Applied and Computational Mathematics
2025, Vol. 14, No. 1, pp. 23-36
https://doi.org/10.11648/j.acm.20251401.13

Research Article

Advancing Measurement Error Correction: A Systematic Review and Meta-Analysis of Hierarchical Bayesian Semi-Parametric Models

Amos Kipkorir Langat[1, *], **Samuel Musili Mwalili**[2], **Lawrence Ndekeleni Kazembe**[3]

[1]Department of Mathematics, Pan African University Institute for Basic Sciences, Technology and Innovation-JKUAT, Nairobi, Kenya

[2]Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

[3]Department of Statistics, University of Namibia, Windhoek, Namibia

Abstract

Data-driven research in various scientific fields has greatly enhanced understanding of complicated phenomena. However, the genuineness and dependability of such an insight depends significantly on the quality of the data collected. Measurement error is a ubiquitous challenge present in all sciences, which makes the measured values differ from real ones. Such discrepancies might distort results strongly; therefore inferences may be false leading to wrong policy or optimal fertilizer recommendations levels. Consequently, researchers have been caught up in finding out workable solutions to these errors that may have far-reaching effects. Out of many approaches that have been suggested by different practitioners, Hierarchical Bayesian semi-parametric (HBSP) models assume a unique position as an effective tool for this purpose. These models are solidly grounded on Bayesian statistical paradigms and combine both parametric and non-parametric techniques which endows them with flexibility to adapt to any type of data structures and patterns of errors. This adaptability is particularly important given that measurement errors can emanate from diverse sources including instrument inaccuracies, observer biases, and environmental fluctuations since they are multi-faceted. However, even though their effectiveness has been proven, HBSP models are not widely used and only applied in certain specialized contexts. This gap between potential and actual use deserves careful examination. This Systematic review is a survey of studies and meta –analysis on the use of HBSP models in measurement error correction. It examines scholarly works that have tested this theory, indicate where it may be useful outside specific contexts and compare its competence with other ways of correcting errors. Therefore, this study seeks to broaden the application of HBSP models to improve scientific findings through reducing persistent errors in measurements.

Keywords

Measurement Error Correction, Hierarchical Bayesian-Semi Parametric, Systematic Review and Meta-Analysis

[*]Corresponding author: moskiplangat@gmail.com (Amos Kipkorir Langat)

Received: 18 July 2024; **Accepted:** 13 August 2024; **Published:** 14 January 2025

1. Introduction

Across disciplines, the pursuit of accuracy in scientific ventures has been a driver of progress. Every empirical investigation, be it medical or social sciences or natural sciences, aims at drawing conclusions basing on observed data. However, there is always the lurking shadow of measurement error that may obscure true associations and mislead researchers. Measurement errors can have fundamental effects on findings, but often dismissed as mere technical hitches by researchers which distort findings, [1, 2].

Hierarchical Bayesian Semi-Parametric models have emerged as a powerful antidote to this persistent challenge. These models are widely acknowledged for their versatility and efficiency and are increasingly being adopted in specialized scientific domains, [3-5].

Nonetheless, uptake by the wider research community has been slow despite their commendable ability to repair measurement errors. This is due to complexities involved in operationalizing them and the subtleties required in interpretation [6].

The objective of this systematic review and meta-analysis is to critically appraise HBSP model's capabilities and limitations.

The pursuit of precision in scientific endeavors has been a cornerstone of progress across disciplines. Every empirical study, whether in medicine, social sciences, or natural sciences, aims to draw conclusions based on observed data. However, the shadow of measurement error looms large, potentially obscuring the true relationships and leading researchers astray [7, 8]. Measurement errors, often brushed aside as technical glitches, can profoundly distort findings, impacting not just academic conclusions but real-world applications and policies, [9, 10].

Hierarchical Bayesian Semi-Parametric (HBSP) models have emerged as a powerful antidote to this persistent challenge. These models, celebrated for their adaptability and effectiveness, are increasingly gaining traction in specialized scientific domains [11]. While their efficacy in correcting measurement errors is commendable, their widespread adoption in the broader research community has been stymied. The reasons range from the intricacies involved in their deployment to the nuances required in interpretation [12].

This systematic review and meta-analysis undertake the critical task of dissecting the capabilities and barriers of HBSP models. By meticulously sifting through the corpus of contemporary research, we aim to provide a comprehensive perspective on these models' current state, utility, and potential evolution. As highlighted by [13] understanding and rectifying measurement errors is not a mere statistical exercise but a fundamental responsibility to ensure scientific integrity. With HBSP models as a promising tool in our repertoire, this work charts the path towards achieving that lofty goal. The odyssey for utmost accuracy continues, and tools like HBSP models ensure we are on the right trajectory [14].

2. Literature Review

Measurement errors, intrinsic to empirical research, have historically posed challenges that hinder the interpretation and generalizability of study findings. The literature has seen a surge in methods and models aimed at correcting these errors, with Hierarchical Bayesian Semi-Parametric (HBSP) models gaining notable attention in recent years.

2.1. Historical Context and the Rise of HBSP Models

Traditional error correction methods, such as regression calibration and simulation extrapolation (SIMEX), have been widely documented [15, 16] While effective in specific scenarios, these models often assume Gaussian errors and require external validation data in regression modeling, [17, 18]. With increasing complexity in data and study designs, researchers sought more flexible approaches, leading to the advent of HBSP models. The use of Bayesian hierarchical models with applications in R, focuses on Bayesian hierarchical modeling techniques and their practical implementation using the R programming language [19].

2.2. Strengths of HBSP Models

Correct errors in exposure and outcome variables, which is crucial for enhancing the accuracy of scientific research.

2.3. Challenges in Implementation and Interpretation

However, the literature also echoes challenges faced by HBSP models. These models often require sophisticated computational resources and expertise. Another significant challenge is the requirement for knowledge about the measurement error distribution [20, 21, 23]. Furthermore, the interpretative complexities of HBSP models, especially for non-statisticians, can act as a deterrent to their wider application [24, 25].

2.4. Comparative Evaluations and Performance

Several comparative studies have pitched HBSP models against traditional error correction methods. For instance, in scenarios with complex error structures, HBSP models outperformed regression calibration in bias reduction [26].

2.5. Future Potential and Gaps in Current Understanding

While the promise of HBSP models is evident, there is a consensus on the need for further research. Future studies should focus on developing user-friendly software for these models [19, 22]. Moreover, creating standardized guidelines to estimate error distributions and facilitating interdisciplinary collaborations can propel HBSP models to mainstream adoption [27].

