	2.88	2.05
Rad et al. (2013) Wheat 36×6	ABB	7.05	6.68	4.04	2.10	6.20	6.10	1.70	1.60	6.95	6.74	2.23	1.93
	MICE+PMM	7.65	6.54	19.62	4.65	7.27	7.00	9.74	4.19	7.60	7.31	7.28	3.12
	MICE+RF	5.58	5.47	1.46	1.50	8.32	8.19	2.40	1.84	4.31	4.28	0.74	1.32
	rSVD01	4.24	4.06	1.14	1.47	4.61	4.48	0.74	1.21	5.34	5.22	1.31	1.47
	rSVD84	3.69	3.64	0.48	0.89	4.00	3.83	0.62	0.90	4.54	4.41	0.75	1.09
	rSVD84-01	3.98	3.92	0.65	0.99	4.16	4.05	0.59	1.05	4.66	4.63	0.94	1.30

Table 24. P_e by percentage and method — Mechanism **MNAR** (REP 11–15)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	1.76e+03	1.71e+03	1.16e+05	398.40	1.65e+03	1.64e+03	8.22e+04	368.03	1.74e+03	1.74e+03	9.76e+04	313.11
	MICE+PMM	2.53e+03	2.47e+03	3.34e+05	700.10	2.18e+03	2.05e+03	2.70e+05	486.91	2.06e+03	1.98e+03	1.95e+05	530.85
	MICE+RF	2.38e+03	2.41e+03	2.34e+05	627.00	1.89e+03	1.93e+03	1.74e+05	521.85	2.31e+03	2.32e+03	2.33e+05	594.55
	rSVD01	1.30e+03	1.30e+03	148.85	16.32	1.30e+03	1.30e+03	162.23	16.24	1.33e+03	1.33e+03	460.23	33.56
	rSVD84	1.30e+03	1.30e+03	74.97	12.17	1.29e+03	1.29e+03	37.04	7.99	1.33e+03	1.32e+03	56.52	9.97
	rSVD84-01	1.31e+03	1.31e+03	326.62	24.52	1.30e+03	1.30e+03	246.16	20.97	1.32e+03	1.32e+03	273.44	18.25
Filho et al. (2008) Cotton 17×23	ABB	2.90	2.72	0.73	0.95	2.92	2.86	0.47	1.05	3.01	2.92	0.45	0.91
	MICE+PMM	3.94	3.96	1.18	1.48	3.87	3.79	1.52	1.28	3.45	3.47	0.63	1.00
	MICE+RF	5.50	5.61	1.07	1.45	2.96	2.92	0.32	0.82	2.90	2.89	0.32	0.91
	rSVD01	1.36	1.36	0.00	0.04	1.27	1.26	0.00	0.02	1.27	1.27	0.00	0.03
	rSVD84	1.36	1.36	0.00	0.01	1.28	1.28	0.00	0.01	1.29	1.28	0.00	0.02
	rSVD84-01	1.34	1.34	0.00	0.04	1.27	1.27	0.00	0.02	1.28	1.27	0.00	0.04
Flores et al. (1998) Faba beans 15×12	ABB	2.58e+03	2.48e+03	3.72e+05	704.29	2.69e+03	2.64e+03	3.45e+05	806.00	2.43e+03	2.35e+03	3.58e+05	895.64
	MICE+PMM	2.69e+03	2.55e+03	1.11e+06	1.47e+03	2.86e+03	2.78e+03	1.07e+06	1.12e+03	2.46e+03	2.39e+03	4.80e+05	786.63
	MICE+RF	3.09e+03	3.09e+03	3.91e+05	842.81	3.99e+03	4.00e+03	8.10e+05	1.40e+03	2.81e+03	2.68e+03	3.22e+05	876.49
	rSVD01	1.48e+03	1.38e+03	7.65e+04	385.96	1.40e+03	1.27e+03	7.22e+04	234.35	1.53e+03	1.35e+03	3.20e+05	424.07
	rSVD84	1.39e+03	1.29e+03	4.72e+04	168.01	1.37e+03	1.27e+03	5.19e+04	160.67	1.36e+03	1.29e+03	6.57e+04	271.06
	rSVD84-01	1.43e+03	1.33e+03	6.89e+04	227.97	1.41e+03	1.37e+03	3.24e+04	268.00	1.40e+03	1.31e+03	1.01e+05	288.94
Mattos et al. (2013) Sugarcane 22×5	ABB	8.97	8.76	7.08	3.40	9.68	9.47	7.37	3.78	10.32	9.85	8.66	3.79
	MICE+PMM	9.30	6.77	39.42	5.55	10.33	9.47	15.82	5.92	9.86	9.10	14.12	4.12
	MICE+RF	6.39	6.29	1.46	1.73	5.71	5.75	1.45	1.57	5.95	6.07	1.99	1.79
	rSVD01	7.84	6.92	8.79	2.87	8.44	7.78	3.86	1.40	8.79	8.53	0.78	1.00
	rSVD84	6.38	6.05	3.54	1.95	6.36	6.26	2.60	2.17	6.81	6.66	4.07	2.60
	rSVD84-01	6.88	6.49	3.36	2.28	5.24	5.11	1.44	1.64	4.99	4.78	1.30	1.36
Rad et al. (2013) Wheat 36×6	ABB	9.60	9.54	2.04	1.96	8.86	8.64	2.15	1.53	9.13	8.89	1.85	1.32
	MICE+PMM	12.31	11.91	18.08	4.72	9.90	10.20	12.62	5.63	9.48	9.12	11.58	3.93
	MICE+RF	10.11	10.12	4.39	2.51	12.60	12.67	4.28	2.42	11.57	11.60	7.73	3.83
	rSVD01	5.93	5.81	0.58	1.01	6.06	5.99	0.62	1.03	6.36	6.19	0.81	0.87
	rSVD84	7.30	7.05	3.57	2.33	6.01	5.90	0.62	0.96	5.76	5.58	0.38	0.57
	rSVD84-01	4.82	4.63	0.84	1.21	5.60	5.38	0.66	0.69	6.02	5.92	0.46	0.93

Table 25. P_e by percentage and method — Mechanism **MAR** (REP 16–17)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	10.07	9.70	11.09	3.79	10.35	10.21	12.51	5.21	12.38	12.15	22.14	6.14
	MICE+PMM	12.14	11.95	54.56	10.18	12.47	10.78	61.64	11.68	12.84	9.72	66.40	10.13
	MICE+RF	8.49	8.70	2.99	2.21	7.50	7.62	2.31	2.18	5.68	5.69	0.98	1.28
	rSVD01	6.62	5.91	33.40	3.02	5.76	5.27	6.68	2.85	4.54	4.34	1.59	1.72
	rSVD84	6.21	5.88	3.87	2.53	5.76	5.47	3.12	2.29	3.87	3.72	1.05	1.37
	rSVD84-01	5.38	4.94	5.34	2.88	4.89	4.44	3.19	1.77	4.22	3.91	1.55	1.55
Zobel et al. (1988) Soybean 7×35	ABB	3.47e+03	3.41e+03	7.90e+05	1.22e+03	3.31e+03	3.29e+03	7.86e+05	1.03e+03	3.84e+03	3.80e+03	1.27e+06	1.21e+03
	MICE+PMM	4.35e+03	3.66e+03	7.47e+06	3.33e+03	4.15e+03	3.69e+03	5.09e+06	3.38e+03	4.85e+03	4.62e+03	6.71e+06	3.67e+03
	MICE+RF	7.32e+03	7.33e+03	2.15e+06	1.96e+03	2.16e+03	2.11e+03	1.89e+05	600.49	1.37e+04	1.36e+04	6.65e+06	2.81e+03
	rSVD01	2.28e+03	2.20e+03	4.72e+05	858.27	2.49e+03	2.51e+03	3.45e+05	872.83	2.37e+03	2.28e+03	9.42e+05	772.95
	rSVD84	1.99e+03	1.95e+03	2.04e+05	505.79	2.14e+03	2.00e+03	4.00e+05	628.66	2.23e+03	2.19e+03	2.47e+05	687.93
	rSVD84-01	2.22e+03	2.12e+03	3.32e+05	809.35	2.21e+03	2.20e+03	3.26e+05	753.75	2.03e+03	2.00e+03	3.15e+05	676.02

Table 26. P_e by percentage and method — Mechanism **MCAR** (REP 16–17)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	7.33	6.61	10.12	4.65	8.56	8.50	7.82	3.80	9.94	9.54	5.76	2.51
	MICE+PMM	7.53	6.19	37.38	6.42	7.38	6.97	17.22	6.56	9.05	8.26	17.63	5.52
	MICE+RF	1.83	1.81	0.15	0.48	7.40	7.31	2.04	2.15	13.48	13.64	8.44	3.65
	rSVD01	3.74	3.55	1.95	1.67	4.94	4.94	2.43	2.19	6.17	5.60	23.00	2.40
	rSVD84	3.10	2.99	0.82	1.12	4.02	3.70	1.84	1.53	4.47	4.27	2.45	1.97
	rSVD84-01	3.49	3.11	1.96	1.27	4.69	4.38	4.22	1.80	5.05	4.86	2.56	2.05
Zobel et al. (1988) Soybean 7×35	ABB	2.67e+03	2.54e+03	5.08e+05	918.71	2.84e+03	2.75e+03	3.13e+05	758.33	2.71e+03	2.66e+03	2.99e+05	774.30
	MICE+PMM	2.90e+03	2.57e+03	1.96e+06	1.97e+03	3.25e+03	3.12e+03	1.69e+06	1.72e+03	3.01e+03	2.89e+03	1.15e+06	1.28e+03
	MICE+RF	2.28e+03	2.28e+03	2.00e+05	581.11	6.31e+03	6.37e+03	1.47e+06	1.73e+03	2.93e+03	2.96e+03	3.78e+05	810.96
	rSVD01	1.17e+03	1.16e+03	6.50e+04	264.46	1.45e+03	1.40e+03	1.55e+05	369.14	1.64e+03	1.51e+03	5.96e+05	475.34
	rSVD84	992.69	944.06	4.26e+04	283.87	1.21e+03	1.14e+03	7.05e+04	283.43	1.30e+03	1.25e+03	8.61e+04	278.48
	rSVD84-01	1.14e+03	1.17e+03	7.57e+04	397.34	1.26e+03	1.21e+03	7.94e+04	359.88	1.43e+03	1.39e+03	1.40e+05	429.03

Table 27. P_e (top) by percentage and method — Mechanism **MNAR** (REP 16–17)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	9.50	9.49	4.90	3.09	10.51	10.42	3.05	2.27	11.97	11.59	4.37	2.03
	MICE+PMM	10.71	10.16	38.77	9.06	10.36	10.10	9.23	3.24	13.03	12.44	19.62	4.81
	MICE+RF	7.79	7.75	2.20	2.01	11.99	11.70	6.93	3.63	9.12	9.27	3.44	2.70
	rSVD01	8.33	5.68	315.53	3.15	3.77	3.33	1.59	1.97	4.95	4.67	2.26	1.89
	rSVD84	6.42	6.03	3.79	2.19	3.39	3.22	0.92	0.76	4.80	4.90	1.37	1.62
	rSVD84-01	5.35	5.08	3.84	2.48	3.60	3.35	1.89	1.02	4.54	4.14	1.91	1.42
Zobel et al. (1988) Soybean 7×35	ABB	3.47e+03	3.47e+03	3.17e+05	602.75	3.63e+03	3.54e+03	2.25e+05	533.08	3.41e+03	3.39e+03	1.77e+05	523.44
	MICE+PMM	3.98e+03	3.74e+03	2.70e+06	2.87e+03	3.90e+03	3.92e+03	1.78e+06	2.10e+03	3.79e+03	3.68e+03	1.15e+06	1.51e+03
	MICE+RF	2.44e+03	2.40e+03	2.88e+05	762.14	5.97e+03	5.81e+03	1.65e+06	1.68e+03	3.22e+03	3.21e+03	3.50e+05	811.41
	rSVD01	2.86e+03	2.67e+03	7.33e+05	897.60	3.47e+03	3.18e+03	1.61e+06	988.89	2.60e+03	2.41e+03	7.12e+05	735.67
	rSVD84	2.23e+03	2.18e+03	1.59e+05	495.96	2.45e+03	2.30e+03	3.35e+05	690.12	2.12e+03	2.03e+03	2.21e+05	488.29
	rSVD84-01	2.59e+03	2.43e+03	4.64e+05	917.82	2.82e+03	2.64e+03	5.94e+05	854.19	2.18e+03	2.06e+03	2.23e+05	547.42

6 Conclusions

At the outset, the objectives of this study were to develop and validate a multiple imputation method based on robust variants of Singular Value Decomposition (rSVD84 and rSVD01) to effectively handle missing data under different mechanisms (MCAR, MAR, and MNAR) and in the presence of various levels of outliers. These objectives were addressed by proposing and evaluating the rSVD84, rSVD01, and rSVD84-01 models, comparing them with benchmark methods such as ABB and MICE, and testing their performance through simulation studies and real-world datasets. In simulated scenarios with moderate contamination, the proposed SVD models achieved competitive performance, although they did not surpass mice in efficiency; nevertheless, they outperformed ABB, with the advantage of not relying on a specific data distribution.

In extreme scenarios involving outliers, SVD-based models demonstrated a clear advantage, maintaining stable performance even as data loss increased, with the rSVD84 model showing particular robustness. Across all scenarios, the Missing Not at Random (MNAR) mechanism proved most challenging, leading to lower predictive accuracy and greater variance between imputations. When applied to real-world datasets, the SVD models displayed inconsistent results with multivariate data, performing considerably worse than mice and ABB, likely due to the bootstrap approach being more suitable for GxE (Genotype by Environment) type data. In very small datasets, such as the 5×4 dataset from Forkman and Piepho (2014), the SVD models were unable to fully leverage their randomization mechanism, resulting in unstable imputations.

Although SVD models performed exceptionally well on some datasets (Yan et al., 2007) with low prediction errors (P_e), they generally approximated the performance of benchmark methods and rarely achieved the best P_e scores. Notably,the imputation by ABB frequently outperformed the SVD models and, in some cases, even mice, indicating that patterns observed in simulated data do not always translate to real-world datasets. A consistent observation across both simulated and real data is that, despite maintaining good predictive accuracy in the presence of extreme outliers, the variance between imputations for SVD models increases substantially, highlighting that the uncertainty associated with their imputations grows considerably and revealing an inherent limitation in imputing matrices with these characteristics.

7 Future Work

The results of this study open several promising lines of research for further exploration. One potential direction is the integration of multiple imputation with robust methods for handling outliers and missing data, but, particularly by exploring alternative randomization strategies beyond the current bidirectional bootstrap approach to enhance the stability and accuracy of imputations. Another avenue involves testing the proposed methods with different types of outliers, both in simulation studies and real-world datasets, to better understand their robustness across varied conditions or even with more percentage of contamination. Further investigation into robust imputation techniques is also recommended to specifically address the challenges posed by the MNAR mechanism, which consistently reduced predictive performance in this study. Finally, evaluating other robust SVD variants, such as rSVD17 (García-Peña et al., 2021), may provide additional improvements in accuracy and reliability, offering new opportunities to refine multiple imputation approaches in complex data scenarios.

8 References

References

[1] Arciniegas-Alarcón, S., dos Santos Dias, C. T. S., and García-Peña, M.: Imputação múltipla livre de distribuição em tabelas incompletas de dupla entrada. Available at https://www.scielo.br/j/pab/a/xBmT4VFN7CxFKwLBdMJbJvj/abstract/?lang=en

[2] Arciniegas-Alarcón, S., García-Peña, M., and Canas Rodrigues, P.: New multiple imputation methods for genotype-by-environment data that combine singular value decomposition and Jackknife resampling or weighting schemes. Computers and Electronics in Agriculture **176** (2020) 105617. https://doi.org/10.1016/j.compag.2020.105617

[3] Arciniegas-Alarcón, S., García-Peña, M., and Krzanowski, W. J.: Cross-Validation for Lower Rank Matrices Containing Outliers. Applied System Innovation **5**(4) (2022) 69. https://doi.org/10.3390/asi5040069

[4] Arciniegas-Alarcón, S., García-Peña, M., Krzanowski, W. J., and Rengifo, C.: Cross-validation to select the optimum rank for a reduced-rank approximation to multivariate data. Journal of Crop Improvement **38**(4) (2024) 344–367. https://doi.org/10.1080/15427528.2024.2349610

[5] Baraldi, A. N., and Enders, C. K.: An introduction to modern missing data analyses. Arizona State University. https://www.uvm.edu/~statdhtx/StatPages/Missing_Data/Baraldi-Enders.html

[6] Bello, A. L.: A bootstrap method for using imputation techniques for data with missing values. Biometrical Journal **36**(4) (1994) 453–464.

[7] Boukerche, A., Zheng, L., and Alfandi, O.: Outlier Detection: Methods, Models, and Classification. ACM Computing Surveys **53**(3) (2021) Article 55. https://doi.org/10.1145/3381028

[8] Brahim, J. G., Chu, H., and Chen, M.-H.: Missing Data in Clinical Studies: Issues and Methods. Journal of Clinical Oncology **30**(26) (2012) 3297–3303. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3948388/

[9] Bro, R., and Smilde, A. K.: Principal component analysis. Analytical Methods **6**(9) (2014) 2812–2831. https://doi.org/10.1039/C3AY41907J

[10] Caliński, T., Czajka, S., Kaczmarek, Z., et al.: Analyzing the genotype-by-environment interactions under a randomization-derived mixed model. Journal of Agricultural, Biological, and Environmental Statistics **14** (2009) 224–241. https://doi.org/10.1198/jabes.2009.0014

[Caliński et al., 2009a] Caliński, T., Czajka, S., Kaczmarek, Z., Krajewski, P., and Pilarczyk, W.: Analyzing the genotype-by-environment interactions under a randomization-derived mixed model. Journal of Agricultural, Biological, and Environmental Statistics **14**(2) (2009) 224–241. https://doi.org/10.1198/jabes.2009.0014

[Caliński et al., 2009b] Caliński, T., Czajka, S., Kaczmarek, Z., Krajewski, P., and Pilarczyk, W.: A mixed model analysis of variance for multi-environment variety trials. Statistical Papers **50**(4) (2009) 735–759. https://doi.org/10.1007/s00362-009-0249-1

[Candes et al., 2011] Candes, E. J., Li, X. D., Ma, Y., and Wright, J.: Robust Principal Component Analysis? Journal of the ACM **58** (2011) Article 11. https://doi.org/10.1145/1970392.1970395

[Carpenter and Kenward, 2005] Carpenter, J., and Kenward, M.: Missing value jargon. Department of Medical Statistics, London School of Hygiene & Tropical Medicine. https://missingdata.lshtm.ac.uk

[Cornelius et al., 1996] Cornelius, P. L., Crossa, J., and Seyedsadr, M. S.: Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. In: Kang, M. S., and Gauch, H. G. (eds.) Genotype-by-Environment Interaction, pp. 199–231. CRC Press, Boca Raton (1996).

[Corsten and Denis, 1990] Corsten, L. C. A., and Denis, J. B.: Structuring interaction in two-way tables by clustering. Biometrics **46**(1) (1990) 207–215. https://doi.org/10.2307/2531644

[Enders, 2023] Enders, C. K.: Missing data: An update on the state of the art. Psychological Methods (2023), advance online publication. https://doi.org/10.1037/met0000563

[Farias, 2005] Farias, F. J. C.: Selection index in upland cotton cultivars. PhD thesis, Universidade de São Paulo (2005).

