Journal of Advanced Research in Natural and Applied Sciences
dergipark.org.tr/en/pub/jarnas
Open Access

ISSN: 2757-5195 | 11 (3) (2025) 224-236 | doi.org/10.28979/jarnas.1671967

New Memory Type Estimators for Systematic Sampling

Eda Gizem Koçyiğit[1]

Article Info
Received: 08 Apr 2025
Accepted: 30 Jun 2025
Published: 30 Sep 2025
Research Article

Abstract − This study aims to enhance estimation accuracy in systematic sampling by proposing a set of novel Exponentially Weighted Moving Average (EWMA)-based memory-type estimators. While memory-type estimators have been explored in other sampling frameworks, they have not yet been adapted to systematic sampling, which is known for its uniform population coverage and greater efficiency compared to simple random sampling. To address this gap, we develop three new estimators: An EWMA-based ratio estimator, an exponential ratio estimator, and a regression estimator. Through comprehensive simulation studies using both synthetic and real-world datasets, we demonstrate that the proposed estimators consistently outperform traditional methods in terms of efficiency. Notably, the ratio and regression-type estimators exhibit superior performance in different distributional settings, particularly when the weight parameter ϑ is set to 0.3 for symmetric distributions. These results offer a practical and robust alternative for survey statisticians and practitioners working with structured populations. The proposed methodology makes both theoretical and empirical contributions to the field of finite population estimation under complex designs.

Keywords − *EWMA, mean, estimation, systematic sampling, simulation*

1. Introduction

Systematic sampling ensures uniform coverage across the entire region for all units, making it generally more efficient than simple random sampling (SRS), as suggested by Madow and Madow [1]. Consequently, researchers and field workers often prefer systematic sampling due to its simplicity and operational convenience [2]. Although many articles have addressed systematic sampling, recent studies have increasingly focused on developing new selection methods and estimators. For example, Pal et al. [3] developed a difference estimator within the framework of systematic sampling. Furthermore, Bello et al. [4] investigated the treatment of missing values using diagonal systematic sampling. In addition, Shahzad et al. [5] proposed a robust regression estimator tailored to systematic sampling, while Azeem et al. [6] introduced a modified sampling design for linearly trending data. Moreover, Pandey and Shukla [7] proposed a clustering method based on stratified linear systematic sampling to identify financial risk groups using big data mining. Azeem [8, 9] also studied proportion estimation within diagonal systematic sampling and developed an alternative sampling scheme. Similarly, Mukherjee and Singh [10] introduced a new sampling design based on the systematic approach.

Shabbir et al. [11] proposed a mean estimator utilizing multiple auxiliary variables. Khan et al. [12] improved estimator efficiency by introducing an optimal pairing strategy within a new systematic sampling

[1]eda.kocyigit@deu.edu.tr (Corresponding Author)
[1]Department of Statistics, Faculty of Science, Dokuz Eylül University, İzmir, Türkiye

scheme. Lastly, Gupta et al. [13] proposed a modified systematic sampling design incorporating random initialization. Research on systematic sampling extends beyond the field of statistics. For instance, Lee et al. [14] applied systematic sampling and linear regression to study cyber-attacks and cybersecurity issues. Similarly, Ansari et al. [15] used systematic sampling to estimate the remaining functional life of lithium-ion batteries.

Memory-type estimators based on Exponentially Weighted Moving Average (EWMA) control chart statistics represent a novel approach to enhance estimator efficiency by incorporating historical data from the same or similar populations [16]. In the existing literature, such estimators have been developed for various sampling methods, including simple random sampling [17], stratified sampling [18–20], two-phase sampling [21], and ranked set sampling [22]. However, no memory-type estimator has yet been proposed for systematic sampling. This study introduces a memory-type estimator for systematic sampling using the EWMA approach, filling a notable gap in the literature and aiming to enhance estimation efficiency in this context. The main original contribution of this paper lies in the development of novel EWMA-based estimators tailored for systematic sampling—a gap previously unaddressed in the literature.

The rest of the paper is organized as follows: Section 2 presents mean estimators for systematic sampling, followed by the recommended method. Section 3 details the simulation studies using both synthetic and real data. Section 4 discusses the simulation results, and the final section outlines potential directions for future research.

2. Mean Estimators for Systematic Sampling

Since systematic sampling may not be familiar to all readers, we briefly recall its basic structure and estimation approach before presenting the proposed methodology. In its simplest and most commonly used form, the standard systematic sampling design selects every k-th unit from a finite population of N units, assuming that the sample size n satisfies $N = nk$ for some integer k. The success of systematic sampling depends on the ordering of the units. If there is information about the y values of the units, they should be ranked accordingly before performing systematic sampling, as this approach can enhance sampling efficiency [2, 23]. From a population of size N, systematic samples of size n can be selected by setting $k = N/n$. In systematic sampling, the simple mean estimator is calculated as shown (2.1), where y_{ij} denotes the j-th unit of the i-th systematic sample. This estimator is unbiased only if $N = nk$.

$$\hat{y}_0 = \frac{\sum_{i=1}^{n} y_{ij}}{n} \tag{2.1}$$

In systematic sampling, the ratio estimator was derived by Swain [24], as shown in (2.2).

$$\hat{y}_1 = \frac{\hat{y}_0}{\hat{x}_0}\bar{X} \tag{2.2}$$

Here, $\hat{x}_0 = \frac{\sum_{i=1}^{n} x_{ij}}{n}$ represents the sample mean obtained through systematic sampling and $\bar{X}$ is the population mean. The exponential estimator for systematic sampling, as proposed by Singh et al. [25], is given in (2.3).

$$\hat{y}_2 = \hat{y}_0\left(\frac{\bar{X} - \hat{x}_0}{\bar{X} + \hat{x}_0}\right) \tag{2.3}$$

2.1. Proposed Memory-Type Estimators

Starting from this section, we introduce our original methodological contributions, which include three newly designed EWMA-based estimators. EWMA statistics included in the proposed estimators are defined for Y and X, as given in (2.4) and (2.5).

