https://doi.org/10.56093/JISAS.V79I3.5
Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 79(3) 2025 247–253

Estimation of Crop Area at District Level from Reduced Sample Size - A Case Study in the State of Uttar Pradesh

Kaustav Aditya, Neha Purwar, Bharti and Safeela Nasrin

ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 06 November 2025; Revised 08 December 2025; Accepted 19 December 2025

SUMMARY

Agriculture plays a vital role in the Indian economy, making the collection and management of agricultural statistics highly significant. Reliable statistics are essential for sustaining growth in the agricultural sector. Area and production estimates form the two primary components of agricultural statistics. Traditionally, area estimates are generated through village listing using a household-based approach followed by complete enumeration. Under both the Timely Reporting Scheme (TRS) and the Establishment of an Agency for Reporting Area Statistics (EARAS), 20% of villages in each state are selected annually for full enumeration. However, due to limited manpower and an overwhelming workload, data quality has been declining, and the expected outcomes are not being fully realized. Moreover, enumerating such a large proportion of villages is time-consuming, making it difficult to produce timely data for policy formulation.

To address these challenges, ICAR-IASRI, New Delhi, with support from the Directorate of Economics and Statistics (DES), Ministry of Agriculture and Farmers' Welfare (MoA & FW), Government of India, conducted a pilot study to generate district-level area estimates for major food grain crops using a reduced village sample. In this approach, villages are not fully enumerated; instead, crop area is estimated from a carefully selected sample, serving as an alternative to the patwari-based system used in TRS. This paper presents the findings of the pilot study conducted in Uttar Pradesh for two major crops—rice and wheat—across the Kharif and Rabi seasons of the 2015–16 agricultural year.

Keywords: Sample size; Agricultural statistics; Area enumeration; Household approach: Pilot survey.

1. INTRODUCTION

Surveys play a crucial role in policy formulation in developing nations, particularly where agriculture remains a central pillar of the economy. In India, for instance, more than half of rural households rely on farming as their primary source of livelihood. Given this dominant role of the agricultural sector, the systematic collection and management of agricultural statistics becomes essential. India's agricultural roots trace back to the Indus Valley Civilization, and although the sector's contribution to national Gross Domestic Product (GDP) has gradually declined with economic diversification, agriculture continues to support the country's socio-economic structure more broadly than any other sector. Since independence, India has made remarkable advancements in achieving food security. While the population has grown threefold, food-grain production has increased more than four times, resulting in higher per-capita availability of grains. Over the decades, the agricultural economy has undergone significant structural transformation: agriculture's share in GDP dropped from about 43% in 1970 to around 13% in 2016. Yet, according to the Economic Survey 2020–21, agriculture emerged as the only sector to record growth during 2020–21, expanding by 3.4% at constant prices while other sectors contracted. Consequently, the sector's share in GDP rose from 17.8% in 2019–20 to 19.9% in 2020–21its highest level since 2003–04. Reliable information on different aspects of agriculture is crucial for effective planning and decision-making. Agricultural statistics provide this essential data, making accurate estimates of area,

Corresponding author: Kaustav Aditya
E-mail address: katu4493@gmail.com

yield, and production often referred to collectively as crop statistics fundamental for shaping strategies and policies in the sector.

Before Independence, when land taxes formed the main source of government income, agricultural statistics in India were largely compiled by village-level officials, except in regions governed by the permanent settlement system. Provincial revenue departments had established standardized formats for recording land use and cropping patterns, supported by periodic inspections from higher level officials to ensure accuracy and completeness.

After Independence, as national policy shifted toward accelerating agricultural growth, there was an increased demand for timely and reliable data on agricultural trends. Over the years, the system has experienced a gradual decline, primarily because state revenue departments placed diminishing emphasis on the accurate compilation of village-level data and on adhering to the established supervision and inspection procedures, as their functions expanded in scope and complexity. The persistent shortcomings in the mechanisms for collecting and verifying data at the village level, coupled with insufficient corrective action by state governments, significantly weakened the system's ability to generate comprehensive, reliable, and timely estimates of crop areas and yields. This deterioration eventually prompted the adoption of sample survey methods to estimate land use and cropping statistics at state and national levels. The current framework Timely Reporting Scheme(TRS)/ Establishment of an Agency for Reporting of Agricultural Statistics (EARAS) was introduced in the 1970s and involves area enumeration through a 20% rotating sample of villages each year. The TRS, first launched in Maharashtra and Uttar Pradesh in 1968-69, was later extended to all temporarily settled regions with central financial support. Its main aim was to reduce delays in producing dependable estimates of sown area for major crops and to create a sampling frame for Crop Cutting Experiments (CCEs) under the General Crop Estimation Surveys (GCES).