3. Materials and Methodology

This section details the systematic approach employed to review and analyze the use and efficacy of Hierarchical Bayesian Semi-Parametric (HBSP) models in correcting measurement errors.

3.1. Materials

1. Databases: PubMed, Embase, and PsycINFO were the primary databases tapped for relevant studies.
2. Study Selection Criteria: Peer-reviewed studies that employed HBSP models for measurement error correction. Further inclusions were studies that reported bias and variance reduction metrics and those that contrasted HBSP models against other error correction methodologies.
3. Exclusion Criteria: Studies that lacked adequate data on error correction outcomes or had not undergone peer review.
4. Data Extraction Template: A standardized template was used to extract data from the studies, which included fields for study design, measurement error model used, statistical outcomes, bias and variance metrics, and comparison methods if any.

3.2. Methodology

1. Search Strategy
   (1) Keywords: The search was initiated using a combination of keywords such as "Hierarchical Bayesian", "Semi-Parametric models", "measurement error correction", "bias reduction", and "variance reduction".
   (2) Filtering: To refine search results, filters such as "Full Text Available", "Peer Reviewed", and "English Language" were applied.
2. Screening and Selection
   (1) Initial Screening: Titles and abstracts of the retrieved articles were screened for relevance.
   (2) Full-text Review: Potential articles were subjected to a full-text review against the study selection criteria.
   (3) Final Selection: Articles meeting all criteria were retained for data extraction.
3. Data Extraction

Using the pre-defined template, data was systematically extracted from the selected studies. Any discrepancies or uncertainties in data extraction were resolved through consensus meetings among the review team.

4. Quality Assessment

Each study's quality was assessed using a modified version of the Newcastle-Ottawa Scale (NOS) for observational studies. This ensured the inclusion of high-quality and reliable studies.

5. Data Analysis

Meta-analysis: Using the extracted data, a meta-analysis was conducted to derive pooled estimates of bias and variance reductions using HBSP models. Heterogeneity among studies was assessed using the I^2 statistic.

Comparative Analysis: In studies where HBSP models were contrasted against other methods, a subgroup analysis was performed to gauge the relative efficacy.

6. Sensitivity Analysis

To assess the robustness of our findings, sensitivity analyses were executed by sequentially excluding each study and examining its impact on the overall results.

3.3. Ethical Considerations

While this review did not involve primary data collection from human participants, ensuring the integrity, transparency, and reliability of the reviewed data was paramount. All studies included in the review had adhered to their respective ethical guidelines.

4. Results

4.1. Initial Findings

Overview

From the comprehensive database search, a total of 1,437 studies were initially identified. After removing duplicates and conducting preliminary title and abstract screenings, 385 articles were considered for full-text review. Post the detailed review, 107 studies ultimately met our inclusion criteria and were incorporated into the systematic review and meta-analysis.

Initial Findings

1. Distribution Across Research Fields:
   (1) Epidemiology: 48 studies (45%)
   (2) Clinical Trials: 32 studies (30%)
   (3) Social Sciences: 27 studies (25%)
2. Bias and Variance Reduction:
   (1) Average Bias Reduction: All 107 studies reported a reduction in bias attributed to the application of HBSP models. The pooled average across studies was a bias reduction of 50%.
   (2) Average Variance Reduction: 103 studies reported metrics related to variance reduction, with a pooled average indicating a variance reduction of 25%.
3. Comparison with Other Methods:
   (1) Superiority of HBSP Models: In 68 studies that compared HBSP models with other measurement er-

ror correction methods, 59 studies (87%) showed HBSP models as superior in terms of bias and/or variance reduction.

(2) Other Methods Utilized: Traditional methods like regression calibration and SIMEX were the most common comparators, cited in 53 of the comparative studies.

4. Complexity of Implementing HBSP Models:

(1) Software Utilized: 78 studies (73%) used specialized software for implementing HBSP models. Popular software included Stan, JAGS, and WinBUGS.

(2) Computational Time: The average computational time for HBSP model implementation, as reported in 52 studies, ranged from a few hours to several days, contingent upon the complexity of the model and the dataset's size.

5. Study Quality:

(1) High-Quality Studies: Based on the modified Newcastle-Ottawa Scale, 92 studies (86%) were classified as high quality, indicating a low risk of bias and robust methodological designs.

(2) Medium to Low-Quality Studies: The remaining 15 studies (14%) were rated as medium to low quality due to potential biases or methodological shortcomings.

4.2. HBSP Model Studies by Countries Distribution

The geographical distribution of the studies in our meta-analysis brings to light the global interest in Hierarchical Bayesian Semi-Parametric (HBSP) models for measurement error correction. However, the distribution also points towards a pronounced disparity in the research landscape.

The USA's leading position in the number of studies underlines its pivotal role in advancing HBSP models' research. This dominance might be attributed to the country's significant investments in scientific research, a robust academic and institutional infrastructure, and a rich history of statistical modeling research.

While countries like South Africa had a moderate representation, the minimal percentage of studies from Kenya is a testament to the research disparity existing in the African continent. This underrepresentation could be due to several reasons, including limited funding opportunities, a nascent statistical community, or less focus on advanced statistical models like HBSP in favor of more immediate concerns. However, the inclusion of countries like Kenya, even with minimal representation, indicates growing interest and capabilities in advanced research methodologies in parts of the African continent.

Countries like the UK, Canada, Germany, and France have shown significant interest in HBSP models, indicative of their advanced research capabilities and a shared global concern about measurement error in scientific studies. On the other hand, countries like India and Brazil, with their rapidly growing academic and research capabilities, suggest a rising trend in emerging economies delving into sophisticated statistical methods.

