International Journal of Mathematics and Statistics
ISSN (Online): 2208-2212
Volume 11 Issue 01 March 2025
DOI: 10.69980/sdqawe10

COMPARATIVE ANALYSIS OF BIAS AND ROOT MEAN SQUARE ERROR IN RESAMPLING AND PARAMETRIC POWER ESTIMATION

Dr. Charlotte Evans[1], Dr. Matthias Keller[2], Dr. Natalie Brooks[3*]

[1]Department of Statistical Computing and Applied Mathematics, Westbridge Institute of Mathematical Sciences, United Kingdom
[2]Department of Probability and Statistical Inference, Rhine Valley Institute of Technology, Germany
[3]Department of Data Science and Quantitative Research, Pacific Research Institute, Australia*

Corresponding Author: Dr. Natalie Brooks

Abstract

Power estimation is very important in proper research planning, especially for small samples and different distributional settings. In this study, the accuracy of the resampling-based and parametric statistical power estimation methods was compared using bias and rmse. Secondary Monte Carlo simulation data for 36 conditions (3 levels of pilot size and 4 levels of effect size with three probability distributions) were subjected to a quantitative comparative analysis. Estimation accuracy was assessed using descriptive statistics and at the scenario level. The overall mean RMSE of the resampling-based estimation method (0.2796) was lower than that of the parametric estimation method (0.2991) by approximately 6.53%. The mean absolute bias for the resampling was also lower (0.1094) than that of the parametric estimation (0.1269). However, Parametric estimation had lower RMSE in 19 of 36 scenarios. As the size of the pilot sample increased, both methods were found to have decreasing estimation error. Resampling worked well under the skewed distributions and in moderate to large effect sizes while parametric estimation had marginal benefits under normal or bounded distributions. This suggests that the accuracy of estimation is a function of sample size, size of the effect and distributional characteristics. The study highlights that when choosing statistical power estimation methods, it is important to consider both the bias and RMSE, and that methodological decisions can be condition specific when planning small sample studies.

Keyword: - *Bootstrap Resampling; Parametric Power Estimation; Statistical Bias; Root Mean Square Error; Monte Carlo Simulation.*

INTRODUCTION

In quantitative research, statistical power estimation plays a crucial role in assessing the likelihood of discovering a true effect given certain experimental condition. Successful power estimation also helps determine the correct sample size, enhances statistical validity, and minimizes the possibility of reaching false conclusions. But, the size of the sample, the type of probability distribution, the size of the effect and the statistical assumptions used in the estimation procedure can influence the accuracy of the estimation. The methodological performance assessment under varying conditions of analysis has been shown to be crucial in comparative studies performed in scientific applications (Acici, 2025). The importance of correct numerical estimation and systematic comparisons of alternative computational approaches is likewise stressed in predictive modelling studies (Allal et al., 2025). Studies on forecasting performance have also emphasized the need for choosing suitable estimation techniques to reduce forecasting error and enhance the reliability of the analysis (Carrera & Kim, 2024).

Two important methods of statistical estimation are resampling and parametric. Resampling procedures quantify a variability and uncertainty associated with the estimation for repeated samples, while parametric approaches rely on assumptions about probability distributions and population characteristics. Alrumaidhi et al. (2023) have shown the methodological relevance of resampling strategies to assess the predictive performance of parametric and nonparametric models. Other sampling approaches have been introduced in computational estimation methods to compensate for measurement uncertainty and numerical stability issues (Cui et al., 2023). The usefulness of simulation-based investigations in linear regression to assess statistical estimation procedures has been demonstrated (Fitrianto & Linganathan, 2022). Additionally, resampling techniques offer a structured approach to quantifying uncertainty, assessing model stability, and evaluating statistical performance, especially in the absence of analytical solutions (James et al., 2023).

For the evaluation of accuracy of statistical estimators, bias and root mean square error (RMSE) are commonly used. Bias is a systematic difference between the estimated and reference values; RMSE is a combining of systematic error and estimation variability and is a measure of overall accuracy. The smaller the absolute bias and RMSE, the better the estimation performance. The validation of the models has shown that comparing numerical results with the reference measurements is crucial to evaluate the numerical reliability of the model (Deville et al., 2024). Likewise, bias correction studies have revealed the significance of systematic estimation errors and enhancing the precision of measurements (Elsebaie et al., 2025). Conducting research using calibration has also highlighted the importance of assessing estimation inaccuracies for the creation of reliable quantitative methods (Han et al., 2025). However, the link between bias and RMSE must be treated with caution as an estimator with low bias can have high variability.

Under small-sample conditions, where the number of observations is small, the reliability of statistical estimation becomes increasingly important because the uncertainty of the estimates and sensitivity to assumptions regarding the distribution of the data may be magnified. Comparative modeling studies have demonstrated the need to assess the estimation performance with respect to different computational procedures (Chaaban & Alfadl, 2024). Uncertainty-aware methods have also been shown to be relevant in quantitative inference in Bayesian estimation studies (Dong et al., 2024). Studies with computational forecasting have focussed on methodological improvement and systematic evaluation of the forecast accuracy (Ferkous et al., 2024). Likewise, comparative analysis of advanced computational models has pointed out the methodological aspect of how choices can affect the performance of estimation (Khan et al., 2024). While these studies provide an overall indication of the importance of comparative accuracy assessment, they do not directly address the relative degree of bias and RMSE of resampling-based and parametric statistical power estimators under controlled small-sample simulation conditions.