[Filho et al., 2008] Filho, J. L. S., Morello, C. L., Farias, F. J. C., Lamas, F. M., Pedrosa, M. B., and Ribeiro, J. L.: Comparação de métodos para avaliar a adaptabilidade e estabilidade produtiva em algodoeiro. Pesquisa Agropecuária Brasileira **43**(3) (2008) 349–355. https://doi.org/10.1590/S0100-204X2008000300009

[Flores et al., 1998] Flores, F., Moreno, M. T., and Cubero, J. I.: A comparison of univariate and multivariate methods to analyze G×E interaction. Field Crops Research **56**(3) (1998) 271–286. https://doi.org/10.1016/S0378-4290(97)00095-6

[Forkman and Piepho, 2014] Forkman, J., and Piepho, H. P.: Parametric bootstrap methods for testing multiplicative terms in GGE and AMMI models. Biometrics **70**(3) (2014) 639–647. https://doi.org/10.1111/biom.12162

[Gabriel, 2002] Gabriel, K. R.: Le biplot – outil d'exploration de données multidimensionelles. Journal de la Société Française de Statistique **143** (2002) 5–55.

[Gabriel and Odoroff, 1984] Gabriel, K. R., and Odoroff, L.: Resistant lower rank approximation of matrices. In: Diday, E., et al. (eds.) Data Analysis and Informatics III, pp. 23–30. North-Holland, Amsterdam (1984).

[García-Peña et al., 2021] García-Peña, M., Arciniegas-Alarcón, S., Krzanowski, W. J., and Duarte, D.: Missing-value imputation using the robust singular-value decomposition: Proposals and numerical evaluation. Crop Science **61** (2021) 3288–3300. https://doi.org/10.1002/csc2.20508

[García-Peña et al., 2025] García-Peña, M., Arciniegas-Alarcón, S., and Basford, K. E.: Application of weighted low rank approximations: outlier detection in a data matrix. BMC Research Notes **18** (2025) 234. https://doi.org/10.1186/s13104-025-07284-2

[Hawkins et al., 2001] Hawkins, D. M., Liu, L., and Young, S. S.: Robust singular-value decomposition (Technical Report 122). National Institute of Statistical Sciences. https://www.niss.org/sites/default/files/technicalreports/tr122.pdf

[Heymans and Eekhout, 2019] Heymans, M., and Eekhout, I.: Applied Missing Data Analysis With SPSS and RStudio. First Draft. https://bookdown.org/mwheymans/bookmi/

[Hu and Yang, 2013] Hu, Z., and Yang, R.-C.: Improved Statistical Inference for Graphical Description and Interpretation of Genotype × Environment Interaction. Crop Science (2013). https://www.researchgate.net/publication/277449535_Improved_Statistical_Inference_for_Graphical_Description_and_Interpretation_of_Genotype_Environment_Interaction

[Ibrahim et al., 2005] Ibrahim, J. G., Chen, M. H., Lipsitz, S. R., and Herring, A. H.: Missing-Data Methods for Generalized Linear Models: A Comparative Review. Journal of the American Statistical Association **100**(469) (2005) 332–346. https://doi.org/10.1198/016214504000001844

[Josse and Husson, 2012] Josse, J., and Husson, F.: Handling missing values in exploratory multivariate data analysis methods. Journal de la Société Française de Statistique **153**(2) (2012) 80–99. https://www.researchgate.net/publication/266856067_Handling_missing_values_in_exploratory_multivariate_data_analysis_methods

[Kim et al., 2005] Kim, H., Golub, G. H., and Park, H.: Missing value estimation for DNA microarray gene expression data: Local least squares imputation. Bioinformatics **21**(2) (2005) 187–198. https://sci2s.ugr.es/keel/pdf/algorithm/articulo/2005-Kim-B.pdf

[Krzanowski, 1988] Krzanowski, W. J.: Missing value imputation in multivariate data using the singular-value decomposition of a matrix. Biometrical Letters **25** (1988) 31–39. https://doi.org/10.1007/BF02295279

[Krzanowski, 2007] Krzanowski, W. J.: Statistical Principles and Techniques in Scientific and Social Research. Oxford University Press, Oxford. https://research.ebsco.com/linkprocessor/plink?id=be1c26ca-20c7-3ce0-b2a9-d36b313994e1

[Lavoranti, 2003] Lavoranti, O. J.: Estabilidade e adaptabilidade fenotípica através da reamostragem "Bootstrap" no modelo AMMI. PhD thesis, Universidade de São Paulo (2003). http://www.teses.usp.br

[Li, 2013] Li, C.: Little's test of missing completely at random. The Stata Journal **13**(4) (2013) 795–809. https://journals.sagepub.com/doi/pdf/10.1177/1536867X1301300407

[Li et al., 2019] Li, X., Deng, S., Li, L., and Jiang, Y.: Outlier detection based on robust Mahalanobis distance and its application. Open Journal of Statistics **9**(1) (2019) 1–8. https://doi.org/10.4236/ojs.2019.91002

[Little and Rubin, 2002] Little, R. J. A., and Rubin, D. B.: Statistical Analysis with Missing Data. John Wiley & Sons, New York (2002). https://ebookcentral.proquest.com/lib/bibliojaveriana-ebooks/detail.action?docID=1775204

[Ma and Zhang, 2021] Ma, C., and Zhang, C.: Identifiable generative models for missing not at random data imputation. arXiv preprint (2021). https://arxiv.org/pdf/2110.14708

[Madley-Dowd et al., 2019] Madley-Dowd, P., Hughes, R., Tilling, K., and Heron, J.: The proportion of missing data should not be used to guide decisions on multiple imputation. Journal of Clinical Epidemiology **110** (2019) 63–73. https://www.sciencedirect.com/science/article/pii/S0895435618308710

[Mattos et al., 2013] Mattos, P. H. C., Oliveira, R. A. J., Filho, C. B., Daros, E., and Veríssimo, M. A. A.: Evaluation of sugarcane genotypes and production environments in Paraná by GGE biplot and AMMI analysis. Crop Breeding and Applied Biotechnology **13**(1) (2013) 83–90. https://doi.org/10.1590/S1984-70332013000100010

[Murray, 2018] Murray, J. S.: Multiple imputation: A review of practical and theoretical findings. arXiv preprint (2018). https://arxiv.org/pdf/1801.04058

[Rad et al., 2013] Rad, M. R. N., Kadir, M. A., Rafii, M. Y., Jaafar, H. Z. E., Naghavi, M. R., and Ahmadi, F.: Genotype × environment interaction by AMMI and GGE biplot analysis in three consecutive generations of wheat (Triticum aestivum) under normal and drought stress conditions. Australian Journal of Crop Science **7**(7) (2013) 956–961.

[Rodrigues et al., 2016] Rodrigues, P. C., Monteiro, A., and Lourenço, V. M.: A robust AMMI model for the analysis of genotype-by-environment data. Bioinformatics **32**(1) (2016) 58–66. https://doi.org/10.1093/bioinformatics/btv533

[Rubin, 1987] Rubin, D. B.: Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons, New York (1987). https://doi.org/10.1002/9780470316696

[Schafer and Olsen, 1998] Schafer, J. L., and Olsen, M. K.: Multiple Imputation for Multivariate Missing-Data Problems: A Data Analyst's Perspective. Multivariate Behavioral Research **33**(4) (1998) 545–571. https://doi.org/10.1207/s15327906mbr3304_5

[Shah et al., 2014] Shah, A. D., Bartlett, J. W., Carpenter, J., Nicholas, O., and Hemingway, H.: Comparison of random forest and parametric imputation models for imputing missing data using MICE: a CALIBER study. American Journal of Epidemiology **179**(6) (2014) 764–774. https://doi.org/10.1093/aje/kwt312

[Stekhoven and Bühlmann, 2012] Stekhoven, D. J., and Bühlmann, P.: MissForest—nonparametric missing value imputation for mixed-type data. Bioinformatics **28**(1) (2012) 112–118. https://doi.org/10.1093/bioinformatics/btr597

[Templ, 2023] Templ, M.: Visualization and Imputation of Missing Values with Applications in R. Springer, Cham (2023). https://link.springer.com/book/10.1007/978-3-031-30073-8

[Van Buuren, 2011] Van Buuren, S.: Multiple Imputation of Multilevel Data. Routledge, 181–204 (2011). https://doi.org/10.1201/b11826

[Yan et al., 2007] Yan, W., Kang, M. S., Ma, B., Woods, S., and Cornelius, P. L.: GGE Biplot vs. AMMI Analysis of Genotype-by-Environment Data. Crop Science **47** (2007) 643–653. https://doi.org/10.2135/cropsci2006.06.0374

[Yang, 2007] Yang, R. C.: Mixed model analysis of crossover genotype-environment interactions. Crop Science **47**(3) (2007) 1051–1062. https://doi.org/10.2135/cropsci2006.09.0611

[Zhang, 2021] Zhang, X.: Tutorial: How to Generate Missing Data For Simulation Studies. (2021). https://doi.org/10.31234/osf.io/rq6yb

[Zobel et al., 1988] Zobel, R. W., Wright, M. J., and Gauch, H. G.: Statistical analysis of a yield trial. Agronomy Journal **80**(3) (1988) 388–393. https://doi.org/10.2134/agronj1988.00021962008000030002x

9 Appendices

Method	Mean	Median	Variance	IQR
$p = 10\%$				
mice+pmm	2.8620	2.8303	0.4624	0.8351
mice+rf	2.8855	2.9512	0.7841	1.2029
rSVD84	5.6811	5.6707	0.6705	1.3741
rSVD01-84	5.6838	5.6713	0.6697	1.3818
rSVD01	5.6868	5.6721	0.6691	1.3836
ABB	6.0121	6.0504	0.7125	1.3549
$p = 20\%$				
mice+pmm	3.2334	3.2028	0.5130	0.8316
mice+rf	3.2003	3.0696	0.8987	1.3392
rSVD84	5.3000	5.3070	0.5776	1.0656
rSVD01-84	5.3229	5.3054	0.5855	1.0372
rSVD01	5.4447	5.3631	0.9872	1.0566
ABB	5.7497	5.7140	0.4655	0.9904
$p = 30\%$				
mice+pmm	3.6481	3.6011	0.4379	0.8936
mice+rf	3.7442	3.6455	0.8654	1.1192
rSVD84	5.1090	4.9960	0.5456	0.9336
rSVD01-84	5.1959	5.0403	0.6590	0.8990
rSVD01	5.5231	5.2000	1.9708	1.1669
ABB	5.6143	5.5837	0.4844	0.9878

Table 28. Prediction error (P_e) under MAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
mice+rf	1.9250	1.8837	0.4771	0.7065
mice+pmm	1.9796	1.9773	0.2811	0.6625
rSVD01	4.4372	4.4206	0.1802	0.5251
rSVD01-84	4.4363	4.4243	0.1805	0.5281
rSVD84	4.4355	4.4266	0.1808	0.5312
ABB	4.8453	4.8538	0.3145	0.7751
$p = 20\%$				
mice+rf	2.1058	2.0697	0.2835	0.6635
mice+pmm	2.1013	2.1151	0.1400	0.4727
rSVD84	4.4064	4.3486	0.1261	0.4150
rSVD01-84	4.4238	4.3598	0.1423	0.4360
rSVD01	4.4640	4.3603	0.2901	0.4340
ABB	4.7646	4.7204	0.2451	0.5518
$p = 30\%$				
mice+rf	2.2264	2.0512	0.3974	0.9499
mice+pmm	2.2234	2.2187	0.1300	0.5305
rSVD84	4.4398	4.4426	0.0816	0.4049
rSVD01-84	4.5416	4.4604	0.1601	0.4805
rSVD01	4.7911	4.4637	0.8875	0.4936
ABB	4.8189	4.7741	0.1920	0.5911

Table 29. Prediction error (P_e) under MCAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
mice+rf	4.3884	4.1408	1.3742	1.4350
mice+pmm	4.4710	4.3212	0.6382	1.0215
rSVD84	8.0558	8.0220	0.2126	0.4675
rSVD01-84	8.0726	8.0335	0.2126	0.4699
rSVD01	8.1087	8.0553	0.2476	0.5106
ABB	8.3340	8.3234	0.1702	0.5503
$p = 20\%$				
mice+pmm	5.0733	5.0129	0.4220	1.0660
mice+rf	5.0719	5.0443	1.0990	1.4776
rSVD84	7.5838	7.5589	0.1658	0.5434
rSVD01-84	7.6011	7.5799	0.1656	0.5487
rSVD01	7.7775	7.6645	0.4694	0.4997
ABB	7.8192	7.8000	0.1407	0.4879
$p = 30\%$				
mice+pmm	6.0631	6.0659	0.4437	0.9258
mice+rf	6.1365	6.1365	1.9479	2.1404
rSVD84	7.3928	7.3069	0.5265	0.4720
rSVD01-84	7.4346	7.3338	0.5306	0.4622
rSVD01	7.7758	7.5022	1.2769	0.6373
ABB	7.5663	7.5170	0.1519	0.4885

Table 30. Prediction error (P_e) under MNAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
mice+pmm	9.9687	8.9343	24.4445	5.5350
mice+rf	9.8693	8.9668	35.7776	6.8370
rSVD84	12.1347	12.0740	4.1648	2.7935
rSVD01-84	12.4203	12.3043	4.3662	2.9668
rSVD01	13.1388	13.0078	4.8750	3.1551
ABB	21.2483	21.3587	14.8275	5.3536
$p = 20\%$				
mice+pmm	12.1315	11.4317	23.8991	6.8304
mice+rf	11.4083	11.4943	34.6470	6.8499
rSVD84	13.8099	13.0806	9.9339	3.4864
rSVD01-84	13.9997	13.3756	8.7620	3.6788
rSVD01	14.6535	13.9893	9.0607	3.9918
ABB	20.4742	19.5758	19.9646	5.2922
$p = 30\%$				
mice+rf	13.3870	10.8026	71.1549	9.1820
mice+pmm	13.1225	11.4731	50.0623	7.9938
rSVD84	15.3462	14.4489	18.4413	3.7950
rSVD01-84	15.3580	14.6669	14.5514	3.8013
rSVD01	15.9238	15.2073	14.0842	3.6384
ABB	19.7753	18.8589	29.2351	7.3453

Table 31. Variance among imputations ($\mathbf{V}$) under MAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
mice+rf	9.5014	8.3908	23.7024	6.3826
mice+pmm	9.5264	9.0806	9.7818	4.5073
rSVD84	10.3288	10.3447	2.6881	2.0877
rSVD01-84	10.4989	10.4959	2.7148	2.0190
rSVD01	10.9484	10.9540	2.8066	2.0386
ABB	23.3302	22.9247	12.8239	4.0302
$p = 20\%$				
mice+pmm	9.8067	9.5094	7.5661	3.5609
mice+rf	9.8471	9.5246	19.2063	6.0469
rSVD84	12.0089	11.8113	3.4727	2.4817
rSVD01-84	12.2013	12.0187	3.4108	2.4840
rSVD01	12.7458	12.6106	3.5278	2.3909
ABB	23.4496	23.3118	9.3295	3.9055
$p = 30\%$				
mice+rf	10.0419	9.6249	22.8505	5.5632
mice+pmm	10.2045	10.0211	7.3664	3.7278
rSVD84	13.9967	13.8457	4.9473	2.6484
rSVD01-84	14.1460	13.9660	4.5566	2.4042
rSVD01	14.7232	14.6248	4.6945	2.2385
ABB	23.3238	23.1471	9.6075	4.6846

Table 32. Variance among imputations ($\mathbf{V}$) under MCAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
mice+rf	7.7759	5.8291	34.9007	5.4298
mice+pmm	7.7517	6.2625	25.3929	5.0749
rSVD84	10.3008	10.1608	3.4258	2.0367
rSVD01-84	10.5512	10.4914	3.4041	2.0224
rSVD01	11.2125	11.1996	3.8424	2.2017
ABB	18.8611	18.4834	9.3655	4.2520
$p = 20\%$				
mice+rf	7.9548	6.8941	19.7678	5.4469
mice+pmm	7.9328	7.2667	13.8413	4.3879
rSVD84	10.7049	10.3512	6.5668	2.6392
rSVD01-84	10.7971	10.5317	5.3231	2.6888
rSVD01	11.2835	11.0430	5.4338	2.5565
ABB	16.2689	16.3000	4.3281	2.7933
$p = 30\%$				
mice+rf	7.9985	7.6932	11.3383	4.7247
mice+pmm	8.3832	8.0959	5.0797	2.6946
rSVD84	11.2733	10.5026	11.0049	2.9818
rSVD01-84	11.2175	10.6602	8.0867	2.8133
rSVD01	11.5863	11.0599	7.9546	2.9953
ABB	13.7840	13.6150	4.2702	2.2590

Table 33. Variance among imputations ($\mathbf{V}$) under MNAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
rSVD84	6.4178	6.3906	1.5149	1.7157
rSVD01-84	6.5939	6.4991	1.8342	1.9868
rSVD01	6.7986	6.6385	2.2221	2.1429
ABB	11.7617	11.8044	11.0902	4.6714
mice+pmm	14.7495	13.5481	50.3947	7.3579
mice+rf	34.3446	28.2718	363.1586	20.5142
$p = 20\%$				
rSVD84	6.4190	6.2340	1.8235	1.6348
rSVD01-84	6.6392	6.4462	2.1913	1.8801
rSVD01	7.0245	6.7819	2.9152	2.1885
ABB	11.9257	11.7352	13.2881	4.1506
mice+pmm	15.7732	13.9130	82.5031	8.6890
mice+rf	40.4272	36.4373	420.4717	24.5992
$p = 30\%$				
rSVD84	6.5229	6.3300	1.7815	2.0893
rSVD01-84	6.7495	6.5578	2.0281	2.0865
rSVD01	7.3665	6.9867	3.3579	2.4175
ABB	11.3085	10.9999	9.1128	3.9961
mice+pmm	15.0096	13.8329	31.0232	6.3528
mice+rf	41.5616	37.8417	433.9890	27.0840

Table 34. Prediction error (P_e) under MAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
rSVD84	4.7061	4.6554	0.2134	0.5669
rSVD01-84	4.7916	4.7690	0.2324	0.6294
rSVD01	4.8972	4.8606	0.2591	0.6791
ABB	10.5496	10.4886	2.6094	1.8859
mice+pmm	12.3523	10.7141	37.0696	5.5523
mice+rf	21.8205	21.8515	41.9789	8.1956
$p = 20\%$				
rSVD84	5.0643	5.0365	0.2396	0.5848
rSVD01-84	5.2529	5.2036	0.2786	0.6704
rSVD01	5.5163	5.4263	1.0494	0.6930
ABB	11.0466	10.7682	3.6045	2.1413
mice+pmm	13.6455	12.8405	18.0121	6.1081
mice+rf	23.3178	22.1241	36.3856	8.3948
$p = 30\%$				
rSVD84	5.5634	5.4334	0.4183	0.7733
rSVD01-84	5.8593	5.7034	0.5131	0.9336
rSVD01	6.3988	6.0370	1.0494	1.4125
ABB	10.9263	10.6829	2.9831	2.5736
mice+pmm	13.0694	12.2495	18.0234	5.3543
mice+rf	23.1352	22.7789	25.4325	7.3421

Table 35. Prediction error (P_e) under MCAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
rSVD84	9.9119	9.7808	0.7218	1.1178
rSVD01-84	10.2388	10.0358	0.9128	1.2934
rSVD01	10.6009	10.3464	1.1985	1.4838
mice+pmm	17.1152	15.3419	37.3328	7.9619
ABB	16.2871	16.1711	1.9391	1.9661
mice+rf	33.4346	31.2091	121.4565	13.6813
$p = 20\%$				
rSVD84	10.3091	10.1658	1.2374	1.3223
rSVD01-84	10.6477	10.5246	1.4575	1.3294
rSVD01	11.1975	11.1050	1.8929	1.5143
ABB	16.6853	16.4506	4.7241	2.5162
mice+pmm	17.0730	16.7440	20.4267	5.7486
mice+rf	32.1005	29.4333	136.8696	11.8640
$p = 30\%$				
rSVD84	10.5019	10.4836	1.2493	1.3873
rSVD01-84	10.7901	10.7185	1.4014	1.5636
rSVD01	11.4395	11.4091	1.8146	1.6493
ABB	15.5507	15.3198	2.6325	2.4092
mice+pmm	18.7844	17.4740	27.2679	6.1571
mice+rf	30.2478	28.1796	94.6135	13.9552

Table 36. Prediction error (P_e) under MNAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
rSVD84	571.9633	513.7169	68987.4476	285.4311
rSVD01-84	719.0283	647.1862	96817.0593	339.8798
rSVD01	929.4995	833.2110	160527.0386	409.3822
ABB	3121.9069	2904.5421	1098811.8578	1427.5179
mice+pmm	2890.1812	2498.9679	2517979.4321	1622.388
mice+rf	103.5472	91.1364	4174.8510	76.9731
$p = 20\%$				
rSVD84	835.0142	750.6510	142333.0450	351.5520
rSVD01-84	1017.6039	906.2284	202834.5590	385.3017
rSVD01	1287.6536	1153.7530	321841.6040	487.4809
ABB	3200.4090	2817.9985	3185470.5790	2132.5780
mice+pmm	3200.4	2817.9	3185470.5	2132.5
mice+rf	171.8711	155.4456	8554.0160	132.6323
$p = 30\%$				
rSVD84	1000.6745	929.9595	170825.2799	488.7156
rSVD01-84	1176.1349	1069.9511	198076.2694	553.0553
rSVD01	1455.5677	1343.0157	286168.6705	637.4053
ABB	2825.6205	2764.7421	850639.5820	1168.0912
mice+pmm	3203.7424	2683.5830	3099831	1903.03
mice+rf	265.6767	250.9006	31380.7595	208.7756

Table 37. Variance among imputations ($\mathbf{V}$) under MAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
rSVD84	367.8464	328.6590	23278.1161	177.0084
rSVD01-84	455.8580	413.3098	32193.5219	219.7210
rSVD01	580.3850	527.4198	49171.1177	303.0942
ABB	3030.1011	3023.8773	418518.4443	932.2408
mice+pmm	2636.6574	2390.36	1570165.07	1365.13
mice+rf	68.6269	61.3011	1200.3549	44.4398
$p = 20\%$				
rSVD84	632.3672	583.8219	50923.8248	260.5645
rSVD01-84	783.8607	737.0170	69492.2458	322.7370
rSVD01	1001.7887	961.9042	104614.7636	415.3586
ABB	3163.1884	3092.9585	434708.5778	821.5389
mice+pmm	3095.9676	2907.9615	1161175.3220	1563.6760
mice+rf	117.8451	107.1058	2772.5299	66.8569
$p = 30\%$				
rSVD84	936.6314	867.8370	126807.9242	385.9205
rSVD01-84	1122.3327	1055.0080	163248.2352	440.1711
rSVD01	1399.1035	1308.3400	232975.1442	550.2081
ABB	3102.6196	3032.1840	424917.7063	911.4790
mice+pmm	2940.5738	2758.079	1338384.4	1423.129
mice+rf	149.5618	140.0740	3981.5319	91.5794

Table 38. Variance among imputations ($\mathbf{V}$) under MCAR (grouped by missingness percentage).