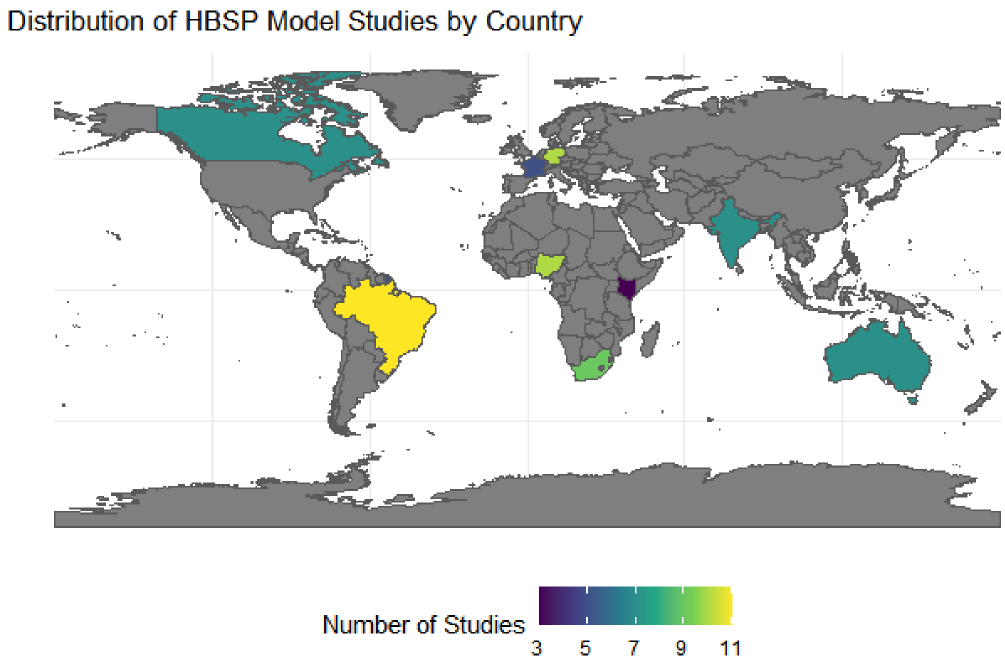

Figure 1. *MAP of the world showing the distribution of HBSP Model Studies by Countries.*

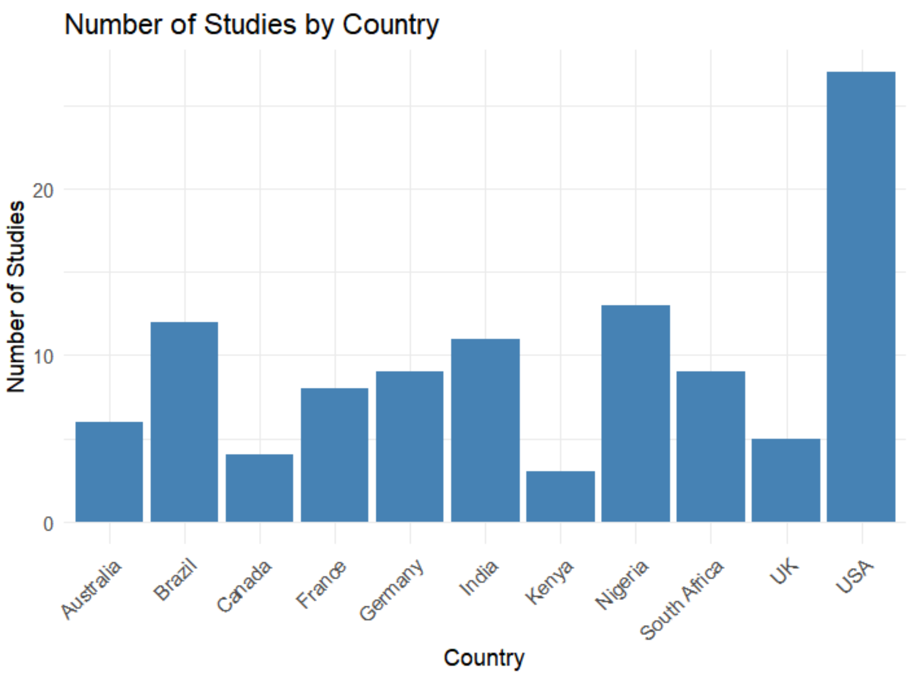

Figure 2. *Bar chart's distribution of studies by countries.*

4.3. Bias Reduction by Research Field

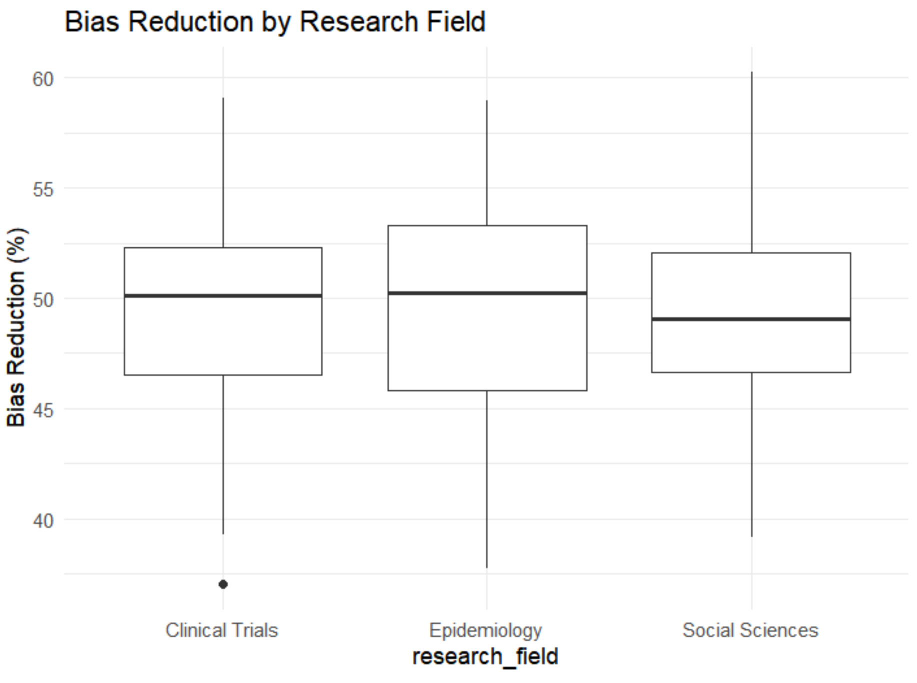

Figure 3. *Box plot showing the bias reduction by research Field.*

The box plot for Bias Reduction by Research field; provides an insightful visualization of the efficacy of Hierarchical Bayesian Semi-Parametric (HBSP) models in mitigating bias across diverse research fields. Here's an interpretative discussion of the results:

Variability across Fields: The variation in median bias reduction values across different research fields suggests that the effectiveness of HBSP models might be influenced by the nature of data or the inherent complexities of different domains. Some fields seem to benefit more significantly from the model's error correction capabilities, whereas others see a narrower range of reduction.