Thus, the current study performs a comparative secondary analysis of the Monte Carlo simulation results to assess bias and RMSE for resampling-based and parametric power estimation. Thirty-six simulation scenarios are analysed, which vary with the pilot sample size, the magnitude of effect, and the probability distribution. The study will investigate systematic estimation errors and the accuracy of the estimation to determine conditions for more reliable statistical power estimation and methodological guidance in planning quantitative research.

Research Objectives

- To compare the bias and root mean square error of resampling-based and parametric statistical power estimation methods across Monte Carlo simulation scenarios.
- To evaluate the influence of pilot sample size, effect magnitude, and probability distribution on the accuracy and variability of statistical power estimates.
- To identify the simulation conditions under which resampling-based or parametric approaches demonstrate lower absolute bias and RMSE, thereby assessing their relative suitability for small-sample statistical power estimation.

2. Materials and Methods

2.1. Study Design

A quantitative, comparative research design was used to compare the accuracy of the resampling-based and parametric statistical power estimation approaches in the present study, and secondary Monte Carlo simulation results were used to evaluate the accuracy. Systematic error (bias) and overall estimation error (RMSE) were the two parameters to which the analysis was done. The two methods were compared using the same simulation conditions with varying size of the pilot samples, effect sizes and probability distributions. The study was designed to determine the conditions under which one or the other method had a smaller error of estimation and was more reliable.

2.2. Data Source and Simulation Characteristics

This study used publicly available secondary-simulated data reported by da Silva et al. (2026). Thirty-six simulation scenarios by 16 variables were included in the analytical file: reference statistical power, bootstrap estimates, parametric estimates, bias, RMSE, estimation variability. A normal, a skewed log normal, and a bounded discrete distribution (with ceiling effects) were all considered. The effect sizes considered were 0.0, 0.3, 0.6, and 0.9 for three different pilot sample sizes: 10, 15, and 20 observations per group. A target number of observations per group of 20 were used for the main analysis. 400 independent pilot samples and 400 bootstrap repetitions per pilot were used in each scenario to systematically investigate the estimation accuracy in different statistical scenarios.

2.3. Study Variables and Outcome Measures

This study used bias and RMSE for resampling-based and parametric power estimation as primary outcome measures. Bias was used to assess the systematic difference between the estimated and reference statistical power and RMSE was used to assess overall estimation error. Estimation method, pilot sample size, effect size and probability distribution were the explanatory variables. Other measures were the reference power, the mean estimated power and variability of estimation. The statistical power was determined using Monte Carlo simulation of a two-sided Mann–Whitney U test for a 0.05 level of significance. The lower value of absolute bias and RMSE showed that the estimation was more accurate.

2.4. Resampling-Based and Parametric Power Estimation

The resampling method estimated statistical power by resampling the pilot samples, which were independently generated, with replacement. A two-sided Mann–Whitney U test was carried out for each repetition, and the statistical power was calculated as the percentage of significant results at the 0.05 level. The parametric approach was based on the standardized mean differences obtained from pilot observations and two-sided approximation of independent-samples t-test to estimate the power. Both methods were compared as to their reference statistical power. When interpreting estimation performance, the differences in their underlying statistical assumptions were taken into account.

2.5. Assessment of Bias and Root Mean Square Error

Bias and RMSE were calculated to quantify estimation accuracy. Bias represented the average difference between estimated and reference power, while RMSE measured the square root of the average squared estimation error. The mathematical expressions were:

Bias:

$$\mathrm{Bias}(\hat{P}) = \frac{1}{R}\sum_{r=1}^{R} (\hat{P}_r - P_{\mathrm{ref}})$$

Root Mean Square Error:

$$\mathrm{RMSE}(\hat{P}) = \sqrt{\frac{1}{R}\sum_{r=1}^{R} (\hat{P}_r - P_{\mathrm{ref}})^2}$$

Here, \(\hat{P}_r\) represents estimated statistical power in repetition \(r\), \(P_{\mathrm{ref}}\) denotes reference statistical power, and \(R\) represents the number of independent pilot repetitions. Positive bias indicated overestimation, whereas negative bias indicated underestimation. Lower absolute bias and RMSE reflected greater estimation accuracy. The analysis used the previously calculated values from the simulation results.

2.6. Statistical Analysis

The bias and RMSE for the 36 simulation scenarios were summarized with descriptive statistics such as mean, standard deviation, and percentage differences. Comparisons of results obtained by the two sampling methods, resampling and parametric estimation, were made at the scenario level. Stratified comparisons were made for the effect of pilot sample size, effect magnitude and probability distribution. Pearson or Spearman correlation analysis combined with simple linear regression, were, where appropriate, used to study relations between aspects of the simulation and the estimation error. Comparisons were mainly descriptive findings, when the predefined simulation conditions were given. Data were presented in four tables and four figures to show the differences in the accuracy of estimation.

2.7. Ethical Considerations

The study did not involve any human participants, personal information, or experimental interventions, but rather relied on publicly available secondary simulation results. This meant that no further ethical approval and informed consent was needed. It recognized the original research source appropriately and interpreted all the findings within the scope of the research results available in the simulation.

3. Results

A total of 36 Monte Carlo scenarios were analysed, which involved a comparison of resampling-based and parametric statistical power estimation. The results were assessed for three pilot sample sizes, four effect sizes, and three

probability distributions. Differences in the accuracy of the estimation were investigated using bias, absolute bias, root mean square error (RMSE), and confidence interval coverage. The findings are presented in five tables and five figures, each of which is referenced in the text.