Method	**Mean**	**Median**	**Variance**	**IQR**
$p = 10\%$				
rSVD84	607.3012	564.1873	57299.4680	310.8881
rSVD01-84	769.6728	706.8804	89973.9906	359.1349
rSVD01	1004.7385	907.9561	162362.3782	454.6370
ABB	2995.0856	2975.8135	269601.8354	681.9477
mice+pmm	2636.65	2390.36	1570165.07	1365.13
mice+rf	125.0897	109.8589	4833.5478	68.9223
$p = 20\%$				
rSVD84	944.0177	818.5178	154420.9960	516.7055
rSVD01-84	1134.3894	1045.6935	258988.1772	538.6440
rSVD01	1426.4520	1320.4389	387286.5895	713.0625
ABB	3217.7594	3180.8013	494035.7179	751.2147
mice+pmm	2702.1274	2654.6499	1221235	1588
mice+rf	162.8023	133.8680	8924.3014	115.2394
$p = 30\%$				
rSVD84	1060.6522	966.0869	154420.9960	486.0697
rSVD01-84	1224.6791	1141.2664	174655.1722	535.2732
rSVD01	1490.7989	1419.0680	242554.3890	588.6854
ABB	2864.9597	2845.5780	283390.0579	662.5097
mice+pmm	3069	2828.32	1465753	1501.3
mice+rf	190.2017	164.3358	13666.8990	108.3489

Table 39. Variance among imputations ($\mathbf{V}$) under MNAR (grouped by missingness percentage).

Table 40. V by percentage and method — Mechanism **MAR** (REP 1–5)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	75.88	63.39	945.09	22.72	63.30	49.37	1.06e+03	23.82	53.40	36.81	1.31e+03	18.72
	MICE+PMM	39.60	20.67	1.60e+03	35.19	43.97	25.15	1.90e+03	26.05	36.31	23.81	955.42	18.06
	MICE+RF	82.99	85.77	721.04	41.13	20.38	20.73	37.69	8.10	28.06	28.00	51.71	11.02
	rSVD01	325.04	318.67	2.77e+03	61.08	452.45	439.39	6.50e+03	58.77	438.69	413.18	7.15e+03	126.13
	rSVD84	393.34	376.95	2.79e+03	33.62	453.89	435.73	6.23e+03	38.83	401.95	390.78	2.03e+03	41.11
	rSVD84-01	236.03	228.09	2.70e+03	69.12	360.21	346.21	4.87e+03	76.69	373.76	369.34	3.65e+03	83.22
Bro and Smilde (2014) Red Wine 14×44	ABB	3.49	2.91	2.61	2.13	2.95	2.53	1.53	1.44	2.65	2.44	1.07	1.72
	MICE+PMM	2.62	2.12	2.43	1.57	2.54	1.97	1.78	1.55	2.49	1.99	2.21	1.14
	MICE+RF	6.18	6.06	3.75	2.23	1.59	1.56	0.21	0.55	3.52	3.61	1.10	1.34
	rSVD01	1.52e+04	1.19e+04	1.40e+08	1.27e+04	3.50e+04	1.79e+04	2.01e+09	2.75e+04	1.01e+05	1.82e+04	3.09e+11	1.28e+04
	rSVD84	19.02	18.69	3.60	1.30	35.97	35.58	2.94	1.84	347.66	313.79	9.90e+03	55.09
	rSVD84-01	4.87e+03	3.42e+03	1.85e+07	4.27e+03	1.20e+04	4.73e+03	1.73e+09	5.62e+03	4.85e+04	8.93e+03	6.87e+10	1.06e+04
Forkman and Piepho (2014) Soybean 7×10	ABB	1.73e+05	1.53e+05	3.60e+09	6.88e+04	1.44e+05	1.30e+05	3.33e+09	9.61e+04	1.35e+05	1.04e+05	6.21e+09	7.30e+04
	MICE+PMM	1.18e+05	1.02e+05	2.02e+09	3.71e+04	1.08e+05	8.23e+04	2.47e+09	6.52e+04	1.11e+05	8.48e+04	3.23e+09	4.60e+04
	MICE+RF	7.06e+04	6.87e+04	4.92e+08	2.66e+04	2.24e+05	2.26e+05	4.57e+09	8.33e+04	6.35e+04	6.54e+04	4.22e+08	2.93e+04
	rSVD01	4.30e+05	4.20e+05	4.55e+09	7.03e+04	4.71e+05	4.50e+05	7.34e+09	5.49e+04	5.36e+05	5.26e+05	5.73e+09	8.66e+04
	rSVD84	4.51e+05	4.43e+05	1.11e+09	3.48e+04	4.44e+05	4.28e+05	3.70e+09	3.39e+04	4.31e+05	4.15e+05	2.61e+09	5.05e+04
	rSVD84-01	3.86e+05	3.84e+05	3.04e+09	6.39e+04	4.41e+05	4.26e+05	5.43e+09	7.04e+04	4.64e+05	4.59e+05	4.64e+09	6.99e+04
Forkman and Piepho (2014) Pea 5×4	ABB	3.02e+05	3.19e+05	3.40e+10	2.89e+05	3.14e+05	3.02e+05	4.26e+10	3.37e+05	2.44e+05	2.25e+05	3.35e+10	3.48e+05
	MICE+PMM	4.39e+05	4.62e+05	8.28e+10	5.26e+05	4.46e+05	4.17e+05	9.73e+10	4.72e+05	4.37e+05	1.15e+05	2.06e+11	6.13e+05
	MICE+RF	4.96e+05	4.93e+05	2.43e+10	2.48e+05	5.08e+05	5.14e+05	2.28e+10	1.74e+05	1.44e+06	1.39e+06	1.70e+11	5.22e+05
	rSVD01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6.99e+05	6.94e+05	3.40e+10	2.35e+05
	rSVD84	2.13e+06	2.00e+06	2.06e+11	5.13e+05	2.37e+06	2.28e+06	1.52e+11	4.07e+05	8.47e+06	8.07e+06	1.28e+13	5.48e+05
	rSVD84-01	7.24e+05	6.03e+05	2.30e+11	2.61e+05	6.71e+05	5.97e+05	1.05e+11	2.22e+05	1.30e+06	1.13e+06	3.73e+11	4.54e+05
Cornelius et al. (1996) Maize 9×20	ABB	1.90e+06	1.81e+06	1.68e+11	4.75e+05	1.79e+06	1.73e+06	2.03e+11	4.96e+05	1.94e+06	1.81e+06	3.00e+11	7.29e+05
	MICE+PMM	1.26e+06	1.17e+06	2.00e+11	5.41e+05	1.20e+06	1.14e+06	1.71e+11	3.84e+05	1.63e+06	1.48e+06	3.47e+11	7.27e+05
	MICE+RF	8.72e+05	8.81e+05	7.27e+10	3.45e+05	2.62e+06	2.69e+06	5.86e+11	1.01e+06	1.70e+06	1.63e+06	2.77e+11	7.12e+05
	rSVD01	2.15e+06	2.02e+06	1.77e+11	4.65e+05	1.66e+06	1.63e+06	3.09e+10	2.20e+05	2.21e+06	2.14e+06	1.27e+11	3.87e+05
	rSVD84	1.85e+06	1.76e+06	1.93e+11	2.11e+05	2.05e+06	2.00e+06	6.45e+10	3.78e+05	2.69e+06	2.53e+06	2.60e+11	4.52e+05
	rSVD84-01	1.96e+06	1.86e+06	2.58e+11	4.70e+05	1.85e+06	1.83e+06	4.40e+10	2.15e+05	2.21e+06	2.10e+06	1.43e+11	3.79e+05

Table 41. V by percentage and method — Mechanism **MCAR** (REP 1–5)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	60.10	52.98	1.12e+03	30.81	62.29	53.87	638.46	25.36	67.84	61.74	729.10	31.39
	MICE+PMM	30.22	22.62	552.58	22.71	35.11	28.14	595.22	23.64	42.98	36.44	516.12	26.70
	MICE+RF	12.57	12.57	18.22	5.78	77.20	79.10	622.21	33.55	54.34	53.13	192.25	17.23
	rSVD01	303.10	302.60	3.61e+03	82.49	373.72	378.28	6.40e+03	95.88	427.96	430.63	9.36e+03	130.55
	rSVD84	246.24	234.74	5.75e+03	75.08	275.95	268.30	4.70e+03	77.91	376.16	356.67	1.31e+04	130.05
	rSVD84-01	239.90	234.29	2.94e+03	70.70	280.07	278.97	2.84e+03	81.80	349.54	335.37	7.88e+03	95.25
Bro and Smilde (2014) Red Wine 14×44	ABB	1.84e+03	1.68e+03	1.04e+06	1.33e+03	1.90e+03	1.64e+03	8.16e+05	1.06e+03	1.99e+03	1.57e+03	1.69e+06	1.18e+03
	MICE+PMM	1.58e+03	1.25e+03	1.55e+06	895.32	1.58e+03	1.33e+03	8.76e+05	854.39	1.72e+03	1.26e+03	1.28e+06	1.06e+03
	MICE+RF	960.59	957.67	6.46e+04	345.87	574.75	560.03	2.42e+04	209.26	1.16e+03	1.16e+03	1.03e+05	406.48
	rSVD01	1.33e+04	7.28e+03	2.96e+08	1.04e+04	3.23e+04	1.22e+04	8.95e+09	1.31e+04	2.85e+04	1.68e+04	1.49e+09	2.43e+04
	rSVD84	23.22	15.32	668.96	8.72	84.90	53.38	7.32e+03	55.30	180.71	142.72	2.27e+04	123.05
	rSVD84-01	5.24e+03	2.15e+03	1.23e+08	3.29e+03	6.89e+03	2.32e+03	2.02e+08	4.82e+03	1.11e+04	3.80e+03	1.05e+09	3.80e+03
Forkman and Piepho (2014) Soybean 7×10	ABB	2.06e+05	1.76e+05	1.29e+10	1.09e+05	2.39e+05	2.00e+05	1.89e+10	9.88e+04	2.08e+05	1.80e+05	7.87e+09	1.11e+05
	MICE+PMM	1.22e+05	1.04e+05	4.32e+09	6.50e+04	1.68e+05	1.37e+05	1.66e+10	8.71e+04	1.62e+05	1.30e+05	1.30e+10	8.77e+04
	MICE+RF	2.45e+05	2.39e+05	4.99e+09	8.89e+04	2.62e+05	2.63e+05	7.84e+09	1.16e+05	2.51e+05	2.44e+05	4.80e+09	1.05e+05
	rSVD01	3.85e+05	3.80e+05	4.43e+09	9.11e+04	4.33e+05	4.31e+05	6.55e+09	1.02e+05	5.63e+05	5.65e+05	1.40e+10	1.71e+05
	rSVD84	3.69e+05	3.62e+05	2.66e+09	6.24e+04	4.45e+05	4.34e+05	7.25e+09	1.04e+05	5.32e+05	5.33e+05	1.18e+10	1.48e+05
	rSVD84-01	3.51e+05	3.48e+05	3.05e+09	8.26e+04	4.13e+05	4.20e+05	5.36e+09	9.85e+04	5.06e+05	4.91e+05	1.19e+10	1.40e+05
Forkman and Piepho (2014) Pea 5×4	ABB	4.55e+05	2.15e+05	3.44e+11	4.53e+05	4.29e+05	3.29e+05	1.28e+11	4.43e+05	5.32e+05	4.58e+05	1.08e+11	4.34e+05
	MICE+PMM	5.58e+05	1.68e+05	7.71e+11	3.75e+05	5.45e+05	2.77e+05	5.70e+11	3.32e+05	6.58e+05	2.56e+05	6.53e+11	7.26e+05
	MICE+RF	3.30e+03	1.26e+03	2.56e+07	2.11e+03	4.63e+03	2.30e+03	4.05e+07	4.50e+03	2.47e+05	2.43e+05	5.45e+09	8.32e+04
	rSVD01	5.88e+05	4.78e+05	1.19e+11	3.20e+05	6.81e+05	5.94e+05	1.53e+11	3.89e+05	6.81e+05	6.27e+05	1.18e+11	4.35e+05
	rSVD84	4.28e+05	3.67e+05	5.28e+10	2.00e+05	5.43e+05	4.91e+05	5.28e+10	2.40e+05	9.06e+05	6.63e+05	1.20e+12	4.71e+05
	rSVD84-01	4.62e+05	4.05e+05	4.03e+10	2.08e+05	5.65e+05	5.39e+05	6.13e+10	2.58e+05	6.15e+05	5.66e+05	7.77e+10	3.10e+05
Cornelius et al. (1996) Maize 9×20	ABB	2.58e+06	2.47e+06	3.17e+11	7.32e+05	2.68e+06	2.62e+06	2.75e+11	5.88e+05	2.73e+06	2.64e+06	2.69e+11	6.18e+05
	MICE+PMM	1.40e+06	1.24e+06	2.66e+11	6.32e+05	1.51e+06	1.46e+06	1.82e+11	5.55e+05	1.87e+06	1.75e+06	2.71e+11	6.39e+05
	MICE+RF	1.20e+06	1.22e+06	1.15e+11	4.56e+05	1.39e+06	1.36e+06	2.39e+11	7.01e+05	1.47e+06	1.48e+06	1.86e+11	5.46e+05
	rSVD01	1.78e+06	1.75e+06	2.54e+10	2.19e+05	1.95e+06	1.90e+06	3.84e+10	2.59e+05	2.02e+06	1.98e+06	8.45e+10	3.26e+05
	rSVD84	1.60e+06	1.59e+06	1.15e+10	1.24e+05	1.81e+06	1.79e+06	1.82e+10	1.81e+05	1.72e+06	1.69e+06	4.54e+10	1.83e+05
	rSVD84-01	1.69e+06	1.68e+06	1.95e+10	1.47e+05	1.77e+06	1.74e+06	2.43e+10	2.21e+05	1.92e+06	1.85e+06	5.91e+10	2.72e+05

Table 42. V by percentage and method — Mechanism **MNAR** (REP 1–5)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	52.66	45.28	361.28	23.12	45.18	41.64	314.10	22.14	41.70	36.13	363.14	25.43
	MICE+PMM	26.85	17.87	502.04	22.73	31.62	28.80	258.53	18.04	33.58	29.78	225.26	24.46
	MICE+RF	13.15	13.06	17.36	5.68	30.63	30.61	83.97	12.61	35.50	36.03	100.31	12.26
	rSVD01	442.88	442.53	2.55e+03	63.07	404.77	392.67	2.74e+03	51.34	584.53	581.57	1.16e+04	90.35
	rSVD84	401.12	397.81	891.80	27.59	276.43	271.68	431.47	26.50	421.68	410.04	1.90e+03	36.72
	rSVD84-01	302.21	289.50	3.56e+03	78.50	326.65	320.85	1.95e+03	54.48	458.41	451.21	4.85e+03	63.90
Bro and Smilde (2014) Red Wine 14×44	ABB	2.04e+03	1.77e+03	8.22e+05	1.50e+03	1.76e+03	1.46e+03	6.77e+05	1.22e+03	1.56e+03	1.00e+03	7.06e+05	1.06e+03
	MICE+PMM	1.42e+03	1.13e+03	9.02e+05	1.14e+03	1.41e+03	1.02e+03	7.90e+05	990.52	1.46e+03	965.80	9.05e+05	969.11
	MICE+RF	1.94e+03	1.99e+03	4.03e+05	883.45	2.86e+03	2.84e+03	7.50e+05	1.12e+03	1.08e+03	1.06e+03	1.05e+05	402.34
	rSVD01	5.27e+04	2.63e+04	2.58e+09	8.78e+04	2.17e+04	9.77e+03	2.23e+09	1.04e+04	1.18e+05	2.93e+04	3.94e+11	3.28e+04
	rSVD84	32.78	32.80	47.83	3.50	40.33	32.69	1.16e+03	1.64	403.81	344.55	6.63e+04	36.00
	rSVD84-01	7.88e+03	3.30e+03	2.21e+08	4.79e+03	1.10e+04	4.43e+03	3.77e+08	7.19e+03	6.00e+04	7.15e+03	1.42e+11	1.42e+04
Forkman and Piepho (2014) Soybean 7×10	ABB	1.61e+05	1.55e+05	2.68e+09	6.32e+04	1.39e+05	1.27e+05	4.04e+09	9.06e+04	1.26e+05	1.02e+05	3.30e+09	7.27e+04
	MICE+PMM	1.05e+05	9.13e+04	1.78e+09	4.08e+04	9.47e+04	8.27e+04	1.71e+09	6.32e+04	8.51e+04	7.13e+04	1.50e+09	4.18e+04
	MICE+RF	7.41e+04	7.50e+04	5.04e+08	2.51e+04	5.53e+04	5.67e+04	1.93e+08	1.52e+04	7.62e+04	7.53e+04	4.81e+08	3.23e+04
	rSVD01	3.84e+05	3.79e+05	1.96e+09	4.72e+04	4.18e+05	4.10e+05	2.94e+09	5.80e+04	4.68e+05	4.56e+05	5.45e+09	5.67e+04
	rSVD84	4.12e+05	4.10e+05	9.37e+08	3.05e+04	4.13e+05	4.03e+05	2.66e+09	3.92e+04	4.24e+05	4.18e+05	1.87e+09	4.72e+04
	rSVD84-01	3.60e+05	3.56e+05	2.38e+09	6.28e+04	4.31e+05	4.26e+05	3.79e+09	7.07e+04	4.22e+05	4.15e+05	3.58e+09	6.00e+04
Forkman and Piepho (2014) Pea 5×4	ABB	3.52e+05	3.33e+05	2.06e+10	1.57e+05	3.23e+05	3.02e+05	1.78e+10	1.80e+05	3.29e+05	2.84e+05	2.48e+10	1.85e+05
	MICE+PMM	4.64e+05	4.28e+05	4.22e+10	2.95e+05	4.24e+05	3.87e+05	3.62e+10	2.33e+05	4.27e+05	3.60e+05	4.34e+10	2.43e+05
	MICE+RF	3.00e+05	2.96e+05	8.60e+09	1.24e+05	4.00e+05	4.04e+05	1.55e+10	1.64e+05	5.71e+05	5.91e+05	2.90e+10	2.29e+05
	rSVD01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	rSVD84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	rSVD84-01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cornelius et al. (1996) Maize 9×20	ABB	2.08e+06	2.03e+06	9.20e+10	3.66e+05	1.74e+06	1.70e+06	9.39e+10	3.79e+05	1.58e+06	1.57e+06	9.78e+10	4.13e+05
	MICE+PMM	1.48e+06	1.40e+06	1.72e+11	5.00e+05	1.22e+06	1.16e+06	1.22e+11	4.77e+05	1.18e+06	1.12e+06	1.53e+11	4.72e+05
	MICE+RF	1.11e+06	1.13e+06	1.23e+11	4.36e+05	1.26e+06	1.25e+06	1.08e+11	4.35e+05	5.26e+05	5.10e+05	2.40e+10	1.76e+05
	rSVD01	1.02e+06	9.92e+05	2.51e+10	1.80e+05	3.97e+05	3.72e+05	1.19e+10	1.22e+05	6.16e+05	5.36e+05	8.43e+10	1.98e+05
	rSVD84	2.01e+06	1.86e+06	2.20e+11	6.77e+05	1.22e+06	1.16e+06	6.35e+10	2.86e+05	1.06e+06	9.96e+05	6.22e+10	2.41e+05
	rSVD84-01	1.39e+06	1.38e+06	3.72e+10	2.38e+05	6.08e+05	5.92e+05	1.69e+10	1.61e+05	6.46e+05	6.27e+05	3.51e+10	1.66e+05