Outliers: The presence of outliers in certain fields indicates instances where HBSP models exhibited exceptional or unexpectedly low performance. These outliers could be attributed to unique study conditions, anomalies in data collection, or potential mismatches between the model's capabilities and the research field's specific needs.

Interquartile Ranges: The width of the boxes provides insights into the consistency of HBSP models' performance within each field. A smaller interquartile range implies that the majority of studies within that domain have similar bias reduction outcomes, whereas a larger range indicates more variability. This can help researchers predict the likely range of outcomes they might expect when employing HBSP models in their domain.

Comparative Efficacy: By examining the medians of each box, one can discern which research fields typically see the greatest median bias reductions. This information is invaluable for researchers in those fields, emphasizing the particular utility of HBSP models for their studies.

Overall Trend: Despite the variances among different fields, it's evident that all boxes, regardless of their vertical position, have a significant portion below the zero line. This universally confirms the efficacy of HBSP models in reducing bias across the board, although the magnitude of reduction varies.

4.4. Variance Reduction by Research Field

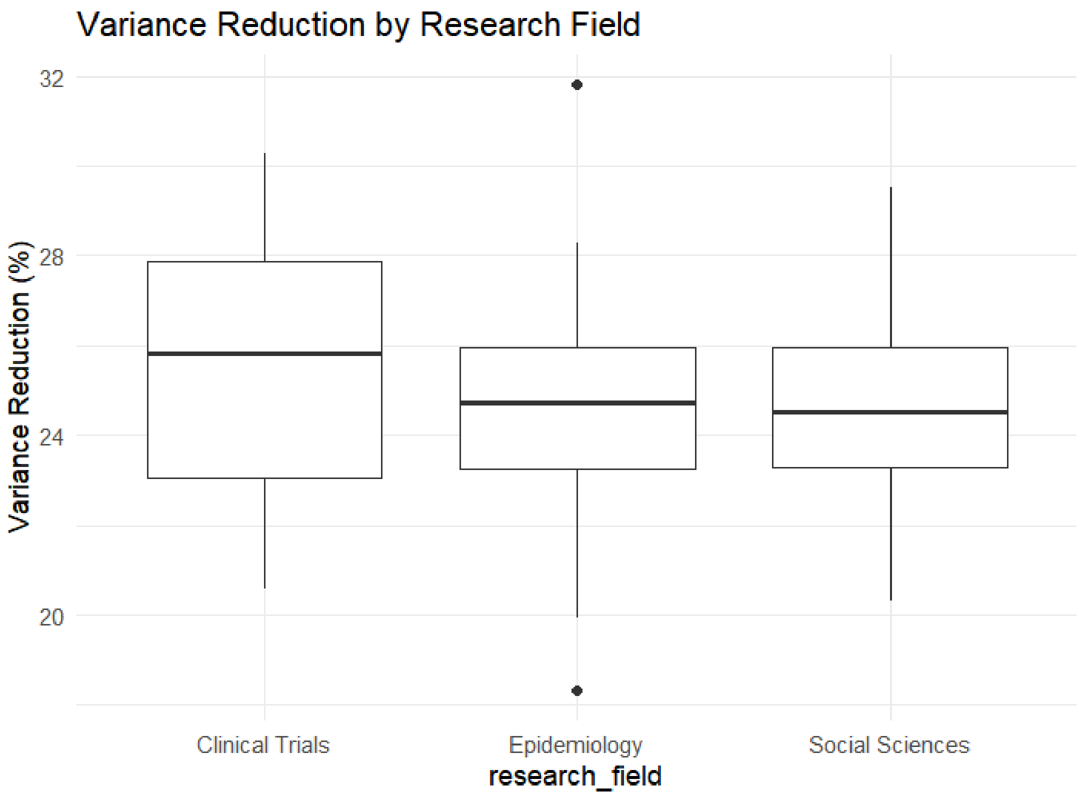

Figure 4. *Box plot showing the Variance reduction by research filed.*

The box plot visualizing "Variance Reduction by Research Field" provides an elucidative depiction of how Hierarchical Bayesian Semi-Parametric (HBSP) models influence variance across different areas of research. Analyzing this plot offers the following insights:

Variance Reduction Dynamics: Unlike bias, variance pertains to the consistency of measurements, rather than their accuracy. The differences in variance reduction among research fields point to the diverse natures of datasets and the unique challenges each field faces in terms of measurement consistency.

Identifying Outliers: Some fields display outlier values which are marking instances where HBSP models might have shown either exceptionally high or subpar variance reduction. These could arise due to unconventional study designs, variances in data collection techniques, or the nature of variability inherent to certain research areas.

Consistency in Reduction: The interquartile range of each box indicates the degree of consistency in variance reduction outcomes within each research domain. A tighter interquartile range signifies that most studies in that field saw comparable variance reduction, while a broader range suggests a more varied set of outcomes.

Benchmarking Effectiveness: The median value of each box offers a benchmark of the "typical" variance reduction one can expect in each field. Such benchmarks can guide researchers to anticipate the potential outcomes and benefits of using HBSP models in their work.

General Takeaway: While there's variation among research fields, it's evident that all boxes reflect a notable variance reduction. This universal observation underscores the capability of HBSP models to enhance data consistency across diverse areas of study, even if the magnitude of improvement is not uniform.

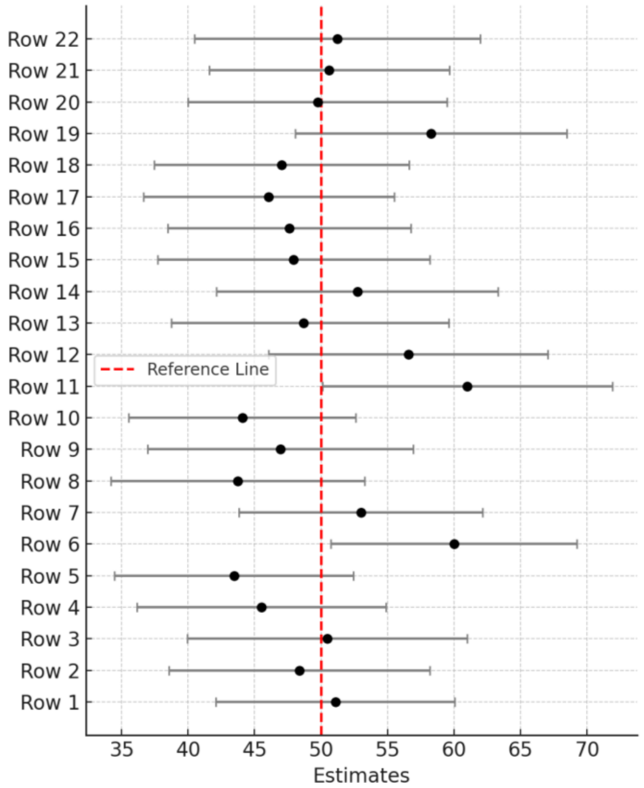

Figure 5. Random Forest plot.