3.1. Characteristics of Monte Carlo Simulation Scenarios

A simulation framework was built based upon 36 scenarios equally distributed over three probability distributions and three pilot sample sizes. Four levels of each distribution and sample-size combination were assessed. The same simulation conditions and reference statistical power values were used for both times of estimation. As presented in Table 1, the design made balanced comparisons of the pre-defined experimental conditions.

Table 1. Characteristics of the Monte Carlo Simulation Scenarios

Simulation characteristic	Category	Number of scenarios	Percentage (%)
Probability distribution	Normal	12	33.33
	Skewed log-normal	12	33.33
	Bounded/discrete ceiling	12	33.33
Pilot sample size (per group)	10	12	33.33
	15	12	33.33
	20	12	33.33
Effect magnitude	0.0	9	25.00
	0.3	9	25.00
	0.6	9	25.00
	0.9	9	25.00

3.2. Comparative Analysis of Bias between Resampling and Parametric Methods

The comparison of the bias in estimation showed differences between resampling based and parametric methods for pilot sample sizes. The average bias was positive for both approaches, suggesting that on average, they tend to overestimate the reference statistical power. Table 2 shows that the mean resampling bias was reduced from 0.0893 to 0.0394 with increase in the pilot size from 10 to 20, and parametric bias was reduced from 0.0575 to 0.0039. Resampling showed lower mean absolute bias at all the pilot sample sizes although parametric estimation had lower signed mean bias. These differences can be seen in Figure 1.

Table 2. Comparison of Bias between Resampling and Parametric Power Estimation

Pilot sample size	Resampling bias	Parametric bias	Resampling absolute bias	Parametric absolute bias
10	0.0893	0.0575	0.1438	0.1500
15	0.0518	0.0174	0.1037	0.1262
20	0.0394	0.0039	0.0806	0.1045
Overall	0.0601	0.0263	0.1094	0.1269

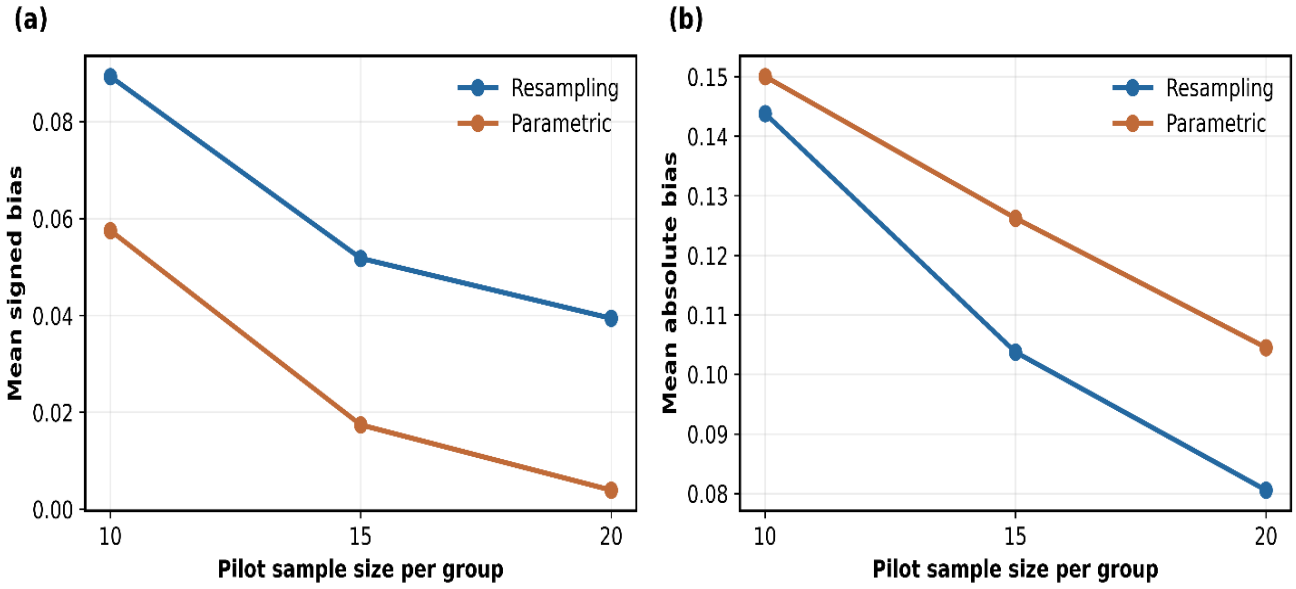


Figure 1. Comparison of signed and absolute bias across pilot sample sizes

Note. Panel (a) illustrates mean signed bias, and panel (b) presents mean absolute bias for resampling and parametric approaches across the three pilot sample sizes.

3.3. Comparative Analysis of Root Mean Square Error

A consistent reduction in average estimation error was shown as the sample size of the pilots increased in the RMSE comparison. The mean RMSE of the resampling method was found to be 0.3281 at a pilot size of 10 and 0.2363 at a pilot size of 20 as presented in Table 3. The parametric RMSE also saw a fall from 0.3414 to 0.2593. The mean RMSE for all 36 scenarios based on resampling is 0.2796, while that for parametric estimation is 0.2991. At the pilot sizes of 15 and 20 the difference in average RMSE was the largest. These trends are shown in Figure 2.