Table 43. V by percentage and method — Mechanism **MAR** (REP 6–10)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	28.18	25.29	86.60	8.77	26.47	23.03	80.84	11.82	26.66	24.46	67.45	9.38
	MICE+PMM	24.26	20.37	190.26	7.82	22.54	19.21	92.13	8.90	22.91	20.91	66.18	8.11
	MICE+RF	21.94	22.62	56.45	9.55	20.91	20.71	38.61	8.63	15.17	15.34	18.25	5.65
	rSVD01	53.96	51.62	76.87	8.97	52.48	50.18	62.13	8.13	86.06	85.69	50.55	9.46
	rSVD84	48.82	46.46	41.11	8.45	48.03	46.93	22.47	5.66	103.33	102.29	47.51	6.34
	rSVD84-01	45.78	44.40	44.50	8.12	48.48	47.18	57.35	7.16	98.31	99.21	174.12	18.26
Yan et al. (2007) Wheat 18×9	ABB	0.26	0.26	0.00	0.07	0.25	0.24	0.00	0.07	0.22	0.22	0.00	0.05
	MICE+PMM	0.26	0.25	0.00	0.08	0.23	0.22	0.00	0.08	0.22	0.22	0.00	0.06
	MICE+RF	0.14	0.14	0.00	0.06	0.20	0.20	0.00	0.07	0.35	0.36	0.01	0.16
	rSVD01	0.83	0.80	0.02	0.13	0.97	0.94	0.02	0.15	0.93	0.92	0.01	0.14
	rSVD84	0.80	0.79	0.00	0.07	0.75	0.72	0.01	0.07	0.86	0.86	0.00	0.05
	rSVD84-01	0.76	0.73	0.01	0.12	0.81	0.78	0.02	0.13	0.80	0.79	0.01	0.09
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	2.04	1.98	0.12	0.36	1.93	1.89	0.12	0.43	1.83	1.77	0.19	0.56
	MICE+PMM	1.45	1.37	0.13	0.34	1.38	1.33	0.10	0.39	1.39	1.28	0.17	0.47
	MICE+RF	1.43	1.46	0.15	0.50	2.06	2.12	0.39	0.97	1.11	1.10	0.10	0.37
	rSVD01	7.05	7.04	0.16	0.51	8.61	8.55	0.21	0.64	9.56	9.48	0.30	0.63
	rSVD84	5.83	5.80	0.05	0.32	9.70	9.67	0.04	0.28	10.91	10.92	0.10	0.31
	rSVD84-01	6.41	6.40	0.15	0.41	8.15	8.13	0.20	0.67	9.92	9.89	0.29	0.65
Calinski et al. (2009a) Pea 18×9	ABB	12.45	11.85	7.78	3.76	13.89	12.48	14.09	5.02	11.12	10.30	7.24	3.37
	MICE+PMM	11.37	10.50	10.12	3.36	13.40	12.49	17.75	5.58	11.07	9.91	10.16	3.90
	MICE+RF	11.87	11.22	13.96	5.00	22.61	22.91	40.05	8.29	11.48	12.17	13.35	5.71
	rSVD01	34.47	33.94	9.49	3.67	55.04	54.88	9.96	3.52	75.19	75.29	15.38	5.07
	rSVD84	35.05	35.14	3.36	2.36	52.00	52.48	6.61	3.34	65.64	65.74	4.01	2.79
	rSVD84-01	34.61	34.87	11.44	4.08	50.27	50.17	22.82	5.94	67.94	67.87	22.93	6.73
Calinski et al. (2009b) Winter rye 18×15	ABB	62.27	61.80	95.22	11.52	52.87	51.52	102.12	14.46	51.76	52.05	76.23	11.30
	MICE+PMM	65.62	64.06	190.12	16.63	56.40	54.63	167.78	18.40	55.98	55.74	102.34	16.10
	MICE+RF	48.16	46.53	213.32	20.28	38.36	39.31	116.88	14.50	72.98	73.81	326.73	24.61
	rSVD01	45.58	47.26	22.33	8.55	70.25	69.61	24.37	5.71	79.53	77.71	73.26	9.28
	rSVD84	48.35	48.24	3.74	1.98	68.95	68.82	16.28	3.14	78.98	78.28	22.87	4.04
	rSVD84-01	50.43	50.21	16.28	5.15	65.86	65.01	34.79	7.41	78.09	76.26	68.81	9.01

Table 44. V by percentage and method — Mechanism **MCAR** (REP 6–10)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	40.01	37.83	263.37	12.46	34.85	32.90	103.69	12.29	35.19	33.58	61.68	11.68
	MICE+PMM	29.43	26.99	115.33	12.46	27.36	26.21	91.33	13.42	29.40	28.29	61.83	10.48
	MICE+RF	15.28	15.18	26.50	6.02	24.39	25.07	65.27	10.96	27.86	27.73	77.31	11.55
	rSVD01	46.87	45.75	68.32	9.46	51.62	50.41	69.89	9.16	59.40	57.78	146.77	12.67
	rSVD84	40.96	40.31	29.60	6.41	46.58	46.23	32.66	7.48	54.99	52.28	231.27	11.39
	rSVD84-01	42.54	41.16	38.55	6.11	46.18	45.00	47.73	7.60	54.28	50.62	165.12	10.24
Yan et al. (2007) Wheat 18×9	ABB	0.38	0.37	0.01	0.10	0.39	0.38	0.01	0.11	0.35	0.34	0.00	0.09
	MICE+PMM	0.30	0.26	0.02	0.12	0.32	0.30	0.01	0.12	0.29	0.29	0.01	0.08
	MICE+RF	0.23	0.23	0.00	0.09	0.16	0.16	0.00	0.07	0.24	0.24	0.00	0.10
	rSVD01	0.48	0.47	0.00	0.06	0.55	0.56	0.00	0.08	0.62	0.62	0.01	0.12
	rSVD84	0.47	0.46	0.00	0.06	0.54	0.53	0.00	0.08	0.58	0.57	0.01	0.11
	rSVD84-01	0.46	0.45	0.00	0.06	0.52	0.52	0.00	0.06	0.57	0.56	0.01	0.11
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	2.23	2.18	0.08	0.32	2.18	2.15	0.04	0.30	2.21	2.17	0.06	0.29
	MICE+PMM	1.50	1.47	0.07	0.36	1.52	1.52	0.06	0.31	1.56	1.53	0.06	0.26
	MICE+RF	1.60	1.63	0.17	0.61	1.82	1.84	0.35	0.83	1.47	1.49	0.20	0.59
	rSVD01	5.31	5.31	0.10	0.37	6.15	6.11	0.18	0.57	7.22	7.22	0.35	0.78
	rSVD84	4.91	4.87	0.07	0.35	5.72	5.66	0.14	0.49	6.82	6.71	0.57	0.80
	rSVD84-01	4.93	4.91	0.08	0.36	5.82	5.73	0.17	0.57	6.92	6.83	0.43	0.75
Calinski et al. (2009a) Pea 18×9	ABB	13.98	13.58	11.50	4.39	14.46	14.31	8.14	3.42	12.99	12.56	7.48	3.11
	MICE+PMM	11.10	10.42	14.58	4.34	11.89	11.24	8.29	4.07	11.35	10.88	6.81	3.50
	MICE+RF	8.22	8.19	5.47	2.56	12.96	13.14	16.00	5.63	12.06	11.93	14.00	5.19
	rSVD01	33.42	33.19	20.09	4.66	41.47	40.93	32.74	6.69	51.71	51.46	43.83	7.36
	rSVD84	32.45	32.66	13.07	5.25	40.38	39.28	27.17	7.77	46.94	46.90	26.78	6.73
	rSVD84-01	31.37	30.90	12.37	4.16	38.88	38.31	29.51	6.99	46.07	45.86	32.02	7.23
Calinski et al. (2009b) Winter rye 18×15	ABB	71.79	69.53	192.59	15.03	67.33	65.89	115.14	13.84	70.47	67.56	164.79	13.51
	MICE+PMM	67.15	63.92	281.14	22.16	66.82	64.44	160.67	10.07	73.23	71.01	199.17	13.18
	MICE+RF	74.23	73.25	643.50	39.35	59.80	56.55	338.41	24.66	73.96	73.10	400.71	28.60
	rSVD01	49.86	49.83	13.61	5.07	54.46	54.11	24.37	6.83	58.52	58.38	42.88	8.95
	rSVD84	50.94	50.54	13.76	5.30	51.38	50.73	21.24	5.17	60.55	59.30	47.59	9.11
	rSVD84-01	49.80	48.93	16.35	4.63	52.97	52.09	17.71	4.85	58.90	59.09	42.58	8.17

Table 45. V by percentage and method — Mechanism **MNAR** (REP 6–10)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	27.38	25.86	41.07	8.08	20.29	18.95	29.79	8.24	18.70	18.21	33.13	8.09
	MICE+PMM	25.06	22.23	114.30	11.41	17.71	16.11	40.06	7.56	15.78	14.20	41.24	8.39
	MICE+RF	30.72	29.36	67.89	10.11	13.60	13.61	14.84	4.91	15.27	15.00	13.49	4.84
	rSVD01	50.00	48.48	107.91	12.33	77.75	77.87	115.76	17.03	102.27	102.85	180.51	16.70
	rSVD84	47.36	45.62	55.46	8.25	121.30	118.58	73.46	9.22	115.48	114.15	70.21	7.20
	rSVD84-01	45.91	44.17	69.82	10.66	97.39	95.00	274.09	21.30	108.06	106.44	287.06	21.48
Yan et al. (2007) Wheat 18×9	ABB	0.28	0.27	0.00	0.07	0.23	0.23	0.00	0.07	0.18	0.17	0.00	0.07
	MICE+PMM	0.26	0.25	0.00	0.09	0.22	0.20	0.00	0.08	0.16	0.14	0.00	0.07
	MICE+RF	0.19	0.19	0.00	0.07	0.26	0.26	0.01	0.09	0.25	0.26	0.01	0.09
	rSVD01	0.66	0.64	0.01	0.10	0.45	0.44	0.00	0.05	0.45	0.45	0.00	0.07
	rSVD84	0.46	0.45	0.00	0.04	1.15	1.13	0.01	0.10	0.83	0.80	0.01	0.09
	rSVD84-01	0.56	0.56	0.01	0.12	0.60	0.59	0.00	0.07	0.55	0.55	0.00	0.08
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	1.58	1.57	0.04	0.27	1.26	1.23	0.03	0.18	1.10	1.10	0.02	0.20
	MICE+PMM	1.16	1.14	0.04	0.23	0.93	0.92	0.03	0.17	0.85	0.82	0.03	0.18
	MICE+RF	1.10	1.11	0.12	0.44	1.01	1.04	0.08	0.40	0.76	0.73	0.05	0.30
	rSVD01	6.85	6.83	0.43	0.73	8.24	8.24	0.16	0.48	9.52	9.58	0.26	0.70
	rSVD84	8.63	8.66	0.16	0.43	12.10	12.06	0.12	0.34	17.00	16.97	0.15	0.47
	rSVD84-01	6.66	6.67	0.27	0.58	9.00	8.98	0.38	0.93	11.85	11.79	0.37	0.83
Calinski et al. (2009a) Pea 18×9	ABB	10.79	10.40	3.13	2.54	9.96	9.57	4.32	2.74	7.69	7.39	2.77	1.98
	MICE+PMM	9.21	8.87	4.81	3.11	8.69	8.59	6.29	3.72	6.61	6.40	2.91	2.18
	MICE+RF	7.72	7.57	7.00	3.50	5.95	5.98	3.25	2.84	10.44	10.45	8.73	3.91
	rSVD01	40.12	39.70	7.12	3.70	62.17	63.12	53.74	11.66	75.45	75.85	42.44	9.89
	rSVD84	36.68	36.90	2.47	1.57	98.76	98.65	11.58	3.53	81.23	80.90	8.02	2.92
	rSVD84-01	34.33	34.46	8.81	4.03	62.31	62.30	36.43	8.29	71.52	71.03	34.40	8.43
Calinski et al. (2009b) Winter rye 18×15	ABB	60.59	59.98	68.18	13.24	49.40	47.76	54.31	8.31	43.28	41.20	52.73	7.94
	MICE+PMM	54.77	53.69	112.05	12.37	59.79	58.96	93.38	12.38	47.97	46.33	84.81	9.22
	MICE+RF	60.01	60.83	296.82	22.47	58.35	58.01	232.84	15.86	48.14	48.18	173.65	19.45
	rSVD01	49.77	49.91	16.67	4.90	64.33	63.06	55.33	9.66	214.77	212.87	605.75	32.82
	rSVD84	47.65	47.82	3.96	2.21	62.56	61.40	33.70	2.96	186.46	183.48	192.54	14.20
	rSVD84-01	48.33	48.31	17.66	5.34	57.78	56.69	59.84	10.96	189.45	192.67	799.73	43.48

Table 46. V by percentage and method — Mechanism **MAR** (REP 11–15)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	1.23e+05	1.20e+05	3.32e+08	2.11e+04	1.23e+05	1.19e+05	4.03e+08	2.08e+04	1.19e+05	1.17e+05	3.92e+08	2.60e+04
	MICE+PMM	1.53e+05	1.52e+05	7.80e+08	3.35e+04	1.56e+05	1.54e+05	4.97e+08	2.50e+04	1.42e+05	1.40e+05	4.45e+08	3.07e+04
	MICE+RF	1.31e+05	1.29e+05	1.58e+09	5.57e+04	1.22e+05	1.23e+05	1.32e+09	4.62e+04	1.20e+05	1.19e+05	1.16e+09	4.41e+04
	rSVD01	4.59e+05	4.60e+05	4.10e+08	2.56e+04	5.20e+05	5.21e+05	4.21e+08	2.62e+04	6.13e+05	6.12e+05	3.73e+08	2.51e+04
	rSVD84	4.31e+05	4.31e+05	1.42e+08	1.60e+04	4.63e+05	4.61e+05	2.07e+08	1.45e+04	5.93e+05	5.92e+05	2.94e+08	1.85e+04
	rSVD84-01	4.34e+05	4.33e+05	4.07e+08	2.45e+04	4.81e+05	4.82e+05	6.17e+08	3.33e+04	5.76e+05	5.78e+05	6.72e+08	3.79e+04
Filho et al. (2008) Cotton 17×23	ABB	0.31	0.30	0.01	0.10	0.30	0.27	0.01	0.11	0.27	0.25	0.01	0.08
	MICE+PMM	0.40	0.35	0.02	0.13	0.37	0.34	0.01	0.12	0.32	0.30	0.01	0.08
	MICE+RF	0.42	0.42	0.01	0.16	0.24	0.24	0.01	0.09	0.31	0.30	0.01	0.11
	rSVD01	0.44	0.43	0.00	0.04	0.52	0.52	0.00	0.03	0.52	0.52	0.00	0.04
	rSVD84	0.43	0.43	0.00	0.02	0.45	0.45	0.00	0.02	0.52	0.51	0.00	0.02
	rSVD84-01	0.42	0.42	0.00	0.04	0.46	0.47	0.00	0.03	0.51	0.51	0.00	0.03
Flores et al. (1998) Faba beans 15×12	ABB	2.17e+05	2.02e+05	3.29e+09	5.14e+04	2.26e+05	2.10e+05	5.66e+09	8.14e+04	1.78e+05	1.68e+05	1.56e+09	3.77e+04
	MICE+PMM	2.47e+05	2.29e+05	3.70e+09	4.14e+04	2.15e+05	1.95e+05	5.76e+09	6.60e+04	1.56e+05	1.51e+05	6.11e+08	3.30e+04
	MICE+RF	2.39e+05	2.41e+05	5.88e+09	1.10e+05	2.10e+05	2.01e+05	4.35e+09	9.63e+04	1.57e+05	1.58e+05	2.33e+09	6.49e+04
	rSVD01	3.41e+05	3.38e+05	2.24e+09	6.49e+04	4.67e+05	4.61e+05	1.91e+09	4.59e+04	5.47e+05	5.34e+05	3.57e+09	3.89e+04
	rSVD84	3.56e+05	3.49e+05	9.77e+08	2.16e+04	4.07e+05	4.03e+05	5.21e+08	2.19e+04	5.07e+05	5.06e+05	1.16e+09	2.77e+04
	rSVD84-01	3.07e+05	3.05e+05	2.16e+09	5.45e+04	4.36e+05	4.34e+05	2.47e+09	6.11e+04	5.01e+05	4.98e+05	2.80e+09	5.50e+04
Mattos et al. (2013) Sugarcane 22×5	ABB	3.12	2.85	0.78	1.10	3.47	3.10	1.30	1.36	3.64	3.19	1.73	1.38
	MICE+PMM	2.86	2.54	1.31	1.07	3.29	2.77	2.54	1.36	3.12	2.72	1.36	1.09
	MICE+RF	3.12	3.19	0.82	1.30	3.68	3.64	1.11	1.62	2.27	2.29	0.49	0.83
	rSVD01	9.56	9.06	3.01	1.52	10.29	9.83	2.81	1.91	11.46	10.67	4.86	2.96
	rSVD84	7.66	7.41	1.17	1.28	9.92	9.54	2.50	1.66	9.16	8.81	2.19	1.97
	rSVD84-01	7.96	7.55	2.45	1.98	9.06	8.76	2.27	1.66	9.80	9.16	3.43	2.53
Rad et al. (2013) Wheat 36×6	ABB	3.93	3.78	0.81	1.24	3.05	2.89	0.73	0.92	3.17	2.99	0.93	1.25
	MICE+PMM	2.63	2.24	1.65	1.28	2.07	1.79	0.97	1.22	2.48	2.15	1.37	1.19
	MICE+RF	1.61	1.58	0.24	0.70	1.36	1.33	0.19	0.56	1.68	1.63	0.21	0.62
	rSVD01	5.35	5.26	0.32	0.62	3.15	3.09	0.08	0.36	3.49	3.37	0.19	0.48
	rSVD84	9.46	9.32	0.89	1.14	4.13	4.10	0.13	0.45	3.52	3.44	0.13	0.42
	rSVD84-01	5.82	5.76	0.35	0.75	3.30	3.26	0.10	0.33	3.32	3.20	0.15	0.44

