Improving The Accuracy of Area Sampling Frame Estimators for Agricultural Surveys Using Unequal Clustered Segment Sampling: The Case of Indonesia

H Zikra[1*], W P Buana[1], Y Bimarta[1], N Paramitasari[1]

[1] BPS-Statistics Indonesia, Jakarta, Indonesia

*Corresponding author's email: hazanul.zikra@bps.go.id

Abstract. Accurate rice production data are vital for maintaining national food security and formulating effective agricultural policies. In Indonesia, the Area Sampling Frame (KSA) method has been widely implemented to estimate rice harvest areas using segments of 300 meters×300 meters represented by nine observation points. However, this approach faces limitations, particularly the risk of undercoverage bias when estimating areas across different rice growth stages, especially if the observation points fall outside the target rice-growing regions as population area. To address this issue, the present study introduces the Unequal Clustered Segment Sampling method as an alternative to the traditional KSA approach. The Unequal Clustered Segment Sampling method improves estimation accuracy by refining the sampling frame and excluding non-target segments, spatial points located outside actual rice-growing regions. Through a design-based estimation framework, the proposed method accounts for unequal cluster sizes, allowing a more representative depiction of field conditions. The empirical results demonstrate that the Unequal Clustered Segment Sampling method significantly reduces bias and enhances the precision of rice area estimates compared to the conventional KSA. These findings suggest that incorporating unequal clustered segment sampling designs into KSA-based surveys can yield more reliable and representative estimates, particularly in heterogeneous or fragmented agricultural landscapes.

Keyword: Agricultural Statistics, Area Sampling Frame, Cluster Sampling, Rice Production, Unequal Cluster Size

1.   Introduction

To generate accurate rice production data, the Area Sample Frame (KSA) method has been implemented in Indonesia since 2018 [1]. In estimating the different phases of rice cultivation, Statistics Indonesia (BPS) employs the KSA method by systematically dividing rice fields into sampling units, known as segments, each measuring 300×300 meters. These segments are further subdivided into nine equally sized subsegments (equal clusters), each measuring 100×100 meters [2].

Although the KSA method relies on objective data collection, further refinement is needed to produce more accurate estimates. The use of square segments may lead bias, as the fixed observation points within these segments can sometimes fall outside the actual paddy field boundaries, leading to inefficiencies and potential frame undercoverage bias [3]. While area frame sampling offers greater protection against non-sampling errors, such as missing or overlapping units in the frame, these types of errors can still occur in area frame surveys, including KSA [4].

Area frames are well protected against undercoverage, but this source of bias can appear if the perimeter of the region excludes some agricultural areas (minor islands) or more often if some strata are excluded from sampling because they are supposed to be non-agricultural, but they contain some agriculture or small agriculture land [3]. Based on literature, the application of Unequal Clustered Sampling with weighting has been shown to be significantly more efficient than its unweighted counterpart. Weighted estimators can adjust for biases resulting from variations in cluster size and the correlation between cluster size and the measured characteristics [5]. In this context, incorporating a correction factor into the estimator, by only points that fall within actual paddy fields when estimating rice growth phases including harvested area, can lead to more efficient and accurate estimates. A cluster refers to a group of observation units that either occur naturally or are purposefully constructed and can function as a single sampling unit [6].
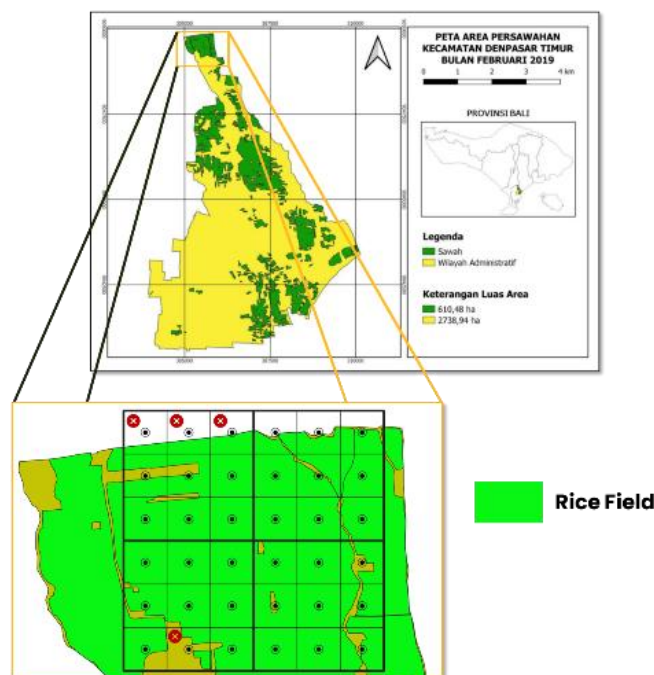

Figure 1. Illustration of Subsegment Areas Excluded from the
Sampling Frame.

The primary objective of this study is to identify an Area Frame Sampling method that offers greater efficiency and is suitable for implementation by BPS. Specifically, the research aims to compare the efficiency of estimators for the proportion of rice growth phases between Equal Clustered Segments and Unequal Clustered Segments.