Under the scheme, village patwaris are expected to conduct complete enumeration of crop areas in the selected 20% of villages and submit crop area summaries within prescribed timelines, enabling the preparation of advance seasonal estimates. Villages are selected randomly so that each village is covered once every five years. While TRS supports advance estimation each season, village officials continue to compile land use and irrigated/unirrigated area records in non-TRS villages. These data, although finalized later, are ultimately incorporated into state-published district-level land use and crop statistics.

For the permanently settled states of Kerala, Odisha, and West Bengal, a separate programme known as the EARAS was launched in the early 1970s and later extended to the northeastern states. Under this scheme, information on land use and cropping patterns is obtained through sample surveys conducted by a designated cadre of field personnel working under the state statistical authority responsible for agricultural statistics, with all expenses borne by the Central Government.

In recent decades, however, the introduction of numerous schemes over the last fifty years has significantly increased the volume of work, while the number of staff assigned to these tasks has steadily declined, adversely affecting the quality of the data. Additionally, manpower shortages across various government departments have made it increasingly difficult to conduct these surveys with the same efficiency as in earlier years.

To address these challenges, a pilot study was undertaken by ICAR–IASRI, New Delhi, with support from the Directorate of Economics and Statistics (DES), Ministry of Agriculture and Farmers' Welfare (MoA&FW), Government of India. The objective was to develop a methodology for generating district-level estimates of major crop areas using a substantially reduced sample of villages specifically, by enumerating only 2% of villages in each state annually for estimating the area under key crops at the district and state levels. In these sampled villages, cluster sampling would replace complete enumeration, thereby reducing both workload and time requirements while gradually enhancing data quality to produce more reliable estimates of cropped area.

Another significant development has been the introduction of remote sensing technologies for estimating the area under major crops and forecasting production by integrating satellite data with weather information. Launched in 1988, this initiative has expanded over time to generate multiple in-season national-level forecasts for key crops. The programme, known as Forecasting Agricultural output using Space, Agro-meteorology and Land based observations, i.e. FASAL) is being implemented by Space Application Centre (SAC), and is fully funded by the Ministry of Agriculture and Farmers' Welfare (MoA&FW), Government of India. The area estimates are derived from multi-temporal satellite imagery, which has steadily improved in resolution and analytical capability.

The MoA&FW is understood to incorporate these remote sensing–based outputs, together with data from the Timely Reporting Scheme (TRS) and various empirical yield models, in formulating advance estimates of crop area and production. Nonetheless, a major limitation is that these estimates depend on model fitting, and their accuracy is influenced by the stochastic behaviour of model parameters. Moreover, such models rely on several underlying assumptions that may not consistently reflect real-world conditions. For this reason, the Food and Agriculture Organization (FAO), Rome, has identified Crop Cutting Experiments for yield estimation and the rope-and-compass method for area estimation as the "gold standard" techniques (Sud *et al.,* 2017).

India's agricultural statistics system traditionally depends on village patwaris to record plot-level information on land use, crop-wise area, and crop yields, the latter being estimated through Crop Cutting Experiments (CCEs) conducted in statistically selected villages and fields. This system has remained largely unchanged for several decades. Consequently, the process of data collection, compilation, and dissemination is time-consuming, and the results often undergo multiple rounds of revision. Moreover, the existing framework has not been able to adequately improve the reliability, accuracy, or timeliness of agricultural statistics.

The primary aim of the present study is to generate rapid district- and state-level estimates of crop area and yield using a reduced sample of villages, while simultaneously reducing the reliance on the patwari-based system of area data collection. The focus was on developing an alternative methodology for estimating crop area at broader administrative levels without involving patwaris in the data collection process.