Table 47. V by percentage and method — Mechanism **MCAR** (REP 11–15)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	1.22e+05	1.19e+05	5.34e+08	2.96e+04	1.21e+05	1.19e+05	3.57e+08	2.40e+04	1.26e+05	1.19e+05	6.12e+08	2.57e+04
	MICE+PMM	1.45e+05	1.46e+05	1.09e+09	4.55e+04	1.51e+05	1.49e+05	4.77e+08	2.59e+04	1.51e+05	1.45e+05	7.28e+08	3.01e+04
	MICE+RF	1.49e+05	1.53e+05	1.76e+09	6.03e+04	1.44e+05	1.46e+05	1.85e+09	5.20e+04	1.26e+05	1.30e+05	1.47e+09	5.25e+04
	rSVD01	4.43e+05	4.44e+05	8.44e+08	4.38e+04	4.92e+05	4.90e+05	1.33e+09	4.97e+04	5.81e+05	5.77e+05	1.70e+09	5.08e+04
	rSVD84	4.29e+05	4.27e+05	5.40e+08	3.05e+04	4.75e+05	4.67e+05	9.76e+08	5.10e+04	5.43e+05	5.39e+05	1.75e+09	5.08e+04
	rSVD84-01	4.16e+05	4.23e+05	6.11e+08	3.91e+04	4.68e+05	4.68e+05	1.25e+09	5.74e+04	5.37e+05	5.34e+05	1.71e+09	5.91e+04
Filho et al. (2008) Cotton 17×23	ABB	0.39	0.38	0.01	0.14	0.38	0.34	0.01	0.09	0.37	0.35	0.01	0.09
	MICE+PMM	0.48	0.48	0.01	0.16	0.48	0.44	0.02	0.13	0.45	0.43	0.01	0.12
	MICE+RF	0.47	0.47	0.02	0.21	0.59	0.59	0.04	0.21	0.37	0.36	0.01	0.13
	rSVD01	0.46	0.46	0.00	0.05	0.52	0.52	0.00	0.05	0.59	0.59	0.00	0.07
	rSVD84	0.45	0.46	0.00	0.05	0.50	0.50	0.00	0.04	0.57	0.57	0.00	0.07
	rSVD84-01	0.44	0.44	0.00	0.05	0.49	0.49	0.00	0.05	0.55	0.55	0.00	0.06
Flores et al. (1998) Faba beans 15×12	ABB	2.76e+05	2.65e+05	5.57e+09	7.43e+04	2.86e+05	2.77e+05	3.01e+09	6.31e+04	3.12e+05	2.99e+05	4.77e+09	1.13e+05
	MICE+PMM	2.80e+05	2.58e+05	9.87e+09	1.17e+05	2.84e+05	2.80e+05	3.44e+09	6.83e+04	3.08e+05	2.89e+05	4.93e+09	1.07e+05
	MICE+RF	3.93e+05	3.84e+05	1.30e+10	1.71e+05	1.77e+05	1.74e+05	2.51e+09	6.88e+04	2.92e+05	2.95e+05	8.84e+09	1.39e+05
	rSVD01	2.81e+05	2.78e+05	1.19e+09	4.48e+04	3.23e+05	3.23e+05	1.94e+09	4.99e+04	4.14e+05	4.07e+05	5.75e+09	1.04e+05
	rSVD84	2.83e+05	2.79e+05	1.15e+09	4.36e+04	3.42e+05	3.37e+05	2.16e+09	6.50e+04	4.28e+05	4.24e+05	4.90e+09	8.99e+04
	rSVD84-01	2.77e+05	2.77e+05	1.03e+09	4.50e+04	3.23e+05	3.19e+05	1.91e+09	4.58e+04	3.98e+05	3.93e+05	4.25e+09	7.97e+04
Mattos et al. (2013) Sugarcane 22×5	ABB	4.10	3.73	1.64	1.25	4.13	3.93	1.07	1.28	4.68	4.26	2.13	1.69
	MICE+PMM	3.50	3.29	1.16	1.37	3.51	3.39	0.68	0.98	3.82	3.40	1.83	1.52
	MICE+RF	3.52	3.50	1.00	1.28	2.87	2.84	0.87	1.23	3.50	3.48	1.44	1.37
	rSVD01	7.86	7.60	2.42	1.49	8.37	8.23	1.63	1.56	9.64	9.23	3.90	2.25
	rSVD84	6.55	6.32	1.26	0.98	7.37	7.12	1.87	1.80	9.47	8.96	4.89	3.00
	rSVD84-01	6.77	6.66	1.68	1.46	7.41	7.35	1.29	1.37	9.02	8.84	3.77	1.90
Rad et al. (2013) Wheat 36×6	ABB	5.15	4.95	1.15	0.97	4.82	4.62	0.78	1.16	4.92	4.87	0.61	1.06
	MICE+PMM	2.22	2.12	0.98	1.02	2.29	2.18	0.87	0.97	2.60	2.51	0.59	1.09
	MICE+RF	1.62	1.57	0.30	0.75	1.84	1.79	0.30	0.74	2.42	2.40	0.32	0.79
	rSVD01	6.92	6.89	0.39	0.84	7.18	7.13	0.52	0.89	7.59	7.47	0.96	1.17
	rSVD84	6.30	6.23	0.39	0.70	6.55	6.50	0.45	0.76	7.08	6.99	0.81	1.33
	rSVD84-01	6.44	6.38	0.41	0.65	6.67	6.57	0.53	0.92	7.15	7.03	0.68	1.01

Table 48. V by percentage and method — Mechanism **MNAR** (REP 11–15)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	8.77e+04	8.38e+04	1.82e+08	1.35e+04	7.53e+04	7.43e+04	1.22e+08	1.68e+04	6.18e+04	6.01e+04	1.79e+08	1.65e+04
	MICE+PMM	1.09e+05	1.05e+05	4.23e+08	2.65e+04	9.69e+04	9.60e+04	1.57e+08	1.89e+04	7.41e+04	7.12e+04	2.43e+08	1.98e+04
	MICE+RF	1.08e+05	1.04e+05	1.09e+09	4.53e+04	9.07e+04	9.04e+04	8.59e+08	4.48e+04	8.52e+04	8.49e+04	6.58e+08	3.50e+04
	rSVD01	4.46e+05	4.45e+05	2.55e+08	2.18e+04	5.13e+05	5.14e+05	3.84e+08	2.73e+04	6.31e+05	6.33e+05	6.64e+08	3.43e+04
	rSVD84	4.08e+05	4.06e+05	1.24e+08	1.47e+04	4.73e+05	4.73e+05	1.51e+08	1.51e+04	5.62e+05	5.62e+05	1.59e+08	1.60e+04
	rSVD84-01	4.25e+05	4.27e+05	3.81e+08	2.85e+04	4.61e+05	4.59e+05	3.41e+08	2.54e+04	5.66e+05	5.66e+05	4.59e+08	2.99e+04
Filho et al. (2008) Cotton 17×23	ABB	0.26	0.25	0.00	0.08	0.21	0.21	0.00	0.08	0.20	0.19	0.00	0.07
	MICE+PMM	0.31	0.29	0.01	0.09	0.25	0.24	0.01	0.09	0.24	0.23	0.00	0.09
	MICE+RF	0.45	0.43	0.02	0.22	0.20	0.20	0.00	0.09	0.18	0.18	0.00	0.08
	rSVD01	0.47	0.47	0.00	0.03	0.57	0.57	0.00	0.03	0.65	0.65	0.00	0.03
	rSVD84	0.43	0.43	0.00	0.01	0.52	0.52	0.00	0.02	0.65	0.64	0.00	0.02
	rSVD84-01	0.43	0.43	0.00	0.03	0.53	0.53	0.00	0.04	0.60	0.60	0.00	0.04
Flores et al. (1998) Faba beans 15×12	ABB	1.98e+05	1.92e+05	1.74e+09	4.75e+04	1.78e+05	1.78e+05	1.71e+09	4.98e+04	1.35e+05	1.27e+05	1.51e+09	5.80e+04
	MICE+PMM	1.63e+05	1.55e+05	2.14e+09	4.62e+04	1.77e+05	1.65e+05	3.03e+09	8.10e+04	1.27e+05	1.21e+05	1.45e+09	5.18e+04
	MICE+RF	1.40e+05	1.40e+05	1.22e+09	4.68e+04	1.51e+05	1.56e+05	2.54e+09	7.29e+04	1.33e+05	1.33e+05	1.29e+09	4.62e+04
	rSVD01	3.83e+05	3.79e+05	1.83e+09	5.23e+04	3.37e+05	3.30e+05	1.85e+09	4.73e+04	7.46e+05	7.41e+05	6.30e+09	8.45e+04
	rSVD84	3.11e+05	3.11e+05	3.06e+08	2.47e+04	3.44e+05	3.38e+05	1.20e+09	2.76e+04	6.49e+05	6.44e+05	1.92e+09	4.12e+04
	rSVD84-01	3.15e+05	3.12e+05	1.99e+09	4.48e+04	3.46e+05	3.46e+05	2.77e+09	5.89e+04	6.31e+05	6.25e+05	4.07e+09	7.84e+04
Mattos et al. (2013) Sugarcane 22×5	ABB	3.46	3.31	0.89	1.10	3.03	2.90	0.66	1.21	2.45	2.27	0.53	0.99
	MICE+PMM	3.09	2.80	0.96	1.14	2.69	2.45	0.70	0.93	2.01	1.90	0.38	0.87
	MICE+RF	2.14	2.16	0.38	0.85	1.69	1.66	0.26	0.62	1.61	1.64	0.22	0.62
	rSVD01	7.39	6.94	2.03	1.48	8.08	7.89	0.98	1.15	8.74	8.64	0.63	0.98
	rSVD84	6.59	6.33	1.15	1.18	10.19	9.88	1.35	1.62	16.07	15.53	2.35	1.88
	rSVD84-01	6.81	6.48	1.54	1.44	8.33	8.09	1.17	1.33	9.97	9.99	0.90	1.20
Rad et al. (2013) Wheat 36×6	ABB	3.46	3.38	0.35	0.71	2.50	2.38	0.33	0.66	2.16	2.08	0.28	0.61
	MICE+PMM	2.68	2.49	0.87	0.95	1.69	1.60	0.27	0.66	1.58	1.42	0.48	0.73
	MICE+RF	2.93	2.96	0.86	1.30	2.31	2.29	0.41	0.92	0.95	0.94	0.09	0.40
	rSVD01	2.62	2.57	0.17	0.42	3.26	3.15	0.20	0.49	3.10	2.95	0.24	0.45
	rSVD84	7.38	7.00	1.18	1.22	3.91	3.81	0.19	0.44	2.53	2.47	0.11	0.25
	rSVD84-01	3.69	3.59	0.26	0.52	3.07	3.01	0.24	0.52	2.75	2.66	0.13	0.37

Table 49. V by percentage and method — Mechanism **MAR** (REP 16–17)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	5.17	4.91	1.75	1.51	4.92	4.70	1.52	1.86	4.74	4.49	1.60	1.83
	MICE+PMM	2.65	2.29	3.05	1.55	2.88	2.65	2.60	1.48	3.19	2.62	2.45	1.90
	MICE+RF	1.24	1.29	0.14	0.54	3.24	3.25	0.99	1.33	1.63	1.59	0.20	0.53
	rSVD01	3.13	3.04	0.61	1.03	2.85	2.76	0.36	0.76	2.23	2.16	0.14	0.44
	rSVD84	4.69	4.51	0.84	1.14	4.02	3.96	0.35	0.77	1.68	1.61	0.09	0.29
	rSVD84-01	3.48	3.45	0.40	0.92	2.95	2.92	0.21	0.59	2.00	1.94	0.13	0.44
Zobel et al. (1988) Soybean 7×35	ABB	6.77e+05	6.42e+05	1.99e+10	1.74e+05	5.68e+05	5.55e+05	1.37e+10	1.49e+05	5.43e+05	5.39e+05	1.97e+10	1.74e+05
	MICE+PMM	3.06e+05	2.57e+05	2.40e+10	1.33e+05	2.78e+05	2.49e+05	1.02e+10	1.17e+05	3.05e+05	2.59e+05	2.80e+10	1.80e+05
	MICE+RF	4.38e+05	4.33e+05	1.87e+10	1.83e+05	2.08e+05	2.13e+05	3.86e+09	7.85e+04	6.71e+05	6.52e+05	4.51e+10	3.03e+05
	rSVD01	5.79e+05	5.70e+05	7.40e+09	9.17e+04	6.00e+05	5.85e+05	7.09e+09	1.01e+05	6.30e+05	6.09e+05	5.83e+09	1.01e+05
	rSVD84	5.35e+05	5.21e+05	3.41e+09	6.72e+04	6.30e+05	6.19e+05	6.21e+09	8.65e+04	7.09e+05	6.99e+05	5.46e+09	1.01e+05
	rSVD84-01	5.46e+05	5.35e+05	4.36e+09	9.02e+04	5.99e+05	5.87e+05	5.25e+09	9.10e+04	6.42e+05	6.40e+05	4.36e+09	7.39e+04

Table 50. V by percentage and method — Mechanism **MCAR** (REP 16–17)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	5.35	5.08	1.64	1.74	5.61	5.53	1.29	1.20	5.90	5.59	1.82	1.59
	MICE+PMM	2.13	1.78	1.61	1.47	2.59	2.24	1.43	1.52	3.28	3.03	1.62	1.46
	MICE+RF	1.95	2.00	0.38	0.84	3.53	3.39	1.10	1.44	4.32	4.28	1.81	1.92
	rSVD01	4.60	4.54	0.33	0.69	4.94	4.72	0.91	0.96	5.14	4.80	0.98	1.44
	rSVD84	4.19	4.09	0.29	0.69	4.42	4.23	0.55	0.72	5.18	4.99	0.70	1.06
	rSVD84-01	4.39	4.28	0.39	0.69	4.73	4.52	0.79	0.96	4.96	4.86	0.72	1.02
Zobel et al. (1988) Soybean 7×35	ABB	8.04e+05	7.97e+05	1.29e+10	1.50e+05	7.95e+05	7.91e+05	6.90e+09	1.19e+05	7.47e+05	7.35e+05	7.94e+09	1.09e+05
	MICE+PMM	3.96e+05	3.99e+05	1.86e+10	1.94e+05	4.17e+05	4.05e+05	1.32e+10	1.34e+05	3.78e+05	3.68e+05	9.08e+09	1.18e+05
	MICE+RF	4.80e+05	4.72e+05	2.49e+10	2.25e+05	4.83e+05	4.85e+05	2.05e+10	1.77e+05	3.04e+05	2.97e+05	7.15e+09	1.33e+05
	rSVD01	5.66e+05	5.60e+05	3.25e+09	5.74e+04	6.17e+05	6.03e+05	5.79e+09	6.15e+04	6.31e+05	6.15e+05	5.56e+09	7.55e+04
	rSVD84	5.35e+05	5.29e+05	1.90e+09	6.26e+04	5.72e+05	5.69e+05	3.80e+09	6.92e+04	6.07e+05	5.91e+05	5.69e+09	6.98e+04
	rSVD84-01	5.57e+05	5.49e+05	3.07e+09	5.99e+04	5.85e+05	5.82e+05	2.27e+09	6.45e+04	6.16e+05	6.10e+05	5.03e+09	8.51e+04

Table 51. V by percentage and method — Mechanism **MNAR** (REP 16–17)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	4.91	4.84	0.69	1.04	4.66	4.54	0.42	0.79	4.62	4.57	0.37	0.81
	MICE+PMM	2.70	2.53	2.16	1.29	3.01	2.86	0.70	1.11	2.91	2.76	0.53	0.89
	MICE+RF	2.70	2.66	0.68	1.23	2.95	3.02	0.92	1.42	3.43	3.39	1.17	1.32
	rSVD01	3.08	2.85	0.75	1.26	0.89	0.77	0.11	0.29	1.72	1.60	0.21	0.44
	rSVD84	5.15	4.96	1.18	1.24	0.87	0.76	0.13	0.34	2.40	2.28	0.18	0.48
	rSVD84-01	3.64	3.57	0.62	1.01	0.86	0.79	0.07	0.33	1.75	1.63	0.21	0.41
Zobel et al. (1988) Soybean 7×35	ABB	6.30e+05	6.33e+05	5.12e+09	8.54e+04	5.75e+05	5.69e+05	3.56e+09	5.76e+04	4.73e+05	4.71e+05	3.89e+09	7.87e+04
	MICE+PMM	2.92e+05	2.72e+05	8.70e+09	1.34e+05	3.34e+05	3.21e+05	4.29e+09	8.26e+04	3.15e+05	3.04e+05	4.67e+09	7.66e+04
	MICE+RF	3.63e+05	3.51e+05	1.19e+10	1.32e+05	4.09e+05	4.17e+05	1.54e+10	1.71e+05	4.27e+05	4.31e+05	1.45e+10	1.71e+05
	rSVD01	4.89e+05	4.83e+05	6.57e+09	9.82e+04	4.60e+05	4.37e+05	9.24e+09	9.52e+04	2.69e+05	2.53e+05	3.82e+09	6.77e+04
	rSVD84	4.96e+05	4.83e+05	2.66e+09	6.30e+04	3.95e+05	3.72e+05	4.82e+09	7.86e+04	3.18e+05	3.04e+05	2.71e+09	5.12e+04
	rSVD84-01	4.96e+05	4.77e+05	5.48e+09	9.61e+04	4.03e+05	3.84e+05	5.89e+09	7.42e+04	2.67e+05	2.57e+05	2.63e+09	5.67e+04