Table 3. Comparison of Root Mean Square Error between Resampling and Parametric Power Estimation

Pilot sample size	Resampling RMSE	Parametric RMSE	Difference (Resampling − Parametric)	Relative difference (%)
10	0.3281	0.3414	−0.0133	−3.90
15	0.2744	0.2968	−0.0224	−7.55
20	0.2363	0.2593	−0.0230	−8.87
Overall	0.2796	0.2991	−0.0196	−6.54

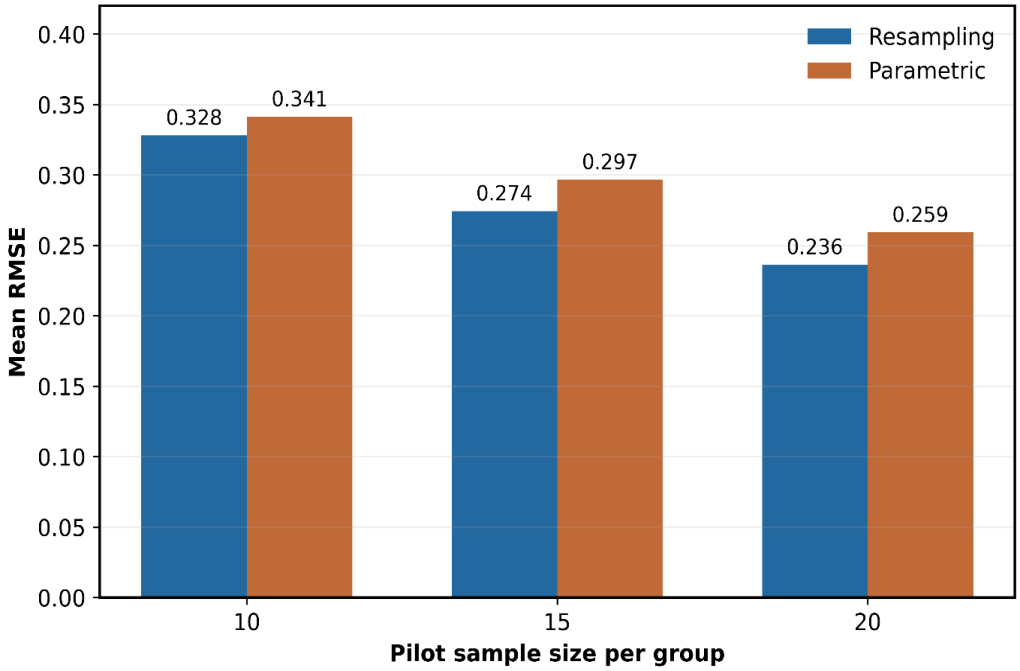


Figure 2. Root mean square error across pilot sample sizes

3.4. Influence of Effect Size on Bias and RMSE

Estimation errors showed direction and magnitude effects of the effect size. Table 4 shows both methods to be positively biased at effect size 0.0 and 0.3, which means that both methods overestimate the power of the reference. Average bias of both methods shifted to negative for an effect size of 0.9, showing underestimation. The mean RMSE were lower for Resampling than for Parametric estimation at effect sizes of 0.6 and 0.9, but the mean RMSE were lower for Parametric estimation at effect sizes of 0.0 and 0.3. The changes in bias and RMSE for each of the four effect sizes are shown in Figure 3.

Table 4. Effect Size-Specific Comparison of Bias and RMSE

Effect size	Resampling bias	Parametric bias	Resampling RMSE	Parametric RMSE
0.0	0.1713	0.1636	0.2608	0.2533
0.3	0.1326	0.0986	0.3011	0.2891
0.6	0.0074	−0.0472	0.3038	0.3479
0.9	−0.0708	−0.1098	0.2527	0.3063