$$Ewma_{Y(T)} = \vartheta \hat{y}_0 + (1-\vartheta) Ewma_{Y(T-1)} \tag{2.4}$$

$$Ewma_{Y(T)} = \vartheta \hat{x}_0 + (1-\vartheta) Ewma_{X(T-1)} \tag{2.5}$$

Here, the weight parameter, denoted by ϑ, should take a value between (0, 1]. For systematic sampling, the first estimator, the EWMA-based memory-type ratio estimator inspired by Aslam et al. [22], is given in (2.6); the second estimator, the EWMA-based memory-type exponential ratio estimator inspired by Singh et al. [25], is given in (2.7); and the third estimator, the EWMA-based memory-type regression estimator inspired by Koçyiğit [16], is given in (2.8).

$$\hat{y}_{P1} = \frac{Ewma_{Y(T)}}{Ewma_{X(T)}} \bar{X} \tag{2.6}$$

$$\hat{y}_{P2} = Ewma_{Y(T)} \exp\left(\frac{\bar{X} - Ewma_{X(T)}}{\bar{X} + Ewma_{X(T)}}\right) \tag{2.7}$$

$$\hat{y}_{P3} = Ewma_{Y(T)} + \acute{b}\left(\bar{X} - Ewma_{X(T)}\right) \tag{2.8}$$

In (2.8), $\acute{b}$ is the slope coefficient obtained from the sample.

The proposed estimators incorporate EWMA to leverage information from both the current and past values of the study and auxiliary variables. This memory-type structure allows the estimator to 'remember' prior trends, improving stability and efficiency, particularly when dealing with correlated data. The weight parameter ϑ controls the relative influence of past versus current data points.

3. Simulation Studies

All simulation studies are conducted using the R program. The estimators presented in Section 2 provide estimates made without ranking to examine the effect of ranking in systematic sampling. In contrast, the estimators denoted by, $\hat{y}_\partial,\ \partial = 11, 22, P11, P22, P33$ represent estimates where the population is ranked according to the auxiliary variable X, respectively. After samples are drawn from both synthetically generated populations and the real data set, the estimated values of the estimators are calculated. Following the calculation of mean square errors, the results are expressed as Relative Efficiency (*REff*), as shown in (3.1). The general simulation steps are summarized in Figure 1.

$$REff_\gamma = \frac{MSE(\hat{y}_0)}{MSE(\hat{y}_\gamma)}, \gamma = 1, 2, P1, P2, P3, 11, 22, P11, P22, P33 \tag{3.1}$$

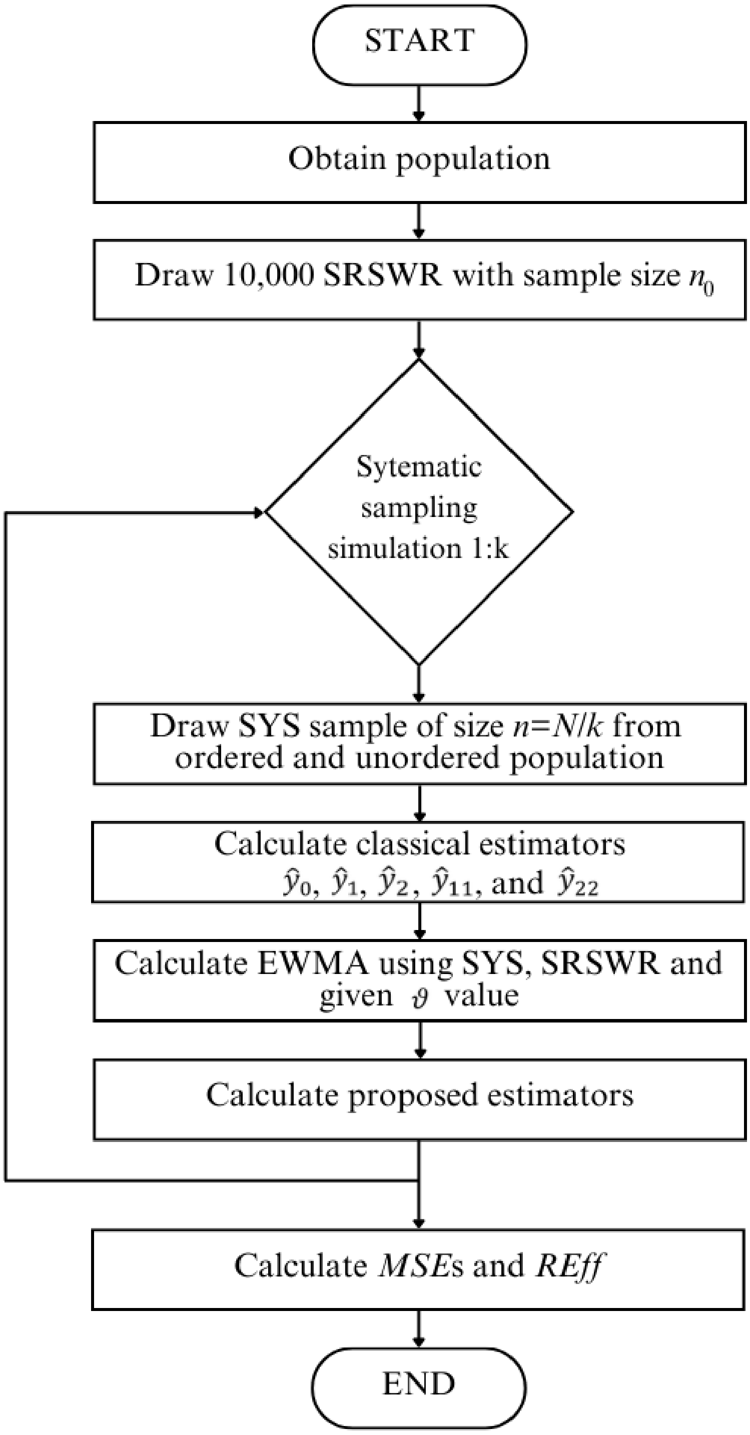

Figure 1. Simulation steps

3.1. Simulation with Synthetic Data

This section takes the bivariate normal distribution derived with random (10,1) parameters as the population, where $N = 1000$ and $\rho = 0.65, 0.75, 0.85, and\ 0.95$. The sample size is set to $n = 10, 20, 25, 50, 100, and\ 200$ to ensure the unbiasedness of the systematic sampling. For the EWMA, $T = 2$, and the sample sizes for obtaining the old statistics are $n_0 = 10\ and\ 50$. The weight parameter ϑ is set to $0.3, 0.5, 0.7, and\ 0.9$. The results are presented in Tables 1-3, ordered according to the correlation coefficients in increasing order.

