https://doi.org/10.56093/JISAS.V80I1.18
Available online at https://isas.live/newsite
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 80(1) 2026 207–220

Modified Sampling Methodology for Estimation of Area and Production of Horticultural Crops

Tauqueer Ahmad, Prachi Misra Sahoo, Ankur Biswas, Kaustav Aditya, Deepak Singh and Raju Kumar

ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 03April 2026; Revised 22 April 2026; Accepted 22 April 2026

SUMMARY

Reliable estimation of horticultural crop area and production is crucial for agricultural planning in India. The earlier Crop Estimation Survey-Fruits & Vegetables (CES-F&V) methodology developed by ICAR–IASRI was discontinued due to operational complexity and high cost, necessitating an alternative approach. This study presents a modified sampling methodology developed under the CHAMAN programme for integrated assessment of fruits and vegetables. Field testing across Maharashtra, Andhra Pradesh, Tamil Nadu, and Himachal Pradesh generated district-level estimates of area, production, and yield. The results showed close agreement with official statistics of the Department of Agriculture & Farmers Welfare, with acceptable precision levels. A significant reduction in sample size, from 150–200 to 80 villages per district, improved cost-effectiveness and operational efficiency without compromising accuracy. The methodology demonstrates strong potential for large-scale implementation, including the use of model-based approaches in low-productivity districts.

Keywords: Horticulture; Sampling design, Area and production estimation; Stratified multistage sampling; CHAMAN; Field testing.

1. INTRODUCTION

Agriculture plays a vital role in the Indian economy. Horticulture sector is an important sub-sector of the agriculture with its contribution of about 33% in the economy of agriculture and allied sectors. Horticulture in India encompasses fruits, vegetables, spices, flowers, and medicinal and aromatic plants, and is a major driver of agricultural growth, contributing nearly one-third to the agricultural and allied economy. India is the second-largest producer of fruits and vegetables globally, with fruits and vegetables accounting for about 90% of total horticultural output. As per recent estimates, total horticultural production in India reached about 370.85 million tonnes from 29.9 million hectares in 2024–25, including 117.65 million tonnes of fruits and 217.79 million tonnes of vegetables (MoA&FW, 2026).

The estimation of area, yield, and production of horticultural crops is statistical complex due to the heterogeneity practices, multi-harvest nature, and spatial variability. Despite being a leading producer, India lacks reliable official estimates for many horticultural crops, creating a significant gap in the agricultural statistical system. Robust, timely, and accurate data are therefore essential for effective policy formulation and development planning.

Traditionally, horticultural area and production were estimated through the Crop Estimation Survey on Fruits and Vegetables (CES-F&V), implemented by the Directorate of Economics and Statistics (DES) based on a stratified three stage random sampling design, developed by ICAR–IASRI. The estimators of production are typically derived using design-based unbiased estimators under the Horvitz–Thompson framework or ratio estimators incorporating auxiliary information (Cochran, 1977; Sukhatme *et al.,* 1984). However, the CES-F&V design is operationally

Corresponding author: Tauqueer Ahmad
E-mail address: tauqueer.khan01@gmail.com

intensive, requiring large sample (150–200 villages per district), multiple visits, and crop-specific surveys, limiting scalability and efficiency (Ahmad *et al.,* 2014).

In contrast, estimates compiled by National Horticulture Board (NHB) were based on reports furnished by the State Directorates and market intelligence, lacking statistically rigorous sampling and thus prone to reporting biases. Conversely, estimates from the Directorate of Economics and Statistics (DES) were derived from pilot surveys using robust statistical methodologies. Notable discrepancies existed between NHB and DES estimates across states and crops. Since 2012–13, the Division of Horticulture under the Ministry of Agriculture & Farmers Welfare has been publishing consolidated data through Horticulture Statistics at a Glance.

Recognizing these limitations, the National Statistical Commission (2001) recommended the developing an alternative sampling methodology integrating auxiliary data such as market arrivals, export data, and grower-level information. In response, ICAR-IASRI conducted pilot studies to design a stratified multistage sampling method with integrated crop coverage. The proposed methodology demonstrated practical applicability (Ahmad *et al.,* 2011) and was subsequently validated at the state level in Himachal Pradesh (Ahmad *et al.,* 2015).

The methodology was stratified multistage sampling design, with selection of taluks and villages by SRSWOR and fields/orchards as intermediate units. It enables integrated estimation of multiple crops, improving efficiency, with production estimated using expansion and ratio estimators and variance computed under standard stratified sampling methods (Cochran, 1977).

This study developed and evaluated a modified sampling methodology with domain estimation of area and production of horticultural crops. It assesses statistical properties (bias, variance, efficiency) and operational feasibility relative to existing methods, providing a scalable framework to strengthen horticultural statistics in India.

The specific objectives of the study are to develop a modified methodology and evaluate the performance of the proposed modified methodology for estimating crop-wise acreage of major fruit and vegetable crops, and to assess its reliability in estimating yield rates and total production at the state level.

2. MODIFIED SAMPLING METHODOLOGY FOR ESTIMATION OF AREA AND PRODUCTION OF HORTICULTURAL CROPS

The study proposes a stratified multistage sampling design to generate reliable estimates of horticultural crops area and production. The methodology incorporates design-based estimation procedures along with auxiliary information to enhance precision and representativeness.

2.1 Sampling design

The study suggested a stratified multistage random sampling design (Fig. 1). Districts were stratified by area under fruits and vegetables into high-productivity (70–80% of total area) and low-productivity strata, with ~40% districts selected from the high-productivity strata and two from the low-productivity strata using SRSWOR. Selected districts were further stratified at the taluk/tehsil level (60–70% area as high-productivity; units <5% excluded), and two taluks from each group were selected. Subsequently, 20 villages per taluk were selected by SRSWOR and completely enumerated. For fruits, five orchards (≥12 bearing trees) were sampled per village, with proportional allocation across major crops, and yield estimated from randomly selected tree clusters. For vegetables, growers meeting minimum area criteria (≥0.1 ha plains; ≥0.01 ha hills) were

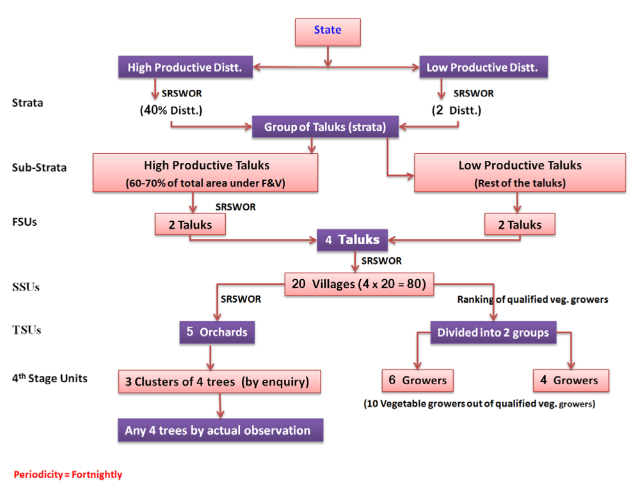

Fig. 1. Flow diagram of the stratified multistage sampling design used in the study

stratified and 10 growers per village were sampled (6:4 allocation), with yield measured using randomly located 5 m × 5 m plots at harvest. Flow diagram of the proposed design is presented in Fig. 1.

2.2 Study area

The study was conducted across four major horticulture-producing states of India, namely Maharashtra, Andhra Pradesh, Tamil Nadu, and Himachal Pradesh. Districts were selected from both high- and low-productivity strata to ensure adequate representation across varying levels of horticultural intensity. The stratum-wise selected districts is listed in Table 1.

2.3 Data used for the study

Both primary and secondary data were used in the study. Primary data was collected to generate reliable estimates of horticultural crops area and production, while secondary data were used for stratification and validation of estimates. The data were collected as part of the field implementation of the modified sampling methodology under the Coordinated programme on Horticulture Assessment and Management using geo informatics (CHAMAN) programme relating to generation of horticultural statistics using survey sampling

2.3.1 Primary data

Primary data was collected through field surveys (2016-2017 across Maharashtra, Andhra Pradesh, Tamil Nadu, and Himachal Pradesh using 10 ICAR–IASRI schedules covering different sampling stages and units. Primary data were digitized through dedicated data entry software, and a dual-mode of data collection mechanism using physical observations by investigator along with farmer's enquiry was adopted to minimize measurement and reporting errors..