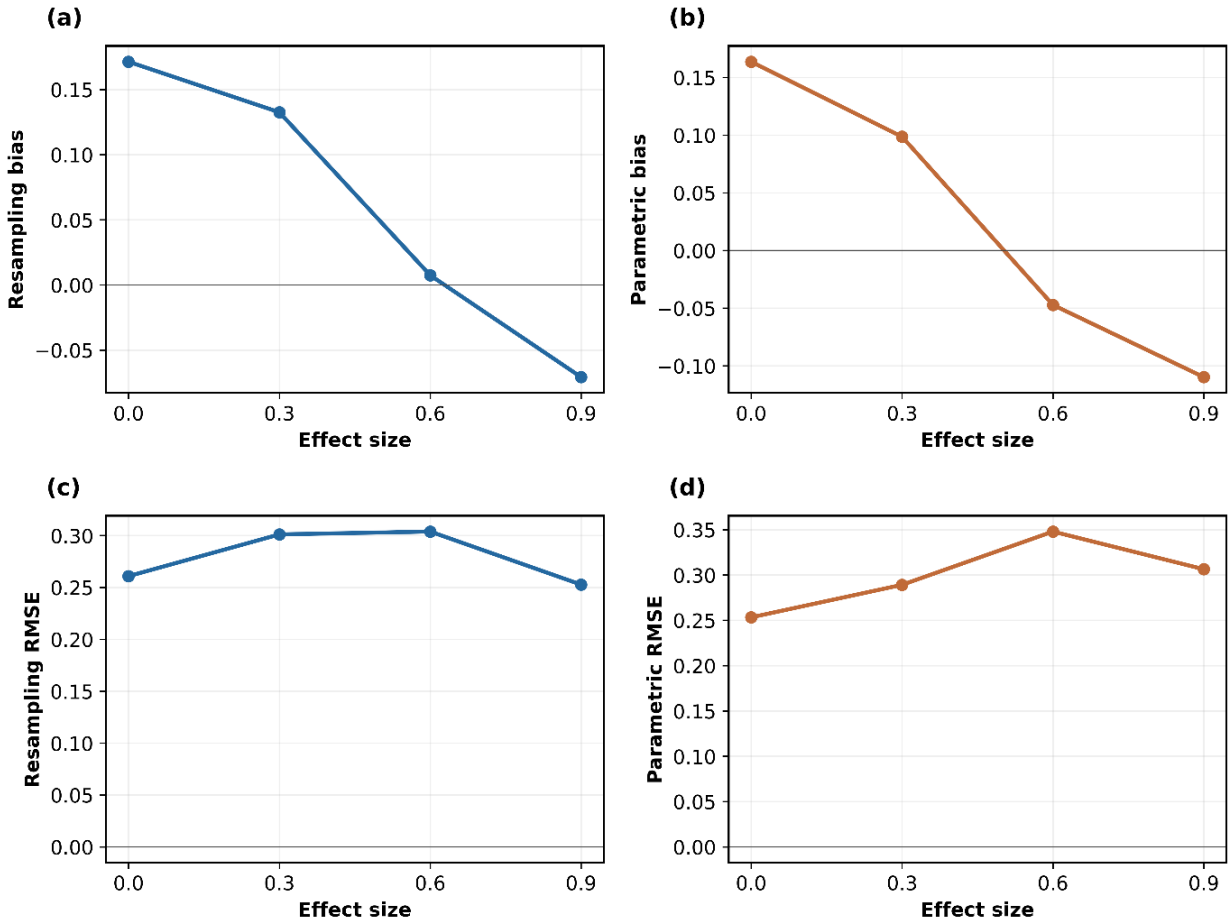


Figure 3. Effect size-dependent variation in statistical estimation accuracy

Note. Panels (a) and (b) show mean signed bias for resampling and parametric estimation, respectively. Panels (c) and (d) show corresponding mean RMSE across four effect sizes. The horizontal zero lines in the bias panels indicate the absence of systematic error.

3.5. Distribution-Specific Comparison of Estimation Accuracy

The accuracy of estimation was different for normal, skewed, and bounded/discrete (ceiling-effect) distributions. The mean absolute bias and RMSE of the resampling for the skewed distribution were significantly lower as shown in Table 5. The mean RMSE for the resampling was 0.2662 while the mean RMSE for the parametric estimation was 0.3280. For normal distribution as well as ceiling effect distribution, parametric estimation results were slightly better in terms of mean RMSE. The patterns for the absolute bias and RMSE for the different distributions are shown in Figure 4 and indicate that the relative performance of the different methods was dependent on the underlying probability distribution.

Table 5. Distribution-Specific Comparison of Bias and RMSE

Distribution type	Resampling absolute bias	Parametric absolute bias	Resampling RMSE	Parametric RMSE
Normal	0.1040	0.0989	0.2820	0.2799
Skewed log-normal	0.1090	0.1707	0.2662	0.3280
Bounded/discrete ceiling	0.1151	0.1111	0.2907	0.2895
Overall	0.1094	0.1269	0.2796	0.2991

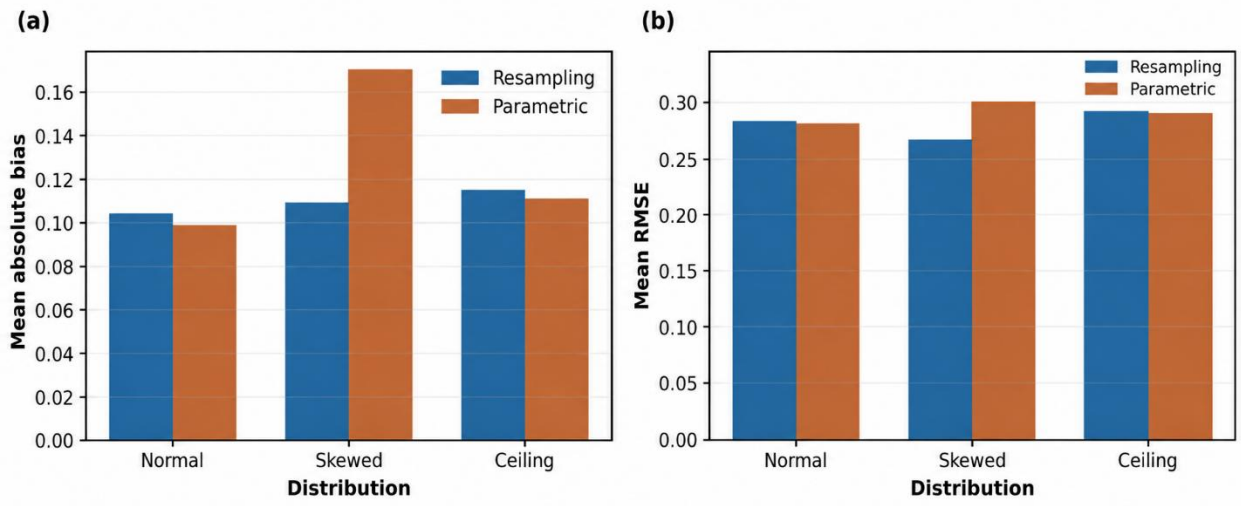


Figure 4. Distribution-specific differences in absolute bias and RMSE

Note. Panel (a) compares mean absolute bias, while panel (b) presents mean RMSE for resampling and parametric estimation across normal, skewed, and ceiling-effect distributions.