Table 1. Simulation results for ρ=0.65

	n	10	20	25	50	100	200
	$REff_1$	1.7602	1.7171	1.3950	1.3466	0.5882	0.5278
$\rho=0.65$	$REff_2$	1.8165	1.7245	1.6395	2.2486	1.1090	2.9062
	$REff_{11}$	2.0244	1.7618	1.3794	1.1826	1.0205	1.5133
	$REff_{22}$	1.8927	1.8425	1.4478	1.2787	1.1416	1.3021

ϑ	n_0	10						50					
	n	10	20	25	50	100	200	10	20	25	50	100	200
	$REff_{P1}$	19.5718	19.0309	15.5293	14.9056	6.5061	5.8528	19.5888	19.0336	15.5316	14.8877	6.5246	5.8820
	$REff_{P2}$	20.1928	19.1625	18.2599	**24.9055**	12.3488	**32.2873**	20.2245	19.1509	18.2582	**25.1050**	12.3524	**31.7218**
0.3	$REff_{P3}$	21.0412	19.1536	**18.4119**	24.6654	10.2676	19.5407	21.1044	19.1434	**18.4154**	24.3556	10.2503	19.3661
	$REff_{P11}$	**22.5437**	19.6228	15.2816	13.1312	11.2814	16.4684	**22.5720**	19.6338	15.2824	13.1155	11.3187	16.7348
	$REff_{P22}$	21.0204	20.5182	16.0598	14.1317	**12.6796**	14.4662	21.0543	20.5036	16.0588	14.1906	**12.6821**	14.3473
	$REff_{P33}$	21.6166	**20.6313**	16.0683	13.6622	12.3822	15.0221	21.6354	**20.6019**	16.0589	13.8038	12.2876	15.0792
	$REff_{P1}$	7.0508	6.8645	5.5896	5.3634	2.3492	2.1166	7.0528	6.8593	5.5899	5.3664	2.3510	2.1174
	$REff_{P2}$	7.2765	6.8964	6.5685	9.0206	4.4430	11.4761	7.2780	6.8959	6.5709	9.0274	4.4435	11.6296
0.5	$REff_{P3}$	7.5959	6.8934	6.6337	8.7492	3.6952	7.0285	7.6011	6.8928	6.6405	8.7922	3.6937	7.0456
	$REff_{P11}$	8.1178	7.0693	5.5053	4.7213	4.0783	6.0453	8.1198	7.0641	5.5074	4.7241	4.0811	6.0414
	$REff_{P22}$	7.5765	7.3789	5.7841	5.1079	4.5653	5.1793	7.5781	7.3785	5.7857	5.1110	4.5656	5.2087
	$REff_{P33}$	7.7891	7.4217	5.7761	4.9835	4.4442	5.4178	7.7885	7.4190	5.7799	4.9709	4.4369	5.4444
	$REff_{P1}$	3.5949	3.5015	2.8491	2.7425	1.1991	1.0775	3.5961	3.5023	2.8501	2.7425	1.2000	1.0793
	$REff_{P2}$	3.7108	3.5187	3.3496	4.5993	2.2648	5.9201	3.7109	3.5191	3.3497	4.5994	2.2657	5.9306
0.7	$REff_{P3}$	3.8775	3.5168	3.3874	4.4860	1.8847	3.5981	3.8773	3.5170	3.3864	4.4836	1.8846	3.5954
	$REff_{P11}$	4.1376	3.6005	2.8113	2.4117	2.0813	3.0802	4.1381	3.6013	2.8117	2.4116	2.0825	3.0863
	$REff_{P22}$	3.8647	3.7627	2.9527	2.6085	2.3289	2.6553	3.8648	3.7631	2.9528	2.6085	2.3297	2.6568
	$REff_{P33}$	3.9738	3.7853	2.9494	2.5356	2.2639	2.7824	3.9740	3.7857	2.9494	2.5363	2.2630	2.7814
	$REff_{P1}$	2.1739	2.1191	1.7228	1.6613	0.7260	0.6518	2.1738	2.1194	1.7230	1.6614	0.7259	0.6520
	$REff_{P2}$	2.2434	2.1288	2.0249	2.7781	1.3696	3.5865	2.2434	2.1289	2.0249	2.7782	1.3695	3.5883
0.9	$REff_{P3}$	2.3457	2.1275	2.0491	2.7121	1.1402	2.1779	2.3457	2.1275	2.0492	2.7120	1.1402	2.1774
	$REff_{P11}$	2.5006	2.1759	1.7022	1.4596	1.2598	1.8678	2.5006	2.1762	1.7023	1.4596	1.2597	1.8683
	$REff_{P22}$	2.3371	2.2751	1.7870	1.5784	1.4093	1.6073	2.3371	2.2752	1.7871	1.5784	1.4093	1.6075
	$REff_{P33}$	2.4039	2.2900	1.7842	1.5341	1.3692	1.6838	2.4039	2.2901	1.7840	1.5341	1.3696	1.6835

Bold marked present the best *REff* value.