2.   Research Method

As mention above, the primary objective of this study is to identify an Area Frame Sampling method that offers greater efficiency and is suitable for implementation by BPS. To fulfill the objective, this study compares several estimation methods in surveys that employ the Area Frame Sampling (KSA) approach to estimate paddy harvest area. The methodology involves evaluating the performance of each method through 100,000 iterations of repeated sampling simulations, in which five samples are drawn repeatedly in each iteration. This technique is employed due to computational limitations in generating estimators from all possible samples; therefore, to still obtain the expected estimator value that approximates the parameter value, a sampling iteration technique of 100,000 times is performed. The data used in this study consist of cross-sectional spatial data, specifically a map driven from the delineation of rice paddy land in East Denpasar District, Denpasar City, collected in February 2019. This map, produced as a part of a field practicum by student of Politeknik Statistika STIS, includes information on field boundaries, spatial distribution, geographic area, rice growth phases, and the coordinates of the paddy field locations [7]. Before conducting the sampling simulation, a systematic segment sampling frame must first be constructed. The square segments selected as population segments are those that contain at least fifty percent paddy fields as the target population [8]. After constructing the square segment sampling frame, the next step is to perform the extraction of observation points using QGIS software. This extraction of observation points aims to build a sampling frame that already contains observation points. This sampling frame will subsequently be used for the sampling iteration.
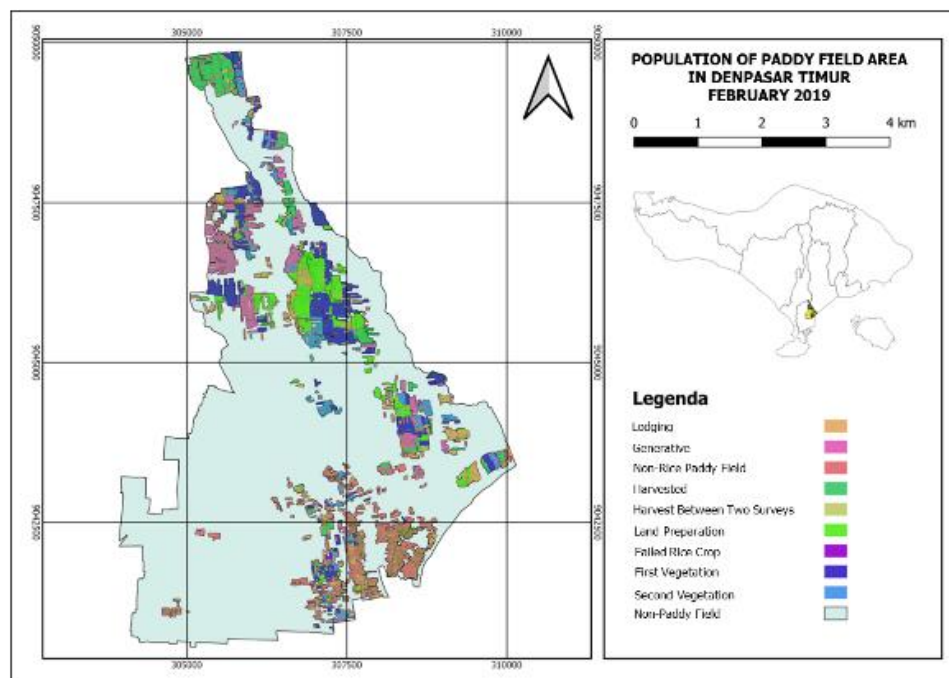

Figure 2. Population Map of Paddy Fields in East Denpasar
District by Rice Growth Phase, February 2019.

Following the simulation, in which five samples were drawn 100,000 times, the next step involves estimating the proportion and area of rice growth stages using four different sampling methods. The specific estimation methods employed for area calculation are presented in Table 1 [5].

Table 1. Estimation Methods Simulated in this Study.

Type of Estimator	Methods Code	Type of Estimator
Equal Clustered Segment Sampling	Method 1	Biased Frame Undercoverage
Unequal Clustered Segment Sampling without Weighting	Method 2	Biased
Unequal Clustered Segment Sampling with Sample Cluster Weighting	Method 3	Consistent Biased from Point of Observation
Unequal Clustered Segment Sampling with Population Cluster Weighting	Method 4	Unbiased from Point of Observation

The formulation for calculating the proportion of rice growth phase i (p_i) and its corresponding between-cluster variance (s_b^2) under Method 1 is as follow [5]:

$$p_i = \frac{1}{n}\sum_{j=1}^{n} p_{ij} \tag{1}$$

$$s_b^2 = \frac{1}{(n\text{-}1)}\sum_{j=1}^{n} (p_{ij} - p_i)^2 \tag{2}$$

Where p_{ij} represents the proportion of phase i in segment j and n is sample size. Meanwhile, the formulation for calculating the proportion of phase i under Method 2 is shown in equation (3). However, the calculation for the between-cluster variance remains the same as in equation (2).

$$p_i = \frac{1}{n}\sum_{j=1}^{n} \frac{M_j}{\overline{M}'} p_{ij} = \frac{\sum_{j=1}^{n} a_{ij}}{\sum_{j=1}^{n} M_j} \tag{3}$$

Where M_j is the number of clusters in sample j and $\overline{M}'$ is the average number of clusters across all samples. For Method 3, the formulation for calculating the proportion of phase i is the same as in equation (3). However, the variance for Method 3 is presented in equation (4):

$$s_b^2 = \frac{1}{\overline{M}'^2 (n\text{-}1)}\sum_{j=1}^{n} M_j^2 (p_{ij} - p_i)^2 \tag{4}$$