3.6. Overall Comparative Performance across Simulation Scenarios

The mean RMSE values for resampling estimation were lower than the mean RMSE values for parametric estimation in all 36 simulation scenarios, and the mean RMSE values for resampling estimation were approximately 6.53% smaller than the mean RMSE values for parametric estimation. A comparison of the two approaches across each individual scenario, however, indicated resampling had lower RMSE in 17 scenarios (47.22%) and parametric estimation in 19 scenarios (52.78%). Similarly, the resampling scenarios showed less absolute bias in 13 scenarios (36.11%) than parametric estimation showed in 23 scenarios (63.89%). The paired RMSE values are shown in Figure 5, and it shows that the mean benefit of resampling was not uniform for each individual scenario. Resampling was most beneficial on average on terms of RMSE under skewed distributional conditions and moderate to large effect magnitudes.

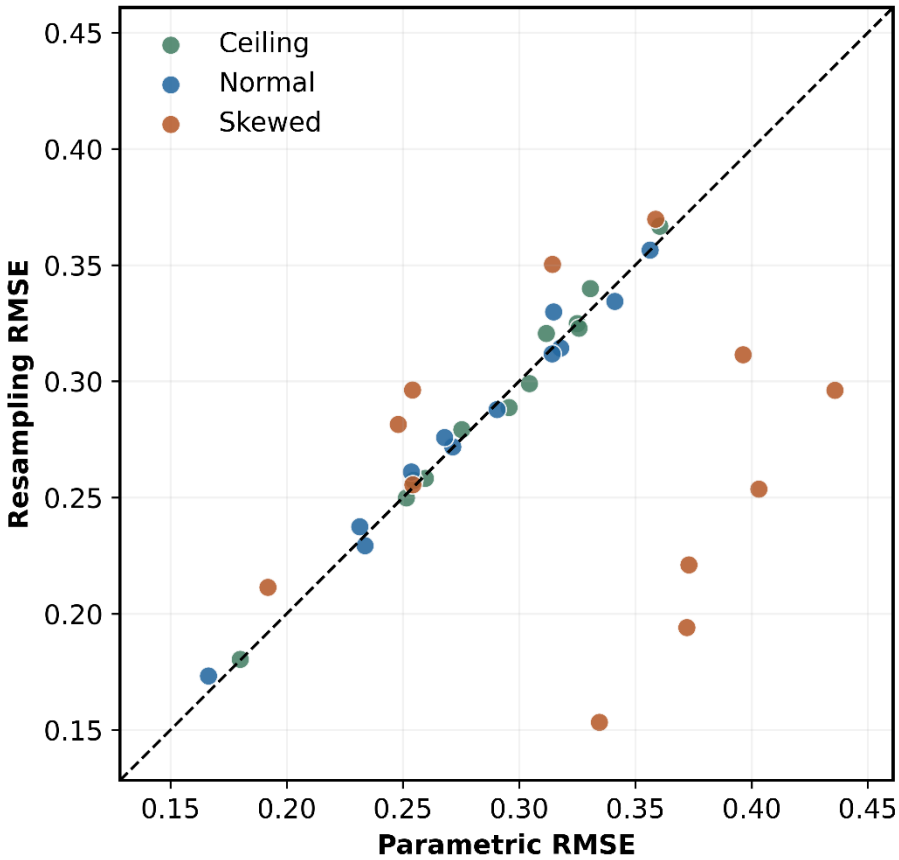


Figure 5. Paired comparison of resampling and parametric RMSE across simulation scenarios

The results collectively indicate that resampling-based power estimation achieved lower average absolute bias and RMSE across the simulated conditions, particularly for skewed distributions and larger effect sizes. Nevertheless, parametric estimation showed advantages in several individual scenarios, especially at smaller effect sizes. These differences emphasize that comparative estimation performance varied with the statistical conditions considered.

4. Discussion

In the present study, the accuracy of resampling-based and parametric methods of statistical power estimation was compared in 36 Monte Carlo simulation scenarios. The results showed that the overall mean root mean square error (RMSE) of resampling (0.2796) was lower than the one of parametric estimation (0.2991), by about 6.53%. Resampling also showed a smaller mean absolute difference (0.1094 versus 0.1269) which represents the average estimation error. For the 36 individual scenarios, parametric estimation was better in 19 of the cases, indicating that the resampling benefits varied with the simulation conditions. Some of the important factors associated with differences in estimation performance were pilot size, size of effect and the probability distribution.

Comparative results show that neither of the estimation methods consistently outperformed the other in all conditions. Lower average RMSE for resampling and lower errors in some individual scenarios for parametric estimation. This variation shows that it is important to consider several accuracy indicators to evaluate the statistical methods instead of one single performance indicator. Differences in predictive performance between alternative modelling techniques have also been found with comparative investigations of machine learning algorithms (Awopejo et al., 2025). Studies that have used hybrid resampling methods have demonstrated their potential applications in enhancing estimation accuracy in data-poor scenarios (Khosravi et al., 2025). Additionally, using heterogeneous or homogeneous ensemble methods, the results have shown the impact of methodological characteristics on the predictive accuracy in analytical settings (Matougui et al., 2024). The results confirm the overall need for systematic methodological comparisons.

The RMSE for both estimation approaches decreased as sample size for the pilots increased. Resampling RMSE for a sample size of 10 dropped from 0.3281 to 0.2363 at 20 observations per group; parametric RMSE from 0.3414 to 0.2593. There was also a significant effect of the magnitude of the effect on comparative performance, with parametric estimation being slightly better when the effect size was 0.0 and 0.3 and resampling being better when the effect size was 0.6 and 0.9. Previous comparative forecasting research has focused on the methodological choice and analytical conditions and their impact on forecasting performance (Koukaras et al., 2024). Likewise, studies with augmented observations have also shown that data characteristics play a crucial role in the accuracy of prediction (Lee et al., 2025). Comparative applications of modelling have also placed an emphasis on comparative performance evaluation of alternative modelling approaches (Martinović et al., 2025).