2.3.2 Secondary data

Secondary data of agricultural year 2015-16 was used for sampling design and stratification, and validation of the estimates were done using district-level statistics from the Ministry of Agriculture & Farmers Welfare, state horticulture departments, and CES-F&V datasets. Auxiliary data such as market arrivals and administrative records were also used for stratification, crop identification, and validation of estimates.

2.3.3 Data Processing and Management

Dedicated data entry and data analysis softwares were developed for the digitization, validation, and processing of survey data. Field data collected through paper based schedules were digitized to a centralized server at ICAR–IASRI, enabling automated validation, consistency checks, and quality control. The processed dataset supports generation of village-, district-, and state-level estimates using design-based estimators under the stratified multistage sampling framework.

2.4 Estimation procedures

2.4.1 Estimation procedures for fruits

Estimation procedures for estimating area, production, productivity, number of bearing trees, number stray bearing trees etc. of fruits at district as well as state level were developed and are explained as under:

a. Estimation of number of trees and area under fruit crops in a district

Notations followed:

K = Total number of district level strata (group of districts)

L_{id} = Total number of sub-strata in d-th district of i-th district level stratum

Table 1. Stratum-wise selected districts in Andhra Pradesh, Himachal Pradesh, Maharashtra, and Tamil Nadu for the study based on horticultural area coverage

State	No. of selected districts	No of districts on the basis of stratum		Selected districts	
		Stratum-1	Stratum-2	Stratum-1	Stratum-2
Andhra Pradesh	13	7	6	Chittoor, Kurnool, Anantapur	Guntur, West Godavari
Himachal Pradesh	12	5	7	Shimla, Kangra	Una, Solan
Maharashtra	33	10	23	Nashik, Ratnagiri, Satara, Jalna	Beed, Nandurbar
Tamil Nadu	31	9	22	Salem, Vellupuram, Krishnagiri, Erode	Ariyalur, Ramnathapurum

D_i = Total number of districts in i-th district level stratum, i = 1,2,…,K

d_i = Number of districts selected from i-th district level stratum, i = 1,2,…,K

d_{if} = Number of districts selected from i-th district level stratum for f-th fruit, i = 1,2,…,K, f = 1, 2, . . .,F and F = Total number of fruit crops

T_{idh} = Total number of taluks/tehsils in h-th sub-stratum of d-th selected district of i-th district level stratum, h = 1, 2,…,L_{id}, d=1,2,…,d_i , i=1,2,…,K

t_{idh} = Number of taluks/tehsils selected from h-th sub-stratum in d-th district of i-th district level stratum

t_{idhf}= Number of taluks/tehsils selected from h-th sub-stratum in d-th district of i-th stratum having f-th fruit

V_{idht} = Total number of villages in t-th selected taluk/tehsil of h-th sub-stratum of d-th selected district in i-th district level stratum, h = 1, 2,..., L_{id} and t = 1, 2, . . .,t_{idh}

v_{idht} = Number of selected villages from t-th selected taluk/tehsil of h-th sub-stratum of d-th district in i-th stratum

v_{idhtf} = Number of selected villages from t-th selected taluk/tehsil of h-th sub-stratum in d-th district of i-th stratum having f-th fruit

R_{idhtvf} = Total number of orchards in v-th selected village from t-th selected taluk/tehsil of h-th sub-stratum in d-th selected district of i-th district level stratum having f-th fruit, v = 1, 2, . . ., v_{idht}

r_{idhtvf} = Number of selected orchards in v-th selected village from t-th selected taluk/tehsil of h-th sub-stratum in d-th selected district of i-th district level stratum for f-th fruit

$B_{idhtvrf}$ = Total number of bearing trees in r-th selected orchard of v-th selected village from t-th selected taluk/tehsil of h-th sub-stratum in d-th selected district of i-th district level stratum for f-th fruit, r = 1, 2, . . .,r_{idhtvf}

$b_{idhtvrf}$ = Number of selected bearing trees in r-th selected orchard of v-th selected village from t-th selected taluk/tehsil of h-th sub-stratum in d-th selected district of i-th district level stratum for f-th fruit

x_{idhtvf} = Value of the character under study e.g. number of trees, bearing or young, area under the orchard etc. recorded on the basis of complete enumeration of v-th village selected from t-th taluk/tehsil of h-th sub-stratum in d-th selected district of i-th district level stratum for f-th fruit.

Following domain estimation approach (Särndal *et al.* 1992), an estimate of total for the character x (briefly we call it total unless otherwise stated) for h-th sub-stratum of d-th district in i-th stratum for f-th fruit is given by

$$\hat{X}_{idhf} = \frac{T_{idh}}{t_{idh}} \sum_{t=1}^{t_{idhf}} \frac{V_{idht}}{v_{idht}} \sum_{v=1}^{v_{idhtf}} x_{idhtvf} .$$

The estimates of all the strata in a district are added to get the estimate at district level as following

$$\hat{X}_{idf} = \sum_{h=1}^{L_{id}} \hat{X}_{idhf} . \tag{1}$$

An estimate of variance of $\hat{X}_{idf}$ is given by

$$\hat{V}\left(\hat{X}_{idf}\right) = \sum_{h}^{L_{id}} \hat{V}(\hat{X}_{idhf})$$

$$= T_{idh}^2 \left[\left(\frac{1}{t_{idh}} - \frac{1}{T_{idh}} \right) \left(p_{idhf} s_{idhf}^2 + p_{idhf} q_{idhf} \overline{x}_{idhf}^2 \right) + \frac{1}{T_{idh} \times t_{idh}} \sum_{t=1}^{t_{idh}} V_{idht}^2 \left(\frac{1}{v_{idht}} - \frac{1}{V_{idht}} \right) \left(p_{idhtf} s_{idhtf}^2 + p_{idhtf} q_{idhtf} \overline{x}_{idhtf}^2 \right) \right]$$

where,

$$p_{idhf} = \frac{t_{idhf}}{t_{idh}}, \quad q_{idhf} = 1 - p_{idhf}, \quad p_{idhtf} = \frac{v_{idhtf}}{v_{idht}},$$

$$q_{idhtf} = 1 - p_{idhtf}, \quad \overline{x}_{idhf} = \frac{1}{t_{idhf}} \sum_{t=1}^{t_{idhf}} V_{idht} p_{idhtf} \overline{x}_{idhtf},$$

$$s_{idhf}^2 = \frac{1}{t_{idhf} - 1} \sum_{t=1}^{t_{idhf}} \left(V_{idht} p_{idhtf} \overline{x}_{idhtf} - \overline{x}_{idhf} \right)^2,$$

$$s_{idhtf}^2 = \frac{1}{v_{idhtf} - 1} \sum_{v=1}^{v_{idhtf}} \left(x_{idhtvf} - \overline{x}_{idhtf} \right)^2, \text{ and}$$

$$\overline{x}_{idhtf} = \frac{1}{v_{idhtf}} \sum_{v=1}^{v_{idhtf}} x_{idhtvf}$$

Estimate of standard error of $\hat{X}_{idf}$ is given by

$$\hat{S.E.}(\hat{X}_{idf}) = \sqrt{\hat{V}(\hat{X}_{idf})}\ .$$

Estimate of percentage standard error of $\hat{X}_{idf}$ is given by

$$\%\,\hat{S.E.}\left(\hat{X}_{idf}\right) = \frac{\hat{S.E.}(\hat{X}_{idf})}{\hat{X}_{idf}} \times 100. \tag{2}$$

b. Estimation of number of trees and area under fruit crops at state level

An estimate of total for the character x for f-th fruit in a state is given by

$$\hat{X}_f = \sum_{i=1}^{K} \hat{X}_{if} = \sum_{i=1}^{K} \frac{D_i}{d_i} \sum_{d=1}^{d_{if}} \hat{X}_{idf} \tag{3}$$

where, $\hat{X}_{idf}$ is district level estimate for the d-th district under i-th stratum for the f-th fruit given in Equation (1).