ICDSOS
The 3rd International Conference
on Data Science and Official Statistics
November 27 - 28, 2025

For Method 4, the proportion of phase i and its corresponding variance are calculated using Equation (5) and (6), as shown below.

$$p_i = \frac{1}{n} \sum_{j=1}^{n} \frac{M_j}{\overline{M}} p_{ij} \tag{5}$$

$$s_b^2 = \frac{1}{(n-1)} \sum_{j=1}^{n} \left(\frac{M_i p_i}{\overline{M}} - p_i \right)^2 \tag{6}$$

Where $\overline{M}$ represents the average number of clusters (subsegments) in the population. The sampling variance, denoted as $v(p_i)$, for all methods across all rice growth phases is calculated using Equation (7), as shown below.

$$v(p_i) = \frac{(1 - \sfrac{n}{N})}{n} s_b^2 \tag{7}$$

To measure accuracy, the next step is to calculate the bias. Bias is the difference between the expected value of the proportion of the i rice growth stage area and the parameter value of the proportion of the i rice growth stage. Accuracy is measured using the Mean Square Error (MSE) value. A higher MSE value suggests that the estimator is less accurate. Mean Square Error is the sum of the sampling variance and the squared value of the bias for each method.

Subsequently, the relative efficiency (RE) value is utilized to identify the most efficient method [9]. Mathematically, RE is the ratio between the MSE value of one estimator and the MSE of another estimator. An RE value below 100 indicates that an estimator is relatively more efficient than other estimators.

3. Result and Discussion

3.1 Population of Paddy Field in Denpasar Timur on February 2019

In general, paddy field landscapes in Bali exhibit irregular shapes, are terraced, and tend to be dispersed, and often adjacent to non-agricultural land, a pattern also observed in the East Denpasar District. According to the Ministry of Agrarian Affairs and Spatial Planning (Badan Pertanahan Nasional/BPN), the administrative area of East Denpasar District covers approximately 2,231 hectares [10]. In 2019, the reported extent of paddy fields in the district was around 690 hectares [11]. However, based on the delineation conducted during the Field Work Practice by Politeknik Statistika STIS in February 2019, the identified eligible irrigated rice field area amounted to 479.14 hectares [12]. This presents a discrepancy of 30.56 percent when compared to the paddy field base area reported by the Ministry. The delineation results also show that approximately 21.48 percent of East Denpasar's total area consist of rice fields, indicating notable agricultural potential within the district [12]. This is further supported by economic data: based on the Gross Regional Domestic Product (GRDP) at Current Prices of Denpasar City in 2019, the Agriculture, Forestry, and Fisheries sector ranked fifth in contribution, accounting for 6.48 percent of the total GRDP [13].

Table 2. Paddy Field Area by Rice Growth Phase Based on Delineation Results from the Field Work Practice of Politeknik Statistika STIS, Februari 2019 [12].

Phase	Proportion	Area (Hectare)
First Vegetation (V1)	0.21	101.87
Second Vegetation (V2)	0.09	41.06
Generative (G)	0.13	63.39
Harvested (P)	0.09	41.95
Farmland Preparation (PL)	0.13	62.30
Lodging (B)	0.11	51.23
Non-Rice Paddy Field (LL)	0.24	114.11
Harvest Between Two Surveys (P-2)	~0	2.12
Failed Rice Crop (PS)	~0	1.11
Total	**1.00**	**479.14**

Note: *~0: Has a very small value or approaches zero*

The delineation of rice paddy land in February 2019 indicated that the largest proportion of rice cultivation in East Denpasar District was in the non-rice paddy (LL) phase, covering approximately 114.11 hectares, which corresponds to a proportion of 0.24. The harvested area (P) during the same period was around 41.95 hectares, representing roughly 0.09 of the total area. The most extensive standing crop phase was the Early Vegetative (V1) stage, accounting for 101.87 hectares or 21 percent of the paddy fields, followed by the Generative (G) phase with 63.39 hectares (approximately 13 percent). These findings indicate that the majority of rice paddies in East Denpasar were in the early stages of the planting season in February 2019. Conversely, the Potential Crop Failure (PS) phase had the smallest area, covering only 1.11 hectares, representing a proportion close to zero.

In accordance with the established protocols for constructing a square segment sampling frame, a square segment is considered eligible for inclusion in the sampling frame if it contains at least 50 percent of the designated target population [4][14][15]. Applying this criterion to the paddy field population map of East Denpasar resulted in 57 segment units. The proportions and areas corresponding to each rice growth stage, based on observation points falling within the target population (paddy fields), are summarized in Table 2.

Table 3. Derived Proportions and Areas of Rice Growth Phases Based on Square Segment Sampling Frame Development.