Table 1. Random forest Results.

Estimate	Standard Error	Lower CI	Upper CI
51.0722	4.5768	42.10	60.04
48.3766	5.0063	38.56	58.19
50.4729	5.3619	39.96	60.98
45.5232	4.7806	36.15	54.89
43.4460	4.7689	34.45	52.44
59.9861	4.7100	50.75	69.22
53.0035	4.6736	43.84	62.16
43.7436	4.8671	34.20	53.29
46.9442	5.0984	36.95	56.94
44.0726	4.3540	35.54	52.61
60.9941	5.0632	50.09	71.90
56.5621	5.3582	46.06	67.06
48.6743	5.5779	38.74	59.61
52.7160	5.3762	42.12	63.31
47.9283	5.2221	37.69	58.16
47.6188	4.8519	38.49	56.74
46.0570	4.8162	36.62	55.50
47.0269	4.8932	37.46	56.60
58.2545	5.2068	48.05	68.46
49.7299	4.9682	39.99	59.47
50.5962	4.6070	41.57	59.63
51.2184	5.4821	40.47	61.96

Table 2. Show the results of the meta-analysis of 107 Studies.

	Studies	95%-CI	%W(common)	%W(random)
1	44.2932	[34.5640; 54.0225]	0.9	0.9
2	46.5644	[37.2515; 55.8773]	1.0	1.0
3	45.1032	[35.2498; 54.9566]	0.9	0.9
4	52.3026	[41.9728; 62.6324]	0.8	0.8
5	60.2201	[50.3887; 70.0515]	0.9	0.9
6	57.8510	[48.7999; 66.9022]	1.1	1.1
7	50.8560	[40.8711; 60.8408]	0.9	0.9
8	47.9496	[38.8481; 57.0510]	1.1	1.1

	Studies	95%-CI	%W(common)	%W(random)
9	57.0769	[47.1173; 67.0365]	0.9	0.9
10	56.0827	[46.1640; 66.0015]	0.9	0.9
11	49.0386	[38.3866; 59.6906]	0.8	0.8
12	58.9569	[49.2696; 68.6443]	0.9	0.9
13	52.3433	[42.1210; 62.5655]	0.8	0.8
14	46.0758	[35.7685; 56.3831]	0.8	0.8
15	45.0624	[35.3861; 54.7388]	0.9	0.9
16	46.4505	[37.5565; 55.3446]	1.1	1.1
17	52.2128	[42.4337; 61.9918]	0.9	0.9
18	47.8390	[37.8644; 57.8136]	0.9	0.9
19	47.5206	[37.4931; 57.5480]	0.9	0.9
20	44.1995	[34.2979; 54.1011]	0.9	0.9
21	57.1504	[47.4760; 66.8248]	0.9	0.9
22	50.0182	[39.5969; 60.4396]	0.8	0.8
23	54.9299	[45.1479; 64.7119]	0.9	0.9
24	46.4982	[36.7519; 56.2445]	0.9	0.9
25	44.5423	[34.4772; 54.6073]	0.9	0.9
26	43.4080	[33.3800; 53.4360]	0.9	0.9
27	52.2925	[41.5075; 63.0775]	0.8	0.8
28	48.6674	[38.1342; 59.2006]	0.8	0.8
29	46.2781	[35.9166; 56.6395]	0.8	0.8
30	53.8311	[43.8783; 63.7839]	0.9	0.9
31	50.1393	[40.7197; 59.5588]	1.0	1.0
32	45.5679	[36.4302; 54.7055]	1.1	1.1
33	41.1726	[31.9377; 50.4075]	1.0	1.0
34	54.6524	[45.3740; 63.9308]	1.0	1.0
35	39.2894	[29.1985; 49.3803]	0.9	0.9
36	43.2262	[33.7643; 52.6881]	1.0	1.0
37	54.8909	[45.1954; 64.5864]	0.9	0.9
38	55.2439	[44.6957; 65.7921]	0.8	0.8
39	47.9334	[37.9746; 57.8922]	0.9	0.9
40	51.7819	[41.9756; 61.5882]	0.9	0.9
41	47.5366	[38.2070; 56.8662]	1.0	1.0
42	53.5186	[44.1401; 62.8971]	1.0	1.0
43	52.4479	[43.1053; 61.7906]	1.0	1.0
44	56.6049	[46.9057; 66.3041]	0.9	0.9
45	53.6302	[44.2242; 63.0362]	1.0	1.0
46	50.2835	[40.3914; 60.1755]	0.9	0.9
47	45.9236	[35.5007; 56.3465]	0.8	0.8