Approximate estimate of variance of $\hat{X}_f$ is given by

$$\hat{V}\left(\hat{X}_f\right) = \sum_{i}^{K} \hat{V}\left(\hat{X}_{if}\right)$$

$$= \sum_{i}^{K} D_i^2 \left(\frac{1}{d_i} - \frac{1}{D_i}\right)\left(p_{if} s_{if}^2 + p_{if} q_{if} \overline{x}_{if}^2\right)$$

where,

$$s_{if}^2 = \frac{1}{d_{if} - 1} \sum_{d=1}^{d_{if}} \left(\hat{X}_{idf} - \overline{x}_{if}\right)^2,\ \ \overline{x}_{if} = \frac{1}{d_{if}} \sum_{d=1}^{d_{if}} \hat{X}_{idf}\,,$$

$p_{if} = \dfrac{d_{if}}{d_i}$ and $q_{if} = 1 - p_{if}$.

c. Estimation of average yield per tree of fruit crops in a district

Let, $y_{idhtvrbf}$ be the yield of b-th selected bearing tree of r-th selected orchard of of v-th selected village from t-th taluk/tehsil of h-th sub-stratum in d-th selected district of i-th district level stratum for f-th fruit. An estimate of average yield of the f-th fruit in the t-th taluk/tehsil of h-th sub-stratum in d-th selected district of i-th district level stratum is given by

$$\overline{y}_{idhtf} = \frac{1}{v_{idhtf}} \sum_{v=1}^{v_{idhtf}} \overline{y}_{idhtvf}$$

where,

$$\overline{y}_{idhtvf} = \frac{1}{r_{idhtvf}} \sum_{r=1}^{r_{idhtvf}} \overline{y}_{idhtvrf} \text{ and}$$

$$\overline{y}_{idhtvrf} = \frac{1}{b_{idhtvrf}} \sum_{b=1}^{b_{idhtvrf}} y_{idhtvrbf}\,.$$

An estimate of average yield of the f-th fruit in the h-th sub-stratum of d-th district in selected i-th stratum is given by

$$\overline{y}_{idhf} = \frac{\sum_{t=1}^{t_{idhf}} \hat{B}_{idhtf}\, \overline{y}_{idhtf}}{\sum_{t=1}^{t_{idhf}} \hat{B}_{idhtf}} \tag{4}$$

where, $\hat{B}_{idhtf}$ is the estimate of number of bearing trees of f-th fruit in the t-th taluk/tehsil of h-th sub-stratum in d-th selected district of i-th stratum and is given by

$$\hat{B}_{idhtf} = \frac{V_{idht}}{v_{idht}} \sum_{v=1}^{v_{idhtf}} b_{idhtvf}$$

where,

b_{idhtvf} is the total number of bearing trees of f-th fruit in v-th selected village (complete enumeration) of t-th selected taluk/tehsil of h-th sub-stratum in d-th selected district of i-th district level stratum.

An estimate of overall average yield of the f-th fruit at district level is given by

$$\overline{y}_{idf} = \sum_{h=1}^{L_{id}} \left(\frac{\hat{B}_{idhf}}{\hat{B}_{idf}}\right) \overline{y}_{idhf} \tag{5}$$

where,

$\hat{B}_{idhf} = \dfrac{T_{idh}}{t_{idh}} \sum_{t=1}^{t_{idhf}} \hat{B}_{idhtf}$ and $\hat{B}_{idf} = \sum_{h=1}^{L_{id}} \hat{B}_{idhf}$ are the estimates of number of bearing trees of f-th fruit in h-th sub-stratum of d-th selected district of i-th stratum and an estimate of total number of bearing trees at district level (pooled over all sub-strata) respectively.

An estimate of variance of $\bar{y}_{idf}$ is given by

$$\hat{V}\left(\bar{y}_{idf}\right)=\sum_{h=1}^{L_{id}}\left(\frac{\hat{B}_{idhf}}{\hat{B}_{idf}}\right)^2\hat{V}\left(\bar{y}_{idhf}\right).$$

where,

$$\hat{V}\left(\bar{y}_{idhf}\right)=\left(\frac{1}{t_{idh}}-\frac{1}{T_{idh}}\right)\frac{1}{t_{idhf}-1}\sum_{t=1}^{t_{idhf}}\left(\frac{\hat{B}_{idhtf}\bar{y}_{idhtf}}{\hat{B}_{idhf}}-\bar{y}_{idhf}\right)^2.$$

d. Estimation of average yield per tree of fruit crops in a state

An estimate of average yield of the f-th fruit for i-th stratum is given by

$$\bar{y}_{if}=\frac{1}{d_{if}}\sum_{d=1}^{d_{if}}\bar{y}_{idf}$$

An estimate of overall average yield of the f-th fruit at state level is given by

$$\bar{y}_{f}=\sum_{i=1}^{K}\left(\frac{\hat{B}_{if}}{\hat{B}_{f}}\right)\bar{y}_{if} \tag{6}$$

where,

$\hat{B}_{if}=\frac{D_i}{d_i}\sum_{d=1}^{d_{if}}\hat{B}_{idf}$ and $\hat{B}_{f}=\sum_{i=1}^{K}\hat{B}_{if}$ are the estimates of number of bearing trees of f-th fruit for i-th stratum and an estimate of total number of bearing trees at state level (pooled over all strata) respectively.

An estimate of variance of $\bar{y}_f$ is given by

$$\hat{V}\left(\bar{y}_{f}\right)=\sum_{i=1}^{K}\left(\frac{\hat{B}_{if}}{\hat{B}_{f}}\right)^2\hat{V}\left(\bar{y}_{if}\right)$$

where,

$$\hat{V}\left(\bar{y}_{if}\right)=\left(\frac{1}{d_i}-\frac{1}{D_i}\right)\frac{1}{d_i-1}\sum_{d=1}^{d_{if}}\left(\bar{y}_{idf}-\bar{y}_{if}\right)^2.$$

e. Estimation of production of fruit crops in a district

An estimate of total production of the f-th fruit at district level is given by

$$\hat{Y}_{idf}=\left(\bar{y}_{idf}\right)\times\left(\hat{B}_{idf}\right). \tag{7}$$

where, $\bar{y}_{idf}$ is the estimate of overall average yield for f-th fruit at district level and $\hat{B}_{idf}$ is an estimate of total number of bearing trees of f-th fruit in the district.

With the assumption that the average yield of bearing trees and total number of bearing trees are independent, an estimate of variance is given by

$$\hat{V}\left(\hat{Y}_{idf}\right)=\left(\bar{y}_{idf}\right)^2\times\hat{V}\left(\hat{B}_{idf}\right)+\left(\hat{B}_{idf}\right)^2\times\hat{V}\left(\bar{y}_{idf}\right)-\hat{V}\left(\hat{B}_{idf}\right)\times\hat{V}\left(\bar{y}_{idf}\right).$$

f. Estimation of production of fruit crops at state level

An estimate of total production of the f-th fruit at state level is given by

$$\hat{Y}_{f}=\left(\bar{y}_{f}\right)\times\left(\hat{B}_{f}\right) \tag{8}$$

where $\bar{y}_f$ is the estimate of overall average yield for f-th fruit at state level and $\hat{B}_f$ is an estimate of total number of bearing trees of f-th fruit in the state which is given by

$$\hat{B}_{f}=\sum_{i=1}^{K}\frac{D_i}{d_i}\sum_{d=1}^{d_{if}}\hat{B}_{idf}.$$

With the assumption that the average yield of bearing trees and total number of bearing trees are independent, an estimate of variance is given by

$$\hat{V}\left(\hat{Y}_{f}\right)=\left(\bar{y}_{f}\right)^2\times\hat{V}\left(\hat{B}_{f}\right)+\left(\hat{B}_{f}\right)^2\times\hat{V}\left(\bar{y}_{f}\right)-\hat{V}\left(\hat{B}_{f}\right)\times\hat{V}\left(\bar{y}_{f}\right).$$