Table 52. V by percentage and method — Mechanism **MAR** (REP 1–5)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	8.78e+03	3.16e+03	1.26e+08	1.21e+04	8.43e+03	3.82e+03	1.08e+08	8.88e+03	1.07e+04	3.61e+03	1.98e+08	8.22e+03
	MICE+PMM	1.11e+04	2.98e+03	4.40e+08	7.81e+03	1.00e+04	3.70e+03	2.39e+08	6.41e+03	1.30e+04	3.54e+03	4.06e+08	9.50e+03
	MICE+RF	8.46e+04	8.48e+04	5.79e+08	2.68e+04	4.68e+03	4.66e+03	2.09e+06	1.92e+03	4.74e+03	4.89e+03	2.27e+06	2.10e+03
	rSVD01	2.46e+04	3.30e+03	2.48e+10	6.93e+03	6.85e+04	4.91e+03	3.67e+10	2.67e+04	8.84e+04	1.04e+04	1.01e+11	2.25e+04
	rSVD84	4.26e+03	1.08e+03	5.00e+07	3.74e+03	5.69e+03	2.27e+03	1.11e+08	4.57e+03	7.03e+03	3.51e+03	6.79e+07	9.43e+03
	rSVD84-01	6.93e+03	1.58e+03	4.21e+08	3.94e+03	5.05e+04	3.61e+03	1.17e+11	1.16e+04	3.98e+04	6.12e+03	5.44e+10	1.28e+04
Bro and Smilde (2014) Red Wine 14×44	ABB	601.74	499.48	3.44e+05	883.35	426.43	333.19	1.95e+05	609.67	449.92	469.12	1.69e+05	648.21
	MICE+PMM	558.21	220.41	6.90e+05	641.71	542.77	140.87	4.30e+05	873.74	606.06	370.90	3.90e+05	986.25
	MICE+RF	3.44e+03	3.41e+03	1.28e+06	1.36e+03	22.40	23.01	35.96	8.19	381.46	390.69	1.20e+04	149.84
	rSVD01	7.17e+06	1.82e+04	3.88e+15	6.85e+04	1.50e+06	9.98e+04	6.33e+13	3.51e+05	6.19e+06	2.15e+05	2.55e+15	9.47e+05
	rSVD84	32.21	25.39	554.14	7.41	65.21	52.59	1.63e+03	21.07	5.73e+03	446.48	8.23e+08	210.21
	rSVD84-01	1.56e+05	4.99e+03	4.01e+11	1.96e+04	2.77e+05	1.75e+04	7.27e+11	8.58e+04	1.53e+07	4.76e+04	1.71e+16	4.67e+05
Forkman and Piepho (2014) Soybean 7×10	ABB	2.87e+07	2.53e+07	6.88e+14	3.27e+07	2.96e+07	3.10e+07	6.54e+14	4.16e+07	2.78e+07	1.86e+07	6.83e+14	4.57e+07
	MICE+PMM	1.73e+07	8.72e+06	3.58e+14	2.52e+07	2.17e+07	1.30e+07	4.50e+14	3.35e+07	1.95e+07	5.79e+06	7.75e+14	3.10e+07
	MICE+RF	3.67e+06	3.59e+06	9.49e+11	1.23e+06	3.57e+07	3.50e+07	1.28e+14	1.66e+07	5.96e+05	5.79e+05	2.97e+10	2.45e+05
	rSVD01	4.66e+06	4.58e+05	1.32e+14	2.28e+06	8.14e+06	5.15e+05	3.79e+14	3.60e+06	1.28e+07	3.33e+06	5.56e+14	1.26e+07
	rSVD84	2.04e+06	4.66e+05	3.05e+13	1.21e+05	3.00e+06	4.64e+05	4.99e+13	1.21e+05	6.60e+06	1.80e+06	1.40e+14	4.79e+06
	rSVD84-01	3.89e+06	4.12e+05	8.20e+13	4.85e+05	5.06e+06	5.97e+05	1.47e+14	2.68e+06	9.62e+06	2.56e+06	4.28e+14	7.46e+06
Forkman and Piepho (2014) Pea 5×4	ABB	1.11e+08	1.28e+08	5.07e+15	8.65e+07	1.19e+08	1.25e+08	6.10e+15	6.83e+07	1.09e+08	1.12e+08	6.32e+15	1.67e+08
	MICE+PMM	1.69e+08	1.68e+08	1.32e+16	1.84e+08	1.73e+08	1.70e+08	1.32e+16	1.70e+08	1.69e+08	1.16e+05	3.73e+16	3.57e+08
	MICE+RF	1.63e+08	1.67e+08	2.58e+15	6.22e+07	2.42e+08	2.47e+08	4.90e+15	8.72e+07	4.33e+08	4.39e+08	1.80e+16	1.62e+08
	rSVD01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.32e+08	7.68e+07	4.52e+17	6.19e+07
	rSVD84	1.57e+08	1.34e+08	9.89e+15	1.24e+08	3.84e+08	1.33e+08	4.28e+17	3.02e+08	4.07e+08	1.86e+08	3.07e+17	3.20e+08
	rSVD84-01	5.84e+07	3.47e+07	2.41e+16	3.48e+07	5.22e+07	2.91e+07	1.69e+16	2.54e+07	2.63e+08	1.24e+08	2.18e+17	1.73e+08
Cornelius et al. (1996) Maize 9×20	ABB	3.08e+08	2.90e+08	2.63e+16	2.44e+08	3.45e+08	3.08e+08	3.82e+16	2.75e+08	3.66e+08	3.26e+08	4.76e+16	3.15e+08
	MICE+PMM	3.41e+08	2.56e+08	9.06e+16	4.28e+08	4.19e+08	3.84e+08	7.75e+16	4.15e+08	4.58e+08	3.83e+08	1.10e+17	4.37e+08
	MICE+RF	5.87e+08	5.89e+08	3.05e+16	2.57e+08	1.04e+09	9.95e+08	1.14e+17	4.93e+08	5.72e+08	5.71e+08	2.62e+16	2.32e+08
	rSVD01	2.48e+08	1.93e+08	6.82e+16	2.09e+08	7.67e+07	6.63e+07	4.29e+15	8.12e+07	2.16e+08	1.36e+08	5.68e+16	1.38e+08
	rSVD84	1.16e+08	7.28e+07	2.86e+16	8.71e+07	1.31e+08	1.19e+08	1.07e+16	1.37e+08	1.95e+08	1.34e+08	3.19e+16	1.61e+08
	rSVD84-01	1.74e+08	1.04e+08	4.65e+16	1.70e+08	7.83e+07	6.10e+07	4.99e+15	7.01e+07	1.59e+08	1.26e+08	1.93e+16	1.13e+08

Table 53. V by percentage and method — Mechanism **MCAR** (REP 1–5)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	6.83e+03	3.40e+03	8.69e+07	8.24e+03	7.57e+03	5.08e+03	6.83e+07	1.10e+04	1.08e+04	9.33e+03	8.78e+07	1.49e+04
	MICE+PMM	6.06e+03	2.61e+03	7.83e+07	7.17e+03	7.18e+03	3.03e+03	8.80e+07	9.34e+03	9.70e+03	7.70e+03	9.37e+07	1.25e+04
	MICE+RF	241.56	248.12	4.52e+03	86.64	1.72e+04	1.76e+04	2.33e+07	6.33e+03	7.55e+03	7.81e+03	6.53e+06	3.72e+03
	rSVD01	4.51e+03	2.64e+03	3.00e+07	4.84e+03	8.17e+03	3.26e+03	2.99e+08	5.66e+03	2.31e+04	6.41e+03	2.34e+09	1.72e+04
	rSVD84	3.25e+03	1.26e+03	2.76e+07	3.03e+03	3.55e+03	2.11e+03	2.47e+07	3.72e+03	5.99e+03	3.79e+03	4.37e+07	7.34e+03
	rSVD84-01	3.50e+03	1.98e+03	2.63e+07	3.41e+03	5.69e+03	2.57e+03	1.04e+08	5.68e+03	2.95e+04	4.92e+03	2.29e+10	1.01e+04
Bro and Smilde (2014) Red Wine 14×44	ABB	2.13e+05	3.17e+04	7.16e+10	4.07e+05	2.48e+05	1.82e+05	8.52e+10	4.35e+05	2.81e+05	1.95e+05	1.71e+11	4.28e+05
	MICE+PMM	2.48e+05	4.08e+03	1.79e+11	4.14e+05	2.33e+05	9.07e+04	1.07e+11	3.64e+05	2.74e+05	2.21e+04	1.46e+11	4.90e+05
	MICE+RF	1.62e+03	1.62e+03	2.60e+05	690.82	8.54e+04	8.56e+04	6.55e+08	3.72e+04	3.81e+05	3.75e+05	1.11e+10	1.30e+05
	rSVD01	2.26e+06	1.69e+04	3.81e+14	1.02e+05	2.90e+06	4.84e+04	4.15e+14	1.82e+05	1.14e+08	4.52e+04	1.08e+18	4.70e+05
	rSVD84	34.70	24.93	1.27e+03	16.75	1.70e+03	69.11	1.49e+08	91.47	2.02e+03	169.20	1.10e+08	147.32
	rSVD84-01	9.71e+04	4.00e+03	8.55e+10	3.27e+04	1.83e+05	1.73e+04	3.69e+11	1.13e+05	3.77e+05	1.39e+04	2.81e+12	5.99e+04
Forkman and Piepho (2014) Soybean 7×10	ABB	2.18e+07	1.91e+05	1.81e+15	2.24e+07	2.82e+07	1.16e+07	1.56e+15	4.87e+07	2.56e+07	1.72e+07	7.73e+14	3.16e+07
	MICE+PMM	1.25e+07	1.17e+05	7.41e+14	1.19e+07	2.35e+07	6.07e+06	1.35e+15	3.18e+07	1.97e+07	6.82e+06	9.16e+14	3.12e+07
	MICE+RF	3.35e+07	3.34e+07	1.01e+14	1.52e+07	3.56e+07	3.48e+07	1.20e+14	1.53e+07	5.41e+07	5.38e+07	3.05e+14	2.36e+07
	rSVD01	2.44e+06	4.21e+05	7.10e+13	1.26e+05	4.82e+06	5.45e+05	9.01e+13	4.69e+06	1.06e+07	5.24e+06	1.77e+14	1.39e+07
	rSVD84	1.04e+06	4.10e+05	5.43e+12	1.17e+05	1.95e+06	5.84e+05	1.61e+13	1.15e+06	4.33e+06	1.51e+06	4.14e+13	4.55e+06
	rSVD84-01	2.20e+06	3.88e+05	3.80e+13	1.88e+05	3.29e+06	1.25e+06	2.43e+13	3.52e+06	7.45e+06	3.34e+06	1.22e+14	9.14e+06
Forkman and Piepho (2014) Pea 5×4	ABB	1.13e+08	2.21e+05	3.96e+16	2.36e+08	9.76e+07	5.10e+07	1.49e+16	1.54e+08	1.15e+08	9.41e+07	8.68e+15	9.74e+07
	MICE+PMM	1.73e+08	1.72e+05	1.05e+17	1.85e+08	1.64e+08	2.63e+05	7.96e+16	3.14e+08	1.85e+08	2.78e+05	1.01e+17	2.71e+08
	MICE+RF	1.14e+06	1.89e+03	5.97e+12	7.13e+05	1.61e+06	8.92e+04	7.00e+12	2.36e+06	2.57e+05	2.48e+05	6.38e+09	8.99e+04
	rSVD01	5.97e+08	1.35e+08	3.76e+18	1.15e+08	8.35e+08	1.79e+08	4.91e+18	1.70e+08	5.17e+08	1.86e+08	9.68e+17	2.51e+08
	rSVD84	2.16e+08	7.57e+07	7.62e+17	6.44e+07	2.70e+08	1.16e+08	1.82e+17	1.61e+08	7.01e+08	2.03e+08	1.56e+18	6.44e+08
	rSVD84-01	1.28e+08	8.49e+07	4.34e+16	6.76e+07	3.13e+08	1.36e+08	2.17e+17	1.19e+08	2.60e+08	1.35e+08	1.04e+17	1.92e+08
Cornelius et al. (1996) Maize 9×20	ABB	3.28e+08	3.05e+08	3.07e+16	2.48e+08	3.77e+08	3.68e+08	2.56e+16	1.62e+08	3.96e+08	3.76e+08	2.23e+16	1.49e+08
	MICE+PMM	3.26e+08	2.75e+08	5.22e+16	3.05e+08	3.07e+08	2.91e+08	2.43e+16	1.71e+08	3.24e+08	3.13e+08	2.23e+16	1.97e+08
	MICE+RF	3.93e+08	3.87e+08	1.26e+16	1.55e+08	2.84e+08	2.73e+08	6.28e+15	9.96e+07	1.31e+08	1.31e+08	1.20e+15	4.57e+07
	rSVD01	3.78e+07	1.87e+07	2.09e+15	5.87e+07	8.36e+07	6.53e+07	4.10e+15	8.22e+07	1.47e+08	1.37e+08	9.84e+15	1.18e+08
	rSVD84	1.66e+07	7.96e+06	4.54e+14	2.05e+07	4.74e+07	3.80e+07	1.67e+15	5.06e+07	8.54e+07	7.36e+07	3.30e+15	6.95e+07
	rSVD84-01	2.39e+07	1.52e+07	8.55e+14	2.97e+07	5.66e+07	4.37e+07	2.46e+15	5.97e+07	1.06e+08	8.63e+07	5.16e+15	9.29e+07

Table 54. V by percentage and method — Mechanism **MNAR** (REP 1–5)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Krzanowski (1988) Oil 5×20	ABB	8.00e+03	6.14e+03	4.44e+07	1.06e+04	8.68e+03	7.26e+03	5.18e+07	1.09e+04	1.02e+04	9.30e+03	5.80e+07	1.32e+04
	MICE+PMM	1.08e+04	4.96e+03	1.70e+08	1.43e+04	8.39e+03	5.73e+03	7.45e+07	1.00e+04	1.17e+04	7.73e+03	1.10e+08	1.68e+04
	MICE+RF	1.39e+03	1.42e+03	1.92e+05	485.51	6.66e+03	6.63e+03	4.53e+06	2.82e+03	1.72e+04	1.77e+04	3.00e+07	6.18e+03
	rSVD01	2.49e+04	4.47e+03	2.15e+10	7.57e+03	9.96e+05	4.54e+03	7.53e+13	9.03e+03	2.90e+04	7.40e+03	5.04e+09	2.13e+04
	rSVD84	3.07e+03	1.40e+03	1.48e+07	3.75e+03	3.65e+03	2.16e+03	2.00e+07	4.26e+03	5.75e+03	3.47e+03	5.61e+07	4.99e+03
	rSVD84-01	3.33e+05	2.58e+03	1.06e+13	5.75e+03	6.42e+04	2.81e+03	3.52e+11	5.80e+03	1.44e+04	4.29e+03	2.34e+09	1.08e+04
Bro and Smilde (2014) Red Wine 14×44	ABB	3.72e+05	3.46e+05	1.34e+11	6.09e+05	2.91e+05	2.80e+05	1.02e+11	4.97e+05	2.89e+05	4.50e+03	1.36e+11	5.03e+05
	MICE+PMM	4.14e+05	1.69e+05	3.23e+11	6.33e+05	3.45e+05	1.91e+04	2.14e+11	6.34e+05	3.46e+05	3.79e+03	2.15e+11	6.64e+05
	MICE+RF	3.88e+05	3.85e+05	1.56e+10	1.73e+05	1.08e+06	1.08e+06	8.39e+10	4.00e+05	142.80	141.89	1.78e+03	52.20
	rSVD01	9.89e+05	8.66e+04	9.95e+12	1.96e+05	1.37e+06	4.46e+04	6.69e+13	1.54e+05	2.45e+06	8.49e+04	1.03e+14	8.61e+05
	rSVD84	75.34	39.00	9.56e+04	9.91	1.87e+03	40.00	6.42e+07	15.87	1.08e+04	408.55	2.61e+09	120.37
	rSVD84-01	1.77e+07	8.76e+03	2.97e+16	6.83e+04	2.32e+05	1.08e+04	1.05e+12	5.46e+04	7.31e+05	2.92e+04	9.48e+12	1.53e+05
Forkman and Piepho (2014) Soybean 7×10	ABB	2.81e+07	2.63e+07	6.59e+14	2.93e+07	3.34e+07	3.23e+07	8.23e+14	4.18e+07	2.76e+07	2.01e+07	7.41e+14	3.49e+07
	MICE+PMM	2.18e+07	1.14e+07	5.93e+14	2.85e+07	2.19e+07	1.97e+07	4.25e+14	2.89e+07	1.69e+07	8.02e+06	3.30e+14	2.41e+07
	MICE+RF	3.33e+06	3.39e+06	1.08e+12	1.35e+06	5.30e+04	5.09e+04	2.33e+08	2.15e+04	1.64e+07	1.61e+07	2.33e+13	5.72e+06
	rSVD01	3.28e+06	4.13e+05	7.62e+13	4.01e+05	9.70e+06	3.61e+06	1.50e+14	1.64e+07	1.07e+07	2.59e+06	2.48e+14	1.28e+07
	rSVD84	1.45e+06	4.30e+05	6.15e+12	1.12e+05	6.10e+06	5.48e+05	7.65e+13	1.24e+07	4.97e+06	1.31e+06	6.00e+13	3.28e+06
	rSVD84-01	2.61e+06	4.06e+05	3.11e+13	1.38e+06	7.39e+06	1.72e+06	1.57e+14	6.18e+06	6.43e+06	9.86e+05	2.04e+14	4.22e+06
Forkman and Piepho (2014) Pea 5×4	ABB	1.29e+08	1.21e+08	1.08e+15	5.93e+07	1.21e+08	1.08e+08	1.12e+15	5.32e+07	1.26e+08	1.18e+08	1.13e+15	5.36e+07
	MICE+PMM	1.72e+08	1.84e+08	2.14e+15	8.46e+07	1.64e+08	1.71e+08	2.24e+15	8.77e+07	1.71e+08	1.79e+08	2.12e+15	8.29e+07
	MICE+RF	1.20e+08	1.19e+08	1.21e+15	4.90e+07	1.20e+08	1.22e+08	1.11e+15	3.77e+07	1.28e+08	1.26e+08	1.60e+15	5.70e+07
	rSVD01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	rSVD84	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	rSVD84-01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cornelius et al. (1996) Maize 9×20	ABB	3.31e+08	3.23e+08	8.75e+15	1.15e+08	3.52e+08	3.50e+08	8.27e+15	1.24e+08	4.01e+08	4.04e+08	8.41e+15	1.15e+08
	MICE+PMM	3.98e+08	3.87e+08	3.60e+16	2.83e+08	3.45e+08	3.36e+08	2.00e+16	1.88e+08	3.67e+08	3.65e+08	2.22e+16	1.62e+08
	MICE+RF	3.10e+08	3.08e+08	7.97e+15	1.32e+08	3.93e+08	4.13e+08	1.50e+16	1.90e+08	2.78e+08	2.81e+08	5.38e+15	8.83e+07
	rSVD01	4.12e+07	4.82e+06	2.38e+15	7.40e+07	3.08e+07	2.05e+07	1.78e+15	4.38e+07	7.43e+07	4.63e+07	1.01e+16	6.13e+07
	rSVD84	1.74e+08	1.58e+08	2.69e+16	2.25e+08	1.05e+08	9.37e+07	9.07e+15	1.30e+08	1.02e+08	7.62e+07	8.72e+15	1.10e+08
	rSVD84-01	5.43e+07	4.05e+07	3.18e+15	7.47e+07	4.46e+07	3.72e+07	1.43e+15	5.22e+07	1.04e+08	5.38e+07	1.26e+17	4.98e+07