The accuracy of the estimation was affected by the probability distribution, especially when it was skewed. The mean RMSE for resampling was 0.2662 under the skewed log-normal distribution while that of parametric estimation was 0.3280. Parametric estimation had a slightly lower RMSE in the normal and bounded distributions, however. Distribution sensitive resampling procedures have been shown to be important when using modified bootstrapping for extreme and imbalanced observations (Muhammad et al., 2025). In the field of prediction interval research, which has primarily focused on bootstrap methods, there is also a strong emphasis on assessing uncertainty in computational estimation (Nguinabé et al., 2025). In the same way, adaptive sampling and reweighting strategies have been studied to handle discrepancies in the observational characteristics (Zeng et al., 2023). These studies offer a methodological perspective to explain distribution differences found in the current analysis.

It is argued that when choosing statistical power estimation procedures, systematic and overall estimation errors should be considered. The practical relevance of the accuracy of estimation for quantitative decision-making has been shown through research using model averaging (Matavel et al., 2024). Likewise, comparative predictive studies in energy consumption, soil moisture, and environmental applications have demonstrated the benefits of using alternative modelling approaches (Pai et al., 2025; Settu & Ramaiah, 2025; Taghavi-Bayat et al., 2024). While these studies are not statistical power estimation, they do show that there are other uses of accuracy assessment in computational research. Based on the current results, for small-sample study planning, estimation procedures should be chosen not only by considering conventional methodological assumptions, but also by taking distributional characteristics, pilot sample size and expected effect magnitude into account (Wang & Gao, 2025; Wei et al., 2024).

The study was restricted to a limited number of predefined simulation scenarios, 3 probability distributions, and limited sample sizes and magnitudes of effects for the pilots. In addition, differences between the parametric estimator and the reference statistical power could be due to the use of the t-test approximation of the parametric estimator. Comparisons across conditions were descriptive and not statistically superior across larger populations of simulation conditions. The generalizability and reliability of these findings should be assessed by more data, more probability distributions, other statistical tests and better estimation methods in future research.

5. Conclusion

Comparative accuracy in terms of bias and root mean square error (RMSE) of resampling-based and parametric statistical power estimation methods was examined over 36 Monte Carlo simulation scenarios. Results showed that resampling based estimation had a lower overall mean RMSE (0.2796) than parametric estimation (0.2991), with an approximate 6.53% reduction in estimating error. The mean absolute bias (0.1094) for resampling was also lower than for parametric estimation (0.1269), which is a measure of the average error in estimation. But, in 19 individual scenarios, parametric estimation had lower RMSE, showing that none of the two methods showed consistent superiority in all situations. Estimation performance was affected by pilot sample size, effect magnitude and probability

distribution. Both methods showed decreasing RMSE with increasing pilot sample size, which demonstrates the need for adequate pilot samples for reasonable estimation of the statistical power. Resampling achieved higher accuracy under skewed distributions and moderate to large effect sizes while parametric estimation achieved minor advantages under normal and bounded distributions and smaller effect sizes. The results highlight the need to consider overall estimation error in addition to systematic bias when choosing statistical power estimation procedures. The study also underscores the importance of making distributional assumptions and sample characteristics when dealing with statistical planning. Whilst analysis was limited to a set of pre-determined simulation conditions it has methodological value for quantitative study. Future research should focus on different simulation conditions, other statistical tests and other estimation methods to increase statistical power estimation reliability and generalizability.