2.4.2 Estimation procedures for vegetables

Estimation procedures for estimating area, production and productivity of vegetables at district level were developed and are explained as under:

a. Estimation of area under vegetables crops in a district

Notations followed:

K = Total number of district level strata (group of districts)

D_i = Total number of districts in i-th stratum

d_i = Number of selected districts i-th stratum

L_{id} = Total number of sub-strata in d-th district of i-th stratum; i=1,2,…,K

T_{idh} = Total number of taluks/tehsils in h-th sub-stratum in d-th selected district of i-th stratum, h = 1, 2, . . . ,L_{id}

t_{idh} = Number of taluks/tehsils selected from h-th sub-stratum in d-th selected district of i-th stratum

V_{idht} = Total number of villages in t-th selected taluk/tehsil of h-th sub-stratum in d-th district of i-th stratum, h =1,2…,L_{id} and t = 1,2,….t_{idh}

v_{idht} = Number of selected villages from t-th selected taluk/tehsil of h-th sub-stratum in d-th district of i-th stratum

$a^{c,s}_{idhtv}$ = Value of the character under study i.e. area under c-th vegetable in s-th season on the basis of complete enumeration of v-th village selected from t-th taluk/tehsil of h-th sub-stratum in d-th district of i-th stratum

An estimate of area under c-th crop in s-th season in t-th selected taluk/tehsil of h-th sub-stratum is given by

$$\hat{A}^{c,s}_{idht} = \frac{V_{idht}}{v_{idht}} \sum_{v=1}^{v_{idht}} a^{c,s}_{idhtv} .$$

An estimate of area under c-th crop in s-th season in h-th sub-stratum is given by

$$\hat{A}^{c,s}_{idh} = \frac{T_{idh}}{t_{idh}} \sum_{t=1}^{t_{idh}} \frac{V_{idht}}{v_{idht}} \sum_{v=1}^{v_{idht}} a^{c,s}_{idhtv} .$$

The estimates of all the strata in a district are added to get the estimate at district level.

$$\hat{A}^{c,s}_{id} = \sum_{h=1}^{L_{id}} \hat{A}^{c,s}_{idh} = T_{id} \sum_{h=1}^{L_{id}} W_{idh} \overline{a}^{\,c,s}_{idh} \tag{9}$$

where,

$$T_{id} = \sum_{h=1}^{L_{id}} T_{idh}, \quad W_{idh} = \frac{T_{idh}}{T_{id}} \text{ and } \overline{a}^{c,s}_{idh} = \frac{1}{t_{idh}} \sum_{t=1}^{t_{idh}} \frac{V_{idht}}{v_{idht}} \sum_{v=1}^{v_{idht}} a^{c,s}_{idhtv}$$

An estimate of variance of $\hat{A}^{c,s}_{id}$ is given by

$$\hat{V}\left(\hat{A}^{c,s}_{id}\right) = T^2_{id} \sum_{h=1}^{L_{id}} W^2_{idh}\, \hat{V}\left(\overline{a}^{c,s}_{idh}\right)$$

where,

$$\hat{V}\left(\overline{a}^{c,s}_{idh}\right) = \left(\frac{1}{t_{idh}} - \frac{1}{T_{idh}}\right) \frac{1}{t_{idh}-1} \sum_{t=1}^{t_{idh}} \left(V_{idht}\overline{a}^{c,s}_{idht} - \overline{a}^{c,s}_{idh}\right)^2 + \frac{1}{T_{idh} \times t_{idh}} \sum_{t=1}^{t_{idh}} \left(\frac{1}{v_{idht}} - \frac{1}{V_{idht}}\right) s^2_{idht},$$

$$\overline{a}^{c,s}_{idh} = \frac{1}{t_{idh}} \sum_{v=1}^{t_{idh}} V_{idht}\, a^{c,s}_{idht} ,$$

$$s^2_{idht} = \frac{1}{v_{idht}-1} \sum_{v=1}^{v_{idht}} \left(a^{c,s}_{idhtv} - \overline{a}^{c,s}_{idht}\right)^2 \text{ and } \overline{a}^{c,s}_{idht} = \frac{1}{v_{idht}} \sum_{v=1}^{v_{idht}} a^{c,s}_{idhtv}$$

Total area under c-th vegetable crop in the district during the year along with estimate of variance are given by

$$\hat{A}^{c}_{id} = \sum_{s=1}^{3} \hat{A}^{c,s}_{id} \text{ and } \tag{10}$$

$$\hat{V}\left(\hat{A}^{c}_{id}\right) = \sum_{s=1}^{3} \hat{V}\left(\hat{A}^{c,s}_{id}\right).$$

b. Estimation of area under vegetables crops in a state

An estimate of total area under c-th crop for the i-th district level strata is given by

$$\hat{A}^{c}_{i} = \frac{D_i}{d_i} \sum_{d=1}^{d_i} \hat{A}^{c}_{id} .$$

An estimate of total area under c-th crop at state level during the year can be obtained as

$$\hat{A}^{c} = \sum_{i=1}^{K} \hat{A}^{c}_{i} . \tag{11}$$

An approximate estimate of variance of $\hat{A}^c$ is given by

$$V\left(\hat{A}^{c}\right) = \sum_{i=1}^{K} D^2_i \left(\frac{1}{d_i} - \frac{1}{D_i}\right) \frac{1}{d_i - 1} \sum_{d=1}^{d_i} \left(\hat{A}_{id} - \hat{\bar{A}}_{id}\right)$$

where,

$$\hat{\bar{\mathrm{A}}}_{\mathrm{id}} = \frac{1}{d_i} \sum_{d=1}^{d_i} A^{c}_{id} .$$

c. Estimation of production of vegetable crops in a district

In continuation to earlier notations for vegetable crops, new notations are as followed:

P_{idhtv}^{c} = Total number of plots in v-th selected village of t-th selected taluk/tehsil from h-th sub-stratum in d-th district of i-th stratum for c-th vegetable crop

p_{idhtv}^{c} = Number of selected plots in v-th selected village of t-th selected taluk/tehsil from h-th sub-stratum in d-th district of i-th stratum for c-th vegetable crop

$A_{idhtv}^{c,s}$ = Area under c-th vegetable crop in v-th village of t-th taluk/thesils in h-th sub-stratum in d-th district of i-th stratum for c-th vegetable crop harvested during s-th season

A_{idhtv}^{c} = Area under c-th vegetable crop in v-th village of t-th taluk/thesil h-th sub-stratum in d-th district of i-th stratum

$\hat{A}_{idh}^{c,s}$ = Estimated area under c-th vegetable crop in h-th sub-stratum in d-th district of i-th stratum for during s-th season

$\hat{A}_{idh}^{c}$ = Estimated area under c-th vegetable crop in h-th sub-stratum in d-th district of i-th stratum

$\hat{A}_{id}^{c}$ = Estimated area under c-th vegetable crop in d-th district of i-th stratum

$y_{idhtvp}^{c,s}$ = Yield harvested from p-th plot of standard size having c-th vegetable crop in v-th selected village of t-th selected taluk/tehsil from h-th sub-stratum in d-th district of i-th stratum during s-th season

$\overline{y}_{idhtv}^{c,s} = \dfrac{1}{p_{idhtv}} \sum_{p=1}^{p_{idhtv}} y_{idhtvp}^{c,s}$ be the average yield of c-th vegetable crop in s-th season in v-th selected village of t-th selected taluk/tehsil from h-th sub-stratum in d-th district of i-th stratum

$\overline{y}_{idht}^{c,s} = \dfrac{1}{v_{idht}} \sum_{v=1}^{v_{idht}} \overline{y}_{idhtv}^{c,s} = \dfrac{1}{v_{idht}} \sum_{v=1}^{v_{idht}} \dfrac{1}{p_{idhtv}} \sum_{p=1}^{p_{idhtv}} y_{idhtvp}^{c,s}$ be the average yield of c-th vegetable crop in s-th season in t-th taluk/tehsil of h-th sub-stratum in d-th district of i-th stratum.