Phase	Proportion	Area (Hectare)
First Vegetation (V1)	0.24	116.34
Second Vegetation (V2)	0.06	29.56
Generative (G)	0.14	66.35
Harvested (P)	0.09	44.40
Farmland Preparation (PL)	0.12	55.47
Lodging (B)	0.11	52.05
Non-Rice Paddy Field (LL)	0.24	113.57
Harvest Between Two Surveys (P-2)	~0	1.40
Failed Rice Crop (PS)	0.00	0.00
Total	**1.00**	**479.14**

Note: *~0: Has a very small value or approaches zero*

Table 2 illustrates that the parameters derived from the square segment sampling frame differ noticeably from those obtained through direct delineation of rice growth phases without restricting to the target population. This disparity is primarily attributed to undercoverage bias in the KSA Survey, which arises when the sampling frame does not fully capture the target population. Such bias can occur due to segments that fall partially or entirely outside the target area, observation points located outside paddy fields, or the use of uniform square segments that do not conform to the irregular and fragmented distribution of paddy fields, particularly in East Denpasar [16][7].

Comparison of Estimated Rice Growth Phases from Area Frame Survey (KSA)
Based on the simulation results, the average estimated area for rice growth phases using Method 1 (*equal clustered segment sampling*), the method currently implemented by BPS for calculating harvested area, tends to be lower than the true parameter values across all growth phases. This indicates a consistent pattern of underestimation. One contributing factor is the inclusion of observation points that fall outside the actual paddy field target population, which are nevertheless included in the rice growth phase area calculations. This results in a negative bias in the estimation process [7].

In contrast, the application of Method 2 (*unequal clustered segment sampling without weighting*) and Method 3 (*unequal clustered segment sampling with sample weighting*) produced estimates that tended to underestimate several rice growth phases—specifically LL, P-2, PL, PS, and V2—while overestimating others. This pattern is attributed to the "cancel in–cancel out" effect, in which overestimation in certain phases offsets underestimation in others, since the total of all phase proportions is constrained to sum to one.

By comparison, Method 1 does not exhibit this balancing effect; the total estimated proportion does not sum to one due to the inclusion of observation points outside the paddy field target population, which are excluded from rice growth phase calculations, leading to systematic underestimation. Similarly, Method 4 (*unequal clustered segment sampling with population weighting*) also resulted in overestimation for some phases—PL and V1—and underestimation for others—B, G, LL, P-2, PS, and V2. Notably, however, the estimation for the P (harvested) phase under Method 4 was relatively consistent with the actual observed area.
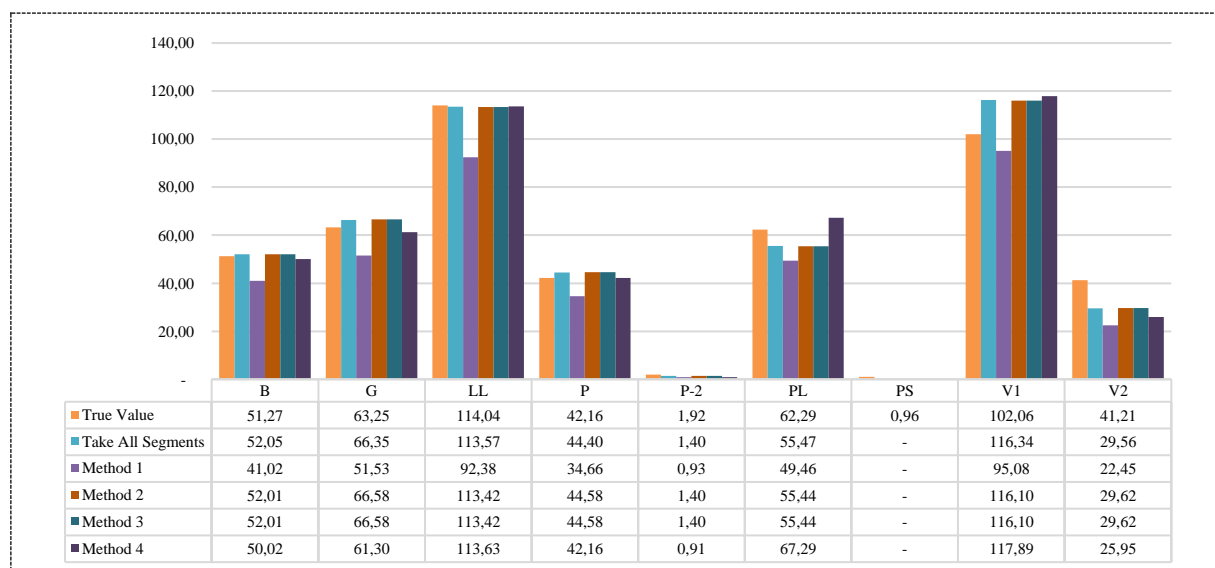

	B	G	LL	P	P-2	PL	PS	V1	V2
■ True Value	51,27	63,25	114,04	42,16	1,92	62,29	0,96	102,06	41,21
■ Take All Segments	52,05	66,35	113,57	44,40	1,40	55,47	-	116,34	29,56
■ Method 1	41,02	51,53	92,38	34,66	0,93	49,46	-	95,08	22,45
■ Method 2	52,01	66,58	113,42	44,58	1,40	55,44	-	116,10	29,62
■ Method 3	52,01	66,58	113,42	44,58	1,40	55,44	-	116,10	29,62
■ Method 4	50,02	61,30	113,63	42,16	0,91	67,29	-	117,89	25,95

Figure 3. Comparison of Estimated Paddy Growth Phase Area by Methods (Hectares).