Table 2. Simulation results for ρ=0.75

	n	**10**	**20**	**25**	**50**	**100**	**200**
	$REff_1$	1.5362	1.4261	2.2395	3.3891	8.7457	3.3935
ρ=0.75	$REff_2$	2.0133	2.2980	2.0104	2.3542	3.9349	2.9132
	$REff_{11}$	1.5326	1.7457	2.1281	1.7016	2.8881	1.7341
	$REff_{22}$	1.4914	1.7754	2.0746	1.7324	3.4460	1.9338

ϑ	n_0	**10**						**50**					
	n	**10**	**20**	**25**	**50**	**100**	**200**	**10**	**20**	**25**	**50**	**100**	**200**
0.3	$REff_{P1}$	17.1882	15.9021	25.2254	37.4644	97.1695	37.5071	17.1905	15.8948	25.2667	37.4380	97.0899	37.2120
	$REff_{P2}$	**22.4067**	**25.5853**	22.3296	26.1754	43.8142	32.3376	**22.4026**	**25.5725**	22.3739	26.1093	43.9354	32.3988
	$REff_{P3}$	18.7838	18.9149	**27.1888**	**41.3532**	**110.5158**	**53.0071**	18.8430	19.0303	**27.2216**	**41.1371**	**111.0763**	**52.5412**
	$REff_{P11}$	17.0911	19.4562	23.6188	18.9165	32.1449	19.2719	17.0990	19.4580	23.6579	18.9031	32.1421	19.1920
	$REff_{P22}$	16.6104	19.7675	23.0015	19.2562	38.2409	21.4666	16.6068	19.7595	23.0472	19.2194	38.3302	21.4917
	$REff_{P33}$	17.7977	20.0293	23.7769	19.4225	35.7239	20.3995	17.8152	20.0282	23.8228	19.3407	35.6539	20.5008
0.5	$REff_{P1}$	6.1781	5.7230	9.0653	13.5060	35.0107	13.5238	6.1774	5.7189	9.0624	13.5055	34.9901	13.5364
	$REff_{P2}$	8.0640	9.2058	8.0574	9.4217	15.7795	11.4585	8.0636	9.2044	8.0554	9.4220	15.7927	11.6584
	$REff_{P3}$	6.7921	6.8223	9.7990	14.8701	40.1948	19.0215	6.7930	6.8536	9.8001	14.8707	40.0197	19.2450
	$REff_{P11}$	6.1502	7.0019	8.5228	6.8081	11.5744	6.9345	6.1492	6.9983	8.5197	6.8078	11.5683	6.9403
	$REff_{P22}$	5.9761	7.1126	8.3036	6.9318	13.7831	7.6474	5.9756	7.1109	8.3018	6.9315	13.7936	7.7365
	$REff_{P33}$	6.4136	7.2104	8.5894	6.9859	12.7967	7.3854	6.4147	7.2171	8.5814	6.9836	12.8330	7.3540
0.7	$REff_{P1}$	3.1449	2.9147	4.6021	6.9025	17.8538	6.9013	3.1454	2.9148	4.6034	6.9005	17.8560	6.9156
	$REff_{P2}$	4.1120	4.6935	4.1074	4.8052	8.0475	5.9312	4.1121	4.6936	4.1076	4.8061	8.0476	5.9485
	$REff_{P3}$	3.4668	3.4986	4.9980	7.5929	20.4634	9.8067	3.4661	3.4981	4.9993	7.5860	20.4576	9.8013
	$REff_{P11}$	3.1333	3.5670	4.3451	3.4741	5.9000	3.5363	3.1336	3.5672	4.3459	3.4728	5.9001	3.5406
	$REff_{P22}$	3.0467	3.6259	4.2351	3.5355	7.0370	3.9402	3.0468	3.6260	4.2354	3.5360	7.0365	3.9475
	$REff_{P33}$	3.2727	3.6823	4.3782	3.5653	6.5389	3.7503	3.2726	3.6820	4.3797	3.5625	6.5409	3.7546
0.9	$REff_{P1}$	1.8989	1.7615	2.7718	4.1809	10.7982	4.1881	1.8990	1.7617	2.7714	4.1811	10.7978	4.1871
	$REff_{P2}$	2.4863	2.8378	2.4830	2.9069	4.8610	3.5973	2.4863	2.8379	2.4829	2.9068	4.8614	3.5972
	$REff_{P3}$	2.0964	2.1153	3.0242	4.5901	12.3629	5.9365	2.0962	2.1145	3.0240	4.5898	12.3683	5.9277
	$REff_{P11}$	1.8934	2.1561	2.6280	2.1008	3.5665	2.1414	1.8934	2.1563	2.6278	2.1008	3.5666	2.1411
	$REff_{P22}$	1.8419	2.1924	2.5616	2.1389	4.2549	2.3877	1.8419	2.1925	2.5615	2.1390	4.2550	2.3875
	$REff_{P33}$	1.9797	2.2272	2.6495	2.1556	3.9589	2.2692	1.9797	2.2271	2.6491	2.1556	3.9581	2.2714

Bold marked present the best *REff* value.

Table 3. Simulation results for ρ=0.85

	n	**10**	**20**	**25**	**50**	**100**	**200**
	$REff_1$	4.8350	4.1849	4.2444	5.5630	4.3901	2.3237
ρ=0.85	$REff_2$	3.0577	2.8581	2.9967	3.6623	4.2975	5.3680
	$REff_{11}$	4.2061	4.3984	3.9697	3.6056	3.4594	1.6245
	$REff_{22}$	3.8579	4.2172	4.3070	4.7745	4.8258	2.0322