The findings presented in this study are based on data collected through the proposed methodology in the state of Uttar Pradesh for two major crops, rice and wheat by the Department of Agriculture and Crop Insurance, Lucknow, Govt. of Uttar Pradesh. The surveys were conducted during Agricultural year 2015-16 and all the data used was received from the respective states during 2017 after completion of data entry operation at the state level. Subsequent analysis was undertaken at ICAR–IASRI, New Delhi, where the results were generated.

2. SAMPLE SIZE AND SAMPLING METHODOLOGY

In this study, a survey was carried out in the state of Uttar Pradesh during the 2015–16 agricultural year, covering both the Kharif and Rabi seasons. The proposed number of villages included for area enumeration in the survey is given in Table 1.

From Table 1, it is evident that we are considering only about 10% of the villages in Uttar Pradesh compared to the sample chosen under the TRS scheme and this represents just 1.9% of all villages in the state. Since TRS selects a 20% rotating WOR sample each year, the proposed sample size corresponds to around 2% of all villages in Uttar Pradesh.

Table 1. Total Number of Villages and Khasra Numbers to be enumerated proposed under pilot survey in Uttar Pradesh

State	Total Number of villages selected under TRS	Number of villages selected in the pilot survey	Survey/Khasra Numbers enumerated for AE
Uttar Pradesh	19565	2000	200000

To efficiently apply the proposed sample size for producing direct estimates of crop area in the respective states, a stratified three-stage random sampling design has been adopted, as described below:

1. Districts were considered as stratum. All the districts will be enumerated.
2. The total number of villages to be selected in the state was distributed among the districts in proportion to their Gross Cropped Area (GCA) under the major crops. For each district, an appropriate minimum sample size was determined

based on standard sample size criteria; in Uttar Pradesh, this minimum was fixed at 15 villages per district. This step was necessary because several districts do not cultivate the major food grain crops such as rice and wheat but still have substantial agricultural activity in other crops (for example, the sugarcane belt in Uttar Pradesh).

3. Within each district, an appropriate number of tehsils was selected using Simple Random Sampling Without Replacement (SRSWOR), with 50% of the tehsils in each district being selected. Subsequently, the allocated villages were distributed among the selected tehsils in proportion to their GCA.
4. Within each stratum, the assigned number of villages was chosen from the complete list of villages using SRSWOR.
5. For area estimation, 100 survey numbers were randomly selected in clusters of five within each selected village. These survey numbers were enumerated across all seasons during AY 2015–16. Thus, in Uttar Pradesh, a total of 2000 × 100 = 200,000 survey numbers were enumerated.
6. Each selected survey number was visited by the primary worker for plot-to-plot crop area measurement using the gold-standard "rope and compass" method along with the Global Positioning System (GPS) throughout different seasons.

The selection of sample units up to the second-stage units (SSUs), i.e., villages, was carried out by ICAR–IASRI, New Delhi, while the third-stage units (TSUs), consisting of survey or sub-survey numbers for area enumeration, were chosen by the respective state field officials. For field data collection, one questionnaire was developed for area enumeration and two separate questionnaires were designed for conducting crop-cutting experiments.

2.1 Sampling Design Adopted for Area Estimation at District Level

The sampling design used for crop area estimation is a Stratified Three-Stage Random Sampling design, wherein districts serve as strata, tehsils/taluks within each stratum act as FSUs, revenue villages within each FSU are treated as SSUs, and survey or sub-survey numbers within each selected village constitute the TSUs. For every selected TSU, the geographic area of the survey number and the area under each cultivated crop were recorded. Following the approaches of Sukhatme *et al.* (1984) and Singh *et al.* (1996), the proposed methodology is outlined below:

Let, y_{ijk} = area under a particular crop in the k^{th} survey/sub-survey number of j^{th} village in i^{th} tehsil/taluk,

N_{ij} = total survey number/sub-survey in the j^{th} village of i^{th} tehsil/taluk,

n_{ij} = total number of selected survey number/sub-survey in the j^{th} village in i^{th} tehsil/taluk,

V_i = total number of villages in the i^{th} tehsil/taluk,

v_i = total number of selected villages in the i^{th} tehsil/taluk,

T = total number of tehsil/taluks in the district, and

t = total number of selected tehsil/taluks in the district.