Simulation sampling results indicate a consistent negative bias in the estimation of rice growth phase area when using Method 1 (equal clustered segment sampling) across all phases. Generally, errors in parameter estimation may arise from two main sources: sampling error and non-sampling error. Sampling error occurs as a natural consequences of selecting a subset (sample) from the population, while non-sampling error arises from factors outside the sampling process. Bias due to non-sampling error can broadly come from specification error, frame error, nonresponse error, measurement error, and processing error [9].

Assuming no errors in observed phases, the shape of rice field polygons, or identification by field officers, the underestimation in Method 1 is predominantly caused by frame undercoverage bias, which is an error in the construction of the sample frame. Several sources of frame undercoverage bias that arise when using Method 1, currently employed by Statistics Indonesia-BPS, are as follows: 1) Points falling outside the rice fields (target population), which leads to the estimated area of rice growth phases tending to be underestimated. This occurs because the calculation of rice growth phase proportions still includes these points due to the use of an equal cluster design, 2) The construction of the segment sample frame always places the sub-segment observation points in the center, making it impossible for them to fall within other target populations within the sample sub-segment, and 3) The segment sample frame not covering the entire target population because, to qualify as a segment sample frame, it must contain at least 50 percent rice fields [15][8].

Similar to Method 1, Method 2 (unequal clustered segment sampling without weighting) also produces biased estimations. However, the estimator generated by Method 2 does not include points that fall outside the target population (rice fields). This means the number of sub-segments used in calculating the area of observed phases can vary for each segment. Assuming sub-segments as cluster elements, the segment sample represents a form of unequal cluster size. This is because the number of eligible points within each segment can differ, making the total number of clusters dependent on the number of points that fall within the target population. Compared to Method 1, Method 2 eliminates the bias stemming from points falling outside the target population. The primary sources of bias in Method 2 are: immovable observation points, where observation points are always fixed at the center of the sub-segment (cluster) and static sample frame, the sample frame itself is fixed and does not change.

Meanwhile, Method 3 (unequal clustered segment sampling with sample weighting) produces parameter estimations that are consistently biased. Similar to Method 2, Method 3 estimates rice growth phases by counting points that fall within the target population. However, Method 3 applies a correction factor for the number of elements within each cluster from sample. This means the resulting estimator remains biased, but its bias decreases as the sample size increases, causing the estimator to approach the "take all" value of the observation points from the segment population [17][18]. The sources of bias in Method 3 are immovable observation points, where observation points are always fixed at the center of the sub-segment (cluster), and a static sample frame, meaning the sample frame itself is fixed and does not change. Consequently, population bias cannot be entirely eliminated. However, the estimator will produce values that approximate the proportion of rice growth phases from all observation points within the segments when samples increased [4].

Method 4 (unequal clustered segment sampling with population weighting) also applies principles similar to Method 3. It excludes points that fall outside the target population, leading to varying sub-segment (cluster) sizes across different segments. Unlike Method 3, Method 4 incorporates a cluster size weight within the population. Estimating the average proportion of rice growth area using Method 4

yields an unbiased estimate of the "take all" proportion of observation points in the segment population, but it requires information on the average number of clusters from the segment population [5][6].

3.2   Comparison of Biased of Estimation Area Growth Phase from Area Frame Survey (KSA)
The average bias across all sampling simulations indicates that Method 1 (equal clustered segment sampling) generally results in underestimation for all phases [15]. The most significant bias was observed in the LL phase, reaching -21.65 hectares, followed by the V2 phase, which experienced an undercoverage bias of -18.75 hectares. The absolute bias for Method 1 is 10.18 hectares. As discussed earlier, Method 1 indeed produces an estimator for rice growth phase area that is biased undercoverage.

Conversely, using Method 2 (unequal clustered segment sampling without weighting) and Method 3 (unequal clustered segment sampling with sample weighting) resulted in generally smaller biases. The most significant undercoverage bias for both Method 2 and Method 3 was observed in the V2 phase, at -11.59 hectares. Meanwhile, the highest over coverage bias for both methods occurred in the V1 phase, with a bias level reaching 14.04 hectares. The biases produced by Method 2 and Method 3 are identical because their distinction lies solely in the calculation of sampling variance and Mean Squared Error (MSE). However, the estimation of rice growth phase area is relatively the same for both methods, as they only consider observation points that fall within the target population, i.e., paddy fields. The absolute bias for Method 2 and Method 3 are 4.56 hectares.