ϑ	n_0	**10**						**50**					
	n	**10**	**20**	**25**	**50**	**100**	**200**	**10**	**20**	**25**	**50**	**100**	**200**
0.3	$REff_{P1}$	**54.5269**	47.0400	47.1559	60.6637	48.3976	25.6734	**54.5276**	47.0835	47.2594	61.3442	48.2231	25.7799
	$REff_{P2}$	33.8854	31.7149	33.3157	40.8038	46.9469	**54.0104**	33.8891	31.7061	33.3245	40.8262	47.8328	**59.2784**
	$REff_{P3}$	44.0908	49.8250	**50.4071**	**82.4707**	**64.5548**	46.6157	44.0083	49.9717	**50.5165**	**82.1131**	**63.5735**	47.6490
	$REff_{P11}$	46.3626	48.7324	43.9668	39.7715	38.4314	17.9736	46.3668	48.7738	44.0463	40.0591	38.3090	18.0188
	$REff_{P22}$	42.7034	46.8201	47.8072	53.0115	52.3708	21.7275	42.7068	46.7984	47.8289	53.0659	53.4882	22.5267
	$REff_{P33}$	48.2789	**51.2645**	46.2603	44.1145	43.3400	19.3091	48.2858	**51.4070**	46.2601	44.3654	42.7834	19.3340
0.5	$REff_{P1}$	19.5710	16.8892	17.0137	22.1452	17.3217	9.2250	19.5599	16.8938	16.9987	22.1397	17.4593	9.3032
	$REff_{P2}$	12.2131	11.4107	11.9961	14.6843	17.0649	21.4573	12.2113	11.4233	11.9962	14.6828	17.2106	21.4706
	$REff_{P3}$	15.8981	17.9800	18.2220	29.5982	23.4352	17.0900	15.8455	17.9837	18.2035	29.6356	23.1379	17.1305
	$REff_{P11}$	16.7388	17.5678	15.8709	14.4265	13.7576	6.4628	16.7323	17.5692	15.8600	14.4239	13.8304	6.4986
	$REff_{P22}$	15.3963	16.8279	17.2214	19.1032	19.0735	8.1256	15.3931	16.8611	17.2211	19.1004	19.2685	8.1293
	$REff_{P33}$	17.3866	18.5129	16.6357	15.9689	15.2995	6.9485	17.3849	18.4874	16.6350	15.9642	15.3967	6.9939
0.7	$REff_{P1}$	9.9369	8.5827	8.6757	11.3171	8.9282	4.7433	9.9377	8.5886	8.6740	11.3166	8.9312	4.7388
	$REff_{P2}$	6.2344	5.8302	6.1196	7.4773	8.7862	10.9518	6.2346	5.8297	6.1194	7.4845	8.7861	10.9524
	$REff_{P3}$	8.0801	9.1622	9.2914	15.1099	11.8978	8.7445	8.0821	9.1741	9.2893	15.1271	11.8874	8.7465
	$REff_{P11}$	8.5551	8.9645	8.0997	7.3562	7.0592	3.3150	8.5557	8.9680	8.0986	7.3579	7.0607	3.3131
	$REff_{P22}$	7.8613	8.6029	8.7903	9.7354	9.8484	4.1470	7.8616	8.6039	8.7898	9.7443	9.8486	4.1468
	$REff_{P33}$	8.8697	9.4233	8.4921	8.1444	7.8618	3.5676	8.8698	9.4252	8.4926	8.1434	7.8686	3.5638
0.9	$REff_{P1}$	5.9849	5.1769	5.2421	6.8606	5.4142	2.8683	5.9841	5.1763	5.2423	6.8615	5.4147	2.8694
	$REff_{P2}$	3.7737	3.5281	3.7004	4.5236	5.3088	6.6265	3.7739	3.5279	3.7004	4.5235	5.3088	6.6269
	$REff_{P3}$	4.8919	5.5511	5.6180	9.1489	7.1897	5.2888	4.8899	5.5500	5.6180	9.1474	7.1861	5.2873
	$REff_{P11}$	5.1873	5.4289	4.9000	4.4514	4.2711	2.0051	5.1871	5.4284	4.9002	4.4516	4.2712	2.0056
	$REff_{P22}$	4.7605	5.2062	5.3171	5.8945	5.9578	2.5086	4.7605	5.2059	5.3172	5.8947	5.9580	2.5089
	$REff_{P33}$	5.3657	5.7040	5.1379	4.9264	4.7591	2.1572	5.3657	5.7028	5.1380	4.9270	4.7596	2.1590

Bold marked present the best *REff* value.

Table 4. Simulation results for ρ=0.95

	n	10	20	25	50	100	200
	$REff_1$	7.2851	5.7901	11.8017	6.2171	4.0074	4.6200
ρ=0.95	$REff_2$	3.4452	3.1905	3.2445	3.7547	3.4292	4.3374
	$REff_{11}$	8.1254	6.8871	13.1487	11.8136	12.0575	17.1696
	$REff_{22}$	6.7464	6.0416	11.0305	14.6982	16.6617	34.5826

ϑ	n_0	10						50					
	n	10	20	25	50	100	200	10	20	25	50	100	200
0.3	$REff_{P1}$	81.0449	65.1341	130.3617	69.0056	44.7170	51.6974	81.1177	65.4500	130.7843	68.9975	44.7077	51.5767
	$REff_{P2}$	38.2971	35.5999	36.0674	41.5308	37.9548	47.7635	38.3351	35.6562	36.1005	41.5531	38.0341	48.0584
	$REff_{P3}$	72.0267	62.7918	125.4078	81.3608	50.8220	63.0224	72.2945	63.3223	125.5002	81.4302	51.0263	62.7529
	$REff_{P11}$	**90.4106**	76.3520	145.1784	131.0259	133.9888	190.6803	**90.4702**	76.7572	145.7031	131.0447	133.7823	189.4438
	$REff_{P22}$	74.9315	66.9617	122.0541	**162.7218**	**183.0658**	**364.5445**	75.0515	67.1864	122.3106	**162.9876**	**184.8529**	**383.0451**
	$REff_{P33}$	89.3404	**79.4083**	**147.8935**	138.5470	143.8050	209.2882	89.5572	**79.7727**	**147.8992**	138.4909	143.8818	206.9460
0.5	$REff_{P1}$	29.1992	23.4626	47.0892	24.8604	16.0773	18.5793	29.1830	23.4632	47.1175	24.8510	16.0829	18.5599
	$REff_{P2}$	13.7965	12.8148	12.9923	14.9637	13.6955	17.2793	13.7958	12.8159	12.9924	14.9644	13.6951	17.2905
	$REff_{P3}$	26.0044	22.8927	45.0154	29.4118	18.3613	22.6930	25.9744	22.8561	45.1047	29.2713	18.3780	22.6454
	$REff_{P11}$	32.5610	27.6152	52.4573	47.2264	48.1758	68.6773	32.5483	27.6235	52.4906	47.1967	48.2438	68.4292
	$REff_{P22}$	27.0100	24.1822	44.0611	58.4746	66.5022	135.5907	27.0074	24.1801	44.0673	58.5683	66.4396	136.5979
	$REff_{P33}$	32.2386	28.7203	53.2159	49.8778	51.7989	75.3717	32.2235	28.7368	53.2609	49.8876	51.9393	74.8123
0.7	$REff_{P1}$	14.8852	11.9119	24.0513	12.6815	8.1941	9.4563	14.8844	11.9100	24.0595	12.6810	8.1953	9.4550
	$REff_{P2}$	7.0354	6.5283	6.6264	7.6483	6.9924	8.8399	7.0357	6.5279	6.6261	7.6499	6.9938	8.8426
	$REff_{P3}$	13.2583	11.6639	22.9852	14.9766	9.3721	11.5646	13.2571	11.6598	23.0171	14.9562	9.3717	11.5614
	$REff_{P11}$	16.5984	14.0800	26.7938	24.0921	24.5876	35.0280	16.5982	14.0777	26.8032	24.0922	24.6135	35.0233
	$REff_{P22}$	13.7746	12.3351	22.4922	29.9309	33.9751	70.3407	13.7749	12.3349	22.4953	29.9745	33.9901	70.5507
	$REff_{P33}$	16.4396	14.6609	27.1598	25.4408	26.4481	38.3437	16.4404	14.6585	27.1820	25.4517	26.4899	38.3150
0.9	$REff_{P1}$	8.9990	7.1668	14.5639	7.6753	4.9503	5.7089	8.9982	7.1671	14.5639	7.6744	4.9503	5.7093
	$REff_{P2}$	4.2547	3.9422	4.0063	4.6329	4.2320	5.3530	4.2543	3.9422	4.0064	4.6329	4.2326	5.3530
	$REff_{P3}$	8.0210	7.0521	13.9203	9.0534	5.6634	6.9950	8.0204	7.0527	13.9189	9.0523	5.6646	6.9956
	$REff_{P11}$	10.0360	8.5066	16.2255	14.5829	14.8849	21.1937	10.0349	8.5069	16.2255	14.5821	14.8855	21.1954
	$REff_{P22}$	8.3308	7.4596	13.6139	18.1420	20.5608	42.6879	8.3303	7.4597	13.6142	18.1417	20.5693	42.6905
	$REff_{P33}$	9.9474	8.8669	16.4467	15.4016	16.0117	23.1855	9.9459	8.8671	16.4448	15.4003	16.0164	23.1891