Estimates of total production for c-th crop in h-th sub-stratum of d-th district of i-th stratum:

Total production in v-th village during s-th season is given by

$\hat{Y}_{idhtv}^{c,s} = 400 \times \overline{y}_{idhtv}^{c,s} \times A_{idhtv}^{c,s}$ (Since size of the plot is 5m×5m).

Total production in v-th village of t-th thesil/taluk in d-th district in i-th stratum during the year is given by

$$\hat{Y}_{idhtv}^{c} = 400 \times \sum_{s=1}^{3} \left(\overline{y}_{idhtv}^{c,s} \times A_{idhtv}^{c,s} \right).$$

Total production in t-th taluk/tehsil in d-th district in i-th strata during the year is given by

$$\hat{Y}_{idht}^{c} = \frac{V_{idht}}{v_{idht}} \sum_{v=1}^{v_{idht}} \hat{Y}_{idhtv}.$$

Total production for c-th crop in h-th sub-stratum in d-th district in i-th strata is given by

$$\hat{Y}_{idh}^{c} = \frac{T_{idh}}{t_{idh}} \sum_{t=1}^{t_{idh}} \hat{Y}_{idht}^{c}. \tag{12}$$

An estimate of variance of $\hat{Y}_{idh}^{c}$ is approximately given by

$$\hat{V}\left(\hat{Y}_{idh}^{c}\right) = \frac{T_{idh}\left(T_{idh} - t_{idh}\right)}{t_{idh}\left(t_{idh} - 1\right)} \sum_{t=1}^{t_{idh}} \left(\hat{Y}_{idht}^{c} - \frac{\hat{Y}_{idh}^{c}}{T_{idh}} \right)^2 + \frac{T_{idh}}{t_{idh}} \sum_{t=1}^{t_{idh}} \left(\frac{1}{v_{idht}} - \frac{1}{V_{idht}} \right) s_{idht}^2.$$

where,

$$s_{idht}^2 = \frac{1}{v_{idht} - 1} \sum_{v=1}^{v_{idht}} \left(\hat{Y}_{idhtv}^{c} - \frac{\hat{Y}_{idht}^{c}}{V_{idht}} \right)^2.$$

Pooling of sub-strata estimates to get district level estimates:

An estimate of total production for c-th crop in the d-th district in i-th stratum during the year is obtained by adding the estimates of all the sub-strata within the district which is given by

$$\hat{Y}_{id}^{\ c} = \sum_{h=1}^{L_{id}} \hat{Y}_{idh}^{c}. \tag{13}$$

An estimate of variance of $\hat{Y}_{id}^{c}$ is given by

$$\hat{V}\left(\hat{Y}_{id}{}^{c}\right)=\sum_{h=1}^{L_{id}}\hat{V}\left(\hat{Y}_{idh}^{c}\right).$$

d. Estimation of production of vegetable crops in a state

Pooling of strata estimates to get state level estimate:

An estimate of total production for c-th crop in a state during the year is obtained by pooling the district level strata estimates in the state which is given by

$$\hat{Y}^{c}=\sum_{i=1}^{K}\hat{Y}_{i}^{c}=\sum_{i=1}^{K}\frac{D_i}{d_i}\sum_{d=1}^{d_i}\hat{Y}_{id}^{c}. \qquad (14)$$

An estimate of variance of $\hat{Y}^c$ is given by

$$\hat{V}\left(\hat{Y}^{c}\right)=\sum_{i=1}^{K}\hat{V}\left(\frac{D_i}{d_i}\sum_{d=1}^{d_i}\hat{Y}_{id}^{c}\right)$$

$$=\sum_{i=1}^{K}D_i^2\left(\frac{1}{d_i}-\frac{1}{D_i}\right)\frac{1}{d_i-1}\sum_{d=1}^{d_i}\left(\hat{Y}_{id}^{c}-\frac{\hat{Y}_{i}^{c}}{D_i}\right)^2.$$

e. Estimation of average yield of vegetable crops in a district

Average yield of c-th crop in a district is given by

$$\bar{y}_{id}{}^{c}=\frac{\hat{Y}_{id}{}^{c}}{\hat{A}_{id}{}^{c}}=\frac{\sum_{h=1}^{L_{id}}\hat{Y}_{idh}^{c}}{\sum_{h=1}^{L_{id}}\hat{A}_{idh}^{c}}. \qquad (15)$$

An approximate estimate of variance of $\bar{y}_{id}^{c}$ can be obtained by a formula of variance of combined ratio estimate

$$\hat{V}\left(\bar{y}_{id}{}^{c}\right)=\frac{1}{\left(\hat{A}_{id}^{c}\right)^2}\Big[\hat{V}\left(\hat{Y}_{id}^{c}\right)-2\,\bar{y}_{id}{}^{c}\,Cov\left(\hat{Y}_{id}^{c},\hat{A}_{id}^{c}\right)+\left(\bar{y}_{id}^{c}\right)^2\hat{V}\left(\hat{A}_{id}^{c}\right)\Big]$$

where, $Cov\left(\hat{Y}_{id}^{c},\hat{A}_{id}^{c}\right)$ denotes covariance between $\hat{Y}_{id}^{c}$ and $\hat{A}_{id}^{c}$ and is given by

$$Cov\left(\hat{Y}_{id}^{c},\hat{A}_{id}^{c}\right)=\sum_{h=1}^{L_{id}}\left[\frac{T_{idh}\left(T_{idh}-t_{idh}\right)}{t_{idh}\left(t_{idh}-1\right)}\sum_{t=1}^{t_{idh}}\left\{\left(\hat{Y}_{idht}^{c}-\frac{\hat{Y}_{idh}^{c}}{T_{idht}}\right)\left(\hat{A}_{idht}^{c}-\frac{\hat{A}_{idh}^{c}}{T_{idh}}\right)\right\}\right]$$

and

$$\hat{V}\left(\hat{A}_{id}^{c}\right)=\sum_{h=1}^{L_{id}}\left(\frac{T_{idh}\left(T_{idh}-t_{idh}\right)}{t_{idh}\left(t_{idh}-1\right)}\sum_{t=1}^{t_{idh}}\left(\hat{A}_{idht}^{c}-\frac{\hat{A}_{idh}^{c}}{T_{idh}}\right)^2+\frac{T_{idh}}{t_{idh}}\sum_{t=1}^{t_{idh}}\left(\frac{1}{v_{idht}}-\frac{1}{V_{idht}}\right)s_{idht,a}^2\right)$$

where, $\hat{A}_{idh}^{c}=\frac{T_{idh}}{t_{idh}}\sum_{t=1}^{t_{idh}}\hat{A}_{idht}^{c}$, $\hat{A}_{idht}^{c}=\frac{V_{idht}}{v_{idht}}\sum_{v=1}^{v_{idht}}a_{idhtv}^{c}$,

$$a_{idhtv}^{c}=\sum_{s=1}^{3}a_{idhtv}^{c,s}\ \text{ and }\ s_{idht,a}^2=\frac{1}{v_{idht}-1}\sum_{v=1}^{v_{idht}}\left(a_{idhtv}^{c}-\frac{\hat{A}_{idht}^{c}}{V_{idht}}\right)^2.$$

f. Estimation of average yield of vegetable crops at state level

Average yield of c-th crop in the state is given by

$$\bar{y}^{c}=\frac{\hat{Y}^{c}}{\hat{A}^{c}}=\frac{\sum_{i=1}^{K}\hat{Y}_{i}^{c}}{\sum_{i=1}^{K}\hat{A}_{i}^{c}} \qquad (16)$$

where,

$$\hat{Y}_{i}^{c}=\frac{D_i}{d_i}\sum_{d=1}^{d_i}\hat{Y}_{id}^{c}\ \text{ and }\ \hat{A}_{i}^{c}=\frac{D_i}{d_i}\sum_{d=1}^{d_i}\hat{A}_{id}^{c}.$$

An approximate estimate of variance of $\bar{y}^c$ is given by

$$\hat{V}\left(\bar{y}^{c}\right)=\frac{1}{\left(\hat{A}^{c}\right)^2}\Big[\hat{V}\left(\hat{Y}^{c}\right)-2\,\bar{y}^{c}\,Cov\left(\hat{Y}^{c},\hat{A}^{c}\right)+\left(\bar{y}^{c}\right)^2\hat{V}\left(\hat{A}^{c}\right)\Big]$$

where,

$$\hat{V}\left(\hat{A}^{c}\right)=\sum_{i=1}^{K}\left[D_i^2\left(\frac{1}{d_i}-\frac{1}{D_i}\right)\frac{1}{d_i-1}\sum_{d=1}^{d_i}\left(\hat{A}_{id}^{c}-\frac{\hat{A}_{i}^{c}}{D_i}\right)^2\right]$$

and

$$Cov\left(\hat{Y}^{c},\hat{A}^{c}\right)=\sum_{i=1}^{K}\left[D_i^2\left(\frac{1}{d_i}-\frac{1}{D_i}\right)\sum_{d=1}^{d_i}\left\{\left(\hat{Y}_{id}^{c}-\frac{\hat{Y}_{i}^{c}}{D_i}\right)\left(\hat{A}_{id}^{c}-\frac{\hat{A}_{i}^{c}}{D_i}\right)\right\}\right].$$

The estimates of standard error and the estimates of percentage standard error of the above estimators were obtained similar to Equation 2.