The average absolute bias for Method 4 is 4.63 hectares. The highest estimator bias was observed in the V1 phase, with a value of 15.83 hectares, while the deepest bias was in the V2 area estimator, at -15.26 hectares. Theoretically, Method 4's application should yield an unbiased estimator for the "take all" observation points. However, any resulting bias is potentially due to undercoverage of the estimator when compared to the total area of all rice field polygons. The bias observed across different observation phases also exhibits a "cancel in, cancel out" effect. This happens because the observation points considered in calculating the area of rice growth phases only include those that fall within the rice fields [18].
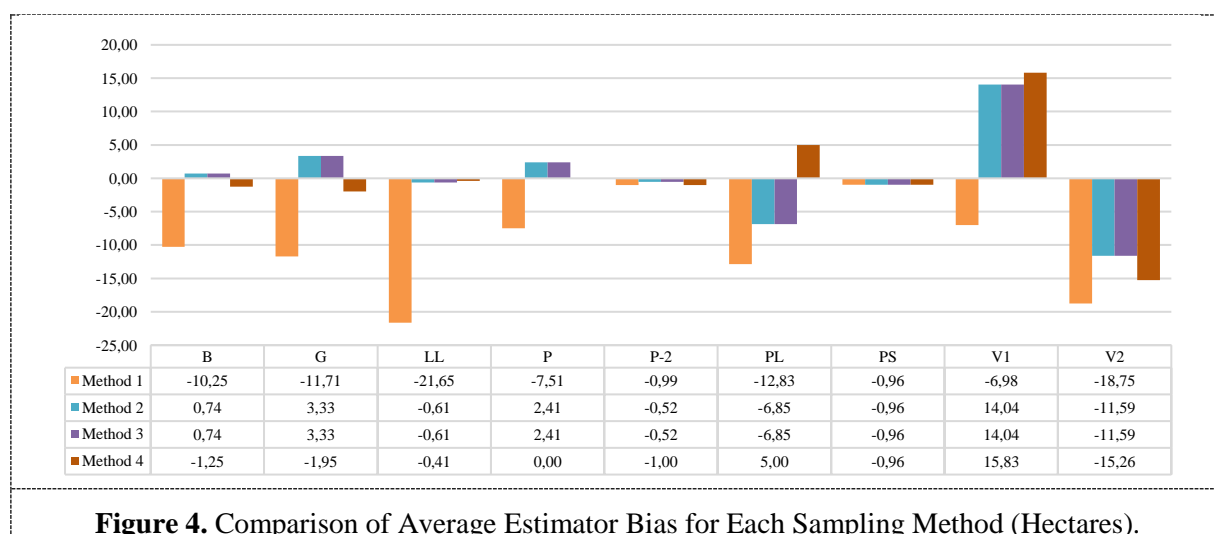

	B	G	LL	P	P-2	PL	PS	V1	V2
Method 1	-10,25	-11,71	-21,65	-7,51	-0,99	-12,83	-0,96	-6,98	-18,75
Method 2	0,74	3,33	-0,61	2,41	-0,52	-6,85	-0,96	14,04	-11,59
Method 3	0,74	3,33	-0,61	2,41	-0,52	-6,85	-0,96	14,04	-11,59
Method 4	-1,25	-1,95	-0,41	0,00	-1,00	5,00	-0,96	15,83	-15,26

Figure 4. Comparison of Average Estimator Bias for Each Sampling Method (Hectares).

Comparison of Mean Square Error (MSE) Growth Phase from Area Frame Survey (KSA)

To compare the accuracy across different sampling designs, the Mean Squared Error (MSE) is utilized. MSE measures the difference between an estimator and the parameter being estimated. Mathematically, MSE is calculated using the following formula [9]:

$$MSE=Bias^2+Variance \tag{8}$$

Based on that formulation, the MSE of all sampling simulations is shown in the following figure.

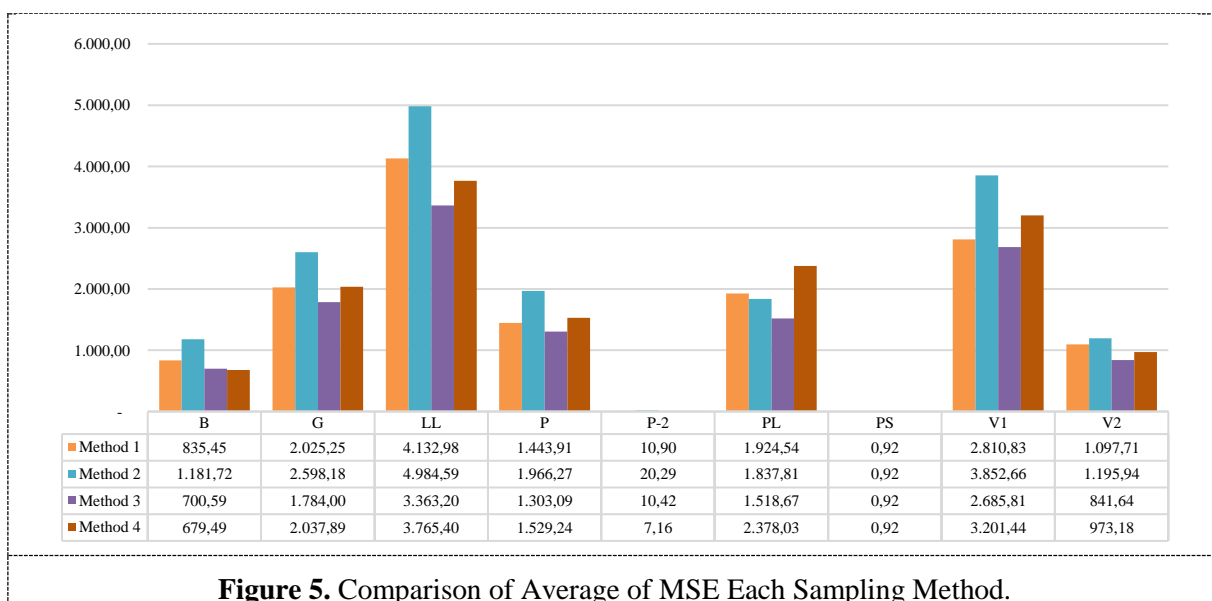|	B	G	LL	P	P-2	PL	PS	V1	V2
Method 1	835,45	2.025,25	4.132,98	1.443,91	10,90	1.924,54	0,92	2.810,83	1.097,71
Method 2	1.181,72	2.598,18	4.984,59	1.966,27	20,29	1.837,81	0,92	3.852,66	1.195,94
Method 3	700,59	1.784,00	3.363,20	1.303,09	10,42	1.518,67	0,92	2.685,81	841,64
Method 4	679,49	2.037,89	3.765,40	1.529,24	7,16	2.378,03	0,92	3.201,44	973,18

Figure 5. Comparison of Average of MSE Each Sampling Method.