Table 55. V by percentage and method — Mechanism **MAR** (REP 6–10)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	5.06e+03	3.88e+03	1.16e+07	4.31e+03	4.08e+03	2.99e+03	1.07e+07	4.58e+03	4.54e+03	3.49e+03	1.48e+07	3.32e+03
	MICE+PMM	5.75e+03	3.81e+03	3.21e+07	4.87e+03	4.34e+03	2.50e+03	1.95e+07	5.28e+03	4.45e+03	3.24e+03	1.79e+07	2.91e+03
	MICE+RF	2.78e+03	2.70e+03	6.79e+05	991.73	2.60e+03	2.72e+03	5.20e+05	839.64	285.55	288.29	6.86e+03	112.61
	rSVD01	2.28e+03	1.52e+03	4.53e+06	2.54e+03	3.05e+03	2.15e+03	7.23e+06	2.85e+03	2.04e+03	1.81e+03	1.99e+06	1.65e+03
	rSVD84	1.80e+03	1.29e+03	2.88e+06	2.63e+03	1.55e+03	1.19e+03	1.64e+06	1.46e+03	1.95e+03	1.76e+03	1.90e+06	1.71e+03
	rSVD84-01	1.74e+03	929.61	3.25e+06	2.29e+03	2.30e+03	1.90e+03	4.28e+06	2.28e+03	1.92e+03	1.33e+03	2.28e+06	1.79e+03
Yan et al. (2007) Wheat 18×9	ABB	44.90	43.36	259.40	21.28	50.75	52.38	333.34	24.72	36.42	37.37	211.17	23.41
	MICE+PMM	51.92	53.75	896.99	43.78	57.89	53.73	1.01e+03	32.87	35.62	30.75	464.41	36.33
	MICE+RF	14.96	15.35	24.64	5.83	58.89	56.87	330.78	23.14	74.79	72.51	376.50	25.86
	rSVD01	46.85	34.22	1.39e+03	41.64	62.00	37.96	6.24e+03	41.35	32.69	27.45	445.61	22.65
	rSVD84	18.60	16.10	170.94	13.57	27.76	18.52	1.02e+03	18.96	15.45	10.87	188.36	13.19
	rSVD84-01	30.31	25.37	627.96	30.82	39.12	28.41	1.62e+03	28.15	23.00	17.58	313.64	17.91
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	330.79	316.37	1.51e+04	178.27	325.84	315.64	1.45e+04	157.68	326.24	293.51	2.82e+04	237.06
	MICE+PMM	289.71	252.02	3.83e+04	160.51	294.75	251.01	3.62e+04	209.82	356.84	273.24	6.88e+04	298.49
	MICE+RF	162.17	163.14	1.76e+03	47.90	586.84	572.59	2.98e+04	198.95	344.18	329.65	1.07e+04	129.25
	rSVD01	146.67	133.19	5.32e+03	93.87	138.18	132.21	2.64e+03	68.50	173.85	161.34	4.73e+03	80.25
	rSVD84	104.86	101.65	1.92e+03	52.33	122.90	117.19	1.85e+03	45.72	145.08	136.49	2.51e+03	58.83
	rSVD84-01	128.35	125.01	2.87e+03	61.07	117.19	114.43	1.49e+03	50.41	147.54	137.41	2.63e+03	58.15
Calinski et al. (2009a) Pea 18×9	ABB	1.70e+03	1.55e+03	9.44e+05	1.57e+03	2.15e+03	2.14e+03	1.47e+06	1.59e+03	1.47e+03	1.11e+03	1.16e+06	1.75e+03
	MICE+PMM	1.92e+03	1.47e+03	2.21e+06	1.96e+03	2.39e+03	2.01e+03	3.15e+06	2.53e+03	1.52e+03	1.07e+03	1.85e+06	1.94e+03
	MICE+RF	3.89e+03	3.90e+03	1.43e+06	1.36e+03	4.09e+03	4.08e+03	1.52e+06	1.64e+03	4.92e+03	4.93e+03	2.34e+06	2.25e+03
	rSVD01	304.16	161.48	1.10e+05	467.33	707.13	600.19	2.15e+05	535.69	935.08	892.73	2.44e+05	792.52
	rSVD84	142.89	81.14	2.17e+04	132.53	314.12	263.71	4.45e+04	225.93	394.06	347.18	4.04e+04	290.01
	rSVD84-01	209.62	111.82	6.10e+04	287.71	482.10	387.54	1.22e+05	436.27	653.24	614.79	1.16e+05	419.84
Calinski et al. (2009b) Winter rye 18×15	ABB	9.02e+03	8.97e+03	1.13e+07	4.57e+03	7.97e+03	7.86e+03	1.32e+07	5.53e+03	9.45e+03	9.63e+03	1.18e+07	5.14e+03
	MICE+PMM	9.58e+03	9.11e+03	2.67e+07	7.18e+03	9.80e+03	9.51e+03	2.97e+07	8.14e+03	1.05e+04	9.28e+03	2.25e+07	6.06e+03
	MICE+RF	7.46e+03	7.63e+03	5.38e+06	2.98e+03	3.77e+03	3.71e+03	9.88e+05	1.48e+03	1.57e+04	1.55e+04	2.65e+07	6.59e+03
	rSVD01	221.77	47.88	2.74e+05	8.88	1.14e+03	840.10	2.01e+06	1.57e+03	2.09e+03	1.35e+03	5.43e+06	2.93e+03
	rSVD84	82.23	48.35	3.79e+04	2.09	587.48	71.40	8.71e+05	436.55	1.32e+03	1.15e+03	1.48e+06	1.47e+03
	rSVD84-01	106.20	50.61	5.93e+04	6.10	775.96	74.07	1.00e+06	1.08e+03	1.44e+03	889.08	2.70e+06	1.91e+03

Table 56. V by percentage and method — Mechanism **MCAR** (REP 6–10)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	5.78e+03	4.78e+03	2.06e+07	4.02e+03	4.47e+03	3.38e+03	1.07e+07	3.64e+03	4.69e+03	4.06e+03	5.78e+06	3.18e+03
	MICE+PMM	4.62e+03	3.64e+03	1.30e+07	4.81e+03	4.03e+03	3.00e+03	1.16e+07	3.83e+03	4.00e+03	3.38e+03	5.22e+06	3.57e+03
	MICE+RF	2.19e+03	2.14e+03	3.66e+05	890.92	1.57e+03	1.59e+03	2.06e+05	642.26	2.11e+03	2.15e+03	3.48e+05	801.62
	rSVD01	2.25e+03	1.72e+03	3.34e+06	2.11e+03	2.48e+03	1.84e+03	6.39e+06	2.34e+03	3.08e+03	2.65e+03	4.35e+06	2.51e+03
	rSVD84	1.21e+03	825.46	1.48e+06	1.12e+03	1.39e+03	1.11e+03	1.61e+06	1.11e+03	1.86e+03	1.51e+03	1.72e+06	1.63e+03
	rSVD84-01	1.74e+03	1.26e+03	2.26e+06	1.78e+03	1.75e+03	1.30e+03	4.72e+06	1.44e+03	2.37e+03	2.02e+03	2.54e+06	1.78e+03
Yan et al. (2007) Wheat 18×9	ABB	50.09	50.07	846.83	36.37	57.13	54.28	560.01	33.37	39.62	36.91	332.89	23.90
	MICE+PMM	49.81	43.44	1.63e+03	49.73	53.48	52.22	929.33	37.84	36.02	32.48	382.03	24.66
	MICE+RF	31.48	30.69	103.82	12.97	24.64	24.79	63.75	9.80	20.99	21.14	44.17	9.24
	rSVD01	7.41	4.01	111.83	10.33	12.94	10.82	88.25	11.59	13.20	10.92	115.21	11.99
	rSVD84	5.69	2.13	104.11	5.36	7.85	6.01	50.85	8.35	8.18	5.82	50.14	7.86
	rSVD84-01	6.37	3.73	64.22	8.00	10.75	8.13	105.43	8.28	10.43	8.98	68.70	10.90
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	307.30	307.53	6.14e+03	97.08	297.77	288.58	3.50e+03	78.20	292.43	284.96	4.91e+03	85.28
	MICE+PMM	305.89	297.23	1.34e+04	160.75	299.95	284.18	1.49e+04	101.72	268.19	259.00	7.23e+03	129.25
	MICE+RF	368.56	371.29	9.38e+03	116.00	416.71	425.95	1.61e+04	165.32	363.41	344.84	1.38e+04	182.15
	rSVD01	91.21	91.44	1.05e+03	43.57	131.11	130.66	2.23e+03	63.11	165.39	160.47	3.78e+03	77.05
	rSVD84	62.70	57.83	871.33	23.25	93.48	87.01	1.11e+03	46.34	119.01	112.73	2.68e+03	51.20
	rSVD84-01	75.05	70.71	1.14e+03	36.78	105.10	102.71	1.55e+03	41.61	144.96	131.68	3.20e+03	67.84
Calinski et al. (2009a) Pea 18×9	ABB	1.84e+03	1.66e+03	1.05e+06	1.34e+03	2.08e+03	1.91e+03	8.09e+05	1.51e+03	1.64e+03	1.54e+03	5.18e+05	1.04e+03
	MICE+PMM	1.91e+03	1.56e+03	2.23e+06	1.83e+03	1.96e+03	1.75e+03	1.12e+06	1.54e+03	1.50e+03	1.35e+03	7.12e+05	906.77
	MICE+RF	6.48e+03	6.40e+03	3.81e+06	2.29e+03	2.73e+03	2.73e+03	7.10e+05	1.11e+03	1.90e+03	1.87e+03	3.14e+05	695.72
	rSVD01	210.12	150.11	3.81e+04	264.40	552.95	423.39	1.85e+05	501.98	635.60	506.42	2.82e+05	534.83
	rSVD84	142.47	69.42	2.37e+04	132.85	322.58	225.66	9.95e+04	291.15	376.03	297.41	1.01e+05	355.48
	rSVD84-01	176.43	100.41	3.35e+04	217.02	387.64	308.25	1.60e+05	357.45	503.41	360.59	1.76e+05	492.80
Calinski et al. (2009b) Winter rye 18×15	ABB	1.03e+04	9.31e+03	2.66e+07	6.32e+03	8.61e+03	8.36e+03	9.14e+06	3.75e+03	1.00e+04	9.62e+03	1.25e+07	4.06e+03
	MICE+PMM	9.97e+03	9.20e+03	3.17e+07	7.47e+03	8.63e+03	7.98e+03	1.31e+07	4.54e+03	1.08e+04	1.10e+04	1.42e+07	3.88e+03
	MICE+RF	1.99e+04	2.01e+04	2.83e+07	7.28e+03	6.82e+03	7.08e+03	3.33e+06	2.67e+03	1.19e+04	1.16e+04	1.25e+07	4.56e+03
	rSVD01	61.30	50.17	5.39e+03	5.29	157.76	55.24	7.98e+04	7.86	418.71	63.03	4.41e+05	474.77
	rSVD84	71.97	50.97	1.61e+04	5.65	125.12	51.36	7.73e+04	6.81	260.29	63.47	1.97e+05	245.85
	rSVD84-01	95.88	49.14	5.80e+04	5.53	142.50	53.42	5.87e+04	7.47	247.06	62.19	1.80e+05	259.22

Table 57. V by percentage and method — Mechanism **MNAR** (REP 6–10)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Corsten and Denis (1990) Corn 20×7	ABB	5.70e+03	5.40e+03	5.86e+06	3.31e+03	4.35e+03	4.04e+03	4.85e+06	2.78e+03	4.50e+03	4.28e+03	4.02e+06	3.27e+03
	MICE+PMM	5.82e+03	4.62e+03	1.97e+07	5.95e+03	3.78e+03	3.38e+03	5.28e+06	3.60e+03	3.94e+03	3.14e+03	6.78e+06	3.46e+03
	MICE+RF	1.41e+04	1.42e+04	1.49e+07	5.58e+03	1.28e+03	1.31e+03	1.33e+05	427.15	2.67e+03	2.69e+03	6.24e+05	1.10e+03
	rSVD01	5.00e+03	4.64e+03	1.41e+07	5.22e+03	1.66e+03	1.24e+03	1.82e+06	1.58e+03	3.48e+03	2.70e+03	5.96e+06	3.09e+03
	rSVD84	2.50e+03	1.81e+03	4.78e+06	3.11e+03	2.21e+03	1.36e+03	4.97e+06	2.03e+03	2.42e+03	2.14e+03	2.95e+06	2.06e+03
	rSVD84-01	3.25e+03	2.74e+03	6.06e+06	3.21e+03	1.55e+03	1.18e+03	1.57e+06	1.54e+03	2.42e+03	2.09e+03	2.65e+06	1.83e+03
Yan et al. (2007) Wheat 18×9	ABB	52.94	50.37	354.45	26.06	62.96	62.30	371.28	23.77	48.44	46.19	272.26	23.30
	MICE+PMM	57.05	54.22	920.36	38.02	61.73	62.45	610.06	35.01	41.24	40.38	359.27	26.08
	MICE+RF	19.72	19.22	38.83	8.46	61.16	60.66	390.49	25.58	67.85	67.41	460.23	30.44
	rSVD01	35.70	29.91	981.68	33.44	9.68	6.55	118.02	13.80	8.81	5.32	84.08	9.82
	rSVD84	15.66	12.76	282.82	11.62	33.91	28.04	823.47	23.72	22.73	17.11	403.61	22.37
	rSVD84-01	29.43	23.41	459.97	29.58	13.70	10.44	140.94	11.05	12.15	8.34	125.41	10.81
Lavoranti (2003) Eucalyptus grandis 20×7	ABB	297.42	301.71	3.91e+03	88.74	289.23	292.07	2.48e+03	68.86	287.43	284.00	2.37e+03	61.79
	MICE+PMM	244.20	222.86	1.14e+04	100.64	256.79	246.43	1.02e+04	134.35	231.85	224.31	7.08e+03	131.64
	MICE+RF	350.05	334.08	1.34e+04	173.36	359.95	361.06	1.01e+04	116.79	164.40	172.87	3.02e+03	68.73
	rSVD01	138.26	132.27	4.58e+03	90.82	142.32	125.83	3.32e+03	76.87	97.04	86.56	1.95e+03	46.54
	rSVD84	130.48	130.87	3.55e+03	68.24	164.82	146.76	4.02e+03	76.67	186.34	175.03	4.07e+03	80.38
	rSVD84-01	115.27	118.05	2.41e+03	54.54	134.30	126.66	2.08e+03	61.00	112.49	101.68	1.93e+03	51.97
Calinski et al. (2009a) Pea 18×9	ABB	1.90e+03	1.90e+03	3.40e+05	827.06	2.25e+03	2.26e+03	4.62e+05	898.43	1.63e+03	1.61e+03	3.31e+05	811.36
	MICE+PMM	1.94e+03	1.78e+03	1.02e+06	1.51e+03	2.10e+03	2.16e+03	7.38e+05	1.25e+03	1.40e+03	1.34e+03	4.52e+05	1.04e+03
	MICE+RF	1.62e+03	1.63e+03	2.24e+05	740.53	1.03e+03	1.02e+03	9.53e+04	315.05	1.71e+03	1.69e+03	3.04e+05	793.44
	rSVD01	343.62	292.10	8.10e+04	383.93	435.97	382.37	1.14e+05	391.33	629.07	553.68	1.57e+05	516.88
	rSVD84	179.58	116.30	4.98e+04	171.66	773.17	730.61	2.86e+05	749.23	542.74	508.82	1.04e+05	418.59
	rSVD84-01	225.44	195.53	4.09e+04	242.36	453.89	389.95	7.94e+04	363.95	431.23	414.02	5.63e+04	269.91
Calinski et al. (2009b) Winter rye 18×15	ABB	9.47e+03	9.07e+03	6.82e+06	3.40e+03	8.36e+03	7.98e+03	4.46e+06	2.51e+03	9.79e+03	9.00e+03	7.37e+06	2.66e+03
	MICE+PMM	9.64e+03	8.89e+03	1.96e+07	5.37e+03	1.11e+04	1.06e+04	1.30e+07	4.83e+03	1.12e+04	1.07e+04	1.48e+07	3.74e+03
	MICE+RF	1.86e+04	1.80e+04	3.32e+07	7.91e+03	9.99e+03	1.00e+04	7.13e+06	3.86e+03	1.08e+04	1.09e+04	9.20e+06	4.40e+03
	rSVD01	279.74	50.58	6.45e+05	6.21	1.55e+03	1.16e+03	2.69e+06	2.18e+03	8.27e+03	6.62e+03	4.60e+07	7.73e+03
	rSVD84	76.07	48.11	1.28e+04	2.39	874.58	64.05	2.99e+06	690.27	6.45e+03	4.59e+03	3.60e+07	5.39e+03
	rSVD84-01	130.59	49.12	8.02e+04	6.25	1.17e+03	761.45	1.93e+06	1.59e+03	6.10e+03	4.33e+03	2.83e+07	5.04e+03

Table 58. V by percentage and method — Mechanism **MAR** (REP 11–15)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	1.67e+07	1.63e+07	3.91e+13	8.09e+06	1.54e+07	1.48e+07	4.31e+13	9.02e+06	1.79e+07	1.77e+07	4.96e+13	1.07e+07
	MICE+PMM	2.12e+07	2.11e+07	4.96e+13	9.11e+06	1.77e+07	1.61e+07	6.31e+13	9.54e+06	1.97e+07	1.87e+07	5.51e+13	1.22e+07
	MICE+RF	2.40e+07	2.35e+07	4.74e+13	1.07e+07	7.62e+06	7.27e+06	5.18e+12	3.03e+06	1.60e+07	1.54e+07	2.81e+13	6.17e+06
	rSVD01	4.86e+05	4.86e+05	5.19e+08	3.26e+04	5.47e+05	5.47e+05	5.33e+08	2.94e+04	6.81e+05	6.51e+05	2.17e+10	4.27e+04
	rSVD84	4.56e+05	4.53e+05	1.98e+08	1.94e+04	4.91e+05	4.87e+05	3.17e+08	2.20e+04	6.39e+05	6.28e+05	5.41e+09	2.97e+04
	rSVD84-01	4.57e+05	4.56e+05	4.42e+08	2.89e+04	5.07e+05	5.07e+05	8.08e+08	3.41e+04	6.34e+05	6.10e+05	1.59e+10	3.96e+04
Filho et al. (2008) Cotton 17×23	ABB	57.57	51.11	1.32e+03	38.98	56.85	46.64	1.28e+03	51.20	45.01	30.95	1.44e+03	33.95
	MICE+PMM	63.66	48.06	2.64e+03	52.87	57.27	37.51	1.97e+03	64.45	46.84	34.22	1.79e+03	38.90
	MICE+RF	86.73	83.67	677.43	29.97	31.24	30.92	84.02	9.59	61.88	59.70	389.38	26.65
	rSVD01	0.46	0.46	0.00	0.04	0.66	0.54	1.28	0.04	0.58	0.55	0.04	0.03
	rSVD84	0.45	0.45	0.00	0.03	0.47	0.47	0.00	0.03	1.16	0.54	8.67	0.03
	rSVD84-01	0.44	0.44	0.00	0.04	0.52	0.49	0.10	0.03	0.57	0.54	0.05	0.04
Flores et al. (1998) Faba beans 15×12	ABB	3.44e+07	2.85e+07	4.64e+14	2.57e+07	4.45e+07	3.99e+07	5.13e+14	3.14e+07	2.96e+07	2.37e+07	2.71e+14	2.51e+07
	MICE+PMM	3.74e+07	3.27e+07	4.57e+14	2.32e+07	4.33e+07	3.59e+07	6.56e+14	2.85e+07	2.43e+07	2.32e+07	1.40e+14	1.63e+07
	MICE+RF	3.45e+07	3.45e+07	1.28e+14	1.52e+07	2.98e+07	2.94e+07	8.63e+13	1.11e+07	3.60e+07	3.49e+07	1.13e+14	1.65e+07
	rSVD01	4.35e+06	3.66e+05	7.71e+13	2.24e+06	6.96e+06	5.54e+06	4.06e+13	9.82e+06	1.47e+07	1.07e+07	3.04e+14	1.20e+07
	rSVD84	3.75e+06	3.59e+05	5.10e+13	2.51e+06	4.47e+06	2.20e+06	3.93e+13	5.52e+06	7.51e+06	4.87e+06	7.62e+13	5.18e+06
	rSVD84-01	2.96e+06	3.20e+05	3.35e+13	1.99e+06	5.35e+06	3.57e+06	4.07e+13	7.18e+06	9.88e+06	6.12e+06	1.22e+14	9.51e+06
Mattos et al. (2013) Sugarcane 22×5	ABB	395.77	293.77	9.68e+04	403.40	485.21	300.78	1.80e+05	683.06	476.81	322.93	1.90e+05	610.94
	MICE+PMM	479.92	310.12	3.34e+05	451.77	619.72	383.37	5.58e+05	754.56	529.15	270.84	3.57e+05	518.33
	MICE+RF	175.15	171.19	2.45e+03	61.54	1.27e+03	1.25e+03	1.22e+05	428.97	2.10	2.10	0.46	0.98
	rSVD01	718.77	512.67	4.22e+05	600.23	726.50	541.23	2.50e+05	518.25	905.22	616.15	4.29e+05	818.41
	rSVD84	353.84	294.35	7.63e+04	362.02	491.32	428.95	1.48e+05	343.97	504.92	452.27	9.05e+04	430.07
	rSVD84-01	504.81	397.43	1.45e+05	420.29	550.18	454.24	1.43e+05	370.67	612.47	471.98	1.58e+05	602.30
Rad et al. (2013) Wheat 36×6	ABB	732.87	631.62	1.20e+05	506.48	648.83	603.27	9.03e+04	347.01	702.67	696.79	1.10e+05	420.81
	MICE+PMM	761.52	565.84	3.68e+05	776.06	887.27	529.86	6.60e+05	949.16	708.01	582.09	2.75e+05	569.31
	MICE+RF	501.63	500.21	2.17e+04	219.30	269.32	254.52	8.96e+03	116.64	303.72	307.12	7.93e+03	115.75
	rSVD01	319.12	284.95	3.66e+04	224.78	155.91	141.02	7.14e+03	119.86	249.54	244.43	1.27e+04	157.31
	rSVD84	417.95	387.40	6.65e+04	343.56	168.20	152.97	1.13e+04	152.33	181.62	161.10	1.26e+04	107.56
	rSVD84-01	275.29	246.36	1.92e+04	181.14	122.07	107.65	6.46e+03	77.39	196.54	175.20	1.09e+04	138.09