An estimator of total area under a crop in the d^{th} district is given by

$$\hat{Y}_D = \frac{T}{t}\sum_{i=1}^{t}\frac{V_i}{v_i}\sum_{j=1}^{v_i}\frac{N_{ij}}{n_{ij}}\sum_{k=1}^{n_{ij}} y_{ijk} \tag{2.1.1}$$

The variance of the estimator $\hat{Y}_D$ is given by

$$V(\hat{Y}_D) = T^2\left(\frac{1}{t}-\frac{1}{T}\right)S_{bu}^2 + \frac{T}{t}\sum_{i=1}^{T}V_i^2\left(\frac{1}{v_i}-\frac{1}{V_i}\right)S_{ui}^2 + \frac{T}{t}\sum_{i=1}^{T}\frac{V_i}{v_i}\sum_{j=1}^{V_i}N_{ij}^2\left(\frac{1}{n_{ij}}-\frac{1}{N_{ij}}\right)S_{ijk}^2 \tag{2.1.2}$$

where, $S_{bu}^2 = \frac{1}{T-1}\sum_{i=1}^{T}(V_i\bar{Y}_{i.}-\bar{Y})^2$,

$$S_{ui}^2 = \frac{1}{V_i-1}\sum_{i=1}(N_{ij}\bar{Y}_{ij}-\bar{Y}_{i.})^2,\ S_{ijk}^2 = \frac{1}{N_{ij}-1}\sum_{i=1}^{N}(y_{ijk}-\bar{Y}_{ij})^2$$

The estimator of variance of $\hat{Y}_D$ is given by,

$$\hat{V}(\hat{Y}_D)=T^2\left(\frac{1}{t}-\frac{1}{T}\right)s_{bu}^2+\frac{T}{t}\sum_{i=1}^{t}V_i^2\left(\frac{1}{v_i}-\frac{1}{V_i}\right)s_{ui}^2+\frac{T}{t}\sum_{i=1}^{t}\frac{V_i}{v_i}\sum_{j=1}^{v_i}N_{ij}^2\left(\frac{1}{n_{ij}}-\frac{1}{N_{ij}}\right)s_{ijk}^2 \quad (2.1.3)$$

where, $s_{bu}^2=\frac{1}{t-1}\sum_{i=1}^{t}(V_i\bar{y}_{i.}-\hat{\bar{y}})^2$,

$s_{ui}^2=\frac{1}{v_i-1}\sum_{j=1}^{v_i}(N_{ij}\bar{y}_{ij}-\bar{y}_{i.})^2$, $s_{ijk}^2=\frac{1}{n_{ij}-1}\sum_{k=1}^{n_{ij}}(y_{ijk}-\bar{y}_{ij})^2$.

2.2 Estimation of Total Crop Area at State Level

The state-level estimate of the total area under a crop is computed by adding together the corresponding area estimates from all individual districts. i.e.

$$\hat{Y}=\sum_{D=1}^{S}\hat{Y}_D \quad (2.1.4)$$

where, S is the total number of districts in the state.

The variance estimate of the total crop area at the state level is calculated by adding together the variance estimates of the total crop area from each district. i.e.

$$\hat{V}(\hat{Y})=\sum_{D=1}^{S}\hat{V}(\hat{Y}_D) \quad (2.1.5)$$

3. RESULTS & DISCUSSION

In Uttar Pradesh, data for two major crops—rice and wheat—were collected by the state agency during the Kharif and Rabi seasons of the 2015–16 agricultural year. Although 2,000 villages were initially planned for selection, with an estimated 200,000 survey numbers/ Khasra numbers to be enumerated, the actual coverage fell short. Only 1,926 villages were surveyed in the Kharif season and 1,845 in the Rabi season. Owing to limited manpower and various administrative constraints, it was not possible to conduct surveys in all the selected villages or districts.