3. RESULTS AND DISCUSSIONS

The district wise estimates of fruit and vegetable area ('000 ha) and production ('000 MT) for the agricultural year 2015–16 were obtained based on the developed methodology and the estimates obtained by ICAR-IASRI and DAFW were compared across four states (Andhra Pradesh, Himachal Pradesh, Maharashtra, Tamil Nadu). ICAR-IASRI estimates are reported with percentage standard errors (% SE), and percentage deviations (% deviation = (ICAR-IASRI – DAFW)/DAFW × 100) are calculated. The comparison evaluates the consistency of ICAR-IASRI estimates against conventional DAFW statistics.

3.1 Fruit Crops

The district-wise estimates of area under fruit crops and production, percentage standard errors (% SE) and percentage deviations between the two sources were calculated. The comparison serves to evaluate the consistency of ICAR-IASRI estimates against conventional DAFW statistics.

3.1.1 Andhra Pradesh

The results indicate that ICAR-IASRI estimates of area and production generally exhibit negative deviations from DAFW values, although the magnitude varies across crops and districts (Table 2). For example, in Anantapur, banana area shows a deviation of –11.96%, whereas production registers a positive deviation of +6.68%, reflecting differences in yield estimation. Substantial positive deviations are observed in districts with intensive cultivation, particularly West Godavari (banana area +32.97%, production +29.48%) and Kurnool (area +20.41%, production +27.22%). These variations may be attributed to spatial clustering and possible inconsistencies in administrative reporting. The precision of ICAR-IASRI estimates, as indicated by percentage standard error (%SE), is moderate, with several estimates exceeding 20%. Higher uncertainty is particularly evident in minor crops, such as guava in Guntur (%SE: 32.67% for area and 43.49% for production), indicating greater spatial variability and limited sample representation. Overall, the results demonstrate that while the methodology provides reasonably consistent estimates, variability persists in heterogeneous cropping systems.

3.1.2 Himachal Pradesh

The results for Himachal Pradesh (Table 2) indicate a consistent pattern wherein survey-based estimates generated by ICAR–IASRI exceed the corresponding figures reported by the DAFW. Substantial positive deviations in production estimates were observed for important fruit crops, notably mango in Kangra (+42.79%), litchi in Kangra (+49.77%), and apple in Shimla (+18.49%), reflecting improved capture of spatial heterogeneity under the sampling methodology. The precision analysis reveals moderate to high variability, with percentage standard errors generally exceeding 20% across crops and districts. The highest uncertainty was observed for lemon production in Una (%SE = 61.95%) and plum in Solan (%SE = 46.25%), likely due to sparse cultivation and limited sample size.

3.1.3 Maharashtra

The results indicate a mixed pattern of agreement between survey-based and official estimates, reflecting both systematic and crop-specific variability. Positive deviations were particularly pronounced for mango in Satara (area: +25.54%; production: +27.27%) and Nandurbar (production: +23.85%), suggesting possible underestimation in administrative records or localized intensification of cultivation (Table 2). In contrast, negative deviations were observed for grapes in Nashik (area: –10.19%) and mosambi in Jalna (area: –13.76%), indicating discrepancies likely arising from reporting differences or spatial heterogeneity. Lower %SE values (≈9%) for mango area in Beed indicate relatively stable and homogeneous crop distribution, whereas higher values exceeding 35%—notably for mosambi production in Jalna and mango area in Nashik—highlight increased variability associated with dispersed cultivation and limited sample representation. A particularly notable case is Ratnagiri mango production, which exhibited a high %SE of 47.81% despite a modest area deviation (+4.39%), suggesting that yield variability rather than area estimation contributed to uncertainty.

3.1.4 Tamil Nadu

The estimates indicate a close agreement between survey-based (ICAR-IASRI) and official (DAFW) statistics (Table 2), with a tendency toward marginal overestimation in several districts. Positive deviations were observed for banana in Ariyalur (+15.31%

area; +14.02% production) and mango in Krishnagiri (+12.84% area; +18.35% production), suggesting robust capture of high-intensity horticultural zones. In contrast, moderate underestimation was evident for banana in Erode (–10.77% area; –8.50% production) and mango in Salem (–7.72% area; –6.58% production). The relatively higher percentage standard errors, particularly for banana in Erode (36.61% for area and 31.09% for production), reflect underlying spatial heterogeneity and sampling variability.

3.2 Vegetable Crops

The district-wise estimates of area under vegetable crops and production, percentage standard errors (% SE) and percentage deviations between the two sources were calculated. The comparison serves to evaluate the consistency of ICAR-IASRI estimates against conventional DAFW statistics.

3.2.1 Andhra Pradesh

The results of Andhra Pradesh indicate heterogeneous agreement between ICAR-IASRI and DAFW estimates across vegetable crops in Andhra Pradesh (Table 3). Tomato exhibited contrasting deviations, with underestimation in Anantpur (area: –18.94%; production: –16.14%) and mixed trends in Chittoor. Substantial positive deviations were observed for cauliflower in West Godawari (+34.9%) and cucumber in Kurnool (+23.07%), reflecting localized variability. Precision levels were moderate to low, with %SE frequently exceeding 20%, particularly for brinjal and tomato. Notably, extremely high uncertainty was observed for cucumber production in Kurnool (%SE = 75.34%), indicating instability in estimates for sparsely sampled crops.

3.2.2 Himachal Pradesh

Estimates of Himachal Pradesh reveals moderate consistency between ICAR-IASRI and DAFW, with positive area deviations for peas (Kangra: +15.35%) and onion (Solan: +20%) (Table 3). However, production estimates showed negative deviations for several crops, including cauliflower (–6.28%) and onion (–24%), suggesting variability in yield estimation. Precision varied considerably across crops, with high %SE values for peas (32.25%), lady finger (32.81%), and onion (47.24%). In contrast, select crops such as tomato in Shimla exhibited high precision (%SE = 2.33%), highlighting the influence of crop distribution and sampling intensity on estimator reliability.

3.2.3 Maharashtra

The results shows good agreement between ICAR-IASRI and DAFW estimates for major vegetable crops in Maharashtra, with most deviations within ±20%. Onion, the dominant crop, showed high consistency in Nashik (area –6.77%; production +2.91%) and Beed (+9.14%; +4.71%), although larger production deviations were observed in Satara (+32.7%) (Table 3). Moderate to high positive deviations were recorded for tomato in Beed (+66.84% area) and peas in Satara (+47.1% area), reflecting localized variability. Precision levels varied considerably, with %SE ranging from low (<10%) to high (>40%), particularly for onion area in Satara (%SE = 44.09%). Overall, the estimates demonstrate robustness for major crops but reduced reliability for minor or spatially dispersed crops.