Bold marked present the best *REff* value.

3.2. Simulation with Real Data

The 2022-2023 data from the Turkish Statistical Institute [26] is used as the population. The population parameters are presented in Table 5, where Y represents the number of fatal injury accidents in 2023, and X is regarded as the number of motor vehicles in 2023. Based on the population size, the systematic sample size is determined as $n = 3, 9, and\ 27$, while the SRSWR sample size is set to $n_0 = 3,\ 5\ and\ 10$. For EWMA, $T = 2$, and since the correlation values in the real data are higher than those in the simulation study, $\vartheta = 0.3, 0.5, 0.7, and\ 0.95$ are selected. The correlation coefficient required for the $\acute{b}$ estimate in the regression estimator is chosen as $\rho_{\text{Y2022- X2022}}$. The simulation results are shown in Table 6.

Table 5. Population parameters for real data set

Parameters	Value	Parameters	Value
Min. (Y_{2023})	175	Min. (Y_{2022})	129
Max. (Y_{2023})	25622	Max. (Y_{2022})	22914
Mean (Y_{2023})	2902.111	Mean (Y_{2022})	2435.321
Std. Dev. (Y_{2023})	3888.364	Std. Dev. (Y_{2022})	3425.946
Skewness (Y_{2023})	3.380185	Skewness (Y_{2022})	3.589712
Kurtosis (Y_{2023})	17.36067	Kurtosis (Y_{2022})	18.93664
$\rho_{\text{Y2023- Y2022}}$	0.9986	$\rho_{\text{Y2022- X2023}}$	0.9689
$\rho_{\text{Y2023- X2023}}$	0.9608	$\rho_{\text{Y2022- X2022}}$	0.9692
$\rho_{\text{Y2023- X2022}}$	0.9611		
Parameters	**Value**	**Parameters**	**Value**
Min. (X_{2023})	9470	Min. (X_{2022})	8802
Max. (X_{2023})	5406820	Max. (X_{2022})	4940010
Mean (X_{2023})	354820.9	Mean (X_{2022})	326948.7
Std. Dev. (X_{2023})	700257.7	Std. Dev. (X_{2022})	641133.3
Skewness (X_{2023})	5.280241	Skewness (X_{2022})	5.257002
Kurtosis (X_{2023})	35.91974	Kurtosis (X_{2022})	35.6198
$\rho_{\text{X2023- X2022}}$	0.9999		