Table 59. V by percentage and method — Mechanism **MCAR** (REP 11–15)

		P10				P20				P30			
Dataset	Method	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	1.55e+07	1.44e+07	5.33e+13	9.18e+06	1.45e+07	1.34e+07	3.77e+13	7.66e+06	1.73e+07	1.68e+07	4.98e+13	8.56e+06
	MICE+PMM	1.81e+07	1.73e+07	7.44e+13	1.05e+07	1.69e+07	1.50e+07	6.98e+13	7.51e+06	1.77e+07	1.82e+07	3.49e+13	8.88e+06
	MICE+RF	3.03e+07	2.99e+07	7.25e+13	1.04e+07	4.29e+06	4.23e+06	2.19e+12	1.89e+06	1.48e+07	1.45e+07	1.66e+13	5.15e+06
	rSVD01	4.67e+05	4.65e+05	8.70e+08	3.88e+04	5.21e+05	5.21e+05	1.38e+09	6.01e+04	6.25e+05	6.23e+05	2.12e+09	5.62e+04
	rSVD84	4.48e+05	4.46e+05	5.75e+08	3.41e+04	5.04e+05	5.00e+05	1.11e+09	4.49e+04	5.79e+05	5.74e+05	2.17e+09	4.75e+04
	rSVD84-01	4.38e+05	4.39e+05	7.88e+08	3.95e+04	4.95e+05	4.95e+05	1.35e+09	5.85e+04	5.75e+05	5.67e+05	3.18e+09	6.15e+04
Filho et al. (2008) Cotton 17×23	ABB	55.39	43.88	1.50e+03	52.22	49.07	40.76	1.04e+03	31.65	41.97	36.02	529.99	28.04
	MICE+PMM	67.78	51.38	2.58e+03	50.27	62.04	50.97	1.73e+03	47.51	43.72	37.18	712.41	27.31
	MICE+RF	38.18	37.85	112.42	15.16	118.41	117.14	1.29e+03	42.45	79.25	80.31	608.28	30.58
	rSVD01	0.49	0.49	0.00	0.05	0.55	0.55	0.00	0.05	0.83	0.62	3.86	0.08
	rSVD84	0.47	0.48	0.00	0.05	0.53	0.53	0.00	0.04	0.59	0.60	0.00	0.07
	rSVD84-01	0.46	0.46	0.00	0.05	0.52	0.51	0.00	0.05	0.59	0.58	0.01	0.07
Flores et al. (1998) Faba beans 15×12	ABB	3.37e+07	2.99e+07	4.63e+14	2.56e+07	3.77e+07	3.65e+07	3.54e+14	2.59e+07	4.66e+07	4.41e+07	4.78e+14	2.70e+07
	MICE+PMM	3.24e+07	2.40e+07	7.80e+14	3.81e+07	3.71e+07	3.42e+07	5.47e+14	2.75e+07	4.38e+07	4.37e+07	4.03e+14	2.98e+07
	MICE+RF	4.11e+07	4.18e+07	1.17e+14	1.81e+07	1.05e+07	1.04e+07	7.93e+12	4.17e+06	3.22e+07	3.09e+07	9.01e+13	1.21e+07
	rSVD01	5.71e+05	2.83e+05	9.24e+11	5.59e+04	2.21e+06	3.44e+05	1.05e+13	2.79e+06	5.71e+06	4.01e+06	4.36e+13	6.76e+06
	rSVD84	6.10e+05	2.88e+05	2.15e+12	5.66e+04	1.79e+06	3.99e+05	7.17e+12	1.58e+06	3.47e+06	1.82e+06	1.64e+13	4.25e+06
	rSVD84-01	7.54e+05	2.87e+05	1.95e+12	5.12e+04	1.89e+06	4.03e+05	8.31e+12	2.16e+06	3.97e+06	1.79e+06	2.55e+13	4.76e+06
Mattos et al. (2013) Sugarcane 22×5	ABB	532.83	458.87	1.38e+05	436.21	566.93	473.40	1.02e+05	473.01	797.67	704.93	1.82e+05	584.10
	MICE+PMM	438.72	302.81	2.03e+05	411.16	477.49	400.95	1.09e+05	360.03	712.20	569.02	2.75e+05	509.36
	MICE+RF	791.90	792.68	6.67e+04	357.34	539.71	554.74	1.60e+04	173.20	151.89	151.83	1.81e+03	50.73
	rSVD01	483.02	387.42	1.12e+05	306.63	511.45	431.98	1.13e+05	406.96	798.78	675.80	2.56e+05	599.91
	rSVD84	249.50	197.84	3.15e+04	189.75	303.29	249.97	4.28e+04	213.46	514.12	429.02	9.09e+04	410.03
	rSVD84-01	348.18	292.67	7.57e+04	233.83	379.35	333.52	5.31e+04	282.12	592.77	520.20	1.33e+05	373.30
Rad et al. (2013) Wheat 36×6	ABB	710.86	641.91	7.60e+04	324.10	575.93	572.38	4.18e+04	261.97	628.85	622.67	3.70e+04	293.19
	MICE+PMM	604.76	549.50	1.36e+05	405.01	531.16	483.54	7.59e+04	369.87	566.59	530.22	6.66e+04	320.95
	MICE+RF	415.16	402.68	1.80e+04	161.65	767.06	789.35	4.55e+04	281.33	330.69	337.81	1.01e+04	119.75
	rSVD01	333.32	293.87	2.66e+04	231.37	387.45	381.41	2.51e+04	216.46	545.85	496.92	7.20e+04	332.21
	rSVD84	217.99	203.29	1.49e+04	160.08	243.34	219.36	1.78e+04	150.20	351.19	311.58	2.49e+04	222.71
	rSVD84-01	278.25	259.52	1.99e+04	164.23	283.01	270.17	1.90e+04	198.27	397.01	373.81	3.17e+04	226.55

Table 60. V by percentage and method — Mechanism **MNAR** (REP 11–15)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Farias (2005) Cotton 15×27	ABB	1.69e+07	1.64e+07	1.76e+13	5.81e+06	1.51e+07	1.52e+07	1.80e+13	6.17e+06	1.55e+07	1.51e+07	2.49e+13	6.11e+06
	MICE+PMM	2.23e+07	2.26e+07	4.12e+13	9.54e+06	1.80e+07	1.85e+07	2.47e+13	5.91e+06	1.74e+07	1.65e+07	3.14e+13	6.22e+06
	MICE+RF	2.07e+07	2.15e+07	3.56e+13	8.80e+06	2.03e+07	2.02e+07	4.23e+13	7.53e+06	2.33e+07	2.25e+07	4.85e+13	9.17e+06
	rSVD01	4.74e+05	4.74e+05	3.11e+08	2.55e+04	5.45e+05	5.45e+05	5.43e+08	3.00e+04	6.97e+05	6.72e+05	3.83e+10	3.67e+04
	rSVD84	4.33e+05	4.32e+05	1.52e+08	1.63e+04	4.99e+05	4.98e+05	2.25e+08	1.90e+04	5.95e+05	5.94e+05	3.38e+08	2.64e+04
	rSVD84-01	4.50e+05	4.52e+05	5.28e+08	3.05e+04	4.87e+05	4.84e+05	4.04e+08	2.58e+04	6.06e+05	6.04e+05	1.62e+09	3.34e+04
Filho et al. (2008) Cotton 17×23	ABB	49.05	41.27	657.03	31.73	50.87	46.93	581.50	34.27	54.08	50.53	563.17	27.13
	MICE+PMM	59.35	52.45	924.15	41.12	62.76	56.85	1.13e+03	45.01	53.38	46.91	550.82	26.82
	MICE+RF	122.91	122.14	1.44e+03	46.67	38.45	36.04	144.17	18.03	36.76	36.31	88.23	13.17
	rSVD01	0.50	0.49	0.00	0.03	0.93	0.60	3.66	0.03	1.20	0.69	4.31	0.05
	rSVD84	0.45	0.45	0.00	0.02	0.57	0.54	0.07	0.03	0.85	0.68	0.25	0.04
	rSVD84-01	0.46	0.46	0.00	0.04	0.57	0.55	0.01	0.05	1.09	0.64	2.56	0.06
Flores et al. (1998) Faba beans 15×12	ABB	3.83e+07	3.65e+07	2.22e+14	1.64e+07	4.08e+07	4.01e+07	2.22e+14	2.09e+07	3.30e+07	3.12e+07	2.20e+14	2.25e+07
	MICE+PMM	3.34e+07	2.89e+07	3.51e+14	2.64e+07	3.85e+07	3.51e+07	3.64e+14	2.36e+07	2.99e+07	2.69e+07	2.41e+14	1.86e+07
	MICE+RF	2.33e+07	2.33e+07	3.96e+13	8.22e+06	5.16e+07	5.12e+07	2.18e+14	1.81e+07	3.52e+07	3.44e+07	1.15e+14	1.54e+07
	rSVD01	9.29e+06	4.62e+06	1.40e+14	1.67e+07	7.95e+06	3.53e+05	1.86e+14	1.16e+07	2.37e+07	1.14e+07	4.06e+15	1.88e+07
	rSVD84	4.06e+06	3.30e+05	4.08e+13	5.37e+06	5.11e+06	1.09e+06	1.14e+14	4.32e+06	1.06e+07	6.60e+06	1.11e+14	1.09e+07
	rSVD84-01	5.57e+06	3.77e+05	1.10e+14	6.50e+06	6.91e+06	3.98e+06	6.84e+13	1.02e+07	1.29e+07	8.61e+06	1.58e+14	1.07e+07
Mattos et al. (2013) Sugarcane 22×5	ABB	629.54	585.68	8.88e+04	414.41	680.18	579.60	1.08e+05	462.46	717.78	612.17	1.15e+05	487.19
	MICE+PMM	493.79	346.75	1.81e+05	428.99	564.47	481.63	1.12e+05	394.08	574.47	502.54	1.24e+05	406.96
	MICE+RF	301.45	301.06	7.93e+03	104.90	248.70	242.85	5.34e+03	90.01	266.97	261.21	6.78e+03	96.79
	rSVD01	820.21	517.76	2.70e+06	480.19	553.95	318.04	1.94e+06	333.84	374.35	317.51	6.28e+04	273.93
	rSVD84	339.03	287.77	7.80e+04	319.19	429.04	357.45	9.50e+04	357.70	622.56	525.70	1.68e+05	519.54
	rSVD84-01	460.34	386.33	9.29e+04	409.14	346.40	278.88	5.11e+04	310.17	360.63	299.18	4.78e+04	286.48
Rad et al. (2013) Wheat 36×6	ABB	678.77	679.77	3.84e+04	267.26	587.97	565.10	2.87e+04	209.69	640.64	633.32	2.34e+04	187.52
	MICE+PMM	781.81	711.89	1.52e+05	492.13	573.49	561.88	1.11e+05	476.28	579.53	530.65	8.84e+04	327.37
	MICE+RF	731.01	743.64	4.01e+04	300.65	798.86	807.66	5.31e+04	297.08	409.50	406.82	1.59e+04	127.76
	rSVD01	128.25	106.87	1.30e+04	158.16	198.00	182.73	1.26e+04	158.57	277.71	237.94	2.17e+04	148.06
	rSVD84	430.86	385.30	9.39e+04	364.29	193.59	172.82	1.23e+04	116.77	153.74	124.21	1.09e+04	85.10
	rSVD84-01	187.46	158.05	1.54e+04	155.17	172.32	146.00	1.10e+04	121.44	199.56	182.60	1.02e+04	129.47

Table 61. V by percentage and method — Mechanism **MAR** (REP 16–17)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	804.70	745.61	1.70e+05	580.48	819.10	830.10	2.08e+05	744.44	959.91	975.08	3.52e+05	938.20
	MICE+PMM	904.23	770.03	5.51e+05	1.13e+03	829.72	634.00	4.48e+05	1.08e+03	924.64	634.46	6.24e+05	1.06e+03
	MICE+RF	413.81	415.23	1.37e+04	161.08	518.25	509.04	2.28e+04	188.33	205.22	204.54	5.21e+03	103.54
	rSVD01	1.87e+03	432.00	1.75e+08	478.62	507.39	384.40	2.19e+05	439.89	363.83	316.83	3.67e+04	258.25
	rSVD84	521.27	485.03	9.09e+04	330.13	396.83	357.56	5.81e+04	269.72	200.31	178.82	1.52e+04	166.03
	rSVD84-01	373.73	334.80	6.17e+04	246.24	316.85	281.79	4.23e+04	204.30	259.31	222.58	2.37e+04	207.97
Zobel et al. (1988) Soybean 7×35	ABB	1.05e+08	1.02e+08	2.22e+15	7.35e+07	9.79e+07	9.77e+07	2.28e+15	6.19e+07	1.20e+08	1.17e+08	3.24e+15	7.00e+07
	MICE+PMM	1.22e+08	9.26e+07	1.03e+16	1.20e+08	1.11e+08	8.88e+07	8.03e+15	1.20e+08	1.33e+08	1.23e+08	8.83e+15	1.34e+08
	MICE+RF	1.64e+08	1.64e+08	1.86e+15	5.56e+07	4.49e+07	4.50e+07	1.34e+14	1.71e+07	3.07e+08	3.13e+08	8.36e+15	1.02e+08
	rSVD01	6.36e+07	5.91e+07	1.38e+15	5.11e+07	8.28e+07	8.10e+07	1.16e+15	4.79e+07	1.11e+08	7.15e+07	1.22e+17	4.45e+07
	rSVD84	3.46e+07	3.19e+07	3.95e+14	2.23e+07	5.10e+07	4.27e+07	1.07e+15	2.94e+07	6.32e+07	6.08e+07	6.60e+14	2.90e+07
	rSVD84-01	5.04e+07	4.92e+07	6.15e+14	3.68e+07	6.11e+07	5.75e+07	7.03e+14	3.51e+07	6.35e+07	5.79e+07	3.48e+15	2.93e+07

Table 62. V by percentage and method — Mechanism **MCAR** (REP 16–17)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	637.42	564.33	1.39e+05	556.50	746.50	712.54	1.29e+05	423.40	905.40	858.88	8.02e+04	314.37
	MICE+PMM	525.62	435.03	2.02e+05	507.65	548.07	454.79	1.66e+05	545.54	692.60	650.05	1.32e+05	532.92
	MICE+RF	113.78	113.08	953.31	42.36	720.23	714.18	4.90e+04	309.28	1.12e+03	1.11e+03	1.15e+05	473.47
	rSVD01	404.88	365.96	6.89e+04	238.52	644.23	599.58	1.10e+05	441.56	1.93e+03	747.14	1.29e+08	559.35
	rSVD84	255.81	231.88	2.55e+04	248.38	374.81	296.05	4.85e+04	264.71	469.06	423.77	7.40e+04	284.65
	rSVD84-01	310.66	270.27	3.60e+04	229.10	635.32	444.74	2.50e+06	334.60	565.73	516.44	7.95e+04	277.27
Zobel et al. (1988) Soybean 7×35	ABB	1.03e+08	9.91e+07	1.28e+15	4.60e+07	1.07e+08	1.07e+08	6.76e+14	3.19e+07	9.18e+07	8.83e+07	5.79e+14	3.12e+07
	MICE+PMM	9.05e+07	8.38e+07	2.61e+15	8.03e+07	9.24e+07	8.87e+07	1.62e+15	4.93e+07	7.77e+07	7.67e+07	8.27e+14	3.75e+07
	MICE+RF	9.16e+07	9.07e+07	7.06e+14	3.40e+07	1.84e+08	1.82e+08	2.67e+15	6.23e+07	5.16e+07	5.22e+07	2.51e+14	2.20e+07
	rSVD01	3.29e+07	3.02e+07	2.80e+14	2.40e+07	5.46e+07	4.95e+07	6.89e+14	2.80e+07	9.01e+07	5.80e+07	6.21e+16	3.72e+07
	rSVD84	1.80e+07	1.54e+07	1.43e+14	1.44e+07	3.07e+07	2.79e+07	2.73e+14	1.75e+07	3.58e+07	3.12e+07	3.40e+14	2.02e+07
	rSVD84-01	2.69e+07	2.35e+07	1.98e+14	1.98e+07	3.63e+07	3.41e+07	2.64e+14	2.00e+07	4.64e+07	4.13e+07	4.39e+14	2.54e+07

Table 63. V by percentage and method — Mechanism **MNAR** (REP 16–17)

Dataset	Method	P10				P20				P30			
		Mean	Median	Variance	IQR	Mean	Median	Variance	IQR	Mean	Median	Variance	IQR
Yang (2007) Barley 6×18	ABB	697.56	665.00	4.61e+04	302.08	797.61	792.39	2.64e+04	230.68	913.63	906.30	2.63e+04	221.37
	MICE+PMM	648.37	577.08	1.93e+05	688.82	750.97	750.50	6.06e+04	332.09	895.28	854.71	1.00e+05	440.59
	MICE+RF	488.66	480.21	1.93e+04	161.29	969.32	954.75	8.50e+04	427.52	592.85	593.84	3.08e+04	214.74
	rSVD01	1.55e+04	396.10	2.10e+10	369.32	154.97	99.35	2.83e+04	247.09	306.08	266.37	5.94e+04	265.83
	rSVD84	533.03	509.94	1.01e+05	466.16	99.29	82.58	1.37e+04	121.79	278.30	273.81	2.18e+04	181.91
	rSVD84-01	358.81	326.95	4.90e+04	220.21	190.93	90.46	7.85e+05	147.23	266.12	182.51	1.19e+05	167.16
Zobel et al. (1988) Soybean 7×35	ABB	9.91e+07	9.80e+07	5.06e+14	2.37e+07	1.06e+08	1.06e+08	2.41e+14	1.93e+07	9.25e+07	9.16e+07	1.95e+14	1.99e+07
	MICE+PMM	9.30e+07	8.57e+07	2.16e+15	7.74e+07	9.92e+07	9.85e+07	1.88e+15	6.45e+07	9.39e+07	8.82e+07	1.28e+15	5.00e+07
	MICE+RF	5.17e+07	5.27e+07	2.32e+14	1.69e+07	2.07e+08	2.09e+08	3.44e+15	7.48e+07	8.74e+07	8.86e+07	7.18e+14	3.23e+07
	rSVD01	8.58e+07	7.67e+07	2.46e+15	4.21e+07	1.75e+08	1.14e+08	2.31e+17	7.70e+07	8.90e+07	5.26e+07	3.76e+16	3.96e+07
	rSVD84	3.62e+07	3.07e+07	4.03e+14	2.22e+07	4.83e+07	3.89e+07	8.78e+14	3.50e+07	3.67e+07	3.16e+07	5.69e+14	2.33e+07
	rSVD84-01	5.95e+07	5.39e+07	9.96e+14	4.40e+07	8.02e+07	6.39e+07	9.94e+15	3.60e+07	4.08e+07	3.50e+07	5.13e+14	2.23e+07