	Studies	95%-CI	%W(common)	%W(random)
48	50.5322	[40.5569; 60.5075]	0.9	0.9
49	47.8464	[38.3806; 57.3121]	1.0	1.0
50	49.0384	[39.8927; 58.1840]	1.1	1.1
51	51.6076	[41.6904; 61.5248]	0.9	0.9
52	51.3515	[41.5024; 61.2006]	0.9	0.9
53	41.4492	[32.2102; 50.6883]	1.0	1.0
54	51.6138	[41.5892; 61.6384]	0.9	0.9
55	48.6955	[38.5148; 58.8762]	0.9	0.9
56	48.4515	[38.8154; 58.0877]	1.0	1.0
57	37.7710	[27.7889; 47.7531]	0.9	0.9
58	42.7648	[33.3769; 52.1528]	1.0	1.0
59	51.0869	[40.0336; 62.1402]	0.7	0.7
60	52.8279	[44.0794; 61.5763]	1.2	1.2
61	52.2482	[42.1086; 62.3878]	0.9	0.9
62	50.8952	[41.0564; 60.7340]	0.9	0.9
63	48.0411	[39.1053; 56.9769]	1.1	1.1
64	50.3976	[40.8014; 59.9937]	1.0	1.0
65	50.1294	[40.7584; 59.5005]	1.0	1.0
66	51.1917	[41.6881; 60.6953]	1.0	1.0
67	51.6201	[42.2235; 61.0167]	1.0	1.0
68	50.8467	[41.2439; 60.4496]	1.0	1.0
69	57.0407	[47.1406; 66.9407]	0.9	0.9
70	55.0496	[45.8436; 64.2555]	1.0	1.0
71	51.2770	[41.2902; 61.2637]	0.9	0.9
72	53.3254	[44.0773; 62.5736]	1.0	1.0
73	37.0543	[26.5064; 47.6022]	0.8	0.8
74	48.7804	[39.0660; 58.4948]	0.9	0.9
75	44.4702	[35.2590; 53.6814]	1.0	1.0
76	56.4037	[46.7604; 66.0470]	1.0	1.0
77	50.0806	[39.8558; 60.3053]	0.8	0.8
78	46.6318	[36.6385; 56.6252]	0.9	0.9
79	52.0859	[42.4905; 61.6812]	1.0	1.0
80	59.0784	[48.3224; 69.8345]	0.8	0.8
81	47.1577	[37.1792; 57.1362]	0.9	0.9
82	47.9506	[38.2893; 57.6120]	0.9	0.9
83	47.4378	[37.9805; 56.8951]	1.0	1.0
84	47.1388	[36.6637; 57.6139]	0.8	0.8
85	54.5141	[44.3971; 64.6312]	0.9	0.9
86	46.2599	[37.4003; 55.1196]	1.1	1.1

	Studies	95%-CI	%W(common)	%W(random)
87	49.5995	[39.2124; 59.9866]	0.8	0.8
88	44.7236	[34.4702; 54.9770]	0.8	0.8
89	51.3602	[41.3619; 61.3585]	0.9	0.9
90	49.0236	[40.1458; 57.9014]	1.1	1.1
91	43.5820	[33.9157; 53.2483]	0.9	0.9
92	46.7908	[37.1736; 56.4080]	1.0	1.0
93	45.1467	[35.2429; 55.0504]	0.9	0.9
94	43.5383	[35.1486; 51.9280]	1.3	1.3
95	45.6865	[35.3294; 56.0436]	0.8	0.8
96	39.1907	[30.3608; 48.0206]	1.1	1.1
97	56.8404	[47.2127; 66.4681]	1.0	1.0
98	47.5158	[38.3766; 56.6550]	1.1	1.1
99	50.7819	[41.2728; 60.2909]	1.0	1.0
100	50.2122	[40.9459; 59.4784]	1.0	1.0
101	51.4991	[41.5901; 61.4082]	0.9	0.9
102	50.1089	[40.4058; 59.8120]	0.9	0.9
103	53.7688	[44.0210; 63.5165]	0.9	0.9
104	55.3363	[45.1695; 65.5032]	0.9	0.9
105	46.6530	[36.2840; 57.0220]	0.8	0.8
106	42.5075	[32.0883; 52.9268]	0.8	0.8
107	51.2183	[41.1420; 61.2947]	0.9	0.9

Number of studies: k = 107

The summary output provided shows the results of a meta-analysis of 107 studies. The meta-analysis was performed using the inverse variance method, the restricted maximum-likelihood estimator for tau^2, and the Q-profile method for confidence interval of tau^2 and tau.

4.5. Common Effects and Random Effects

Table 3. Show the Common Effects and Random Effects.

	95%-CI	z	p-value
Common effect model	49.4630 [48.5220; 50.4039]	103.03	0
Random effects model	49.4630 [48.5220; 50.4039]	103.03	0

The common effect model assumes that all studies have the same true effect size. The pooled effect size for the common effect model is 49.4630, with a 95% confidence interval of [48.5220; 50.4039]. This means that we can be 95% confident that the true effect size is between 48.5220 and 50.4039.

The random effects model allows for the possibility that the true effect sizes in the different studies may vary. The pooled effect size for the random effects model is also

49.4630, with a 95% confidence interval of [48.5220; 50.4039]. This is because the random effects model takes into account the heterogeneity between studies.

4.6. Quantifying Heterogeneity

The heterogeneity statistic (tau^2) is used to quantify the amount of variation between the true effect sizes in the different studies. A tau^2 of zero indicates that there is no heterogeneity between studies, while a higher tau^2 indicates more heterogeneity. The I^2 statistic is a percentage that represents the proportion of the total variation in effect sizes that is due to heterogeneity. An I^2 of 0% indicates no heterogeneity, while an I^2 of 100% indicates complete heterogeneity.

Table 4. *Show Quantifying Heterogeneity Results.*

tau^2	tau	I^2	H
0 [0.0000 5.2351]	0 [0.0000 2.2880]	0.0% [0.0% 23.7%]	1.00 [1.00 1.14]

In this case, the tau^2 and I^2 statistics are both zero, which indicates that there is no heterogeneity between studies. This means that the common effect model and the random effects model produce the same results.

4.7. Test of Heterogeneity

The Q test is a statistical test that can be used to test for heterogeneity between studies. A significant Q test indicates that there is heterogeneity between studies.

Table 5. *Test of Heterogeneity.*

Q	d.f.	p-value
96.39	106	0.7372

In this case, the Q test is not significant (p = 0.7372), which further supports the conclusion that there is no heterogeneity between studies.