Based on the MSE calculations for each method at every rice growth stage, it is evident that applying Method 2 results in a high MSE. This is triggered by the variability in the proportion of rice growth stages, which stems from applying an unequal clustered size to each segment without a correction factor for the number of cluster elements (subsegments).

Meanwhile, Method 3 yields an average MSE value that is lower than the other methods. This is because variability leads to a lower sampling variance due to a correction factor for the differences in the number of clusters in each segment, thus resulting in a lower sampling variance.

The application of Method 4 yields a relatively higher MSE compared to the other methods, as the estimated proportions generated are relatively diverse even though a weighting factor for cluster elements within the population has been applied. Theoretically, the variance of an average estimator with population weights is often greater than the variance with sample weights, especially when the cluster sizes vary significantly and there is a correlation between cluster size and the average cluster value. [5]. For Method 1, the elevated Mean Squared Error (MSE) is attributed not only to sampling variance but also to a consistent underestimation of the parameter's true value. While this leads to a higher relative MSE, it remains considerably lower than the MSE observed in Method 2, which is a result of employing an equal cluster size design.

3.3 Relative Efficiency

To determine the efficiency of the estimation results for each sampling method relative to Method 1 or the Statistics Indonesia-BPS KSA method, relative efficiency (RE) is calculated by comparing the Mean Squared Error (MSE) of each KSA sampling method with the MSE of the Method 1.

The optimal sampling option is determined by the lowest average relative efficiency (RE). Based on the simulation results, the average relative efficiency across all rice growth phases using Method 2 (unequal clustered segment sampling without weighting) is greater than 100. This indicates that the estimator is not more efficient compared to the current Statistics Indonesia-BPS KSA method. The average RE for Method 2 is 128.25.

Table 4. Average Relative Efficiency Across All Phases.

Methods	Average of RE
Method 2	128.25
Method 3	87.82
Method 4	96.75

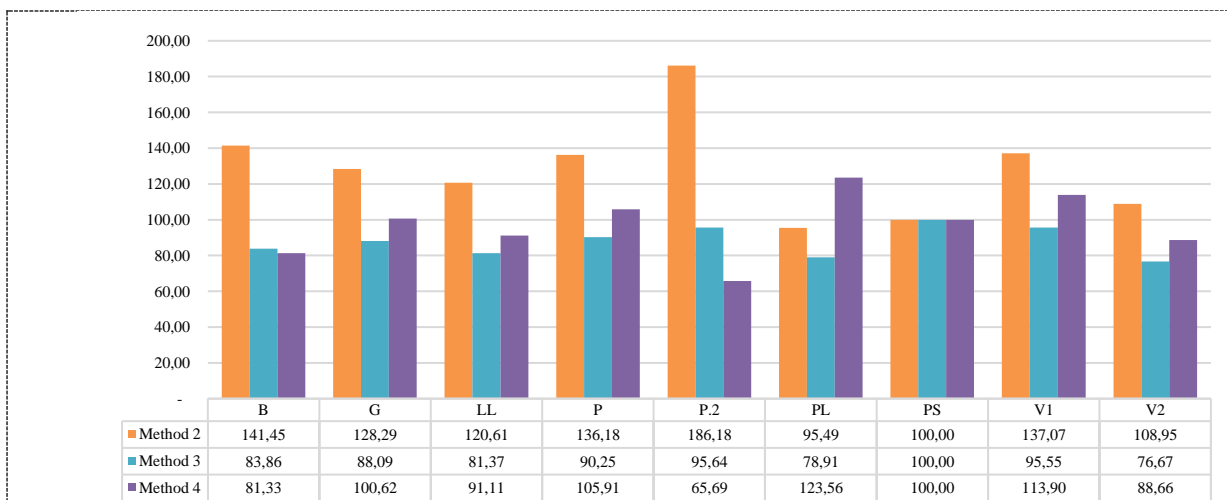|	B	G	LL	P	P.2	PL	PS	V1	V2
Method 2	141,45	128,29	120,61	136,18	186,18	95,49	100,00	137,07	108,95
Method 3	83,86	88,09	81,37	90,25	95,64	78,91	100,00	95,55	76,67
Method 4	81,33	100,62	91,11	105,91	65,69	123,56	100,00	113,90	88,66

Figure 6. Relative Efficiency of Each Sampling Method to Method 1.

In contrast to Method 2 (unequal clustered segment sampling without weighting), Method 3 (unequal clustered segment sampling with sample weighting) yielded an average Relative Efficiency (RE) of 87.82. Generally, the RE values for all phases using Method 3 were less than 100. This indicates that, overall, Method 3 offers significantly better efficiency compared to the current KSA method employed by Statistics Indonesia-BPS.