3.2.4 Tamil Nadu

The findings show mixed patterns of agreement between ICAR-IASRI and DAFW estimates across vegetable crops in Tamil Nadu, with both positive and negative deviations observed (Table 3). Moderate deviations were noted for brinjal and potato, while substantial discrepancies were evident for tapioca in Ariyalur (production –92.3%), indicating possible reporting or sampling inconsistencies. Tomato estimates showed contrasting trends, with underestimation in Krishnagiri (area –17.43%) but overestimation in Villupuram (+47.07%). Precision varied widely, with relatively low %SE for certain crops (e.g., potato in Erode) and high uncertainty for others, particularly production estimates (>30% SE). These findings highlight the influence of crop heterogeneity and sampling intensity on estimation accuracy.

Across all states, the modified stratified multistage sampling framework produced consistent and unbiased estimates of area and production for major horticultural crops, with deviations from official statistics largely confined within ±10–20%. An important advantage of this modified methodology is that it can generate estimates for multiple fruits and vegetables from an integrated sample taken randomly. The methodology performed robustly for dominant crops and intensively cultivated regions (e.g., mango, banana, tomato, onion), where spatial continuity and higher sampling density

Table 2. District-wise comparison of ICAR-IASRI and DAFW estimates of fruit area and production in Andhra Pradesh, Himachal Pradesh, Maharashtra and Tamil Nadu States (2015-16)

State	District	Fruit	Area ('000 ha)			Production ('000 MT)		
			IASRI (±% SE)	DAFW	Deviation (%)	IASRI (±% SE)	DAFW	Deviation (%)
Andhra Pradesh	Anantapur	Banana	10.31 (±18.42)	11.71	–11.96	837.09 (±22.38)	784.7	6.68
		Mango	42.67 (±26.02)	47.69	–10.53	258.95 (±25.82)	286.12	–9.49
		Mosambi/ Orange	43.72 (±21.60)	46.63	–6.24	711.58 (±16.01)	720	–1.17
		Pomegranate	5.2 (±5.66)	5.66	–8.19	77.50 (±15.49)	84.96	–8.78
	Chittoor	Mango	77.23 (±16.51)	73.97	4.41	973.75 (±21.65)	931.64	4.52
	Guntur	Banana	3.87 (±20.46)	4.26	–9.20	208.01 (±24.66)	191.52	8.61
		Guava	0.13 (±32.67)	0.12	10.07	1.82 (±43.49)	1.77	3.02
		Lemon	2.21 (±24.96)	2.28	–2.92	37.84 (±25.43)	40.83	–7.34
		Mango	0.99 (±21.63)	1.09	–8.51	11.16 (±27.56)	11.81	–5.51
	Kurnool	Banana	4.60 (±20.18)	3.82	20.41	243.74 (±17.15)	191.15	27.22
		Mango	10.65 (±13.62)	10.14	5.02	172.95 (±18.05)	152.17	13.15
	West Godawari	Banana	12.58 (±31.95)	9.46	32.97	944.24 (±29.13)	729.23	29.48
		Mango	10.17 (±20.56)	8.51	19.52	102.91 (±25.06)	85.05	21
Himachal Pradesh	Kangra	Mango	21.82 (±4.88)	21.29	2.5	35.56 (±4.53)	24.9	42.79
		Litchi	3.42 (±17.16)	3.14	8.75	6.17 (±16.71)	4.12	49.77
		Pear	0.39 (±21.87)	0.38	3.44	1.83 (±23.31)	1.36	35.13
	Shimla	Apple	46.29 (±19.23)	39.73	16.52	571.58 (±36.21)	482.39	18.49
		Pear	1.53 (±20.16)	1.62	–5.43	6.76 (±45.90)	6.3	7.33
	Solan	Mango	1.74 (±13.28)	1.67	4.12	2.45 (±13.40)	1.97	24.77
		Plum	0.72 (±41.61)	0.62	15.53	1.87 (±46.25)	1.44	30.05
	Una	Mango	2.41 (±20.05)	2.07	16.8	8.36 (±25.47)	6.1	37.21
		Lemon	0.56 (±43.70)	0.49	13.44	1.47 (±61.95)	1.35	8.65
		Kinnu+Orange	1.46 (±26.27)	1.06	38.39	4.36 (±47.23)	3.01	44.93
Maharashtra	Beed	Lemon	1.12 (±19.22)	1.21	–7.44	11.17 (±17.62)	10.19	9.58
		Mango	4.52 (±9.29)	4.26	6.1	22.21 (±19.57)	20.44	8.67
		Mosambi	2.18 (±21.10)	2.01	8.46	20.67 (±31.56)	18.95	9.06
		Pomegranate	2.6 (±21.26)	2.31	12.55	11.16 (±31.50)	9.98	11.81
	Jalna	Mosambi	12.35 (±18.75)	14.32	–13.76	173.50 (±35.72)	193.31	–10.25
		Pomegranate	6.96 (±27.88)	7.38	–5.66	51.52 (±38.21)	55.34	–6.91
	Nandurbar	Banana	3.05 (±17.99)	2.91	4.81	74.86 (±19.07)	64.1	16.8
		Mango	1.75 (±23.62)	1.58	10.95	5.86 (±16.38)	4.73	23.85
		Papaya	1.97 (±22.58)	2.13	–7.56	67.56 (±28.30)	74.66	–9.50
	Nasik	Grape	50.53 (±21.79)	56.27	–10.19	1353.97 (±25.84)	1257.07	7.71
		Mango	5.72 (±33.14)	7	–18.29	12.84 (±35.96)	14.22	–9.70
		Pomegranate	40.10 (±22.47)	38.8	3.36	687.81 (±35.57)	628.11	9.5
	Ratnagiri	Mango	62.74 (±21.55)	60.1	4.39	220.32 (±47.81)	190	15.96
	Satara	Banana	0.4 (±26.50)	0.35	14.55	14.32 (±23.63)	12.31	16.4
		Mango	3.66 (±11.63)	2.92	25.54	14.86 (±17.14)	11.68	27.27
		Pomegranate	3.37 (±14.40)	3.97	–15.17	39.37 (±43.80)	43.67	–9.84
Tamil Nadu	Ariyalur	Banana	0.24 (±23.04)	0.21	15.31	10.72 (±29.05)	9.41	14.02
		Mango	0.70 (±18.14)	0.63	10.99	5.98 (±17.69)	5.02	19.05
	Erode	Banana	8.85 (±36.61)	9.92	–10.77	297.12 (±31.09)	324.73	–8.50
	Krishnagiri	Mango	40.47 (±18.68)	35.87	12.84	162.99 (±19.06)	137.72	18.35
	Salem	Banana	1.43 (±16.46)	1.3	10.2	39.69 (±20.83)	35.63	11.41
		Mango	5.70 (±26.59)	6.18	–7.72	19.04 (±29.18)	20.38	–6.58

Table 3. District-wise comparison of ICAR-IASRI and DAFW estimates of vegetable area and production in Andhra Pradesh, Himachal Pradesh, Maharashtra and Tamil Nadu states (2015-16)