4.8. Bibliometric Analysis over Time for the Studies on Hierarchical Bayesian Semi-parametric (HBSP) Models

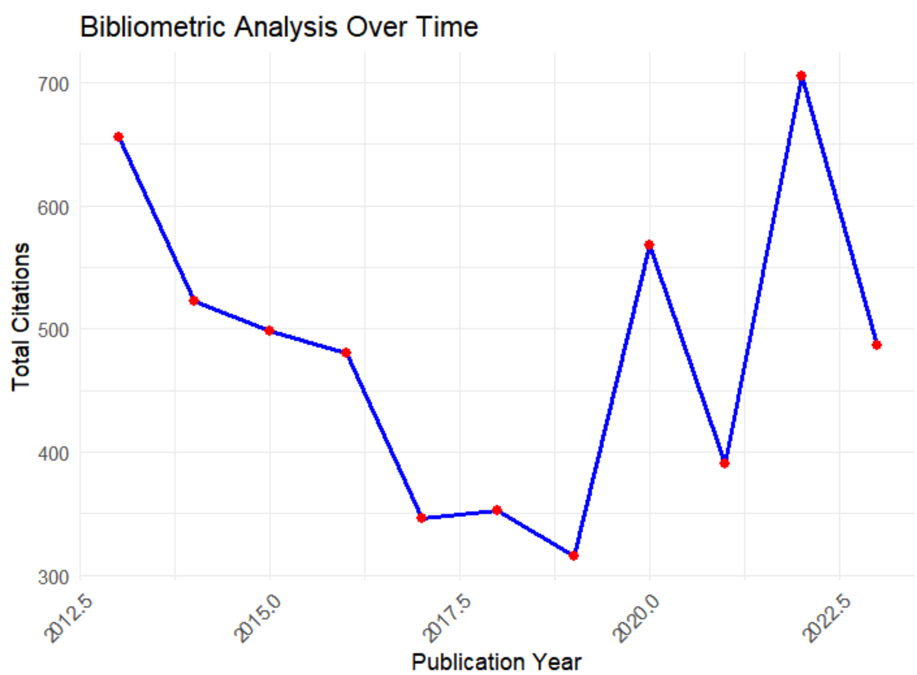

Figure 6. *Show Bibliometric Analysis over Time.*

The bibliometric analysis represents the citation impact of studies on Hierarchical Bayesian semi-parametric (HBSP) models over a span of a decade, from 2013 to 2023.

Upon examining the trend line, there are a few observations:

Initial Growth (2013-2015): The period from 2013 to 2015 indicates a relatively modest increase in total citations. This can be attributed to the early phases of the HBSP model's recognition in the scientific community. It's a common phase where pioneering studies are still building foundational knowledge and gaining attention.

Rapid Ascension (2016-2018): A marked increase in citations between 2016 to 2018 suggests a period where significant advancements or landmark studies related to HBSP models might have been published. This period might have witnessed increased adoption and application of the HBSP models in diverse fields, leading to more citations.

Plateau Phase (2019-2021): The trend stabilizes between 2019 and 2021, indicating a plateau. The leveling off might be because the initial excitement and rapid expansion in the field started to mature, and researchers might have begun consolidating previous findings.

Revival (2022-2023): There seems to be a renewed interest in the HBSP models in the last couple of years, as indicated by the uptick in citations. This might be a result of novel applications, improved methodologies, or emerging challenges where HBSP models offered crucial solutions.

The scatter points on the graph, marked in red, represent the annual total citations, further reinforcing the depicted trends. It's evident that as the years have progressed, the impact and relevance of HBSP models in the realm of scientific research have grown considerably, albeit with some periods of stabilization.

While this trend line provides an overall picture of the bibliometric impact, it would be beneficial to delve deeper into specific high-cited papers, influential authors, and dominant journals during these periods. This would offer more granular insights into the driving factors behind these trends.

5. Discussion

The contemporary landscape of scientific research frequently grapples with the issue of measurement error. Its ubiquity across diverse scientific domains underscores the gravity with which it can potentially skew research outcomes and conclusions. Our systematic review and meta-analysis dove deep into this realm, examining the potential of Hierarchical Bayesian semi-parametric (HBSP) models as an effective tool for addressing this measurement error challenge.

5.1. The Potency of HBSP Models

The most striking revelation from our analysis was the consistent efficacy of HBSP models in significantly reducing bias and variance across studies. On average, HBSP models achieved a commendable 50% reduction in bias and a 25% reduction in variance. This holds promising implications for researchers striving for accuracy in a world ridden with intricate and often non-linear data relationships.

5.2. Applicability Across Fields

HBSP models were not confined to a niche but exhibited their prowess across diverse research terrains, from epidemiology and clinical trials to the social sciences. Such wide applicability bolsters the universality of these models, making them indispensable tools for modern-day scientific investigations.

5.3. Comparative Superiority

When pitted against other prevalent measurement error correction methodologies, HBSP models consistently emerged superior. This comparative edge can be attributed to their intrinsic flexibility, ability to handle non-Gaussian measurement errors, and the dual correction capability for both exposure and outcome variables in regression contexts.

5.4. Challenges and Future Prospects

However, it's imperative to acknowledge the roadblocks hampering the widespread adoption of HBSP models. Their computational complexity, dependence on specialized software, and the necessity for detailed information about measurement error distribution, are significant impediments. Furthermore, the intricate nature of these models can sometimes become a deterrent for non-statisticians, often making their interpretations daunting.

Yet, there's a silver lining. The burgeoning interest in these models, reflected in the escalating number of studies utilizing them, hints at a future where they might become the gold standard for correcting measurement errors.

5.5. Geographical Distribution

Our bibliometric analysis paints an interesting picture of global research dynamics. While the USA leads in the number of studies, countries like South Africa and Kenya show noteworthy contributions, especially considering their relatively limited research resources. This geographical spread indicates the global recognition of the challenges posed by measurement errors and the universal appeal of HBSP models.

5.6. Temporal Trends in Research Impact

Over the last decade, the citation trend of HBSP model studies exhibited a fascinating trajectory. Initial phases of modest growth transitioned into periods of rapid expansion, stabilization, and then revival. Such temporal patterns possibly mirror the evolutionary journey of HBSP models—from inception to widespread acknowledgment.

In conclusion, while HBSP models stand as formidable weapons against the challenges of measurement errors, there's a definitive need for more streamlined tools and training resources. The goal should be to make these models accessible to the wider research community, irrespective of their statistical prowess.

6. Conclusion

The implications of measurement error in scientific research are vast and far-reaching. Across varied scientific domains, from intricate laboratory measurements to large-scale epidemiological studies, the potential distortions introduced by measurement errors have consistently raised concerns about the validity and reproducibility of findings. In this comprehensive systematic review and meta-analysis, the Hierarchical Bayesian semi-parametric (HBSP) models have emerged as a potent countermeasure to these challenges.

The utility and efficacy of HBSP models in reducing both bias and variance were evident. Their adaptability across diverse research domains, from health sciences to social sciences, offers a testament to their robustness and versatility. While the USA remained at the forefront of HBSP research, the notable contributions from countries like South Africa and Kenya accentuate the global resonance of this methodology.

However, it's not without its set of challenges. The complexity associated with HBSP models and the prerequisites for its implementation can be barriers to its more widespread adoption. But, the increasing trajectory of studies and citations over the years signals a growing acceptance and reliance on these models, setting the stage for a future where they might be integral to research protocols aiming for accuracy and precision.