Table 6. Simulation results for real data

	n	3			9			27		
	$REff_1$	3.2128			3.8395			2.9698		
	$REff_2$	5.3690			4.6130			6.1812		
	$REff_{11}$	6.4959			4.6420			8.7457		
	$REff_{22}$	6.7544			15.4167			221.9643		
ϑ	n_0	3			5			10		
	n	3	9	27	3	9	27	3	9	27
0.3	$REff_{P1}$	73.5095	24.8468	12.5293	74.2009	23.7025	13.7282	71.9329	24.3906	12.7751
	$REff_{P2}$	32.2278	13.4182	7.5438	32.8431	13.1718	7.4988	32.5316	13.6294	7.4044
	$REff_{P3}$	44.6884	49.6717	17.5646	45.1393	46.5100	18.7660	46.3305	48.8954	17.6394
	$REff_{P11}$	75.8903	23.9631	14.8148	76.6364	22.9360	16.7388	74.1379	23.5592	15.2090
	$REff_{P22}$	34.9392	16.2091	8.5396	35.6483	15.8594	8.4836	35.2833	16.5033	8.3644
	$REff_{P33}$	**77.9665**	44.7127	14.9305	**78.9279**	42.5718	15.8814	**79.6523**	44.4400	14.9699
0.5	$REff_{P1}$	45.0150	19.2724	12.5293	45.0020	19.4062	11.6347	44.6292	19.3148	11.6096
	$REff_{P2}$	20.2986	12.5020	7.5438	20.0332	12.9683	9.9514	20.2907	12.7266	9.8875
	$REff_{P3}$	19.5910	**68.5304**	17.5646	19.4919	**70.6357**	34.7091	19.2923	**69.5229**	34.4437
	$REff_{P11}$	48.4580	17.9451	14.8148	48.4133	18.0806	20.4697	48.1028	18.0132	20.3665
	$REff_{P22}$	23.8411	22.6729	8.5396	23.5042	23.9962	17.0485	23.8359	23.2846	16.8722
	$REff_{P33}$	36.6442	43.9244	14.9305	36.4721	45.7953	35.8325	36.0847	45.4178	35.4401
0.7	$REff_{P1}$	19.9648	11.3566	7.2193	19.8933	11.4361	7.2808	19.5453	11.3718	7.2522
	$REff_{P2}$	11.5141	9.1499	9.7146	11.5083	9.1222	9.6980	11.4943	9.1281	9.7924
	$REff_{P3}$	10.0595	37.6856	30.1054	10.0331	38.0860	30.5758	9.9129	37.7780	30.4127
	$REff_{P11}$	23.0224	10.7573	18.4291	22.9503	10.8232	18.3825	22.5993	10.7695	18.4607
	$REff_{P22}$	14.0307	25.6945	44.0091	14.0246	25.5403	43.6964	14.0114	25.5847	45.3636
	$REff_{P33}$	18.3008	24.1004	84.6790	18.2503	24.3489	84.6231	18.0161	24.1409	86.3716
0.9	$REff_{P1}$	6.9355	5.7145	3.9788	6.7892	5.7905	3.9931	6.8554	5.7482	3.9949
	$REff_{P2}$	6.8382	5.8143	7.4488	6.8333	5.8201	7.4531	6.8408	5.8112	7.4454
	$REff_{P3}$	5.8405	17.4020	14.8170	5.7940	17.5778	14.8736	5.8161	17.4975	14.8974
	$REff_{P11}$	9.9192	6.0857	11.4250	9.7670	6.1465	11.4599	9.8341	6.1128	11.4630
	$REff_{P22}$	8.5049	19.2196	192.6337	8.5015	19.2192	193.7554	8.5081	19.2085	191.2603
	$REff_{P33}$	10.2734	12.2361	57.8197	10.1835	12.3688	58.2028	10.2239	12.2985	58.3644
0.95	$REff_{P1}$	4.8852	4.7779	3.4392	4.9118	4.7062	3.4558	4.8983	4.7535	3.4395
	$REff_{P2}$	6.0501	5.1806	6.7958	6.0531	5.1749	6.8016	6.0514	5.1782	6.7997
	$REff_{P3}$	5.1742	14.8574	12.4527	5.1829	14.7104	12.5015	5.1785	14.8092	12.4421
	$REff_{P11}$	7.9631	5.3570	10.0057	7.9894	5.2997	10.0494	7.9762	5.3375	10.0067
	$REff_{P22}$	7.5661	17.2615	**236.7119**	7.5685	17.2651	**237.6795**	7.5671	17.2625	**237.8079**
	$REff_{P33}$	9.0317	10.6574	45.8580	9.0475	10.5439	46.1411	9.0396	10.6192	45.7693

Bold marked present the best *REff* value.

4. Results and Discussion

Simulation results based on synthetic data show that the proposed EWMA-based memory-type estimators consistently outperform existing methods across all correlation levels, particularly when the parameter ϑ is set to 0.3. This value provides a balanced weight to past and current information, making it especially effective under symmetric distribution settings. At a correlation of 0.75, non-ranked systematic sampling yielded superior performance, whereas at a higher correlation of 0.95, the ranked systematic sampling with auxiliary variable X was more effective. Interestingly, variations in the first-phase sample size n_0 did not have a significant impact on estimator performance.

In the real data analysis, characterized by a high correlation structure, ranked systematic sampling with auxiliary information generally produced more efficient estimates. However, no consistent improvement in REff was observed with increasing values of n_0. While the existing best-performing estimator in the literature $\hat{y}_2$ remains competitive, the proposed EWMA-based estimators surpassed its performance under several scenarios. For instance, when n =3, the regression-type EWMA estimator with ranked systematic sampling and ϑ = 0.3 provided the best results. For n =9, the non-ranked version with ϑ = 0.5 was preferable, and for n =27, the exponential EWMA estimator with ranked sampling and ϑ = 0.95 delivered the highest efficiency.

As a result of the simulation studies, it was observed that the proposed estimators provide more effective results than the estimators in the literature under all conditions. The proposed estimators served as alternatives to each other. However, except for the lowest n values at correlations of 0.65 and 0.85, the $\hat{y}_{P1}$ estimator did not perform well in both synthetic and real data studies. While $\hat{y}_{P1}$ may still be useful in specific contexts, $\hat{y}_{P2}$ and $\hat{y}_{P3}$ should be prioritized for better performance across a wider range of conditions. Based on the synthetic data simulation results, using the proposed estimators with ϑ = 0.3 for symmetric distributions is recommended.

The main limitation of the proposed approach lies in its reliance on the correct selection of the parameter ϑ. Although ϑ = 0.3 is generally effective for symmetric distributions, optimal values may vary for skewed or multimodal distributions. Moreover, the current evaluation is limited to continuous variables; further testing is needed for categorical or binary data.

Future studies may explore adaptive strategies to select ϑ based on data characteristics or investigate the performance of these estimators in more complex sampling schemes (e.g., stratified or cluster sampling). Moreover, extending the method to multivariate settings or incorporating robust estimation techniques could further enhance its practical utility.

Based on our findings, practitioners, particularly those working in official statistics, health surveys, and market research, can benefit from using the proposed EWMA-based estimators for improved accuracy when dealing with structured populations. Decision-makers are encouraged to adopt the regression-type or exponential-type estimators depending on the sample size and correlation level, with careful tuning of the ϑ parameter to align with the data structure.

5. Conclusion

This study makes a novel methodological contribution by extending memory-type estimation to the context of systematic sampling for the first time. This study aimed to address a methodological gap in memory-type estimators specifically designed for systematic sampling by introducing three novel EWMA-based estimators: A ratio-type, an exponential-type, and a regression-type estimator. Systematic sampling is widely used in practice due to its operational simplicity and efficient coverage of the population, yet no prior research has incorporated exponentially weighted memory structures into estimators under this framework. The simulation and empirical results demonstrated that the proposed estimators significantly improve estimation efficiency over traditional alternatives, particularly when the weight parameter is appropriately chosen. The contribution of this work lies in extending memory-type estimation to systematic designs, offering both theoretical innovation and empirical performance gains. These estimators are especially useful in structured sampling scenarios where auxiliary information is partially available, and are well-suited for fields such as official statistics, agricultural surveys, environmental monitoring, and market research, where systematic designs are commonly employed.