State	District	Fruit	Area ('000 ha)			Production ('000 MT)		
			IASRI (±% SE)	DAFW	Deviation (%)	IASRI (±% SE)	DAFW	Deviation (%)
Andhra Pradesh	Anantpur	Green Chilli	4.18 (±14.46)	4.2	–0.47	119.43 (±29.37)	126	–5.21
		Tomato	11.29 (±20.47)	13.43	–18.94	450.38 (±17.66)	537.04	–16.14
	Chittoor	Brinjal	3.76 (±29.27)	3.6	4.22	81.23 (±34.25)	71.96	12.9
		Tomato	22.71 (±13.93)	22	3.11	927.75 (±18.56)	990	–6.29
	Guntur	Brinjal	1.11 (±19.06)	1.1	2.62	27.86 (±29.25)	27.07	2.91
		Green Chilli	0.40 (±3.20)	0.35	13	7104.45 (±22.52)	7000	1.49
		Tomato	1.25 (±16.39)	1.2	3.41	85.02 (±19.88)	72.18	17.8
	Kurnool	Cauliflower	0.10 (±3.20)	0.09	7.5	1.38 (±32.21)	1.16	19.01
		Cucumber	0.32 (±6.30)	0.27	23.07	2.46 (±75.34)	2.65	–7.00
		Tomato	4.11 (±13.95)	5.09	–23.95	79.32 (±9.92)	76.33	3.92
	West Godawari	Cauliflower	0.32 (±16.58)	0.21	34.9	39.51 (±46.15)	4.18	–5.60
		Cucumber	0.23 (±28.37)	0.25	–6.02	7.15 (±25.94)	6.15	16.36
		Bottle Gourd	0.12 (±28.09)	0.19	–51.06	2.76 (±56.32)	2.24	23
		Beans	0.33 (±29.82)	0.32	2.13	4.13 (±29.58)	4.77	–13.26
Himachal Pradesh	Kangra	Cauliflower	0.58 (±22.48)	0.56	3.8	12.91 (±24.58)	13.78	–6.28
		Peas	0.87 (±20.65)	0.75	15.35	9.41 (±32.25)	8.63	9.01
		Potato	1.98 (±18.25)	1.81	9.35	17.56 (±18.21)	16.88	4.01
	Shimla	Peas	6.62 (±33.62)	6.51	1.7	70.79 (±3.59)	71.7	–1.26
		Potato	6.27 (±21.91)	6.25	0.39	73.81 (±19.02)	62.52	18.05
		Tomato	0.58 (±2.33)	0.57	1.75	28.84 (±26.42)	23.89	20.69
	Solan	Lady Finger	0.16 (±2.30)	0.15	6.66	1.96 (±32.81)	1.79	9.76
		Onion	0.06 (±1.60)	0.05	20	0.88 (±47.24)	1.16	–24.00
		Potato	0.31 (±6.88)	0.3	3.84	2.83 (±35.42)	3.3	–14.33
		Tomato	4.81 (±24.68)	4.61	4.37	225.80 (±33.36)	230.4	–1.99
	Una	Cauliflower	2.20 (±5.20)	2	10	1.78 (±4.00)	2	–10.72
		Potato	1.04 (±0.12)	1	4	11.52 (±7.55)	12.5	–7.83
		Tomato	0.14 (±0.28)	0.12	16.67	3.52 (±13.29)	4.07	–13.54
Maharashtra	Beed	Onion	17.82 (±0.06)	16.33	9.14	209.46 (±14.09)	200.03	4.71
		Tomato	0.50 (±1.05)	0.3	66.84	3.93 (±5.09)	3.81	3.15
	Jalna	Green chilli	1.44 (±10.32)	1.23	16.64	15.094 (±27.36)	12.03	25.47
		Tomato	0.33 (±15.94)	0.36	–8.93	8.91 (±23.56)	7.55	17.98
	Nandurbar	Green chilli	3.08 (±5.25)	3.04	1.12	17.31 (±9.85)	14.92	16.01
		Onion	7.42 (±25.95)	7.47	–0.63	155.99 (±8.11)	127.23	22.61
		Tomato	0.06 (±21.03)	0.05	19	1.68 (±19.21)	1.12	49.74
	Nasik	Onion	129.72 (±29.99)	139.13	–6.77	2495.21 (±15.48)	2424.63	2.91
	Satara	Brinjal	0.77 (±26.97)	0.64	20.12	14.60 (±8.62)	13.19	10.7
		Green chilli	0.42 (±18.45)	0.45	–7.52	8.69 (±13.27)	9.08	–4.25
		Onion	11.88 (±44.09)	11.25	5.65	238.75 (±16.87)	179.92	32.7
		Peas	0.55 (±13.74)	0.37	47.1	2.29 (±35.82)	1.87	22.5
		Potato	1.05 (±23.13)	0.91	14.79	21.73 (±14.94)	22.85	–4.90
		Tomato	1.38 (±18.86)	1.28	8.21	29.48 (±12.29)	26.81	9.95
Tamil Nadu	Ariyalur	Bitter gourd	0.07 (±18.21)	0.07	0	1.54 (±16.45)	1.43	7.72
		Brinjal	0.36 (±14.72)	0.34	5.88	7.60 (±9.76)	6.9	10.19
		Lady finger	0.12 (±15.71)	0.12	0	1.32 (±35.14)	1.73	–23.64

		Tapioca	0.25 (±14.46)	0.29	–13.79	0.79 (±18.25)	10.25	–92.30
		Tomato	0.01 (±3.60)	0.011	–9.00	0.18 (±22.35)	0.15	19.19
	Erode	Potato	0.70 (±2.60)	0.59	18.64	11.88 (±5.60)	10.19	16.64
	Krishnagiri	Tomato	7.85 (±0.24)	6.86	–17.43	120.65 (±31.25)	100.49	20.07
	Salem	Tapioca	13.68 (±11.71)	12.32	11.04	32.18 (±31.25)	30.24	6.42
	Villupuram	Lady finger	0.35 (±0.08)	0.24	42.85	1.90 (±32.59)	1.74	9.15
		Onion	0.13 (±0.18)	0.15	–11.06	1.80 (±26.29)	1.92	–6.48
		Tomato	0.041 (±0.78)	0.03	47.07	0.43 (±19.25)	0.36	17.96
		Tomato	1.38 (±18.86)	1.28	8.21	29.48 (±12.29)	26.81	9.95

reduced variance. However, estimation uncertainty increased for minor, scattered, or heterogeneous crops, as reflected in elevated %SE values (>30% in several cases). The results further demonstrate that the approach effectively captures regional crop specialization patterns**,** while achieving substantial gains in operational efficiency through reduced sample size and integrated crop coverage.

The observed patterns highlight a fundamental trade-off between sampling efficiency and estimation precision in heterogeneous agricultural systems. While the proposed design improves cost-effectiveness and scalability, its performance is inherently influenced by spatial variability and crop distribution. High standard errors in low-prevalence crops suggest limitations of purely design-based inference under fragmented cultivation conditions. In this context, the integration of remote sensing-derived auxiliary variables (e.g., vegetation indices, crop masks) within a model-assisted framework could significantly enhance precision by reducing sampling variance and improving stratification. Thus, a hybrid approach combining ground-based sampling with geospatial data offers a promising pathway for developing a more robust, scalable, and spatially explicit methodology.

The estimates obtained in this study were validated against released official statistics by the Ministry of Agriculture & Farmers Welfare. Comparable studies employing similar large-scale, statistically rigorous sampling designs for horticultural crops are limited; therefore, official statistics were used as the primary reference standard for evaluation.

4. CONCLUSIONS

The study demonstrates that the modified sampling methodology provides reliable and operationally efficient estimates of horticultural crop area and production, with deviations from official statistics generally within tolerance level. It reduces sample size significantly (150–200 of CES&FV to ~80 villages per district) and enables integrated estimation of multiple crops, improving cost and survey efficiency. Higher variability in minor crops indicates scope for improvement using auxiliary data, including remote sensing within a model-assisted framework. Overall, the methodology offers a scalable and statistically robust approach for strengthening horticultural statistics in India.

REFERENCES

Ahmad, T., Bathla, H.V.L., Rai, A., Mathur, D.C. and Sood, R.M. (2011). An alternative methodology for estimation of area and production of horticultural crops. *Model Assisted Statistics and Applications*, **6**(4), 325-336. https://doi.org/10.3233/MAS-2011-0191

Ahmad, T., Bathla, H.V.L., Rai, A. and Sahoo, P.M. (2015). Estimation of area and production of horticultural crops in Himachal Pradesh. *Advances in Applied Research*, **7**(1), 1-8. **DOI:** 10.5958/2349-2104.2015.00001.7

Ahmad, T., Rai, A., Bathla, H.V.L. and Sharma S.D. (2014). Methodological issues relating to horticultural surveys in India. *Advances in Applied Research*, **6**(1), 164-172. **DOI:** 10.5958/j.2349-2104.6.1.001

Cochran, W.G. (1977). *Sampling Techniques*. 3rd Ed., Wiley Eastern Limited, New Delhi.

Government of India. (2017). *Horticultural Statistics at a Glance*. DAFW. (https://agriwelfare.gov.in/en/PublicationReports)

Ministry of Agriculture & Farmers Welfare (2026). Area and production of horticulture crops for 2024–25 (Final Estimates). Government of India, New Delhi.

Särndal C.E., Swensson B. and Wretman J.H. (1992). *Model Assisted Survey Sampling*. Springer: Verlag, New York.

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