References

1. Acici, K. (2025). Comparative analysis of machine and deep learning algorithms for Bragg peak estimation in polymeric materials for tissue-sparing radiotherapy. Polymers, 17(15), 2068.
2. Allal, Z., Noura, H. N., Salman, O., & Chahine, K. (2025). Machine learning-based prediction model of wind turbine power generation. Cleaner Energy Systems, 100218.
3. Alrumaidhi, M., Farag, M. M., & Rakha, H. A. (2023). Comparative analysis of parametric and non-parametric data-driven models to predict road crash severity among elderly drivers using synthetic resampling techniques. Sustainability, 15(13), 9878.
4. Awopejo, T. E., Adigun, P. O., Oyekanmi, T. T., Azeez, N. A. A., Adekanye, M. A., & Obisesan, A. (2025). Machine learning-based prediction of magnetic properties from hysteresis curves: A comparative study of Random Forest, Gradient Boosting, XGBoost, LightGBM and Support Vector Regressions. International Journal of Research Publications in Engineering, Technology and Management (IJRPETM), 8(6), 13456-13479.
5. Carrera, B., & Kim, K. (2024). Comparative analysis of machine learning techniques in predicting wind power generation: A case study of 2018–2021 data from Guatemala. Energies, 17(13), 3158.
6. Chaaban, A. K., & Alfadl, N. (2024). A comparative study of machine learning approaches for an accurate predictive modeling of solar energy generation. Energy Reports, 12, 1293-1302.
7. Cui, B., Chen, W., Weng, D., Wei, X., Sun, Z., Zhao, Y., & Liu, Y. (2023). Observability-constrained resampling-free cubature Kalman filter for GNSS/INS with measurement outliers. Remote sensing, 15(18), 4591.
8. da Silva, E. A., Rocha, D. S., Naves, E. L. M., & Rodrigues, T. B. (2026). Resampling-based power simulation in small samples: a Monte Carlo calibration study (reproducibility package) (Version v1) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.20767954.
9. Deville, L., Theristis, M., King, B. H., Chambers, T. L., & Stein, J. S. (2024). Open-source photovoltaic model pipeline validation against well-characterized system data. Progress in Photovoltaics: Research and Applications, 32(5), 291-303.
10. Dong, G., Shen, F., Sun, L., Zhang, M., & Wei, J. (2024). A Bayesian inferred health prognosis and state of charge estimation for power batteries. IEEE Transactions on Instrumentation and Measurement, 74, 1-12.
11. Elsebaie, I. H., Kawara, A. Q., Alharbi, R., & Alnahit, A. O. (2025). Bias correction methods applied to satellite rainfall products over the western part of Saudi Arabia. Atmosphere, 16(7), 772.
12. Ferkous, K., Guermoui, M., Bellaour, A., boulmaiz, T., & Bailek, N. (2024). Enhancing photovoltaic energy forecasting: a progressive approach using wavelet packet decomposition. Clean Energy, 8(3), 95-108.
13. Fitrianto, A., & Linganathan, P. (2022). Comparisons between resampling techniques in linear regression: A simulation study. CAUCHY: Jurnal Matematika Murni dan Aplikasi, 7(3), 345-353.
14. Han, X., Li, X., Yang, J., Tao, W., Han, G., Wang, J., ... & Li, W. (2025). Evaluation and deep learning-based calibration of nearshore sea surface wind speeds from FY-3E GNOS-II and Tianmu-1 missions. Geo-Spatial Information Science, 28(6), 2709-2720.
15. James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). Resampling methods. In An introduction to statistical learning: With applications in python (pp. 201-228). Cham: Springer International Publishing.
16. Khan, S., Mazhar, T., Khan, M. A., Shahzad, T., Ahmad, W., Bibi, A., ... & Hamam, H. (2024). Comparative analysis of deep neural network architectures for renewable energy forecasting: enhancing accuracy with meteorological and time-based features. Discover Sustainability, 5(1), 533.
17. Khosravi, K., Farooque, A. A., Karbasi, M., Ali, M., Heddam, S., Faghfouri, A., & Abolfathi, S. (2025). Enhanced water quality prediction model using advanced hybridized resampling alternating tree-based and deep learning algorithms. Environmental Science and Pollution Research, 32(11), 6405-6424.
18. Koukaras, P., Mustapha, A., Mystakidis, A., & Tjortjis, C. (2024). Optimizing building short-term load forecasting: A comparative analysis of machine learning models. Energies, 17(6), 1450.
19. Lee, J., Kwon, D. H., Jeong, H., Nam, G., Hwang, E., Kim, J. H., ... & Kim, H. G. (2025). Improving chlorophyll-a estimation using Sentinel-2 data: a comparative analysis of augmented datasets. GIScience & Remote Sensing, 62(1), 2496551.
20. Martinović, M., Dokic, K., & Pudić, D. (2025). Comparative analysis of machine learning models for predicting innovation outcomes: an applied AI approach. Applied Sciences, 15(7), 3636.
21. Matavel, C. E., Meyer-Aurich, A., & Piepho, H. P. (2024). Model-averaging as an accurate approach for ex-post economic optimum nitrogen rate estimation. Precision Agriculture, 25(3), 1324-1339.

22. Matougui, Z., Djerbal, L., & Bahar, R. (2024). A comparative study of heterogeneous and homogeneous ensemble approaches for landslide susceptibility assessment in the Djebahia region, Algeria. Environmental Science and Pollution Research, 31(28), 40554-40580.
23. Muhammad, M., Ul-Saufie, A. Z., Radi, N. F. A., Noor, N. M., & Gusnanto, A. (2025). Modified resampling strategy for extreme values in imbalanced air pollution data using moving block bootstrapping approach with relevance weighting (MBB-RW). Scientific Reports, 16(1), 290.
24. Nguinabé, J., Rockefeller, R., Khan, N. M., & Rodrigues, P. C. (2025). Bootstrap prediction intervals for the long short term memory (LSTM) recurrent neural network. Expert Systems with Applications, 284, 127728.
25. Pai, A., Mishra, K. K., Jeyan, J. M. L., & Sayal, A. (2025). Enhanced household energy consumption forecasting using multivariate long short-term memory (LSTM) networks with weather data integration. Results in Engineering, 106512.
26. Settu, P., & Ramaiah, M. (2025). A data driven comparison of hybrid machine learning techniques for soil moisture modeling using remote sensing imagery. Scientific Reports, 15(1), 43170.
27. Taghavi-Bayat, A., Gerke, M., & Riedel, B. (2024). Soil moisture retrieval at high spatial resolution over alpine ecosystems on Nagqu-Tibetan plateau: A comparative study on semiempirical and machine learning approaches. Science of Remote Sensing, 9, 100135.
28. Wang, L., & Gao, Y. (2025). Estimating and downscaling ESA-CCI soil moisture using Multi-Source remote sensing images and Stacking-Based ensemble learning algorithms in the Shandian River Basin, China. Remote Sensing, 17(4), 716.
29. Wei, Y., Hashim, H. B., Lai, S. H., Chong, K. L., Huang, Y. F., Ahmed, A. N., ... & El-Shafie, A. (2024). Comparative analysis of artificial intelligence methods for streamflow forecasting. IEEe Access, 12, 10865-10885.
30. Zeng, Z., Tao, L., Su, R., Zhu, Y., Meng, L., Tuheti, A., ... & Chen, C. (2023). Unsupervised transfer learning approach with adaptive reweighting and resampling strategy for inter-subject EOG-based gaze angle estimation. IEEE Journal of Biomedical and Health Informatics, 28(1), 157-168.