Similar to Method 3, Method 4 (unequal clustered segment sampling with population weighting) also offers improved estimator efficiency compared to Method 1 (equal clustered segment sampling). This is evident from its average Relative Efficiency of less than 100, and it is generally even lower than Method 3, with an overall average of 96.75.

When considering the potential application of methods for estimating rice growth phase area within Statistics Indonesia-BPS, Method 3 (unequal clustered segment sampling with sample weighting) is more suitable. This is because it only requires weighting information from the sample itself [19].

Although Method 4 (unequal clustered segment sampling with population weighting) offers better efficiency, its implementation for estimating harvest area is less practical due to the general unavailability of population segment information. To effectively apply Method 4, the exact number of observation points in the population must be known to achieve a higher level of estimator precision.

4. Conclusion

Based on the simulation results, the application of Method 3 (unequal clustered segment sampling with sample weighting) offers a relatively more efficient estimation of rice growth phase areas compared to the current KSA Method used by BPS. This is supported by a Relative efficiency (RE) value of less than 100, indicating improved precision over the traditional method.

Moreover, the use of unequal clustered segment sampling with sample weighting is particularly effective in reducing frame undercoverage bias, a key limitation observed in KSA survey estimations. The incorporation of a correction factor based on cluster size allows this method to adjust for uneven distributions of observation points and is practically feasible for implementation by Statistics Indonesia (BPS), especially given the limited availability of complete segment population data across all regions.Scientifically, this approach provides estimates with a lower error rate than equal cluster sampling, making it a viable alternative for improving the accuracy of rice area statistics.

However, despite the rigorous simulation design, this research is subject to several limitations. The data utilized are specific to East Denpasar, Bali, and may not reflect the characteristics of rice cultivation in other regions or in non-paddy field environments. Therefore, further validation through expanded simulations and field testing in diverse rice-producing regions is recommended to ensure the method's generalizability and robustness.

References

[1] Badan Pusat Statistik, *Pedoman Teknis Pendataan Statistik Pertanian Tanaman Pangan Terintegrasi dengan Metode Kerangka Sampel Area.* Jakarta: BPS, 2018.

[2] Badan Pusat Statistik, *Upaya Perbaikan Data Padi dengan Metode Kerangka Sampel Area (KSA)*. Jakarta: BPS, 2018.

[3] J. Gallego, *Area Sampling Frames for Agricultural and Environmental Surveys*. EUR 27595 EN, doi:10.2788/88253 2015.

[4] F. J. Gallego and J. Delincé, *The European Land Use and Cover Area-frame Statistical Survey*. New York: Wiley, 2010.

[5] W. G. . Cochran, *Sampling Techniques.* Wiley, 2005.

[6] A. Asra and A. Prasetyo, *Pengambilan Sampel dalam Penelitian Survei*. Jakarta: PT RajaGrafindo Persada, 2015.

[7] H. Zikra and W. P. Buana, "Analisis perbandingan desain sampling Survei Kerangka Sampel Area (KSA)," in Seminar Nasional Official Statistics, Jakarta: Politeknik Statistika STIS, Sep. 2020, pp. 1326–1336.

[8] F. Javier, G. Pinilla, and E. Commission, *Sampling Frames of Square*, Report EUR 16317, Office for Publications of the E.C., Luxembourg 1995.

[9] P. P. Biemer and L. E. Lyberg, "Introduction to Survey Quality," 2003. [Online]. Available: www.copyright.com.

[10] BPS Kota Denpasar, *Kota Denpasar Dalam Angka Denpasar 2020*. Denpasar: BPS Kota Denpasar, 2020.

[11] BPS Kota Denpasar, *Kecamatan Denpasar Timur dalam Angka 2020*. BPS Kota Denpasar, 2020.

[12] PKL STIS 58, *Studi Akurasi KSA dengan Pendekatan Deliniasi*. Jakarta: Politeknik Statistika STIS, 2019.

[13] Badan Pusat Statistik Kota Denpasar, *Produk Domestik Regional Bruto Kota Denpasar Menurut Lapangan Usaha 2018-2022*. 2025.

[14] Mubekti and L. Sumargana, "Pendekatan kerangka sampel area untuk estimasi dan peramalan produksi padi," *Pangan*, vol. 25, no. 2, pp. 71–146, 2016.

[15] F. J. . Gallego, *Sampling Frames of Square Segments : An Agricultural Information System for the European Union.* OPOCE : European Commission, 1995.

[16] H. Zikra and W. P. Buana, "Penggunaan systematic point sample sebagai area master frame dalam mengestimasi luas panen padi (Studi simulasi sampling di Kecamatan Denpasar Timur Tahun 2019)," J. *Stat. dan Apl.*, vol. 6, no. 1, 2022.

[17] S. L. Lohr, *Sampling: Design and Analysis*. Brooks/Cole, 2010.

[18] L. Kish, *Survey Sampling*. New York: Wiley-Interscience, 1965.

[19]   C. Ferraz, J Delincé. and J Gallego, *Technical Report Series GO-39-2017 Agricultural Master Sampling Frames in Practice Lessons Learned from International Field Experiments and Case Studies,*" 2018, doi: 10.13140/RG.2.2.31374.20808.

ICDSOS
The 3ʳᵈ International Conference
on Data Science and Official Statistics
November 27 - 28, 2025