Future research may focus on adapting the proposed estimators to more complex designs such as modified or stratified systematic sampling, as well as exploring their robustness under skewed populations, outlier contamination, or when the exact condition $N=nk$ is not satisfied. Overall, the proposed approach offers a practical and flexible tool for improving the precision of finite population mean estimation under systematic sampling schemes.

Author Contributions

The author read and approved the final version of the paper.

Conflict of Interest

The author declares no conflict of interest.

Ethical Review and Approval

No approval from the Board of Ethics is required.

References

[1] W. G. Madow, L. H. Madow, *On the theory of systematic sampling I*, Annals of Mathematical Statistics 15(1) (1944) 1–24.

[2] W.G. Cochran, Sampling Techniques, 3rd Edition, Wiley, 1977.

[3] S. K. Pal, S. A. Mahmud, H. P. Singh, *An efficient estimation of finite population mean through difference estimator in systematic sampling*, Afrika Matematika 36 (1) (2025) 1–20.

[4] A. A. Bello, A. Audu, A. B. Zoromawa, M. M. Hamza, *Imputation methods for missing values in estimation of population mean under diagonal systematic sampling scheme*, Asian Journal of Probability and Statistics 26 (11) (2024) 110–123.

[5] U. Shahzad, I. Ahmad, N. H. Al-Noor, M. Hanif, I. M. Almanjahie, *Robust estimation of the population mean using quantile regression under systematic sampling*, Mathematical Population Studies 30 (3) (2023) 195–207.

[6] M. Azeem, S. Hussain, M. Ijaz, N. Salahuddin, A. Salam, *An improved version of systematic sampling design for use with linear trend data*, Heliyon 9 (6) (2023) e17121.

[7] K. K. Pandey, D. Shukla, *Stratified linear systematic sampling based clustering approach for detection of financial risk group by mining of big data*, International Journal of System Assurance Engineering and Management 13 (3) (2022) 1239–1253.

[8] M. Azeem, *A modified version of diagonal systematic sampling in the presence of linear trend*, Plos One, 17 (3) (2022) e0265179.

[9] M. Azeem, *On estimation of population proportion in diagonal systematic sampling*, Life Cycle Reliability and Safety Engineering 10 (3) (2021) 249–254.

[10] A. Mukherjee, A. Singh, *Optimal systematic sampling when sample size is odd in the presence of linear trend and two-way linear trend*, Sankhya A 83 (2021) 128–142.

[11] J. Shabbir, S. Masood, S. Gupta, *A new improved difference-cum-exponential ratio type estimator in systematic sampling two auxiliary variables*, Journal of the National Science Foundation of Sri Lanka 48 (1) (2020).

[12] Z. Khan, J. Shabbir, S. Gupta, A. Shamim, *An optimal systematic sampling scheme*, Journal of Statistical Computation and Simulation 90 (11) (2020) 2023–2036.

[13] S. Gupta, Z. Khan, J. Shabbir, *Modified systematic sampling with multiple random starts*, REVSTAT-Statistical Journal 16 (2) (2018) 187–212.

[14] J. H. Lee, I. H. Ji, S. H. Jeon, J. T. Seo, *Generating ICS anomaly data reflecting cyber-attack based on systematic sampling and linear regression*, Sensors 23 (24) (2023) 9855.

[15] S. Ansari, A. Ayob, M. H. Lipu, A. Hussain, M. H. M. Saad, *Particle swarm optimized data driven model for remaining useful life prediction of lithium-ion batteries by systematic sampling*, Journal of Energy Storage 56 (2022) 106050.

[16] E. G. Koçyiğit, *Using past sample means in exponential ratio and regression type estimators under a simple random sampling*, Soft Computing 29 (2025) 1389–1406.

[17] S. Bhushan, A. Kumar, A. Alrumayh, H. A. Khogeer, R. Onyango, *Evaluating the performance of memory type logarithmic estimators using simple random sampling*, PloS One 17 (12) (2022) e0278264.

[18] I. Aslam, M. Noor-ul-Amin, U. Yasmeen, M. Hanif, *Memory type ratio and product estimators in stratified sampling*, Journal of Reliability and Statistical Studies 13 (1) (2020) 1–20.

[19] M. U. Tariq, M. N. Qureshi, O. A. Alamri, S. Iftikhar, B. S. Alsaedi, M. Hanif, *Variance estimation using memory type estimators based on EWMA statistic for time scaled surveys in stratified sampling*, Scientific Reports 14 (1) (2024) 26700.

[20] S. K. Yadav, G. K. Vishwakarma, R. Varshney, A. Pal, *Improved memory type product estimator for population mean in stratified random sampling under linear cost function*, SN Computer Science 4 (3) (2023) 235.

[21] N. Shahzad, A. Zaidi, S. Zia, *Memory type estimator of population mean using exponentially weighted moving averages in two-phase sampling*, Journal of Positive School Psychology 6 (10) (2022) 1176–1192.

[22] I. Aslam, M. Noor-ul-Amin, M. Hanif, P. Sharma, *Memory type ratio and product estimators under ranked-based sampling schemes*, Communications in Statistics-Theory and Methods 52 (4) (2023) 1155–1177.

[23] S. L. Lohr, Sampling: Design and Analysis, 2nd Edition, Chapman and Hall/CRC, 2010.

[24] A. K. P. C. Swain, *The use of systematic sampling in ratio estimate*, Journal of the Indian Statistical Association 2 (213) (1964) 160–164.

[25] H. P. Singh, R. Tailor, N. K. Jatwa, *Modified ratio and product estimators for population mean in systematic sampling*, Journal of Modern Applied Statistical Methods 10(2) (2011) 4.

[26] TÜİK-Turkish Statistical Institute, Motor vehicle statistics. https://biruni.tuik.gov.tr/medas/, Accessed 8 Apr 2025.