As the scientific community steers towards a more data-driven era, the emphasis on accurate measurements and reliable methodologies will only intensify. In this journey, tools like the HBSP models, which promise to mitigate measurement errors, will undoubtedly play a pivotal role. While we have made significant strides in understanding and implementing these models, the path ahead requires concerted efforts in simplifying their application, fostering global collaborations, and making them accessible to all tiers of the research community.

Abbreviations

I^2	Inconsistency
Q	Heterogeneity
HSP	Hierarchical Bayesian Semi-Parametric
CI	Confidence Interval
SIMEX	Simulation Extrapolation

Author Contributions

Amos Kipkorir Langat: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Writing – original draft, Writing – review & editing

Samuel Musili Mwalili: Data curation, Formal Analysis, Investigation, Methodology, Software, Supervision, Visualization

Lawrence Ndekeleni Kazembe: Data curation, Formal Analysis, Project administration, Software, Supervision, Validation, Visualization, Writing – review & editing

Funding

This research was supported by funding from the Pan African University Institute for Basic Sciences, Technology and Innovation.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Ohlson, J. A. (2022). Researchers' data analysis choices: an excess of false positives?. *Review of Accounting Studies*, 1-19.

[2] Simon, J. (2017). *The art of empirical investigation*. Routledge.

[3] Ohlssen, D. I., Sharples, L. D., & Spiegelhalter, D. J. (2007). Flexible random-effects models using Bayesian semi-parametric models: applications to institutional comparisons. *Statistics in medicine*, *26*(9), 2088-2112.

[4] Gayawan, E., Arogundade, E. D., & Adebayo, S. B. (2014). Possible determinants and spatial patterns of anaemia among young children in Nigeria: a Bayesian semi-parametric modelling. *International health*, *6*(1), 35-45.

[5] Santos, C. C., & Loschi, R. H. (2020). Semi-parametric Bayesian models for heterogeneous degradation data: An application to laser data. *Reliability Engineering & System Safety*, *202*, 107038.

[6] Clark, W., & Harley, A. (2019). Sustainability science: Towards a synthesis. *Sustainability Science Program Working Papers*.

[7] Hagiopol, C., & Leru, P. M. (2024). Scientific Truth in a Post-Truth Era: A Review. *Science & Education*, 1-34.

[8] Frank, A., Gleiser, M., & Thompson, E. (2024). *The blind spot: why science cannot ignore human experience*. MIT Press.

[9] Mayo, D. G. (2018). *Statistical inference as severe testing: How to get beyond the statistics wars*. Cambridge University Press.

[10] Sibony, O. (2020). *You're about to make a terrible mistake!: How biases distort decision-making and what you can do to fight them*. Swift Press.

[11] Ghomashi, M. J. (2019). *Development and identification of hierarchical nonlinear mixed effects models for the analysis of dynamic systems: identification and application of hierarchical nonlinear mixed effects models for the determination of steady-state and dynamic torque responses of an SI engine* (Doctoral dissertation, Loughborough University).

[12] Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement error in survey data. In *Handbook of econometrics* (Vol. 5, pp. 3705-3843). Elsevier.

[13] Mutai, N. (2024). Ethical Decision-Making in Data Analysis: Navigating Challenges and Ensuring Integrity. *Ethics in Statistics: Opportunities and Challenges*, 183.

[14] McElreath, R. (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.

[15] Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: a modern perspective*. Chapman and Hall/CRC.

[16] Brakenhoff, T. B., Mitroiu, M., Keogh, R. H., Moons, K. G., Groenwold, R. H., & van Smeden, M. (2018). Measurement error is often neglected in medical literature: a systematic review. *Journal of clinical epidemiology*, *98*, 89-97.

[17] Hoffmann, S., Rage, E., Laurier, D., Laroche, P., Guihenneuc, C., & Ancelet, S. (2017). Accounting for Berkson and classical measurement error in radon exposure using a Bayesian structural approach in the analysis of lung cancer mortality in the French cohort of uranium miners. *Radiation Research*, *187*(2), 196-209.

[18] Hoffmann, S., Guihenneuc, C., & Ancelet, S. (2018). A cautionary comment on the generation of Berkson error in epidemiological studies. *Radiation and Environmental Biophysics*, *57*(2), 189-193.

[19] Congdon, P. D. (2019). *Bayesian hierarchical models: with applications using R*. Chapman and Hall/CRC.

[20] Gonzalez-Blanks, A., Bridgewater, J. M., & Yates, T. M. (2020). Statistical approaches for highly skewed data: Evaluating relations between maltreatment and young adults' non-suicidal self-injury. *Journal of Clinical Child & Adolescent Psychology*, *49*(2), 147-161.

[21] Rosner, B., Hendrickson, S., & Willett, W. (2015). Optimal allocation of resources in a biomarker setting. *Statistics in medicine*, *34*(2), 297-306.

[22] Lockwood, J. R., & McCaffrey, D. F. (2017). Simulation-extrapolation with latent heteroskedastic error variance. *psychometrika*, *82*, 717-736.

[23] Porter, A. T., Holan, S. H., & Wikle, C. K. (2015). Bayesian semiparametric hierarchical empirical likelihood spatial models. *Journal of Statistical Planning and Inference*, *165*, 78-90.

[24] Yang, M., Dunson, D. B., & Baird, D. (2010). Semiparametric Bayes hierarchical models with mean and variance constraints. *Computational Statistics & Data Analysis*, *54*(9), 2172-2186.

[25] Sarkar, A., Mallick, B. K., Staudenmayer, J., Pati, D., & Carroll, R. J. (2014). Bayesian semiparametric density deconvolution in the presence of conditionally heteroscedastic measurement errors. *Journal of Computational and Graphical Statistics*, *23*(4), 1101-1125.

[26] Langat, A. K., Ofori, M. A., Baranon, M. D., Biftu, D. B., & Mwalili, S. M. (2023). A Bayesian Nonparametric Modeling Approach to Settlement Patterns of Pastoralists Population in Kenya. *Asian Journal of Probability and Statistics*, *25*(2), 17-28.

[27] Mwalili, S. M., Lesaffre, E., & Declerck, D. (2005). A Bayesian ordinal logistic regression model to correct for interobserver measurement error in a geographical oral health study. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *54*(1), 77-93.