Table 2. District-wise Estimates of Area with % CV for Paddy crop in Uttar Pradesh for AY 2015-16

S. No.	District	Crop	Total Area Estimate (ha)	%CV
1	Aligarh	Paddy	341824	25.99
2	Allahabad	Paddy	431805	11.01
3	Ambedkar Nagar	Paddy	256580	18.45
4	Amethi	Paddy	363658	11.12
5	Amroha	Paddy	61454	20.95
6	Baghpat	Paddy	24905	27.37
7	Bahraich	Paddy	296052	7.53
8	Ballia	Paddy	139126	22.15
9	Balrampur	Paddy	217989	14.66
10	Banda	Paddy	236353	15.63
11	Barabanki	Paddy	313632	12.53
12	Basti	Paddy	488448	19.93
13	Bhadrohi	Paddy	63274	16.22
14	Bijnor	Paddy	100721	23.67
15	Bulandshahr	Paddy	230099	8.79
16	Chandauli	Paddy	238286	18.52
17	Deoria	Paddy	257896	16.61
18	Faizabad	Paddy	463065	18.55
19	Ghazipur	Paddy	235204	11.39
20	Gorakhpur	Paddy	755692	22.15
21	Kaushambi	Paddy	201229	21.01
22	Kheri	Paddy	346155	15.94
23	Kushinagar	Paddy	200691	14.82
24	Meerut	Paddy	113432	12.91
25	Moradabad	Paddy	239094	12.16
26	Muzaffarnagar	Paddy	105697	14.87
27	Pilibhit	Paddy	192939	21.92
28	Pratapgarh	Paddy	73309	22.00
29	Rae Bareli	Paddy	165173	17.54
30	Rampur	Paddy	189456	8.95
31	Sambhal	Paddy	73964	19.02
32	Sant Kabir Nagar	Paddy	349732	24.34
33	Shamli	Paddy	34149	26.36
34	Siddharthnagar	Paddy	271722	9.25
35	Shrawasti	Paddy	65350	15.26
36	Sultanpur	Paddy	130193	10.05
37	Unnao	Paddy	289486	22.62
38	Hapur	Paddy	89114	10.90
39	Maharajganj	Paddy	375479	14.01
40	Saharanpur	Paddy	140823	13.30
41	Sitapur	Paddy	311599	28.21

Table 3. District-wise Estimates of Area with % CV for Wheat crop in Uttar Pradesh for AY 2015-16

S. No.	District	Crop	Total Area Estimate (ha)	%CV
1	Agra	Wheat	214858	10.08
2	Aligarh	Wheat	637576	14.51
3	Allahabad	Wheat	432355	16.24
4	Ambedkar Nagar	Wheat	282101	15.46
5	Amethi	Wheat	365466	12.80
6	Amroha	Wheat	184772	21.25
7	Baghpat	Wheat	91380	25.74
8	Bahraich	Wheat	364210	8.38
9	Ballia	Wheat	217935	19.88
10	Balrampur	Wheat	567980	18.41
11	Barabanki	Wheat	256152	10.81
12	Basti	Wheat	564774	13.86
13	Bhadrohi	Wheat	83460	18.92
14	Bijnor	Wheat	226198	10.31
15	Budaun	Wheat	673240	18.85
16	Bulandshahr	Wheat	416952	6.95
17	Chandauli	Wheat	272647	13.59
18	Chitrakoot	Wheat	166785	18.08
19	Deoria	Wheat	324718	25.16
20	Etawah	Wheat	227998	14.84
21	Faizabad	Wheat	528938	18.45
22	Ghaziabad	Wheat	164703	16.30
23	Ghazipur	Wheat	372749	24.18
24	Gonda	Wheat	143855	22.51
25	Hapur	Wheat	141666	10.42
26	Hardoi	Wheat	516698	9.65
27	Hathras	Wheat	475729	19.58
28	Jaunpur	Wheat	368479	10.62
29	Kannauj	Wheat	100821	20.24
30	Kanpur Nagar	Wheat	392573	18.85
31	Kaushambi	Wheat	363846	17.77
32	Kheri	Wheat	331318	10.82
33	Mirzapur	Wheat	446121	22.95
34	Muzaffarnagar	Wheat	304587	10.58
35	Pilibheet	Wheat	186798	17.54
36	Rampur	Wheat	218822	10.12
37	Saharanpur	Wheat	206199	23.12
38	SantKabir Nagar	Wheat	316448	18.47
39	Shrawasti	Wheat	100762	21.97
40	Siddharthnagar	Wheat	264105	9.35
41	Sultanpur	Wheat	121389	11.96
42	Unnao	Wheat	603218	18.49
43	Gorakhpur	Wheat	608764	24.12
44	Kashganj	Wheat	136003	14.31
45	Moradabad	Wheat	220311	12.11
46	Sitapur	Wheat	335611	17.24
47	Sambhal	Wheat	129895	17.21
48	Raebareily	Wheat	214533	10.93
49	Maharajganj	Wheat	384646	14.16
50	Meerut	Wheat	275341	14.14
51	Jhansi	Wheat	289871	21.06
52	Firozabad	Wheat	90840	26.10
53	Bareily	Wheat	486454	19.63

A careful examination of Tables 2 and 3 indicates that design-based district-level crop area estimates are available for only 41 districts for paddy and 53 districts for wheat, out of the total 75 districts in the agricultural year 2015–16. For paddy, the percentage coefficient of variation (%CV) ranges from 7.53 to 28.21, while for wheat it ranges from 6.95 to 26.10. In several districts, the %CV exceeds 20%, as observed in both tables. This high variability is primarily attributable to errors made by field investigators during data collection, including misinterpretation of instructions and the use of the highest survey number instead of the total number of survey numbers while preparing the sampling frame. Consequently, design weights were incorrectly calculated in some districts, deviating from those specified in the sampling design. Additionally, various non-sampling errors during data entry further reduced the effective sample size, thereby inflating the %CV of the district-level estimates.

4. CONCLUSIONS

The primary objective of the pilot project was to reduce the workload of both state and national agencies in generating agricultural statistics and to phase out the traditional patwari system for collecting area statistics in the TRS states, as this age-old method has shown declining data quality. The methodology proposed under the pilot project for collecting crop area statistics is entirely stochastic and independent of any redundant system, such as the patwari system, while producing crop area estimates within an acceptable range of precision. When implemented accurately without deviations, this approach is expected to yield more precise estimates. Additionally, the proposed methodology is easily adaptable and less time-consuming compared to TRS surveys, owing to the relatively small sample size. The application of advanced sampling techniques, such as Small Area Estimation, can further enhance the precision of estimates and enable the generation of reliable estimates for non-sampled districts, where conventional design-based estimates are limited due to insufficient sample size.

5. LIMITATIONS AND CONSTRAINTS OF THE STUDY

For area estimation, sample survey approach was considered based on enumeration of 100 survey numbers randomly selected from each selected village instead of complete enumeration of selected villages. With the given sample size, the major crops were captured in the selected villages, if it was available. For area enumeration one of the major issues emerged was selection of 100 survey numbers in place of complete enumeration. Due to reduction of sample sizes, estimates can only be produced for selected major crops only. If the stratums were agroclimatic zones instead of districts, the study will have been easier.

REFERENCES

Report of Expert Committee on Agricultural Statistics (2011). Directorate of Economics and Statistics, Ministry of Agriculture and Farmers welfare, Govt. of India.

Singh, R., and Mangat, N.P.S. (1996). *Elements of Survey Sampling*. Kluwer Academic publishers.

Sud, U.C., Ahmad, T., Gupta, V.K., Chandra, H., Sahoo, P.M., Aditya, K., Singh, M. and Biswas, A. (2017). Crop 1 & 2: Measuring Crop Area and Yield under Pure Stand, Mixed and Continuous Cropping. Working Paper No. 16. Global Strategy, FAO, Rome Publication. http://gsars.org/wp-content/uploads/2017/01/WP-16.01.2017-Findings-from-the-Field-Tests-Conducted-in-Three-Countries.pdf.

Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok, C. (1984). *Sampling Theory of Surveys with Applications*. Third Revised Edition, Iowa State University Press, USA.